International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

Water Usage Pattern in Urban Areas

Mithil.M[1], Prof. H L Shilpa[2], Dhanush K[3], Dhanu[4]
[1,3,4]PG, Department of Master of computer applications, PES College of Engineering, Mandya, 571401, Karnataka, India.
[2]Assistant Professor, Department of Master of computer applications, PES College of Engineering, Mandya, 571401, Karnataka, India.
Emails: *mithil.m0609@gmail.com[1], shilpa@pesce.ac.in[2], dhanushbnr2003@gmail.com[3], dhanyadhanu200317@gmail.com[4]*

Abstract

Urban water demand forecasting is a critical component of short-term supply planning and real-time operational decision-making in smart city water management. This study presents a scenario-aware machine learning framework designed to predict daily urban water consumption across selected cities in Karnataka, India. The proposed framework integrates meteorological, temporal, and demand-related attributes to model consumption behavior effectively. A robust preprocessing pipeline is implemented, encompassing missing-value imputation, temporal sorting, lag and rolling feature construction, categorical encoding, and correlation-based feature selection. Three predictive models—XGBoost, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU)—are developed and compared. The optimal model is deployed via a Flask-based web application featuring SQLite storage, real-time weather data retrieval through Open-Meteo APIs, public holiday detection, prediction history management, and interactive scenario simulation capabilities for rainfall, heatwave, and holiday demand variations. The integrated framework offers water utilities a practical decision-support tool for proactive resource allocation and contingency planning. By enabling what-if analysis under extreme weather events and special calendar days, the system enhances municipal preparedness and contributes to sustainable urban water resource management. This work demonstrates the operational viability of combining machine learning with scenario-based planning to address the complexities of water supply systems in developing urban regions.

Keywords: *Urban Water Demand Forecasting, Smart Water Management, XGBoost Regressor, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Weather-Aware Forecasting, Scenario-Based Decision Support.*

1. Introduction

URBAN water demand prediction has become an essential component of smart city water resource management due to increasing population growth, urbanization, and climate variability. Accurate forecasting of daily water consumption enables water utilities and government agencies to optimize water supply planning, storage management, distribution scheduling, and operational decision-making. In rapidly developing regions, daily water demand is influenced by multiple factors, including weather conditions, seasonal variations, public holidays, weekends, and city-specific consumption patterns. Consequently, reliable forecasting models are required to support sustainable and efficient water resource management. Traditional water demand estimation methods primarily rely on historical averages, statistical analysis, or manual planning approaches. Although these methods provide basic estimates, they often fail to capture short-term fluctuations caused by dynamic environmental and social factors. Urban water consumption exhibits nonlinear and time-dependent behavior, making conventional forecasting techniques less effective in complex real-world scenarios. Recent advances in

e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

machine learning and deep learning have provided powerful alternatives for modeling such complex relationships and improving prediction accuracy [1]–[4]. This study focuses on predicting daily urban water demand in selected cities of Karnataka using weather, temporal, and consumption-history information. The input features include temperature, humidity, rainfall, weekend status, public holiday status, monsoon indicators, city information, cyclic month representations, and lag-based consumption variables. To evaluate forecasting performance, both machine learning and deep learning approaches are implemented. The machine learning model is based on the XGBoost Regressor, while the deep learning models include Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks. The proposed methodology incorporates several preprocessing and feature engineering techniques, including missing-value handling, city-wise temporal ordering, lag feature generation, rolling consumption statistics, cyclic month encoding, categorical feature encoding, correlation-based feature filtering, and recursive feature selection. These techniques improve data quality and enhance the predictive capability of the forecasting models. Abines [1] demonstrated that advanced deep learning architectures effectively capture temporal dependencies in water consumption data. Lall et al. [2] revealed that consumption patterns are influenced by academic calendars and holidays using IoT-based monitoring systems. Abesua et al. [3] showed that pattern similarity-based techniques extend forecast horizons while removing seasonality determination requirements. Abdel Nasser et al. [4] demonstrated that integrated IoT systems with deep learning enable real-time water consumption analysis and forecasting. Despite these contributions, existing approaches often lack scenario simulation capabilities for extreme weather events and special calendar days. This study addresses this gap by presenting a scenario-aware forecasting framework deploying the optimal model through a Flask-based web application featuring real-time weather data integration, holiday detection, and interactive scenario simulation for rainfall, heatwave, and holiday demand variations. The remainder of this paper is organized as follows: Section II describes the methodology. Section III presents experimental results. Section IV discusses the deployment framework. Section V concludes the paper.

2. Literature Review

Water usage pattern analysis in urban areas has evolved as a specialized research domain connecting smart meter data collection with computational time series forecasting and consumption behavior classification. Recent studies from 2020 onward demonstrate a progressive shift from traditional statistical methods such as ARIMA and exponential smoothing toward machine learning and deep learning-based approaches. These studies are relevant to the present work as they establish that water consumption prediction can be approached either by extracting handcrafted features or by using deep learning models that learn temporal dependencies directly from sequential consumption data. Abines [1], Lall et al. [2], and Abesua et al. [3] collectively demonstrate the effectiveness of advanced deep learning architectures and IoT-based systems for water consumption forecasting. Abines [1] showed that recurrent and attention-based models effectively capture temporal dependencies and seasonal patterns in consumption data. Lall et al. [2] revealed that consumption patterns are significantly influenced by academic calendars and holidays, establishing a precedent for scalable, non-invasive monitoring solutions in intermittent water supply systems using IoT-based smart meters. Abesua et al. [3] demonstrated that pattern similarity-based techniques can simultaneously address two major difficulties in water demand prediction: removing the need for annual seasonality determination through signal normalization and extending forecast horizons to 24 hours. Abdel Nasser et al. [4] further reinforced these findings by demonstrating that integrated IoT systems with micro-services architecture and deep learning can effectively collect and analyze water consumption data in real-time, providing a scalable testbed for water demand management and resource optimization. Jha et al. [5] and Abu Waraga et al. [6]

explored complementary aspects of water management through integrated monitoring and clustering approaches. Jha et al. [5] implemented a Smart Water Monitoring System for real-time water quality and usage monitoring, demonstrating that integrated systems for simultaneous water quality and quantity monitoring with real-time remote access and automated billing can effectively promote water conservation through tiered pricing and prevent health hazards. Abu Waraga et al. [6] investigated water consumption patterns through time series clustering, showing that grouping communities with similar consumption patterns significantly improves forecasting accuracy, with forecasting performance correlating with the magnitude of observations in each cluster. Wang et al. [7] and Deng et al. [8] focused on pattern analysis and hybrid modeling approaches. Wang et al. [7] demonstrated that combining time series clustering, seasonal analysis, and entropy analysis provides a comprehensive understanding of urban water consumption patterns and their drivers, enabling development of personalized water conservation strategies tailored to specific household types. Deng et al. [8] showed that hybrid deep learning models combining Convolutional Neural Networks' feature extraction capabilities with GRU's sequence learning ability significantly outperform single models in capturing complex temporal patterns, with user classification and historical consumption patterns enhancing prediction accuracy. Gorenekli and Gulbag [9], Garcia-Soto et al. [10], and Arsene et al. [12] are particularly relevant to the present work as they applied machine learning and deep learning techniques to water consumption prediction and pattern analysis in urban areas. Gorenekli and Gulbag [9] demonstrated that ensemble methods are most effective for water consumption prediction, with historical consumption data and weather parameters identified as significant predictors. Garcia-Soto et al. [10] showed that deep learning models specifically designed for water consumption forecasting can achieve high accuracy for real-world data, with nested cross-validation providing a robust evaluation methodology. Arsene et al. [12] demonstrated that clustering consumption events enables effective characterization of consumer behavior, with deep learning models offering superior prediction capabilities for identifying consumption sources. El Hanjri et al. [11] addressed data privacy concerns by exploring federated learning for water consumption forecasting, demonstrating that personalization improves predictions for participating clients while preserving consumption data privacy and reducing networking load compared to centralized training. Their work addresses both privacy concerns and data diversity challenges simultaneously. These studies collectively establish that data preprocessing, feature selection, and model architecture are critical factors for prediction performance. The present work is positioned within this research direction by developing a comprehensive water consumption prediction and pattern analysis system for different subscriber categories, integrating structured feature-based prediction with temporal dependency modeling to support decision-making for water resource management and planning.

2.1.Research Gap

The existing literature shows strong progress in urban water demand forecasting using statistical models, machine learning, probabilistic forecasting, and recurrent deep learning. However, most previous studies remain limited to standalone forecasting models and do not provide an integrated operational platform for smart water resource management. The key gap is the limited combination of weather-aware demand prediction, public holiday detection, monsoon-based seasonality, lag-based consumption modelling, heatwave and rainfall scenario simulation, prediction history storage, and operational modules such as dam monitoring, gate operations, water distribution, inspection reporting, emergency reporting, alerts, and report generation. The proposed system is positioned to address this gap by combining XGBoost-based urban water demand forecasting with a Flask-based smart water management application that supports both prediction and practical decision-making.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

3. Methodology

The proposed methodology aims to forecast daily urban water demand for selected cities in Karnataka using weather, calendar, seasonal, city-specific, and historical consumption features. The overall framework consists of data preprocessing, feature engineering, model training, performance evaluation, and deployment within a web-based smart water management platform. The target variable considered in this study is consumption_mld, representing daily water demand in Million Litres per Day (MLD).

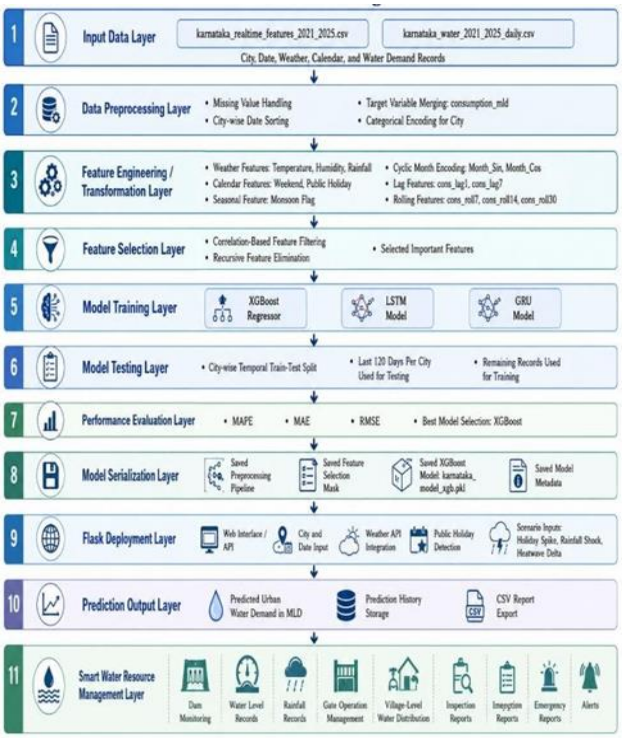

Figure 1 Workflow of the proposed urban water demand prediction system showing the complete pipeline from input data loading to smart water resource management

3.1. Data Set Description

The dataset utilized in this study is designed for daily urban water demand prediction across selected cities in Karnataka, India. The data is organized in a tabular CSV format and comprises two primary files: the feature dataset containing city-wise real-time explanatory variables and the target dataset providing daily water consumption measurements. The feature dataset includes meteorological attributes such as temperature, humidity, and rainfall; calendar-based indicators including weekend status, public holiday status, and monsoon flags; city identifiers; cyclic month representations; and lag-based consumption variables. The target variable, denoted as consumption_mld, represents daily urban water demand measured in million litres per day (MLD) To enhance predictive capability, additional engineered features are generated, including previous-day consumption, seven-day lag consumption, and rolling consumption averages over 7-day, 14-day, and 30-day windows. These features collectively capture environmental conditions, calendar effects, seasonal variations, location-specific characteristics, and historical consumption patterns that influence daily water demand [13]. The original feature dataset contains approximately 20,340 records. After merging with the target consumption data and removing records with missing target values, the dataset reduces to approximately 20,316 records. Following the generation of lag and rolling features and the removal of initial warm-up rows required for feature computation, the final usable dataset comprises approximately 19,956 records. Since the project is formulated as a regression problem with continuous water demand measured in MLD as the prediction target, there is no class distribution in the dataset.

3.2. Data Preprocessing

Data preprocessing is performed to prepare the dataset for machine learning and deep learning models. Missing values in input features are handled using appropriate imputation techniques. Numerical variables such as temperature, humidity, rainfall, lag consumption values, and rolling statistics are processed for model training. Figure 1. Proposed pipeline architecture for urban water demand forecasting and smart resource management, comprising data ingestion, preprocessing, feature engineering, model training (XGBoost, LSTM, GRU), evaluation, serialization, Flask API deployment, and operational decision-support

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

modules Categorical attributes, particularly the city variable, are encoded to capture city-specific demand characteristics. The dataset is organized in chronological order for each city to preserve temporal relationships. Records containing missing target values are removed to ensure data consistency and improve model reliability. The preprocessing pipeline is preserved and stored alongside the trained forecasting model to ensure identical transformations during deployment and real-time prediction [14-19].

3.3.Feature Engineering

Feature engineering is conducted to capture temporal, seasonal, and consumption-history patterns that influence urban water demand.

3.3.1. Temporal and Seasonal Features

- The month variable is transformed into cyclic representations using sine and cosine encoding to preserve seasonal continuity. Monsoon periods are represented using a binary indicator variable. Weekend and public holiday information are incorporated as calendar-based features to capture variations in consumption behavior.

3.3.2. Historical Consumption Features

Several lag-based and rolling-window features are generated from historical water demand records, including:

- Previous-day consumption (Lag-1)
- Seven-day lag consumption (Lag-7)
- Seven-day rolling average
- Fourteen-day rolling average
- Thirty-day rolling average

These features enable the models to learn temporal demand patterns and recurring consumption trends.

3.3.3. Feature Selection

To reduce redundancy and improve model efficiency, correlation-based filtering is applied to remove highly correlated variables. Recursive Feature Elimination (RFE) is subsequently employed to identify the most relevant predictive features for model training.

3.3.4. Model Development

XGBoost Regressor serves as the primary machine learning model for daily urban water demand prediction. It is well-suited for structured tabular datasets due to its capacity to learn nonlinear relationships between input variables and the target output. The model utilizes city identifiers, temperature, humidity, rainfall, weekend status, public holiday status, monsoon flags, cyclic month features, lag features, and rolling consumption features to predict consumption_mld. XGBoost constructs multiple regression trees sequentially, where each new tree aims to minimize the error from its predecessors. The final prediction is obtained by aggregating the outputs of all trees. In the proposed system, XGBoost is trained following preprocessing, feature transformation, correlation filtering, and Recursive Feature Elimination. The trained model is saved as karnataka_model_xgb.pkl and deployed within the Flask-based application. The prediction function is:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i)$$

where $\hat{y}_i$ is the predicted water demand for the ith sample, x_i is the input feature vector, K is the number of trees, and f_k is the prediction function of the kth tree. The objective function is:

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Three forecasting models are implemented and **Compared:**

- XGBoost Regressor
- Long Short-Term Memory (LSTM) Network
- Gated Recurrent Unit (GRU) Network

The XGBoost model is selected due to its ability to effectively handle structured tabular data and nonlinear feature interactions. LSTM and GRU architectures are included to evaluate deep learning approaches capable of modeling temporal dependencies within sequential consumption data.

3.4.Training and Testing Strategy

A city-wise temporal train-test split is adopted to ensure realistic forecasting evaluation. For each city,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

the most recent 120 days of observations are reserved for testing, while all preceding records are used for training. Unlike random data partitioning, this approach evaluates model performance on future unseen observations, closely resembling real-world forecasting scenarios.

3.4.1. Algorithms

The proposed Urban Water Demand Prediction system employs three forecasting models: XGBoost Regressor, Long Short-Term Memory (LSTM) Network, and Gated Recurrent Unit (GRU). These algorithms are applied to predict the continuous target variable consumption_mld, representing daily urban water demand in million litres per day. The system also utilizes feature engineering, correlation-based filtering, Recursive Feature Elimination, and regression evaluation metrics including MAPE, MAE, and RMSE.where $l\ (yi,\ \hat{yi})$ is the prediction loss and $\Omega(fk)$ is the regularization term. The regularization term is:

$$\Omega(f) = \gamma T + \lambda \sum_{j=1}^{T} wj^2$$

where T is the number of leaves, wj is the weight of the jth leaf, γ controls leaf complexity, and λ controls weight regularization.

3.4.2. Long Short-Term Memory Network

Long Short-Term Memory (LSTM) is a recurrent deep learning model employed for sequence-based water demand forecasting. Since water consumption is influenced by historical usage patterns, LSTM is utilized to learn temporal dependencies from sequential demand data. In this project, LSTM is trained using a 30-day sequence length, where preceding observations are used to predict subsequent daily water demand. LSTM incorporates memory cells and gating mechanisms to regulate information flow. The forget gate discards irrelevant past information, the input gate determines which new information should be stored, and the output gate generates the hidden representation used for prediction. These mechanisms make LSTM particularly effective for time-series forecasting tasks where prior demand behavior influences future consumption. The forget gate is:

$$ft = \sigma\ (Wf\ [ht{-}1, xt] + bf)$$

The input gate is:

$$it = \sigma\ (Wi\ [ht{-}1, xt] + bi)$$

The candidate cell state is:

$$\tilde{C}t = \tanh\ (WC\ [ht{-}1, xt] + bC)$$

The cell state update is:

$$Ct = ft \times Ct{-}1 + it \times \tilde{C}t$$

The output gate is:

$$ot = \sigma\ (Wo\ [ht{-}1, xt] + bo)$$

The hidden state is:

$$ht = ot \times \tanh\ (Ct)$$

where xt is the input at time step t, $ht{-}1$ is the previous hidden state, $Ct{-}1$ is the previous cell state, Ct is the updated cell state, and ht is the current hidden state.

3.4.3. Gated Recurrent Unit

Gated Recurrent Unit (GRU) is a recurrent deep learning model employed for time-series water demand prediction. GRU shares similarities with LSTM but features a simplified architecture with fewer gates. It utilizes an update gate and a reset gate to regulate the retention and discarding of past information. In this project, GRU is trained using sequential water demand data to assess whether a simpler recurrent architecture can effectively predict daily urban water demand. The update gate determines how much previous information should be preserved, while the reset gate controls the extent to which past information should be forgotten. GRU is advantageous for time-series regression due to its capacity to learn temporal relationships with fewer parameters compared to LSTM. The update gate is:

$$zt = \sigma\ (Wzxt + Uzht{-}1 + bz)$$

The reset gate is:

$$rt = \sigma\ (Wr\ xt + Urht{-}1 + br)$$

The candidate hidden state is:

$$\tilde{h}t = \tanh\ (Whxt + Uh\ (rt \times ht{-}1) + bh)$$

The final hidden state is:

$$ht = (1 - zt) \times ht{-}1 + zt \times \tilde{h}t$$

where xt is the input at time step t, $ht{-}1$is the

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

previous hidden state, $zt$is the update gate, $rt$is the reset gate, $\tilde{h}t$is the candidate hidden state, and $ht$is the final hidden state.

3.5.Performance Evaluation

Model performance is assessed using three widely adopted regression evaluation metrics:

- Mean Absolute Percentage Error (MAPE)
- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)

These metrics provide a comprehensive evaluation of forecasting accuracy and model reliability. The model exhibiting the lowest prediction error is selected as the final forecasting model.

3.6.System Deployment

The selected forecasting model is deployed using a Flask-based web application integrated with SQLite database storage. The deployment framework supports real-time water demand prediction, weather information retrieval, public holiday detection, prediction history management, and scenario-based simulations. The system enables water resource planners to evaluate the impact of rainfall variations,heatwave conditions, and holiday demand sp ikes, there by supporting informed operational decision makin g and sustainable urban water management.

4. Results and Discussions

This section presents the experimental results obtained from the proposed urban water demand forecasting framework. Since the objective of the study is to predict the continuous variable consumption_mld (Million Litres per Day), the problem is formulated as a regression task. Therefore, regression evaluation metrics including Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) are used to assess model performance. The forecasting models were trained using a comprehensive set of features engineered to capture the multifaceted nature of urban water demand. These features encompass weather-related variables including temperature, humidity, and rainfall; calendar-based indicators such as weekend status and public holiday information; seasonal attributes represented through monsoon flag and cyclic month encoding; city-specific categorical features; and historical consumption patterns captured through lag features (previous-day and seven-day lag consumption) and rolling statistics (7-day, 14-day, and 30-day moving averages). This rich feature set was designed to model the complex relationships between environmental conditions, temporal patterns, and water consumption behavior.

4.1.Model Performance Comparison

Three forecasting models, namely XGBoost Regressor, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU), were implemented and evaluated. Table 1 summarizes the forecasting performance of the developed models.

Table 1 Performance Comparison of Forecasting Models

Model	MAPE	MAE	RMSE
XGBoost Regressor	0.0230	9.467	20.187
GRU	0.067578	25.589	53.817
LSTM	0.068990	26.468	54.437

The results demonstrate that the XGBoost Regressor achieved the best forecasting accuracy among all evaluated models, obtaining the lowest MAPE, MAE, and RMSE values. The model achieved a MAPE of 2.30%, indicating that the average forecasting error is very low relative to actual water demand values. Among the deep learning approaches, the GRU model slightly outperformed the LSTM model. However, both recurrent neural network models produced significantly higher forecasting errors compared with XGBoost. The superior performance of XGBoost can be attributed to its ability to effectively model nonlinear relationships in structured tabular datasets containing weather, calendar, seasonal, and lag-based features. Average Water Consumption by City Shown in Figure 2 - 8.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

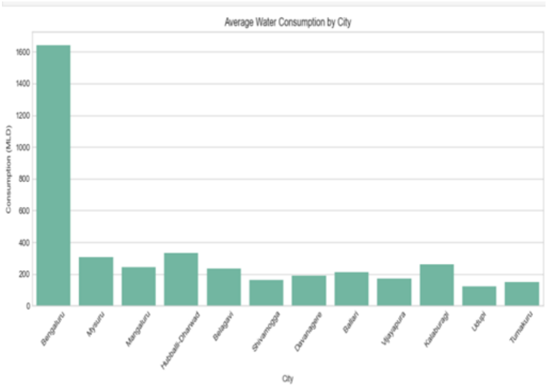

Figure 2 Illustrates the Average Daily Water Consumption Across the Selected Karnataka Cities

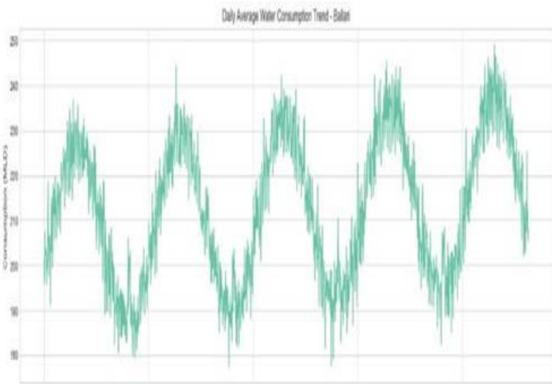

Figure 3 Daily Average Water Consumption Trend Forballari Showing Seasonal Fluctuations with Demand Ranging Between 18 0 MLD and 250 MLD

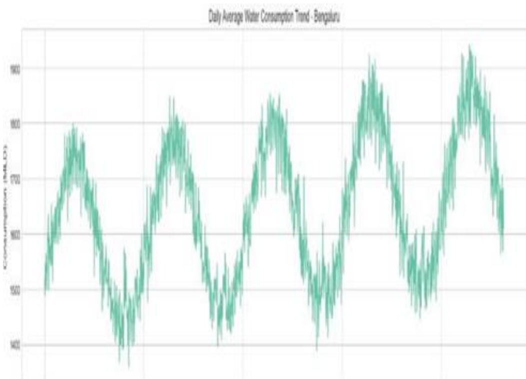

Figure 4 Daily average water consumption trend for Belagavi showing seasonal demand patterns with consumption ranging approximately between 200 MLD and 270 MLD

Hubballi-Dharwad and Mysuru show the next highest consumption levels, followed by Kalaburagi, Mangaluru, Belagavi, Ballari, and Davanagere. Smaller cities such as Shivamogga, Vijayapura, the results indicate significant variation in water demand among cities. Bengaluru exhibits the highest average water consumption, exceeding 1600 MLD, due to its large population and extensive urban infrastructure. Tumakuru, and Udupi exhibit comparatively lower water demand. These differences highlight the importance of incorporating city-specific characteristics into forecasting models.

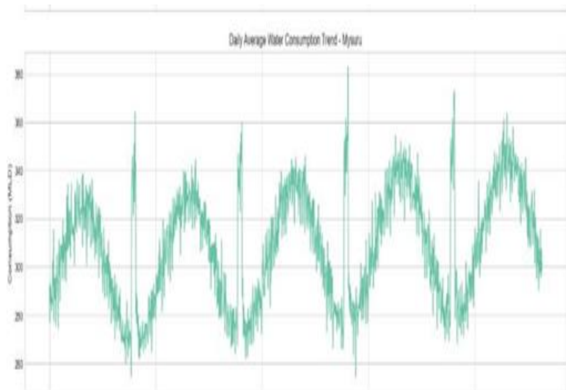

Figure 5 Daily average water consumption trend for Bengaluru showing the highest demand ranging from approximately 1400 MLD to over 1900 MLD

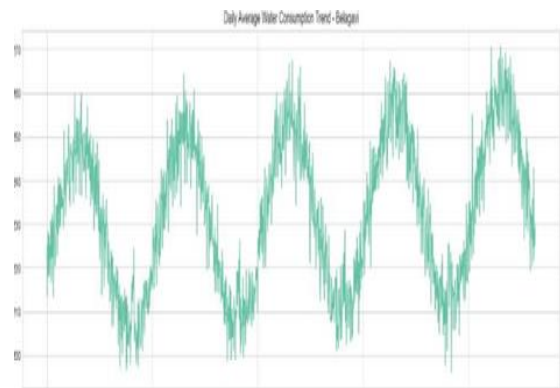

Figure 6 Daily average water consumption trend for Mysuru with demand ranging between approximately 260 MLD and 380 MLD

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

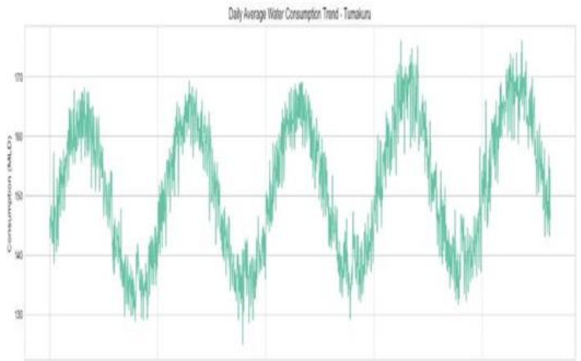

Figure 7 Daily average water consumption trend for Tumakuru showing the lowest demand among all cities, ranging between approximately 130 MLD and 175 MLD with clear seasonal patterns

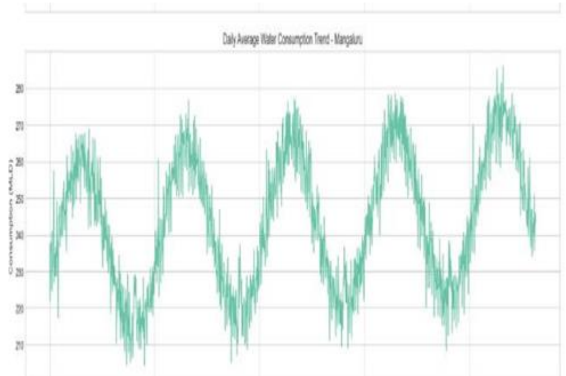

Figure 8 Daily average water consumption trend for Mangaluru showing moderate seasonal variability with demand ranging between approximately 210 MLD and 285 MLD

Conclusion

This study presented a machine learning-based urban water demand prediction and smart water resource management system for selected cities in Karnataka. The main objective was to predict daily urban water demand in million litres per day using weather, calendar, seasonal, city-specific, and historical consumption features [13]. The proposed methodology encompassed data preprocessing, city-wise temporal ordering, missing-value handling, categorical encoding, feature engineering, correlation-based feature filtering, and recursive feature elimination. Weather attributes including temperature, humidity, and rainfall were combined with calendar indicators such as weekend and public holiday status, seasonal behavior captured through monsoon flags and cyclic month encoding, and historical demand patterns represented using lag and rolling consumption features. Three models were implemented and evaluated: XGBoost Regressor, LSTM, and GRU. Performance assessment using standard error metrics demonstrated that XGBoost was more suitable for the structured tabular feature set compared with the recurrent deep learning models. The selected XGBoost model was deployed in a Flask-based web application supporting city-wise demand prediction, weather API integration, public holiday detection, scenario-based simulation, prediction history storage, and CSV report export. The scenario module enabled analysis of demand variations under holiday spikes, rainfall shocks, and heatwave conditions. The system also integrated smart water resource management modules including dam monitoring, water level records, rainfall records, gate operation management, village-level water distribution, inspection reports, emergency reports, and alerts. The major contribution of this project is the integration of urban water demand forecasting with an operational decision-support platform, connecting machine learning-based predictions with practical water management functionalities. This integration makes the system valuable for planning, monitoring, reporting, and scenario-based decision-making in smart water resource management.

Future Enhancements

Future work will focus on incorporating real municipal water consumption records and live IoT-based flow meter data to enable real-time monitoring and dynamic forecasting. Population density and land-use features will be included to capture demographic influences on consumption patterns. The deployment architecture will be extended from SQLite to scalable databases such as PostgreSQL or MySQL to handle larger datasets and concurrent users. Cloud deployment will facilitate remote accessibility and automatic updates. Hybrid forecasting models combining XGBoost with deep learning architectures will be explored to further improve prediction accuracy. Automated real-time lag feature generation will be implemented to eliminate manual preprocessing steps. Demand-level classification into low, medium, and high consumption categories will provide actionable insights for water resource planning. Integration with real-time reservoir sensor data will enable accurate supply-demand balancing. Additional enhancements include expanding scenario simulation to cover drought conditions and infrastructure failure events, implementing anomaly detection for leak identification, and developing mobile application interfaces to promote community participation in water conservation efforts.

Acknowledgment

The author expresses sincere gratitude to the Department of Master of Computer Applications (MCA), P.E.S. College of Engineering, Mandya, for providing the infrastructure, resources, and academic support required for this research.

References

[1]. Abines, A. L. (2025). Machine learning-based time series forecasting of Cauayan City water consumption. In 2025 8th International Conference on Vocational Education and Electrical Engineering (ICVEE).
[2]. Lall, A. K., Terala, A., Goyal, A., Chaudhari, S., Rajan, K. S., & Chouhan, S. S. (2024). Behavioural analysis of water consumption using IoT-based smart retrofit meter. IEEE Access, 34, 36889.
[3]. Abesua, R., Ortiz-Caraballo, C., Preciado, J. C., Figueroa, F. S., Rubio-Largo, A., & Corchado, J. M. (2019). Water demand forecasting using pattern similarity-based techniques. Water (MDPI).
[4]. Abdel Nasser, A., Rashad, M. Z., & Hussein, S. E. (2020). A two-layer water demand prediction system in urban areas based on micro-services and LSTM neural networks. IEEE Access, 8, 147649-147663.
[5]. Jha, M. K., Sah, R. K., Rashmitha, M. S., Sinha, R., Sujatha, B., & Suma, K. V. (2017). Smart water monitoring system for real-time water quality and usage monitoring. In IEEE Conference Proceedings.
[6]. Abu Waraga, O., Abu Talib, M., Abdeljaber, A., & Abdallah, M. (2021). Investigating water consumption patterns through time series clustering. In 2021 14th International Conference on Developments in eSystems Engineering (DeSE).
[7]. Wang, R., Zhao, X., Qiu, H., Cheng, X., & Liu, X. (2024). Uncovering urban water consumption patterns through time series clustering and entropy analysis. Water Research, 122085.
[8]. Deng, L., Chang, X., & Wang, P. (2022). Daily water demand prediction driven by multi-source data. In 7th International Conference on Intelligent, Interactive Systems and Applications (Elsevier Procedia).
[9]. Gorenekli, K., & Gulbag, A. (2024). Comparative analysis of machine learning techniques for water consumption prediction: A case study from Kocaeli Province. Sensors (MDPI), 24(17), 5846.
[10]. Gar cia-Soto, C. G., Torres, J. F., Zamora-Izquierdo, M. A., Palma, J., & Troncoso, A. (2024). Water consumption time series forecasting in urban centers using deep neural networks. Springer.
[11]. El Hanjri, M., Kabbaj, H., Kobbane, A., & Abouaomar, A. (2023). Federated learning for water consumption forecasting in smart cities. In 2023 IEEE International Conference on Communications (ICC): IoT and Sensor Networks Symposium.
[12]. Arsene, D., Predescu, A., Pahontu, B., Chiru, C. G., Apostol, E.-S., & Truică, C.-O. (2022). Advanced strategies for monitoring water consumption patterns in households based on IoT and machine learning. Water (MDPI), 14(14), 2187.
[13]. Ji, Y., Li, Y., Yang, J., Kuang, K., Hu, Y., & Qi, W. (2015).

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0365
e ISSN: 2584-2854
Volume: 04 Issue: 07 July 2026
Page No: 2551 - 2561

[14]. Analysis of urban residential water consumption based on smart meters and fuzzy clustering. In 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing.

[15]. Malialis, K., Mavri, N., Vrachimis, S. G., Kyriakou, M. S., Eliades, D. G., & Polycarpou, M. M. (2024). Urban water consumption forecasting using deep learning and correlated district metered areas. arXiv preprint arXiv:2501.00158.

[16]. Papacharalampous, G., & Langousis, A. (2021). Probabilistic water demand forecasting using quantile regression algorithms. arXiv preprint arXiv:2104.07985.

[17]. VanBerlo, B., Ross, M. A. S., & Hsia, D. (2021). Univariate long-term municipal water demand forecasting. arXiv preprint arXiv:2105.08486.

[18]. Karnataka Water Demand Dataset, "Daily urban water consumption and real-time features for selected cities in Karnataka, India," 2021–2025. [Online]. Available: [Insert DOI or URL if publicly available]

[19]. "Karnataka urban water demand dataset," unpublished dataset, 2021–2025.