CPSM | Conference Proceedings in Science and Management

Proceedings of the 2nd International Engineering Data Analytics and Management Conference (EAMCON 2026), 8 to 10 July 2026, Shinawatra University, Thailand

Published: October 2026
DOI: https://doi.org/10.68337/cpsm.v1.i2.2026-044

Vol. 1, Issue 2 | Proceeding Paper

River Water Quality Prediction Using a Temporal Convolutional Network

Selvin Retna Raj Thavasimuthu[1,*], P. V. Kumar[2], Ramarao Nayapamu[3], M. Srikanth Sagar[2], Illa Mahesh Kumar Swamy[2], Manyam Thaile[2], G. Vasavi[4], Medikonda Asha Kiran[2], Ramesh Babu Pittala[5], and D. Anupama[2]

[1]*Sreenidhi University, Hyderabad, Telangana, India*

[2]*School of Engineering, Anurag University, Hyderabad, Telangana, India*

[3]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (Deemed to be University), Vaddeswaram, Andhra Pradesh, India*

[4]*MLR Institute of Technology, Hyderabad, Telangana, India*

[5]*C.R. Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad, Telangana, India*

[*]Corresponding author: selvin.r@suh.edu.in

Abstract. This paper presents a temporal convolutional network (TCN) method for predicting river water quality parameters. Unlike the recurrent neural network models used previously, the TCN applies causal and dilated one-dimensional convolutions, which capture long-range dependencies in time series without the vanishing gradient problem. The paper targets river water quality parameters such as pH, dissolved oxygen (DO), turbidity, and biochemical oxygen demand (BOD). Water quality management is essential to maintain ecological balance and to keep water safe for human consumption and agriculture. Pollution and various environmental interactions change water quality continuously, and existing methods lack the capacity to monitor and predict these changes dynamically. The TCN model was developed for this monitoring and prediction task. Experiments showed that the proposed TCN model was efficient, accurate, and stable and performed better than the long short-term memory (LSTM) model used for comparison, with an R^2 score of 0.9027 against 0.8923. Its ability to handle sequential data adds to its effectiveness, so the model can be used in practice and extended in future research.

Keywords. Water quality prediction; temporal convolutional network; time-series data; environmental monitoring

I. Introduction

Water is one of the core natural resources that sustain life on Earth. Surface water bodies such as rivers and lakes are the primary source of freshwater. Their quality directly affects the survival of aquatic organisms and human health, and safeguarding it is an environmental challenge shared by every country [1].

In the past few decades, rapid urbanization, industrialization, population growth, and agricultural development have increased the pressure on freshwater resources. The release of untreated or partly treated industrial effluents, domestic sewage, and agricultural runoff has brought pollutants such as organic matter, nutrients, and suspended solids into water bodies [2]. These pollutants affect key water quality parameters such as pH, dissolved oxygen, turbidity, biochemical oxygen demand, and temperature. Uncontrolled variation in these parameters can lead to eutrophication, oxygen depletion, loss of aquatic biodiversity, and water that is less fit for human use [3].

Traditional statistical methods cannot handle the nonlinear nature of water quality data. Water quality variables depend on many environmental factors and vary strongly over time, which makes them difficult to predict with traditional methods [4]. To overcome these limitations, data-driven and machine learning approaches are increasingly used for water quality analysis. A recent review explains how advances in computational methods and the availability of large-scale environmental data have made it possible to apply deep learning models to water quality management [5]. Deep learning can learn complex patterns and correlations from data automatically and can meet the nonlinear, time-dependent modeling requirements of water environments [6]. Four architectures, the convolutional neural network (CNN), the recurrent neural network (RNN), long short-term memory (LSTM), and the gated recurrent unit (GRU), have been used to predict core parameters including dissolved oxygen, temperature, and the water quality index [7]. Reviews of the field point out, however, that recurrent time-series models have high computational complexity, train slowly, and can be unstable on long time series [8].

II. Proposed Framework

This study developed a deep learning model for river water quality prediction that uses time-series data to predict five parameters: pH, dissolved oxygen, biochemical oxygen

Copyright © 2026 The Authors. Published by SHORAM Editorial Services, India, under the Creative Commons Attribution 4.0 License (CC BY 4.0).

demand, turbidity, and the water quality index (WQI). These parameters can be used to judge the overall quality of a river and its suitability for water use. Accurate prediction is necessary because of the growing water pollution caused by industrial and environmental factors. The proposed system is based on a temporal convolutional network (TCN) model [9], which captures temporal dependencies in water quality data. Unlike recurrent models such as the RNN and LSTM, the TCN uses causal and dilated convolutions, which learn long-term dependencies at low computational cost. The TCN model was trained on the preprocessed multivariate time-series dataset with a sliding window. Its performance was evaluated with the R^2 score, root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) [10] and compared with that of an LSTM model. The trained TCN model [11] can then predict future water quality trends and indicate whether river water is likely to become polluted (Fig. 1).

The proposed system is intended as a preventive tool that supports proactive decisions on environmental protection, in line with earlier water quality prediction studies [12], [13]. Using the trained TCN model, it extracts patterns from historical water quality data to forecast future water quality conditions and pollution risks. It helps environmental monitoring agencies identify water quality degradation and manage water resources efficiently. The system is built as a sequence of data collection, water quality prediction, and analysis.

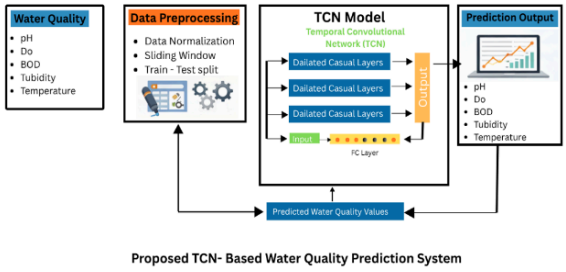

Fig. 1. Proposed TCN-based water quality prediction system.

III. Mathematical Formulation

The water quality monitoring dataset used in this study contains time series of physical and chemical parameters, including pH, DO, BOD, turbidity, and temperature, all of which are essential indicators for water quality assessment. The time-series dataset is represented as

$$X = \{x_1, x_2, x_3, \ldots, x_t\}, \tag{1}$$

where x_t is the observation at time step t. These parameters are key indicators of the level of river water quality and of its fitness for use. Each observation is defined as

$$x_t = [\mathrm{pH}_t, \mathrm{DO}_t, \mathrm{BOD}_t, \mathrm{Turbidity}_t, \mathrm{Temperature}_t]. \tag{2}$$

The objective of the proposed model is to predict future water quality values from historical data, with the TCN model acting as a function $f(\cdot)$ of the k previous time steps. To train the model, the input data are structured as sequences of past observations. The input sequence is defined as

$$X_t = [x_{t-k}, x_{t-k+1}, \ldots, x_t], \tag{3}$$

and the corresponding target output is

$$Y_t = x_{t+1}. \tag{4}$$

IV. Performance Analysis

A. Performance Metrics

The performance of the proposed TCN model was measured with four regression metrics, the R^2 score, RMSE, MAE, and MAPE, which together give a fuller view of prediction accuracy.

Table 1. Performance Comparison

Metric	TCN	LSTM	Improvement
R^2 score	0.9027	0.8923	+0.0104
RMSE	7.9338	8.3494	4.98%
MAE	5.8587	6.2985	6.98%
MAPE	16.9320	18.8209	10.04%

Table 1 compares the proposed TCN and the LSTM model on the R^2 score, RMSE, MAE, and MAPE, metrics widely used to assess prediction models. The TCN had the higher R^2 score, 0.9027 against 0.8923 for the LSTM, which shows that it predicted better than the LSTM. The second metric is the RMSE. The TCN had the lower RMSE, 7.9338, so it reduced errors more than the LSTM did. The third metric is the MAE, and the TCN again had the lower value, which shows that its predictions deviated less from the actual data. The last metric is the MAPE, which was also lower for the TCN than for the LSTM.

B. Training Convergence and Loss Analysis

Fig. 2 shows the training of the TCN and LSTM models over time in terms of the Huber loss, which is not affected by outliers in the water quality dataset. The TCN converged quickly: its training loss started at 1,200 at epoch 1 and fell to well below 100 by epoch 20. The training loss and the validation loss (dashed blue line) stayed close, with no sign of overfitting. TCN training converged at epochs 40 to 50. The early stopping callback ended training at epoch 52, after 20 epochs without improvement in the validation loss. The final validation loss of 48.5 was the best reached by the model and was saved automatically by the model checkpoint callback.

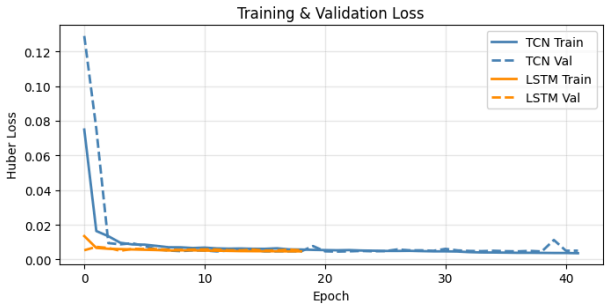

Fig. 2. Training and validation Huber loss of the TCN and LSTM models.

The LSTM network converged faster and reached its optimal point sooner, at epoch 15. Its convergence rate was also good, but its loss never fell below the TCN validation loss during training. Four observations follow. Training and validation losses behaved similarly and did not diverge, so the

models did not overfit. The losses were smooth, without sharp drops, which suggests stable convergence under the Adam optimizer with gradient clipping. Both networks approached their best results within 15 to 50 epochs. The smaller loss values of the TCN show that it captured changes in water quality better.

C. MAE Progression

This study uses the MAE as the main interpretable indicator. It can be read alongside the WQI and complements the Huber loss used for model optimization. Fig. 3 shows that the validation MAE of the TCN (blue) decreased from an initial value of 25 to 5.86. The training MAE (solid line) and the validation MAE (dashed line) converged in parallel, with a small gap of 1 to 2 points. This small gap suggests that the model learned general patterns rather than memorizing the training samples. The MAE curve of the LSTM (orange) is very similar, but its validation MAE is somewhat higher, at 6.30.

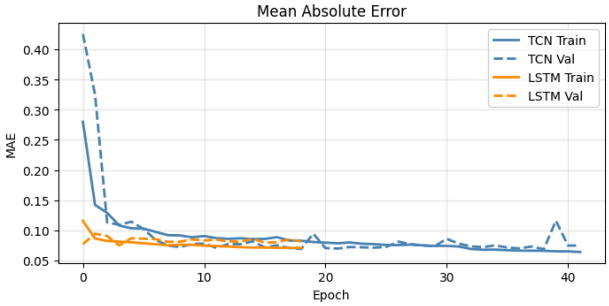

Fig. 3. Training and validation MAE of the TCN and LSTM models.

The difference of 0.44 points between the two models may seem small. In water quality classification, though, where WQI thresholds define safety categories (safe up to 500, acceptable from 500 to 2,000, and unsafe above 2,000), such an improvement means fewer classification errors and more accurate alerts. On average, the prediction error is about 6 WQI points for the TCN and 6.3 WQI points for the LSTM, an improvement of 7%. The rapid fall in MAE during the first 10 to 15 epochs shows how quickly both deep learning models learn the dominant trends in the water quality data. The slow but consistent improvement of the TCN from epoch 15 to 50 indicates that it learned the underlying time dependencies through dilated convolutions reaching up to 63 steps back. The stable MAE in the later epochs indicates that early stopping found a suitable epoch and avoided wasted computation.

V. Conclusion and Future Work

This paper proposed a deep learning method for predicting river water quality and compared a TCN with an LSTM model. The TCN learned temporal features from water quality time series through causal and dilated convolutions, and key water quality indicators, including pH, DO, turbidity, and BOD, were predicted well. The experiments showed high predictive power for the proposed TCN model, with an R^2 score of about 0.90, and lower RMSE and MAE than the LSTM model. The error distribution analysis also indicated that the prediction errors of the TCN were more stable and had significantly lower variance than those of the LSTM. The training analysis showed that the TCN converges faster than recurrent networks and does not suffer from vanishing gradients. The proposed algorithm therefore predicts water quality with high accuracy.

Several aspects can be improved in further work. The model can take other parameters into account, including precipitation, air temperature, and seasonality. Hybrid models combining the TCN with an attention mechanism or other deep learning algorithms may give better results. The model could be extended to real-time data from IoT sensor networks on rivers, or deployed as a web or mobile application so that environmental agencies can use it in decisions on water resource management. Finally, the model may be validated on larger and more diverse datasets collected in different geographical locations, which would improve its ability to generalize.

Author Contributions

All authors contributed to the conception and preparation of this work and approved the final version of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Generative AI Use

The authors declare that no generative artificial intelligence tools were used in the preparation of this manuscript. The authors take full responsibility for the content of the publication.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Li Lin, Haoran Yang, and Xiaocang Xu, "Effects of water pollution on human health and disease heterogeneity: A review," *Front. Environ. Sci.*, vol. 10, Art. no. 880246, 2022, doi: 10.3389/fenvs.2022.880246.

[2] Naseem Akhtar, Muhammad Izzuddin Syakir Ishak, Showkat Ahmad Bhawani, and Khalid Umar, "Various natural and anthropogenic factors responsible for water quality degradation: A review," *Water*, vol. 13, no. 19, Art. no. 2660, 2021, doi: 10.3390/w13192660.

[3] Morgane Le Moal, Chantal Gascuel-Odoux, Alain Ménesguen, Yves Souchon, Claire Étrillard, Alix Levain, Florentina Moatar, Alexandrine Pannard, Philippe Souchu, Alain Lefebvre, and Gilles Pinay, "Eutrophication: A new wine in an old bottle?," *Sci. Total Environ.*, vol. 651, pp. 1-11, 2019, doi: 10.1016/j.scitotenv.2018.09.139.

[4] Tiyasha, Tran Minh Tung, and Zaher Mundher Yaseen, "A survey on river water quality modelling using artificial intelligence models: 2000-2020," *J. Hydrol.*, vol. 585, Art. no. 124670, 2020, doi: 10.1016/j.jhydrol.2020.124670.

[5] Wei Zhi, Alison P. Appling, Heather E. Golden, Joel Podgorski, and Li Li, "Deep learning for water quality," *Nat. Water*, vol. 2, no. 3, pp. 228-241, 2024, doi: 10.1038/s44221-024-00202-z.

[6] Rahim Barzegar, Mohammad Taghi Aalami, and Jan Adamowski, "Short-term water quality variable prediction using a hybrid CNN-LSTM deep learning model," *Stoch. Environ. Res. Risk Assess.*, vol. 34, no. 2, pp. 415-433, 2020, doi: 10.1007/s00477-020-01776-2.

[7] Zhuhua Hu, Yiran Zhang, Yaochi Zhao, Mingshan Xie, Jiezhuo Zhong, Zhigang Tu, and Juntao Liu, "A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture," *Sensors*, vol. 19, no. 6, Art. no. 1420, 2019, doi: 10.3390/s19061420.

[8] Hansika Hewamalage, Christoph Bergmeir, and Kasun Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *Int. J. Forecast.*, vol. 37, no. 1, pp. 388-427, 2021, doi: 10.1016/j.ijforecast.2020.06.008.

[9] Colin Lea, Michael D. Flynn, Rene Vidal, Austin Reiter, and Gregory D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 1003-1012, doi: 10.1109/CVPR.2017.113.

[10] Davide Chicco, Matthijs J. Warrens, and Giuseppe Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, Art. no. e623, 2021, doi: 10.7717/peerj-cs.623.

[11] Yuexin Fu, Zhuhua Hu, Yaochi Zhao, and Mengxing Huang, "A long-term water quality prediction method based on the temporal convolutional network in smart mariculture," *Water*, vol. 13, no. 20, Art. no. 2907, 2021, doi: 10.3390/w13202907.

[12] Kangyang Chen, Hexia Chen, Chuanlong Zhou, Yichao Huang, Xiangyang Qi, Ruqin Shen, Fengrui Liu, Min Zuo, Xinyi Zou, Jinfeng Wang, Yan Zhang, Da Chen, Xingguo Chen, Yongfeng Deng, and Hongqiang Ren, "Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data," *Water Res.*, vol. 171, Art. no. 115454, 2020, doi: 10.1016/j.watres.2019.115454.

[13] Sang-Soo Baek, Jongcheol Pyo, and Jong Ahn Chun, "Prediction of water level and water quality using a CNN-LSTM combined deep learning approach," *Water*, vol. 12, no. 12, Art. no. 3399, 2020, doi: 10.3390/w12123399.