**Design of deep learning application for prediction of permeance and rejection rate using engineering features of membranes: A strategy for clean water applications**

Imtiaz Afzal Khan1, Almustapha A. Wakili2, Syed Muzzamil Hussain Shah1, Jamilu Usman1, Lukka Thuyavan Yogarathinam 1, A.G. Usman3,4, Sani I. Abba2\*, Isam H. Aljundi1,5

1Interdisciplinary Research Centre for Membrane and Water Security, King Fahd University of Petroleum and Minerals Dhahran 31261, Saudi Arabia

2Department of Computer and Information Sciences, Towson University, MD, USA, [almustaphawakili@gmail.com](mailto:almustaphawakili@gmail.com)

3Operational Research Centre in Healthcare, Near East University, Mersin 10, 99138 Nicosia, Turkey

4Department of Analytical Chemistry, Faculty of Pharmacy, Near East University,

TRNC, Mersin 10, 99138, Nicosia, Turkey

5Department of Chemical Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

Corresponding author: [sani.abba@kfupm.edu.sa](mailto:sani.abba@kfupm.edu.sa)

**Abstract**

Reliable, cutting-edge tools to optimize water treatment processes and ensure environmental sustainability and regulatory compliance are paramount. This study aims to evaluate the performance of various machine learning (ML) models in predicting the permeance and rejection rates of RO membranes. The models assessed include convolutional neural network (CNN), long short-term memory (LSTM), multilayer perceptron (MLP), and random forest (RF). A comprehensive dataset comprising key membrane characteristics swelling weight, molecular weight cutoff (MWCO), contact angle, and zeta potential was utilized to train and validate these models. The several evaluation criteria, statistical matrices and 2D visualization were used to assess the model prediction accuracy. Furthermore, reliability and stationarity analyses were conducted using the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test to ensure data consistency and the robustness of model predictions. The RF model demonstrated superior performance in predicting permeance, achieving mean squared error (MSE) of 0.0024, mean absolute percentage error (MAPE) of 2.2490%, mean absolute error (MAE) of 0.0226, percent bias (PBIAS) of -0.0011, and root mean squared error (RMSE) of 0.0489. The CNN, LSTM, and MLP models exhibited higher errors with MAPEs of 191.9223%, 156.1755%, and 200.9758%, respectively. For rejection prediction, the RF model again outperformed the others with MSE of 0.0328, MAPE of 88.1296%, MAE of 0.1105, PBIAS of 0.0445, and RMSE of 0.1810, while the CNN, LSTM, and MLP models showed MAPEs of 136.4040%, 133.9138%, and 108.9924%. The high accuracy and reliability of the RF model highlights its potential in optimizing RO membrane performance, contributing to improved water quality management, regulatory compliance, and sustainable development. Future research should focus on enhancing these models by incorporating additional variables, refining algorithms, and conducting real-world pilot studies. Integrating ML with advanced technologies such as the Internet of Things (IoT) and blockchain can further enhance data collection and processing, promoting transparency and efficiency in water management systems.

**Keywords:** Artificial intelligence; Desalination, Membrane; Clean water; Deep learning

**Introduction**

Urbanization, climate change, and economic transformations, all human-driven factors, are increasingly jeopardizing the availability of freshwater and exacerbating water scarcity [1]. A recent World Meteorological Organization (WMO) report emphasizes a distressful predict that the number of people experiencing water scarcity for at least one month each year is expected to increase significantly by the 2050 [2]. Membrane separation methods, including reverse osmosis (RO) and nanofiltration (NF), are increasingly acknowledged as sustainable and efficient approaches for treating contaminated water sources and producing clean water [3,4]. The thin film composite (TFC) membrane is the primary component of the RO and NF processes, playing a crucial role in improving freshwater production capacity. Membrane fouling, caused by the accumulation of foulants on the membrane surface, significantly hinders water permeability and salt rejection in RO and NF processes. The surface characteristics of membranes, such as hydrophilicity, surface charge, and roughness, play a significant role in determining the efficiency of water transport and the ability to rejection of salts [5]. In membrane desalination, machine learning (ML) models are gaining significant interest for their capacity to improve filtration performance, predict flux dynamics, and elucidate the underlying filtration mechanisms.

ML, a component of artificial intelligence (AI), utilizes various frameworks and algorithms to enhance the analysis and interpretation of complex water purification as well as desalination systems, improving data efficiency, capacity, sensitivity, and consistency while reducing processing time [6]. Artificial neural networks (ANN) have attracted significant attention in the field of membrane separation based wastewater treatment and desalination, inspired by the operational principles of the biological brain [7]. Convolutional neural networks (CNN), recurrent neural networks (RNNs), feedforward neural networks (FNNs), and radial basis function networks (RBFNs) represent the principal architectural types of neural networks. These architectures facilitate accurate predictions of output responses in the evaluation of flux, rejection rates, fouling, and optimization of energy consumption in desalination [8–10]. Lately, deep learning architectures like CNNs and RNNs use multi-layered neural networks to progressively extract abstract features from raw data, enhancing the analysis and optimization of membrane processes [11,12]. Alardhi et al., studied the optimization of hydrodynamic parameters on RO membrane using ANN for the treatment of thermal power station wastewater [13]. Park et al., investigated the use of CNN ML tools to assess *in situ* membrane fouling in nanofiltration (NF) and RO membranes, utilizing optical coherence tomography [14]. The CNN model was shown to be a more effective ML tool for analyzing the formation of organic fouling layers compared to the cake layer mathematical model. Luo et al. explored the use of computational fluid dynamics combined with multilayer artificial neural networks (MLN) to design and optimize module configurations and operating conditions in seawater reverse osmosis (SWRO) system [15]. The optimized design and conditions, determined through a supercomputing simulation approach, demonstrated reduced energy consumption under conditions of high-performance SWRO membranes.

Ensemble modeling is a robust ML approach that enhances the optimization of membrane separation processes through techniques like bagging, boosting, and stacking, thereby improving predictive accuracy and operational efficiency. Tayyebi et al., utilized Random Forest (RF) and CatBoost ML algorithms, enhanced by explainable artificial intelligence (XAI), to analyze the impact of monomers on the performance of TFC RO membranes [16]. The application of Shapley Additive exPlanations (SHAP) analysis in monomer selection surpassed the trade-off between water permeability and NaCl rejection. Yeo et al., employed the gradient boosting trees (GBT) ML algorithm to explore the effects of nanomaterials on customizing polyamide membranes, aiming to enhance both water permeability and salt rejection [17]. The bivariate partial dependence plots analysis identified that the most effective parameters for enhancing desalination efficiency include hydrophilic nanoparticles with a diameter of 5-7 Å, measuring 150 nm in size, and at a loading concentration of 0.1 wt%. Talhami et al. investigated both single and ensemble ML algorithms to predict the desalination performance and membrane flux of TFC and cellulose triacetate (CTA) membranes [18]. Among the ML tools, extreme gradient boosting demonstrated superior performance, achieving more accurate predictions and lower error metrics.

Research comparing the effectiveness of deep learning neural networks and ensemble methods in assessing RO membrane desalination performance is scarce. This study was designed to evaluate the performance of various ANN tools, including multilayer perceptron (MLP), CNN, long short-term memory (LSTM), and ensemble RF-ML algorithm, in predicting the flux and rejection efficiency. The input parameters selected for predicting the target response of RO membrane flux and salt rejection efficiency were swelling (SW), contact angle (CA), zeta potential (ZP), and molecular weight cutoff (MWCO). The study primarily focused on analyzing the features of input parameters and comparing the predictions of ML tools on desalination efficiency, specifically in terms of flux and salt rejection. Furthermore, the study conduct comprehensive reliability and stationarity analyses to ensure the consistency and robustness of the data used in predicting permeance and rejection rates of RO membranes. This involved applying the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test. The ADF test was utilized to check for the presence of unit roots in the time series data, thereby confirming stationarity by testing the null hypothesis that a unit root is present in the series. Similarly, the PP test was employed as an alternative to the ADF test, providing robustness against heteroscedasticity and serial correlation by adjusting the test statistics using non-parametric correction. Conducting these tests ensured that the input data for the ML models was stable and reliable over time, thereby enhancing the accuracy and reliability of the model predictions. This step was crucial in building a robust predictive framework for optimizing water treatment processes and ensuring environmental sustainability

**2. Proposed ML Methodology and Data Source**

The proposed ML methodology for predicting the permeance and rejection rates of RO membranes involves normalization, 10-fold cross-validation, and data partitioning to ensure efficient and reliable model training [19]. Normalization scales the input data to a standard range, improving the convergence speed and ensuring equal contribution from all input features like SW, MWCO, CA, and ZP. To evaluate model performance, a 10-fold cross-validation technique is employed, partitioning the dataset into 10 equal-sized folds where each fold is used once as a validation set while the remaining nine folds are used for training [20]. This method provides a comprehensive assessment of the model’s performance, reducing overfitting risks and ensuring generalization to unseen data. Furthermore, the dataset is divided into 70% for training and 30% for testing, ensuring sufficient data for learning while providing a substantial portion to evaluate the model’s predictive power. This combination of normalization, cross-validation, and data partitioning enhances the reliability and effectiveness of the ML models in predicting RO membrane performance, thereby optimizing water treatment processes. We initially conducted data mining from more than 100 articles (total number of 135,616 data points), to analyze the performance of membranes. These articles were sourced from popular search engines such as Web of Science, Google Scholar, and specific publishers including Elsevier (Scopus), ACS, Wiley, and Springer. The collated information includes details on membrane performance, properties, and process conditions, all of which are thoroughly summarized in **Table 1.**

Given the substantial interest of end-user industries in commercially available membranes, our primary focus was on gathering relevant data in this domain. It’s important to note that achieving accurate predictions of commercial membrane performance remains a significant milestone yet to be fully realized. The study’s target to compare different ML methods necessitates a comprehensive understanding of how these input characteristics influence the output parameters. By analyzing SW, CA, ZP, and MWCO, the models can be trained to recognize patterns and predict the membrane’s performance accurately. The variability in the output parameters, especially permeance, suggests that there are significant factors influencing membrane efficiency, making it an ideal figure for deep learning approaches that excel in handling complex, non-linear relationships. The summarized data highlights the critical input characteristics and their influence on the RO membrane performance, providing a solid basis for evaluating and improving predictive models in clean water applications.

**Table 1:** Summary of characteristic parameters used for the analysis with basic input parameters range

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Summarized data** | | | |
|  | **MIN** | **MAX** | **AVG** | **SD** |
| **Characteristics of membranes** | | | | |
| Swelling thickness (mm) | 0.841167 | 1.888928 | 1.090378 | 0.274628 |
| Swelling weight (mg) (SW) | 1.19 | 1.888928 | 1.467204 | 0.191824 |
| MWCO (g/mol) | 150 | 500 | 283.6081 | 107.3136 |
| Contact angle (o) (CA) | 34 | 87 | 62.48428 | 15.73444 |
| Zeta potential (meV) (ZP) | -37.5 | -1 | -18.6885 | 14.14667 |
| **Output parameters** | | | | |
| Permeance (LMH/bar) | 0.066138 | 64.75434 | 3.312112 | 9.010616 |
| Rejection (%) | -0.41265 | 1 | 0.493678 | 0.364376 |

**2.1 Multi-layer Perceptron (MLP)**

The MLP are a common type of artificial neural network (ANN) known for its ability to handle complex problems. Unlike ANNs, MLPs have specific layers, for instance, an input layer for data, hidden layers to process information, and an output layer for results [21]. Specifically, the input layer nodes establish connections primarily with the hidden and output layers [22]. Signals undergo processing from the input to the output layer, facilitated by sequential mathematical operations involving biases and weights [23]. The schematic of MLP is shown in Fig. 1Training algorithms, like the Levenberg-Marquardt algorithm, help the MLP learn by minimizing the difference between what it predicts and the actual results. The weights and biases are adjusted iteratively using techniques like backpropagation to minimize the error between predicted and actual [24]. The commonly used equation to describe the output of a neuron in a MLP can be given as:

(1)

where, σ is the activation function, such as the sigmoid function or the rectified linear unit, ​ are the weights of the connections between the neuron’s input and the previous layer’s neurons. are the inputs from the previous layer’s neurons, b is the bias term and n is the number of inputs to the neuron.

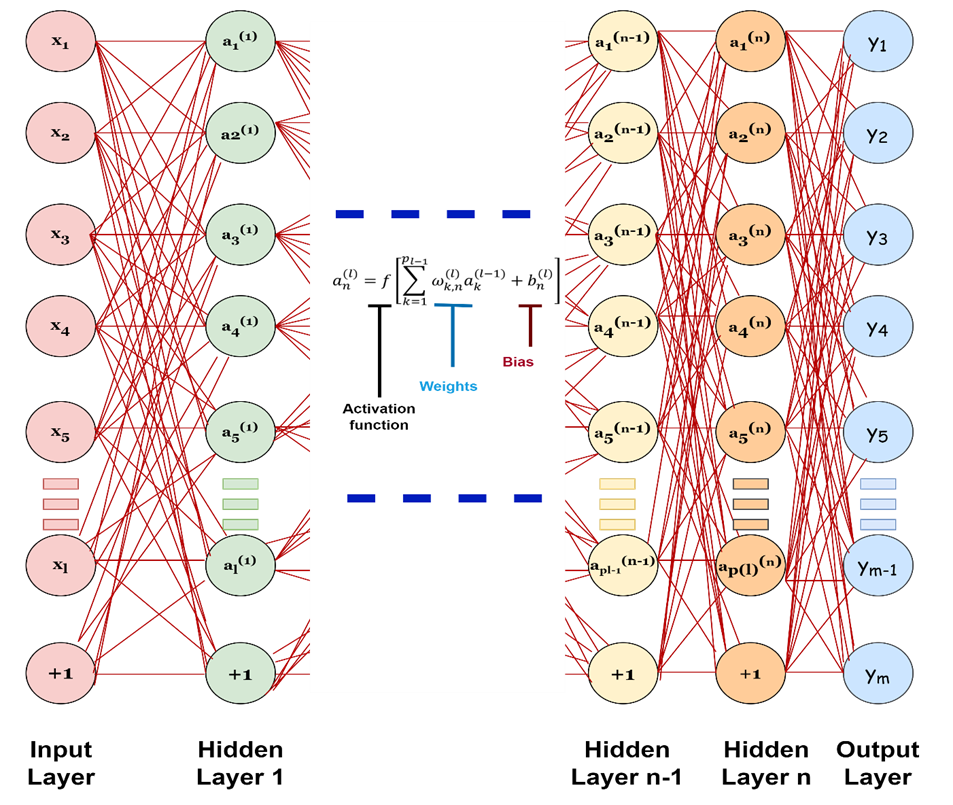


Fig. 1: Three-layer multi-layer perceptron structure

**2.2 Long Short-Term Memory networks (LSTM)**

LSTMs are a special type of neural network that tackles the problems of vanishing and exploding gradients during training, common issues in regular RNNs (Recurrent Neural Networks) [25]. This allows LSTMs to learn from longer sequences than traditional RNNs. Fig.2 shows the schematic of LSTM architecture. Like RNNs, LSTMs are built using repeating modules, but each module in an LSTM contains a special memory cell that helps it remember information for longer periods [26]. Due to its multilayer perception and strong nonlinear mapping capabilities, recurrent neural networks (RNNs) have found wide applications in solving various practical problems such as recognition and prediction [27,28]. However, traditional RNNs often face challenges related to vanishing gradients and long-term dependency issues. In response to these challenges, the LSTM network, a specialized type of RNN, has been developed. LSTM addresses these problems by introducing special gates that regulate the flow of information within the network. This design significantly reduces the likelihood of vanishing gradients, thereby improving the network’s ability to retain long-term memory. An LSTM network comprises three types of gates: the forget gate, input gate, and output gate [29]. The forget gate determines whether information from the previous time step should be retained, while the input gate decides which information from the current input vector should be stored in the cell state. The cell state is then updated based on these decisions. Finally, the output gate controls the information output at the current time step. Through these mechanisms, LSTM networks can effectively handle long-term dependencies and are widely used in tasks requiring sequential data processing, such as natural language processing and time series prediction[30]. LSTM contains three types of gates, i.e., the forget-gate, input-gate, and output-gate. The first layer of forget-gate determines whether the information could pass through the cell state, for instance:

(2)

Where, is the forget gate at time t; is the sigmoid activation function; denotes the weight; denotes the input value; denotes the output value at time t − 1; denotes the bias term.

The second input gate decides which information should be stored in the cell state from the current input vector. The specific expressions are listed as follows:

= (3)

t = (4)

where stands for the input gate at timet; and denote the weights; (⋅) is the activation function; and denote the bias terms. The cell state at time t can be updated by:

= ⊙ + ⊙ t (5)

where ⊙ represents the element-wise multiplication.

The third layer can provide the output information in the current time step. The specific expression is presented as follows:

= (6)

where is the output gate at time t, denotes the weight, denotes the bias term. Then, the output value of the cell can be written as.

= ⊙ () (7)

= (8)

where f (⋅) is the activation function of the output layer.

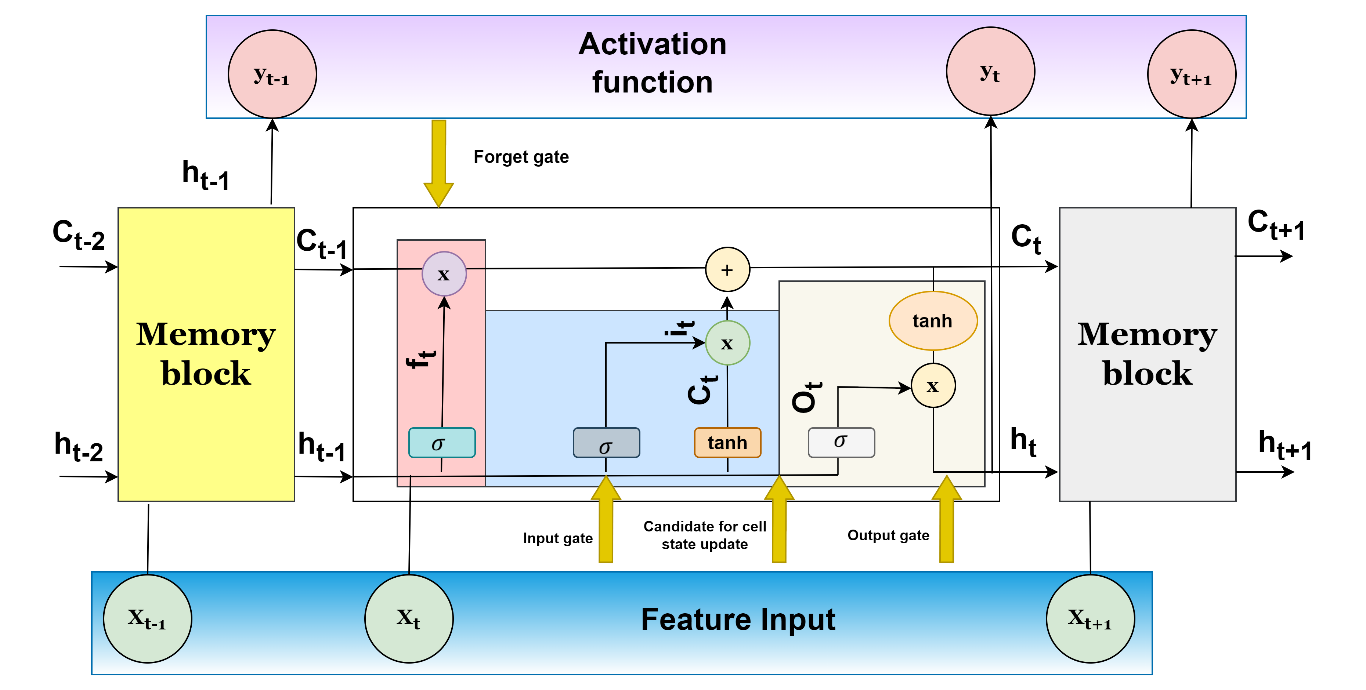
****

Fig. 2: Long Short-Term Memory networks (LSTM) Structure [29]

**2.3 Convolutional Neural Network (CNN)**

The CNN represents a key component of deep learning methodology, featuring a multi-layered feed-forward neural network architecture [31]. Its fundamental structure comprises several key elements, including convolution layers, rectified linear unit (ReLU) layers, pooling layers, and fully connected (FC) layers, as illustrated in Fig. 3. The primary role of the convolution layer is to extract local features from preceding layers [32]. Meanwhile, the ReLU layer conducts element-wise activation, enhancing the network’s nonlinear capabilities [33]. Pooling layers are incorporated for down sampling, with max-pooling often employed to condense feature maps by integrating linguistically similar features[34,35]. To address overfitting concerns, dropout mechanisms are implemented, selectively removing neurons from the CNN architecture. Ultimately, the FC layer computes class score values ranging from 0 to 1, facilitating classification decisions. Notably, SoftMax layers are typically integrated by default for this purpose.

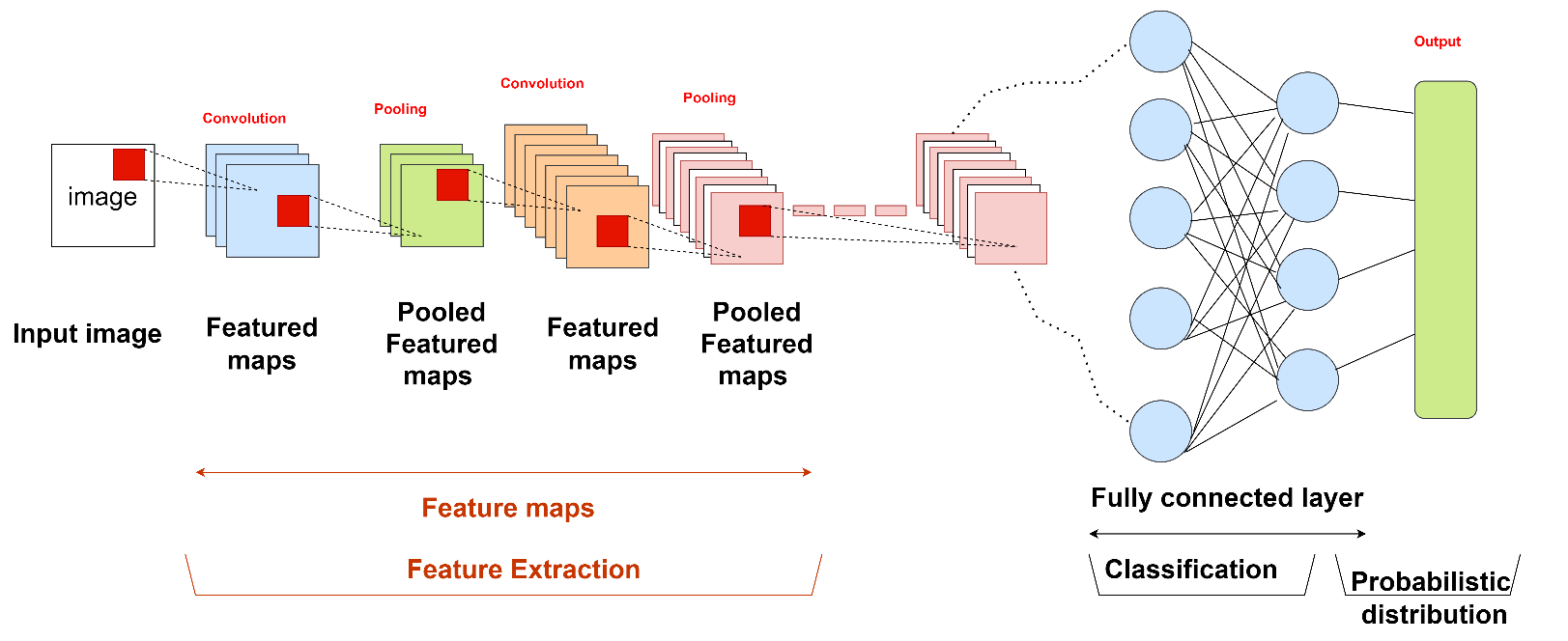
****

Fig. 3: The Convolutional Neural Network (CNN) Structure [31]

**2.4 Random Forest (RF)**

The RF algorithms is a powerful ensemble learning technique used in artificial intelligence for classification and regression tasks. It operates by constructing a multitude of decision trees during the training phase. Each decision tree is trained on a random subset of the dataset and makes independent predictions. Fig. 4 shows the schematic of RF. During inference, the predictions from each tree are aggregated to produce a final prediction, either through voting (for classification) or averaging (for regression) [36]. This ensemble approach enhances the robustness and accuracy of predictions, while also mitigating overfitting. RF are known for their versatility, scalability, and capability to handle high-dimensional data with complex interactions [37]. They are widely utilized in various domains including finance, healthcare, and marketing for tasks such as customer segmentation, anomaly detection, and risk assessment [38]. The RF algorithm predicts the average output for new input data based on a training set comprised of observed input-output data. Compared to conventional statistical techniques, RF exhibit lower susceptibility to overfitting, particularly beneficial when working with limited sample sizes and multiple potential predictors [39,40]. Notably, when confronted with a large number of predictor variables, RF proves particularly advantageous. In the simulation process of the RF model, determining two key parameters is essential: the number of trees constituting the forest (ntree) and the number of predictors evaluated at each node (mtry), as depicted below [41].

(9)

Where M denotes the number of input variables specified in the original dataset.

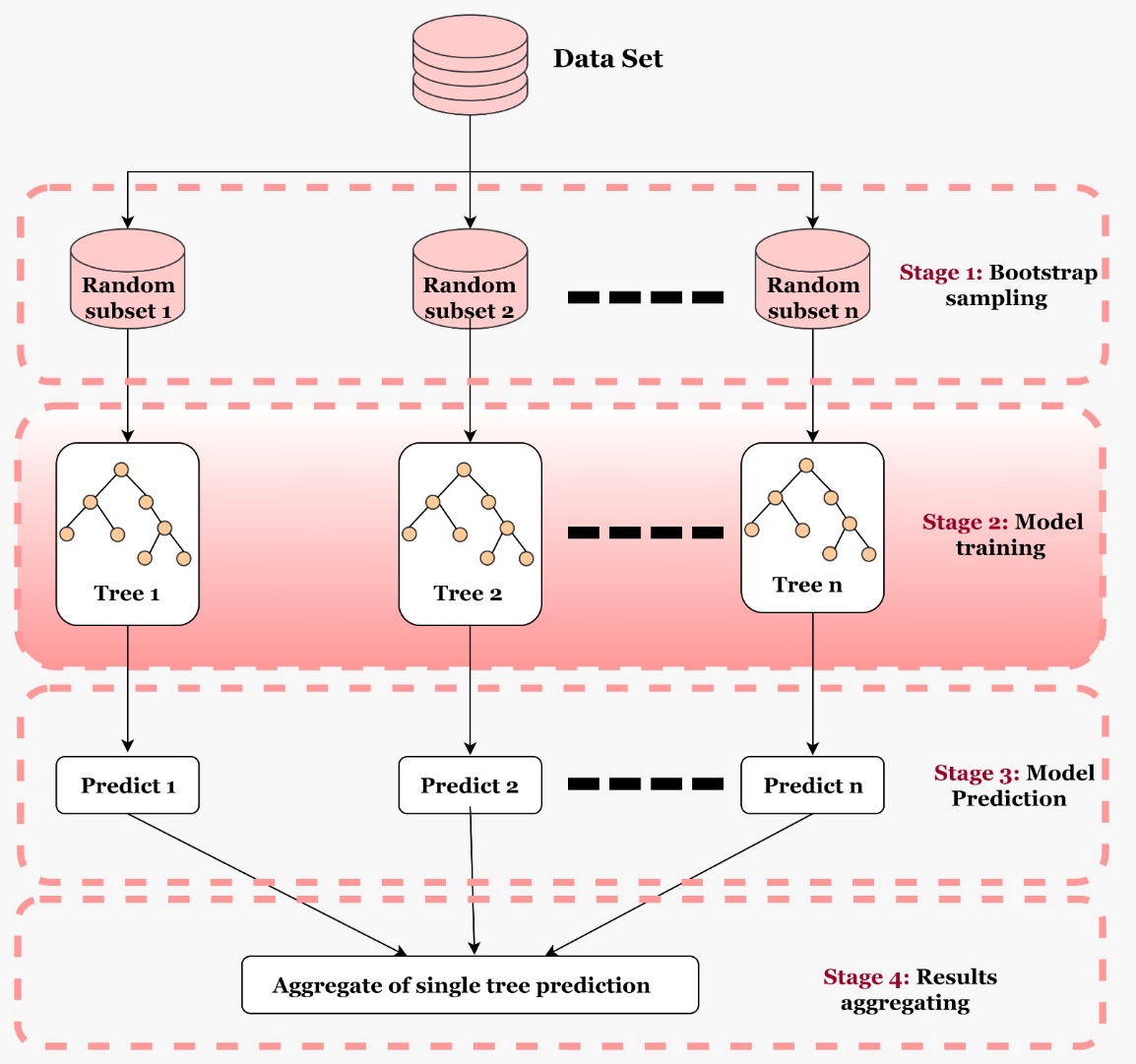


Fig. 4: Random Forest (RF) Structure [38]

**3. Application of Result and Discussions**

**3.1 Exploratory and Dependency**

Table 1 serves as the foundation for developing a deep-learning model to predict permeance and rejection rates of RO membranes. The wide range of values in both input and output parameters indicates the variability in membrane performance, which is essential for training robust predictive models [42]. Table 1 provides crucial insights into the characteristics of membranes and their performance in RO desalination processes. The input parameters include SW thickness (0.841 to 1.889 mm, AVG: 1.090 mm, SD: 0.275 mm), SW weight (1.19 to 1.889 mg, AVG: 1.467 mg, SD: 0.192 mg), MWCO (150 to 500 g/mol, AVG: 283.608 g/mol, SD: 107.314 g/mol), contact angle (34° to 87°, AVG: 62.484°, SD: 15.734°), and ZP (-37.5 to -1 mV, AVG: -18.689 mV, SD: 14.147 mV). These characteristics influence the membrane’s ability to absorb water, selectivity, hydrophilicity, and surface charge, affecting its permeability and fouling resistance. The output parameters include permeance (0.066 to 64.754 Lm-2h-1bar-1, AVG: 3.312 Lm-2h-1bar-1, SD: 9.011 LMH/bar) and rejection rate (-0.413% to 1%, AVG: 0.494%, SD: 0.364%), which measure the membrane’s efficiency in allowing water passage and solute removal. This variability in input and output parameters forms the basis for developing and evaluating deep learning models, such as MLP, CNN, LSTM, and ensemble RF, to predict permeance and rejection rates, thereby enhancing the performance assessment of RO membranes in clean water applications.

Fig.5 reveals the dependency matrix reveals the relationships between the input characteristics of membranes and the target output parameters. The SW shows a positive dependency with MWCO (0.308) and zeta potential (0.336), indicating slight increases in these values with higher SW. Similarly, SW has a negligible correlation with contact angle (-0.071), a weak positive correlation with permeance (0.134), and a moderate negative correlation with rejection (-0.297), suggesting that higher SW is associated with increased permeance and decreased rejection efficiency. However, MWCO is strongly correlated with contact angle (0.708), and weakly with SW (0.308) and ZP (0.052). MWCO has very weak correlations with permeance (0.063) and a moderate negative correlation with rejection (-0.358), indicating lower rejection rates with higher MWCO. CA shows a strong positive correlation with MWCO (0.708), weak positive correlations with zeta potential (0.252) and SW (-0.071), a weak positive correlation with permeance (0.082), and a moderate negative correlation with rejection (-0.239), suggesting higher contact angles lead to lower rejection rates. Furthermore, ZP exhibits weak positive correlations with SW (0.336), MWCO (0.052), CA (0.252), and permeance (0.077), and a weak negative correlation with rejection (-0.190), indicating higher ZP slightly increases permeance and decreases rejection efficiency. Permeance shows weak positive correlations with SW (0.134), MWCO (0.063), CA (0.082), and ZP (0.077), suggesting it is influenced by a combination of these factors. On the other hand, rejection displays moderate negative correlations with SW (-0.297), MWCO (-0.358), CA (-0.239), and ZP (-0.190), indicating that membranes with higher values in these parameters are less effective at solute rejection. These insights are valuable for designing deep learning models to optimize RO membrane performance [43,44].



Fig. 5: dependency between the membrane characteristic parameters

**3.2 Models Hyperparameter Tuning**

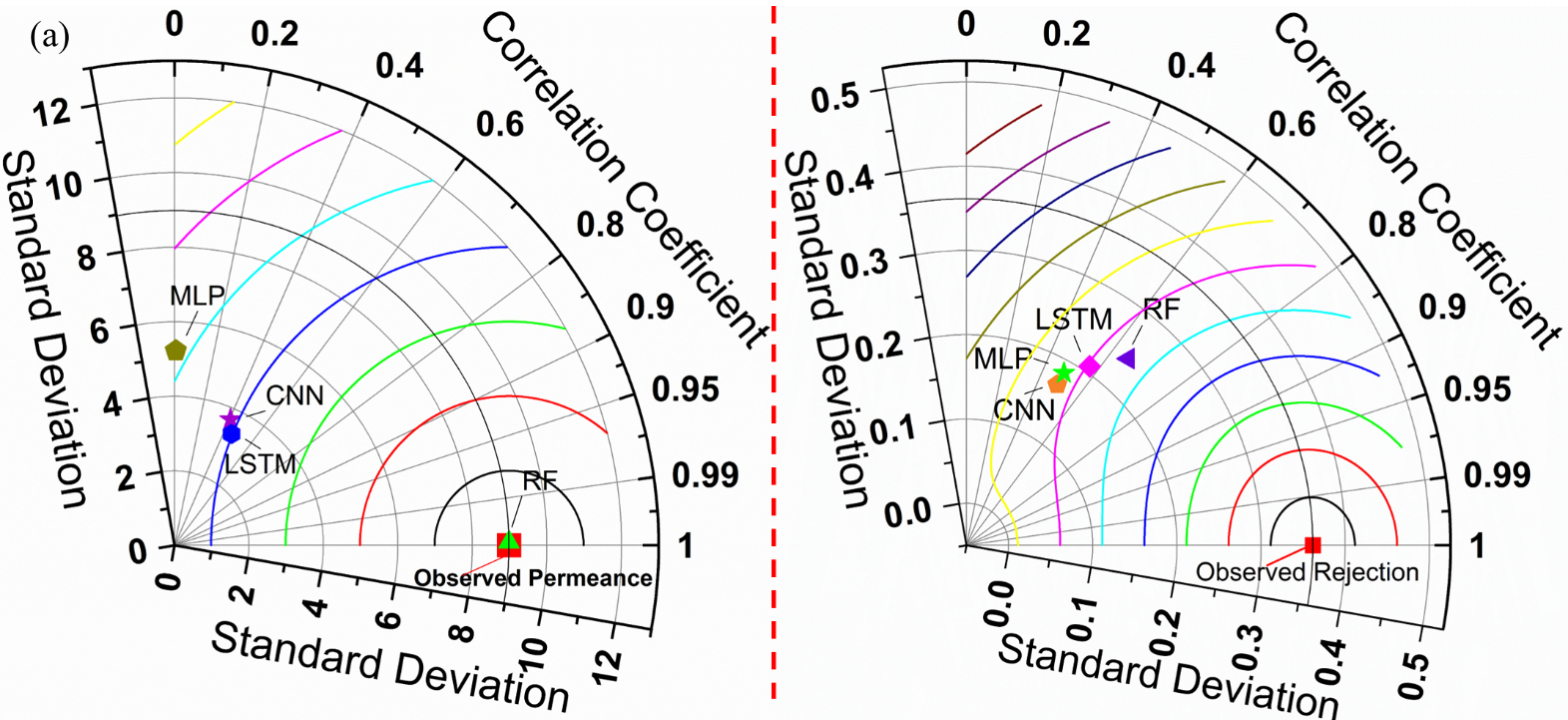
Hyperparameter tuning is crucial for optimizing the performance and accuracy of machine learning models. It helps in finding the right balance to prevent overfitting and underfitting, ensuring that the model generalizes well to unseen data [45,46]. Proper tuning enhances model efficiency, speed, and robustness, leading to stable and consistent training processes. The performance of various machine learning models, including CNN, LSTM, MLP, and RF, was optimized through careful hyperparameter tuning to ensure accurate prediction of permeance and rejection rates of RO membranes. For the RF model, the primary hyperparameters included setting the number of epochs to 5000 to ensure adequate learning iterations. The dataset was partitioned with 80% allocated for training and 20% for validation, providing a robust mechanism to evaluate the model's performance on unseen data and prevent overfitting. The MLP model was tuned with a training size of 80% and a validation size of 20%. The model underwent 101 epochs with an initial learning rate of 0.005, optimized using the Adam optimizer. The architecture varied based on the task: for rejection prediction, it consisted of an input layer with 4 neurons, a hidden layer with 10 neurons, and an output layer with 1 neuron; for permeance prediction, the hidden layer had 20 neurons, and for multiple-input multiple-output (MIMO) tasks, it had 30 neurons. The activation function used was ReLU. After 5000 epochs, the validation loss was 0.0744. Similarly, for rejection prediction, the LSTM model included a training size of 80% and a validation size of 20%, with an input size of 4, a hidden size of 128, 2 layers, and an output size of 1. The learning rate was set at 0.0008, and the model was trained for 20000 epochs, achieving a training loss of 0.0937 and a validation loss of 0.0943.

For permeance prediction, the LSTM architecture had a hidden size of 256, maintained an input size of 4, with a learning rate of 0.001, and was trained for 5000 epochs, achieving a training loss of 64.5796 and a validation loss of 62.0315. However, the CNN model for rejection prediction was developed using the TensorFlow Keras Sequential API, including layers such as Conv1D with 32 filters, MaxPooling1D, Flatten, and two Dense layers (32 and 1 neurons). The kernel size was set to 3, and the ReLU activation function was used. The model was trained with 80% of the data and validated with 20%, over 2000 epochs with a batch size of 64. The training loss was 0.1002, with a Mean Absolute Error (MAE) of 0.2711, while the validation loss was 0.0992, with an MAE of 0.2687. For permeance prediction, a similar CNN architecture was used, achieving a training loss of 67.7911 with an MAE of 3.2671 and a validation loss of 67.2836 with an MAE of 3.1899. The hyperparameter tuning for each model was carefully adjusted to enhance performance for predicting permeance and rejection rates in RO membranes. These carefully selected parameters ensured that the models provided accurate and reliable predictions, significantly contributing to the optimization and efficiency of water treatment processes[47].

**3.3 Predictive ML Results**

The validation phase results for permeance prediction using various ML models, namely CNN, LSTM, MLP, and RF, are summarized in Table 2. The CNN model showed a moderate positive correlation (PCC=0.4593) with an MSE of 39.4746, MAPE of 191.9223, MAE of 1.7024, PBIAS of 0.0357, and RMSE of 6.2829. The LSTM model slightly improved with a PCC of 0.4966, MSE of 38.6019, MAPE of 156.1755, MAE of 1.6115, PBIAS of 0.1960, and RMSE of 6.2130. The MLP model exhibited a strong positive correlation (PCC=0.7027) but higher error rates with an MSE of 44.1834, MAPE of 200.9758, MAE of 1.8574, PBIAS of 0.1170, and RMSE of 6.6471. The RF model outperformed all others, achieving a perfect correlation (PCC=1.0000), minimal MSE (0.0024), MAPE (2.2490), MAE (0.0226), PBIAS (-0.0011), and RMSE (0.0489), indicating excellent prediction accuracy. This modelling is essential for optimizing reverse osmosis membrane performance, as it allows for precise permeance prediction, thereby enhancing desalination processes’ efficiency and effectiveness (Fig. 6a). Similarly, PBAIS results from the validation phase provide critical insight into the systematic bias present in the predictions for permeance prediction. The CNN model has a small positive bias with a PBIAS of 0.0357, indicating that it slightly overestimates permeance values by about 3.57%, suggesting reasonably accurate predictions with minimal overestimation. The LSTM model, with a PBIAS of 0.1960, shows a higher positive bias, suggesting a greater tendency to overestimate permeance values by 19.60%. The outcomes indicate that LSTM predictions are consistently higher than the actual values and less reliable for precise predictions.

Furthermore, the MLP model has a moderate positive bias with a PBIAS of 0.1170, overestimating permeance values by 11.70%, smaller than LSTM but higher than CNN, indicating a moderate prediction overestimation. In contrast, the RF model demonstrates an almost negligible negative bias with a PBIAS of -0.0011, indicating a very slight underestimation of permeance values by 0.11%, which suggests that the RF model’s predictions are very close to the actual values, both in terms of underestimation and overestimation. CNN has the lowest positive bias among the models, indicating better accuracy in predicting permeance with minimal overestimation compared to LSTM and MLP. LSTM exhibits the highest positive bias, suggesting that it consistently overestimates permeance more than the other models, making it less reliable for accurate predictions. MLP, with a positive bias, falls in between CNN and LSTM, showing a moderate overestimation tendency. RF stands out with the closest bias to zero, reflecting almost unbiased predictions and demonstrating its superior accuracy in estimating permeance values. While CNN and MLP have moderate positive biases and LSTM tends to overestimate permeance significantly, the RF model exhibits negligible bias, highlighting its exceptional predictive performance and reliability in permeance estimation (Fig 6b). The outcomes of the permeance prediction modelling, particularly the superior performance of the RF model, have significant positive implications for both EPA regulations and the United Nations Sustainable Development Goals (SDGs). The RF model’s high accuracy and negligible bias can help optimize RO membrane performance, leading to better water quality management and regulatory compliance. The results align with the EPA’s goals of ensuring safe and clean drinking water. Accurate prediction models assist water treatment facilities in maintaining compliance with EPA standards, improving operational efficiency, reducing energy consumption, and minimizing chemical use, thus enhancing overall resource efficiency and reducing environmental impact. In the context of the SDGs, these advancements contribute directly to Goal 6 (Clean Water and Sanitation) by increasing access to clean water through more efficient desalination processes.



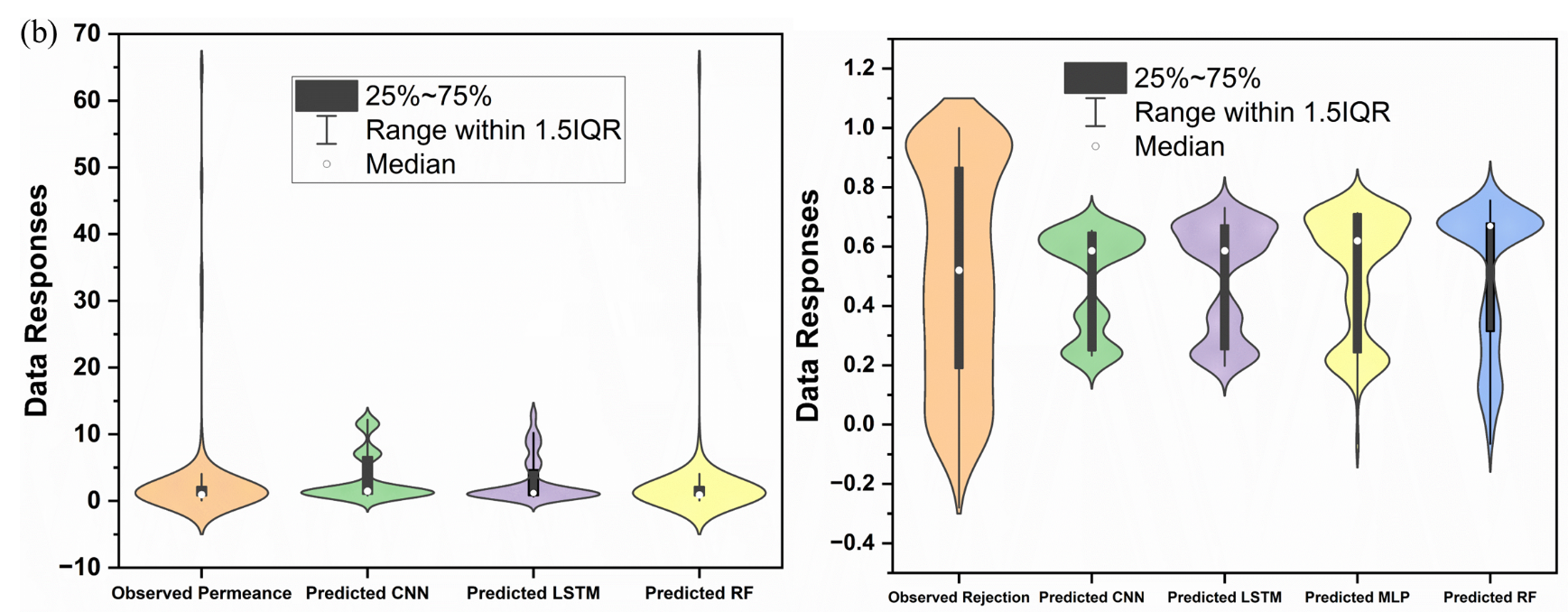


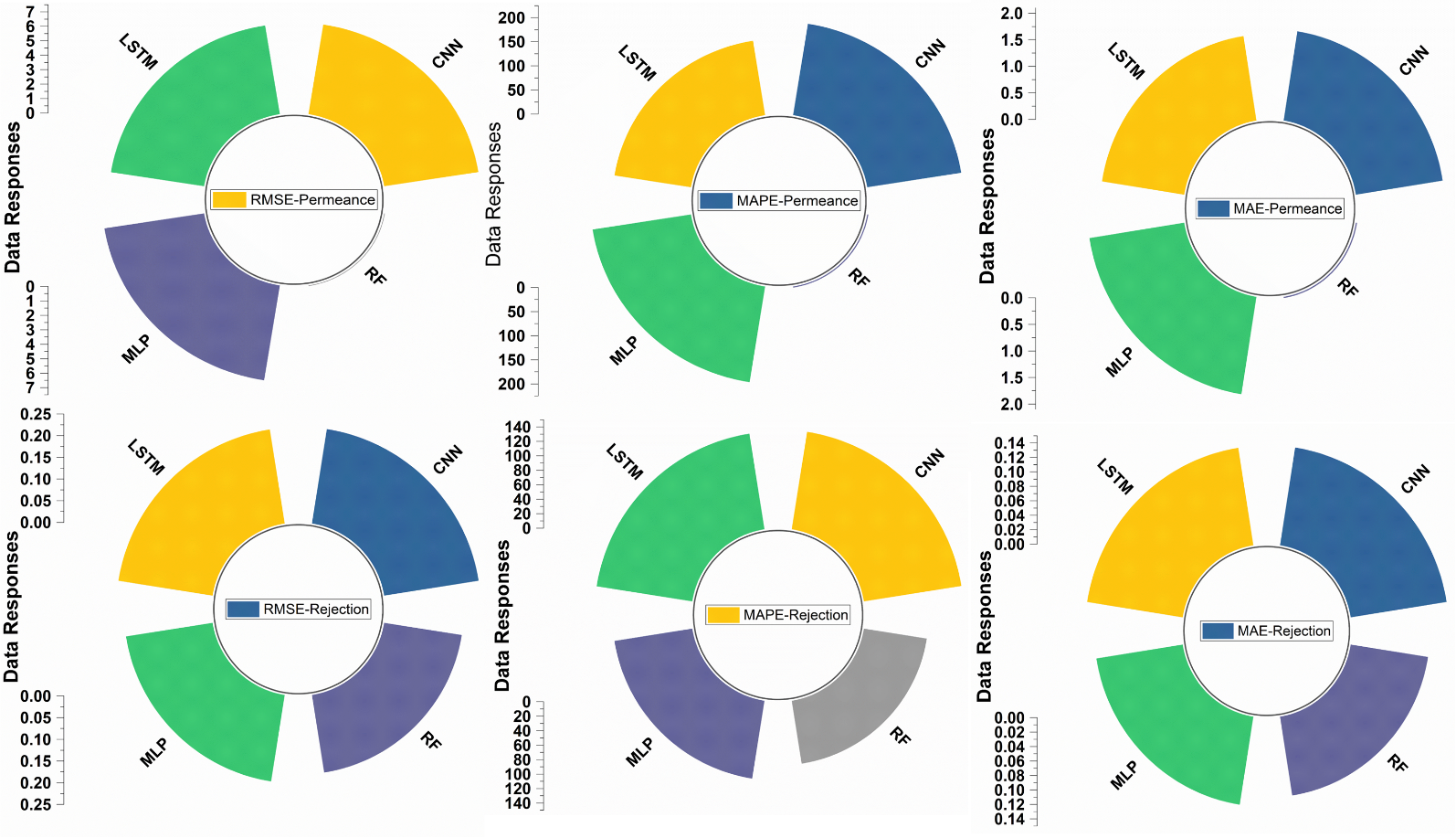
Fig. 6: comparison between the experimental and predicted permeance and rejection (a) Taylor two-dimensional diagram (b) violin plots

Further analysis of the validation phase results for rejection prediction using CNN, LSTM, MLP, and RF highlighted the performance metrics for each model as indicated in Table 2. The CNN model shows a moderate positive correlation (PCC: 0.5001) with an MSE of 0.0486, MAPE of 136.4040, MAE of 0.1371, PBIAS of 0.0717, and RMSE of 0.2204. The LSTM model has a similar moderate positive correlation (PCC: 0.4999) with a slightly lower MSE of 0.0482, MAPE of 133.9138, MAE of 0.1366, PBIAS of 0.0454, and RMSE of 0.2197. The MLP model shows an improved correlation (PCC: 0.6042) with an MSE of 0.0407, MAPE of 108.9924, MAE of 0.1234, PBIAS of 0.0331, and RMSE of 0.2018, indicating better prediction accuracy and lower error rates compared to CNN and LSTM. The RF model outperforms all other models with the highest correlation (PCC: 0.7010), lowest MSE (0.0328), lowest MAPE (88.1296), lowest MAE (0.1105), PBIAS of 0.0445, and lowest RMSE (0.1810), reflecting its superior predictive performance and minimal error. These outcomes suggest that the RF model is the most reliable for predicting rejection rates, followed by MLP, LSTM, and CNN (see, Fig. 7 a and Fig. 7b). The positive implications of these results include enhanced efficiency in water treatment processes, better compliance with regulatory standards, and significant contributions to SDGs by improving access to clean water, promoting innovation in water treatment technologies, and ensuring responsible consumption and production patterns.

Besides, the predictive skills were compared using MAPE criteria; it is worth mentioning that MAPE measures the accuracy of a model’s predictions by expressing the average absolute percentage error between predicted and actual values, which is crucial for understanding the model’s performance and reliability. The CNN model has the highest MAPE at 136.4040%, indicating that its predictions deviate from the actual values by an average of 136.404%, reflecting relatively poor predictive accuracy. The LSTM model has a slightly lower MAPE of 133.9138%, showing an average deviation of 133.914% from actual values, representing a marginal improvement over CNN. The MLP model significantly outperforms CNN and LSTM with a MAPE of 108.9924%, indicating that its predictions deviate from the actual values by 108.992% on average, representing a substantial reduction in prediction error. The RF model achieves the lowest MAPE of 88.1296%, indicating that its predictions deviate from the actual values by 88.130% on average, reflecting the highest predictive accuracy and the lowest error rate among the models. When comparing predictive skill percentages, LSTM improves over CNN by approximately 1.82%, MLP improves over CNN by approximately 20.10%, and RF improves over CNN by approximately 35.39%. Comparing LSTM to MLP, MLP improves predictive skills by approximately 18.60%, while RF improves over LSTM by approximately 34.18%. Finally, RF improves over MLP by approximately 19.09%. These comparisons highlight the superior performance of the RF model, which significantly minimizes prediction errors and enhances the reliability of rejection rate predictions. The high predictive accuracy of the RF model highlights its potential in optimizing reverse osmosis membrane performance, leading to more efficient water treatment processes. It is important to note that the outcomes were compared using a new radial diagram, as presented in Fig 7. The importance of a radial plot lies in its ability to visually compare multiple variables or categories simultaneously, highlighting differences and patterns in data across a circular layout, which aids in identifying relationships and trends.

Table 2: Predictive validation results for single and deep learning model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Permeance** | |  |  |  |  |  |
|  | PCC | MSE | MAPE | MAE | PBIAS | RMSE |
| CNN | 0.4593 | 39.4746 | 191.9223 | 1.7024 | 0.0357 | 6.2829 |
| LSTM | 0.4966 | 38.6019 | 156.1755 | 1.6115 | 0.1960 | 6.2130 |
| MLP | 0.7027 | 44.1834 | 200.9758 | 1.8574 | 0.1170 | 6.6471 |
| RF | 1.0000 | 0.0024 | 2.2490 | 0.0226 | -0.0011 | 0.0489 |
| **Rejection** |  |  |  |  |  |  |
|  | PCC | MSE | MAPE | MAE | PBIAS | RMSE |
| CNN | 0.5001 | 0.0486 | 136.4040 | 0.1371 | 0.0717 | 0.2204 |
| LSTM | 0.4999 | 0.0482 | 133.9138 | 0.1366 | 0.0454 | 0.2197 |
| MLP | 0.6042 | 0.0407 | 108.9924 | 0.1234 | 0.0331 | 0.2018 |
| RF | 0.7010 | 0.0328 | 88.1296 | 0.1105 | 0.0445 | 0.1810 |

****

**Fig. 7:** Radial plot showing the comprehensive performance visualization

**3.3 Stationarity and Reliability Analysis**

The essential aspect of stationarity and reliability analysis in the present study is to ensure that the input data and model predictions remain consistent and dependable over time, which is crucial for accurate and stable performance in predicting permeance and rejection rates in water treatment applications. For this purpose, Table 3 presents the results of stationarity tests for various variables, including CA, MWCO, SW, ZP, Permeance, and Rejection. The tests used are the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test, with the selection criterion being the Schwarz Information Criterion (SIC) for ADF and the Bartlett kernel for PP. The results show the test values, critical values at different significance levels (1%, 5%, and 10%), the probability (Prob) values, and the decisions regarding stationarity. For CA, the ADF test values of -3.800427 and -72.19683 with probabilities of 0.0029 and 0.0001, respectively, indicate stationarity at levels I(0) and I(1). The PP test results with similar test values and probabilities confirm these findings. For MWCO, the ADF test with values of -3.418436 and -72.19476, and probabilities of 0.0104 and 0.0001, indicate stationarity at I(0) and I(1), supported by the PP test with probabilities of 0.0097 and 0.0001. SW shows non-stationarity at level I(0) with ADF and PP test values not meeting the critical values, but stationarity at I(1) with probabilities of 0.0001 for both tests. However, the ZP exhibits strong stationarity with ADF test values of -4.413233 and -72.19418, and PP test values of -4.476638 and -72.19418, all showing probabilities less than 0.0003, indicating stationarity at both levels I(0) and I(1). Permeance and Rejection are also stationary with ADF test values of -7.661522 and -7.19586, respectively, both showing probabilities of 0, indicating stationarity at level I(0). Table 3 demonstrates that most variables exhibit stationarity at least at the first differencing level (I(1)), with several variables, including CA, MWCO, ZP, Permeance, and Rejection, showing stationarity at level I(0). This stationarity is crucial for ensuring the reliability and consistency of the data used in the predictive models, enhancing the accuracy and stability of permeance and rejection rate predictions in water treatment applications.

Table 3: Comparison of different reliability tests using ADF and PP

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variables | TV | 1% CV | 5% CV | 10% CV | Prob | Decision | SC | TT |
| CA | -3.8004 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Schwarz | ADF |
|  | -72.1968 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Schwarz | ADF |
|  | -3.8379 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Bartlett kernel | PP |
|  | -72.1968 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Bartlett kernel | PP |
|  | -3.4184 | -3.4314 | -2.8619 | -2.57 | 0.01 | I (0) | Schwarz | ADF |
| MWCO | -72.1948 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Schwarz | ADF |
|  | -3.4413 | -3.4314 | -2.8619 | -2.57 | 0.01 | I (0) | Bartlett kernel | PP |
|  | -72.1948 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Bartlett kernel | PP |
| SW | -2.5317 | -3.4314 | -2.8619 | -2.57 | 0.11 | I (0) | Schwarz | ADF |
|  | -61.2996 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Schwarz | ADF |
|  | -2.6945 | -3.4314 | -2.8619 | -2.57 | 0.08 | I (0) | Bartlett kernel | PP |
|  | -94.4767 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Bartlett kernel | PP |
| ZP | -4.4132 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Schwarz | ADF |
|  | -72.1942 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (1) | Schwarz | ADF |
|  | -4.4766 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Bartlett kernel | PP |
| Permeance | -7.6615 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Schwarz | ADF |
| Rejection | -7.1959 | -3.4314 | -2.8619 | -2.57 | 0.00 | I (0) | Schwarz | ADF |

**3.3 Environmental Implication**

Applying machine learning models to predict permeance and rejection rates in RO membranes has significant environmental implications, particularly in compliance with EPA regulations and the United Nations (SDGs). Among the various models assessed, the RF model consistently demonstrated superior performance. This high accuracy and low error rate emphasize its robustness in predicting RO membrane performance, and using MAPE as a key metric highlights the prediction accuracy of these models. MAPE measures the average absolute percentage error between predicted and actual values, clearly indicating a model’s performance and reliability. Ensuring stationarity and reliability in data analysis is crucial as it guarantees that input data and model predictions remain consistent and dependable over time. This consistency is vital for making accurate and stable predictions of permeance and rejection rates in water treatment applications. The RF model’s high accuracy and negligible bias can significantly optimize RO membrane performance, leading to better water quality management and regulatory compliance with EPA standards. This optimization aligns with the EPA’s objectives of ensuring safe and clean drinking water. Accurate prediction models enable water treatment facilities to maintain regulatory compliance, improve operational efficiency, reduce energy consumption, and minimize chemical usage. These improvements enhance overall resource efficiency and reduce the environmental footprint of water treatment operations.

Globally, these advancements are critical in addressing water scarcity and ensuring sustainable water management. By improving the efficiency of desalination processes, these models contribute to achieving SDG 6 (Clean Water and Sanitation) by increasing access to clean water. Furthermore, adopting advanced machine learning models like RF supports SDG 9 (Industry, Innovation, and Infrastructure) by fostering technological innovation in water treatment. Efficient water treatment processes support SDG 12 (Responsible Consumption and Production) by promoting sustainable consumption patterns and reducing waste. Additionally, reducing energy consumption and emissions associated with optimized water treatment operations contributes to SDG 13 (Climate Action). Comparing these advancements globally, regions facing acute water scarcity can benefit significantly from applying these models. Countries in arid and semi-arid areas, such as the Middle East and parts of Africa, can enhance their desalination capabilities, ensuring a reliable clean water supply. Similarly, in developed countries, these models can lead to more sustainable water management practices, aligning with stringent environmental regulations and promoting green technologies. The ecological implications of using machine learning models for RO membrane performance prediction are profound. These models enhance the efficiency and reliability of water treatment processes and support broader environmental and sustainability goals. The study emphasizes the importance of integrating advanced technologies into ecological management practices by contributing to the sustainable management of water resources, compliance with regulatory standards, and achieving global environmental targets. The international adoption of such predictive models can be pivotal in addressing water challenges, promoting sustainable development, and ensuring a healthier planet for future generations.

**4. Conclusion**

The primary aim of this study was to evaluate the performance of various machine learning models, including CNN, LSTM, MLP, and RF, in predicting permeance and rejection rates of RO membranes, with the objective of optimizing water treatment processes for environmental sustainability and regulatory compliance. The study utilized a comprehensive dataset comprising critical membrane characteristics such as SW, MWCO, CA, and ZP. Performance metrics included PCC, MSE, MAPE, MAE, PBIAS, and RMSE. The RF model demonstrated superior performance with a PCC of 1.0000, MSE of 0.0024, MAPE of 2.2490%, MAE of 0.0226, PBIAS of -0.0011, and RMSE of 0.0489 for permeance prediction, significantly outperforming the CNN, LSTM, and MLP models, which showed MAPEs of 191.9223%, 156.1755%, and 200.9758%, respectively. For rejection prediction, the RF model again outperformed others with a PCC of 0.7010, MSE of 0.0328, MAPE of 88.1296%, MAE of 0.1105, PBIAS of 0.0445, and RMSE of 0.1810, compared to the CNN (MAPE: 136.4040%), LSTM (MAPE: 133.9138%), and MLP (MAPE: 108.9924%). Future research should focus on enhancing predictive capabilities by incorporating additional variables, refining algorithms, and conducting real-world pilot studies to validate these models under operational conditions. Comparative studies across different regions and integration with advanced technologies like the Internet of Things (IoT) and blockchain could further improve data collection, processing, and transparency. This study highlights the significant potential of the RF model in optimizing RO membrane performance, contributing to enhanced water quality management, regulatory compliance, and the achievement of SDGs related to clean water, industry innovation, and responsible consumption. Implementing these advanced models can lead to improved operational efficiency, reduced energy consumption, and minimized chemical usage, thereby supporting environmental sustainability.

**References**

1. Wang, M.; Bodirsky, B.L.; Rijneveld, R.; Beier, F.; Bak, M.P.; Batool, M.; Droppers, B.; Popp, A.; van Vliet, M.T.H.; Strokal, M. A Triple Increase in Global River Basins with Water Scarcity Due to Future Pollution. *Nat. Commun.* **2024**, *15*, 880, doi:10.1038/s41467-024-44947-3.

2. *WMO, World Meteorological Organization, 2022 State of Global Water Resources 2021*; 2022;

3. Zouhri, N.; Addar, F.Z.; Tahaikt, M.; Elamrani, M.; ELmidaoui, A.; Taky, M. Techno-Economic Study and Optimization of the Performance of Nanofiltration and Reverse Osmosis Membranes in Reducing the Salinity of M′rirt Water City (Morocco). *Desalin. Water Treat.* **2024**, *317*, 100042, doi:10.1016/j.dwt.2024.100042.

4. Alonso, E.; Sanchez-Huerta, C.; Ali, Z.; Wang, Y.; Fortunato, L.; Pinnau, I. Evaluation of Nanofiltration and Reverse Osmosis Membranes for Efficient Rejection of Organic Micropollutants. *J. Memb. Sci.* **2024**, *693*, 122357, doi:10.1016/j.memsci.2023.122357.

5. Liu, Y.; Xin, Z.; Wang, M.; Wang, X.; Zhang, H.; Wang, Z. Optimizing Separation Layer Structure of Polyamide Composite Membrane for High Permselectivity Based on Post-Treatment: A Review. *Desalination* **2024**, *580*, 117585, doi:10.1016/j.desal.2024.117585.

6. Ray, S.S.; Verma, R.K.; Singh, A.; Ganesapillai, M.; Kwon, Y.N. A Holistic Review on How Artificial Intelligence Has Redefined Water Treatment and Seawater Desalination Processes. *Desalination* **2023**, *546*, 116221, doi:10.1016/j.desal.2022.116221.

7. Jawad, J.; Hawari, A.H.; Javaid, S. Artificial Neural Network Modeling of Wastewater Treatment and Desalination Using Membrane Processes : A Review. *Chem. Eng. J.* **2021**, *419*, 129540, doi:10.1016/j.cej.2021.129540.

8. Taloba, A.I. An Artificial Neural Network Mechanism for Optimizing the Water Treatment Process and Desalination Process. *Alexandria Eng. J.* **2022**, *61*, 9287–9295, doi:10.1016/j.aej.2022.03.029.

9. Sayed, E.T.; Olabi, A.G.; Elsaid, K.; Al Radi, M.; Semeraro, C.; Doranehgard, M.H.; Eltayeb, M.E.; Abdelkareem, M.A. Application of Artificial Intelligence Techniques for Modeling, Optimizing, and Controlling Desalination Systems Powered by Renewable Energy Resources. *J. Clean. Prod.* **2023**, *413*, 137486, doi:10.1016/j.jclepro.2023.137486.

10. Niu, C.; Li, X.; Dai, R.; Wang, Z. Artificial Intelligence-Incorporated Membrane Fouling Prediction for Membrane-Based Processes in the Past 20 Years: A Critical Review. *Water Res.* **2022**, *216*, 118299, doi:10.1016/j.watres.2022.118299.

11. Alghamdi, A. A Novel IEF-DLNN and Multi-Objective Based Optimizing Control Strategy for Seawater Reverse Osmosis Desalination Plant. *Heliyon* **2023**, *9*, e13814, doi:10.1016/j.heliyon.2023.e13814.

12. Xu, J.; Meng, K.; Niu, Y.; Zhang, C.; Xu, K.; Rong, J.; Wei, Y.; Yu, X. Deciphering the Electronic-Level Mechanism of Na+ Transport in a Graphdiyne Desalination Membrane with Periodic Nanopores. *Desalination* **2023**, *546*, 116183, doi:10.1016/j.desal.2022.116183.

13. Alardhi, S.M.; Salman, A.D.; Breig, S.J.M.; Jaber, A.A.; Fiyadh, S.S.; AlJaberi, F.Y.; Duc Nguyen, D.; Van, B.; Le, P.C. Artificial Neural Network and Response Surface Methodology for Modeling Reverse Osmosis Process in Wastewater Treatment. *J. Ind. Eng. Chem.* **2024**, *133*, 599–613, doi:10.1016/j.jiec.2024.02.039.

14. Park, S.; Baek, S.S.; Pyo, J.C.; Pachepsky, Y.; Park, J.; Cho, K.H. Deep Neural Networks for Modeling Fouling Growth and Flux Decline during NF/RO Membrane Filtration. *J. Memb. Sci.* **2019**, *587*, 117164, doi:10.1016/j.memsci.2019.06.004.

15. Luo, J.; Li, M.; Hoek, E.M.V.; Heng, Y. Supercomputing and Machine Learning-Aided Optimal Design of High Permeability Seawater Reverse Osmosis Membrane Systems. *Sci. Bull.* **2023**, *68*, 397–407, doi:10.1016/j.scib.2023.01.039.

16. Tayyebi, A.; Alshami, A.S.; Tayyebi, E.; Buelke, C.; Talukder, M.J.; Ismail, N.; Al-Goraee, A.; Rabiei, Z.; Yu, X. Machine Learning – Driven Surface Grafting of Thin-Film Composite Reverse Osmosis (TFC-RO) Membrane. *Desalination* **2024**, *579*, 117502, doi:10.1016/j.desal.2024.117502.

17. Yeo, C.S.H.; Xie, Q.; Wang, X.; Zhang, S. Understanding and Optimization of Thin Film Nanocomposite Membranes for Reverse Osmosis with Machine Learning. *J. Memb. Sci.* **2020**, *606*, 118135, doi:10.1016/j.memsci.2020.118135.

18. Talhami, M.; Wakjira, T.; Alomar, T.; Fouladi, S.; Fezouni, F.; Ebead, U.; Altaee, A.; AL-Ejji, M.; Das, P.; Hawari, A.H. Single and Ensemble Explainable Machine Learning-Based Prediction of Membrane Flux in the Reverse Osmosis Process. *J. Water Process Eng.* **2024**, *57*, 104633, doi:10.1016/j.jwpe.2023.104633.

19. Mohammed, M.A.A.; Kaya, F.; Mohamed, A.; Alarifi, S.S.; Abdelrady, A.; Keshavarzi, A.; Szabó, N.P.; Szűcs, P. Application of GIS-Based Machine Learning Algorithms for Prediction of Irrigational Groundwater Quality Indices. *Front. Earth Sci.* **2023**, *11*, 1–19, doi:10.3389/feart.2023.1274142.

20. Yang, K.; Xu, X.; Yang, B.; Cook, B.; Ramos, H.; Bauchy, M. Prediction of Silicate Glasses’ Stiffness by High-Throughput Molecular Dynamics Simulations and Machine Learning. *arXiv* **2019**, 1–20.

21. Jahani, A.; Rayegani, B. Forest Landscape Visual Quality Evaluation Using Artificial Intelligence Techniques as a Decision Support System. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1473–1486, doi:10.1007/s00477-020-01832-x.

22. Yang, Z. Competing Leaders Grey Wolf Optimizer and Its Application for Training Multi-Layer Perceptron Classifier. *Expert Syst. Appl.* **2024**, *239*, doi:10.1016/j.eswa.2023.122349.

23. Abubakar, A.; Jibril, M.M.; Almeida, C.F.M.; Gemignani, M.; Yahya, M.N.; Abba, S.I. Photovoltaic Arrays and Inverters Using AI and Statistical Learning Techniques : A Focus on Sustainable Environment. **2023**.

24. Afan, H.A.; Ibrahem Ahmed Osman, A.; Essam, Y.; Ahmed, A.N.; Huang, Y.F.; Kisi, O.; Sherif, M.; Sefelnasr, A.; Chau, K. wing; El-Shafie, A. Modeling the Fluctuations of Groundwater Level by Employing Ensemble Deep Learning Techniques. *Eng. Appl. Comput. Fluid Mech.* **2021**, *15*, 1420–1439, doi:10.1080/19942060.2021.1974093.

25. Elomiya, A.; Křupka, J.; Jovčić, S.; Simic, V. Enhanced Prediction of Parking Occupancy through Fusion of Adaptive Neuro-Fuzzy Inference System and Deep Learning Models. *Eng. Appl. Artif. Intell.* **2024**, *129*, 107670, doi:10.1016/j.engappai.2023.107670.

26. Ehteram, M.; Afshari Nia, M.; Panahi, F.; Farrokhi, A. Read-First LSTM Model: A New Variant of Long Short Term Memory Neural Network for Predicting Solar Radiation Data. *Energy Convers. Manag.* **2024**, *305*, 118267, doi:10.1016/j.enconman.2024.118267.

27. Damavandi, H.G.; Shah, R.; Stampoulis, D.; Wei, Y.; Boscovic, D.; Sabo, J. Accurate Prediction of Streamflow Using Long Short-Term Memory Network: A Case Study in the Brazos River Basin in Texas. *Int. J. Environ. Sci. Dev.* **2019**, *10*, 294–300, doi:10.18178/ijesd.2019.10.10.1190.

28. Karimanzira, D.; Rauschenbach, T. Deep Learning Based Model Predictive Control for a Reverse Osmosis Desalination Plant. *J. Appl. Math. Phys.* **2020**, *08*, 2713–2731, doi:10.4236/jamp.2020.812201.

29. Xin, J.; Zhou, C.; Jiang, Y.; Tang, Q.; Yang, X.; Zhou, J. A Signal Recovery Method for Bridge Monitoring System Using TVFEMD and Encoder-Decoder Aided LSTM. *Meas. J. Int. Meas. Confed.* **2023**, *214*, 112797, doi:10.1016/j.measurement.2023.112797.

30. Nosair, A.M.; Shams, M.Y.; AbouElmagd, L.M.; Hassanein, A.E.; Fryar, A.E.; Abu Salem, H.S. Predictive Model for Progressive Salinization in a Coastal Aquifer Using Artificial Intelligence and Hydrogeochemical Techniques: A Case Study of the Nile Delta Aquifer, Egypt. *Environ. Sci. Pollut. Res.* **2022**, *29*, 9318–9340, doi:10.1007/s11356-021-16289-w.

31. Sarkar, A.; Maniruzzaman, M.; Alahe, M.A.; Ahmad, M. An Effective and Novel Approach for Brain Tumor Classification Using AlexNet CNN Feature Extractor and Multiple Eminent Machine Learning Classifiers in MRIs. *J. Sensors* **2023**, *2023*, doi:10.1155/2023/1224619.

32. Ikram, R.M.A.; Mostafa, R.R.; Chen, Z.; Parmar, K.S.; Kisi, O.; Zounemat-Kermani, M. Water Temperature Prediction Using Improved Deep Learning Methods through Reptile Search Algorithm and Weighted Mean of Vectors Optimizer. *J. Mar. Sci. Eng.* **2023**, *11*, doi:10.3390/jmse11020259.

33. Yuan, X.; Huang, L.; Ye, L.; Wang, Y.; Wang, K.; Yang, C.; Gui, W.; Shen, F. Quality Prediction Modeling for Industrial Processes Using Multiscale Attention-Based Convolutional Neural Network. *IEEE Trans. Cybern.* **2024**, *54*, 2696–2707, doi:10.1109/TCYB.2024.3365068.

34. Pan, M.; Zhou, H.; Cao, J.; Liu, Y.; Hao, J.; Li, S.; Chen, C.H. Water Level Prediction Model Based on GRU and CNN. *IEEE Access* **2020**, *8*, 60090–60100, doi:10.1109/ACCESS.2020.2982433.

35. Khosravi, K.; Golkarian, A.; Tiefenbacher, J.P. Using Optimized Deep Learning to Predict Daily Streamflow: A Comparison to Common Machine Learning Algorithms. *Water Resour. Manag.* **2022**, *36*, 699–716, doi:10.1007/s11269-021-03051-7.

36. Sun, Z.; Wang, G.; Li, P.; Wang, H.; Zhang, M.; Liang, X. An Improved Random Forest Based on the Classification Accuracy and Correlation Measurement of Decision Trees. *Expert Syst. Appl.* **2024**, *237*, 121549, doi:10.1016/j.eswa.2023.121549.

37. Khan, M.A.; Shah, M.I.; Javed, M.F.; Khan, M.I.; Rasheed, S.; El-Shorbagy, M.A.; El-Zahar, E.R.; Malik, M.Y. Application of Random Forest for Modelling of Surface Water Salinity. *Ain Shams Eng. J.* **2022**, *13*, 101635.

38. Tao, H.; Salih, S.; Oudah, A.Y.; Abba, S.I.; Ameen, A.M.S.; Awadh, S.M.; Alawi, O.A.; Mostafa, R.R.; Surendran, U.P.; Yaseen, Z.M. Development of New Computational Machine Learning Models for Longitudinal Dispersion Coefficient Determination: Case Study of Natural Streams, United States. *Environ. Sci. Pollut. Res.* **2022**, doi:10.1007/s11356-022-18554-y.

39. Breiman, L. Random Forests. *Mach. Learn.* **2001**, 5–32, doi:10.3390/rs10060911.

40. He, S.; Wu, J.; Wang, D.; He, X. Predictive Modeling of Groundwater Nitrate Pollution and Evaluating Its Main Impact Factors Using Random Forest. *Chemosphere* **2022**, *290*, 133388, doi:10.1016/j.chemosphere.2021.133388.

41. Jibril, M.M.; Bello, A.; Aminu, I.I.; Ibrahim, A.S.; Bashir, A.; Malami, S.I.; Habibu, M.A.; Magaji, M.M. An Overview of Streamflow Prediction Using Random Forest Algorithm. **2022**, 0–7.

42. Kiiza, C.; Pan, S. qi; Bockelmann-Evans, B.; Babatunde, A. Predicting Pollutant Removal in Constructed Wetlands Using Artificial Neural Networks (ANNs). *Water Sci. Eng.* **2020**, *13*, 14–23, doi:10.1016/j.wse.2020.03.005.

43. Brentan, B.M.; Meirelles, G.; Herrera, M.; Luvizotto, E.; Izquierdo, J. Correlation Analysis of Water Demand and Predictive Variables for Short-Term Forecasting Models. *Math. Probl. Eng.* **2017**, *2017*, doi:10.1155/2017/6343625.

44. Abdullahi, J.; Rotimi, A.; Malami, S.I.; Jibrin, H.B.; Tahsin, A.; Abba, S.I. Feasibility of Artificial Intelligence and CROPWAT Models in the Estimation of Uncertain Combined Variable Using Nonlinear Sensitivity Analysis. *2021 1st Int. Conf. Multidiscip. Eng. Appl. Sci. ICMEAS 2021* **2021**, 2–8, doi:10.1109/ICMEAS52683.2021.9692357.

45. Bacanin, N.; Stoean, C.; Zivkovic, M.; Rakic, M.; Strulak-Wójcikiewicz, R.; Stoean, R. On the Benefits of Using Metaheuristics in the Hyperparameter Tuning of Deep Learning Models for Energy Load Forecasting. *Energies* **2023**, *16*, 1–21, doi:10.3390/en16031434.

46. Yao, L.; Li, Y.; Cheng, Q.; Chen, Z.; Song, J. Modeling and Optimization of Metal-Organic Frameworks Membranes for Reverse Osmosis with Artificial Neural Networks. *Desalination* **2022**, *532*, 115729, doi:10.1016/j.desal.2022.115729.

47. Mustafa, H.M.; Hayder, G.; Abba, S.I.; Algarni, A.D.; Mnzool, M.; Nour, A.H. Performance Evaluation of Hydroponic Wastewater Treatment Plant Integrated with Ensemble Learning Techniques : A Feature Selection Approach. **2023**.