INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 1425-1432
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

# REAL-TIME GANGA RIVER WATER QUALITY FORCASTING USING AI ENABLED DSS, SATELLITE DATA, IOT, AND DYNAMIC MODELS

**S. Sai Varun[1], G. Sai Venkatakrishna[2], G. Sai Vignesh[3], D. Sai Vikas[4], A. Sai Vivek[5], Prof. Mairaj Unisa Begum[6]**

[1,2,3,4,5,6]Department of AI & ML Malla Reddy University, Hyderabad, India.

## ABSTRACT

Background: Namami Gange is a flagship program of the government of India for the rejuvenation of Ganga and its tributaries. NMCG Authority order of Oct 2016 States the the pollution in River Ganga and its tributaries shall also be monitored by use of satellite and other remote sensing technologies. As pupulations increase in the Ganga Basin, there are growing water demands, and hence higher levels of sewage flow into rivers of the Ganga Basin from both rural and urban areas. Excess untreated water entering the Ganga river systems transports high organic and pathogen loads. Their use generates high levels of pollution with significant organic loads, with very high Biochemical Oxygen Demand. This will inevitably lower Dissolved Oxygen (DO) concentrations in rivers, thereby threatening fisheries and biodiversity (maacroinvertebrates). At the same time, the high sewage discharges affect the suitability of the river for bathing due to the higher levels ofpathogens in rivers. With climate change increasing the frequency and intensity of rainfall events, this is becoming a significant problem, threatening the water quality and ecology of Ganga river system. The societal impact of providing forecast data that is easily and openly accessible will change how we see and value our rivers and enable citizens to make better decisions based on better data for our health and the health of the rivers of Ganga Basin. Description: An AIenabled decision support system may be developed to Integrate data from multiple sources such as satellite data, IOT- based sensor-generated data, intrumental meteorological measurements, in-situ flow, water quality observations, and miltiple hydrological & hydrodynamic models to work together in real time to generate historical patterns of behaviour and water quality forecasts for Ganga River. Model source code can be seamlessly incorporated into a cloud platform infrastructure

## 1. INTRODUCTION

1.Problem

**Statement**

Rivers play a vital role in sustaining ecosystems, supporting agriculture, and providing water resources for millions of people. However, the deterioration of water quality due to pollution, climate change, and unsustainable human activities has emerged as a critical environmental challenge. In countries like India, major rivers such as the Ganga, Yamuna, and Godavari face severe contamination from industrial effluents, untreated sewage, agricultural runoff, and illegal waste disposal. This pollution not only affects public health and aquatic biodiversity but also has significant socio-economic repercussions, including reduced agricultural productivity, disrupted livelihoods, and increased water treatment costs. Despite the alarming decline in water quality, the existing monitoring systems are inadequate for detecting pollution events in real time, limiting the capacity of authorities to take prompt corrective actions. Traditional river monitoring practices rely heavily on manual sampling and laboratory testing, which are laborintensive, slow, and geographically limited. These methods often fail to capture real-time fluctuations in water quality, making it difficult to identify sudden pollution events, such as industrial discharges or sewage overflows. Furthermore, river water quality is influenced by numerous dynamic factors, including rainfall, seasonal variations, temperature changes, and human activities, which cause continuous and unpredictable changes in parameters such as biochemical oxygen demand (BOD), dissolved oxygen (DO), pH levels, turbidity, and nitrate concentration. Monitoring and predicting these complex interactions using traditional techniques is highly challenging, leading to delayed responses and ineffective water management strategies. The lack of real-time data and predictive capabilities creates significant gaps in the management of river systems. Water authorities struggle to detect pollution hotspots, forecast contamination risks, or anticipate the impacts of climatic variations on water quality. Additionally, most existing monitoring networks have limited geographic coverage and rely on static measurement points, failing to capture the spatial and temporal variability of the river system. As a result, there is an urgent need for an automated, AI-driven system capable of providing real-time monitoring, data driven insights, and predictive analytics to support sustainable water management practices.

## 2. OBJECTIVES OF THE PROJECT

The primary objective of the project is to develop a realtime river water quality monitoring and forecasting system that provides accurate, actionable insights for water management authorities and environmental agencies. The system aims to: • Detect and forecast water pollution events by analyzing real-time satellite and IoT sensor data. • Enable proactive decision-making by delivering timely alerts on water quality fluctuations, aiding authorities in mitigating pollution risks. • Enhance water resource management by providing detailed reports on nitrate levels, BOD, DO, and other parameters to support evidence-based policies. • Improve public safety and environmental protection by informing the public about water conditions for safe recreational activities and drinking water use. • Utilize machine learning for predictive accuracy by implementing KNN algorithms to anticipate water quality changes based on historical trends. • Facilitate large-scale river basin management by ensuring the system is scalable and adaptable to diverse geographical and environmental conditions.

**Scope and Significance of the study**

The scope of this project encompasses the development of an AI-powered river monitoring and management system designed to enhance water quality tracking, flood forecasting, and disaster preparedness, with a specific focus on the Ganga River basin. The system integrates satellite imagery, IoT sensor data, GIS technology, and deep learning models to provide real-time insights into river conditions. It aims to deliver actionable information to water authorities, environmental agencies, disaster management organizations, and researchers by offering capabilities such as 17 water quality monitoring, flood prediction, and risk assessment. The platform processes data from IoTenabled sensors that measure key parameters such as pH, dissolved oxygen (DO), biochemical oxygen demand (BOD), turbidity, and nitrate levels, providing real-time water quality analytics. Additionally, flood dynamics models use rainfall, terrain, and hydrological data to simulate river flow patterns and predict flood events. The system incorporates deep learning algorithms for anomaly detection, enabling the identification of sudden pollution spikes or water quality deterioration. The project features a user-friendly web interface with interactive dashboards, real-time visualizations, and customizable alerts. Stakeholders can access automated reports, historical trends, and predictive analytics to support decision-making. The platform also offers automated alert notifications when pollution thresholds are breached or flood risks are detected. With cloud based deployment, the system ensures scalability and accessibility, allowing real-time data processing and remote monitoring.

## 3. LITERATURE SURVEY

Peri Akiva et al. (2020) [1] introduced H2O-Net, a selfsupervised deep learning model for flood segmentation using satellite and aerial imagery. Their model employed adversarial domain adaptation and label refinement techniques to enhance the accuracy of flood segmentation. The study demonstrated a 10% improvement in pixel accuracy and a 12% increase in mean Intersection over Union (mIoU) compared to existing methods. This work highlights the potential of self supervised learning for enhancing flood detection accuracy. However, the model's performance varied with different sensor data, and it required extensive computational resources, posing scalability challenges. Muhammed Sit et al. (2020) [2] conducted a comprehensive review of deep learning applications in hydrology and water resource management. The study analyzed monitoring, prediction, and classification tasks, identifying key areas where deep learning significantly enhances flood prediction, water quality assessment, and resource management. The authors emphasized the importance of high-quality data and highlighted interpretability challenges in deep learning models. Their work offers a valuable reference for future research aiming to optimize deep learning techniques for hydrological applications. Jonathan Giezendanner et al. (2023) [3] developed a CNN-LSTM-based deep learning framework for fusing satellite data to generate historical inundation maps. By integrating convolutional neural networks (CNNs) with long short-term memory (LSTM) networks, the model reconstructed past flood events with improved accuracy over Bangladesh, outperforming traditional methods. The study demonstrated the effectiveness of combining spatial and temporal data for historical flood mapping. However, the model’s applicability was limited in regions with sparse historical data and required significant computational resources. Claudia Corradino et al. (2019) [4] utilized multispectral Sentinel-2 imagery and machine learning techniques to map recent lava flows at Mount Etna. Their model effectively identified and mapped lava flows, demonstrating the potential of combining satellite imagery with machine learning for accurate disaster monitoring. While the study focused on volcanic activity, the 20 methodology has potential applications in flood mapping and other environmental monitoring tasks. However, further investigation is required to adapt the model for hydrological applications

**Existing System**

The project domain focuses on river system monitoring and management with a particular emphasis on water quality assessment, pollution detection, and flood dynamics in the Ganga River basin. It leverages Artificial Intelligence (AI), remote sensing, and IoT-based monitoring to develop a Decision Support System (DSS) capable of real-time water quality analysis and forecasting. The system integrates machine learning models, satellite imagery, hydrological simulations, and IoT sensors to detect pollution trends and predict contamination sources. The DSS provides real-time monitoring and 3–5 day forecasts for critical water quality parameters, including Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), fecal coliform, 22 ammonia, and phosphorus levels, aiding researchers, policymakers, and water management authorities in making data-driven decisions. It supports multiple data input formats, such as sensor data, satellite images, and hydrological models, offering a comprehensive view of the

river's health. To enhance usability, the system offers interactive visualizations and dashboards accessible through web and mobile applications. It uses advanced evaluation metrics, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), to measure prediction accuracy. Additionally, the project incorporates automated anomaly detection to identify sudden pollution spikes, enhancing its effectiveness for disaster preparedness. The Streamlit-based interface ensures an intuitive and accessible user experience, with caching mechanisms (e.g., @st.cache_data, @st.cache_resource) for performance optimization. The system addresses the growing environmental concerns of the Ganga River, contributing to sustainable river management, climate resilience, and evidence-based policymaking, making it highly relevant for environmental conservation and water resource management domains. 2.3 Existing System: Traditional Traditional river monitoring and pollution assessment systems rely heavily on manual water sampling, laboratory testing, and statistical analysis, making them slow, labor- intensive, and inefficient for realtime monitoring. Conventional methods involve periodic field surveys, where water samples are collected and tested for pollutants, but the time lag between sample collection and result generation reduces their effectiveness for realtime decision- making. Satellite-based monitoring has been used to observe large-scale river dynamics, but it suffers from low temporal resolution and lacks the granularity needed for detecting local pollution sources. Traditional hydrological models are often limited by incomplete or outdated data, making them less reliable for flood forecasting and water quality prediction. Additionally, these models lack automation and adaptability, requiring manual intervention for data processing, model calibration, and validation. 23 Existing systems also face challenges in data integration and multi-source compatibility. Many lack the ability to fuse IoT sensor data, satellite imagery, and hydrological simulations, resulting in fragmented insights. Furthermore, static dashboards and limited user interactivity restrict the accessibility and usability of these systems, making them less practical for large-scale or real-time applications. In water quality assessment, traditional methods rely on regression models and basic statistical techniques that struggle to handle non-linear and complex relationships between water quality parameters. These models offer limited predictive accuracy and cannot detect anomalies or sudden pollution spikes effectively. Additionally, manual reporting systems and the lack of automated alerts reduce the efficiency of disaster preparedness and response measures. Moreover, the absence of AI-based anomaly detection and real-time data visualizations limits the ability of existing systems to provide timely insights, hindering their application in large-scale river management initiatives like the Namami Gange program..

## 4. METHODOLOGY

**Proposed System:**

The proposed system offers a robust and automated solution for real-time water quality monitoring and forecasting of the Ganga River, addressing the inefficiencies of traditional systems reliant on manual sampling and lab-based testing. It integrates advanced machine learning (ML) techniques, such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Logistic Regression, and ensemble classifiers (e.g., Gradient Boosting, Voting Classifier), with IoT sensors, satellite data, and dynamic hydrological models to predict water quality parameters like Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), and nitrate levels over a 3-5 day horizon. The system supports diverse data inputs, including real-time sensor readings, satellite imagery, and meteorological data, enabling comprehensive environmental analysis. It also incorporates anomaly detection to issue early warnings for pollution events (e.g., sewage overflows, nitrate spikes), enhancing its utility for proactive management under initiatives like Namami Gange. Performance metrics such as accuracy, precision, recall, and F1-score are implemented to ensure reliable predictions, with visualizations using Matplotlib and Seaborn providing stakeholders with intuitive insights into water quality trends and forecast accuracy. Built using Django, the proposed system provides a user-friendly, interactive web interface with dashboards displaying realtime data, forecasts, and alerts, along with download options for reports, making it accessible to non-technical users such as policymakers and the public. Features like a centered layout and progress bars for long-running tasks (e.g., data processing) improve usability, while caching mechanisms optimize performance by reducing computation time for repetitive tasks. Robust error handling, logging, and security validations (e.g., input data sanitization) ensure reliability and safety, overcoming the scalability and security limitations of traditional systems. This focus on automation and accessibility streamlines water quality management, supporting evidence based decision-making and public awareness in environmental and public health domains, significantly reducing the manual effort and costs associated with conventional monitoring. The proposed system overcomes the shortcomings of existing IoT- based systems, which lack advanced analytics, by eliminating reliance on static models and manual processes, offering scalability and adaptability to diverse river conditions without extensive retraining. Unlike 27 traditional frequency-based or statistical approaches, it captures dynamic environmental patterns through ML and evaluates predictions comprehensively, ensuring accuracy and timeliness. The integration of multi-source data and real-time processing provides a holistic approach to water quality analysis, addressing the lack of integration in existing systems. By incorporating advanced ML techniques and offering potential for domain-specific customization (e.g., adapting to other rivers) and multilingual support, the system enhances its capability to manage complex environmental data, positioning it as a valuable tool for advancing ecological conservation and research efficiency, particularly for real-time and large-scale applications**.**

**System Architecture:**

The three-tier software architecture (a three-layer architecture) emerged in the 1990s to overcome the limitations of the two-tier architecture. The third tier (middle tier server) is between the user interface (client) and the data management (server) components. This middle tier provides process management where business logic and rules are executed and can accommodate hundreds of users (as compared to only 100 users with the two tier architecture) by providing functions such as queuing, application execution, and database staging. 35 The three tier architecture is used when an effective distributed client/server design is needed that provides (when compared to the two tier) increased performance, flexibility, maintainability, reusability, and scalability, while hiding the complexity of distributed processing from the user. These characteristics have made three layer architectures a popular choice for Internet applications and net- centric information systemstrends.

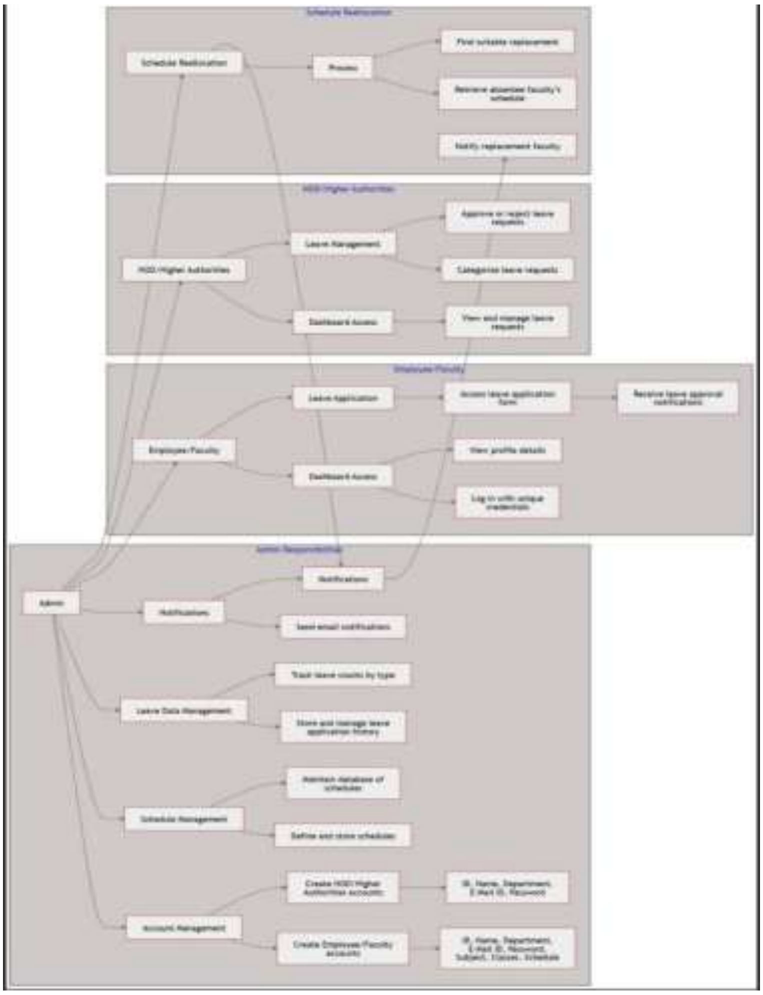


**Fig 1** Block Diagram of Architecture

At the core of the system, a MySQL database stores structured data related to faculty leave records, schedules, and approval workflows. The system incorporates real-time notification services, allowing faculty members to receive updates via email or dashboard notifications regarding the status of their leave requests and any assigned substitute classes. An automated scheduling engine analyzes the availability of other faculty members and assigns substitute lecturers dynamically, ensuring uninterrupted academic sessions. The system also implements role- based access control (RBAC), where faculty members, administrators, and higher authorities have different permission levels to access and modify data. To ensure data security and integrity, authentication mechanisms such as encrypted login credentials and session management are integrated. This robust system architecture enables an efficient, intelligent, and scalable leave management solution tailored to the needs of educational institutions.

## 5. RESULTS AND DISCUSSIONS

The implementation of the AI-driven Leave Management System has significantly improved the efficiency of leave processing within educational institutions. The system's semantic analysis module successfully classifies leave requests based on their textual content, allowing higher authorities to quickly assess and prioritize them. The predictive model for leave approval probability has

demonstrated high accuracy in determining the likelihood of leave acceptance, reducing the time required for

decision-making. Additionally, the automated workflow has minimized manual intervention, ensuring that faculty members can submit and track their leave requests

seamlessly. The integration of email notifications and realtime alerts has further enhanced communication between faculty members and administrators, ensuring

transparency and accountability in the leave management process.

**Fig 2** User profile information


**Fig 3** User login page

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 1425-1432
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

One of the most impactful features of the system is its dynamic class reassignment functionality, which ensures that academic activities are not disrupted due to faculty leave. The scheduling engine efficiently identifies available faculty members with overlapping free hours and assigns substitute classes accordingly. This feature has significantly reduced administrative workload while maintaining smooth academic operations. Furthermore, real-time leave tracking has enabled faculty members and administrators to monitor leave balances accurately, ensuring compliance with institutional leave policies. The system's user-friendly interface has also received positive

feedback from faculty members, who find it intuitive and easy to navigate, further improving adoption rates.


**Fig 4** Detecting water quality testing


**Fig 5** Ganga river water quality testing

Despite its numerous advantages, the system has certain limitations. The accuracy of the AI models depends heavily on the quality and quantity of historical data available for training. In cases where there is insufficient data, the predictive model's performance may be affected. Additionally, while the system efficiently handles structured leave requests, it may face challenges in interpreting complex or ambiguous leave descriptions that require human judgment. Future enhancements could include refining the NLP model for better contextual understanding and expanding integration with other institutional management systems. Overall, the AI- driven Leave Management System presents a scalable and efficient solution for modernizing leave management in educational institutions, streamlining administrative tasks while ensuring uninterrupted academic activities.

## 6. CONCLUSION

Real-time water quality forecasting using AI-enabled decision support systems (DSS), satellite data, IoT, and dynamic models represents a transformative approach to managing water resources effectively. The integration of advanced machine learning algorithms with real-time data from IoT sensors and satellite imagery provides accurate and timely predictions of water quality parameters. This system enhances decision-making for pollution control, resource allocation, and policy implementation. By automating data collection and analysis, it reduces manual efforts and ensures faster response to environmental challenges. Moreover, the dynamic models adapt to changing environmental conditions, offering robust forecasting capabilities. The application of AI ensures scalability and the ability to handle vast datasets, making it ideal for large-scale river systems like the Ganga. However, the success of these systems depends on proper infrastructure, data standardization, and stakeholder collaboration.

## 7. FUTURE SCOPE

The future of water quality forecasting lies in further enhancing AI-enabled systems with hybrid and ensemble machine learning models to improve accuracy and reliability. The integration of predictive analytics with blockchain technology can ensure secure and transparent data sharing among stakeholders. Advancements in IoT sensors and satellite imaging will enable higher- resolution data collection, further refining the forecasting process. The application of deep learning models, such as convolutional neural networks (CNNs), can provide deeper insights into complex patterns in water quality data. Expanding these systems to include real-time alert mechanisms for early warning of critical pollution levels can prevent environmental disasters.

## 8. REFERENCES

[1] S. Diaz, A. Molano, C. Erazo, and J. C. Monroy, "Wqms: water qualitymonitoring station for iot," Int. J. Sens. Netw., vol. 35, no. 2, pp. 79–87,2021.

[2] A. A. Abokifa, A.M.ASCE, K. Haddad, C. Lo, and P. Biswas, "Real-timeidentification of cyber-physical attacks on water distribution systems viamachine learning– based anomaly detection techniques," Journal of WaterResources Planning and Management, vol. 145, no. 1, 2018.

[3] C. M. Ahmed, V. R. Palleti, and A. P. Mathur, "Wadi: A water distributiontestbed for research in the design of secure cyber physical systems,"in Proceedings of the 3rd International Workshop on Cyber-PhysicalSystems for Smart Water Networks, ser. CySWATER '17. New York, NY,USA: Association for Computing Machinery, 2017, p. 25–28. [Online]. Available: https://doi.org/10.1145/3055366.3055375

[4] S. Adepu and A. Mathur, "An investigation into the response of a watertreatment system to cyber attacks," in 2016 IEEE 17th InternationalSymposium on High Assurance Systems Engineering (HASE), 2016, pp.141– 148.

[5] R. Khelif, M. Kharrat, and M. Abid, "A novel design of the multi-wire-based solution for water leak detection and localisation in buried pipes,"Int. J. Sens.

[6] Netw., vol. 39, no. 2, pp. 83–92, 2022 Cho, Byeong-joo, Jang-ho Yun, and Kyeong-ho Lee. "Study of effectiveness for the network separation policy of financial companies." Journal of the Korea Institute of Information Security & Cryptology 25.1 (2015): 181-195.

[7] Bagui, Sikha, et al. "Detecting reconnaissance and discovery tactics from the MITRE ATT&CK framework in Zeek Conn Logs using Spark's machine learning in the big data framework." Sensors 22.20 (2022): 7999.

[8] Anjum, Md Monowar, Shahrear Iqbal, and Benoit Hamelin. "Analyzing the usefulness of the DARPA OpTC Dataset in cyber threat detection research." Proceedings of the 26th ACM Symposium on Access Control Models