Journal of Artificial Intelligence and Metaheuristic

♦ ISSN: 2833-5597 Vol: 09, No. 01, 2025 ♦ PP. 08–14

Article DOI: https://doi.org/10.54216/JAIM.090102

FULL LENGTH ARTICLE

Predictive Analysis of Groundwater Resources Using Random Forest Regression

Khaled Sh. Gaber[1,1,*] **Manish Kumar Singla**[1]

[1] Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA

[2] Department of Interdisciplinary Courses in Engineering, Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India

Emails: khsherif@jcsis.org · manish.singla@chitkara.edu.in

Received: October 04, 2024 Revised: December 02, 2024 Accepted: January 13, 2025 ⋆ Corresponding author

ABSTRACT

The lack of water is one of the most crucial problems of our day; therefore, optimized water resource management and predictions gathered by patrons are of utmost importance. In the turmoil of a country like India, which lives a variety of lifestyles and has a complicated network of rivers, the urgent need for an active point of view to take care of water shortages becomes exceptionally vital. In this study, India's groundwater, available at the district level for the year 2017, was the area of focus, with this analysis utilizing a dataset of 689 rows, each representing a district, and 16 columns describing the various groundwater extraction and recharge metrics. The study involves five regression models adapting RandomForestRegressor, DecisionTreeRegressor, MLPRegressor, KNeighborsRegressor, and SupportVectorRegression for water resource evaluation and prediction. Every model is appraised by using a thorough metrics set where we incorporate Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Explained Variance Score (EVS), Max Error, Median Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), R-squared (R2), among others. Our results put the spotlight on RandomForestRegressor, making MSE measures the same as 0.000206624, endorsing its better performance versus the criteria considered. The approach used in this model provides us with an ensemble effect that makes it more robust in the sense that we can capture the interrelationships within the dataset in a comprehensive way. DecisionTreeRegressor also provides nice options for precision and transparency. The use of such models depicts the potential value of predictive analytics, which has the role of improving resource management and planning because we can all agree that the impending water crisis is also a fact. These research outcomes provide us with important data for well-informed decision-making and strategic management of water reserves through all avenues and most affected areas to air most of the impact of water scarcity.

Keywords: Groundwater ▪ Random forest Regression Model ▪ Groundwater production ▪ Water security

1. INTRODUCTION

Water is the foundation of life, an indispensable resource that supports ecosystems, drives economic growth, and sustains human civilizations. It is essential for drinking, sanitation, agriculture, energy production, and industrial processes. However, in recent decades, the world has been grappling with an escalating water crisis driven by a combination of rising demand, population growth, urban expansion, industrial development, and the impacts of climate change. These factors have placed unprecedented pressure on global water resources, with groundwater playing a central role in mitigat-

© 2025 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

ing the growing demand for water [1].

Groundwater serves as a vital source of freshwater for billions of people worldwide. It contributes significantly to agricultural irrigation, drinking water supplies, and industrial processes. In many regions, particularly in arid and semi-arid areas, groundwater is often the only reliable water source during periods of drought or reduced surface water availability. Yet, despite its importance, groundwater faces significant challenges, including over-extraction, contamination, and mismanagement. Over-extraction, in particular, has led to severe consequences, such as declining water tables, land subsidence, and the intrusion of saline water into freshwater aquifers. Pollution from industrial discharges, agricultural runoff, and urban wastewater further exacerbates the problem, compromising water quality and reducing its usability; machine-learning reviews describe the growing role of predictive models in assessing groundwater quality [2].

The growing complexity of groundwater issues underscores the urgent need for a deeper understanding of its dynamics at the regional and local levels. Studies have applied GIS, remote sensing, and machine learning to map groundwater potential in specific Indian regions, including Bundelkhand. However, potential-zone maps do not by themselves provide an integrated, district-level account of groundwater availability, extraction, and recharge. This distinction is important in rapidly developing nations, where uneven distribution and unregulated usage of groundwater resources can create significant disparities in access and sustainability. Policymakers, resource managers, and local authorities need detailed data to implement effective groundwater management strategies and support long-term water security [3].

This study seeks to address these critical gaps by conducting a comprehensive analysis of groundwater resources at the district level. Our objectives are to provide a detailed understanding of groundwater availability, identify patterns of usage and recharge, and evaluate the impacts of both natural and anthropogenic factors on groundwater dynamics. The research leverages advanced geospatial technologies, hydrological modeling, and robust data analytics to generate actionable insights into the state of groundwater resources. By doing so, this study aims to inform the development of targeted, evidence-based policies and strategies for sustainable groundwater management [4].

A key aspect of this research is its emphasis on spatial and temporal variability. Groundwater levels and quality can vary significantly within and across districts due to factors such as geological formations, land use patterns, rainfall distribution, and human activities. This study examines these variations to provide a nuanced understanding of groundwater dynamics. Additionally, the research explores the role of climate change, which is expected to intensify water-related challenges through changes in precipitation patterns, increased evaporation rates, and greater frequency of extreme weather events. Understanding these impacts is crucial for developing adaptive management strategies that can mitigate the effects of climate change on groundwater resources.

To guide this investigation, the study poses several critical questions:

- What is the current status of groundwater resources at the district level, including variations in quality, quantity, and recharge rates across specific regions?
- What are the natural and human-induced factors that influence groundwater availability, movement, and sustainability?
- How can advanced tools, such as geospatial information systems (GIS) and groundwater models, enhance the assessment and monitoring of groundwater resources over time?
- What are the potential impacts of climate change on groundwater availability and sustainability, and what strategies can be implemented to address these challenges at the district level?

This study's findings have broad implications for water resource management and policy development. By providing a detailed, district-wise analysis of groundwater resources, the research aims to equip decision-makers with the knowledge and tools needed to address water-related challenges effectively. The insights gained from this study will help in designing sustainable groundwater management strategies, improving resource allocation, and ensuring water security for future generations. Additionally, the research seeks to foster greater awareness and collaboration among stakeholders, including policymakers, resource managers, scientists, and local communities, to promote sustainable water resource management.

The structure of this paper is as follows: Section 2 provides a review of the relevant literature, highlighting existing research and identifying gaps. Section 3 describes the dataset and methodology used in the study, including data collection, preprocessing, and analysis techniques. Section 4 presents the results of the study, including district-level analyses of groundwater resources and the performance of various predictive models. Finally, Section 5 discusses the implications of the findings and offers recommendations for sustainable groundwater management, along with a summary of the study's contributions and future research directions.

By addressing the critical issues of groundwater availability, sustainability, and management, this study contributes to the broader goal of achieving water security in the face of growing demand and environmental uncertainties. It is hoped that the findings of this research will serve as a valuable resource for stakeholders and policymakers, guiding efforts to ensure the sustainable use and protection of groundwater resources in the years to come.

2. LITERATURE REVIEW

By 2020, half of India's population will live in big cities with substantial groundwater stress. Overexploitation, urbanization, population increase, and climate change have depleted urban groundwater resources. [5] uses machine learning to analyze how these variables affect groundwater levels in Bengaluru using Singular Spectrum Analysis (SSA), Mutual Information Theory (MI), Genetic Algorithm (GA), Artificial Neural Network (ANN), and Support Vector Machine (SVM). The pre-processing of input data consists of the following: monthly population growth rate, groundwater levels,

precipitation, temperature, Net Overland Infiltration (NOI), Southern Oscillation Index (SOI), and El Niño 3 (NI´aO3). ANN and SVM algorithms provide reliable monthly groundwater level forecasts ($R > 85\%$) using optimized input sets. Pre-processing improves 66% of monitored wells' accuracy. A favorable association between population expansion and diminishing groundwater levels provides a forecasting tool for metropolitan regions without leakage data for physical-based modeling.

Likewise, the conventional method for evaluating irrigation water quality in developing societies is both uneconomic and strenuous. These tasks contribute to the designing of an integrated approach using physical components and artificial intelligence algorithms to interpret in situ water quality indicators for irrigation from aquifers. The model depends on using Electrical Conductivity (EC), Temperature (T), and pH for predicting parameters like Total Dissolved Solids (TDS), Potential Salinity (PS), Sodium Adsorption Ratio (SAR), and other values. The sentences with Morocco's Berrechid aquifer data, Adaboost, and RF are ranked higher than Random Forest, Artificial Neural Network, and Support Vector Regression. Nevertheless, ANN and SVR models are also more generalized, and they are more immune to variables from input. These models greatly help farmers and decision-makers by predicting a low-cost, genuine, real-time irrigated water quality forecasting system based on improvement in physical attributes [6].

Globally, groundwater salinization threatens water availability and quality. Because salinity reflects interacting natural and human influences, its prediction benefits from careful feature selection and model evaluation. Yassin et al. studied groundwater salinization in the Eastern Province of Saudi Arabia by combining feature-engineering methods with Gaussian Process Regression, Support Vector Regression, regression trees, and robust linear regression. Their results show the potential of data-driven models for salinity assessment while underscoring the value of selecting informative input features [7].

For irrigation, water supply, and land development, groundwater level forecasting (GWL) is essential. Even while eXtreme Gradient Boosting and Random Forests, in particular, are good at predicting GWL, multiscale developments like urbanization present difficulties. [8] resolves these issues by combining ML with wavelet transformations (WT). However, real-world usability suffers when important WT concerns are ignored. In Kumamoto City, Southern Japan, this study shows how modern ML approaches may be properly coupled with WT to forecast GWL (1–3 months ahead), addressing important concerns, including land subsidence and groundwater quality.

The implementation of better freshwater management, especially in arid areas, requires the regeneration of groundwater. [9] studies the SVM, MARS, and RF MLAs to predict more accurately while creating surface water recharge potential maps (GRPMs). The eighteen factors incorporated by GRPMs include elevation, slope angle, TWI, fault density, roughness degree, drainage density, wetness index, lithology, land use, annual evaporative transposition, minimum temperature, maximum temperature, and 24-hour rainfall based on soil samples classified into training sets (70%) and validation sets (30%). The LASSO method for estimating the influences of various independent factors is employed in this model. Validation metrics include ROC-AUC, TPR, FPR, F-measures, fallout, sensitivity, specificity, TSS, and CCI. Results show RF outperforms SVM and MARS (AUC = 0.962 and 0.987, respectively). Since GRPM has unmatched reliability, policymakers could use it as a dependable guide in sustained planning and management of groundwater resources.

Sustainable groundwater resource management requires accurate vulnerability mapping. The GALDIT groundwater vulnerability mapping system is improved by combining Bootstrap Aggregating and Disjoint Aggregating resampling methods with ML models (XGBoost, LGBM, AdaBoost, CatBoost, and RF). GALDIT includes groundwater occurrence (G), hydraulic conductivity (A), depth to groundwater level (L), seashore distance (D), saltwater intrusion status impact (I), and aquifer thickness. ML techniques and resampling handle GALDIT's weightings and ratings' subjectivity and tiny datasets. Total dissolved solid concentrations were used to alter vulnerability indicators in the Iranian Shabestar Plain aquifer using ML models. For solo (XGBoost-GALDIT) and linked resampling (BA-XGBoost-GALDIT, DA-RF-GALDIT) ML models, correlation coefficients (r) and distance correlation (DC) metrics improved significantly compared to GALDIT. Integrating resampling and ML models improves GALDIT vulnerability accuracy [10].

Both rapid groundwater loss quality and natural sources are pressing global issues. Groundwater is lost through aquifer abstractions treated for agriculture, industry, and potable water and degraded by pollution. A smart, efficient, and sustainable framework for groundwater management is dependent on local to regional-level data and modeling. With the advent of AI in hydrological technologies, they tend to surpass classical models in recent years. [11] highlights the role of AI in groundwater modeling through an Indian national example, illustrating how AI can solve present problems and predict future challenges. Groundwater reservoirs in agricultural regions are affected differently by nonlinear dynamics due to the complicated relations between surface water supply, irrigation, groundwater extraction, and climate.

A machine learning-based groundwater ensemble model, including artificial neural networks, support vector machines, and response surface regression, is presented to address uncertainty. Such a method improves groundwater storage prediction accuracy due to its computational efficiency compared to physical models. Bayesian model averaging measures and accounts for parameter and structural uncertainty of ML models, delivering prediction distributions. The ensemble model is effective for the over-diverted San Joaquin River Basin, which works with great efficiency and improved computational economy. The results demonstrate groundwater storage is mainly affected by agricultural irrigation.

Furthermore, climate change and overexploitation are increasing groundwater demand, essential for agriculture, society, and the economy. [12] defines groundwater potential zones (GWPZ) in Kerala, India, over several months. GIS, Analytical Hierarchical Process (AHP), and Empirical Orthogonal Function (EOF) are integrated into a hybrid model that takes four layers of climate change and eleven layers of surface characteristics into account. Kerala has 13% vulnerable zones

(very low and low potential). With over 82% accuracy, a large random sample validation using measured groundwater levels outperforms the fundamental ranking approach with less than 60%. Agriculturists, geologists, and hydrologists can plan better with hybrid-coupled model results.

The state of Bengal Basin's aquifers is becoming terminally poorer because of the contamination of shallow water. Without previous aquifer development, pathways and durations of field geologic conditions should be studied. This model, which is three-dimensional, imitates satellite transport on the ground and shows that hydrogeologic factors exert the main influence on flow lines and transit time. Repeated horizontal aquifer irrigation may prevent flows to deeper ones. Water through aquitards and windows might flow, but groundwater that is neither arsenic nor chloride-rich can enter deep aquifers under a cumulative effect [13].

3. DATASET

3.1 Dataset Description

The "District-wise Ground Water Resources by July 2017" dataset is a relatively voluminous data collection that analyzes groundwater extraction and replenishment technology in various Indian districts. With this dataset, the exclusive details of water regulation, district by district, can be analyzed. This data provides an interesting statistic about the water status in July of 2017 during the meteorological conditions, various geographical regions, and the misuse of water illegally.

For the assessment of numerical data on groundwater extraction for different purposes and the dependence of such extraction recharging rates on both natural and anthropogenic factors, the data is carefully laid out and dispatched by district. Assessment of the sustainability of previous water consumption patterns, creation of water-saving policies, and identification of regions in danger of groundwater depletion are better achieved using this knowledge. Finally, the dataset is a helpful source of information for scholars, policymakers, and stakeholders aiming to advance effective solutions in the face of water challenges and foster a sustainable future in India.

3.2 Preprocessing Steps

An essential aspect of working with the "District-wise Ground Water Resources by July 2017" dataset is data preprocessing, which ensures the accuracy and dependability of subsequent studies. After intensive data cleansing, data integrity was ensured. Missing values were completed using imputation, while duplicates were eliminated.

Normalization was applied to account for different rates of groundwater withdrawal and the amount of water replenishment across districts, enabling meaningful comparisons. District names were standardized in alignment with other data attributes such as population and rainfall ratings, facilitating better data quality.

Groundwater management patterns and distinctive features were analyzed using class numbering and timing coordination. The final stage of the preprocessing involved verification and validation by comparing the processed data to official records, ensuring accuracy. These preprocessing steps constitute the groundwork for reliable and valuable analyses of water management and resources in India.

3.3 Descriptive Statistics

Descriptive statistics summarize data efficiently and comprehensively, revealing trends and patterns in district-wise groundwater resources across India. The "July 2017" dataset provides insights through measures such as mean, median, and standard deviation, identifying trends in attributes, variations, and distributions.

This foundation supports advanced statistical analyses and informed water resource management. For instance, the water resource extraction rate across Indian states is visualized using histograms, such as in Figure 1. These plots highlight urban districts with over-critical water extraction rates, illustrating unsustainable water usage patterns. Groundwater extraction trends appear unbalanced, with peaks and valleys signifying the variability in resource usage. Such visualizations underscore the need for tailored water management policies.

In summary, the synthetic description of complex data parameters is achieved by computing the mean, variance, and percentiles of groundwater extraction and recharge rates. Statistical graphs further illustrate distributions across districts, providing concise yet comprehensive summaries for future analyses. Descriptive statistics play a vital role in simplifying data and offering insights into groundwater challenges and opportunities across regions in India. Such processes assist stakeholders in projecting strategies for sustainable groundwater management and conflict resolution.

4. RESULTS

This section presents the findings of our study of district-level groundwater resources using regression models. The dataset includes 689 districts in India, each represented by 16 variables related to annual water exposure and recharge activities for the 2017 financial year. The study aimed to analyze the performance of various regression models in assessing groundwater resources based on this dataset.

We evaluated five regression models: Random Forest Regressor, Decision Tree Regressor, MLP Regressor, K-Nearest Neighbors Regressor, and Support Vector Regressor. Each model was trained and tested using the dataset, and performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Explained Variance Score (EVS), Max Error, Median Absolute Error (MedAE), Mean Absolute Percentage Error (MAPE), and R-squared (R2) were computed. This section provides a detailed summary of each regression model's performance, interpretations, and implications for groundwater use predictions at the district level in India.

4.1 Machine Learning Models

In this study, we implemented various regression models to infer and predict groundwater resources in Indian districts for 2017. Each regression model was selected based on its ability to capture diverse dimensions of water resource dynamics. Each model is described below.

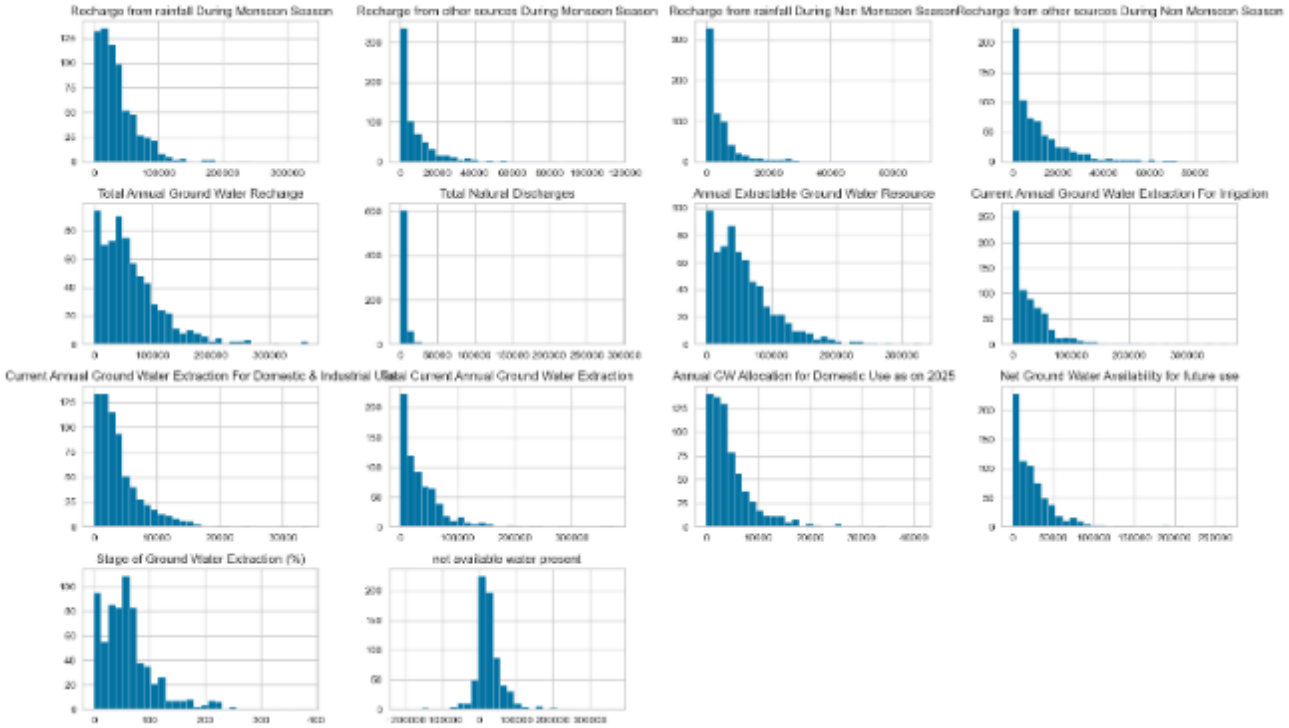


Figure 1. Histogram of Groundwater Extraction Rates Across Districts.

Model	MSE	RMSE	MAE	EVS	Max Error
Random Forest Regressor	0.000207	0.014374	0.006704	0.978487	0.127185
Decision Tree Regressor	0.000312	0.017663	0.011296	0.967859	0.113979
MLP Regressor	0.001803	0.042461	0.024618	0.812035	0.238271
K-Nearest Neighbors Regressor	0.001987	0.044575	0.020945	0.808075	0.295732
Support Vector Regressor	0.003282	0.057293	0.053145	0.917645	0.125943

Model	MedAE	MAPE	R^2	MTD	RRMSE
Random Forest Regressor	0.003019	1.122193	0.978401	0.000207	0.02741
Decision Tree Regressor	0.007392	3.721257	0.967388	0.000312	0.03368
MLP Regressor	0.014781	21.25771	0.811528	0.001803	0.08096
K-Nearest Neighbors Regressor	0.009586	2.397753	0.792295	0.001987	0.08499
Support Vector Regressor	0.055464	76.74634	0.656868	0.003282	0.10925

Model	Willmott	MBE	SD	r	NSE
Random Forest Regressor	0.988861	0.000908	0.014346	0.989141	0.978401
Decision Tree Regressor	0.967388	-0.00212	0.017535	0.983559	0.967388
MLP Regressor	0.977853	0.002202	0.042404	0.900849	0.811528
K-Nearest Neighbors Regressor	0.128322	-0.01229	0.042849	0.89011	0.792295
Support Vector Regressor	0.60962	0.049947	0.028068	0.810474	0.656868

Table 1. Regression Model Performance Metrics.

- **Random Forest Regressor:** An ensemble learning method that constructs multiple decision trees during training and averages the outputs to calculate the final regression value.
- **Decision Tree Regressor:** A non-linear model that splits data into subsets using a tree-like structure to predict the target variable.
- **MLP Regressor:** A type of artificial neural network with multiple layers of nodes (neurons) designed to capture complex, non-linear interactions in the data.
- **K-Nearest Neighbors Regressor:** A non-parametric algorithm that predicts values based on the closest data points to the test point.
- **Support Vector Regressor (SVR):** A multidimensional algorithm that transforms data into a higher-dimensional space and identifies an optimal plane for regression.

The following sections detail the performance of each regression model and their effectiveness in predicting groundwater resources.

4.2 Performance Metrics

The performance of each regression model was assessed using the following metrics:

1. **Mean Squared Error (MSE):** Measures the mean squared differences between actual and predicted values. A smaller MSE indicates better model performance.
2. **Root Mean Squared Error (RMSE):** The square root of MSE, providing an interpretable measure of error in the same units as the target variable.
3. **Mean Absolute Error (MAE):** The average absolute difference between actual and predicted values, indicating model accuracy.
4. **Explained Variance Score (EVS):** Represents the proportion of variance in the target variable explained by the model.
5. **Max Error:** The maximum difference between predicted and actual values in the dataset.
6. **Median Absolute Error (MedAE):** The median of absolute differences, less sensitive to extreme values than MAE.
7. **Mean Absolute Percentage Error (MAPE):** The average percentage difference between actual and predicted values.
8. **R-squared (R2):** Indicates the proportion of the target variable's variance that is explained by the model.

4.3 Key Findings

Table 1 summarizes the results of each regression model. Key findings are as follows:

- **Random Forest Regressor and Decision Tree Regressor:** These ensemble models outperform others with the lowest MSE, RMSE, and MAE values, indicating precise predictions and robust handling of data complexity.
- **MLP Regressor:** Provides good results but exhibits higher error values (MSE, RMSE, MAE) compared to ensemble models. Its complex architecture enables it to capture non-linear interactions, although training time increases.
- **K-Nearest Neighbors and SVR:** While providing useful insights, these models show higher error values and lower R2 scores, indicating difficulty in capturing patterns in the dataset.

The Random Forest Regression model demonstrated the best performance, with the lowest MSE, RMSE, and MAE values, and the highest R2 score. This model effectively forecasts district-wise groundwater resources. The results emphasize the superiority of ensemble models, particularly the Random Forest Regressor, in predicting district-wise groundwater resources. These findings provide valuable insights for policymakers and stakeholders to address groundwater management challenges effectively.

5. CONCLUSION

In summary, our study highlights the significant role of statistical analysis and the design of regression models in the decision-making process to adequately address the challenging water resource problems in the country. By analyzing a wide domain of groundwater recharge and usage, we employed five regression variants as part of the study.

Our findings indicate that the Random Forest Regressor is the best-performing machine learning model, exhibiting a low Mean Squared Error and high accuracy. This approach effectively captures the complex network of relationships within the dataset. Additionally, the Decision Tree Regressor not only ensures high accuracy rates but also provides interpretability, making it a valuable tool for groundwater resource management.

The results affirm the necessity of comprehensive, evidence-based water policy management as a critical determinant for ensuring the sustainable use of water. To enhance the precision and reliability of these models, subsequent trial runs should be conducted for various circumstances and time periods. Furthermore, additional groundwater-affecting factors should be considered to refine the models.

Improving accuracy will also benefit from the incorporation of real-time data and advanced modeling techniques, which can further enhance the predictive power of the models. Such advancements are vital for promoting sustainability and improving the effectiveness of water resource management in the face of increasing demand and environmental challenges.

REFERENCES

[1] E. M. Almetwally and M. A. Meraou, "Application of environmental data with new extension of nadarajah-haghighi distribution," *Computational Journal of Mathematical and Statistical Sciences*, vol. 1, no. 1, pp. 26–41, 2022.

[2] R. Haggerty, J. Sun, H. Yu, and Y. Li, "Application of machine learning in groundwater quality modeling: A comprehensive review," *Water Research*, vol. 233, p. 119745, 2023.

[3] M. Kumar, P. Singh, and P. Singh, "Machine learning and GIS-RS-based algorithms for mapping the groundwater potentiality in the Bundelkhand region, India," *Ecological Informatics*, vol. 74, p. 101980, 2023.

[4] K. M. Hamdia, X. Zhuang, and T. Rabczuk, "An efficient optimization approach for designing machine learning models based on genetic algorithm," *Neural Computing and Applications*, vol. 33, no. 6, pp. 1923–1933, Mar. 2021.

[5] Y. Ma, R. Han, and W. Wang, "Portfolio optimization with return prediction using deep learning and machine learning," *Expert Systems with Applications*, vol. 165, p. 113973, Mar. 2021.

[6] J. Li, X. Du, and J. R. R. A. Martins, "Machine learning in aerodynamic shape optimization," *Progress in Aerospace Sciences*, vol. 134, p. 100849, Oct. 2022.

[7] M. A. Yassin, A. G. Usman, S. I. Abba, D. U. Ozsahin, and I. H. Aljundi, "Intelligent learning algorithms integrated with feature engineering for sustainable groundwater salinization modelling: Eastern Province of Saudi

Arabia," *Results in Engineering*, vol. 20, p. 101434, 2023.

[8] A. T. M. S. Rahman, T. Hosono, J. M. Quilty, J. Das, and A. Basak, "Multiscale groundwater level forecasting: Coupling new machine learning approaches with wavelet transforms," *Advances in Water Resources*, vol. 141, p. 103595, 2020.

[9] H. R. Pourghasemi, N. Sadhasivam, S. Yousefi, S. Tavangar, H. Ghaffari Nazarlou *et al.*, "Using machine learning algorithms to map the groundwater recharge potential zones," *Journal of Environmental Management*, vol. 265, p. 110525, 2020.

[10] R. Barzegar, S. Razzagh, J. Quilty, J. Adamowski, H. Kheyrollah Pour *et al.*, "Improving GALDIT-based groundwater vulnerability predictive mapping using coupled resampling algorithms and machine learning models," *Journal of Hydrology*, vol. 598, p. 126370, 2021.

[11] P. Malakar, S. Sarkar, A. Mukherjee, S. Bhanja, and A. Y. Sun, "Use of machine learning and deep learning methods in groundwater," in *Developments in Earth Surface Processes*, Elsevier, ch. 40, pp. 545–557, 2021.

[12] K. A., A. Nair, and G. Singh, "An improved hybrid-coupled model for delineation of groundwater potential zones using surface and climatological factors," *Theoretical and Applied Climatology*, vol. 151, no. 3, pp. 2001–2022, 2023.

[13] A. Zahid, M. Q. Hassan, and K. M. U. Ahmed, "Simulation of flowpaths and travel time of groundwater through arsenic-contaminated zone in the multi-layered aquifer system of bengal basin," *Environmental Earth Sciences*, vol. 73, no. 3, pp. 979–991, 2015.