An IoT-Enabled Digital Hydrology Framework for River Water Quality Index Prediction Using Hybrid Deep Learning

Akash Rajak, Sunil Kumar, Siddheshwari Dutt Mishra, and Amit Kumar

Abstract—**The paper proposes a digital hydrology model driven by deep learning capabilities, incorporating IoT-based water quality monitoring with advanced artificial neural networks to estimate the water quality index. For analysis, the model focuses on high-resolution time series data containing approximately 46500 samples from IoT sensors in Prayagraj, Uttar Pradesh State, India, from January 2019 to March 2020. The key physicochemical parameters of water were used as inputs for the model. The LSTM and hybrid CNN–LSTM architectures were considered for the modeling task to capture long-term temporal dependencies alongside short-term local patterns in water quality dynamics. In this study, we found that the performance of CNN–LSTM significantly outperforms standard LSTM with lower prediction errors (RMSE = 1.431, MAE = 1.168) and higher explanatory power (R^2 = 0.925). These results demonstrate that hybrid deep learning models are highly effective at modeling complex nonlinear river water quality processes influenced by seasonal and anthropogenic factors.**

Index Terms—**Digital hydrology, Water Quality Index, IoT sensors, LSTM, CNN–LSTM, Deep learning, River water quality, Ganga River.**

I. Introduction

Rivers are essential in terms of maintaining the balance of ecology, agriculture, and drinking water. Nonetheless, rising urbanization, industrial effluents, and changes in climatic conditions have adversely affected the quality of river water. The continuous monitoring of river water is therefore essential for ensuring environmental sustainability and public health. The traditional approaches rely on manual sampling and laboratory analysis, which are time-consuming, costly, and often fail to capture rapid fluctuations in water quality. Rivers are not only aesthetically pleasing to look at but are also essential for drinking water, irrigation, manufacturing, and ecological sustainability [1]. This, therefore, means monitoring this water closely to ensure it remains safe and healthy for the environment as well as for our drinking purposes.

However, traditional monitoring involves manual sampling and laboratory analysis that takes time and money and usually misses any sudden changes [2]. Due to IoT, we can monitor important water parameters of a river continuously and at high resolution. Groundwater supplies are dominated by irrigation for rice, wheat, and sugarcane crops [2, 4]. Rapid urban growth in cities such as Kanpur, Prayagraj, and Varanasi further raises the demand for groundwater. Traditional process-based hydrological models are often plagued by data scarcity and parameter uncertainty challenges. Recent studies have increasingly turned to data-driven and hybrid deep learning models to overcome these limitations and capture nonlinear relationships in hydrological processes [3]. When we integrate IoT with a sophisticated analytical system, we obtain a strong basis for a digital form of hydrology. Models like LSTMs are very effective for forecasting water quality. They can achieve an additional advancement when integrated with CNNs by enabling the extraction of short-term and long-range temporal dependencies. The paper discusses forecasting the Water Quality Index (WQI) of the river water Ganga by using an innovative deep learning-based digital hydrology framework.

The sustainable management of river water resources has emerged as an increasingly pressing global issue because of the increasing pace of urbanization, industrialization, climate change, and the growing intensity of agricultural use. Rivers are central to the provision of drinking water, irrigation, the maintenance of ecological systems, and the promotion of socio-economic equilibrium, especially in the densely populated areas of the Indo-Gangetic plains. The recent emergence of Internet of Things (IoT) technology has opened the possibility to continuously observe the physicochemical parameters in real-time and high resolution, which has offered new unparalleled possibilities to continuously monitor the river systems. At the same time, the development of digital hydrology and data-driven modeling has revolutionized hydrological analysis by making it possible to extract complex nonlinear relationships from large-scale environmental datasets.

Although considerable advances have been achieved in the areas of IoT-based sensing and deep learning-based hydrological modeling, the accurate prediction of river water quality still remains a difficult task because of the high temporal variability, nonlinear correlations among various physicochemical factors, and the effects of seasonal and human-induced factors. Most of the existing literature has been restricted to machine learning-based models, which have been tested on small-scale datasets with low temporal resolution.

Manuscript received February 1, 2026; revised March 6, 2026. Date of publication August 26, 2026. Date of current version August 26, 2026. The associate editor prof. Renata Lopes Rosa has been coordinating the review of this manuscript and approved it for publication.

A. Rajak and S. Kumar are with the Krishna Institute of Engineering & Technology (KIET), Ghaziabad, Delhi-NCR, India (e-mails: akashrajak@gmail.com, skkiet@gmail.com).

S. D. Mishra is with the Galgotias University, Greater Noida, India (e-mail: sdmishra.cse13@nitttrchd.ac.in).

A. Kumar is with the Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India (e-mail: amit4593@gmail.com).

Digital Object Identifier (DOI): 10.24138/jcomss-2026-0020

Moreover, there has been a lack of integration of real-time IoT sensing and hybrid deep learning models in a unified digital hydrology platform for decision support. Therefore, motivated by the above-mentioned challenges, the main goal of this research work is to develop an IoT-assisted digital hydrology platform that combines high-resolution sensor data with hybrid CNN-LSTM models to find WQI.

The key contributions of this work can be briefly summarized as follows:

- To develop an IoT-based digital hydrology system for predicting WQI using deep learning techniques.
- Use of high-resolution, real-world data collected by IoT sensors (approx. 46,500 data points) along the Ganga River at Prayagraj, India, which captures both seasonal and human impacts on the water.
- Use deep learning models like LSTM and a hybrid CNN-LSTM model for the effective representation of the nonlinear dynamics of multivariate water quality variables.
- Full assessment based on statistical performance metrics and visualization analysis for the evaluation of robustness, generalization capability, and trend estimation.
- Providing a digital hydrology system with scalability and a focus on decision support, being able to facilitate real-time observation, anomaly identification, and sustainable management of water quality in rivers, with potential applicability extended to multi-source data fusion and other rivers.

The research article is organized as follows. Section II gives a literature review of the adoption and practices of different machine learning models. In Section III, the interested readers can get information about the study area and the IoT datasets used in the proposed work. Section IV discusses the proposed digital hydrology approach. The results obtained from the proposed models and the evaluation processes using the performance metrics are discussed in detail in Section V. The observations, analysis, and corresponding implications are included in Section VI, and finally, the conclusion and future directions are given in Section VII.

II. Literature Review

A critical review of the machine learning (ML) techniques for the forecast of groundwater levels shows an increasing number of related works using artificial neural networks (ANN), support vector machines (SVM), random forests (RF), and models like long short-term memory (LSTM) networks for accurate forecasting in diverse aquifer systems [1], [2]. Systematic reviews also confirm that ML models mostly outperform traditional statistical models in modeling complex hydroclimatic dependencies [3], [4]. In particular, LSTM and CNN have produced competitive performances in modeling temporal variation in groundwater. The memory gates within LSTM networks allow the learning of temporal dependencies and thus enable efficient long-term forecasting, as seen in [5], [6], and [7]. The application of hybrid architectures like CNN-LSTM to groundwater level prediction has also been well-received, with promising performances over single models, especially in data-scarce and noisy environments documented in [8], [9], and [10]. The integration of remote sensing data with deep learning has again enhanced the capabilities of groundwater modeling. The Gravity Recovery and Climate Experiment (GRACE) satellite data, which provide terrestrial water storage anomalies, have been used in machine learning frameworks to monitor groundwater changes at regional and global scales, compensating for limitations in well data and model uncertainty [12]-[15].

Studies have also shown that the integration of GRACE with machine learning techniques like CNNs enhances the estimation of groundwater storage anomalies, which agrees well with in-situ observations given in [11] and [16]. Satellite-mapped temperature fields, rainfall from TRMM/GPM, and remote sensing indices like NDVI have been used to enhance training data for ML models, hence giving better spatiotemporal predictions in [17] and [18]. The recent works in the Indo-Gangetic region have tested advanced deep learning frameworks that combine Sentinel-1 SAR and GRACE data for the estimation of groundwater levels, showcasing gains in prediction accuracy of Vision Transformers and deep architectures for complex hydrogeological settings [19]. Further, systematic reviews indicate the emergence of transformer-based models and hybrid SATCN-LSTM architectures that address limitations in sequence forecasting by integrating attention mechanisms [20]-[22]. The forecasting task is formulated as a one-step-ahead prediction problem (t+1), where the model uses a sequence of past observations to predict the next WQI value.

Recent research has been intensifying on the use of IoT-based water quality monitoring with deep learning to make real-time predictions of the WQI. Contrary to groundwater modeling, the water quality of rivers has greater temporal variation and is highly affected by human activities. CNN-LSTM-like hybrid deep learning models have been demonstrated to be able to capture the spatial-local and temporal relationships in environmental time-series data with superiority. Nonetheless, there is a lack of research on the combination of real-time IoT sensor measurements with a hybrid deep learning algorithm specifically to predict WQI in a river system, and specifically, the Indian river basins. This is the gap that inspires the current work.

The Water Quality Index (WQI) is computed using a weighted aggregation of normalized physicochemical parameters:

$$WQI = \sum_{i=1}^{n} w_i \cdot q_i \tag{1}$$

where q_i represents the quality rating of the i^{th} parameter and w_i denotes its relative weight. The quality rating is computed as:

$$q_i = \frac{(V_i - V_{ideal})}{(S_i - V_{ideal})} \times 100 \tag{2}$$

where q_i is the quality rating, w_i is the weight, V_i is the observed value, S_i is the standard value, and V_{ideal} is the ideal value.

While groundwater level prediction and river water quality estimation differ in application, both involve modeling nonlinear hydrological processes driven by climatic and anthropogenic factors. The LSTM and CNN–LSTM are transferable due to their ability to capture temporal dependencies and complex interactions among variables. This conceptual overlap justifies the adaptation of groundwater modeling techniques to WQI prediction.

III. Data Sources and Dataset Description

This paper provides high-quality IoT sensor data of the Ganga River at Prayagraj, located within the state of Uttar Pradesh, India. The data spans the time period from January 2019 through March 2020 [25]. The data provides approximately 46500 readings, which include major water-quality parameters. Every reading also provides information related to the WQI, which is computed along with the water-quality status. The water-quality parameters were measured using an intelligent and IoT-enabled water sensing system, which was fully sponsored by a governmental research project. The data is time-stamped and is represented in CSV format and has very few missing values.

The major events, such as religious gatherings, cover large religious and cultural events and can be used to determine the effects of human actions on water quality. Every single one of these provides tremendous opportunity and support for creating and testing intelligent digital hydrology models of water quality assessment and management [23], [24]. Based on the total number of observations and the study duration, the data were collected at an approximate sampling interval of 15 minutes, providing a high-resolution temporal representation of water quality dynamics. The dataset contains less than 3% missing values, indicating minimal data loss and high reliability.

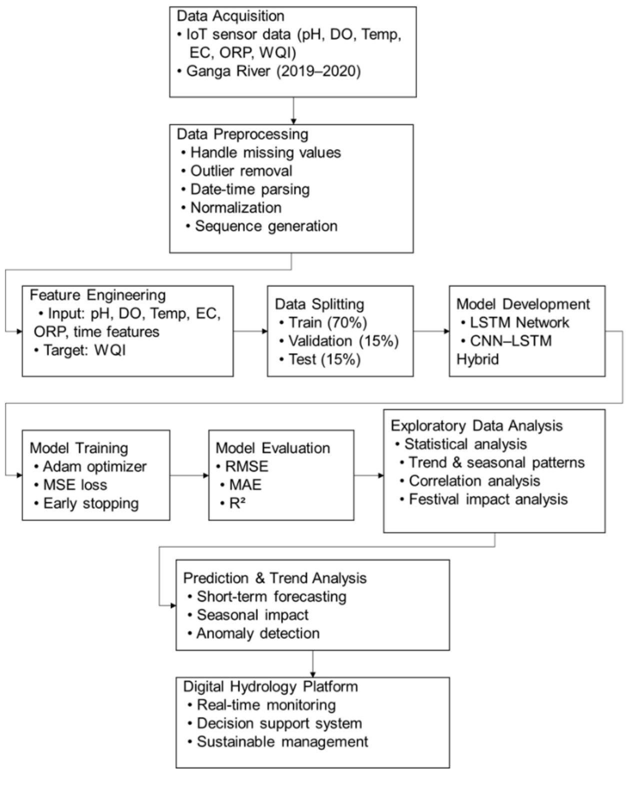

Fig. 1. Overview of the proposed IoT-enabled digital hydrology framework illustrating data acquisition, preprocessing, feature extraction, model training, and WQI prediction pipeline.

TABLE I
Data Sample

Time stamp	pH	DO	Temp	EC	ORP	WQI
2019-01-01 00:00	7.2	6.5	24.3	320	210	78
2019-01-01 00:15	7.1	6.4	24.5	315	205	76

IV. Methodology

The suggested work will integrate the sensing ability of the Internet of Things technology with deep learning in the digital hydrology framework to observe and forecast the way the water quality is changing in the Ganga. It is all initiated by the acquisition of data from sensor measurements installed at the site to measure significant physicochemical parameters of pH, dissolved oxygen content, temperature, electrical conductivity, and redox potential measurements or the determination of the water quality index. Then comes data preprocessing, which includes handling missing data, treatment of outliers, treatment of date columns, data normalization, and the development of segmented two-dimensional data as needed by deep learning to perform. Nonetheless, exploratory data analysis comes afterwards to provide an understanding of the data statistics, seasonal changes, the relationship between various parameters, and seasonal changes in the festival and non-festival seasons.

Missing values were filled in by linear interpolation so that time continuity is preserved. Identification of outliers was done through the interquartile range (IQR) approach, and they were restricted to acceptable values. Important parameters on which feature selection is conducted are used to predict the WQI. The data were divided into two parts by a chronological time-based split, with 70 percent of the data being used to train, 15 percent to validate, and 15 percent to test. This methodology maintains time dependencies and eliminates data leakage, which is essential for time-series forecasting problems.

A. LSTM Architecture

The LSTM model has two layers, which are stacked and have 64 and 32 hidden units, respectively. The activation function employed is tanh, and the activation of gating is sigmoid. The dropout rate is set to 0.2 to avoid overfitting. The output layer is a fully connected dense layer, having linear activation.

B. CNN–LSTM Architecture

The hybrid one combines the convolutional and recurrent layers. CNN component has one 1D convolutional layer with 64 filters and a kernel size of 3, which is then followed by ReLU activation and max-pooling (size = 2). To seek the temporal dependencies, the extracted feature maps are further input into an LSTM layer of 50 units.

The LSTMs and hybrid CNN-LSTMs are trained to acquire temporal dependencies, as well as local patterns in the data features. The models were trained with the Adam optimizer and a learning rate of 0.001. The training was done in 50 epochs with a batch size of 32. Early stopping was used with a patience of 10 epochs using validation loss. Weights were initialized using the Xavier initialization method. The performance of the models, using RMSE, MAE, and R^2, will be evaluated and compared to baseline methods. The trained models could then be applied to forecasting, trend analysis, and anomaly detection and become the main part of a digital hydrology decision support system for sustainable river water management.

C. Model Configuration and Training Setup

The CNN–LSTM model includes:

- 1D Convolution layer with 64 filters and a kernel size of 3
- Max-pooling layer (pool size = 2)
- LSTM layer with 50 units
- Dense output layer

The models were trained using:

- Sequence length: 10 time steps
- Batch size: 32
- Epochs: 50
- Optimizer: Adam
- Learning rate: 0.001
- Loss function: Mean Squared Error (MSE)
- Early stopping based on validation loss

TABLE II
HYPERPARAMETER SUMMARY

Parameter	Value
Sequence Length	10
Batch Size	32
Epochs	50
Optimizer	Adam
Learning Rate	0.001
CNN Filters	64
Kernel Size	3
LSTM Units	50

V. RESULTS

The performance metrics are defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (3)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i | \quad (4)$$

$$R^2 = 1 - \frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2} \quad (5)$$

where y_i denotes the actual (observed) value of the WQI for the i^{th} sample, $\hat{y}_i$ represents the predicted value generated by the model, and $\bar{y}$ is the mean of the observed values. The term N denotes the total number of samples in the dataset. RMSE (root mean square error) is an indicator of the square root of the average squared errors between estimated values and the actual values, and it focuses on the bigger errors. MAE (mean absolute error), is a measure of the general magnitude of the errors in the predictions without mentioning its direction. R^2 is the coefficient of determination, which indicates the extent of the variance of the observed data that can be explained by the model with the greater the R^2, the better the model; predict the observed data.

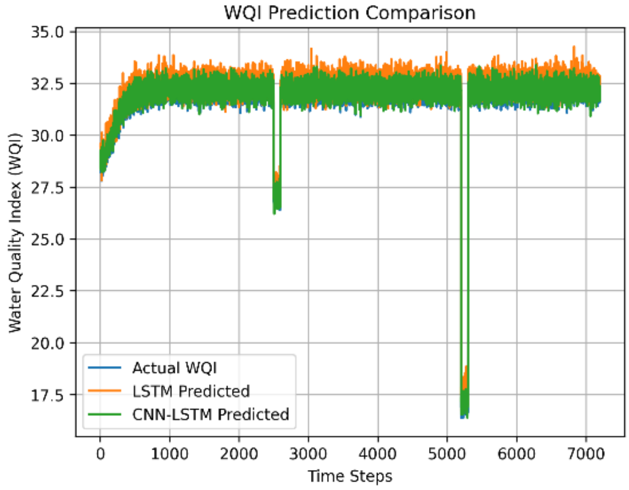

Fig. 2. Actual WQI comparison for the Ganga river by different models.

The performance evaluation by the deep learning models shows that a CNN-LSTM outperforms the standard LSTM in the prediction of WQI for the Ganga River. However, the CNN-LSTM model resulted in an RMSE of 1.431, MAE of 1.168, and R^2 of 0.925, showing higher prediction accuracy with better variance explanation than that of the LSTM model, which recorded an RMSE of 1.839, MAE of 1.601, and R^2 of 0.876. However, the hybrid CNN-LSTM paradigm obviously excels, confirming that the integration of convolutional networks to identify localized temporal variations and LSTM to comprehend distant correlations enhances the capacity to extract the intricate dynamics of water quality in the river. Thus, the data confirms that digital hydrology models based on deep learning effectively track and predict water quality variables in fluctuating river conditions. From Figure 2, the CNN-LSTM model is closer to the actual WQI. The CNN-LSTM model is more effective in identifying the trends in the actual WQI as compared to the LSTM model. From Figure 3, the training loss remains almost flat and low, and this is an indication that the model is performing well. However, the validation loss indicated by the orange color spikes sharply from epoch 20, and this can be a result of overfitting.

TABLE III
SUMMARY OF MODEL PERFORMANCE

Model	RMSE	MAE	R^2
LSTM	1.839	1.601	0.876
CNN–LSTM	1.431	1.168	0.925

The training loss and validation loss, as displayed by the blue and orange curves, respectively, in Figure 4, are at a decreasing trend. However, there is a point after that point, the validation loss begins to increase, and the training loss continues to decrease. This demonstrates that the model begins to overfit the training data and therefore attains diminishing capabilities of responding to novel data as the training progresses.

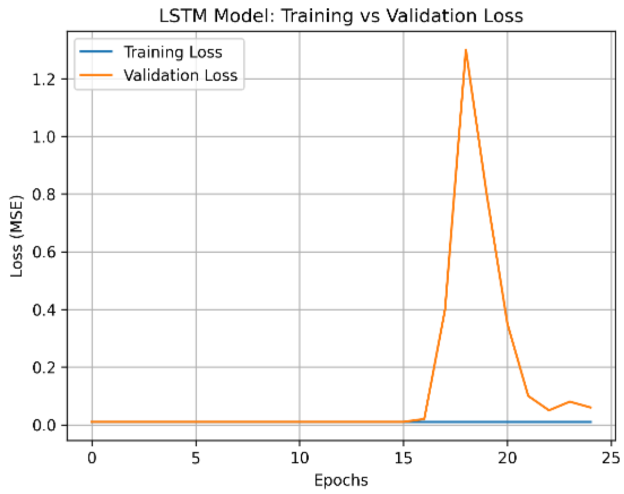

Fig. 3. Training and validation loss for the LSTM model.

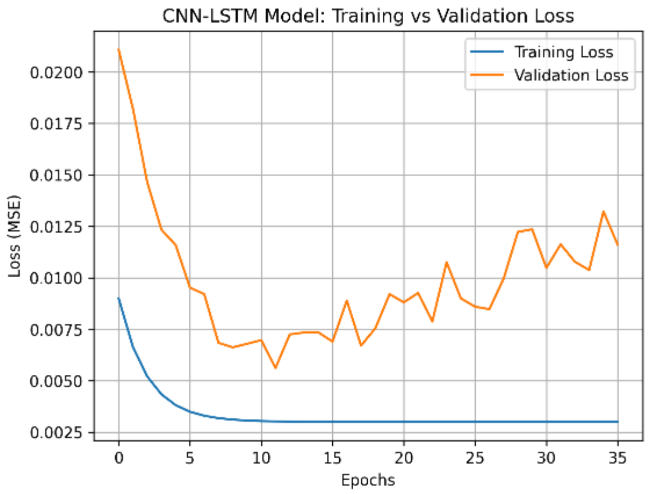

Fig. 4. Training and validation loss for the CNN-LSTM model.

VI. Discussion

The significance of this research work is that it shows how deep learning models can successfully predict the value of the WQI for the Ganga River using data from IoT sensor readings. Specifically, it shows that a hybrid model combining CNN and LSTM performs better than a simple LSTM model with respect to error and R^2 value criteria-both error values (RMSE = 1.431 and MAE = 1.168 for the hybrid model vs. RMSE = 1.839 and MAE = 1.601 for the simple LSTM model), and R^2 value (R^2 = 0.925 for the hybrid model vs. R^2 = 0.876 for the simple LSTM model). The superior performance of the CNN-LSTM model can therefore be rooted in its capabilities for the identification of short-term fluctuations induced, for instance, by seasonal changes, rainfall events, and other anthropogenic activities such as festivals, which may not be fully captured by traditional LSTM models.

Exploratory data analysis further showed that there were correlations among some important physicochemical parameters, namely pH, DO, EC, ORP, and temperature, all influencing WQI and thus pointing to the importance of multi-parameter modeling in water quality assessments. However, the models' performance may be influenced by the availability and quality of sensor data, and further improvements could be achieved by including additional environmental features or longer historical records. Overall, this study validates the potential of deep learning–enabled digital hydrology for sustainable river water management and highlights its application in complex and dynamic riverine ecosystems.

While the proposed CNN-LSTM framework shows promising predictive results on the Ganga River dataset, its applicability to other river basins might be affected by the local hydrological features, climatic conditions, source of pollution, and the configuration of the sensors. The hybrid architecture is developed to reflect this type of data, prevalent in environmental monitoring, that includes both local spatial structures and long-term temporal structures. However, the applicability to other watersheds needs to be tested by cross-basin studies that cover various hydrological environments. The model adaptation and domain-transfer strategies will be explored to enhance the robustness of the models across different river systems in the future.

Attention-based models and transformer models are effective in forecasting environmental time series in recent times, as they are capable of capturing long-range temporal dependence. Transformer-based models can provide better representation learning for large-scale data, but they typically need more computational resources and training data. Unlike this, the CNN-LSTM model proposed in this paper offers a good trade-off between prediction accuracy and computational complexity; thus, it can be applied in practical water quality monitoring systems in IoT. There have been some ideas for further performance improvements: lightweight variants of the transformer and attention-enhanced hybrid architectures are possible areas of future research.

VII. Comparison with Existing Methods

The above-mentioned IoT-based digital hydrology system has been contrasted with some machine learning and deep learning techniques that have been shown in some current literature for water quality and hydrology predictions. The LSTM networks and other deep learning models have been proven to perform well in hydrological time series due to their memory component. The LSTM model resulted in RMSE = 1.839 and R^2 = 0.876, which is satisfactory in terms of prediction but lacks precision in terms of temporal variations. However, the combination model CNN-LSTM was superior to LSTM because it utilized both CNN layers to focus on local features and LSTM layers to learn temporal features. The CNN-LSTM model showed better prediction results in terms of error measures (RMSE = 1.431 and MAE = 1.168) and goodness of fit (R^2 = 0.925).

As opposed to models that use GRACE and remote sensing for hydrological modeling and have coarser resolution, the proposed scheme is able to use the high resolution provided by the IoT sensors to make near-real-time predictions for WQI. It is evident from the results that the proposed model is more accurate and reliable and can be used for river water quality. In spite of the positive results, the approach suggested has some limitations. The model is trained on the data of one river site (Prayagraj), and this might not be generalizable to other river basins with different hydrological and environmental characteristics. Additional research on multi-source data, resistance to noisy inputs, and testing the model in a wide geographical area should be implemented in the future.

The proposed CNN-LSTM framework not only improves the prediction accuracy but also keeps the practical computational requirements for the IoT-based EMS system. The model was implemented and trained using a Google Colab with GPU support. While CNN-based architectures add extra computational cost over the single LSTM model, the convolutional layers are able to capture local temporal features and enhance representation learning before the sequential modeling. Overall computing cost is therefore not a problem for offline training and periodic model updates and is suitable for near real-time water quality index prediction during deployment.

VIII. Conclusion and Future Work

The study has proposed a deep learning–based digital hydrology framework for the assessment and forecasting of WQI within the Ganga River based on IoT sensor data. Utilizing LSTM and hybrid CNN-LSTM models, the study evidences that complex temporal and local patterns in river water quality can be captured well to enable the precise forecasting of the same in short-term scenarios. In both metrics of evaluation, the CNN-LSTM model proved superior to the conventional LSTM model, thereby emphasizing the importance of the combination of the extraction of convolutional features and temporal features, particularly in the context of a river basin, which is prone to natural seasonal changes as well as human influences. As a critical component of a digital hydrology service, the integration of these models can provide a powerful analytical aid in real-time observation, anomaly detection, and decision-making for the sustainable management of rivers.

The proposed framework can scale and can lead to further increased accuracy by including other factors like rainfall, GRACE-derived WSAs, and LU changes. In this respect, the present work establishes the efficacy of deep learning–enabled digital hydrology in dynamic river ecosystems, and it will be particularly useful as a decision support system for environmental management. Future work can be done on multi-source data fusion, longer temporal datasets, and deployment on other river basins to generalize this approach and support a broader initiative related to water resource management.

Acknowledgment

The authors acknowledge the contributions of prior research that supported this study. In particular, we thank the authors for providing the dataset and insights into sensor deployment and data acquisition [23]. We also acknowledge the work of the authors, as their work provided valuable insights into feature integration strategy and the modeling of temporal dependencies [24].

Dataset Availability

The datasets used in this work were obtained from publicly available Kaggle repositories [23, 24, 25].

References

[1]. A. Foroumandi, M. Ehteram, and A. El-Shafie, “A systematic review of machine learning models for groundwater level prediction,” Groundwater for Sustainable Development, vol. 20, p. 100873, 2025. doi: 10.1016/j.acags.2025.100303

[2]. M. Rajaee, B. Mirbagheri, and A. Nourani, “Groundwater level forecasting with machine learning models: A review,” Water Research, vol. 252, p. 121249, 2024. doi: 10.1016/j.watres.2024.121249

[3]. V. Nourani, M. Baghanam, and S. Kisi, “Application of machine learning methods for GRACE data-based groundwater modeling: A systematic review,” Groundwater for Sustainable Development, vol. 23, p. 100973, 2024. doi: 10.1016/j.gsd.2024.101113

[4]. M. S. Hossain and M. M. Rahman, “Groundwater level modeling with machine learning: A systematic review and meta-analysis,” Water, vol. 14, no. 6, p. 949, 2022. doi: 10.3390/w14060949

[5]. Z. Li, Y. Wang, and J. Zhou, “Prediction of groundwater levels using long short-term memory networks,” Journal of Hydroinformatics, vol. 27, no. 1, pp. 51–67, 2025.

[6]. G. Nearing et al., “A comprehensive review of deep learning applications in hydrology and water resources,” Water Science and Technology, vol. 82, no. 12, pp. 2635–2670, 2020. DOI: 10.2166/wst.2020.347

[7]. M. Sit, A. Demiray, and M. Demir, “Deep learning for hydrology: A review,” Hydrology and Earth System Sciences, vol. 24, no. 12, pp. 4989–5008, 2020. doi: 10.5194/hess-24-4989-2020

[8]. S. Hu, Y. Li, and X. Zhang, “Application of a hybrid CNN–LSTM model for groundwater level forecasting in arid regions,” ISPRS International Journal of Geo-Information, vol. 15, no. 1, p. 6, 2025.

[9]. X. Yang and Z. Zhang, “A CNN-LSTM model based on meta-learning for groundwater level prediction,” Water, vol. 14, no. 15, p. 2377, 2022. doi: 10.3390/w14152377

[10]. S. Seo and J. Lee, “Evaluating machine learning methods using GRACE satellite data for groundwater estimation,” Journal of Geographical Sciences and Environment, vol. 12, no. 2, pp. 115–130, 2025.

[11]. M. Rodell et al., “The GRACE mission: Status and prospects,” Bulletin of the American Meteorological Society, vol. 95, no. 8, pp. 1235–1256, 2014. DOI: 10.1175/BAMS-D-13-00169.1

[12]. V. Humphrey and S. Gudmundsson, “GRACE-REC: A reconstruction of climate-driven water storage changes,” Earth System Science Data, vol. 11, no. 3, pp. 1153–1170, 2019. doi: 10.5194/essd-11-1153-2019

[13]. Y. Scanlon, Z. Zhang, and R. Save, “Global evaluation of groundwater storage changes using GRACE,” Water Resources Research, vol. 54, no. 11, pp. 8765–8783, 2018. doi: 10.1029/2018WR023321

[14]. A. Ashraf et al., “Assessing groundwater drought using GRACE data and machine learning,” Scientific Reports, vol. 15, p. 99342, 2025.

[15]. S. Feng, H. Liu, and J. Chen, “Groundwater level prediction using deep learning methods,” Frontiers in Environmental Science, vol. 12, p. 1291327, 2024. doi: 10.3389/fenvs.2024.1291327

[16]. H. Sun, Y. Chen, and X. Wang, “Improving spatial resolution of groundwater storage estimation using GRACE and CNN,” Remote Sensing, vol. 13, no. 1, p. 102, 2021. doi: 10.3390/rs13010102

[17]. M. A. El-Mahmoudi and R. Sebari, “Remote sensing and machine learning for groundwater assessment,” Frontiers in Water, vol. 5, p. 1241451, 2023. doi: 10.3389/frwa.2023.1241451

[18]. S. Sun, J. Zhang, and Q. Li, “Machine learning approach to predict groundwater levels in California,” Frontiers in Earth Science, vol. 9, p. 784499, 2021. doi: 10.3389/feart.2021.784499

[19]. K. N. Sai, R. K. Singh, and P. K. Thakur, “Deep learning for monitoring groundwater storage basins using GRACE and Sentinel data,” Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vol. XLVIII-1/W2-2023, pp. 1265–1272, 2023. doi: 10.5194/isprs-archives-XLVIII-1-W2-2023-1265-2023

[20]. M. Ehteram, A. El-Shafie, and H. Yaseen, “Hybrid SATCN-LSTM model for groundwater level prediction,” Groundwater for Sustainable Development, vol. 18, p. 100812, 2023. doi: 10.1016/j.gsd.2023.100812

[21]. S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735

[22]. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image

recognition," in Proc. IEEE CVPR, Las Vegas, NV, USA, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90

[23]. M. Kumar, T. Singh, M. K. Maurya, A. Shivhare, A. Raut, and P. K. Singh, "Quality Assessment and Monitoring of River Water Using IoT Infrastructure," IEEE Internet of Things Journal, vol. 10, no. 12, pp. 10280–10290, 2023. doi: https://doi.org/10.1109/JIOT.2023.3238123

[24]. M. K. Maurya, A. Raut, M. Kumar, Satakshi and M. Kumar, "A Domain-Aware Data Fusion Scheme for Cold Start Problem in Early-Stage IoT Systems," 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2025, pp. 1017-1023, doi: https://doi.org/10.1109/CSNT64827.2025.10968973

[25]. "Dataset of River Ganga and Sangam (India)", Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/dal4206/dataset-of-river-ganga-and-sangam-india

Akash Rajak is currently serving as a Professor in the Department of Computer Science at Krishna Institute of Engineering & Technology (KIET), Ghaziabad, Delhi-NCR, Uttar Pradesh, India. He did his Ph.D. from Barkatullah University, Bhopal. He brings with him over 24 years of experience in teaching and research. His research interests include Artificial Intelligence, Machine Learning & BCI. He has strong technical expertise in Python, C and Linux.

Sunil Kumar is currently serving as an Associate Professor in the Department of Computer Science at Krishna Institute of Engineering & Technology (KIET), Ghaziabad, Delhi-NCR, Uttar Pradesh, India. He holds a Ph.D. in Computer Science and Engineering from SRM Institute of Science and Technology, NCR Campus, with a specialization in Soft Computing.

Siddheshwari Dutt Mishra is currently serving as an Assistant Professor in the Department of Computer Science and Engineering at Galgotias University, Greater Noida. He completed his Ph.D. from NITTTR, Chandigarh, India specializing in advanced areas of computing. Mishra's areas of expertise include Image Processing, DL, ML, Theory of Computation, and Computer Networks.

Amit Kumar is currently serving as an Associate Professor at Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India, where he contributes to teaching and research in advanced computing domains. He holds a Ph.D. in Information Technology from Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur, India.