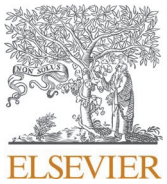
Engineering Applications of Artificial Intelligence 138 (2024) 109426



Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Graph neural networks-based dynamic water quality state estimation in water distribution networks

Aly K. Salem [a,b], Ahmad F. Taha [c], Ahmed A. Abokifa [a,*]

[a] Department of Civil, Materials, and Environmental Engineering, The University of Illinois Chicago, Chicago, IL, 60607, USA
[b] Faculty of Engineering, Cairo University, Giza, 12613, Egypt
[c] Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, 37235, USA

A B S T R A C T

Monitoring chlorine concentrations (CCs) plays an essential role in water quality management of water distribution networks (WDNs). However, placing water quality sensors at every junction is infeasible due to their high costs. Consequently, sensors are typically placed at a subset of junctions, which challenges the pursuit of a comprehensive assessment of chlorine concentrations throughout the network, which is necessary to ensure a safe water supply to end users. In this study, we tackle this challenge by introducing a framework for water quality state estimation (WQSE) using Graph Neural Networks (GNNs). WQSE reconstructs unmeasured CCs throughout the network based on measurements from a limited number of sensors distributed across the WDN. This study developed two GNN models to estimate CCs at all junctions. In the first model, a GNN model is trained to conduct Static Prediction (SP) of CCs based on data collected from a specific sensor network design (i.e., sensor placement configuration). In the second model, a GNN model is trained using data from various sensor designs to produce a generalized GNN model capable of conducting Dynamic Prediction (DP) of CCs. That is, the model can reconstruct CCs throughout the WDN based on data collected from any sensor network, even if different from those it was trained. The two models were applied to the C-Town benchmark network, considering that only 3% of the junctions were equipped with sensors. The results of the two models highlighted their ability to produce accurate predictions for intermediate junctions while struggling to predict CCs at dead-end junctions. The SP model outperformed the DP model in terms of accuracy. In addition, the SP model was shown to be robust against noisy measurements and produced better predictions than the physical model. The DP model stood out for its flexibility in being applicable to different sensor network designs. Furthermore, the DP model accuracy was shown to be highly dependent on the input sensor design, allowing for its implementation in sensor placement optimization.

1. Introduction

Water distribution networks (WDNs) are crucial infrastructures that sustain the well-being of communities by providing reliable conveyance of treated water. Due to various factors, including pipe corrosion, the formation of biofilms, and the potential for contaminant intrusion, water quality notably declines throughout WDNs (García-Ávila et al., 2021). To preserve satisfactory water quality all the way to demand locations, chlorine is typically employed as a disinfectant (Hallam et al., 2003). This is achieved by maintaining a minimum free chlorine residual to allow the chlorine to inhibit microbial growth and prevent waterborne diseases effectively (García-Ávila et al., 2021). Failing to uphold the chlorine concentrations (CCs) above the minimum threshold poses a health hazard to consumers (Monteiro et al., 2020). On the other hand, excessive chlorine can lead to the formation of harmful disinfection byproducts (Islam et al., 2017). For this reason, CCs are typically recommended to be below a maximum concentration of 4 mg/L as set by (USEPA, 1998).

In order to achieve the balance between effective disinfection and avoiding the formation of harmful byproducts, monitoring and controlling CCs is an integral part of WDNs management (Aisopou et al., 2012). The monitoring process involves continuous measurement of CCs throughout the network, which are then analyzed to determine if an intervention is needed to control these concentrations, such as injecting more chlorine through booster stations (Drewa and Brdys, 2007). Recently, technological advancements allowed operators to remotely measure CCs and other water quality parameters, such as pH and dissolved oxygen, through water quality sensors (Suresh et al., 2014). However, the high cost of the sensor makes it infeasible to place a sensor at every junction, limiting the ubiquitous monitoring of CCs throughout the network (Rajakumar et al., 2019). Similar limitations are encountered in monitoring various systems, such as power and transportation systems.

To tackle this issue, a State Estimation (SE) approach is commonly used to estimate the unmeasured variables based on limited measurements provided by available sensors (Yu and Powell, 1994). Implementing the SE approach is vital, as obtaining a precise estimation of a

* Corresponding author.
E-mail address: abokifa@uic.edu (A.A. Abokifa).

https://doi.org/10.1016/j.engappai.2024.109426
Received 25 March 2024; Received in revised form 1 September 2024; Accepted 1 October 2024
Available online 11 October 2024

Abbreviations

CCs	Chlorine Concentrations
DMAs	District Meter Areas
DP	Dynamic Prediction
GAT	Graph Attention Networks
GCN	Graph Convolutional Networks
GNNs	Graph Neural Networks
MAPE	Mean Absolute Percentage Error
PE	Percentage Error
SD	Standard Deviation
SE	State Estimation
SP	Static Prediction
TAGCN	Topology Adaptive Graph Convolution Networks
WDNs	Water Distribution Networks
WNTR	Water Network Tool for Resilience
WQSE	Water Quality State Estimation

particular system variable (CC in our case) assists in making informed decisions and accordingly, effectively managing the system of concern. For instance, determining the optimal locations and injection rates for chlorine booster stations requires a comprehensive understanding of chlorine levels throughout the network. These decisions are critical as they impact both the initial and operational costs of WDNs and ensure the safety of the water supplied to consumers. Generally, SE approach was applied in the field of WDNs to estimate flows, pressures, and water quality parameters (Ashraf et al., 2023; D'Souza and Kumar, 2010; Xing and Sela, 2022). SE was also used as an optimization metric in solving the sensor placement problem to enhance the observability of the water quality dynamics (Taha et al., 2021).

Several techniques have been proposed to solve the SE problem in various systems, which can be broadly classified into either model-based or data-driven SE approaches. A comprehensive review of SE approaches proposed for WDNs can be found in (Tshehla et al., 2017). For the majority of model-based approaches, the SE problem is typically formulated as an inverse modeling problem in which an optimization algorithm is implemented to estimate the unmeasured variables that minimize the differences between the measurements and the predictions at sensor locations (Andersen and Powell, 2000; Preis et al., 2011). However, this optimization requires conducting numerous evaluations of computationally expensive numerical models, which prohibits real-time SE. Other model-based studies attempted to get around this issue by mathematically formulating the numerical solution of the partial differential equations describing the WDN dynamics into a system of equations (Vrachimis et al., 2021; Wang et al., 2022). Nevertheless, such approaches are limited to simplified systems and are incapable of fully capturing the complexity of WDN dynamics.

On the other hand, data-driven SE approaches typically rely on machine learning methods to learn the patterns exhibited by the variable of concern at various WDN junctions (D'Souza and Kumar, 2010; May et al., 2008; Soyupak et al., 2011). While data-driven SE approaches do not require numerous computations at the prediction stage, they can only be trained to conduct predictions based on a specific sensor design. That is, they are incapable of producing predictions based on data collected from sensor designs other than that on which they were trained. This can be attributed to the fact that data-driven approaches merely learn the relationship between the inputs and outputs, without developing a representation of the topological structure (i.e., junction-pipe connectivity) of the underlying WDN.

Recently, Graph Neural Networks (GNNs) were introduced as an innovative technique to operate on complex graph-structured data. GNN was essentially introduced to perform social network analysis and recommender systems. However, the GNN's ability to learn the underlying relationships between the graph's nodes and edges contributed to expanding the GNN applications in various fields, this includes medicine to predict drug-drug interaction (Wang et al., 2021), economics to predict economic growth (Hui et al., 2020) and engineering to predict power flow (Donon et al., 2020), and traffic flow (Bao et al., 2024; Huang et al., 2024). In the field of water distribution, GNNs were applied for estimating head and flow in WDNs considering supervised and semi-supervised approaches (Xing and Sela, 2022), estimating nodal pressure in WDNs (Ashraf et al., 2023), and estimating water loss in WDNs (Fu et al., 2024).

Despite the superiority of GNNs over other neural networks in handling graph-structured data, it has only been used by (Li et al., 2024) to perform CC predictions, in which a static prediction (SP) model was developed to predict CCs for a specific sensor placement configuration. Building on the previous work, which applied an SP GNN model for CC state estimation (i.e., considering fixed sensor configuration), this study proposes a dynamic prediction (DP) GNN model for estimating chlorine concentrations in water distribution networks, adaptable to any sensor placement. We begin by developing an SP GNN model to predict concentrations at unobserved junctions for a specific sensor design, serving as a proof of concept. We then evaluate the model's robustness against noisy data. Next, we introduce the DP GNN model that can perform CC state estimation for any sensor design, thereby extending the applicability of GNN in chlorine state estimation and paving the way for sensor placement optimization using GNN. In this study, we thoroughly analyzed and compared the performance, advantages, and limitations of both models.

2. Methodology

In this study, we aim to estimate CCs at every junction in the WDN, leveraging CC data at any subset of junctions (i.e., sensors), considering steady state conditions and a single species model. To attain this goal, we introduce two distinct GNN models: i) SP and ii) DP GNN models. GNN was employed in this study as it is the only efficient way to estimate CC knowing information from any subset of junctions. This is due to its ability to represent graph-structured data and understand node connectivity. In contrast, other traditional models, such as Multilayer Perceptron Neural Networks, would only be applicable when the spatial relationship between sensor and non-sensor junctions is constant. This highlights the necessity of using GNN in this context. Furthermore, one can't practically compare the GNN results to those of traditional models as doing so would necessitate training a vast number of independent models, each trained to perform predictions considering a specific sensor design.

As illustrated in Fig. 1, the development of the SP and DP models encompasses two stages: i) the pre-processing stage and ii) the GNN training and testing stage. This section provides a detailed explanation of these stages, while a pseudocode outlines the processes within these stages can be found in the supplemental materials.

2.1. GNN model formulation

The primary objective of a GNN model is to capture latent node representations of a given graph. These representations are intended to comprehensively encapsulate the underlying information and complex relationships presented in the graph to proficiently make a graph or nodal-level predictions. In the context of this research, the WDN is represented as a graph $G = (V, E, X, Z)$, where $V = \{v_1, v_2, v_3, \ldots, v_N\}$ is the set of the N nodes (i.e., junctions), $E = \{e_{uv} | \forall v \in V, u \in N(v)\}$ is the set of M edges (i.e., pipes), $X \in R^{N \times D}$ is the set of node features with D dimensions (e.g., junction demand, CC), $Z = R^{M \times L}$ is the set of edge features with L dimensions (e.g., pipe length, diameter). Since water flows through pipes in both directions, WDNs are presented as undirected graphs, in which the junctions' connectivity is represented by a

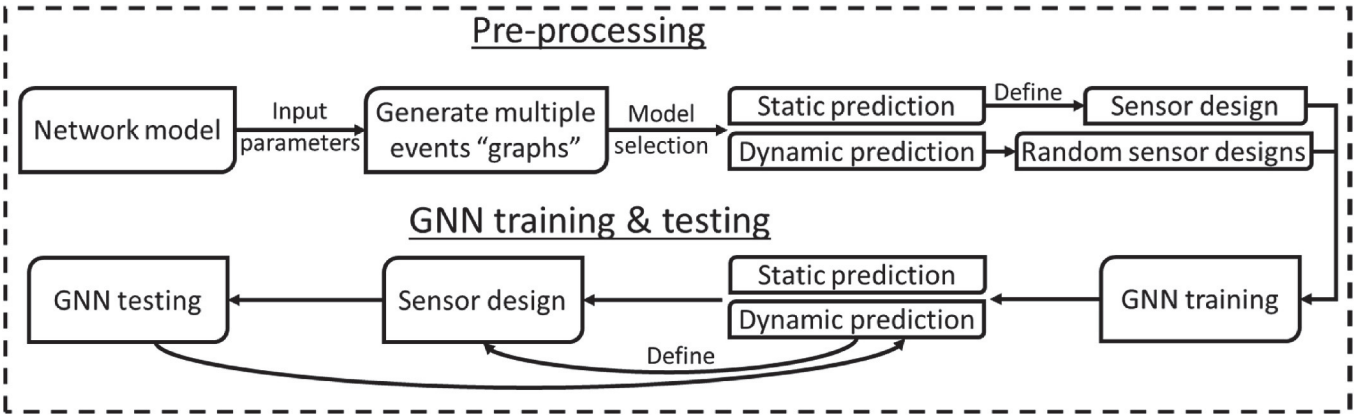


Fig. 1. The workflow of the proposed framework.

bidirectional (i.e., symmetrical) edge adjacency matrix (A). In this study, GNN is used to perform node-level predictions to predict CCs employing the features of the neighboring junctions while utilizing the junctions' connectivity.

2.1.1. GNN model variants

Two GNN models are proposed in this study, the SP model and the DP model. In the SP model, GNN is utilized to predict CCs considering a fixed set of sensors, where the junctions with available CC information (i.e., sensors) are constant across all events (i.e., graphs). In this model, the user defines the IDs of the sensors, while the CCs of all other junctions are considered unknown and are sought to be predicted by the GNN model. In comparison, the DP model is trained to make predictions based on measurements collected from any set of sensors, such as the case of mobile water quality sensors. This is done by employing a random set of sensors to train the GNN model. Furthermore, the number and locations of the sensors used in the prediction are dynamic. This is achieved by training the model to make predictions based on a range of sensor-to-non-sensor ratios defined to the model. This ratio is then randomly converted by the model to a certain sensor design (specific sensors' number and location) on an event basis, allowing for a different sensor design for each training event. A comparison of the training and prediction stages of the SP and DP models is visualized in Fig. 2. In addition, the DP model allows the user to assign a specific set of junctions to either sensors or non-sensor junctions, which gives flexibility to account for fixed sensors.

2.1.2. Dataset generation

The training and testing of the GNN models necessitate a dataset of a large number of events (i.e., graphs); vital to this process is the incorporation of a diverse array of events in this dataset, empowering the model to predict unforeseen events accurately. In this study, two key aspects distinguish each individual event within the dataset: i) *junctions' demands* ($d(n)$): controlled by the demand parameters which define the total demand (D) and the range of how many junctions in the network with demand (z_{min}, z_{max}), ii) *chlorine injection rates* ($I(n)$): controlled by the injection parameters which define the injection locations (m), and the range of injection rates at these locations (i_{min}, i_{max}). The determination of the number, location, and demand allocation for the network junctions in addition to the determination of the injection rates at the injection location follow a random selection process according to the user-defined parameters as shown in Eq. (1), and Eq. (2).

$$d(n) = \frac{r_n}{\sum r_n} \times D, r_n \begin{cases} U(0,1), n \in L \subset \{1,2,\ldots,N\} \mid |L| = U(z_{min}, z_{max}) \\ 0, n \notin L \subset \{1,2,\ldots,N\} \mid |L| = U(z_{min}, z_{max}) \end{cases} \quad (1)$$

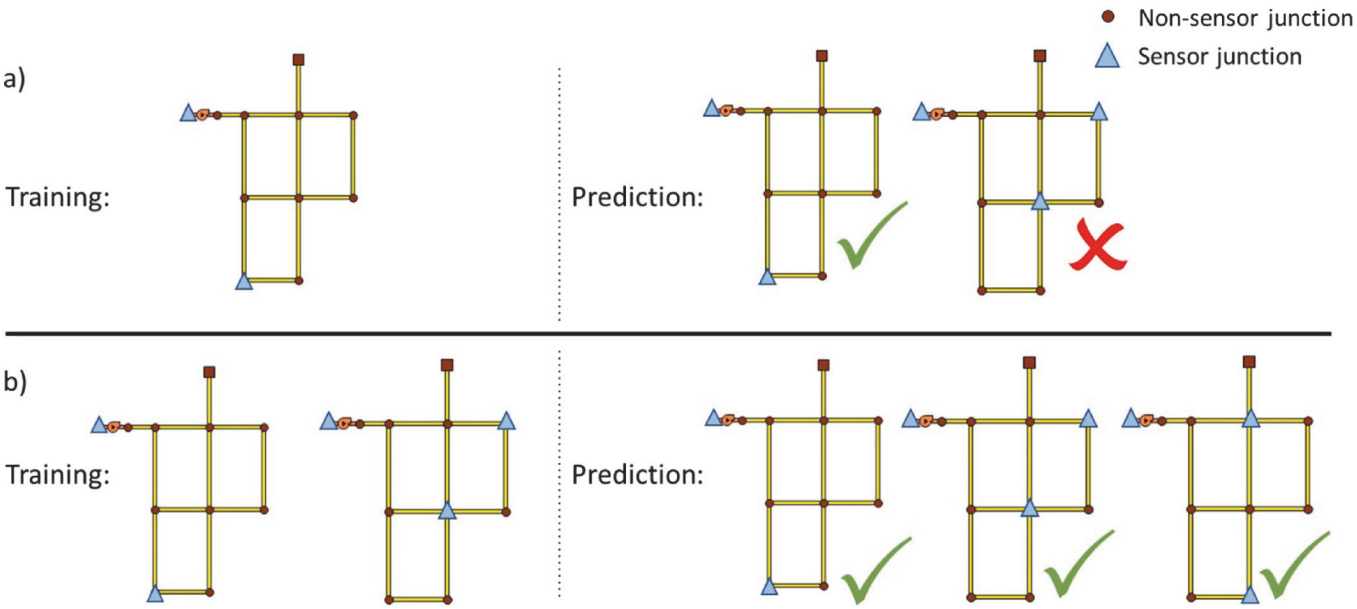


Fig. 2. A comparison of the training and prediction stages of the a) SP and the b) DP models.

$$I(n)=\begin{cases} U(i_{min},i_{max}), n\in m\subset\{1,2,\ldots,N\} \\ 0, n\notin m\subset\{1,2,\ldots,N\} \end{cases} \quad (2)$$

Once $d(n)$ and $I(n)$ are defined for all events, the Python interface of EPANET (WNTR) is employed to perform the hydraulic and water quality simulation of these events and to extract CCs at all network junctions (Klise et al., 2017). Subsequently, a masking process is implemented to remove the CC values of certain junctions. In the SP model, all the CC values are masked in all events except for sensor junctions defined by the user. On the other hand, in the DP model, different sets of junctions are masked in each event, and this masking process is performed randomly based on the sensor-to-non-sensor ratio range defined by the user (50% of junctions are masked at 0.5 ratio). In addition, the user can enforce a set of junctions to be sensors or non-sensors across all events in the DP model. A sample database representing the data generation process step by step for a small network covering both the SP and DP models is provided in the supplemental materials.

2.1.3. GNN model architecture

In this study, the Topology Adaptive Graph Convolution Networks (TAGCN) was used as the GNN model architecture (Du et al., 2017). TAGCN was chosen due to its superiority over other GNN architectures such as the Graph Convolutional Networks (GCN) and the Graph Attention Networks (GAT). Unlike GCN, TAGCN incorporates filters that act as attention coefficients that tailor the contributions of neighboring nodes during the aggregation process to effectively capture distinctive node features and prevent over-smoothing. Although GAT was introduced to address the same GCN limitation, it is not well-suited for large-scale graphs (Du et al., 2017). In contrast, TAGCN showed to scale well while allowing for dynamic adjustment of aggregation parameters on a local scale within each graph region. TAGCN effectively models complex relationships within diverse graphs by learning K number of graph filters as in Eq. (3), then it uses these graph filters along with learnable bias to perform predictions as in Eq. (4). The process begins by representing the WDN as a graph $G=(V,E,X,Z)$, where all junctions in the network are mapped to nodes (V), junction features to node features (X), junction IDs to edge indices (E), and pipe features to edge features (Z). In the next step, the graph is input into a dense layer to transform X for all junctions. The output of this dense layer is then combined with the edge adjacency matrix (A) to reconstruct $\acute{G}$ with transformed features. Subsequently, $\acute{G}$ is processed through multiple layers of TAGCN, and the output from the TAGCN is passed through another dense layer to predict the CCs for all junctions. The proposed model workflow is presented in Fig. 3.

$$\mathrm{G}_{f,x}^{(\ell)}=\sum_{k=0}^{K} g_{f,x,k}^{(\ell)}\,\mathrm{A}^k \quad (3)$$

$$\mathrm{Y}_{f}^{(\ell)}=\sigma\left(\sum_{x=1}^{X}\mathrm{G}_{f,x}^{(\ell)}\,\mathrm{x}_{x}^{(\ell)}+b_f 1_N\right) \quad (4)$$

where $\mathrm{G}_{f,x}^{(\ell)}$ is the f^{th} graph filter applied to the x^{th} feature in the ℓ^{th} layer; k and K are the filter index and the total number of filters; $g_{f,x,k}^{(\ell)}$ is the graph filter polynomial coefficients (i.e., learnable weights); A^k is a matrix that represents all length-k paths between all graph vertices (e.g., length-4 paths between node 1 to node 2); $\mathrm{Y}_{f}^{(\ell)}$ is the f^{th} output feature of the ℓ^{th} layer; σ denotes a rectified linear unit (*ReLU*); $\mathrm{x}_{x}^{(\ell)}$ is the input data of the ℓ^{th} layer for all nodes for x^{th} feature; b_f is a learnable bias, and 1_N is a unity vector of size N. Fig. 4 visualizes the main processes and equations employed in the TAGCN model. A detailed explanation of the TAGCN mathematical model can be found in Du et al., 2017.

The primary input of a TAGCN GNN model is the adjacency matrix (A) that is used to calculate the A^k term in Eq. (4). The adjacency matrix (A) represents the network's topology and is equal to $[a_{ij}]^{N\times N}$ with a_{ij} equals 1 if node i is connected to node j, and 0 otherwise. Additionally, a TAGCN GNN model incorporates node features, which in this study were composed of flows, CCs, and junction indicators. Positive flow values represent junction demands, whereas negative values indicate supply from source junctions (e.g., reservoirs). For CCs, the CC values are used wherever this information is available (e.g., at sensors), and 0 is held elsewhere. For the junction indicator, a value of 1 was used for junctions with known CC (i.e., a sensor), and 0 for junctions with unknown CC. The resulting junction input vector is $[Q,\ CC,\ J_{indicator}]$. Despite the TAGCN GNN model's exclusion of edge features, their inclusion would not have added value due to their consistency across different events.

During the training process, the GNN model utilizes a loss function to predict CCs based on some inputs. In this study, this loss function was defined as the normalized root mean square error $(nRMSE)$, as presented in Eq. (5). The *nRMSE* effectively encapsulates the overall CC prediction performance across all junctions and events. However, an additional error metric becomes necessary to assess individual junction prediction errors across various events. Hence, the percentage error (PE) and mean absolute percentage error $(MAPE)$ presented in Eq. (6) and Eq. (7) were used to evaluate these particular prediction errors.

$$nRMSE=\frac{\left[\sum_{g=1}^{G}\sum_{n=1}^{N}(C_n-\widehat{C}_n)^2/N\right]^{1/2}}{\sum_{g=1}^{G}\left(\sum_{n=1}^{N}C_n\right)/N} \quad (5)$$

$$PE_n=\frac{C_n-\widehat{C}_n}{C_n}\times 100 \quad \rightarrow C_n>0.01;\ PE_n \text{ is undefined otherwise} \quad (6)$$

$$MAPE_n=\frac{\sum_{g=1}^{G}|PE_n|_g}{G}\times 100 \quad (7)$$

where *nRMSE* is the normalized root mean square error; C_n and $\widehat{C}_n$ are the actual and predicted chlorine concentrations; n is the junction index

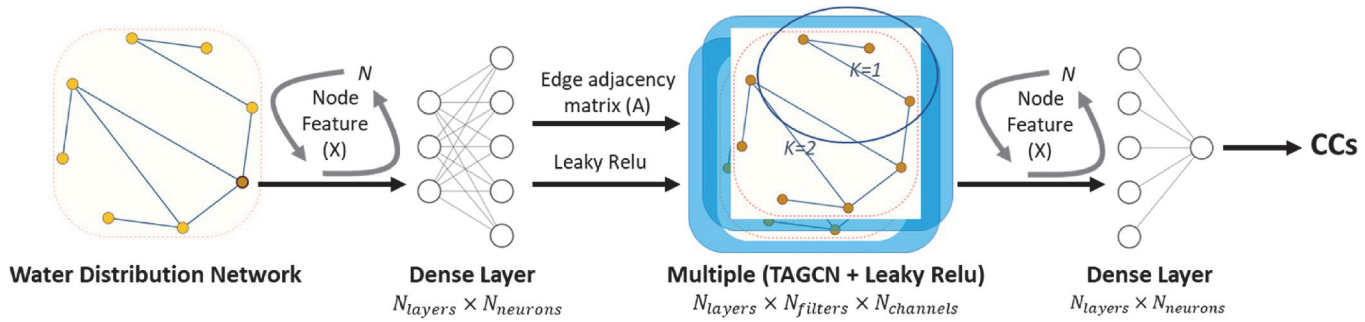


Fig. 3. The proposed model workflow.

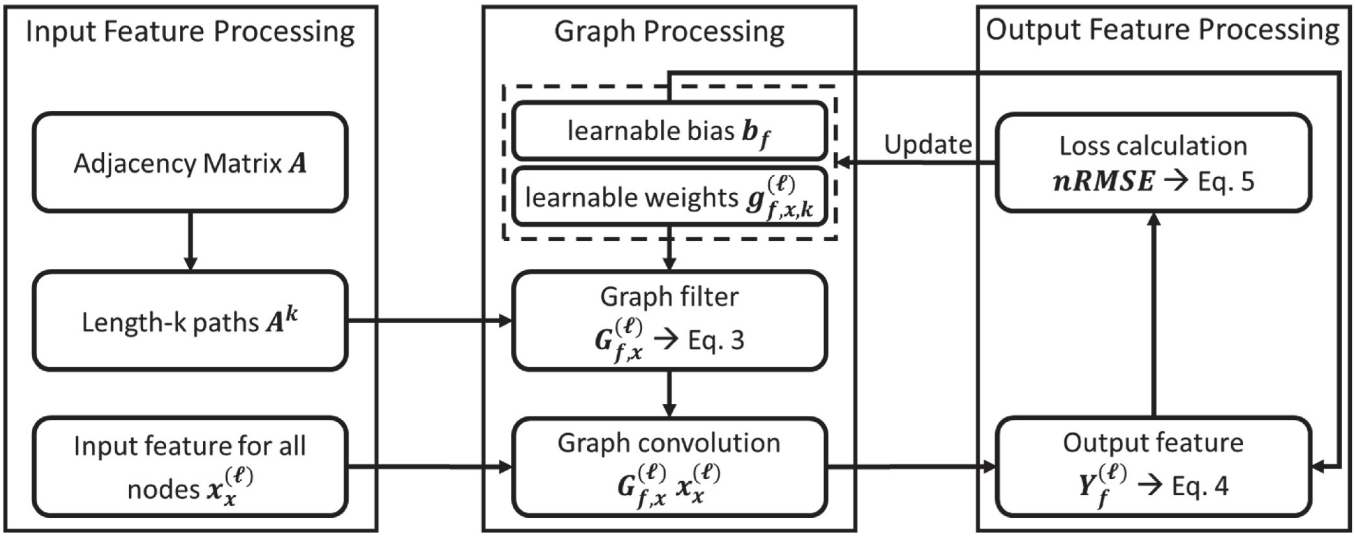


Fig. 4. Overview of the main processes and equations involved in the TAGCN model.

out of the N total number of junctions; g is the graph index out of G total number of graphs.

3. Case study

The C-town benchmark network was used to test the performance of the developed GNN model due to its extensive use in WDN-related studies (Brahmbhatt et al., 2023; Ostfeld et al., 2012; Rajabi and Tabesh, 2024; Tornyeviadzi et al., 2024). The network has 5 district meter areas (DMAs), each with a pumping system, with a monthly average total demand of 175 L/s. The network consists of 388 junctions, 429 pipes, four valves, one reservoir, and seven tanks (Fig. S2). The C-town network model has multiple controls to manage the value operations based on the tank status (i.e., filling or emptying). Since the steady-state condition is assumed in this study, the tanks and the associated controls were eliminated. The C-Town network was employed in this study to generate a synthetic dataset for training and testing the proposed GNN models. To accomplish this, it is assumed that chlorine is being introduced to the network at the main water source (i.e., reservoir) with a minimum and maximum concentration of 1 mg/L and 4 mg/L, respectively. The actual value of the chlorine injection varies from event to another and is selected randomly as previously explained in the dataset generation section. For the chlorine reaction kinetics, a first-order decay rate is assumed with global bulk and wall coefficients of 0.55 day^{-1} and 0.3 m/day respectively (Rossman et al., 1994).

3.1. Test scenarios

In this study, four different investigations were conducted to (i) establish the performance of the SP model for CC state estimation by comparing its predictions to the actual chlorine concentrations, (ii) check the influence of measurement noise on the performance of the SP model by introducing normally distributed noise, (iii) examine the performance of the DP model by comparing its predictions to that of the SP model, (iv) test the sensitivity of the DP model to the input sensor design and its applicability in performing sensor optimization by comparing the DP model predictions of three different sensor designs.

3.2. Sensor designs

The SP model requires a prior definition of the sensors' number and location, integrating this data as its input. On the other hand, the DP model can undergo training without this information. Nevertheless, it is important to adopt an effective strategy for positioning water quality sensors, as sensor placement significantly impacts the performance of chlorine concentration state estimation. Although sensor optimization does not fall within the primary focus of this study, we aimed to employ a decent sensor design that ensures good representation for all junctions. This was achieved by utilizing the Spectral Clustering approach (Von Luxburg, 2007) to produce sensor design A, in which the junctions were grouped into 9 clusters, and then, a sensor was placed in the center of each cluster, while the 10th sensor was placed at the injection source. By relating the resulting sensor design to the network district meter areas (Fig. 5), we observe that at least one sensor was placed in each DMA, which provides the desired spread.

Sensor designs B and C were also introduced to test the DP model sensitivity to different sensor designs. The objective in these two designs was to achieve better monitoring at localized DMA (e.g., 1 and 2, respectively) by concentrating the sensors at these DMAs. Fig. 5 shows the location of the different sensor designs with respect to the DMAs. All taken together, sensor design A will be used to train the SP model and will be applied along with sensor designs B and C to test the DP model.

3.3. Used parameters

As elaborated in section 2.1.2, a substantial volume of events is essential for the creation of the training and testing datasets of the GNN models. The parameters employed to construct these datasets are presented in Table 1. Furthermore, the GNN training is controlled by various hyperparameters, such as the number of GNN layers and hidden channels within each layer. Hyperparameter tuning was conducted to find the optimal set of hyperparameters for this study. The used hyperparameters are summarized in Table 2.

4. Results and discussion

4.1. Static prediction (SP) model

In this section, the SP model performance is presented by comparing its predictions to the actual CCs. Fig. 6 depicts a one-to-one plot featuring this comparison for (a) sensors and (b) non-sensor junctions, with the actual CC on the x-axis and the predicted CC on the y-axis, along with the 45-degree line which symbolizes a perfect match. The SP model predictions almost perfectly matched the actual CCs for the sensor junctions (Fig. 6a) with a *nRMSE* of 0.007. This very high accuracy can be attributed to the availability of actual CC data at these junctions,

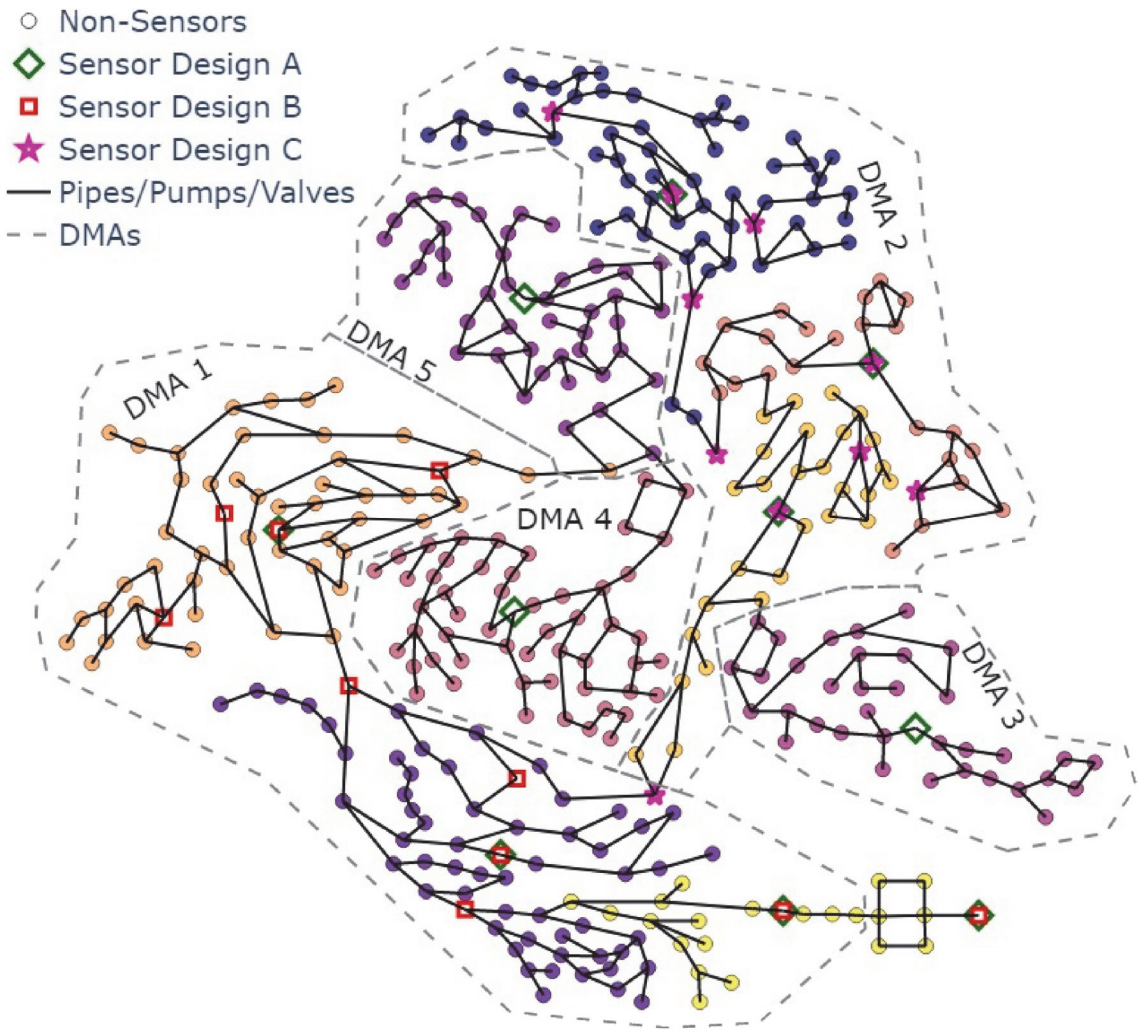


Fig. 5. Sensor designs used in this study.

Table 1
Generation and masking parameters.

Generation Parameter	Value	SP model Masking Parameter	
Total number of events	50,000	Follows sensor designs in Figure 3	
Total demand	175 L/s		
Range of demand junctions' number	[338-388]	DP model Masking Parameter	Value
Injection locations	R1	Minimum and maximum masking ratio	(0, 0.98)
Injection rates range	[1,4]	Forced masked/unmasked junctions	None

Table 2
GNN hyperparameters.

Number of:	Value	Parameter	Value
Dense layers	1	Learning rate	0.001
TAGCN layers	4	Train/Test split ratio	0.8
Filters	3	Total epochs	100
Hidden channels	100		

which was provided as input to the model. For the non-sensor junctions (Fig. 6b), the SP model showed notably lower accuracy in predicting the actual CCs compared to sensor junctions. Nevertheless, the error was still fairly low (*nRMSE* 0.029) reflecting the reliability of the SP model. The fact that CCs slightly diverged from the actual CCs at the sensor junctions is also worth highlighting as it reflects a non-trivial relationship between inputs and outputs within the developed GNN model. This is also reflected in the predictions for the non-sensor junctions, where input CCs are zeros while output CCs are fairly accurate.

Fig. 6b shows that the SP model predictions of high CCs (e.g., the upper left portion of the plot) were relatively more accurate than those with low CCs. This is because high CCs are associated with the junctions near the injection source where the decay is not very prominent. On the other side, the SP model predictions of zero CCs (e.g., the lower right portion of the plot) associated with the dead-end junctions were

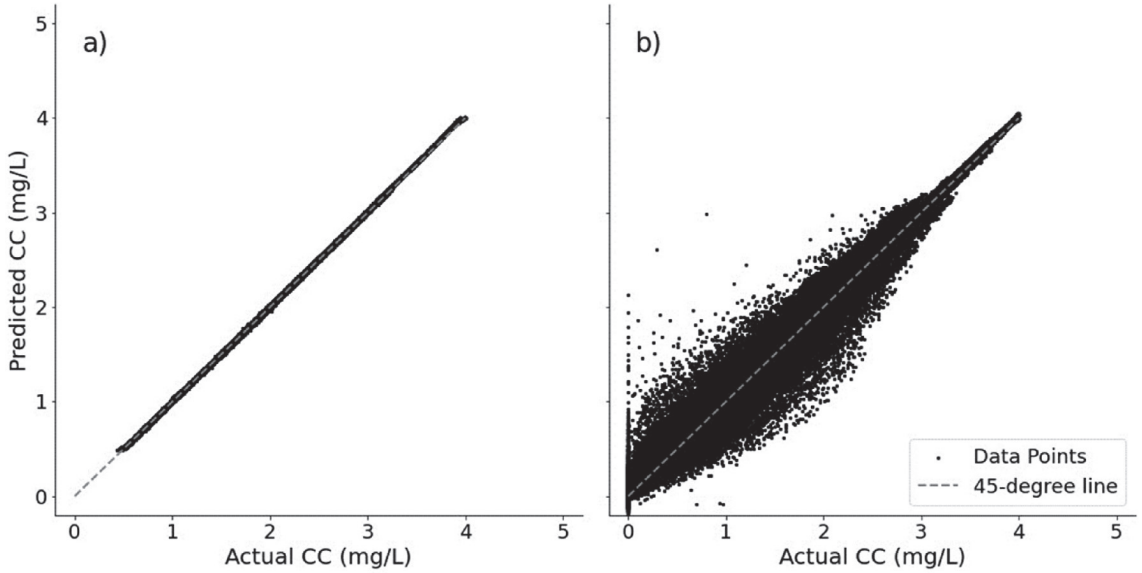


Fig. 6. One-to-one plot for SP model predictions for a) sensors, and b) non-sensors.

significantly less accurate. This weak performance highlights the SP model's struggle to differentiate between zero demands at intermediate junctions and dead-end junctions. In the case of dead-end junctions with zero demands, the water is stagnant, and accordingly, the CCs are zero (i.e., for the case of steady state), whereas for intermediate junctions with zero demands, the water is flowing, yet CCs are non-zero, leading the SP model to predict non-zero CCs for the dead-end junctions.

Switching from the overall performance and focusing on the inter-junction performance among all the events by investigating the mean absolute percentage error (*MAPE*), Fig. 7 shows a heatmap of the SP model predictions *MAPE* for all junctions. This figure shows very small errors were achieved for the majority of the intermediate junctions ($MAPE < 5\%$). Fig. 7 also shows that the *MAPE* is correlated to the distance to the nearby sensors, where junctions closer to the sensors tend to have smaller errors compared to further junctions. Similar to Fig. 6b, the SP model performed poorly at the dead-end junctions, resulting in high *MAPE*.

Looking at the distribution of the Percentage Error (*PE*) across all

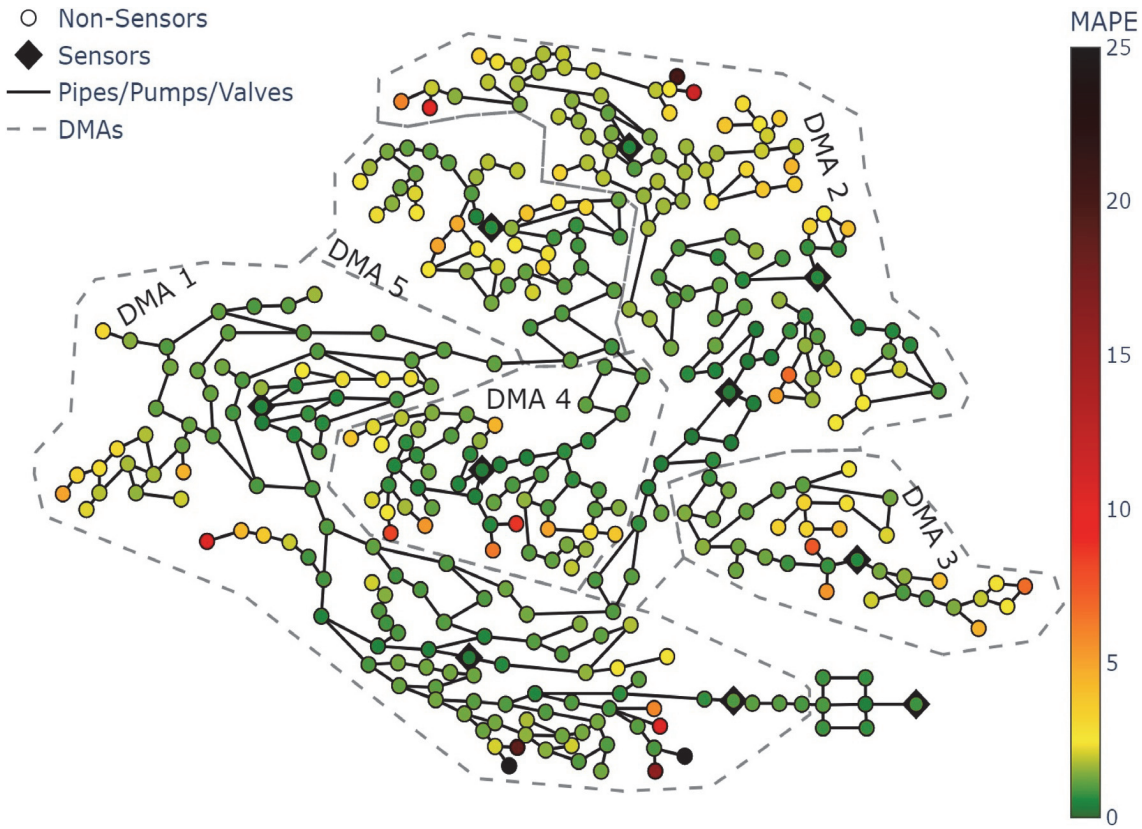


Fig. 7. Heatmap of the SP model predictions MAPE.

tested events (10,000 events), Fig. S3 shows that 96% of the junctions had their median *PE* between ±3% (J144* to J27). This highlights the consistency of the SP model prediction across different events. Moreover, only 2% of the data points (data point: junction prediction of an event) had errors of more than ±30%. Fig. 8 shows the *PE* distribution of a subset of junctions, in which they are classified into three groups based on the prediction trend; group A: represents the ordinarily predicted junctions with median *PE* between ±3%., group B: features overly predicted junctions with median $PE < -3\%$, and group C: represents underly-predicted junctions with median $PE > 3\%$.

All intermediate junctions were included in Group A with the majority of dead-end junctions, resulting in a total of 373 junctions, whereas groups B and C only included eight junctions each, featuring some of the dead-end junctions. By projecting the three groups on the network layout in Fig. 8, it was found that the overly and underly predicted junctions form pairs, where one dead-end junction lies in group B, and the other lies in group C. More interestingly, the SP model predictions seemed to follow a systematic over/under estimation pattern. Fig. S4 focuses on two branching regions and shows the lengths of the pipes connecting the dead-end junctions to the main junction, and the one-to-one plot for these junctions. The figure shows that the SP model systematically overestimated the CCs of the longer pipes while underestimating the CCs of the shorter pipes.

4.2. Noisy sensor measurements

In this section, we test the SP model robustness against noisy sensor measurements by enforcing a normally distributed measurement noise with a mean of 0 and a standard deviation (SD) of 0.5 mg/L in the training data. After training the SP model, it was then tested on different noise distributions with SD of 0, 0.2 and 0.5 mg/L to examine its performance on unforeseen noise distributions.

Fig. 9 shows the heatmaps of the SP model predictions *MAPE* for the three different SD. This figure shows that accounting for noisy measurements with an SD of 0.2 mg/L didn't significantly increase the *MAPE* compared to the case of zero noise. In contrast, the same figure shows that the *MAPE* significantly increased when the noise SD increased to 0.5 mg/L and most of the junctions had a *MAPE* between 10% and 30%. Moreover, by comparing the zero noise predictions of this SP model (i.e., trained on noisy data) shown in Fig. 9a to those of the previous SP model (i.e., trained without noise) shown in Fig. 7, it was noticed that the latter model predictions were more accurate since it was trained on non-noisy measurements.

Considering the case with the highest noise (SD = 0.5 mg/L), the SP model predictions were compared to the physical model results (EPANET simulations). This was done by re-running the physical model considering the noisy measurement as the injection CC at the injection source (R1). Interestingly, the SP model predictions were more accurate than the physical model (Fig. S5), with *nRMSE* of 0.162 and 0.209, respectively. This highlights the advantage of using the SP model over the physical model as it takes advantage of the measurements from all sensors, unlike the physical model that only utilizes the measurement at the injection source to predict the CCs of all network junctions.

4.3. Dynamic prediction (DP) model

In this section, the DP model predictions are compared to the actual CCs and those of the SP model. Fig. 10 depicts a one-to-one for the predicted and actual CCs for sensors and non-sensor junctions. This figure shows that the DP model predictions were reasonably accurate for both sensors and non-sensor junctions with *nRMSE* of 0.026 and 0.078, respectively. Comparing the DP model to the SP model, Fig. 10a shows that the DP model predictions for the sensors' junctions were not as accurate as those of the SP model (Fig. 6a). This can be attributed to the

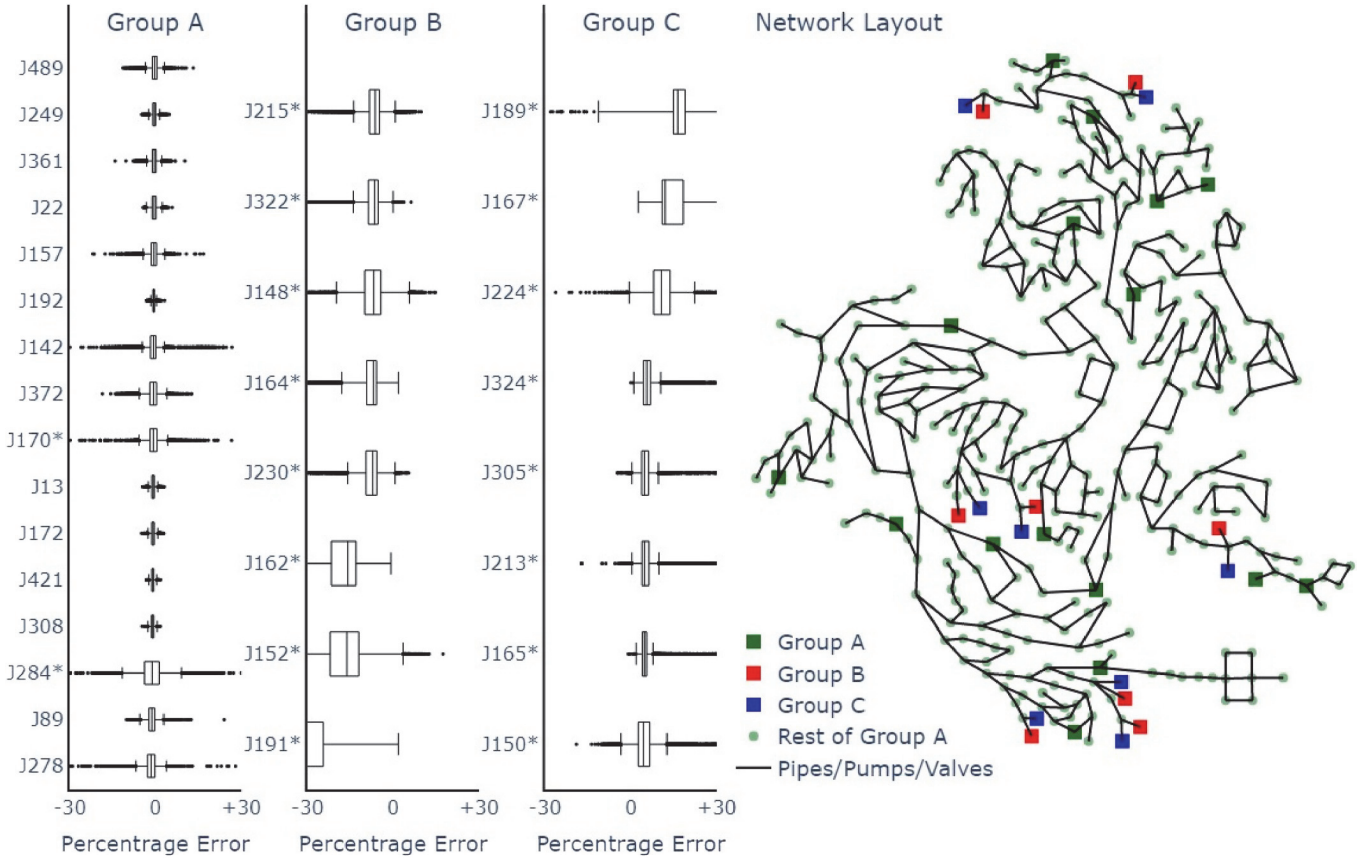


Fig. 8. The percentage error distribution of the SP model predictions for the three groups projected on the network layout (*: dead-end junction).

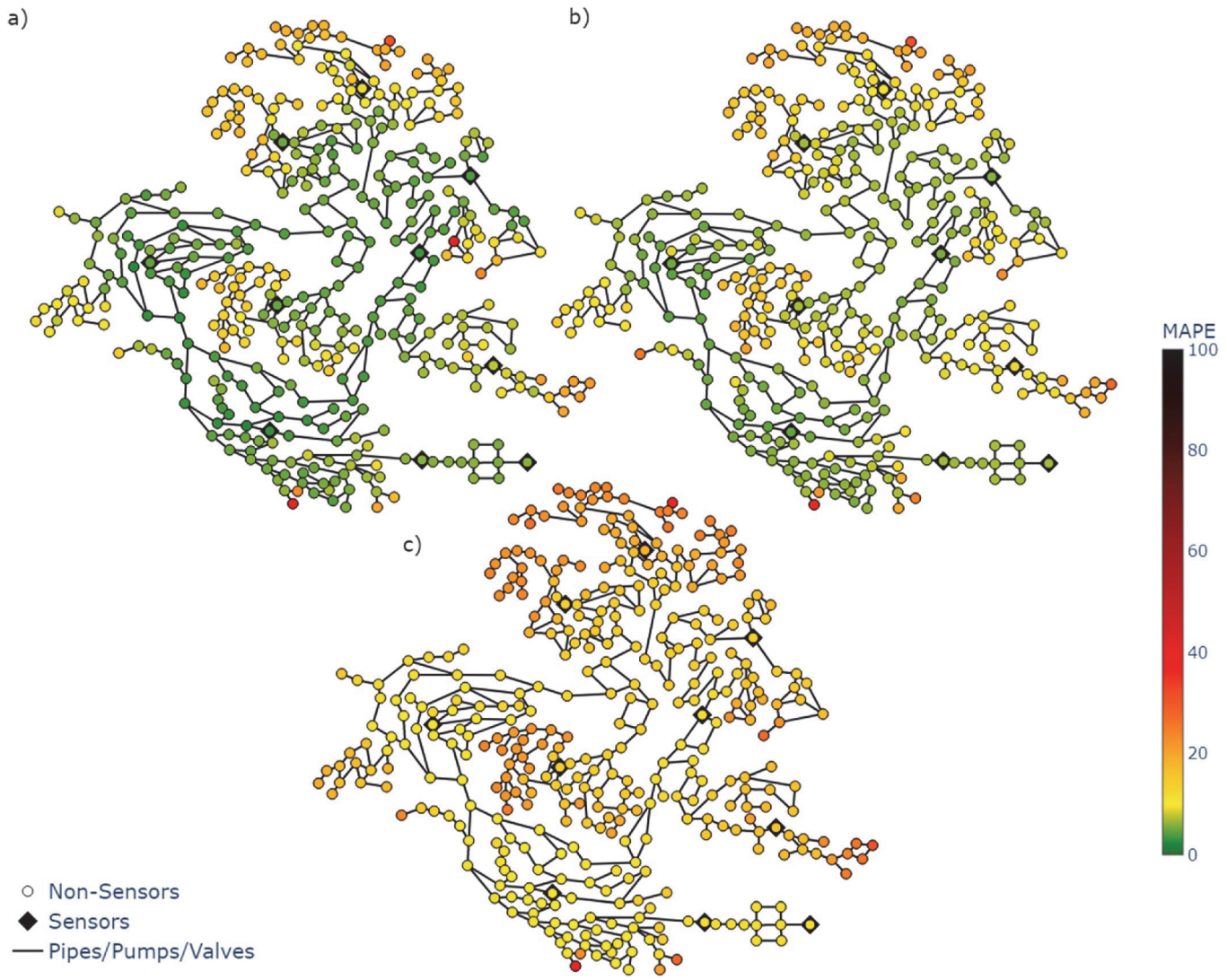


Fig. 9. Heatmaps of the SP model predictions for noise with SD of a) zero, b) 0.2 mg/L and c) 0.5 mg/L.

fact that the DP model was trained on different sensor designs, unlike the SP model which was trained on the same sensor design. A similar observation can be noticed for the non-sensor junctions (Figs. 10b and 6b).

Investigating the inter-junctions prediction errors by looking at the *MAPE* heatmaps of the DP model and the SP model shown in Fig. 11, a noticeable accuracy difference can be seen in favor of the SP model, which aligns with the one-to-one results. However, the two models were similar in having relatively higher errors at dead-end junctions. Looking at the *PE* distribution of the DP model predictions (Fig. S6) and following the same grouping criteria as earlier, group A (median *PE* between $\pm$ 3%) contained 43% of the junctions, whereas group B (median *PE* $< -$ 3%) and group C (median *PE* $> 3\%$) contained 48% and 9% of the junctions, respectively. These results show that the DP model predictions are less consistent across different events, featuring significantly wider *PE* distribution. Moreover, no over/under estimation pattern was observed for the DP model predictions (Fig. S7).

4.4. Effect of sensor design

Although the DP model predictions were outperformed by those of the SP model, the DP model has the advantage of being applicable to different sensor designs. Thus, the DP can be useful in the optimization of sensor network design. In this section, the DP model sensitivity to the input sensor design is tested by comparing the previous DP model predictions that utilize sensor design A (presented above) to those that utilize sensor designs B and C (Fig. 5). Fig. 12 shows the heatmaps of the DP model predictions *MAPE* after applying sensor designs B and C. Even though the same DP model was used, significantly different error patterns were realized. This highlights the impact of the sensor design on the prediction accuracy, and proves the DP model's ability to be used in sensor optimization as it had different responses to the different input designs,

Fig. 12 also shows that concentrating the sensors in one DMA has significantly enhanced the prediction accuracy of this DMA while deteriorating the prediction accuracy of the other DMAs. These results show the importance of using appropriate sensor design in order to gain an overall accurate prediction for all network junctions, such as sensor design A (Fig. 11). Taken together, the results revealed that the DP model is superior to the SP model in the aspect of having the flexibility to evaluate the performance of various sensor designs, demonstrating its applicability in the context of sensor placement optimization.

5. Conclusions

In this study, we tackle the problem of estimating CCs in WDNs by utilizing GNNs. This was achieved by predicting CCs at all network junctions relying solely on the information from a limited set of sensors.

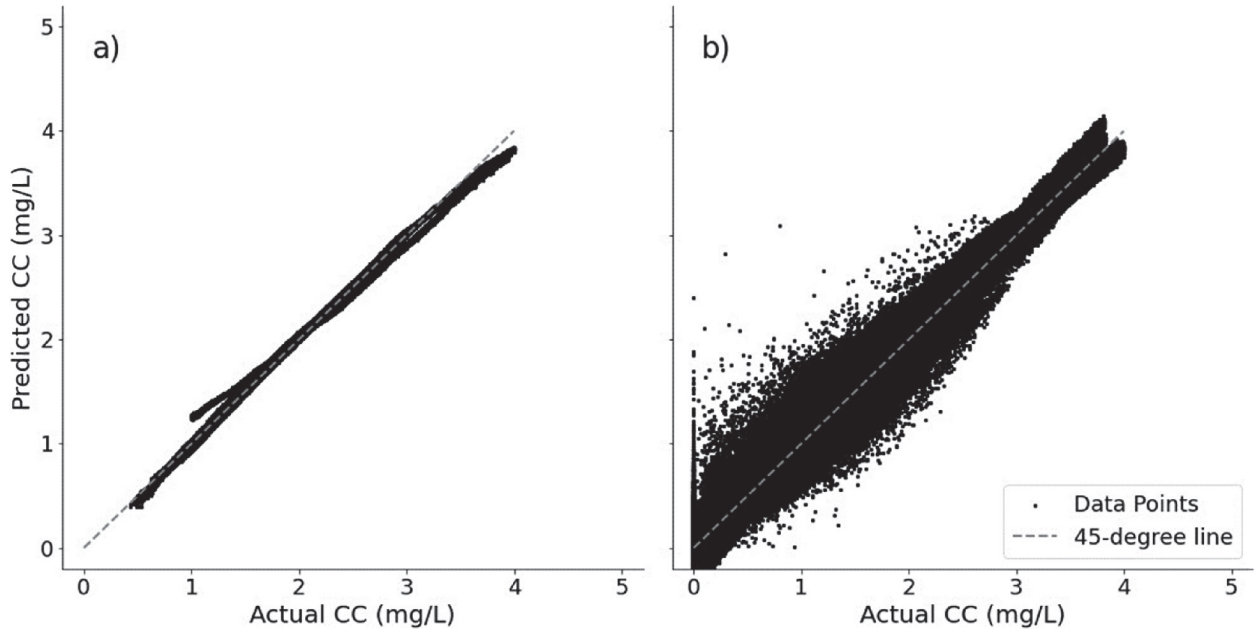


Fig. 10. One-to-one plot for DP model predictions for a) sensors, and b) non-sensors.

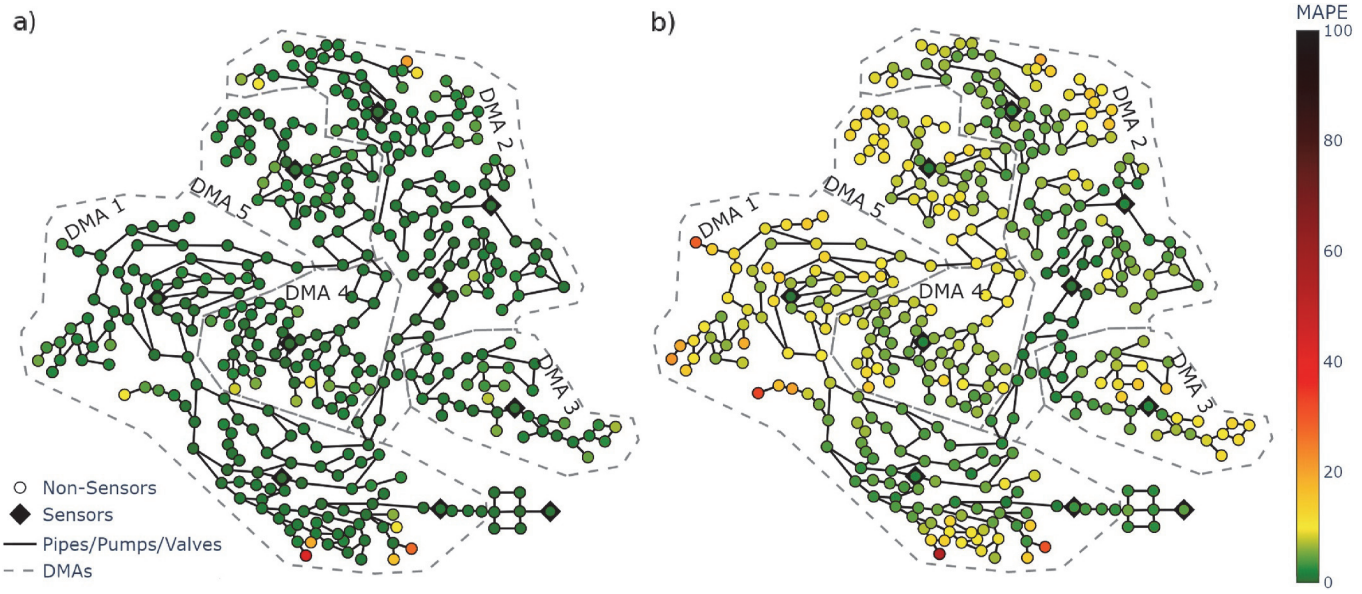


Fig. 11. Heatmaps of the MAPE for a) the SP model and b) the DP model.

Two GNN models were developed to predict CCs, namely, i) Static Prediction (SP) model, and ii) Dynamic Prediction (DP) model. The SP model predicts CCs at non-sensor junctions considering the information from a specific sensor configuration (i.e., sensor design) on which it was initially trained, and hence, it requires the sensor design to be defined in advance. On the other hand, the DP model has the capability to predict CCs by utilizing information from different sensor designs, omitting the need to specify the sensor design in the training process. The performance of the two models was tested by applying them to the C-Twon benchmark network.

The SP model results revealed the model's ability to accurately predict CCs for the non-sensor junctions with minor errors. It was evident that the SP model predictions of the intermediate junctions were significantly more accurate than the dead-end junctions. The investigation of the error distribution of each junction among all events highlighted the SP model's ability to precisely predict CCs across different events. However, the SP model predictions were shown to systematically overestimate dead-end junctions connected to longer pipes and underestimate those connected to shorter pipes. Furthermore, the SP model showed a high robustness to noisy measurements, and its predictions were shown to be more accurate than the physical model when high noise was implemented. In terms of accuracy and prediction consistency, the DP model delivered fairly accurate results, although it was outperformed by the SP model as it had higher errors and broader error distribution. Nevertheless, the DP model stood out for its flexibility to be applied to different sensor designs. The DP model also showed a rational response when applied to different sensor designs, showing its potential utilization in sensor placement optimization.

Future studies are encouraged to address the current study's limitation of steady-state conditions to better reflect WDN dynamics by

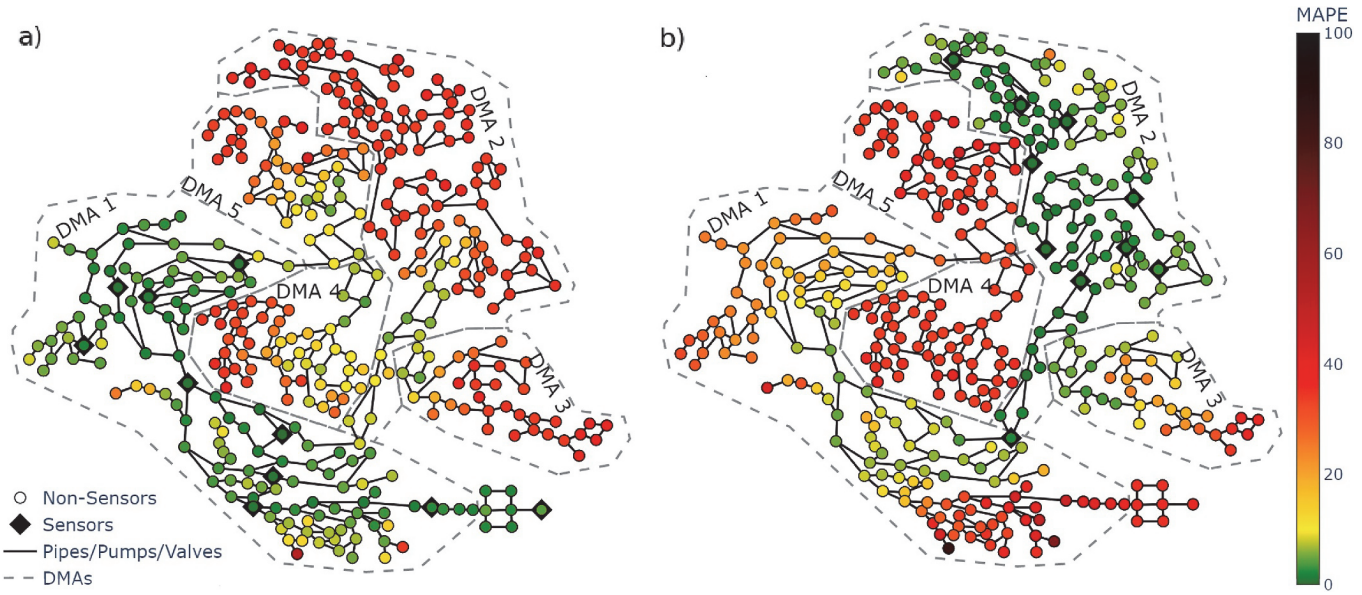


Fig. 12. Heatmaps of the DP model predictions MAPE for sensor designs a) B, and B) C.

incorporating the variation of demands and injection concentrations over time. Other studies can also explore future GNN architectures when they become available and compare their performance to the one used in this study. Finally, future research can utilize the introduced SP model in performing multi-species water quality state estimation, in addition to employing the DP model in evaluating the performance of different sensor design alternatives to obtain the optimal sensor design (Salem and Abokifa, 2024).

CRediT authorship contribution statement

Aly K. Salem: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ahmad F. Taha:** Writing – review & editing, Funding acquisition, Conceptualization. **Ahmed A. Abokifa:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Ahmed Abokifa and Ahmad Taha report financial support was provided by the National Science Foundation under awards number 2015603 and 2015671. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.engappai.2024.109426.

References

Aisopou, A., Stoianov, I., Graham, N.J.D., 2012. In-pipe water quality monitoring in water supply systems under steady and unsteady state flow conditions: a quantitative assessment. Water Res. 46, 235–246. https://doi.org/10.1016/j.watres.2011.10.058.

Andersen, J.H., Powell, R.S., 2000. Implicit state-estimation technique for water network monitoring. Urban Water 2, 123–130. https://doi.org/10.1016/s1462-0758(00)00050-9.

Ashraf, I., Hermes, L., Artelt, A., Hammer, B., 2023. Spatial graph convolution neural networks for water distribution systems. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 29–41. https://doi.org/10.1007/978-3-031-30047-9_3.

Bao, Y., Shen, Q., Cao, Y., Ding, W., Shi, Q., 2024. Residual attention enhanced time-varying multi-factor graph convolutional network for traffic flow prediction. Eng. Appl. Artif. Intell. 133, 108135. https://doi.org/10.1016/j.engappai.2024.108135.

Brahmbhatt, P., Maheshwari, A., Gudi, R.D., 2023. Digital twin assisted decision support system for quality regulation and leak localization task in large-scale water distribution networks. Digit. Chem. Eng. 9, 100127. https://doi.org/10.1016/j.dche.2023.100127.

D'Souza, C.D., Kumar, M.S.M., 2010. Comparison of ANN models for predicting water quality in distribution systems. J. Am. Water Works Assoc. 102, 92–106. https://doi.org/10.1002/j.1551-8833.2010.tb10152.x.

Donon, B., Clément, R., Donnot, B., Marot, A., Guyon, I., Schoenauer, M., 2020. Neural networks for power flow: graph neural solver. Elec. Power Syst. Res. 189, 106547. https://doi.org/10.1016/j.epsr.2020.106547.

Drewa, M., Brdys, M.A., 2007. Optimized allocation of chlorination stations for integrated quantity and quality control in drinking water distribution systems. IFAC Proceedings Volumes (IFAC-PapersOnline). IFAC. https://doi.org/10.3182/20070723-3-pl-2917.00011.

Du, J., Zhang, S., Wu, G., Moura, J.M.F., Kar, S., 2017. Topology Adaptive Graph Convolutional Networks, pp. 1–13.

Fu, M., Zhang, Q., Rong, K., Mundher, Z., Zheng, L., 2024. Engineering Applications of Artificial Intelligence Integrated dynamic multi-threshold pattern recognition with graph attention long short-term neural memory network for water distribution network losses prediction : an automated expert system. Eng. Appl. Artif. Intell. 127, 107277. https://doi.org/10.1016/j.engappai.2023.107277.

García-Ávila, F., Avilés-Añazco, A., Ordoñez-Jara, J., Guanuchi-Quezada, C., Flores del Pino, L., Ramos-Fernández, L., 2021. Modeling of residual chlorine in a drinking water network in times of pandemic of the SARS-CoV-2 (COVID-19). Sustain. Environ. Res. 31. https://doi.org/10.1186/s42834-021-00084-w.

Hallam, N.B., Hua, F., West, J.R., Forster, C.F., Simms, J., 2003. Bulk decay of chlorine in water distribution systems. J. Water Resour. Plann. Manag. 129, 78–81. https://doi.org/10.1061/(asce)0733-9496(2003)129:1(78.

Huang, C., Chen, D., Fan, T., Wu, B., Yan, X., 2024. Incorporating environmental knowledge embedding and spatial-temporal graph attention networks for inland vessel traffic flow prediction. Eng. Appl. Artif. Intell. 133, 108301. https://doi.org/10.1016/j.engappai.2024.108301.

Hui, B., Yan, D., Ku, W.S., Wang, W., 2020. Predicting economic growth by region embedding: a multigraph convolutional network approach. Int. Conf. Inf. Knowl. Manag. Proc. 555–564. https://doi.org/10.1145/3340531.3411882.

Islam, N., Sadiq, R., Rodriguez, M.J., 2017. Optimizing locations for chlorine booster stations in small water distribution networks. J. Water Resour. Plann. Manag. 143, 1–16. https://doi.org/10.1061/(asce)wr.1943-5452.0000759.

Klise, K.A., Hart, D., Moriarty, D., Bynum, M.L., Murray, R., Burkhardt, J., Haxton, T., 2017. Water Network Tool for Resilience (WNTR) User Manual Disclaimer. United State Environmental Protection Agency. https://doi.org/10.2172/1376816.

Li, Z., Liu, H., Zhang, C., Fu, G., 2024. Real-time water quality prediction in water distribution networks using graph neural networks with sparse monitoring data. Water Res. 250, 121018. https://doi.org/10.1016/j.watres.2023.121018.

May, R.J., Dandy, G.C., Maier, H.R., Nixon, J.B., 2008. Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. Environ. Model. Software 23, 1289–1299. https://doi.org/10.1016/j.envsoft.2008.03.008.

Monteiro, L., Carneiro, J., Covas, D.I.C., 2020. Modelling chlorine wall decay in a full-scale water supply system. Urban Water J. 17, 754–762. https://doi.org/10.1080/1573062X.2020.1804595.

Ostfeld, A., Salomons, E., Ormsbee, L., Uber, J.G., Bros, C.M., Kalungi, P., Burd, R., Zazula-Coetzee, B., Belrain, T., Kang, D., Lansey, K., Shen, H., McBean, E., Yi Wu, Z., Walski, T., Alvisi, S., Franchini, M., Johnson, J.P., Ghimire, S.R., Barkdoll, B.D., Koppel, T., Vassiljev, A., Kim, J.H., Chung, G., Yoo, D.G., Diao, K., Zhou, Y., Li, J., Liu, Z., Chang, K., Gao, J., Qu, S., Yuan, Y., Prasad, T.D., Laucelli, D., Vamvakeridou Lyroudia, L.S., Kapelan, Z., Savic, D., Berardi, L., Barbaro, G., Giustolisi, O., Asadzadeh, M., Tolson, B.A., McKillop, R., 2012. Battle of the water calibration networks. J. Water Resour. Plann. Manag. 138, 523–532. https://doi.org/10.1061/(asce)wr.1943-5452.0000191.

Preis, A., Whittle, A.J., Ostfeld, A., Perelman, L., 2011. Efficient hydraulic state estimation technique using reduced models of urban water networks. J. Water Resour. Plann. Manag. 137, 343–351. https://doi.org/10.1061/(asce)wr.1943-5452.0000113.

Rajabi, M., Tabesh, M., 2024. Pressure sensor placement for leakage detection and calibration of water distribution networks based on multiview clustering and global sensitivity analysis. J. Water Resour. Plann. Manag. 150, 1–16. https://doi.org/10.1061/jwrmd5.wreng-6262.

Rajakumar, A.G., Mohan Kumar, M.S., Amrutur, B., Kapelan, Z., 2019. Real-time water quality modeling with ensemble kalman filter for state and parameter estimation in water distribution networks. J. Water Resour. Plann. Manag. 145, 1–12. https://doi.org/10.1061/(asce)wr.1943-5452.0001118.

Rossman, L.A., Clark, R.M., Grayman, W.M., 1994. Modeling chlorine residuals in drinking-water distribution systems. J. Environ. Eng. 120, 803–820.

Salem, A.K., Abokifa, A.A., 2024. Optimal Sensor Placement in Water Distribution Networks Using Dynamic Prediction Graph Neural Networks. Eng. Proc. 69, 171. https://doi.org/10.3390/engproc2024069171.

Soyupak, S., Kilic, H., Karadirek, I.E., Muhammetoglu, H., 2011. On the usage of artificial neural networks in chlorine control applications for water distribution networks with high quality water. J. Water Supply Res. Technol. - Aqua 60, 51–60. https://doi.org/10.2166/aqua.2011.086.

Suresh, M., Manohar, U., Anjana, G.R., Stoleru, R., Mohan Kumar, M.S., 2014. A cyber-physical system for continuous monitoring of Water Distribution Systems. Int. Conf. Wirel. Mob. Comput. Netw. Commun 570–577. https://doi.org/10.1109/WiMOB.2014.6962227.

Taha, A.F., Wang, S., Guo, Y., Summers, T.H., Gatsis, N., Giacomoni, M.H., Abokifa, A.A., 2021. Revisiting the water quality sensor placement problem: optimizing network observability and state estimation metrics. J. Water Resour. Plann. Manag. 147, 1–13. https://doi.org/10.1061/(asce)wr.1943-5452.0001374.

Tornyeviadzi, H.M., Owusu-Ansah, E., Mohammed, H., Seidu, R., 2024. Node search space reduction for optimal placement of pressure sensors in water distribution networks for leakage detection. Alex. Eng. J. 94, 325–338. https://doi.org/10.1016/j.aej.2024.03.037.

Tshehla, K.S., Hamam, Y., Abu-Mahfouz, A.M., 2017. State estimation in water distribution network: a review. Proc. - 2017 IEEE 15th Int. Conf. Ind. Informatics, INDIN 2017, 1247–1252. https://doi.org/10.1109/INDIN.2017.8104953.

USEPA, 1998. National primary drinking water regulations: disinfectants and disinfection byproducts notice of data availability. Fed. Regist. 63 (61). https://www.govinfo.gov/content/pkg/FR-1998-03-31/pdf/98-8215.pdf.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416. https://doi.org/10.1007/s11222-007-9033-z.

Vrachimis, S.G., Eliades, D.G., Polycarpou, M.M., 2021. Calculating chlorine concentration bounds in water distribution networks: a backtracking uncertainty bounding approach. Water Resour. Res. 57, 1–22. https://doi.org/10.1029/2020WR028684.

Wang, S., Taha, A.F., Gatsis, N., Sela, L., Giacomoni, M.H., 2022. Probabilistic state estimation in water networks. IEEE Trans. Control Syst. Technol. 30, 507–519. https://doi.org/10.1109/TCST.2021.3066102.

Wang, Y., Min, Y., Chen, X., Wu, J., 2021. Multi-view graph contrastive representation learning for drug-drug interaction prediction. Web Conf. 2021 - Proc. World Wide Web Conf. WWW 2021, 2921–2933. https://doi.org/10.1145/3442381.3449786.

Xing, L., Sela, L., 2022. Graph neural networks for state estimation in water distribution systems: application of supervised and semisupervised learning. J. Water Resour. Plann. Manag. 148, 1–14. https://doi.org/10.1061/(asce)wr.1943-5452.0001550.

Yu, G., Powell, R.S., 1994. Optimal design of meter placement in water distribution systems. Int. J. Syst. Sci. 25, 2155–2166. https://doi.org/10.1080/00207729408949342.