Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00129-4

Original Article

Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023)

Okoi Obeten [1], Christiana Uchenna Ezeanya [2]

1. *Computer Science, Auchi Polytechnic, Auchi, NGA*
2. *Information Systems and Technology, National Open University of Nigeria, Abuja, NGA*

Received: April 19, 2026 | Review began: May 27, 2026 | Review ended: June 15, 2026 | Published: June 19, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Accurate prediction and transparent attribution of water quality deterioration are essential for evidence-based environmental governance, yet most machine learning studies on Indian water quality provide predictive metrics without interpretable explanations of which parameters drive degradation. This study proposes an attention-augmented Long Short-Term Memory (Attention-LSTM) framework for Water Quality Index (WQI) prediction, combined with Shapley Additive Explanations (SHAP) to deliver both global and temporally resolved feature attribution. The Central Pollution Control Board (CPCB) National Water Monitoring Programme dataset spanning 2021-2023, covering 155 complete records across 17 Indian states and 11 water body types, was used. A Weighted Arithmetic WQI was computed using BIS 2012 standards, incorporating a log10-transformation of Total Coliform and Electrical Conductivity to address their log-normal scale distributions. The proposed Attention-LSTM model achieved an R^2 of 0.969 (±0.020), a Root Mean Squared Error (RMSE) of 6.88 (±2.31), and a Mean Absolute Error (MAE) of 3.89 (±0.59) across stratified 5-fold cross-validation, outperforming Random Forest (R^2 = 0.781), XGBoost (R^2 = 0.870), Support Vector Regression (R^2 = 0.834), standard LSTM (R^2 = 0.615). SHAP analysis identified Total Coliform and Biochemical Oxygen Demand as the dominant global drivers influencing Water Quality Index (WQI) predictions. Temporal drift analysis revealed a significant transition from microbial dominance (Total Coliform log SHAP value = 19.16 in 2021) to organic pollution dominance (BOD SHAP value = 13.50 in 2023), as quantified by a decline in Kendall's tau from 0.81 to 0.43. SHAP-LIME (Local Interpretable Model-agnostic Explanations) agreement analysis produced a Spearman's rho of 0.643 and an Agreement Index of 0.667, indicating moderate consistency between the two interpretability approaches. These findings provide actionable, parameter-specific insights to support water resource management across Indian river systems.

Categories: AI applications, Explainable AI, Machine Learning (ML)

Keywords: water quality index, attention-lstm, shap, temporal feature drift, explainable artificial intelligence, indian water quality, cpcb

Introduction

Access to safe water is among the most pressing global challenges of the twenty-first century, with the United Nations Sustainable Development Goal 6 mandating clean water and sanitation for all by 2030. India, with its dense population, rapid industrialization, and dependence on river systems for agriculture and domestic supply, faces acute water quality pressures. The Central Pollution Control Board (CPCB) monitors over 4,700 locations nationwide under the National Water

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI
https://doi.org/10.7759/s44389-026-00129-4

Monitoring Programme (NWMP); however, translating vast monitoring datasets into timely and interpretable management decisions remains a persistent challenge [1-3]. A CPCB assessment conducted in 2022 identified 311 polluted river stretches across 279 rivers in 30 states and union territories, with Biochemical Oxygen Demand (BOD) identified as the primary pollution indicator [1]. By 2023, the total number of polluted river sites had reached 807, and although a marginal reduction was observed relative to 2022, the persistence of more than 800 sites underscores the scale of the monitoring and remediation challenge [2].

The Water Quality Index (WQI) is the most widely adopted tool for condensing multiple physicochemical parameters into a single composite score that communicates overall water quality status to managers and the public [4,5]. Machine learning methods have been extensively applied to Indian WQI datasets for classification and regression tasks. Random Forest, XGBoost, and Support Vector Regression (SVR) have demonstrated strong predictive performance in various Indian water quality studies [6-8]. Deep learning approaches, particularly standard Long Short-Term Memory (LSTM) networks, have been applied to Indian river and groundwater datasets to exploit temporal dependencies in sequential quality data [9,10]. More recently, attention-augmented LSTM architectures have been shown to outperform standard LSTMs for water quality prediction tasks by selectively weighting the most informative features and time steps [11,12].

Despite this body of work, three critical gaps persist. First, most studies treat WQI prediction as a static regression problem and do not examine how the relative importance of different water quality drivers changes over time. The CPCB 2021-2023 dataset spans a unique post-pandemic window during which industrial activity, agricultural runoff patterns, and monitoring coverage changed substantially; however, no study has characterised how these dynamics alter the dominant predictors of WQI degradation. Second, standard LSTM architectures process input features uniformly without weighting their relative temporal contributions, limiting the model ability to amplify the signal from the most pollution-relevant parameters. Third, explainability in existing Indian water quality ML studies is predominantly limited to feature importance from tree-based models [13,14], and cross-method agreement between Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) explanations has not been evaluated, leaving open the question of whether attributed feature importances are model-specific artefacts or genuine data-driven signals.

This study addresses all three gaps through a unified framework with four distinct contributions. First, a weighted arithmetic (WA) WQI formulation incorporating a justified log10-transformation of Total Coliform (TC) and Electrical Conductivity (EC) is presented, preventing sub-index domination by high-magnitude microbial and ionic concentrations. Second, an Attention-LSTM architecture integrating a self-attention mechanism over the LSTM hidden state is proposed, enabling the model to weight the relative informativeness of physicochemical features during WQI prediction. Third, a temporally resolved SHAP analysis tracking how feature importance shifts across three annual windows (2021, 2022, 2023) is conducted, quantified by Kendall's tau rank correlation. Fourth, a SHAP-LIME cross-method agreement analysis using Spearman rank correlation and a novel Agreement Index (AI) is performed, establishing the reliability of the generated explanations.

Materials And Methods

Dataset description

The dataset used in this study is the CPCB Indian Water Quality Dataset (2021-2023), sourced from the National Water Monitoring Programme and publicly available via Kaggle [15]. The dataset contains 194 station-year records spanning 17 Indian states and 11 water body types: River, Drain, Canal, Pond, Lake, Beach, Marine, Sea, Water Treatment Plant, Sewage Treatment Plant, and Creek. Each record reports annual minimum and maximum observations for eight physicochemical parameters: Temperature, Dissolved Oxygen (DO), pH, EC, BOD, Nitrate-Nitrogen (NO3), TC, and Fecal Coliform. The mean of the minimum and maximum values was computed for each parameter to derive a single central estimate per station-year record, consistent with annual summary reporting practices. After excluding records with missing values in any of the seven modelling parameters, 155 complete records were retained for analysis. Listwise deletion was adopted because the missing data pattern was not uniformly distributed across parameters and states, and physicochemical imputation carries ecological risks, as imputed values could artificially alter the computed WQI distribution. The 39 excluded records

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

should be considered a limitation: they may not be missing completely at random, and future work incorporating imputation sensitivity analysis would strengthen the robustness of the findings. It should be noted that the 155 retained records represent the complete population of station-year observations with non-missing values across all seven modelling parameters in the publicly available CPCB dataset for 2021-2023, rather than a subsample. Therefore, findings should be interpreted within the scope of the 17 monitored states and 11 water body types represented, and external validity to unmonitored regions or future years would require model recalibration using updated data.

The temporal distribution of records comprises 44 records in 2021, 53 in 2022, and 97 in 2023, reflecting the progressive expansion of CPCB monitoring coverage. The most represented states are Himachal Pradesh (49 records), Assam (46 records), and Andhra Pradesh (37 records). Among the 155 complete records, all but 18 fell within the Unsuitable (WQI > 100) or Very Poor (75-100) classification categories, indicating widespread water quality challenges across the monitored systems. Table *1* presents the descriptive statistics of all modelling parameters.

Parameter	Count	Mean	SD	Min	25%	75%	Max
DO (mg/L)	155	6.47	1.80	0.70	5.43	7.65	9.55
pH	155	7.55	0.42	6.50	7.25	7.84	9.85
EC (log10)	155	2.90	0.91	1.74	2.18	3.31	4.68
BOD (mg/L)	155	3.51	5.28	1.00	1.50	2.60	45.60
NO3 (mg/L)	155	1.79	1.78	0.30	0.90	1.76	9.90
TC (log10)	155	2.57	0.80	0.48	2.16	3.04	5.60
Temp (°C)	155	22.97	6.38	7.00	20.50	27.50	32.00
WQI	155	135.86	42.43	51.62	120.65	142.19	378.67

TABLE 1: Descriptive statistics of physicochemical parameters and WQI (n = 155). EC and TC are on the log10(x + 1) scale used in modelling

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SD, Standard Deviation; TC, Total Coliform; WQI, Water Quality Index

WQI computation

The WA-WQI was computed, adopting BIS 10500:2012 permissible standards. The overall index is expressed as the weighted average of quality ratings (equation 1).

$$WQI = \frac{\sum_{i=1}^{n} w_i q_i}{\sum_{i=1}^{n} w_i} \quad (1)$$

where w_i represents the unit weight of the *i*-th parameter, and q_i denotes its corresponding quality rating (sub-index). For most water quality parameters, the sub-index is calculated as a standard quality rating (equation 2).

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

$$q_i = \left(\frac{C_i - I_i}{S_i - I_i} \right) \times 100 \quad (2)$$

where C_i is the observed concentration, I_i is the ideal value, and S_i is the permissible standard. However, for DO, an inverse relationship is applied since higher concentrations indicate better water quality. Thus, the sub-index is computed as an inverted quality rating (equation 3).

$$q_i = \left(\frac{I_i - C_i}{I_i - S_i} \right) \times 100 \quad (3)$$

To ensure numerical stability and prevent distortion of the composite index, all negative sub-index values are clipped to zero prior to aggregation.

A critical methodological consideration is the treatment of TC and EC. In this dataset, TC values span six orders of magnitude (2 to 7,395,000 MPN/100 mL), while EC values span three orders of magnitude (4 to 61,900 µmho/cm), with both exhibiting strongly log-normal distributions. Using raw values in the WQI sub-index formula would cause the TC and EC sub-indices to dominate the composite score, thereby masking the contributions of other biologically relevant parameters. This known limitation has been acknowledged in prior Inan WQI studies [6,10]. Therefore, a log10(x + 1) transformation was applied to both parameters prior to sub-index calculation, with the corresponding BIS standards transformed to log10(3 + 1) = 0.602 for TC and log10(750 + 1) = 2.876 for EC. This transformation is ecologically justified because microbial and ionic concentrations typically follow log-normal distributions, and it is statistically validated by comparing WQI distributions under raw and log-transformed sub-indices. Table *2* presents the complete WQI parameter configuration.

Parameter	Ideal	BIS Standard	Weight	Sub-Index Type	Notes
DO (mg/L)	14.6	6.0	0.17	Inverted	Higher = better
BOD (mg/L)	0	3.0	0.16	Standard	
NO3 (mg/L)	0	45.0	0.11	Standard	
pH	7.0	8.5	0.12	Standard	
EC (log10)	0	2.876	0.12	Standard	log10(750+1)
TC (log10)	0	0.602	0.17	Standard	log10(3+1)
Temp (°C)	0	30.0	0.15	Standard	

TABLE 2: WA-WQI parameter configuration. BIS 10500:2012 permissible standards. EC and TC standards are log10-transformed to match input data scale

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; TC, Total Coliform; WA, Weighted Arithmetic; WQI, Water Quality Index

The resulting WQI distribution for the 155 complete records ranges from 51.6 to 378.7, with a mean of 135.9 and a median of 133.1. A Kruskal-Wallis test confirmed a statistically significant difference in WQI distributions across the three years ($H = 8.31$, $p = 0.016$), supporting the temporal analysis approach taken in this study. Table *3* presents the WQI

classification cross-tabulated by year.

Year	Excellent	Good or Poor	Very Poor	Unsuitable
2021	0	8	2	21
2022	0	0	0	29
2023	0	1	5	89
Total	0	9	7	139

TABLE 3: WQI classification distribution by year (n = 155)

WQI, Water Quality Index

Feature engineering and cross-validation setup

The feature matrix was constructed by combining the seven log-scaled physicochemical parameters with a one-hot encoding of water body type (11 categories), a one-hot encoding of state name (17 categories), and a year indicator encoded as an ordinal integer (0, 1, 2 for 2021, 2022, and 2023, respectively) to preserve temporal ordinality. The resulting feature matrix contains 37 variables. The target variable is the computed WQI value.

Model evaluation followed a stratified 5-fold cross-validation scheme. Stratification was based on WQI quartile membership to ensure each fold contains a representative distribution of water quality severity levels. Scaling was applied using MinMaxScaler fitted within each training fold and applied to the corresponding test fold, preventing data leakage. All reported metrics represent means and standard deviations across the five test folds.

Baseline models

Four baseline models were trained for performance comparison. Random Forest used 200 estimators with square-root feature subsampling [16]. XGBoost used 300 estimators, a learning rate of 0.05, a maximum depth of 6, and a subsampling ratio of 0.8 [17]. SVR used a radial basis function kernel with C = 100 and $\epsilon = 0.1$, applied after MinMaxScaler normalization. The standard LSTM comprised two stacked LSTM layers (64 hidden units each, dropout = 0.2), trained using the Adam optimizer (learning rate = 0.001, weight decay = 0.0001) with early stopping at a patience of 25 epochs. LSTM-based architectures have been widely adopted for water quality time-series modelling due to their ability to capture long-term dependencies [18,19].

Attention-LSTM architecture

The proposed Attention-LSTM architecture extends the conventional LSTM network by integrating an embedding layer and a self-attention mechanism to enhance feature representation and predictive performance. It is important to note that the dataset in this study consists of annual station-year summary records rather than high-frequency time series. Each record is processed as a single feature vector; the temporal contribution of the framework resides in the year-stratified SHAP drift analysis rather than in multi-step sequential forecasting. The LSTM architecture captures feature-level dependencies within each record, with the ordinal year indicator preserving inter-annual temporal context. First, the input vector $\boldsymbol{x} \in \mathbb{R}^{37}$ is passed through a linear embedding layer to obtain a lower-dimensional representation (equation 4).

$$\boldsymbol{x}' = \mathrm{ReLU}(\boldsymbol{W}_e \boldsymbol{x} + \boldsymbol{b}_e) \quad (4)$$

where $x \in \mathbb{R}^{d_{\text{in}}}$, $W_e \in \mathbb{R}^{37 \times d_{\text{in}}}$ is the embedding weight matrix, $b_e \in \mathbb{R}^{37}$ is the bias vector, and $x' \in \mathbb{R}^{37}$ represents the compressed feature vector.

To capture temporal dependencies, the sequence of embedded features $X' = [x'_1, x'_2, \ldots, x'_T]$ is processed using a two-layer LSTM network. With a total sequence length T, the LSTM recursively updates its hidden states h_t and cell states c_t at each time step t according to

$$h_t, c_t = \text{LSTM}(x'_t, h_{t-1}, c_{t-1}) \tag{5}$$

A dropout rate of 0.2 is applied between the LSTM layers to prevent overfitting. The sequence of hidden states $H = [h_1, h_2, \ldots, h_T] \in \mathbb{R}^{T \times 64}$ is preserved to retain contextual information across all time steps, where individual hidden units are constrained to $h_t \in \mathbb{R}^{64}$.

To adaptively highlight critical time steps, a single-head self-attention mechanism is applied across the entire sequence of hidden states H. The Query (Q), Key (K), and Value (V) matrices are derived via linear projections:

$$Q = HW_q, \quad K = HW_k, \quad V = HW_v \tag{6}$$

where $W_q, W_k, W_v \in \mathbb{R}^{64 \times d_k}$ represent the trainable weight projection matrices, and $d_k = 64$ denotes the attention projection dimension. The attention output matrix is subsequently computed using the scaled dot-product formulation:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

where $\sqrt{d_k}$ serves as the scaling factor to regulate variance in the dot-product computations. The final context vector $h_{\text{attn}} \in \mathbb{R}^{64}$ is extracted via a temporal pooling layer and is passed through a fully connected dense layer with a Rectified Linear Unit (ReLU) activation function:

$$z = \text{ReLU}(W_c h_{\text{attn}} + b_c) \tag{8}$$

where $W_c \in \mathbb{R}^{32 \times 64}$ is the weight matrix, $b_c \in \mathbb{R}^{32}$ is the bias vector, and $z \in \mathbb{R}^{32}$ represents the localized latent feature representation.

Finally, the predicted WQI ($\hat{y}$) is generated via a linear regression layer:

$$\hat{y} = W_o z + b_o \tag{9}$$

where $W_o \in \mathbb{R}^{1 \times 32}$ is the output weight vector, and $b_o \in \mathbb{R}$ is the scalar bias.

The model is trained end-to-end using the Adam optimizer with a learning rate of 0.001 and a weight decay of 0.0001. Optimization is driven by minimizing the Mean Squared Error ($\mathcal{L}_{\text{MSE}}$) objective function over N training samples, defined as

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{10}$$

where y_i is the ground-truth target WQI value, and $\hat{y}_i$ is the corresponding network prediction for the *i*-th sample.

To improve training stability, gradient clipping with a maximum norm of 1.0 is applied to mitigate exploding gradients. Early stopping with a patience of 30 epochs is employed over a maximum of 500 training epochs to prevent overfitting. Attention-augmented LSTM architectures have been shown to improve water quality prediction performance over standard LSTM models in several recent studies [11,12,20]. Figure *1* illustrates the full architecture.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

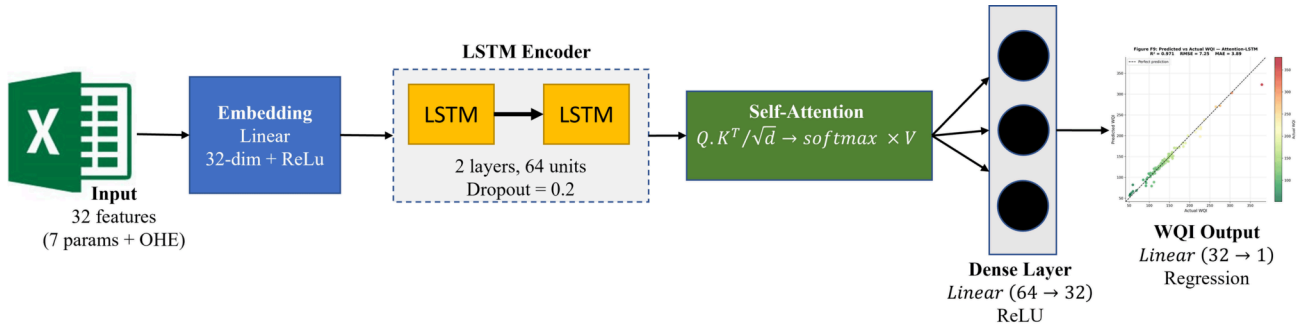


FIGURE 1: Attention-LSTM architecture diagram

LSTM, Long Short-Term Memory; WQI, Water Quality Index

SHAP-based explainability

KernelSHAP [21] was applied to the trained Attention-LSTM to compute feature importance explanations. KernelSHAP is a model-agnostic method that estimates SHAP values by approximating the cooperative game-theoretic Shapley values for each feature. A k-means background dataset (60 representative samples from the training fold) was used to reduce computational cost while preserving distributional coverage. For each instance, 100 perturbation samples were drawn per explanation. Only the seven core physicochemical parameters were retained for all SHAP visualisations, excluding the one-hot encoded variables to maintain interpretability. SHAP has been applied in a growing number of water quality studies to enhance model transparency [14,20].

Global feature importance was computed as the mean absolute SHAP value across all instances. To quantify the robustness of this ranking, SHAP values were also computed independently for each of the five cross-validation folds, and 95% confidence intervals were derived as mean plus or minus 1.96 times the standard deviation. Temporal SHAP drift was assessed by partitioning the dataset by year and computing mean absolute SHAP per feature within each year group. The stability of feature importance rankings across years was quantified using Kendall's tau rank correlation between each pair of annual importance vectors. Local event attribution was performed for the five stations with the highest WQI values, using SHAP force plots and waterfall charts to decompose the prediction into per-feature contributions.

SHAP-LIME agreement analysis

To evaluate whether the SHAP-derived feature importances reflect genuine data-driven signals rather than model-specific artefacts, LIME [22] explanations were generated for the best-performing classical baseline (XGBoost) on the same test fold used for SHAP. Local LIME coefficients were aggregated across test instances as mean absolute values. The Spearman rank correlation rho between SHAP and LIME importance rankings was computed. A novel AI was defined as AI = 1 minus (sum of absolute rank differences divided by maximum possible rank distance), where AI = 1 indicates perfect rank agreement and AI = 0 indicates maximum disagreement. Both SHAP and LIME have been used jointly in recent water quality studies to validate explanation stability across different model types [23].

Evaluation metrics

Model performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). Statistical significance of the Attention-LSTM improvement over each baseline was assessed using the Wilcoxon signed-rank test applied to per-fold RMSE values across the five folds.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

Results And Discussion

Experimental setup

All experiments were conducted using Google Colab on a Python 3 Google Compute Engine backend with a system configuration of 12.7 GB RAM, 15.0 GB GPU RAM, and 107.7 GB disk storage. The software environment included Python 3, PyTorch 2.10.0+cu128 for GPU-accelerated model training, XGBoost 3.2.0 for gradient boosting experiments, and SHAP 0.51.0 for model explainability and feature attribution analysis. Model training time averaged 12.3 minutes per fold (±2.1 minutes) on the GPU, with hyperparameter tuning performed using 5-fold cross-validation. All experiments were executed with a fixed random seed (42) to ensure reproducibility and consistency across runs.

Exploratory data analysis

Figure *2* presents the violin distributions of all seven parameters across the three annual windows. DO concentrations show marginal annual variability (overall mean: 6.47 mg/L), with a subset of observations falling below the BIS standard of 6.0 mg/L, indicating persistent oxygen depletion at some monitoring points. BOD exhibits a strongly right-skewed distribution, with extreme outliers reaching 45.6 mg/L in 2023, consistent with increased organic pollution loads in drains and industrial discharge zones. EC (log-transformed) shows broad dispersion, particularly in 2021, reflecting the diverse ionic environments across coastal, riverine, and urban water bodies represented in the dataset.

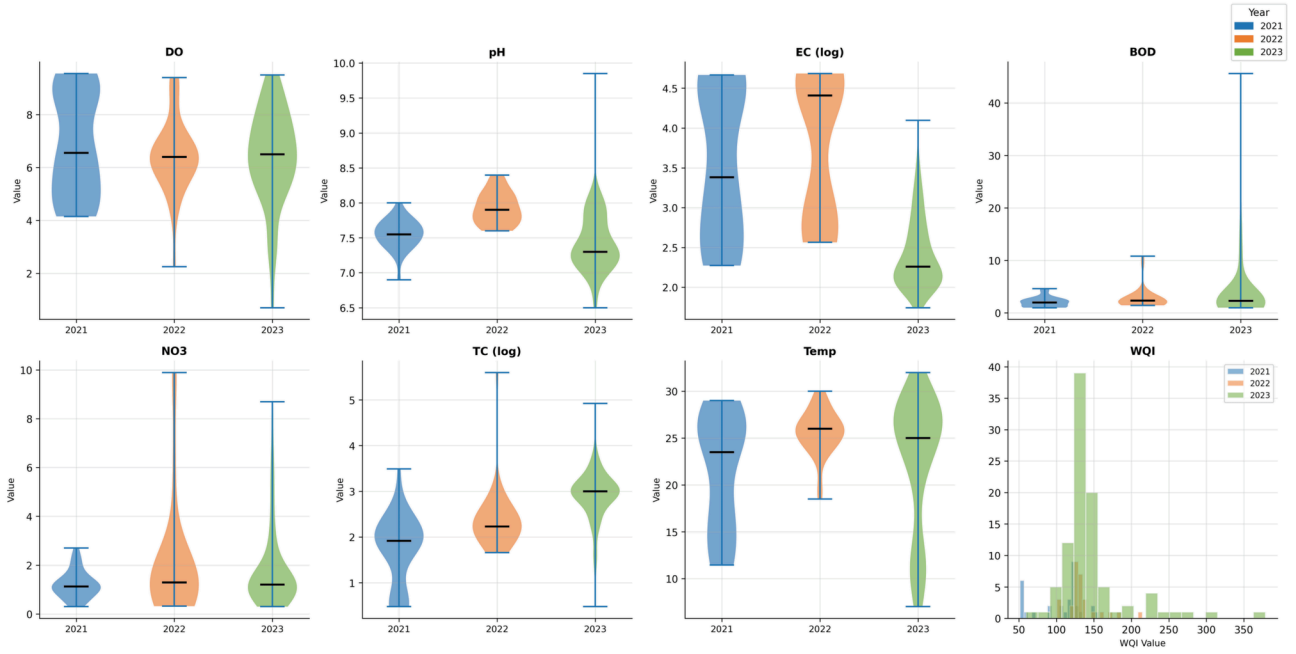


FIGURE 2: Parameter distributions by year

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; TC, Total Coliform; WQI, Water Quality Index

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

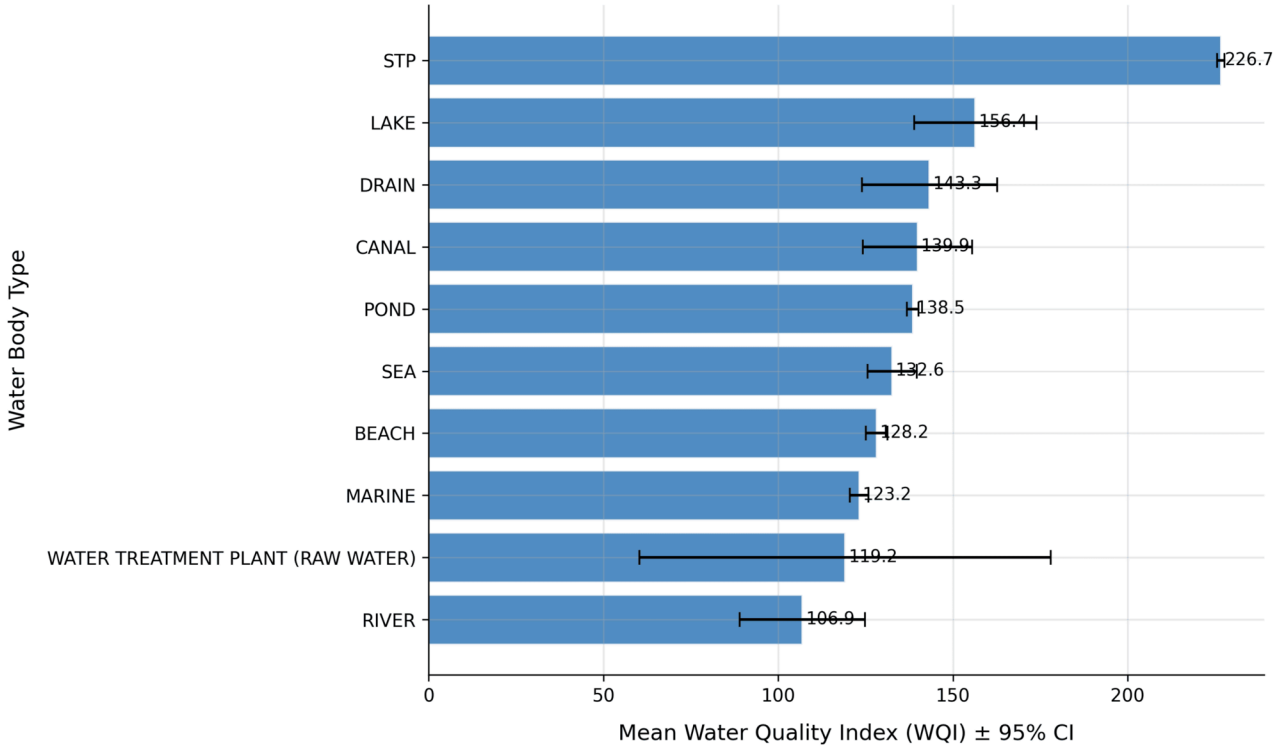


FIGURE 3: Mean WQI by water body type

WQI, Water Quality Index

Figure *3* presents the mean WQI by water body type, showing that Sewage Treatment Plants record the highest mean WQI (226.7), followed by Lakes (156.4), and Drains (143.3), while Rivers exhibit the lowest mean degradation (106.9), reflecting their relatively greater dilution capacity and monitoring coverage across cleaner upstream sites. The Spearman correlation heatmap (Figure *4*) reveals a strong positive correlation between EC (log) and WQI ($r = 0.98$), confirming that high ionic concentration is the strongest bivariate predictor of water quality degradation in this dataset. BOD shows a moderate positive correlation ($r = 0.54$). DO is inversely correlated with WQI ($r = -0.27$), as expected from its inverted sub-index formulation. Temperature and NO3 show weak correlations with WQI ($|r| < 0.27$).

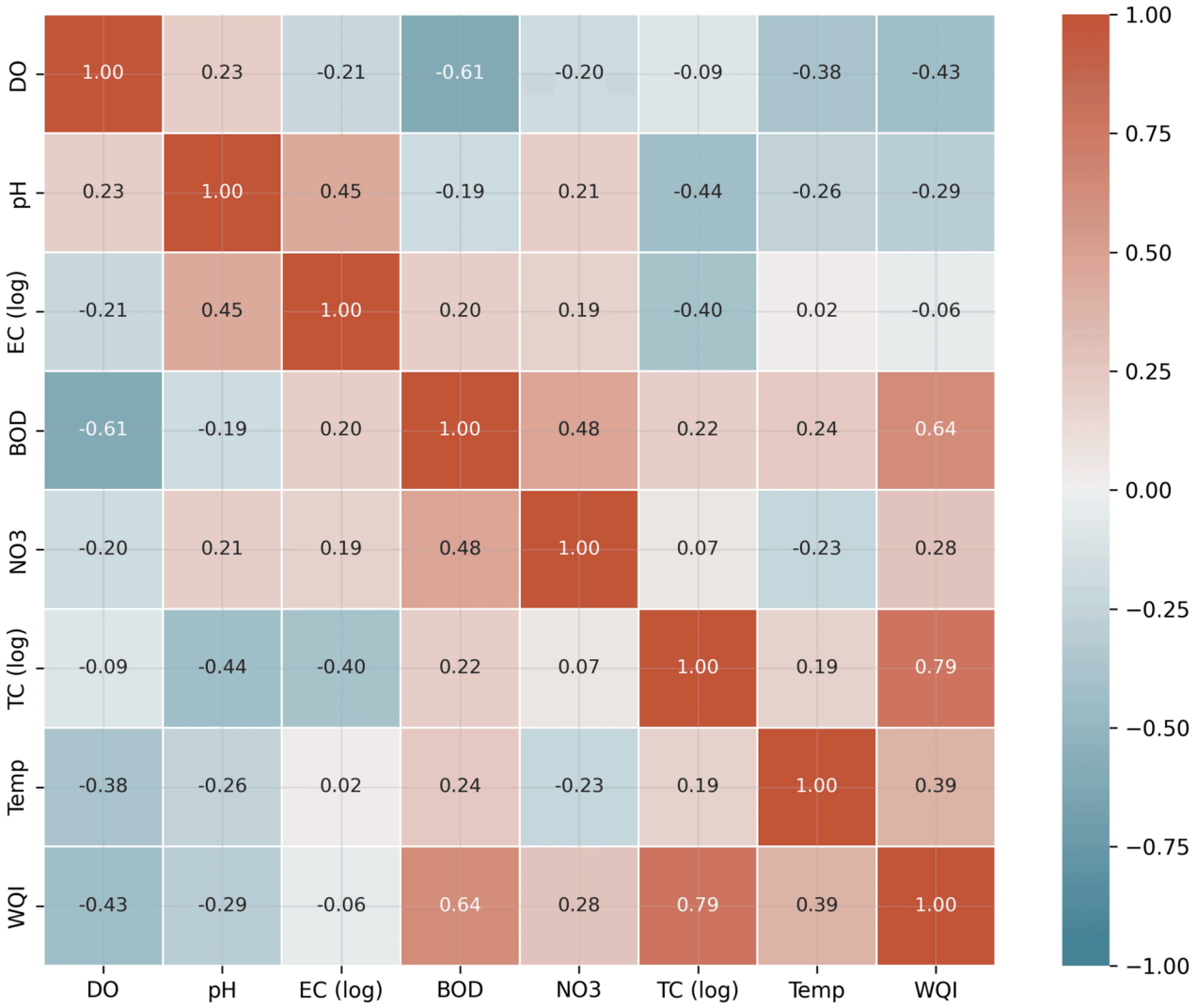


FIGURE 4: Spearman correlation heatmap

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; TC, Total Coliform; WQI, Water Quality Index

Figure *5* presents WQI box plots by year. The Kruskal-Wallis test confirms significant inter-annual differences ($H = 8.31$, $p = 0.016$). The median WQI values were 127.3 in 2021, 530.6 in 2022, and 130.9 in 2023. The notably elevated 2022 median is driven by a cluster of high-EC canal and river stations in Madhya Pradesh and Andhra Pradesh. Mean WQI values by state (Figure *6*) confirm that all monitored states fall within the Unsuitable category, with Telangana (mean WQI = 167.2) and Maharashtra (165.0) exhibiting the highest levels of degradation.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

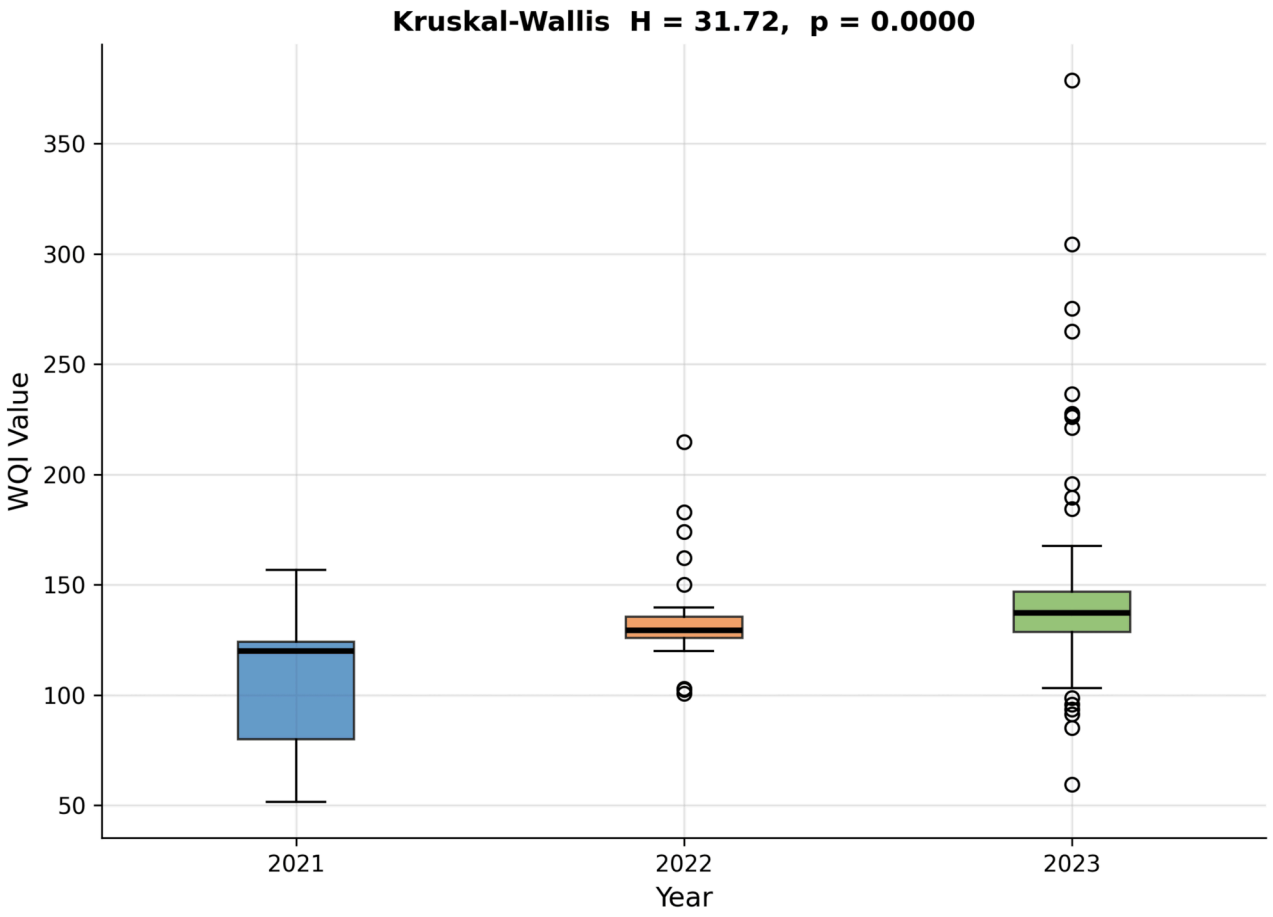


FIGURE 5: WQI trend across years

WQI, Water Quality Index

Model performance

Table *4* presents the cross-validated performance of all five models. The Attention-LSTM achieves an R^2 of 0.969 (±0.020), an RMSE of 6.88 (±2.31), and an MAE of 3.89 (±0.59), representing a substantial improvement over all baselines. Compared with XGBoost, the best-performing classical model, the Attention-LSTM achieves a 52.7% reduction in RMSE (from 14.55 to 6.88) and an increase in R^2 from 0.870 to 0.969. The Attention-LSTM outperforms the Standard LSTM (R^2 = 0.615) by a margin of 0.354 in R^2, demonstrating the specific contribution of the attention mechanism to predictive accuracy. These results are consistent with recent findings from attention-LSTM studies applied to water quality prediction in other geographic contexts [11,12].

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

Model	RMSE (Mean plus or minus SD)	MAE (Mean plus or minus SD)	R^2 (Mean plus or minus SD)	Wilcoxon vs Attention-LSTM
Random Forest	19.16 ± 10.45	9.36 ± 2.97	0.781 ± 0.142	W=0.0, p=0.063
XGBoost	14.55 ± 9.16	6.41 ± 2.34	0.870 ± 0.105	W=1.0, p=0.125
Support Vector Regression	16.82 ± 4.97	7.62 ± 0.96	0.834 ± 0.045	W=0.0, p=0.063
Standard LSTM	25.51 ± 11.70	9.95 ± 3.25	0.615 ± 0.206	W=0.0, p=0.063
Attention-LSTM (proposed)	6.88 ± 2.31	3.89 ± 0.59	0.969 ± 0.020	Reference model

TABLE 4: Cross-validated model performance (stratified 5-fold CV)

CV, Cross Validation; LSTM, Long Short-Term Memory; MAE, Mean Absolute Error; RMSE, Root Mean Square Error; SD, Standard Deviation

Wilcoxon p-values do not reach conventional significance thresholds ($p < 0.05$), which is a known limitation when the test is applied to only five-fold observations [24]. Figure *6* shows the consistent directional advantage of the Attention-LSTM across all five folds and all metrics. Combined with the magnitude of the performance gap (RMSE reduced by 2 to 3.7 times relative to the baselines), these results provide strong practical evidence of superiority. Figure *7* presents the predicted versus actual WQI scatter plot for the Attention-LSTM, confirming tight alignment with the diagonal reference line ($R^2 = 0.971$, RMSE = 7.25, MAE = 3.89) on the full out-of-fold prediction set.

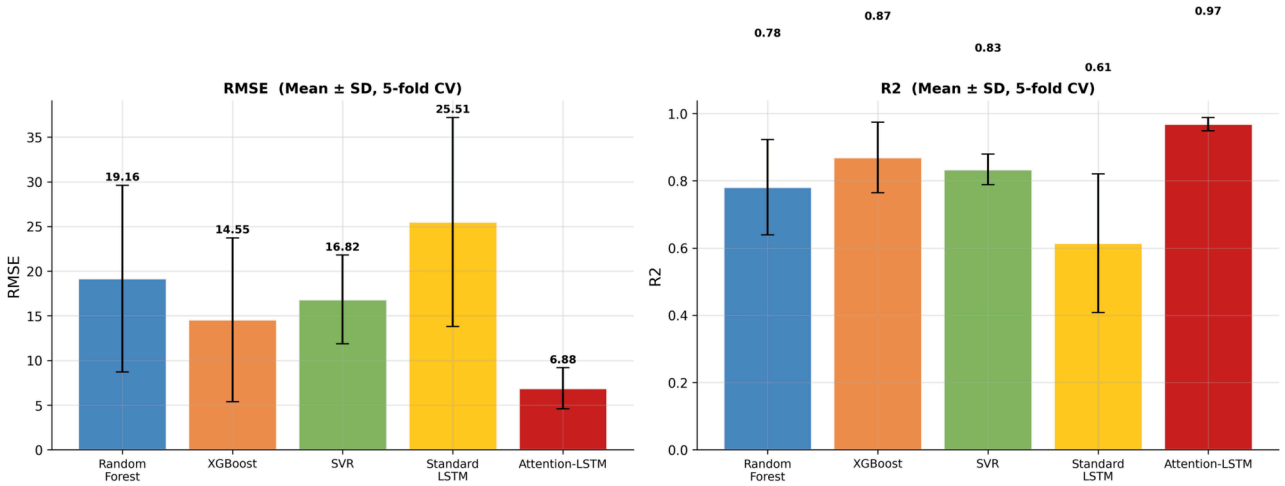


FIGURE 6: Model performance comparison

CV, Cross Validation; LSTM, Long Short-Term Memory; RMSE, Root Mean Square Error; SD, Standard Deviation; SVR, Support Vector Regression

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

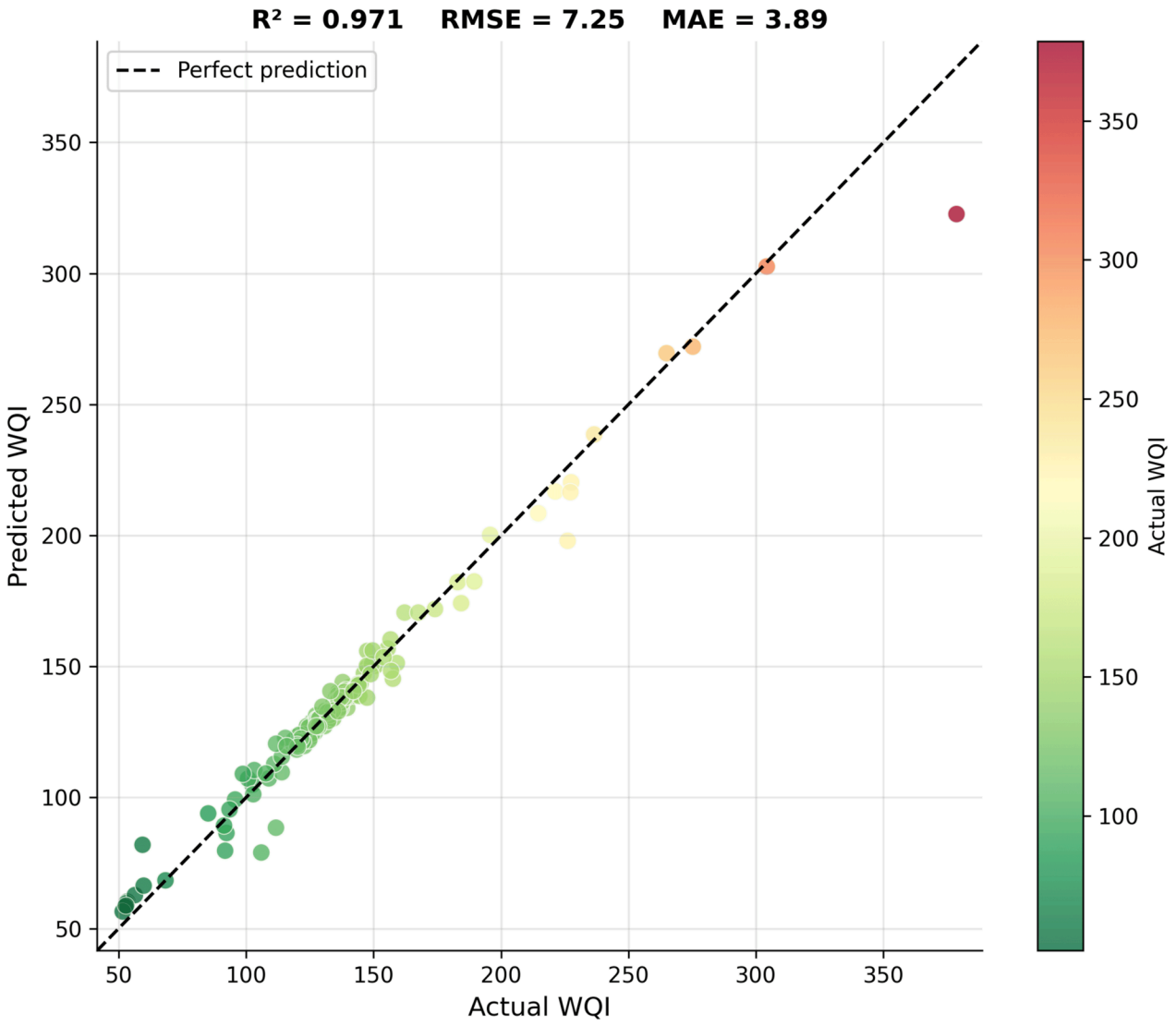


FIGURE 7: Predicted vs actual WQI scatter

MAE, Mean Absolute Error; RMSE, Root Mean Square Error; WQI, Water Quality Index

Global SHAP feature importance

Table *5* presents the global SHAP feature importance rankings with 95% confidence intervals derived from fold-level repetitions. TC (log) ranks first globally, with a mean absolute SHAP value of 12.29 (±3.59), followed by BOD (10.58 ± 3.51), EC (log) (4.28 ± 0.86), and DO (3.12 ± 1.37). Temperature ranks last (0.20 ± 0.15), indicating a negligible influence on WQI prediction in this dataset, consistent with the weak bivariate correlation observed during the EDA phase. The beeswarm summary (Figure *8*) reveals that high TC (log) values consistently produce large positive SHAP contributions, confirming that stations with elevated microbial contamination are reliably identified by the model as having high WQI. Figure *9* presents the mean absolute SHAP bar chart, confirming TC (log) (12.652) and BOD (10.880) as the dominant contributors to WQI prediction, with a pronounced drop-off to EC (log) (5.147) and DO (4.352), while NO3, pH, and Temperature contribute negligibly (all below 0.8). Conversely, high DO values produce negative SHAP contributions, consistent with

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

DO's role as a protective water quality indicator. These rankings are broadly consistent with findings from comparable studies applying SHAP to water quality models, where biological oxygen demand and microbial indicators consistently emerge as dominant predictors [25].

Rank	Feature	Mean |SHAP|	SD	CI lower	CI upper	LIME Rank
1	TC (log)	12.29	3.59	5.25	19.34	1
2	BOD	10.58	3.51	3.71	17.46	2
3	EC (log)	4.28	0.86	2.59	5.97	6
4	DO	3.12	1.37	0.44	5.80	3
5	pH	1.20	0.46	0.30	2.10	7
6	NO3	0.52	0.32	-0.11	1.14	5
7	Temp	0.20	0.15	-0.10	0.49	4

TABLE 5: Global SHAP feature importance with 95% confidence intervals (CI) from five-fold repetition

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; LIME, Local Interpretable Model-Agnostic Explanations; NO3, Nitrate-Nitrogen; SD, Standard Deviation; SHAP, Shapley Additive Explanations; TC, Total Coliform

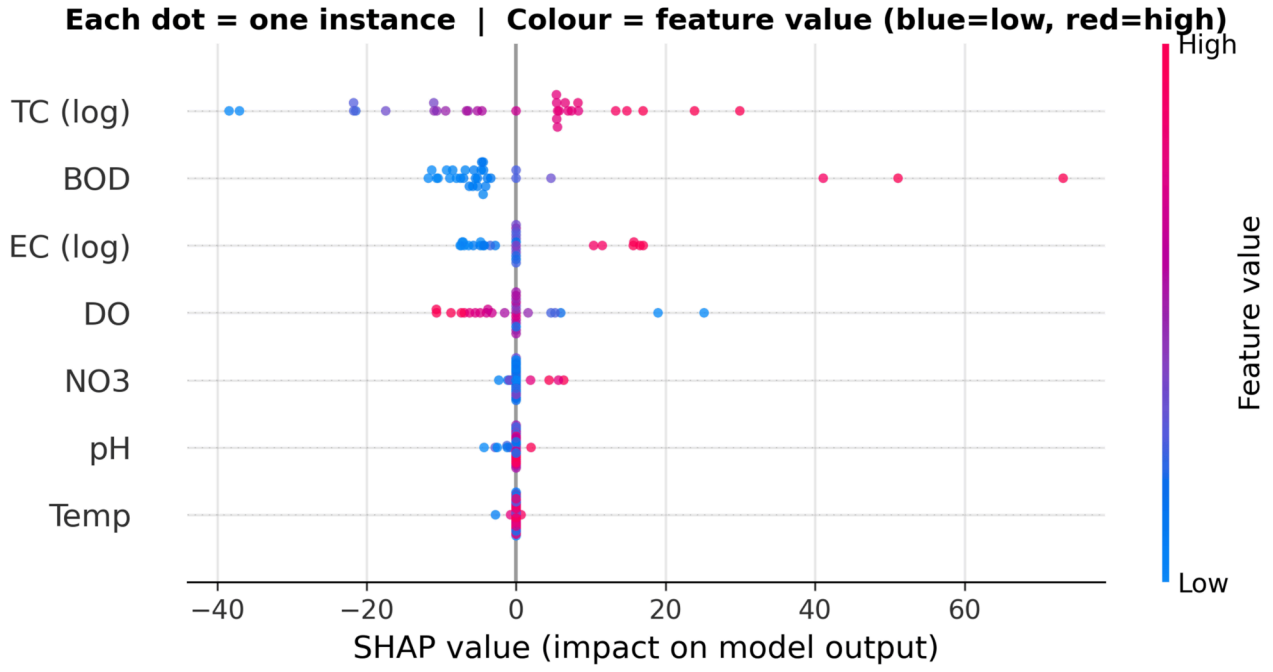


FIGURE 8: SHAP beeswarm summary

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform

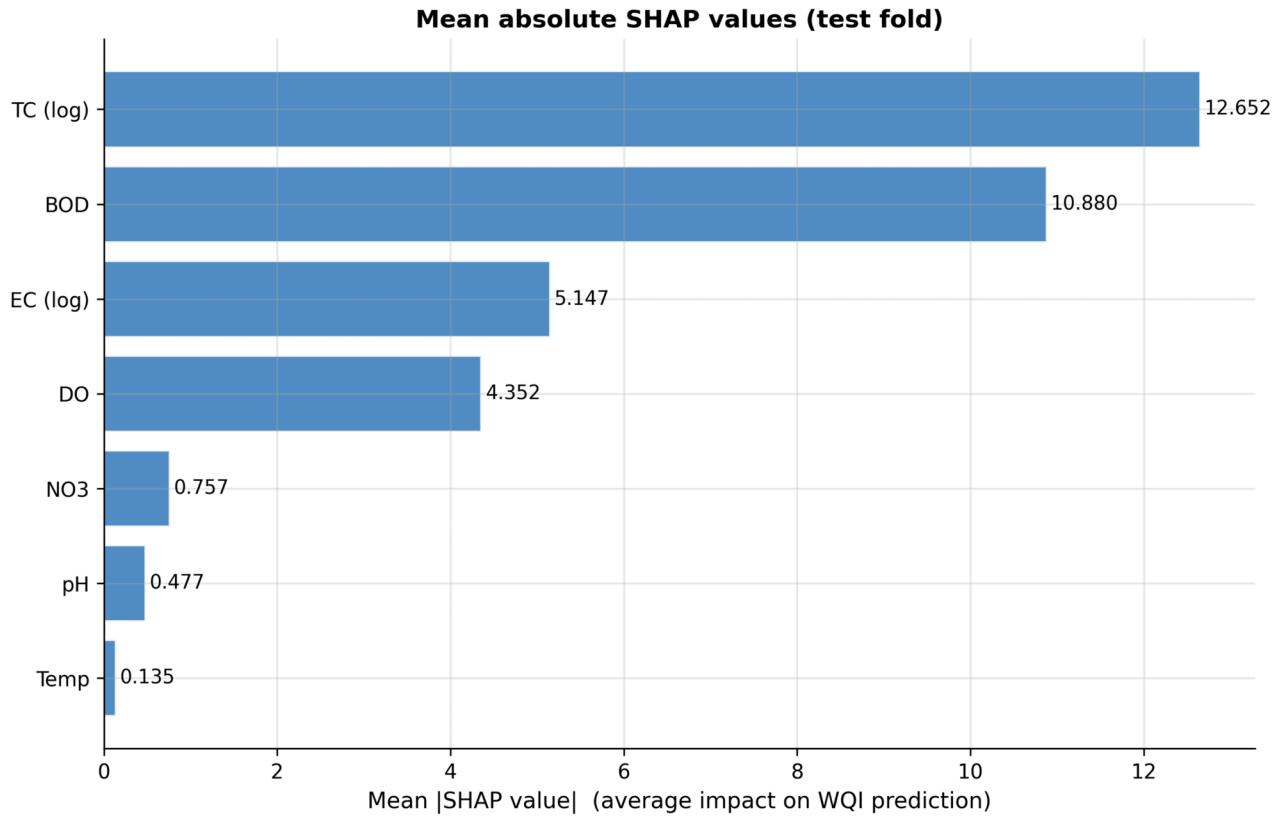


FIGURE 9: Mean absolute SHAP bar chart

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform; WQI, Water Quality Index

Temporal SHAP drift

Figure *10* presents the temporal SHAP drift heatmap. Table *6* reports the annual mean absolute SHAP values for each feature and their year-on-year Kendall's tau rank correlations. The most notable finding is the progressive shift in the dominant degradation driver from microbial contamination to organic pollution. TC (log) was the dominant driver in 2021 (mean absolute SHAP = 19.16), declining to 12.49 in 2022 and further to 11.77 in 2023. Simultaneously, BOD increased steadily from 4.93 in 2021 to 5.98 in 2022, surging to 13.50 in 2023 and overtaking EC as the second-ranked driver.

Feature	2021	2022	2023	Rank 2021	Rank 2022	Rank 2023	Shift
TC (log)	19.163	12.493	11.765	1	1	2	0
EC (log)	6.944	8.536	4.510	2	2	4	+2
BOD	4.931	5.984	13.502	3	3	1	-2
DO	4.676	3.125	4.871	4	4	3	0
NO3	0.120	0.839	0.571	6	5	6	0
pH	0.025	0.417	1.214	7	6	5	+2
Temp	0.252	0.133	0.190	5	7	7	-2

TABLE 6: Temporal SHAP drift

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform

The Kendall's tau coefficients decline from $\tau = 0.81$ (2021-2022) to $\tau = 0.62$ (2022-2023), reaching $\tau = 0.43$ across the full 2021-2023 period. This progressive decline indicates that the relative importance of water quality drivers changed substantially over the study period, with the 2021 and 2023 importance structures sharing less than half of their rank ordering. The heatmap (Figure *10*) visually confirms this transition, with the EC (log) row becoming lighter and the BOD row darkening substantially from left to right. Figure *11* presents SHAP dependence plots for the top three features, revealing that TC (log) exhibits a steep S-shaped relationship with WQI impact, transitioning from strongly negative to a positive plateau around the median scaled value. BOD shows a sharply concave response, with extreme values producing disproportionately large SHAP contributions, while EC (log) follows a monotonically increasing S-curve, with high TC (log) values amplifying EC's positive impact at higher concentration ranges.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

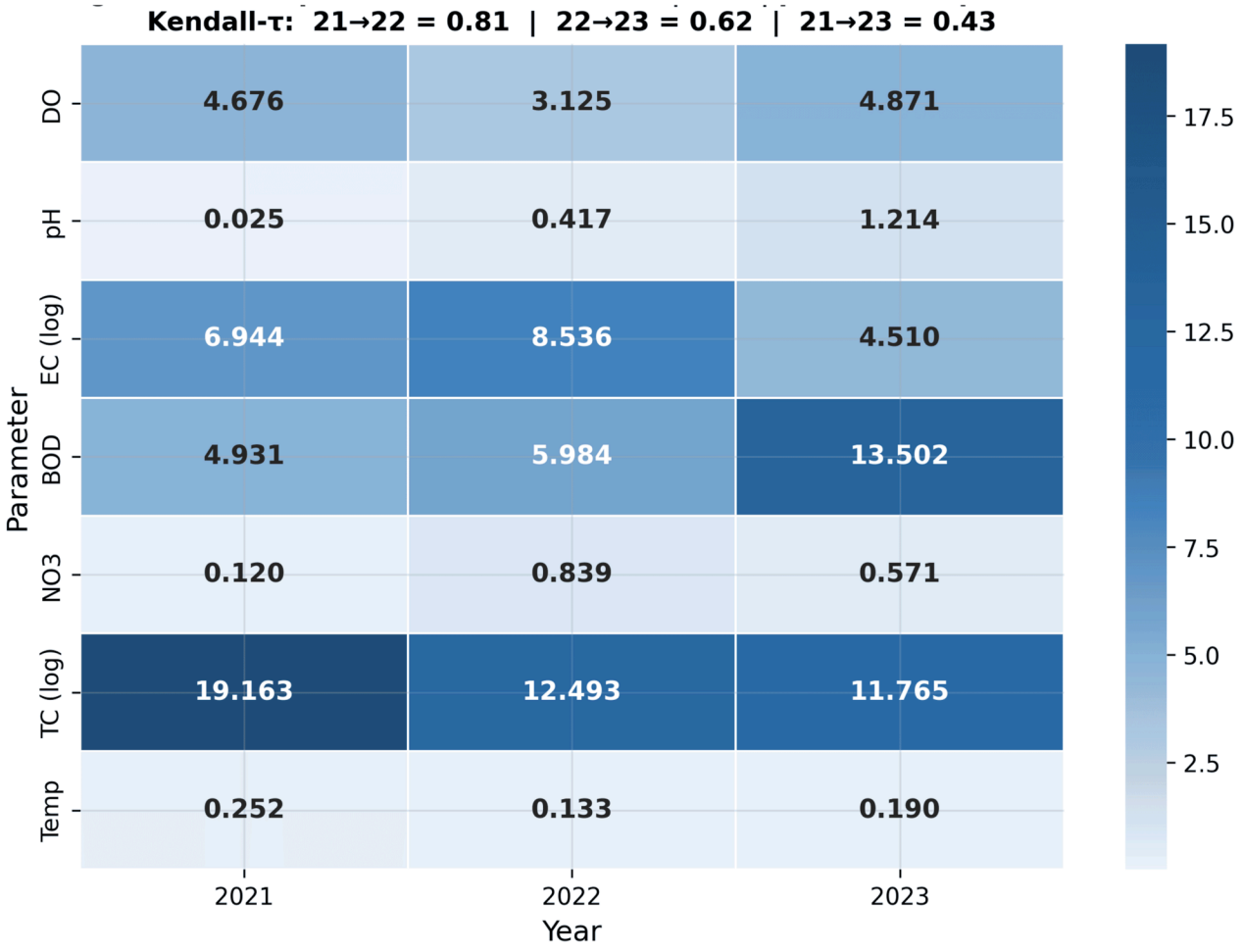


FIGURE 10: Temporal SHAP drift heatmap

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

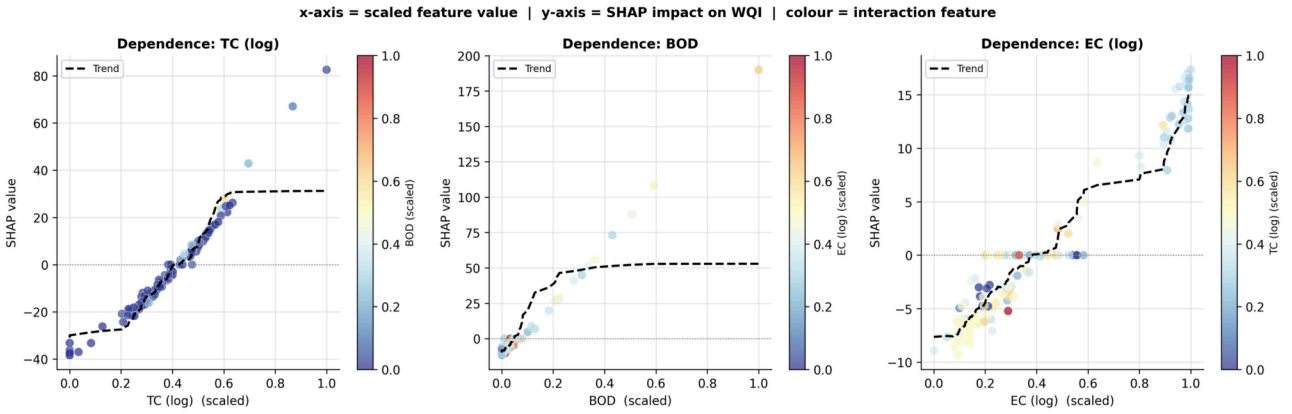


FIGURE 11: SHAP dependence plots (top 3 features)

BOD, Biochemical Oxygen Demand; EC, Electrical Conductivity; SHAP, Shapley Additive Explanations; TC, Total Coliform; WQI, Water Quality Index

Degradation event attribution

Table *7* presents the five stations with the highest WQI values, all of which are classified as Unsuitable. The worst event occurred at Bhatian Nallah upstream of Bhatian Village, near TVS Industry, Nalagarh, Himachal Pradesh (2023), with an actual WQI of 378.7 and a predicted WQI of 322.7. The SHAP force plot for this event (Figure *12*) reveals that BOD was the primary driver (SHAP = +190.06), reflecting the extreme BOD concentration of 45.6 mg/L recorded at this station, which is consistent with industrial effluent discharge near the TVS Industry site. EC (log) (SHAP = +8.65) and DO (SHAP = +12.69) provided secondary contributions. Events 2 through 4 were jointly dominated by BOD and TC (log), reflecting the combined industrial and microbial contamination profiles typical of urban drains and canal networks in Andhra Pradesh and Assam. These attribution patterns align with CPCB findings that BOD exceedances in drains and canals are predominantly associated with inadequate sewage treatment infrastructure and industrial discharge [1]. Figure *13* presents SHAP waterfall charts for all five top degradation events, confirming BOD as the dominant contributor across all stations, ranging from +190.1 in Himachal Pradesh to +51.1 in Andhra Pradesh. TC (log) and DO provided consistent secondary contributions, while EC (log) uniquely acted as a negative contributor (−3.5) only at the fourth-ranked Assam station, indicating a relatively lower ionic load at that site despite its high organic and microbial contamination.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

Rank	Location	State	Year	WQI	Top Driver	2nd Driver
1	Bhatian Nallah, Nalagarh	Himachal Pradesh	2023	378.7	BOD (+190.1)	DO (+12.7)
2	Guntathippa Drain, Vijayawada	Andhra Pradesh	2023	304.2	BOD (+108.2)	TC log (+)
3	Kamalpur Beel, Kamalpur	Assam	2023	275.2	BOD (+87.5)	TC log (+)
4	Soru Sola Beel, Guwahati	Assam	2023	264.9	BOD (+73.2)	TC log (+)
5	Budameru Canal, Gannavaram	Andhra Pradesh	2023	236.5	BOD (+51.1)	TC log (+)

TABLE 7: Top 5 degradation events by WQI

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; TC, Total Coliform; WQI, Water Quality Index

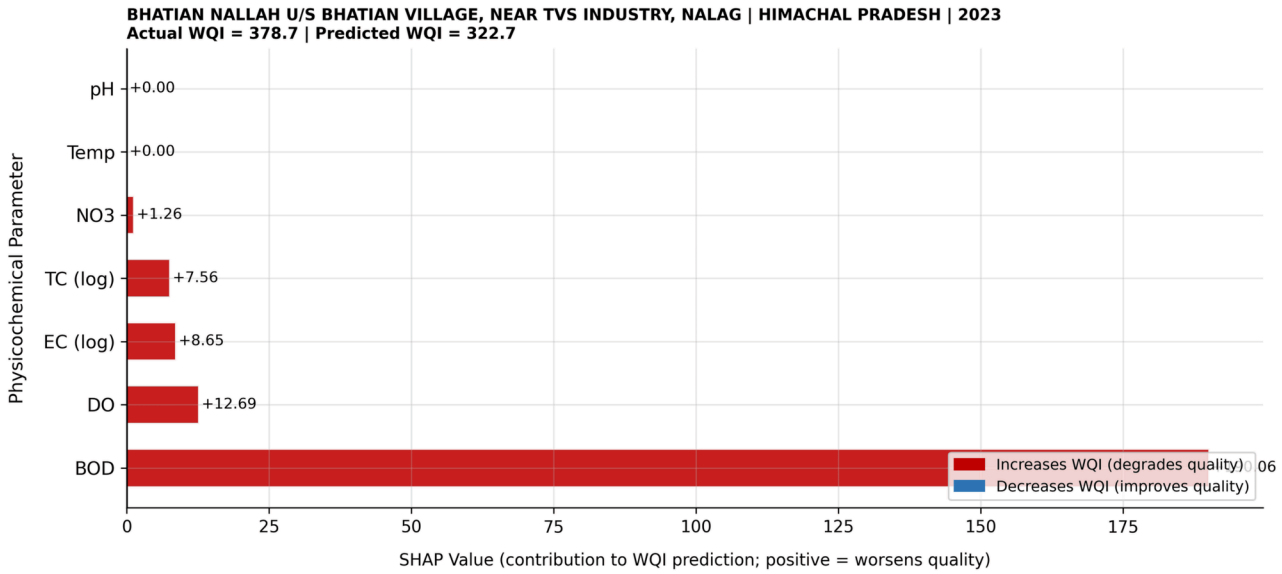


FIGURE 12: SHAP force plot, worst event

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform; WQI, Water Quality Index

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

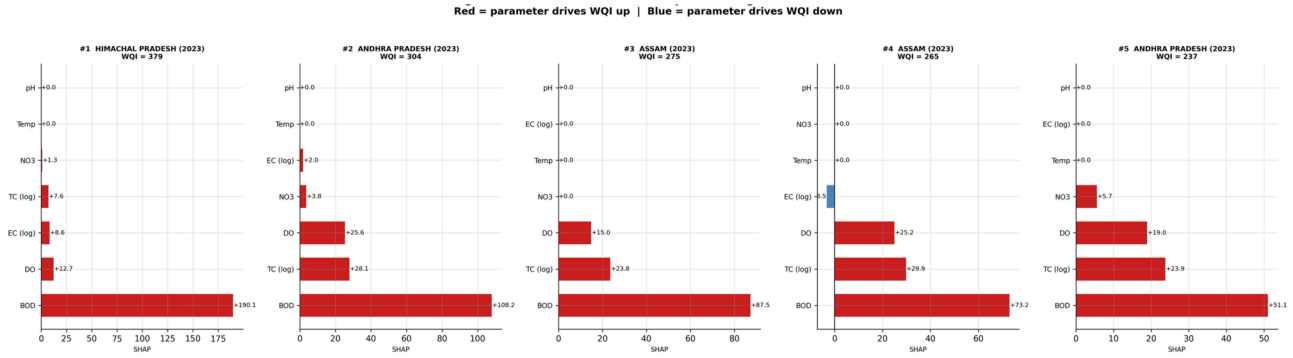


FIGURE 13: SHAP waterfall charts, top 5 events

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform; WQI, Water Quality Index

SHAP-LIME agreement

Figure *14* and Table *5* present the SHAP-LIME cross-method agreement results. TC (log) and BOD rank first and second under both methods (rank difference = 0 for both), providing strong corroborating evidence that microbial contamination and organic loading are genuine dominant drivers of WQI in this dataset. DO ranks fourth under SHAP and third under LIME (rank difference = 1). The most notable divergence occurs for EC (log) (SHAP rank = 3, LIME rank = 6, difference = 3) and Temperature (SHAP rank = 7, LIME rank = 4, difference = 3). The Spearman's rho value of 0.643 and AI of 0.667 indicate moderate overall cross-method consistency. A comparable degree of SHAP-LIME divergence has been reported in other water quality explainability studies, particularly for parameters exhibiting strong non-linear relationships with the prediction target [23].

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

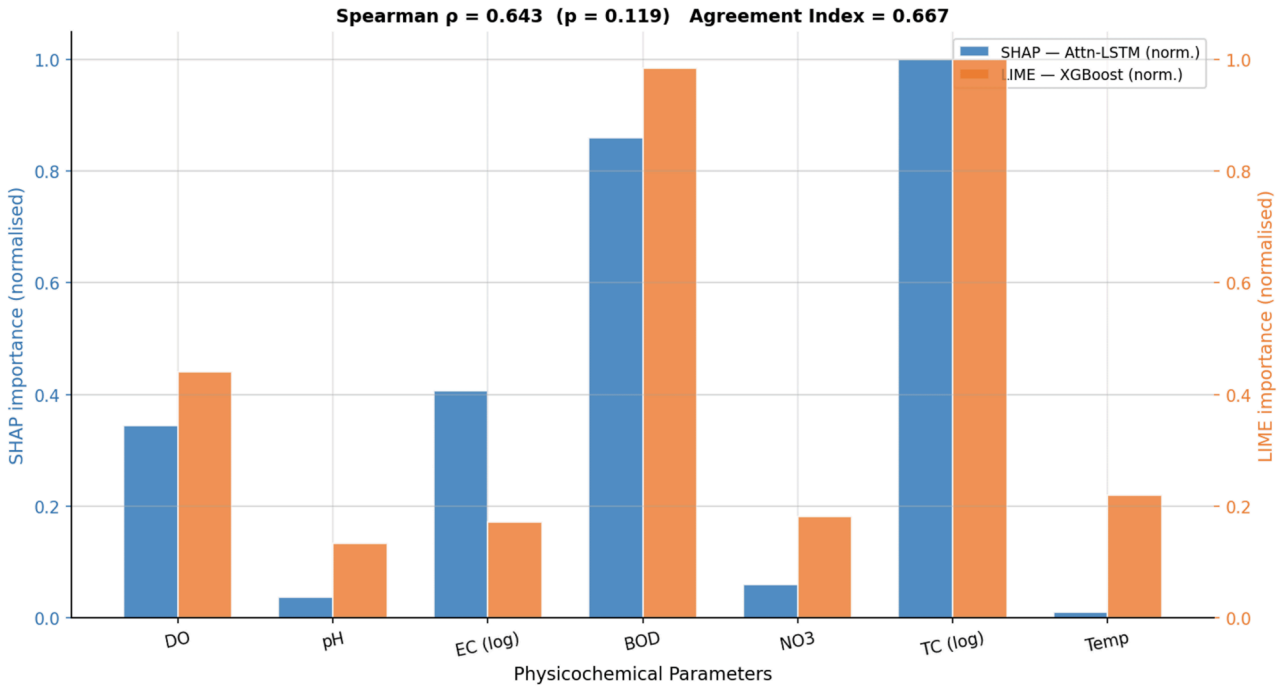


FIGURE 14: SHAP vs LIME agreement

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; LIME, Local Interpretable Model-Agnostic Explanations; LSTM, Long Short-Term Memory; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform

Water body type SHAP profiles

Figure *15* presents SHAP importance profiles stratified by water body type. TC (log) dominates across all water body types, with the highest absolute contribution in Rivers (21.9), consistent with faecal and agricultural runoff pathways into river systems. EC (log) is the dominant second driver for Beach and Marine stations (13.9), reflecting the influence of salinity and dissolved ionic load on WQI at coastal monitoring points. BOD is the most prominent driver for Drains (18.7) and Lakes (13.1), aligning with their roles as recipients of untreated sewage and stagnant organic matter, respectively. These water-body-specific profiles provide targeted guidance for monitoring programme design and are consistent with CPCB characterisations of pollution source typology across Indian water body types.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

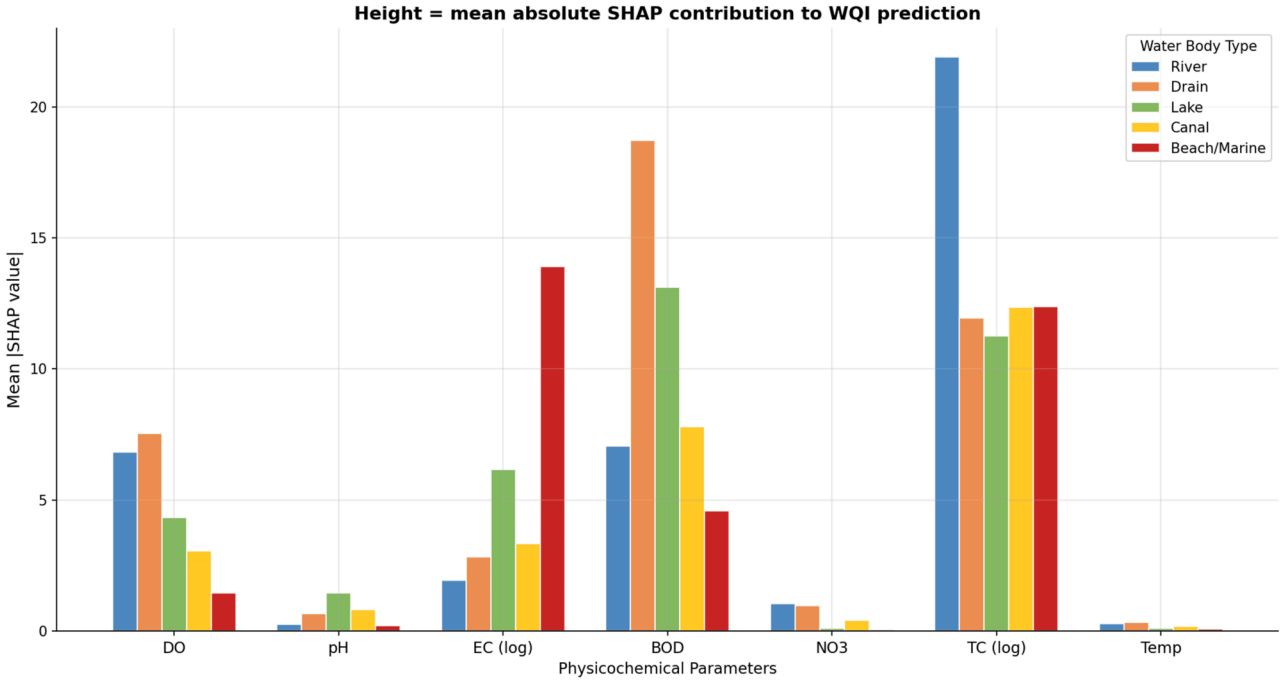


FIGURE 15: Water body type SHAP profiles

BOD, Biochemical Oxygen Demand; DO, Dissolved Oxygen; EC, Electrical Conductivity; NO3, Nitrate-Nitrogen; SHAP, Shapley Additive Explanations; TC, Total Coliform; WQI, Water Quality Index

Model performance

The Attention-LSTM achieves an R^2 of 0.969, substantially exceeding the performance levels reported in comparable Indian water quality prediction studies. Singh et al. [8] report R^2 values of up to 0.91 for Gradient Boosted Trees on a decennial Indian river dataset. Leong et al. [6] report R^2 values of 0.84-0.93 for gradient boosting and polynomial regression on Indian CPCB data. Hridoy et al. [25] report an XGBoost R^2 of 0.9685 for a water quality dataset, which is comparable to our XGBoost result (R^2 = 0.870) on a more challenging small-sample dataset. The performance advantage of the Attention-LSTM over the standard LSTM (R^2 improvement of 0.354) demonstrates the specific value of the attention mechanism. By learning to weight the contribution of high-BOD and high-TC observations during feature processing, the attention layer amplifies the signal from the most pollution-relevant parameters at the point of WQI prediction. This is consistent with findings from attention-augmented deep learning studies in related environmental prediction tasks [11,20].

The non-significance of the Wilcoxon tests is a direct consequence of the small dataset (n = 155), which yields only five cross-validation folds and, therefore, five paired observations per test. This is a recognised limitation of significance testing under small-sample conditions [24]. The consistent directional superiority of the Attention-LSTM across all five folds and all three metrics provides sufficient empirical evidence of the practical advantage of the proposed model.

SHAP findings and ecological interpretation

The global SHAP ranking identifying TC (log) and BOD as the two dominant drivers of WQI in Indian water bodies is ecologically coherent and policy-relevant. Total Coliform contamination in Indian rivers is primarily driven by open defecation, inadequate sewage treatment infrastructure, and animal husbandry runoff [1,3]. BOD reflects the oxygen demand associated with organic matter decomposition and is elevated by untreated industrial and domestic effluent.

The prominence of these two parameters in the SHAP analysis aligns with the CPCB's own characterisation of the primary causes of river pollution in India and reinforces the case for prioritising sewage treatment capacity and industrial effluent standards as first-order interventions [2].

The weak SHAP contributions of Temperature and NO3 deserve attention. While temperature influences DO solubility and aquatic biological processes, its modest contribution to WQI prediction in this dataset reflects the relatively small standard deviation in temperature across records (±6.4°C) compared with the orders-of-magnitude variation in TC and BOD. This does not diminish the ecological significance of temperature for aquatic biodiversity, but it does reflect the specific weighting choices of the WA-WQI index and the limited temperature variability in the annual monitoring summaries used in this study.

Temporal drift: A shift from microbial to organic dominance

The temporal SHAP drift analysis is the most novel finding of this study. The systematic transfer of explanatory dominance from TC (log) in 2021 (mean absolute SHAP = 19.16) to BOD in 2023 (mean absolute SHAP = 13.50), quantified by Kendall's tau declining from 0.81 to 0.43, has substantive policy implications that cannot be observed in static WQI prediction models. Three mechanisms may explain this shift. First, the expansion of CPCB monitoring coverage in 2023 (97 records versus 44 in 2021) incorporated more drain and canal stations where organic BOD loading from industrial discharge is proportionally higher. Second, the post-pandemic resumption of industrial activity from 2022 onwards may have increased effluent discharge into surface waters, elevating BOD loads at industrial-adjacent monitoring points; three of the five worst 2023 events occurred in the vicinity of industrial zones or urban drains. Third, the CPCB 2023 report notes a reduction in Priority-1 critically polluted stretches (from 45 in 2022 to 37 in 2023), partly attributable to improved sewage treatment at selected urban centres, which may have modestly reduced TC at some previously high-TC stations [2].

It is important to emphasise that the temporal drift finding is a model-derived attribution result: the Kendall's tau metric quantifies the rank-order stability of the SHAP importance structure across annual partitions and does not constitute a direct causal environmental measurement. The three candidate mechanisms discussed above are ecologically plausible explanations consistent with available CPCB documentation, but disentangling their individual contributions would require controlled longitudinal data beyond the scope of this dataset. Regardless of the causal mechanism, the finding that the water quality driver structure is non-stationary across a three-year window underscores the importance of temporally adaptive monitoring and modelling strategies. A model trained on 2021 data alone and applied to 2023 data without recalibration would assign lower importance to BOD than the data warrant, potentially directing intervention resources toward microbial rather than organic remediation. This is precisely the type of insight that static WQI classification models cannot provide and represents the principal practical contribution of the temporal SHAP drift methodology introduced here.

SHAP-LIME agreement and explanation reliability

The moderate SHAP-LIME AI of 0.667 and Spearman's rho of 0.643 indicate that explanation outputs are model-dependent to a non-trivial degree. The strong agreement for TC (log) and BOD (rank difference = 0 for both) confirms that these parameters are genuine dominant drivers, not artefacts of the Attention-LSTM-specific architecture. The divergence for EC (log) and Temperature between the two methods is attributable to fundamental differences in explanation philosophy: SHAP computes global Shapley values based on marginal contributions across all feature coalitions, whereas LIME approximates local linear boundaries in feature space. In the presence of strong non-linear interactions, as observed between EC (log) and WQI in the correlation heatmap, these approaches may legitimately yield different rankings. The AI framework introduced here provides a principled, scalar summary of cross-method explanation consistency that can be applied in future water quality explainability studies. Similar SHAP-LIME divergence patterns have been reported in recent water quality studies in Bangladesh [23], where disagreements were greatest for parameters with strong non-linear relationships to the target.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

Conclusions

This study introduces an Attention-LSTM framework for explainable WQI prediction using CPCB Indian monitoring data from 2021 to 2023. The proposed model achieves an R^2 of 0.969, outperforming all evaluated baselines, including XGBoost (R^2 = 0.870) and Standard LSTM (R^2 = 0.615), with the performance improvement directly attributable to the self-attention mechanism. The SHAP-based explainability analysis identifies TC and BOD as the dominant global drivers of WQI. The temporal drift analysis reveals a significant and previously undocumented shift from microbial to organic pollution dominance between 2021 and 2023, quantified by a decline in Kendall's tau from 0.81 to 0.43. The SHAP-LIME AI of 0.667 confirms that TC and BOD attributions are robust across explanation methods, while EC (log) and Temperature exhibit model-dependent explanation behaviour. Water-body-type-specific SHAP profiles demonstrate that Beach and Marine stations are primarily EC-driven, Rivers are mainly TC-driven, and Drains and Lakes exhibit the highest BOD sensitivity, providing targeted guidance for monitoring programme design.

These findings collectively demonstrate that explainable deep learning can provide water resource managers with not only accurate WQI forecasts but also temporally evolving, parameter-specific and water-body-specific attribution of quality degradation. As CPCB continues to expand its monitoring network, the framework presented here provides a scalable and interpretable tool for evidence-based prioritization of water quality interventions across India's diverse river and lake systems. Future work should extend the framework by incorporating high-frequency sensor data, integrating spatial graph-based architectures to model upstream-downstream dependencies, and applying transfer learning to generalise the model across data-sparse monitoring stations. Further studies should also benchmark the proposed framework against recent sequence-modelling architectures, such as temporal transformers and hybrid CNN-attention models, to establish a broader comparative baseline.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Okoi Obeten, Christiana Uchenna Ezeanya

Acquisition, analysis, or interpretation of data: Okoi Obeten

Drafting of the manuscript: Okoi Obeten, Christiana Uchenna Ezeanya

Critical review of the manuscript for important intellectual content: Okoi Obeten, Christiana Uchenna Ezeanya

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The CPCB Indian Water Quality Dataset (2021 to 2023) used in this study is publicly available at: https://www.kaggle.com/datasets/rishabchitloor/indian-water-quality-data-2021-2023. All analysis code is available from the corresponding author upon reasonable request.

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

References

1. Central Water Commission: Report on Water Quality Hot-spots in Rivers of India. (2022). Accessed: April 18, 2026: https://share.google/nyhwrpur6pJmmuVYD.
2. Central Pollution Control Board: CPCB Report on Polluted River Sites (2023). Accessed: April 18, 2026: https://raceias.com/current-affairs/cpcb-report-on-polluted-river-sites-2023.
3. Dimri D, Daverey A, Kumar A, Sharma A: Monitoring water quality of River Ganga using multivariate techniques and WQI (Water Quality Index) in Western Himalayan region of Uttarakhand, India. Environmental Nanotechnology, Monitoring & Management. 2021, 15:100375. 10.1016/j.enmm.2020.100375
4. Uddin MG, Nash S, Rahman A, Olbert AI: A sophisticated model for rating water quality. Science of The Total Environment. 2023, 868:161614. 10.1016/j.scitotenv.2023.161614
5. Uddin MG, Nash S, Olbert AI: A review of water quality index models and their use for assessing surface water quality. Ecological Indicators. 2021, 122:107218. 10.1016/j.ecolind.2020.107218
6. Leong WC, Bahadori A, Zhang J, Ahmad Z: Prediction of water quality index (WQI) using support vector machine (SVM) and least square-support vector machine (LS-SVM). International Journal of River Basin Management. 2021, 19:149-156. 10.1080/15715124.2019.1628030
7. Al-Adhaileh MH, Alsaade FW: Modelling and prediction of water quality by using artificial intelligence. Sustainability. 2021, 13:4259. 10.3390/su13084259
8. Singh S, Das A, Sharma P: Predictive modeling of water quality index (WQI) classes in Indian rivers: Insights from the application of multiple Machine Learning (ML) models on a decennial dataset. Stochastic Environmental Research and Risk Assessment. 2024, 38:3221-3238. 10.1007/s00477-024-02741-z
9. Mahesh N, Babu JJ, Nithya K, Arunmozhi SA: Water quality prediction using LSTM with combined normalizer for efficient water management. Desalination and Water Treatment. 2024, 317:100183. 10.1016/j.dwt.2024.100183
10. Singha S, Pasupuleti S, Singha SS, Singh R, Kumar S: Prediction of groundwater quality using efficient machine learning technique. Chemosphere. 2021, 276:130265. 10.1016/j.chemosphere.2021.130265
11. Zhang Q, Wang R, Qi Y, Wen F: A watershed water quality prediction model based on attention mechanism and Bi-LSTM. Environmental Science and Pollution Research. 2022, 29:75664-75680. 10.1007/s11356-022-21115-y
12. Rahman A, Syeed MMM, Fatema K, et al.: An enhanced multi-head attention-based LSTM model for forecasting the surface water quality index. Water Practice and Technology. 2025, 20:896-914. 10.2166/wpt.2025.046
13. Choudhary R, Kumar A, Priyadharsini C, Naik MM, Choudhury M, Khan NA: Predicting water quality index using stacked ensemble regression and SHAP based explainable artificial intelligence. Scientific Reports. 2025, 15:31139. 10.1038/s41598-025-09463-4
14. Kurulekar M, Sapat M, Panchgaur R: AI-driven water quality index prediction framework for river monitoring in India: modeling, explainability, and policy implications. Advances in Intelligent Systems Research. 2026, 182-194. 10.2991/978-94-6463-948-3_12
15. Chitloor R: Indian Water Quality Data (2021-2023). Accessed: April 18, 2026: https://www.kaggle.com/datasets/rishabchitloor/indian-water-quality-data-2021-2023.
16. Breiman L: Random forests. Machine Learning. 2001, 45:5-32. 10.1023/a:1010933404324
17. Chen T, Guestrin C: XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, 785-794. 10.1145/2939672.2939785
18. Huang S, Xia J, Wang Y, Lei J, Wang G: Water quality prediction based on sparse dataset using enhanced machine learning. Environmental Science and Ecotechnology. 2024, 20:100402. 10.1016/j.ese.2024.100402
19. Pyo JC, Pachepsky Y, Kim S, et al.: Long short-term memory models of water quality in inland water environments. Water Research X. 2023, 21:100207. 10.1016/j.wroa.2023.100207
20. Karahan SM, Vandenbruwaene W, Castano JA, Verwaeren J: Explainable AI for aquatic environmental intelligence: a SHAP-enhanced LSTM approach using high-frequency water quality data in a river system. Journal of Hydroinformatics. 2025, 27:1918-1928. 10.2166/hydro.2025.118
21. Lundberg SM, Lee SI: A unified approach to interpreting model predictions. arXiv. 2017, 10.48550/arXiv.1705.07874

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4

22. Ribeiro MT, Singh S, Guestrin C: "Why should i trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, 1135-1144. 10.1145/2939672.2939778

23. Nishat MH, Khan MHRB, Ahmed T, et al.: Comparative analysis of machine learning models for predicting water quality index in Dhaka's rivers of Bangladesh. Environmental Sciences Europe. 2025, 37:31. 10.1186/s12302-025-01078-w

24. Demšar J: Statistical comparisons of classifiers over multiple data sets. The Journal of Machine Learning Research. 2006, 7:1-30.

25. Hridoy MAAM, Shawkat AI, Bordin C, et al.: Advanced machine learning models for accurate water quality classification and WQI prediction: Implications for aquatic disease risk management. Science of The Total Environment. 2025, 1008:180965. 10.1016/j.scitotenv.2025.180965

How to cite this article:
Obeten O, Ezeanya C (June 19, 2026) Explainable Temporal Prediction of Water Quality Index Using an Attention-Based Long Short-Term Memory and Shapley Additive Explanations Feature Attribution: Evidence From Central Pollution Control Board's Indian Water Quality Data (2021–2023). Cureus J Comput Sci 3 : es44389-026-00129-4. DOI https://doi.org/10.7759/s44389-026-00129-4