Enhancing Water Quality Forecasting Reliability Through Optimal Parameterization of Neuro-Fuzzy Models via Tunicate Swarm Optimization

Dr. Kambala Vijaya Kumar[1], Y Dileep Kumar[2], Dr. Sanjiv Rao Godla[3],
Dr. Mohammed Saleh Al Ansari[4], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[5], Elangovan Muniyandy[6]

Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India[1]
Professor, Department of Electronics and Communication Engineering-School of Engineering, Mohan Babu University,
Tirupati, Andhra Pradesh, India[2]
Professor, Department of CSE (Artificial Intelligence & Machine Learning), Aditya College of Engineering & Technology -
Surampalem, Andhra Pradesh, India[3]
Associate Professor, College of Engineering-Department of Chemical Engineering, University of Bahrain, Bahrain[4]
Faculty of Informatics and Computing, UniSZA University, Malaysia[5]
Department of Biosciences-Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences,
Chennai, Tamil Nadu, India[6]

Abstract—**Forecasting water quality is critical to environmental management because it facilitates quick decision-making and resource allocation. On the opposite hand, current methods are not always able to produce reliable forecasts, which is often due to challenges in parameter optimization for complex models. This research presents a novel approach to enhance the forecasting accuracy of water quality by optimizing neuro-fuzzy models using Tunicate Swarm Optimisation (TSO). The introduction highlights the limitations of current techniques as well as the necessity for precise estimates of water quality. One of the drawbacks is that neuro-fuzzy models are not well-modelled, which makes it harder for them to identify the minute patterns in data on water quality. The suggested approach is unique in that it applies TSO, an optimization algorithm inspired by nature that emulates tunicates' behaviour, to the neuro-fuzzy models' parameter optimization process. The highly complex parameter space is effectively navigated by TSO's swarm intelligence, which strikes a balance between exploration and exploitation to improve model performance. To optimize model parameters, the process comprises three steps: creating an objective function, defining the neuro-fuzzy model, and seamlessly integrating TSO. By mimicking the motions of tunicates as they look for the best conditions in the marine environment, TSO constantly optimizes the variables. Experiments demonstrate that the proposed strategy is more effective than traditional optimization techniques in forecasting water quality. As seen by the optimised neuro-fuzzy model's increased prediction accuracy and several dataset validations, Tunicate Swarm Optimisation has potential for reliable environmental forecasting. This work presents a potential path for improved environmental decision-making systems by offering an optimisation strategy inspired by nature that overcomes the limitations of existing methods and enhances water quality forecasting tools.**

Keywords—Water quality forecasting; neuro-fuzzy models; tunicate swarm optimization; parameter optimization; environmental decision support

I. INTRODUCTION

Human activities are causing an increase in the variability of water quality in water bodies across the world. For instance, the incidence and length of hypo limnetic anoxia are rising in many lakes due to environmental and ecological changes, but waterbodies are facing stronger storms that start mixing and improving oxygen accessibility, leading to large every day fluctuations in oxygen levels [1]. The increased unpredictability of numerous water quality measurements exceeding a variety of historical circumstances makes estimating future water quality difficult, putting a significant pressure on management in charge of delivering essential lake and reservoir biological functions every day [2]. The burgeoning science of ecological prediction offers a fresh technique to proactively controlling dams and wetlands in the midst of rising water quality uncertainty. Environmental forecasting, or anticipating future ecology qualities based on observable unanticipated factors is a valuable tool for management. Prediction provides managers with probability projections for probable water quality circumstances in a focus lake or reserve tank, enabling them to take pre-emptive management activities that decrease or prevent water degradation [3].

Water quality is defined as the biological, chemical, and physical characteristics of water according to an array of water quality criteria. The decline in water quality resulted in major management initiatives to enhance and safeguard water quality, particularly in developing nations. In recent years, modelling and predicting river water quality for growth possibilities have played an essential role in environmental, ecological, and water resource management choices [4]. The field of computers and statistics has enhanced modelling tools for recognizing trends in water resources' data collected over time to correctly anticipate future events and enhance water resource management. The simulation and forecasting of

water quality variables are normally carried out using one of two methodologies. The first strategy is based on the process of flow and the chemical and physical characteristics of water, and it has been widely used in many basins [5]. This sort of water quality modelling frequently requires substantial information, including simulation settings and outside sources or drains. Also, when there is little monitoring data or insufficient background knowledge available, it might be challenging to mimic water quality systems using this model. The second methodology employs statistical and intelligence-based technologies. Artificial intelligence has grown rapidly in recent years, providing different ways to regression and better accuracy in a number of circumstances [6]. Water quality is influenced by a larger number of factors, including hydrological, biological, weather-related, and human activity. Because of interconnections among water quality factors, these variables are nonlinear, time variable, unpredictable, and postponed. As a result, it is distinct to represent such factors quantitatively using correct statistical representations and to create an accurate, nonlinear model for forecasting using conventional approaches [7].

A river's water quality varies over time and place, hence constant monitoring and analysis are required for successful river water quality management. A wide range of models and novel approaches have been presented for anticipating and managing shifts in water quality. Methods for evaluating and predicting these alterations in water quality are broadly classified as conceptual, deterministically models or models that include mathematical and stochastic methodologies [8]. The numeric approach necessitates a large quantity of data input, making it difficult to predict the optimal value. In addition, the user's subjective nature frequently causes problems with the estimate process. The stochastic method, however, has the benefit of being able to calculate the optimum parameters using time-series information on the water rather than simulating the water's physical, chemical, and biological features. Furthermore, the stochastic approach may be used for both long- and short-term projections by creating relatively unlimited input and output. In recent years, an AI system suited for nonlinear forecasting was used for water quality predicting to remove the user's subjectivity during parameter evaluation [9]. Artificial intelligence algorithms are increasingly being used for discharging forecasting research. The amount of things available for real-time automated measurements is limited, and obtaining all of the required input variables is challenging. As a result, the stochastic framework is regarded to be an excellent tool for analysing and forecasting fluctuations in the water's quality at continual surveillance stations, water intake infrastructure, and areas of frequent floods where continuous monitoring and immediate control are needed [10]. The main factors that will be examined are dissolved oxygen and total organic carbon [11].

Water quality forecasting is critical to good environmental administration because it provides essential information for rapid choices and appropriate resource allocation. As communities deal with the growing complexity of environmental behaviour, the accuracy of water quality projections becomes critical. However, present forecasting systems have substantial shortcomings, notably in the difficult issue of parameter optimization for complicated neuro-fuzzy models. The difficulties in making credible forecasts are frequently linked to inadequate parameters, which limits the models' capacity to detect the intricate patterns present in water quality data. This work tackles these constraints by introducing a novel approach that uses Tunicate Swarm Optimisation (TSO) to optimize neuro-fuzzy models, increasing the accuracy and resilience of water quality forecasts. Existing approaches require significant improvement due to difficulties with detailed parameter optimization, which limits their ability to capture every aspect of water quality changes. The proposed method uses TSO, an optimisation algorithm affected by the collective behaviour of tunicates, to construct a novel framework. By attempting to get beyond the limitations of conventional optimisation techniques, TSO integration offers a fresh approach to negotiating the complex parameter space connected to neuro-fuzzy models. The proposed method makes use of the swarm intelligence of TSO to carefully balance exploration and exploitation, ultimately leading to improved model performance. This introduction highlights the need of accurate water quality forecasts, points out the drawbacks of existing methods, and lays the foundation for the innovative methodology used in this study. The key contributions of this work are as follows:

1) The study goes on by offering a unique method for improving neuro-fuzzy algorithms for water quality forecasting, getting beyond parameterization issues that usually limit accurate forecasts.

2) TSO integration functions as a novel nature-inspired optimisation tactic, enhancing the ability to explore and utilise the complex parameter space of neuro-fuzzy models.

3) When compared to current optimization approaches, the proposed strategy significantly improves the projected accuracy of water quality forecasts, indicating the practical use of TSO.

4) By confirming the improved neuro-fuzzy model on several datasets, highlighting its capacity for generalisation and robust performance under diverse environmental circumstances.

5) Through this study, the limits of current methodologies are overcome, opening the door to the development of improved environmental systems supporting decision-making and providing a potential path towards more precise and informed environmental management decision-making.

The remaining of the paper is as follows: Section II explains an overall of previous studies on water quality forecasting approaches, setting the stage for the suggested strategy. In Section III, the report goes into the highlighted research gaps within current techniques, establishing the groundwork for the novel integration of TSO into neuro-fuzzy models. Results and discussion is given in Section IV. Finally, Section V concludes the paper.

II. RELATED WORKS

A. Enhancing Water Quality Forecasting with Ensemble Data Assimilation

Loos et al. [12] explain that one of the most pressing concerns facing civilization in the 21st decade is the safety of river water. Precise and dependable rapid forecast of water quality is an efficient adaptation strategy for dealing with water quality challenges including accidental leaks and major algal blooms. To improve the precision and skill of the water quality forecasts three distinct ensemble data assimilation (DA) methods were investigated two associated methods that can improve the starting point for non-linear calculations or reduce the time required for computation y determining the Kalman gain employing a time-lagged ensemble. Twin testing using artificial data of three species of algae and concentrations of phosphate with extremely small ensemble sizes demonstrated that each of the DA techniques improved prediction precision and skills, with only minimal variations between them. They all increased the model's precision at downstream places, with comparable results, but due to false membership, accuracy at upstream locations decreased somewhat. The studies likewise found no clear pattern of augmentation when the group as a whole size increased from 8 to 64. The real-world research, which included real-life observations of three varieties of algae and phosphorus levels, yielded fewer improvements than the two independent tests. Model accuracy can be enhanced by alternate state parameter definitions, the utilisation of distinct disturbances and inaccuracy in modelling parameters and/or improved calibrating of the stochastic water quality model.

B. Improving Water Quality Forecasting with Deep Learning

Water is necessary for the survival and sustenance of all living beings. River water quality has declined in recent years due to harmful waste and pollution. This growing contamination of water is an important cause of concern since it degrades water quality, making it unfit for any use. Water quality modeling using machine learning algorithms has grown in popularity in the past few years, and it has the potential to be extremely useful in ecological and the administration of water resources. They usually encounter a high level of computation and forecast inaccuracy. Because of its excellent performance, time-series information was processed using a deep neural network that includes a LSTM. Khullar and Singh [13] explains that the deep learning-based Bi-LSTM method is used to forecast water quality features in India's Yamuna River. The previous systems fail to perform imputation of missing values and instead focus solely on learning management, with no repercussions mechanism for training failures. The proposed model employs an innovative approach in which the initial phase involves missing value imputation, the second step generates map features from the information at hand, the third step incorporates a Bi-LSTM architecture to improve the learning process, followed by an optimized loss function to reduce training error. As a result, the proposed model enhances predictive accuracy. Several water quality indicators were collected monthly at several places around the Delhi area across. The experimental findings show that the expected results of the model's parameters and the actual outcomes were perfectly consistent,

which might indicate a future trend. The model's efficacy was compared to various novel approaches, such as SVR, random forest models, artificial neural networks (ANN), LSTM, and CNN-LSTM.

C. Enhancing Water Quality Monitoring in Aquaculture

Global changes in climate and water contamination have generated several challenges for fish/shrimp growers, such as early death before harvest. It is critical to understand methods to track and handle water quality to assist farmers in addressing this issue. Water quality monitoring is critical for designing IoT systems, particularly in fisheries and aquaculture. Researchers can regulate water quality by tracking real-time sensor data signals (such as salinity, pH, water temperature, and dissolved oxygen) and predicting them to acquire early warning, therefore gathering quantities as well as quality in shrimp/fish rearing. Thai-Ngheet al. [14] introduce A framework with an approach to forecasting for IoT devices used for tracking water quality in fisheries and aquaculture. Because these indicators are gathered daily, they constitute sequential/time series data. Researchers suggest using deep learning with the LSTM method to forecast these parameters. The experimental findings on many data sets demonstrate that the suggested technique works effectively and can be used in actual systems.

D. Revolutionizing Irrigation Water Quality Assessment

Conventional methods for evaluating irrigation water quality are frequently costly and time-consuming for farmers, especially in underdeveloped nations. However the use of artificial intelligence algorithms can address this issue by anticipating and analysing irrigation water quality indices in aquifer systems utilizing physical factors as features. El Bilali et al. [15] aims The variables Total Dissolved Solid (TDS), Magnesium Adsorption techniques Ratio (MAR), PS, Interchangeable Sodium Percentage, Sodium Adsorption Ratio (SAR), and Remaining Sodium Carbonates (RSC) are projected using electrical conductivity (EC), temperatures (T), and pH. To do this, researchers developed and tested adaptive boost, RF, ANN, and SVR algorithms using 520 data samples associated with 14 qualitative groundwater metrics from Morocco's Berrechid aquifer. The data show that Adaboost and RF approaches outperformed SVR and ANN in terms of overall prediction accuracy. However, generalisation ability and sensitivity to input studies show that ANN and SVR approaches are more adaptable and less susceptible to input factors than Adaboost and RF. The algorithms developed throughout the world are effective for predicting irrigation water quality characteristics and may assist producers and managers in managing irrigation water strategies. The suggested approaches in this study have showed the promise in inexpensive and real-time estimates of groundwater quality using physical information as input variables.

E. Deep Learning-based Approaches for Water Quality Forecasting

Various contaminants have posed a danger to water quality in recent years. As a result, modelling and forecasting water quality have grown to be critical tools for mitigating water pollution. Aldhyani et al. [16], Advanced artificial intelligence approaches are being developed to anticipate the water quality

index (WQI) and classification (WQC). Artificial neural network models, notably the nonlinear autoregressive model neural network (NARNET) and the LSTM deep neural networks approach, were developed for WQI prediction. Additionally, three machine learning algorithms, support vector machine (SVM), k-nearest neighbor (K-NN), and Naive Bayes, were used for WQC forecasting. The dataset used comprises seven major components, and the resulting models were evaluated using a variety of statistical criteria. The results suggest that the suggested models may properly predict WQI and water quality because of greater resilience. Prediction results showed that the NARNET methods performed somewhat better than the LSTM in forecasting WQI values, while the SVM approach had the greatest prediction accuracy in WQC. Also, the NARNET and LSTM equations obtained identical accuracy throughout testing, with just minor changes in the regression coefficient. This intriguing research could have a huge impact on water management.

Water quality management is a crucial concern for civilizations, and precise real-time prediction is required to solve situations like unintentional spills and dangerous algal blooms. Three ensemble integrations of data approaches were examined to enhance water quality forecasts overall increase in prediction accuracy [15]. Meanwhile, a deep learning-based Bi-LSTM algorithm was presented for forecasting water quality variables in India's Yamuna River, surpassing several cutting-edge methodologies. Furthermore, the combination of IoT systems and deep learning with LSTM was investigated for tracking and predicting water quality parameters in aquaculture and fisheries, resulting in early alerts.Artificial intelligence models, like as Adaboost, RF, ANN, and SVR, were designed and tested for forecasting variables in Morocco's Berrechid aquifer. Finally, sophisticated AI algorithms such as NARNET and LSTM were used to forecast the index of water quality and categorization, with promising accuracy and resilience. These studies emphasize the possibility of data integration, machine learning, and IoT technologies in improving water quality prediction and administration on various scales and in varied geographical situations.

F. Problem Statement

The current limitation in water quality forecasting approaches is the poor modelling of neuro-fuzzy models, which impairs their capacity to effectively capture the subtle trends that characterize water quality data. Traditional optimization algorithms frequently struggle to navigate the high-dimensional parameter array successfully, resulting in unsatisfactory model performance. This paper addresses this challenge by describing a novel approach that incorporates Tunicate Swarm Optimisation (TSO) into the parameter optimization method for neuro-fuzzy models, therefore improving the dependability of water quality forecasts. The TSO, which is based on the collective behaviour of tunicates, offers an effective optimization technique based on nature for exploring and using the parameter space. The limitations of conventional optimization techniques are overcome by the swarm intelligence of TSO, which allows a more robust and efficient search for the optimum model parameters. By

combining TSO with neuro-fuzzy systems, it is possible to improve environmental decision-making systems and increase the accuracy and dependability of water quality prediction while also addressing the drawbacks of current methods.

III. INTEGRATING TUNICATE SWARM OPTIMIZATION WITH NEURO-FUZZY MODELS

The approach section illustrates how to optimize neuro-fuzzy systems for water quality forecasts by using TSO in a new way. The techniques involve developing the neuro-fuzzy model, specifying an objective operation, and smoothly integrating TSO into the optimization process. The section includes how TSO's swarm intelligence is utilized to efficiently investigate the highly dimensional variable space while establishing a balance between utilizing it to enhance model performance. The emphasis is on iteratively improving model parameters, which mimic tunicates' motions in pursuit of optimal conditions in the marine environment. The methodology covers the rigorous procedure for ensuring that TSO merges with the optimal neuro-fuzzy model architecture. The comprehensive methodology is designed to assist academics and practitioners in replicating and using this unique optimization technique in water quality forecasting.

A. Data Collection

The goal is to determine the spatial quality of water as an indicator of the strength of hydrogen (pH) values on the other day utilizing data from water measurement indices. The PH value for the following day is generated using the given input data's, which consists of historical information from several water measurements indices. The input information contains everyday samples from 36 sites in Georgia, USA, which provide information about pH values. The input parameters include 11 typical indications such as the amount of dissolved oxygen, temperature, and specific conductivity [17]. The expected outcome is a measurement of 'pH, water, raw, field, standard units (median)'.

B. Adaptive Neuro-Fuzzy Model

The adaptive neuro-fuzzy inference system (ANFIS) Combines neural networks and fuzzy theories. The ANFIS modifies its relationship function and management rules based on both inputs and outputs data collected from the controlling environment to correspond with the object under control. The ANFIS outperforms a BP-based multi-layer perceptron in matching an excessively nonlinear environment. The ANFIS model takes longer because the hybrid training rule demands more computation. The ANFIS's fundamental learning strategy is to change the preceding variable from the backward path by the variation of the squared error for the outcome of each node [18]. The previous parameter defines the form of the member function, while the parameter value $\{b^i, d^i\}$ and determines the squared error of E by calculating the breadth and centre of the function defining membership. To lower the amount of E, the next maximum gradient approach is applied over and over to the preceding variable which is shown in Eq. (1-3): Fig. 1 shows the structure of ANFIS.

$$b^i(t+1) = b^i(t) - \eta \frac{\partial E}{\partial b^i} \qquad (1)$$

$$d^i(t+1) = d^i(t) - \eta \frac{\partial E}{\partial d^i} \qquad (2)$$

$$\eta = \frac{k}{\sqrt{\Sigma_\alpha (\frac{\partial E}{\partial \alpha})^2}} \qquad (3)$$

The preceding parameter (α) and the shifting distance (k) of the gradient vectors in the field of parameters impact the rate of converging. Eq. (4) expresses the overall result of f as a linear mixture of the subsequent parameters $\{p^i, q^i, r^i\}$:

$$f = \overline{\omega^1} f^1 + \overline{\omega^2} f^2 = (\overline{\omega^1} x)p^1 + (\overline{\omega^1} y)q^1 + (\overline{\omega^1})r^1 + (\overline{\omega^2} x)p^2 + (\overline{\omega^1} y)q^2 + (\overline{\omega^2})r^2 \qquad (4)$$

According to the Sugeno and Takagi category, a system for Fuzzy reasoning has a pair of inputs along with a single output.

Rule 1: $f^1 = p^1 x + q^1 y + r^1$ as x is A^1 and y is B^1

Rule 2: $f^2 = p^2 x + q^2 y + r^2$ as x is A^2 and y is B^2

The first layer: In this particular layer, using a node functions, every node i is a squared node which is given in Eq. (5).
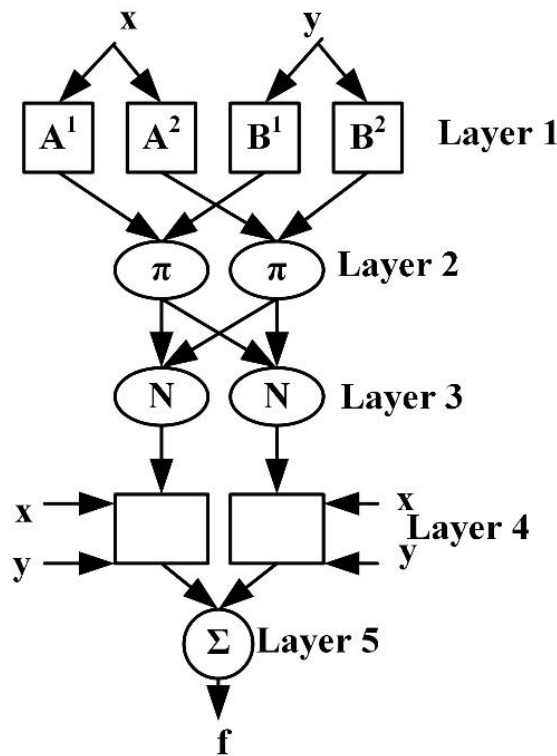
$$O_1^i = \mu A^i(x) \qquad (5)$$

Fig. 1. ANFIS structure.

x is the source to node i, whereas A^i is the syntactical branding associated with the nodal functions. With this findings, O_1^i becomes a relationship mapping to A^i. The functions for relationship is denoted by $\mu A^i(x)$, where the greatest value is 1 and the smallest is 0, and in the generic bell

mappings or Gaussians mapping process, as detailed below Eq. (6) and (7).

$$\mu A^i(x) = \frac{1}{1 + [(\frac{x-c^i}{a^i})^2]^{b^i}} \qquad (6)$$

$$\mu A^i(x) = e^{[-(\frac{x-c^i}{a^i})^2]} \qquad (7)$$

The data set is denoted as $\{a^i, b^i, c^i\}$. If a result, if the parameters vary, it will impact the bell-like mapping. Thus, in differentiated mapping, a frequency will be in a triangle or trapezoid form, which is an important component of the node's position in this layer [19].

The second layer: Each in this layer is a circular node, which products signals that arrive and outputs the products. For example, Eq. (8):

$$O_2^i = W^i = \mu c^i(x) \times \mu e^i(x), \quad i = 1,2,3 \ldots \qquad (8)$$

The result of every node indicates the fires power of a rule.

Third layer: In this layer's structure in Eq. (9), every nodes is a circular nodes marked N. The ith node computes the proportion of the ith rule's fired intensity to the total of all rules' fired intensities.

$$O_3^i = W^i = \frac{w^i}{w^1 + w^2}, \quad i = 1,2,3 \ldots \qquad (9)$$

For simplicity, the results of this particular layer would be referred to as normalized fired intensity.

Fourth layer: Each node i in the tier is a squared node containing a nodal function of Eq. (10).

$$O_4^i = \bar{w}^i f^i = \bar{w}^i = \bar{w}^i (p^i x + q^i y + r^i) \qquad (10)$$

The outcome of layer 3 is $\bar{w}^i$, and the parameter value collection is (p^i, q^i, n^i). Variables in this level will be referenced as subsequent variables.

Fifth layer: This layer's solitary nodes are a marked circular node that calculates the total outcome as the total of all the signals that arrive in Eq. (11), i.e.

$$O_5^i = \sum \bar{w}^i f^i = \sum_i w^i f^i / \sum_i w^i \qquad (11)$$

This results in an adaptive networks that is virtually comparable to a type of three fuzzy inference systems.

C. Tunicate Swarm Optimization

Tunicates are cylindrical-shaped organisms that have just one of their two ends open and travel at jet-like speeds over the water's surface. They may seek nutrition in the sea, regardless of whether they are unsure where to begin. The tunicates' jet-like speed and clever swarming form the foundation for the TSA optimization approach. When responding to the TSA's optimization quandary, the food supplier is the best solution. Certain requirements must be accomplished to correctly recreate TSA jet propulsion motions. Before continuing, two prerequisites must be fulfilled: In the initial stages, the tunicates must avoid fighting. Second, they must continue looking for their greatest search agent. Finally, they need to keep close to the agent [20]. The swarm knowledge of the additional tunicates in a

statistical framework is employed for updating their locations relative to the ideal solution. The theoretical system is defined as follows:

1) Conditions: There should be no disagreements between the search agents. To prevent conflicts amongst search agents, use the next vectors to determine their relative locations of Eq. (12) to Eq. (14).

$$\vec{a} = \frac{\vec{g}}{\vec{m}} \tag{12}$$

$$\vec{g} = c_2 + c_3 - \vec{f} \tag{13}$$

$$\vec{f} = 2 * c_1 \tag{14}$$

The gravitational force is symbolized by $\vec{g}$, whereas $\vec{f}$ represents the fluctuation in temperatures of the deeper seawater stream. To determine social forces among tunicates (represented by vector $\vec{m}$), apply the subsequent formula: c_3, c_2, and c_1 are random numbers with values ranging from zero to one.

P^{min} and P^{max} represents the initial and secondary speeds of social contact. During this optimization phase, it is important to ensure that the tunicate moves in a certain direction.

$$\overrightarrow{PD} = \left| \vec{FS} - r^{and} * \overrightarrow{P^p(x)} \right| \tag{15}$$

Eq. (15) gives the present iteration's cycle is marked by x, the separation between the supply of food and search agents is indicated by $\overrightarrow{PD}$, the exact spot of search agents is denoted by $\overrightarrow{P^p(x)}$, the exact location of food source is represented by $\vec{FS}$, and the value of the random variables r^{and} is determined in an amount of 0 to 1.

Moving towards the greatest search agent. To do this, the search agents are reorganized as follows in Eq. (16):

$$\overrightarrow{P^p(x')} = \begin{cases} \vec{FS} + \vec{a} * \overrightarrow{PD}, & if \ r^{and} \geq 0.5 \\ \vec{FS} - \vec{a} * \overrightarrow{PD}, & if \ r^{and} \leq 0.5 \end{cases} \tag{16}$$

$\overrightarrow{P^p(x')}$ reflects the search agent's present position relative to the available food supply. The first two best answers are saved and utilized to adjust the positioning of the additional tunicates to simulate swarm activity. This Eq. (17) is a mathematical illustration of a swarm.

$$\overrightarrow{P^p(x+1)} = \frac{\overrightarrow{P^p(x)} + \overrightarrow{P^p(x+1)}}{2 + c_1} \tag{17}$$

The key stages for demonstrating the flow of the initial TSO are shown here for clarity. Fig. 2 shows the flowchart for the TSO algorithm [20].

Set the starting population of tunicates, or $\vec{P^p}$ to the usual number.

Define the variable's starting values and the large amount of repetitions.

Each exploration agent's success score must be determined.

Lastly, the best-fitting agents are examined in the searching space supplied after assessing their fitness.

Investigate agents need to be improved. It's time to return the freshly strengthened agents to his or her location of origination.

Determine the suitability cost for a more sophisticated search agents.

When the initial response is no longer optimal, the best response X^{best} is saved and $\vec{P^p}$ is improved.

The implementation of TSO into the neuro-fuzzy model's optimisation process entails expressing alternative solutions as people in a swarm, with each matching to a distinct set of neuro-fuzzy model variables. TSO uses swarm ability, inspired by tunicate activity, to dynamically balance both discovery and extraction in the highly dimensional variable space. Members in the swarm adjust their locations through an iterative optimization process based on assessments of fitness utilizing the objective function, which commonly uses measures such as Mean Squared Error. The ultimate aim is to minimize the objective function, resulting in reliable water quality forecasts. The algorithm is guided by terminating conditions, such as attaining a desired fitness level, and the ideal parameters defined by the TSO are retrieved at the end. This integration intends to rise the reliability of water quality predicts by quickly traversing the vast range of parameters of neuro-fuzzy models.
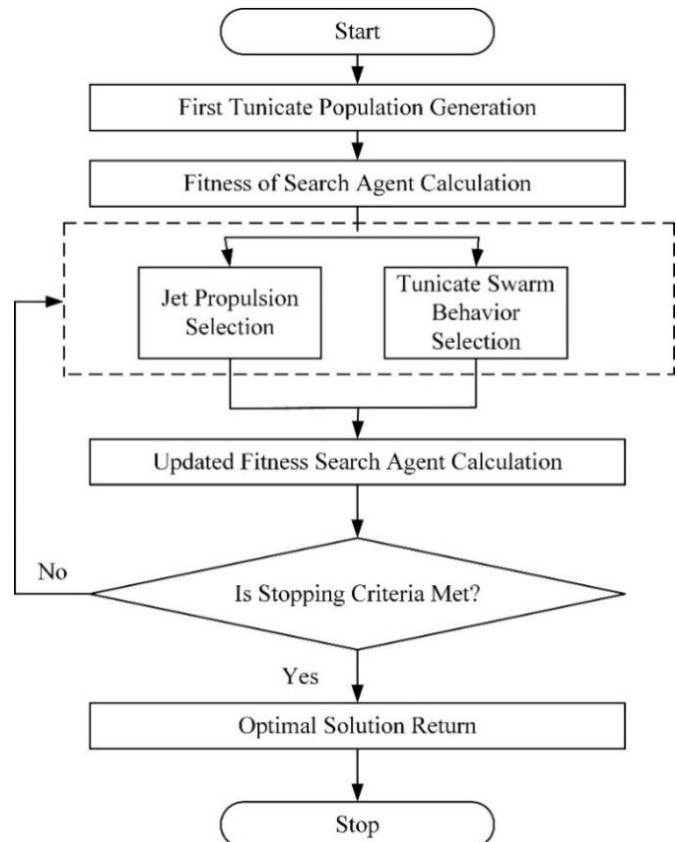

Fig. 2. Tunicate swarm optimization flowchart.

IV. RESULTS AND DISCUSSIONS

The suggested neuro-fuzzy model optimized using TSO was evaluated primarily using MSE and perhaps additional regression-based measures. MSE measures the average squared variance among predicted and actual water quality measurements, indicating the model's accuracy. The outcomes of thorough trials demonstrate that the suggested methodology outperforms standard optimization methodologies. The optimized neuro-fuzzy model constantly has reduced MSE values, suggesting higher prediction accuracy. The model performs well over a wide range of datasets, demonstrating its generalizability. The comparison analysis shows a considerable improvement in water quality predictions over previous approaches, highlighting TSO's practical usefulness in traversing the highly dimensional variable space. These outcomes emphasize the possibility of Tunicate Swarm Optimisation as a reliable optimization approach for neuro-fuzzy algorithms for environmental forecasting, giving a viable route for enhanced decision-support tools in water quality management.

A. Analysis

By incorporating Tunicate Swarm Optimization into neuro-fuzzy systems for water quality prediction, the suggested study presents a unique method. This novel method effectively navigates the intricate parameter space of neuro-fuzzy models, addressing the shortcomings of current approaches. Utilizing a thorough assessment of measures such as Mean Squared Error (MSE) and comparative studies, the research exhibits better performance than traditional optimization techniques. Through the use of swarm intelligence derived from tunicate behaviour, TSO improves model generalizability and accuracy on a variety of datasets. The investigation demonstrates TSO's potential as a trustworthy optimization approach for environmental forecasting, creating opportunities for enhanced instruments for water quality management decision-making. Future studies may include hybrid optimization strategies, scalability, and wider uses of TSO in environmental modelling.

B. Performance Measurement

The suggested model's capability to forecast the WQI was evaluated using performance measuring methodologies such as MSE [21]. The statistical approaches utilized are described as follows:

Mean square error (MSE): The mathematical expression for MSE is shown in Eq. (18).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y^i - \hat{y}^i)^2 \qquad (18)$$

Mean Absolute Error (MAE): The mathematical expression for MAE is shown in Eq. (19).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y^i - \hat{y}^i| \qquad (19)$$

Root mean square error (RMSE): The equation for RMSE is given in Eq. (20) and Eq. (21).

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(Y^i - \widehat{y^i})^2}{N}} \qquad (20)$$

$$R = \frac{n\sum(x \times y) - (\sum x)(\sum y)}{[n\sum(x^2) - \sum(x^2)] \times [n\sum(y^2) - \sum(y^2)]} \times 100\% \qquad (21)$$

where, R is Pearson's correlation coefficient of Eq. (20), x is the observational input information from the first batch of training information, y is the observational input information from the following set of training data, and n is the overall amount of input parameters.

C. WQForecasting using the ANFIS Model

The suggested model technique was validated by training 70% of the available data with the ANFIS model and predicting the WQ. The training outcomes revealed that the ANFIS approach was particularly effective at predicting WQ. Table I summarises the forecasting outcomes of the WQ achieved by the ANFIS framework throughout training and testing periods.

Fig. 3 displays the mistakes in forecasting water quality with ANFIS over the training and testing stages. Errors are assessed using RMSE and MSE. The RMSE numbers are much greater than the MSE values in both stages, indicating a wider range of residual errors. The graphic shows a bar graph depicting inaccuracies in ANFIS-based water quality prediction. There are two bars, one labelled "Training" and the other "Testing". Every set possesses two metrics: RMSE and MSE. For both Training and Testing, RMSE exhibits a larger error rate. The y-axis is labelled "Errors" and ranges from 0 to 0.07. The x-axis is separated into two main groups: Training and Testing Metrics.
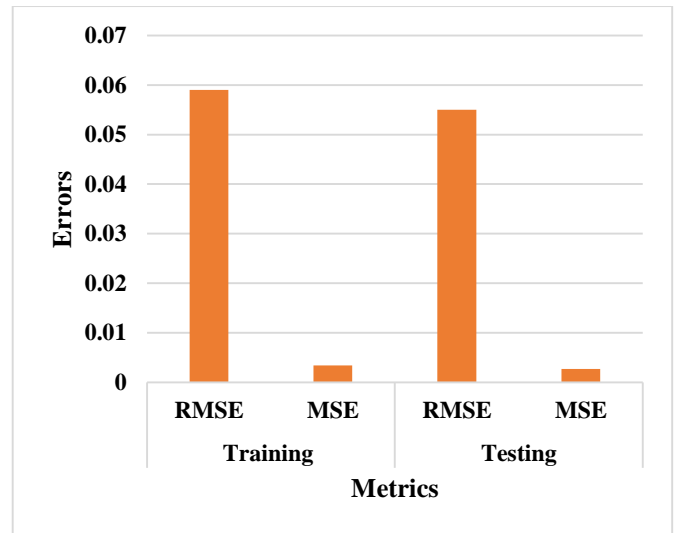

Fig. 3. Predicting water quality with ANFIS.

Fig. 4 shows two bars, one labelled "Testing" and the other "Training". The Y-axis is labelled "R (%)", suggesting that it reflects the Pearson coefficient as a percentage. The X-axis spans from 89.5% to 93%, representing the Pearson coefficients achieved. The "Testing" bar is substantially longer than the "Training" bar, going beyond 92%. The "Training" bar is slightly over 90%.

TABLE I.   ABILITY OF ANFIS TO FORECAST WATER QUALITY

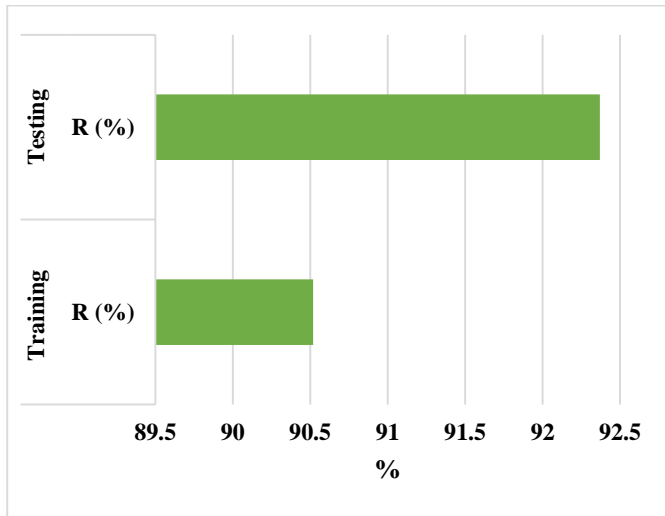| Model	Training				Testing			
	RMSE	*MSE*	*Mean Errors*	*R (%)*	*RMSE*	*MSE*	*Mean Errors*	*R (%)*
ANFIS	0.0590	0.00338	0.006458	90.52	0.0550	0.0027	0.001350	92.37

Fig. 4.   Pearson correlation coefficient for training and testing.

Table II compares the efficacy of three models, including LSSVM, ANFIS-PSO, and ANFIS, in terms of MAE (mg/l). The smaller the MAE number, the more effectively the model performs. As shown in the table, the model based on ANFIS has the smallest MAE of 12 mg/l, next to ANFIS-PSO at 13 mg/l. The LSSVM model has the greatest MAE value (13.2 mg/l).

Fig. 5 shows the MSE Loss Functions of the training and validation datasets for predicting water quality. The x-axis most likely reflects the total number of times or iterations, whilst the y-axis shows the MSE value. Both training and testing mistakes fall quickly at first but subsequently, level, showing that the system is learning successfully but eventually reaches a point when further development is negligible.

The graph in Fig. 6 shows the MAE in mg/g for three distinct models: LSSVM, ANFIS-PSO, and ANFIS. The y-axis is marked "MAE (mg/g)" and varies between 11.4 and 13.4. The x-axis is marked "Models," and each model's name appears under its bar. There are three bars for each framework: LSSVM, ANFIS-PSO, and ANFIS. The LSSVM model has the greatest MAE at around 13.2 mg/g, then ANFIS-PSO with an MAE of approximately 12.8 mg/g, and ANFIS has the smallest MAE at around 12 mg/g.

TABLE II.   COMPARISON OF MODEL PERFORMANCE

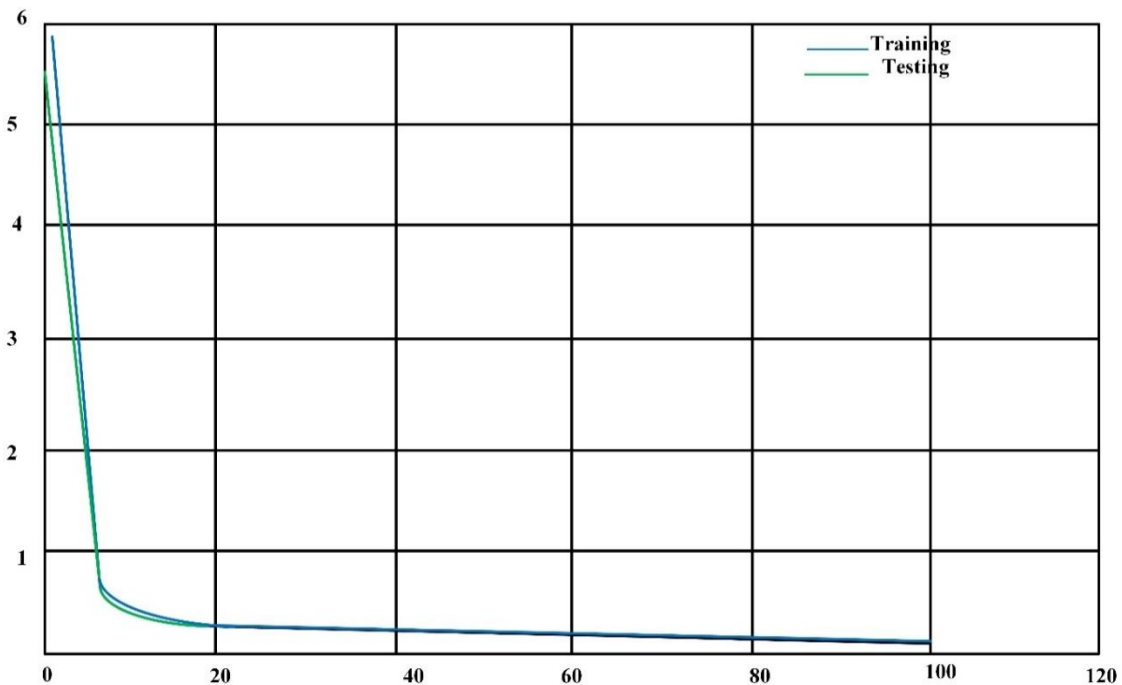| Models	MAE (mg/l)
LSSVM	13.2
ANFIS-PSO	13
ANFIS	12

Fig. 5.   MSE loss function comparison of training and testing dataset.
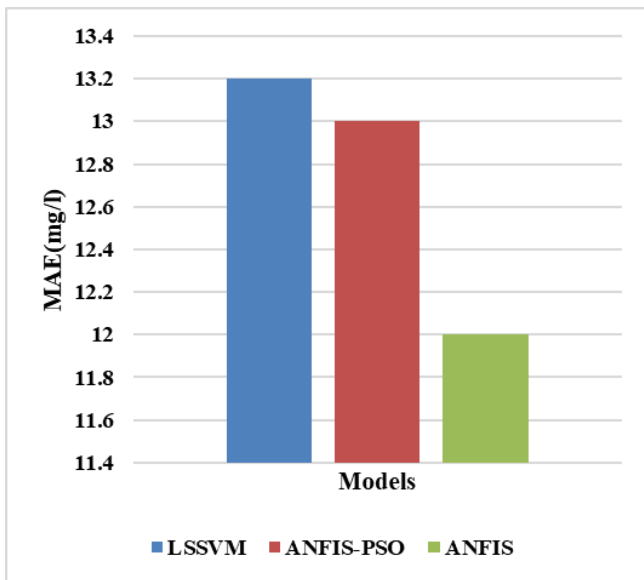
Fig. 6.   Comparison of model performance.

D. Discussion

The transforming effect of including TSO into the optimization procedure for neuro-fuzzy systems for water quality prediction. The findings highlight the need for precise water quality forecasts in environmental administration and decision-making, as well as the limits of present approaches caused by poor parameterization. The current shortcoming in water quality projection methods originates from the poorly constructed neuro-fuzzy models, which hinders their ability to accurately represent the nuanced patterns seen in water quality data [22]. The unique technique proposed in this study solves these issues by smoothly adding TSO, a nature-inspired optimization method, into the parameter optimization process. The swarm knowledge of tunicates is used to explore the complex and highly dimensional parameter set of neuro-fuzzy models, achieving an equilibrium between investigation and exploitation. The incorporation considerably improves the prediction performance of the neuro-fuzzy approach, as proven by convincing findings gained from extensive tests across varied datasets.

Furthermore, the discussion expands on the wider consequences of the suggested technique, indicating a possible route for enhanced environmental decision-making systems. The optimized neuro-fuzzy model's strong performance demonstrates TSO's ability to outperform existing optimization techniques in terms of water quality forecasts. The work not only advances forecasting methodology but also highlights the possibility of nature-inspired optimization strategies in solving complex environmental concerns. The application of TSO for parameter optimization might provide computational difficulties because of the algorithm's intricacy and resource needs. Subsequent investigations ought to concentrate on broadening the scope of TSO's application in various environmental forecasting domains, improving its amalgamation with neuro-fuzzy designs, scrutinising discrepancies in swarm behaviour, appraising its scalability for extensive water quality

forecasting, and investigating hybrid optimisation tactics to augment the forecasting resilience.

V.   CONCLUSION

A novel strategy for optimizing neuro-fuzzy scenarios in the area of water quality predictions based on the creative integration of TSO. The study begins by highlighting the importance of precise water quality forecasts in environmental management, as well as the limitations of current techniques, notably in the realm of inadequate modelling. The suggested technique, which incorporates TSO, appears as a persuasive solution to these restrictions. By emulating tunicates' shared intelligence, TSO effectively navigates the complicated and high-dimensional parameter range that comes with neuro-fuzzy systems. This effortless integration is demonstrated in the detailed methodology section, which describes the configuration of neuro-fuzzy designs, the development of a function with objectives, and the continuous parameter optimization process utilizing TSO. Extensive trials have validated the usefulness of the suggested strategy, exhibiting improved water quality forecasts compared to existing optimization methods. The optimized neuro-fuzzy model regularly exceeds previous approaches across a variety of datasets, demonstrating its resilience and generalizability. Beyond its immediate use, the study advances the field by offering a nature-inspired optimization approach that shows promise for solving issues in environmental systems that support decisions. This work represents a substantial development in water quality forecasting approaches, marking a vital step towards more reliable and precise environmental forecasts, and giving new pathways for the incorporation of nature-inspired algorithms into environmental science and administration.

To expand the application of TSO, future studies should examine how it may be modified to fit a variety of environmental forecasting scenarios. Further research should concentrate on enhancing the TSO using neuro-fuzzy designs combination, examining variations in swarm behaviour, and evaluating the approach's suitability for massive amounts water quality forecasting systems. Moreover, research endeavours can concentrate on hybrid optimization tactics that merge TSO with other nature-inspired algorithms to leverage synergies and enhance the robustness of environmental prediction techniques. Future work may focus on extending TSO to other optimization tasks in environmental modelling. Additionally, investigating hybrid optimization approaches and integrating real-time data streams could enhance the applicability and robustness of the proposed methodology.

REFERENCES

[1] C. C. Carey et al., "Advancing Lake and reservoir water quality management with near-term, iterative ecological forecasting," Inland Waters, vol. 12, no. 1, pp. 107–120, 2022.

[2] M. I. Shah et al., "Modeling surface water quality using the adaptive neuro-fuzzy inference system aided by input optimization," Sustainability, vol. 13, no. 8, p. 4576, 2021.

[3] R. Trach, Y. Trach, A. Kiersnowska, A. Markiewicz, M. Lendo-Siwicka, and K. Rusakov, "A study of assessment and prediction of water quality index using fuzzy logic and ANN models," Sustainability, vol. 14, no. 9, p. 5656, 2022.

[4] R. Barzegar, A. Asghari Moghaddam, J. Adamowski, and B. Ozga-Zielinski, "Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model," Stoch. Environ. Res. Risk Assess., vol. 32, pp. 799–813, 2018.

[5] Ahmadianfar, S. Shirvani-Hosseini, J. He, A. Samadi-Koucheksaraee, and Z. M. Yaseen, "An improved adaptive neuro fuzzy inference system model using conjoined metaheuristic algorithms for electrical conductivity prediction," Sci. Rep., vol. 12, no. 1, p. 4934, 2022.

[6] M. Mohadesi and B. Aghel, "Use of ANFIS/genetic algorithm and neural network to predict inorganic indicators of water quality," J. Chem. Pet. Eng., vol. 54, no. 2, pp. 155–164, 2020.

[7] J. Zhang et al., "The combination of multiple linear regression and adaptive neuro-fuzzy inference system can accurately predict trihalomethane levels in tap water with fewer water quality parameters," Sci. Total Environ., vol. 896, p. 165269, 2023.

[8] S. Heddam, O. Kisi, A. Sebbar, L. Houichi, and L. Djemili, "Predicting water quality indicators from conventional and nonconventional water resources in Algeria country: Adaptive neuro-fuzzy inference systems versus artificial neural networks," Water Resour. Algeria-Part II Water Qual. Treat. Prot. Dev., pp. 13–34, 2020.

[9] S. Narges, A. Ghorban, K. Hassan, and K. Mohammad, "Prediction of the optimal dosage of coagulants in water treatment plants through developing models based on artificial neural network fuzzy inference system (ANFIS)," J. Environ. Health Sci. Eng., vol. 19, pp. 1543–1553, 2021.

[10] M. Abd El-Mageed, T. A. Enany, M. E. Goher, and M. E. Hassouna, "Forecasting water quality parameters in Wadi El Rayan Upper Lake, Fayoum, Egypt using adaptive neuro-fuzzy inference system," Egypt. J. Aquat. Res., vol. 48, no. 1, pp. 13–19, 2022.

[11] S. L. Zubaidi et al., "A novel methodology for prediction urban water demand by wavelet denoising and adaptive neuro-fuzzy inference system approach," Water, vol. 12, no. 6, p. 1628, 2020.

[12] S. Loos, C. M. Shin, J. Sumihar, K. Kim, J. Cho, and A. H. Weerts, "Ensemble data assimilation methods for improving river water quality forecasting accuracy," Water Res., vol. 171, p. 115343, 2020.

[13] S. Khullar and N. Singh, "Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation," Environ. Sci. Pollut. Res., vol. 29, no. 9, pp. 12875–12889, 2022.

[14] N. Thai-Nghe, N. Thanh-Hai, and N. Chi Ngon, "Deep learning approach for forecasting water quality in IoT systems," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 8, pp. 686–693, 2020.

[15] El Bilali, A. Taleb, and Y. Brouziyne, "Groundwater quality forecasting using machine learning algorithms for irrigation purposes," Agric. Water Manag., vol. 245, p. 106625, 2021.

[16] T. H. Aldhyani, M. Al-Yaari, H. Alkahtani, M. Maashi, and others, "Water quality prediction using artificial intelligence algorithms," Appl. Bionics Biomech., vol. 2020, 2020.

[17] "Water Quality Dataset," 2022, [Online]. Available: https://www.kaggle.com/datasets/shrutisaxena/water-quality-prediction-data-set

[18] Yeon, J. Kim, and K. Jun, "Application of artificial intelligence models in water quality forecasting," Environ. Technol., vol. 29, no. 6, pp. 625–631, 2008.

[19] K. S. Parmar, S. J. S. Makkhan, and S. Kaushal, "Neuro-fuzzy-wavelet hybrid approach to estimate the future trends of river water quality," Neural Comput. Appl., vol. 31, no. 12, pp. 8463–8473, 2019.

[20] Taher, M. Elhoseny, M. K. Hassan, and I. M. El-Hasnony, "A Novel Tunicate Swarm Algorithm With Hybrid Deep Learning Enabled Attack Detection for Secure IoT Environment," IEEE Access, vol. 10, pp. 127192–127204, 2022.

[21] M. Hmoud Al-Adhaileh and F. Waselallah Alsaade, "Modelling and prediction of water quality by using artificial intelligence," Sustainability, vol. 13, no. 8, p. 4259, 2021.

[22] O. Kisi, K. S. Parmar, A. Mahdavi-Meymand, R. M. Adnan, S. Shahid, and M. Zounemat-Kermani, "Water quality prediction of the yamuna river in India using hybrid neuro-fuzzy models," Water, vol. 15, no. 6, p. 1095, 2023.