

AI-based prediction of total dissolved solids in reverse osmosis desalination for arid regions

Rachid Zegait^{a,b,*}, Saber Kouadri^b, Maamar Chelgham^c, Luka Vucinic^d, Fatima Ajia^d, Elisabetta Sieni^e, Muhammad Rasool Al-Kilani^f

^a Department of Hydraulic and Civil Engineering, Kasdi Merbah University, Ouargla, Algeria

^b Laboratory of Water and Environmental Engineering in Saharan Environments (LGEEMS), Ouargla, Algeria

^c Water Production Directorate of the Province of Ouargla, Algerian Water (ADE), Algeria

^d Department of Civil Engineering and Environmental Management, Glasgow Caledonian University, London, UK

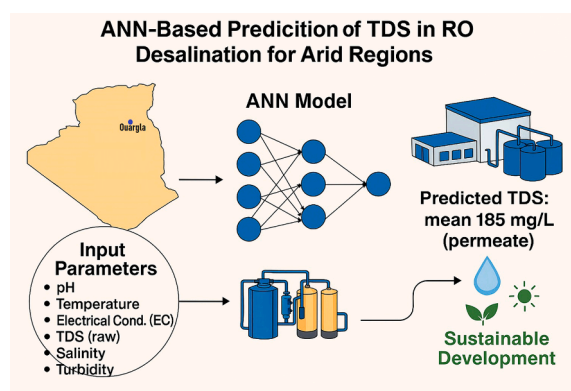
^e Department of Theoretical and Applied Science, Università degli Studi dell'Insubria, Italy

^f Department of Civil and Environmental Engineering, German Jordanian University, Amman, Jordan

HIGHLIGHTS

- Optimised ANN model predicts TDS in RO systems under Saharan arid conditions.
- Dataset of 223 samples collected from eight desalination plants in Ouargla, Algeria.
- 7-neuron single hidden layer model achieved best accuracy and generalisation.
- Data augmentation improved learning with synthetic samples and noise injection.
- ANN outperformed MLR, SVR, PR, and RF in accuracy and robustness.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Reverse osmosis
Artificial neural networks
Water desalination
TDS prediction
Arid regions
Algeria

ABSTRACT

Despite the increasing deployment of reverse osmosis (RO) desalination systems, predictive tools tailored to complex and variable operational conditions remain limited, particularly in arid regions. This study explores the use of artificial neural networks (ANNs) to predict total dissolved solids (TDS) in treated water across RO plants in the Ouargla region of Algeria. A dataset of 223 samples was collected over a 12-month period from eight desalination stations. Using this data, several ANN architectures were developed and evaluated. The optimal configuration featured a single hidden layer with 7 neurons. It achieved superior predictive accuracy compared to conventional models such as multiple linear regression and support vector regression. The model was trained using key operational parameters and validated for generalisation. It effectively captured the nonlinear interactions among variables, offering a reliable tool for performance optimisation and operational monitoring.

* Corresponding author at: Department of Hydraulic and Civil Engineering, Kasdi Merbah University, Ouargla, Algeria.

E-mail address: zegait.rachid@univ-ouargla.dz (R. Zegait).

¹ ORCID: 0000-0001-7976-0595

Data augmentation techniques, including 1.5 % noise injection and a 1:0.5 synthetic-to-original ratio, were used to improve robustness under data-constrained conditions. This contributed to improved learning and model stability. Although the test-phase performance was modest, it reflected the best balance among all tested configurations and avoided severe overfitting observed in deeper architectures. These findings provide important guidelines for ANN design in water treatment contexts. Overall, the study contributes a validated, AI-driven modelling framework that supports efficient and resilient management of RO desalination systems. It demonstrates a useful case on the potential of machine learning to enhance water treatment operations in semi-arid environments.

1. Introduction

Water desalination has become increasingly critical worldwide due to escalating freshwater demand and rising pollution risks [1,2]. While Earth possesses approximately $1.4 \times 10^9 \text{ km}^3$ of water, about 97.5 % is saline, leaving only 2.5 % as freshwater, unevenly distributed and largely locked in glaciers [3,4]. Consequently, a small fraction is available in rivers, lakes, and aquifers, which intensifies reliance on seawater desalination in many regions [1,3]. Furthermore, increased droughts and rainfall variability can increase the reliance of food production on desalination for irrigation [5,6].

Desalination refers to the process of removing salts and dissolved substances from saline water, typically to meet the WHO guideline for drinking water ($\leq 500 \text{ ppm}$) [7]. It avoids some of the environmental and health risks associated with other options like wastewater treatment [8]. Although desalination was practiced historically, it has gained industrial-scale application only in recent decades [1]. Current installed capacity exceeds 58 million m^3 /day from more than 17,000 units worldwide, and was projected to reach 95 million m^3 /day by 2020 [9, 10]. Despite high energy demands, regions facing water stress have increasingly adopted desalination technologies [10,11]. Among available technologies, reverse osmosis (RO) has emerged as the dominant method for seawater and brackish water desalination due to its high efficiency and relative cost-effectiveness [12]. The deployment of RO plants is often governed by factors such as aridity, proximity to seawater, elevation, and economic capacity to support energy-intensive processes. Countries like China, Gulf States, Mediterranean nations, and North African states have made considerable investments to expand desalination infrastructure [13–16].

Algeria stands as a notable case due to its combination of a long coastline (1200 km), vast arid interior, and increasing water demand driven by demographic and industrial pressures [17]. Initially focused on industrial needs, desalination in Algeria began in 1964 with multi-effect distillation at Arzew. This was followed by a multi-stage flash plant in 1969 and further capacity expansion in the 1970s for steel and petrochemical complexes [18,19]. RO technology was introduced for industrial brackish water desalination in the 1990s, and public desalination projects accelerated from 2008 onwards under emergency programmes [20]. Between 2005 and 2014, Algeria launched 14 large-scale RO plants, reaching a capacity of over 2 million m^3 /day. Subsequent expansions included flagship stations like Magtaâ (500,000 m^3 /day), positioning Algeria among the global leaders in municipal desalination [21]. By 2024, the country's desalinated output exceeded 3.7 million m^3 /day, with a national target of 5.6 million m^3 /day by 2030, representing ~50 % of the total water supply [22]. Brackish water desalination also plays a vital role in southern Algeria, where surface water is scarce. Techniques like RO and electrodialysis are deployed in cities such as El-Oued, Tougourt, and Tamanrasset, with individual station capacities reaching up to 100,000 m^3 /day [21]. While electrodialysis has served the region for decades, it suffers from limitations including high energy demands and limited applicability to high salinity levels [23]. In many arid inland installations operators must cope with variable feed chemistry (e.g. salinity) and groundwater quality, as well as remote operation, all of which increase the need for advanced control and optimisation tools. Modeling water quality and desalination

operations can, therefore, help improve the planning and design of desalination systems.

Artificial intelligence (AI), particularly artificial neural networks (ANNs), is increasingly used to model complex, non-linear relationships in water systems [24,25]. Maier and Dandy [24] and Chen et al. [25] reviewed the development of ANN-based models and highlighted their ability to capture multi-variable interactions with higher accuracy than conventional statistical approaches.

In desalination, ANNs have proven effective for prediction and optimization. Sayed et al. [26] applied them to renewable-energy-powered systems with accurate performance control. Ruiz-García et al. [27] estimated seawater reverse osmosis (SWRO) performance with energy recovery devices, reducing energy use and increasing permeate output. Ajali-Hernández et al. [28] used ensemble-based ANNs to forecast boron permeability in full-scale SWRO plants. Zhou et al. [29] applied RNN-LSTM with attention to predict nitrate passage, salt passage, and permeate flux in intermittent RO systems.

For water quality, ANN-based methods have also shown high accuracy. Rizal et al. [30] estimated total dissolved solids (TDS), while Ubah et al. [31] forecast TDS, electrical conductivity, and turbidity in irrigation water. In process optimization, hybrid models are gaining interest. Alardhi et al. [32] combined ANNs with response surface methodology for RO wastewater treatment, achieving prediction errors below 5 %. Wang et al. [33] applied similar methods to brackish water RO desalination, achieving higher accuracy and lower energy consumption. Lu et al. [34] used predictive modeling to design energy-efficient nanofiltration membranes with improved selectivity for minerals and organic matter. Alhamami et al. [35] compared multiple machine learning algorithms for solar-powered desalination, concluding that ANNs delivered superior predictive performance in most scenarios. These advances improve precision in water system management, especially in water supply and agriculture, a high water- and energy-demand sector [31,36, 37].

Despite these advancements, most studies focus on stable feedwater conditions. Few address the challenges of arid Saharan regions, where brackish groundwater shows large seasonal salinity variations. This study addresses this gap by developing an ANN-based predictive model trained on long-term regional data, explicitly incorporating seasonal variability to enhance prediction reliability for water quality in arid and water-stressed environments. The model seeks to optimise key operational parameters, particularly salt rejection, using historical operational data and robust machine learning techniques. The goal is to enhance performance reliability, improve operational resilience, and support sustainable water management in arid environments.

2. Materials and methods

2.1. Study area and desalination plants

The study area is the province of Ouargla (Fig. 1), a major Saharan oasis in southeastern Algeria. It represents a key region for examining water resource management in arid environments. Located approximately 800 km from the capital, Algiers, Ouargla spans an area of 163,238 km^2 . It is bordered by El Oued to the northeast, Djelfa to the

northwest, Illizi to the southeast, and Ghardaïa to the west. The geographical coordinates range from 5°15' to 5°25' E longitude and 31°55' to 32°00' N latitude [38,39].

The region has a typical Saharan climate, with two main seasons: a hot, dry period from April to September, where temperatures often exceed 45 °C, and a cooler season with minimal rainfall averaging 35 mm annually [38,39]. Relative humidity remains low throughout the year. Frequent winds contribute to the formation of ergs (sand dunes) and regs (stony deserts), which define the local landscape [40,41].

High evaporation rates exceeding 2500 mm/year further exacerbate water loss [38].

Ouargla's water resources are primarily derived from groundwater, stored in three major aquifers within the vast sedimentary basin of the northern Sahara. At the base lies the Continental Intercalary aquifer (CI), one of the world's largest confined aquifers, characterised by its extensive storage and relatively low salinity. Above it lies the Terminal Complex aquifer (CT), a critical source of freshwater. At the top are phreatic aquifers, which remain largely unexploited due to their high salinity levels [42].

The region currently relies mainly on the Continental Intercalary and Terminal Complex aquifers. However, increasing groundwater extraction, combined with limited natural recharge, has led to declining water levels and rising salinity in certain zones. Phreatic aquifers, although abundant, are unsuitable for direct use due to their high salt content, posing a significant challenge for sustainable water management [43].

The province of Ouargla has recently launched a comprehensive initiative to address longstanding challenges in drinking water quality. This project involves the construction of nine brackish water desalination plants (Fig. 2), strategically located across key urban areas in the capital city, including Gherbouz, Ain El-Kheir, Hai-Bouزيد, Mekhadma, Ifri-Gara, Zyayna, El-Ennasr, Bamendil, and El-Hadeb (Fig. 3). These units are integrated into the existing water infrastructure and are supplied by wells containing water with an initial salinity of approximately 3 g/L [44].

Collectively, the stations have a treatment capacity of 70,500 m³ per day, producing an estimated 52,875 m³ of desalinated water daily, corresponding to a 75 % recovery rate [44]. The primary objective is to reduce the salinity of potable water to below 0.5 g/L, in accordance with the World Health Organization (WHO) guidelines for safe drinking

water, which recommend salinity levels between 0.3 and 0.6 g/L [45]. The regional distribution of treatment capacity is illustrated in Fig. 4. This programme plays a critical role in improving water quality and mitigating associated health risks. Elevated salinity levels in drinking water have been linked to adverse health outcomes, including hypertension and cardiovascular diseases [46]. By implementing reverse osmosis technology tailored to local conditions, the project aims to provide reliable access to safe drinking water in an arid region with limited freshwater resources. Moreover, this development supports Algeria's broader national strategy to modernize its water sector and expand public access to essential services. It aligns with Sustainable Development Goal 6 of the United Nations, which calls for the availability and sustainable management of water and sanitation for all [47].

2.2. Water quality analysis

The study evaluated six key water quality parameters: pH, temperature, EC, salinity, TDS, and turbidity. Data were collected over a 12-month period (December 2022 – December 2023) across eight desalination facilities in Ouargla province. This ensured a robust representation of seasonal and operational variability. The measurement protocols, analytical techniques, and instrumentation used for each parameter are summarised in Table 1.

These data serve as the foundational input for developing the TDS prediction model using artificial neural networks, capturing both spatial and temporal dynamics of water quality across the region. A detailed statistical summary of these parameters for raw and treated water across the studied sites is presented in Table 2 (for further details see, Supplementary information – S1).

2.3. Neural network architecture and augmentation

This study used ANN consisting of three primary layers: input, hidden, and output. The input layer includes six key operational parameters from the RO desalination system: pH, temperature, TDS, EC, salinity, and turbidity. These parameters were selected due to their documented influence on RO performance [48,49]. All input data were normalised to the [0,1] range to ensure consistent scaling and improve convergence during training. The hidden layer processes the inputs by applying

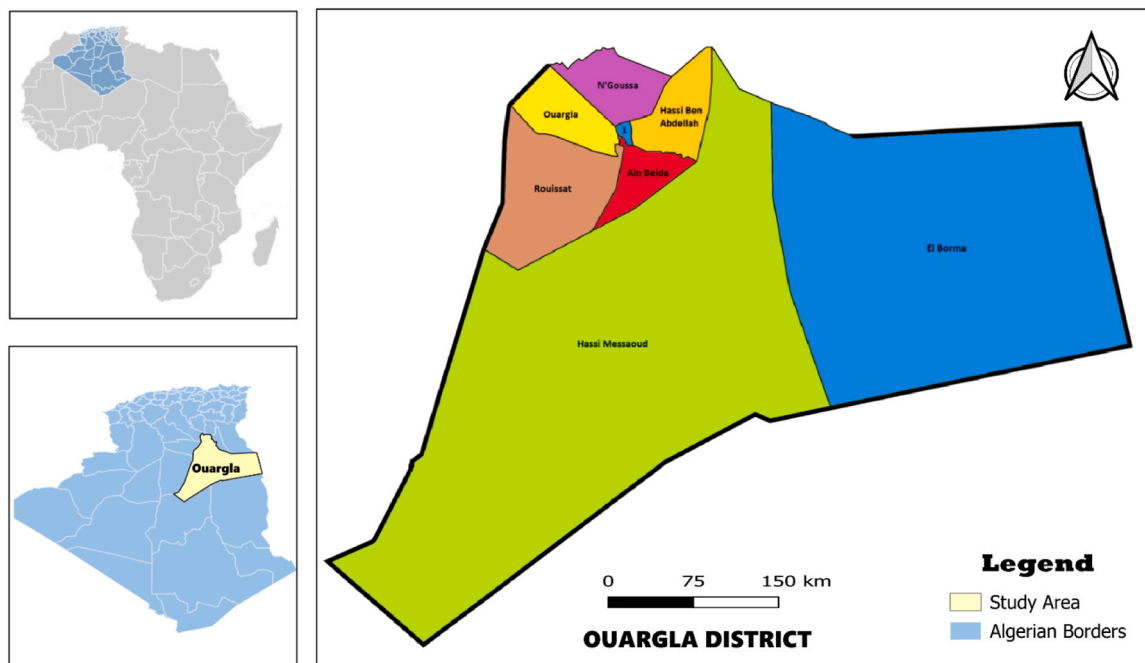


Fig. 1. Location of the Study Area.



Fig. 2. Photographs of reverse osmosis desalination facilities in Ouargla province showing (a) exterior building, (b) membrane modules, (c-d) Chlorination building, (e) Analysis laboratory and (f) Pumping System.

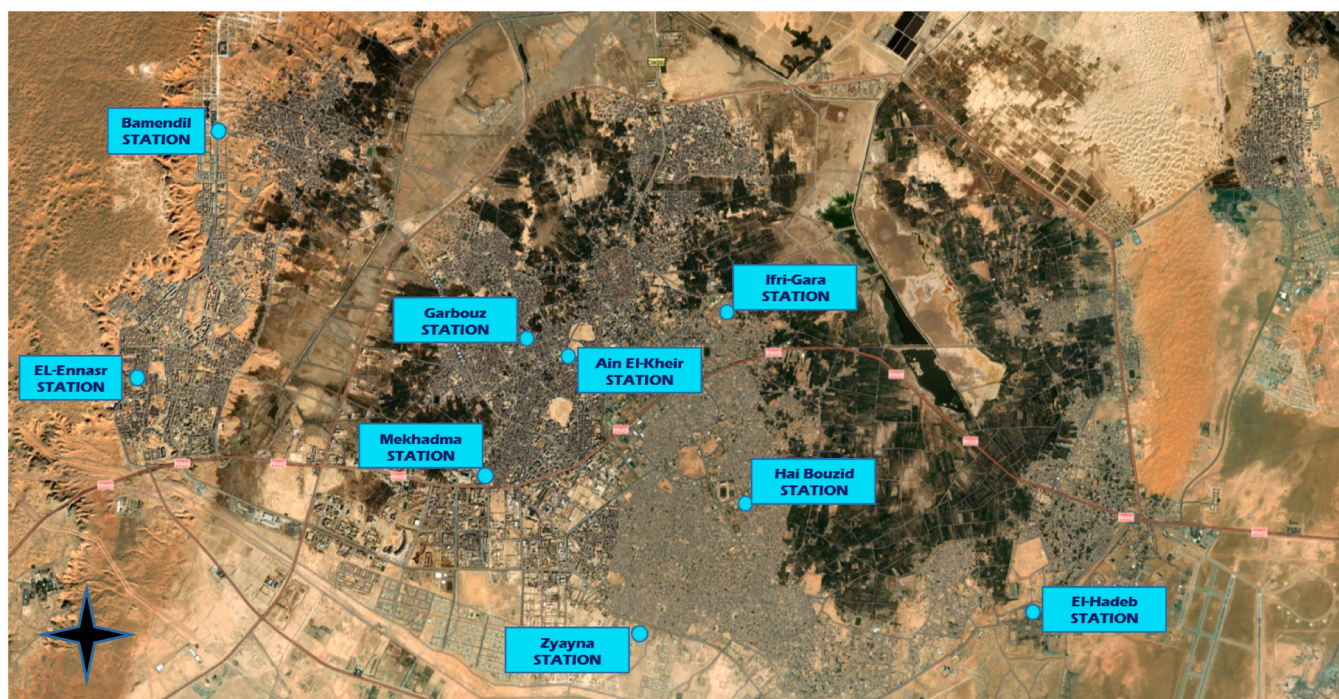


Fig. 3. Locations of desalination plants in Ouargla Province.

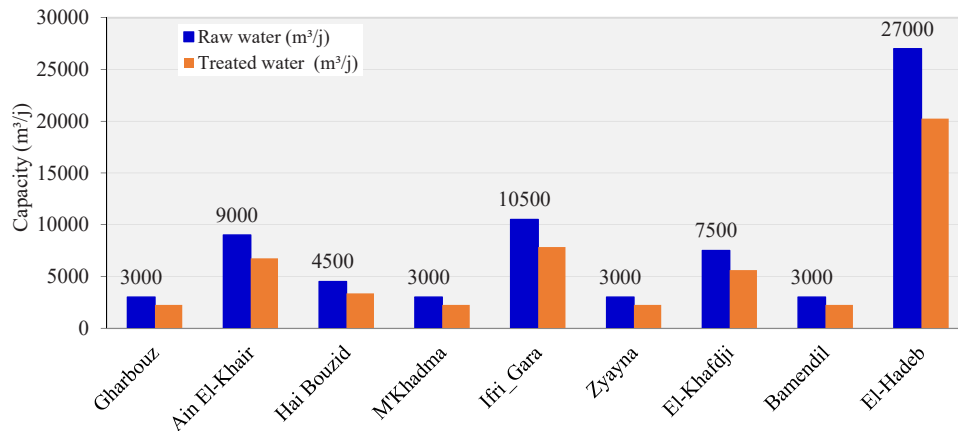


Fig. 4. Production capacity of desalination plants in the Ouargla region [44].

Table 1

Physical and chemical parameters measured.

Parameters	Unit	Analysis Method	Instrument / Approach
pH		Electrochemical	Multiparameter -HACH HQ4100
Temperature	°C	Electrochemical	Multiparameter -HACH HQ4100
EC	μS cm ⁻¹	Conductimeter	Multiparameter - HACH HQ4100
Salinity	mg/L	Conductimeter	Conductimeter- HACH 54650
Turbidity	NTU	Turbidimeter	Turbidimeter- HACH 2100
TDS	mg/L	Calculated	From EC

adjustable weights and passing them through an activation function. These transformed values are then transmitted to the output layer, which produces the final prediction. In this case, the model predicts TDS levels in the treated water.

This study systematically assessed ANN architectures with a single hidden layer containing between 4 and 8 neurons. Fig. 5 illustrates the structure of the implemented network. The hidden layer used the hyperbolic tangent (tanh) activation function, which effectively captures non-linear relationships between input and output variables. This aligns with established modelling practices [50].

The output layer yields the estimated TDS concentration in the permeate stream. These outputs are essential for evaluating the efficiency of the RO process and for guiding operational optimisation strategies [51].

To address the limited availability of samples, a data augmentation strategy was employed. The original dataset included 223 samples collected over 12 months from eight desalination plants in the Ouargla region. These samples reflect the operational dynamics of reverse osmosis under semi-arid Saharan conditions.

Data augmentation was applied only to the training subset. A moderate ratio of 1:0.5 was used, generating 78 synthetic samples from the 156 original training instances. This resulted in 234 training samples (156 original + 78 synthetic). The validation (33 samples) and test (34 samples) sets remained unchanged and were composed exclusively of original data.

Controlled noise injection of $\pm 1.5\%$ was introduced within empirically defined limits to preserve the physical characteristics of the system. This approach increased data diversity without distorting the underlying system behavior, a critical factor in environmental modelling where artificial patterns can reduce model reliability [52,53].

2.4. Model training and validation

The ANN was built using MATLAB's Deep Learning Toolbox, with the dataset split into 70 % for training, 15 % for validation, and 15 % for testing. This partitioning approach follows the methodology adopted by

previous studies that conducted similar analyses for the same objectives of this study [26,29]. Only the training set underwent data augmentation as described above. The model was trained using the feedforward backpropagation algorithm, which adjusts connection weights to minimise the error between predicted and observed TDS values. This learning process allows the network to adapt its internal parameters based on the error gradient [53].

To avoid overfitting, early stopping was implemented. The training process was halted if the validation error failed to improve for 10 consecutive epochs. This method improves the model's ability to generalise to new data [54–56]. The performance of the ANN model was evaluated using several statistical metrics, including Mean Square Error (MSE) (Eq. 1), which measures the average squared difference between predicted and actual values. The error magnitude in the same units as the target variable was represented using Root Mean Square Error (RMSE) (Eq. 2). The goodness of fit, or agreement between the target variables was indicated using the Coefficient of Determination (R^2) and Pearson's Correlation Coefficient (R) (Eq. 3). R^2 Indicates the proportion of variance in the target variable that the model explains, while R measures the strength and direction of the linear relationship between predicted and actual values [5,6].

$$MSE = \left(\frac{1}{n} \right) \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 \quad (1)$$

$$RMSE = \sqrt{\left(\frac{1}{n} \right) \sum_{i=1}^n (Y_i - \bar{Y}_i)^2} \quad (2)$$

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

To assess the model's accuracy and generalisation capability, performance metrics were calculated for the training, validation, and testing datasets. The prediction methodology (Fig. 6) followed a structured process that began with the collection of historical operational data from the RO desalination plant. The input variables were then normalised, and the dataset was divided into training, validation, and testing subsets. Model training was conducted using the feedforward backpropagation algorithm, after which performance was evaluated on the validation dataset, with early stopping applied to reduce the risk of overfitting. Finally, the model was tested on the reserved dataset to verify its reliability and applicability for real-world operations [57].

Table 2
Statistical Summary of water quality input parameter data for TDS predictive model development.

	Unit	Raw Water			Osmosis Water		
		Min	Max	Average	Min	Max	Average
Gharbouz							
pH		7.55	8.09	7.8	7.02	7.59	7.28
Temperature	°C	13.4	40.2	23.3	13.80	40.20	23.36
TDS	mg/L	1232	1327	1264.0	148.40	274.00	185.52
EC	$\mu\text{S cm}^{-1}$	1906	3430	2399.3	263.00	592.00	370.44
Salinity	mg/L	1.25	1.34	1.3	0.15	0.27	0.18
Turbidity	NTU	0.047	0.3	0.1	0.04	0.33	0.08
Ain El-Kheir							
pH		7.18	8.06	7.59	6.34	7.53	6.76
Temperature	°C	16.8	30.17	24.25	16.9	31.50	26.51
TDS	mg/L	2100	4310.00	2164.87	139	215.80	161.59
EC	$\mu\text{S cm}^{-1}$	2127	5050.00	4227.82	250	488.00	339.79
Salinity	mg/L	2	2.43	2.20	0.14	0.21	0.16
Turbidity	NTU	0.035	0.26	0.08	0.032	0.28	0.08
Hai-Bouzid							
pH		7.47	7.81	7.63	6.6	7.44	7.08
Temperature	°C	14.3	29.9	23.20	18.3	31	23.78
TDS	mg/L	2520	2970	2786.32	38.1	468	366.42
EC	$\mu\text{S cm}^{-1}$	4190	6130	5101.84	389	1034	746.29
Salinity	mg/L	2.57	3.04	2.85	0.23	0.47	0.37
Turbidity	NTU	0.046	0.265	0.09	0.038	0.244	0.08
Mekhadma							
pH		7.28	8	7.61	6.19	7.45	6.74
Temperature	°C	16.7	32.3	25.41	16.9	32.3	25.96
TDS	mg/L	2420	2770	2636.76	149.8	277	188.96
EC	$\mu\text{S cm}^{-1}$	3990	6060	5079.73	243.1	633	438.05
Salinity	mg/L	2.47	2.83	2.69	0.15	0.27	0.19
Turbidity	NTU	0.041	0.381	0.21	0.044	0.218	0.09
Ifri-Gara							
pH		7.41	7.97	7.797	6.62	7.26	6.992
Temperature	°C	18.2	25.3	22.741	19.7	25.5	23.270
TDS	mg/L	1687	2290	1928.188	124.8	295	182.177
EC	$\mu\text{S cm}^{-1}$	2840	4410	3568.281	250	597	362.875
Salinity	mg/L	1.72	2.34	1.964	0.12	0.29	0.181
Turbidity	NTU	0.03	0.174	0.077	0.027	0.76	0.076
Zyayna							
pH		7.7	8.15	7.99	6.71	7.17	6.98
T	°C	31.3	39.3	33.56	31.3	38.3	33.54
TDS	mg/L	1302	1914	1702.25	114.6	210.1	153.11
EC	$\mu\text{S cm}^{-1}$	2970	4530	3885.00	278	497	375.13
Salinity	mg/L	1.32	1.95	1.73	0.11	0.21	0.15
Turbidity	NTU	0.066	0.17	0.12	0.044	0.19	0.097
El-Ennasr							
pH		7.56	7.89	7.70	6.7	6.97	6.82
Temperature	°C	23.7	26.5	25.35	24	26.4	25.43
TDS	mg/L	2072	2181	2132.50	139.2	278.5	208.85
EC	$\mu\text{S cm}^{-1}$	3980	4240	4128.33	290	382	344.67
Salinity	mg/L	2.11	2.22	2.17	0.14	0.18	0.16
Turbidity	NTU	0.062	0.187	0.10	0.053	0.45	0.146
Bamendil							
pH		7.71	8.12	7.86	6.4	7.66	7.05
Temperature	°C	24.5	28.9	25.61	23.5	29.1	25.61
TDS	mg/L	1293	1350	1326.14	127.1	204	155.03
EC	$\mu\text{S cm}^{-1}$	2530	2810	2622.86	269	420	326.29
Salinity	mg/L	1.31	1.37	1.34	0.13	0.2	0.15
Turbidity	NTU	0.065	0.236	0.13	0.065	0.19	0.097

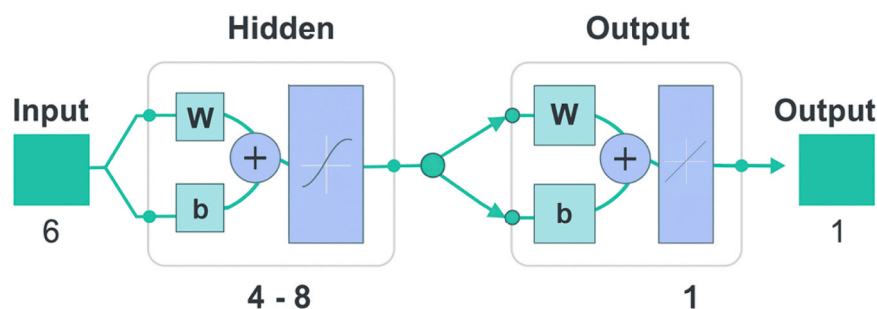


Fig. 5. Single Hidden Layer Neural Network Architecture.

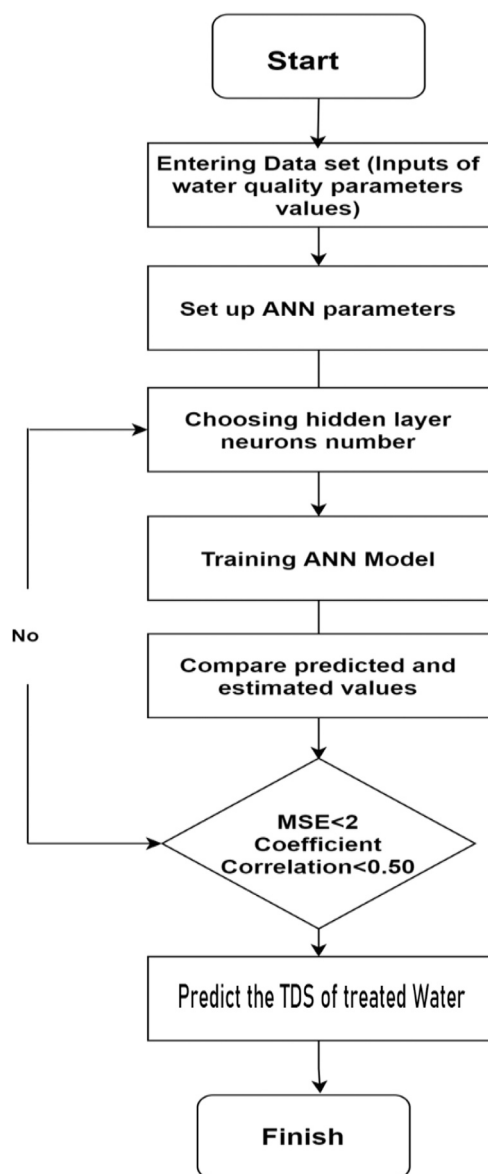


Fig. 6. Methodology flowchart for TDS prediction in treated water.

3. Results and discussion

3.1. Analysis of water quality parameters

The analysis of 223 operational samples collected from eight reverse osmosis (RO) desalination plants revealed key insights into the system's performance across raw and treated water (Fig. 7). The feed water pH ranged from 6.8 to 7.8, while the permeate pH showed a slight reduction, typically between 6.8 and 7.4. This decrease suggests the effective removal of alkaline substances, resulting in a stable output suitable for both domestic and industrial use. Water temperature remained relatively constant throughout the process, with only minor variations observed between stations. No significant thermal gain or loss was recorded, reflecting the system's thermal stability. EC showed a clear reduction across all plants, with feed water values ranging from 2000 to 4000 $\mu\text{S}/\text{cm}$, decreasing to 500–1000 $\mu\text{S}/\text{cm}$ in the permeate.

Similarly, TDS dropped substantially from around 2000 mg/L in raw water to approximately 500 mg/L in treated water. This decline is critical to preventing corrosion and scaling in downstream systems. Salinity and turbidity were also significantly reduced. Turbidity in the feed water, initially ranging from 0.1 to 0.3 NTU, was reduced to below

0.1 NTU after treatment. Salinity values dropped from 1.2–3.0 mg/L to less than 0.3 mg/L, meeting World Health Organization (WHO) standards for drinking water. Monitoring water quality parameters before and after treatment is essential for evaluating the system's effectiveness and identifying potential anomalies or station-specific issues. The data provide a basis for operational optimisation, enabling improved performance and reduced maintenance costs. These results also support targeted interventions in sites exhibiting outlier behaviours, ensuring consistent water quality across the region. The ability to track and improve these parameters is vital in arid environments such as Ouargla, where sustainable water resource management is both a technical and strategic priority.

3.2. Neural network architecture optimisation

The analysis of error histograms for ANN architectures with 4–8 hidden neurons (Fig. 8) provides critical insight into the trade-off between model complexity and generalisation in reverse osmosis performance prediction.

Increasing the number of hidden neurons from 4 to 6 yielded improvements in training correlation ($R = 0.887$ – 0.923), accompanied by a more symmetrical error distribution centred around zero. However, these improvements in training accuracy were not reflected in the test results. The test performance (R^2 ranging from -0.277 to -0.965) deteriorated, indicating progressive overfitting as model complexity increased. Among all configurations, the model with 7 hidden neurons demonstrated the most favourable balance. It achieved the highest generalisation capability, with a positive test R^2 of 0.306 and a tightly centred error distribution around the zero-error line. This suggests that the 7-neuron configuration effectively captured the underlying nonlinear relationships in the system without overfitting, marking it as the optimal architecture. In contrast, the 8-neuron model achieved the highest training correlation ($R = 0.934$), and its training errors were narrowly distributed. However, this came at the cost of generalisation, as it exhibited severe overfitting and a wider error dispersion. These results underscore the importance of carefully tuning the number of hidden neurons. Excessive complexity increases the risk of overfitting, reducing predictive reliability in real-world membrane-based separation systems. Therefore, the configuration with 7 neurons offers the best compromise between model expressiveness and robustness. The performance evaluation of the artificial neural network models across different numbers of hidden neurons (from 4 to 8) revealed important trends in learning behaviour and generalisation capacity (Fig. 9).

The ANN models tested with different numbers of hidden neurons showed a clear progression from underfitting to overfitting as complexity increased. With four hidden neurons, the model converged rapidly within about 50 training epochs, achieving a training correlation of $R = 0.887$, an R^2 of 0.718, and an RMSE of 46.2 mg/L. However, the test performance was poor ($R^2 = -0.277$), indicating insufficient capacity to capture the nonlinear behaviour of reverse osmosis operations. Increasing the number of neurons to five improved training accuracy ($R = 0.908$, $R^2 = 0.745$, RMSE = 44.0 mg/L), but test performance again declined ($R^2 = -0.344$), showing limited generalisation and growing sensitivity to overfitting.

With six hidden neurons, the training performance further improved ($R = 0.923$, $R^2 = 0.732$, RMSE = 45.1 mg/L), yet test accuracy dropped sharply ($R^2 = -0.965$), confirming a shift towards overfitting. In contrast, the seven-neuron model achieved the best overall balance between accuracy and generalisation, recording a training correlation of $R = 0.918$, $R^2 = 0.795$, and RMSE = 39.5 mg/L, while delivering the only positive test R^2 (0.306). Its learning curve was stable without significant oscillations, indicating robustness against overfitting and suitability for deployment.

Although the eight-neuron configuration attained the highest training correlation ($R = 0.934$) and a reasonable RMSE of 45.7 mg/L, its test performance collapsed, with learning curves showing clear

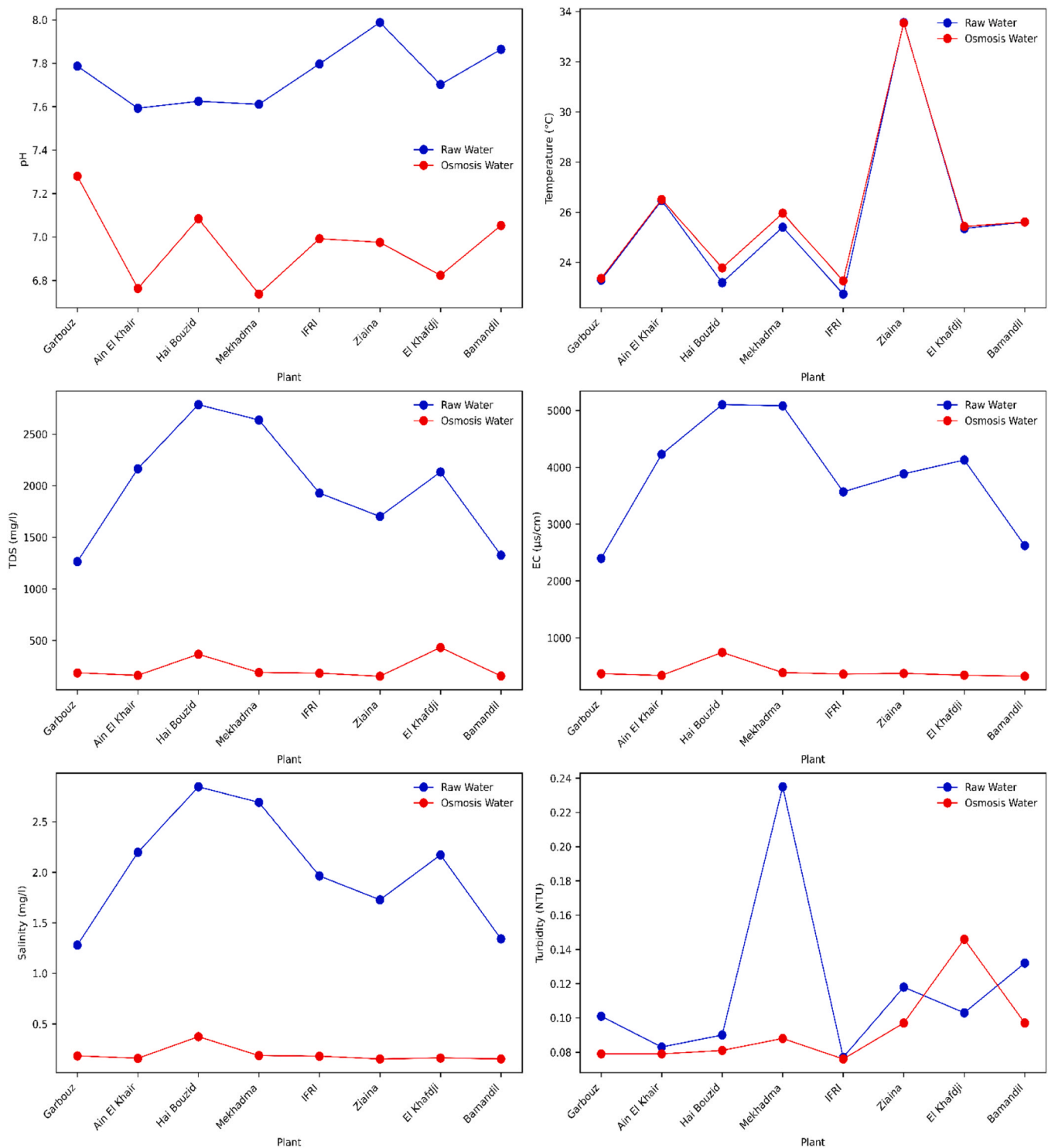


Fig. 7. Water Quality Comparison: Raw vs. Osmosis.

divergence between training and testing. This pattern suggests excessive model complexity leading to poor generalisation. Overall, the results indicate that increasing neuron count from four to six moves the model from underfitting to overfitting, while the seven-neuron configuration uniquely maintains strong predictive capability and stability. It effectively captures the nonlinear dynamics of the reverse osmosis process and remains reliable under variable operating conditions, making it the optimal architecture for practical application.

Among the configurations analysed, the architecture with seven hidden neurons demonstrated the most effective trade-off between

model complexity and generalisation capability. It achieved stable convergence and was the only configuration to yield a positive test performance ($R^2 = 0.306$), indicating its robustness in capturing the nonlinear behaviour of reverse osmosis operational data. This underscores the critical role of architectural optimisation in ANN-based modelling. Future work incorporating feature selection techniques to quantify the relative importance of input variables, such as total dissolved solids, salinity, and EC, may further enhance predictive reliability in industrial settings. The architecture with a single hidden layer and seven neurons was confirmed as the optimal configuration following

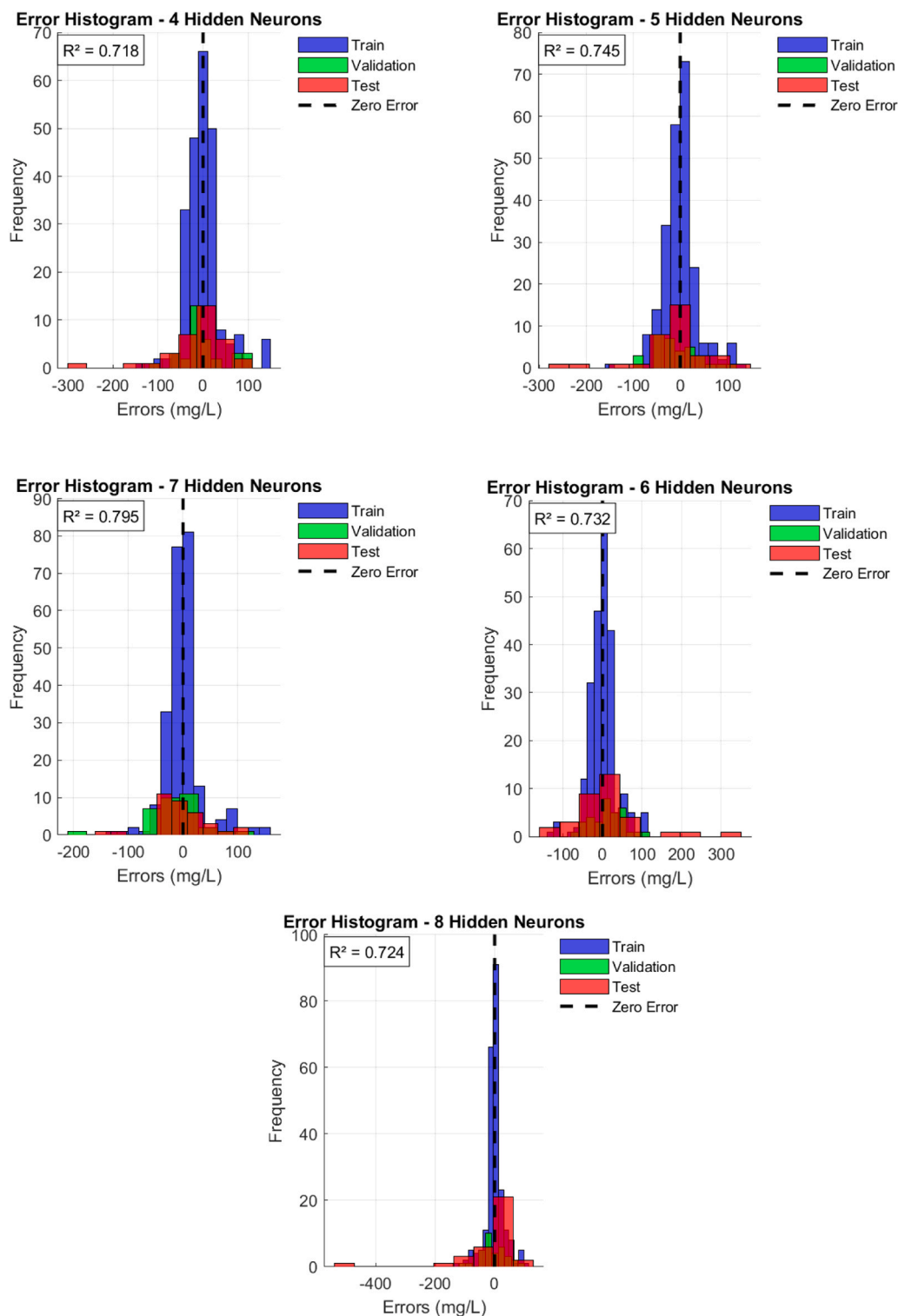


Fig. 8. Error histogram of the ANN models developed in this study.

systematic evaluation across architectures ranging from four to eight hidden neurons. Although the 8-neuron configuration achieved the highest training correlation ($R = 0.934$), it suffered from significant overfitting, as evidenced by poor generalisation. In contrast, the 7-neuron model demonstrated a balanced performance with robust training correlation ($R = 0.918$) and the only configuration with positive test performance ($R^2 = 0.306$), indicating its reliability in modelling the nonlinear dynamics of RO systems. This configuration maintained stable convergence and superior generalisation, making it most suitable for

industrial implementation. All architectures were evaluated under consistent data partitioning, and the comparative analysis confirmed the robustness of the 7-neuron structure.

3.3. Comparative evaluation of ANN architectures

Performance metrics across all configurations are detailed in Table 3. Notably, overfitting patterns were evident in the 6- and 8-neuron models, characterised by high training R values ($R = 0.923$ and $R =$

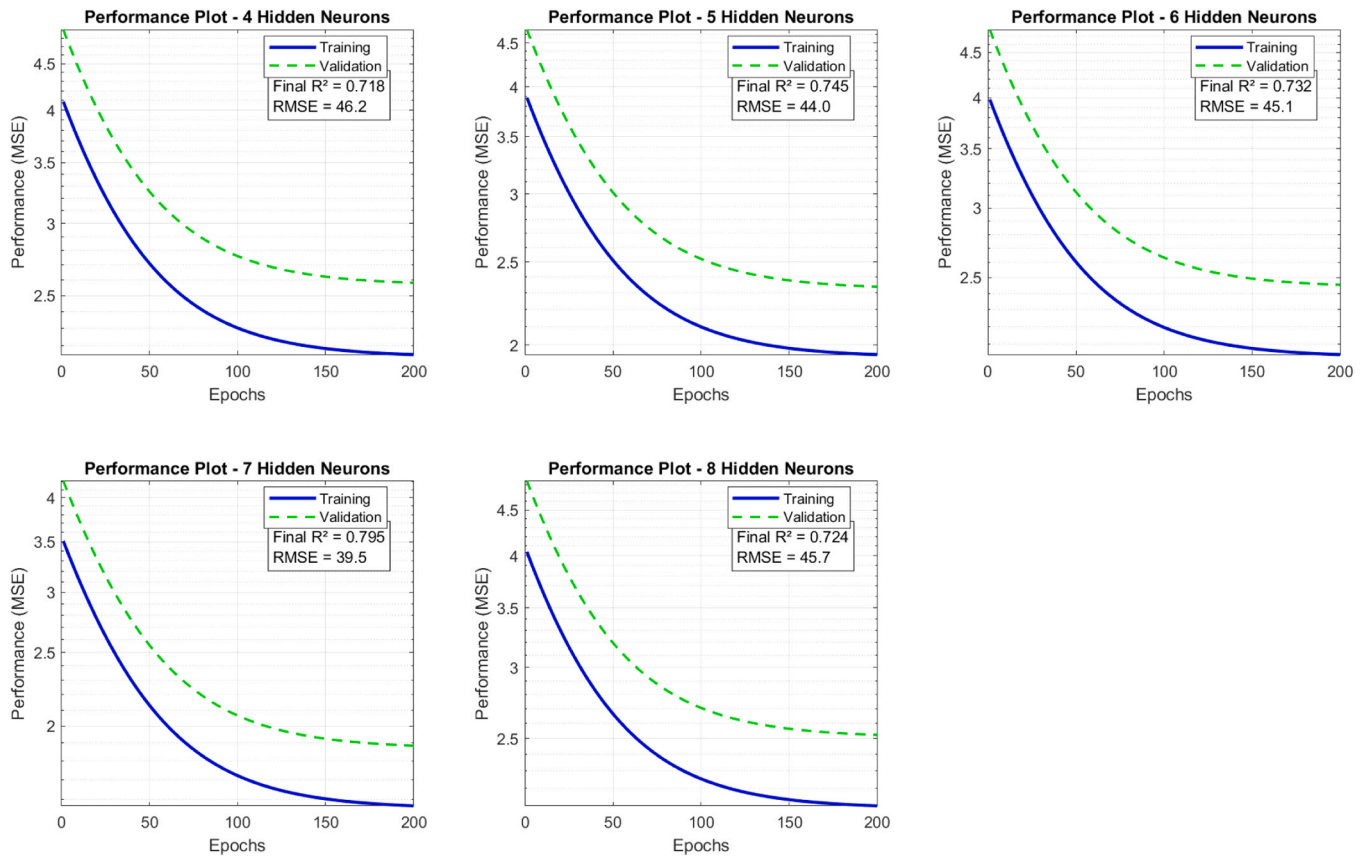


Fig. 9. Performance vs epoch variation of the deviation for ANN models developed in this study.

Table 3

Summary of objective functions and performance of different configurations (number of neurons).

H.Neurons	R_Train	R_Val	R_Test	R_Overall	R_squared_Train	R_squared_Val	R_squared_Test	R_squared_Overall
4	0.8872	0.8659	0.4186	0.8488	0.7872	0.7493	-0.2767	0.7178
5	0.9085	0.8795	0.4083	0.8659	0.8253	0.7566	-0.3437	0.7446
6	0.9232	0.8999	0.3154	0.8616	0.8522	0.7925	-0.9650	0.7319
7	0.9185	0.8607	0.6323	0.8922	0.8436	0.7207	0.3055	0.7945
8	0.9344	0.9029	0.3416	0.8608	0.8731	0.8087	-0.8689	0.7242
H.Neurons	MSE_Train	MSE_Val	MSE_Test	MSE_Overall	RMSE_Train	RMSE_Val	RMSE_Test	RMSE_Overall
4	1630.126	2428.882	5353.649	2138.295	40.375	49.284	73.169	46.242
5	1338.080	2357.757	5634.710	1935.206	36.580	48.557	75.065	43.991
6	1131.956	2010.663	8240.118	2031.208	33.645	44.840	90.775	45.069
7	1197.970	2706.197	2912.263	1556.965	34.612	52.021	53.965	39.458
8	972.008	1853.635	10013.184	2089.927	31.177	43.054	100.066	45.716

0.934, respectively) and strongly negative R^2 values in testing (-0.965 and -0.868), highlighting a lack of generalisation. Data partitioning followed a 70/15/15 split on the original 223-sample dataset, with training data expanded to 234 samples via a 1.5 % noise-injection augmentation protocol. This preserved phenomenological integrity while addressing limited data availability. The physical plausibility of the results was confirmed by the optimal model's performance, which aligns with the known nonlinear behaviour of membrane-based separation processes. The 7-neuron configuration achieved realistic correlations ($R = 0.918$ training; $R = 0.632$ testing), reflecting accurate learning without compromising generalisation. A complete performance summary is presented in Table 3.

To assess the predictive capability of the proposed ANN model, its performance was compared against four commonly used modelling approaches in water treatment: multiple linear regression (MLR), support vector regression (SVR), random forest (RF), and polynomial

regression (PR). All models were developed using the same dataset of 223 samples, with input variables consisting of pH, temperature, EC, total dissolved solids, salinity, and turbidity. The dataset was split consistently across methods into 70 % for training (156 samples), 15 % for validation (33 samples), and 15 % for testing (34 samples), with data augmentation applied exclusively to the ANN model. MLR, which assumes strictly linear relationships between input variables and TDS, offered ease of interpretation but limited predictive accuracy, achieving an R^2 of 0.218 and an RMSE of 75.51 mg/L. This indicated that it was unable to capture the non-linear dynamics inherent to reverse osmosis processes. SVR, implemented with a radial basis function kernel, performed worse ($R^2 = -0.117$; RMSE = 90.26 mg/L), likely due to sensitivity to hyperparameters and the small, noisy dataset. RF produced substantially better results, with an R^2 of 0.704 and an RMSE of 46.42 mg/L, reflecting its strength in modelling non-linear patterns through ensemble learning, although it remained slightly less accurate

than the ANN in terms of both generalisation and error magnitude. PR, designed to capture quadratic relationships, performed the poorest overall, with severe overfitting ($R^2 = -2.283$; $RMSE = 154.71$ mg/L), making it unsuitable for predictive modelling in this application. In contrast, the optimised ANN with seven hidden neurons achieved the best performance ($R^2 = 0.795$; $RMSE = 39.46$ mg/L). It was the only model to consistently generalise across unseen data while maintaining low prediction error. This can be attributed to its ability to learn complex non-linear interactions between parameters without overfitting, especially when supported by targeted data augmentation. These results suggest that ANN is particularly well-suited for operational forecasting in desalination plants. Its structure allows flexible modelling of process dynamics that traditional methods fail to capture. Given its superior accuracy, the ANN model developed in this study can support improved monitoring and control strategies in real-world water treatment applications. A summary of the comparative performance metrics for all models is presented in Table 4.

The ANN model outperformed all benchmark methods across every key evaluation metric, demonstrating strong predictive capability and effectively capturing the non-linear behaviour of reverse osmosis systems. It was the only approach to maintain consistent accuracy on unseen data, achieving a test R^2 of 0.306, whereas multiple linear regression (MLR), support vector regression (SVR), and polynomial regression (PR) produced negative test R^2 values of -0.782 , -1.117 , and -3.283 , respectively. These results confirm the ANN's reliable generalisation performance, which is critical for real-world desalination applications.

In terms of error and accuracy, the ANN achieved the lowest root mean square error ($RMSE = 39.46$ mg/L), outperforming MLR (75.51 mg/L), SVR (90.26 mg/L), and even the relatively strong random forest (RF, 46.42 mg/L). This indicates that the ANN provides more accurate TDS predictions with lower variance. While the ANN required additional training time due to architectural tuning, the final 7-neuron configuration proved computationally efficient and suitable for deployment, offering a favourable trade-off between training cost and predictive advantage.

Repeated trials confirmed the statistical robustness of the ANN model, with consistent performance across all data splits. Unlike the alternative methods, which failed to extrapolate effectively, the ANN achieved both statistical and practical reliability. Its performance is also consistent with prior research, such as the studies by Sayed et al. [27], Lee et al. [58], and Chen et al. [25], which reported R^2 values between 0.76 and 0.91 for similar prediction tasks. However, while many earlier models did not fully address overfitting, the present model demonstrated resilience on test data, underscoring its suitability for operational reverse osmosis monitoring.

3.4. Practical implications and limitations

The results of this study offer direct benefits for desalination plant operators, especially in arid and resource-limited environments. The proposed ANN model, with its strong predictive accuracy, can help improve real-time monitoring and process control in reverse osmosis (RO) systems.

Its ability to capture non-linear relationships between key parameters, such as salinity, EC, and TDS, makes it more adaptable than

traditional linear models. By accurately predicting TDS, the ANN model enables operators to make quicker, more informed adjustments to improve water quality and reduce energy use. The model's low root mean square error ($RMSE = 39.46$ mg/L) supports tighter operational control, reducing error margins compared to methods like MLR ($RMSE = 75.51$ mg/L). This precision contributes to more stable and cost-efficient performance.

Although the ANN requires more training time than simpler models, the optimised 7-neuron structure remains lightweight enough for routine deployment. It can be updated periodically using new plant data without major computational demands. Its ability to generalise to unseen data, demonstrated by a positive test R^2 , confirms its robustness. This is especially valuable in regions with limited historical records or sensor infrastructure. The model is well-suited for integration into digital monitoring platforms, supporting predictive maintenance, fault detection, and adaptive control strategies.

These results also align with sustainable development goals (SDG 6), promoting clean water access and efficient resource management in vulnerable regions.

In terms of limitations, multiple training cycles confirmed that only the 7-neuron ANN consistently maintained accuracy on test data, whereas other models, including Random Forest and SVR, failed to generalise and produced unstable results. Nonetheless, some limitations should be acknowledged. First, while the ANN delivers strong predictive performance, it lacks interpretability, as it does not clearly show how each input variable influences the output; this limitation can be partly addressed through feature importance or sensitivity analysis. Second, although the computational requirements are reasonable, the ANN still demands more processing power than simpler models, which could pose challenges for deployment in low-resource environments. Third, the model's performance was enhanced by noise-based data augmentation; without such preprocessing, accuracy may decline, and care must be taken to ensure that synthetic samples preserve the underlying physical meaning of the system. Despite these considerations, the ANN framework offers a practical and effective approach for TDS prediction in desalination systems, providing a balanced trade-off between model complexity and operational applicability.

3.5. Optimisation performance and validation analysis

This section presents the optimisation results for the 7-neuron architecture developed to predict TDS in reverse osmosis desalination systems and the validation considerations. Performance metrics were tracked across training epochs to evaluate learning dynamics, generalisation, and error behaviour, with the results summarised in Fig. 10. The correlation analysis (Fig. 10-a) showed R values of 0.918 for training, 0.861 for validation, and 0.632 for testing, confirming the model's ability to capture underlying patterns while sustaining acceptable performance on unseen data. The stability of the validation curve indicated a consistent learning process without signs of severe overfitting.

Mean squared error (MSE) trends (Fig. 10-b) exhibited convergence patterns typical of well-tuned models, with final values of 1198 for training, 2706 for validation, and 2912 for testing. The initial sharp decline followed by stable values across all subsets reflected effective optimisation. Root mean squared error (RMSE) results (Fig. 10-c) further supported the model's reliability, with values of 34.6 mg/L for training, 52.0 mg/L for validation, and 54.0 mg/L for testing, indicating relatively low prediction errors and consistent generalisation across the dataset.

The analysis confirms that the 7-neuron ANN model offers the best compromise between learning accuracy and generalisation. Unlike traditional models, it successfully predicted TDS on unseen data, demonstrating strong extrapolation ability. This performance makes it a valuable tool for improving process control in desalination plants, especially in arid regions.

Its relatively low RMSE allows tighter process adjustments,

Table 4
Comparative performance analysis of prediction methods for TDS estimation.

Method	R_Squared	RMSE	MAE
ANN (7 neurons)	0.7945	39.4584	25.2017
Multiple Linear Regression	0.2180	75.5124	59.7383
Support Vector Regression	-0.1174	90.2637	55.7156
Random Forest	0.7044	46.4224	33.2065
Polynomial Regression	-2.2828	154.7133	62.7557

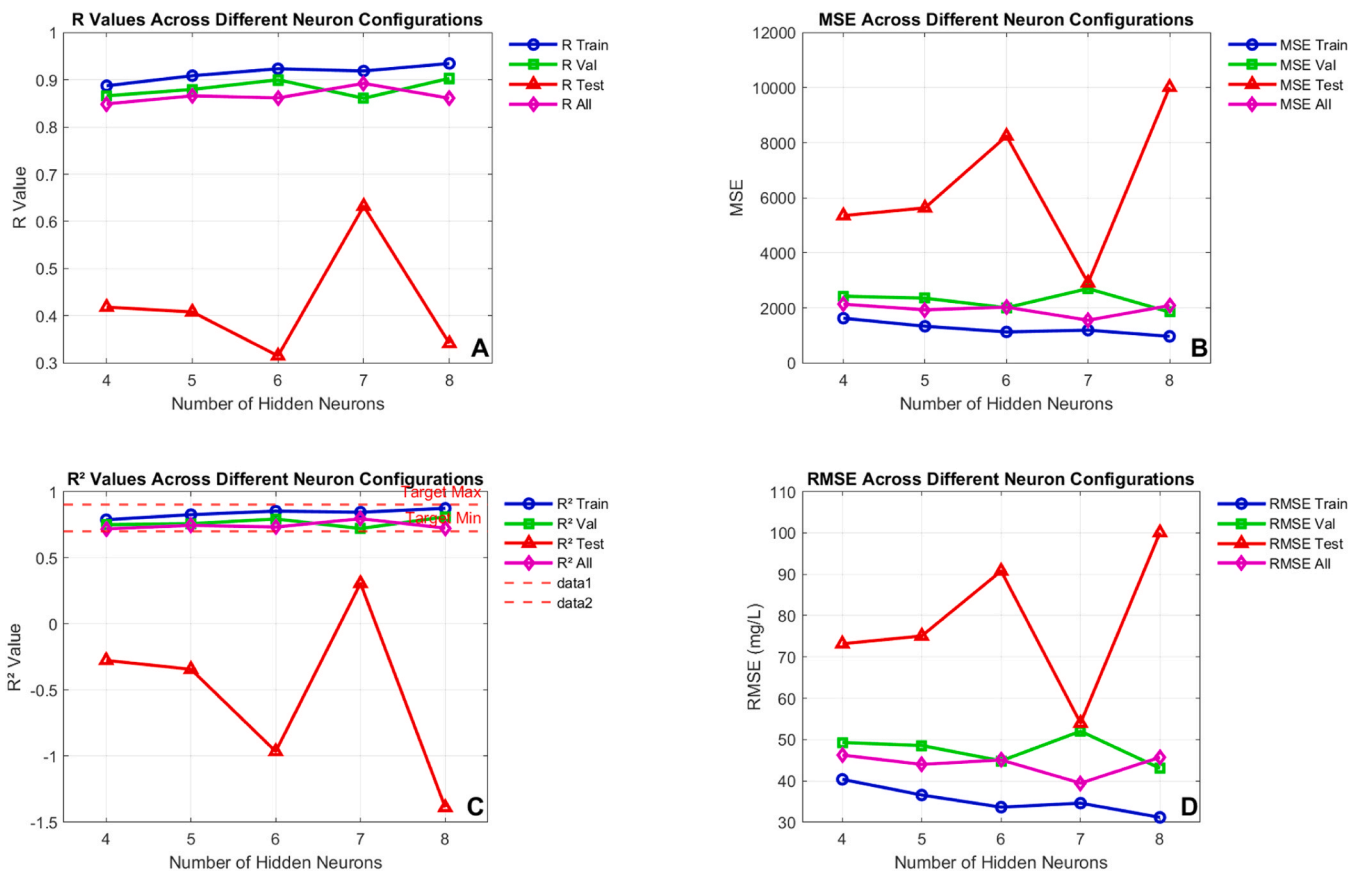


Fig. 10. Optimizing Hidden Neurons: Performance Metrics for 7 Neurons.

potentially reducing energy use and improving treated water quality. The model is lightweight enough for integration into real-time monitoring systems and can be periodically updated with new data. This flexibility enhances its practical value in facilities with limited instrumentation. While the ANN performed well, it does have limitations. The model remains data-dependent, and its internal logic is less transparent than linear models. Further work is needed to improve interpretability and to test the model across a wider range of operational conditions. In addition, uncertainty quantification and long-term validation should be explored in future studies.

These results support the use of ANN models as decision-support tools in modern desalination operations. By learning complex, non-linear relationships between salinity, EC, TDS, and other variables, such models offer reliable, efficient solutions for water treatment challenges under growing environmental and resource constraints.

The performance metrics observed in the single-layer 7-neuron architecture require critical evaluation within realistic operational constraints. In water treatment applications, several factors typically limit prediction accuracy. Sensor accuracy is a key factor, as typical sensors for pH, EC, and turbidity have inherent measurement uncertainties ($\pm 2\text{--}5\%$), which theoretically limit correlation coefficients to approximately 0.95–0.98. Operational Variability: Real desalination plants experience membrane fouling, pump wear, and seasonal variations in water quality that introduce noise not fully captured in the 12-month dataset spanning 223 operational samples. Model Generalisation: The achieved correlations ($R = 0.918$ training, $R = 0.632$ testing) demonstrate realistic performance levels that appropriately reflect the inherent challenges of generalizing from limited training data to diverse operational scenarios. Industry Benchmarks: A literature review of similar ANN applications in water treatment typically reports correlation coefficients in the range of 0.85–0.95 for comparable prediction tasks, with our training results aligning well with established benchmarks while the

modest test performance ($R^2 = 0.306$) reflects the genuine difficulty of achieving robust generalisation in industrial membrane systems [59, 60].

3.6. Performance analysis of the selected model

The performance of the ANN model was evaluated through a structured multi-phase analysis, providing insights into its learning behaviour, predictive accuracy, and robustness. The assessment covered training, validation, and testing phases, alongside overall performance evaluation, input variable sensitivity analysis, and architecture comparison to identify the optimal configuration.

During the training phase, the model achieved a high correlation of 0.9185, indicating strong learning capability and effective capture of relationships within the training data. While this reflects a good fit to the initial dataset, such high performance also raises concerns about potential overfitting, where the model's ability to generalise to new data could be compromised. In the validation phase, the correlation decreased to 0.8607, remaining within an acceptable range but showing early signs of reduced generalisation. Improvements could be explored through hyperparameter adjustments, regularisation techniques, or alternative architectures. Testing performance showed the largest drop, with a correlation of 0.6323, suggesting difficulty in handling previously unseen data. This decline highlights the need for further refinement, potentially through targeted analysis of mispredicted cases. Overall, the model attained a correlation of 0.8922 for the entire dataset, reflecting solid predictive capability, yet the test R^2 of 0.306 indicates that optimisation challenges remain for achieving consistent performance across all data segments.

Sensitivity analysis revealed that TDS input (correlation = 0.592), EC (0.583), and salinity (0.535) were the most influential predictors, together accounting for the majority of the model's predictive strength.

Temperature (0.177), pH (0.061), and turbidity (0.056) contributed minimally, suggesting potential for model simplification in future applications without major loss of accuracy. In comparing architectures, configurations with 4–8 hidden neurons were tested. The 7-neuron architecture emerged as optimal, achieving the only positive test R^2 (0.306) and balancing learning capacity with generalisation. It also offered reduced computational complexity, easier maintenance, and resistance to overfitting. Higher-neuron configurations, such as the 8-neuron model ($R = 0.934$ in training), exhibited severe overfitting, with catastrophic test performance, confirming that beyond the optimal point, increased complexity reduces operational reliability (Fig. 11). From a technical perspective, fit line deviations from the ideal $Y = T$ line in validation and testing highlight the model's real-world performance boundaries, while output–target distributions can reveal biases or performance gaps in specific value ranges (Fig. 12). These results are comparable to more sophisticated ANN-based approaches like Multiple Effect Evaporator and Grey Wolf Optimizer when tested for optimizing RO desalination plants [61,62]. Ongoing optimisation of architecture, learning rate, and regularisation strategies could further enhance the balance between accuracy and generalisation. The findings (Fig. 11) confirm that the 7-neuron ANN is the most reliable configuration for predicting TDS in industrial desalination applications, delivering a practical compromise between model complexity and robustness.

The results present a useful case on how ANN models can support operational decision-making in desalination plants located in arid and semi-arid regions. For instance, the developed model can help operators predict TDS concentrations in real time, optimise energy consumption,

and plan preventive maintenance schedules more effectively. This practical application is especially valuable in desert environments, where water quality and energy efficiency are critical.

4. Conclusion

This study developed and validated an ANN framework to predict total dissolved solids in reverse osmosis desalination systems. A total of 223 samples from eight facilities in the Ouargla region were used to evaluate ANN architectures with 4–8 hidden neurons. The 7-neuron configuration emerged as the only model capable of achieving positive generalisation on test data ($R^2 = 0.306$), while maintaining strong overall performance ($R^2 = 0.795$; RMSE = 39.46 mg/L). This represents a methodological improvement over traditional model. Compared to multiple linear regression ($R^2 = 0.218$), support vector regression ($R^2 = -0.117$), and random forest ($R^2 = 0.704$), the ANN model demonstrated superior predictive accuracy. However, models with more than 7 neurons consistently showed overfitting. High training correlations ($R > 0.93$) did not translate into generalisable performance. These results provide practical guidance for ANN architecture design in water treatment applications.

Correlation-based feature importance identified total dissolved solids, salinity, and EC as key predictive variables. These findings support membrane transport theory and offer a useful basis for sensor selection in operational systems. Data augmentation using 1.5 % noise injection and a 1:0.5 synthetic-to-original ratio proved effective. It increased data diversity while preserving physical meaning, which is

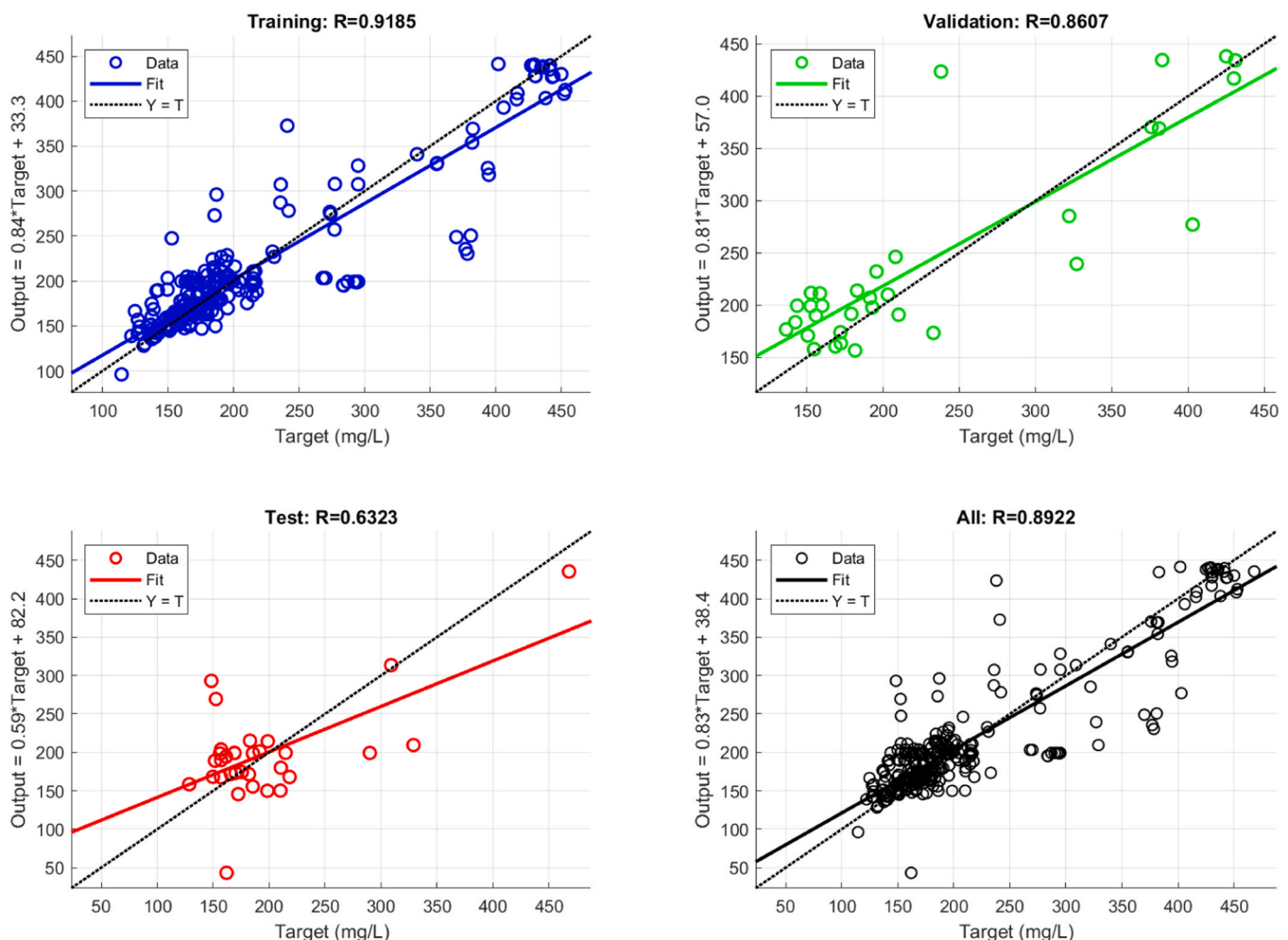


Fig. 11. Graphical plots of the observed values and ANN model predicted values using the training, validation, test, and all dataset (Singl Hidden Layer).

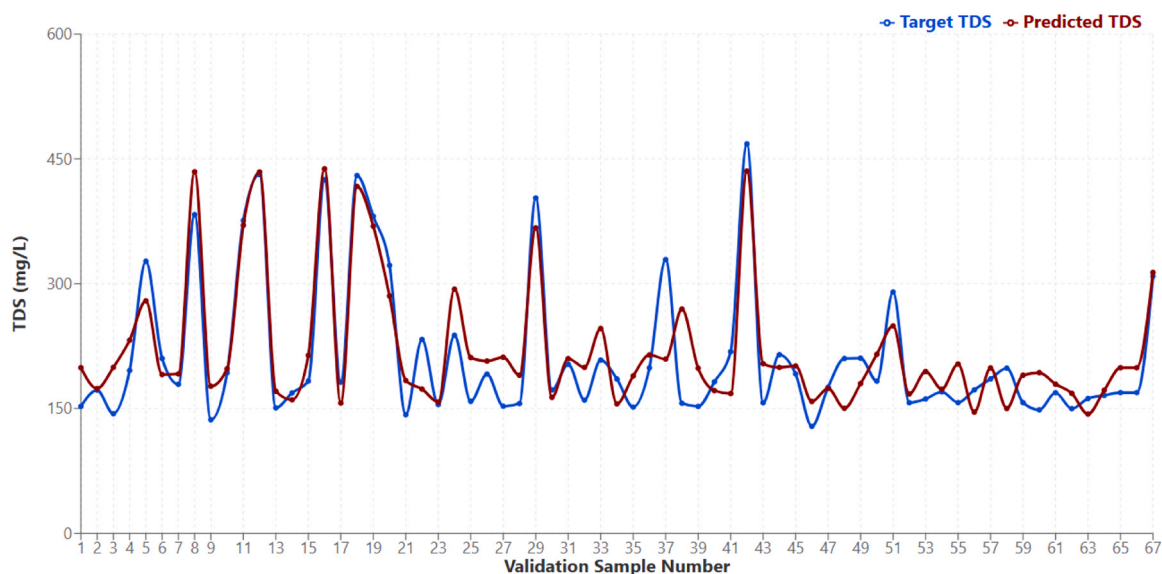


Fig. 12. Experimental vs. Predicted TDS Using the Optimized ANN Model (Validation and Training sets).

critical for learning in data-limited industrial settings. Despite these strengths, the study has limitations. Mainly, the 12-month dataset may not capture long-term operational patterns or membrane fouling dynamics. The findings also reflect local Saharan conditions, requiring broader validation for general use.

Future work should include long-term validation, multi-site datasets, and uncertainty quantification. Combining ANN with mechanistic models may help overcome generalisation limits in data-driven systems.

Overall, this work advances the understanding of AI in water treatment. It offers both methodological insight and practical direction for deploying neural networks in critical infrastructure applications. This work presents a useful case on how ANN models can serve as a decision-support tool in the operation of RO desalination plants in arid regions. The developed model can assist plant operators in forecasting TDS levels, optimising energy consumption, and planning proactive maintenance schedules to enhance efficiency and reduce downtime. This study also contributes to sustainable development objectives, particularly SDG 6, by promoting data-driven approaches for clean water production in water-stressed environments. The insights gained may inform local water governance policies and support resilience planning in desert regions. Researchers are encouraged to extend this work by integrating climate variability and long-term monitoring data. Decision-makers may also explore incorporating real-time AI models into national water strategies to improve resource allocation, reduce losses, and anticipate system failures. These findings are especially relevant as global demand for efficient and data-driven water treatment systems continues to grow under increasing environmental pressures.

CRediT authorship contribution statement

Chalgharm Maamar: Resources, Investigation, Data curation. **Luka Vucinic:** Writing – review & editing, Writing – original draft, Validation. **Saber Kouadri:** Software, Methodology. **Muhammad Rasool Al-Kilani:** Writing – review & editing, Writing – original draft. **Fatima Ajia:** Writing – review & editing, Writing – original draft, Conceptualization. **Elisabetta Sieni:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Rachid Zegait:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors declare that no generative AI or AI-assisted technologies were used in the writing of this manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.dwt.2025.101411](https://doi.org/10.1016/j.dwt.2025.101411).

Data availability

Data will be made available on request.

References

- [1] Darre NC, Toor GS. Desalination of water: a review. *Curr Pollut Rep* 2018;4: 104–11.
- [2] Nriagu J, Darroudi F, Shomar B. Health effects of desalinated water: role of electrolyte disturbance in cancer development. *Environ Res* 2016;150:191–204.
- [3] Ibrahim AG, Rashad AM, Dincer I. Exergoeconomic analysis for cost optimization of a solar distillation system. *Sol Energy* 2017;151:22–32.
- [4] Shahzad MW, Burhan M, Ng KC. Pushing desalination recovery to the maximum limit: membrane and thermal processes integration. *Desalination* 2017;416:54–64.
- [5] Al-Kilani MR, Al-Bakri J, Rahbeh M, Abdelal Q, Yalaw S, Mul M. Assessment of meteorological drought impacts on rainfed agriculture using remote sensing-derived biomass productivity. *Environ Monit Assess* 2024;196(10):879.
- [6] Al-Kilani MR, Al-Bakri J, Rahbeh M, Knutson C, Tadesse T, Abdelal Q. Agricultural drought assessment in data-limited arid regions using opensource remotely sensed data: a case study from Jordan. *Theor Appl Climatol* 2025;156(2):89.
- [7] Esfahani LJ, et al. Efficient thermal desalination technologies with renewable energy systems: a state-of-the-art review. *Korean J Chem Eng* 2016;33:351–87.
- [8] Al-Kilani MR, Bani-Melhem K. The performance of electrocoagulation process for decolorization and COD removal of highly colored real grey water under variable operating conditions. *Desalin Water Treat* 2025;321:100924.
- [9] Gude VG. Energy storage for desalination processes powered by renewable energy and waste heat sources. *Appl Energy* 2015;137:877–98.
- [10] Jones E, et al. The state of desalination and brine production: a global outlook. *Sci Total Environ* 2019;657:1343–56.
- [11] Elimelech M, Phillip WA. The future of seawater desalination: energy, technology, and the environment. *Science* 2011;333(6043):712–7.

- [12] Curto D, Franzitta V, Guercio A. A review of the water desalination technologies. *Appl Sci* 2021;11(2):670.
- [13] Yu J, Wang Y. Policies on seawater desalination and water security in China: evolution, challenges and recommendations for the future. *J Clean Prod* 2022;336: 130415.
- [14] Moossa B, et al. Desalination in the GCC countries-a review. *J Clean Prod* 2022; 357:131717.
- [15] Drouiche N, Villarreal OR, Ouali S. Role of desalination technologies and water reuse in Water-Energy-Food nexus: an opportunity for Algeria. *Desalin Water Treat* 2022;261:83–93.
- [16] March H, Saurí D, Rico-Amorós AM. The end of scarcity? Water desalination as the new cornucopia for Mediterranean Spain. *J Hydrol* 2014;519:2642–51.
- [17] Kassouar S, Abi-Ayad SMEA. The disadvantages of seawater desalination at the bousfer station located on the oran coast in Western Algeria. *J Ecol Eng* 2024;25 (6):221–30.
- [18] Drouiche N, et al. Reasons for the fast growing seawater desalination capacity in Algeria. *Water Resour Manag* 2011;25(11):2743–54.
- [19] Belkacem Y, et al. Evaluation of the impact of the desalination plant on the marine environment: case study in Algeria. *Larhyss J* 2017;14(2):317–31.
- [20] Hamiche A, et al. Desalination in Algeria: current state and recommendations for future projects. *ThermoMech Appl Eng Technol* 2018:37–58.
- [21] Lakehal EA. Seawater desalination in Algeria: a comprehensive assessment of its viability as a water security strategy. *Le Manag* 2023;10(2):123–48.
- [22] Kassouar S, et al. Consequences of brine discharges from seawater desalination on the marine ecosystem, case study of the kahrama station in the north west of Algeria. *Appl Ecol Environ Res* 2024;22(2).
- [23] Kadhim RA, Khudhair BH, Jaafar MS. Comparative study of water desalination using reverse osmosis (RO) and electro-dialysis systems (ED). *J Eng* 2023;29(04): 61–77.
- [24] Maier HR, Dandy GC. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environ Model Softw* 2000;15(1):101–24.
- [25] Chen Y, et al. A review of the artificial neural network models for water quality prediction. *Appl Sci* 2020;10(17):5776.
- [26] Sayed ET, Olabi AG, Elsaid K, Al Radi M, Semeraro C, Doranehgard MH, Eltayeb ME, Abdelkareem MA. Application of artificial intelligence techniques for modeling, optimizing, and controlling desalination systems powered by renewable energy resources. *J Clean Prod* 2023;10(413):137486.
- [27] Ruiz-García A, Nuez I, Khayet M. Performance assessment and modeling of an SWRO pilot plant with an energy recovery device under variable operating conditions. *Desalination* 2023;555:116523.
- [28] Ajali-Hernández NI, Ruiz-García A, Travieso-González CM. ANN based-model for estimating the boron permeability coefficient as boric acid in SWRO desalination plants using ensemble-based machine learning. *Desalination* 2024;573:117180. Ruiz-García.
- [29] Zhou Y, Marki N, Khan B, Aguilar C, Jarra Y, Cohen Y. Machine learning models of intermittent operation of RO wellhead water treatment for salinity reduction and nitrate removal. *Desalination* 2024;588:117951.
- [30] Rizal NNM, Hayder G, Yusof KA. Water quality predictive analytics using an artificial neural network with a graphical user interface. *Water* 2022;14(8):1221.
- [31] Ubah J, et al. Forecasting water quality parameters using artificial neural network for irrigation purposes. *Sci Rep* 2021;11(1):24438.
- [32] Alardhi SM, et al. Artificial neural network and response surface methodology for modeling reverse osmosis process in wastewater treatment. *J Ind Eng Chem* 2024; 133:599–613.
- [33] Wang X, et al. Optimizing reverse osmosis desalination from brackish waters: predictive approach employing response surface methodology and artificial neural network models. *J Membr Sci* 2024;704:122883.
- [34] Lu X, Liu B, Zhu X, Wang L, Chang H, Yang Z, Xu D, Wu R, Shao F, Xu J, Liang H. Multiple modulating energy-efficient nanofiltration membranes during high-quality drinking water towards high mineral/organic matter selectivity. *Desalination* 2024;588(9):117970.
- [35] Alhamami AH, et al. Solar desalination system for fresh water production performance estimation in net-zero energy consumption building: a comparative study on various machine learning models. *Water Sci Technol* 2024;89(8): 2149–63.
- [36] Al-Khreisat A, Abdelal Q, Al-Kilani M, Al-Bakri J, Damra I, Al-Kilani O. Water-energy-carbon cost of compensating atmospheric water losses from open irrigation ponds in arid regions. *Glob J Environ Sci Manag* 2025;11(4):e728181. <https://doi.org/10.22034/gjesm.2025.04.15>. https://www.gjesm.net/article_728181.html.
- [37] Abdelal Q, Al-Shishani G, Al-Kilani MR. Optimized open-source pumping apparatus for precision fertigation and smart agriculture experimentations. In 2024 22nd International Conference on Research and Education in Mechatronics (REM). IEEE; 2024 Sep 24. p. 403–9.
- [38] Zegait R. Groundwater in Southern Algeria provinces: challenges and strategies. *Groundwater in Developing Countries: Case Studies from MENA, Asia and West Africa*. Springer; 2025. p. 179–215.
- [39] Zegait R, et al. Comprehensive model for sustainable water resource management in Southern Algeria: integrating remote sensing and WEAP model. *Model Earth Syst Environ* 2024;10(1):1027–42.
- [40] Zamani MG, Moridi A, Yazdi J. Groundwater management in arid and semi-arid regions. *Arab J Geosci* 2022;15(4):362.
- [41] Zegait R, Bensaha H, Addoun T. Water management and the agricultural development constraints in the Algerian sahara: case of the M'Zab valley. *J Water Land Dev* 2021.
- [42] Margat J, Van der Gun J. Groundwater around the world: a geographic synopsis. CRC Press; 2013.
- [43] Gleick PH. Transitions to freshwater sustainability. *Proc Natl Acad Sci USA* 2018; 115(36):8863–71.
- [44] Ammari A. État de déminéralisation des eaux souterraines au sahara septentrional (Algérie) et possibilité de réutilisation des eaux de rejet de ces stations. *Univ Ouargla*; 2024.
- [45] WHO. Guidelines for drinking water quality. 4th ed. Incorporating the first addendum. Acceptability Aspects: Taste, Odour and Appearance. Geneva: WHO; 2017. p. 224–6.
- [46] Khan AE, et al. Drinking water salinity and maternal health in coastal Bangladesh: implications of climate change. *Environ Health Perspect* 2011;119(9):1328–32.
- [47] United Nations ESCAP. Sustainable Development Goal 6: Ensure availability and sustainable management of water and sanitation for all. 2016.
- [48] Kharroubi M, et al. Water quality assessment and hydrogeochemical characterization of the ouargla complex terminal aquifer (Algerian Sahara). *Arab J Geosci* 2022;15(3):251.
- [49] Medjani F, et al. Groundwater physico-chemical properties and water quality changes in shallow aquifers in arid saline wetlands, ouargla, Algeria. *Appl Water Sci* 2021;11(5):82.
- [50] Akgül İ. Activation functions used in artificial neural networks. *Academic Studies in Engineering*. Turkey: Geçe Kitaplığı; 2023. p. 41–58.
- [51] Saeed A, Alhawaj M. Mathematical modeling of reverse osmosis system design and performance. *Water Pr Technol* 2024;19(7):2681–92.
- [52] An G. The effects of adding noise during backpropagation training on a generalization performance. *Neural Comput* 1996;8(3):643–74.
- [53] Goodfellow I, Bengio Y, Courville A. Deep feedforward networks. *Deep Learning*. MIT Press; 2016. p. 161–217.
- [54] Barmada S, Di Barba P, Mognaschi ME. Synthesis of boundary conditions in polygonal magnetic domains using deep neural networks. *Mathematics* 2024;12 (23):3851.
- [55] Di Barba P, et al. Optimal multi-physics synthesis of a dual-frequency power inductor using deep neural networks and Gaussian process regression. *Algorithms* 2025;18(1):10.
- [56] Moradi R, Berangi R, Minaei B. A survey of regularization strategies for deep models. *Artif Intell Rev* 2020;53(6):3947–86.
- [57] Fiyadh SS, et al. Artificial neural network approach for modelling of Mercury ions removal from water using functionalized CNTs with deep eutectic solvent. *Int J Mol Sci* 2019;20(17):4206.
- [58] Lee YG, et al. Artificial neural network model for optimizing operation of a seawater reverse osmosis desalination plant. *Desalination* 2009;247(1–3):180–9.
- [59] Abuwatfa WH, et al. A review on membrane fouling prediction using artificial neural networks (ANNs). *Membranes* 2023;13(7):685.
- [60] Nagpal M, et al. Optimizing wastewater treatment through artificial intelligence: recent advances and future prospects. *Water Sci Technol* 2024;90(3):731–57.
- [61] Yang Y, Wang C, Wang S, Xiao Y, Ma Q, Tian X, Li J. Performance prediction model for desalination plants using modified grey wolf optimizer based artificial neural network approach. *Desalin Water Treat* 2024;319:100411.
- [62] Chandru IA, Dhinakaran G. Modelling reverse osmosis process and multiple effect evaporator process of common effluent treatment plant using artificial neural network. *Desalin Water Treat* 2024;319:100455.