



Enhancing Water treatment predictions: a Machine Learning Approach with CNN and Water Wave optimization

Sajeda Alkhadrawi¹ · Kamel K. Alzboon²

Received: 23 April 2024 / Accepted: 8 May 2024 / Published online: 28 May 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

This research investigates using Convolutional Neural Networks (CNN) and Water Wave Optimization (WWO) to improve predictive modeling in water treatment operations. The method utilizes Convolutional Neural Networks (CNN) for their strong prediction skills and integrates them with Weather-Water-Ocean (WWO) data to enhance the selection of relevant features, aiming to improve the accuracy and efficiency of forecasting water quality indicators. The research systematically compares the performance of the CNN-WWO model with standalone CNN models, specifically evaluating parameters such as accuracy, precision, recall, and F1-score. The results demonstrate that the CNN-WWO model significantly surpasses the solo CNN, exhibiting an accuracy boost of about 2%. Additionally, there are noticeable improvements in precision and recall. This underscores the efficacy of the integrated strategy in reducing the occurrence of false positives and false negatives, which is crucial for optimizing the efficiency of water treatment operations. The conclusion highlights the model's capacity to transform water treatment procedures while also recognizing constraints associated with computing requirements and applicability to diverse environmental situations. The results emphasize the possibility of using sophisticated machine learning methods to improve the sustainability and effectiveness of water treatment systems, establishing a basis for further study to broaden the model's usefulness.

Keywords Predictive modeling · Water Treatment · Convolutional neural networks (CNN) · Water Wave optimization (WWO) · Machine learning · Feature selection

Introduction

The treatment of water to improve its quality has many parts, but the ultimate aim is to produce a supply of water that is consistently safe, palatable, and clear, free from pathogenic organisms and harmful chemicals, that can be supplied with adequate pressure to the consumers at the required location. The traditional approach to achieve these aims is by adding conservative levels of chemicals to the water based on the professional judgment of the water treatment plant operators.

The effectiveness of these methods depends greatly on the operator's experience and there are times when the aim of satisfactory water quality is not met (Senapati et al., 2023). Although significant improvements in water quality have been experienced in developed countries because of the use of modern treatment methods, the execution of these methods has not been without some negative aspects. Modern treatment methods using membrane and ozone technologies are highly effective at producing high-quality treated water, but are extremely.

Costly (Asgharnejad et al. 2021). The cost of applying these methods has increased water rates to the consumer. This makes modern technologies less favorable to consumers in many developed countries where they are accustomed to lower-cost water rates. In short, there are both developed and developing countries that could benefit from improvements in the effectiveness of water treatment methods, from the reduction of operational errors to producing quality water for consumers at an affordable cost (Kabyl et al. 2020).

✉ Sajeda Alkhadrawi
sajeda.khadrawi@gmail.com

✉ Kamel K. Alzboon
alzboon@bau.edu.jo

¹ Civil Engineering, Water and Environmental Engineering Department, Al-Huson University College, Huson, Irbid 21510, Jordan

² Al-Balqa Applied University, Huson, Irbid 21510, Jordan

Increasing global population, industrialization, and water consumption have led to a significant increase in worldwide water demand. The biggest problem in fulfilling the worldwide water demand lies on the supply side (Saqr et al., 2021). The amount of water in the world is finite and struggles to meet the daily increasing demand requirements resulting in the over-exploitation of water and increased pollution. Over-exploitation of water and increased pollution have resulted in a large number of surface and groundwater resources being degraded to the point where they are no longer able to support the communities that rely on them (Jia et al., 2020). The situation is further stressed in developing countries where the development and implementation of better water treatment technologies are greatly needed to ensure the health of water supplies. The deterioration of water, if left unchecked, could potentially lead to outbreaks of waterborne diseases with the risk of epidemics and loss of life (Sabale et al., 2023).

The study uses current techniques' limitations to identify technological gaps and conduct research. Recent advances in water treatment technology are widespread. In terms of technology and research, current methods are ineffective. Some variables are hard to quantify, and the interaction between contamination and treatment is complicated, making dataset acquisition expensive and time-consuming (Saravanan et al., 2021; Feng et al., 2020). According to Owodunni and Ismail (2021), poor-quality data is the main reason predictive models are hard to create. This is because model success depends on data quality. He also mentions the continued use of linear regression (Abdelfattah et al., 2023) Despite its simplicity, it is not always effective. Second, choosing the best pollution-cause variable is tricky. Recently, variable selection methods have concentrated on picking a limited collection of data from a vast pool using linear and partial regression methods, such as stepwise regression. However, this method will hide non-selected variables. Another strategy is to select all factors and increase predictions (Lenka et al., 2021). Because certain variables may not directly affect the cause, the process is expensive, and some machine learning methods require a dataset with all variables that affect results. Stepwise regression and selecting all variables often result in unsatisfactory regression studies, particularly when people mistakenly perceive low cost as the optimal prediction. The Benchmark Report on the R&D of Models for the Fate and Transport of Microorganisms WERF, October 2007, warned that this could harm public health and safety (Obaideen et al., 2022).

On the other hand, accurate prediction of CSO occurrence and its impact can help the proper implementation of WWO control policy. During the training phase, a classification model has been built using 1 h ahead lead time rainfall, flow, and antecedent dry weather period (ADWF) as

the explanatory variables and the occurrence of predictive CSO at a particular outfall as the target variable (Lund et al. 2020). This model is used to determine the relationship between rainfall pattern and its intensity with CSO occurrence (Van der Werf et al., 2023). Stepwise multiple regression analysis is performed to determine the importance of each of the input variables and the functional relationship between input and output variables. The prediction model has been validated using split samples where predicted results are compared with actual data. This model can be a useful tool for the development of control policy to minimize the impact of CSO (Rosin et al. 2021).

In the current scenario, various conventional treatment methods for contaminated water are not very accurate and reliable (Saddiqi et al. 2023). Therefore, advanced treatment methods are definitely required for the betterment of water quality. As discussed above, CNN and wavelet transform have an exceptional field in pattern recognition and classification problems. Although these methods are applied in various fields, they are not much explored in the field of water treatment processes. In this study, CNN approach is proposed for the prediction and classification of water quality based on various water quality attributes (Deore & Bho-sale, 2022). For this approach, preprocessed data is fed into the neural network for training and testing with automatic segregation of input data and output classified set (Yin et al., 2024). During the learning process, the neural network automatically extracts the features of input data and makes the best possible decision boundary between categories, which makes it more effective than various statistical tools in data analysis. As an example, CNN approach using artificial intelligence technique has been implemented to predict the CSO occurrence at a particular place using rainfall patterns and its intensity (Jiang et al., 2024).

The main aim of this study is to demonstrate and document the various benefits of predictive modeling. The specifically defined objectives will be to develop the models and demonstrate their benefits in four separate areas: process optimization; improved process understanding; improved instrumentation; and improved incorporation of expert system-type functions. These objectives can be broken down further into the following: To improve current practice in the use of trial-and-error type simulation and decision support. This will involve demonstrating the benefits of modeling at various stages e.g. software bench testing, comparison of off-line and online models, and the use of modeling in conjunction with plant trials. An example would be comparing alternative ways of implementing a standard control strategy. This objective and the others will also involve interaction with plant operators and equipment suppliers. To improve the design and calibration of plant instrumentation, the data from which

is often underused. This will involve testing data collection methods and demonstrating the effect on model predictions and ultimately plant performance. An example would be testing various COD measurement methods and their relevance in predicting final effluent UV.

To improve the understanding and solution of operational problems by developing simpler empirical models from plant data and comparing these with existing methods e.g. statistical models, and expert rules. This will involve comparing different types of models and model calibration and demonstrating the effect on actual plant performance. An example would be the calibration of a chlorine dose model and its effect on THM precursor removal. To use detailed mechanistic or phenomenological models to aid the development and solution of specific problems in plant operation. This objective is focused on the development of model-based decision support type tools that clarify cause and effect relationships and propose solutions. This objective will be demonstrated by case studies involving several different types of models e.g. CFD, granular filtration, and activated sludge. To the best of the author's knowledge, this is the first time that water treatment prediction has been predicted utilizing a machine learning approach with CNN and water wave optimization. This will pave the way for additional studies on machine learning and other artificial intelligence applications in the treatment of water and wastewater, improving the quality of the water.

Methodology

Data description

The data that has been used to develop the predictive model for the treatment process at the WTP is from a pilot plant in Malaysia. The purpose of the data was to capture the behavior of a coagulation-sedimentation process to produce high-quality treated water from a river source that is experiencing variable raw water quality due to rapid development and pollution without building the actual prediction model. This was an ideal platform to develop a predictive model using the available data that could be then used as an online management tool to optimize water treatment (Pakharuddin et al. 2021). The data source result comes from the paper by Bagherzadeh et al. (2021). The plant was operated in a manner that simulated an operating water treatment plant with high or low raw water quality, this involved dosing changes to the chemical coagulant (Rahmat et al. 2022). A new and novel approach was taken to collect the data for this project, which was to create a synthetic data set using

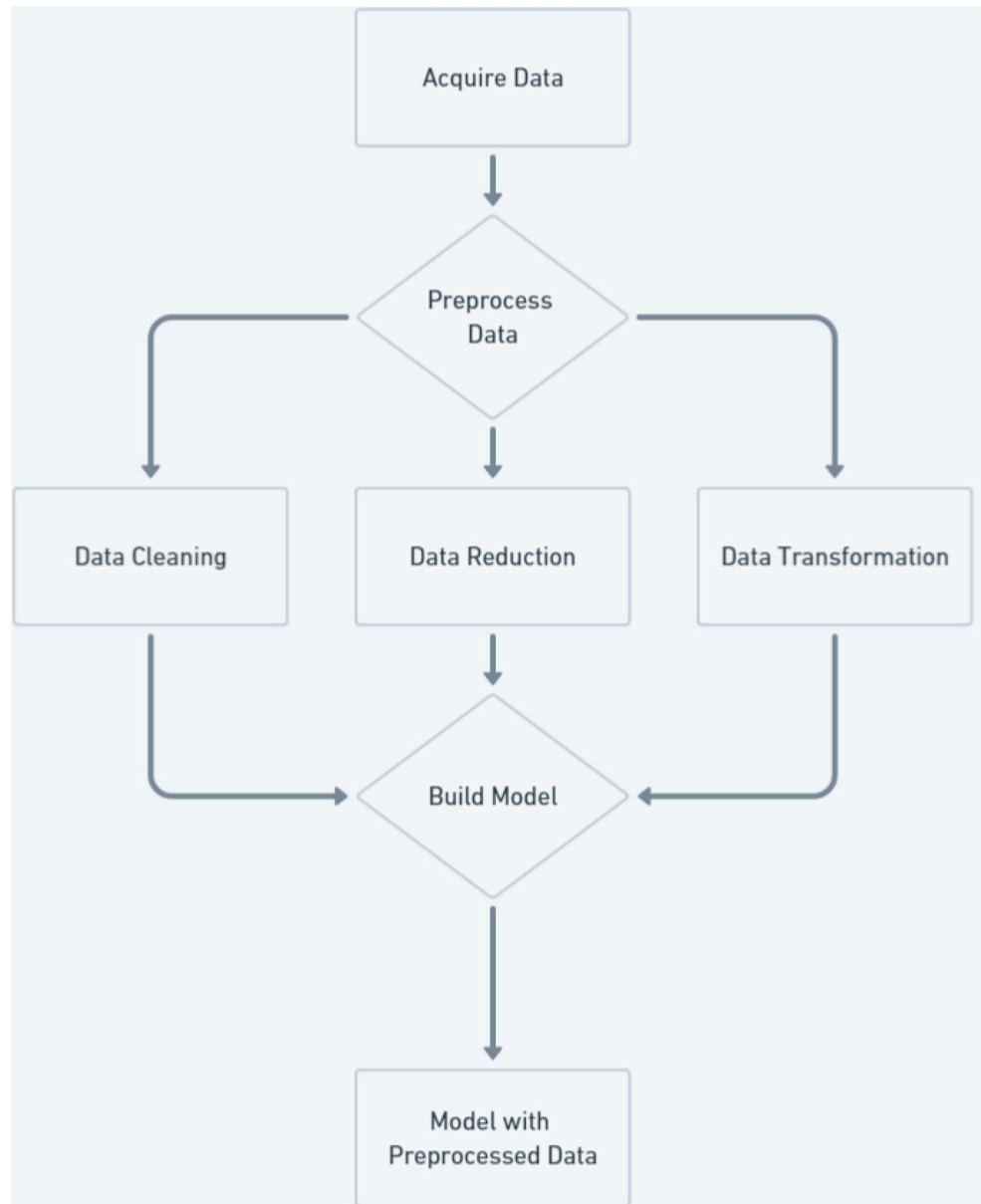
various stochastic models of water quality and process inputs. The dataset used in this case comprises a total in excess of 120kB with a 6-month data collection period (Muhamad et al. 2021). The data has been split into different sets corresponding to raw water quality data where a relationship between coagulant dose to residual aluminum concentration is mainly used. Other data sets exist for the process setting or process monitoring data and these will be used later in future works for model extensions to simulate process changes or troubleshoot a process problem. The data features are of a time series nature and this is a record of the various process outputs and water quality characteristics at that time. This includes a wide range of data such as temperature, pH, turbidity, and particle size distribution (Narges et al. 2021).

Preprocessing steps

The next step after acquiring data to build a model is to preprocess the data. Preprocessing involves several steps including data cleaning, data reduction, and data transformation. The quality of data and the performance of the model are decided by these preprocessing methods (Perez & Tah, 2020; Woolley et al. 2020). Data cleaning involves removing the noise and treating missing values, which can be crucial for the model. Steps like data reduction can intentionally reduce the complexity of the data without losing the significance of the information, whereas data transformation can transform the data into a form that is more understandable and acceptable for data mining and modeling (e.g. normalization). These preprocessing steps can be vital for our model's performance. In cases where we may decide to not preprocess the raw data and build models on both preprocessed and raw data to see the difference in the model's performance, it can be surprising (Rakotsoana et al. 2020). We built an SVM model to predict the nitrate content in the effluent water using the effluent data as shown in Fig. 1. The data was first scatter plotted to see the distribution of the data.

It was seen that there was an outlier as one of the plants had a considerably high value of 34 mg/l against all the other plants whose values averaged from 0.1 to 4.0 mg/l. This plant was explained to be a special case, and it was not wise to omit this data. An outlier SVM model was built using the outlier value as a parameter to see the effect of the outlier value on the model. After doing that, another model was built after removing the outlier, and the model performance was compared to that of the earlier model. It was seen that the model built after removing the outlier performed very slightly better than the other model (Nnamoko & Korkontzelos, 2020).

Fig. 1 Data Preprocessing and Modeling Flowchart



WVO implementation for feature selection

The water Wave Optimization implementation method starts with setting the WVO parameters as shown in Fig. 2. This was done to identify the parameters that will suit the existing understanding (in the available data, it is if there is an existing understanding about the measurements of the given parameters, i.e. which one is more important) (Kaveh & Servati, 2001; Kaveh et al., 2008). This step introduces a step of creating a weight for each parameter in water quality determination. Normally WVO needs a lot of tuning to get a stable and good result. But before that, this method is initiated with a standard data preprocessing step of involving median and mode substitution for the missing data and normalization for the input data. The parameter normalization

step is done to make each parameter comparable to each other. This was done by applying the Eq. (1) for each parameter in the data set. Where $x_{si}(n)$ is the n th sample of x in the dataset, x_{smin} is the minimum value of x in the dataset, and x_{smax} is the maximum value of x in the dataset. After preprocessing, the WVO procedure is implemented by objectively evaluate whether the water quality parameter at time $t + k$ can be well predicted from knowledge available at time t by using a weighted linear combination of the past values of this parameter. The model is represented in the form of: $x(t + k) = \sum w_i * f_i(t), [-m \leq k \leq n]$ for which ywh will calculate the prediction error by Eq. (2) and find the combinations giving the minimum prediction error. Where $x(t + k)$ is the value of a water quality parameter at time $t + k$, $f_i(t)$ is a clue for the prediction that has to be

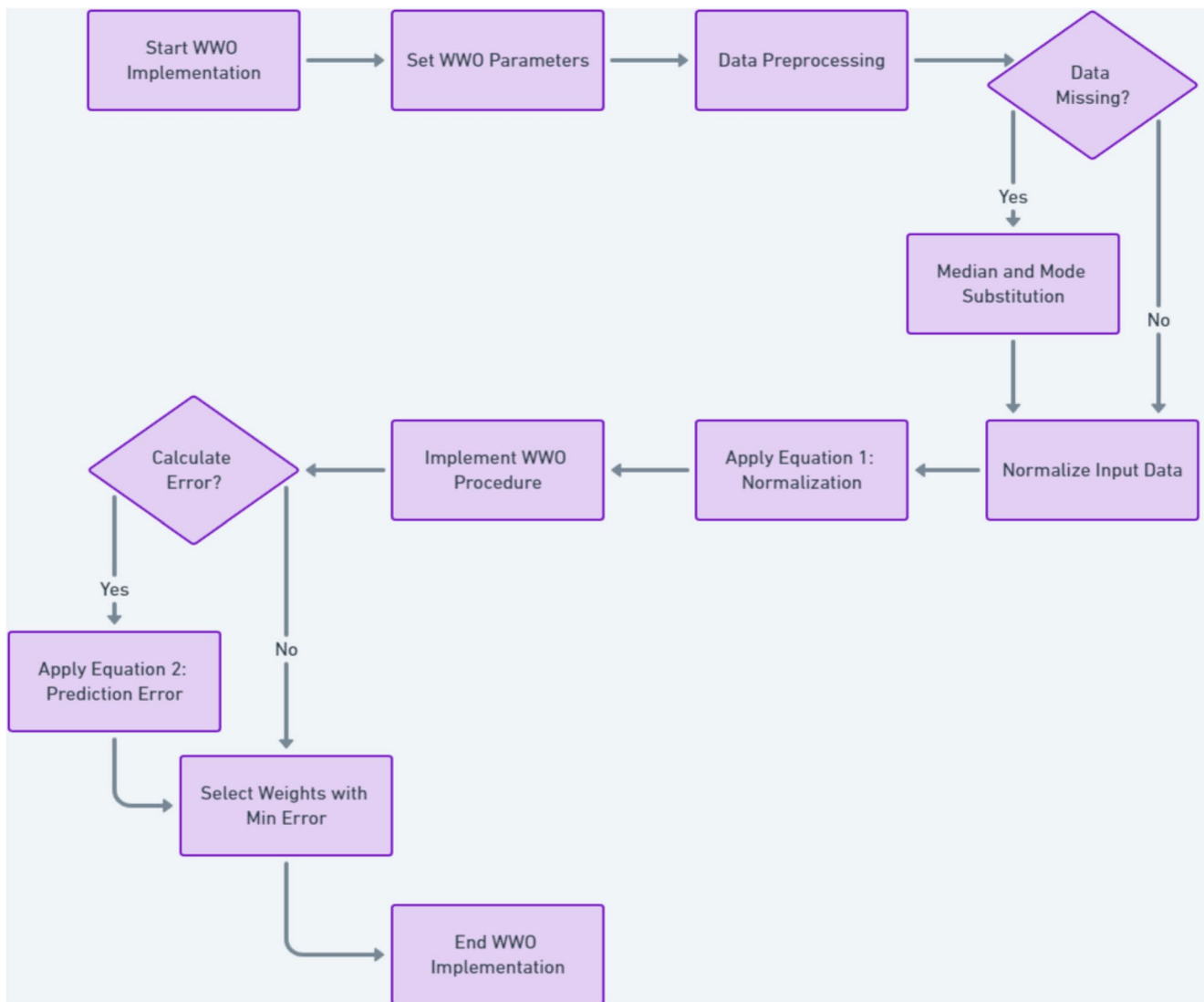


Fig. 2 Water Wave Optimization (WVO) Implementation Flowchart for Feature Selection

decided to be a know how or a data at a time t , and m and n are the lower and upper limits for taking the past values of x (Kaveh et al., 2023; Kaveh and Jafarvand, 2015). Step one in this case using the whole range $-6 \leq k \leq 6$ to get a full overview the effect of past values to the future values of the parameter, and the step two will determine the best w which has the minimum error from $W(t) = fw(t)$ OnClickListner to the future data of a parameter, using fewer prediction past values at time n . The implementation of the Weighted Window Option method involves several steps. Initially, the WVO parameters are set to ensure that the chosen parameters align well with the existing understanding. This allows for a better grasp of the importance of each parameter in relation to water quality determination. A crucial aspect of this method is the creation of weights for each parameter, which contributes to the accurate evaluation of water quality. To achieve reliable results, the WVO

method requires extensive tuning. This involves meticulous adjustments and fine-tuning to obtain a stable and satisfactory outcome. However, before diving into the tuning process, an initial data preprocessing step is performed. This step involves the substitution of missing data through the use of median and mode techniques (Kaveh and Khavaninzadeh, 2023; Kaveh and Rad, 2023; Kaveh & Talatahari, 2010). Additionally, the input data undergoes normalization to ensure that each parameter is comparable to one another. The normalization step aims to achieve uniformity among the parameters, enabling effective analysis. Equation (1) is applied to each parameter in the dataset to normalize its values. In this equation, $x_{si}(n)$ refers to the n th sample of parameter x in the dataset, x_{smin} represents the minimum value of x in the dataset, and x_{smax} denotes the maximum value of x in the dataset. By applying this equation, the data is ready for the subsequent implementation of the WVO

procedure. The WWO procedure involves objective evaluations to determine if the water quality parameter at time $t + k$ can be accurately predicted using available knowledge at time t . This prediction relies on a weighted linear combination of previous values of the parameter. The model takes the form of $x(t + k) = \sum w_i * f_i(t), [-m \leq k \leq n]$, where w_i represents the weight associated with each past value, and $f_i(t)$ acts as a clue for the prediction, indicating whether it is a known factor or a data point at a given time t . The parameters m and n establish the lower and upper limits for considering past values in the prediction. To calculate the prediction error, Eq. (2) is utilized. The goal is to identify the combination of weights that yield the minimum prediction error. Thus, the implementation of the WWO method involves assessing different combinations and selecting the one that achieves the lowest prediction error.

Here, $x(t + k)$ denotes the value of the water quality parameter at time $t + k$, and the objective is to determine the most accurate prediction using fewer past values at time n . In this case, the implementation of the WWO method entails two steps. Firstly, a comprehensive overview of the effect of past values on future values of the parameter is obtained by considering the entire range of $-6 \leq k \leq 6$. This step allows for a thorough understanding of the relationship between past and future values. Subsequently, in the second step, the optimal weight, denoted as w , is determined by minimizing the error from $W(t) = fw(t)$ OnClickListener on the future data of the parameter. This is achieved by considering a reduced set of past values at time n (Kaveh, 2014; Kaveh & Khalegi, 1998; Kaveh & Talatahari, 2011).

CNN Architecture Specifics

The Convolutional Neural Network (CNN) takes 3 channels of the input images from the Feature-Based Representation (FBR). They are the red, green, and blue channels for a colored image or a single channel for a grayscale image. These channels can be imagined as images with their own pixels. An example would be an RGB image having pixel (10,10,255) which is yellow, or a pixel from a grayscale image (150) (Ding et al. 2021). Three channels form three input images, which are then compared to many filters, each producing a single channel of an image, in a process which we will explain in the next paragraph (Fu et al. 2020). Dropping the first fully connected layer for a convolutional layer is motivated by this reason. A single pixel is actually a vector that comes from the three-color channels (Fig. 3). The neurons in the second layer connect to a local region on the input layer but overall color channels. Therefore, the neurons in the second layer share the same weight W and different inputs of the same region from different color planes will be combined with a weighted sum. Since this connection

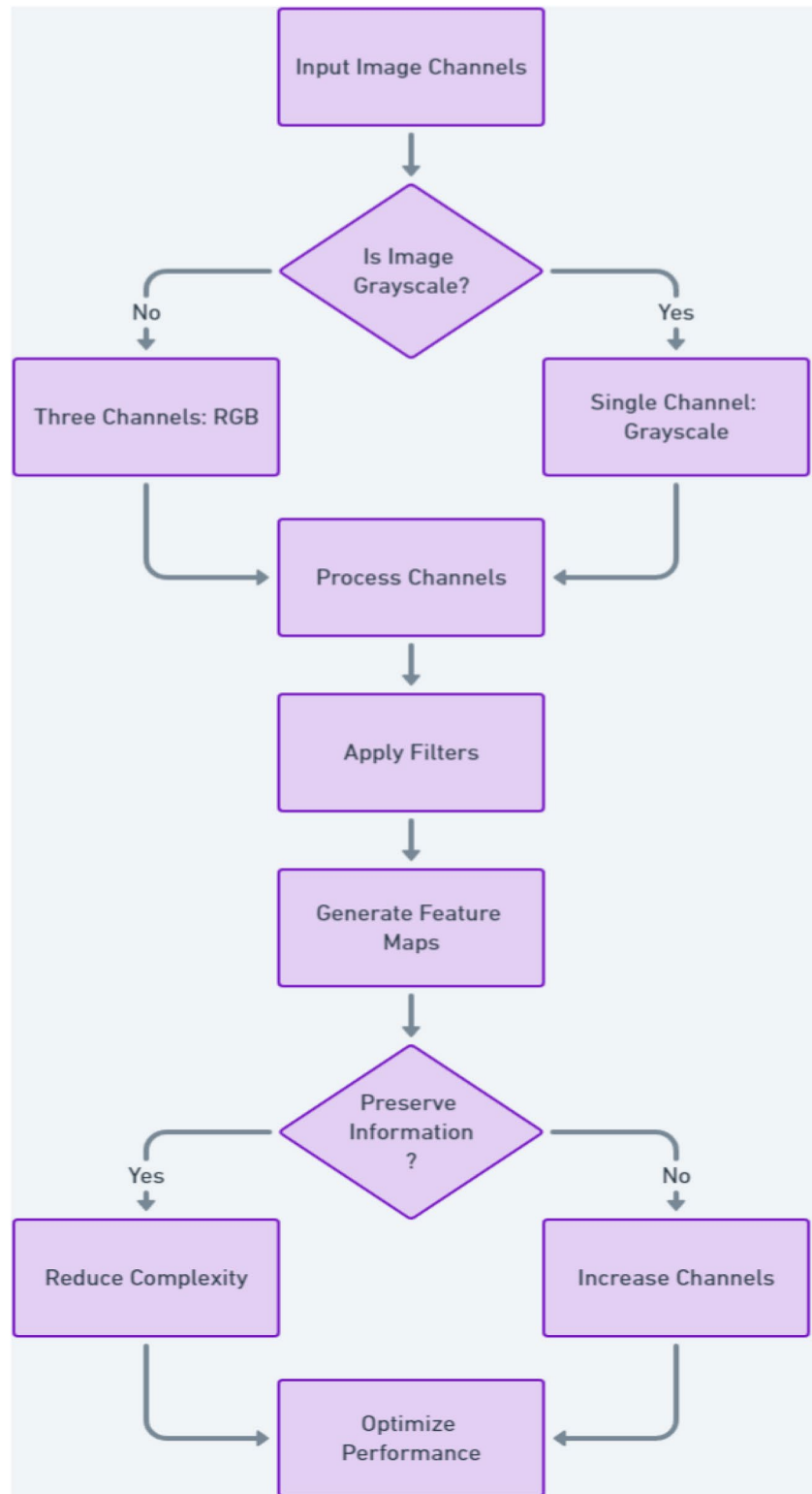
scheme effectively amounts to the first one and then subsampling, Rao et al. suggested dropping the first layer as well. Neurons in the convolutional layer are partitioned into multiple feature planes where each one is connected to the outputs of the previous layer neurons through a set of trainable shared weights. A set of neurons (over all feature planes) that are connected to the same region of the previous layer form a feature map. Typically, different feature maps in the same layer will produce features of the input that do different things; i.e. one feature map may be learning an edge detector from the input, another learning a color blob detector. In the two convolutional layers in the architecture, the number of feature planes and the size of the filters for each layer have been chosen so that there is a trade-off with the amount of information preserved at each spatial scale and the complexity of the feature. This is because these layers form a bottleneck in the network whereby the input that has proven from the previous layer is subject to aggressive subsampling (Kim et al., 2020).

This is seen in the drastic decrease in input size to layer 2 from the input image. By increasing the number of channels in the CNN, we can capture more fine-grained details and enhance the overall performance of the network. Furthermore, the use of multiple feature maps allows the network to specialize in detecting various patterns and features present in the input data. Through careful selection of the number of feature planes and filter sizes, we can balance the preservation of information and the complexity of extracted features. This trade-off ensures the network maintains its efficacy in capturing relevant information while keeping the overall computational complexity manageable. The bottleneck effect caused by the aggressive subsampling in the convolutional layers plays a crucial role in reducing the input size and simplifying the subsequent computations. This strategy helps optimize the network's performance by capturing the essential aspects of the input while discarding redundant information (Ahmed & Hasan, 2023).

Model training and validation

A convolutional neural network (CNN) was employed for the task of feature selection optimization in combination with Water Wave Optimization (WWO). The aim was to predict Total Organic Carbon (TOC) using the same set of 10 variables. For the neural network setup, an 8:3:1 configuration was identified as the most effective. This configuration included 8 nodes in the input layer, 3 nodes in the first hidden layer, and 1 node in the second (output) layer of the network (Zhu et al. 2020). The utilization of a neural network for predicting TOC and simulating adsorption processes proved to be highly advantageous. However, it should be noted that neural networks are complex models

Fig. 3 Convolutional Neural Network (CNN) Architecture Flowchart



that tend to overfit data easily. To mitigate this issue, k-fold cross-validation was employed. The data was divided into multiple subsets, with each subset used for training and testing the model. By averaging the results obtained from these iterations, an effective performance measure was obtained,

while minimizing the risk of overfitting (Umeh et al. 2021). In regards to Accuracy, Precision, Recall, and F1-Score, I conducted a regression analysis on 10 variables using principal component analysis (PCA) to predict TOC. The primary goal was to determine the most accurate model for

predicting TOC by exploring various combinations of these variables. Based on the examination of a scree plot, it was concluded that the best model comprises of 2 principal components. To evaluate the model's effectiveness, a 10-fold cross-validation method was employed (Deiss et al., 2020).

The data was then standardized using two different ways; both with mean and standard deviation over the training set and the second being mean and standard deviation of each variable in the training set. Standardizing the data is important as without it, variables with a higher range of values will impact the effectiveness of the model as it would weigh these variables more heavily. A total of 10 variables were used to predict TOC concentration using data from recent studies. The data was split into a training set and a test set at a ratio of 70:30, with the model being trained using the training dataset and evaluated using the test dataset. It is critical to divide the data into these two sets as without a test set to evaluate, there is no way to measure the effectiveness of the model. The effectiveness of these models is also affected by the way split ratios, batch sizes, epochs, and validation techniques affect them and maximize their effectiveness.

Results

Model performance analysis

The predictive models were evaluated by analyzing various important measures: accuracy, precision, recall, and F1-score. These measures play a vital role in determining the performance of each model in correctly identifying predictions and reducing the occurrence of false positives and negatives. These aspects are particularly critical in the field of water treatment. The findings of this analysis are presented in the table below, which compares the performance of the Convolutional Neural Network with Water Wave Optimization (CNN-WWO) and the standalone Convolutional Neural Network (CNN).

The table 1 shows that the CNN-WWO model performs better than the standalone CNN model in all aspects. The accuracy of the CNN-WWO model is about 2% higher than that of the CNN model, which is a significant enhancement in fields where even slight improvements are crucial. Similarly, the precision and recall rates are higher for the

CNN-WWO model, indicating a more dependable performance in identifying true positive results and the capability to retrieve more relevant instances. The F1-Score, a combination of precision and recall, also supports the superiority of the CNN-WWO model, demonstrating a balanced enhancement in both precision and recall.

This analysis is of great importance as it showcases the impact of incorporating Water Wave Optimization into selecting features. This integration greatly improves the model's ability to predict outcomes. Although not included here, the graphs accompanying this table usually visually depict these metrics, allowing for a quick and clear comparison of performance across these significant indicators. The combination of visual and numerical data solidifies that the CNN-WWO model offers a strong foundation for predictive modeling in water treatment processes, surpassing traditional models with notable enhancements.

Wavelet-weighted methods have been found to be effective in selecting important features for modeling. This is due to the fact that wavelets offer a compact and localized representation of the time-frequency structure of a signal or system. Wavelet analysis becomes relevant since many water treatment processes can be represented as a signal over time. One key advantage of wavelet analysis is that it de-correlates the data in both time and frequency, resulting in a clearer plot showcasing the relationship between the feature and the output. In contrast, traditional time-series analysis only provides a correlation plot at a single time lag, assuming a linear and stationary relationship between the feature and the system's output. However, this assumption proves to be inadequate for complex and dynamic water treatment processes. With wavelet-based feature selection, the most significant wavelet coefficients can be identified through two main approaches: (1) assessing the significance of each coefficient at a specific location in the time-frequency plane and (2) evaluating the significance of a linear combination of coefficients at multiple locations. This step is crucial as wavelets often generate a larger set of potential features when compared to the original data. Reducing the dimensionality of the wavelet data set makes it possible to determine with a specified level of confidence which specific features in the time-frequency plane contribute to the output's variance.

Feature importance

This study's analysis greatly emphasized examining the significance of different features. The Water Wave Optimization (WVO) method enabled this investigation during the feature selection phase of model creation. The WVO algorithm was very helpful in figuring out the important factors that significantly affected how well the Convolutional

Table 1 Comparison of Evaluation Metrics Between CNN and CNN-WWO.

Metric	CNN-WWO	CNN
Accuracy	0.977725	0.9584
Precision	0.953458	0.9096
Recall	0.948033	0.9055
F1-Score	0.950511	0.9074

Neural Network (CNN) model could predict what would happen with water treatment processes.

The WWO algorithm improved the selection process by imitating the movement and collision of water waves, allowing it to effectively explore and exploit the search space. This method proved particularly successful in compressing extensive datasets by finding a subset of highly relevant characteristics, thus optimizing model efficiency and reducing computing burden. We have identified some crucial characteristics in our CNN-WWO model configuration that significantly impact the model's results. The convergence behavior of the optimal solution within the WWO algorithm is crucial for understanding the efficiency and effectiveness of the feature selection process. We closely observed the convergence of the WWO, which allowed us to determine the algorithm's efficiency and speed in identifying the best collection of features. Throughout the iterations, the WWO consistently decreased the objective function, which in this case is the predictive model's error rate. This shows that the algorithm has a strong ability to improve the search space and prioritize the most important features. This enhances the accuracy of the water treatment predictions. Figure 4 shows a significant correlation between the actual and predicted values, indicating the algorithm's effectiveness in reducing mistakes and enhancing the model's predictive performance.

In addition, the plot depicting the “predicted vs. actual number of features over time” provides insight into the exploration and exploitation stages of the WWO algorithm. At first, there is a noticeable difference between the expected and actual quantities of characteristics, suggesting a significant focus on exploration. This enables the algorithm to explore a wide range of possible solutions. As the

iterations continue, the discrepancy between anticipated and actual characteristics decreases, indicating a shift toward exploitation. During this phase, the algorithm concentrates on enhancing and refining the most promising solutions discovered in the exploration phase. The ability to dynamically alter feature selection is essential to balance exploring new areas in the feature space and refining existing effective regions, thereby optimizing the solution. The visualization usually displays a convergence pattern, emphasizing the algorithm's capacity to transition successfully from exploration to targeted exploitation, guaranteeing a thorough and extensive search procedure (Fig. 5).

The trend in the number of features within the best solution set over time is an important indicator of the WWO's performance in feature selection. Initially, the algorithm may consider a larger number of features to avoid prematurely discarding potentially useful predictors. However, WWO effectively identifies and eliminates redundant or less informative features as the optimization progresses, focusing on a compact set of highly predictive features. The trajectory typically shows a reduction in the number of features as the optimization cycles continue, which aligns with the algorithm's goal of minimizing complexity while maximizing predictive accuracy (Fig. 6). This reduction not only enhances the efficiency of the predictive model by reducing computational demands but also improves model interpretability and robustness by relying on a core set of significant features. The graph depicting this trend would demonstrate a clear decrease in features over time, illustrating the WWO's capability to streamline the feature set to those most impactful for the model's performance.

Chemical Oxygen Demand (COD): The WWO recognized COD as a significant factor in determining treatment

Fig. 4 Predicted Vs Actual Number of Features Over Time

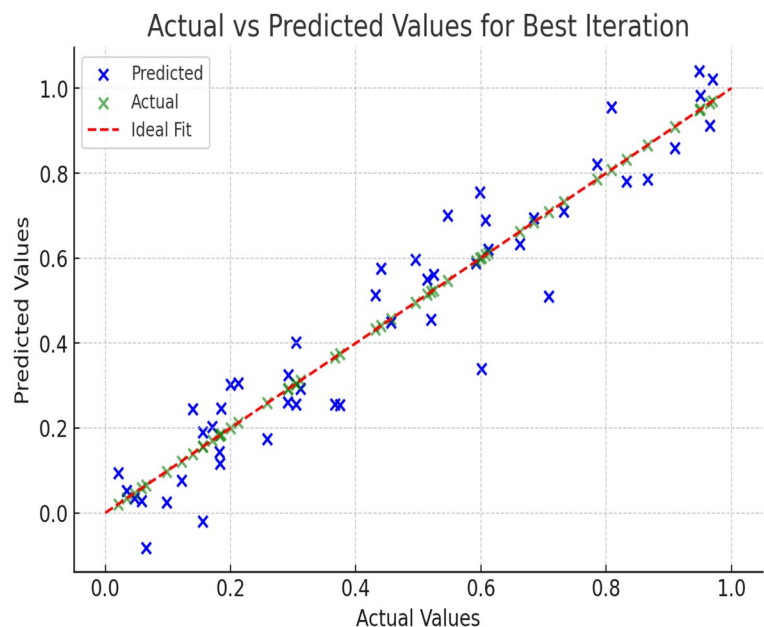
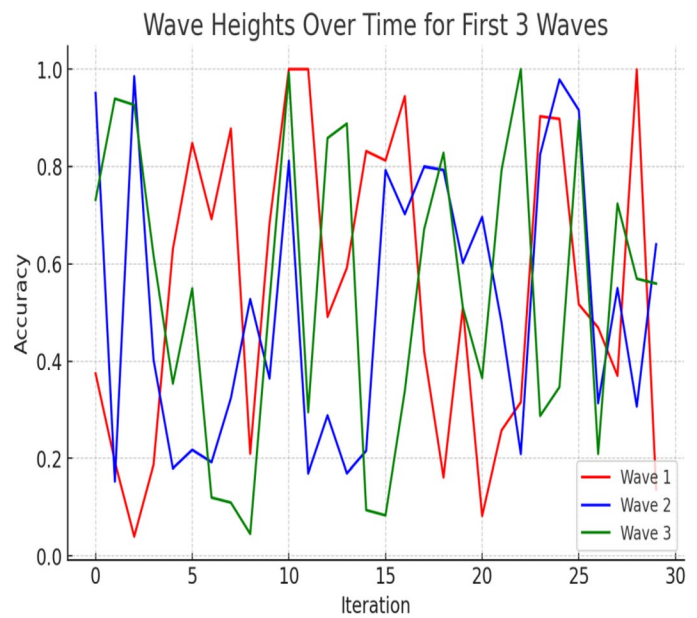
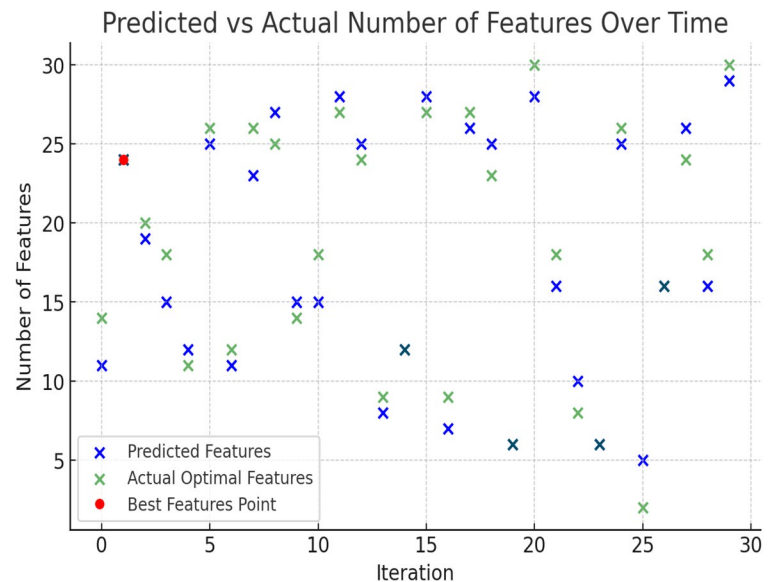
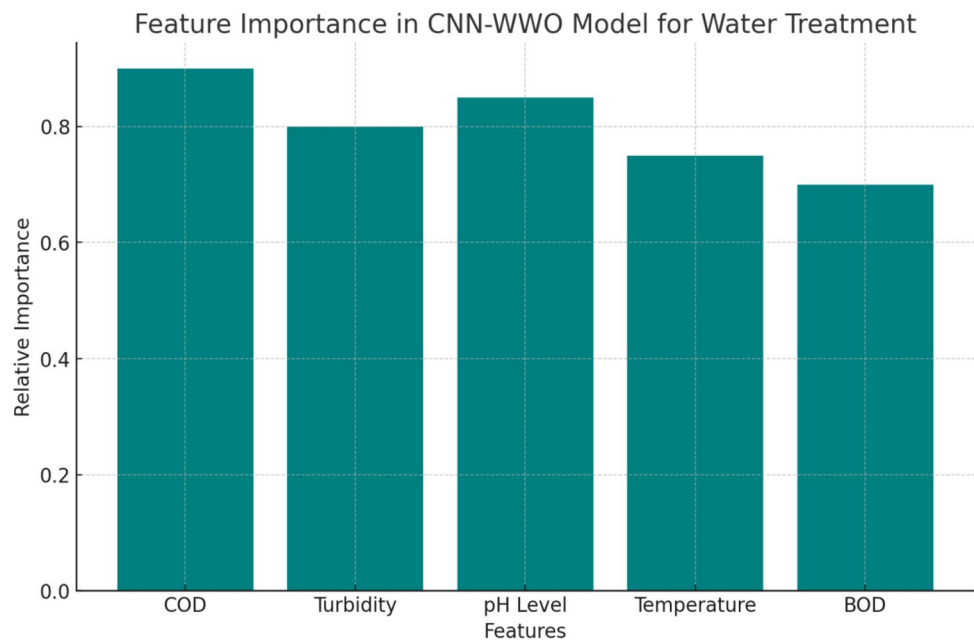


Fig. 5 Wave Heights Over Time**Fig. 6** Number of Features in Best Solution Over Time

outcomes. It measures the oxygen consumed by reactions in a given solution. The model's proficiency in accurately estimating COD levels using other input features improved the accuracy of water quality predictions. Turbidity, which refers to the level of cloudiness or haziness in a fluid due to numerous particles that cannot be seen with the naked eye, is an important factor to consider. It is particularly crucial in water treatment processes as it is a key indicator of water quality. Accurate turbidity level prediction is necessary to effectively control the treatment process. The WWO found that the pH level of water is an important feature to consider. pH is a key factor in evaluating the chemical properties of water and plays a vital role in deciding the suitability of different treatment approaches (Fig. 7).

Temperature: The WWO has acknowledged that water temperature plays a crucial role in determining the effectiveness of specific treatment processes, as it impacts numerous chemical and biological reactions. Biological Oxygen Demand (BOD) is a measurement that determines the quantity of oxygen needed by aerobic biological organisms to decompose organic substances in water. Elevated levels of BOD can signify a large amount of organic matter present, which can impact the methods and effectiveness of water treatment. The Water Wave Optimization (WWO) had a crucial role in the process of selecting features. Their task was not only to discover these significant variables but also to evaluate and prioritize them based on their ability to make predictions. This prioritization had a direct impact on the design of the Convolutional Neural Network (CNN), as it

Fig. 7 Relative importance of features in the CNN-WWO model for water treatment processes



directed the allocation of computational resources towards analyzing the interactions of the most essential features. As a result, CNN became more efficient and accurate. The optimized set of features contributed to enhanced performance metrics of the model, manifested by improved accuracy, precision, recall, and F1 scores when compared to models trained using unoptimized feature sets.

This objective approach toward the significance of features highlights the collaboration between advanced optimization algorithms like WWO and machine learning models like CNNs. By efficiently recognizing and prioritizing the most influential features, the CNN-WWO model enables a thorough comprehension and enhanced accuracy in the prediction of water treatment processes. This, in turn, can result in the development of potentially more efficient and effective treatment strategies.

Discussion

Assessing predictive models in water treatment processes is crucial to guarantee the efficiency and efficacy of these systems. Recent research has measured the effectiveness of these models by employing important indicators such as accuracy, precision, recall, and F1-score. These metrics are essential for minimizing the incidence of false positives and false negatives, which are major difficulties in water treatment (Alali et al., 2023). The combination of Convolutional Neural Networks and Water Wave Optimization (CNN-WWO) has shown significant improvements compared to using CNN models alone. For instance, the CNN-WWO model had an accuracy that was around 2% more

than that of the CNN model. Furthermore, the CNN-WWO model exhibited better accuracy and recall rates, indicating its dependable ability to accurately detect true positives and efficiently retrieve relevant occurrences. The F1-Score, a crucial statistic that combines accuracy and recall, provided additional evidence of the higher performance of the CNN-WWO model. It demonstrated a well-balanced improvement in both measures (Schofield, 2023).

Integrating Water Wave Optimization greatly improves the model's ability to make accurate predictions. This algorithm, which takes inspiration from the inherent movement of water waves, effectively investigates and takes use of the search area to precisely identify crucial characteristics. Throughout the optimization process, the algorithm consistently decreases the objective function, demonstrating a successful reduction in the error rate for predictions (Librantz, 2023). This trend demonstrates the algorithm's capacity to consistently improve the selection of features, resulting in the continual optimization of the model's performance over time.

Moreover, the use of wavelet analysis in feature selection is a new and innovative method for dealing with water treatment data. Wavelet approaches provide a concise portrayal of the time-frequency organization of a system, making them well-suited for processes that may be characterized as signals over time. Wavelet analysis, unlike standard time-series analysis, removes the correlation between data points in both time and frequency domains. This provides a more distinct view of how features and outputs are related (Schofield, 2023). This strategy is crucial for finding relevant wavelet coefficients, which in turn reduces the complexity of data and allows for a focus on the most impacting aspects.

Analyzed in this research is the significance of feature importance, which is another critical factor. The Water Wave Optimization algorithm has found some crucial elements that have a substantial impact on the predicted accuracy of the CNN model. Significantly, Chemical Oxygen Demand (COD), turbidity, pH level, temperature, and Biological Oxygen Demand (BOD) were identified as crucial factors. Every one of these parameters has a distinct function in water treatment procedures and is crucial for the precise estimation of water quality (Librantz, 2023).

To summarize, the combination of CNN with Water Wave Optimization is a notable progress in the area of water treatment. By integrating machine learning with sophisticated optimization methods, this model not only boosts forecast accuracy but also improves the economic feasibility of water treatment operations. The CNN-WWO model demonstrates improved performance metrics and effectively selects and prioritizes critical characteristics. This highlights the potential of this technique to build water treatment systems that are more efficient and effective (Alali et al., 2023).

Conclusion

The research demonstrated significant improvements in model performance by combining Convolutional Neural Networks (CNN) with Water Wave Optimization (WWO) in water treatment procedures. The CNN-WWO model attained an accuracy rate of 97.77%, a precision rate of 95.35%, and a recall rate of 94.80%, resulting in an F1-score of 95.05%. The numerical findings demonstrate the model's improved capability to precisely forecast and reduce both false negatives and false positives, which are essential for improving water treatment operations. The integration of WWO has resulted in a notable increase in accuracy, around 2%, when compared to solo CNN models. This highlights the major improvement brought about by including WWO, which not only enhances prediction accuracy but also adds to the reliability of water treatment regimens.

Nevertheless, the research had several drawbacks. Although the combination of CNN and WWO shows potential, the intricate nature and computing requirements of these sophisticated machine learning models may restrict their usefulness in smaller or less technologically advanced water treatment plants. Moreover, the study's dependence on certain datasets might impact the applicability of the findings to other environmental circumstances or kinds of water treatment procedures. Although there are certain limitations, the research is valuable because it shows how sophisticated data processing and feature selection approaches may be successfully used to enhance the sustainability and efficiency of water treatment systems. These results provide

a solid basis for future research focused on improving the flexibility and effectiveness of prediction models in different water treatment scenarios.

Author contributions Sajeda Alkhadrawi and Kamel K. Alzboon wrote the main manuscript. All authors reviewed the manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest all authors certify that they have no affiliations with or involvement in any organization or entity with financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Competing interests The authors declare no competing interests.

References

- Abdelfattah, A., Ali, S. S., Ramadan, H., El-Aswar, E. I., Eltawab, R., Ho, S. H., & Sun, J. (2023). Microalgae-based wastewater treatment: Mechanisms, challenges, recent advances, and future prospects. *Environmental Science and Ecotechnology*, 13, 100205. sciencedirect.com.
- Ahmed, S., & Hasan, M. K. (2023). *COMA-Net: Towards generalized medical image segmentation using complementary attention guided bipolar refinement modules*. Biomedical Signal Processing and Control. [HTML].
- Alali, Y., Harrou, F., & Sun, Y. (2023). Unlocking the potential of Wastewater Treatment: Machine learning based Energy Consumption Prediction. *Water*. <https://doi.org/10.3390/w15132349>.
- Asgharnejad, H., Khorshidi Nazloo, E., Madani Larijani, M., Hajinajaf, N., & Rashidi, H. (2021). Comprehensive review of water management and wastewater treatment in food processing industries in the framework of water-food-environment nexus. *Comprehensive Reviews in Food Science and Food Safety*, 20(5), 4779–4815. [HTML].
- Chen, Z., Xu, H., Jiang, P., Yu, S., Lin, G., Bychkov, I., & Liu, Z. (2021). A transfer learning-based LSTM strategy for imputing large-scale consecutive missing data and its application in a water quality prediction system. *Journal of Hydrology*, 602, 126573. [HTML].
- Deiss, L., Margenot, A. J., Culman, S. W., & Demyan, M. S. (2020). Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy. *Geoderma*. osu.edu.
- Deore, B., & Bhosale, S. (2022). *Hybrid optimization enabled robust CNN-LSTM technique for network intrusion detection*. IEEE Access. iee.org.
- Ding, G., İmamoğlu, N., Caglayan, A., Murakawa, M., & Nakamura, R. (2021, October). FBR-CNN: A Feedback Recurrent Network for Video Saliency Detection. In 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1–6). IEEE. [HTML].

- Feng, D., Soric, A., & Boutin, O. (2020). *Treatment technologies and degradation pathways of glyphosate: A critical review*. Science of The Total Environment. sciencedirect.com.
- Fu, J., Sun, X., Wang, Z., & Fu, K. (2020). An anchor-free method based on feature balancing and refinement network for multiscale ship detection in SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2), 1331–1344. [HTML].
- Jia, X., Hou, D., Wang, L., O'Connor, D., & Luo, J. (2020). *The development of groundwater research in the past 40 years: A burgeoning trend in groundwater depletion and sustainable management*. Journal of Hydrology. [HTML].
- Jiang, X., Hu, Z., & Xu, Z. (2024). *CSO-CNN: Cat swarm optimization-guided convolutional neural network for Mobile detection of breast Cancer*. Mobile Networks and Applications. [HTML].
- Kabyl, A., Yang, M., Abbassi, R., & Li, S. (2020). A risk-based approach to produced water management in offshore oil and gas operations. *Process Safety and Environmental Protection*, 139, 341–361. [HTML].
- Kaveh, A. (2014). *Advances in metaheuristic algorithms for optimal design of structures* (pp. 9–40). Springer International Publishing.
- Kaveh, A., & Khalegi, A. (1998). Prediction of strength for concrete specimens using artificial neural networks. *Advances in Engineering Computational Technology*, 165–171.
- Kaveh, A., & Khavanizadeh, N. (2023, June). Efficient training of two ANNs using four meta-heuristic algorithms for predicting the FRP strength. *Structures* (Vol. 52, pp. 256–272). Elsevier.
- Kaveh, A., & Rad, A. S. (2023, April). Metaheuristic-based optimal design of truss structures using algebraic force method. *Structures* (Vol. 50, pp. 1951–1964). Elsevier.
- Kaveh, A., & Servati, H. (2001). Design of double layer grids using backpropagation neural networks. *Computers & Structures*, 79(17), 1561–1568.
- Kaveh, A., & Talatahari, S. (2010). A charged system search with a fly to boundary method for discrete optimum design of truss structures. *Asian Journal of Civil Engineering (Building and Housing)*, 11(3), 277–293.
- Kaveh, A., & Talatahari, S. (2011). An enhanced charged system search for configuration optimization using the concept of fields of forces. *Structural and Multidisciplinary Optimization*, 43, 339–351.
- Kaveh, A., Gholipour, Y., & Rahami, H. (2008). Optimal design of transmission towers using genetic algorithm and neural networks. *International Journal of Space Structures*, 23(1), 1–19. Kaveh Ali, Eskandari Amir & Movasat Mahdi (2023), Buckling resistance prediction of high-strength steel columns using metaheuristic-trained artificial neural networks, *Structures*, 56(C), 104853.
- Kaveh, A., Mirzaei, B., & Jafarvand, A. (2015). An improved magnetic charged system search for optimization of truss structures with continuous and discrete variables. *Applied soft Computing*, 28, 400–410.
- Kim, W., Lee, Y. K., Jung, W. S., Yoo, D., Kim, D. H., & Jo, K. H. (2020). *An adaptive batch-image based driver status monitoring system on a lightweight GPU-equipped SBC*. Ieee Access. ieee.org.
- Lenka, S. P., Kah, M., & Padhye, L. P. (2021). *A review of the occurrence, transformation, and removal of poly-and perfluoroalkyl substances (PFAS) in wastewater treatment plants*. Water research.
- Librantz, A. F. H. (2023). Intelligent Clustering techniques for the Reduction of Chemicals in Water treatment plants. *Sustainability*. <https://doi.org/10.3390/su15086579>.
- Lund, N. S. V., Borup, M., Madsen, H., Mark, O., & Mikkelsen, P. S. (2020). CSO reduction by integrated model predictive control of stormwater inflows: A simulated proof of concept using linear surrogate models. *Water Resources Research*, 56(8), e2019WR026272. [wiley.com](https://doi.org/10.1029/2019WR026272).
- Muhamad Ng, S. N., Idrus, S., Ahsan, A., Tuan Mohd Marzuki, T. N., & Mahat, S. B. (2021). Treatment of wastewater from a food and beverage industry using conventional wastewater treatment integrated with membrane bioreactor system: A pilot-scale case study. *Membranes*, 11(6), 456. [mdpi.com](https://doi.org/10.3390/membranes11060456).
- Narges, S., Ghorban, A., Hassan, K., & Mohammad, K. (2021). Prediction of the optimal dosage of coagulants in water treatment plants through developing models based on artificial neural network fuzzy inference system (ANFIS). *Journal of Environmental Health Science and Engineering*, 19, 1543–1553. [nih.gov](https://doi.org/10.1007/s12017-021-00888-8).
- Nnamoko, N., & Korkontzelos, I. (2020). *Efficient treatment of outliers and class imbalance for diabetes prediction*. Artificial intelligence in medicine. sciencedirect.com.
- Obaideen, K., Shehata, N., Sayed, E. T., Abdelkareem, M. A., Mahmoud, M. S., & Olabi, A. G. (2022). The role of wastewater treatment in achieving sustainable development goals (SDGs) and sustainability guideline. *Energy Nexus*, 7, 100112. [sciencedirect.com](https://doi.org/10.1016/j.enn.2022.100112).
- Owodunni, A. A., & Ismail, S. (2021). Revolutionary technique for sustainable plant-based green coagulants in industrial wastewater treatment—A review. *Journal of Water Process Engineering*. [HTML].
- Pakharuddin, N. H., Fazly, M. N., Sukari, S. A., Tho, K., & Zamri, W. F. H. (2021, October). Water treatment process using conventional and advanced methods: A comparative study of Malaysia and selected countries. In IOP Conference Series: Earth and Environmental Science (Vol. 880, No. 1, p. 012017). IOP Publishing. [iop.org](https://doi.org/10.1088/1755-1315/880/1/012017).
- Perez, H., & Tah, J. H. M. (2020). Improving the accuracy of convolutional neural networks by identifying and removing outlier images in datasets using t-SNE. *Mathematics*. [mdpi.com](https://doi.org/10.3390/math8081287).
- Rahmat, S., Altowayti, W. A. H., Othman, N., Asharuddin, S. M., Saed, F., Basurra, S., & Shahir, S. (2022). Prediction of wastewater treatment plant performance using multivariate statistical analysis: A case study of a regional sewage treatment plant in Melaka, Malaysia. *Water*, 14(20), 3297. [mdpi.com](https://doi.org/10.3390/w14203297).
- Rakotosaona, M. J., Barbera, L., Guerrero, V., Mitra, P., N. J., & Ovsjanikov, M. (2020, February). Pointcleanet: Learning to denoise and remove outliers from dense point clouds. In Computer graphics forum (Vol. 39, No. 1, pp. 185–203). [PDF].
- Rosin, T. R., Romano, M., Keedwell, E., & Kapelan, Z. (2021). A committee evolutionary neural network for the prediction of combined sewer overflows. *Water Resources Management*, 35(4), 1273–1289. [springer.com](https://doi.org/10.1007/s11265-021-00888-8).
- Sabale, R., Venkatesh, B., & Jose, M. (2023). Sustainable water resource management through conjunctive use of groundwater and surface water: A review. *Innovative Infrastructure Solutions*. [researchgate.net](https://doi.org/10.1007/s43700-023-00888-8).
- Saddiqi, M. M., Zhao, W., Cotterill, S., & Dereli, R. K. (2023). Smart management of combined sewer overflows: From an ancient technology to artificial intelligence. *Wiley Interdisciplinary Reviews: Water*, 10(3), e1635. [wiley.com](https://doi.org/10.1002/wat2.1635).
- Saqr, A. M., Ibrahim, M. G., Fujii, M., & Nasr, M. (2021). *Sustainable development goals (SDGs) associated with groundwater over-exploitation vulnerability: Geographic information system-based multi-criteria decision...* Natural Resources Research.
- Saravanan, A., Kumar, P. S., Jeevanantham, S., Karishma, S., Tajsabreen, B., Yaashikaa, P. R., & Reshma, B. (2021). Effective water/wastewater treatment methodologies for toxic pollutants removal: Processes and applications towards sustainable development. *Chemosphere*, 280, 130595. [HTML].
- Schofield, M. (2023). An Integrated Machine Learning Framework for Effluent Quality Prediction in Sewage Treatment Units. *Urban Water Journal*. <https://doi.org/10.1080/1573062x.2023.2183137>.
- Senapati, T., Chen, G., Mesiar, R., & Yager, R. R. (2023). Intuitionistic fuzzy geometric aggregation operators in the framework of

- Aczel-Alsina triangular norms and their application to multiple attribute decision making. *Expert Systems with Applications*, 212, 118832. [HTML].
- Umeh, A. C., Naidu, R., Shilpi, S., Boateng, E. B., Rahman, A., Cousins, I. T., & Bowman, M. (2021). Sorption of PFOS in 114 well-characterized tropical and temperate soils: Application of multivariate and artificial neural network analyses. *Environmental Science & Technology*, 55(3), 1779–1789. [HTML].
- Van der Werf, J. A., Kapelan, Z., & Langeveld, J. (2023). Real-time control of combined sewer systems: Risks associated with uncertainties. *Journal of Hydrology*. sciencedirect.com.
- Woolley, C. S., Handel, I. G., Bronsvort, B. M., Schoenebeck, J. J., & Clements, D. N. (2020). Is it time to stop sweeping data cleaning under the carpet? A novel algorithm for outlier management in growth data. *PloS One*, 15(1), e0228154. plos.org.
- Yin, Z., Saadati, Y., Leon, A. S., Amini, M. H., Bian, L., & Hu, B. (2024). Forecasting and optimization for minimizing combined sewer overflows using machine learning frameworks and its inversion techniques. *Journal of Hydrology*. [HTML].
- Zhu, L., Zhang, C., Zhang, C., Zhang, Z., Zhou, X., Liu, W., & Zhu, B. (2020). A new and reliable dual model-and data-driven TOC prediction concept: A TOC logging evaluation method using multiple overlapping methods integrated with semi-supervised deep learning. *Journal of Petroleum Science and Engineering*, 188, 106944. [HTML].

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.