



# Prioritizing water distribution pipelines rehabilitation using machine learning algorithms

Nehal Elshaboury<sup>1</sup> · Mohamed Marzouk<sup>2</sup>

Accepted: 28 February 2022 / Published online: 25 March 2022  
© The Author(s) 2022

## Abstract

The majority of water pipelines are subjected to serious deterioration and degradation challenges. This research examines the application of optimized neural network models for estimating the condition of water pipelines in Shaker Al-Bahery, Egypt. The proposed hybrid models are compared against the classical neural network, adaptive neuro-fuzzy inference system, and group method of data handling using four evaluation metrics. These metrics are; Fraction of Prediction within a Factor of Two (FACT2), Willmott's index of agreement (WI), Root Mean Squared Error (RMSE), and Mean Bias Error (MBE). The results show that the neural network trained using Particle Swarm Optimization (PSO) algorithm (FACT2 = 0.93, WI = 0.96, RMSE = 0.09, and MBE = 0.05) outperforms other machine learning models. Furthermore, three multi-objective swarm intelligence algorithms are applied to determine the near-optimum intervention strategies, namely PSO salp swarm optimization, and grey wolf optimization. The performances of the aforementioned algorithms are evaluated using Generalized Spread (GS), Spread ( $\Delta$ ), and Generational Distance (GD). The results yield that the PSO algorithm (GS = 0.54,  $\Delta$  = 0.82, and GD = 0.01) exhibits better results when compared to the other algorithms. The obtained near-optimum solutions are ranked using a new additive ratio assessment and grey relational analysis decision-making techniques. Finally, the overall ranking is obtained using a new approach based on the half-quadratic theory. This aggregated ranking obtains a consensus index and a trust level of 0.97.

**Keywords** Optimized neural network · Machine learning · Swarm intelligence · Ensemble ranking · Half-quadratic theory

## 1 Introduction

Water distribution networks are responsible for providing water to consumers (Sophia et al. 2020). Water pipelines are the primary components of water networks (Fontana and Morais 2016). More than 16% of water pipes have surpassed their useful lives, being subjected to serious aging and deterioration challenges (Folkman 2018).

Degradation of water systems leads to frequent leakage and failures, water supply discontinuance, impaired water quality, and damage to the surrounding infrastructure (Han et al. 2015; El-Abbasy et al. 2016; Zangenehmadar and Moselhi 2016; Aşchilean and Giurca 2018). For example, six billion gallons of treated water are lost every day due to pipe leakage (ASCE 2017). Besides, the water main break rates in the USA and Canada increased from 11 to 14 breaks/100 miles/year over the past six years (Folkman 2018).

The deteriorated water pipelines require enormous investment (Mohamed and Zayed 2013). The American Society of Civil Engineers infrastructure report card (ASCE 2017) rated the performance of water networks a fair grade of “D” (poor/at risk) on a scale of “A” (exceptional: fit for the future) to “F” (failing/critical: unfit for purpose). The Canadian Infrastructure Report Card (CIRC 2019) stated that approximately 30% of water infrastructure is in very good condition, 40% is in good condition,

---

✉ Nehal Elshaboury  
nehal\_ahmed\_2014@hotmail.com

Mohamed Marzouk  
mm\_marzouk@yahoo.com

<sup>1</sup> Housing and Building National Research Centre,  
Construction and Project Management Research Institute,  
Giza, Egypt

<sup>2</sup> Construction Engineering and Management, Structural  
Engineering Department, Faculty of Engineering, Cairo  
University, Giza, Egypt

and 25% is in fair, poor, or very poor condition. The Environmental Protection Agency (EPA 2018) reported that an investment of \$472.60 billion would be needed over the next 20 years to ensure the provision of safe drinking water. Out of \$472.60 billion, \$312.6 billion is needed to replace and maintain deteriorated water distribution and transmission pipelines.

The above discussion highlights the deterioration problem of water infrastructure assets. This dilemma could be solved by developing a deterioration model that forecasts the future condition of water pipelines. Furthermore, this model could be linked to a budget allocation model to prioritize the maintenance and replacement plans for water pipelines based on their condition and deterioration rates. The proposed model provides infrastructure asset managers and practitioners with an ensemble decision regarding the optimum time and type of the required intervention strategies. This leads to upgrading the asset performance, increasing the customer service level, reducing the operation and maintenance costs, and improving the municipality's reputation (Aikman 2015; Elshaboury et al. 2021a).

## 2 Literature review

Machine learning models have been used extensively for water systems modeling. For instance, Zangenehmadar and Moselhi (2016) predicted the residual life of water pipelines by applying the Feed-Forward Neural Network (FFNN) with the Levenberg–Marquardt algorithm. Several typologies of FFNN models (i.e., different number of hidden neurons) were tested and compared using the coefficient of determination ( $R^2$ ), Mean Absolute Error (MAE), Relative Absolute Error (RAE), Root-Relative Square Error (RRSE), and Mean Absolute Percentage Error (MAPE). The results showed the robustness and accuracy of neural network models in estimating the remaining useful life of water pipelines. Tavakoli (2018) developed a model that estimated the residual life of water pipelines using FFNN and Adaptive Neural Fuzzy Inference System (ANFIS). It was concluded that these models could be utilized in predicting the remaining useful life of water pipelines.

However, recent research studies showed that stand-alone machine learning models do not yield accurate results because of over-fitting, long training times, and premature convergence. Besides, their performances are significantly affected by their structure design and parameter selection (Zhou et al. 2019). That is why some studies have applied evolutionary algorithms for optimizing parameters in machine learning models. For example, Meirelles et al. (2017) applied the FFNN model to estimate the nodal pressure in a water network. The model was

integrated with a Particle Swarm Optimization (PSO) algorithm to minimize the difference between the simulated and forecasted pressure values. The proposed hybrid strategy increased the calibration accuracy when compared to the standard procedure. Yalçın et al. (2018) applied a hybrid ANFIS model for detecting water leakage locations in water distribution systems. The model comprised least-squares and backpropagation learning algorithms. The effectiveness of the proposed model was demonstrated by comparing its results against those of the most popular methods used in this field.

Several studies have been conducted to optimize maintenance and replacement for water infrastructure assets. Surco et al. (2018) developed an optimization model to rehabilitate and expand water distribution networks using PSO. The model accounted for the change in the pipe's internal roughness, water velocities, and nodal pressures using Epanet hydraulic simulator. The results showed the efficiency of the proposed model for water network optimization. Zhou (2018) optimized the rehabilitation of water pipelines using a modified Non-dominated Sorting Genetic Algorithm (NSGA-II). The different intervention actions for pipelines comprised no action, relining, or full replacement. The model aimed at minimizing life cycle cost and burst number and maximizing hydraulic reliability taking into consideration financial and hydraulic constraints. Elshaboury et al. (2020) optimized the rehabilitation of water networks using multi-objective GA and PSO. The decision variables incorporated no action, minor repair, major repair, and full replacement. The main objectives of the model were maximizing the network condition and minimizing the total costs of intervention actions. The results yielded a better performance of PSO in terms of the Ratio of Non-dominated Individuals (RNI), Generational Distance (GD), Spacing (S), Maximum Pareto Front Error (MPFE), and Spread ( $\Delta$ ).

Numerous efforts have been exerted to investigate the application of Multi-Criteria Decision-Making (MCDM) in rehabilitating water networks. El-Chanati et al. (2016) evaluated the performance index of water networks using four MCDM methods; Analytic Network Process (ANP), Fuzzy ANP (FANP), Analytic Hierarchy Process (AHP), and Fuzzy AHP (FAHP). The FANP method was found to be the most accurate method because it accounted for the uncertainties and interdependencies among the assessment factors. Tscheikner-Gratl et al. (2017) compared five MCDM techniques (i.e., AHP, Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE), Weighted Sum Model (WSM), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), and Elimination and Choice Expressing Reality (ELECTRE)) for prioritizing water systems rehabilitation. These techniques yielded different results and thus it was

recommended to apply several methods to improve the reliability of results. Elshaboury et al. (2020) employed two MCDM techniques to rank the near-optimum intervention solutions for water networks. These techniques are Multi-Objective Optimization on the basis of Ratio Analysis (MOORA) and TOPSIS. The results showed that there was a very strong relationship between the aforementioned techniques using the Spearman correlation coefficient.

The main objective of this research is developing a practical framework that prioritizes water distribution pipelines maintenance and rehabilitation strategies. To achieve this objective, the following sub-objectives are carried out:

- (1) Implementing a FFNN model trained using meta-heuristic algorithms to estimate the condition of water pipes.
- (2) Utilizing the forecasted condition to determine the near-optimum intervention actions using PSO, Salp Swarm Optimization (SSO), and Grey Wolf Optimization (GWO) algorithms.
- (3) Ranking the maintenance and rehabilitation strategies using a new Additive Ratio Assessment (ARAS) and Grey Relational Analysis (GRA) techniques.
- (4) Acquiring the aggregated ensemble ranking using an approach based on the half-quadratic theory.

### 3 Machine learning algorithms

In this research, five machine learning algorithms are applied to predict the condition of pipelines, which are the ANFIS, Group Method of Data Handling (GMDH), classical FFNN, FFNN-GA, and FFNN-PSO. Each of these algorithms is illustrated in detail as in the below sub-sections.

#### 3.1 Adaptive neuro-fuzzy inference system

ANFIS inherits the capabilities of neural networks and fuzzy logic to provide powerful non-linear modeling of the problem (Azad et al. 2019). The basic Sugeno ANFIS structure comprises five layers. The first layer provides membership grades of the crisp input nodes. The second layer involves multiplying the membership functions to obtain the output of fuzzy rules. The third layer normalizes the strength of all rules. The fourth layer computes the contribution of different rules towards the overall output. The fifth layer defuzzifies the fuzzy results of different rules into a crisp output (Tiwari et al. 2018). There exist three ANFIS methods to generate the basic fuzzy inference system namely grid partitioning, subtractive clustering, and Fuzzy C-Means (FCM) clustering (Azad et al. 2019). In

this research, the ANFIS-FCM method is used because it generates better performance compared to other methods.

#### 3.2 Group method of data handling

GMDH is a self-organized approach that was developed for solving complex nonlinear problems (Ivakhnenko 1971). It is characterized by automatically determining the number of layers and neurons in the hidden layers and optimum topology. It is possible to consider all different combinations of inputs. Then, using one of the available minimizing techniques, polynomial coefficients are calculated (with training data). The neurons with a higher external criterion value (for testing data) are retained, whereas those with a lower value are discarded. The network architecture and mathematical prediction function are determined when a stopping criterion is achieved. Otherwise, the process continues and the next layer is created (Azimi et al. 2018).

#### 3.3 Feed-forward neural networks

An Artificial Neural Network (ANN) is capable of modeling the nonlinear and complex behavior of water networks (Lawrence 1994). It is typically composed of a large number of neurons that are arranged in layer(s) and connected through weights and biases (Zou et al. 2009). ANN has two phases of learning and recalling (Sbarufatti et al. 2016). The learning phase trains the network to figure out a relationship between input(s) and output(s). The recalling phase predicts the output(s) from the input(s) based on the trained network. As for the advantages of the ANN, it uses the historical data to modify the network until the output values reach the target ones. On the other side, the training speed of ANN is slow when the network structure and design are not precise (Golnaraghi et al. 2019).

#### 3.4 Neural network model trained using metaheuristic algorithms

The neural network is applied in this research to estimate the future condition of water pipes given no intervention action is applied. The utilized backpropagation learning algorithm adjusts its weights and biases depending on the differences between anticipated and target values. However, the initial values of these parameters largely impact the network results (Devikanniga et al. 2019). Accordingly, the neural network can be trained to determine the optimum values of weights and biases. In this research, the GA and PSO algorithms are used to train the FFNN model for achieving better performance (Feng 2006). These algorithms are regarded as one of the most popular and efficient algorithms for training FFNN (Garg et al. 2014; Chiroma et al. 2017). More details about the GA and PSO

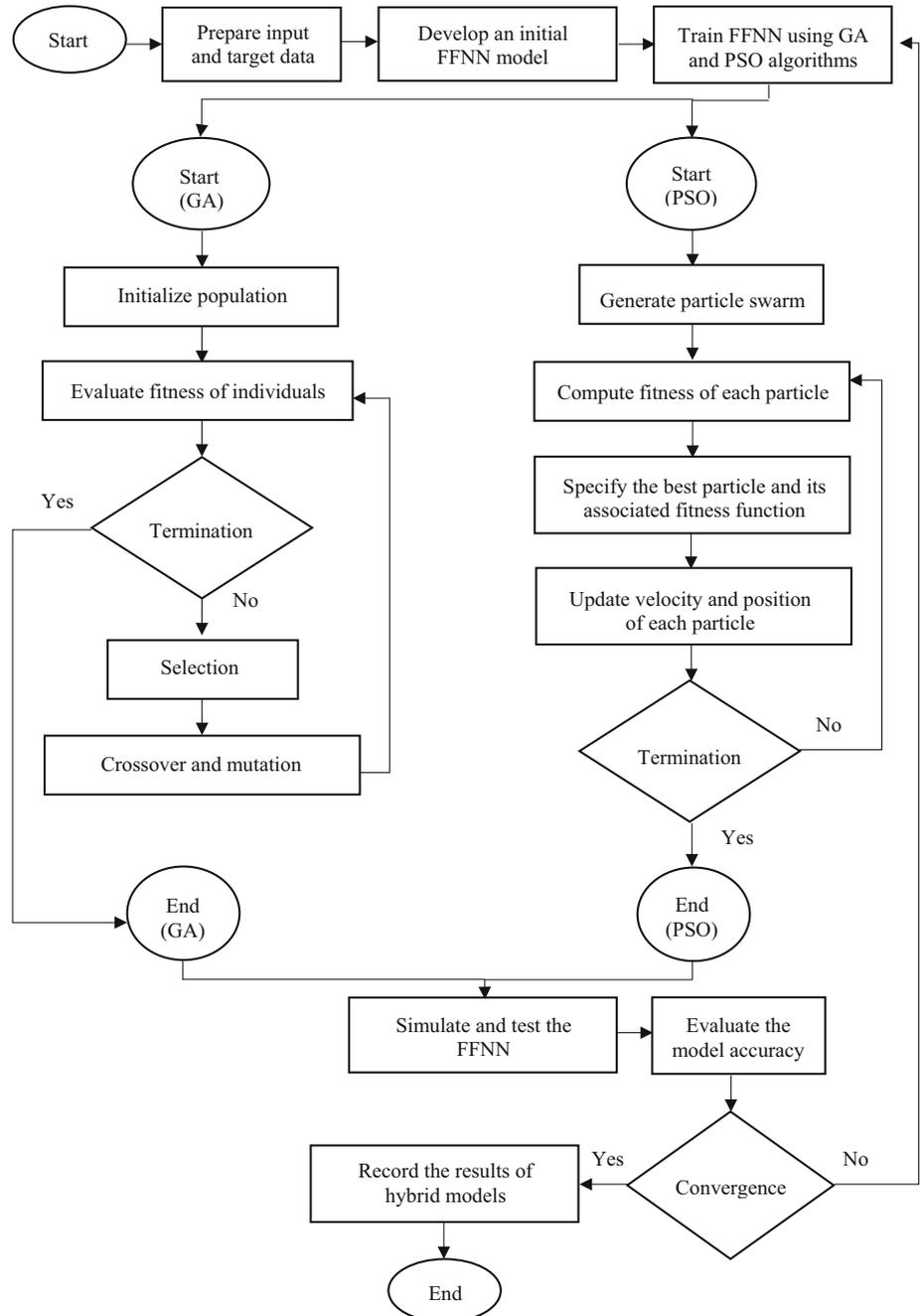
algorithms can be found in the literature (Holland 1975; Eberhart and Kennedy 1995). Combining neural networks with metaheuristic algorithms enhances their capabilities for solving real problems while preventing overfitting or local minima during training (Pater 2016). The flowchart of the optimized FFNN model procedure is illustrated in Fig. 1. The metaheuristic algorithms initialize the weights and calculate their fitness functions to start training the network. In this research, the network fitness is interpreted by estimating the error as per Eq. 1. The optimization

process stops when the global best solution (i.e., minimum error function) is achieved (Lazzús 2013).

$$MSE = \frac{\sum_{i=1}^{N_D} (y_i^{calc} - y_i^{exp})^2}{N_D} \tag{1}$$

where MSE refers to the mean squared error,  $N_D$  refers to the number of data points, and  $y_i^{calc}$  and  $y_i^{exp}$  refer to the calculated and expected values, respectively.

**Fig. 1** Flowchart of a hybrid FFNN model trained using metaheuristic algorithms



### 3.5 Performance metrics

Many metrics could be used to measure the performance of machine learning algorithms (Mishra 2018). In this research, Fraction of Prediction within a Factor of Two (FACT2), Index of Agreement (WI), Root Mean Square Error (RMSE), and Mean Bias Error (MBE) metrics are applied to evaluate the algorithms. A brief description of each metric is presented in the following sub-sections.

#### 3.5.1 Fraction of prediction within a factor of two

FACT2 examines the degree of closeness between the observed and modeled values as per Eq. 2. The closer this value is to one, the better this model is performing (Sayegh et al. 2014).

$$FACT2 = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{o_i - \bar{o}_i}{\sigma_o} \right) \left( \frac{p_i - \bar{p}_i}{\sigma_p} \right), 0.5 \leq \frac{o_i}{p_i} \leq 2 \tag{2}$$

where  $o_i$  represents the observed value,  $p_i$  represents the predicted value,  $\bar{o}_i$  represents the mean observed value,  $\bar{p}_i$  represents the mean predicted value,  $\sigma_o$  represents the standard deviation of the observed values, and  $\sigma_p$  represents the standard deviation of the predicted values, and  $n$  represents the number of observations.

#### 3.5.2 Willmott’s index of agreement

WI is calculated by multiplying the ratio of mean square error to potential error by the number of data points and deducting one, as seen in Eq. 3. It shall be mentioned that a higher WI value implies a good agreement between the predicted and target values and vice versa (Elshaboury et al. 2021b).

$$WI = 1 - \left[ \frac{\sum_{i=1}^n (o_i - p_i)^2}{\sum_{i=1}^n (|p_i - \bar{o}| + |o_i - \bar{o}|)^2} \right] \tag{3}$$

#### 3.5.3 Root mean squared error

RMSE calculates the distance/closeness between observed and predicted data points, as per Eq. 4. The lower RMSE value is associated with a higher prediction accuracy of the model (Elshaboury and Marzouk 2020).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - p_i)^2} \tag{4}$$

#### 3.5.4 Mean bias error

MBE measures the average bias in the predicted values as per Eq. 5. The lower value of this metric indicates a stronger forecasting accuracy of the model (Sharu and Ab Razak 2020).

$$MBE = \frac{1}{n} \sum_{i=1}^n |p_i - o_i| \tag{5}$$

## 4 Swarm intelligence algorithms

Swarm intelligence algorithms mimic the behavior of plants, insects, and animals as they strive to survive. These algorithms have attracted popularity in recent years because of their self-learning capabilities, self-organization, simplicity, flexibility, co-evolution, versatility, and adaptability to external variations (Chakraborty and Kar 2017; Lim and Leong 2018). In this research, the PSO, SSO, and GWO algorithms are utilized to determine the near-optimum intervention strategies. Descriptions of the SSO and GWO algorithms are provided in the following sub-sections.

### 4.1 Salp swarm optimization

SSO is inspired by the swarming behavior of salps (Mirjalili et al. 2017). Salps belong to the family of Salpidae that has a transparent body and tissues like jellyfishes (Henschke et al. 2016). They live in deep oceans and change positions by pumping water through their bodies. They are organized in swarms called salp chains. The salp chain comprises leaders and followers. The leaders are found at the front of the chain and the other salps are called followers. The target of the swarm is the sources of food (Ibrahim et al. 2018).

### 4.2 Grey wolf optimization

GWO is inspired by the hunting process of grey wolves (Panda and Das 2019). This unique algorithm follows a hierarchical pack hunting behavior. The alphas are authorized to decide the hunting time and resting place for the whole group. The betas advise the leaders in their decisions and maintain discipline for the group. The delta wolves follow the orders of alphas and betas and dominate omegas. The omegas follow the orders of all other dominant wolves (Mirjalili et al. 2014). The group hunting behavior of grey wolves includes three phases of tracking and chasing the prey, encircling and harassing the target, and attacking the prey (Jitkongchuen et al. 2016).

### 4.3 Performance metrics

Many metrics could be employed to evaluate the performance of evolutionary algorithms (Yu et al. 2018). In this research, three measures are investigated to compare the different swarm intelligence algorithms namely, Generalized Spread (GS), Spread ( $\Delta$ ), and Generational Distance (GD). It shall be noted that lower values of these metrics indicate a better performance of the algorithm. A brief description of each metric is presented in the following sub-sections.

#### 4.3.1 Generalized spread

The GS metric measures the distribution of the obtained near-optimum solutions using Eq. 6 (Zhou et al. 2006).

$$\Delta^*(A, P) = \frac{\sum_{m=1}^M d_m^P + \sum_{i=1}^{|A|} |d_i - \bar{d}|}{\sum_{m=1}^M d_m^e + |A|\bar{d}} \tag{6}$$

where  $d_i$  refers to the Euclidean distance between neighboring solutions in the non-dominated front,  $\bar{d}$  denotes the average of these distances, and  $d_m^P$  represents the distance between the extreme solutions of true Pareto front ( $P$ ) and approximate Pareto front ( $A$ ) with respect to the  $m$ th objective function.

#### 4.3.2 Spread

The delta indicator ( $\Delta$ ) evaluates the spread of the non-dominated solutions as per Eq. 7 (Deb et al. 2002).

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_f + d_l + (N - 1)\bar{d}} \tag{7}$$

where  $d_f$  and  $d_l$  are the Euclidean distances between the obtained non-dominated set's extreme and boundary solutions.

#### 4.3.3 Generational distance

The GD metric examines the diversity and convergence of the obtained solutions compared to the true Pareto front as per Eq. 8 (Veldhuizen 1999).

$$GD = \frac{(\sum_{i=1}^n d_i^m)^{1/m}}{n} \tag{8}$$

where  $d_i$  is the Euclidean distance between a non-dominated solution obtained by an algorithm and the closest Pareto front solution.

## 5 Weights of criteria

The weights of criteria reflect their relative significance from the decision maker's perspective such that larger weights indicate higher importance of criteria and vice versa. In this research, an Indifference Threshold-based Attribute Ratio Analysis (ITARA) method is applied to compute the weights of criteria. The computation methodology of this method is provided below (Hatefi 2019):

The indifference threshold value for each criterion is computed based on the difference between the mean and standard deviation of criteria (Mladineo et al. 2016). The normalized indifference threshold value is then determined using Eq. 9.

$$NIT_j = \frac{IT_j}{\sum_{i=1}^m a_{ij}} \tag{9}$$

where  $IT_j$  and  $NIT_j$  refer to the indifference threshold value and the normalized indifference threshold value for the  $j$ th criteria, respectively,  $m$  refers to the number of alternatives, and  $a_{ij}$  refers to the measure of performance of the  $i$ th alternative with respect to the  $j$ th attribute.

The normalized scores ( $\beta_{ij}$ ) are sorted in ascending order and the ordered distances between these scores ( $\gamma_{ij}$ ) are computed. The difference between  $\gamma_{ij}$  and  $NIT_j$  is computed using Eq. 10.

$$\delta_{ij} = \begin{cases} \gamma_{ij} - NIT_j & \text{for } \gamma_{ij} > NIT_j \\ 0 & \text{for } \gamma_{ij} \leq NIT_j \end{cases} \quad \forall i \in M, \forall j \in N \tag{10}$$

Finally, the weights of criteria are assigned using Eqs. 11 and 12, respectively.

$$V_j = \left( \sum_{i=1}^{m-1} \delta_{ij}^p \right)^{1/p}, \quad \forall j \in N \tag{11}$$

$$w_j = \frac{V_j}{\sum_{j=1}^n V_j} \tag{12}$$

## 6 Multi-criteria decision-making techniques

In this research, ARAS and GRA decision-making techniques are applied to rank the intervention strategies and develop the most optimal budget allocation plan. Each of these techniques is illustrated in the below sub-sections.

### 6.1 ARAS method

ARAS method compares the utility functions of each alternative to the best alternative. The application steps of

this method are proposed below (Zavadskas and Turskis 2010):

The normalized decision matrix for beneficial and non-beneficial attributes is equated using Eqs. 13 and 14, respectively.

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \tag{13}$$

$$r_{ij} = \frac{\frac{1}{x_{ij}}}{\sum_{i=1}^m \frac{1}{x_{ij}}} \tag{14}$$

where  $r_{ij}$  represents the normalized decision matrix and  $x_{ij}$  represents the measure of performance of the  $i$ th alternative with respect to the  $j$ th attribute.

The weighted normalized decision matrix is determined using Eq. 15.

$$Y_{ij} = r_{ij} \times w_j \tag{15}$$

where  $Y_{ij}$  represents the weighted normalized decision matrix and  $w_j$  represents the weight of each attribute.

The utility degree for each alternative is calculated using Eq. 16. It shall be noted that the best alternative is associated with the highest utility degree.

$$U_i = \frac{\sum_{j=1}^n Y_{ij}}{\max \cdot \sum_{j=1}^n Y_{ij}} \tag{16}$$

where  $U_i$  represents the utility degree for each alternative.

### 6.2 Grey relational analysis

GRA method computes the grey relational grade which describes the relationships among different alternatives. This method comprises four steps as described below (Kuo et al. 2008):

For beneficial and non-beneficial attributes, the normalized decision matrix is calculated using Eqs. 17 and 18, respectively.

$$y_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \tag{17}$$

$$y_{ij} = \frac{\max(x_{ij}) - x_{ij}}{\max(x_{ij}) - \min(x_{ij})} \tag{18}$$

where  $y_{ij}$  represents the normalized decision matrix.

The reference alternative ( $y_{0j}$ ) is defined based on its performance values. It is associated with performance values closest to or equal to one and zero in the case of beneficial and non-beneficial attributes, respectively.

The grey relational coefficient is determined between the reference alternative and all comparable alternatives using Eq. 19.

$$\gamma(y_{0j}, y_{ij}) = \frac{\Delta_{\min} + \xi \Delta_{\max}}{\Delta_{ij} + \xi \Delta_{\max}} \tag{19}$$

where  $\gamma(y_{0j}, y_{ij})$  represents the grey relational coefficient between  $y_{0j}$  and  $y_{ij}$ ,  $\Delta_{ij} = |y_{0j} - y_{ij}|$ ,  $\Delta_{\min}$  and  $\Delta_{\max}$  are the minimum and maximum values of  $\Delta_{ij}$ , respectively, and  $\xi$  is the distinguishing coefficient and is taken generally as 0.5.

The grey relational grade which reflects the level of correlation between the reference sequence and the comparability sequence is computed using Eq. 20. The best alternative is the one with the highest relational grade because it is most similar to the reference sequence.

$$r(y_0, y_i) = \sum_{j=1}^n w_j \times \gamma(y_{0j}, y_{ij}) \tag{20}$$

where  $r(y_0, y_i)$  represents the grey relational grade between  $y_{0j}$  and  $y_{ij}$ .

## 7 Aggregated ranking of alternatives

Decision-making techniques use different mechanisms and yield distinct rankings. Therefore, it is essential to provide an aggregated ranking to determine the optimal solution. In this research, a new approach based on the half-quadratic theory is adopted (Mohammadi and Rezaei 2020). The ensemble ranking of MCDM methods is computed using Eqs. 21–23, respectively.

$$\alpha_m = \delta(R^m - R_2^*) \tag{21}$$

where  $\alpha_m$  refers to the half-quadratic auxiliary variable,  $R^m$  refers to the ranking of the  $m$ th MCDM method,  $m$  refers to the number of MCDM methods, and  $R^*$  refers to the final aggregated ranking.

$$w_m = \alpha_m / \sum_j \alpha_j \tag{22}$$

where  $w_m$  refers to the weight of each MCDM method.

$$R^* = \sum_m w_m \times R^m \tag{23}$$

The consensus index which reflects the level of agreement among MCDM methods on the final ranking is computed using Eq. 24.

$$C(R^*) = \frac{1}{KM} \sum_{k=1}^K \sum_{m=1}^M \frac{\mathcal{N}_\sigma(R_k^* - R_k^m)}{\mathcal{N}_\sigma(0)} \tag{24}$$

where  $C(R^*)$  refers to the consensus index of the final ranking  $R^*$ ,  $K$  refers to the number of alternatives, and  $\mathcal{N}_\sigma$  refers to the probability density function of the Gaussian

distribution whose standard deviation is  $\sigma$  and mean is zero.

Finally, the trust level which indicates the level at which the ensemble ranking can be accredited is evaluated using Eq. 25.

$$T(R^*) = \frac{1}{K} \sum_{k=1}^K \sum_{m=1}^M w_m \times \left( \frac{\mathcal{N}_\sigma(R_k^* - R_k^m)}{\mathcal{N}_\sigma(0)} \right) \quad (25)$$

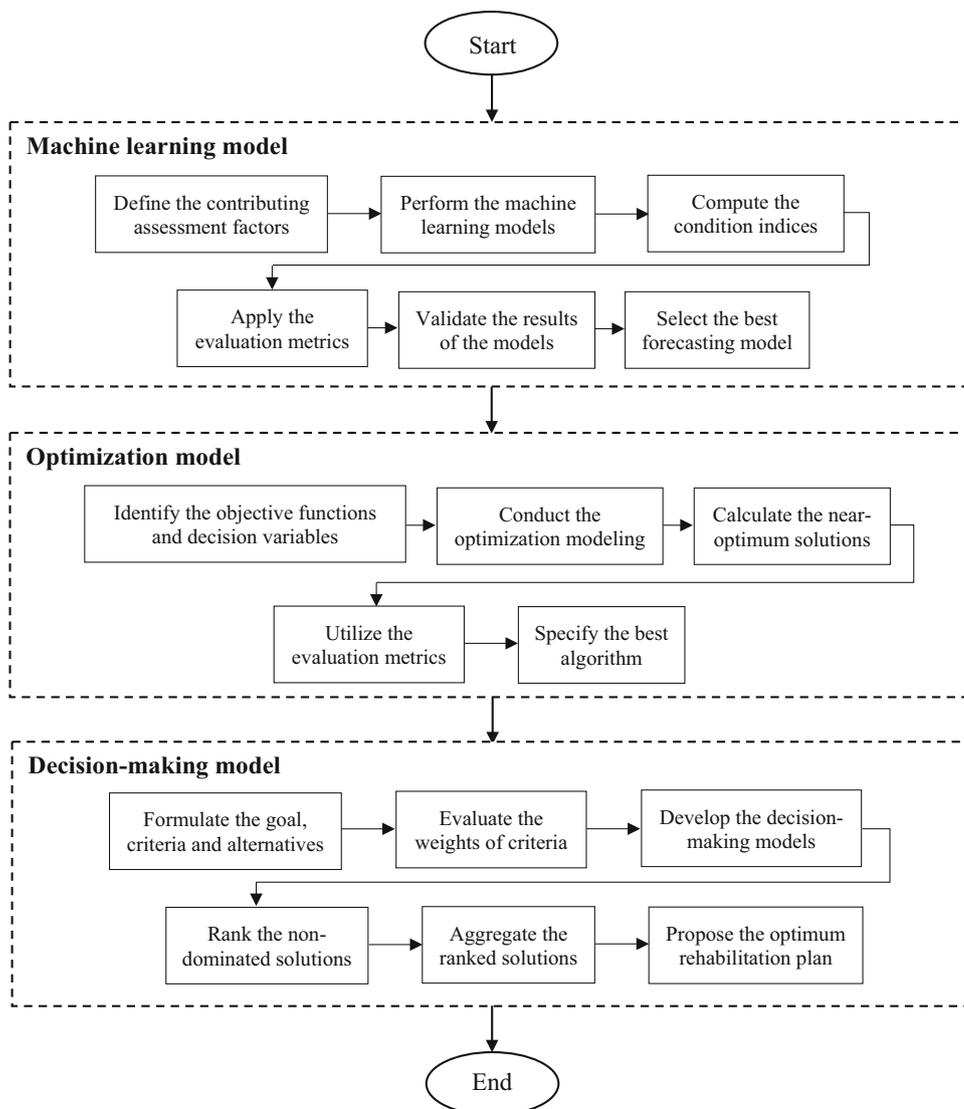
where  $T(R^*)$  represents the trust level of the ensemble ranking  $R^*$ .

### 8 Model development and implementation

The proposed flowchart to prioritize water pipeline rehabilitation is illustrated in Fig. 2. The framework is composed of three major components namely, machine

learning, optimization, and decision-making. The machine learning model involves: (a) predicting the condition indices of pipelines using several models, (b) comparing the results using evaluation metrics, and (c) verifying the models. The optimization model comprises: (a) formulating the optimization problem, (b) conducting the optimization modeling to calculate the near-optimum solutions, and (c) utilizing the evaluation metrics to specify the best algorithm. The decision-making model includes: (a) structuring the decision-making problem, (b) evaluating the weights of criteria, (c) developing the decision-making models to rank the non-dominated solutions, and (d) aggregating the ranked solutions and developing the budget allocation plan.

Fig. 2 Components of the proposed framework



### 8.1 Machine learning model

The machine learning models relate the pipe characteristics such as length, age, diameter, and wall thickness to its condition. After identifying the input and output variables, the next step comprises implementing ANFIS, GMDH, FFNN, FFNN-GA, and FFNN-PSO models. Approximately 70 and 30% of the data are used for training and testing purposes, respectively. For ANFIS, the number of clusters is set as 15 while the number of epochs and iterations is adjusted to 200. The optimum values of initial step size, step size decrease rate, and step size increase rate are selected as 0.01, 0.9, and 1.1, respectively. For GMDH, FFNN, and optimized FFNN, the number of hidden neurons is assumed to be 10 to provide a fair comparison of the models. Moreover, the Levenberg–Marquardt algorithm is employed to implement neural networks because of its strong performance in solving nonlinear problems (Zangenehmadar and Moselhi 2016). The code is written in MATLAB R2019a to build the machine learning models.

The five models are developed using the training and testing data to compare their predictive performances using FACT-2, WI, RMSE, and MBE evaluation metrics. These calculations are implemented in this research using Microsoft Excel. The outcomes of the prediction models are verified using the Taylor diagram which is developed using Mathematica v12.0.

### 8.2 Optimization model

The optimization model incorporates two objective functions which are; maximizing the condition of water pipelines (Eq. 26) and minimizing the costs of intervention strategies (Eq. 27).

$$CIP = \sum_{j=1}^N \sum_{x=1}^Z CIP_{ij} \tag{26}$$

$$CP = \frac{\sum_{j=1}^N \sum_{x=1}^Z CP_{ij}}{(1+r)^t} \tag{27}$$

where  $CIP_{ij}$  represents the improved condition index of the  $j$ th pipeline after applying an  $x$ th intervention strategy,  $Z$  represents the number of applied strategies,  $N$  represents the number of pipelines,  $CP_{ij}$  represents the cost of the  $x$ th intervention strategy applied to the  $j$ th pipeline,  $r$  refers to the discount rate, and  $t$  refers to the study period. In this research, the discount rate is taken as 7% and the study period is assumed to be three years.

The decision variables comprise the possible intervention actions for pipelines namely full replacement, major repair, minor repair, or no action. Minor and major repairs are applied to restore sections of pipelines using

compression coupling and telescopic coupling, respectively. The future condition of water pipelines before adopting an intervention action is forecasted using a neural network model coupled with a PSO algorithm. Meanwhile, the improved condition of pipelines is estimated based on the chosen intervention strategy, as shown in Table 1 (El-Masoudi 2016).

The costs associated with the no-action, minor repair, major repair, and full replacement are assumed to be 0, 20, 50, and 100% of the replacement costs, respectively (El-Masoudi 2016). The replacement costs per unit length for different sizes of unplasticized polyvinyl chloride (uPVC) pipes are depicted in Table 2.

After formulating the optimization problem, the PSO, SSO, and GWO algorithms are applied to determine the near-optimum intervention strategies. For PSO, personal and global learning coefficients are both equal to 2. Besides, the mutation rate, inertia weight, and inertia weight dumping rate are assumed to be 0.1, 1, and 0.99, respectively (Elshaboury et al. 2020). For SSO, the values of the random parameters lie in the interval of [0, 1] (Mirjalili et al. 2017). For PSO and GWO, the number of grids per dimension, grid inflation parameter, leader selection parameter, and deletion selection parameter are assumed to be 10, 0.1, 2, and 2, respectively (Lai et al. 2019). To provide a fair comparison of the optimization algorithms, it is assumed that the population size, maximum repository size, and maximum number of iterations are set to be 200, 100, and 100, respectively.

The implementation steps of these algorithms are summarized as follows: (a) define the candidate solutions (i.e., pipelines) in the current population, (b) encode the solutions whose length is 519 (i.e.,  $173 \times 3$ ) and can hold four variables (i.e., intervention actions), (c) initialize the population and select the parameters, (d) compute the fitness functions (i.e., maximizing condition and minimizing cost) of each candidate solution, (e) forming a new population from solutions with higher fitness values, (f) updating the parameters of the algorithm till satisfying the objective functions, and (g) terminating the algorithm when

**Table 1** Improved condition after applying intervention actions

Original condition	Intervention action		
	Minor repair	Major repair	Full replacement
Critical	Moderate	Good	Excellent
Poor	Moderate	Good	Excellent
Moderate	Good	Good	Excellent
Good	–	–	Excellent
Excellent	–	–	–

**Table 2** Cost data for uPVC pipes

Diameter (millimeters)	Cost per unit length (\$/meter)
100	7.44
150	11.06
200	15.58
300	21.68
400	31.02

achieving the stopping criterion. The outcomes of these algorithms are evaluated using the GS, delta, and GD metrics. The code is written in MATLAB to perform the multi-objective algorithms and evaluation metrics.

### 8.3 Decision-making model

The weights of condition and cost criteria are computed using the ITARA method. The ARAS and GRA techniques are then employed to rank the non-dominated solutions obtained from the optimization model. These calculations are implemented in this research using Microsoft Excel. The different rankings obtained from the decision-making techniques are aggregated using a half-quadratic-based method. Finally, the consensus index and the trust level of this ensemble ranking are evaluated. The aggregated rankings are computed in the MATLAB environment.

## 9 Case study

A water distribution network in Shaker Al-Bahery, Egypt has been selected as the application case study (see Fig. 3). The collected data include factors such as length, material, age, diameter, depth, and wall thickness. The network consists of 173 pipelines with a total length of 10.3 km. All the network pipes are made of uPVC and installed at a depth of 1.3 m. They were installed in this residential area at the age of 12 years. Their diameter range between 100 and 400 mm, and the corresponding wall thickness is extracted from the manufacturer's technical specifications.

## 10 Results and discussion

Four neural network models with one hidden layer and different number of neurons are developed. In the hidden layers of the FFNN1, FFNN2, FFNN3, and FFNN4 models, there are 5, 10, 15, and 20 neurons, respectively. As summarized in Table 3, the performance of the FFNN models is assessed using four different evaluation metrics,

which are FACT-2, WI, RMSE, and MBE. In general, the high values of FACT-2 and WI indicate better performance for any model. On the other side, the low values of RMSE and MBE reflect a good prediction accuracy of the model. In comparison to other models, FFNN2 has the highest FACT-2 and WI values of 0.87 and 0.93, respectively. Besides, this model is associated with the minimum RMSE (i.e., 0.12) and MBE (i.e., 0.06) values. Therefore, the number of hidden neurons is set to 10 based on the results of the evaluation metrics.

A summary of the observed and predicted condition indices using the developed machine learning models is illustrated in Fig. 4. The mean observed condition indices is 5.97 while that of the models ranges between 3.65 and 5.98. Meanwhile, the standard deviation of the observed indices is 0.24 while that of the prediction models lies within a range of 0.22 and 0.29.

The forecasting results of the machine learning models are evaluated as depicted in Table 4. The FFNN model yields a FACT-2 value of 0.87 compared to 0.76 in ANFIS and 0.73 in GMDH. As for the WI metric, FFNN has a value of 0.93, higher than the reported value of 0.86 by ANFIS and substantially higher than 0.17 reported by GMDH. Finally, the neural network model exhibits better performance than the other two models in terms of the RMSE (0.12) and MBE metrics (0.06). This emphasizes that there is a substantial improvement in the values of metrics for the FFNN model compared to ANFIS and GMDH models.

In an attempt to improve the performance of the conventional neural network, it is trained using GA and PSO algorithms. Most of the FFNN-PSO predictions (i.e., 93%) lie within a factor of two of the observed values, whereas FACT2 values for the rest of the models range from 87 to 91%. Based on the WI, the FFNN-PSO model with a WI value of 0.96 outperforms the other models, exhibiting WI values of less than 0.95. Meanwhile, the proposed model is characterized by the lowest values of RMSE (0.09) and MBE (0.05), outperforming the other models with values of more than 0.10 and 0.06, respectively. It can be concluded that incorporating the PSO algorithm into the classical FFNN model enhances the model robustness for predicting the pipe condition.

As shown in Fig. 5, the Taylor diagram illustrates that the correlation coefficient values of the prediction models lie in the range of 0.73–0.93. The GMDH model shows the lowest correlation coefficient value (i.e., 0.73), while the FFNN-PSO model has the highest correlation coefficient value (i.e., 0.93). The standard deviation of the FFNN-PSO model is 0.22, whereas the standard deviations of the other models range between 0.22 and 0.29. Finally, in terms of the root mean square error, the FFNN-PSO shows the lowest RMSE value (i.e., 0.09). Therefore, it can be

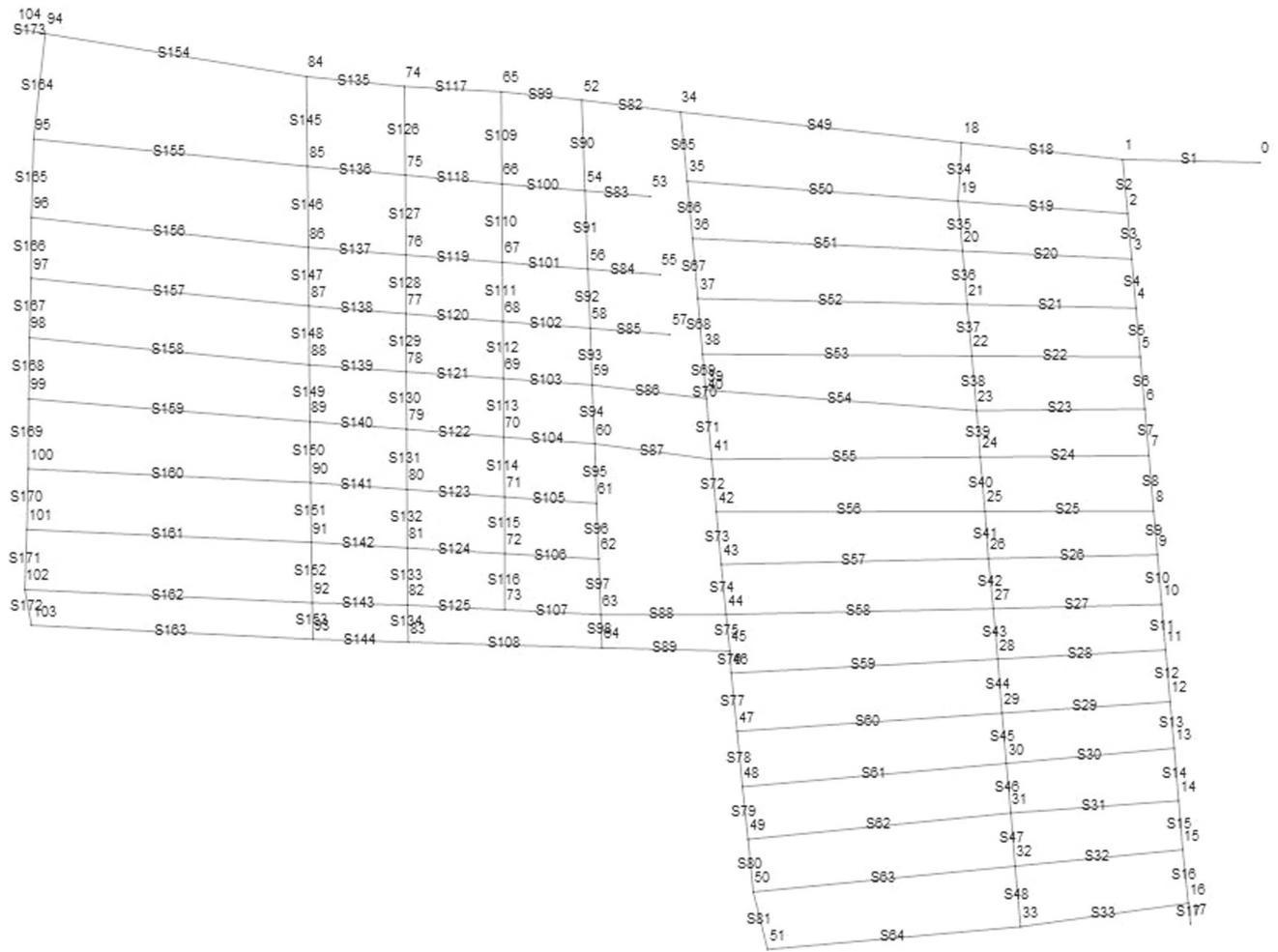


Fig. 3 Layout of Shaker Al-Bahery water distribution network

Table 3 Comparison of performances of the neural network models

Performance measure	Neural network model			
	FFNN1	FFNN2	FFNN3	FFNN4
FACT-2	0.82	0.87	0.81	0.84
WI	0.88	0.93	0.87	0.89
RMSE	0.18	0.12	0.18	0.17
MBE	0.13	0.06	0.13	0.12

concluded that the FFNN-PSO model provides more consistent forecasts in terms of the correlation coefficient, standard deviation, and root mean square error.

Swarm intelligence algorithms yield a set of different Pareto-optimal solutions. Therefore, the obtained solutions are evaluated to assess the optimization algorithms. As depicted in Table 5, the PSO algorithm is associated with the lowest GS, Δ, and GD values. This emphasizes that the

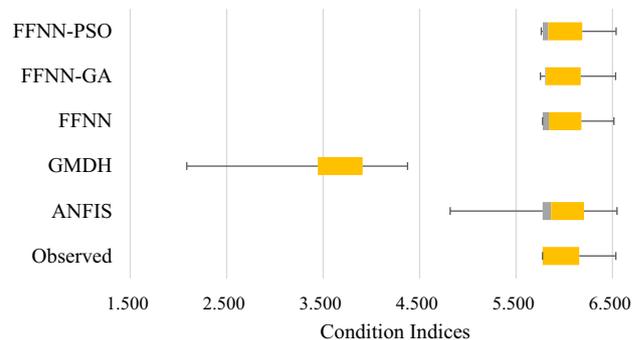


Fig. 4 Comparison of the actual and predicted condition indices

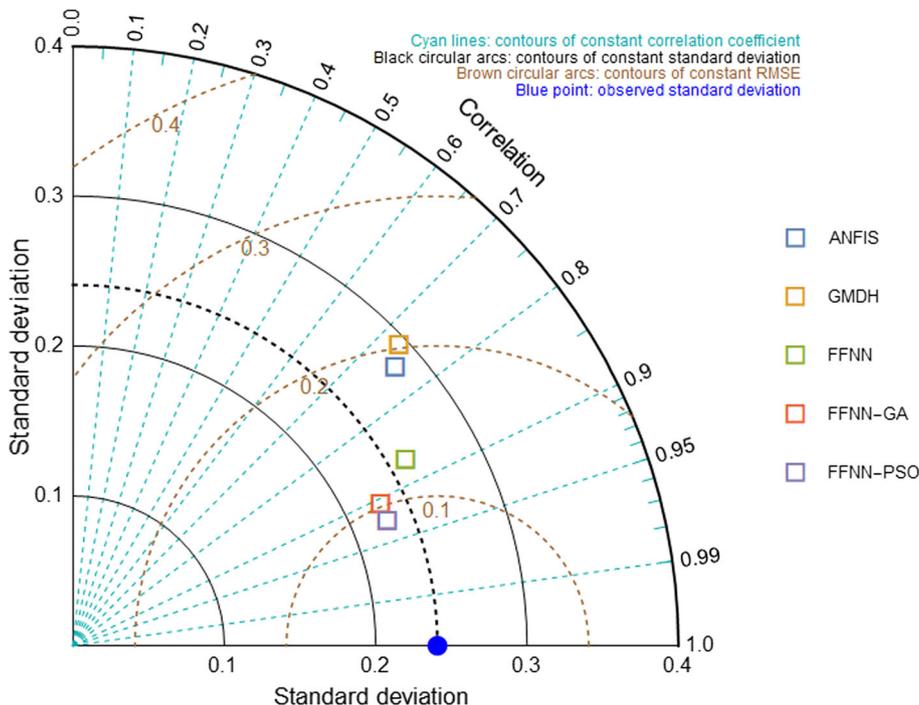
PSO algorithm is more suitable for optimizing the rehabilitation of water distribution pipelines.

The obtained non-dominated intervention solutions are ranked using ARAS and GRA techniques to rank the optimal solutions. The criteria of MCDM techniques are the improved condition of pipelines and costs of intervention strategies. As shown in Table 6, the improved

**Table 4** Comparison of performances of the machine learning models

Performance measure	Machine learning model				
	ANFIS	GMDH	FFNN	FFNN-GA	FFNN-PSO
FACT-2	0.76	0.73	0.87	0.91	0.93
WI	0.86	0.17	0.93	0.95	0.96
RMSE	0.19	2.33	0.12	0.10	0.09
MBE	0.07	2.32	0.06	0.06	0.05

**Fig. 5** Verification of different prediction models using Taylor diagram



**Table 5** Evaluation of performances of the swarm intelligence algorithms

Performance measure	Swarm intelligence algorithm		
	PSO	SSO	GWO
GS	0.54	0.72	0.69
$\Delta$	0.82	0.97	0.90
GD	0.01	0.24	0.02

**Table 6** Weights of the criteria using the ITARA method

Index	Cost of interventions	Improved condition
$v_j$	0.007	0.003
$w_j$	0.730	0.270

condition and total cost represent 27% and 73%, respectively, using the ITARA method.

As shown in Table 7, each MCDM technique follows a certain methodology and thus yields different rankings for most of the solutions. However, both techniques assign the solution [1,559,222 7.20] as the best-ranked solution. The rankings obtained from the decision-making methods are aggregated using a half-quadratic-based method. The ensemble ranking obtains a consensus index and a trust level of 0.97. This means that the rankings have a strong degree of consensus.

### 11 Conclusion

Water distribution network pipelines are approaching the end of their service life. Therefore, it is essential to predict their condition and deterioration rates to perform the necessary intervention plans at the right time and prevent disastrous failures. It is imperative to establish a relationship between the condition and the influencing parameters

**Table 7** Aggregated rankings of optimum intervention strategies

Alternative	ARAS ranking	GRA ranking	Aggregated ranking
[1,559,222–7.20]	1	1	1
[1,723,825–7.48]	14	18	16
[1,754,630–7.52]	17	19	18
[1,848,487–7.60]	21	17	19

(i.e., length, age, diameter, and wall thickness). This research forecasted the condition of water pipelines using an Adaptive Neuro-Fuzzy Inference System (ANFIS), Group Method of Data Handling (GMDH), Feed-Forward Neural Network (FFNN), and a hybrid FFNN model trained using Genetic Algorithms (GA) and Particle Swarm Optimization (PSO). It was concluded that evolving FFNN with PSO algorithm (FACT2 = 0.93, WI = 0.96, RMSE = 0.09, and MBE = 0.05) enhanced the performance of modeling water pipelines condition. It is advisable to explore the degree of condition improvement of the different proposed intervention solutions. Therefore, the PSO, Salp Swarm Optimization (SSO), and Grey Wolf Optimization (GWO) were employed to obtain the non-dominated solutions. The results yielded that the PSO algorithm (GS = 0.54, Δ = 0.82, and GD = 0.01) exhibited better results when compared to other algorithms. The Pareto-front solutions of the optimization models were assessed using a new Additive Ratio Assessment (ARAS) and Grey Relational Analysis (GRA) decision-making techniques. These techniques assigned the solution [1,559,222–7.20] as the best-ranked solution. Since there was a difference in some of the rankings obtained from both techniques, these rankings were aggregated using a new approach based on the half-quadratic theory. The ensemble ranking obtained a consensus index and a trust level of 0.97. This implied that the ensemble ranking could be accredited due to the high degree of consensus among the ranks. The developed framework was demonstrated using a water distribution network in Shaker Al-Bahery, Egypt. This research was expected to assist the water municipality in allocating the available budget efficiently and effectively as well as scheduling the needed intervention strategies. This research could be extended in future by considering and comparing the performance of known neuro-fuzzy based methodologies (e.g., fuzzy relational neural network) for estimating the condition of water pipelines.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

**Data availability** All data and material generated or used during the study appear in the submitted article.

**Declarations**

**Conflict of interest** The authors declare that they have no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**References**

Aikman DI (2015) Water services asset management: An international perspective. *Infrastruct Asset Manag* 1(2):34–41

ASCE (American Society of Civil Engineers) (2017) Report card for America’s infrastructure. <http://www.infrastructurereportcard.org>. Accessed 23 April 2020

Aşchilean I, Giurca I (2018) Choosing a water distribution pipe rehabilitation solution using the analytical network process method. *Water* 10(4):484–507

Azad A, Manoochehri M, Kashi H, Farzin S, Karami H, Nourani V, Shiri J (2019) Comparative evaluation of intelligent algorithms to improve adaptive neuro-fuzzy inference system performance in precipitation modelling. *J Hydrol* 571:214–224

Azimi H, Bonakdari H, Ebtehaj I, Gharabaghi B, Khoshbin F (2018) Evolutionary design of generalized group method of data handling-type neural network for estimating the hydraulic jump roller length. *Acta Mech* 229(3):1197–1214

Chakraborty A, Kar AK (2017) Swarm intelligence: a review of algorithms. In: Patnaik S, Yang X-S, Nakamatsu K (eds) *Nature-Inspired Computing and Optimization: Theory and Applications*. Springer, Cham, pp 475–494. [https://doi.org/10.1007/978-3-319-50920-4\\_19](https://doi.org/10.1007/978-3-319-50920-4_19)

Chiroma H, Abdulkareem S, Abubakar A, Herawan T (2017) Neural networks optimization through genetic algorithm searches: a review. *Appl Math Inf Sci* 11(6):1543–1564

CIRC (Canadian Infrastructure Report Card) (2019) Monitoring the state of Canada’s core public infrastructure: The Canadian infrastructure report card 2019. <http://canadianinfrastructure.ca/downloads/canadian-infrastructure-report-card-2019.pdf>. Accessed 23 April 2020

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE T Evolut Comput* 6(2):182–197

- Devikanniga D, Vetrivel K, Badrinath N (2019) Review of meta-heuristic optimization based artificial neural networks and its applications. *J Phys Conf Ser* 1362(1):012074
- Eberhart RC, Kennedy J (1995) A new optimizer using particle swarm theory. In: Proceedings of the 6th international symposium on micro machine and human science, pp 39–43. Nagoya, Japan: IEEE
- El-Abbasy M, El-Chanati H, Mosleh F, Senouci A, Zayed T, Al-Derham H (2016) Integrated performance assessment model for water distribution networks. *Struct Infrastruct E* 12(11):1505–1524
- El-Chanati HE, El-Abbasy MS, Mosleh F, Senouci A, Abouhamad M, Gkountis I, Zayed T, Al-Derham H (2016) Multi-criteria decision making models for water pipelines. *J Perform Constr Fac* 30(4):04015090
- El-Masoudi I (2016) Condition assessment and optimal repair strategies of water networks using genetic algorithms. Dissertation, Mansoura University
- Elshaboury N, Attia T, Marzouk M (2020) Application of evolutionary optimization algorithms for rehabilitation of water distribution networks. *J Constr Eng M* 146(7):04020069
- Elshaboury N, Attia T, Marzouk M (2021a) Reliability assessment of water distribution networks using minimum cut set analysis. *J Infrastruct Syst* 27(1):04020048
- Elshaboury N, Marzouk M (2020) Comparing machine learning models for predicting water pipelines condition. In: Proceedings of the 2nd novel intelligent and leading emerging sciences conference (NILES). Giza, Egypt: IEEE
- Elshaboury N, Elshourbagy M, Al-Sakkaf A, Abdelkader E (2021b) Rainfall forecasting in arid regions using an ensemble of artificial neural networks. In: Proceedings of the 1st international conference on fundamental, applied sciences and technology (ICoFAST 2021b). Al Mukalla, Yemen
- EPA (Environmental Protection Agency) (2018) Drinking water infrastructure needs survey and assessment - sixth report to congress. [https://www.epa.gov/sites/production/files/2018-10/documents/corrected\\_sixth\\_drinking\\_water\\_infrastructure\\_needs\\_survey\\_and\\_assessment.pdf](https://www.epa.gov/sites/production/files/2018-10/documents/corrected_sixth_drinking_water_infrastructure_needs_survey_and_assessment.pdf). Accessed 23 April 2020
- Feng HM (2006) Self-generation RBFNs using evolutionary PSO learning. *Neurocomputing* 70(1–3):241–251
- Folkman S (2018) Water main break rates in the USA and Canada: A comprehensive study. [https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1173&context=mae\\_facpub](https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1173&context=mae_facpub). Accessed 23 April 2020
- Fontana M, Morais D (2016) Decision model to control water losses in distribution networks. *Prod* 26(4):688–697
- Garg S, Patra K, Pal SK (2014) Particle swarm optimization of a neural network model in a machining process. *Sadhana* 39(3):533–548
- Golnaraghi S, Zangenehmadar Z, Moselhi O, Alkass S (2019) Application of artificial neural network(s) in predicting formwork labour productivity. *Adv Civ Eng*. <https://doi.org/10.1155/2019/5972620>
- Han S, Chae MJ, Hwang H, Choung YK (2015) Evaluation of customer-driven level of service for water infrastructure asset management. *J Manage Eng* 31(4):04014067
- Hatefi MA (2019) Indifference threshold-based attribute ratio analysis: a method for assigning the weights to the attributes in multiple attribute decision making. *Appl Soft Comput* 74:643–651
- Henschke N, Everett JD, Richardson AJ, Suthers IM (2016) Rethinking the role of salps in the ocean. *Trends Ecol Evol* 31(9):720–733
- Holland J (1975) *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor
- Ibrahim RA, Ewees AA, Oliva D, Elaziz MA, Lu S (2018) Improved salp swarm algorithm based on particle swarm optimization for feature selection. *J Amb Intel Hum Comp* 10(8):3155–3169
- Ivakhnenko AG (1971) Polynomial theory of complex systems. *IEEE Trans Syst Man Cybern SMC-1*:364–378
- Jitkongchuen D, Phaidang P, Pongtawevirat P (2016) Grey wolf optimization algorithm with invasion-based migration operation. In: Proceedings of the 15th Int. Conf. Comput. Inf. Sci. (ICIS). Okayama: IEEE
- Kuo Y, Yang T, Huang GW (2008) The use of grey relational analysis in solving multiple attribute decision-making problems. *Comput Ind Eng* 55(1):80–93
- Lai X, Li C, Zhang N, Zhou J (2019) A multi-objective artificial sheep algorithm. *Neural Comput Appl* 31:4049–4083
- Lawrence J (1994) *Introduction to neural networks design, theory and applications*. California Scientific Software Press, USA
- Lazzús JA (2013) Neural network-particle swarm modeling to predict thermal properties. *Math Comput Model* 57(9–10):2408–2418
- Lim SM, Leong KY (2018) A brief survey on intelligent swarm-based algorithms for solving optimization problems. In: Del Ser J, Osaba E (eds) *Nature-inspired methods for stochastic, robust and dynamic optimization*. InTech. <https://doi.org/10.5772/intechopen.76979>
- Meirelles G, Manzi D, Brentan B, Goulart T, Luvizotto E (2017) Calibration model for water distribution network using pressures estimated by artificial neural networks. *Water Resour Manag* 31(13):4339–4351
- Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61
- Mirjalili S, Gandomi AH, Mirjalili SZ, Saremi S, Faris H, Mirjalili SM (2017) Salp swarm algorithm: a bio-inspired optimizer for engineering design problems. *Adv Eng Softw* 114:163–191
- Mishra A (2018) Metrics to evaluate your machine learning algorithm. <https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234>. Accessed 16 January 2020
- Mladineo M, Jajac N, Rogulj K (2016) A simplified approach to the PROMETHEE method for priority setting in management of mine action project. *Croat Oper Res Rev* 7(2):249–268
- Mohamed E, Zayed T (2013) Modeling fund allocation to water main rehabilitation projects. *J Perform Constr Facil* 27(5):646–655
- Mohammadi M, Rezaei J (2020) Ensemble ranking: aggregation of rankings produced by different multi-criteria decision-making methods. *Omega* 96:102254
- Panda M, Das B (2019) Grey wolf optimizer and its applications: a survey. In: Proceedings of the 3rd International Conference on Microelectronics, Computing and Communication Systems. Singapore: Springer
- Pater Ł (2016) Application of artificial neural networks and genetic algorithms for crude fractional distillation process modeling. Nicolaus Copernicus University Toruń, Poland, Faculty of Mathematics and Computer Science
- Sayegh AS, Munir S, Habeebullah TM (2014) Comparing the performance of statistical models for predicting PM10 concentrations. *Aerosol Air Qual Res* 14(3):653–665
- Sbarufatti C, Corbetta M, Manes A, Giglio M (2016) Sequential monte-carlo sampling based on a committee of artificial neural networks for posterior state estimation and residual lifetime prediction. *Int J Fatigue* 83(1):10–23
- Sharu EH, Ab Razak MS (2020) Hydraulic performance and modelling of pressurized drip irrigation system. *Water* 12(8):2295
- Sophia SG, Sharmila VC, Suchitra S, Muthu TS, Pavithra B (2020) Water management using genetic algorithm-based machine learning. *Soft Comput* 24(22):17153–65

- Surco D, Vecchi T, Ravagnani M (2018) Rehabilitation of water distribution networks using particle swarm optimization. *Desalin Water Treat* 106:312–329
- Tavakoli R (2018) Remaining useful life prediction of water pipes using artificial neural network and adaptive neuro-fuzzy inference system models. Dissertation, Texas University
- Tiwari S, Babbar R, Kaur G (2018) Performance evaluation of two ANFIS models for predicting water quality index of river Satluj (India). *Adv Civ Eng* 2018:1–10
- Tscheikner-Gratl F, Egger P, Rauch W, Kleidorfer M (2017) Comparison of multi-criteria decision support methods for integrated rehabilitation prioritization. *Water* 9(2):68
- Veldhuizen DV (1999) Multi-objective evolutionary algorithms: classifications, analyses, and new innovations. PhD Thesis, Air Force Institute of Technology, Ohio, USA
- Yalçın BC, Demir C, Gökçe M, Koyun A (2018) Water leakage detection for complex pipe systems using hybrid learning algorithm based on ANFIS method. *J Comput Inf Sci Eng* 18(4):041004
- Yu X, Lu Y, Yu X (2018) Evaluating multiobjective evolutionary algorithms using MCDM methods. *Math Probl Eng* 2018:1–13
- Zangenehmadar Z, Moselhi O (2016) Assessment of remaining useful life of pipelines using different artificial neural networks models. *J Perform Constr Facil* 30(5):04016032
- Zavadskas EK, Turskis Z (2010) A new additive ratio assessment (ARAS) method in multi criteria decision-making. *Technol Econ Dev Econ* 16(2):159–172
- Zhou Y (2018) Deterioration and optimal rehabilitation modelling for urban water distribution systems. CRC Press, London
- Zhou A, Jin Y, Zhang Q, Sendhoff B, Tsang E (2006) (2006) Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion. *Evolutionary Computation*. IEEE Press, pp 892–899
- Zhou X, Zhang M, Xu Z, Cai C, Huang Y, Zheng Y (2019) Shallow and deep neural network training by water wave optimization. *Swarm Evol Comput* 50:100561
- Zou J, Han Y, So SS (2009) Overview of artificial neural networks. Humana Press, USA

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.