Applied Water Science (2025) 15:110
https://doi.org/10.1007/s13201-025-02452-y

ORIGINAL ARTICLE

Automated interpretation of deep learning-based water quality assessment system for enhanced environmental management decisions

Javed Mallick[1] · Saeed Alqadhi[1] · Majed Alsubih[1] · Mohamed Fatahalla Mohamed Ahmed[2] · Hazem Ghassan Abdo[3]

Received: 2 February 2024 / Accepted: 21 March 2025

Abstract
Water quality assessment is a critical issue in the Aseer region of Saudi Arabia, where environmental and anthropogenic factors pose a major challenge to both drinking water and irrigation systems. The aim of this study was to carry out a detailed assessment of the water resources in the region, focussing on the most important aspects affecting water quality. The main objectives were to calculate various water quality indices for drinking and irrigation purposes, to develop an automated system using convolutional neural networks (CNN) to predict these indices and to increase the transparency of these models using explainable artificial intelligence (XAI) methods. Methodologically, the study used CNN algorithms optimised by Bayesian techniques for the prediction of eight water quality indices, coupled with SHAPley Additive exPlanations (SHAP) analysis under XAI to interpret the complex decision-making processes of these models. This dual approach enabled a comprehensive and insightful assessment of water quality. Using a robust dataset from the Aseer region, eight water quality indices were calculated, revealing significant variations and highlighting areas of concern. In this study, the entropy weight-based DWQI averaged 77.90 with a high standard deviation (std) of 39.08, reflecting considerable variability. The automated CNN models demonstrated robust performance in predicting water quality indices, with high accuracy ($R^2 = 0.959$ in training and 0.945 in testing) for sodium percentage (Na%). However, the Magnesium Hazard (MH) index showed lower accuracy, suggesting possible overfitting and the need for further optimisation. SHAP analysis highlighted chloride, sulphate, and total dissolved solids as key contributors to the WQI, while sodium and calcium were significant for the sodium adsorption ratio. These insights enhance understanding of parameter influence on water quality assessments. This study introduces an advanced computational approach integrating CNN and XAI techniques, improving water quality evaluation and supporting informed environmental management in the Aseer region.

Keywords Water quality assessment · Drinking water quality index (DWQI) · Irrigation water quality · Convolutional neural networks (CNN) · Explainable artificial intelligence (XAI) · Water resource management

✉ Hazem Ghassan Abdo
Hazemabdo@tartous-univ.edu.sy

Javed Mallick
jmallick@kku.edu.sa

Saeed Alqadhi
sdalqadi@kku.edu.sa

Majed Alsubih
malsubih@kku.edu.sa

Mohamed Fatahalla Mohamed Ahmed
dr.nitaifa@gmail.com

[1] Department of Civil Engineering, College of Engineering, King Khalid University, P.O. Box: 394, 61411 Abha, Kingdom of Saudi Arabia

[2] Department of Environment and Ecology, Faculty of Geographical and Environmental Sciences, University of Khartoum, Khartoum, Sudan

[3] Geography Department, Faculty of Arts and Humanities, Tartous University, Tartous, Syria

Published online: 29 April 2025

Introduction

Water resources play a crucial role in domestic, industrial, and agricultural activities (Busico et al. 2020). High-quality water not only reduces treatment costs but also enhances agricultural productivity (Mbjzvo et al. 2019). However, increasing demand due to population growth, urbanisation, and industrialisation is exerting immense pressure on water resources (United Nations Environment Programme (UNEP) 2021). This challenge is particularly severe in arid and semi-arid regions like Saudi Arabia, where water scarcity is persistent (El Bilali and Taleb 2020). Both manmade and geogenic sources of contamination contribute to water quality degradation, necessitating effective management and monitoring (UNEP 2020). In these regions, overexploitation of groundwater leads to declining water levels and deterioration of quality (Asante-Annor et al. 2018). Contaminated water poses severe health risks, including waterborne diseases and other serious conditions (Salifu et al. 2017; Aghazadesh and Mogaddam 2010). Effective water resource management requires a structured approach involving continuous monitoring from source to consumption (Ahmed et al. 2019). Regular sampling and assessment are essential to evaluate water suitability for human consumption, agriculture, and industry (Gupta and Gupta 2020). Sustainable water management is crucial for public health, economic development, and environmental conservation, especially in water-scarce regions (Iqbal et al. 2020).

Water quality monitoring is often hindered by high costs, labour-intensive procedures, and limited access to advanced laboratories (Uddin et al. 2022a, b). To simplify assessment, the water quality index (WQI) was developed, which aggregates multiple water quality parameters into a single interpretable score (Uddin et al. 2022a). This facilitates informed decision-making on water usability (Sakizadeh 2016). The WQI classifies water quality into categories such as good, fair, marginal, or poor (Uddin et al. 2022a, b). For irrigation, various indices assess water suitability, including sodium adsorption ratio (SAR), Kelly's ratio (KR), residual sodium carbonate (RSC), and magnesium adsorption ratio (MAR) (Islam and Mostafa 2022a, b, c; Elsayed et al. 2020; Jahin et al. 2020; Batarseh et al. 2021). While these indices effectively determine irrigation water quality, they require extensive sampling, lab analysis, and data handling, making them time-consuming and costly (Uddin et al. 2022a, b). Efficient, automated assessment methods are needed for sustainable water resource management (Kouadri et al. 2021).

Recent advancements in machine learning (ML) offer promising solutions for predicting and analysing water quality parameters (Khoi et al. 2022; Kouadri et al. 2021). ML techniques such as artificial neural networks (ANNs), fuzzy logic, neuro-fuzzy models, and decision trees have improved prediction accuracy for surface and groundwater quality (Lap et al. 2023; Uddin et al. 2021; Ahmed et al. 2021). These models process complex datasets, identify patterns, and provide predictive insights (Ram et al. 2021; Krishnan et al. 2023). Despite these advancements, conventional ML models face limitations in handling highly complex water quality data. Deep learning (DL), a subset of ML, has demonstrated superior performance in extracting intricate patterns from large datasets (Talukdar et al. 2023). DL has been successfully applied in various fields, including image classification, speech recognition, and natural language processing (Nava et al. 2023; Zeng et al. 2023).

Among DL architectures, convolutional neural networks (CNNs) are particularly effective for recognising subtle patterns in diverse datasets, making them suitable for water quality prediction (Talukdar et al. 2023; Chen and Fan 2023). However, CNN models often function as "black boxes," making their decision-making process opaque (Talukdar et al. 2023, 2024). This lack of transparency limits their applicability in environmental decision-making, where interpretability is crucial (Mia et al. 2023). To enhance model transparency, explainable artificial intelligence (XAI) techniques have been introduced, including Local Interpretable Model-Agnostic Explanations (LIME) and SHAPley Additive exPlanations (SHAP) (Talukdar et al. 2024). LIME approximates a complex DL model by generating an interpretable, locally faithful representation to explain individual predictions, while SHAP provides a global perspective by decomposing predictions into individual feature contributions, offering insights into parameter influence on model outcomes (Das et al. 2023; Sarkar et al. 2023). These techniques enhance model interpretability, enabling stakeholders to make informed decisions in water resource management (Ahmed et al. 2024).

Existing research has predominantly focused on drinking water quality, often neglecting irrigation water quality, which is equally critical. A comprehensive approach that integrates both drinking and irrigation water indices is needed. Additionally, current assessment methods are costly and time-consuming, requiring repetitive, labour-intensive procedures. Developing an automated decision support system that learns intricate relationships between water quality parameters and indices can address these challenges.

This study aims to:

1. calculate water quality indices for drinking and irrigation water to enable a comprehensive assessment.
2. develop an automated, optimised system using DL algorithms to predict water quality indices.

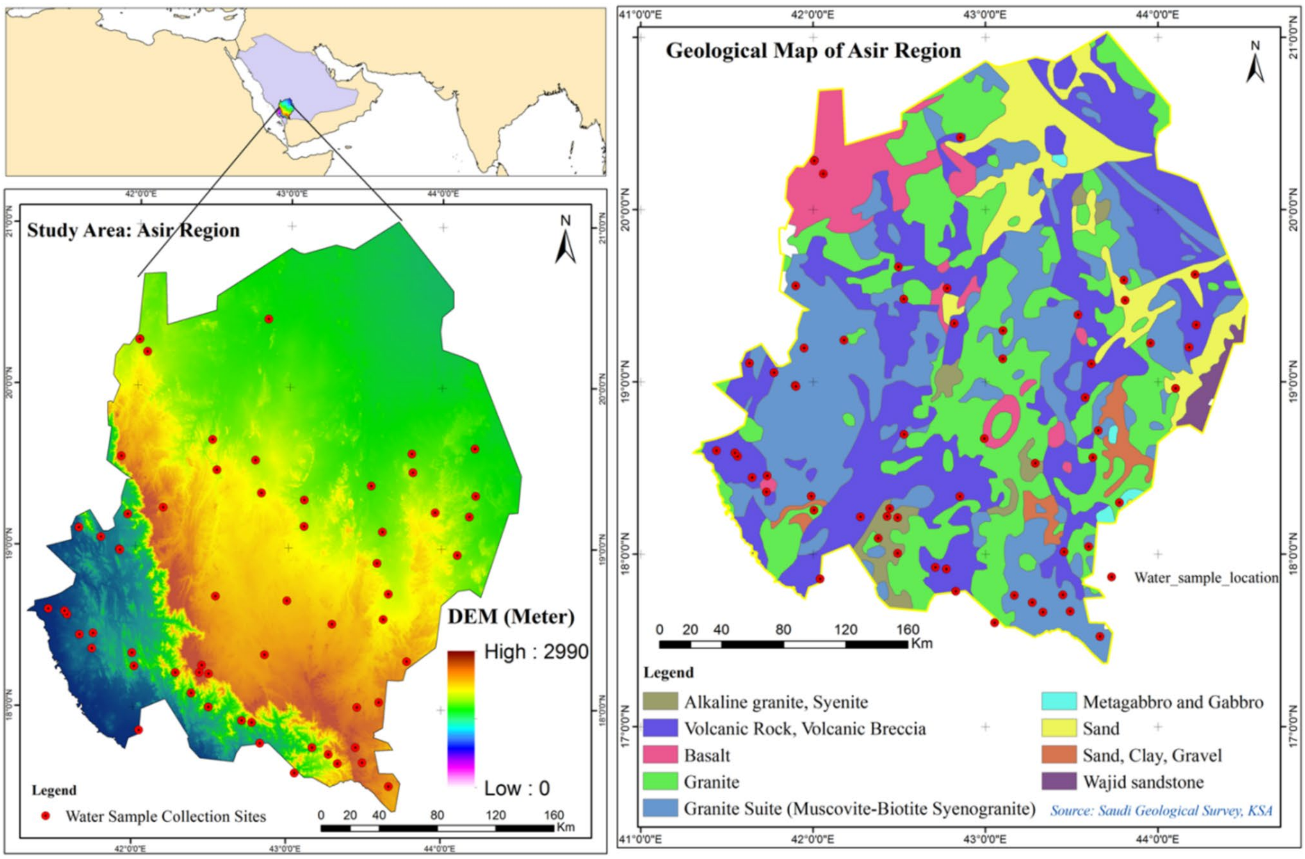

Fig. 1 Study area and geological map

3. enhance model transparency by applying XAI techniques to interpret the DL model's predictions.

By integrating CNN with XAI, this study advances water quality assessment, model interpretability, and decision-making transparency, contributing to effective water resource management.

Materials and methods

Study area

The Asir Province, situated in Saudi Arabia, falls within the coordinates of 18°12′029.355″N to 18°12′051.436″ N latitude and 42°29′05.157″ E to 42°29′019.795″ E longitude, encompassing an area of 84,250 square kilometres (as indicated in Fig. 1b). Asir exhibits a diverse range of climatic conditions, including hot desert, cold desert, cold semi-arid, and hot semi-arid zones. It receives an annual precipitation of 350 mm, making it a significant region for agriculture. In terms of geological composition, Asir comprises aquifers such as quaternary alluvium, quartz sandstone, and conglomerates, with secondary aquifers primarily composed of calcareous deposits undergoing lateral diagenetic modifications. These aquifers display notable porosity and karstification, as discussed by Mallick et al. (2021). Figure 1B provides a geological map of the Asir region. An essential source of groundwater in the area is the unconfined quaternary alluvial aquifers, which receive replenishment from runoff originating from the Asir highlands. These shallow aquifers possess an estimated annual recharge volume of 1196×106 m^3 and exhibit varying water quality, ranging from poor to good, as documented by Dabbagh and Abderrahman (1997). The presence of a 100-m-thick layer of alluvial fill contributes to the high-quality water discovered in Wadi-al-Dawassir.

The Asir region is rapidly enhancing its rainwater harvesting initiatives by building check dams. These dams facilitate the accumulation of a substantial amount of water, which can be used to cultivate 15,000 hectares of agricultural land. Currently, if only a quarter of the runoff water that is presently wasted could be efficiently captured, it would meet all of Saudi Arabia's existing agricultural water needs. The majority of Saudi Arabia's runoff occurs along the escarpment in the Asir region, where wadis flow towards the coastal area, contributing approximately 60% of the nation's total runoff. Most wadi structures consist of sand and gravel, and after traveling a short distance, the runoff infiltrates subsurface water bodies (wadis), creating a sub-flow that recharges the groundwater. Storm runoff can happen in the Asir region at any time during the year. Saudi Arabia has a long-standing tradition of constructing dams, particularly in the Hijaz and Asir regions. As of 2018, as reported by the Saudi Arabia Ministry of Environment, Water, and Agriculture, there are 509 dams spread across the kingdom, with 117 of them situated in the Asir region. The primary objective of building these dams is to capture runoff and replenish the groundwater system, although some dams also serve as sources of drinking water and direct irrigation for agriculture.

Data processing

A total of 62 water samples were collected from different wells within the study area, as shown in Fig. 1. To ensure data reliability and consistency, all samples underwent a rigorous quality control process before analysis. No missing values were present in the dataset, as each sample was thoroughly validated before inclusion in the study. To identify outliers and noisy data, a histogram with probability distribution analysis was used, ensuring that anomalous values did not skew model predictions. To evaluate water quality, various parameters, including electrical conductivity (EC), pH, and total dissolved solids (TDS), were measured in the field using HANNA handheld sensors. Prior to each sampling session, sensors were calibrated daily using standard solutions with pH values of 4.0, 7.0, and 10.0, as well as EC standards at 84 µS/cm, 1413 µS/cm, and 12.8 mS/cm. For each sampling location, two water samples were collected in high-density polyethylene (HDPE) bottles. One sample was acidified in the field with a 1:1 nitric acid solution for cation analysis, while the second remained unaltered for anion analysis. Cation concentrations, including calcium (Ca^{2+}), sodium (Na^{+}), magnesium (Mg^{2+}), potassium (K^{+}), and iron (Fe), were measured using an atomic absorption spectrophotometer (Thermo Scientific M series). Anion concentrations, including chloride (Cl^{-}), fluoride (F^{-}), nitrate (NO_3^{-}), and sulphate (SO_4^{2-}), were analysed using an ion chromatograph (Dionex) in gradient mode. Additionally, bicarbonate (HCO_3^{-}) was determined using a titrimetric method, and total alkalinity and hardness were assessed following APHA 1995 standard procedures. All reagents, standards, and chemicals used in these analyses were analytical grade and procured from Merck.

Method for drinking water quality index (DWQI)

There are several methods for developing a drinking water quality index (DWQI), including the weighted arithmetic water quality index (WAWQI), Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI), National Sanitation Foundation Water Quality Index (NSFWQI), and entropy weight water quality index (EWQI). Each method has its advantages, but we selected the entropy-based DWQI because it provides an objective, data-driven approach that does not rely on subjective weight assignments (Han et al. 2023). Unlike traditional methods that use predetermined weights, entropy weighting quantifies the degree of randomness or disorder in the information provided by each parameter, ensuring that more influential parameters receive higher weights (Han et al. 2023; Alfaleh et al. 2023). This enhances the accuracy and reliability of water quality assessment, making it particularly useful for regions where water quality varies significantly due to natural and anthropogenic factors.

The method for developing an entropy-based DWQI involves a systematic approach to assessing overall drinking water quality using specified parameters. The physical parameters include pH, electrical conductivity (EC), total dissolved solids (TDS), and turbidity. The physicochemical parameters include ammonia, nitrite, nitrate, chloride, sulphate, total hardness, total calcium, magnesium, iron, fluoride, alkalinity, and residual chlorine. Table 1 presents the physical and physicochemical parameters used in the DWQI, along with their permissible limits as per the World Health Organization (WHO) and other regulatory agencies. The process begins with the collection and normalisation of water quality data for each parameter to ensure comparability across different units and scales (Gupta and Mishra 2023). The entropy weighting method is then applied, which quantifies the degree of disorder in the information provided by each parameter (Alfaleh et al. 2023). The entropy value for each parameter is calculated to determine its weight, reflecting its relative importance in water quality assessment (Han et al. 2023). Parameters with lower entropy values have greater influence on water quality and are assigned higher weights. The final DWQI score is computed by aggregating these weighted parameters (Talukdar et al. 2024). Each parameter’s contribution is multiplied by its respective weight, and the sum of these products gives the final index. The DWQI is a quantitative measure that allows easy interpretation and comparison of water quality, typically ranging from 0 to 100, with higher values indicating better quality (Talukdar et al. 2024). The entropy-based DWQI offers a robust, objective, and comprehensive assessment of water quality, ensuring a more reliable evaluation compared to traditional approaches (Zhe et al. 2021; Verma et al. 2022; Han et al. 2023).

Table 1 Drinking water quality parameters and standard limits

Parameter	Type	Standard Limit	Unit	Source
pH	Physical	6.5–8.5	-	WHO (2017)
Electrical conductivity (EC)	Physical	200–300	μS/cm	WHO (2017)
Total dissolved solids (TDS)	Physical	500	mg/L	WHO (2017)
Turbidity	Physical	5	NTU	WHO (2017)
Ammonia	Physicochemical	1.5	mg/L	WHO (2017)
Nitrite	Physicochemical	0.2	mg/L	WHO (2017)
Nitrate	Physicochemical	50	mg/L	WHO (2017)
Chloride	Physicochemical	250	mg/L	EPA (2023)
Sulphate	Physicochemical	250	mg/L	EPA (2023)
Total hardness (as $CaCO_3$)	Physicochemical	200	mg/L	EPA (2023)
Calcium	Physicochemical	75	mg/L	EPA (2023)
Magnesium	Physicochemical	50	mg/L	EPA (2023)
Iron	Physicochemical	0.3	mg/L	EPA (2023)
Fluoride	Physicochemical	1.5	mg/L	WHO (2017)
Alkalinity (as $CaCO_3$)	Physicochemical	200	mg/L	CDC (2022)
Residual chlorine	Physicochemical	5	mg/L	WHO (2017)

Methods for irrigation water quality index (IWQI)

Method for developing automated Bayesian optimised Convolutional Neural network algorithm-based system to predict DWQI and IWQIs

The development of an automated system based on a Bayesian optimised CNN algorithm for the prediction of DWQI and IWQI requires a sophisticated and methodical approach. The system integrates the principles of convolutional neural networks and Bayesian optimisation to create a model tailored for the prediction of water quality indices based on different water parameters.

Convolutional neural networks (CNNs) CNNs in the field of water quality index prediction are designed to process and interpret data characterised by multiple input variables as they occur in water quality assessment (e.g. sodium absorption ratio, magnesium hazard) (Talukdar et al. 2023). The architecture of a CNN is particularly well suited to this task due to its unique composition. It is specifically tailored to recognise and learn complex patterns and relationships in multi-dimensional datasets (Yang et al. 2021). At the core of a CNN are convolutional layers that apply a series of learnable filters to the input data. These filters are designed to automatically and adaptively extract relevant features from the data, such as recognising specific chemical concentrations or relationships between different water quality parameters (Sheikh Khozani et al. 2022). The convolution process creates feature maps that represent these recognised features (Mei et al. 2022). The convolutional layers are followed by pooling layers, which are used to reduce the spatial dimensions (i.e. width and height, but not depth) of the feature maps (Talukdar et al. 2023). This reduction is critical for reducing the computational load and mitigating the risk of overfitting by abstracting the features to a more manageable level (Tan et al. 2022). Pooling (often implemented as max-pooling) simplifies the information in the output of the convolutional layers by summarising their main features and retaining only the most important information. Finally, the network culminates in one or more dense layers, which are themselves fully connected neural networks. These layers integrate the high-level features extracted from the convolutional layers and the pooling layers to fulfil the final task of predicting water quality indices. The dense layers work on the principle of recognising complex patterns formed from the combinations of features extracted in the earlier layers. Through training, the network learns to assign different importance to different features and thus recognise the complicated relationships and dependencies between the different water quality parameters. Therefore, CNNs with their hierarchical structure of convolutional layers, pooling layers and dense layers are exceptionally good at analysing complex, multivariate datasets as they occur in water quality monitoring (Tan et al. 2022; Talukdar et al. 2023). They are excellent at revealing underlying patterns and dependencies between different parameters, making them a powerful tool for predicting water quality indices with high accuracy and reliability (Mei et al. 2022).

Bayesian optimisation Bayesian optimisation plays a crucial role in fine-tuning the CNN model. It is an approach for optimising model hyperparameters, such as the number of filters in convolutional layers, the size of dense layers, the learning rate, batch size and epochs. In Bayesian optimisation, a probabilistic model of the function is created that maps the values of the hyperparameters to the target, which is evaluated against a validation set. In this case, the objective is to minimise the mean square error (MSE) between the predicted and actual water quality indices. Bayesian optimisation iteratively selects the hyperparameters to be evaluated by striking a balance between exploration (trying out new hyperparameters) and exploitation (using hyperparameters that have worked well in the past).

Implementation The implementation of an automated Bayesian-optimised CNN for predicting DWQI and IWQI starts with a careful pre-processing phase. In this phase, the dataset, which contains a variety of water quality parameters, is scaled and normalised to ensure uniformity and compatibility with the CNN architecture. The data are then remodelled to meet the input requirements of the CNN. This usually involves converting the dataset into a format suitable for convolution operations. After data preparation, the Bayesian optimisation process is initiated. This sophisticated algorithm systematically navigates through the hyperparameter space of the CNN and uses a probabilistic model to guide the search for the optimal set of hyperparameters. This model balances exploration and exploitation strategies and efficiently searches for the hyperparameter combination that minimises the prediction error, which is usually assessed during the validation process with MSE. After identifying the optimal hyperparameters, the final phase is to train the CNN model with these parameters over the entire dataset. This training is an iterative process in which the model learns to recognise and extract complicated patterns and relationships between the water quality parameters to improve its prediction accuracy. After training, the efficiency of the model is rigorously evaluated using new, unseen data. Key performance metrics such as root mean square error (RMSE), mean absolute error (MAE) and R-squared values (R^2) are used to assess the precision and reliability of the model in predicting the water quality indices. RMSE and MAE provide information on the average size of the prediction error, while R^2 is a measure of the variance of the water quality indices that the model can explain. Through this comprehensive process, the system not only ensures accurate predictions of water quality indices, but also embodies a robust, data-driven approach to advanced water quality analyses in different environmental and agricultural contexts.

To evaluate the computational efficiency of the CNN model, training and inference times were recorded. The models were trained on a MSI Alpha 15 A3DD laptop with an AMD Ryzen 7 3750H processor, 16 GB RAM, and Radeon Vega Mobile Graphics. The training process, which involved Bayesian optimisation over 50 iterations, required approximately 1 h and 45 min per model. To enhance computational efficiency, batch normalisation and parallel processing were implemented, allowing for faster convergence.

Interpretation of automated DL model-based system using automated explainable artificial intelligence algorithm for better decision-making

The interpretation of an automated, DL-model-based system for predicting water quality indices using automated explainable artificial intelligence algorithms, such as SHAP (SHAPley Additive exPlanations), represents a significant advance in the field of environmental data science (Talukdar et al. 2024). SHAP, a game-theoretic approach to explaining the output of any machine learning model, provides insight into how each feature in the dataset contributes to the prediction of the model (Talukdar et al. 2023). This methodology is particularly important for complex models such as DL, where the decision-making process is often opaque and difficult to interpret (Ahmed et al. 2024).

In this context, SHAP values are calculated for each feature in the dataset, quantifying the influence of each feature on the outcome of the model (Talukdar et al. 2024). The trained DL model and the reshaped training data are used to calculate the SHAP values to ensure consistency with the data format used in model training. These values provide a detailed breakdown of the contribution of each water quality parameter (such as ammonia, nitrite, nitrate, and TDS) to the predicted water quality index. The higher the SHAP value of a characteristic, the greater its influence on the prediction of the model for a particular case. The visualisation tools provided by SHAP, such as summary plots, offer an intuitive representation of these contributions (Das et al. 2023). They allow researchers and decision makers to recognise which features are most influential in predicting water quality indices and how different values of these features change the prediction. For example, a summarised representation can show whether higher values of parameters such as total dissolved solids (TDS) or chloride drive the prediction of a poor water quality index, providing actionable insights. Integrating SHAP into the modelling process not only increases the transparency and interpretability of DL models, but also supports better decision-making in water quality management. By understanding the driving factors behind the predictions, stakeholders can develop more targeted strategies for water treatment, resource allocation and environmental policy. This interpretive approach thus bridges the gap between complex DL modelling results and practical, data-driven decision-making in water resource management.

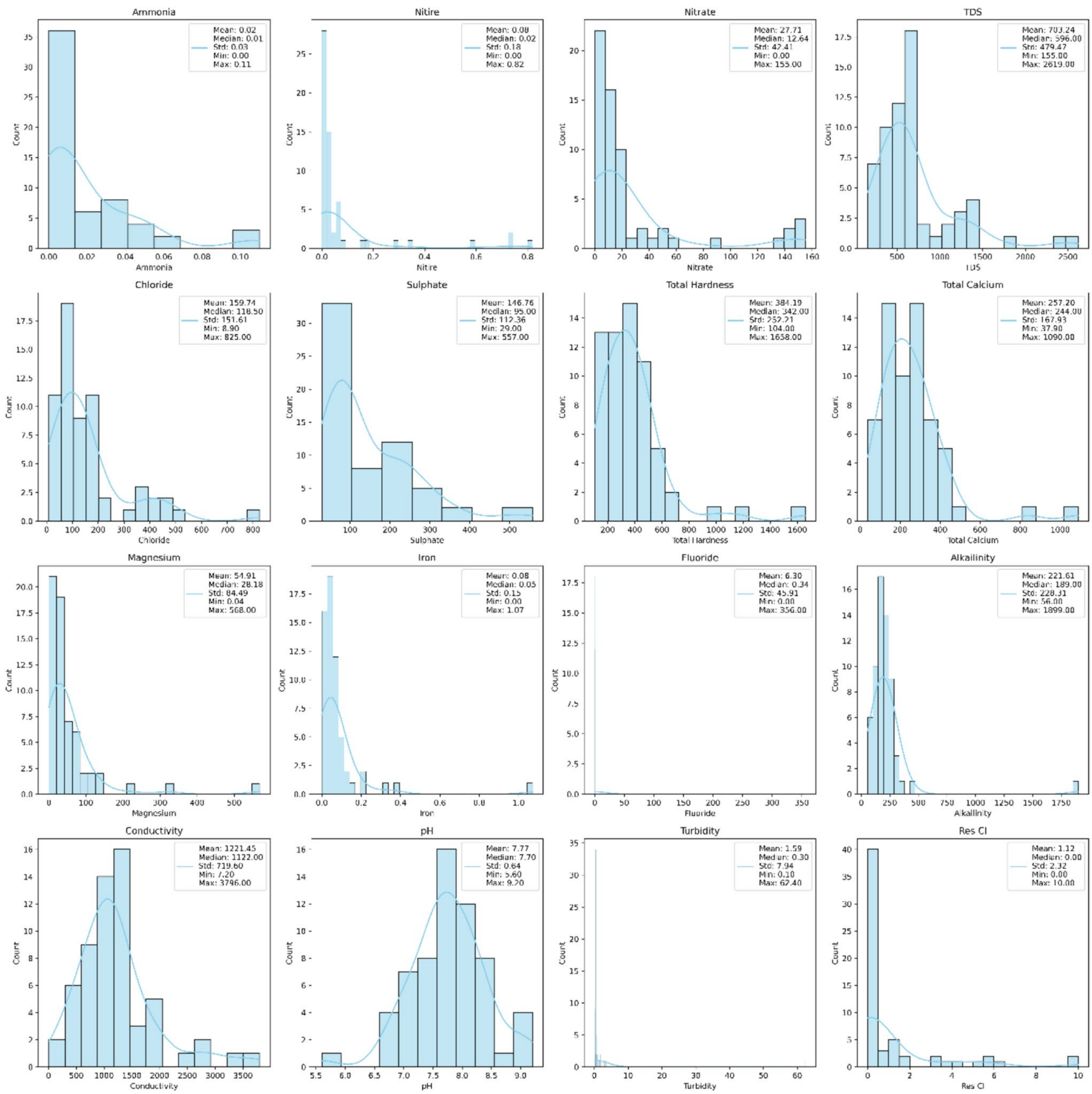

Fig. 2 Histograms depicting the distribution of various water quality parameters used in assessing DWQI and IWQI, showcasing mean values, standard deviations, and range of concentrations

Results

Statistical characteristics of water quality parameters

The statistical analysis of water quality parameters indicates various effects on water pollution, both in terms of drinking water and suitability for irrigation. Figure 2 shows a series of histograms, each corresponding to a different water quality parameter and providing a visual representation of their distribution across a dataset. The ammonia, nitrite and nitrate parameters show a right skewed distribution, indicating a concentration of lower values with fewer high outliers, which could indicate point source pollution from higher concentrations. TDS, chloride and sulphate concentrations show a wider spread, with TDS showing a particularly wide range, that could indicate different geological sources or industrial discharges. Total hardness and total calcium also show considerable variation, with some values ranging into regions that could lead to deposits in pipework and potentially affect soil permeability during irrigation. Magnesium and iron concentrations show mostly low to moderate values, but also contain some extreme values that could have a negative impact on water utilisation. Fluoride shows a predominantly low range with some extreme values that indicate isolated contamination. Alkalinity and conductivity show a relatively central tendency, indicating consistent water properties possibly due to stable geological or environmental conditions. The pH values hover around neutrality, but show a dispersion that could affect metal solubility and microbial activity. Turbidity and residual chlorine are mostly low, which is good for drinking water clarity and control of disinfection by-products, but their higher values require attention to avoid potential health risks and irrigation problems.

The ammonia values with a mean value of 0.02 mg/L and a relatively low standard deviation (SD) of 0.03 mg/L indicate minimal contamination, which is favourable for both drinking and irrigation purposes. Nitrite and nitrate, with mean values of 0.08 mg/L and 27.71 mg/L, respectively, are of concern when compared to drinking water standards where nitrate levels should not exceed 10 mg/L due to the risk of methaemoglobinaemia or'blue baby syndrome'TDS, with a mean value of 703.24 mg/L, is within the acceptable limit for drinking water but could affect sensitive crops if used for irrigation. Chloride and sulphate concentrations are within safe drinking water limits with a mean of 159.74 mg/L and 146.76 mg/L, respectively, but may require careful management when used for irrigation to avoid problems with soil salinity and toxicity. Total hardness and total calcium have high mean values of 384.19 mg/L and 257.20 mg/L, respectively, which can lead to deposits in pipes and reduced soap effectiveness, but are generally not harmful to human health. In irrigation, high hardness could affect soil structure and plant growth. The mean magnesium concentration of 54.91 mg/L indicates possible adverse effects on soil physical properties and plant nutrient uptake during irrigation. Iron and fluoride with mean values of 0.08 mg/L and 6.30 mg/L could pose risks; iron affects the taste and discolouration of water, and fluoride levels are well above the recommended level of 1.50 mg/L for drinking water, posing a risk of dental fluorosis. Alkalinity, conductivity, pH, turbidity and residual chlorine are within the typical ranges for drinking water with mean values of 221.61 mg/L, 1221.45 µS/cm, 7.77, 1.59 NTU and 1.12 mg/L, respectively, but their higher values may affect the suitability of the water for irrigation and potentially cause damage to soil structure and nutrient uptake problems for plants.

Computation of DWQI and IWQIs and their statistical characteristics

The scientific analysis of DWQI and IWQI for 62 samples in the Aseer region of Saudi Arabia shows remarkable quantitative characteristics (Fig. 3). For good irrigation quality, we look for a low SAR value, which indicating a lower risk of sodium-related soil permeability problems, a higher PI value, which indicates better soil permeability, a lower MH value, reducing the risk of magnesium toxicity, a lower KR value, which indicating a lower sodium content, and a lower RSC value, reducing the risk of soil alkalisation. The DWQI, an important indicator of potability, shows a mean value of 77.90 with a significant standard deviation (std) of 39.08. This large standard deviation indicates significant variability in water quality between samples. The minimum DWQI value is 22.56, indicating that some samples are of relatively high quality, while the maximum value of 222.99 indicates extremely poor water quality in other samples. Considering that a higher DWQI value means poorer water quality, these values indicate a wide range of drinkability concerns in the region.

The overall indication of IWQIs in the Aseer region of Saudi Arabia, as derived from the data provided, shows a multi-layered picture of water quality and its impact on irrigation. These indices, including SAR, PI, MH, KR and RSC, collectively provide insight into the suitability of the water for agricultural purposes. Let's start with the SAR value, which has a mean value of 0.48 and a range of 0.06 to 1.33. It is an important measure of the ratio of sodium to calcium and magnesium in the water. High SAR values can lead to dispersion of the soil and reduce its permeability and aeration. This can have a negative impact on seed germination and reduce crop yields. In the Aseer region, the variation in SAR indicates that while some areas have suitable water for irrigation, others may be at risk of soil degradation due to high sodium content. The PI, with a mean of 25.89 and a range of 10.88 to 50.76, is an indicator of the impact of water on soil permeability. Water with a low PI can cause soil compaction, reducing the soil's ability to allow air and water to pass through, which is essential for healthy plant growth. The wide range of PI values in the samples suggests that irrigation water could lead to soil compaction in some parts of the region, thereby affecting plant growth. MH and KR) are other important indices. MH, with a mean value of 18.55 and a range of 1.70 to 48.57, reflects the concentration of magnesium in the water. A high magnesium content can change the soil properties and make it more difficult for plants to absorb water and nutrients. This change can lead to lower crop yields and, in severe cases, even to crop failure. Kelly's ratio, with a mean value of 0.10 and a range of 0.01 to 0.28, is another measure of

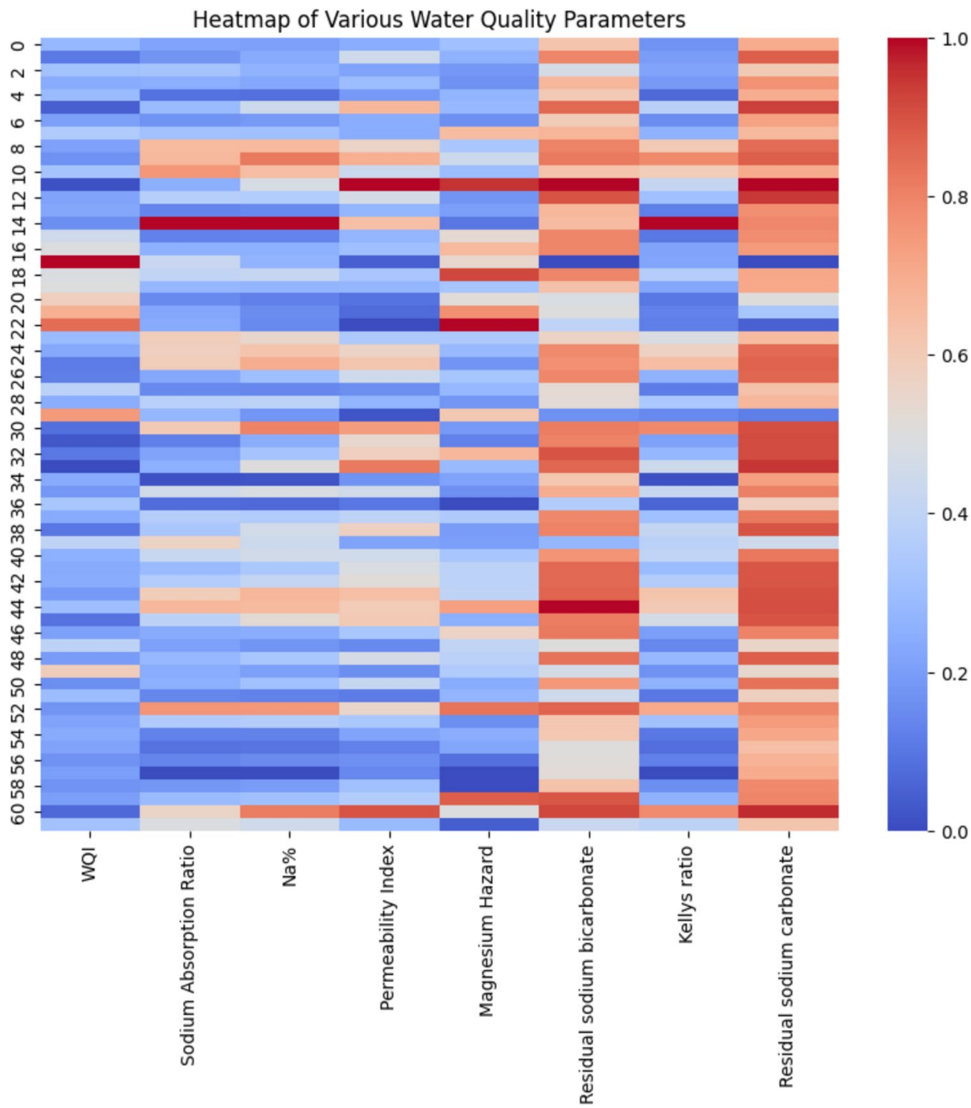

Fig. 3 Heatmap visualisation of normalised DWQI and IWQIs from the Aseer region, highlighting the variability and distribution of values that inform water suitability for drinking and agriculture

the sodium content in water. A higher KR value indicates an excess of sodium, which can lead to alkalisation of the soil, affecting soil structure and fertility. The variations in the LR values in the samples in the Aseer region indicate that some areas may have problems with soil alkalisation due to irrigation practises. Finally, the RSC value, with a mean of −8.54 and a range of −30.98 to −0.28, indicates the concentration of carbonate and bicarbonate in relation to calcium and magnesium. High RSC values can lead to the precipitation of calcium and magnesium and increase the sodium concentration in the soil. This can lead to alkalisation and sodification of the soil, making it hard and impermeable, which is detrimental to plant growth. The negative mean value of RSC across the region is a positive sign and indicating a generally low risk of these problems, although the range of variation suggests that some areas may still be at risk.

Based on these criteria, regions such as Alaziziyah-1, Radoom, Alakaas, Alheema and Sadd Abha have favourable conditions. They have relatively low SAR, moderate to high PI and lower MH, KR and RSC. On the other hand, regions with poor irrigation quality typically have high SAR, low PI, high MH, high KR and high RSC. Places such as Wadi Hayat, Alhajjaj, Alnasa, Khayoor and Alyanood have these characteristics. They have high SAR values, which can lead to softening of the soil and reduced permeability, higher MH values, which indicate possible magnesium toxicity, and higher KR and RSC values, which can lead to alkalisation of the soil.

Relationship between water quality parameters and DWQI and IWQI

Figure 4 shows the correlation coefficients between various water quality parameters and both the DWQI and the

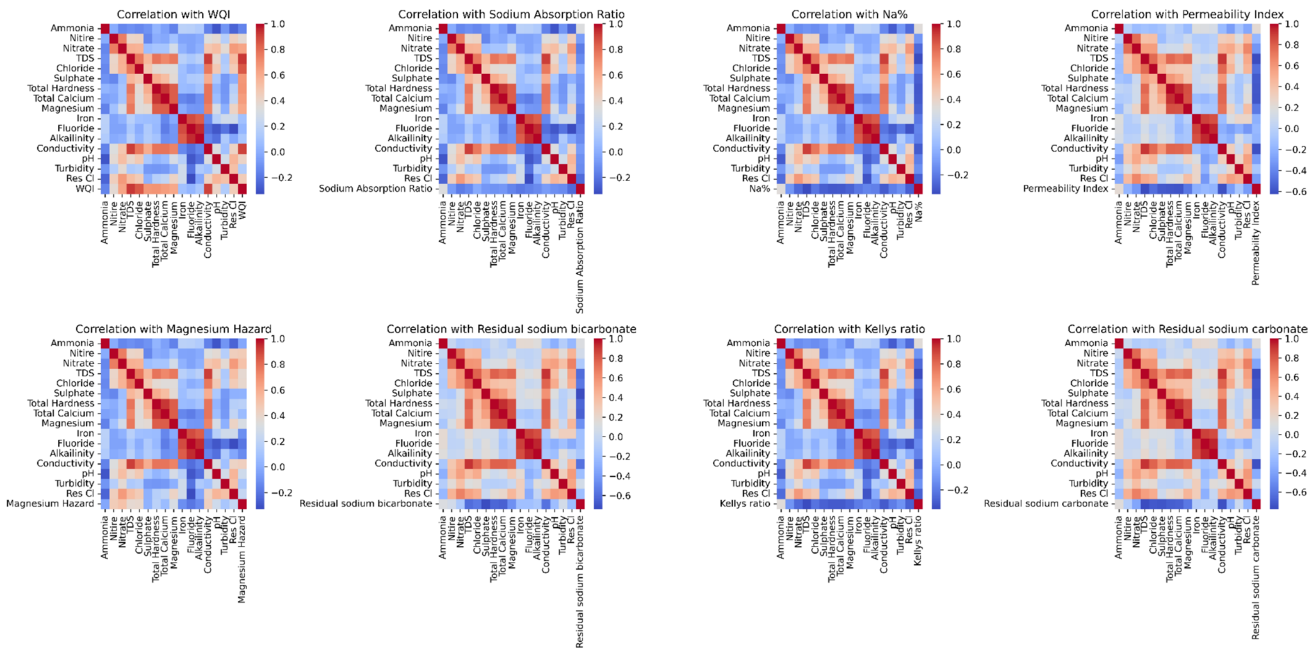

Fig. 4 Heatmap visualisations depicting the correlation coefficients between various water quality parameters and the DWQI as well as multiple IWQIs. Each panel represents a matrix of correlation values ranging from − 1 (perfect negative correlation, shown in blue) to + 1 (perfect positive correlation, shown in red), with the intensity of the colour indicating the strength of the correlation

various IWQIs. These coefficients range from − 1 to 1, where 1 stands for a perfect positive correlation, − 1 for a perfect negative correlation and 0 for no correlation. For the DWQI, which reflects the drinkability of the water, there are several notable positive correlations. Parameters such as TDS, chloride, sulphate, total hardness and conductivity show strong positive correlations with the DWQI, suggesting that higher concentrations of these parameters can lead to poorer drinking water quality. A high positive correlation with these parameters is an indication of potential health risks and aesthetic problems, such as water hardness and salinity, which can affect taste and drinkability. Looking at the IWQIs, SAR shows a significant positive correlation with sodium content (Na%), which is to be expected as SAR is a direct measure of the sodium hazard in irrigation water. High SAR values can lead to dispersion in the soil, which has a negative effect on soil structure and plant growth. Other indices, such as MH, show strong positive correlations with parameters such as magnesium and total hardness. This suggests that higher levels of magnesium and calcium in water contribute to magnesium vulnerability and may have an impact on soil quality and crop yield. PI shows a positive correlation with TDS and sulphate, indicating that higher concentrations of solutes can affect soil permeability and consequently the movement of water and air through the soil, which is critical for crop health. Conversely, KR and RSC have positive correlations with parameters such as sodium and bicarbonate, which at high concentrations can lead to soil alkalisation and sodicity, affecting the soil's ability to support plant life. On the negative side of the correlations, parameters such as pH and fluoride show inverse relationships with indices such as SAR and RSC for some IWQIs. This could indicate that the risk to irrigation quality decreases in some aspects as water becomes more alkaline or fluoride levels increase, although the specific impact depend on the overall balance of all water quality parameters. Therefore, the correlation heat maps provide an integrated overview of how individual water quality parameters can collectively affect the overall quality of water for drinking and irrigation purposes. The strength and direction of these correlations are critical to understanding the potential impacts on human health and agricultural productivity. Management strategies must take these correlations into account to ensure both safe drinking water and effective irrigation practises.

Development of optimised DL algorithm-based automated system for predicting DWQI and IWQIs

This study focuses on the creation and refinement of an automated system that uses DL algorithms for the prediction of DWQI and IWQI. It comprises several steps: First, the design of this DL-based system is described, detailing the architecture and initial configuration of the neural network models to process water quality data for prediction purposes.

The DL system is then optimised using Bayesian optimisation, a strategy for selecting the most effective hyperparameters to improve model performance and prediction accuracy. Following the optimisation, we explain the implementation of the DL system with these optimal hyperparameters to ensure that the model performs with the highest efficiency. Finally, we evaluate the performance of the system, which is likely to include the assessment of prediction accuracy, robustness and reliability against a set of test data and a comparison with existing prediction models.

Construction of automated DL-based system

The automated deep learning system was developed to predict DWQI and various IWQIs using a CNN, which is known for its ability to recognise patterns and relationships in multi-dimensional data. Automation of the system is achieved through a Python-based code that iteratively trains separate models for each water quality index. The data are first standardised using a standard scaler to ensure that the model inputs have a mean of zero and a standard deviation of one—a crucial step for the effective training of neural networks. The data are then reshaped to fit the requirements of the CNN, with each feature converted to a single input channel, as CNNs are designed to work with multi-dimensional data. The quantification of the hyperparameters is critical to the performance of the model and is approached using Bayesian optimisation. This is a strategy in which the parameter space is searched sequentially using Bayes'theorem to find the set of hyperparameters that minimises the loss function of the model, in this case the MSE. The hyperparameter search space is defined with ranges for the convolution philtre size (32 to 256 for the first layer and 32 to 256 for the second layer), the dense layer size (32 to 512), the learning rate (from 1e−5 to 1e−2), the stack size (32, 64 or 128) and the number of epochs (10 or 20). The Bayesian optimisation runs over 50 calls with 10 random starts in order to strike a balance between exploring the search space and exploiting the best parameters found. The best hyperparameters are then used to retrain the final model. This is saved and used to make predictions for both the training and test set. These predictions are then used to calculate evaluation metrics such as RMSE, MAE and R2 for both the training and test data to provide a quantitative measure of the predictive accuracy of the model.

Optimisation of DL-based system using Bayesian optimisation

The optimisation of hyperparameters using Bayesian optimisation for the DL-based system for predicting water quality indices is demonstrated using the target plots for the WQI, SAR, Na%, PI, MH, RSB, KR and RSC. The optimisation landscape for the WQI, as shown in the provided target plot (Fig. 5 and Supplementary Fig. 1–7), shows a multi-dimensional search space in which the Bayesian optimisation process attempts to minimise the loss function, probably the mean squared error. The colour gradient represents the size of the objective function in the search space, with cooler colours (e.g. shades of green) indicating lower values and thus better hyperparameter combinations, while warmer colours (e.g. shades of yellow and red) represent higher values of the objective function. The distribution of the points in the plots indicates the regions explored by the optimisation algorithm. The red stars mark the locations where the best hyperparameter sets were found during the search. The line graph in each row reflects the one-dimensional marginal distribution of the objective function with respect to each hyperparameter, with the red vertical line indicating the position of the best value found for the respective hyperparameter (Table 2). For WQI, the ideal configuration includes convolution philtre sizes of 152 and 158, a dense layer of 147 units, a learning rate of 0.006293117, a stack size of 32 and 20 epochs. The Sodium Absorption Ratio model requires the largest convolution philtre size of 256 and the densest layer size of 246, with a lower learning rate of 0.004014267, indicating the need for a highly complex model with a cautious learning approach. The model for Na% prediction also uses large convolutional filters, but with a smaller dense layer size of 119 and a slightly lower learning rate of 0.003959418, indicating a refined search within a broad feature extraction. For the permeability index, a contrasting setup is chosen with smaller convolution filters (32 and 256) and a minimum dense layer size (32), but a higher learning rate of 0.005840441, which may reflect a different data distribution pattern for this index. The model for magnesium hazard requires the maximum capacity of dense layers (512) with the highest learning rate (0.008512638) and the largest stack size (128), indicating its complexity and the need for extensive data processing. The models for residual sodium bicarbonate and Kelly's ratio both favour moderate configurations with unique convolution philtres and dense layer sizes, while the prediction for residual sodium carbonate uses a high learning rate of 0.01, the maximum in the group, for potentially faster convergence. These customised hyperparameters demonstrate the model’s ability to adapt to the nuances of each specific water quality index for accurate prediction.

The convergence plots for the Bayesian optimisation of the CNN models for different water quality indices show a quantitative reduction of the validation loss over iterations, which indicates the learning efficiency during the training process (Fig. 6). For the WQI model, the validation loss starts at about 8 (probably scaled by 102) and drops rapidly to below 2 within the first 10 iterations, indicating a fast learning phase. This rapid drop is followed by a plateau,

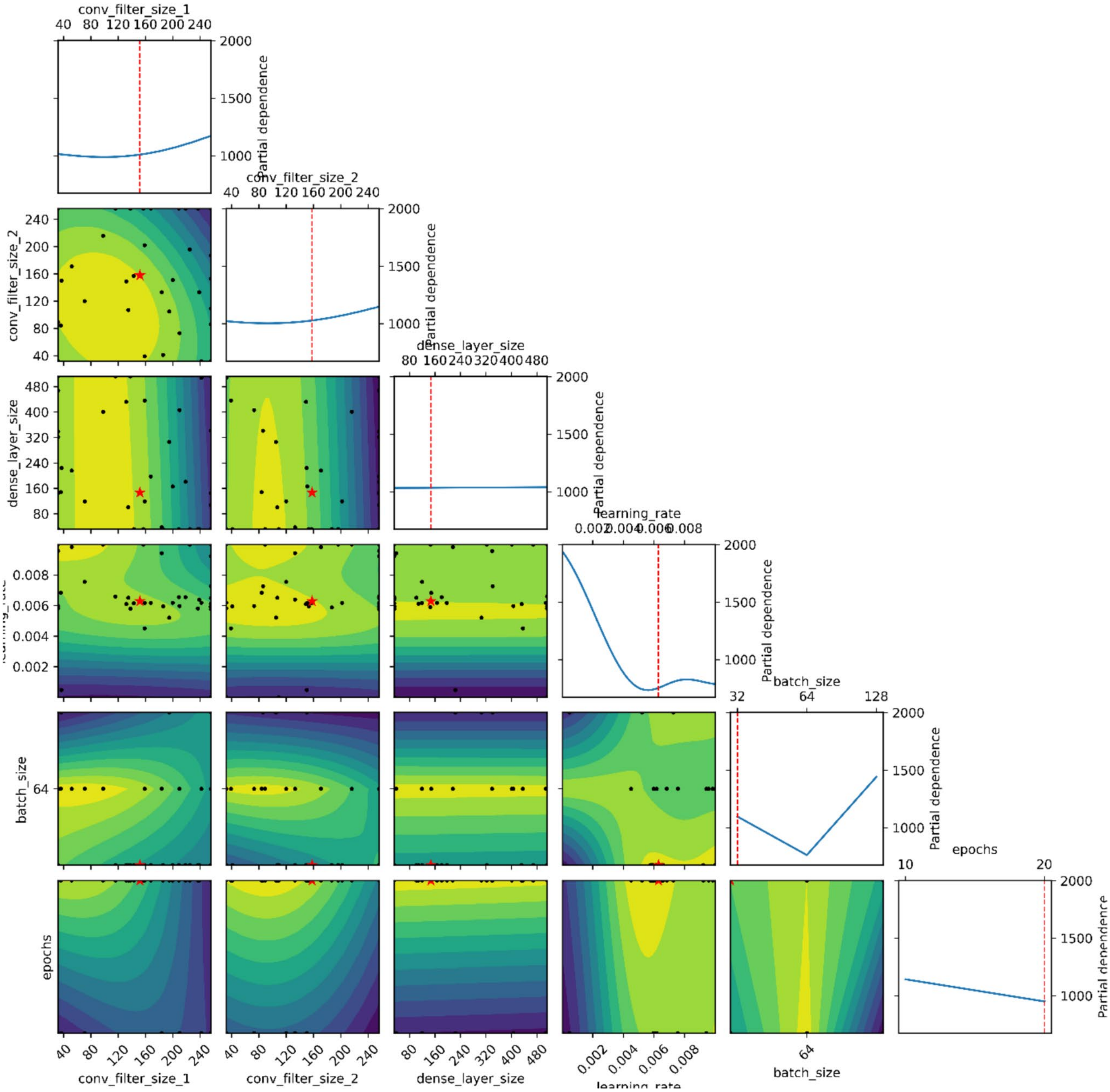

Fig. 5 Contour and marginal distribution plots from Bayesian optimisation showing the relationship between hyperparameter values and model performance for the WQI. The contour maps display the multi-dimensional optimisation space with cooler colours representing lower objective function values and warmer colours indicating higher values

indicating that the model quickly identified a promising region in hyperparameter space and then refined its parameters within that region. The sodium absorption ratio, with an initial loss of just over 2.5 (scaled by 10 − 1), and the permeability index, which starts at around 4 (scaled by 10 − 1), both show a similarly steep decline in loss within the first 10 iterations and stabilise quickly, suggesting rapid convergence to an optimal solution. In contrast, the loss of the Na% model starts close to 9 and shows a more gradual decline, levelling off just above 1, indicating a more complex parameter landscape that requires more iterations to optimise. The Magnesium Hazard and residual sodium carbonate models start with significantly higher initial losses, perhaps indicating greater complexity or less optimal starting parameters, but both show an impressive decline, especially the Magnesium Hazard, which reduces its loss by more than half. The residual sodium bicarbonate and Kelly's ratio models, on the other hand, start with lower initial losses (scaled by

Table 2 Summary of optimal hyperparameter values determined through Bayesian optimisation for deep learning models predicting various water quality indices

	conv_filter_size_1	conv_filter_size_2	dense_layer_size	learning_rate	batch_size	epochs
WQI	152	158	147	0.006293	32	20
Sodium absorption ratio	256	32	246	0.004014	32	20
Na%	256	32	119	0.003959	32	20
Permeability index	32	256	32	0.00584	32	20
Magnesium hazard	32	256	512	0.008513	128	20
Residual sodium bicarbonate	38	83	483	0.002766	32	20
Kelly's ratio	171	32	32	0.006981	32	20
Residual sodium carbonate	151	157	32	0.01	32	20

10 − 1 and 10 − 2, respectively) and show a very gradual decline, suggesting either a smoother optimisation landscape or fewer gains from hyperparameter adjustments. Overall, the models show a common pattern of fast initial learning, where large gains in prediction accuracy are achieved, followed by a levelling off as the models approach the limits of what hyperparameter tuning can achieve given the data and model architecture. This pattern is typical of machine learning, where early leaps in performance are achieved by moving from poor initial settings to more appropriate regions of hyperparameter space, before fine-tuning and marginal improvements characterise the later stages of learning.

Implementation of DL-based system with optimal hyperparameters and their evaluations

After determining the optimal hyperparameters, the CNN-based automated system was retrained on the training datasets and evaluated on separate test datasets. The performance metrics—RMSE, MAE, and R^2—provide a quantitative assessment of model accuracy and generalisation (Table 3 and Fig. 7). Most models demonstrated strong predictive capabilities, indicating effective learning from training data, with minimal overfitting. For example, the WQI model has an R^2 of 0.866 in training and 0.808 in test, with the RMSE increasing slightly from training to test, which could indicate very slight overfitting, but generally indicates good generalisation ability. The sodium absorption ratio model achieves high R^2 values at both training (0.934) and testing (0.851), with a relatively small increase in RMSE and MAE, suggesting that the model is well fitted and generalises effectively to new data. The Na% model shows exceptional performance and generalisability with an R^2 of 0.959 on training and an impressive 0.945 on testing, with a minimal increase in RMSE and MAE, indicating optimal model behaviour with high predictive accuracy. However, the model for MH exhibited a noticeable drop in R^2 from 0.711 (training) to 0.568 (testing), with a corresponding increase in RMSE and MAE. To address overfitting, dropout layers (20%) and L2 regularisation (λ = 0.001) were applied, which introduced a controlled degree of randomness in the learning process and prevented excessive reliance on specific patterns within the training data. Additionally, early stopping was incorporated to halt training once validation loss plateaued, ensuring optimal generalisation. Despite these mitigations, the complexity of MH prediction remains a challenge due to the nonlinear relationships between magnesium concentration, soil interactions, and water quality variables, which require further refinement in future research.. In contrast, the models for residual sodium bicarbonate, Kelly's ratio and residual sodium carbonate retain high R^2 values from training to test, with Kelly's ratio showing exemplary performance and robustness with a test R^2 of 0.940. Overall, most models exhibit strong predictive capabilities, indicating effective learning from the training data. However, the variations in the performance metrics, particularly for the magnesium hazard model, highlight the need for further refinement of the model to ensure consistent prediction accuracy at different water quality indices.

Interpretation of optimised DL-based automated system using XAI

The development of an optimised CNN-based system for rapid monitoring of water quality indices represents a significant advance in environmental management and public health. This system utilises the power of deep learning and can process complex, multi-dimensional water quality data to provide real-time assessments of various indices. To improve the transparency and interpretability of these highly accurate but often opaque models, the application of XAI methods, in particular SHAP values and variable importance plots, is used (Figs. 8 and 9). These XAI techniques decipher the decision-making process of the model by quantifying the impact of each feature on the model's predictions, providing insight into the relationship between specific water quality parameters and the indices in question.

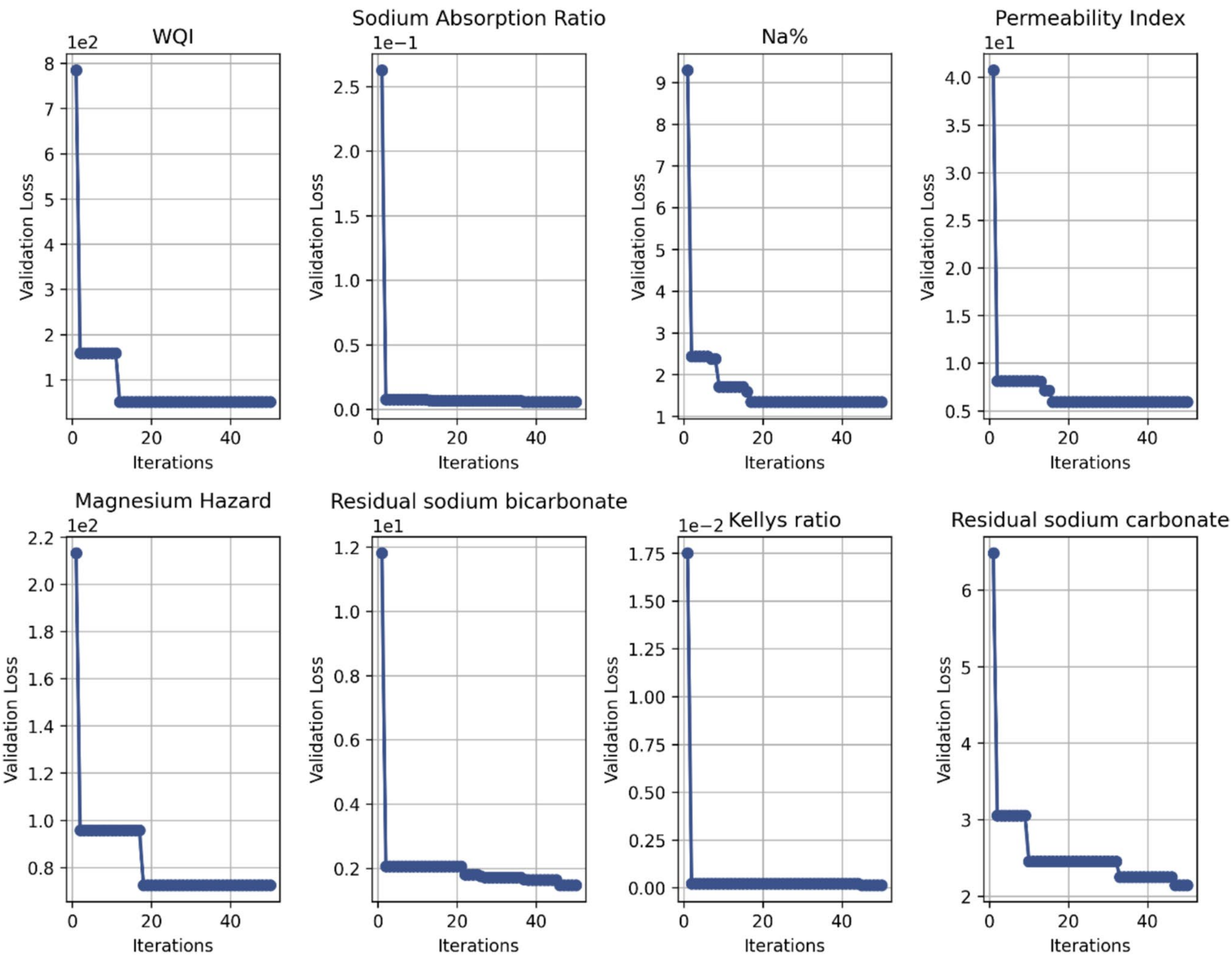

Fig. 6 Convergence plots for Bayesian optimisation of hyperparameters in CNN models predicting water quality indices. Each subplot represents the optimisation progression for a specific index—WQI, sodium absorption ratio, Na%, permeability index, magnesium hazard, residual sodium bicarbonate, Kelly's ratio, and residual sodium carbonate—over 50 iterations. The vertical axes are scaled to the respective magnitudes of validation loss, illustrating the decrease in loss as the optimisation algorithm iteratively refines hyperparameters. The rapid decline in validation loss during early iterations followed by a plateau indicates efficient learning and convergence to optimal hyperparameter values across the different models

Table 3 Performance metrics of CNN models on training and testing datasets, presented for various water quality indices

Indices	Train			Test		
	RMSE	MAE	R^2	RMSE	MAE	R^2
WQI	15.072	11.988	0.866	10.517	9.949	0.808
Sodium absorption ratio	0.068	0.051	0.934	0.125	0.109	0.851
Na%	0.910	0.689	0.959	1.195	0.951	0.945
Permeability index	2.733	2.146	0.907	2.770	2.106	0.885
Magnesium hazard	10.706	8.382	0.711	11.820	9.216	0.568
Residual sodium bicarbonate	1.102	0.847	0.945	1.324	0.909	0.833
Kelly's ratio	0.018	0.014	0.902	0.016	0.014	0.940
Residual sodium carbonate	0.977	0.794	0.979	1.711	1.218	0.856

Metrics include RMSE, MAE, and R^2, reflecting the models' accuracy and predictive power

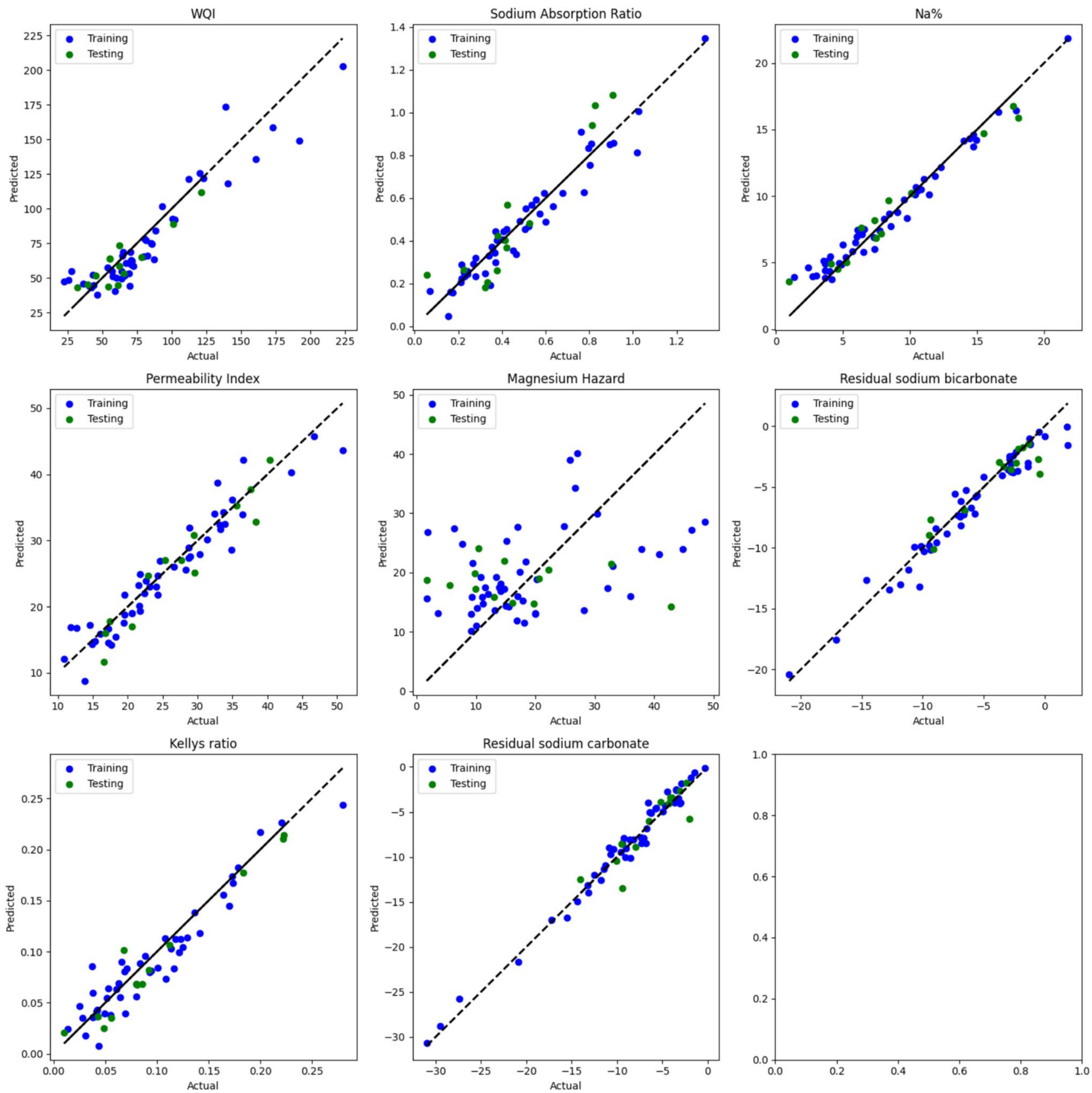

Fig. 7 Scatter plots displaying the performance of CNN models with optimal hyperparameters on training and testing datasets for eight different water quality indices: WQI, Sodium Absorption Ratio, Na%, Permeability Index, Magnesium Hazard, residual sodium bicarbonate, Kelly's ratio, and residual sodium carbonate. Points represent the predicted versus actual values, with training data in blue and testing data in green, clustered around the dashed line that represents perfect prediction accuracy

The SHAP analysis of the WQI model indicates a strong influence of features such as chloride, sulphate and TDS, as indicated by their higher SHAP values. These parameters show higher positional variance along the SHAP value axis, indicating their differential influence on the model's results for different observations. For the SAR model, features such as sodium fraction and calcium have significant SHAP values, indicating that the concentrations of these ions are crucial for the prediction of SAR by the model. The scientific quantification of these influences allows prioritisation of features based on their importance and shows which water quality parameters most influence the water quality indices and should therefore be closely monitored. For example, if features such as TDS (Total Dissolved Solids) or chloride have high SHAP values in the Na% model, this indicates that these variables are key in determining the sodium content

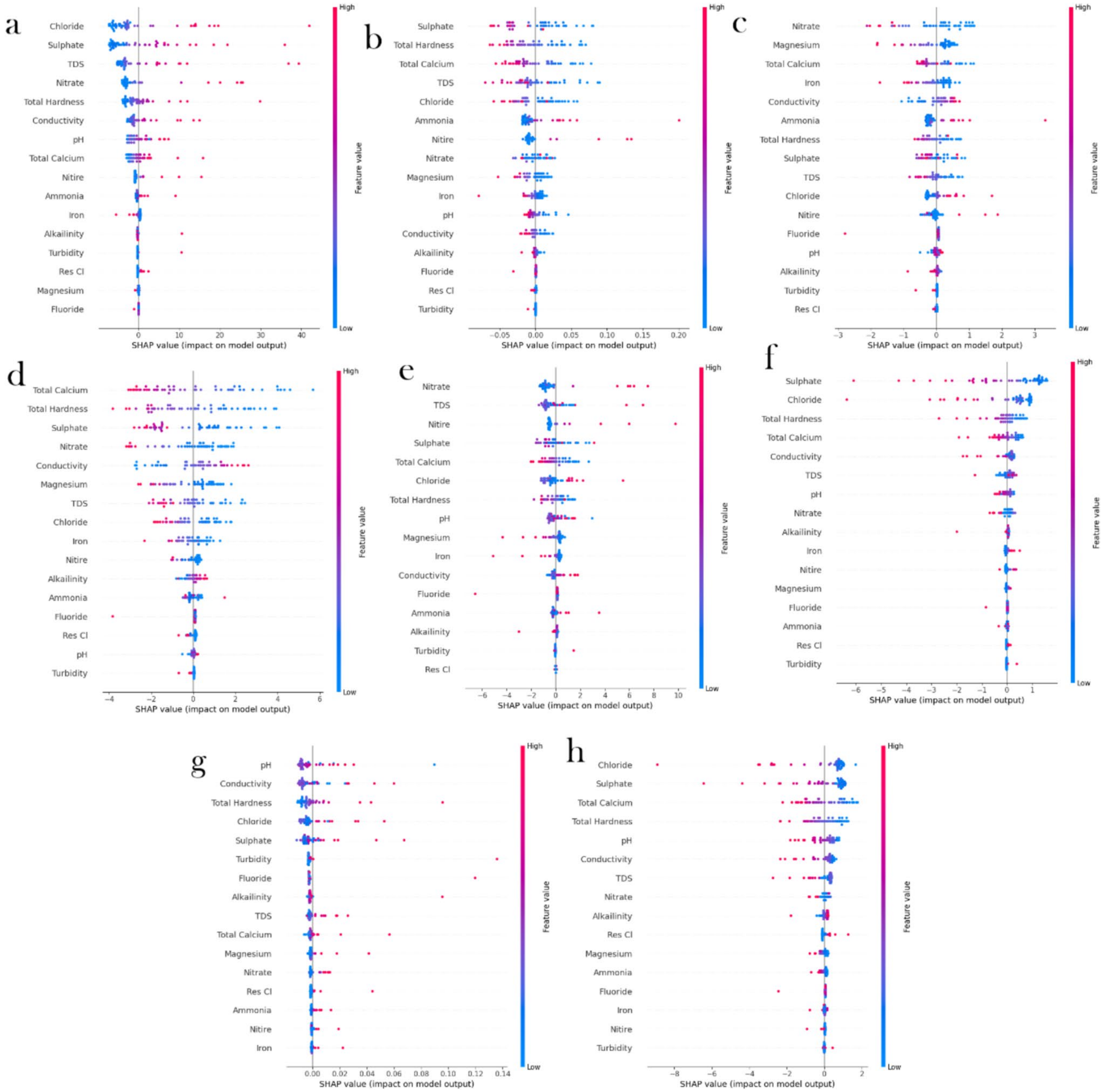

Fig. 8 SHAP value summary plots for eight different water quality indices. Each subplot (a-h) corresponds to a specific index and shows the distribution of SHAP values for each feature within the predictive models. The colour indicates the feature value, with blue representing high and red representing low values. Points to the right of the vertical zero line indicate an increase in the predicted index value, whereas points to the left indicate a decrease. These plots provide a visual representation of the impact of individual water quality parameters on each index, highlighting the features that are most influential in the model's predictions

of the water, with higher values of these features likely to drive up the predicted Na%. Similarly with the permeability index, features with high SHAP values, which may include particle size or organic matter content, indicate the key factors influencing soil permeability predicted by the model. Negative SHAP values indicate a feature that pushes the prediction down, while positive values push it up. A wide distribution of SHAP values for a single trait across predictions indicates that the impact of the trait varies depending on its interaction with other traits. For MH, traits such as nitrate, magnesium, total calcium and iron are the top four influential factors with the highest SHAP values, suggesting that they are the most important predictors of MH in the water samples. In contrast, the characteristics with

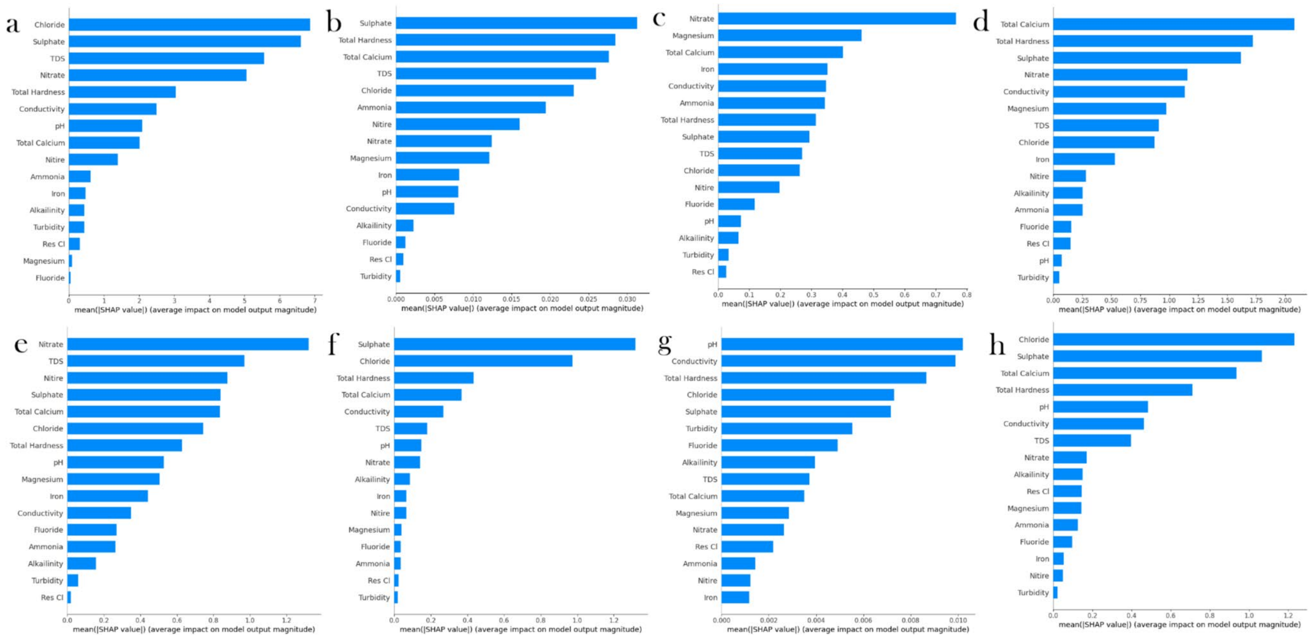

Fig. 9 Variable importance plots derived from the mean absolute SHAP values for eight different water quality indices. Each bar chart (a-h) represents the average impact on the model output magnitude for each feature, ranking them in order of importance for a specific water quality index. These plots summarise the contribution of each parameter to the model's prediction

the lowest influence include ammonia, fluoride, alkalinity and turbidity, which have only a minimal influence on the predictions of the model. This distribution of SHAP values provides a clear quantitative measure of the contribution of each feature to the model's decisions, with higher absolute values indicating a stronger influence. Similarly, for the RSB model, sulphate, chloride, total hardness and total calcium are the most influential parameters, while fluoride, ammonia, residual chloride and turbidity are less influential parameters. The SHAP value plots for KR and RSC show the effects of the various water quality parameters on the model results. For KR, which assesses the sodium hazard for agricultural purposes, parameters such as pH, conductivity, total hardness and chloride appear to be the most influential, as they have the highest mean absolute SHAP values. This indicates that the model weights these factors heavily in the estimation of the LR. The least influential parameters with the lowest SHAP values could include iron, nitrite, ammonia and fluoride, indicating that they play a minor role in the model's predictions of LR. For RSC, the parameters sulphate, total calcium, chloride and total hardness could be important if they have the highest SHAP values, while factors such as turbidity, iron, nitrite and magnesium could have only a minimal influence on the RSC output of the model.

Proposed decision-making based on XAI for all water quality indices

Management strategies to improve WQI and SAR should focus on the parameters identified as most influential in the SHAP analysis. For WQI, efforts to reduce chloride, sulphate and TDS levels could be prioritised as these are the variables that have the greatest impact on the index value. Similarly, management for SAR should focus on regulating sodium and calcium concentrations in the water. This could include implementing treatment processes that target these specific ions or adopting measures aimed at reducing their introduction into water sources. By focusing on these key variables, water quality management can be carried out more efficiently by directing resources and actions to where they are most needed and where they will have the greatest impact on maintaining or improving water quality. Management strategies to improve Na% and PI should be based on these findings. If high TDS levels are a major cause of elevated Na%, efforts could focus on desalination or dilution processes to reduce TDS levels in water sources. If organic content is found to have an impact on PI, strategies to improve PI could include adding organic matter to the soil to improve its structure or adjusting irrigation practises to better manage soil permeability. By targeting management strategies to the highest impact features as identified by the SHAP analysis, interventions can be more targeted and effective, potentially improving water quality for agriculture and drinking water use overall. For management strategies aimed at improving magnesium vulnerability and residual sodium bicarbonate levels, these findings are invaluable. To mitigate MH, measures could focus on regulating activities that lead to high concentrations of nitrate, magnesium, calcium and iron in water sources, such as agricultural runoff or industrial discharges. Measures could include improving wastewater treatment, promoting the use of fertilisers with lower nitrate content or introducing soil improvement measures that bind magnesium and calcium. For the RSB, which is influenced by various parameters, the strategies would differ. For example, if sulphate and chloride are key drivers, then controlling the sources of these ions, perhaps through improved filtration techniques or remediation of contaminated sources, would be effective. Management actions should prioritise these key drivers while monitoring the least influential factors to ensure a holistic approach to improving water quality. Management strategies to optimise the Kelly's ratio should focus on maintaining a balanced pH, monitoring and controlling conductivity and managing hardness and chloride levels in the water as these factors significantly influence the KR. Strategies could include the use of pH adjusters, water softeners or chloride removal processes. To improve RSC, which is important for preventing soil sodicity, focus on controlling sulphate, calcium and chloride levels. This can be done through treatment processes that reduce the concentration of these ions or by using alternative water sources with lower levels of these constituents. By identifying and targeting the key parameters that the model deems most critical, water quality management can be more precise and effective, leading to better results in agricultural water use.

Discussion

Water quality assessment is a crucial component of sustainable water resource management, particularly in regions facing water scarcity and pollution challenges. The Aseer region of Saudi Arabia presents a complex water quality landscape due to environmental and anthropogenic factors, necessitating a robust, automated system for evaluating drinking and irrigation water quality. This study utilised CNNs optimised through Bayesian techniques, coupled with XAI methodologies such as SHAP to assess eight WQIs. By integrating deep learning with interpretability techniques, the study ensures that model predictions are not only accurate but also comprehensible for policymakers and stakeholders.

The results indicate significant spatial variability in water quality across the Aseer region. The DWQI exhibited a mean value of 77.90 with a standard deviation of 39.08, reflecting disparities in water potability. For irrigation, indices such as the SAR, PI, and MH highlighted potential concerns regarding soil degradation and crop yield. High SAR values indicate sodium accumulation, which can deteriorate soil structure and permeability, ultimately affecting agricultural productivity (Islam and Mostafa 2022a). Similarly, variability in PI values suggests the potential for soil compaction, which may hinder root penetration and water retention (Thapa et al. 2020). The MH index, which measures magnesium concentration, identified regions where high magnesium levels could reduce soil fertility and crop health (Muniz et al. 2020). These findings align with previous research demonstrating that irrigation water quality directly influences soil sustainability and long-term agricultural viability (Simsek and Gunduz 2007). The automated CNN-based system for water quality assessment proved to be a methodological breakthrough, significantly reducing analysis time and resource consumption while enabling the simultaneous evaluation of multiple indices. The model demonstrated high predictive accuracy, with the Na% model achieving an R^2 of 0.959 in training and 0.945 in testing, confirming its robustness. However, the MH model's lower accuracy ($R^2 = 0.568$ in testing) indicates challenges in capturing magnesium-related variations, suggesting the need for more diverse datasets, refined architectures, or advanced optimisation techniques. Despite this, the system effectively handled complex environmental data, emphasising the importance of continuous model improvement. The findings contribute to both scientific knowledge and practical water resource management in the Aseer region.

To highlight the contributions of this study, a comparative analysis with previous machine learning applications in water quality assessment is necessary. Table 4 provides an overview of key methodologies, highlighting the CNN-XAI approach as a more interpretable and automated system.

This comparison highlights that while previous studies have successfully applied machine learning to water quality assessment, many lack interpretability and automation. Our CNN-XAI framework ensures transparent decision-making, reducing the "black-box" nature of deep learning models and allowing stakeholders to understand key influencing factors in water quality predictions (Patel et al. 2024). While the study focused on the Aseer region, the CNN-based system has broader applicability for water quality assessment in other hydrological settings. However, variations in water chemistry, land use patterns, and environmental factors across different geographic locations necessitate further model validation. Studies suggest that transfer learning techniques or re-training models on diverse datasets can improve model adaptability (Ibrahim et al. 2023). For instance, applying the model to regions with distinct geological formations or seasonal variations in water quality could help assess its generalisability.

Table 4 Comparison of different water quality prediction methodologies and advancements of the current approach

Methodology	Key features	Advantages	Limitations	Citations
Traditional water quality indices (WQI, IWQI, etc.)	Uses fixed-weight formulas to classify water quality based on physicochemical parameters	Simple and widely used, easy to interpret for decision-making	Lacks adaptability to complex spatial and temporal variations, subjective weighting can introduce bias	Uddin et al. (2021), Verma et al. (2022), Simsek and Gunduz (2007)
Statistical models (PCA, MLR, etc.)	Uses regression and factor analysis to establish relationships between water quality parameters	Identifies key influencing variables, useful for exploratory analysis	Assumes linearity, may not capture nonlinear dependencies in complex datasets	Ahmed et al. (2019), Gupta and Mishra (2023)
Machine learning (ANN, decision trees, SVM, RF)	Employs computational models that learn from historical data to predict water quality indices	Higher accuracy than statistical models, captures nonlinear relationships	Requires large labelled datasets, model tuning can be complex	Ahmed et al. (2021), Kouadri et al. (2021), Lap et al. (2023)
Deep learning (CNN, LSTM, hybrid models)	Uses multi-layer neural networks to automatically extract patterns from water quality datasets	Superior performance in recognising complex interactions, can process large datasets efficiently	High computational cost,"black-box"nature reduces interpretability	Talukdar et al. (2023), Zeng et al. (2023), Mia et al. (2023)
Explainable AI (SHAP, LIME, XAI techniques)	Interpretable ML/DL models that quantify parameter influence on predictions	Enhances trust in predictions, provides transparency in decision-making	Requires additional computational resources, explainability is model-dependent	Talukdar et al. (2024), Das et al. (2023), Sarkar et al. (2023)
Hybrid models (CNN-LSTM, attention-based models)	Combines different ML/DL architectures for improved predictive accuracy	Captures both spatial (CNN) and temporal (LSTM) dependencies in water quality trends	Requires advanced expertise to implement, risk of overfitting without proper regularisation	Yang et al. (2021), Mei et al. (2022), Tan et al. (2022)
Fuzzy logic & GIS-based techniques	Uses rule-based systems and spatial mapping for assessing water quality	Handles uncertainty in data, effective in regional-scale assessments	Requires expert-defined rules, can be subjective in decision thresholds	Patel and Chitnis (2022), Islam and Mostafa (2022a, b, c)
Current Approach (CNN Optimised with Bayesian + SHAP for Explainability)	Uses Bayesian optimisation to enhance CNN model performance and SHAP to improve interpretability	High predictive accuracy, robust parameter tuning, interpretable AI model for decision-making	Computationally intensive, requires expertise in deep learning and XAI techniques	This Study

Future studies could also incorporate regional hydrological parameters to improve model transferability and predictive accuracy in varying climatic conditions (Mo et al. 2024).

While the CNN-based model demonstrated strong predictive accuracy, uncertainty quantification techniques such as confidence intervals (CI) or ensemble learning methods were not incorporated in this study. Ensemble-based methods such as bagging, boosting, or Bayesian neural networks could be explored in future research to enhance model stability and reliability (Patel et al. 2024). These techniques have been effectively applied in hydrological and climate change studies to reduce prediction variance and improve generalisation. Additionally, Bayesian uncertainty estimation techniques can provide a probabilistic assessment of predictions, improving decision-making in water resource management (Islam and Mostafa 2022b). This study incorporated widely accepted IWQIs such as SAR, PI, and MH. However, newly developed indices, such as the integrated irrigation water quality index (IIWQI), provide a comprehensive evaluation of irrigation water suitability by integrating multiple physicochemical parameters into a single score (Islam and Mostafa 2022c). Similarly, GIS-integrated IWQI techniques have been proposed to spatially visualise irrigation water risks (Simsek and Gunduz 2007). Future research should incorporate these advanced indices to ensure a more holistic evaluation of irrigation water quality, particularly in agricultural planning and water management strategies (Jahin et al. 2020).

The integration of CNN models with XAI techniques represents a significant advancement in environmental science, enabling automated, accurate, and interpretable water quality assessments. This study successfully quantified spatial variations in drinking and irrigation water quality while ensuring transparent decision-making through SHAP analysis. Although the model demonstrated high predictive accuracy, further validation across different regions is necessary to ensure wider applicability. Future studies should incorporate ensemble learning, Bayesian uncertainty estimation, and advanced IWQIs to further enhance model robustness and adaptability. By leveraging these advancements, future research can contribute to sustainable water resource management and policy-making in regions facing water quality challenges.

Conclusion

This comprehensive study in the Aseer region of Saudi Arabia represents a significant advancement in water quality assessment by integrating CNN optimised with Bayesian techniques. The predictive modelling of eight key water quality indices, including the DWQI with an average value of 77.90 ($\pm$ 39.08), demonstrated the diverse nature of water quality in the region. Notably, the Na% exhibited the highest accuracy ($R^2 = 0.959$ in training, 0.945 in testing), highlighting its predictive robustness. The integration of DL and XAI techniques, particularly SHAP analysis, enhanced the interpretability of model predictions by identifying key influencing parameters such as chloride, sulphate, and TDS. This study underscores the potential of advanced computational methods to improve water quality management and decision-making, particularly in regions facing both environmental and anthropogenic challenges. However, limitations remain, including the dependency on data quality and the inherent complexity of water quality interactions despite interpretability enhancements through SHAP analysis. Future research should incorporate a broader range of environmental factors and explore the adaptability of these models in diverse geographical contexts. This study serves as a valuable framework for researchers and policymakers, demonstrating the power of AI-driven approaches in addressing real-world environmental challenges and ensuring sustainable water resource management.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13201-025-02452-y.

Acknowledgements The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Research Group under grant number RGP2/412/45.

Author's contribution Javed Mallick, Saeed Alqadhi, and Majed Alsubih were involved in the conceptualisation; Javed Mallick and Saeed Alqadhi contributed to the data curation, formal analysis and project administration; Saeed Alqadhi acquired the funding; Javed Mallick was involved in the methodology; Majed Alsubih contributed to the resources; Javed Mallick contributed to the software and supervision; Saeed Alqadhi and Javed Mallick performed the validation; Saeed Alqadhi, Javed Mallick, Mohamed Fatahalla Mohamed Ahmed, and Hazem Ghassan Abdo assisted in writing—original draft; Majed Alsubih contributed to the writing—review and editing.

Funding Funding for this research was given under award numbers RGP2/412/45 by the Deanship of Scientific Research; King Khalid University, Ministry of Education, Kingdom of Saudi Arabia.

Availability of data and materials The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors declare that they have no competing interests. No potential conflict of interest was reported by the authors.

Ethics approval and consent to participates Not applicable.

Consent for publication Not applicable.

References

Aghazadeh N, Mogaddam AA (2010) Assessment of groundwater quality and its suitability for drinking and agricultural uses in the Oshnavieh area, Northwest of Iran. J Environ Prot 1(01):30

Ahmed AN, Othman FB, Afan HA, Ibrahim RK, Fai CM, Hossain MS et al (2019) Machine learning methods for better water quality prediction. J Hydrol 578:124084

Ahmed M, Mumtaz R, Hassan Zaidi SM (2021) Analysis of water quality indices and machine learning techniques for rating water pollution: a case study of Rawal Dam, Pakistan. Water Supply 21(6):3225–3250

Ahmed IA, Talukdar S, Baig MRI, Ramana GV, Rahman A (2024) Quantifying soil erosion and influential factors in Guwahati's urban watershed using statistical analysis, machine and deep learning. Remote Sens Appl Soc Environ 33:101088

Alfaleh A, Ben Khedher N, Alviz-Meza A (2023) Is the entropy-weighted water quality index a suitable index for evaluating the groundwater quality in Ha'il, Saudi Arabia? Water Sci Technol 88(3):778–797

Asante-Annor A, Bewil PN, Boateng D (2018) Evaluation of groundwater suitability for irrigation in the Lambussie-Karni district of Ghana. Ghana Min J 18(1):9–19

Batarseh M, Imreizeeq E, Tilev S, Al Alaween M, Suleiman W, Al Remeithi AM et al (2021) Assessment of groundwater quality for irrigation in the arid regions using irrigation water quality index (IWQI) and GIS-Zoning maps: case study from Abu Dhabi Emirate, UAE. Groundw Sustain Dev 14:100611

Busico G, Kazakis N, Cuoco E, Colombani N, Tedesco D, Voudouris K, Mastrocicco M (2020) A novel hybrid method of specific vulnerability to anthropogenic pollution using multivariate statistical and regression analyses. Water Res 171:115386

Centers for Disease Control and Prevention (CDC). (2022). Drinking water standards and regulations: an overview. Accessed from https://www.cdc.gov/drinking-water/about/drinking-water-standards-and-regulations-an-overview.html

Chen C, Fan L (2023) An attribution deep learning interpretation model for landslide susceptibility mapping in the Three Gorges Reservoir Area. IEEE Trans Geosci Remote Sens 61:1

Dabbagh AE, Abderrahman WA (1997) Management of groundwater resources under various irrigation water use scenarios in Saudi Arabia. Arab J Geosci Eng 22(1):47–64

Das T, Talukdar S, Shahfahad, Baig MRI, Hang HT, Siddiqui AM, Rahman A (2023) Assessing vulnerability to cyclones in coastal Odisha using fuzzy logic integrated AHP: towards effective risk management. Spat Inf Res 32: 277–295

El Bilali A, Taleb A (2020) Prediction of irrigation water quality parameters using machine learning models in a semi-arid environment. J Saudi Soc Agric Sci 19(7):439–451

Elsayed S, Hussein H, Moghanm FS, Khedher KM, Eid EM, Gad M (2020) Application of irrigation water quality indices and multivariate statistical techniques for surface water quality assessments in the Northern Nile Delta, Egypt. Water 12(12):3300

Gupta S, Gupta SK (2021) A critical review on water quality index tool: genesis, evolution and future directions. Eco Inform 63:101299

Gupta D, Mishra VK (2023) Development of entropy-river water quality index for predicting water quality classification through machine learning approach. Stoch Env Res Risk Assess 37(11):4249–4271

Han X, Tang F, Liu AL (2023) Drinking water quality evaluation in supply systems in Wuhan, China: application of entropy weight water quality index and multivariate statistical analysis. Environ Sci Pollut Res 31:1–13

Ibrahim H, Yaseen ZM, Scholz M, Ali M, Gad M, Elsayed S et al (2023) Evaluation and prediction of groundwater quality for irrigation using an integrated water quality indices, machine learning models and GIS approaches: a representative case study. Water 15(4):694

Iqbal AB, Rahman MM, Mondal DR, Khandaker NR, Khan HM, Ahsan GU et al (2020) Assessment of Bangladesh groundwater for drinking and irrigation using weighted overlay analysis. Groundw Sustain Dev 10:100312

Islam MS, Mostafa MG (2022a) Comparison of classical and developed indexing methods for assessing the groundwater suitability for irrigation. J Sustain Agric Environ 1(3):226–239

Islam MS, Mostafa MG (2022b) Development of an integrated irrigation water quality index (IIWQIndex) model. Water Supply 22(2):2322–2337

Islam MS, Mostafa MG (2022c) Suitability of water quality index methods for assessing groundwater quality in the Ganges River basin area. H2Open J 5(2):198–220

Jahin HS, Abuzaid AS, Abdellatif AD (2020) Using multivariate analysis to develop irrigation water quality index for surface water in Kafr El-Sheikh Governorate, Egypt. Environ Technol Innov 17:100532

Khoi DN, Quan NT, Linh DQ, Nhi PTT, Thuy NTD (2022) Using machine learning models for predicting the water quality index in the La Buong River, Vietnam. Water 14(10):1552

Kouadri S, Elbeltagi A, Islam ARMT, Kateb S (2021) Performance of machine learning methods in predicting water quality index based on irregular data set: application on Illizi region (Algerian southeast). Appl Water Sci 11(12):190

Krishan G, Kumar M, Rao MS, Garg R, Yadav BK, Kansal ML et al (2023) Integrated approach for the investigation of groundwater quality through hydrochemistry and water quality index (WQI). Urban Clim. 47:101383

Lap BQ, Du Nguyen H, Hang PT, Phi NQ, Hoang VT, Linh PG, Hang BTT (2023) Predicting water quality index (WQI) by feature selection and machine learning: a case study of An Kim Hai irrigation system. Eco Inform 74:101991

Mallick J, Kumar A, Almesfer MK, Alsubih M, Singh CK, Ahmed M, Khan RA (2021) An index-based approach to assess groundwater quality for drinking and irrigation in Asir region of Saudi Arabia. Arab J Geosci 14:1–17

Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RF (2019) Epilepsy-related and other causes of mortality in people with epilepsy: a systematic review of systematic reviews. Epilepsy Res 157:106192

Mei P, Li M, Zhang Q, Li G (2022) Prediction model of drinking water source quality with potential industrial-agricultural pollution based on CNN-GRU-Attention. J Hydrol 610:127934

Mia MY, Haque ME, Islam ARMT, Jannat JN, Jion MMMF, Islam MS et al (2023) Analysis of self-organizing maps and explainable artificial intelligence to identify hydrochemical factors that

drive drinking water quality in Haor region. Sci Total Environ 904:166927
Mo Y, Xu J, Liu C, Wu J, Chen D (2024) Assessment and prediction of Water Quality Index (WQI) by seasonal key water parameters in a coastal city: application of machine learning models. Environ Monit Assess 196(11):1008
Muniz DH, Malaquias JV, Lima JE, Oliveira-Filho EC (2020) Proposal of an irrigation water quality index (IWQI) for regional use in the Federal District, Brazil. Environ Monit Assess 192(9):607
Nava L, Carraro E, Reyes-Carmona C, Puliero S, Bhuyan K, Rosi A et al (2023) Landslide displacement forecasting using deep learning and monitoring data across selected sites. Landslides 20(10):2111–2129
Patel A, Chitnis K (2022) Application of fuzzy logic in river water quality modelling for analysis of industrialization and climate change impact on Sabarmati river. Water Supply 22(1):238–250
Patel A, Yadav SM, Teegavarapu R (2024) Enhancing real-time flood forecasting and warning system by integrating ensemble techniques and hydrologic model simulations. J Water Clim Change 15(9):4307–4327
Ram A, Tiwari SK, Pandey HK, Chaurasia AK, Singh S, Singh YV (2021) Groundwater quality assessment using water quality index (WQI) under GIS framework. Appl Water Sci 11:1–20
Sakizadeh M (2016) Artificial intelligence for the prediction of water quality index in groundwater systems. Model Earth Syst Environ 2:1–9
Salifu M, Aidoo F, Hayford MS, Adomako D, Asare E (2017) Evaluating the suitability of groundwater for irrigational purposes in some selected districts of the Upper West region of Ghana. Appl Water Sci 7:653–662
Sarkar SK, Rudra RR, Sohan AR, Das PC, Ekram KMM, Talukdar S et al (2023) Coupling of machine learning and remote sensing for soil salinity mapping in coastal area of Bangladesh. Sci Rep 13(1):17056
Sheikh Khozani Z, Iranmehr M, Wan Mohtar WHM (2022) Improving Water Quality Index prediction for water resources management plans in Malaysia: application of machine learning techniques. Geocarto Int 37(25):10058–10075
Simsek C, Gunduz O (2007) IWQ index: a GIS-integrated technique to assess irrigation water quality. Environ Monit Assess 128:277–300
Talukdar S, Ahmed S, Naikoo MW, Rahman A, Mallik S, Ningthoujam S et al (2023) Predicting lake water quality index with sensitivity-uncertainty analysis using deep learning algorithms. J Clean Prod 406:136885
Talukdar S, Bera S, Naikoo MW, Ramana GV, Mallik S, Kumar PA, Rahman A (2024) Optimisation and interpretation of machine and deep learning models for improved water quality management in Lake Loktak. J Environ Manag 351:119866
Tan W, Zhang J, Wu J, Lan H, Liu X, Xiao K et al (2022) Application of cnn and long short-term memory network in water quality predicting. Intell Autom Soft Comput 34(3):1943–1958
Thapa R, Gupta S, Kaur H (2020) Introducing an irrigation water quality index (IWQI) based on the case study of the Dwarka River basin, Birbhum, West Bengal, India. Sustain Water Resour Manag 6(5):86
U.S. Environmental Protection Agency (EPA) (2023) National primary drinking water regulations. Accessed from https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations
Uddin MG, Nash S, Olbert AI (2021) A review of water quality index models and their use for assessing surface water quality. Ecol Ind 122:107218
Uddin MG, Nash S, Diganta MTM, Rahman A, Olbert AI (2022b) Robust machine learning algorithms for predicting coastal water quality index. J Environ Manag 321:115923
Uddin G, Nash S, Olbert AI (2022a) Optimization of parameters in a water quality index model using principal component analysis. In Proceedings of the 39th IAHR world congress, vol 19, p 24. International Association for Hydro-Environment Engineering and Research (IAHR) Spain
United Nations Environment Programme (UNEP) (2021) Progress on integrated water resour manage. Tracking SDG 6 series: global indicator 6.5.1 updates and acceleration needs. United Nations Environment Programme. https://www.unwater.org/publications/progress-on-integrated-water-resources-management-651/. Accessed Feb 2022
Verma M, Loganathan VA, Bhatt VK (2022) Development of entropy and deviation-based water quality index: case of River Ganga, India. Ecol Indic 143:109319
World Health Organization (WHO) (2017) Guidelines for drinking-water quality, 4th edition, incorporating the first addendum. WHO, Geneva
Yang Y, Xiong Q, Wu C, Zou Q, Yu Y, Yi H, Gao M (2021) A study on water quality prediction by a hybrid CNN-LSTM model with attention mechanism. Environ Sci Pollut Res 28(39):55129–55139
Zeng T, Glade T, Xie Y, Yin K, Peduto D (2023) Deep learning powered long-term warning systems for reservoir landslides. Int J Disaster Risk Reduct 94:103820
Zhe W, Xigang X, Feng Y (2021) An abnormal phenomenon in entropy weight method in the dynamic evaluation of water quality index. Ecol Ind 131:108137