

An Intelligent Data-Driven Framework for Water Leakage Detection and Maintenance Decision Support in Water Distribution Networks

Sunanth H¹, Dr. K. Akila²

^{1,2}Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

Abstract

Water leakage in distribution networks results in significant water loss and makes maintenance more difficult, especially when the leakage occurs in underground pipelines. Pressure measurements provide useful information about changes in network behaviour, but interpreting measurements from multiple locations can be difficult when demand conditions also vary. This paper presents an integrated data-driven framework for leakage detection, leak-location prediction, severity classification, water-loss estimation, and maintenance decision support. The framework uses the EPANET NET3 benchmark network and the Water Network Tool for Resilience (WNTR) to generate simulated leakage scenarios and corresponding pressure measurements. Pressure values from ten selected monitoring nodes, together with a network demand factor, are used as inputs to three Random Forest models. The first model predicts the probable leak location, the second predicts leak severity, and the third estimates daily water loss. A rule-based layer converts the prediction results into maintenance priority and recommendation. To demonstrate sensor-to-application communication, an ESP32 with virtual sensor data is simulated using Wokwi and connected through MQTT to a Python prediction application. A Streamlit dashboard is used to display the final results. The final dataset contains 1000 simulated records with 16 variables. Using the saved models and the evaluation procedure described in this work, the leak-location model achieved 97.50% accuracy, the severity model achieved 98.50% accuracy, and the water-loss regression model achieved an R^2 value of 0.9979. The results show that the proposed prototype can connect hydraulic simulation, machine learning, simulated IoT communication, and maintenance-oriented decision support in a single workflow. Since the current dataset and sensor layer are simulated, further validation using physical sensors and real water-distribution data is required before practical deployment.

Keywords: Water Leakage Detection, Water Distribution Network, Leak Localization, Random Forest, EPANET, WNTR, MQTT, IoT, Water Loss Estimation, Maintenance Decision Support

1. Introduction

Water distribution networks are important infrastructure systems that transport treated water from supply sources to consumers. These networks contain interconnected pipes, junctions, reservoirs, pumps,

valves, and demand nodes. Over time, ageing, corrosion, joint failures, pressure variations, and other forms of deterioration can result in leakage. Since a large part of a distribution network is located underground, identifying the exact location of a leak is often difficult. Leakage is not only a problem of lost water. A continuing leak can also affect network pressure, increase operating costs, reduce the efficiency of the distribution system, and create additional maintenance requirements. For this reason, leakage detection and localization have received considerable attention in water-distribution research. Recent reviews show that both hydraulic model-based and data-driven methods are widely used, with each approach having different advantages and limitations [1, 4].

Pressure is particularly useful for leakage analysis because a leakage event changes the hydraulic state of a network. A pressure reduction or a change in the pressure pattern can be observed at several monitoring locations. Therefore, a collection of pressure measurements can be used as an indirect representation of a leakage condition. However, the relationship between pressure patterns, leakage location, leakage severity, and water loss is not always simple, especially when normal water demand changes.

Data-driven machine-learning methods can learn such relationships from labelled examples. Lucin et al. demonstrated the use of Random Forest for leak-location prediction using simulated pressure measurements and sparse sensor placement [2]. Their later work considered pipe segmentation and showed that demand variation and leakage uncertainty can make precise localization difficult [3]. These studies indicate that pressure-based machine learning can be useful, but also highlight the importance of the training data and sensor configuration.

Other recent studies have investigated different data-driven approaches. Tornyeviadzi and Seidu proposed a one-dimensional convolutional neural-network autoencoder for leakage detection and localization using multivariate SCADA data [5]. Wu et al. investigated leakage detection using multiple features extracted from high-frequency pressure data [6]. Yang and Wang studied pressure-sensor deployment for leak identification using the NET3 water-distribution network and considered practical factors such as sensor coverage and sensor failures [7].

Although these studies provide useful individual solutions, a practical monitoring system needs to go beyond simply predicting a leak location. An operator may also need information about the probable severity, estimated water loss, maintenance priority, and recommended action. In addition, sensor information must eventually be transferred from the monitoring layer to the prediction system.

This work therefore focuses on developing an integrated prototype that connects these stages. The system uses hydraulic simulation to generate leakage scenarios, machine learning to analyse pressure patterns, simulated IoT communication to demonstrate data transfer, and rule-based decision support to convert model outputs into maintenance information.

The main contributions of this work are:

- generation of a labelled leakage dataset using the EPANET NET3 benchmark network and WNTR;
- use of pressure measurements from ten selected monitoring nodes and a network demand factor;
- separate Random Forest models for leak-location prediction, severity classification, and water-loss estimation;
- integration of model outputs with a rule-based maintenance decision-support layer;
- demonstration of an ESP32/Wokwi and MQTT communication path; and
- visualization of the final prediction and maintenance information through a Streamlit dashboard.

The objective is not to introduce a new Random Forest algorithm or a new communication protocol. Instead, the work demonstrates how established components can be combined into a complete leakage-management prototype.

2. Related Work

Water leakage detection and localization have been studied using hydraulic modelling, statistical techniques, machine learning, sensor-based approaches, and combinations of these methods.

Hu et al. reviewed model-based and data-driven approaches for leak detection and localization and discussed the dependence of model-based approaches on hydraulic-model accuracy and measurement quality [1]. The review also identified limitations associated with changing water demand and measurement uncertainty. Romero-Ben et al. later provided a broad review of leakage detection and localization and discussed the transition from model-based approaches toward data-driven methods and the challenges involved in applying these methods to real water utilities [4].

Random Forest has been directly applied to pressure-based leak localization. Lucin et al. used Random Forest classifiers to predict leak locations in water distribution networks using simulated leakage scenarios and sparse pressure measurements [2]. The study included variation in leakage parameters and network demand, making it relevant to data-driven learning from hydraulic simulations.

Lucin et al. subsequently extended this direction by considering pipe segmentation for more detailed leak localization [3]. Their study showed that node-level prediction may not be sufficient when the actual leak can occur between two network nodes. It also showed that different leakage and demand conditions can produce similar pressure measurements. This limitation is important for the present work because the proposed framework predicts leakage at predefined location classes rather than attempting exact pipe-level localization.

Recent research has also moved toward deep-learning approaches. Tornyeviadzi and Seidu developed a semi-supervised 1D convolutional autoencoder using multivariate supervisory control and data acquisition (SCADA) data for leakage detection and localization [5]. Their work demonstrates the potential of time-series learning when real or benchmark SCADA measurements are available. In comparison, the present work uses simulated pressure features and a Random Forest model because the current objective is to develop an interpretable tabular-data prototype.

Wu et al. investigated leakage detection using multiple features extracted from high-frequency pressure measurements [6]. Their results demonstrate that pressure signals contain useful information for distinguishing leakage conditions. The present framework uses pressure measurements as direct input features but does not use high-frequency signal processing.

Sensor placement is another important research area. Yang and Wang proposed an approach for pressure-sensor deployment that considered detection coverage, sensitivity, budget constraints, and sensor failures and demonstrated the method on the NET3 network [7]. The current work also uses the NET3 benchmark network, but the selected ten monitoring nodes are treated as the project's monitoring configuration rather than being claimed as an optimally selected sensor set.

Research has also considered the connection between leakage analysis and maintenance. Tornyeviadzi et al. proposed a dynamic segment criticality framework that can support maintenance scheduling in water distribution networks [8]. Their work illustrates the importance of converting network-analysis results into information that can support operational decisions.

The literature therefore shows that pressure-based machine learning, hydraulic simulation, sensor placement, and maintenance analysis are all established research areas. The gap addressed in this work is at the system-integration level. The proposed prototype combines leakage location prediction, severity classification, water-loss estimation, maintenance prioritization, simulated sensor communication, and dashboard visualization within one workflow.

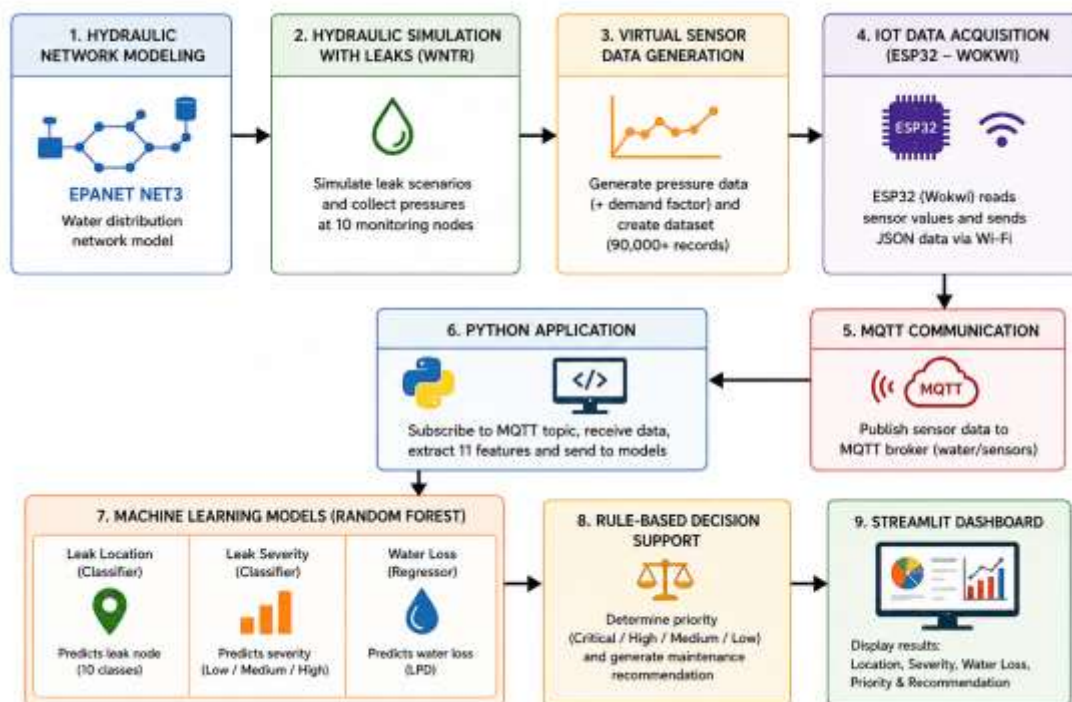
3. Proposed System

The proposed framework is designed to convert simulated pressure observations into leakage and maintenance information.

The system has five main functional parts. First, WNTR is used with the EPANET NET3 network to generate leakage scenarios and pressure observations. Second, the generated dataset is used with three Random Forest models. Third, a simulated ESP32 and MQTT communication path demonstrates how pressure data can be transferred to the Python application. Fourth, the prediction results are processed by a rule-based maintenance layer. Finally, a Streamlit dashboard presents the results.

The overall system architecture is shown in Figure 1.

Figure 1: Overall Architecture of the Proposed Water Leakage Detection and Maintenance Decision-Support Framework.



The architecture begins with the hydraulic network and leakage simulation. Pressure values are collected from selected monitoring nodes and combined with the network demand factor. These values form the input vector for the prediction models.

Three prediction tasks are performed independently. The location model predicts the probable leakage location. The severity model predicts whether the leakage condition belongs to the Low, Medium, or High class. The water-loss model estimates the amount of water lost per day.

The outputs are then passed to a rule-based decision-support layer. This layer assigns a maintenance priority and produces a recommendation. The final information is displayed through the dashboard.

4. Dataset Generation and Hydraulic Simulation

4.1 EPANET NET3 Network

The Environmental Protection Agency Network Analysis Tool (EPANET) is widely used for modelling the hydraulic behaviour of drinking-water distribution networks [11]. The present work uses the NET3 benchmark network as the hydraulic environment.

The Water Network Tool for Resilience (WNTR) is a Python-based framework developed for analysing the behaviour and resilience of water distribution networks [10]. It provides a convenient interface for modifying network conditions and running hydraulic simulations.

In this work, WNTR is used to introduce leakage conditions into the NET3 network and obtain pressure responses for different scenarios.

4.2 Monitoring Nodes

Pressure measurements are collected from ten selected monitoring nodes:

{10, 20, 35, 50, 60, 101, 105, 111, 119, 601} (1)

These measurements represent the pressure inputs available to the machine-learning models.

The possible leakage-location classes used in the dataset are:

{15, 20, 35, 50, 60, 101, 105, 111, 119, 120} (2)

The monitoring nodes and leakage-location classes are therefore not identical. The model learns the relationship between the combined pressure pattern and the simulated leakage condition instead of simply selecting the sensor with the largest or smallest pressure.

4.3 Dataset

The final Phase 8 dataset contains 1000 records and 16 columns.

The input features are the ten pressure measurements and the Demand Factor. The target variables are Leak Location, Severity, and Water Loss LPD. Leak Pressure and Priority are also retained in the dataset for analysis and decision support.

The severity classes used in the study are Low, Medium, and High.

The dataset is generated from simulation rather than collected from physical water pipelines. Therefore, the labels represent the simulated leakage conditions used during dataset generation.

5. Machine-Learning Models

Random Forest is an ensemble learning method that combines multiple decision trees and can be used for both classification and regression [9]. It was selected for this project because the input consists of structured numerical pressure and demand features and because feature importance can be extracted from the trained models.

Three independent Random Forest models are used.

5.1 Leak-Location Model

The leak-location model is a multiclass classifier. Its input consists of the ten pressure measurements and Demand Factor. The prediction can be represented as

$$\hat{L} = fL(x) \quad (3)$$

where $\hat{L}$ is the predicted leakage location and x is the input feature vector.

5.2 Leak-Severity Model

The severity model predicts one of three classes:

$$\hat{S} = f_S(x) \quad (4)$$

where

$$S \in \{\text{Low, Medium, High}\} \quad (5)$$

5.3 Water-Loss Model

The water-loss model is a Random Forest regression model. It predicts daily water loss in litres per day:

$$\hat{W} = f_W(x) \quad (6)$$

The regression output is compared with the simulated water-loss value in the dataset.

6. Evaluation Method

The classification models are evaluated using accuracy:

$$\text{Accuracy} = N_{\text{correct}} \div N_{\text{total}} \quad (7)$$

For water-loss estimation, the coefficient of determination is used:

$$R^2 = 1 - [\Sigma(y_i - \hat{y}_i)^2 \div \Sigma(y_i - \bar{y})^2] \quad (8)$$

Mean absolute error is calculated as:

$$\text{MAE} = (1 \div n) \times \Sigma|y_i - \hat{y}_i| \quad (9)$$

Root mean square error is calculated as:

$$\text{RMSE} = \sqrt{[(1 \div n) \times \Sigma(y_i - \hat{y}_i)^2]} \quad (10)$$

The final model-analysis notebook used the final Phase 8 dataset and the saved Random Forest models developed during the project. A 20% evaluation subset was used to calculate the reported metrics.

7. Simulated IoT Communication

The Internet of Things (IoT) component is implemented as a software simulation rather than a physical deployment.

An ESP32 is simulated in Wokwi and is used to represent the sensor-side device. Virtual pressure values are packaged as a JSON message. The message is transferred through Message Queuing Telemetry Transport (MQTT) to a Python subscriber application.

The Python application receives the MQTT message, extracts the pressure values and Demand Factor, and passes them to the saved Random Forest models. The resulting predictions are then processed by the decision-support layer.

The IoT implementation is intended to demonstrate the communication concept. It should not be interpreted as a physical sensor deployment.

8. Maintenance Decision Support

The prediction models provide technical outputs, but an operator also needs a simple way to interpret these outputs. A rule-based decision-support layer is therefore used after the prediction stage. The layer considers the predicted severity and estimated water loss and assigns a project-defined maintenance priority.

The priority levels used in the prototype are Low, Medium, High, and Critical.

The decision-support layer also produces a short maintenance recommendation. For example, a Low-priority condition may result in a recommendation to continue monitoring, while a higher-priority condition can result in a recommendation for scheduled or immediate inspection.

These rules are project-specific. They are not intended to replace maintenance standards or professional engineering judgement.

This separation between prediction and decision support also allows the maintenance thresholds to be modified without retraining the machine-learning models.

9. Results and Discussion

9.1 Overall Model Performance

The results obtained from the saved models are summarized in Table 1.

Table 1: Performance of the Existing Saved Models

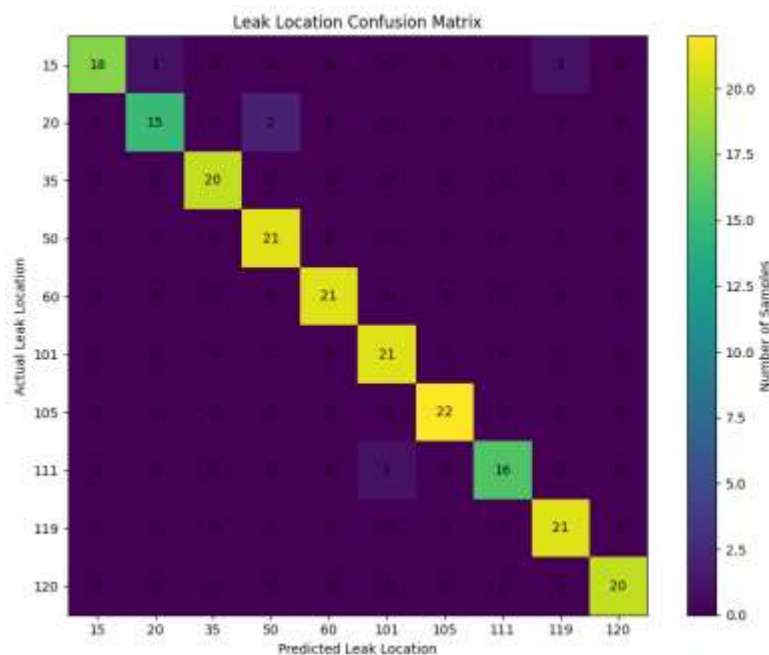
Model	Result
Leak Location	97.50% Accuracy
Leak Severity	98.50% Accuracy
Water Loss	$R^2 = 0.9979$
Water Loss	MAE = 141,368.90 LPD
Water Loss	RMSE = 310,723.70 LPD

The location model correctly classified 195 of the 200 evaluation samples, resulting in 97.50% accuracy. The severity model correctly classified 197 of the 200 evaluation samples, resulting in 98.50% accuracy. The water-loss model achieved an R^2 value of 0.9979, together with an MAE of 141,368.90 LPD and an RMSE of 310,723.70 LPD.

These results indicate that the saved models capture strong relationships between the simulated pressure/demand features and the target variables. However, the values should be interpreted within the simulation-based evaluation setting.

9.2 Leak-Location Classification

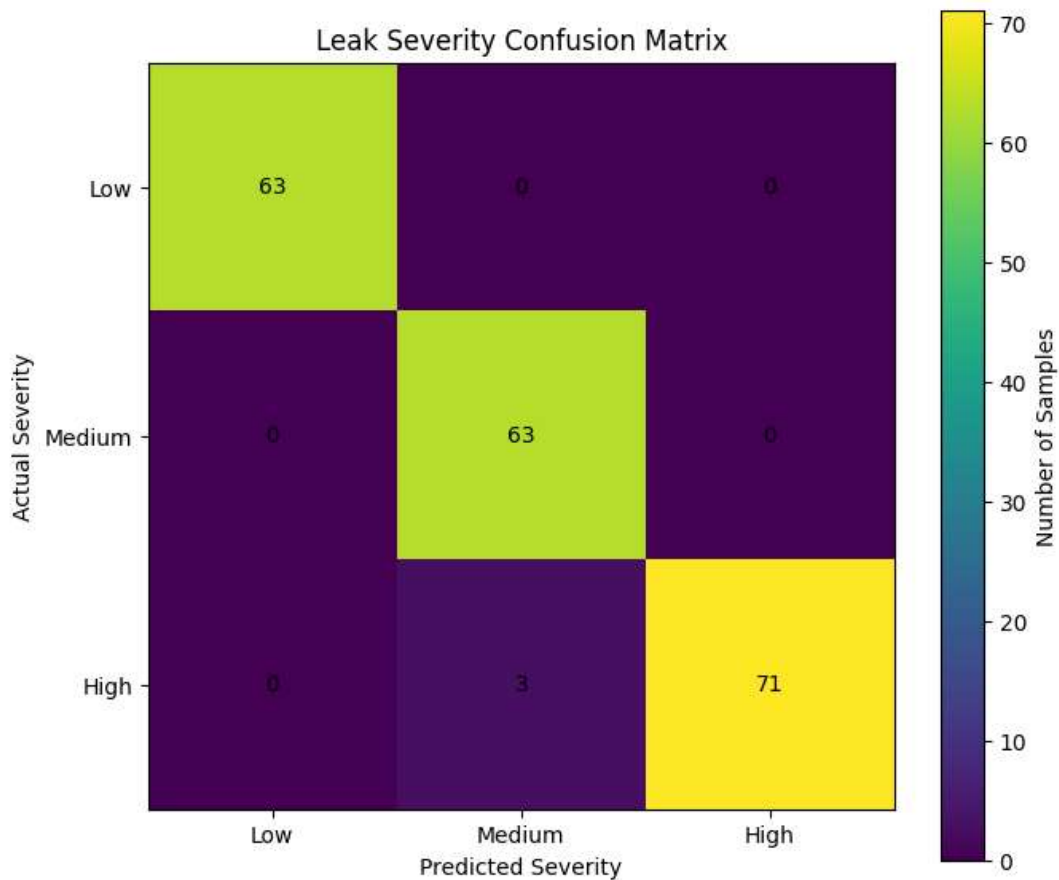
Figure 2: Confusion Matrix for Leak-Location Prediction.



The model correctly classified 195 of the 200 evaluation samples. Five samples were misclassified. The main errors occurred between locations 15 and 20, between locations 15 and 119, between locations 20 and 50, and between locations 111 and 101. Such errors indicate that some leakage scenarios can produce similar pressure patterns. The result supports the usefulness of pressure measurements for location prediction but also shows why the model should be treated as a decision-support component rather than as an unquestionable physical measurement of leak position.

9.3 Leak-Severity Classification

Figure 3: Confusion Matrix for Leak-Severity Classification.



The severity confusion matrix is shown in Figure 3. The model correctly classified all 63 Low-severity samples and all 63 Medium-severity samples in the evaluation subset. Among the High-severity samples, 71 were correctly classified and three were classified as Medium. The resulting accuracy is 98.50%. The small number of errors occurred between neighbouring severity categories, which is reasonable for simulated conditions where pressure responses can overlap.

9.4 Water-Loss Regression

Figure 4: Actual Versus Predicted Water Loss for the Evaluation Data.

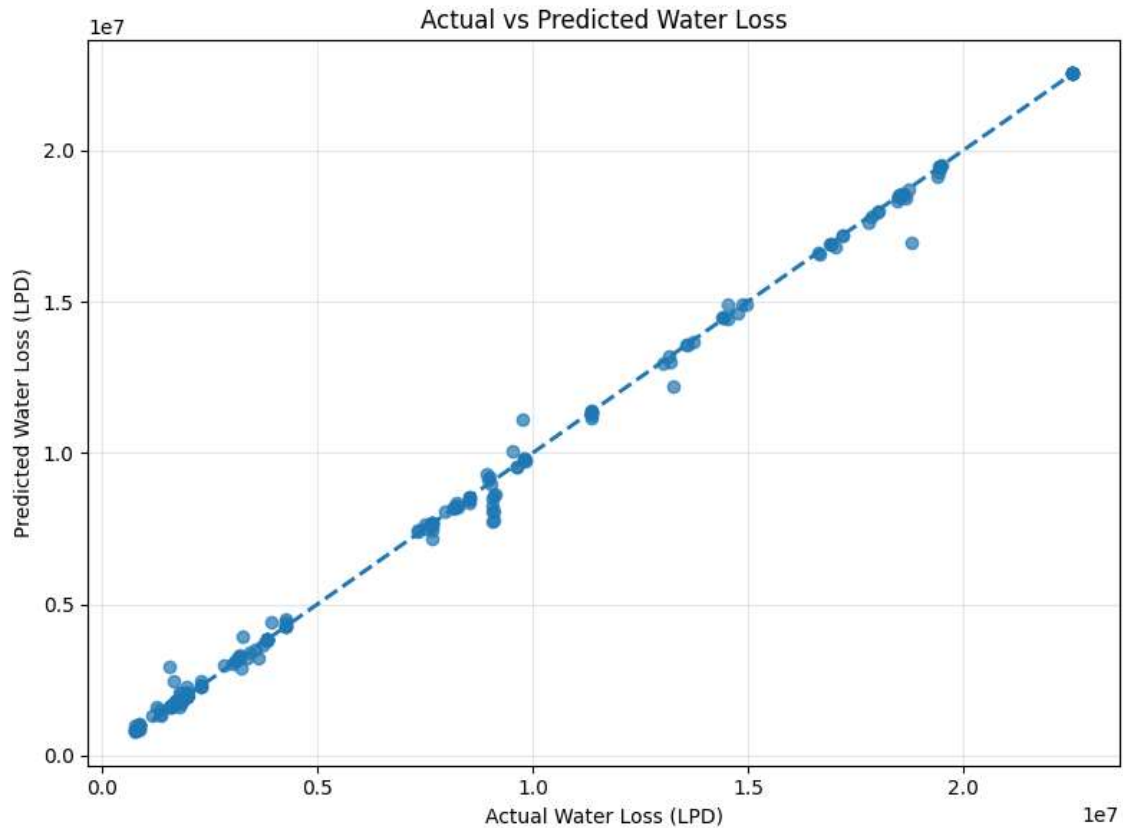
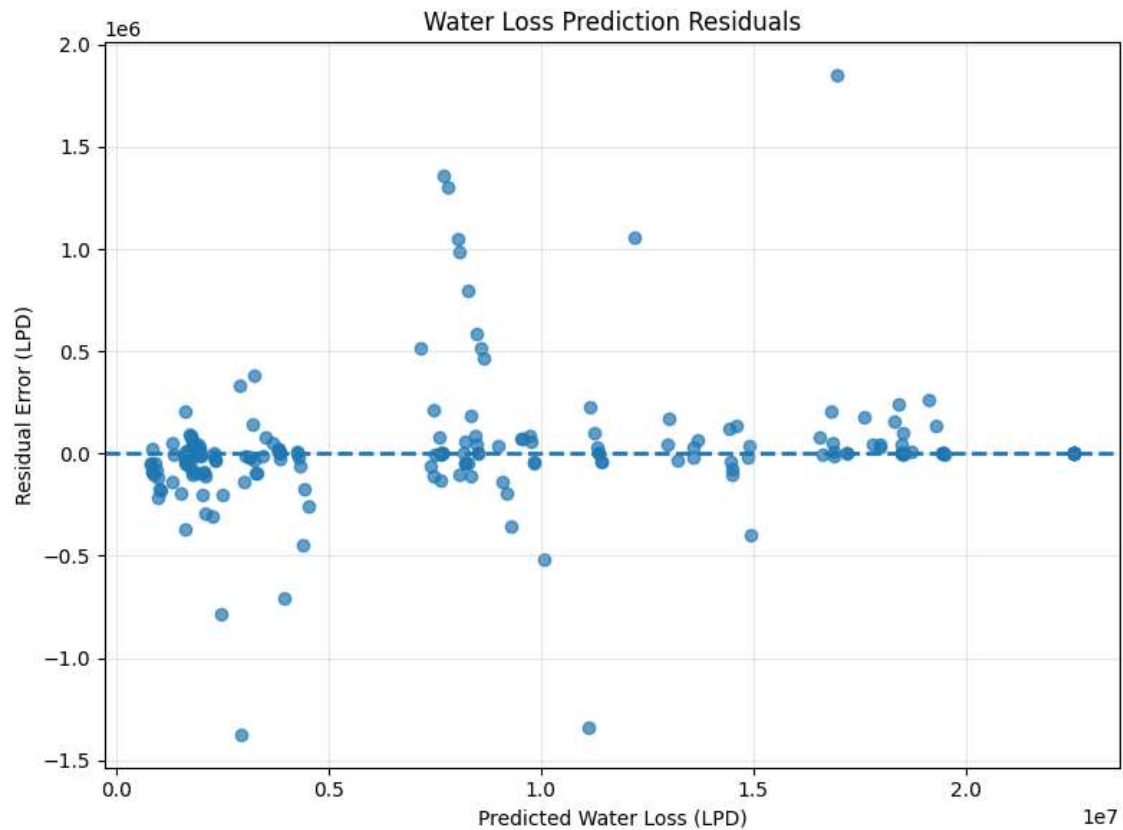


Figure 5: Residual Errors of the Water-Loss Regression Model.



The actual-versus-predicted water-loss graph is shown in Figure 4.

Most points are located close to the diagonal reference line, which represents equal actual and predicted values. This visual pattern is consistent with the high R^2 value of 0.9979.

The residual plot is shown in Figure 5.

The residuals are generally distributed around zero, although some larger errors can be observed for particular water-loss ranges. The MAE and RMSE provide additional information about these prediction errors.

9.5 Feature Importance

Figure 6: Feature Importance for the Water-Loss Prediction Model.

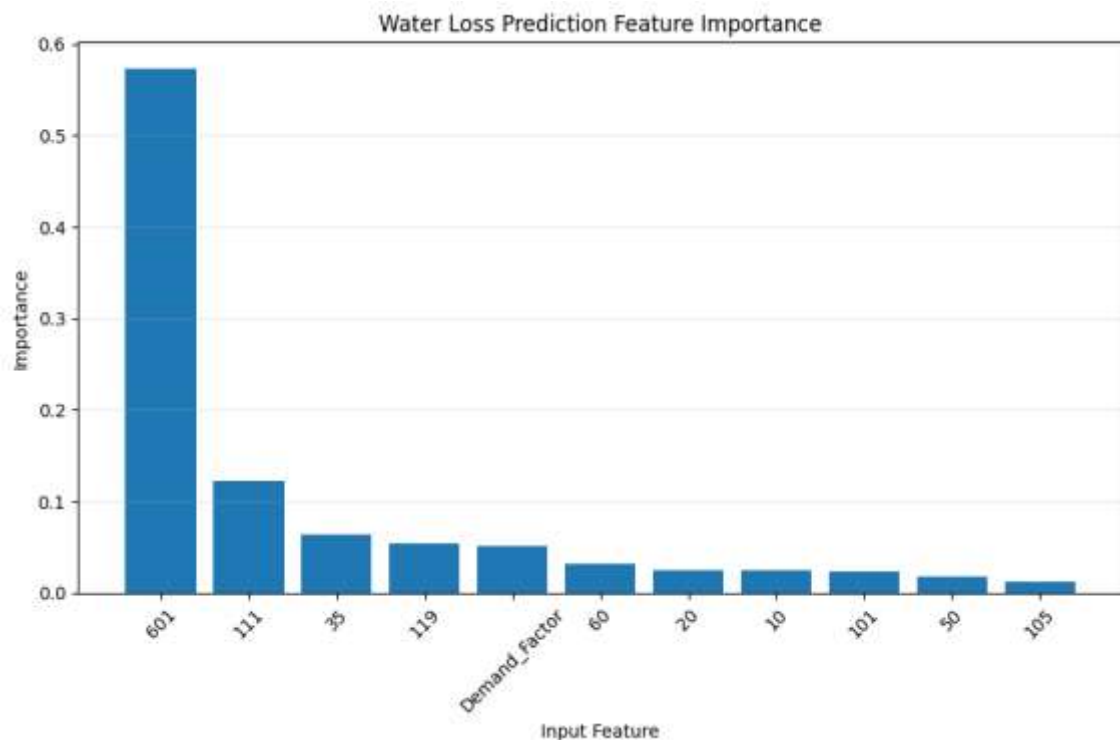
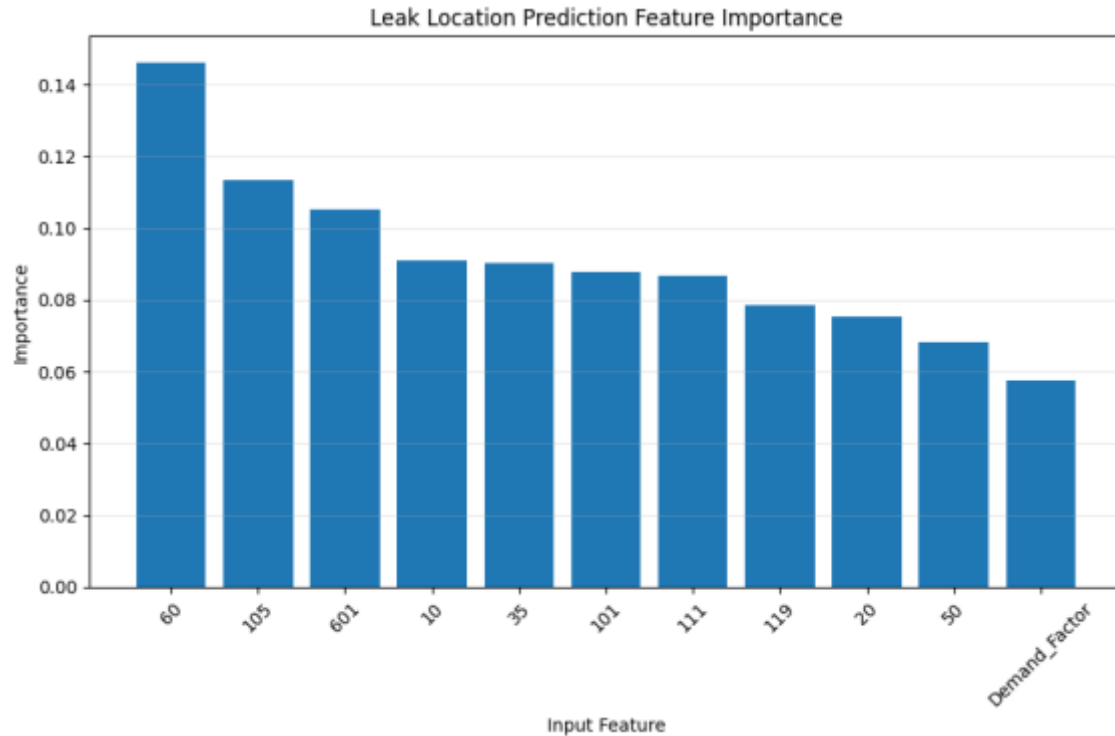


Figure 7: Feature Importance for the Leak-Location Prediction Model.



Feature importance was examined for the saved Random Forest models.

For the water-loss model, Node 601 has the highest feature importance, followed by Node 111, Node 35, Node 119, and Demand Factor.

The high importance of Node 601 does not mean that Node 601 is the physical leak location. Feature importance indicates how strongly the model uses an input feature when making predictions. It does not directly identify the physical source of a leak.

For the leak-location model, Node 60 has the highest feature importance, followed by Node 105, Node 601, Node 10, and Node 35.

The difference between the two feature-importance graphs is useful because the two models solve different prediction tasks. A feature can be important for estimating water loss without being the most important feature for identifying leak location.

9.6 MQTT-Based Prediction

Figure 8: Example Output Produced After Receiving Simulated Sensor Data Through MQTT.

```
=====
SMART WATER LEAKAGE DETECTION SYSTEM
=====

Leak Location      : 111
Leak Severity      : Low
Estimated Water Loss : 1,804,078.97 LPD
Priority           : Low
Recommendation      : Continue monitoring.
                    No immediate maintenance required.

=====
```

During the prototype demonstration, the ESP32 simulation successfully connected to Wi-Fi and the MQTT broker. Pressure values and Demand Factor were published as a JSON message and received by the Python subscriber.

The received data were passed to the existing machine-learning models, and the application generated leakage location, severity, water-loss estimation, priority, and recommendation.

An example of the model output is shown in Figure 8.

This demonstrates that the communication and prediction components can operate as a connected prototype rather than as isolated experiments.

9.7 Dashboard

Figure 9: Streamlit Dashboard Displaying Leakage Predictions and Maintenance Information.



The final predictions are presented through a Streamlit dashboard. An example dashboard view is shown in Figure 9.

The dashboard combines the predicted location, severity, estimated water loss, maintenance priority, and recommendation in one interface. This makes the output easier to interpret than examining individual model results separately.

10. Discussion

The results show that the proposed prototype can perform three related prediction tasks from the same pressure and demand inputs.

The location model achieved 97.50% accuracy, showing that the simulated pressure patterns provide useful information for distinguishing the predefined leakage-location classes. The remaining errors demonstrate that some locations can produce similar pressure responses.

The severity model achieved 98.50% accuracy. The confusion matrix shows that the few errors occurred between Medium and High severity, while Low and Medium cases were correctly classified in the evaluation subset.

The water-loss model achieved an R^2 value of 0.9979. The actual-versus-predicted graph also shows a strong alignment between the predicted and actual values.

The feature-importance results provide an additional interpretation of the models. Node 601 is the most important feature for water-loss prediction, whereas Node 60 is the most important feature for location prediction. This difference supports the use of separate models for different prediction tasks.

However, these results should not be interpreted as direct evidence of equivalent performance on a physical water network. The final dataset was generated using hydraulic simulation, and the sensor layer is simulated. Real pressure sensors can introduce noise, calibration errors, missing values, communication failures, and other uncertainties.

11. System-Level Contribution

The individual technologies used in this work are established. Random Forest has been widely studied, pressure-based leakage analysis is an established research direction, EPANET and WNTR are established hydraulic modelling tools, and MQTT is a commonly used IoT communication protocol.

The contribution of this work is therefore not presented as a new algorithm.

Instead, the work integrates several established components into a single prototype. The system connects hydraulic leakage simulation, pressure-based machine learning, severity classification, water-loss estimation, maintenance decision support, simulated sensor communication, and dashboard visualization.

This integration provides a complete path from simulated network behaviour to maintenance-oriented information.

A further practical contribution is the separation between prediction and decision support. The machine-learning models produce technical predictions, while the rule-based layer converts those predictions into a priority and recommendation. This allows maintenance rules to be modified without retraining the prediction models.

12. Limitations

The first limitation is that the dataset is simulation-based. The machine-learning models therefore learn relationships generated by the WNTR hydraulic model rather than directly from field measurements.

The second limitation is that the pressure sensors and ESP32 are simulated. Physical sensors may behave differently because of noise, calibration drift, sensor failure, and installation conditions.

The third limitation is the use of a single benchmark network. The NET3 topology does not represent every possible real-world water distribution network.

The fourth limitation is that the leakage-location model predicts predefined location classes. It does not currently perform exact pipe-segment localization.

The fifth limitation is that the severity classes are based on the simulated leakage scenarios. They are not physical field measurements of leak severity.

The sixth limitation is that the maintenance priority rules are project-defined. They should be validated with water-utility experts before operational use.

Finally, the current evaluation is based on the final Phase 8 dataset and saved models. Further experiments with independent datasets, cross-validation, additional demand conditions, and real measurements would provide stronger evidence of generalization.

13. Future Work

The next stage of the work is physical validation. Pressure sensors can be connected to ESP32 devices and tested on a small laboratory water-distribution setup. This would allow the communication and prediction pipeline to be evaluated using real sensor measurements.

Another direction is sensor reduction. The current framework uses ten monitoring nodes. A systematic sensor-placement and feature-selection study can investigate whether a smaller number of sensors can maintain acceptable prediction performance. Previous research has shown that sensor placement can strongly affect leakage identification [7].

The dataset can also be expanded with more leakage sizes, demand conditions, sensor noise, missing measurements, and different network states. Such variations could improve robustness.

Future work can also compare Random Forest with gradient-boosting models, neural networks, and graph-based learning. Graph neural networks are particularly relevant because water distribution networks naturally contain nodes and connections.

The maintenance layer can also be improved by incorporating additional operational information such as pipe age, pipe material, repair history, and estimated repair cost.

Finally, the simulated MQTT communication can be extended to a physical or cloud-based monitoring platform capable of generating alerts when a high-priority leakage condition is detected.

14. Conclusion

This paper presented an integrated data-driven framework for water leakage detection and maintenance decision support in water distribution networks.

The proposed system uses the EPANET NET3 benchmark network and WNTR to generate simulated leakage conditions and pressure observations. Pressure values from ten selected monitoring nodes and the network Demand Factor are used by three Random Forest models to predict leakage location, classify leakage severity, and estimate daily water loss.

The final Phase 8 dataset contains 1000 records and 16 variables. Using the saved project models and the evaluation procedure described in this paper, the leak-location model achieved 97.50% accuracy, the leak-severity model achieved 98.50% accuracy, and the water-loss model achieved an R^2 value of 0.9979. The water-loss model also produced an MAE of 141,368.90 LPD and an RMSE of 310,723.70 LPD.

The system also demonstrated simulated sensor communication through an ESP32/Wokwi environment and MQTT. The received pressure data were passed to the Python prediction application and converted

into leakage predictions and maintenance recommendations. A Streamlit dashboard was used to present the results.

The main contribution is the integration of hydraulic simulation, machine learning, simulated IoT communication, water-loss estimation, and maintenance decision support into a single prototype. The work does not claim that the individual algorithms or technologies are new.

The results are encouraging within the simulated environment, but real-world deployment requires additional validation. Physical pressure sensors, real water-distribution datasets, broader network conditions, and expert-validated maintenance rules are important directions for future work.

References

1. Hu Z., Chen B., Chen W., Tan D., Shen D., “Review of model-based and data-driven approaches for leak detection and location in water distribution systems”, *Water Supply*, 2021, 21 (7), 3282–3306. <https://doi.org/10.2166/ws.2021.101>
2. Lucin I., Lucin B., Carija Z., Sikirica A., “Data-driven leak localization in urban water distribution networks using big data for random forest classifier”, *Mathematics*, 2021, 9 (6), 672. <https://doi.org/10.3390/math9060672>
3. Lucin I., Carija Z., Dru zeta S., Lucin B., “Detailed leak localization in water distribution networks using random forest classifier and pipe segmentation”, *IEEE Access*, 2021, 9, 155113–155122. <https://doi.org/10.1109/ACCESS.2021.3129703>
4. Romero-Ben L., Alves D., Blesa J., Cembrano M.G., Puig V., Duviella E., “Leak detection and localization in water distribution networks: Review and perspective”, *Annual Reviews in Control*, 2023, 55, 392–419. <https://doi.org/10.1016/j.arcontrol.2023.03.012>
5. Tornyeviadzi H.M., Seidu R., “Leakage detection in water distribution networks via 1D CNN deep autoencoder for multivariate SCADA data”, *Engineering Applications of Artificial Intelligence*, 2023, 122, 106062. <https://doi.org/10.1016/j.engappai.2023.106062>
6. Wu X., Peng S., Zheng G., Fang X., Tian Y., “Leakage detection in water distribution networks based on multi-feature extraction from high-frequency pressure data”, *Water*, 2023, 15 (6), 1187. <https://doi.org/10.3390/w15061187>
7. Yang G., Wang H., “Optimal pressure sensor deployment for leak identification in water distribution networks”, *Sensors*, 2023, 23 (12), 5691. <https://doi.org/10.3390/s23125691>
8. Tornyeviadzi H.M., Mohammed H., Seidu R., “Dynamic segment criticality analysis: A precursor to scheduling of maintenance routines in water distribution networks”, *Alexandria Engineering Journal*, 2022, 61 (12), 9261–9272. <https://doi.org/10.1016/j.aej.2022.03.012>
9. Breiman L., “Random forests”, *Machine Learning*, 2001, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
10. Klise K.A., Bynum M., Moriarty D., Murray R., “A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study”, *Environmental Modelling & Software*, 2017, 95, 420–431. <https://doi.org/10.1016/j.envsoft.2017.06.022>
11. Rossman L.A., *EPANET 2 Users Manual*, U.S. Environmental Protection Agency, Cincinnati, OH, USA, 2000.