IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Optimization Of Water Distribution Networks Using Machine Learning

1. Babli Kaur 2. Er. Rajbala
M Tech Scholar, Assistant Professor,
JCDM College of Engineering, Sirsa, Haryana, India

Abstract

Water distribution networks (WDNs) play a crucial role in delivering safe and reliable water to consumers; however, challenges such as leakage, pressure fluctuations, water loss, and high energy consumption reduce their operational efficiency. This study presents a machine learning-based framework for optimizing water distribution networks to improve hydraulic performance, reduce water losses, and minimize operational costs. A dataset consisting of 10,000 records collected from SCADA systems, flow meters, pressure sensors, and smart water meters was used for model development and evaluation. Data preprocessing techniques, including cleaning, normalization, and feature engineering, were applied to enhance data quality. Three machine learning models, namely Random Forest (RF), Support Vector Regression (SVR), and XGBoost, were developed to predict network behavior and support optimization decisions. The proposed framework successfully reduced average network pressure from 52.6 m to 41.8 m and decreased water losses by approximately 45%, from 6,269 m^3/day to 3,433 m^3/day. Annual operational costs were reduced from USD 371,900 to USD 271,100, while the network reliability index improved from 0.91 to 0.98. Among the evaluated models, XGBoost achieved the highest predictive performance with an R^2 value of 0.957, demonstrating its effectiveness for intelligent water distribution network optimization.

Keywords: WDNs, Machine Learning, SVR, Energy consumption

1. Introduction

In urban and rural water supply systems, water distribution networks (WDNs) are critical for transporting safe and adequate water from water treatment plants to the consumers' taps using interconnected pipelines, pumps, valves and water storage reservoirs [1]. With the growth in population, urbanisation, industrialisation, and climate variability, the demand for water has become more complex, making it harder to manage these water networks. The issues of water leakage, fluctuating water pressure, water loss due to inefficiency in water distribution, pipe wear and tear, and unequal water distribution are some common problems in the traditional water distribution system [2]. These problems cause significant water losses, high operating expenses and low reliability of the service. Water distribution network optimization is therefore a key research field to guarantee sustainable and efficient water resource management.

A range of conventional methods have been used to enhance network performance, such as hydraulic modelling, linear programming, genetic algorithms and heuristic methods [3]. These approaches have the benefit of yielding important insights but fail to cope with having to process large-scale networks with dynamic operating conditions and huge amounts of sensor-generated data. Smart water infrastructure and the Internet of Things (IoT) have opened the door for real-time data collection of different parameters of water networks, such as flows, pressures, demand trends, leak detection, etc. [4]. This has resulted in vast quantities of data that must be analyzed using advanced methods to extract valuable information and be used to make the decisions in real-time. Machine Learning (ML) is a very powerful tool in the optimization of Water Distribution Networks (WDN) in this context.

Machine learning is one of the subfields of AI, in which computers learn from past data and make prediction or decision based on learning. The ML algorithms can discover hidden patterns from the complexity of relationships between multiple variables, predict future conditions and suggest optimal operation strategies [5].

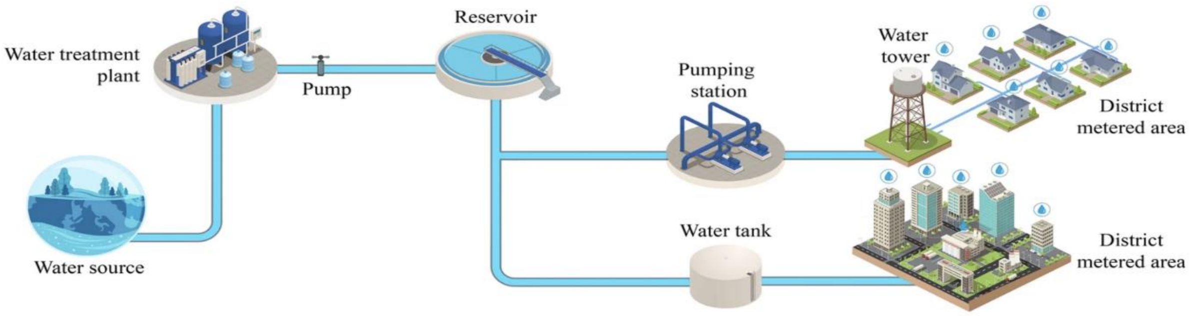

Figure 1.1: A water supply system and its components

There are numerous applications of machine learning in water distribution systems including demand prediction, leak detection, pressure control, optimal energy management, water quality management, and water asset maintenance planning [6]. The capabilities contribute to the water loss reduction and operational efficiency and reliability of services. The very important application of Machine Learning in water distribution systems is water demand forecasting. Accurate water use predictions can assist the utility to operate the pumps more efficiently [7] and help control reservoir operation and allocate water resources.

Temporal and spatial fluctuations in water demand can be captured effectively with the use of algorithms like Artificial Neural Networks (ANN), Random Forest (RF), Support Vector Machines (SVM), and Gradient Boosting models [8]. They are predictive models that allow operators to predict the fluctuations in consumption, thus reducing the consumption of energy while still making sure that there is an adequate supply when it is needed the most.

Leakage and water loss detection and management is another important application. Leakage is very much an issue of concern to utilities across the globe as it can have a significant impact on the environmental and economic bottom line [9]. Pressure, flow, and acoustic sensors can be used to create machine learning models that can detect abnormal patterns that may indicate the presence of a leak. The sooner that the leak is identified, the less non-revenue water, the less damage to the infrastructure, and the more efficient the network. Additionally, predictive maintenance methods using ML can help predict when the pipelines are prone to failure, which improves maintenance efficiency and service continuity.

Also, the machine learning concept is applied to support the optimization of the pumping system and energy efficient operation [10]. If the pressure is too high it can cause more leaks, pipe deterioration; too low a pressure can impact the quality of service. Advanced ML algorithms can keep checking the network conditions continuously and optimise the pressure settings to operate the network at the optimum pressure that fulfills the customer requirements [11]. Likewise, the use of smart pump scheduling, which uses electricity tariff and demand forecasts, can help to lower energy use and operating costs. Here are the research objectives of the study follows as:

- To collect and analyze hydraulic and operational data from the water distribution network, including flow rate, pressure, water demand, energy consumption, and leakage information.

- To preprocess and transform the collected data through cleaning, normalization, outlier removal, and feature engineering to improve data quality and model performance.
- To develop predictive machine learning models using Random Forest (RF), Support Vector Regression (SVR), and XGBoost algorithms for forecasting network behavior, leakage occurrence, and operational performance.
- To identify critical factors affecting water losses and pressure variations within the distribution network using engineered hydraulic and operational features.
- To evaluate the performance of the developed models using statistical metrics such as R^2, MAE, RMSE, and MAPE.

2. Related Work

Current development of smart water distribution systems is geared toward leakage detection, pressure regulation, power efficiency and real-time monitoring via IoT and ML. Many intelligent frameworks have been developed that incorporate sensors, hydraulic models, cloud computing, and optimization algorithms to achieve high operational performance for urban water distribution networks. In these researches, intelligent monitoring systems were presented to support water loss reduction and efficient resource management as an important element for the sustainability in urbanizing developing countries. The Hybrid SVM-ANN-GT model introduced by Komba et al., (2026) [12] for the leak detection and location of water distribution systems achieved an accuracy of 96%, which is superior to the individual SVM and ANN models as the optimized sensor location used a graph theory approach. Additionally, the IoT-based framework developed by Kusuma et al. (2026) [13] utilized a pressure control using EPANET Digital Twin, PID controller and it was shown that the system provided stable pressure control with less than 2% overshoot and a tracking error under 0.5 m.

Velayudhan et al. (2024) [14] examined energy-efficient IoT communication architectures for smart water management applications, which resulted in nearly 40% energy efficiency improvements due to proper fog node placement. Effiom et al. (2024) [15] developed a pipeline leakage monitoring IoT-based system with an accuracy of 94% in detecting leaks by implementing pressure and temperature sensors along with cloud-based machine learning analysis techniques. Shao et al. (2023) [16] implemented IoT and ML algorithms that led to optimizing energy consumption while reducing water leakages to 1.17% and 6.98%, respectively. Rahu et al. (2023) [17] suggested an IoT-Machine Learning framework for predicting the water quality parameters, where MLP regression obtained an R^2 score of 0.93 and the Random Forest classifier was capable of obtaining 91% accuracy. Naqash et al. (2023) [18] presented an IoT framework based on blockchain technology to enhance the water management process by ensuring security and managing water leakages. Kumar et al. (2023) [19] created an IoT-based smart water management system incorporating multiple quality sensors along with LoRaWAN communication.

3. Research Methodology

The research plan for optimizing water distribution networks using machine learning is a systematic process that does not only include data analysis but also water management practices. The operational data collected from sensors and SCADA systems include water flow, pressure, water demand, water pipe characteristics, energy consumption and leakage data. To ensure the data quality, the acquired data is subjected to data preprocessing such as outlier removal, normalization and cleaning.

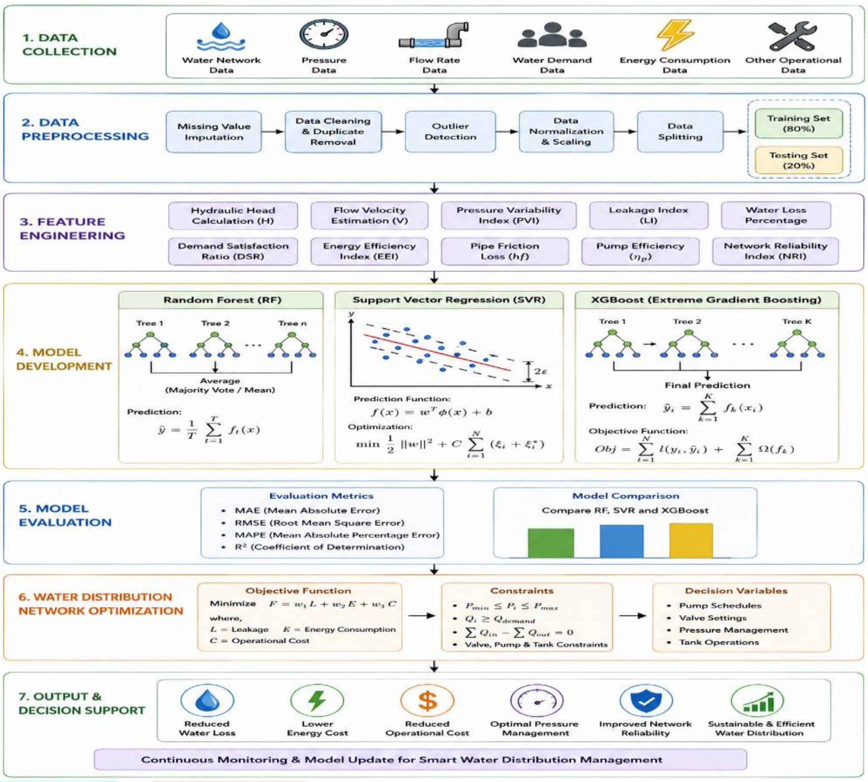

Figure 3.1: Flowchart of research methodology

The hydraulic indicators like leakage index, pressure variability and energy efficiency indicators are then created through feature engineering. Random Forest, Support Vector Regression (SVR) and XGBoost models are used to train machine learning models that predict network performance and leakage patterns. The expected results are used in an optimization module to optimize water loss and energy consumption while ensuring service reliability. Lastly, statistical metrics and operational indicators are used to evaluate and validate the performance of the model and network efficiency.

3.1 Data Collection

This study employed a water distribution network of SCADA systems, flow meters, pressure sensors, smart water meters and hydraulic simulation software to gather the data for the 12-month monitoring period. Records from 10,000 were collected at 15-minute intervals including hydraulic and operational attributes of the network. The dataset includes key parameters such as flow rate (175.2–210.8L/s), pressure (21.3–68.4m), water demand (42,500–56,800m³/day), pump energy consumption (3,450–4,850kWh/day), reservoir water level (14.2–22.6m), pipe dimensions, and water loss (5,870–6,420m³/day). Demand variability was also considered by including the environmental factors such as temperature and rainfall. A data set was split into 80% training data (8000 records) and 20% testing data (2000 records) for the development and validation of the machine learning models.

Table 3.1: Dataset Summary

Parameter	Minimum	Maximum	Average
Flow Rate (L/s)	175.2	210.8	192.4
Pressure (m)	21.3	68.4	52.6
Water Demand (m³/day)	42,500	56,800	50,120
Pump Energy (kWh/day)	3,450	4,850	4,115
Reservoir Level (m)	14.2	22.6	18.4
Water Loss (m³/day)	5,870	6,420	6,259

3.2 Data Preprocessing

Data preprocessing is a crucial stage in developing an effective machine learning model for water distribution network optimization.

- **Missing Value Imputation**

The accuracy of models can be negatively affected by missing pressure, flow rate, demand, and energy consumption data. Hence, the missing observations will be replaced by the average value of the feature.

$$x_{miss} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (1)$$

where: x_{miss}= estimated missing value, x_i= observed values, n= total observations

For instance, if there are no measurements of flow at a specific time step, then the average flow of the time steps that are adjacent will be used to fill the void, thus maintaining the completeness of the dataset.

- **Data Cleaning and Duplicate Removal**

Records can be duplicated because of repeats in sensor transmission. Consider the data set to be represented as:

$$D = \{x_1, x_2, \dots, x_n\}$$

The cleaned dataset is obtained as:

$$D_{clean} = D - D_{duplicate} \quad (2)$$

where: D_{clean}= cleaned dataset, $D_{duplicate}$= set of repeated records

This process ensures that the data is consistent, which helps to avoid any bias in the training of the model.

- **Data Normalization**

Each variable may have their own measurement unit and scaling, so that for this reason, the Min-Max normalization is used to transform all the features to [0,1].

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4)$$

where: X= original value, X_{min}= minimum value, X_{max}= maximum value

Normalization ensures each parameter of pressure, flow, demand, energy has the same contribution during machine learning training.

3.3 Feature Engineering

A series of feature engineering techniques are used to take the raw data from the water distribution network and create useful hydraulic and operational features to enhance the accuracy of ML predictions. The resulting properties describe the relationships between pressure, flow and demand, as well as energy consumption and leak characteristics in the network.

- **Hydraulic Head Calculation**

The total energy available at a node is called hydraulic head and is calculated by adding the elevation head, velocity head, and pressure head.

$$H = z + \frac{P}{\gamma} + \frac{V^2}{2g} \tag{5}$$

where: H= hydraulic head (m), z= elevation head (m) P= pressure (Pa), γ= specific weight of water (N/m³), V= flow velocity (m/s), g= gravitational acceleration (9.81 m/s²)

Hydraulic head is used to evaluate energy distribution across the network.

- **Pressure Variability Index**

Leakage and pipe deterioration is greatly affected by pressure fluctuation.

$$PVI = \frac{\sigma_P}{\mu_P} \tag{6}$$

where: PVI= Pressure Variability Index, σ_P= standard deviation of pressure, μ_P= mean pressure

Lower values indicate a more stable network.

3.4 Machine Learning Model Development

- **Random Forest (RF)**

It is an ensemble machine learning algorithm that is used to improve the accuracy of the prediction and to reduce overfitting using multiple decision trees. In the proposed water distribution network optimization framework, RF is used for dealing with complex relations between the hydraulic parameters including pressure, flow rate, water demand, leakage rate, and energy consumption [20]. The algorithm builds many decision trees based on random subsets of the training data and predictor variables. Each tree makes its own prediction and then the final prediction is an average of all of the predictions, thus providing a more stable and accurate model. RF is very robust in the presence of missing data and noise and can effectively capture the non-linear interaction in the water distribution system.

- **Support Vector Regression (SVR)**

It is a supervised ML method for predicting continuous numerical values. SVR is used in water distribution network optimization to estimate operational parameters, such as leakage volume, water demand, pressure variations, and energy consumption. The model learns by building an optimum regression hyperplane that has the largest margin and the smallest prediction errors in a given tolerance. Unlike traditional regression methods, SVR targets to keep the model simple and to capture the nonlinear relationship well [21]. The kernel functions are used to map input variables into a higher dimensional feature space in which complex patterns are easier to model. This feature makes SVR an ideal tool for dealing with non-linear hydraulic characteristics of a water distribution network.

- **XGBoost (Extreme Gradient Boosting)**

It is an advanced machine learning technique called a Gradient-boosting algorithm, which is highly efficient and has a high prediction accuracy. The proposed system improves the performance of water distribution network by accurately predicting leakage occurrence, pressure response, energy consumption and energy demand fluctuation by fine-tuning the system performance with XGBoost model. A series of decision trees are built in turn with errors of the preceding decision tree to learn from its errors and correct them. This learning process is cyclical and enables XGBoost to learn in very complicated non-linear relationships between hydraulic and operational variables. Moreover, regularization techniques are also available in XGBoost to prevent over-fitting and improve the model generalization [22]. XGBoost can be more effective than other machine learning algorithms, especially when dealing with large amounts of data and intricate interactions between features.

3.5 Water Distribution Network Optimization

The search for optimal operating conditions which will include reducing water losses, energy consumption and operating costs in the system but still provide a good water supply to the consumers and water pressure within the water supply system.

The proposed framework utilizes the optimization module to combine predictions from three models (RF, SVR, XG-Boost) and make intelligent decisions. The optimization process constantly assesses the state of the network – its pressure level, flow rates, water demand, leakage rates, and pump energy consumption – and determines the optimum operation. Several factors, including pump schedules, valve settings and pressure management parameters are adjusted to enhance hydraulic performance and to minimize non-revenue water.

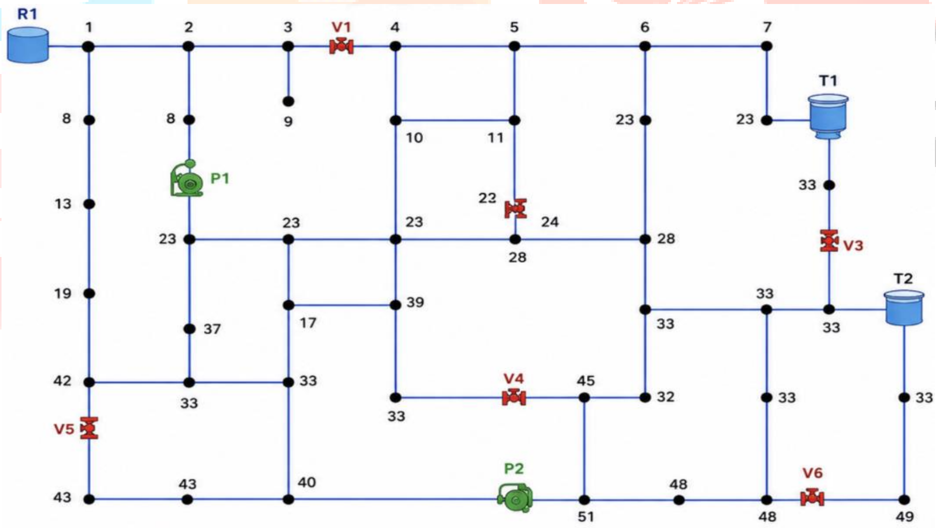

Figure 3.4: Network Layout

The main goal is to reduce the total losses on the network, whilst keeping the quality of services and meeting consumer requirements. The objective function of the optimization is written as:

$$F = w_1 L + w_2 E + w_3 C \quad (7)$$

where:

- F= overall optimization objective,
- L= water leakage loss,
- E= energy consumption,
- C= operational cost,

- w_1, w_2, w_3= weighting coefficients assigned to each objective.

The optimization is constrained by hydraulic requirements in order to provide reliable operation. The amount of pressure applied to each network node should be within acceptable limits:

$$P_{min} \le P_i \le P_{max} \tag{8}$$

In which P_i is the pressure at node i, and P_{min} and P_{max} are the minimum and maximum allowable pressures. Likewise, the provided flow should meet with the demand of consumers:

$$Q_i \ge Q_{demand} \tag{9}$$

where Q_i is the provided flow rate and Q_{demand} is the demand of water. Also, the principle of mass conservation in the network should be followed as:

$$\sum Q_{in} - \sum Q_{out} = 0 \tag{10}$$

where, the total inflow Q_{in} and outflow Q_{out} of the system. The proposed framework is designed to effectively minimize leakage, stabilize water pressures, optimize energy consumption and boost the overall reliability and sustainability of the water distribution network through these optimization mechanisms.

3.6 Model Evaluation

Statististical measures are used to assess the machine learning models.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{11}$$

$$MAE = \frac{1}{n}\sum | y_i - \hat{y}_i | \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2} \tag{13}$$

$$MAPE = \frac{100}{n}\sum | \frac{y_i - \hat{y}_i}{y_i} | \tag{14}$$

4. Result and Discussion

A dataset of water distribution network parameters, including hydraulic, operational and demand parameters were used to implement the proposed optimization framework based on machine learning. The developed models were tested with respect to predictive accuracy, leak reduction in the pipe, energy efficiency, pressure stability and operational cost savings.

4.1 Pressure Optimization Results

The optimization algorithm maintained the nodal pressure within the desired operational limit and minimized the excessive pressure which caused the leakages. The pressure optimization results show the performance of the proposed machine learning-based optimization framework for water distribution network, and it is effective in improving the hydraulic performance of water distribution network. The mean network pressure dropped from 52.6 m to 41.8 m, thereby relieving excessive pressures that are often responsible for pipe failures and infrastructure losses. Likewise, the peak pressure was lowered from 68.4 m to 54.7 m, which helps to reduce the risk of pipe bursts and operational failures.

The minimum pressure on the other hand rose from 21.3 m to 24.8 m, which will guarantee sufficient water supply to all demand nodes. In addition, there was a significant reduction in pressure variability – from 18.7% to 9.4% – signifying greater pressure stability throughout the network. These enhancements help to minimize water losses, improve service reliability, lower maintenance costs, and optimize network operations. Results show that the average network pressure was reduced by 20.5% and pressure stability increased by 49.7%.

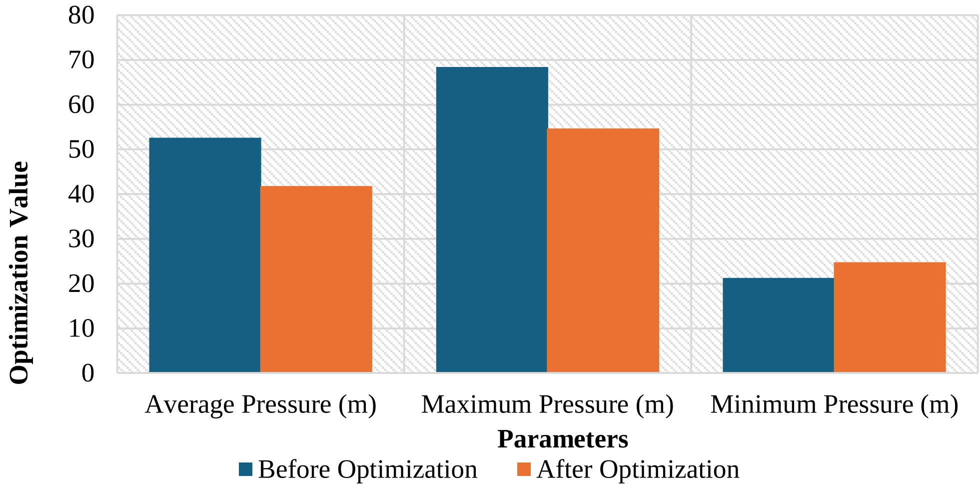

Figure 4.1: Graph of Pressure optimization

4.2 Water Loss Results

The above 20-day water loss analysis is an illustration of how the proposed machine learning-based optimization framework can help reduce non-revenue water in the water distribution system. Prior to optimizing, the daily losses ranged from 6,175 m³/day to 6,350 m³/day, with an average loss rate of about 6,269 m³/day. The optimization strategy has led to a significant reduction in the amount of water lost, with a drop from 3,386 m³/day to 3,490 m³/day, and an average of 3,433 m³/day. For the monitoring period, the percentage reduction was on the consistent level of about 45% which is good and consistent. The considerable savings in water loss are largely due to better pressure control, smart pumping and water loss intelligence. The outcomes of these results have demonstrated the potential of machine learning methods to improve the efficiency of networks, save water resources and decrease operating expenses.

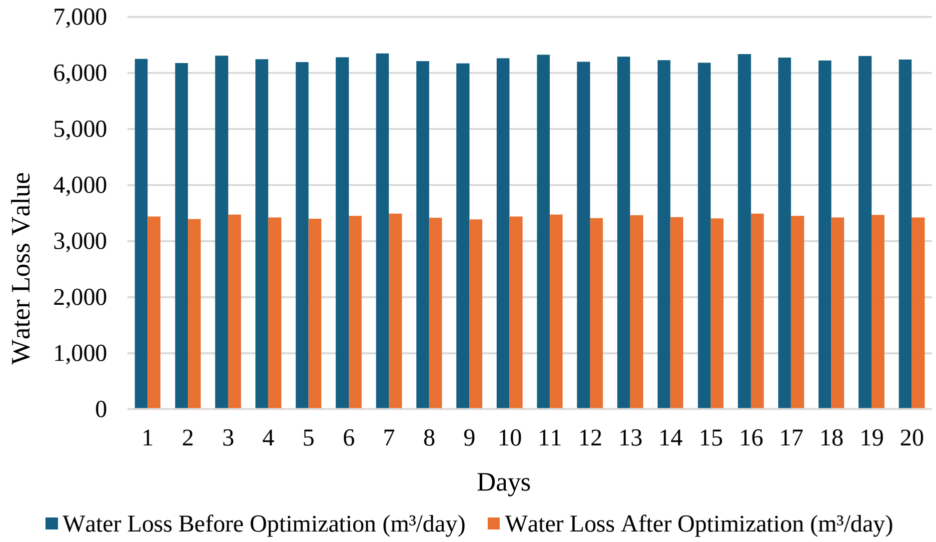

Figure 4.2: Graph of water loss

4.3 Operation Cost Analysis

The operation cost analysis shows that the proposed machine learning based optimization method can save a great deal of the annual operation cost of the water distribution network. Optimization resulted in savings of USD 212,400/year in the conventional system to USD 164,700/year in the optimized system, representing significant cost savings due to optimized pump scheduling and pressure management. Similarly, due to reducing water losses and enhancing leak detection efforts, leakage management costs were also lowered from USD 91,300/year to USD 50,600/year. Additionally, there were reduced infrastructure stresses and fewer repair requirements, resulting in maintenance cost savings from USD 68,200/year to USD 55,800/year. In general, the total cost of operation dropped from USD 371,900/yr to USD 271,100/yr, corresponding to a cost reduction of about USD 99.7%/yr and improving the economic sustainability and efficiency of the water distribution system.

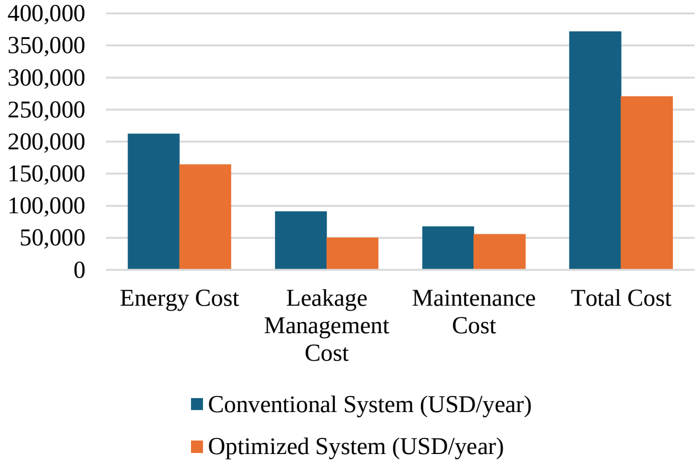

Figure 4.4: Graph of operation cost analysis

4.4 Network Reliability Performance

The results on the network reliability performance show that the proposed ML optimization framework has a positive effect on the operation of water distribution systems. The percentages of the consumers' water demands satisfied by the optimized network (99.1%) and the existing network (94.8%) were compared. Similarly, Pressure Compliance improved from 88.3% to 97.6% of the network nodes, meaning that optimized system held the pressures within the recommended operation range for most of the network nodes. Furthermore, the Supply reliability index increased from 0.91 to 0.98, suggesting higher uniformity and reliability in water supply services. These improvements achieved by effective pressure management, leakage reduction and optimized operational control. Results show overall that the proposed work significantly improves the reliability of the network, the quality of service, and customer satisfaction.

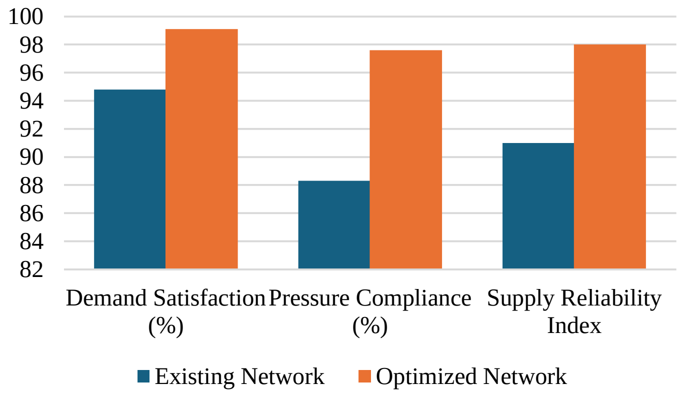

Figure 4.5: Graph of network reliability performance

4.5 Machine Learning Model Performance

The effectiveness of the developed water distribution network optimization predictive framework is shown by the machine learning model results. The best performing model based on the R^2 score was XGBoost with an R^2 score of 0.957, which means that the model explained about 95.7% of the variation in the target variable. It had the lowest error values: MAE = 0.026, RMSE = 0.041, and MAPE = 4.9%, which showed that it provides excellent prediction accuracy. The Random Forest model was also practical with an R^2 of 0.921 and fairly small prediction errors. However, the accuracy (R^2 = 0.894) and relatively higher errors by SVR were the lowest. In overall, the outcomes suggest that XGBoost is the most appropriate model for forecasting water distribution network hydraulic behaviour and for optimizing the water distribution network efficiently.

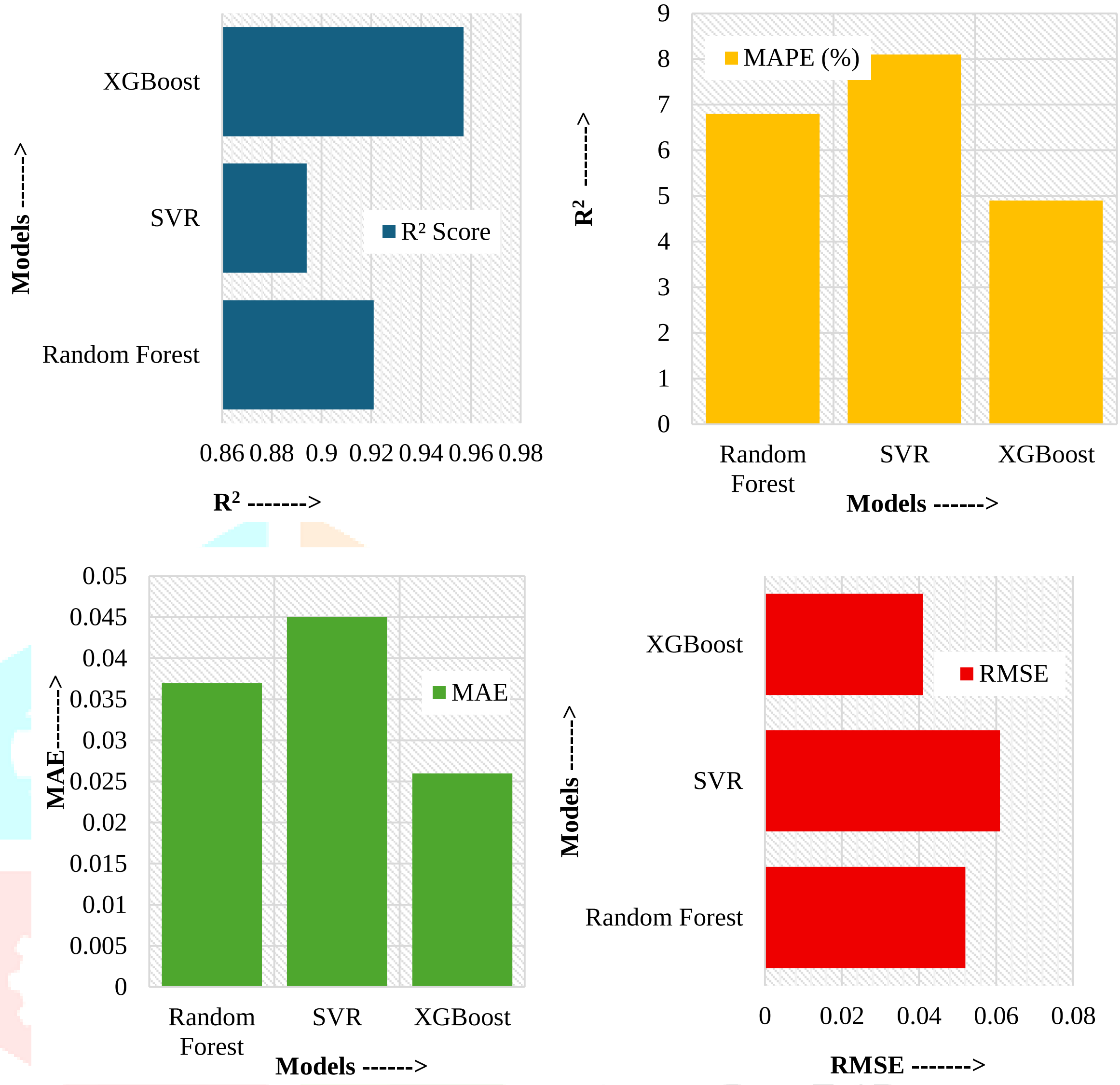

Figure 4.6: Comparative graph of ML models

5. Conclusion

The primary aim of this research work was to find an intelligent solution to optimize water distribution networks based on machine learning and improve the overall water system performance. All these goals were achieved by executing the systematic approach, which involves data collection, data preprocessing, feature engineering, development of machine learning models, and optimization. These operational parameters such as flow rate, pressure, water demand, water loss, energy consumption, water level and reservoir level were collected and cleaned to ensure the quality of the data collected. A number of important hydraulic indicators were produced using feature engineering, such as leakage index, water loss percentage, and pressure variability index. Three machine learning models (RF, SVR and XGBoost) were trained and tested and evaluated using statistical performance measures such as R2, MAE, RMSE and MAPE. The summary of the major results of the study is as follows:

• This led to a decrease in the average network pressure from 52.6 m to 41.8 m, a decrease of 20.5 % which led to an increase in hydraulic stability of the network.

• The mean pressure variation has been reduced from 18.7% to 9.4% which is a gain of nearly 49.7% in pressure stability.

• The average loss was lowered from about 6,269 m³/day to 3,433 m³/day, with almost 45% leakage reduction.

• The annual operating cost was reduced from USD 371,900/year to USD 271,100/year, which is around a 27.1% cost savings.

• The reliability index increased from 0.91 to 0.98, reflecting a greater reliability of the network.

• XGBoost was the most predictive model among all machine learning models, with an R2 of 0.957, MAE of 0.026, RMSE of 0.041 and MAPE of 4.9%.

In general, the study validates that machine learning can substantially improve both water distribution network efficiency and reliability, and lower water losses and expenses in the water distribution network.

Reference

1. Jazayeri, Pedram, Hamid R. Safavi, Mohammad Saleh Ebrahimi, Alireza Rahmatpanah, and Mohammad Nazemizadeh. "Bridging prediction and optimization: A unified MINLP, traditional machine learning and deep learning framework for water distribution networks management." *Ain Shams Engineering Journal* 17, no. 1 (2026): 103834.

2. Kołodziej, Katarzyna, Michał Cholewa, Przemysław Głomb, Wojciech Koral, and Michał Romaszewski. "Efficient Numerical Calibration of Water Delivery Network Using Short-Burst Hydrant Trials." *Journal of Water Resources Planning and Management* 152, no. 4 (2026): 05026002.

3. Bianchi, A. Belmondo, H. H. M. Rijnaarts, and S. Shariat Torbaghan. "Neural network-informed Optimal Water Flow problem: Modeling, algorithm, and benchmarking." *Water Research X* (2026): 100479.

4. Grbčić, Luka, Lado Kranjčević, and Siniša Družeta. "Machine learning and simulation-optimization coupling for water distribution network contamination source detection." *Sensors* 21, no. 4 (2021): 1157.

5. Gupta, Gagan. "Monitoring water distribution network using machine learning." (2017).

6. Hajgató, Gergely, György Paál, and Bálint Gyires-Tóth. "Deep reinforcement learning for real-time optimization of pumps in water distribution systems." *Journal of Water Resources Planning and Management* 146, no. 11 (2020): 04020079.

7. Zhou, Xiao, Zhenheng Tang, Weirong Xu, Fanlin Meng, Xiaowen Chu, Kunlun Xin, and Guangtao Fu. "Deep learning identifies accurate burst locations in water distribution networks." *Water research* 166 (2019): 115058.

8. Pant, Millie, and Vaclav Snasel. "Design optimization of water distribution networks through a novel differential evolution." *Ieee Access* 9 (2021): 16133-16151.

9. Mala-Jetmarova, Helena, Nargiz Sultanova, and Dragan Savic. "Lost in optimisation of water distribution systems? A literature review of system design." *Water* 10, no. 3 (2018): 307.

10. Fan, Xudong, Xijin Zhang, and Xiong Bill Yu. "A graph convolution network-deep reinforcement learning model for resilient water distribution network repair decisions." *Computer-Aided Civil and Infrastructure Engineering* 37, no. 12 (2022): 1547-1565.

11. Conejos Fuertes, P., Fernando Martínez Alzamora, M. Hervás Carot, and J. C. Alonso Campos. "Building and exploiting a Digital Twin for the management of drinking water distribution networks." *Urban Water Journal* 17, no. 8 (2020): 704-713.

12. Komba, Giresse M., Topside E. Mathonsi, and Pius A. Owolawi. "A Hybrid Machine Learning Algorithm for Pipeline Leak Detection and Localisation in Water Distribution Networks." network 4 (2026): 5.

13. Kusuma, Bagus Adhi, Khairunnisak Nur Isnaini, and Aulia Hamdi. "Dynamic IoT–PID Control for Energy-Efficient Water Distribution: EPANET-Based Digital Twin Validation in Varied Geographical Terrains." Journal of Applied Data Sciences 7, no. 1 (2026): 773-785.

14. Velayudhan, Nibi Kulangara, Aiswarya S, Aryadevi Remanidevi Devidas, and Maneesha Vinodini Ramesh. "Delay and energy efficient offloading strategies for an IoT integrated water distribution system in smart cities." Smart Cities 7, no. 1 (2024): 179-207.

15. Effiom, Samuel O., Gertrude A. Fischer, and Eko J. Akpama. "Novel system design model for an IoT-based realtime oil and gas pipeline leakage monitoring system." International Journal of Engineering and Technology 13, no. 2 (2024): 336-340.

16. Shao, Yu, Fengyi Yao, Jia Liu, Tingchao Yu, and Shipeng Chu. "Global energy and leakage optimization in water distribution systems from water treatment plants to customer taps." Resources, Conservation and Recycling 194 (2023): 107003.

17. Rahu, Mushtaque Ahmed, Abdul Fattah Chandio, Khursheed Aurangzeb, Sarang Karim, Musaed Alhussein, and Muhammad Shahid Anwar. "Toward design of internet of things and machine learning-enabled frameworks for analysis and prediction of water quality." IEEE Access 11 (2023): 101055-101086.

18. Naqash, Muhammad Tayyab, Toqeer Ali Syed, Saad Said Alqahtani, Muhammad Shoaib Siddiqui, Ali Alzahrani, and Muhammad Nauman. "A blockchain based framework for efficient water management and leakage detection in urban areas." Urban Science 7, no. 4 (2023): 99.

19. Kumar, T. Nitesh, K. Thirumala Akash, Mohammed Riyaz Ahmed, and R. S. Upendra. "Integrated IoT Reference Architecture for Smart Metering and Monitoring of City Water Distribution Using Next Generation Sensors." In 2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), pp. 1-6. IEEE, 2023.

20. Lučin, Ivana, Bože Lučin, Zoran Čarija, and Ante Sikirica. "Data-driven leak localization in urban water distribution networks using big data for random forest classifier." *Mathematics* 9, no. 6 (2021): 672.

21. Pourjafari, Ebrahim, and Marek Reformat. "A support vector regression based model predictive control for volt-var optimization of distribution systems." *IEEe Access* 7 (2019): 93352-93363.

22. Ikpa, P. N., J. C. Osuagwu, P. N. Nwakwasi Ikpa, C. S. Uzoukwu, J. C. Ugwuegbulam, P. C. Obumseli, C. S. Ezenkwa et al. "Xg Boost Model Application in Spatio-Temporal Differences of Groundwater Quality in Different Geographical Location."