Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sensor-Efficient Machine Learning for Leak Detection, Localization, and Loss Estimation in an Experimental Water Distribution Network

Awni Al-otoom

awni_otoom@just.edu.jo

Jordan University of Science and Technology

Research Article

Keywords:

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11156716/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Water losses in distribution networks are difficult to identify because pressure and flow variations caused by leakage can be confounded with changes in normal hydraulic operating conditions. This study develops and experimentally evaluates a machine-learning framework for detecting, localizing, and quantifying water leakage under variable valve configurations, including both single and simultaneous leaks. Experiments were conducted on a laboratory-scale hydraulic network instrumented with pressure and flow sensors, generating 2,447 steady-state observations comprising 373 no-leak observations, 576 observations with leakage at location L1, 760 at location L2, and 738 with simultaneous L1 + L2 leakage. Hydraulic measurements and actuator states were used to classify the four operating conditions and estimate total water loss. To reduce information leakage between hydraulically similar observations and provide a more rigorous assessment of generalization, model performance was evaluated using grouped cross-validation. Random Forest achieved a leak-state localization accuracy of 0.974, while Extra Trees estimated total leakage with $R^2 = 0.974$, MAE = 0.159 L min^{-1}, and RMSE = 0.287 L min^{-1}. A reduced-input analysis using only four pressure measurements retained substantial diagnostic capability, achieving a localization accuracy of 0.859. Matched operating-condition analysis further demonstrated that leaks produced spatially differentiated pressure responses across the network, supporting the physical basis of the learned classification patterns. The results demonstrate that single and simultaneous leaks can be distinguished from normal hydraulic variations and that leakage magnitude can be estimated accurately using a relatively small sensor set. The proposed approach provides an experimentally validated basis for sensor-efficient, data-driven monitoring of water losses and supports the development of scalable leak-detection systems for water distribution networks.

1. Introduction

Water losses remain a persistent operational, economic, and resource-management challenge in water distribution systems (WDSs). Leakage results in the loss of treated water, increases treatment and pumping requirements, and may remain concealed for extended periods. Reliable leakage monitoring therefore requires more than identifying an abnormal hydraulic condition. For effective diagnosis and intervention, it is desirable to determine whether leakage is present, identify its probable location, and estimate the magnitude of the associated water loss. Data-driven methods are increasingly being investigated for these purposes, although the availability and quality of representative leakage data remain important limitations to their development and application [1].

Conventional leakage detection includes field inspection, acoustic techniques, pressure and flow analysis, water-balance methods, and hydraulic-model-based diagnosis. Increasing availability of operational measurements has encouraged data-driven approaches that extract leakage information directly from measurable hydraulic variables. Machine-learning (ML) techniques are particularly attractive because they can represent nonlinear relationships between leakage and hydraulic response without requiring all relationships to be prescribed explicitly. Previous studies have demonstrated the use of pressure and flow measurements for data-driven leak detection and localization using classifiers, artificial neural networks, and other ML approaches [2–4].

The hydraulic response to leakage, however, is not unique. Pressure and flow also vary in response to demand, valve operation, pump conditions, boundary conditions, and interactions among different flow paths. A pressure reduction or increase in system flow therefore cannot automatically be interpreted as evidence of leakage. The central diagnostic problem is to distinguish hydraulic changes associated with abnormal water loss from those produced by legitimate variations in network operation.

This distinction also has important implications for model validation. A data-driven model may produce apparently high predictive performance when hydraulically or temporally related observations occur in both training and testing datasets. Ajoodani et al. [5] demonstrated that conventional random data partitioning can substantially overestimate leak-localization performance relative to evaluation strategies that preserve the temporal structure of hydraulic data. Consequently, evaluation procedures should be designed to test whether learned leakage signatures generalize beyond observations that closely resemble those used for model training.

Machine learning provides a useful framework for this problem because leakage can generate multivariable hydraulic patterns that may be difficult to identify using individual thresholds. Rather than interpreting each sensor independently, ML models can learn relationships among pressure, flow, operating conditions, and leakage state. Tyagi et al. [3], for example, developed a two-stage data-driven framework based on pressure and flow information from a limited number of sensor locations. Their method used changes in prediction-error distributions to identify and localize leakage and was evaluated using EPANET-generated data for the Hanoi, Net3, and C-Town benchmark networks. Sun et al. [2] similarly demonstrated pressure-based data-driven localization using a limited measurement architecture.

Experimental investigations provide a complementary perspective because they evaluate leakage signatures using measurements affected by actual sensor response, hydraulic variability, and physical system behavior. Mahdi et al. [4] used measurements from a laboratory-scale water distribution test rig with looped and branched configurations and demonstrated ML-based leakage detection and localization using dynamic pressure and vibration-derived features. Such experimental studies are important for determining whether diagnostic patterns identified computationally remain distinguishable in measured physical systems.

A major advantage of multivariable diagnosis is its ability to exploit the spatial hydraulic response rather than interpreting individual pressure measurements independently. A leak may cause only a modest response at one sensor, whereas the combined response across several monitoring locations can form a distinguishable pattern. For a network instrumented with n pressure sensors, the pressure state can be represented conceptually as

$$P = [P1, P2, ..., Pn]$$

and, relative to a comparable no-leak operating condition, the pressure perturbation can be represented as

$$\Delta P = [\Delta P1, \Delta P2, ..., \Delta Pn].$$

If leakage at different locations generates reproducibly different pressure-response vectors, localization becomes physically plausible because the classification model can learn location-dependent manifestations of the network hydraulics rather than simply detecting a generic pressure decrease. This concept is consistent with previous pressure-based localization studies in which spatial pressure information is used to distinguish candidate leakage locations [2, 6].

The amount of instrumentation required to obtain this information is an important practical consideration. Increasing the number of sensors generally increases hydraulic information but also increases equipment, installation, calibration, communication, and maintenance requirements. Pressure-sensor placement has consequently become an important research topic. Khorshidi et al. [7] developed a hybrid information-entropy framework for pressure-sensor placement, while Santos-Ruiz et al. [6] used information theory to maximize leak-location information while minimizing redundancy among pressure sensors. These studies demonstrate that

diagnostic performance should be considered together with sensing requirements rather than assuming that denser instrumentation is necessarily the preferred solution.

This motivates an important sensor-efficiency question: how much leakage information can be retained when diagnosis is restricted to a small set of spatial pressure measurements?

A pressure-only architecture is attractive because pressure sensors can capture the distributed hydraulic effects of leakage without requiring flow measurement at every branch. However, removing flow and actuator information also removes predictors that may help distinguish different operating conditions. Sensor reduction therefore creates a trade-off between instrumentation burden and available diagnostic information. The relevant engineering objective is not necessarily to reproduce the performance of a fully instrumented architecture, but to determine whether useful detection and localization capability can be retained with substantially reduced sensing.

The diagnostic challenge becomes more complex when multiple leaks occur simultaneously. Different combinations of leakage locations and magnitudes may generate overlapping hydraulic responses, increasing the nonuniqueness of the localization problem. Multiple-leak scenarios have therefore received increasing research attention. Li et al. [8] developed a framework for detecting and localizing multiple leaks by combining hydraulic simulation and machine learning. More recently, Wang [9] proposed a hybrid hydraulic-simulation and graph-transformer framework capable of addressing concurrent abrupt and incipient leaks. Bandreddi et al. [10] developed a multistage localization methodology specifically for an unknown number of simultaneous leaks under hydraulic uncertainty. These studies demonstrate that multiple-leak localization is an emerging extension of conventional single-leak diagnosis.

Simultaneous leakage is not necessarily equivalent to considering two independent single-leak events. An additional leakage pathway changes network flow distribution and the pressure field and may consequently alter hydraulic conditions elsewhere in the system.

Despite these advances, an important gap remains between the methodological capability demonstrated in computational studies and the experimental evidence needed to establish whether single and simultaneous leaks can be reliably distinguished under variable physical operating conditions. Recent multiple-leak studies have demonstrated increasingly sophisticated localization approaches [8–10], but have relied predominantly on hydraulic simulation, benchmark networks, or model-generated data. Comparatively less attention has been given to controlled experimental evaluation in which single and simultaneous leaks are physically introduced while the normal hydraulic operating state is systematically varied. This distinction is important because changes in valve configuration can themselves produce substantial variations in pressure and flow, potentially resembling or masking the hydraulic response to leakage. A robust diagnostic framework must therefore distinguish the signature of leakage from the variability associated with normal system operation rather than merely discriminate among leakage scenarios generated under fixed or simulated conditions.

The present study addresses this gap using a controlled experimental hydraulic network in which no-leak, two individual leak locations, and simultaneous leakage are investigated across a wide range of valve configurations. The study is designed to examine three related questions: whether single and simultaneous leakage states can be distinguished from each other and from normal hydraulic variability; whether the resulting total water loss can be quantitatively estimated; and how much localization capability can be retained when the available information is reduced to pressure measurements alone. In addition to predictive modeling, leak observations are compared with hydraulically comparable no-leak conditions to determine whether the different leakage states produce identifiable

spatial pressure signatures. The study therefore combines experimental leak-state classification, leakage quantification, sensor-reduction analysis, and hydraulic interpretation within a common framework.

A combined leakage condition, used in this study, can therefore produce a multivariable response that differs from the response associated with either individual leak.

For the two leakage locations considered in the present study, the diagnostic state space is therefore defined as

$C \in \{\text{No Leak}, L1, L2, L1 + L2\}$,

where L1 and L2 represent individual leakage locations and L1 + L2 represents their simultaneous occurrence. Treating these states separately allows the model to determine whether measured hydraulic information contains sufficient information not only to identify leakage but also to distinguish individual from combined leakage conditions.

Leak monitoring also involves more than classification. Detection establishes whether an abnormal condition exists, localization identifies its probable source or region, and quantification estimates the magnitude of the associated loss. These functions can be regarded as progressively richer levels of hydraulic diagnosis:

Detection → Localization → Quantification.

Localization can reduce the region requiring inspection, while quantification provides information on leakage severity and the amount of water being lost. In the present study, classification and regression are therefore treated as complementary tasks. Classification is used to identify the leakage state, while regression is used to estimate total water loss.

Another methodological challenge concerns the availability of representative leakage data. Wu et al. [1] identified data scarcity as a major obstacle to data-driven leakage detection, particularly because observed and labeled leakage data are difficult and costly to obtain. They highlighted data creation, reduction of model data requirements, and knowledge transfer as important research directions. Controlled physical experiments provide one means of generating labeled leakage observations while maintaining knowledge of the operating condition, leakage location, and leakage magnitude.

Physical interpretation of the resulting models is also important. High classification accuracy alone does not establish that an ML model has learned hydraulically meaningful leakage behavior. Feature-importance measures can identify influential predictors, but they do not directly demonstrate how leakage modifies pressure across the network. A complementary approach is to compare leak and no-leak observations obtained under identical or closely comparable operating configurations.

For pressure sensor i, the leakage-associated perturbation can be expressed as

$\Delta P_i = P_{i,leak} - P_{i,control}$.

The resulting vector across all pressure sensors represents a spatial leakage fingerprint. Comparing these fingerprints for L1, L2, and L1 + L2 provides a physical interpretation of whether different leakage states produce distinguishable spatial effects. Matched-condition analysis is particularly useful when valve configurations vary across the experimental program because it helps separate pressure changes associated with leakage from those resulting from normal actuator changes.

Despite substantial progress in data-driven leakage detection and localization, several methodological issues therefore remain relevant. Representative leakage data are limited [1]; diagnostic performance depends strongly on the information available from the sensing architecture [6, 7]; validation procedures can yield optimistic results when the structure of hydraulic observations is not adequately considered [5]; and simultaneous leakage introduces additional localization complexity [8–10]. Furthermore, much of the published work uses hydraulic simulations or benchmark networks [2, 3, 6, 8–10], while experimental studies provide an important complementary basis for assessing whether diagnostic signatures remain identifiable in measured physical systems [4].

The present study addresses these issues through controlled experiments on a hydraulic network operated across multiple valve configurations under four conditions: no leak, leakage at L1, leakage at L2, and simultaneous leakage at L1 and L2. The resulting dataset contains 2,447 steady-state observations covering normal and abnormal hydraulic states. Multiple ML algorithms are evaluated for two complementary tasks: multiclass leakage-state localization and continuous estimation of total water loss. Generalization is assessed using grouped cross-validation so that observations associated with related operating configurations are kept together during model evaluation, reducing the potential for optimistic performance estimates arising from closely related training and validation observations.

A reduced pressure-only architecture using four pressure measurements, P1–P4, is additionally evaluated to determine the localization capability retained after substantial reduction of the available hydraulic information. Leakage observations are also compared with matched or closely comparable no-leak operating conditions to examine the spatial pressure perturbations associated with L1, L2, and simultaneous L1 + L2 leakage.

Accordingly, the contribution of this study lies in experimentally integrating four related diagnostic functions within a common data-driven framework: distinguishing leakage from variable hydraulic operation, localizing single and simultaneous leakage states, quantitatively estimating total water loss, and evaluating the diagnostic capability retained under pressure-only sensing. Rather than claiming the first use of ML for leakage detection or the first treatment of multiple leaks, the study examines whether these functions can be achieved together using controlled physical measurements under variable operating conditions. The combination of experimental validation, grouped evaluation across operating configurations, explicit simultaneous-leak classification, leakage-flow regression, sensor reduction, and matched-condition hydraulic interpretation provides a framework for assessing both the predictive performance and physical basis of sensor-efficient water-loss monitoring.

2. Materials and Methods

2.1 Experimental Framework

A laboratory-scale hydraulic network was used to investigate the detection, localization, and quantification of water leakage under variable operating conditions. The experimental system was designed to reproduce controlled changes in network hydraulic resistance through three independently adjustable control valves while simultaneously recording flow, pressure, and pump operating variables. Two leakage locations, denoted L1 and L2, were incorporated into the network to permit experiments with individual and simultaneous leakage.

Four experimental conditions were considered: normal operation without intentional leakage (CONTROL), leakage at L1, leakage at L2, and simultaneous leakage at both locations (L1 + L2). During the experiments, the three control-valve openings were varied to generate a broad range of hydraulic operating states. This design was

intended to prevent leakage diagnosis from being restricted to a single hydraulic condition and to test whether leakage signatures could be distinguished from pressure and flow variations generated by normal changes in valve configuration.

The experimental framework therefore incorporated two sources of hydraulic variability: controlled changes in network operation through valve adjustment and abnormal water loss through L1, L2, or their simultaneous activation. This distinction was fundamental to the study because the objective was to identify leakage within a variable operating envelope rather than under a single fixed hydraulic condition.

A schematic diagram of the experimental system, showing the pump, control valves, flowmeters, pressure measurement locations, and leakage points L1 and L2, is presented in Fig. 1.

2.2 Experimental Hydraulic System and Leak Configuration

The experiments were conducted using a laboratory-scale recirculating hydraulic network designed to provide controlled variation in hydraulic resistance, branch-flow distribution, and leakage conditions. The system comprised a storage tank, a recycle tank, a main circulation pump, a recycle pump, three controlled flow branches, and approximately 7 m of nominal 3/4-in. hydraulic piping. The main and recycle pumps were centrifugal units rated at 380 W and 240 V. Water was supplied from the storage tank through the main circulation line, distributed through the experimental network, collected in the recycle tank, and returned through the recirculation system. The resulting closed-loop platform enabled hydraulic operating conditions and intentional water losses to be investigated systematically.

Three 3/4-in. gate valves, denoted CV1, CV2, and CV3, were installed in the controlled branches to vary hydraulic resistance and redistribute flow within the network. Each control valve was mechanically coupled to a 12-V bipolar stepper motor rated at 4.2 N·cm and operated through an L298N motor-driver module. Valve actuation was managed through an ESP32-based control architecture, with valve position represented on a normalized scale from 0% (fully closed) to 100% (fully open). Nominal valve openings of 0, 25, 50, 75, and 100% were used to generate a broad range of operating configurations.

The hydraulic network was instrumented to measure spatial flow and pressure responses together with pump electrical loading and control-valve positions. Four flow measurements (F1–F4) and four pressure measurements (P1–P4) were distributed across the experimental network. Pump operation was represented by RMS electrical current. The physical actuated control valves are designated CV1–CV3 in Fig. 1, while their recorded opening positions are denoted V1–V3 in the experimental dataset and subsequent machine-learning analysis. These variables constitute the hydraulic and operating-state information subsequently used by the machine-learning models, consistent with the input vector defined in the present analysis. The instrumentation used in the experimental system is summarized in Table 1.

Table 1
Instrumentation and actuator specifications of the experimental hydraulic system

Measurement/component	Designation	Device/specification	Measurement range / rating	Reported accuracy	Recorded quantity
Flow sensors	F1–F4	YF-S201	1–30 L min⁻¹	± 1%	Flow rate (L min⁻¹)
Leak Flow sensors	FL1, FL2	YF-S201	1–30 L min⁻¹	± 1%	Flow rate (L min⁻¹)
Pressure sensors	P1–P4	MPS20N0040D	1–40 kPa	± 0.3%	Pressure (kPa)
Pump-current sensor	I_{pump}	ACS712	Up to 5 A	± 5%	RMS electrical current (A)
Control valves	CV1–CV3	3/4-in. gate valves	0–100% opening	—	Valve opening (%)
Valve actuators	—	12-V bipolar stepper motors	4.2 N·cm rated torque	—	Mechanical valve actuation
Motor drivers	—	L298N	—	—	Stepper-motor control
Leak valves	L1, L2	1/2-in. manually adjustable ball valves	Variable manual opening	—	Controlled external leakage
Two intentional leakage locations, denoted L1 and L2, were physically incorporated into the network. Each leak was implemented using a manually adjustable 1/2-in. ball valve installed as a lateral outlet protruding from an existing branch pipe. Thus, the leakage outlets discharged water directly from the corresponding branch to the exterior of the hydraulic network and did not constitute additional recirculating branches. L1 was installed on the same hydraulic branch as CV1 and positioned upstream of CV1, whereas L2 was installed on the same branch as CV3 and positioned upstream of CV3. Leakage discharge was measured independently using dedicated flowmeters installed to quantify the water discharged through the L1 and L2 leak outlets. These leak-flow measurements were separate from the four network flow measurements (F1–F4) used as predictor variables in the machine-learning analysis For experiments involving an individual leak, the measured discharge through the corresponding leak outlet represented the total water loss. For simultaneous L1 + L2 experiments, total water loss was calculated as the sum of the independently measured leak discharges:					
Qloss=QL1 + QL2					

where Q_{L1} and Q_{L2} are the independently measured discharge rates through leak outlets L1 and L2, respectively. Consequently, the regression target, Total True Loss, was obtained independently of the network flow measurements F1–F4 used as model predictors.

Leak magnitude was controlled manually by adjusting the opening of the corresponding 1/2-in. ball valve. Four diagnostic conditions were investigated. Under CONTROL operation, both leak valves were closed. For the L1 condition, leakage was introduced through L1 while L2 remained closed; for L2, leakage was introduced through L2 while L1 remained closed. During the simultaneous L1 + L2 condition, both leak valves were opened concurrently. The simultaneous-leak condition therefore represented two physically occurring leakage pathways rather than a numerically constructed combination of independent single-leak measurements. This distinction is important because concurrent discharge through L1 and L2 modifies network flow distribution and pressure and can

therefore generate a hydraulic response that is not necessarily a simple superposition of the two individual leakage responses.

During the CONTROL and leakage experiments, CV1–CV3 were systematically varied to generate different combinations of network hydraulic resistance and flow distribution. The experimental design therefore incorporated two distinct sources of hydraulic variation: normal operational variation produced by control-valve adjustment and abnormal water loss produced by L1, L2, or simultaneous L1 + L2 leakage. This design was central to the study because the diagnostic models were required to distinguish leakage from hydraulic changes caused by normal valve operation rather than identify leakage under a single fixed network state.

The experimental platform employed a networked microcontroller architecture for valve control, measurement acquisition, and experimental-state management. An ESP32-based master controller coordinated the experimental sequence with the valve-control system, acquired the hydraulic measurements, and associated each measurement record with the corresponding operating configuration. Valve status was confirmed before measurements were associated with the requested operating state, reducing the possibility of assigning hydraulic observations to an actuator configuration that had not yet been achieved.

Measurements collected during changes in valve position or before hydraulic stabilization were distinguished from steady operation. Records designated as SETTLING or TRANSIENT were excluded from the present statistical and machine-learning analyses, and only observations classified as STEADY were retained. This filtering was intended to ensure that the models characterized stabilized hydraulic leakage signatures rather than short-duration responses associated with valve motion or system adjustment. The resulting analysis dataset contained 2,447 steady-state observations: 373 CONTROL, 576 L1, 760 L2, and 738 simultaneous L1 + L2 observations.

For each retained observation, the predictor vector was defined as

$$X = [F_1, F_2, F_3, F_4, P_1, P_2, P_3, P_4, I_{\mathrm{pump}}, V_1, V_2, V_3] \, ,$$

where F1–F4 are the measured flow rates, P1–P4 are the pressure measurements, $I_{\mathrm{pump}}$is the pump RMS current, and V1–V3 are the recorded opening positions of the physical actuated control valves CV1–CV3, respectively. This combined representation enabled subsequent analyses to determine whether leakage information originated primarily from direct hydraulic measurements or whether actuator and pump-state information provided substantial additional diagnostic information.

2.3 Instrumentation and Measured Variables

The hydraulic network was instrumented to record four flow variables, four pressure variables, pump electrical current, and the positions of the three actuated control valves CV1–CV3. The measured hydraulic variables were denoted F1–F4 for flow rate and P1–P4 for pressure. Flow rates were expressed in L min⁻¹ and pressure measurements in kPa. Pump operation was represented by the root-mean-square electrical current. The commanded opening positions of CV1–CV3 were expressed as percentages from 0% to 100%. For consistency with the experimental dataset, these three control-valve opening variables are denoted V1, V2, and V3 in the subsequent analysis. For each observation, the recorded information can therefore be represented by the hydraulic and operating-state vector

X = [F1, F2, F3, F4, P1, P2, P3, P4, Ipump, V1, V2, V3],

where F1–F4 are the measured flow rates, P1–P4 are the pressure measurements, Ipump is the pump RMS current, and V1–V3 represent the recorded opening positions of the actuated control valves CV1–CV3, respectively.

The combination of hydraulic measurements and actuator information allowed the ML models to distinguish changes caused by normal valve operation from changes associated with leakage. The pressure measurements were additionally evaluated independently as a reduced sensing architecture, as described in Section 2.14.

2.4 Experimental Conditions and Data Collection

Experiments were conducted separately for the four diagnostic states: CONTROL, L1, L2, and L1 + L2. The CONTROL experiments represented normal hydraulic operation without intentional leakage and provided the baseline operating dataset. The leakage experiments were subsequently conducted with leakage introduced at L1, L2, or simultaneously at both locations.

Leakage discharge was measured independently using dedicated flowmeters installed to quantify the water discharged through the L1 and L2 leak outlets. These leak-flow measurements were separate from the four network flow measurements (F1–F4) used as predictor variables in the machine-learning analysis. For experiments involving an individual leak, the measured discharge through the corresponding leak outlet represented the total water loss. For simultaneous L1 + L2 experiments, total water loss was calculated as the sum of the independently measured leak discharges:

$$Q_{\text{loss}} = Q_{L1} + Q_{L2}$$

where Q_{L1} and Q_{L2} are the independently measured discharge rates through leak outlets L1 and L2, respectively. Consequently, the regression target, Total True Loss, was obtained independently of the network flow measurements F1–F4 used as model predictors.

Leakage magnitude was varied experimentally by manually adjusting the 1/2-in. ball valves installed at the L1 and L2 leak outlets. The resulting leakage severity was characterized by the measured discharge through the dedicated leak-flow meters rather than by nominal ball-valve position. For each imposed leakage condition, the hydraulic system was allowed to reach a steady operating state before observations were retained for analysis. The recorded leak-flow measurements were used to characterize the actual leakage magnitude associated with each observation. This procedure allowed the experimental dataset to incorporate variation in both network operating configuration and leakage magnitude.

The three valve openings were systematically varied during the experimental program to generate different combinations of network hydraulic resistance and flow distribution. Consequently, the recorded dataset contained observations spanning a range of flow rates and pressure states rather than repeated measurements from only one operating point.

Only observations identified as steady-state records were retained for the statistical and ML analyses. Transient and settling observations were excluded to prevent short-duration responses following actuator adjustment or changes in experimental condition from being interpreted as representative leakage signatures.

Following filtering, the final dataset contained 2,447 steady-state observations. Of these, 373 represented CONTROL conditions, 576 represented leakage at L1, 760 represented leakage at L2, and 738 represented simultaneous L1 + L2 leakage. The resulting class distribution is summarized in Table 1.

Table 2
Distribution of steady-state observations used in the analysis.

Hydraulic state	Description	Number of observations
CONTROL	No intentional leakage	373
L1	Leakage at location L1	576
L2	Leakage at location L2	760
L1 + L2	Simultaneous leakage at L1 and L2	738
Total		2,447

The multiclass response variable used for leak-state classification was therefore defined as

$C \in \{CONTROL, L1, L2, L1 + L2\}$.

This formulation enabled the classification models to distinguish normal operation from leakage while simultaneously evaluating whether the hydraulic measurements could differentiate the two individual leakage locations from their combined occurrence.

2.5 Leakage Quantification

In addition to classification, the experimental data were used to formulate a continuous leakage-quantification problem. Total water loss was expressed in L min^{-1} and treated as the target variable for regression analysis. CONTROL observations were assigned zero leakage, whereas observations associated with L1, L2, and L1 + L2 were assigned the corresponding experimentally determined total loss.

The regression problem was formulated as

$Q_{loss} = f(X)$,

where Q_{loss} represents the total leakage rate and X represents the available hydraulic and operating variables.

This formulation was intentionally separated from the classification problem. The classification task identifies the leakage state, whereas the regression task estimates the magnitude of water loss. The two models therefore provide complementary diagnostic outputs: location/state information and leakage severity.

For simultaneous leakage, the regression target represents the total measured water loss associated with the combined L1 + L2 condition rather than treating the two leakage locations as independent classification outputs.

2.6 Data Preparation

The datasets obtained from the no-leak and leakage experiments were standardized into a common data structure before analysis. Variable names and units were harmonized across the experimental datasets, with P1–P4 consistently treated as pressure measurements in kPa and F1–F4 as flow rates in L min^{-1}.

Only records classified as steady state were included in subsequent analysis. Each retained observation was assigned one of the four diagnostic labels according to the experimental condition under which it was collected. The datasets were then combined into a common analysis dataset while preserving information describing the original experiment and hydraulic operating condition.

No-leak observations were assigned a total leakage rate of 0 L min^{-1}. Leakage observations retained their corresponding experimentally determined loss values for regression analysis.

This preprocessing produced a unified dataset containing the hydraulic measurements, actuator states, diagnostic class, leakage magnitude, and experimental identifiers required for subsequent statistical and ML analysis.

2.7 Machine-Learning Tasks

Two supervised ML tasks were developed.

The first task was multiclass leak-state classification. The objective was to predict

$C \in \{CONTROL, L1, L2, L1 + L2\}$

from the available hydraulic and operating variables.

The second task was regression-based leakage quantification. The objective was to predict the continuous total water-loss rate, Qloss, in L min^{-1}.

Leak detection and localization were formulated as a single four-class classification problem comprising CONTROL, L1, L2, and L1 + L2. Within this formulation, discrimination between CONTROL and any of the three leakage classes represents leak detection, whereas discrimination among L1, L2, and L1 + L2 represents identification of the leakage configuration. Thus, detection and localization were performed within a unified multiclass framework rather than through separate binary-detection and localization models.

Using separate classification and regression models allowed the study to address two distinct operational questions: where or under what leakage configuration is water being lost, and the magnitude of loss.

Multiple ML algorithms were evaluated rather than relying on a single model family. For classification, the evaluated algorithms included logistic regression, Random Forest, and Extra Trees. The same comparative principle was applied to the regression problem, allowing nonlinear ensemble methods to be compared with simpler baseline approaches.

Model selection was based on out-of-sample performance obtained using grouped cross-validation rather than model fit to the complete dataset.

2.9 Machine-Learning Algorithms

Candidate algorithms were selected to provide a comparison between relatively simple linear models and nonlinear tree-based ensemble methods. This comparison was considered appropriate because the hydraulic response to leakage may involve nonlinear interactions among pressure, flow, valve configuration, and leakage magnitude. Logistic Regression and Ridge Regression were therefore used as regularized linear reference models for classification and regression, respectively, while Random Forest and Extra Trees were evaluated as nonlinear ensemble-learning approaches.

Logistic Regression was used as the linear baseline for four-class leak-state classification. Logistic Regression models class probabilities as functions of a linear combination of the predictor variables and can be extended to multiclass classification through a multinomial formulation [11]. Its inclusion provided a useful reference for determining whether the four hydraulic states could be separated adequately using predominantly linear decision

relationships. Substantially better performance from the tree-based models would therefore indicate that nonlinear relationships and interactions among the hydraulic variables contribute important discriminatory information.

Random Forest is an ensemble-learning method that combines predictions from multiple decision trees. Individual trees are constructed using bootstrap samples of the training data, while random subsets of predictors are considered during node splitting. The resulting diversification among trees reduces the dependence on any individual tree and generally improves predictive stability and resistance to overfitting [12]. For classification, predictions from the individual trees are aggregated to determine the final class assignment. Random Forest was particularly relevant to the present application because tree-based models can represent nonlinear thresholds and interactions among pressure, flow, valve-opening, and pump-state variables without requiring these relationships to be specified explicitly.

Extra Trees (Extremely Randomized Trees) is also an ensemble of decision trees but introduces additional randomization during tree construction. In particular, both the candidate predictors and splitting thresholds are randomized more strongly than in conventional Random Forests [13]. Predictions are subsequently aggregated across the ensemble. This additional randomization can reduce variance and provide an effective representation of complex nonlinear relationships. Extra Trees was evaluated for both leak-state classification and continuous leakage estimation in the present study.

Ridge Regression was used as the linear reference model for total water-loss estimation. Ridge Regression extends ordinary least-squares regression by adding an L_2penalty to the magnitude of the regression coefficients [14]. The regularization term reduces sensitivity to correlated predictors and can improve numerical stability when multiple explanatory variables contain overlapping information. This characteristic was relevant to the present dataset because flow and pressure measurements obtained at different network locations may be hydraulically correlated. Ridge Regression therefore provided a regularized linear benchmark against which the nonlinear Random Forest and Extra Trees regression models could be evaluated.

The candidate models consequently represented increasing flexibility in the relationships that could be learned from the experimental data. Logistic Regression and Ridge Regression provided interpretable regularized linear baselines, whereas Random Forest and Extra Trees allowed nonlinear response surfaces, threshold effects, and interactions among predictors. All candidate algorithms were evaluated using the same grouped cross-validation framework described in the following section so that model comparisons reflected performance on withheld valve-defined operating configurations rather than differences in data partitioning.

Python Implementation and Computational Workflow

All data preprocessing, machine-learning model development, cross-validation, performance evaluation, and feature-importance analysis were implemented in Python. The analysis workflow was designed to preserve the experimental grouping structure throughout model training and evaluation. Following import of the experimental dataset, only observations classified as steady state were retained, and the predictor variables, leakage-state labels, total measured water loss, and valve-configuration information were extracted for subsequent analysis.

The machine-learning models were implemented using the Python scientific-computing environment and the scikit-learn library [15]. Classification models included Logistic Regression, Random Forest, and Extra Trees, whereas leakage-quantification models included Ridge Regression, Random Forest, and Extra Trees. The same

predictor definitions and experimental observations were used when comparing candidate algorithms so that differences in performance reflected the learning algorithm rather than changes in the underlying dataset.

To reduce information leakage between hydraulically related observations, cross-validation groups were generated from the exact (V_1, V_2, V_3)valve-opening triplets. The resulting 617 unique operating-condition groups were partitioned using five-fold grouped cross-validation. All observations belonging to the same valve-opening configuration were therefore retained within a single fold. For each cross-validation iteration, models were fitted using four folds and evaluated exclusively on the withheld valve-configuration groups. Out-of-fold predictions from the five validation folds were then combined to calculate the reported classification and regression metrics. This implementation ensured that performance was evaluated for valve-defined operating configurations that were not represented identically during model training.

For classification, out-of-fold predictions were evaluated using accuracy, balanced accuracy, macro-averaged F1 score, and the multiclass confusion matrix. For regression, predicted total water loss was compared with the experimentally measured leakage rate using R^2, mean absolute error (MAE), and root mean squared error (RMSE). Candidate algorithms were first compared using the complete predictor architecture. The selected Random Forest classifier and Extra Trees regressor were subsequently used for the sensor-architecture analyses.

Sensor-ablation analysis was performed by repeating the grouped-validation procedure using predefined subsets of the available predictors, including valve positions and pump current, flow measurements alone, pressure measurements alone, combined flow and pressure measurements, and the complete measurement architecture. This allowed the contribution of different sensing modalities to be evaluated under a common validation framework rather than by comparing models obtained from different data partitions.

For the selected tree-based models, predictor importance was extracted using the native impurity-based feature-importance measure implemented by the fitted ensemble estimators. Importance values were normalized and ranked separately for leak-state classification and total water-loss estimation. The computational workflow therefore proceeded from experimental-data filtering and predictor construction through grouped partitioning, candidate-model evaluation, model selection, sensor-ablation analysis, and final performance and feature-importance assessment, as summarized in Figure 2.

2.10 Grouped Cross-Validation

A central methodological consideration was preventing observations obtained at the same hydraulic operating point from appearing simultaneously in the training and validation subsets. Because multiple steady-state observations were recorded at identical valve configurations, conventional random observation-level partitioning could place hydraulically related observations in both subsets and thereby produce an optimistic estimate of model generalization.

Grouped five-fold cross-validation was therefore used, with validation groups defined by the exact combination of the three control-valve openings, (V_1, V_2, V_3). Each unique valve-opening triplet constituted one operating-condition group. The 2,447 steady-state observations comprised 617 unique valve-configuration groups. All observations associated with a given valve-opening triplet were assigned to the same fold, irrespective of leakage state, ensuring that an identical actuator operating point could not occur in both the training and validation subsets within a given cross-validation iteration.

The five folds contained 489–490 observations and 122–124 unique valve-configuration groups each, with all four diagnostic states (CONTROL, L1, L2, and L1 + L2) represented in every fold. For each iteration, the model was trained using four folds and evaluated on the withheld groups. Out-of-fold predictions from the five validation folds were subsequently combined to calculate the reported classification and regression performance metrics.

This validation strategy provides a more demanding assessment of whether the learned leakage signatures generalize across unseen valve-defined operating configurations, rather than merely reproducing patterns from repeated observations collected at operating points already represented during model training.

2.11 Classification Performance Metrics

Multiclass leak-localization performance was evaluated using accuracy, balanced accuracy, and macro-averaged F1 score.

Accuracy was calculated as

Accuracy = Ncorrect / Ntotal,

where Ncorrect is the number of correctly classified observations and Ntotal is the total number of evaluated observations.

Because the four diagnostic classes contained different numbers of observations, balanced accuracy was additionally calculated to reduce the influence of class-frequency differences. Macro-F1 was used to provide an overall measure of precision and recall in which each leakage class contributed equally to the final score.

The confusion matrix was also examined to determine which hydraulic states were most frequently confused, particularly whether simultaneous L1 + L2 leakage was misclassified as either individual leakage state.

2.12 Leakage-Quantification Performance

Regression performance for total water-loss estimation was evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE).

These metrics were calculated as

$$MAE = (1/n) \Sigma |y_i - \hat{y}_i|$$

and

$$RMSE = \sqrt{[(1/n) \Sigma (y_i - \hat{y}_i)^2]},$$

where y_i is the experimentally determined leakage rate and $\hat{y}_i$ is the corresponding out-of-sample model prediction.

R^2 was used to quantify the proportion of observed variability in leakage magnitude represented by the model, whereas MAE provided the average absolute prediction error directly in L min^{-1}. RMSE was additionally reported because it gives greater weight to larger prediction errors.

Predicted-versus-observed leakage plots were used to examine agreement over the full range of measured water-loss rates and to identify any systematic underprediction or overprediction at particular leakage magnitudes.

2.13 Pressure-Only Sensor-Reduction Analysis

To investigate the extent to which leakage localization could be achieved with reduced instrumentation, a separate model architecture was evaluated using only the four pressure measurements:

Xpressure = [P1, P2, P3, P4].

Flow measurements, pump current, and valve-position information were excluded from this reduced architecture.

The objective of this analysis was not to assume that pressure-only sensing would provide equivalent performance to the complete information architecture. Instead, it quantified the diagnostic information retained when the model was restricted to four spatial pressure measurements.

The pressure-only model was evaluated using the same grouped cross-validation principle applied to the complete architecture, allowing its performance to be compared on a consistent basis. The resulting reduction in classification performance was interpreted relative to the reduction in sensing and operating-state information.

2.14 Feature Importance

Feature importance was evaluated for the selected tree-based classification and regression models using their native impurity-based feature-importance scores. For each fitted model, the importance assigned to a predictor represents the normalized total reduction in the model-specific splitting criterion attributable to that predictor across the ensemble of decision trees. Predictor importance values were subsequently ranked from highest to lowest separately for leak-state classification and total water-loss estimation. These scores quantify the extent to which the fitted models relied on individual predictors and should not be interpreted as causal effects or as direct measures of the physical sensitivity of the hydraulic system.

2.15 Matched Operating-Condition Pressure Analysis

A separate matched-condition analysis was conducted to examine the physical pressure response associated with leakage while accounting for differences in control-valve configuration. For each leakage observation, the nearest available CONTROL observation was identified in three-dimensional valve-opening space defined by V_1, V_2, and V_3. The distance between leakage observation $L$and candidate CONTROL observation $C$was calculated using the Euclidean distance

$$d_{LC} = \sqrt{(V_{1,L} - V_{1,C})^2 + (V_{2,L} - V_{2,C})^2 + (V_{3,L} - V_{3,C})^2}$$

where $d_{LC}$is the Euclidean distance between leakage observation $L$and CONTROL observation C, and $V_{j,L}$and $V_{j,C}$are the corresponding opening percentages of control valve $j$under the leakage and CONTROL conditions, respectively.

The CONTROL observation having the minimum Euclidean distance was selected as the matched baseline for each leakage observation. Because valve openings were expressed as percentages on the same 0–100% scale, no additional scaling of the three valve-position dimensions was applied. For the primary matched-pressure analysis, only leakage observations with a nearest-neighbor distance of $d_{LC} \leq 20$percentage points were retained. This criterion retained 553 of 576 L1 observations (96.0%), 736 of 760 L2 observations (96.8%), and 718 of 738 simultaneous L1 + L2 observations (97.3%).

For each matched observation and pressure sensor i, the leakage-associated pressure change was calculated as

ΔPi = Pi,leak − Pi,control.

Positive values therefore indicate that the measured pressure during the leakage experiment exceeded the corresponding matched CONTROL value, whereas negative values indicate a pressure reduction relative to the matched baseline.

The resulting pressure-change vector was defined as

ΔP = [ΔP1, ΔP2, ΔP3, ΔP4].

Separate pressure fingerprints were obtained for L1, L2, and L1 + L2. For each pressure location and leakage state, the mean pressure change, standard deviation, standard error, 95% confidence interval, and median pressure change were calculated.

Absolute pressure differences in kPa were used as the primary basis for physical interpretation. Percentage pressure changes were considered supplementary because percentage differences can become unstable when the matched baseline pressure approaches zero.

This analysis was used to determine whether the three leakage states produced distinguishable spatial pressure responses and to provide a physical interpretation complementary to the ML classification results.

Because repeated observations were available for some valve configurations, conventional observation-level confidence intervals could underestimate uncertainty. Accordingly, 95% confidence intervals for the mean matched pressure changes were calculated using cluster-robust standard errors, with observations clustered according to the exact V_1-V_2-V_3 valve-opening triplet. This approach retained the observation-level mean pressure change while accounting for dependence among repeated measurements obtained under the same hydraulic operating configuration.

2.16 Overall Analysis Framework

The complete methodological framework consisted of four interconnected analyses. First, the hydraulic measurements were used to classify normal, single-leak, and simultaneous-leak states. Second, regression models were used to estimate the magnitude of total water loss. Third, the pressure-only architecture was evaluated to quantify the diagnostic capability retained under reduced sensing. Fourth, matched-condition pressure analysis was used to determine whether the predictive patterns were supported by physically interpretable spatial pressure responses.

The overall workflow can therefore be summarized as

Experimental hydraulic states

→ steady-state data extraction

→ grouped model validation

→ leak-state classification

→ leakage quantification

→ pressure-only sensor reduction

→ matched-condition hydraulic interpretation.

This framework was designed to evaluate both predictive performance and the physical basis of the leakage signatures while explicitly accounting for variability in normal hydraulic operating conditions.

3. Results and Discussion

3.1 Dataset Characteristics and Hydraulic Operating Range

The final dataset contained 2,447 steady-state observations, comprising 373 CONTROL observations, 576 observations with leakage at L1, 760 with leakage at L2, and 738 with simultaneous leakage at L1 and L2. Because the experimental program deliberately varied the three control-valve positions, observations within each leakage class represented a range of hydraulic operating conditions rather than repeated measurements at a single operating point.

Table 2, therefore, provides descriptive characterization of the experimental operating envelope. The reported standard deviations represent the dispersion of all steady-state observations within each experimental class and consequently include variability associated with changing valve configurations in addition to short-term experimental and measurement variability. They should not be interpreted as sensor uncertainty or as the uncertainty associated with the class mean.

Table 3
Descriptive hydraulic characteristics of the experimental dataset

State	n	F1 (L/min)	F2 (L/min)	F3 (L/min)	F4 (L/min)	P1 (kPa)	P2 (kPa)	P3 (kPa)	P4 (kPa)	Total loss (L/min)
CONTROL	373	15.973 ± 2.163	5.493 ± 2.536	4.835 ± 2.832	4.432 ± 2.955	4.200 ± 2.867	3.148 ± 3.456	4.442 ± 2.862	3.523 ± 0.736	0
L1	576	16.126 ± 1.241	5.232 ± 2.962	3.882 ± 2.807	1.633 ± 2.385	5.057 ± 2.678	4.734 ± 2.503	5.144 ± 2.786	3.424 ± 0.115	0.425 ± 0.541
L2	760	15.542 ± 1.532	4.634 ± 2.641	3.731 ± 3.113	2.647 ± 2.415	4.838 ± 2.714	4.246 ± 2.237	4.925 ± 2.812	3.374 ± 0.015	0.536 ± 0.637
L1 + L2	738	14.630 ± 1.385	3.693 ± 2.467	4.926 ± 2.346	3.186 ± 1.966	4.917 ± 2.629	4.414 ± 2.504	4.983 ± 2.651	3.357 ± 0.012	3.379 ± 1.885

Values are mean ± sample SD across all retained steady-state observations. SD describes the overall within-class hydraulic dispersion across the investigated operating configurations and is not a measure of sensor uncertainty.

Several broad differences are evident. Mean total water loss increased from zero under CONTROL conditions to 0.425 L/min for L1, 0.536 L/min for L2, and 3.379 L/min for simultaneous L1 + L2 leakage. The simultaneous-leak condition also exhibited substantially greater variability in loss magnitude, with an SD of 1.885 L/min.

The flow response depended strongly on leakage state. F4, for example, decreased from a mean of 4.432 L/min under CONTROL operation to 1.633 L/min under L1, whereas different redistribution occurred under L2 and L1 +

L2. These differences indicate that leakage altered the distribution of flow through the network rather than producing a uniform response across all branches.

The raw pressure means require more cautious interpretation. P1–P3 were, on average, higher in the leakage datasets than in the CONTROL dataset. These differences should not be interpreted directly as leak-induced pressure increases because the four datasets did not contain identical distributions of V1/V2/V3 operating configurations. The raw class means therefore combine the effects of leakage and operating state.

For this reason, no inference regarding the physical pressure effect of leakage was based solely on Table 2. Leakage-associated pressure changes were instead evaluated using the operating-condition-matched analysis described below.

This distinction materially strengthens the analysis: Table 3 describes the experimental domain; it does not establish the leakage fingerprint.

3.2 Operating-Condition-Matched Pressure Fingerprints

Because the CONTROL and leakage experiments were performed independently, exact V1/V2/V3 matches were not available for all leakage observations. Each leakage observation was therefore associated with the nearest CONTROL condition in three-dimensional valve-state space. The primary analysis retained observations for which the valve-state distance was no greater than 20 percentage points according to the predefined matching criterion.

This procedure retained 553 of 576 L1 observations (96.0%), 736 of 760 L2 observations (96.8%), and 718 of 738 L1 + L2 observations (97.3%). Thus, the matched analysis retained the great majority of the experimental leakage data.

For each pressure location,

$$\Delta P_i = P_{i,\mathrm{leak}} - P_{i,\mathrm{matched}\backslash\mathrm{hspace}\{0.17\mathrm{em}\}\mathrm{CONTROL}}.$$

Table 4 summarizes the resulting pressure differences.

Table 4
Matched leakage-associated pressure changes relative to CONTROL conditions. Values are mean ΔP (kPa) with 95% cluster-robust confidence intervals, with clustering defined by the exact V1–V2–V3 valve-opening configuration.

Leak case	Sensor	Mean ΔP (kPa)	Cluster-robust 95% CI
L1	P1	0.861	**[0.484, 1.239]**
	P2	0.823	**[0.462, 1.184]**
	P3	0.561	**[0.088, 1.034]**
	P4	0.004	**[-0.006, 0.015]**
L2	P1	0.485	**[0.165, 0.805]**
	P2	0.274	**[0.060, 0.489]**
	P3	0.181	**[-0.225, 0.587]**
	P4	-0.049	**[-0.057, -0.041]**
L1 + L2	P1	0.563	**[0.299, 0.828]**
	P2	0.392	**[0.173, 0.611]**
	P3	0.200	**[-0.140, 0.539]**
	P4	-0.065	**[-0.071, -0.059]**

This presentation is preferable to mean ± SD for the matched analysis because the confidence interval addresses uncertainty in the estimated mean pressure response, whereas the SD primarily describes the considerable observation-to-observation variability.

The three leakage states produced different spatial pressure patterns. L1 generated the largest mean responses at P1 and P2, with increases of 0.861 and 0.823 kPa, respectively, while P4 was essentially unchanged. In contrast, L2 produced smaller positive responses at P1–P3 and a mean decrease of 0.049 kPa at P4. Simultaneous L1 + L2 leakage produced positive mean responses at P1 and P2 but the largest reduction at P4, − 0.065 kPa.

The P4 response was particularly distinctive. Its 95% confidence interval included zero for L1 but was entirely negative for both L2 and L1 + L2. This spatial differentiation provides a plausible hydraulic basis for the subsequent finding that P4 was the most important individual predictor in the localization model.

The interpretation of P3 is more restrained. Although the mean ΔP3 was positive for all leakage states, the 95% confidence intervals for L2 and L1 + L2 crossed zero. The data therefore provide weaker evidence for a systematic P3 response to those two leakage conditions than for the P1, P2, and P4 responses.

3.3 Selection of the Machine-Learning Algorithms

Before comparing sensing architectures, candidate algorithms were evaluated using the complete candidate predictor set and grouped five-fold cross-validation. Results are shown in Table 4

Table 5
Algorithm comparison for four-class leakage localization

Model	Accuracy	Balanced accuracy	Macro-F1
Logistic Regression	0.879	0.878	0.876
Random Forest	0.974	0.975	0.973
Extra Trees	0.953	0.957	0.953

Random Forest provided the highest classification accuracy, 0.974, compared with 0.953 for Extra Trees and 0.879 for Logistic Regression. Balanced accuracy and macro-F1 followed the same pattern, indicating that the result was not driven solely by class imbalance.

The advantage of the ensemble models over Logistic Regression indicates substantial nonlinear structure in the relationship between hydraulic measurements and leakage state. Random Forest was therefore selected as the classification algorithm for subsequent sensing-architecture comparisons. For leakage quantification, Extra Trees provided the strongest regression performance, as shown in Table 5

Table 6
Algorithm comparison for total water-loss estimation

Model	R^2	MAE (L/min)	RMSE (L/min)
Ridge	0.530	0.931	1.224
Random Forest	0.949	0.178	0.403
Extra Trees	0.974	0.159	0.287

Extra Trees achieved R^2 = 0.974, MAE = 0.159 L/min, and RMSE = 0.287 L/min and was therefore retained as the regression algorithm for the subsequent sensor-ablation analysis.

3.4 Effect of Sensing Architecture on Leakage Localization

Having selected the model families, the available input information was systematically reduced to determine how localization performance depended on sensing architecture.

Table 7
Effect of input architecture on four-class leakage localization

Input architecture	Accuracy	Macro-F1
Valves + pump current	0.586	0.629
Flows only	0.796	0.802
P1–P4 only	0.859	0.853
Flow + P1–P4	0.962	0.957
All measurements + valves	0.972	0.971

The complete architecture achieved 0.972 accuracy and macro-F1 of 0.971. These values should be regarded as the principal full-architecture results when discussing sensor efficiency.

The difference between the algorithm-selection Random Forest accuracy of 0.974 and the full-architecture ablation accuracy of 0.972 is only 0.002, but the two values originate from distinct analytical comparisons. Accordingly, 0.974 is reported only as the best performance observed during algorithm selection, whereas 0.972 is used when comparing sensing architectures.

The pressure-only architecture is one of the most important results. Using only P1–P4 retained an accuracy of 0.859 and macro-F1 of 0.853 for the four-state problem. Pressure measurements alone therefore contained substantial spatial information concerning whether the system was operating under CONTROL, L1, L2, or L1 + L2 conditions.

Flows alone produced lower localization accuracy of 0.796. However, combining flow and pressure measurements increased accuracy to 0.962, only one percentage point below the complete architecture.

The valves-plus-pump-current architecture achieved only 0.586 accuracy. Actuator and pump-state information alone was therefore insufficient to reproduce the diagnostic information contained in direct hydraulic measurements.

Figure 3 demonstrates the effect of sensing architecture on four-class leak localization performance. The results show a clear increase in diagnostic capability as direct hydraulic measurements are incorporated into the feature set. Using only valve positions and pump current produced relatively limited performance, with an accuracy of 0.586 and macro-F1 of 0.629. Flow measurements alone improved the accuracy to 0.796, while the four pressure measurements (P1–P4) achieved a higher localization accuracy of 0.859 and macro-F1 of 0.853, indicating that spatial pressure information provides substantial discriminatory information for identifying leak location. Combining flow and pressure measurements increased the accuracy to 0.962 and macro-F1 to 0.957, approaching the performance of the complete measurement architecture (accuracy = 0.972; macro-F1 = 0.971). As shown in Fig. 4, the relatively small improvement obtained by additionally incorporating valve positions and pump current suggests that most of the information required for leak localization is already captured by the combined flow and pressure measurements. This finding is particularly relevant to sensor-efficient monitoring, as it indicates that high localization performance may be achieved without requiring the complete set of available operational measurements.

3.5 Localization of Simultaneous Leakage

The ability to distinguish L1 + L2 from individual leakage states was examined using the grouped out-of-fold predictions from the selected Random Forest model.

The confusion matrix showed that 719 of 738 simultaneous-leak observations were correctly classified, corresponding to approximately 97.4% correct recognition. Twelve L1 + L2 observations were classified as L2, seven as CONTROL, and none as L1.

An important consideration in interpreting the simultaneous-leak classification is that leakage magnitude was not identical across the three leakage configurations. The L1 + L2 experiments generally exhibited greater total water loss than the individual L1 and L2 experiments. Consequently, the high recognition rate for the simultaneous-leak class may reflect information associated with both leakage configuration and leakage severity. The present results therefore demonstrate reliable discrimination among the experimentally investigated leak states, but they should not be interpreted as establishing magnitude-independent localization of simultaneous leaks. Further validation

using overlapping or magnitude-matched leakage ranges would provide a more stringent assessment of configuration-specific discrimination.

As shown in Table 7, the four-state localization model demonstrated strong discrimination among the CONTROL, L1, L2, and simultaneous L1 + L2 conditions, with most observations concentrated along the diagonal of the confusion matrix. The model correctly classified 366 of 373 CONTROL observations (98.1%), 563 of 576 L1 observations (97.7%), 736 of 760 L2 observations (96.8%), and 719 of 738 simultaneous L1 + L2 observations (97.4%). Importantly, the high recognition rate for the simultaneous-leak condition demonstrates that the model was able to distinguish combined leakage from the corresponding individual leak states rather than simply identifying the presence of leakage. The strongest residual confusion occurred between L2 and L1 + L2, with 15 L2 observations classified as L1 + L2 and 12 L1 + L2 observations classified as L2. This pattern suggests that these two states occasionally produced similar hydraulic signatures under certain valve configurations. Nevertheless, the low overall number of off-diagonal classifications indicates that the measured hydraulic variables contained sufficiently distinct information to differentiate normal operation, individual leaks, and simultaneous leakage across the investigated operating conditions.

Table 8
Confusion matrix for four-state leakage localization.

Actual \ Predicted	CONTROL	L1	L2	L1 + L2
CONTROL	366	5	1	1
L1	7	563	4	2
L2	8	1	736	15
L1 + L2	7	0	12	719

The strongest residual confusion occurred between L2 and L1 + L2: 15 L2 observations were classified as combined leakage, while 12 combined observations were classified as L2. This pattern is consistent with the matched-pressure results, because both L2 and L1 + L2 produced negative P4 responses, whereas L1 produced almost no mean P4 change.

The classification errors therefore show a degree of hydraulic structure rather than appearing randomly distributed among classes.

3.6 Quantification of Total Water Loss

Sensor-ablation analysis was also performed for leakage magnitude using the selected Extra Trees regression framework.

The sensing architecture also had a pronounced effect on the estimation of total water loss. Using only valve positions and pump current provided limited predictive capability (R^2 = 0.256, MAE = 1.073 L min^{-1}, and RMSE = 1.540 L min^{-1}), indicating that actuator states and pump loading alone were insufficient to reconstruct leakage magnitude reliably. Pressure-only sensing improved the prediction substantially (R^2 = 0.679), but remained less effective than flow-based architectures. In contrast, flow measurements alone achieved R^2 = 0.899, with MAE = 0.340 L min^{-1} and RMSE = 0.568 L min^{-1}, demonstrating that flow variables contain strong information regarding the magnitude of water loss. Combining flow and pressure measurements further increased performance to R^2 = 0.966, MAE = 0.166 L min^{-1}, and RMSE = 0.328 L min^{-1}, approaching the performance of the complete sensing

architecture (R^2 = 0.973, MAE = 0.161 L min^{-1}, and RMSE = 0.291 L min^{-1}). These results complement the localization findings by revealing different diagnostic roles for the measured variables: pressure measurements were particularly informative for distinguishing the spatial location of leakage, whereas flow measurements contributed more strongly to estimating its magnitude. The relatively small improvement obtained when valve positions and pump current were added to the combined flow-pressure architecture further suggests that accurate water-loss quantification can be achieved with a reduced hydraulic sensor set, providing a practical basis for sensor-efficient monitoring systems.

Table 9
Effect of sensing architecture on total water-loss estimation.

Input architecture	R^2	MAE (L/min)	RMSE (L/min)
Valves + pump current	0.256	1.073	1.540
P1–P4 only	0.679	0.588	1.012
Flows only	0.899	0.340	0.568
Flow + P1–P4	0.966	0.166	0.328
All measurements + valves	0.973	0.161	0.291

This suggests a functional distinction between the two measurement types:

$$\mathrm{spatial\backslash:pressure\backslash:response} \rightarrow \ \mathrm{strong\backslash:localization\backslash:information},$$

whereas;

$$\mathrm{flow\backslash:response} \rightarrow \ \mathrm{strong\backslash:loss\text{-}magnitudeinformation}.$$

Combining the two measurement types was highly effective. Flow + P1–P4 achieved R^2 = 0.966, MAE = 0.166 L/min, and RMSE = 0.328 L/min, approaching the complete architecture without requiring valve positions or pump current as predictive inputs.

This is arguably more important for the practical sensor-efficiency argument than the marginal difference between R^2 = 0.966 and 0.973.

The observed-versus-predicted plot, as shown in Fig. 4, demonstrates a strong agreement between the measured and machine-learning-estimated total water loss across most of the investigated range. The concentration of observations along the expected one-to-one relationship indicates that the model captures both low and high leakage magnitudes with relatively small prediction errors, consistent with the strong overall regression performance (R^2 = 0.973, MAE = 0.161 L min^{-1}, and RMSE = 0.291 L min^{-1}). The relationship remains approximately linear over the full range of observed losses, indicating that model performance is not restricted to a narrow leakage regime. Some larger deviations are evident at isolated operating conditions, particularly at higher loss rates, where individual observations are noticeably under- or over-predicted. Nevertheless, these points are relatively sparse compared with the dominant trend. Overall, the figure provides direct visual evidence that the selected model can quantitatively reconstruct total water loss over a broad range of experimental hydraulic conditions, supporting its use not only for identifying leakage states but also for estimating the magnitude of water loss.

3.7 Feature Importance and Relationship to the Hydraulic Response

Feature-importance analysis provided further insight into the different information used for localization and quantification.

Table 10 reveals a clear distinction between the variables contributing to leak localization and those governing the estimation of total water loss. For localization, pressure measurements dominate the highest-ranked predictors, with P4 showing the largest importance (0.241), followed by P2 (0.098), while P1 and P3 also appear among the six most influential variables. This result supports the premise that the spatial distribution of pressure responses provides important information for distinguishing among CONTROL, L1, L2, and simultaneous L1 + L2 conditions. In contrast, loss estimation is dominated by flow measurements: F1 has the highest importance (0.341), followed by F4 (0.151) and F3 (0.119), with F2 also appearing among the six highest-ranked predictors. P4 remains influential for both tasks, ranking first for localization and fourth for loss estimation, suggesting that this measurement location captures information relevant to both leak state and magnitude. The different importance profiles are consistent with the sensor-ablation results, in which pressure-only measurements provided substantial localization capability, whereas flow measurements were considerably more effective for quantitative loss estimation. Taken together, these findings indicate that pressure and flow sensors perform complementary diagnostic functions: spatial pressure responses are particularly informative for identifying where leakage occurs, while flow measurements provide stronger information regarding how much water is being lost.

Table 10
Six highest-ranked predictors for localization and loss estimation

Rank	Localization	Importance	Loss estimation	Importance
1	P4	0.241	F1	0.341
2	P2	0.098	F4	0.151
3	F1	0.092	F3	0.119
4	P1	0.079	P4	0.109
5	P3	0.078	V2	0.055
6	V1	0.077	F2	0.052

P4 was the strongest individual localization predictor, with an importance of 0.241. This result is consistent with the matched-condition pressure analysis: This result is consistent with the matched-condition pressure analysis: P4 showed essentially no mean change for L1, whereas its 95% cluster-robust confidence intervals were entirely below zero for both L2 and L1 + L2.

In contrast, total-loss estimation was dominated by flow measurements. F1 accounted for an importance of 0.341, followed by F4 and F3. This agrees with the sensor-ablation results, where flow-only measurements retained much greater quantitative leakage information than pressure-only measurements.

The convergence of three independent analyses—matched pressure response, sensor ablation, and feature importance—provides stronger evidence than feature importance alone. Together, they indicate that spatial

pressure behavior is particularly informative for identifying leakage configuration, whereas the flow field contains stronger information about leakage magnitude.

Feature importance should nevertheless not be interpreted as a causal decomposition because the predictors are hydraulically coupled and potentially correlated.

3.8 Sensor-Efficiency Implications

The results show that the sensing requirements for localization and quantification are not identical. Four pressure measurements alone retained 85.9% four-class localization accuracy, whereas flow-only measurements achieved stronger leakage-magnitude prediction with $R^2 = 0.899$. Combining flow and pressure measurements produced 96.2% localization accuracy and $R^2 = 0.966$ for leakage quantification.

This finding suggests that the practical question is not simply how many sensors are required, but what information each sensor type contributes.

A monitoring architecture designed primarily for distributed leak screening and localization could rely more heavily on a limited set of spatially distributed pressure measurements. Where accurate water-loss quantification is also required, complementary flow measurements would provide substantial additional information.

Importantly, the addition of valve positions and pump current to the combined flow-pressure architecture yielded only modest improvement: localization increased from 0.962 to 0.972, while loss R^2 increased from 0.966 to 0.973. Thus, most of the diagnostic information in this experimental system was already contained in the hydraulic measurements.

This result strengthens the sensor-efficiency argument because high performance did not require every available operating variable.

4. Conclusions

This study demonstrated that machine-learning models can distinguish normal operation, individual leakage at L1 and L2, and simultaneous L1 + L2 leakage across variable hydraulic operating conditions in a laboratory-scale water distribution network. Under grouped five-fold cross-validation across valve-defined operating configurations, Random Forest achieved a four-class leak-state classification accuracy of 0.974, balanced accuracy of 0.975, and macro-F1 of 0.973. The simultaneous L1 + L2 condition was correctly identified in 719 of 738 observations (97.4%). However, because the combined-leak experiments generally exhibited greater total leakage than the individual-leak experiments, this result should be interpreted as discrimination among the experimentally investigated leakage states rather than as evidence of magnitude-independent simultaneous-leak localization.

Total water-loss estimation showed a different dependence on sensor type. During algorithm selection, the selected Extra Trees regression model achieved $R^2 = 0.974$, MAE = 0.159 L min^{-1}, and RMSE = 0.287 L min^{-1} using the complete predictor set. In the separate sensing-architecture analysis, flow measurements alone retained strong predictive capability ($R^2 = 0.899$), whereas pressure-only measurements achieved $R^2 = 0.679$.

The results also demonstrate that high localization performance does not necessarily require the complete measurement architecture. Pressure measurements alone achieved a localization accuracy of 0.859, while combining flow and pressure measurements increased the accuracy to 0.962 and macro-F1 to 0.957, approaching the performance obtained using all available measurements. The operating-condition-matched pressure analysis

provided complementary physical evidence for this behavior by revealing spatially differentiated pressure responses associated with L1, L2, and simultaneous L1 + L2 leakage. These findings support the use of distributed pressure information as an effective basis for sensor-efficient leak localization.

Overall, the study establishes an experimentally validated framework that integrates leak detection, localization, simultaneous-leak recognition, and quantitative water-loss estimation while explicitly considering variability in normal hydraulic operation. The results further indicate that a reduced architecture based on pressure and flow measurements may retain much of the diagnostic performance of the complete instrumentation system. The reported cross-validation performance therefore represents generalization across withheld valve-defined operating configurations within the present experimental system and should not be interpreted as direct evidence of transferability to different network geometries, sensor layouts, or field operating environments.

Future work should evaluate the framework on larger and more complex networks, additional and previously unseen leak locations, different leakage magnitudes, transient operating conditions, measurement noise, and changing demand patterns. Such validation will be necessary to determine the extent to which the sensor-efficient relationships identified in the present experimental network can be transferred to field-scale water distribution systems.

Declarations

Author Contribution

A.A. Wrote all the mansucript, gnerated the idea and executed the experimental work

Acknowledgement

I would like to acknoledge Jordan University of Science and Technology for thier support

References

1. Wu Y, Liu S, Kapelan Z (2024) Addressing data limitations in leakage detection of water distribution systems: Data creation, data requirement reduction, and knowledge transfer. Water Res 267:122471. https://doi.org/10.1016/j.watres.2024.122471
2. Sun C, Parellada B, Puig V, Cembrano G Leak localization in water distribution networks using pressure and data-driven classifier approach. Water 2020, 12, 54. https://doi.org/10.3390/w12010054
3. Tyagi V, Pandey P, Jain S, Ramachandran P A two-stage model for data-driven leakage detection and localization in water distribution networks. Water 2023, 15, 2710. https://doi.org/10.3390/w15152710
4. Mahdi NM, Jassim AH, Abulqasim SA, Basem A, Al-Haddad LA, Al-Ogaili AAF (2024) Leak detection and localization in water distribution systems using advanced feature analysis and an artificial neural network. Desalin Water Treat 320:100685. https://doi.org/10.1016/j.dwt.2024.100685
5. Ajoodani A, Nazif S, Ramazi P (1988) Towards a realistic data-driven leak localization in water distribution networks. Water 2025 17. https://doi.org/10.3390/w17131988
6. Santos-Ruiz I, López-Estrada F-R, Puig V, Valencia-Palomo G, Hernández H-R (2022) Pressure sensor placement for leak localization in water distribution networks using information theory. Sensors 22:443.

https://doi.org/10.3390/s22020443

7. Khorshidi MS, Nikoo MR, Taravatrooy N, Sadegh M, Al-Wardy M, Al-Rawas GA (2020) Pressure sensor placement in water distribution networks for leak detection using a hybrid information-entropy approach. Inf Sci 516:56–71. https://doi.org/10.1016/j.ins.2019.12.043
8. Li Z, Wang J, Yan H, Li S, Tao T, Xin K (2022) Fast detection and localization of multiple leaks in water distribution network jointly driven by simulation and machine learning. J Water Resour Plan Manag 148:04022050. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001574
9. Wang H (2026) AI for detecting and localizing concurrent abrupt and incipient leaks in water distribution networks. Water Res 295:125538. https://doi.org/10.1016/j.watres.2026.125538
10. Bandreddi R, Farmani R, Melville-Shreeve P (2026) Pressure gradient driven multistage leak localization for multiple simultaneous leaks in water distribution networks under uncertainty. J Water Resour Plan Manag 152. https://doi.org/10.1061/JWRMD5.WRENG-7561
11. Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer, New York. 10.1007/978-0-387-84858-7
12. Breiman L, Random Forests (2001) Mach Learn 45:5–32. 10.1023/A:1010933404324
13. Geurts P, Ernst D, Wehenkel L (2006) Extremely Randomized Trees Mach Learn 63:3–42. 10.1007/s10994-006-6226-1
14. Hoerl AE, Kennard RW (1970) Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 12:55–67. 10.1080/00401706.1970.10488634
15. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V et al (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830

Figures

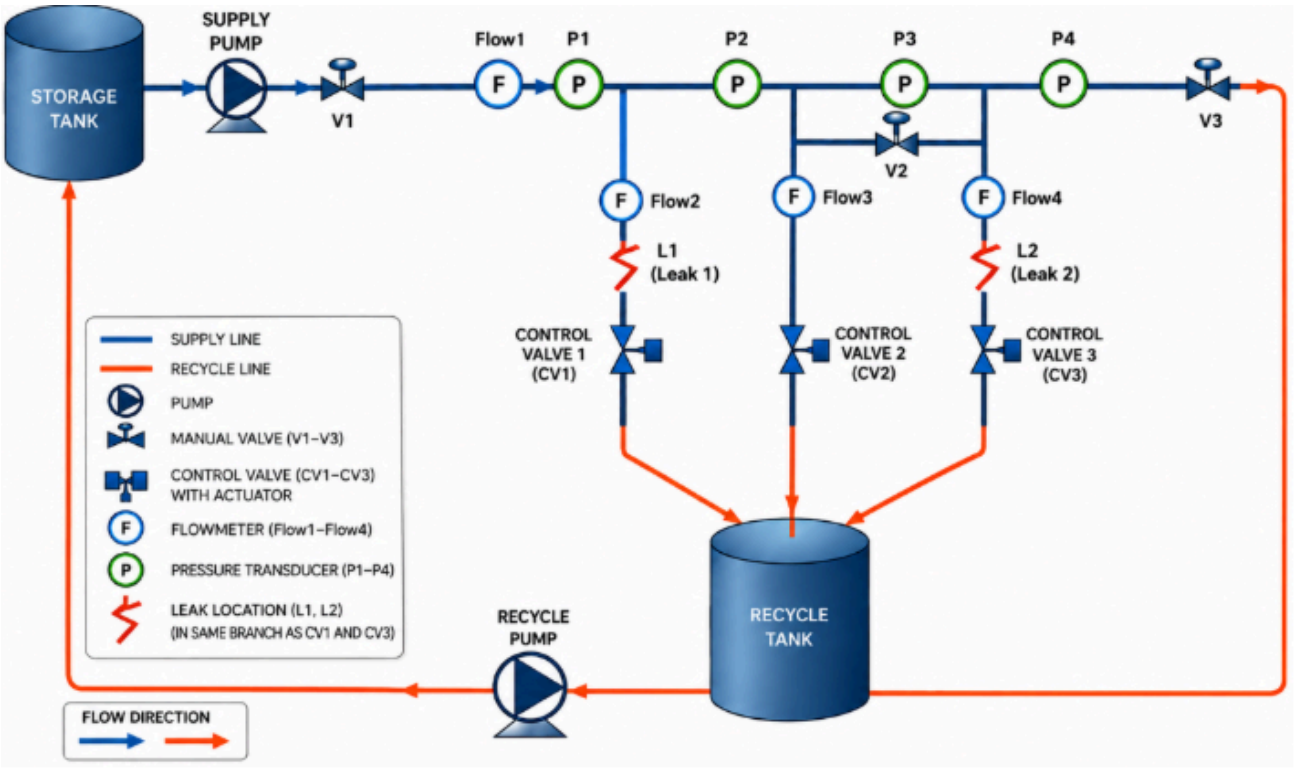


Figure 1

Schematic representation of the experimental hydraulic network showing the storage and recycle tanks, supply and recycle pumps, flowmeters (Flow1–Flow4), pressure transducers (P1–P4), manual valves (V1–V3), and actuated control valves (CV1–CV3).

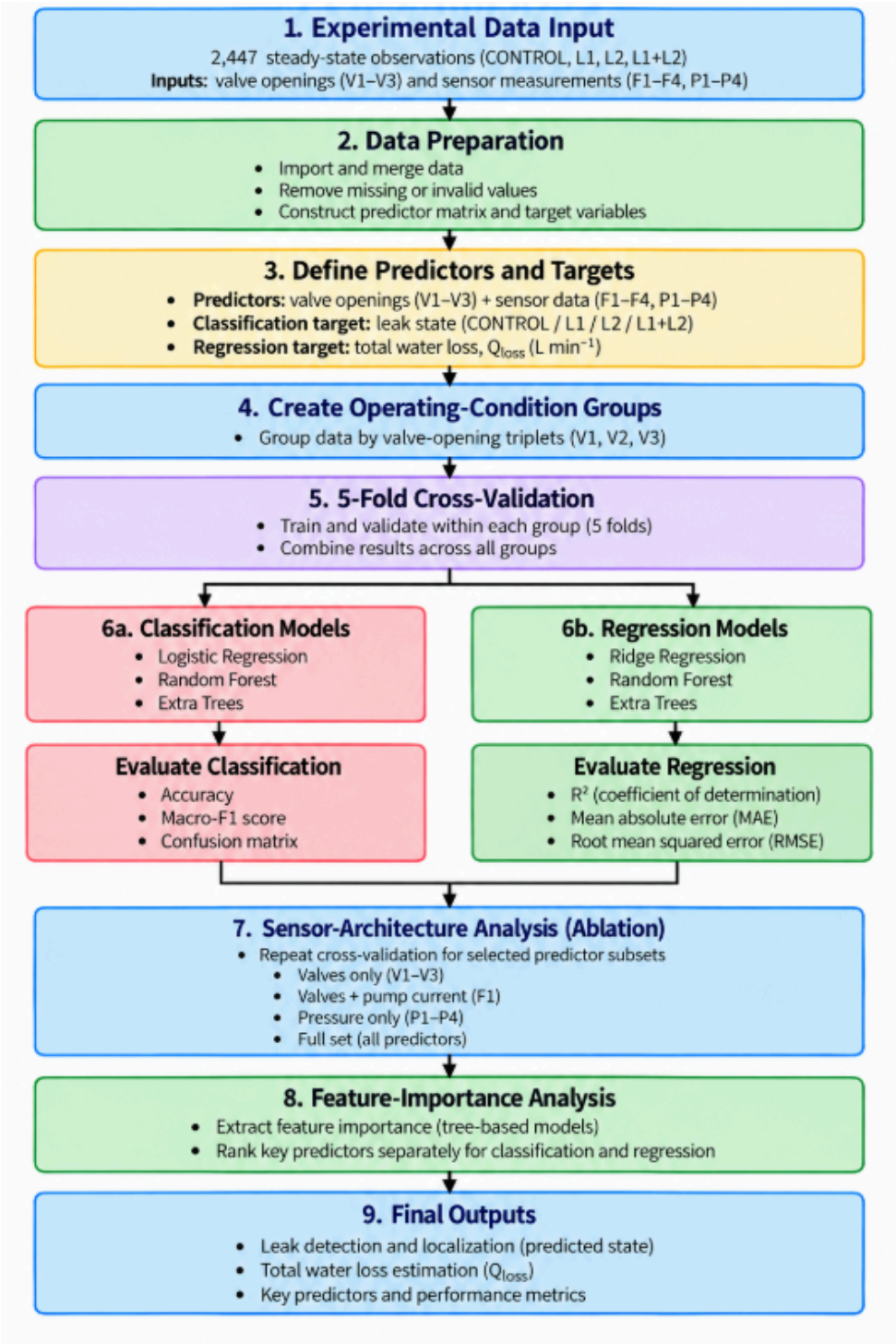


Figure 2

Computational workflow of the Python-based machine-learning framework for leak-state classification and total water-loss estimation, including grouped cross-validation, model selection, sensor-ablation analysis, and feature-

importance assessment.

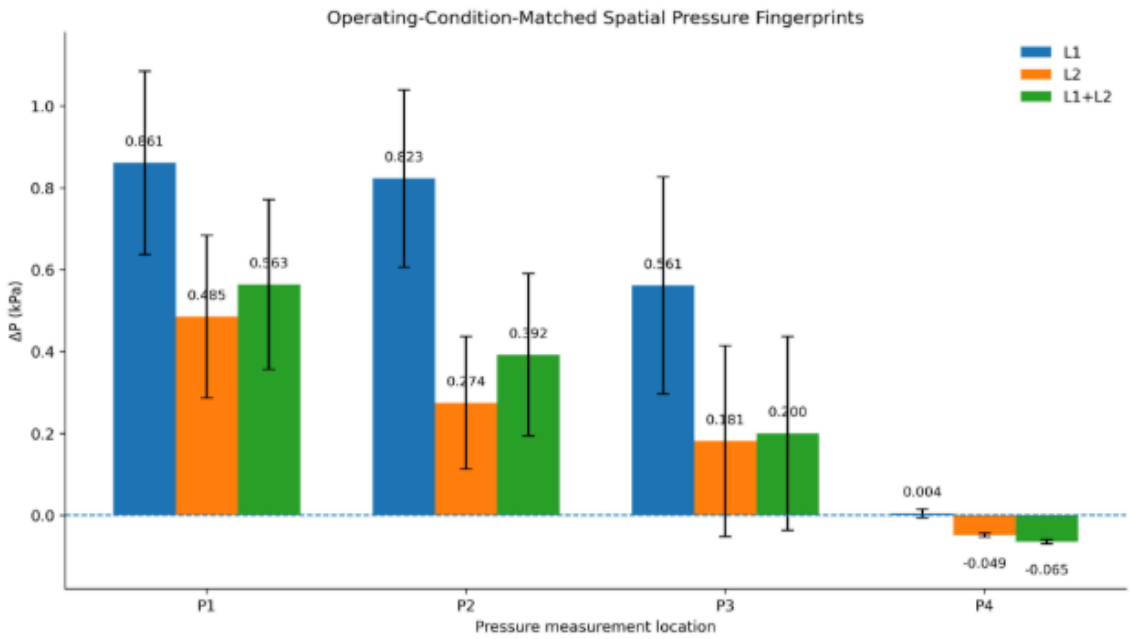


Figure 3

Operating-condition-matched spatial pressure fingerprints for L1, L2, and simultaneous L1+L2 leakage. Bars represent mean ΔP and error bars represent 95% cluster-robust confidence intervals, with clustering defined by the exact V1–V2–V3 valve-opening configuration.

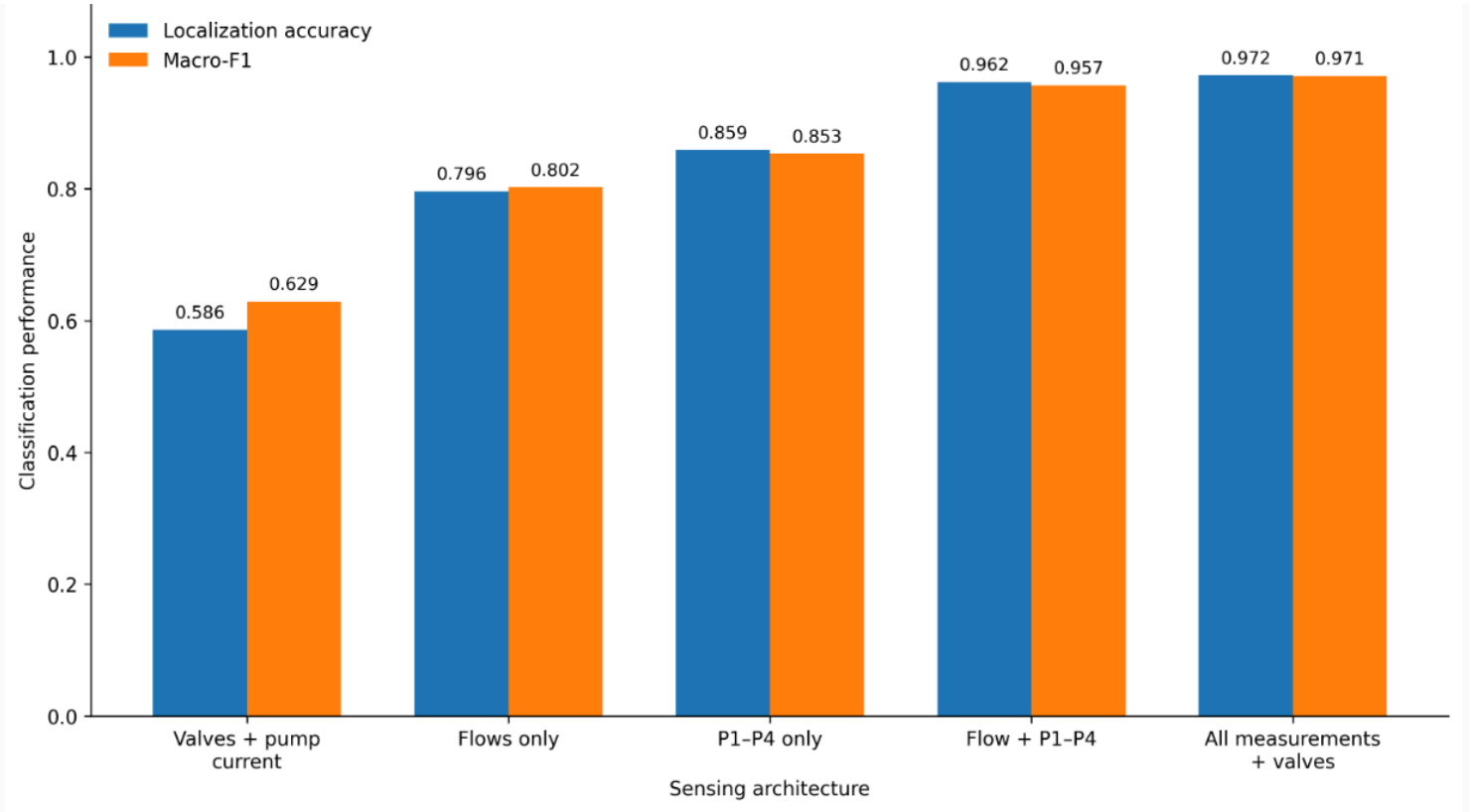


Figure 4

Effect of sensing architecture on four-class leakage localization performance under grouped five-fold cross-validation.

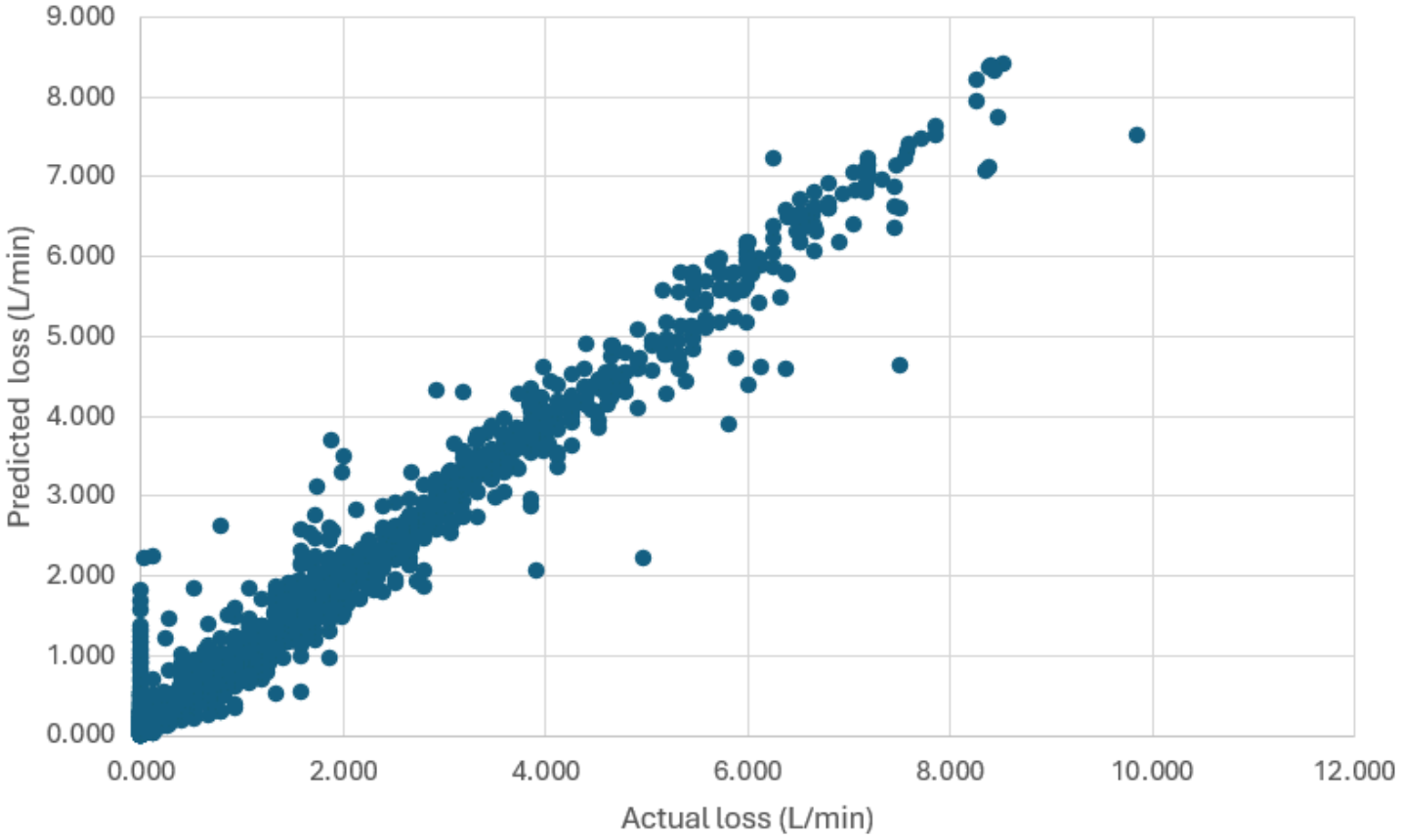


Figure 5

Observed versus predicted total water loss for the selected machine-learning regression model under grouped cross-validation.