

## Index

S.NO	TOPIC	PAGE NO
1.	Introduction	7-9
	1.1 Motivation	8
	1.2 Objective	9
2.	Existing Work	10-12
3.	Topic of the work	13-28
	3.1 Introduction	13
	3.2 All About Data	13
	3.3 Mapping to ML/DL	15
	3.3.1 Type of Machine Learning Problem	15
	3.3.2 Metrics	15
	3.4 Exploratory Data Analysis	
4.	Conclusion & Future Work	28-30
5.	References	30-31

## **1. INTRODUCTION**

Efficient water management is a critical aspect of sustainable urban development, particularly as global populations rise and climatic patterns shift. Accurate forecasting of water demand plays a pivotal role in ensuring the optimised distribution of this essential resource. By predicting future consumption patterns, utilities and policymakers can mitigate the risks of water shortages, reduce operational costs, and enhance service reliability.

This report explores the methodologies and tools employed in forecasting water demand, emphasizing their application in optimizing water distribution systems. It highlights the importance of integrating advanced technologies, such as machine learning algorithms and real-time data analytics, to enhance prediction accuracy. Additionally, the report examines case studies and best practices from various regions to provide actionable insights for water utility managers and planners.

The discussion underscores the interdependence of demand forecasting and distribution network efficiency, illustrating how precise projections can lead to significant resource conservation, energy savings, and improved water access for communities. Ultimately, this analysis aims to contribute to the broader objective of achieving sustainable water resource management in the face of evolving challenges.

## **1.1.Motivation**

Water is a vital resource that supports life, sustains ecosystems, and drives economic and social development. However, the increasing global population, rapid urbanization, and the unpredictable impacts of climate change are placing immense pressure on water resources. These challenges necessitate the efficient management of water supply systems to ensure equitable access, minimize waste, and support sustainable development.

The process of forecasting water demand allows utility providers, policymakers, and engineers to anticipate future water needs and plan accordingly. This is crucial in minimizing the risks of over-supply, which leads to resource wastage and increased operational costs, as well as under-supply, which can disrupt daily life, hamper economic activities, and harm public health.

Moreover, the increasing adoption of smart water technologies and Internet of Things (IoT)-based solutions has created opportunities for more precise and dynamic demand forecasting. These innovations allow real-time monitoring of water consumption patterns and integration of diverse datasets, such as weather forecasts, population growth rates, and urban development plans.

The motivation for focusing on optimized water distribution through demand forecasting also stems from its environmental and economic benefits. Efficient distribution reduces energy consumption associated with water treatment and pumping, thereby lowering greenhouse gas emissions.

From an economic perspective, accurate demand forecasting minimizes financial losses by reducing costs associated with emergency responses, infrastructure overhauls, and water scarcity penalties.

In conclusion, forecasting water demand for optimized distribution is not merely a technical challenge; it is a critical component of sustainable water resource management. By addressing this need, we can contribute to the broader goals of environmental conservation, economic efficiency, and societal well-being, while ensuring that future generations have access to this indispensable resource.

## 1.2.Objective

The primary objective of this report is to explore and outline methodologies for forecasting water demand to achieve optimized distribution, ensuring the efficient, equitable, and sustainable management of water resources.

The focus is to develop a framework that integrates advanced predictive tools, innovative technologies, and actionable insights for addressing the challenges posed by growing water demand and constrained supply systems.

This report aims to identify and evaluate various forecasting models that cater to both short-term operational needs and long-term strategic planning.

The project aims to:

1. **Anticipate Future Water Needs:** Develop accurate models to predict short-term and long-term water demand based on various influencing factors.
2. **Optimize Resource Allocation:** Enhance the efficiency of water distribution systems by aligning supply with demand forecasts.
3. **Support Sustainable Development:** Minimize water wastage and ensure equitable distribution to address the needs of growing populations and urban areas.
4. **Enhance System Resilience:** Prepare for potential disruptions, such as climate change impacts, through dynamic and adaptable forecasting approaches.
5. **Promote Environmental Conservation:** Reduce the over-extraction of natural water sources and lower energy consumption related to water distribution.
6. **Ensure Economic Efficiency:** Decrease operational costs and mitigate financial risks associated with water scarcity or infrastructure inefficiencies.
7. **Foster Public Trust:** Provide consistent and reliable water supply, building confidence in water management systems and authorities.

Furthermore, the objective includes identifying strategies to implement these forecasting tools within water management frameworks to achieve sustainable and efficient resource allocation.

## **2. EXISTING WORK**

Numerous studies and technological advancements have focused on forecasting water demand for optimized distribution. These efforts span statistical, machine learning, and hybrid modelling approaches, as well as the integration of innovative tools like IoT and GIS-based systems.

### **1. Dataset Preprocessing**

1. Dataset Source: Groundwater consumption and rainfall data.
2. Feature Selection: Chose Water\_Consumption as the target variable and Rainfall as an influencing factor.
3. Data Cleaning:
  - Renamed columns (Water\_Consumption, Rainfall).
  - Handled missing values using fillna() and dropna().
  - Normalized data using MinMaxScaler() to scale values between 0 and 1.

### **2. Time-Series Data Preparation**

1. Converted dataset into sequences with a window size of 10 time steps.
2. Splitting Data: 80% for training, 20% for testing.

### **3. LSTM Model Architecture**

1. Why LSTM? Designed for sequential data and time-series forecasting.
2. Layers Used:
  - Two LSTM layers (each with 100 neurons).
  - Dense layers for output prediction.
3. Hyperparameters:
  - Optimizer: Adam
  - Loss Function: Mean Squared Error (MSE)
  - Epochs: 100
  - Batch Size: 16

### **4. Model Training & Evaluation**

1. Model Training: The LSTM model was trained using model.fit() for 100 epochs.
2. Performance Metrics:
  - Training Loss vs. Validation Loss curve was analyzed.
  - Actual vs. Predicted values were plotted for comparison.

### **5. Web Dashboard Deployment**

1. Flask + Plotly used to create an interactive dashboard.
2. Predictions stored in predictions.csv and displayed dynamically.
3. Users can view real-time forecasts via a web interface.

### 3. Proposed Work

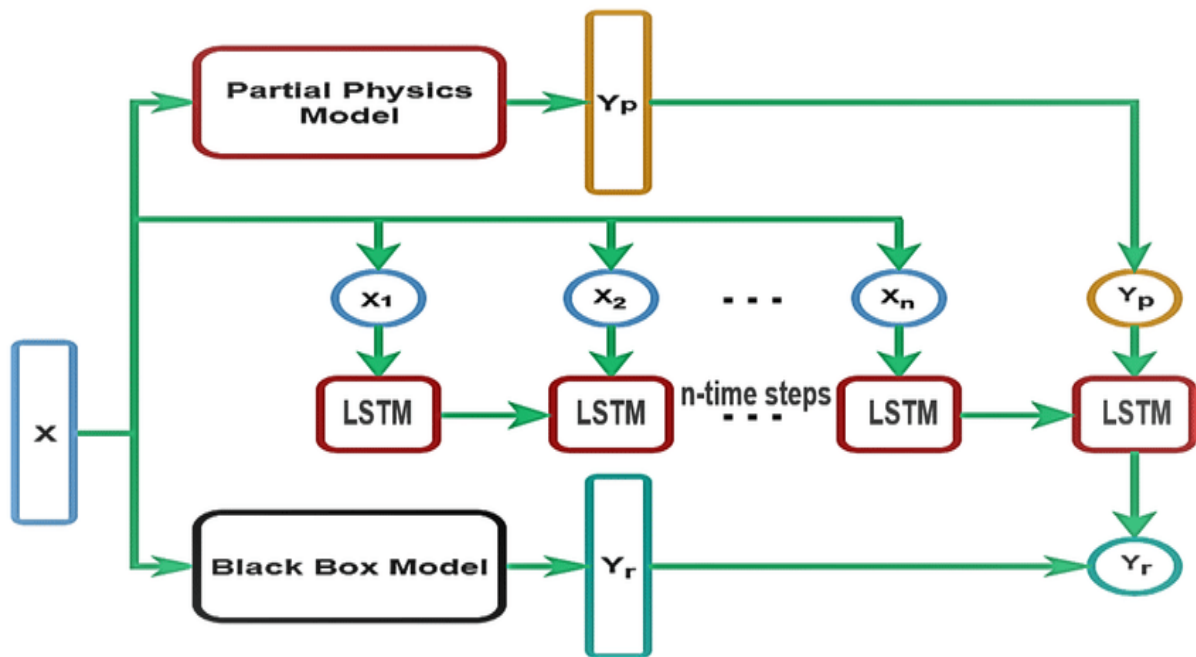


Figure 1 BLOCK DIAGRAM OF EXISTING PROJECTS

## **Topic of the Work**

### **3.1 Introduction**

Water management in rural areas has long faced challenges due to inefficiencies, limited resources, and a lack of data-driven decision-making tools. This project introduces a Predictive Analytics System for Water Demand and Distribution, which leverages historical data, artificial intelligence (AI), and machine learning (ML) to forecast water demand and optimize water distribution. By addressing issues such as unpredictable water consumption patterns and seasonal demand fluctuations, the proposed system ensures equitable and efficient water resource allocation.

The predictive system focuses on analyzing real-time and historical data, including population trends, agricultural usage, and weather conditions. Using advanced machine learning models, the project aims to minimize wastage, reduce shortages, and improve overall water management outcomes for rural regions.

### **3.2 All About Data**

The success of any machine learning model depends on the quality and diversity of data. This section discusses the sources, nature, and preprocessing techniques applied to the data used in the project.

#### **3.2.1 Data Sources The project utilizes a variety of data sources to ensure accurate predictions:**

1. **Historical Water Consumption Data:** Data on past water usage patterns for domestic, agricultural, and industrial purposes in rural areas.
2. **Weather Data:** Temperature, humidity, rainfall, and other climatic variables influencing water demand.
3. **Population and Socio-Economic Data:** Population growth, agricultural needs, and socio-economic trends impacting water requirements.
4. **Seasonal Trends:** Analysis of demand variations caused by seasonal changes.
5. **IoT and Sensor Data:** Real-time data collected from water meters and sensors for dynamic adjustments.

#### **3.2.2 Data Preprocessing** Data preprocessing is a critical step to prepare raw data for model training and analysis. The following techniques are applied:

1. **Handling Missing Values:** Imputation methods like mean/median replacement for missing records.
2. **Normalization:** Scaling data to ensure uniformity in units and magnitude.
3. **Outlier Detection and Treatment:** Identifying anomalies using statistical methods and ensuring data consistency.
4. **Feature Engineering:** Creating meaningful features such as water demand per capita, seasonal demand ratios, and weather-based indicators.

5. Time Series Formatting: Organizing data into time series format for accurate trend analysis and predictions.

### **3.3 Mapping to ML/DL**

The system maps the problem of water demand forecasting and distribution optimization to machine learning and deep learning techniques.

#### **3.3.1 Type of Machine Learning Problem**

The water demand prediction problem is framed as a Supervised Learning Regression Problem. Given historical and real-time data inputs, the model predicts continuous values, such as daily or weekly water demand.

#### **Key characteristics of the problem:**

1. Input Features: Population size, weather parameters, historical water usage, seasonal trends, and real-time sensor readings.
2. Target Variable: Predicted water demand (measured in liters/day or equivalent).

#### **3.3.2 Metrics**

To evaluate the performance of the predictive models, the following metrics are used:

For Artificial Neural Networks (ANN):

1. Root Mean Squared Error (RMSE): Measures the standard deviation of the residuals (errors) between predicted and actual values. A lower RMSE indicates better model performance.
2. Mean Absolute Error (MAE): Computes the average absolute difference between predicted and actual values, providing a direct measure of prediction accuracy.
3. R-Squared ( $R^2$ ): Evaluates how well the model explains the variance in the target variable.

#### **For Regression Models:**

1. Mean Squared Error (MSE): Measures the average squared differences between predicted and actual values, penalizing larger errors more heavily.
2. Adjusted R-Squared: Similar to  $R^2$  but adjusted for the number of predictors in the model, ensuring no overfitting occurs.
3. Mean Absolute Percentage Error (MAPE): Provides error as a percentage, useful for understanding relative prediction accuracy.

#### **For Support Vector Regression (SVR):**

1. Epsilon-Insensitive Loss: Measures how predictions fall within a margin of tolerance (epsilon), ignoring small deviations.
2. RMSE: Evaluates prediction accuracy by measuring error magnitude.



- Correlation Coefficient (r): Assesses the strength and direction of the relationship between predicted and actual values.


These metrics ensure that the system delivers reliable and precise forecasts, which are critical for optimizing water distribution.

### 3.4 Exploratory Data Analysis (EDA)

Water Consumption Data

(1)	(2)
<b>Rivers &amp; Canals (length in km)</b>	195210
<b>Other Water Bodies (area in Mha)</b>	
Reservoirs	2.91
Tanks & Ponds	2.41
Flood Plain Lakes & Derelict Water bodies	0.80
Brackish Water	1.24
<b>Total</b>	<b>7.36</b>
Source : Department of Animal Husbandry, Dairying & Fisheries, M/o Agriculture	

Figure 2 Inland Water resources in the country



<b>Gujarat</b>	3865	2.43	0.71	0.12	1	4.26
<b>Maharashtra</b>	16000	2.79	0.59	-	0.1	3.48
<b>Arunachal Pradesh</b>	2000	-	2.76	0.42	-	3.18
<b>Rajasthan</b>	5290	1.2	1.8	-	-	3.00
<b>Madhya Pradesh</b>	17088	2.27	0.6	-	-	2.87
<b>Others incl UTs</b>	84415	3.53	3.24	1.39	1.2	9.36
<b>TOTAL</b>	195210	29.07	24.14	7.98	12.4	73.59
Source : Department of Animal Husbandry, Dairying & Fisheries, M/o Agriculture						

Figure 3 Madhya Pradesh Individual Data

Table T3: States by total length of rivers and canals	
Length (km) (1)	Name of States/UT (2)
<500	Andaman & Nicobar Is, Chandigarh, Dadra & Nagar Haveli, Daman & Diu, Delhi, Goa, Lakshadweep, Puducherry
500-999	Sikkim
1000-1999	Mizoram, Nagaland, Tripura
2000-4999	Arunachal Pradesh, Assam, Bihar, Chhattisgarh, Gujarat, Himachal Pradesh, Jharkhand, Kerala, Manipur, Orissa, Uttarakhand, West Bengal
5000-9999	Haryana, Karnataka, Meghalaya, Rajasthan, Tamil Nadu,
10000-14999	Andhra Pradesh
15000-19999	Madhya Pradesh, Maharashtra, Punjab,
20000-24999	-
25000 & above	Jammu & Kashmir, Uttar Pradesh
Source: Department of Animal Husbandry, Dairying & Fisheries, M/o Agriculture	

Figure 4 States by total length of rivers and canals

Madhya Pradesh total length of river and canal is 16000 with 2.79 reservoirs 0.59 Tanks, and Pond, 0.1 Brackish Water with total of 8.11.

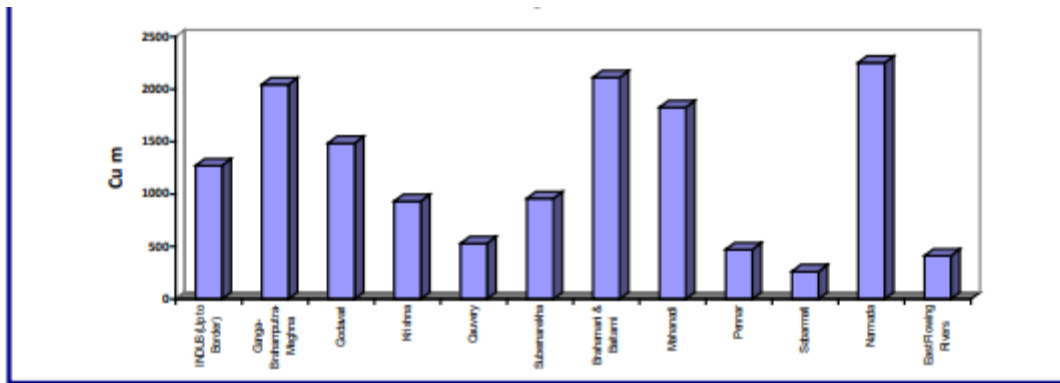


Figure 5 Estimated per capita availability of water in Different River Basins during 2010

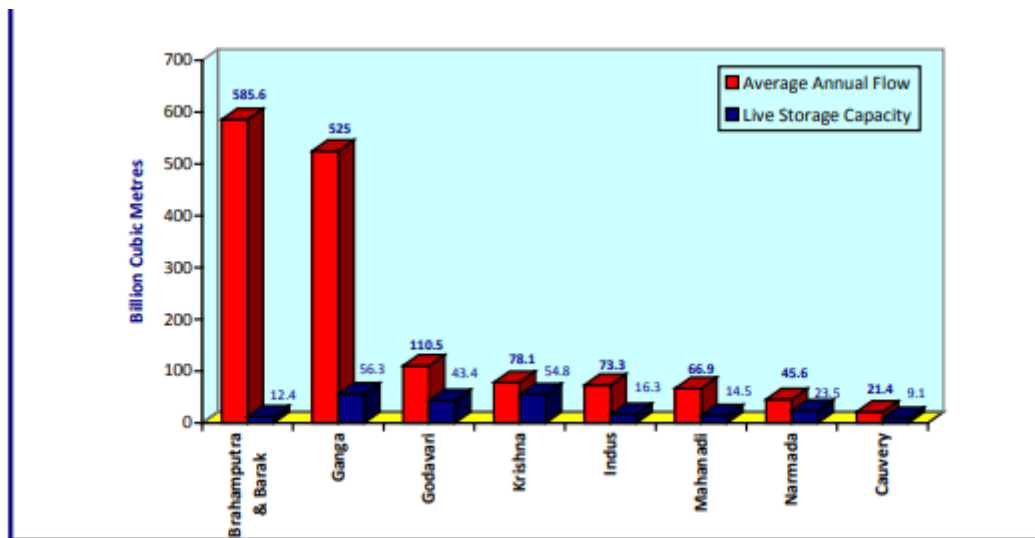


Figure 6 Basin Wise flow and storage potential in India

The States of Andhra Pradesh, Gujarat, Karnataka, Madhya Pradesh, Maharashtra, Orissa and Uttar Pradesh together account for about 72 % of total live storage capacity in the country.

State (1)	Annual Replenishable Ground Water Resources	
	(BCM / Year) (2)	%
Andhra Pradesh	33.83	7.8
Assam	30.35	7.0
Bihar	28.63	6.6
Chhattisgarh	12.22	2.8
Gujarat	18.43	4.3
Karnataka	16.81	3.9
Madhya Pradesh	33.95	7.9
Maharashtra	35.73	8.3
Orissa	17.78	4.1
Punjab	22.56	5.2
Rajasthan	11.86	2.8
Tamil Nadu	22.94	5.3
Uttar Pradesh	75.25	17.5
West Bengal	30.50	7.1
Others	40.19	9.3
Total	431.02	100.0

Sources : Central Ground Water Board, Min. of Water Resources

Figure 7 Annual Replenishable Ground Water Resources

It shows 14 States comprise 91% of ground water potential. Among the States, in terms of share of replenishable ground water resources Madhya Pradesh (7.9%).

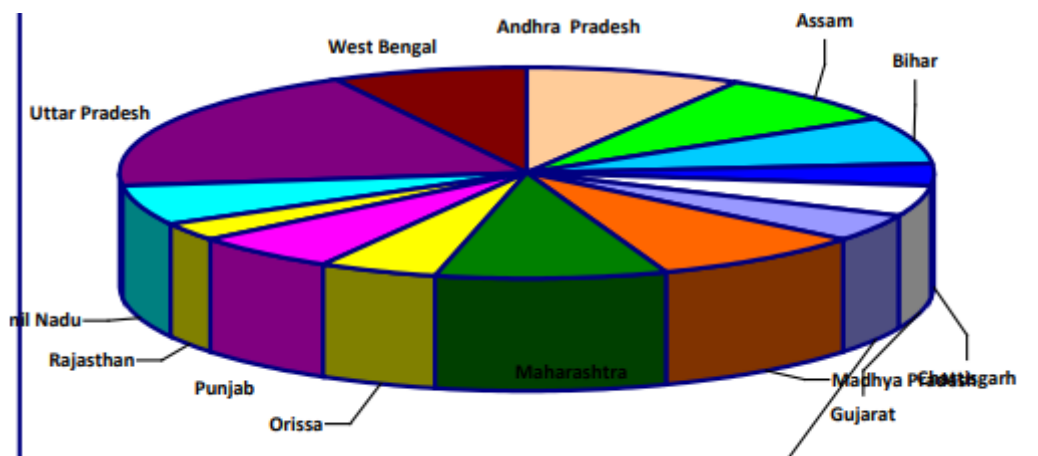


Figure 8 Annual Replenishable Ground Water Resources

Orange Denotes For Annual Replenishable Ground Water Resources of Madhya Pradesh

Table T11: Number of Dams by State (as on 30.11.2012)							
State/UT	Completed	Under Construction	Total	State/UT	Completed	Under Construction	Total
(1)	(2)	(3)	(4)				
A&N Islands	2	0	2	Maharashtra	1693	152	1845
Andhra Pradesh	290	44	334	Manipur	3	2	5
Arunachal Pr.	1	0	1	Meghalaya	5	1	6
Assam	3	2	5	Mizoram	0	0	0
Bihar	24	4	28	Nagaland	0	0	0
Chhattisgarh	243	14	257	Orissa	198	6	204
Goa	5	0	5	Punjab	14	1	15
Gujarat	621	45	666	Rajasthan	201	10	211
Haryana	1	0	1	Sikkim	2	0	2
Himachal Pr.	13	6	19	Tamil Nadu	116	0	116
J& K	12	2	14	Tripura	1	0	1
Jharkhand	49	28	77	Uttar Pradesh	114	16	130
Karnataka	230	1	231	Uttarakhand	13	6	19
Kerala	58	1	59	West Bengal	28	0	28
Madhya Pradesh	899	7	906				
				Total	4839	348	5187
Source: Central Water Commission, (Dam Safety Monitoring Directorate) National Register of Large Dam							

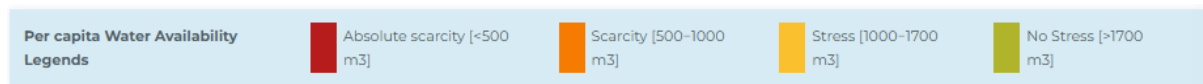
Figure 9 Number of Dams By state

The maximum number of dams completed in the country in Madhya Pradesh are 899 finished with 7 Under Construction.

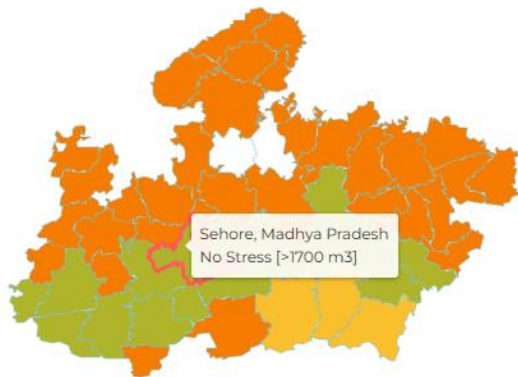
Sl.No.	Name of State	TOTAL LIVE STORAGE CAPACITY ( BCM)		
		Completed Projects	Under Construction Projects	Total
1	2	3	4	5
1	ANDHRA PRADESH	28.716	7.062	35.778
2	ASSAM	0.012	0.547	0.559
3	ARUNACHAL PRADESH	0.000	0.241	0.241
4	BIHAR	2.613	0.436	3.049
5	CHHATTISGARH	6.736	0.877	7.613
6	GOA	0.290	0.000	0.290
7	GUJARAT	18.359	8.175	26.534
8	HIMACHAL PRADESH	13.792	0.100	13.891
9	JAMMU AND KASHMIR	0.029	.000	0.029
10	JHARKHAND	2.436	6.039	8.475
11	KARNATAKA	31.896	0.736	32.632
12	KERALA	9.768	1.264	11.032
13	MADHYA PRADESH	33.075	1.695	34.770
14	MAHARASHTRA	37.358	10.736	48.094
15	MANIPUR	0.407	8.509	8.916
16	MEGHALAYA	0.479	0.007	0.486
17	MIZORAM	0.000	0.663	0.663

Figure 10 Total Live Storage Capacity

## Water Availability





Per capita Water Availability 2050



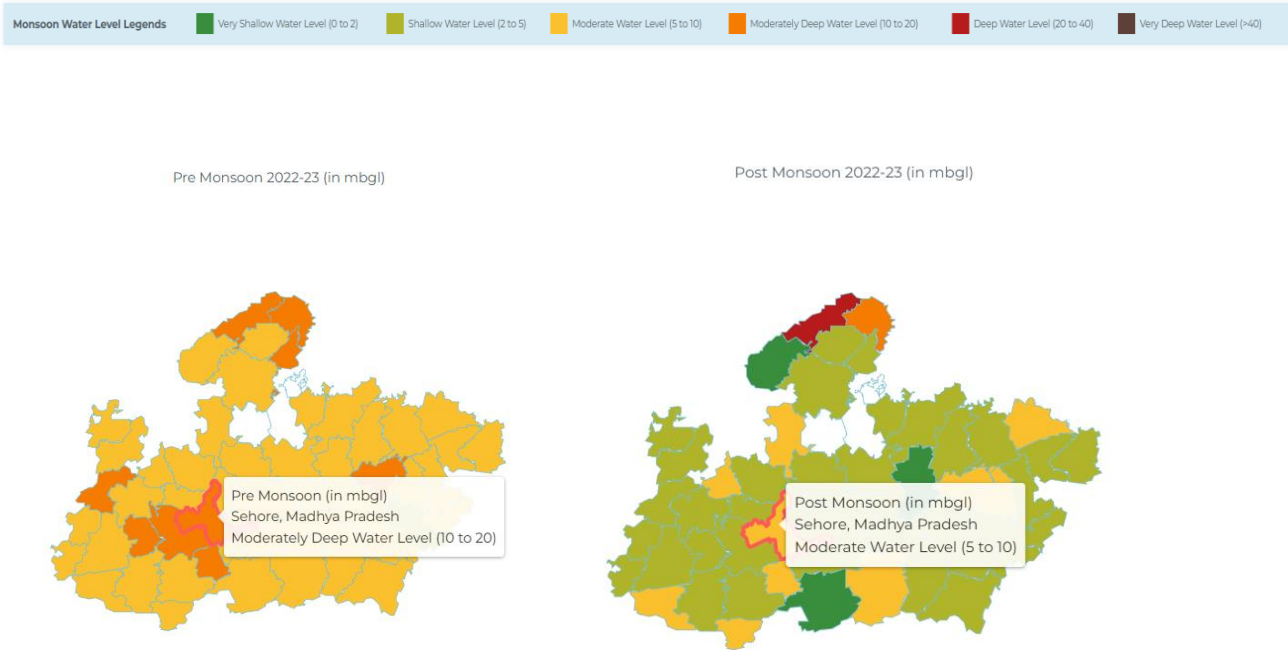
Per capita Water Availability 2025




**Per Capita Water availability for Sehore, Madhya Pradesh in the Year 2050 lies Under Stress category of 1000-1700m3** 

**Per Capita Water availability for Sehore, Madhya Pradesh in the Year 2025 lies Under Stress category of 1000-1700m3** 

## Pre and Post Monsoon Water Levels



**Pre-Monsoon Water availability for Sehore, Madhya Pradesh in the Year 2022-23 lies Under Moderately Deep Water level (10 to 20)** 

**Post Monsoon Water availability for Sehore, Madhya Pradesh in the Year 2022-23 lies Under Moderate Deep-Water level (5 to 10)** 

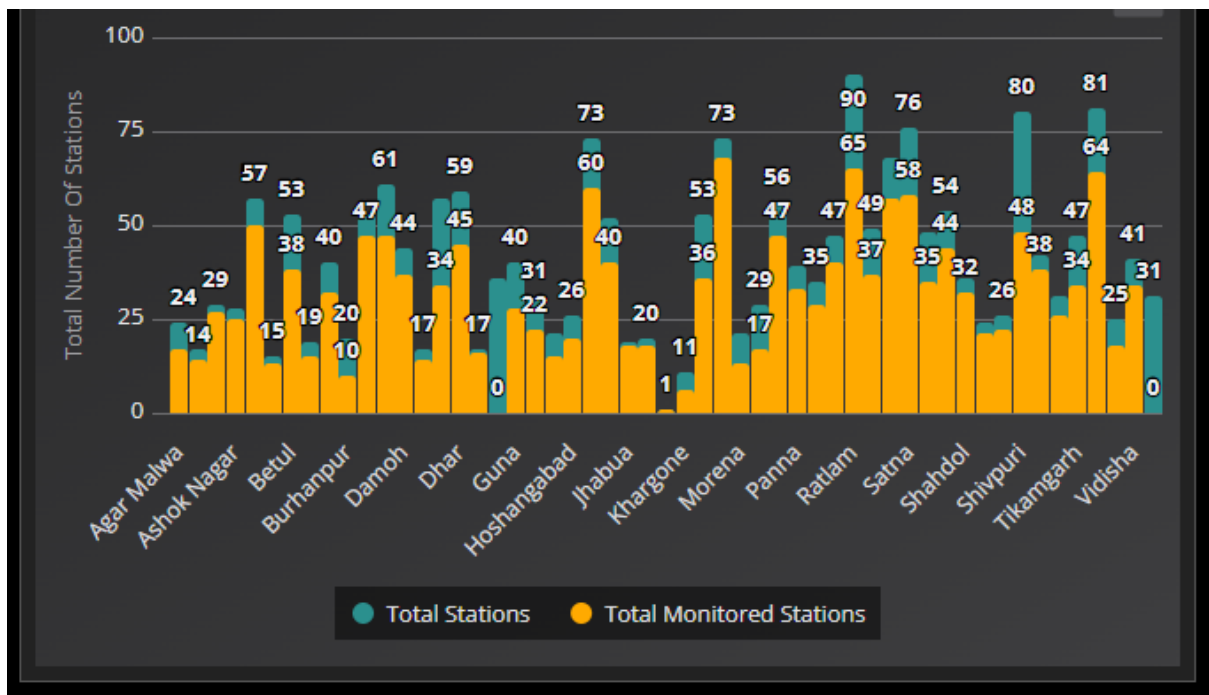


Figure 11 District Wise Station Count in Sehore, Madhya Pradesh

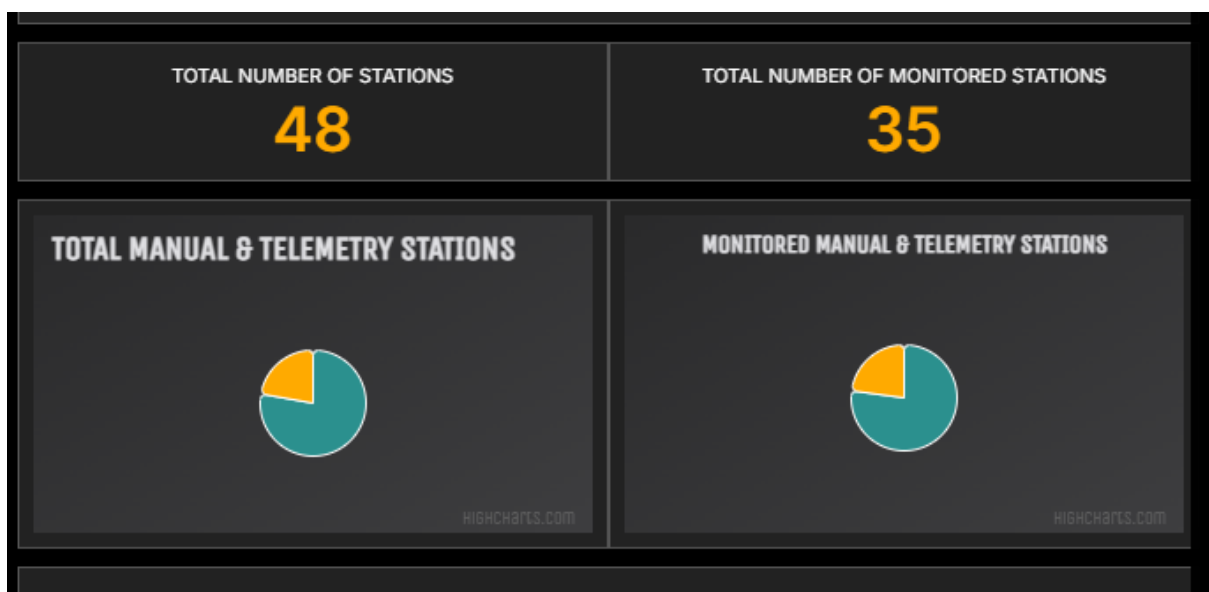


Figure 12 Monthly Ground Water Level Information

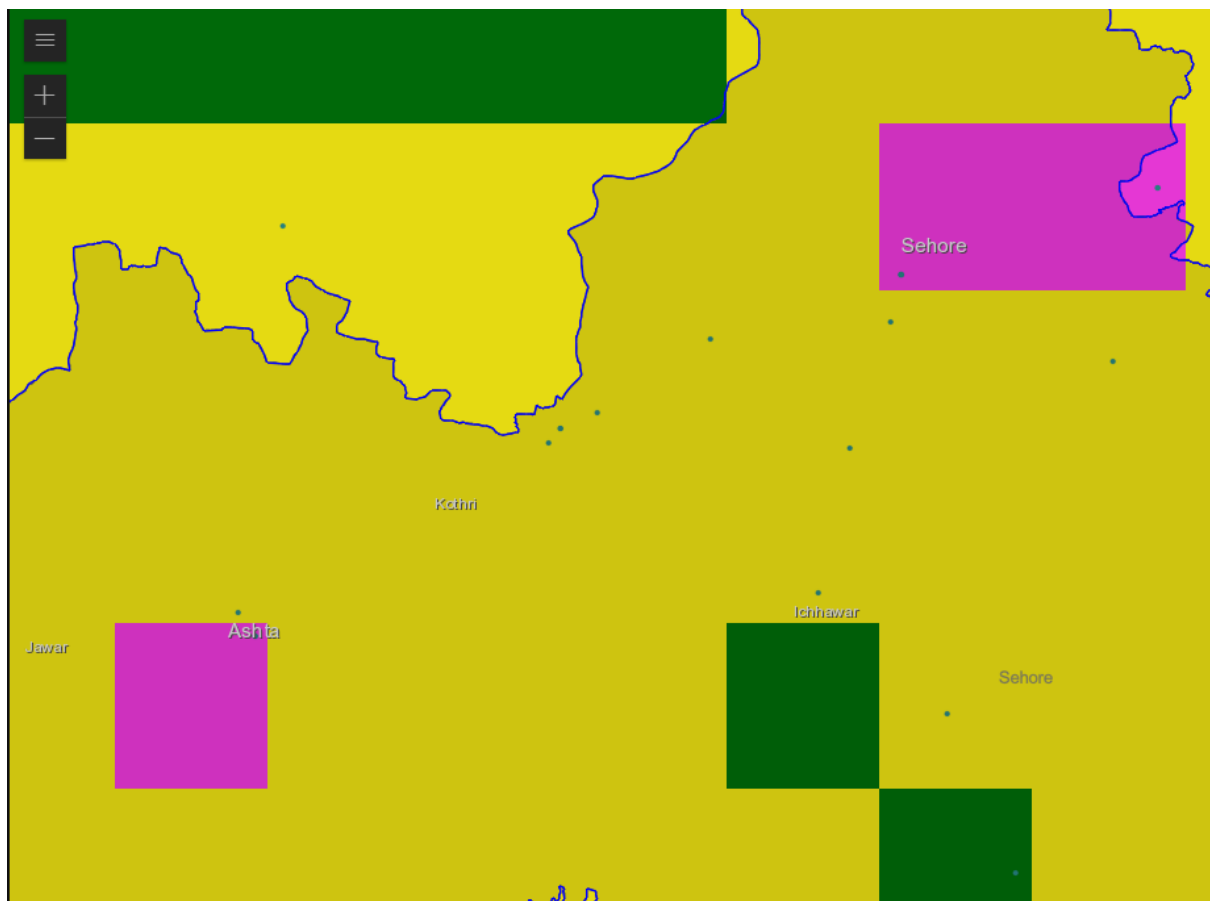
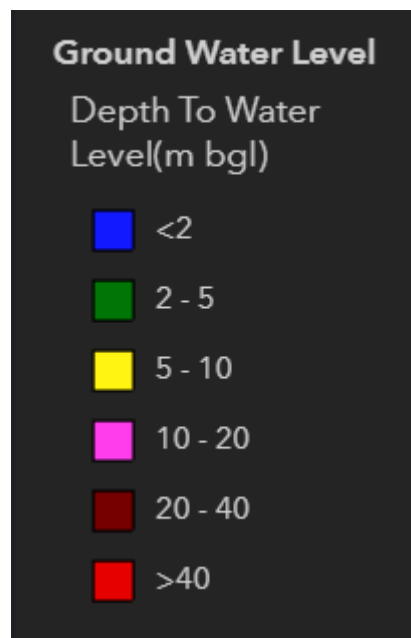


Figure 13 Ground Water Level in MP





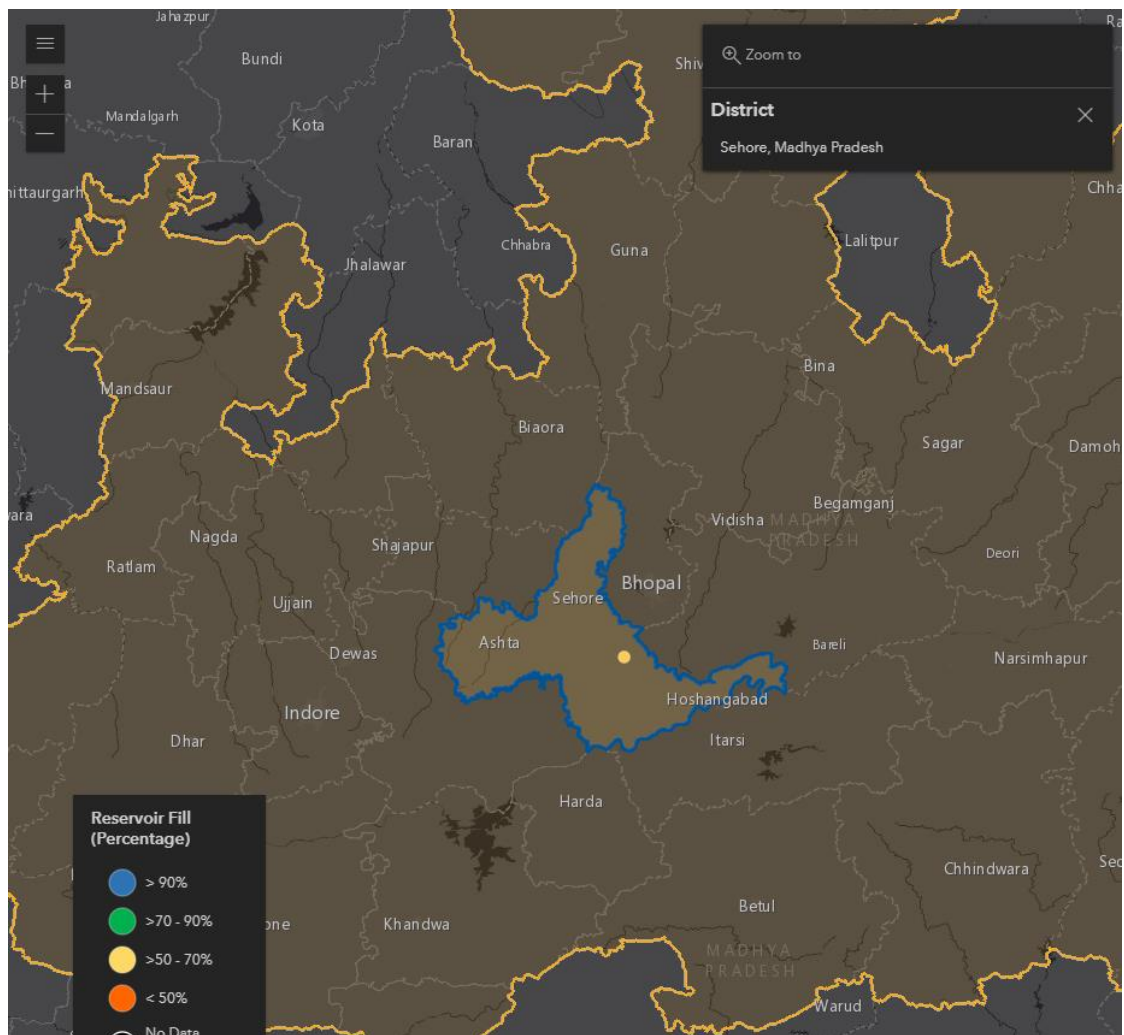
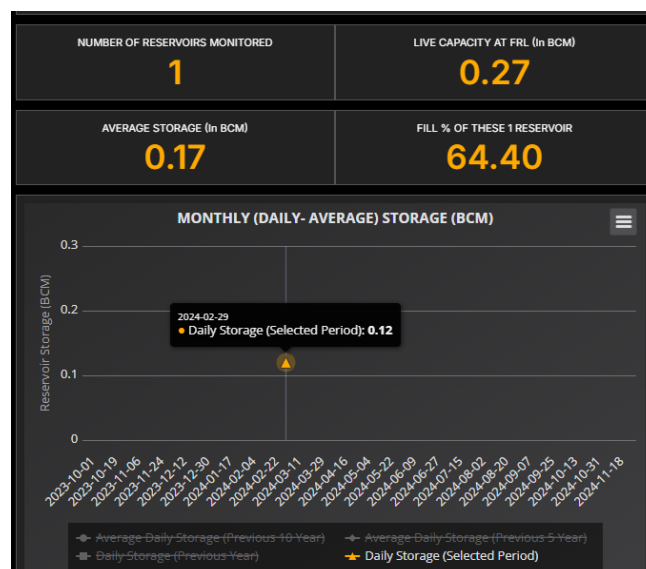


Figure 14 Reservoir Fill Percentage in Sehore, Madhya Pradesh lie Between 50-70%



## 4. Conclusion & Future Work

### 4.1 Conclusion

1. The LSTM-based forecasting model effectively predicts water demand, allowing for optimized resource distribution.
2. The Flask dashboard enables real-time visualization, making predictions accessible to decision-makers.

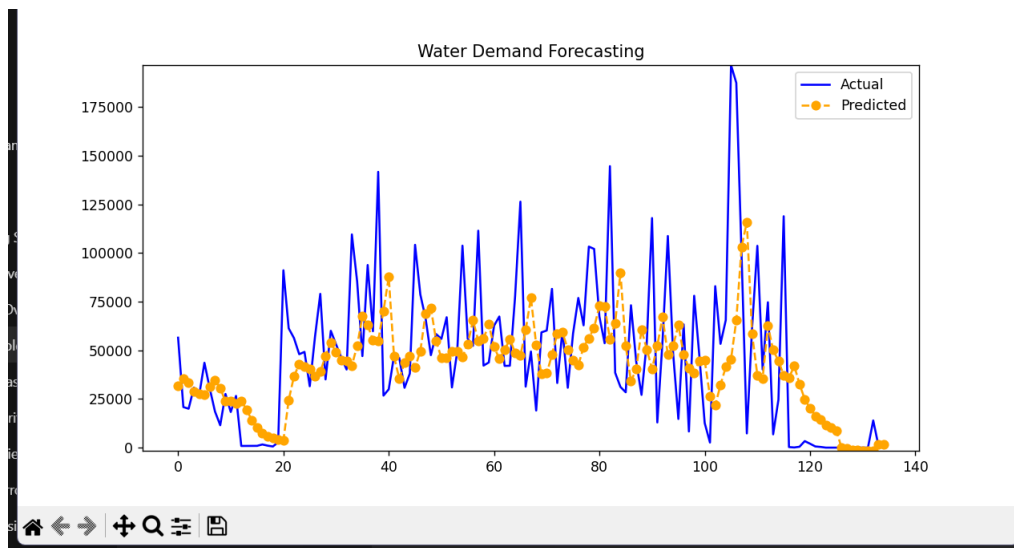


Figure 15 Water Demanding Forecasting



Figure 16 Loss Curve

## 4.2 Future Improvements

1. Enhance dataset: Include temperature, humidity, and population growth for more accurate predictions.
2. Optimize hyperparameters for improved model performance.
3. Deploy on a cloud platform like AWS/GCP for large-scale real-time forecasting.
4. Compare with other models like GRU, Transformer networks, or hybrid approaches.

## 5.Reference

1. H. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
2. J. Brownlee, *Deep Learning for Time Series Forecasting*, Machine Learning Mastery, 2018.
3. F. Chollet, *Deep Learning with Python*, Manning Publications, 2017.
4. G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
5. Y. S. Abu-Mostafa, M. Magdon-Ismail, and H. T. Lin, *Learning from Data*, AMLBook, 2012.
6. T. Asefa, et al., "Multivariate short-term water demand forecasting using artificial neural networks," *J. Water Resour. Plann. Manage.*, vol. 132, no. 1, pp. 10–20, 2006.
7. M. Herrera, et al., "Short-term water demand forecasting using artificial neural networks," *J. Hydrol.*, vol. 387, no. 1–2, pp. 28–38, 2010.
8. J. Adamowski and C. Karapataki, "Comparison of MLR and ANN models for urban water demand forecasting," *J. Hydrol.*, vol. 408, no. 1, pp. 28–34, 2011.
9. S. Mohan and S. Ahilan, "Water Demand Forecasting using Deep Learning," *Int. J. Sci. Res. Sci. Technol.*, vol. 5, no. 6, pp. 456–461, 2019.
10. S. Pande and M. Sivapalan, "Data and modeling innovations for water management," *Hydrol. Earth Syst. Sci.*, vol. 20, pp. 3541–3556, 2016.
11. A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Springer, 2012.
12. F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, 2000.
13. J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.
14. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

15. S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Proc. 32nd Int. Conf. Mach. Learn.*, 2015.
16. Central Ground Water Board (CGWB), India, "Annual Ground Water Report," 2020.
17. India Meteorological Department (IMD), "Rainfall Statistics of India," Govt. of India, 2021.
18. Ministry of Jal Shakti, "National Compilation on Dynamic Ground Water Resources," Govt. of India, 2022.
19. Kaggle, "Water Demand Forecasting Dataset," [Online]. Available: <https://www.kaggle.com>
20. WHO/UNICEF JMP, "Progress on Drinking Water, Sanitation and Hygiene," 2021.
21. TensorFlow, "TensorFlow Official Documentation," [Online]. Available: <https://www.tensorflow.org>
22. Keras, "Keras Documentation," [Online]. Available: <https://keras.io>
23. Wes McKinney, *Python for Data Analysis*, O'Reilly Media, 2017.
24. NumPy Developers, "NumPy Documentation," [Online]. Available: <https://numpy.org>
25. Matplotlib Developers, "Matplotlib Documentation," [Online]. Available: <https://matplotlib.org>
26. Scikit-learn Developers, "Scikit-learn Documentation," [Online]. Available: <https://scikit-learn.org>
27. Plotly, "Plotly for Python," [Online]. Available: <https://plotly.com/python/>
28. Flask, "Flask Documentation," [Online]. Available: <https://flask.palletsprojects.com>
29. GitHub, "Open-Source Water Forecasting Repositories," [Online]. Available: <https://github.com>
30. Python Software Foundation, "Python 3 Documentation," [Online]. Available: <https://docs.python.org/3/>

