dx.doi.org/10.22093/wwj.2024.454419.3414

Original Paper

Water and Wastewater, Vol. 34, No. 6, pp: 55-73

Predicting Effluent Quality Parameters Using Ensemble Models, Artificial Neural Networks and Naked Mole-Rat Algorithm

Elham Ghanbari Adivi[1*], Ali Raeisi[2]

1. Assoc. Prof., Dept. of Water Engineering, Faculty of Agriculture, Shahrekord University, Iran
(Corresponding Author) ghanbariadivi@sku.ac.ir
2. Assist. Prof., Dept. of Water Engineering, Faculty of Agriculture, Shahrekord University, Iran

(Received Dec. 13, 2023 Accepted Feb. 21, 2024)

To cite this article:
Ghanbari Adivi, E. and Raeisi, A., 2024. Predicting effluent quality parameters using ensemble models, artificial neural networks and naked mole-rat algorithm. *Water and Wastewater*, 34(6), 55-73.
https://doi.org/10.22093/wwj.2024.454419.3414.

Abstract

A reliable simulation of wastewater effluent parameters is important for reducing the operational costs of a wastewater treatment plant. In this study, optimized artificial neural networks and inclusive multiple models are used to predict the effluent biochemical oxygen demand, chemical oxygen demand and total suspended solids of a WWTPs in Shahrekord basin, Iran. The influent quality parameters (COD_{inf}, BOD_{inf}, TSS_{inf} and PH_{inf}) are used as inputs to the models. The naked mole-rat algorithm is used to tune ANN parameters. This investigation compares the capabilities of ANN-NMRA with those of ANN-firefly, ANN-sine cosine algorithm, ANN-genetic algorithm and ANN models. The output of hybrid and standalone models is incorporated into an ANN model as an IMM model. Several individual models are used in an IMM model for predicting outputs. Hence, an IMM model increases the accuracy of individual models. In this study, ANN models are modified by using a goodness factor to reduce computational time. This paper presents a new preprocessing method for selecting the best input combination and analyzes the uncertainty of model and input parameters. IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA and ANN models achieve MAEs of 0.789, 0.998, 1.19, 1.26, 1.34 and 1.40 mg/L for predicting BOD_{eff} in the testing stage, respectively. The IMM model has the highest accuracy for predicting COD_{eff} and TSS_{eff}. A good factor reduces the computational time of ANN models by removing redundant hidden neurons. The uncertainty analysis results show that model parameters provide higher uncertainty than input parameters.

Keywords: Effluent Quality Parameters, Artificial Neural Network Models, Uncertainty, Optimization Algorithms, Wastewater Treatment Plant.

1. Introduction

It is important to treat industrial wastewater because it contains various contaminants, which lead to environmental pollution and water shortages (Nadiri et al., 2018). Wastewater consists of water, organic matter, and minerals. (Asghari et al., 2022). Water treatment plants[1] are influenced by several factors, including the nature of the environment, technology, and the economy. A WWTP's main goal is to reduce pollutants by reducing operations costs and improving effluent quality (Elmaadawy et al., 2021). Modeling a WWTP's performance and controlling its quality parameters may be necessary for managing a WWTP. Because the wastewater treatment process involves many simultaneous nonlinear mechanisms, so predicting the wastewater effluent parameters is challenging.

The water treatment process is complex and nonlinear, so predicting effluent quality parameters is challenging (Nadiri et al., 2018). Using direct measurements to measure effluent quality parameters can be costly and complex. Other tools for predicting effluent quality parameters are physical-based models. It is possible to apply physical-based models to predict effluent quality, but these models may take a long time and be expensive. In recent years, researchers have widely applied machine learning models[2] to predict effluent quality parameters. In Table 1, the applications of different MLMs are presented for predicting effluent quality parameters. Even though MLMs successfully predicted effluent quality parameters, several research gaps exist:

1) Different input parameters affect effluent quality parameters, so finding the best input scenario for predicting effluent quality parameters remains a challenge.

2) Also, previous papers did not attempt to employ robust ensemble models for predicting effluent quality parameters. An ensemble model is an effective strategy for reducing computational errors because outputs of an ensemble model are obtained using the advantages of multiple individual models (Sharafati et al., 2020). By combining multiple individual models, ensemble models are effective strategies for reducing computational errors (Sharafati et al., 2020).

3) Another shortcoming of previous studies is that they did not provide an effective strategy for selecting model inputs.

4) During the modeling process, the computational time of MLMs is an important consideration. Previous studies found no effective solution for reducing the computational time of MLMs. In previous studies, uncertainty sources were not considered in the modeling process.

Thus, this paper tries to address research gaps by utilizing advanced models. This paper aims to predict effluent quality parameters in a case study.

This study suggests a novel ensemble model, known as the inclusive multiple model[3]. The first step in this research is to develop a new Artificial Neural Network[4] model, called the ANN-naked mole-rat algorithm[5], to predict effluent quality parameters. NMRA is applied to adjust ANN parameters, including weight and bias parameters. The potential of NRMA for training ANN parameters is benchmarked against the sine cosine algorithm[6] and the firefly algorithms[7] and genetic algorithms[8]. We insert the outputs of ANN-FFA, ANN-GA, ANN-SCA, ANN-NRMA, and ANN models into the IMM model after acquiring their outputs. An IMM model is a group of multiple ANN models. NRMA has several advantages in this study.

Salgotra and Singh (2019), introduced NMRA for solving optimization problems. The NRMA had high capabilities, such as high precision, fast convergence, and potential for global integration.

A variety of mathematical functions and engineering problems were used to evaluate NMRA (Salgotra and Singh, 2019); NRMA outperformed other optimization algorithms such as GA, PSO, BOA and FFA. The main contributions of this paper are:

1) In this study, the effects of different uncertainty resources on the outputs are assessed using generalized likelihood uncertainty estimation[9].

2) A new preprocessing method, namely gamma test-NRMA, is used to determine the best input scenario in this study. Several fields have applied GT for choosing the best input scenarios, including natural gas forecasting (Salehnia et al., 2013), streamflow prediction (Noori et al., 2011), daily discharge prediction (Rahbar et al., 2022), suspended sediment load prediction (Panahi et al., 2021), etc. The GT method is a nonlinear method for selecting the best input data. Choosing the best input combination will be time-consuming when a modeler confronts many input data. A new version of GT is applied in this study to determine the best input combination using NRMA. When NRMA is integrated with GT, the complexity of the modeling process is reduced.

3) This study uses a goodness factor to modify the structure of ANNs. By reducing the number of redundant neurons, this coefficient can improve the performance of ANNs.

[1] Wastewater Treatment Plants (WWTPs)
[2] Machine Learning Models (MLMs)
[3] Inclusive Multiple Model (IMM)
[4] Artificial Neural Network (ANN)
[5] Naked Mole-Rat Algorithm (NMRA)
[6] Sine Cosine Algorithm (SCA)
[7] Firefly Algorithms (FFA)
[8] Genetic Algorithms (GA)
[9] Generalized Likelihood Uncertainty Estimation (GLUE)

Table 1. Literature review for predicting effluent quality parameters

Description	Analysis
Guo et al. applied an ANN and support vector machine[1] for predicting total effluent nitrogen (T-N) (Guo et al., 2015).	ANN model outperformed the SVM model for predicting total effluent nitrogen (T-N).
Heddam et al. applied generalized regression neural network[2] and multiple linear regression[3] for predicting effluent COD and BOD_5 (Heddam et al., 2016).	For predicting COD and BOD_5, GRNN performed better than MLR. One of the challenges of their study was tuning the GRNN parameters.
Nadiri et al. applied a supervised committee fuzzy logic[4] for predicting effluent water quality. Using a combination of individual fuzzy logic models, SCFL forecasted water quality (Nadiri et al., 2018).	SCFL had a mean absolute percentage error[5] between 10 and 13% for BOD, COD and TSS prediction.
Lotfi et al. coupled the stochastic linear model[6] with extreme learning machine techniques for predicting effluent BOD, COD and Total Suspended Solids[7] (Lotfi et al., 2019).	This hybrid model had a high ability to predict effluent quality parameters in a WWTP.
Sharafati et al. applied Gradient Boost Regression[8] and a random forest model for predicting effluent TSS, BOD_5 and COD. They applied a gamma test for predicting the best input variables (Sharafati et al., 2020).	The RMSEs of GDR for predicting TSS, BOD_5 and COD were 30.3, 4.6 and 9.6 mg/L, respectively. One of the difficulties of their study was selecting the best input variables.
Niu et al. integrated deep belief networks[9] with GA for predicting effluent quality parameters (Niu et al., 2020).	It was found that DBN-GA was highly accurate for predicting effluent quality parameters. In this study, one of the challenges was preparing the DBN-GA structure.
Nourani et al. used ANN, SVM and ANFIS for predicting BOD and COD of Tabriz. Inputs were influent quality parameters, including BOD_5, COD_5 and pH. The outputs were effluent BOD_5 and COD_5 (Nourani et al., 2021).	The results of the SVM model were better than those of other models. One of the challenges of their study was adjusting ANN, SVM and ANFIS parameters.
Meng et al. used an adaptive task-oriented radial basis function (ATO-RBF) for predicting effluent BOD and DO_5 (Meng et al., 2021).	They reported that ATO-RBF using a second error algorithm was more accurate than RBF.
Elmaadawy et al. combined Random Vector Functional Link[10] Network and foraging optimizer (MRFO[11]) for predicting TSS and BOD_5. In this study, RVFL-MRFO was benchmarked against ANN models (Elmaadawy et al., 2021).	R^2 values of MRFO-RVFL model for predicting BOD_5 and TSS were 0.92 and 0.91, respectively. MRFO-RVFL had better accuracy for predicting BOD_5 and TSS. Finding the best input combination was one of the main issues in their study.
Hejabi et al. applied SVM and ANN for predicting effluent quality of Tabriz WWTP in Iran. They used quality parameters of BOD, COD and TSS to predict effluent parameters of BOD, COD and TSS (Hejabi et al., 2021).	For predicting effluent quality, ANN models were superior to SVM models. For obtaining accurate results, the study's main challenge was to prepare the structure of an ANN and a SVM model.

[1] Support Vector Machine (SWM)
[2] Generalized Regression Neural Network (GRNN)
[3] Multiple Linear Regression (MLR)
[4] Supervised Committee Fuzzy Logic (SCFL)
[5] Mean Absolute Percentage Error (MAPE)
[6] Stochastic Linear Model (LSM)
[7] Total Suspended Solids (TSS)
[8] Gradient Boost Regression (GBR)
[9] Deep Belief Networks (DBN)
[10] Random Vector Functional Link (RVFL)
[11] Manta Rays Foraging Optimization (MRFO)

The theory and structure of algorithms are explained in Section 2. The data are described in Section 3. Discussion and analysis of results are presented in Section 4. In Section 5, the conclusion is presented.

2. Materials and methods

2.1. Mathematical model of ANNs

ANNs mimic the human brain and are made up of processors, mathematical functions, weights, biases, and computational layers. There are three types of computation layers in ANNs (Ehteram et al., 2021a). The input layer is the first layer in an ANN model. Here, input data is received and transformed into middle layers. Weigh connections connect the first layer to the middle layers. The middle layers contain mathematical functions called activation functions. The outputs of middle layers are computed by introducing weighted inputs and bias into activation functions. The outputs of the middle layers are sent to the last layer (Ehteram et al., 2021b). The final layer analyzes received data and provides final outputs based on its activation function. Fig. 1 shows the mathematical model of ANN. Neurons are the smallest units in the ANN model. If the weight connections of neurons are not large, then those neurons are considered inactive and have low contributions to final outputs (Watada, 2006). Thus, removing inactive neurons can reduce computational costs. This study uses the goodness factor to remove the redundant neurons (the most useless neurons). For ANN models, it is important to set proper values for ANN parameters. Optimization algorithms handle the process. Optimization algorithms use objective function and initial population of solutions to find the optimal value of ANN parameters, including the weight and bias.

2.2. Structure of NMRA

NMRA mimics the life of naked mole rats in swarms. Each swarm has a queen (Salgotra and Singh, 2019). Each swarm contains two groups of males. In the first group, males mate with queens (Salgotra and Singh, 2019). The males in the second group carry out different tasks, such as maintenance, defense, or provisioning. Workers can be divided into two groups based on their objective function: simple workers and breeders. Workers with the best performance are called breeders, while others are called simple workers (Salgotra and Singh, 2019). Initially, naked mole-rats are initialized as follows

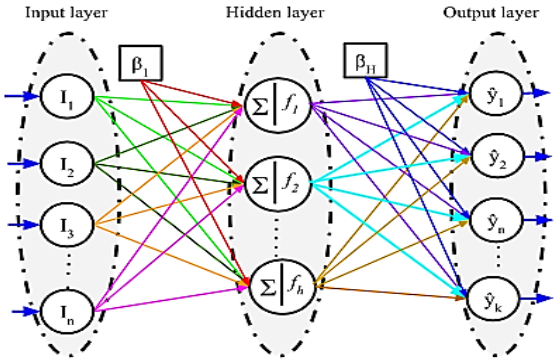
$$NM_{i,j} = NM_{i,j} + u(0,1) \times (NM_{max,j} - NM_{min,j}) \quad (1)$$

Where
the location of i^{th} rat in j^{th} dimension, the maximum value of decision variable in j^{th} dimension, the minimum value of decision variable in j^{th} dimension, and random value. The objective function is computed to determine the best solution, workers, and breeders in the next level. The queen is the optimal solution with the best objective function value. The simple workers try to enhance their efficiency (Salgotra and Singh, 2019). Equation 2 is used to update the workers. Thus, provided workers are retained if their quality is better than previous ones (Salgotra and Singh, 2019). Equation 2 lets workers improve their quality to become breeders and eventually mate with a queen. A simple worker may replace a breeder if it shows high performance (a better objective function)

$$wo_i^{t+1} = wo_i^t + \lambda(wo_j^t - wo_k^t) \quad (2)$$

Where
new location of workers in t+1 iteration, location of the worker in t iteration, and the mating frequency are defined. In the next step, breeders update themselves based on Equation 3 to be chosen for mating

$$br_i^{t+1} = (1-\lambda)b_i^t + \lambda(d - b_i^t) \quad (3)$$



(I: input, f: activation function, y: output, β : bias)

Fig. 1. Mathematical model of ANN

Where,
new i^{th} breeder in t+1 iteration, the best breeder, and i^{th} breeder in t^{th} iteration are defined. In the next phase, workers and breeders are evaluated to choose the best solutions.

2.3. Structure of CSA

Mirjalili introduced SCA to solve complex problems. It is inspired by sine and cosine functions. The best solution can guide other agents in the search space (Mirjalili, 2018). SCA updates the solutions in each iteration using sine and cosine functions. SCA has several benefits, including ease of computation, flexibility, and high efficiency for finding global solutions. SCA's solutions are updated using the following equation

$$SO_i^{t+1} = \begin{bmatrix} SO_i^t + ra_1 \times \sin(ra_2) * \left| ra_3 P_i^t - SO_i^t \right| \leftarrow ra_4 < 0.5 \\ O_i^t + ra_1 \times \sin(ra_2) * \left| ra_3 P_i^t - SO_i^t \right| \leftarrow ra_4 \geq 0.5 \end{bmatrix} \quad (4)$$

Where
SO_i^t is i^{th} solution in iteration t, SO_i^{t+1} is i^{th} solution in iteration t+1, P_i^t is the location of the destination point, ra_1, ra_2, ra_3, and ra_4 are random parameters. ra_1 is obtained for each iteration based on the following equation

$$ra_1 = x - t\frac{x}{T} \quad (5)$$

Where,
x is constant value, T is maximum number of iterations[1], and t is the number of current iteration. In SCA, the initial population of agents is first provided. The objective function is computed to determine the best solution (Mirjalili, 2018). Based on Equation 2, the solutions update themselves. Optimization proceeds until a stop condition is reached (Mirjalili, 2018).

2.4. Structure of FFA for solving the optimization problem

FFA is one of the most popular optimization methods. FFA was inspired by the life of fireflies. A firefly emits light (Ghorbani et al., 2018). The fireflies with higher light are considered more attractive (Ghorbani et al., 2018). Higher brilliancy fireflies have better objective function values (Shaibani et al., 2021). A firefly's position represents a decision variable in search space. Fireflies have a brightness equivalent to the value of the objective function. First, firefly positions are initialized (Shaibani et al., 2021). The objective function is used to identify fireflies with higher brightness Fireflies update their locations based on the following equation (Ghorbani et al., 2018)

$$fi_i^{t+1} = f_i^t + \beta_0 e^{-\gamma r^2}(fi_j^t - fi_i^t) + \psi\varepsilon \quad (6)$$

Where
fi_i^{t+1} is location of $i+1^{th}$ firefly in t^{th} iteration, f_i^t isposition of i^{th} firefly in t^{th} iteration, β_0 is the attractiveness, ε and ψ are random number, and r is the distance between two fireflies. The optimization process continues until the stop condition is met.

2.5. Structure of GA for solving optimization problems

For nonlinear and complex problems, GA is an effective optimization method. In GA, chromosomes play a major role in solving optimization problems (Mirjalili, 2018). Initializing a population of chromosomes is the first step of the algorithm. Three operators are used to update the initial chromosomes. One of the operators is named selection. By computing the objective function for each chromosome, the chromosomes with the best quality are chosen to produce offspring in the next generation. The crossover operator is the second operator. By combining the genetic data of two parents, new offspring can be produced. A mutation operator is the third operator. Mutations increase the genetic diversity of GA by modifying some of its genes. An effective way of escaping from the local optimum is to use the mutation operator.

2.6. Optimized ANN models

The choice of optimization algorithms is determined by considering parameters such as speed of convergence and accuracy. For this study, optimization algorithms were considered that exhibit both high accuracy and fast convergence.

This study uses robust optimization algorithms to adjust ANN parameters since finding optimum ANN parameters is an important issue. ANN parameters (weight and bias) are considered unknown parameters. These parameters are defined as the decision variables of the optimization algorithm. In the first level, ANN models are trained using the initial values of ANN parameters. The root mean square error is used to evaluate the performance of ANN models in the training stage. Next, the termination criterion is assessed. The ANN models run in the testing stage if the termination criteria are met; otherwise, optimization algorithms should be used to determine the best ANN parameters. A matrix of the initial population for each algorithm should be defined to link optimization algorithms and ANN models. The matrix includes the initial weights and

[1] Maximum Number of Iterations (MNI)

biases. Encoding of ANN parameters is required for linking ANN models with optimization algorithms. ANN parameters are represented by the locations of rats, fireflies, chromosomes, and agents of the SCA. ANN models using training data are run to evaluate the quality of the algorithm population, including the value of ANN parameters. As discussed in the previous section, each algorithm has different operators to update solutions. By updating solutions, the algorithms achieve new values of ANN parameters. The optimization cycle continues until the convergence criterion is met. At the end of the optimization process, the ANN parameters' values are decoded and inserted into the structure of the ANN model for testing use.

2.7. Structure of an Inclusive multiple model

Optimized ANN models can produce accurate results, but they are not necessarily superior to all soft computing models. Although MLMs have different advantages, they can also have disadvantages (Liang et al., 2021). It may be possible to enhance the efficiency of an individual ANN model by creating an ensemble model using the output of individual ANN models. There are different types of ensemble models (Liang et al., 2021). Bayesian models are one of the most important ensemble models, but they are complex models. One way to establish ensemble models is to use IMMs. This study uses hybrid and standalone ANN models for predicting output variables. A simple ANN is used as an ensemble model for integrating individual models (Liang et al., 2021). An IMM is constructed based on the outputs of both hybrid and standalone ANN models. Ensemble systems not only increase accuracy but they also provide synergy among multiple models. In the first level, each of the individual models is used to provide output.

2.8. Uncertainty analysis of models

Different sources of uncertainty exist in the modeling process, such as model and input parameters. Hence, it is essential to determine the uncertainty of outputs based on uncertainty resources. Various methods can be used to estimate uncertainty. A robust method for computing uncertainty is GLUE. In different fields such as river flow (Ragab et al., 2020), hydrological simulation (Kan et al., 2020), flood estimation (Cu Thi et al., 2018), urban flood modeling (Liu et al., 2020), sediment load prediction (Panahi et al., 2021), etc., GLUE is widely applied for uncertainty analysis. Model parameters and inputs are considered uncertainty sources in this study. Therefore, two scenarios are used to compute the uncertainty of inputs and model parameters separately. We estimate uncertainty as follows:

In GLUE, the first level is to determine the prior distribution[1] of parameters. The previous studies used a normal distribution for parameters (Panahi et al., 2021). The distribution of input parameters follows the normal distribution, but the distribution of model parameters does not. Due to the absence of physical notions for model parameters, using normal distributions could result in computation errors. Based on the variability of parameters during the calibration level, Srivastava et al. presented the characteristics of the distribution of ANN parameters (Srivastava et al., 2007). For this study, 4000 calibrated ANN models were used to obtain the PD of ANN parameters. This number is sufficient because the obtained type of probability distributions and their parameter values did not change between the calibrated 1000 ANN models and the calibrated 4000 ANN models.

Monte Carlo and PD are employed to generate a large number of samples of input parameters and model parameters.

In this step, likelihood is calculated. For computing likelihood, different functions may be used. Nash Sutcliff efficiency[2] is commonly used as the likelihood function (Panahi et al., 2021; Ragab, 2020). The parameters based on the likelihood calculation are divided into two categories. Parameters with likelihood values below a threshold are considered non-behavioral. They are ignored when modeling. The rest are categorized as behavioral parameters.

The posterior distribution is computed based on the likelihood and PD

$$p(k_i|eqp) = \frac{p(k_i) * p(eqp|k_i)}{\sum_{i=1}^{N} p(eqp|k_i)} \quad (7)$$

Where

Eqp is effluent equality parameters $p(k_i|eqp)$ is the posterior distribution, $p(k_i)$ is PD, and $p(eqp \mid k_i)$ is posterior distribution.

The mean and variance of data are computed based on the following equations

$$\varsigma = \sum_{i=1}^{N} p(\kappa_i \mid eqp).\kappa_i \quad (8)$$

$$\sigma^2 = \sum_{i=1}^{N} p(\kappa_i \mid eqp).(\kappa_i - \varsigma)^2 \quad (9)$$

Where

ς : average, σ^2 : variance, N: number of samples.

[1] Prior Distribution (PD)

[2] Nash Sutcliff Efficiency (NSE)

Table 2. Statistical details of input and output data

Parameter	Average	Maximum	Minimum	Standard deviation
Input				
BOD_{inf} (mg/L)	214.62	432.231	115.321	41.23
TSS_{in} (mg/L)	210.65	290.12	165.12	207
PH_{inf} (mg/L)	7.41	7.9	6.8	0.18
COD_{inf} (mg/L)	407.12	591.23	222.3	71.36
Output				
BOD_{eff} (mg/L)	22.7	54.22	4.00	10.39
TSS_{eff} (mg/L)	22.7	45	6.00	12.23
COD_{eff} (mg/L)	26.1	60.00	14.00	5.23

3. Case study

This study aims to predict effluent equality parameters of WWTPs in the Shahrekord basin, one of the most important plains in Iran. The area of the basin is 1,211 km^2. 330 MCM of the plain's water resources is harvested annually. Agriculture has the highest water demand in this basin. Recently, the plain has been experiencing successive droughts. A major challenge for the basin is the depletion of the groundwater resource. In the basin, the WWTP plays an important role. It can treat the water for use in agriculture. Moreover, the treated water will recharge the groundwater aquifers. An inlet channel is used for transporting wastewater to different parts of the Shahrekord WWTP. Shahrekord WWTP acts based on diffuser aerated activated sludge with a capacity of 765 m^3/day. The average annual discharge of sewage is 1.5 m^3/s. Fig. 2a shows the location of the basin. Fig. 2b, 2c and 2d show data points for BOD, COD, TSS. Daily data were collected from 2016 to 2019.

Hejabi et al. and Nourani et al. stated that influent parameters of BOD_{inf}, COD_{inf}, TSS_{inf}, and PH_{inf} had a high correlation with effluent quality parameters of BOD_{eff}, COD_{eff}, TSS_{eff}, PH_{eff} (Hejabi et al., 2021; Nourani et al., 2021). In this study, BOD_{inf}, COD_{inf}, TSS_{inf}, and PH_{inf} are used to predict the outputs of COD_{eff}, TSS_{eff}, BOD_{eff}. Table 2 shows statistical details of input and output data.

The concentration of different input data was measured by the sensors. Afterwards, the quality of the data was checked using various methods.

Regarding Table 2, there are four input variables for predicting each output variable (BOD_{inf}, COD_{inf}, TSS_{inf}, and PH_{inf}). The total number of input combinations for predicting output variables is 2^4-1. Thus, it is essential to use an effective test to select the best input combination. The Gamma test is used to choose the best input combination.

The selection of inputs was informed by a comprehensive review of the literature. In addition, a sensitivity analysis[1] was conducted to identify the inputs that provide the most significant information.

- **Gamma test (GT) for the selection of inputs**

GT is a nonlinear analysis tool for selecting the best input combination. GT provides a smooth relationship between inputs and outputs (Panahi et al., 2021). Consider this data sample

$$\{in_1,...,in_m,out_i\}=\left[(in_i,out_i)\,|\,1\le i\le M\right] \quad (10)$$

Where

in_i: i^{th} input, out_i: i^{th} output, and M: number of patterns. The GT algorithm builds the k^{th} nearest neighbor lists of input data ($in_{N[i,k]}$) based on the given sample data. The delta function is calculated as follows

$$\Delta_M(k)=\frac{1}{M}\sum_{i=1}^{M}\left|in_{N[i,k]}-in_i\right|^2,1\le k\le p \quad (11)$$

Where

p is number of nearest neighbors, and $\Delta_M(k)$ is delta function. The Gamma function of outputs is calculated as follows (Panahi et al., 2021)

$$\gamma_M(k)=\frac{1}{2M}\sum_{i=1}^{M}\left|out_{N[i,k]}-out_i\right|^2,1\le k\le p \quad (12)$$

Where

$out_{N[I,k]}$ is k^{th} nearest neighbor lists of output data. Finally, Γ as the gamma index of GT is computed as follows (Panahi et al., 2021)

$$out=A\delta+\Gamma \quad (13)$$

Where

A is gradient, Γ is intercept, and out. The best input combination has the lowest Γ. GT is useful for selecting input data, but it may be time-consuming to compute Γ

[1] Sensitivity Analysis (SA)

dx.doi.org/10.22093/wwj.2024.454419.3414

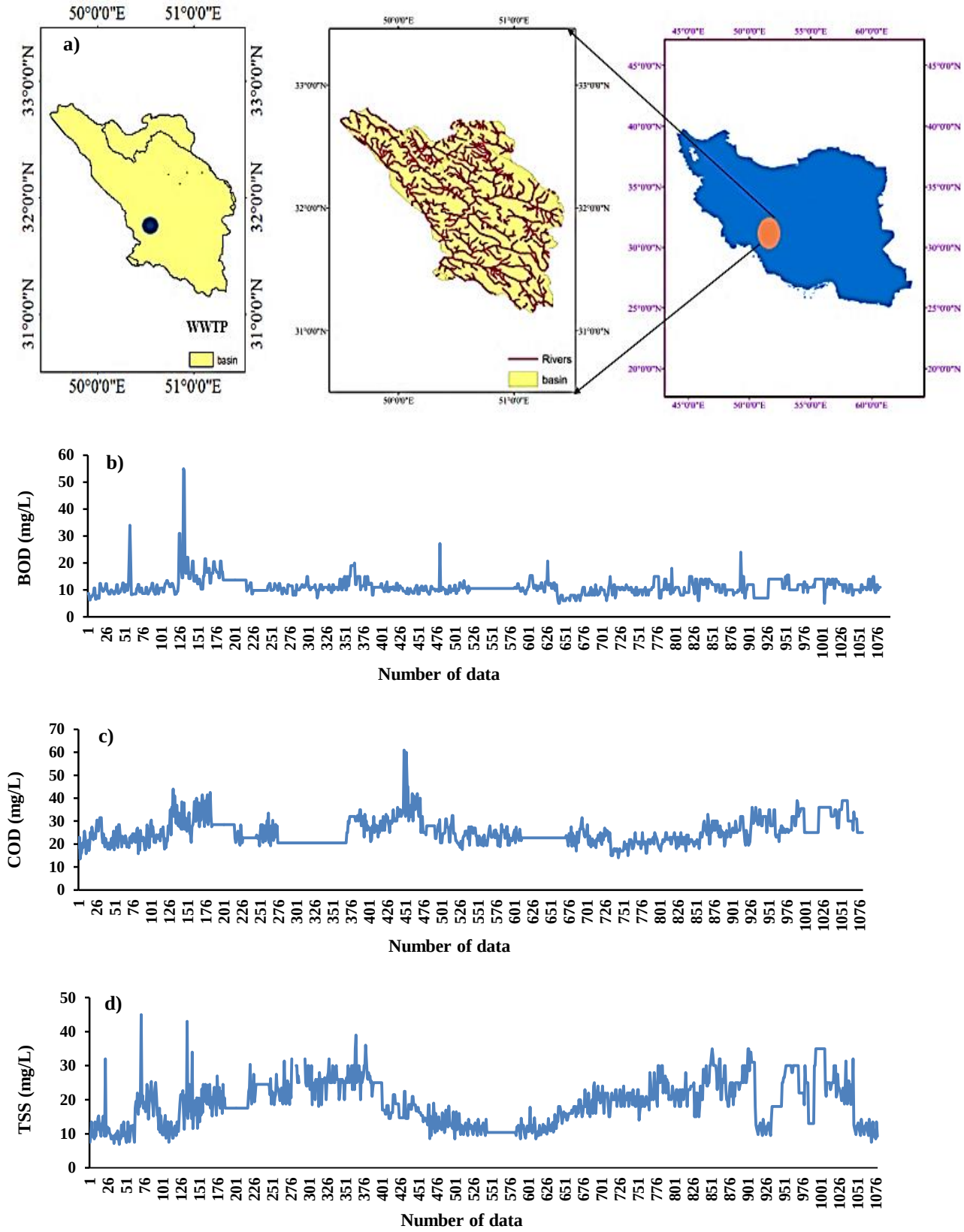


Fig. 2. a) Location of case study, b) data points for BOD, c) data points for COD and d) data points for COD (period 2016-2019)

Table 3. First to third best input combinations

Input	Γ
Best input combinations for predicting BOD	
COD_{inf}, TSS_{inf}, and BOD_{inf}	0.0234
TSS_{inf} and BOD_{inf}	0.0345
BOD_{inf}, PH_{inf}, COD_{inf}	0.0456
Best input combinations for predicting COD	
COD_{inf}, TSS_{inf}, and BOD_{inf}	0.0123
TSS_{inf} and COD_{inf}	0.0321
BOD_{inf}, PH_{inf}, COD_{inf}	0.0423
Best input combinations for predicting TSS	
COD_{inf}, TSS_{inf}, and BOD_{inf}, PH	0.0123
TSS_{inf} and COD_{inf}	0.0245
TSS_{inf} and COD_{inf}	0.0267

for 2^4-1. This study integrates NMRA with GT to reduce the computational time and difficulties associated with the GT model. In the first step, names of input variables, including BOD_{inf}, COD_{inf}, TSS_{inf}, and PH_{inf} are inserted into NMRA. Then, NMRA creates random input combinations based on the names of the input variables. An optimization algorithm encodes the names of input variables. Γ is considered the objective function. Various input combinations are evaluated by computing Γ. The location of each rat shows the input combination.

4. Results and discussion

4.1. Determination of data size for training and testing stages

It is necessary to determine the data size for training and testing. Fig. 3 shows the variation of objective function value versus sample size for predicting BOD_{eff}. ANN-NMRA's objective functions for sizes of 50%, 55%, 60%, 65%, 70%, 75%, 80%, and 85% of data in the training stage are 0.925, 1.223, 1.234, 1.345, 0.86, 1.112, 0.965 and 0.978, respectively A training stage using 70% and a testing stage using 30% of the available data produces the lowest objective function (Fig. 3). ANN-SCA's objective functions for sizes of 50%, 55%, 60%, 65%, 70%, 75%, 80%, and 85% of data in the training stage are 1.872, 1.455, 1.345, 1.554, 1.2, 1.323, 1.456, and 1.345, respectively. ANN-SCA provides the lowest objective function value using 70% and 30% data. In the training and testing stages, 70% and 30% of data are suitable for ANN-FFA, ANN-GA, and ANN models. Similarly, the best data size for predicting COD_{eff} and TSS_{eff} is determined.

4.2. Determination of the best inputs

This study uses four input variables for predicting each effluent quality parameter. Hence, 2^4-1 combinations of input variables are available for predicting output variables. NMRA reduces the computation cost of GT for identifying the best input scenario. Table 3 shows the first to third-best input combinations for predicting COD_{eff}, TSS_{eff}, and BOD_{eff}. For BOD_{eff} and COD_{eff} prediction, the Input combination of COD_{inf}, TSS_{inf}, and BOD_{eff} has the lowest Γ value.

For TSS_{eff} prediction, the Input combination of COD_{inf}, TSS_{inf}, PH_{inf} and BOD_{inf} has the lowest alpha value. The correlation coefficient only determines important input variables, while the hybrid GT determines the best input combination. Among 15 input combinations, Table 2 suggests the best input combinations. These input combinations have the lowest Γ among others. Hence, COD_{eff}, TSS_{eff}, and BOD_{eff} are predicted by using the input combinations with the lowest Γ value.

4.3. Use of SA to determine random parameters

Different factors influence the efficiency of an optimization algorithm. For optimization algorithms, random parameters such as the size of the population and the number of iterations are important. An easy way to determine the value of random parameters is to perform a SA. SA computes the objective function value (RMSE[1]) for different population sizes and different iterations. Fig. 4a, 4b and 4c show SA for different population sizes. Fig. 4a, 4b and 4c show the results of the SA for predicting BOD_{eff}, COD_{eff} and TSS_{eff}. For population sizes of 50, 100, 150, 200, and 250, the RMSEs of NMRA were 0.912, 0.876, 1.134, 1.234 and 1.245 mg/L, respectively (Fig. 4a). Therefore, the best size for NRMA was 100. Fig. 4b shows the SA process for predicting COD_{eff}.

As depicted in Fig. 4b, the RMSEs of NMRA for the population size of 50, 100, 150, 200, and 250 were 0.989, 0.812, 1.123, 1.224, and 1.344 mg/L, respectively. Fig. 4c shows the SA process for predicting TSS_{eff}. As depicted in Fig. 4c, RMSEs of NMRA for population

[1] Root Mean Square Error (RMSE)

dx.doi.org/10.22093/wwj.2024.454419.3414

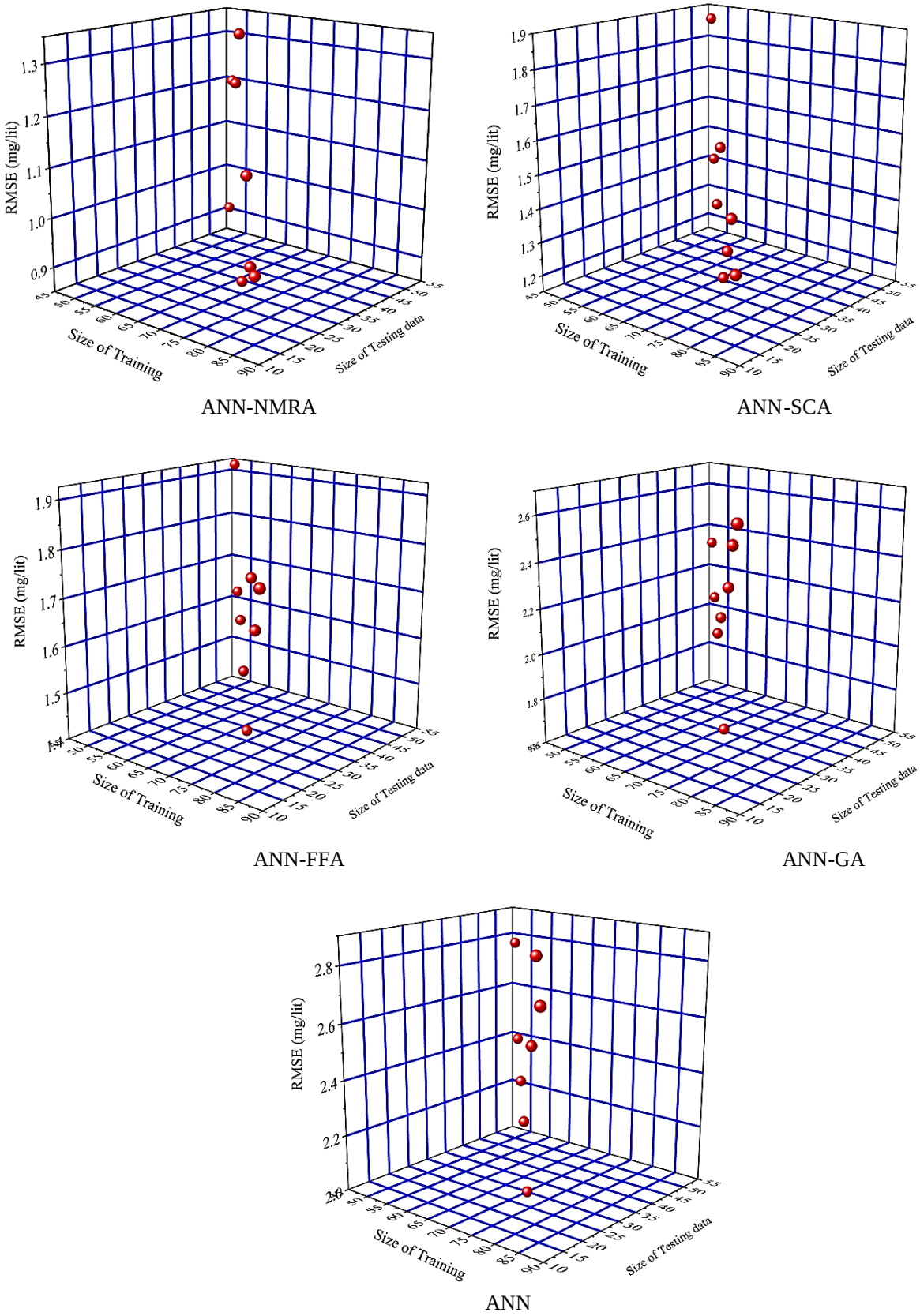


Fig. 3. Selection of the best data size for predicting BOD_{eff}

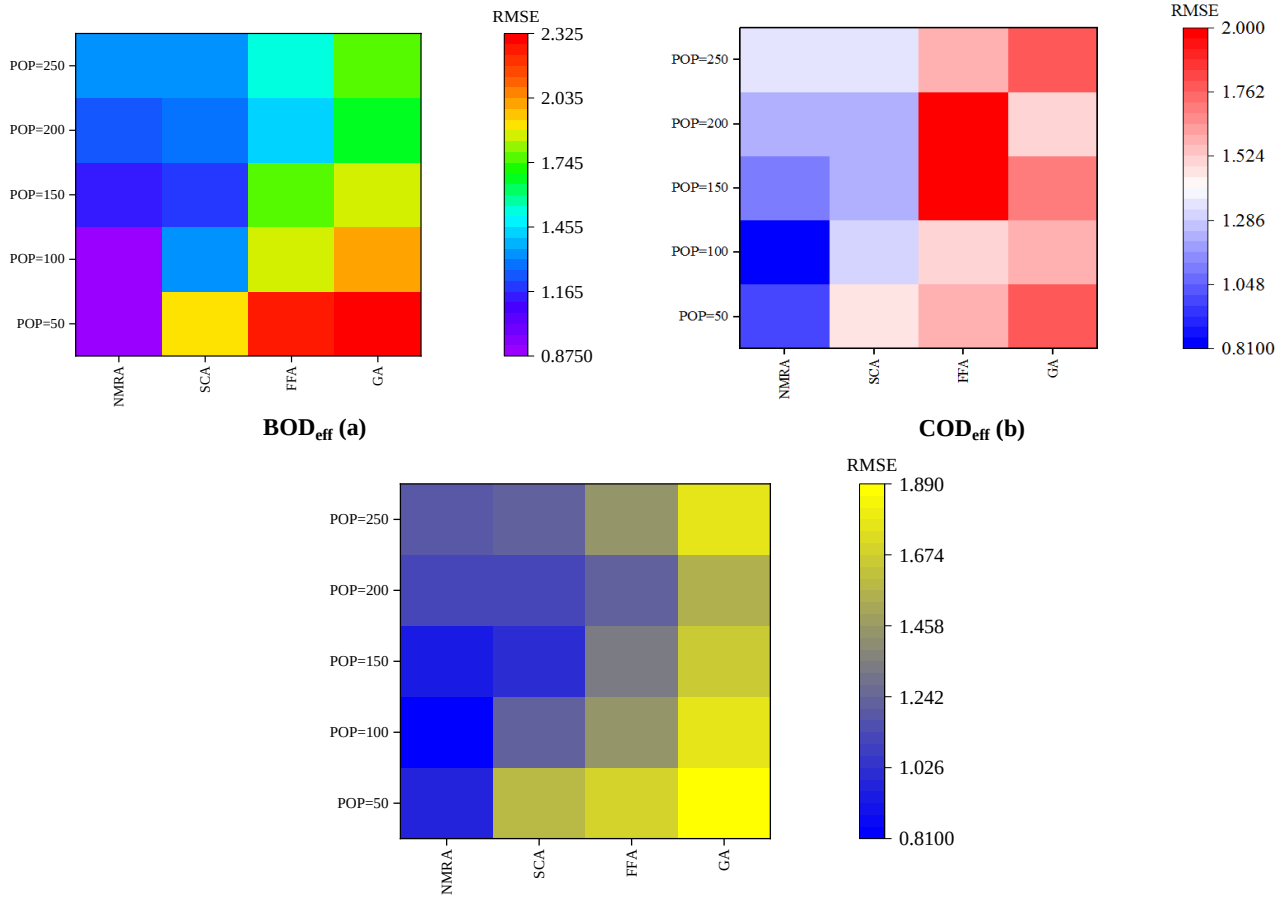
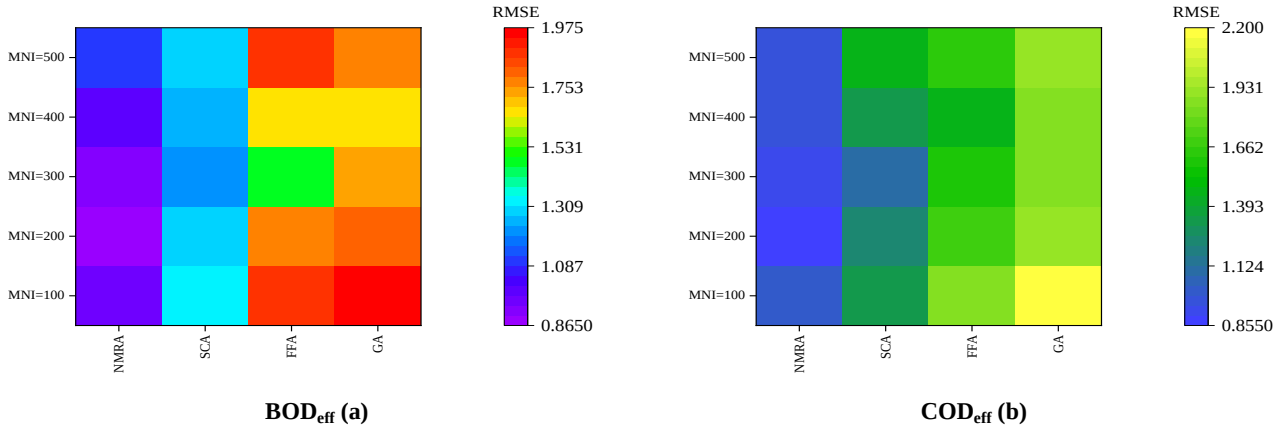


Fig. 4. SA for population size



dx.doi.org/10.22093/wwj.2024.454419.3414

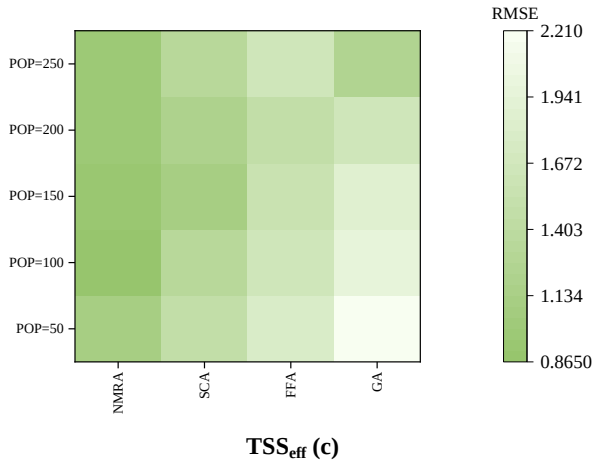


Fig. 5. SA for MNI

sizes of 50, 100, 150, 200 and 250 were 0.982, 0.814, 0.944, 1.12 and 1.2 mg/L, respectively. The best population size to predict COD_{eff} and TSS_{eff} based on NMRA was 100. Fig. 5a, 5b and 4c show SA for obtaining the best value of the MNI. Fig. 5a shows the SA process for predicting BOD_{eff} As depicted in Fig. 5a, RMSEs of NMRA for the MNI of 100, 200, 300, 400, and 500 were 0.955, 0.867, 0.914, 0.986, and 1.12 mg/L, respectively. Fig. 4b shows the SA process for predicting COD_{eff}. As depicted in Fig. 5b, RMSEs of NMRA for the MNI of 100, 200, 300, 400, and 500 were 0.998, 0.856, 0.934, 0.945, and 0.956 mg/L, respectively. Fig. 10c shows the SA process for predicting TSS_{eff}.

As depicted in Fig. 5c, RMSEs of NMRA for the MNI of 100, 200, 300, 400, and 500 were 1.1, 0.867, 0.934, 0.956, and 0.967 mg/L, respectively. In the modeling process of COD_{eff} and TSS_{eff}, 200 was the best value for MNI. Similarly, the best values for population size and MNI for other optimization algorithms were obtained based on sub-Fig. 4 and 5.

4.4. Investigation of the accuracy of models

This section evaluates the accuracy of various models. It presents the Mean Absolute Error[1], NSE, and Percentage of Bias[2] metrics for different models' predictions of BOD_{eff}, as depicted in Fig. 6a, 6b and 6c.

Here, various statistical analyses are utilized to examine the performance of the models and their accuracy in predicting outcomes. A Heat map of error indices was employed, which compares the model results with actual data based on statistical criteria such as MAE, NSE, and P BIAS. Additionally, a Box Plot was used to compare the mean values obtained from the model with the actual values to demonstrate the model's performance. The Taylor Diagram, which includes NSD, Correlation Coefficient, and CRMSD, was also used to assess the model's accuracy.

For predicting BOD_{eff} at the testing phase, the MAEs of the IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models were 78, 0.998, 1.19, 1.26, 1.34, and 1.40 mg/L, respectively. For predicting BOD_{eff} in the testing stage, PBIASs of 12%, 15%, 17%, 20%, 21%, and 23% were obtained by IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN, respectively. For predicting BOD_{eff} in the testing stage, NSEs of 0.93, 0.90, 0.89, 0.86, 0.82, and 0.80 were obtained by IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN. The results indicate that IMM and ANN-NMRA provided the highest levels of precision.

IMM ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models provided PBIASs of 16, 19, 23, 26, 27, and 28%. For predicting COD_{eff}, NSEs of 0.94, 0.93, 0.91, 0.89, 0.85, and 0.82 were obtained by IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models, respectively. IMM model's MAE was 18%, 21%, 22%, 34, and 40% lower than that of ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models, respectively.

For predicting TSS_{eff} at the testing stage, IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models produced MAEs of 0.712, 0.912, 1.12, 1.25, 1.29, and 1.32 mg/L. IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models yielded NSEs of 0.92, 0.90, 0.87, 0.85, 84, and 0.80 for predicting TSS_{eff}. PBIASs of 16 %, 18 %, 20 %, 22 %, 23 %, and 24 % were obtained by IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models, respectively, for predicting TSS_{eff} in the testing stage.

To check the quality of data and remove outliers, boxplots were computed for each input variable.

For predicting BOD_{eff}, the means of observed data, IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA,

[1] Mean Absolute Error (MAE)
[2] Percent Bias (PBIAS)

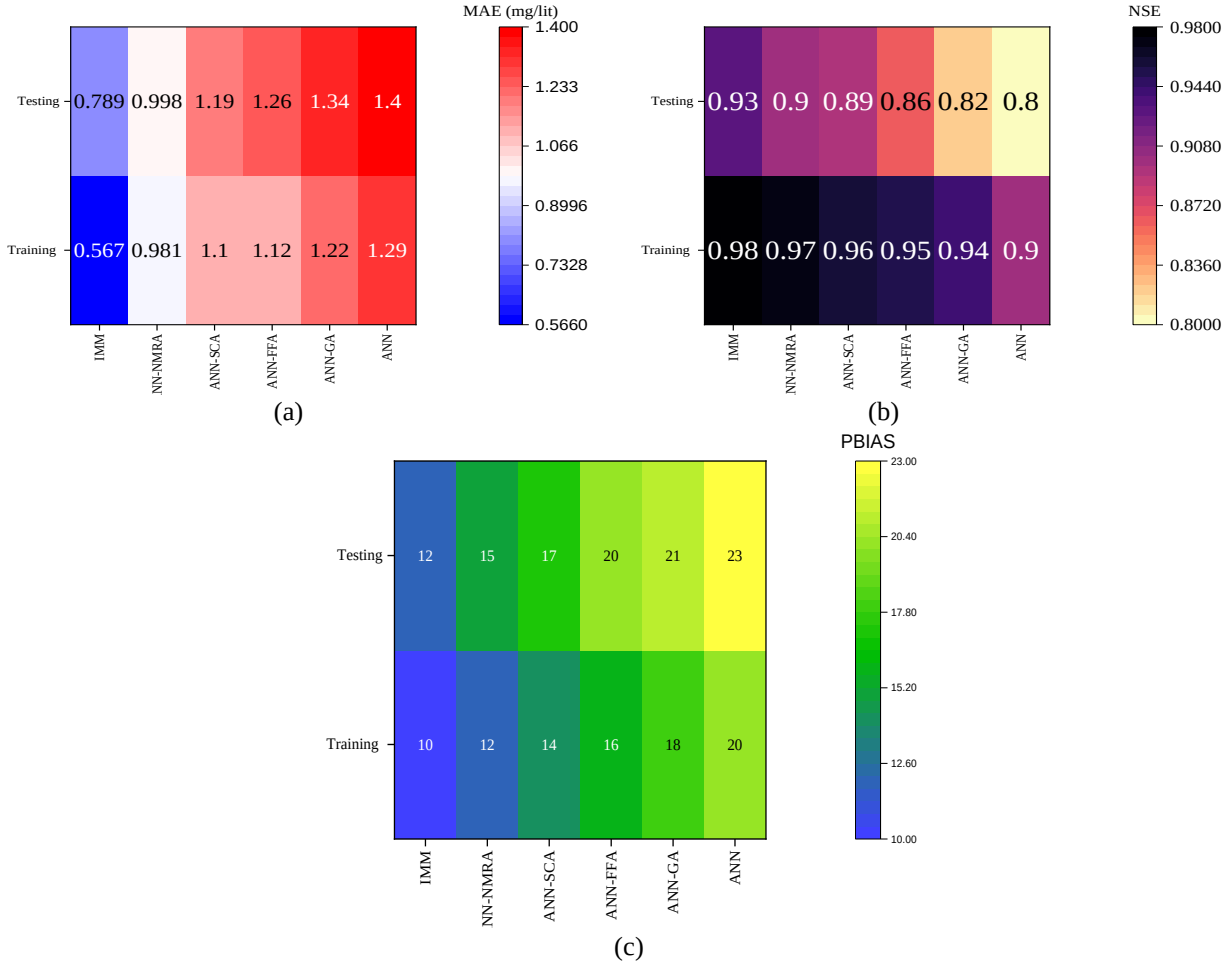


Fig. 6. Heat map of error indices for predicting BOD_{eff} based on a) MAE, b) NSE and c) PBIAS

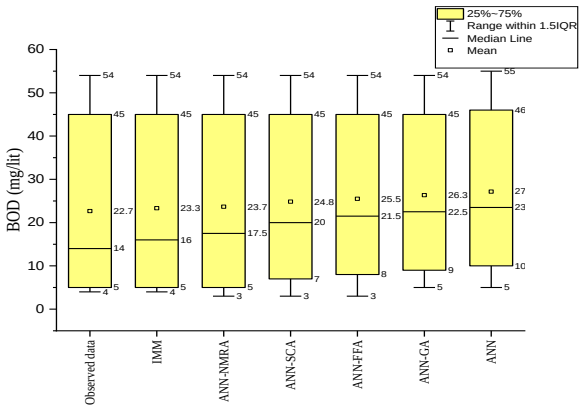


Fig. 7. Box plots of models for predicting BOD_{eff}

and ANN models were 22.7, 23.3, 23.7, 24.8, 25.5, 26.3, and 27 mg/L, respectively. For predicting COD_{eff}, medians of observed data, IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models for predicting COD_{eff} were 16, 17, 18, 19, 19, 21, and 21 mg/L, respectively. For predicting TSS_{eff}, the medians of observed data, IMM, ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models were 22.7, 23.0, 23.4, 24.1, 25.30, 26.30, and 27 mg/L respectively. Fig. 7 illustrates boxplot for predicting BOD_{eff}. Taylor diagrams are one of the tools for evaluating the accuracy of models. The diagram

dx.doi.org/10.22093/wwj.2024.454419.3414

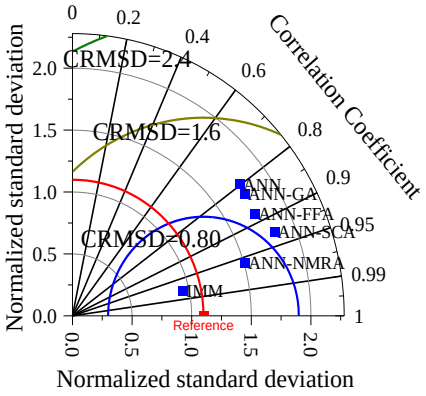


Fig. 8. Taylor diagram (BOD)

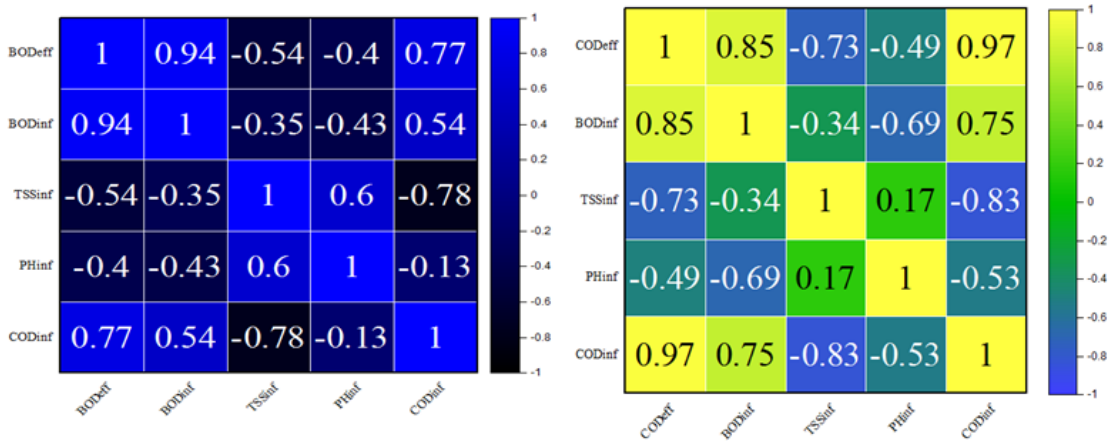


Fig. 9. Correlation heat map between inputs and BOD_{eff}, COD_{eff}

evaluates the models' efficiency. Each model with a smaller distance from the observed data (reference data) is more accurate. The performance of models based on Taylor diagrams can be seen in Fig. 8. As shown in Fig. 8, the CRMSDs of IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models for predicting BOD_{eff} were 0.24, 0.51, 0.83, 0.86, 0.96, and 1.02, respectively.

For predicting BOD_{eff}, the correlation coefficients of IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models were 0.97, 0.95, 0.92, 0.88, 0.82, and 0.79, respectively. For predicting COD_{eff}, CRMSDs of IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models were 0.17, 0.45, 1.04, 1.13, 1.18, and 1.25, respectively. For predicting TSS_{eff}, CRMSDs of IMM, ANN-NMRA, ANN-FFA, ANN-GA, ANN-FFA, and ANN models were 0.11, 0.45, 1.01, 1.17, and 1.31, 1, 51, respectively.

IMM and ANN-NMRA have the highest accuracy among the different models. The benefits of different ANN models were embedded into the IMM model. The IMM model uses hybrid and standalone models to achieve the best accuracy. The IMM model also provides synergy among different models. NMRA improves the accuracy of models by updating common workers and breeders with Equations 2 and 3. These two equations can be used to provide new solutions and escape from optimum local situations. Therefore, NMRA provides better solutions than other algorithms. In general, the integration of optimization algorithms with ANN models improves the efficiency of ANN models.

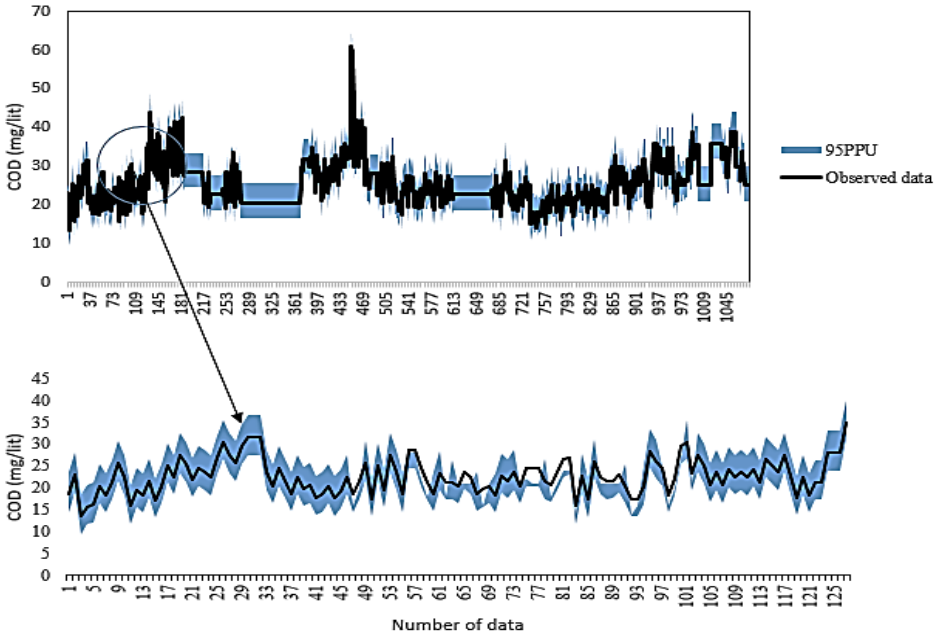
4.5. Further discussion

In the previous section, hybrid GT was used for selecting input scenarios. However, the performance of hybrid GT needs to be verified. In the best scenario of GT, available parameters impact outputs significantly, so these parameters should naturally have a high correlation with outputs. Fig. 9 shows the correlation heat map between inputs and outputs.

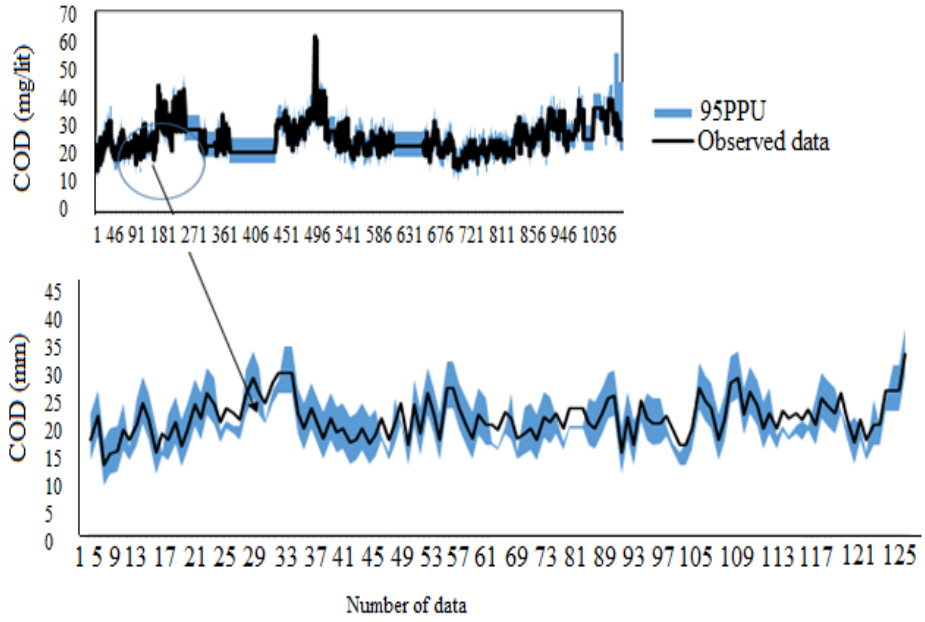
BOD_{inf}, COD_{inf}, and TSS_{inf} are highly correlated with BOD_{eff}. BOD_{inf} has the most correlation with BOD_{eff} Fig. 9 shows that BOD_{inf}, COD_{inf}, and TSS_{inf} have a high correlation with COD_{eff}. The best scenario chosen by Gamma includes BOD_{inf}, COD_{inf}, and TSS_{inf}. TSS_{eff} has a high correlation with BOD_{inf}, COD_{inf}, TSS_{inf}, and PH_{inf}.

The Gamma test selected these parameters as the best scenario for predicting TSS_{eff}. Therefore, the correlation heat map confirms the performance of hybrid GT.

In this article, uncertainty analysis examines the effects of uncertainty of inputs and model parameters on outputs. Here is a sample of the IMM uncertainty bound for predicting COD_{eff}. Fig. 10a and 10b show the uncertainty bounds of IMM by considering input and parameter uncertainty. In Fig. 10a, 15 data points do not fall within the bounds. Hence, P is 99% (1-15/1080). The width of the bound is 0.12. In Fig. 10b, 30 data points do not fall within the bounds. Therefore, P for Fig. 10b is 98% (1-30/1080). The width of the bound for Fig. 10b is 0.15. Hence, input uncertainty (IU) is lower than parameter uncertainty.



(a)



(b)

Fig. 10. Uncertainty analysis of IMM model based on a) input uncertainty and b) parameter uncertainty

dx.doi.org/10.22093/wwj.2024.454419.3414

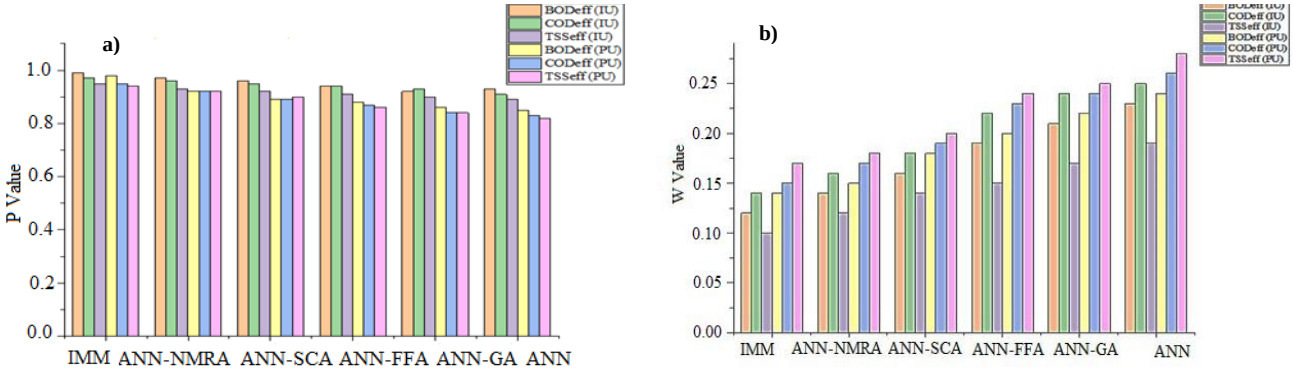


Fig. 11. a) P values for models based on uncertainty analysis and b) W values for models based on parameter uncertainty

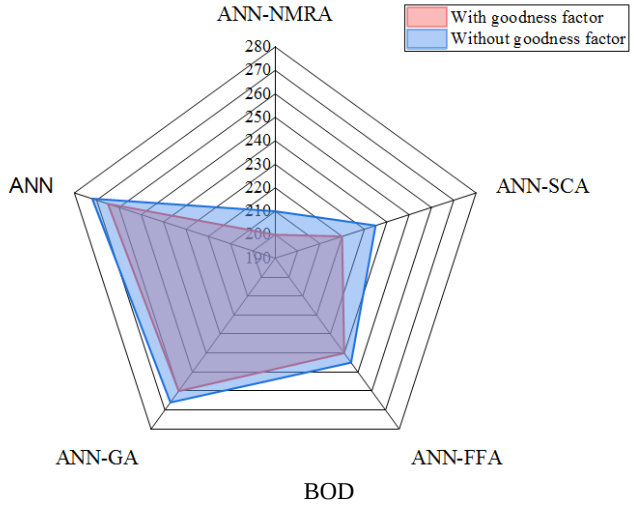


Fig. 12. Computational time of models

We use the same approach to compute the P and W of different models for predicting COD_{eff}, TSS_{eff}, and BOD_{eff}. Fig. 11a and 11b show the P and W for predicting COD_{eff}, TSS_{eff}, and BOD_{eff}. Among other models, IMM and ANN-NMRA provide the lowest values of w and P. ANN, based on input and parameter uncertainty (PU_, provides the highest uncertainty level. Fig. 11a and 11b indicate that IU is lower than PU. IMM's IUs for predicting COD_{eff}, TSS_{eff}, and BOD_{eff} are 0.97, 0.95, and 0.99, respectively, while its PUs for predicting COD_{eff}, TSS_{eff}, and BOD_{eff} are 0.95, 0.94, and 0.98, respectively.

Fig. 12 compares the computational times of models with and without goodness factors.

Each model requires a different amount of time for processing and producing output. In this research, with the help of a goodness factor, an attempt has been made to minimize this duration in order to reduce computational costs.

The CPU time of modified ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models with a goodness factor was 200, 220, 240, 260, and 265 s, whereas the CPU time of ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA, and ANN models without goodness factor was 210, 235, 245, 265, and 272 s, respectively. Models with a goodness factor reduce the computational time of models without a goodness factor since the goodness factor removes ineffective neurons from the structure of ANN models.

The results of this study are also compared with those of other studies. Pai et al. used adaptive neuro-fuzzy interface system[1] and ANN models for predicting effluent quality parameters. The ANFIS model produced correlation coefficients of 0.88 and 0.93 in the testing and training stages. The correlation coefficient of the IMM model in the current study is 0.99 (Pai et al., 2011).

Nourani et al. applied ANN and SVM and the ensemble method for predicting BOD_{eff} and COD_{eff}. The ensemble yielded correlation coefficients of 0.95 and 0.90 for predicting COD_{eff} and BOD_{eff}, respectively. Correlation coefficients of IMM are 0.99 and 0.97 for predicting COD_{eff} and BOD_{eff} (Nourani et al., 2018).

[1] Adaptive Neuro- Fuzzy Interface System (ANFIS)

Heddam et al. applied ANFIS with the grid partition[1] method for predicting COD_{eff}. The study reported an MAE of 5.712 for ANFIS-GP, but the MAEs for IMM models at the training and testing stages are 0.567 and 0.789 for the present study (Heddam et al., 2020).

Sharafati et al. used GBR for predicting effluent quality parameters. They reported correlation coefficients of 0.97 and 0.75 for predicting BOD and COD, while the correlation coefficients of IMM were 0.97 and 0.99 for predicting BOD and COD (Sharafati et al., 2020).

5. Conclusion

Agriculture and groundwater resource management require reclaimed water. A WWTP is used to treat wastewater. Evaluation of effluent quality parameters of a WWTP is an important topic for decision-makers. This study predicts BOD_{eff}, COD_{eff} and TSS_{eff} using hybrid ANN models and an ensemble model. This study integrates GT with NMRA to determine the best input. Although there are 24-1 possible input combinations for predicting outputs, GT-NMRA automatically determines the best input combination. The IMM model in the current study is used as an ensemble model to improve output accuracy. IMM's advantage is that it uses the potential of various ANN models and provides synergy amongst them. For predicting COD_{eff}, the MAE of the IMM model is 18%, 21%, 22%, 34, and 40% lower than the MAE of ANN-NMRA, ANN-SCA, ANN-FFA, ANN-GA and ANN models, respectively. ANN model has the highest MAE and the lowest NSE for predicting BOD_{eff} and TSS_{eff}, while IMM and ANN-NMRA have the lowest MAE and the highest NSE for predicting BOD_{eff} and TSS_{eff}. The study indicates that model parameters lead to higher uncertainty compared to inputs. It indicates that identifying redundant hidden neurons is important for reducing CPU time in ANN models. In ANN models, the goodness factor reduces CPU time significantly. NMRA indicates a high potential for training soft computing models since it uses advanced operators for escaping from local optimum solutions. Water resource managers can use the results of this study to provide spatial and temporal maps of water quality parameters in large basins.

Additionally, the proposed models can estimate various metrological parameters in hydrology. NMRA can be integrated with other soft computing models to predict effluent quality parameters in future studies. The capability of the IMM model in the current study can also be compared with that of other similar models. It's essential to acknowledge that since the model accepts diverse inputs, it's adaptable to a range of environmental scenarios. Consequently, this new model is regarded as universally applicable.

Given the critical importance of water quality monitoring in industrial applications, the existing model can be effectively employed within this sector for evaluating water quality.

6. Acknowledgments

This study was funded by the University of Shahrekord, Iran. The financial support of this organization is appreciated. The authors are also grateful to the Chaharmahal and Bakhtiari Water and Wastewater Company of Iran, for providing the data for this research.

[1] Grid Partition (GP)

References

Asghari, P., Nourani, V., Sharghi, E. and Behfar, N., 2022. Using ensemble model to improve ANN, ANFIS, SVR models in predicting effluent BOD and COD. *Amirkabir Journal of Civil Engineering*, 53(11), 4683-4702. https://doi.org/10.22060/CEEJ.2020.18441.6873.

Cu Thi, P., Ball, J. E. and Dao, N. H., 2018. Uncertainty estimation using the Glue and Bayesian approaches in flood estimation: a case study-Ba River, Vietnam. *Water*, 10(11), 1641. https://doi.org/10.3390/w10111641.

Ehteram, M., Banadkooki, F. B., Fai, C. M., Moslemzadeh, M., Sapitang, M., Ahmed, A. N., et al. 2021a. Optimal operation of multi-reservoir systems for increasing power generation using a seagull optimization algorithm and heading policy. *Energy Reports*, 7, 3703-3725. https://doi.org/10.1016/j.egyr.2021.06.008.

Ehteram, M., Sammen, S. S., Panahi, F. and Sidek, L. M., 2021b. A hybrid novel SVM model for predicting CO_2 emissions using Multiobjective Seagull Optimization. *Environmental Science and Pollution Research*, 28, 66171-66192. https://doi.org/10.1007/s11356-021-15223-4.

Elmaadawy, K., Abd Elaziz, M., Elsheikh, A. H., Moawad, A., Liu, B. and Lu, S., 2021. Utilization of random vector functional link integrated with manta ray foraging optimization for effluent prediction of wastewater treatment plant. *Journal of Environmental Management*, 298, 113520. https://doi.org/10.1016/j.jenvman.2021.113520.

Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., H. Kashani, M. and Mohammadi, B., 2018. Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: case study in North Iran. *Theoretical and Applied Climatology*, 133, 1119-1131. https://doi.org/10.1007/s00704-017-2244-0.

Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. P., et al. 2015. Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences*, 32, 90-101. https://doi.org/10.1016/j.jes.2015.01.007.

Heddam, S., Kisi, O, Sebbar, A., Houichi, L., and Djemili, L., 2020. Predicting water quality indicators from conventional and nonconventional water resources in Algeria country: adaptive neuro-fuzzy inference systems versus artificial neural networks. *Water Resources in Algeria-Part II: Water Quality, Treatment, Protection and Development*, 13-34. https://doi.org/10.1007/698_2019_399.

Heddam, S., Lamda, H. and Filali, S., 2016. Predicting effluent biochemical oxygen demand in a wastewater treatment plant using generalized regression neural network based approach: a comparative study. *Environmental Processes*, 3, 153-165. https://doi.org/10.1007/s40710-016-0129-3.

Hejabi, N., Saghebian, S. M., Aalami, M. T. and Nourani, V., 2021. Evaluation of the effluent quality parameters of wastewater treatment plant based on uncertainty analysis and post-processing approaches (case study). *Water Science and Technology*, 83(7), 1633-1648. https://doi.org/10.2166/wst.2021.067.

Kan, G., He, X., Ding, L., Li, J., Hong, Y. and Liang, K., 2020. Heterogeneous parallel computing accelerated generalized likelihood uncertainty estimation (GLUE) method for fast hydrological model uncertainty analysis purpose. *Engineering with Computers*, 36, 75-96. https://doi.org/10.1007/s00366-018-0685-4.

Liang, G., Panahi, F., Ahmed, A. N., Ehteram, M., Band, S. S. and Elshafie, A., 2021. Predicting municipal solid waste using a coupled artificial neural network with archimedes optimisation algorithm and socioeconomic components. *Journal of Cleaner Production*, 315, 128039. https://doi.org/10.1016/j.jclepro.2021.128039.

Liu, J., Shao, W., Xiang, C., Mei, C. and Li, Z., 2020. Uncertainties of urban flood modeling: influence of parameters for different underlying surfaces. *Environmental Research*, 182, 108929. https://doi.org/10.1016/j.envres.2019.108929.

Lotfi, K., Bonakdari, H., Ebtehaj, I., Mjalli, F. S., Zeynoddin, M., Delatolla, R., et al. 2019. Predicting wastewater treatment plant quality parameters using a novel hybrid linear-nonlinear methodology. *Journal of Environmental Management*, 240, 463-474. https://doi.org/10.1016/j.jenvman.2019.03.137.

Meng, X., Zhang, Y. and Qiao, J., 2021. An adaptive task-oriented RBF network for key water quality parameters prediction in wastewater treatment process. *Neural Computing and Applications*, 33(17), 11401-11414. https://doi.org/10.1007/s00521-020-05659-z.

Mirjalili, S., 2018. Genetic algorithm. *Evolutionary Algorithms and Neural Networks: Theory and Applications*, 43-55. https://doi.org/10.1007/978-3-319-93025-1_4.

Nadiri, A. A., Shokri, S., Tsai, F. T. C. and Moghaddam, A. A., 2018. Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model. *Journal of Cleaner Production*, 180, 539-549. https://doi.org/10.1016/j.jclepro.2018.01.139.

Niu, G., Yi, X., Chen, C., Li, X., Han, D., Yan, B., et al. 2020. A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. *Journal of Cleaner Production*, 265, 121787. https://doi.org/10.1016/j.jclepro.2020.121787.

Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A., et al. 2011. Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology*, 401(3-4), 177-189. https://doi.org/10.1016/j.jhydrol.2011.02.021.

Nourani, V., Asghari, P. and Sharghi, E., 2021. Artificial intelligence based ensemble modeling of wastewater treatment plant using jittered data. *Journal of Cleaner Production*, 291, 125772. https://doi.org/10.1016/j.jclepro.2020.125772.

Nourani, V., Elkiran, G. and Abba, S. I., 2018. Wastewater treatment plant performance analysis using artificial intelligence–an ensemble approach. *Water Science and Technology*, 78(10), 2064-2076. https://doi.org/10.2166/wst.2018.477.

Pai, T. Y., Yang, P. Y., Wang, S. C., Lo, M. H., Chiang, C. F., Kuo, J. L., et al. 2011. Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Applied Mathematical Modelling*, 35(8), 3674-3684. https://doi.org/10.1016/j.apm.2011.01.019.

Panahi, F., Ehteram, M. and Emami, M., 2021. Suspended sediment load prediction based on soft computing models and Black Widow Optimization Algorithm using an enhanced gamma test. *Environmental Science and Pollution Research*, 28(35), 48253-48273. https://doi.org/10.1007/s11356-021-14065-4.

Ragab, R., Kaelin, A., Afzal, M. and Panagea, I., 2020. Application of Generalized Likelihood Uncertainty Estimation (GLUE) at different temporal scales to reduce the uncertainty level in modelled river flows. *Hydrological Sciences Journal*, 65(11), 1856-1871. https://doi.org/10.1080/02626667.2020.1764961.

Rahbar, A., Mirarabi, A., Nakhaei, M., Talkhabi, M. and Jamali, M., 2022. A comparative analysis of data-driven models (SVR, ANFIS, and ANNs) for daily karst spring discharge prediction. *Water Resources Management*, 36(2), 589-609. https://doi.org/10.1007/s11269-021-03041-9.

Salehnia, N., Falahi, M. A., Seifi, A. and Adeli, M. H. M., 2013. Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis. *Journal of Natural Gas Science and Engineering*, 14, 238-249. https://doi.org/10.1016/j.jngse.2013.07.002.

Salgotra, R. and Singh, U., 2019. The naked mole-rat algorithm. *Neural Computing and Applications*, 31, 8837-8857. https://doi.org/10.1007/s00521-019-04464-7.

Shaibani, M. J., Emamgholipour, S. and Moazeni, S. S., 2022. Investigation of robustness of hybrid artificial neural network with artificial bee colony and firefly algorithm in predicting COVID-19 new cases: case study of Iran. *Stochastic Environmental Research and Risk Assessment*, 36(9), 2461-2476. https://doi.org/10.1007/s00477-021-02098-7.

Sharafati, A., Asadollah, S. B. H. S. and Hosseinzadeh, M., 2020. The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty. *Process Safety and Environmental Protection*, 140, 68-78. https://doi.org/10.1016/j.psep.2020.04.045.

Srivastav, R. K., Sudheer, K. P. and Chaubey, I., 2007. A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. *Water Resources Research*, 43(10). https://doi.org/10.1029/2006WR005352.

Watada, J., 2006. Structural learning of neural networks for forecasting stock prices. *In Knowledge-Based Intelligent Information and Engineering Systems: 10th International Conference, KES 2006, Bournemouth, UK, October 9-11, 2006. Proceedings, Part III 10* (pp. 972-979). Springer Berlin Heidelberg. https://doi.org/10.1007/11893011_123.