International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com

# Advanced Data-Driven Approach for Forecasting Reservoir Water Levels and Optimizing Water Resource Management Through ARIMA and Geospatial Analytics

**Aniket V. Walunj[1], Sampada S. Swami[2]**

[1,2]B.Sc. Data Science Students, Department of Technology, Savitribai Phule Pune University, Pune, Maharashtra, India

**Abstract**
Effective management of water resources is crucial for sustainable development. This paper presents a comprehensive framework for analyzing and predicting reservoir water levels using time-series modeling and geospatial visualization. By leveraging historical water-level data, ARIMA models, and interactive visualization through Python and Streamlit, we provide actionable insights for policymakers and water resource managers to optimize water distribution, planning, and conservation strategies.

**Keywords:** Water Resource Management, Time-Series Forecasting, Geospatial Analysis, Streamlit Dashboard, ARIMA, Reservoir Level Prediction, Data Science, Python, Streamlit.

## I. INTRODUCTION

Water is a critical natural resource, fundamental for agriculture, industry, and domestic use. The increasing pressures of population growth, urbanization, and climate variability have amplified the need for efficient water resource management. Reservoirs play a central role in this ecosystem by storing water, regulating irrigation, and ensuring a stable supply for various consumption needs. The ability to accurately predict reservoir water levels allows authorities to proactively plan for droughts, mitigate flood risks, optimize irrigation schedules, and manage hydropower production. While traditional management methods often rely on historical observations and heuristics, these approaches may fail to capture complex temporal patterns. Consequently, data-driven approaches using
statistical models have become essential for modern water management.
This study focuses on developing a robust forecasting tool that is not only accurate but also accessible to decision-makers, thereby bridging the gap between raw data and actionable strategy.

## II. LITERATURE REVIEW

Time-series analysis is commonly used to model and forecast hydrological data, including reservoir water levels. The Auto-Regressive Integrated Moving Average (ARIMA) model is widely applied due to its simplicity and effectiveness in capturing linear trends, as detailed in Forecasting: Principles and Practice. While ARIMA performs well, advanced machine learning and deep learning models like Random Forests, XGBoost, and LSTM networks have been explored to capture nonlinear patterns and long-term dependencies. Geospatial visualization is also essential for analyzing reservoir locations and identifying areas with critical water levels. This study combines ARIMA with interactive geospatial tools using data from the Central Water Commission (CWC) of India.

## III. DATASET DESCRIPTION

The dataset used in this study consists of daily water level measurements for multiple reservoirs across India, with data spanning from January 2024 to March 2025. The data was sourced from public portals such as the Government of India's Central Water Commission (CWC) and India-WRIS. Each record in the dataset includes the following features:

- *Reservoir_name:* Name of the reservoir
- *Date:* Date of measurement
- *Level:* Water level in meters
- *Lat:* Latitude
- *Long:* Longitude

**TABLE I: Sample Reservoir Dataset**

| Reservoir | Date | Level | Lat | Long |
|---|---|---|---|---|
| Panshet | 01-01-2025 | 45.3 | 18.509 | 73.74 |
| Khadakwasla | 01-01-2025 | 52.1 | 18.530 | 73.802 |
| Temghar | 01-01-2025 | 38.7 | 18.600 | 73.75 |
| Mandira Dam | 14-01-2025 | 210.31 | 22.2667 | 84.6667 |
| Koyana | 29-01-2025 | 654.74 | 17.599 | 73.74 |

## IV. DATA PREPROCESSING

Data preprocessing is a crucial step to ensure accurate forecasting. Steps include:

1. **Date Conversion:** Converting 'Date' column to datetime format.
2. **Missing Value Handling**: Dropping rows with missing critical values.
3. **Reservoir Filtering:** Selecting a single reservoir for focused analysis.
4. **Chronological Sorting:** Sorting data by date to preserve temporal order.

```
import pandas as pd

# Load dataset
df = pd.read_csv("data.csv")

# Convert 'Date' column to datetime format
df['Date'] = pd.to_datetime(df['Date'], format='%d-%m-%Y', errors='coerce')

# Drop missing values
df.dropna(subset=['Date', 'Reservoir_name', 'Level'], inplace=True)

# Select a specific reservoir
selected_reservoir = 'Panshet'
```

```
df_reservoir = df[df['Reservoir_name'] = = selected_reservoir]
```

## V. METHODOLOGY

ARIMA was selected due to its effectiveness in capturing linear temporal dependencies in univariate time-series data.

ARIMA combines:

- **AR (p):** Autoregression
- **I (d):** Differencing for stationarity
- **MA (q):** Moving Average

We train the model using historical water-level data and forecast future values for a specified number of days.

**Listing 2: ARIMA Model Training**

```
from statsmodels.tsa.arima.model import ARIMA

data = d_reservoir['Level']
model = ARIMA(data, order=(5, 1, 0))
model_fit = model.fit()
forecast = model_fit.forecast(steps=30)
```

### B. Evaluation Metrics

Performance is evaluated using:

- **MAE:** Mean Absolute Error
- **RMSE:** Root Mean Squared Error
- **MAPE:** Mean Absolute Percentage Error
- **Accuracy:** 100 − MAPE

**Equations:**

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\, y_i - \hat{y}_i \,|$$

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

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n} \frac{|\, y_i - \hat{y}_i \,|}{y_i}$$

## VI. SYSTEM ARCHITECTURE

The entire system was developed as an interactive web application using Python, with Streamlit for the user interface and Plotly for visualizations. This dashboard provides a user-friendly platform for stakeholders to analyze reservoir data.

**System Components:**

1. **Reservoir Selection Dropdown** Allows users to choose specific reservoirs for analysis


**Figure 1: Selecting Reservoirs from Dropdown List**

2. **Time-Series Visualization** Displays historical water level trends interactively.

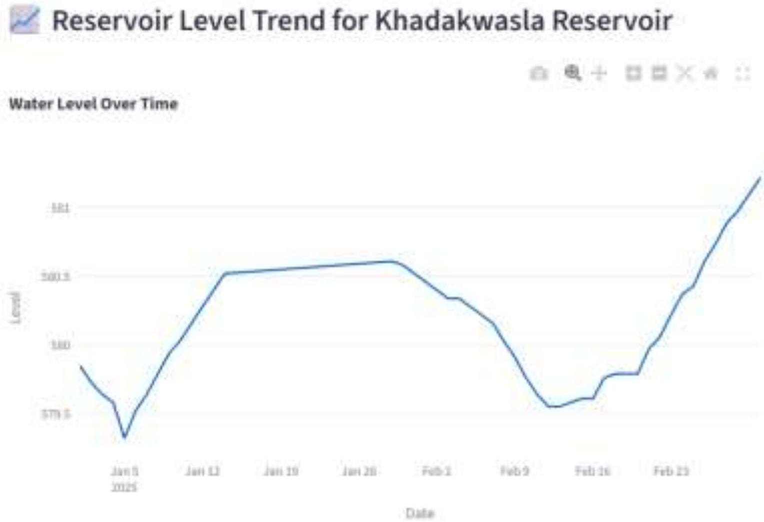


**Figure 2: Time-Series Line Plot of Historical Reservoir Water Levels**

3. **Geospatial Reservoir Map** Shows real-time locations of reservoirs with water level indicators.


**Figure 3: Interactive Map Showing Reservoir Locations**

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

4. **Model Tuning Panel** Sidebar sliders enable adjustment of ARIMA parameters (p, d, q).


**Figure 4: Sidebar Sliders for Adjusting ARIMA Model Parameters (p, d, q)**

## VII. RESULTS AND DISCUSSION

The system successfully integrates predictive modeling with visualization to provide meaningful insights.

1. **Forecast Visualization:** The model forecasts reservoir levels effectively, combining past and predicted data in a single chart.

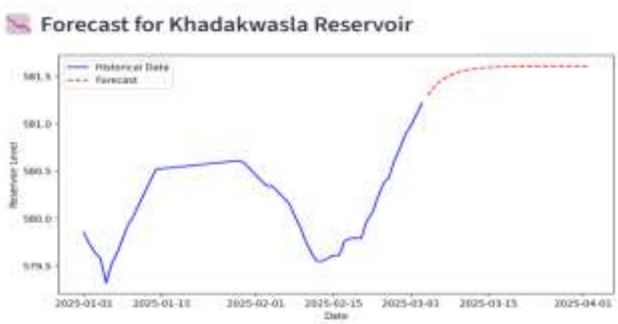


***Figure 5: Combined Historical + Forecast Water Level Plot***

2. **User Interface Output:**Streamlit displays the forecast results and provides an option to download data.

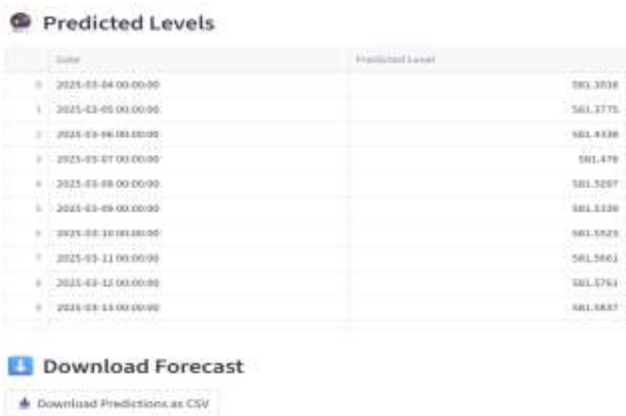


***Figure 6: Screenshot Showing Result & Download Button in Streamlit UI***

3. **Model Performance Metrics:** The evaluation metrics (MAE, RMSE, MAPE, Accuracy) confirm that the ARIMA model performs effectively for short-term water level prediction.

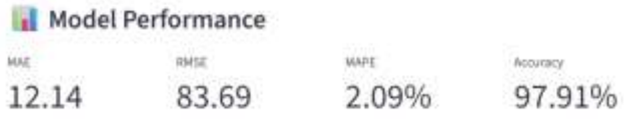


*Figure 7: Visual Display of Model Performance Metrics (MAE, RMSE, MAPE, Accuracy)*

## VIII. Conclusion and Future Work

This study presents a comprehensive framework for water resource management and forecasting using historical reservoir data. By integrating time-series analysis with the ARIMA model and interactive geospatial visualization, the project allows users to monitor water levels, forecast future availability, and make informed decisions for agriculture, urban water supply, flood management, and hydropower planning.

The Streamlit-based dashboard provides an interactive and user-friendly interface, enabling stakeholders to select reservoirs, explore historical trends, adjust forecasting parameters, and download predictions. Performance evaluation metrics such as MAE, RMSE, MAPE, and Accuracy demonstrate that the ARIMA model performs effectively for short-term water level prediction, making it a practical tool for policymakers and water resource managers.

By combining predictive analytics with visualization, this project bridges the gap between data-driven forecasting and actionable decision-making, ensuring better resource allocation and proactive water management strategies.

### Future Work

- Integrate real-time IoT sensor data from reservoirs to provide live water-level monitoring.
- Explore advanced models such as SARIMA, Prophet, or LSTM networks for improved long-term prediction accuracy.
- Incorporate rainfall, temperature, and climate data to enhance forecasting and understand environmental impacts.
- Develop a web-based dashboard for government agencies, enabling multi-reservoir analysis and automated alerts for low or critical water levels.
- Add predictive insights for irrigation planning, helping farmers optimize crop water usage and reduce waste.

## IX. References

1. Central Water Commission (CWC), Government of India. *Reservoir Level Data Portal*. https://www.cwc.gov.in/
   *Access to live and historical reservoir water level data across India.*
2. Open Government Data (OGD) Platform India. *Central Water Commission Datasets*. https://www.data.gov.in/keywords/CWC
   *Comprehensive datasets on reservoir levels and water storage.*
3. India-WRIS (Water Resources Information System). *Reservoir Data*. https://indiawris.gov.in/wris
   *Integrated water resources data, including reservoir levels.*

4. Water Resources Research Journal. *Recent Articles on Water Forecasting*. https://agupubs.onlinelibrary.wiley.com/journal/19447973
*Peer-reviewed journal featuring studies on water resource management and forecasting.*
5. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice (3rd Edition)*. https://otexts.com/fpp3/
*Comprehensive guide on time series forecasting methods. Foundational text on ARIMA modeling techniques.*
6. Streamlit Documentation. *Streamlit for Python*. https://docs.streamlit.io
*Official documentation for building interactive web applications.*
7. Plotly Documentation. *Interactive Graphing with Plotly in Python*. https://plotly.com/python/
*Resources for creating interactive visualizations.*