water

MDPI

Article

Enhancing the Prediction of Influent Total Nitrogen in Wastewater Treatment Plant Using Adaptive Neuro-Fuzzy Inference System–Gradient-Based Optimization Algorithm

Misbah Ikram [1], Hongbo Liu [1,*], Ahmed Mohammed Sami Al-Janabi [2], Ozgur Kisi [3,4,5], Wang Mo [6,*], Muhammad Ali [7] and Rana Muhammad Adnan [6,*]

1 School of Environmental Science and Technology, Tianjin University, Tianjin 300072, China; rai.ai@gzhu.edu.cn
2 Department of Civil Engineering, Cihan University-Erbil, Erbil 44001, Iraq; ahmed.mohammedsami@cihanuniversity.edu.iq
3 Department of Civil Engineering, Lübeck University of Applied Science, 23562 Lübeck, Germany; ozgur.kisi@th-luebeck.de
4 Department of Civil Engineering, School of Technology, Ilia State University, 0162 Tbilisi, Georgia
5 School of Civil, Environmental and Architectural Engineering, Korea University, Seoul 02841, Republic of Korea
6 College of Architecture and Urban Planning, Guangzhou University, Guangzhou 510006, China
7 College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; alirana233@gmail.com
* Correspondence: llhhbb@tju.edu.cn (H.L.); saupwangmo@gzhu.edu.cn (W.M.); rana@gzhu.edu.cn (R.M.A.)

Abstract: For the accurate estimation of daily influent total nitrogen of sewage plants, a novel hybrid approach is proposed in this study, where a gradient-based optimization (GBO) algorithm is employed to adjust the hyper-parameters of an adaptive neuro-fuzzy system (ANFIS). Several benchmark methods for optimizing ANFIS parameters are compared, which include particle swarm optimization (PSO), gray wolf optimization (GWO), and gradient-based optimization (GBO). The prediction accuracy of the ANFIS-GBO model is evaluated against other models using four statistical measures: root-mean-squared error (RMSE), mean absolute error (MAE), and Nash–Sutcliffe efficiency (NSE), and coefficient of determination (R^2). Test results show that the suggested ANFIS-GBO outperforms the standalone ANFIS, hybrid ANFIS-PSO and ANFIS-GWO methods in daily influent total nitrogen prediction from the sewage treatment plant. The ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO models are evaluated using seven distinct input combinations to predict daily TN_{inf}. The results from both the testing and training periods demonstrate that these models, namely ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO, exhibit the highest level of accuracy for the seventh input combination (Q_w, pH, SS, TP, NH_3-N, COD, and BOD5). ANFS-GBO-7 reduced the RMSE in the prediction of ANFIS-7, ANFIS-PSO-7, and ANFIS-GWO-7 by 21.77, 10.73, and 6.81%, respectively, in the test stage. Results from testing and training further demonstrate that increasing the number of parameters (NH_3-N, COD, and BOD) as input improves the models' ability to make predictions. The outcomes show that the ANFIS-GBO model can potentially be suggested for the daily prediction of influent total nitrogen (TN_{inf}) in full-scale wastewater treatment plants.

Keywords: sewage treatment plant; influent total nitrogen prediction; adaptive network-driven fuzzy inference system; particle swarm optimization; gray wolf optimization; gradient-based optimization

Citation: Ikram, M.; Liu, H.; Al-Janabi, A.M.S.; Kisi, O.; Mo, W.; Ali, M.; Adnan, R.M. Enhancing the Prediction of Influent Total Nitrogen in Wastewater Treatment Plant Using Adaptive Neuro-Fuzzy Inference System–Gradient-Based Optimization Algorithm. *Water* **2024**, *16*, 3038. https://doi.org/10.3390/w16213038

Academic Editor: Christos S. Akratos

Received: 23 September 2024
Revised: 16 October 2024
Accepted: 22 October 2024
Published: 23 October 2024

1. Introduction

Wastewater is a diluted mixture of contaminants originating from residential, commercial, industrial, and municipal sources. Its characteristics can vary based on the source of discharge [1]. Wastewater is contaminated with a variety of materials, including bacteria, organic waste, colloids, suspended particles, and nutrients [2,3].

Wastewater treatment plants (WWTPs) are made up of a variety of tools and procedures for treating wastewater, which generally permits the disposal of industrial effluents and anthropogenic activities without endangering public health or causing unfavorable harm to the environment [4]. Wastewater management and maintaining a clean environment are crucial in promoting the effective development of and protecting public health [5]. Because a WWTP is a very complex and dynamic system, maintaining effective operation and control over it is crucial in preventing public and environmental health problems. Any WWTP's performance and operation are significantly impacted by a number of factors, including influent fluctuations [6]. To design WWTPs, it is important to obtain the influent characteristics [7].

Furthermore, the input flow dramatically impacts the wastewater treatment plant's energy consumption and overall treatment process. Certain key factors, such as total nitrogen, are challenging to assess. In addition, nitrogen is a significant pollutant in wastewater that must be reduced to a specific level before discharge. The main forms of total nitrogen (TN) in wastewater include ammonia, nitrite, nitrate, and organically bound nitrogen [8]. Monitoring TN in the influent of wastewater treatment plants (WWTPs) is crucial for optimizing nutrient removal systems, controlling sludge production, and managing various wastewater treatment processes [9]. Engineers are required to understand and measure the properties of wastewater, particularly the nutrient components, both at the beginning and end of the treatment process. In order to gather the essential information, the operator must collect sensor data or take samples of the wastewater and analyze the flow of influent and effluent in the treatment plant to determine the characteristics of the untreated waste. Introducing inadequately treated wastewater containing nutrients into water bodies like groundwater systems can lead to various health problems [10]. Predicting the incoming water quality for future periods is essential to efficiently handle the plant's functions and regulate the quality of the effluent. Controlling wastewater treatment plants is difficult because of their complex mechanisms and varying wastewater quality. Many traditional methods rely on assumptions and simplified estimates, which limits their effectiveness [11,12]. However, traditional methods are complex and laborious [13]. Mechanistic models require extensive data, making calibration and validation challenging. The model's identification remains uncertain, making only approximations of mathematical expressions [14]. Data-driven AI techniques do not require this specification [15]. Over the last few years, the utilization of AI technology for modeling and forecasting environmental phenomena has increased due to its ability to address practical challenges in wastewater treatment [16], river quality monitoring [17], and management of water resources [18]. Artificial intelligence (AI) approaches are suitable for modeling complicated WWTPs because they can learn to generate non-linear linkages that can explain the complicated connections within the dataset without requiring the difficult effort of addressing deterministic nonlinear mathematics [19]. Additionally, when given inputs that they have never seen before, AI models can generalize the input–output relationship and generate an output. Recently, numerous research studies have been carried out using intelligent techniques in wastewater treatment. Tumer and Edebali [20] investigated the Konya wastewater treatment plant using artificial neural network (ANN) models, evaluating its treatment efficiency using input values like pH, temperature, chemical oxygen demand (COD), total suspended solids (TSSs), and biological oxygen demand (BOD). Arismendy and Gomez [21] proposed a multilayer perceptron neural network for predicting COD in bioreactors. Alsulaili and Refaie [22] studied the use of ANNs for forecasting influent BOD_5 concentration and evaluating the effectiveness of wastewater treatment plants (WWTPs). The ANN model outperformed other variables, with R^2 values of 0.752, 0.612, and 0.631 for BOD, COD, and TSS concentrations. Yaqub and Asif [23] developed a neural network based on long short-term memory to forecast ammonium removal efficiency in an anaerobic–anoxic–oxic membrane bioreactor system. Hejabi and Aalami [24] used feedforward neural network (FFNN) and support vector machine (SVM) models for the performance prediction of a Tabriz WWTP, finding that AI models outperformed linear models in terms of efficiency.

Nadiri et al. [25] proposed an advanced fuzzy logic (FL) model and compared it with a supervised committee FL model for a WWTP, which outperformed individual FL and CFL models in predicting effluent water parameters. Cheng et al. [26] developed six deep learning models using long short-term memory (LSTM) and gated recurrent unit (GRU) architectures to forecast key features of a wastewater treatment plant, including influent flow, temperature, BOD, effluent chloride, and power consumption. Granata et al. [27] used SVM and regression trees to predict the fluent concentrations of BOD, COD, TSSs, and total dissolved solids (TDSs) in a wastewater treatment plant, showing robustness, reliability, and high generalization capability, with SVM showing a better performance. Safdar et al. [28] utilized partial least squares (PLS), MLR, multilayer perceptron (MLP), LSTM, GRU, and multihead-attention-based GRU (MAGRU) models for predicting hourly based TN effluent. They found that MAGRU provided more accurate results in comparison to other models. Bagherzadeh et al. [29] adopted the ANN, Random Forest (RF), and Gradient Boosting Machine (GBM) models for daily influent TN forecasting. GBM-based models outperformed the other benchmark models. The study conducted by Nourani et al. [30] used three AI methods, FFNN, adaptive neuro-fuzzy inference system (ANFIS), and SVM, to assess the performance of the Nicosia wastewater treatment plant. The ANFIS model was found to be more effective than the other models. Araromi et al. [31] conducted a study on the activated sludge process in an industrial wastewater treatment plant, utilizing the ANFIS and generalized linear model (GLM) regression to develop predictive models for effluent COD and BOD. The ANFIS models showed superior performance in accurately predicting effluent variables. Pai et al. [32] utilized three different types of ANFIS to forecast the levels of effluent suspended solids (SSs), COD, and pH from a wastewater treatment plant located in an industrial park. In order to make a comparison, they also employed an ANN. The study's findings demonstrated that ANFIS outperformed ANN in accurately predicting the effluent levels. ANFIS achieved minimum mean absolute percentage errors of 2.67%, 2.80%, and 0.42% for SSs, COD, and pH, respectively. Additionally, the maximum correlation coefficients for SSs, COD, and pH were found to be 0.96, 0.93, and 0.95. The architecture of ANFIS was shown to overcome the limitations of traditional ANNs. Qiao et al. [33] presented a study to examine the effectiveness of the ANFIS model in predicting the reduction in significant pollutants in a WWTP. The researchers compared and analyzed the measured removal and predicted values of key pollutants, such as COD, BOD, ammonia nitrogen (NH_3-N), and suspended solids (SSs). The study's results showed satisfactory linear outcomes with better R^2 values. Although the above literature shows successful applications of the ANFIS model, to achieve more accurate results, the optimal structure of the ANFIS model is required. To achieve the optimal structure, many optimization algorithms have been utilized in the literature. It has been shown that implementing AI optimization approaches, such as particle swarm optimization (PSO), can save operating expenses in wastewater treatment systems [34]. PSO has been extensively employed in wastewater treatment as a training algorithm for ANNs and to estimate the optimum process parameters. For instance, Akbaş et al. [35] developed an integrated model MLP-PSO for predicting and optimizing the biogas production in a wastewater treatment facility. Similarly, Sarkar et al. [36] utilized an ANN-PSO model for optimizing the process parameters for the maximum biosorptive removal from wastewater. Mahadeva et al. [37] proposed an optimized PSO-based artificial neural network (PSO-ANN) technique to predict reverse osmosis water treatment desalination plant performance. The model's simulations showed superiority over existing ANN models, with an R^2 of 99.1% and an error of 0.006. In recent years, the gray wolf optimizer (GWO) algorithm is successfully utilized in different environmental problems for the determination of the optimal structures of the models. Tikhamarine et al. [38] used the GWO algorithm to enhance monthly streamflow predictions, outperforming alternative techniques like particle swarm optimization and multi-verse optimization. Xia et al. [39] optimized SVM for non-linear regression predictions and water quality classifications using four strategies. The Differential Evolution GWO (DE-GWO) algorithm outperformed others due to

its ability to prevent local optimal solutions. The performance of various evolutionary algorithms, such as the Fire-Fly Algorithm (FFA), Genetic Algorithm (GA), GWO, PSO, and DE hybridized with ANFIS, was trained and tested by Riahi-madvar et al. [40]. Findings indicated that ANFIS-GWO was the best hybrid model for streamflow forecasting, which is a significant improvement above traditional black box machine learning models. In the past couple of years, the gradient-based optimizer algorithm (ANN-GBO) received much attention in finding different model's optimal structures due to its main advantages of easy implementation and lesser parameter tuning compared to other optimization algorithms [41]. Konakoglu et al. [42] studied the effectiveness of an artificial neural network with a gradient-based optimizer algorithm (ANN-GBO) in identifying local geoids. The method outperformed metaheuristic-based ANN models, conventional ANNs, and other interpolation techniques using performance metrics like R^2, MAE, MARE, and RMSE, resulting in the smallest prediction error. Kadkhodazade et al. [43] used a new LSSVM (least square support vector machine) model and a gradient-based optimization technique to assess water quality metrics across three Karun River basin monitoring stations. The LSSVM-GBO method performed better than other benchmark datasets and methods. Adnan et al. [44] used the GBO algorithm to develop a hybrid approach for precise streamflow estimation in a hilly river basin, enhancing predictions and outperforming other models.

The above literature review shows that GBO is not used in the evaluation of wastewater treatment plants. The aim of the work presented in this paper is to provide a systematic and thorough approach to the development of artificial intelligence techniques in modeling and monitoring wastewater treatment plants and show the potential of the hybrid systems of these techniques to deal with the complexity and uncertainty in the process. This research aims to predict influent total nitrogen from a sewage treatment plant in Dalian City, China, using different hybrid ANFIS models, with the best predicting model structures resulting in minimum performance errors. This paper introduces a novel hybrid method utilizing the GBO approach to fine-tune the hyper-parameters of the ANFIS model. This approach facilitates the accurate prediction of a wastewater treatment plant's monitoring capacity. This study's results accurately estimate the daily influent total nitrogen concentration in wastewater plants, offering crucial data support for developing water pollution control programs. The major contribution of this paper is the advancement in the predicting models that have been developed and tested for predicting TN_{inf}. The hybrid adaptive neuro-fuzzy system–gradient-based optimization algorithm is, in fact, a new and novel modeling paradigm, and for the first time, these approaches have been applied for forecasting wastewater treatment plant variables. In fact, this is also the first reported study that has used GBO and ANFIS models together to predict TN_{inf} data. The data-driven models are more accurate, reliable, and have greater predictive capability than the standalone models. The following are the overall contributions from this paper:

1. Preprocessing techniques, such as the normalization technique in the case of ANFIS models, help to provide better prediction results.
2. There are no existing studies in the literature that compare ANFIS, ANFIS-PSO, and ANFIS-GWO methods for predicting TN_{inf} from sewage plants.
3. No available studies in the literature evaluate daily TN predictions using both ANFIS-GBO and ANFIS-GWO models.
4. Hybrid models provide better prediction results than the standalone models. The hybrid of GBO with ANFIS is the first application for forecasting TN.

The purpose of this study is to better understand how different hybrid approaches may improve influent TN prediction performance in WWTPs. The following are the paper's particular goals: (i) to define various input combinations based on a review of the literature; (ii) develop ANFIS-based hybrid models by comparing input combinations using algorithms like PSO, GWO, and GBO to determine the optimal TN forecasting model; and (iii) assess the viability of utilizing an evolutionary algorithm GBO as a new ML model for TN prediction in comparison to alternative approaches (ANFIS, PSO, and GWO).

2. Materials and Methods

2.1. Utilized Data and Study Area

The Dkai plant is located on No. 6 Huanghai West Road, Jinzhou New District, Dalian City. It is a municipal sewage treatment plant that treats domestic and industrial wastewater (industrial wastewater exceeds 40%). It mainly accepts the municipal management of Maqiaozi Street, Dalian Economic and Technological Development Zone. Sewage from the network extends to Northeast 4th Street in the east, Tieshan West Road in the north, Binhai Tourist Road in the south, and Zhenxing Road in the west within about 23 km^2. After treatment, it is discharged into the Yellow Sea. The built-up scale is 100,000 tons/day, and the total area is 70,774 square meters. The franchise period is from 21 August 2011 to 20 August 2041, spanning 30 years. This factory was built in 1988, more than 30 years ago, with a scale of 80,000 tons per day, and the design water discharge standard is secondary. In 2010, Henderson took over the expansion, upgrade, and transformation of the factory, and the scale was expanded to 100,000 tons/day; the water discharge standard was Class A, and the budget audit value was CNY 132 million. After upgrading in 2010, the process flow chart shown in Figure 1 was finally formed. According to Chinese authority, the Class A discharge standard volumes of pollutants for municipal wastewater treatment plants in terms of COD, BOD, SSs, TP, and TN are 50, 10, 10, 5 and 15 mg/L, respectively. The average daily treatment capacity is 63,000–66,000 tons/day. In addition to domestic sewage, there are 14 major sewage enterprises, with discharges ranging from 200 to 5000 m^3/d, which are mainly electrical, machinery, building materials, pharmaceuticals, food, special glass, and dyeing processes. In this study, the influent TN parameter is selected to be predicted using other influential variables. The precise prediction of the influent total nitrogen variable is crucial and is selected as a case study because TN from the influent of WWTPs plays a significant role in the performance of nutrient removal systems, controlling sludge production, and the operation of different parts of wastewater treatment processes.

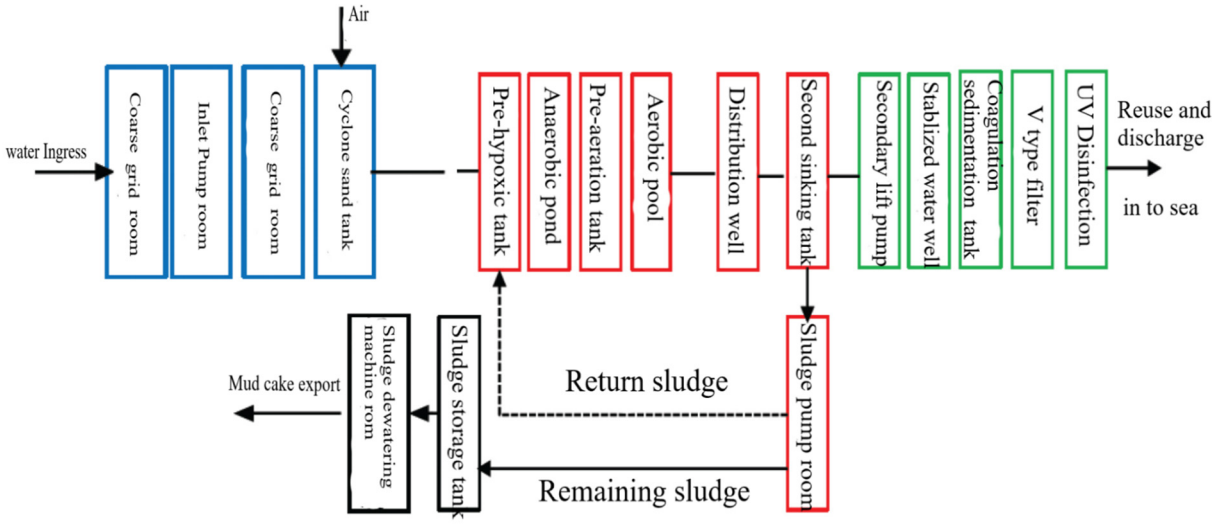

Figure 1. Process flow chart of DK sewage treatment plant.

The daily data for this study were obtained from the Dkai plant. The daily measured dataset comprised 7 parameters of influents in the input layer, namely Q_w, $PH_{influent}$, $SS_{influent}$, total phosphorus ($TP_{influent}$), NH_3-$N_{influent}$, $COD_{influent}$, and $BOD_{influent}$, and the output parameter was $TN_{influent}$. To develop an ANFIS-GBO model for estimating TN_{inf}, the available dataset (730 daily water quality data from the year 2017–2018) was used. The dataset was split into two sections, i.e., training and the testing datasets, for the application of models. For selecting suitable training and testing datasets, initially, different splitting strategies were adopted (50–50, 60–40, 70–30, 75–25 and 80–20%). After initial model applications, it was found that 75–25% of the training–testing dataset partition provided

better results. Therefore, this partition was later adopted in our study to apply all model and data input combinations. Statistical properties of the water quality parameters (mean, max, min, skewness, and standard deviation) used in this study are shown in Table 1. Water discharge (Q_w) has high skewness especially during training.

Table 1. Statistical properties of the water quality parameters used in the study.

Dataset	Statistics	Qw	PH	SS	TP	NH3-N	COD	BOD	TN
Whole dataset	Mean	64,123	7.40	146	5.79	47.4	564	153	54.8
	Min.	32,954	7.00	103	2.04	12.0	313	107	20.0
	Max.	106,660	7.80	195	10.8	67.4	931	205	75.0
	Skewness	0.708	−0.337	−0.131	0.260	−0.371	0.121	−0.293	−0.558
	Std.dev.	9204.6	0.124	10.9	1.11	6.85	93.7	18.1	6.55
Training	Mean	63,949	7.41	147	5.87	47.3	575	155	54.8
	Min.	32,954	7.00	103	2.43	23.3	313	109	28.0
	Max.	103,136	7.80	195	10.8	67.4	931	205	75.0
	Skewness	0.446	−0.337	−0.106	0.389	0.061	0.004	−0.306	−0.206
	Std.dev.	9144.9	0.126	11.1	1.16	6.50	98.8	18.1	6.26
Testing	Mean	64,665	7.38	143	5.53	47.5	529	149	54.7
	Min.	43,973	7.00	106	2.04	12.0	332	107	20.0
	Max.	106,660	7.60	175	7.36	64.3	690	186	70.0
	Skewness	1.475	−0.440	−0.486	−1.123	−1.150	−0.585	−0.388	−1.219
	Std.dev.	9367.9	0.116	9.9	0.90	7.84	64.2	17.0	7.37

2.2. Data Normalization

One important preprocessing step in dealing with machine learning models is the data normalization of input parameters. It guarantees that every feature makes an equal contribution to the learning process of the model and enhances the model's performance.

At the initial stage before the model's training, it is necessary to normalize input and output data, ensuring they fall within the 0 to 1 range [45,46]:

$$X_i = \frac{x_u - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where 'X_i' represents the normalized data, 'x_u' i denotes the observed data, and 'x_{min}' and 'x_{max}' refer to the minimum and maximum values in the dataset, respectively. Statistical examination of input–output parameters is a fundamental aspect of AI modeling, as it determines the type and strength of interactions between inputs and outputs.

2.3. Adaptive Neuro-Fuzzy Inference System

Adaptive network-based fuzzy inference system (ANFIS) is a fuzzy system within a network consisting of five layers that is Sugeno-type [47,48]. One intriguing feature of the ANFIS is that a Sugeno model can map any non-linear function, provided there is no restriction on the number of rules [49]. In the multilayer feedforward network known as ANFIS, every node processes incoming signals according to a specific purpose. To obtain a desired input–output mapping, the parameters associated with these nodes are updated based on a provided training set and a gradient-based learning technique [49–52]. The ANFIS model can benefit from the advantages of both in a single framework because it can include both ANNs and fuzzy logic concepts. According to Kisi et al. [53], its inference system is a collection of fuzzy If-Then rules that can precisely estimate non-linear functions. ANFIS architecture with two input variables (x, y) and one output (f) is conceptually described inside the five layers depicted in Figure 2 to explain the general principles of the ANFIS model. As a typical rule set containing two fuzzy If-Then rules, the first-order Sugeno fuzzy approach is expressed as follows:

$$\text{Rule1}: \text{ If } x \text{ is A1 and } y \text{ is B1, Then } f_1 = p_1 x + q_1 y + r_1 \quad (2)$$

$$\text{Rule1}: \text{ If } x \text{ is A1 and } y \text{ is B2, Then } f_2 = p_2x + q_2x + r_2 \quad (3)$$

In which A1, A2 and B1, B2 are the membership functions (MFs) for x and y vectors, respectively, and $p1$, $q1$, $r1$ and $p2$, $q2$, $r2$ are the output function parameters. The neural network structure contains five layers, excluding the input layer (Layer 0):

1. The input layer

Layer 0, the input layer, has n nodes where n is the number of inputs to the system.

2. The fuzzification layer

Layer 1 is the fuzzification layer in which each node represents a membership value to a linguistic term as a Gaussian function with the mean:

$$\mu Ai(x) = \frac{1}{1 + \left[\frac{x - ci}{ai}\right]^{2bi}} \quad (4)$$

where the function's parameters are ai, bi, and ci. These settings are adaptable. During the learning phase, the back-propagation technique is used to modify their values. The membership function for the linguistic term *'Ai'* varies in accordance with the values of the parameters.

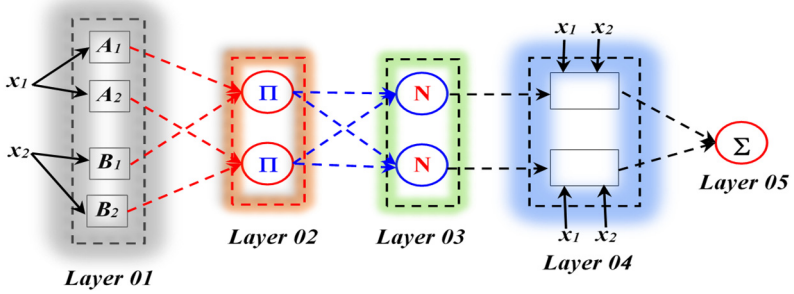

Figure 2. ANFIS structure diagram.

3. The Rules Layer

Using a multiplication operator, every node in Layer 2 contributes to the rule's strength. Min (AND) operation is carried out by it. For each variable in the antecedent component of Rule i, where x_0 and xi have linguistic values Ai and Bi, respectively, the rule's firing strength is calculated by the product of membership values μAi (x_0) and μBi (x_1). The quantity of Layer 2 rules is shown through the p^n nodes. If a rule is an If-Then rule, then each node represents the rule's antecedent component. The number of input variables in this case is n, and the number of membership functions is p.

$$Wi = \mu Ai(x_0) \times \mu Bi(x_1) \quad (5)$$

4. The Standardization Layer

The third layer is the normalization layer, which uses the following equation to equalize the strength of each rule:

$$\overline{W}i = \frac{Wi}{\sum_{j=1}^{R} Wj} \quad (6)$$

where wi, calculated in Layer 2, is the ith rule's degree of firing. Node i computes the ratio of its rule's firing strength to the total firing strength of all rules. This layer has p^n nodes. The layer of adaptable nodes is called the 4th layer. Each node with this layer generates a linear function, and the multilayer feedforward neural network's error function is used to modify the function coefficients:

$$\overline{W}ifi = \overline{W}i(p_{0 \times 0} + p_{1 \times 1} + p_2) \quad (7)$$

The parameters are denoted as pi's, where $i = n + 1$ represents the input numbers into the system, or the number of nodes in Layer 0. Two inputs were provided for this equation. Last but not least, Layer 3's output is $\overline{w}i$. A learning step is applied to update these settings. ANFIS takes advantage of the approximation of the least squares. The back-propagation technique is used in temporal models to train them.

5. The Output Layer or Defuzzification Layer

Layer 5 is the output layer, the function of which is the summation of net outputs of the nodes in Layer 4. The output is computed as

$$\sum_i \overline{w}ifi = \frac{\sum_i wifi}{\sum_i wi} \tag{8}$$

The output of node i in Layer 4 is represented by $\overline{w}ifi$. This signifies the portion of Rule I that follows. All the rule consequences added together are the neuro-fuzzy system's total output.

5.1. Particle Swarm Optimization

The PSO algorithm, inspired by Eberhart and Kennedy's [54] stochastic optimization technique, models the social behavior of insects and animal herds in search of food. It combines the evolutionary algorithm, which explores a broad solution space, with the synthetic life algorithm, which mimics life-like behaviors. PSO particles update their positions based on five key principles of artificial life: proximity, responsiveness to change, unrestricted movement, stability in dynamic environments, and adaptability [55–57]. These concepts guide the cooperative behavior of swarms in the algorithm.

The fourth and fifth hypotheses are opposing viewpoints on the same issue. These five concepts represent the main elements of artificial life systems and serve as the guiding principles for the artificial life system of the swarm. Particles in PSO can alter their locations and velocities according to proximity and efficiency demands influenced by environmental changes. Moreover, the PSO swarm does not have movement constraints, but rather seeks the optimum solution. PSO particles maintain constant motions by responding to changes in the surroundings. As a result, the PSO networks adhere to the five notions described above.

The basic calculations are as follows:

$$v_{ij}(t+1) = v_{ij}\,(t) + c_1 r_1\,(P_{ij}(t) - X_{ij}(t)) + c_2 r_2\,(P_{gj}(t) - X_{gj}(t)) \tag{9}$$

$$X_{ij}(t+1) = X_{ij}(t) + v_{ij}(t+1) \tag{10}$$

where c_1 and c_2 serve as acceleration constants, r_1 and r_2 are uniformly distributed random variables within the interval (0, 1), and t is the number of repetitions. The particle swarm's location across the space of likely solutions is represented by X_i. It is possible to vary the particle's position X_{i0} and velocity V_{i0} freely and then dynamically depending on how well it has performed in the past. The ideal outcome that particles seek within the global search space is the p_g, but the particle's optimal local location is p_i.

5.2. Grey Wolf Optimizer (GWO)

GWO is a recent swarm intelligence algorithm proposed by Mirjalili et al. [58], whose approaches mimic the social structure and hunting techniques of the gray wolf (GW) herd. GW, which includes α, β, δ, and ω, are strictly hierarchical social animals (Figure 3). To achieve global optimization, the group head is responsible for assigning various tasks to individuals at varying levels. Because of its straightforward design, low requirement for parameter changes, and excellent performance, the GWO algorithm has found widespread use in optimization [59–63].

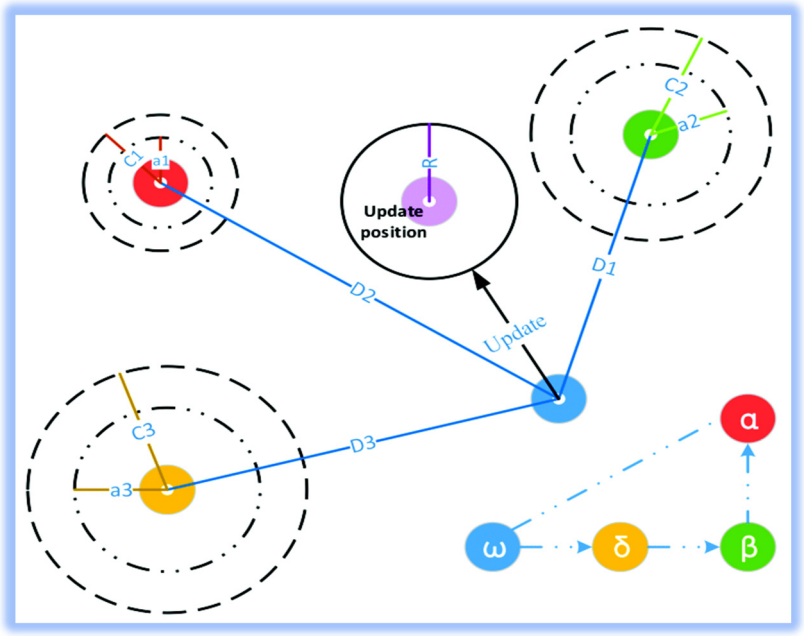

Figure 3. GWO model structure.

The location of the *i-th* wolf is represented by $X_i = X_1, X_2, \ldots, X_d$. X_d refers to the position of the *i-th* wolf in D-dimensional space within a collection of N gray wolves ($X = X_1, X_2, \ldots X_N$). The specific function of hunting is

$$\vec{D} = |\vec{C} \cdot \vec{X}_P(t) - \vec{X}(t)| \tag{11}$$

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A}.\vec{D} \tag{12}$$

where t indicates the current iteration, $\vec{A}$ and $\vec{c}$ are coefficient vectors, ~X_p is the position vector of the prey, and $\vec{X}_P$ indicates the position vector of a gray wolf. The vectors $\vec{A}$ and $\vec{c}$ are calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \tag{13}$$

$$\vec{C} = 2.\vec{r}_2 \tag{14}$$

where components of $\vec{a}$ are linearly decreased from 2 to 0 throughout iterations and r_1, r_2 are random vectors in [0, 1]. For all algorithms in this investigation, the population size, iteration count, and run count were 30, 150, and 10, respectively.

GW is quite promising in terms of food search. As the leader of all operations, α frequently has the participation of β and δ. Although β and δ can offer advice or suggest a potential solution, α is designated in the GWO as the ideal answer. Consequently, the best answers are α, β and δ. It is possible to express the solutions' updated positions as follows:

$$\vec{D}\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \tag{15}$$

$$\vec{X}_1 = \vec{X}_2 - \vec{A}_1 \cdot (\vec{D}x), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}\delta) \tag{16}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{17}$$

where D_α, D_β, and D_δ define the spatial distances between the prey's position and α, β, and δ; $X(t + 1)$ indicates the vector of position; A and C refer to the random vectors; and X_α, X_β, and X_δ denote the current locations of α, β, δ, etc.

5.3. Gradient-Based Optimizer (GBO)

The search route for domain exploration is created using a set of vectors using the Newton method. According to [41], the suggested GBO is mainly driven by population volume and gradient-based approaches. The optimization challenges involve reducing the objective function's value (Figure 4).

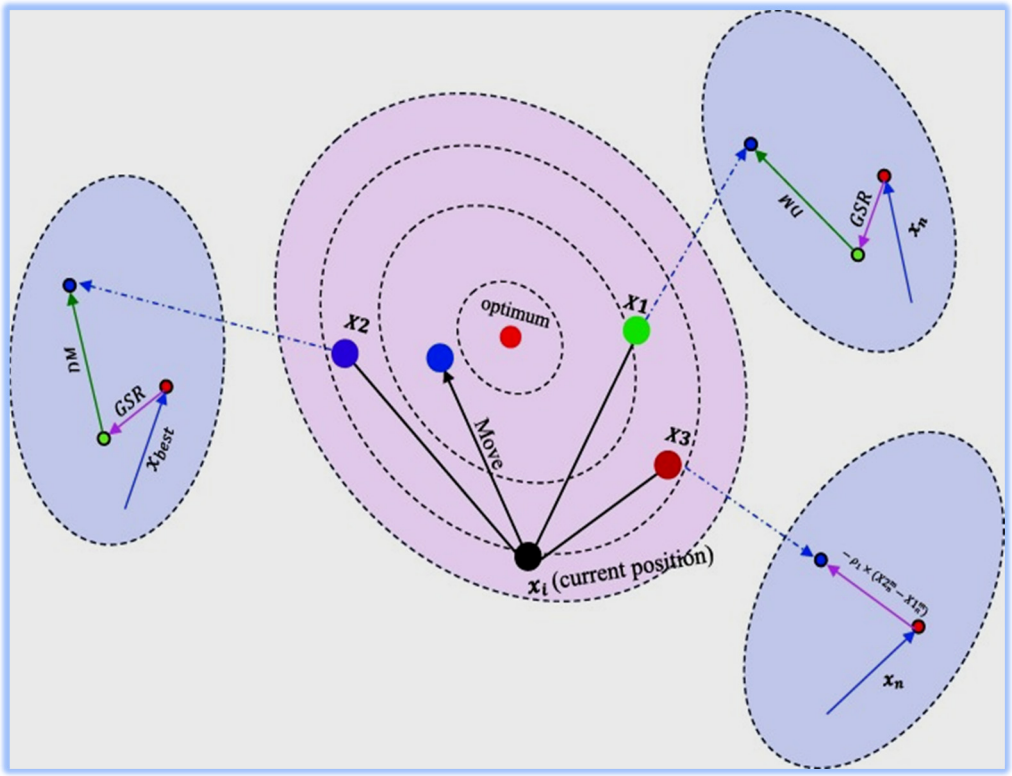

Figure 4. Sketch map of the GBO algorithm.

5.3.1. Initialization

Optimization involves a set of constraints, objective functions, and decision variables. The GBO has two control parameters: the transition parameter (∝) that governs the shift from discovering to exploiting, and the likelihood rate. The population scale and the number of iterations needed are determined according to how complex the problem is. The proposed approach refers to each member of the population as a "vector." Thus, in a D-dimensional search space, the N-vectors make up the GBO. Thus, a vector can be explained as follows:

$$X_{n,d} = [\, X_{n-1}, X_{n-2}, \ldots\ldots\ldots X_{n,D}] \tag{18}$$

$$d = 1, 2, \ldots, D \tag{19}$$

The initial GBO vectors in the D-dimensional search region are typically produced randomly, which may be described as

$$X_n = X_{min} + rand(0,\ 1) \times (X_{max} - X_{min}) \tag{20}$$

where the limits for the decision variables X are given by X_{min} and X_{max}, with rand (0, 1) representing a random value.

5.3.2. Gradient Search Rule

The gradient search rule (CSR) manages the mobility of the vectors to improve locations and facilitate feasible space exploration. The GSR is recommended to strengthen the GBO's convergence and improve the discovery pattern using the GB process principle. Ypma et al. [64] note that Newton's gradient-based method is the source of this methodology. For deriving the non-distinguishable optimization problems, a numerical gradient method might be used.

A projected initial solution is usually the starting point of the GB technique, which proceeds into the next site following a gradient-specified path. The GSR is derived from the first-order derivative by employing the Taylor series, which may be expressed as follows for functions $(x + \Delta x)$ and $(x - \Delta x)$ [64]:

$$f(\mathrm{x}+\Delta \mathrm{x}) = f(\mathrm{x}) + f'(\mathrm{x}_0)\ \Delta \mathrm{x} + \frac{f''(x_0)\Delta x^2}{2!} + \frac{f^3(x_0)\Delta x^3}{3!} + \ldots \tag{21}$$

$$f(\mathrm{x}-\Delta \mathrm{x}) = \mathrm{f}(\mathrm{x}) - f'(\mathrm{x}_0)\Delta \mathrm{x} + \frac{f''(x_0)\Delta x^2}{2!} - \frac{f^3(x_0)\Delta x^3}{3!} \tag{22}$$

$$f'(\mathrm{x}) = \frac{f(\mathrm{x}+\Delta \mathrm{x}) - f(\mathrm{x}-\Delta \mathrm{x})}{2\Delta \mathrm{x}} \tag{23}$$

$$x_{n+1} = x_n - \frac{2\Delta x \times f(x_n)}{f(x_n+\Delta x) - f(x_n-\Delta x)} \tag{24}$$

A few adjustments are required to control the search because the GSR is the method's core component. Considering Equations (3)–(18), $x_n + \Delta x$ and $x_n - \Delta x$ are denoted by the adjacent positions of x_n. Within the GBO algorithm, position $x_n + \Delta x$ is less fit than $\mathrm{x_n}$, while $x_n - \Delta x$ is more fit than x_n. A stronger location within the positional vicinity (x_n) is substituted for the position by the $x_n - \Delta x$.

In contrast, $\mathrm{x_{worst}}$, which denotes a worse condition close to x_n, is used instead of $x_n + \Delta x$. The following is an explanation of the proposed GSR:

$$\mathrm{GSR} = rand\ n \times p1 \times \frac{2\Delta x \times x_n}{(x_{worst} - x_{best} + \varepsilon)} \tag{25}$$

where ε is a fraction in the interval [0, 0.1] and randn is a random number. The $\mathrm{x_{worst}}$ and $\mathrm{x_{best}}$ consider the lowest and most significant results during optimization. In the following stage, GBO uses GSR to find a new positional solution (x_n^m) to replace the existing one:

$$x_n^{m+1} = r_a \times (r_b \times X1_n^m + (1 - r_b) \times X2_n^m + (1 - r_a) \times X3_n^m \tag{26}$$

The updated current location solutions of x_n^m are $X1_n^m$; $X2_n^m$ and $X3_n^n$, respectively.

5.3.3. Local Escaping Operator (LEO)

GBO is a tool used by LEO to explain complex phenomena. It can significantly alter the location of the solution x_n^{m+1} [41]. LEO applies several solutions, including the best position $(\mathrm{x_{best}})$, the solutions $(x1_n^m)$ and $(x2_n^m)$, and the two random values $\mathrm{x_{r_1}^m}$ and $(x_{r_2}^m)$, as well as a newly developed solution x_k^m. This results in more accurate outcomes (x_{LEO}^m) as stated below:

$$X_{LEO}^m = \begin{Bmatrix} if\ rand < 0.5 \\ x_n^{m+1} + f1 \times \left(u1 \times x_{best} - u2 \times x_k^m\right) + f2 \times \rho 1 \times \frac{u3 \times (x2_n^m - x1_n^m) + u2 \times \left(x_{r1}^m - x_{r2}^m\right)}{2} \\ +u2 \times \left(x_{r1}^m - x_{r2}^m\right))/2 \\ otherwise \\ x_{best} + f1 \times \left(u1 \times x_{best} - u2 \times x_k^m\right) + f2 \times \rho 1 \times (u3 \times (x2_n^m - x1_n^m) \\ +u2 \times \left(x_{r1}^m - x_{r2}^m\right))/2 \end{Bmatrix} \tag{27}$$

In this instance, *pr* reflects the probability, while $u1$, $u2$, and $u3$ are three randomly generated numbers. $f1$ denotes the standardized random number within the range of -1 to 1, and $f2$ refers to a random number drawn from a normal distribution with a mean of 0 and a standard deviation of 1.

5.4. GBO-Based ANFIS Parameter Optimization

The suggested model, ANFIS-GBO, is presented in this section. It optimizes ANFIS's control settings by using the GBO algorithm. The model's accuracy is contingent upon the degree to which the ANFIS parameter tweaking is executed optimally. Thus, choosing the best ANFIS parameters is the most important part of the ANFIS-GBO algorithm. The following summarizes the ANFIS-GBO's primary stages [44]:

1. Determine the factors that affect the problem's input–output dynamics that are being studied. Assign training and testing sets to the input and the output datasets.
2. To generate the FIS, derive a collection of rules that characterize the whole data using the FCM clustering technique. The dimensions of each search agent are specified by deriving the parameters for ANFIS, including those of the membership function. By following these procedures, you may employ GBO to train and tune the FIS.

- First stage: Create the starting MF parameter population.
- Second stage: Using the formula in Equation (28), determine the fitness value of every search agent in the population.
- Use the GBO algorithm to update the search agent's position in stage three.
- Fourth Stage: Proceed to stage 2 if the stoppage condition is not fulfilled; otherwise, terminate and deliver the optimal ANFIS settings. Use the best answers to anticipate the output during the testing phase when the training is over. Evaluate ANFIS-GBO's effectiveness using performance measures. Compare ANFIS-GBO's performance against that of ANFIS-GWO, ANFIS-PSO, and ANFIS. Four statistical indicators (Equations (28)–(31)) are used to numerically assess ANFIS-GBO performance:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(Q_0)_i - (Q_C)_i]^2} \tag{28}$$

$$\text{MAE(MeanAbsoluteError)} = \frac{1}{N}\sum_{i=1}^{N}|(Q_0)_i - (Q_C)_i| \tag{29}$$

$$\text{NSE (Nash–Sutcliffe Efficiency)} = 1 - \frac{\sum_{i=1}^{N}[(Q_0)_i - (Q_C)_i]^2}{\sum_{i=1}^{N}\left[(Q_0)_i - (\overline{Q}_0)\right]^2} \tag{30}$$

$$R^2\ \text{(Coefficient of Determination)} = 1 - \frac{\sum_{i=1}^{N}\left[((Q_0)_i - (\overline{Q}_0))((Q_C)_i - (\overline{Q}_C))\right]}{\sum_{i=1}^{N}\left[(Q_C)_i - (\overline{Q}_C)\right]^2\sum_{i=1}^{N}\left[(Q_0)_i - (\overline{Q}_0)\right]^2} \tag{31}$$

where N is the total number of data points and Q_0, Q_c, $\overline{Q}_0$ represent observed TN, estimated TN, and mean of TN, respectively. The primary statistics that are frequently used to assess the performance of wastewater treatment plants are these four statistical indices [65].

6. Results and Discussion

6.1. Input Combinations

The first part of this study focuses on input combinations for predicting inlet TN. Table 2 shows seven different input combinations and then four models (ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO) applied on each combination to determine which input combination gives good prediction results. The first combination consists of one input water discharge, Q_w. The second combination shows two inputs: water discharge (Q_w) and pH. The third combination includes three inputs: Q_w, pH, and suspended solids (SSs). The fourth combination has four inputs Q_w, pH, SSs, and total phosphorus (TP). Five inputs, Q_w, pH, SSs, TP, and ammonia nitrogen (NH_3-N), are used in the fifth combination. Six inputs, Qw, pH, SSs, TP, NH_3-N, and chemical oxygen demand (COD), are present in

the sixth combination, and the seventh combination shows seven inputs: Q_w, pH, SSs, TP, NH_3-N, and biological oxygen command (BOD) (see Table 2).

Table 2. Different input combinations used for the prediction of sediment loads.

	Input Combinations	Models			
		ANFIS	**ANFIS-PSO**	**ANFIS-GWO**	**ANFIS-GBO**
1	Q_w	ANFIS-1	ANFIS-PSO-1	ANFIS-GWO-1	ANFIS-GBO-1
2	Q_w, pH	ANFIS-2	ANFIS-PSO-2	ANFIS-GWO-2	ANFIS-GBO-2
3	Q_w, pH, SS	ANFIS-3	ANFIS-PSO-3	ANFIS-GWO-3	ANFIS-GBO-3
4	Q_w, pH, SS, TP	ANFIS-4	ANFIS-PSO-4	ANFIS-GWO-4	ANFIS-GBO-4
5	Q_w, pH, SS, TP, NH_3-N	ANFIS-5	ANFIS-PSO-5	ANFIS-GWO-5	ANFIS-GBO-5
6	Q_w, pH, SS, TP, NH_3-N, COD	ANFIS-6	ANFIS-PSO-6	ANFIS-GWO-6	ANFIS-GBO-6
7	Q_w, pH, SS, TP, NH_3-N, COD, BOD	ANFIS-7	ANFIS-PSO-7	ANFIS-GWO-7	ANFIS-GBO-7

6.1.1. Prediction Results of ANFIS Model

Table 3 presents the training and testing results of the ANFIS model for various input combinations. This shows the performance of the ANFIS model concerning each input combination and explains the best input combinations that give satisfactory results of TN_{inf}. Table 3 clearly shows that input combinations (5), (6), and (7) outperformed input combinations (1–4) significantly. The ANFIS-7 model proved that input combination (7) displayed marginally superior performance compared to input combination (5) and (6). In the training and testing phases, this model with the seventh input combination generated the smallest RMSE (2.2442 and 3.4275) and MAE (1.7273 and 2.2720), as well as the greatest NSE (0.8722 and 0.7824) and R^2 (0.8722 and 0.7848).

Table 3. Training and test statistics of the models for TN prediction—ANFIS.

Input Combinations	Models	Training Period				Test Period			
	ANFIS	**RMSE**	**MAE**	**R^2**	**NSE**	**RMSE**	**MAE**	**R^2**	**NSE**
1	ANFIS-1	5.8576	4.4930	0.2793	0.2665	6.9199	5.4010	0.1284	0.1131
2	ANFIS-2	5.7675	4.3069	0.3631	0.3607	6.4672	5.1905	0.2359	0.2253
3	ANFIS-3	5.5218	4.2199	0.3996	0.3888	6.2589	4.8927	0.2819	0.2744
4	ANFIS-4	5.1630	3.9628	0.5659	0.5363	5.8786	4.4947	0.3795	0.3599
5	ANFIS-5	2.2944	1.8727	0.8672	0.8647	3.6263	2.3876	0.7673	0.7564
6	ANFIS-6	2.3585	1.9873	0.8567	0.8547	4.0166	2.6683	0.7239	0.7012
7	ANFIS-7	2.2442	1.7273	0.8722	0.8722	3.4275	2.2720	0.7848	0.7824

Table 3 illustrates that when parameters BOD, COD, and NH_3-N were utilized as inputs (the 5th to 7th input combinations), the ANFIS model's performance increased. The results of input combinations (4 and 7) reduced RMSE from 5.1630 to 2.2442 (56.53%) and MAE from 3.9628 to 1.7273 (56.41%), while enhanced NSE from 0.5363 to 0.8722 (62.63%) and R^2 from 0.5659 to 0.8722 (54.12%) in the training phase. For the testing stage, the RMSE decreased by 41.69%, from 5.8786 to 3.4275, and MAE by 49.45%, from 4.4947 to 2.2720. However, the NSE ascended from 0.3599 to 0.7824 and the coefficient of determination (R^2) from 0.3795 to 0.7848.

6.1.2. ANFIS-PSO Model Results

The performance of the ANFIS-PSO model for estimating TN_{inf} is shown in Table 4. After seeing the training and test results of Table 4, the results proved that ANFIS-PSO

models provide distinct predictions for various input combinations. Here, again, input combinations (1–4) showed weaker performance in comparison to input combinations (5th–7th). In this model, the 7th input combination (Q_w, pH, SS, TP, NH_3-N, COD, BOD) provided the lowest RMSE (2.1440 and 3.0034) and MAE (1.6184 and 1.8581), and the greatest NSE (0.8798 and 0.8329) and R^2 (0.8833 and 0.8514) for the training and testing stages. After adding some parameters (NH_3-N, COD, BOD) in the 7th combination, ANFIS-PSO's prediction accuracy increased.

Table 4. Training and test statistics of the models for TN prediction—ANFIS-PSO.

Input Combinations	Models	Training Period				Test Period			
	ANFIS-PSO	**RMSE**	**MAE**	**R^2**	**NSE**	**RMSE**	**MAE**	**R^2**	**NSE**
1	ANFIS-PSO-1	5.8387	4.4543	0.3129	0.3051	6.3512	4.7836	0.2719	0.2529
2	ANFIS-PSO-2	5.6782	4.3764	0.3817	0.3672	6.1846	4.6737	0.3206	0.2916
3	ANFIS-PSO-3	5.4716	4.2782	0.4362	0.4139	5.9539	4.5868	0.3761	0.3434
4	ANFIS-PSO-4	5.0034	3.9776	0.5742	0.5479	5.5387	4.3547	0.4562	0.4318
5	ANFIS-PSO-5	2.2117	1.6955	0.8758	0.8711	3.0626	1.8792	0.8385	0.8304
6	ANFIS-PSO-6	2.2944	1.7418	0.8705	0.8671	3.1513	2.0271	0.8307	0.8161
7	ANFIS-PSO-7	2.1440	1.6184	0.8833	0.8798	3.0034	1.8581	0.8514	0.8329

The 4th and 7th input combinations showed that RMSE decreased from 5.0034 to 2.1440 (57.14%) and MAE from 3.9776 to 1.6184 (59.31%), while NSE enhanced from 0.5479 to 0.8798 (60.57%) and R^2 from 0.5742 to 0.8833 (53.83%) during the training phase. During the testing stage, the value of RMSE decreased from 5.5387 to 3.0034 (45.77%) and MAE from 4.3547 to 1.8581 (53.81%), while the NSE enhanced from 0.4318 to 0.8329 (92.89%) and R^2 from 0.4562 to 0.8514 (86.63%). After including some parameters (NH_3-N, COD, BOD) in the fifth, sixth, and seventh input combinations, the efficiency of the ANFIS-PSO method was enhanced. However, for predicting influent total nitrogen from the sewage plant, the seventh input combination showed greater performance than the fifth and sixth combination.

6.1.3. ANFIS-GWO Model Results

The prediction results of the ANFIS-GWO hybrid model using different input combinations are provided in Table 5. For the 7th input combination (Q_w, pH, SS, TP, NH3-N, COD, BOD), the ANFIS-GWO model achieved the lowest RMSE values (2.0714 and 2.8772) and MAE values (1.5695 and 1.8030), along with the highest NSE (0.8828 and 0.8467) and R^2 (0.8876 and 0.8542) during the training and testing stages. Comparing the 4th and 7th input combinations, the ANFIS-GWO model's performance improved significantly: RMSE decreased from 4.8026 to 2.0714 (56.86%) and MAE from 3.7394 to 1.5695 (58.02%), while NSE increased from 0.5619 to 0.8828 (57.10%) and R^2 from 0.5881 to 0.8876 (50.92%) during training. In testing, RMSE dropped from 5.0426 to 2.8772 (42.94%), MAE from 3.8069 to 1.8030 (52.63%), and NSE improved from 0.5290 to 0.8467 (60.05%), with R^2 increasing from 0.5566 to 0.8542 (53.46%). The results of the 7th input combination showed a slight difference between PSO and GWO models.

6.1.4. ANFIS-GBO Model Results

Table 6 provides the training and testing results of the ANFIS-GBO model using different input combinations. The model using the 7th combination achieved the lowest RMSE values (1.7334 and 2.6810) and MAE values (1.3201 and 1.6713), along with the highest NSE (0.9224 and 0.8574) and R^2 (0.9249 and 0.8675) during training and testing, respectively. The ANFIS-GBO model performed significantly better with the 7th combination than the 4th. During training, RMSE dropped from 4.3951 to 1.7334 (60.5%), and MAE from 3.0178

to 1.3201 (56.25%), while NSE increased from 0.6038 to 0.9224 (52.7%) and R^2 from 0.6276 to 0.9249 (47.3%). In the testing phase, RMSE decreased from 5.0059 to 2.6810 (46.4%) and MAE from 3.7927 to 1.6713 (55.9%), while NSE rose from 0.5359 to 0.8574 (59.9%) and R^2 from 0.5644 to 0.8675 (53.7%).

Table 5. Training and test statistics of the models for TN prediction—ANFIS-GWO.

Input Combinations	Models	Training Period				Test Period			
	ANFIS-GWO	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE
1	ANFIS-GWO-1	5.7160	4.4034	0.3210	0.3082	6.2055	4.7184	0.2961	0.2868
2	ANFIS-GWO-2	4.9357	3.7871	0.3909	0.3817	6.1194	4.6451	0.3456	0.3358
3	ANFIS-GWO-3	5.2109	3.9294	0.4402	0.4240	5.9152	4.5361	0.4248	0.3990
4	ANFIS-GWO-4	4.8026	3.7394	0.5881	0.5619	5.0426	3.8069	0.5566	0.5290
5	ANFIS-GWO-5	2.1609	1.6339	0.8815	0.8762	2.9150	1.8556	0.8503	0.8426
6	ANFIS-GWO-6	2.2338	1.7131	0.8734	0.8717	3.0389	1.9188	0.8435	0.8290
7	ANFIS-GWO-7	2.0714	1.5695	0.8876	0.8828	2.8772	1.8030	0.8542	0.8467

Table 6. Training and test statistics of the models for TN prediction—ANFIS-GBO.

Input Combinations	Models	Training Period				TEST PERIOD			
	ANFIS-GBO	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE
1	ANFIS-GBO-1	5.3760	4.1532	0.3349	0.3135	6.0533	4.6611	0.3058	0.2957
2	ANFIS-GBO-2	5.5798	4.3196	0.3917	0.3708	5.9388	4.5565	0.3723	0.3468
3	ANFIS-GBO-3	4.9075	3.8129	0.4723	0.4662	5.7478	4.3072	0.4438	0.4164
4	ANFIS-GBO-4	4.3951	3.0178	0.6276	0.6038	5.0059	3.7927	0.5644	0.5359
5	ANFIS-GBO-5	1.9420	1.4455	0.9069	0.9043	2.8097	1.7127	0.8571	0.8538
6	ANFIS-GBO-6	2.1390	1.5408	0.8929	0.8839	2.8996	1.8062	0.8536	0.8443
7	ANFIS-GBO-7	1.7334	1.3201	0.9249	0.9224	2.6810	1.6713	0.8675	0.8574

Tables 5 and 6 show the testing outcomes for the most effective ANFIS-GWO and ANFIS-GBO methods. The performance of both models was robust for input combinations (5) to (7) and demonstrated slightly superior accuracy compared to the input combinations (1–4) across all input cases. Just like the ANFIS, ANFIS-PSO, and ANFIS-GWO models, the ANFIS-GBO model provided the worst performance for input combinations (1–4) due to the absence of BOD, COD, and NH_3-N parameters, which are most important for the predictions of TN. Input combination (7) produced the most effective ANFIS-GWO and ANFIS-GBO models. In four models, input combinations 5 to 7 achieved marginally improved performance compared to the input combinations 1 to 4. On average, input combination 7 produced marginally superior results compared to the input combinations 5 and 6 for all models.

The ANFIS-GBO model revealed the most outstanding performance for the estimation of TN_{inf} as compared to the ANFIS, ANFIS-PSO, and ANFIS-GWO models with Qw, pH, SSs, TP, NH3-N, COD, and BOD as inputs (see Tables 5 and 6).

6.2. Comparison of ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO Models

Comparing all the models (Tables 3–6) showed that the ANFIS-GBO method demonstrated superior performance relative to the other methods in daily influent total nitrogen predictions for DKWWTP. The RMSE values of the ANFIS-GBO-7 model were less than those of the different models in training and testing. The ANFIS-GBO-7 model achieved higher R^2 values than the other models for predicting the TNinf in the test stage. The

average MAE and RMSE values of the ANFIS-GBO model on the testing results were 17.57% and 14.91% greater than the average MAE and RMSE values of the ANFIS model, respectively. The average NSE and R^2 of ANFIS-GBO showed 29.17% and 32.26% greater efficiency than the ANFIS model. However, ANFIS-GBO was recognized as the most effective model for predicting the TN_{inf} in wastewater because it had the least MAE and the highest R^2 for the 7th input combination, which reveals that the ANFIS-GBO had superior overall performance compared to the other three methods (see Tables 3–6).

The time variation graphs of all ANFIS models are shown in Figure 4 for the test period using the best input combinations. It is evident from the time variation graph (Figure 5) that the ANFIS-GBO estimates closely follow the corresponding observed values. Scatterplot comparisons of the implemented methods are presented in Figure 6 for the best input combinations. Scatter plots show that input combination 7 (Figure 6) has a high R^2 for all hybrid ANFIS models. It is clear from the graphs of Figure 6 that the ANFIS-GBO has fewer scattered estimates with a higher R^2 (0.8675) value and its fit line equation approaches the precise line (y = x) more closely.

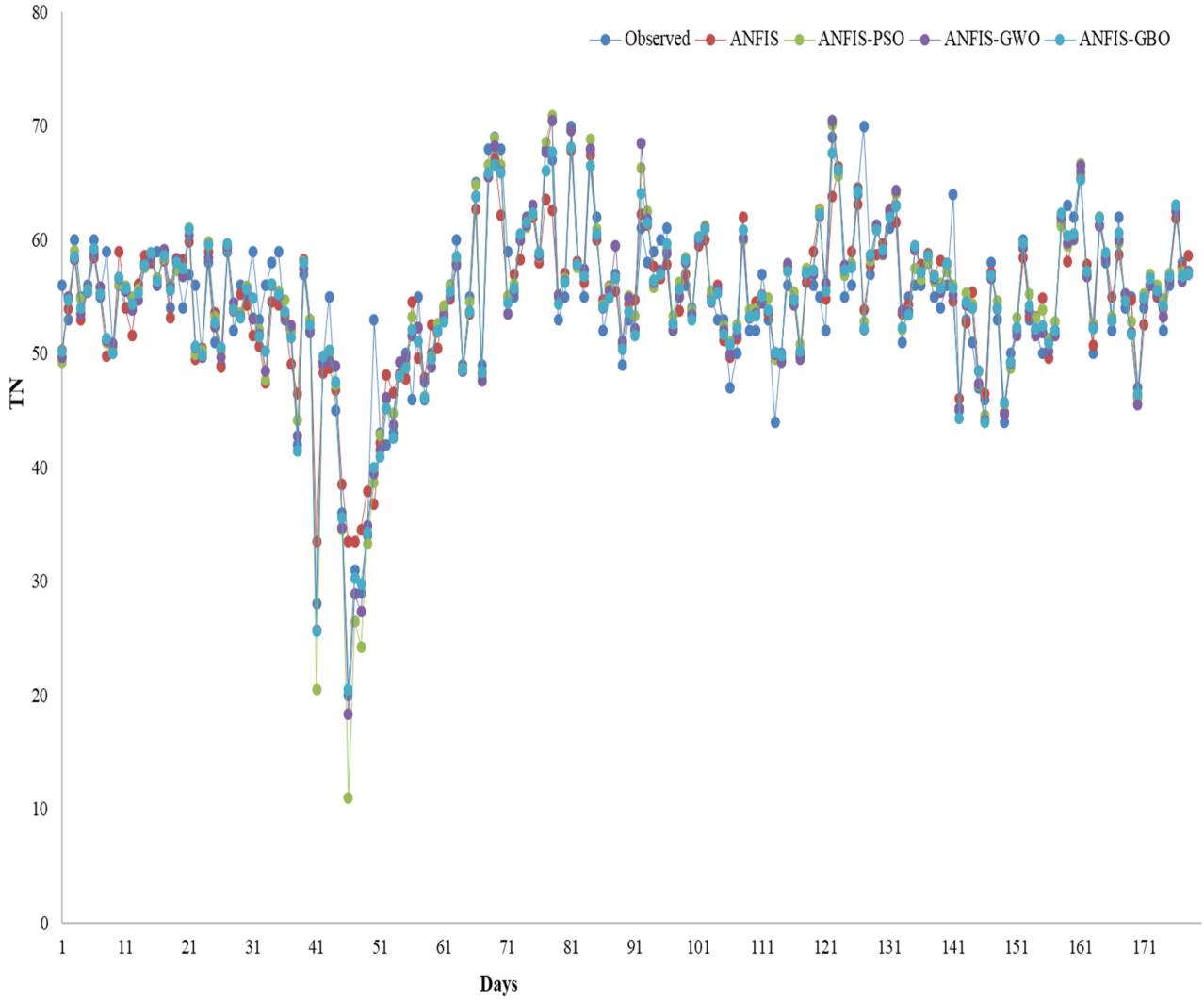

Figure 5. Time variation graphs of the observed and predicted TN by different ANFIS-based models in the test period using the best input combination: 7.

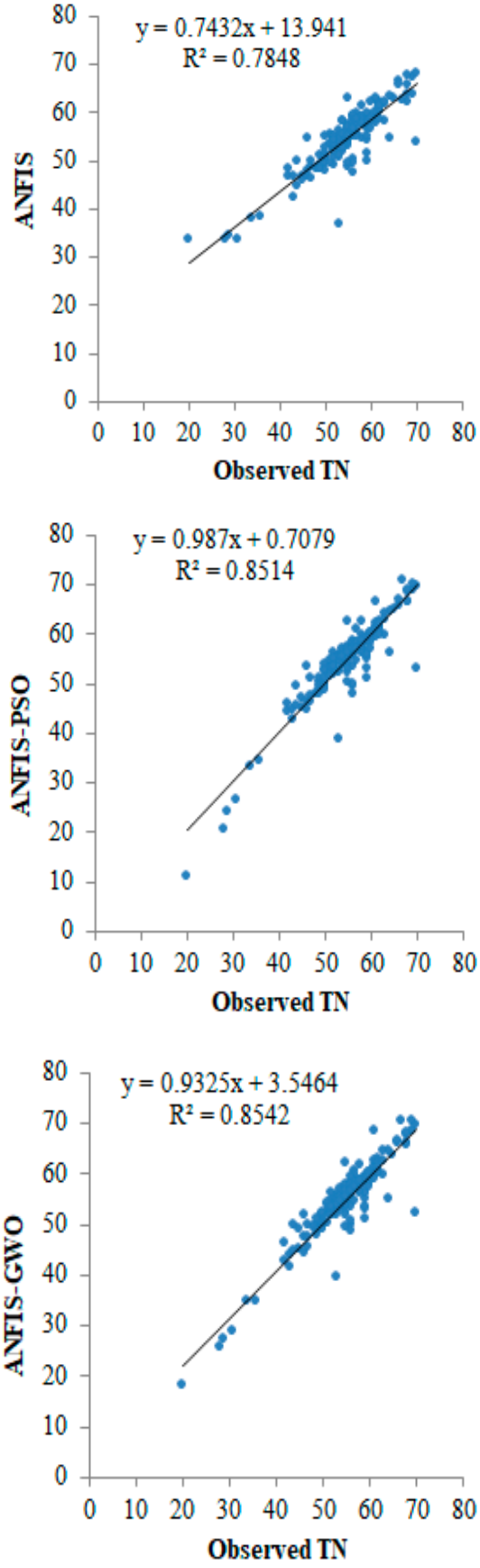

Figure 6. *Cont.*

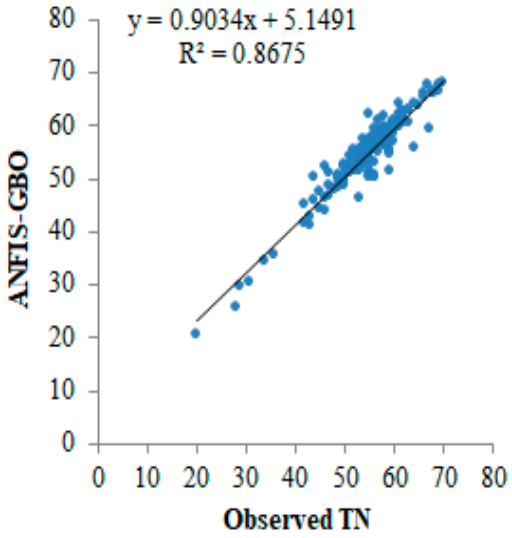

Figure 6. Scatterplots of the observed and predicted TN by different ANFIS-based models in the test period using the best input combination: 7.

A Taylor diagram of the predicted TN by various ANFIS-based methods during the testing stage with the best input combinations is shown in Figure 7. This graphical technique, which displays three statistics, i.e., correlation coefficient, standard deviation, and root-mean-squared error, on a single graph is beneficial for evaluating model performance. It is clear from the diagram that the 7th input shows good prediction results for all models. As can be observed from the Taylor diagrams for the 7th input case, the hybrid ANFIS-GBO has a lower standard deviation and RMSE to the observed one and a higher correlation than the other ANFIS-based methods. It was followed by the ANFIS-GBO, ANFIS-GWO, and ANFIS-PSO methods in terms of accuracy.

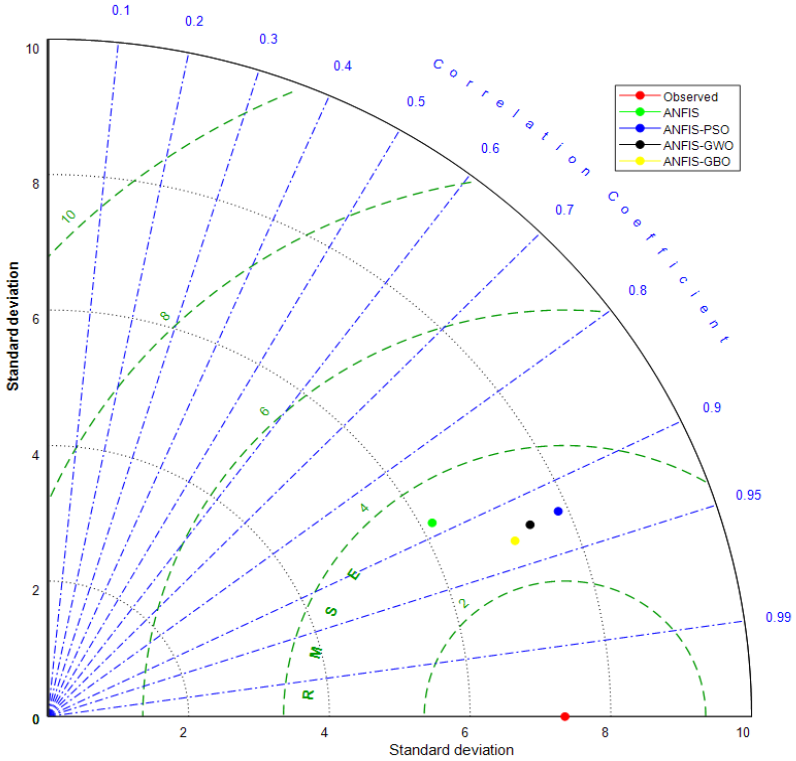

Figure 7. Taylor diagrams of the predicted TN by different ANFIS-based models in the test period using the best input combinations.

Figure 8 shows violin diagrams of each model for the best input combination. After comparing all the models, the suggested hybrid approach (ANFIS-GBO) showed supremacy on the 7th input combination by exhibiting a distribution closer to the actual value than the other input cases (1st to 6th). The results produced using the statistical measures shown in the preceding tables were validated by a graphical comparison. The hybrid ANFIS–GBO performed better than the standalone and hybrid ANFIS methods for estimating daily TNinf from the Dkai plant.

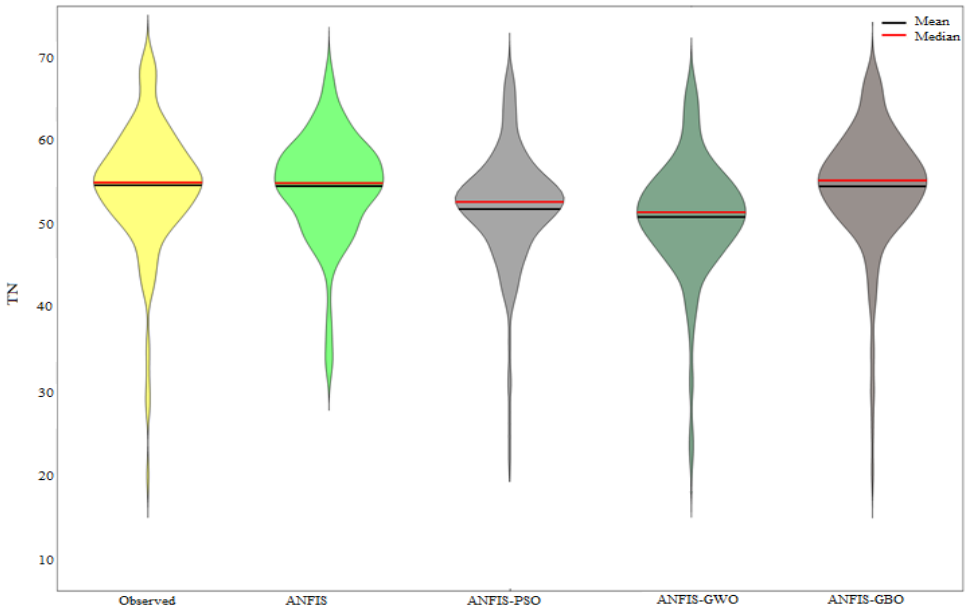

Figure 8. Violin charts of the predicted TN by different ANFIS-based models in the test period using the best input combinations.

Comparative plots are shown in Figure 9, which demonstrate the performance of ANFIS-based models for all input combinations using testing data. The figures show that hybrid ANFIS-GBO demonstrated greater performance over standalone and other hybrid ANFIS models using the 7th input combination for estimating the influent TN by having lower cumulative error percentage.

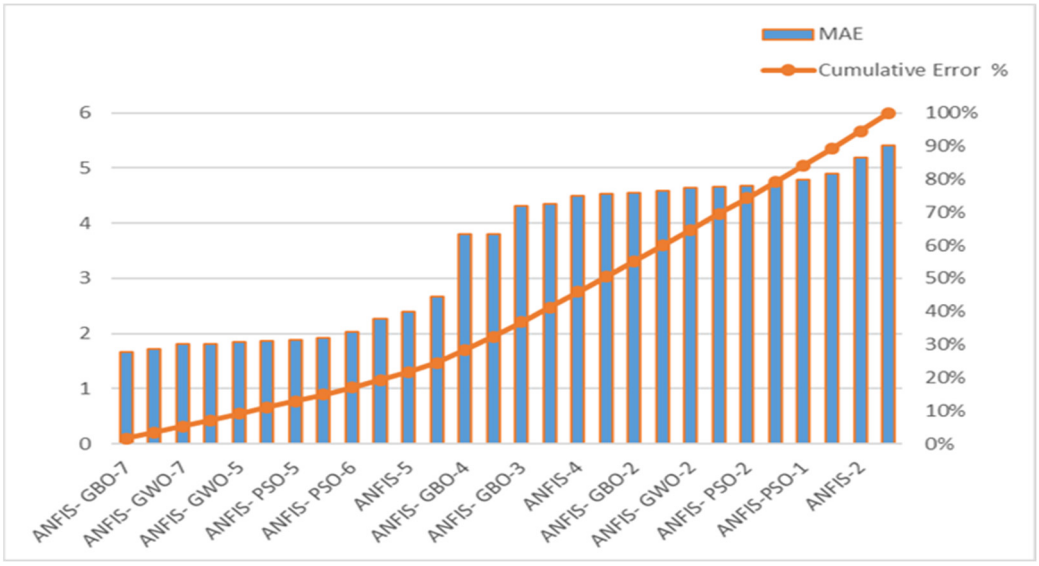

Figure 9. Pareto chart of all ANFIS-based models for the all input combinations using testing data.

6.3. Comparison with Previous Research

Many studies, like those by Pai et al. [32], Nadiri et al. [25], and Nourani et al. [30], have successfully used standalone ANFIS or hybrid ANFIS models in wastewater treatment plant modeling, focusing on parameters like COD, BOD, SSs, and pH. For example, Pai et al. [32] demonstrated that ANFIS models predicted effluent parameters like SSs, COD, and pH with low mean absolute percentage errors (2.67%, 2.80%, and 0.42%). However, the study's introduction of GBO optimization for TN prediction is unique, showing greater performance than models in the literature. In studies like Mahadeva et al. [35] and Akbaş et al. [33], ANFIS-PSO models have been effectively used for various wastewater-related predictions. For example, Mahadeva et al. [37] optimized ANN models for a desalination plant using PSO, achieving an R^2 value of 99.1%, and Sarkar et al. [36] used ANN-PSO for biosorption modeling, achieving high predictive accuracy. However, this study's ANFIS-GBO surpasses the performance of these PSO-optimized models, as shown by improved RMSE, MAE, and NSE values.

Studies such as Riahi-Madvar et al. [40] and Tikhamarine et al. [38] have applied GWO in conjunction with ANFIS for river flow forecasting and streamflow prediction. Riahi-Madvar et al. found that ANFIS-GWO outperformed other models like PSO and DE, demonstrating that GWO is an effective optimizer. However, the GBO-enhanced ANFIS in this study exhibits superior results, especially for predicting complex influent nitrogen parameters. In the literature, ANN and ANFIS models often rely on input variables such as pH, temperature, BOD, COD, and TSSs [20]. Many studies, like those by Nasr et al. [66] and Alsulaili et al. [22], also include COD and BOD as essential inputs for performance prediction. However, this study's focus on influent total nitrogen, coupled with its thorough exploration of input parameter combinations, adds a layer of sophistication not seen in many prior studies.

This study is the first to apply the gradient-based optimization (GBO) algorithm in combination with ANFIS for wastewater treatment plant modeling. The GBO algorithm offers a more accurate tuning of ANFIS hyper-parameters, resulting in enhanced prediction accuracy for complex non-linear systems like wastewater TN levels. While algorithms like PSO and GWO have been widely applied in wastewater modeling, the GBO is a relatively new metaheuristic optimization algorithm.

7. Conclusions

Early detection of fluctuating influent and effluent concentrations is essential for effectively operating and managing wastewater treatment plants. The presented study investigated the precision of a novel hybrid machine learning technique, ANFIS-GBO, in predicting the daily influent total nitrogen from the Dkai sewage plant. The standalone ANFIS and two hybrid ANFIS-PSO and ANFIS-GWO models using MAE, RMSE, R^2, NSE, and graphical techniques like time variation, violin, Taylor, and scatter diagrams were compared with ANFIS-GBO results. Different input combinations were used as inputs for the aforementioned models. The application and results lead to the following conclusions.

The ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO models were tested using seven input combinations for daily TN_{inf} prediction. The prediction results of testing and training periods showed that the ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO models displayed the highest performance for the seventh input scenario (7th) in the testing and training periods. The models whose inputs are Q_w, pH, SSs, TP, NH_3-N, COD_{cr}, and BOD5 have the best performance criteria among the input combinations tried in the study. It was found that all the models provided the best accuracy with input combination 7.

Assessment criteria and graphics agree that the ANFIS-GBO-7 method was more successful than the hybrid ANFIS-GWO-7 and ANFIS-PSO-7 and standalone ANFIS-7 methods in daily TN_{inf} prediction. The improvement in RMSE was 6.81, 10.73, and 21.77%, in the test stage, respectively. There was a slight variation between the hybrid ANFIS-PSO and ANFIS-GWO. However, these techniques outperformed the solo ANFIS. The RMSE improved by 12.37 and 16.05% throughout the test stage for the ANFIS-PSO and

ANFIS-GWO models. All models showed the greatest performance after adding some parameters (NH3-N, COD, and BOD) as inputs. The advancements in RMSE of ANFIS, ANFIS-PSO, ANFIS-GWO, and ANFIS-GBO are 41.69, 45.77, 42.9, and 46.44% in the test stage, respectively. The results of ANFIS-based models in estimating the daily TN_{inf} revealed that the new ANFIS-GBO-7 method presented the smallest mean absolute error (MAE) compared to other models. The improvements of MAE achieved by ANFIS-GBO-7 were 26.4, 10, and 7.3% in test stage for the ANFIS-7, ANFIS-PSO-7, and ANFIS-GWO-7 models, respectively. The research proposed a hybrid ANFIS-GBO approach in predicting daily TN_{inf} for the DKWWTP.

Following the successful testing of the ANFIS-GBO model, the next steps involve scaling the technology for broader wastewater treatment applications. This will include further pilot testing at multiple sites to validate the model's robustness, integration with real-time monitoring systems, and enhancement for computational efficiency to handle larger datasets. Additionally, the development of a user-friendly interface and operator training will facilitate practical adoption. Compliance with data security regulations and certification for regulatory standards will ensure the model's reliability. By building partnerships with industry stakeholders, this technology can be effectively commercialized, providing wastewater treatment facilities with a cost-effective solution for nitrogen management.

Using the ANFIS-GBO model to predict influent total nitrogen and other water quality parameters offers significant benefits for wastewater treatment plant operations. By providing accurate, real-time predictions, the model allows operators to optimize treatment processes proactively, which can lead to improved nutrient removal, reduced energy consumption, and cost savings. Additionally, predictive insights into water quality parameters enable better regulatory compliance and reduced risk of environmental pollution, contributing to more sustainable plant operations and better protection of surrounding ecosystems.

While the ANFIS-GBO model demonstrates promising accuracy in predicting influent total nitrogen, the study has certain limitations. First, the model is validated on data from a single wastewater treatment plant, which may limit its generalizability across facilities with significantly different operational parameters or environmental conditions. Additionally, the study focuses on specific input variables, and other potentially influential factors like dissolved oxygen (DO), total dissolved solids (TDSs), and total suspended solids (TSSs) are not considered. Future research should test the model across diverse treatment plants and include additional variables to further enhance prediction accuracy and ensure the model's broader applicability. In future studies, it is also recommended to use discrete wavelet transform, singular spectrum analysis, and empirical model decomposition preprocessing techniques for denoising the water quality data.

Author Contributions: Conceptualization, M.I. and H.L.; Methodology, W.M. and R.M.A.; Formal analysis, M.I., A.M.S.A.-J., O.K. and R.M.A.; Data curation, W.M.; Writing—original draft, M.I., O.K. and R.M.A.; Writing—review & editing, A.M.S.A.-J. and M.A.; Visualization, M.A.; Supervision, H.L. and W.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would also like to express their sincere appreciation to the associate editor and the anonymous reviewers for their comments and suggestions. This work was supported by the National Natural Science Foundation of China (52350410465), and the General Projects of Guangdong Natural Science Research Projects (2023A1515011520).

Data Availability Statement: The data presented in this study will be available on request from the corresponding author.

Conflicts of Interest: There are no conflicts of interest in this study.

References

1. Aydın Temel, F.; Özyazıcı, G.; Uslu, V.R.; Ardalı, Y. Full scale subsurface flow constructed wetlands for domestic wastewater treatment: 3 years' experience. *Environ. Prog. Sustain. Energy* **2018**, *37*, 1348–1360. [CrossRef]
2. Abba, S.I.; Elkiran, G.; Nourani, V. Improving novel extreme learning machine using PCA algorithms for multi-parametric modeling of the municipal wastewater treatment plant. *Desalin. Water Treat* **2021**, *215*, 414–426. [CrossRef]

3. Mahmoud, M.; Mossad, M.; Mahanna, H. Degradation of levofloxacin using electro coagulation residuals-alginate beads as a novel heterogeneous electro-fenton composite. *J. Environ. Manag.* **2024**, *359*, 120972. [CrossRef] [PubMed]
4. Gómez, T.; Gémar, G.; Molinos-Senante, M.; Sala-Garrido, R.; Caballero, R. Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach. *J. Clean. Prod.* **2017**, *164*, 315–324. [CrossRef]
5. Zhou, M.; Zhang, Y.; Wang, J.; Shi, Y.; Puig, V. Water quality indicator interval prediction in wastewater treatment process based on the improved BES-LSSVM algorithm. *Sensors* **2022**, *22*, 422. [CrossRef]
6. Ansari, M.; Othman, F.; El-Shafie, A. Optimized fuzzy inference system to enhance prediction accuracy for influent characteristics of a sewage treatment plant. *Sci. Total Environ.* **2020**, *722*, 137878. [CrossRef]
7. Wang, J.H.; Zhao, X.L.; Guo, Z.W.; Yan, P.; Gao, X.; Shen, Y.; Chen, Y.P. A full-view management method based on artificial neural networks for energy and material-savings in wastewater treatment plants. *Environ. Res.* **2022**, *211*, 113054. [CrossRef]
8. Vilela, P.; Safder, U.; Heo, S.; Nguyen, H.T.; Lim, J.Y.; Nam, K.; Oh, T.S.; Yoo, C. Dynamic calibration of process-wide partial-nitritation modeling with airlift granular for nitrogen removal in a full-scale wastewater treatment plant. *Chemosphere* **2022**, *305*, 135411. [CrossRef]
9. Salgot, M.; Folch, M. Wastewater treatment and water reuse. *Curr. Opin. Environ. Sci. Health* **2018**, *2*, 64–74. [CrossRef]
10. Karunanidhi, D.; Aravinthasamy, P.; Subramani, T.; Roy, P.D.; Srinivasamoorthy, K. Risk of fluoride-rich groundwater on human health: Remediation through managed aquifer recharge in a hard rock terrain, South India. *Nat. Resour. Res.* **2020**, *29*, 2369–2395. [CrossRef]
11. Hameed, M.; Sharqi, S.S.; Yaseen, Z.M.; Afan, H.A.; Hussain, A.; Elshafie, A. Application of artificial intelligence (AI) techniques in water quality index prediction: A case study in tropical region, Malaysia. *Neural Comput. Appl.* **2017**, *28*, 893–905. [CrossRef]
12. Gao, F.; Nan, J.; Zhang, X. Simulating a cyclic activated sludge system by employing a modified ASM3 model for wastewater treatment. *Bioprocess Biosyst. Eng.* **2017**, *40*, 877–890. [CrossRef] [PubMed]
13. Chen, Y.; Han, D. Water quality monitoring in smart city: A pilot project. *Autom. Constr.* **2018**, *89*, 307–316. [CrossRef]
14. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. II: Hydrologic applications. *J. Hydrol. Eng.* **2000**, *5*, 124–137. [CrossRef]
15. Çinar, Ö. New tool for evaluation of performance of wastewater treatment plant: Artificial neural network. *Process Biochem.* **2005**, *40*, 2980–2984. [CrossRef]
16. Fan, M.; Hu, J.; Cao, R.; Ruan, W.; Wei, X. A review on experimental design for pollutants removal in water treatment with the aid of artificial intelligence. *Chemosphere* **2018**, *200*, 330–343. [CrossRef]
17. Elkiran, G.; Nourani, V.; Abba, S.I. Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach. *J. Hydrol.* **2019**, *577*, 123962. [CrossRef]
18. Zhao, L.; Dai, T.; Qiao, Z.; Sun, P.; Hao, J.; Yang, Y. Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. *Process Saf. Environ. Prot.* **2020**, *133*, 169–182. [CrossRef]
19. Hamed, M.M.; Khalafallah, M.G.; Hassanien, E.A. Prediction of wastewater treatment plant performance using artificial neural networks. *Environ. Model. Softw.* **2004**, *19*, 919–928. [CrossRef]
20. Tumer, A.E.; Edebali, S. An artificial neural network model for wastewater treatment plant of Konya. *Int. J. Intell. Syst. Appl. Eng.* **2015**, *3*, 131–135. [CrossRef]
21. Arismendy, L.; Cárdenas, C.; Gómez, D.; Maturana, A.; Mejía, R.; Quintero, M.C.G. Intelligent system for the predictive analysis of an industrial wastewater treatment process. *Sustainability* **2020**, *12*, 6348. [CrossRef]
22. Alsulaili, A.; Refaie, A. Artificial neural network modeling approach for the prediction of five-day biological oxygen demand and wastewater treatment plant performance. *Water Supply* **2021**, *21*, 1861–1877. [CrossRef]
23. Yaqub, M.; Asif, H.; Kim, S.; Lee, W. Modeling of a full-scale sewage treatment plant to predict the nutrient removal efficiency using a long short-term memory (LSTM) neural network. *J. Water Process Eng.* **2020**, *37*, 101388. [CrossRef]
24. Hejabi, N.; Saghebian, S.M.; Aalami, M.T.; Nourani, V. Evaluation of the effluent quality parameters of wastewater treatment plant based on uncertainty analysis and post-processing approaches (case study). *Water Sci. Technol.* **2021**, *83*, 1633–1648. [CrossRef]
25. Nadiri, A.A.; Shokri, S.; Tsai, F.T.C.; Moghaddam, A.A. Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model. *J. Clean. Prod.* **2018**, *180*, 539–549. [CrossRef]
26. Cheng, T.; Harrou, F.; Kadri, F.; Sun, Y.; Leiknes, T. Forecasting of wastewater treatment plant key features using deep learning-based models: A case study. *IEEE Access* **2020**, *8*, 184475–184485. [CrossRef]
27. Granata, F.; Papirio, S.; Esposito, G.; Gargano, R.; De Marinis, G. Machine learning algorithms for the forecasting of wastewater quality indicators. *Water* **2017**, *9*, 105. [CrossRef]
28. Safder, U.; Kim, J.; Pak, G.; Rhee, G.; You, K. Investigating Machine Learning Applications for Effective Real-Time Water Quality Parameter Monitoring in Full-Scale Wastewater Treatment Plants. *Water* **2022**, *14*, 3147. [CrossRef]
29. Bagherzadeh, F.; Mehrani, M.J.; Basirifard, M.; Roostaei, J. Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance. *J. Water Process Eng.* **2021**, *41*, 102033. [CrossRef]
30. Nourani, V.; Elkiran, G.; Abba, S.I. Wastewater treatment plant performance analysis using artificial intelligence—An ensemble approach. *Water Sci. Technol.* **2018**, *78*, 2064–2076. [CrossRef]

31. Araromi, D.O.; Majekodunmi, O.T.; Adeniran, J.A.; Salawudeen, T.O. Modeling of an activated sludge process for effluent prediction—A comparative study using ANFIS and GLM regression. *Environ. Monit. Assess.* **2018**, *190*, 495. [CrossRef]
32. Pai, T.Y.; Yang, P.Y.; Wang, S.C.; Lo, M.H.; Chiang, C.F.; Kuo, J.L.; Chang, Y.H. Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Appl. Math. Model.* **2011**, *35*, 3674–3684. [CrossRef]
33. Qiao, L.; Yang, P.; Leng, Q.; Xu, L.; Bi, Y.; Xu, J.; Ding, C. Exploring ANFIS application based on actual data from wastewater treatment plant for predicting effluent removal quality of selected major pollutants. *J. Water Process Eng.* **2023**, *56*, 104247. [CrossRef]
34. Malviya, A.; Jaspal, D. Artificial intelligence as an upcoming technology in wastewater treatment: A comprehensive review. *Environ. Technol. Rev.* **2021**, *10*, 177–187. [CrossRef]
35. Akbaş, H.; Bilgen, B.; Turhan, A.M. An integrated prediction and optimization model of biogas production system at a wastewater treatment facility. *Bioresour. Technol.* **2015**, *196*, 566–576. [CrossRef]
36. Sarkar, B.; Sharma, U.; Adhikari, K.; Lahiri, S.K.; Baltrėnaitė, E.; Baltrėnas, P.; Dutta, S. Application of artificial neural network and particle swarm optimization for modelling and optimization of biosorption of lead (II) and nickel (II) from wastewater using dead cyanobacterial biomass. *J. Indian Chem. Soc.* **2021**, *98*, 100039. [CrossRef]
37. Mahadeva, R.; Kumar, M.; Patole, S.P.; Manik, G. An optimized PSO-ANN model for improved prediction of water treatment desalination plant performance. *Water Supply* **2022**, *22*, 2874–2882. [CrossRef]
38. Tikhamarine, Y.; Souag-Gamane, D.; Ahmed, A.N.; Kisi, O.; El-Shafie, A. Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. *J. Hydrol.* **2020**, *582*, 124435. [CrossRef]
39. Xia, J.; Zeng, J. Environmental factors assisted the evaluation of entropy water quality indices with efficient machine learning technique. *Water Resour. Manag.* **2022**, *36*, 2045–2060. [CrossRef]
40. Riahi-Madvar, H.; Dehghani, M.; Memarzadeh, R.; Gharabaghi, B. Short to long-term forecasting of river flows by heuristic optimization algorithms hybridized with ANFIS. *Water Resour. Manag.* **2021**, *35*, 1149–1166. [CrossRef]
41. Ahmadianfar, I.; Bozorg-Haddad, O.; Chu, X. Gradient-based optimizer: A new metaheuristic optimization algorithm. *Inf. Sci.* **2020**, *540*, 131–159. [CrossRef]
42. Konakoglu, B.; Aydemir, S.B.; Kutlu Onay, F. Application of a metaheuristic gradient-based optimizer algorithm integrated into artificial neural network model in a local geoid modeling with global navigation satellite systems/leveling measurements. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e7017. [CrossRef]
43. Kadkhodazadeh, M.; Farzin, S. A novel LSSVM model integrated with GBO algorithm to assessment of water quality parameters. *Water Resour. Manag.* **2021**, *35*, 3939–3968. [CrossRef]
44. Adnan, R.M.; Mostafa, R.R.; Elbeltagi, A.; Yaseen, Z.M.; Shahid, S.; Kisi, O. Development of new machine learning model for streamflow prediction: Case studies in Pakistan. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 999–1033. [CrossRef]
45. Nourani, V.; Hakimzadeh, H.; Amini, A.B. Implementation of artificial neural network technique in the simulation of dam breach hydrograph. *J. Hydroinform.* **2012**, *14*, 478–496. [CrossRef]
46. Nourani, V.; Baghanam, A.H.; Gebremichael, M. Investigating the Ability of Artificial Neural Network (ANN) Models to Estimate Missing Rain-gauge Data. *J. Environ. Inform.* **2012**, *19*, 38–50. [CrossRef]
47. Sugeno, M.; Kang, G.T. Structure identification of fuzzy model. *Fuzzy Sets Syst.* **1988**, *28*, 15–33. [CrossRef]
48. Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *15*, 116–132. [CrossRef]
49. Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
50. Jang, J.S.; Sun, C.T. Neuro-fuzzy modeling and control. *Proc. IEEE* **1995**, *83*, 378–406. [CrossRef]
51. Jang, J.S.R.; Gulley, N. *MATLAB/Fuzzy Logic Toolbox*; MathWorks. Inc.: Natick, MA, USA, 1997.
52. Adnan, R.M.; Liang, Z.; El-Shafie, A.; Zounemat-Kermani, M.; Kisi, O. Prediction of Suspended Sediment Load Using Data-Driven Models. *Water* **2019**, *11*, 2060. [CrossRef]
53. Kisi, O.; Shiri, J.; Karimi, S.; Adnan, R.M. Three different adaptive neuro fuzzy computing techniques for forecasting long-period daily streamflow. *Big Data Eng. Appl.* **2018**, *44*, 303–321.
54. Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995*; IEEE: Piscataway, NJ, USA, 1995; pp. 39–43.
55. Montalvo, I.; Izquierdo, J.; Pérez, R.; Tung, M.M. Particle swarm optimization applied to the design of water supply systems. *Comput. Math. Appl.* **2008**, *56*, 769–776. [CrossRef]
56. Wang, D.; Tan, D.; Liu, L. Particle swarm optimization algorithm: An overview. *Soft Comput.* **2018**, *22*, 387–408. [CrossRef]
57. Ikram, R.M.A.; Cao, X.; Sadeghifar, T.; Kuriqi, A.; Kisi, O.; Shahid, S. Improving Significant Wave Height Prediction Using a Neuro-Fuzzy Approach and Marine Predators Algorithm. *J. Mar. Sci. Eng.* **2023**, *11*, 1163. [CrossRef]
58. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
59. Saremi, S.; Mirjalili, S.Z.; Mirjalili, S.M. Evolutionary population dynamics and grey wolf optimizer. *Neural Comput. Appl.* **2015**, *26*, 1257–1263. [CrossRef]
60. Wang, H.; Bai, Y.; Li, C.; Guo, Z.; Zhang, J. Time series prediction model of grey wolf optimized echo state network. *Data Sci. J.* **2019**, *18*, 16. [CrossRef]

61. Ikram, R.M.A.; Mostafa, R.R.; Chen, Z.; Islam, A.R.M.T.; Kisi, O.; Kuriqi, A.; Zounemat-Kermani, M. Advanced Hybrid Metaheuristic Machine Learning Models Application for Reference Crop Evapotranspiration Prediction. *Agronomy* **2023**, *13*, 98. [CrossRef]
62. Gao, Z.M.; Zhao, J. An improved grey wolf optimization algorithm with variable weights. *Comput. Intell. Neurosci.* **2019**, *2019*, 2981282. [CrossRef]
63. Himanshu, N.; Kumar, V.; Burman, A.; Maity, D.; Gordan, B. Grey wolf optimization approach for searching critical failure surface in soil slopes. *Eng. Comput.* **2021**, *37*, 2059–2072. [CrossRef]
64. Ypma, T.J. Historical development of the Newton–Raphson method. *SIAM Rev.* **1995**, *37*, 531–551. [CrossRef]
65. Ahmadi, M.M.; Mahdavirad, H.; Bakhtiari, B. Multi-criteria analysis of site selection for groundwater recharge with treated municipal wastewater. *Water Sci. Technol.* **2017**, *76*, 909–919. [CrossRef] [PubMed]
66. Nasr, M.S.; Moustafa, M.A.; Seif, H.A.; El Kobrosy, G. Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. *Alex. Eng. J.* **2012**, *51*, 37–43. [CrossRef]