MAYI-EDGE: An Offline-First Edge-AI and Digital-Twin Architecture for Intermittent Water Systems

Hervé Kinkete Mfumabi[1*]; Gracia Grace Disanka[1]; Daniel Kibala Mfumangimbi[1]; Evodie Ikene Mushid[1]; Kenayah Ebonge Maya[1]; Kéren Mayele Loba[1]; Jean-Beni Lokuli Isekongo[1]; Sephora Deko Kila[1]; Ferdinand Sangunza[1]

[1]Faculty of Computer Science and Artificial Intelligence, Université William Booth, Kinshasa, Democratic Republic of the Congo

Corresponding Author: Hervé Kinkete Mfumabi[*]

Publication Date: 2026/08/31

Abstract: Intermittent water supply creates nonstationary pressure, uncertain availability, leakage risk, and operational blind spots, while cloud-only monitoring is fragile where connectivity and power are unreliable. This article presents MAYI-EDGE as an architecture-led reference study with a reproducible algorithmic demonstration, an executable offline transaction contract, and a prespecified confirmatory protocol. The proposed offline-first cyber-physical architecture couples flow, pressure, and tank-level sensing with ESP32-S3-class edge processing, explicitly frozen quality control and inference, transactional local buffering, MQTT-TLS communication, server-side services, and a versioned EPANET/WNTR hydraulic-twin design. A deterministic synthetic fixture demonstrates the frozen detector logic under a declared surrogate-reference substitution, alert persistence within the prespecified 60-s window, backlog behavior, checksum-confirmed acknowledgement, and server de-duplication. It does not claim that an EPANET twin, complete physical prototype, MQTT deployment, or field system has already been executed. The defensible novelty is the joint treatment of intermittent supply, offline-first edge inference, transactional store-and-forward, a versioned hydraulic-twin design, resource-constrained deployment, and confirmatory evaluation prespecified before testing. Four statistical hypotheses (H1–H4) and a separate engineering validation gate (H5) define the future campaign. Parameters, temporal splits, baseline selection, ablations, calibration, missingness, network stress, energy profiling, human evaluation, multiplicity, and fixed sample sizes are specified in advance.

Keywords: *Digital Twin; Edge AI; EPANET; Intermittent Water Supply; Internet of Things; Leak Detection; Offline-First Systems; Reproducible Protocol.*

How to Cite: Hervé Kinkete Mfumabi; Gracia Grace Disanka; Daniel Kibala Mfumangimbi; Evodie Ikene Mushid; Kenayah Ebonge Maya; Kéren Mayele Loba; Jean-Beni Lokuli Isekongo; Sephora Deko Kila; Ferdinand Sangunza (2026) MAYI-EDGE: An Offline-First Edge-AI and Digital-Twin Architecture for Intermittent Water Systems. *International Journal of Innovative Science and Research Technology*, 11(8), 2135-2147. https://doi.org/10.38124/ijisrt/26aug874

I. INTRODUCTION

Intermittent water distribution is not merely continuous supply operated for fewer hours. Filling and emptying cycles, pressure transients, entrained air, variable demand, and uncontrolled storage produce a time-varying hydraulic regime in which conventional steady-state assumptions can fail. These dynamics increase uncertainty for users and complicate detection of leaks or abnormal consumption. In low-resource settings, the same networks often operate with sparse instrumentation, unstable connectivity, and limited technical capacity.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

Recent research has advanced three complementary directions. IoT platforms enable dense sensing but raise questions of communication reliability, energy use, calibration, and scale [1]. Machine-learning methods support demand forecasting and anomaly detection, yet their reported performance can be optimistic when random splits leak temporally adjacent observations across training and test sets [2]. Digital twins combine measurements with hydraulic simulation and "what-if" analysis, but require persistent synchronization and explicit validation against the physical system [3–5]. Edge–cloud frameworks reduce latency and bandwidth dependence, although most published work concerns water quality or continuously pressurized networks rather than intermittent supply [6].

The scientific gap is therefore not an absolute absence of IoT, machine learning, leak detection, or digital twins; recent systems already combine several of these elements. The narrower, defensible gap addressed here is the lack of a clearly specified combination centered on intermittent supply, offline-first edge inference, transactional store-and-forward recovery, a versioned hydraulic twin, resource-constrained deployment, and confirmatory evaluation prespecified before test-set access. MAYI-EDGE operationalizes that combination and defines how each claimed increment will be isolated through frozen baselines and ablations.

➢ *Research Questions and Testable Objectives*

Table 1 Research Questions and Testable Objectives

ID	Research question	Primary testable objective
RQ1	Can offline-first edge processing maintain monitoring and reduce anomaly-to-alert latency during degraded connectivity?	Compare edge-enabled versus cloud-only latency, continuity, backlog integrity, and recovery.
RQ2	Does hybrid physics–data inference improve leak detection and zone localization?	Compare MAYI-EDGE with fixed thresholds, CUSUM, Isolation Forest, and twin-residual baselines.
RQ3	Can water-availability windows be forecast without temporal leakage and with calibrated uncertainty?	Evaluate persistence, seasonal-naïve, logistic/gradient-boosted models using blocked temporal splits.
RQ4	Does the calibrated digital twin reproduce pressure and flow sufficiently for scenario analysis?	Assess held-out pressure/flow error, coverage, and parameter stability across operating regimes.

➢ *Contributions*

- A five-layer, offline-first reference architecture for intermittent water distribution, with explicit separation of sensing, edge, transport, twin, and decision functions.
- A confirmatory protocol that freezes outcomes, hypotheses, scenario labels, temporal splits, baselines, ablations, exclusion rules, and statistical tests before evaluation.
- A hybrid leak-detection design combining robust edge statistics, lightweight machine learning, and pressure/flow residuals from an EPANET-based twin.
- A reproducibility and governance package covering configurations, versioned models, calibration records, seeds, audit logs, privacy, and secure MQTT operation.

II. RELATED WORK AND SCIENTIFIC POSITIONING

Large-scale IoT reviews emphasize deployment density, link budget, energy, and simulation at network scale [1]. Machine-learning reviews document increasingly sophisticated data-driven and hybrid methods but also limited public field data and heterogeneous benchmarks [2,17]. Digital-twin work now includes probabilistic leak localization [18], real-time leakage detection and localization with connected hardware [19], multimodal prediction [20], and TwinAI graph-reinforcement-learning control [21]. Recent leakage research covers robust stochastic-process-control/CUSUM variants [22] and concurrent abrupt and incipient leaks using simulated SCADA and hybrid inference [23]. Intermittent-supply research has also progressed from hydraulic optimization [15] to mixed EPANET and machine-learning workflows [24]. These studies rule out any broad claim that hydraulic twins or hybrid AI are themselves new.

MAYI-EDGE therefore claims novelty only for the following conjunction: intermittent operating regimes; offline-first edge inference; transactional store-and-forward with sequence, acknowledgement, checksum, and de-duplication semantics; a versioned hydraulic twin; deployment constraints made explicit at the device level; and a confirmatory comparison prespecified before test-set access. No single row in the current literature comparison is asserted to be inferior overall; the table identifies which dimensions each work makes central.

Table 2 Related Work and Scientific Positioning

System / evidence	IoT sensing	Edge inference	Offline backlog	Hydraulic twin	Intermittent supply	Confirmatory baselines
Pagano et al. [1]	Yes	Discussed	Partial	Simulation tools	Not central	Survey
Probabilistic DT [18]	Pressure data	Cloud/deep model	No	Yes	No	Benchmark study

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

Travaš et al. [19]	Connected hardware	Controller-side	Not central	Real-time	No	Physical-loop tests
Syed et al. [20]	Multimodal	Transformer	No	Yes	No	Predictive comparison
TwinAI [21]	Network state	Graph RL	Not offline-first	Yes/versioned policy	No	Control baselines
Steins & Cominola [22]	Pressure	SPC/CUSUM	No	No	No	Seven SPC variants
Wang et al. [23]	Simulated SCADA	Hybrid AI	No	Hydraulic simulator	No	Abrupt + incipient leaks
Shah & Patel [24]	Water-quality sensors	ML	No	EPANET	Yes	Mixed tool/data method
MAYI-EDGE (this protocol)	Flow/pressure/level	Yes/frozen	Transactional + tested	Versioned EPANET/WNTR	Primary context	Validation-frozen rules, ML, cloud-only + ablations

III. MATERIALS AND METHODS

At manuscript submission, MAYI-EDGE is reported as a system conception, reference architecture, reproducible algorithmic demonstration, executable offline transaction contract, and prespecified experimental protocol. Architecture diagrams specify the intended full system; the supplied artifact executes only the frozen detector, surrogate-reference residual, trace export, and SQLite acknowledgement/de-duplication contract. It does not claim an assembled bill of materials, deployed MQTT/backend chain, executed EPANET/WNTR twin, dashboard, or measured field effectiveness. Numerical values that will arise from the confirmatory campaign are not predicted or fabricated. The protocol will be described as preregistered only after the protocol, power script, fixed sample sizes, analysis code, and repository snapshot receive a permanent timestamp before test-label access. Pilot observations used for debugging will be excluded from confirmatory inference.

➢ *MAYI-EDGE System Architecture*

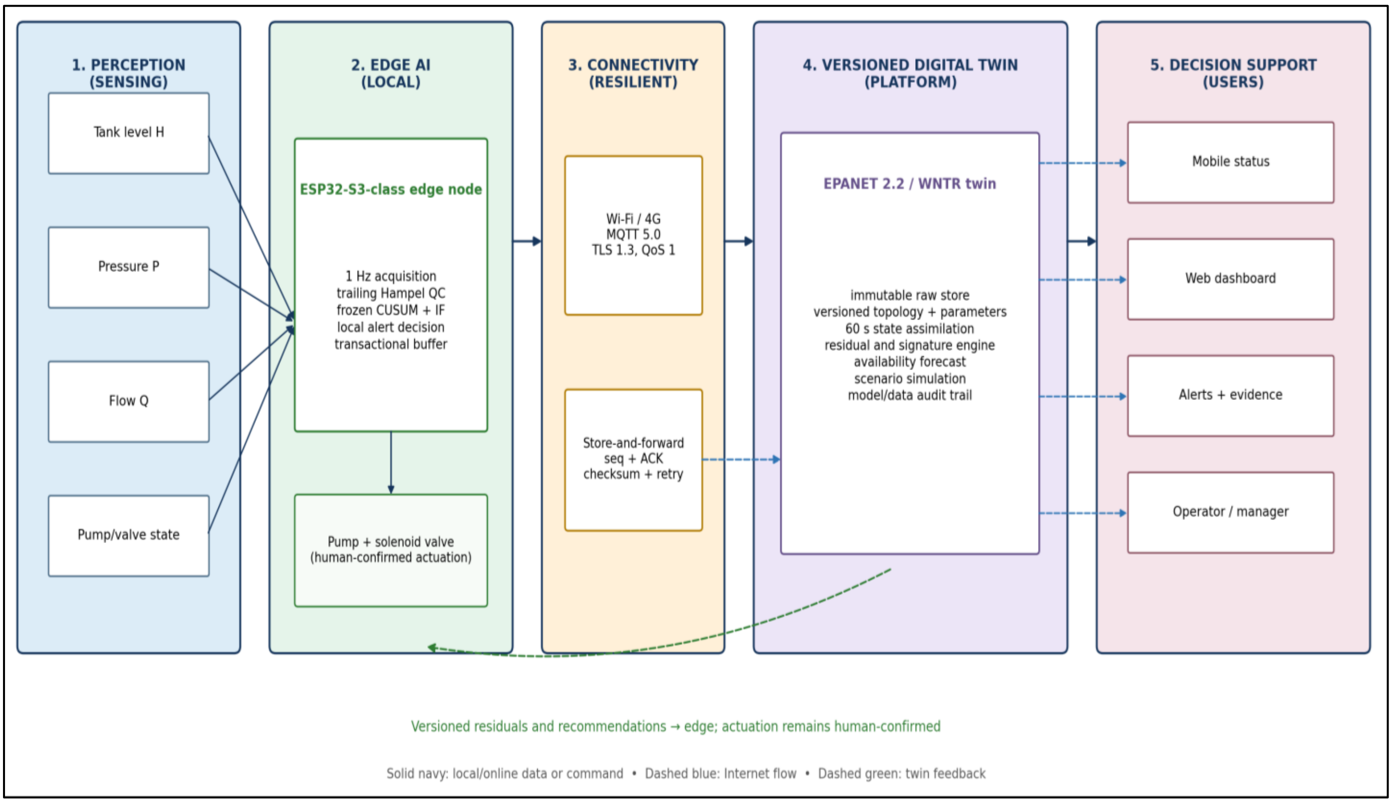

Fig 1 Improved MAYI-EDGE IoT–Edge AI–Digital-Twin Architecture Derived from the Authors' Original Five-Layer Concept. The Redesign Makes Local Actuation, Transactional Recovery, Twin Versioning, Evidence Delivery, and Human-Confirmed Control Explicit. Author-Generated Figure.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

➢ *Reference Hardware and Measurement Chain*

Table 3 Reference Hardware and Measurement Chain

Component	Reference specification	Measurement role	Verification / substitution rule
Edge controller	ESP32-S3, dual-core 240 MHz, ≥8 MB PSRAM, microSD, RTC	Acquisition, QC, inference, encrypted buffer	Resource profile logged; equivalent board permitted only via amendment
Flow sensing	Hall-effect inline meter, nominal 1–30 L/min; pulse output	Pipe flow Q (L/min)	Gravimetric calibration against timed mass/volume
Pressure sensing	0–1.2 MPa transducer, 0.5–4.5 V, external 16-bit ADC	Node pressure P (kPa)	Reference gauge at 20/50/80% span
Tank level	Waterproof ultrasonic or hydrostatic probe	Level H (cm; converted to volume)	Manual ruler/reference points; temperature compensation
Actuation state	Dry-contact/isolated inputs for pump and valve	Context and ground-truth transitions	Timestamped manual log
Power	12 V supply + DC backup; inline power monitor	Energy, outage autonomy	Cross-check with bench multimeter

This table is a reference hardware specification, not evidence that a complete physical prototype has already been assembled. Before confirmatory collection, the actually procured manufacturer, part number, serial number, range, accuracy class, supply voltage, firmware build, and calibration coefficients must be frozen in a machine-readable bill of materials. Any later substitution is a protocol deviation and triggers device-stratified sensitivity analysis.

➢ *Firmware Contract, Data Schema, and Quality Control*

The executable demonstration implements the same acquisition contract intended for hardware deployment: raw channels are sampled or replayed at 1 Hz, checked for range and rate of change, and aggregated into 10-s features (median, interquartile range, slope, minimum, maximum, pulse count, state transitions, and missingness). Every record carries UTC timestamp, monotonic sequence number, device ID, executable/firmware hash, calibration version, sensor-quality flags, model version, battery voltage, and connectivity status. Local time is display-only. Clock drift is estimated at reconnection and retained as metadata.

A microSD-backed SQLite ring buffer stores at least 14 days of 10-s records. Writes use transactions and checksums. When connectivity returns, unsent records are published oldest-first; the server de-duplicates on device_id + sequence_number. No record is deleted until broker acknowledgement and server checksum confirmation. Buffer occupancy, rejected rows, retry count, and clock corrections are themselves telemetry.

➢ *MQTT, Storage, and Security*

MQTT 5.0 over TLS 1.3 is planned with per-device credentials, QoS 1 for telemetry and alerts, retained device-status messages, bounded payloads, and topic-level authorization. The topic convention is mayi/ {site}/ {device}/ {stream}/ {schema_version}. Mosquitto acts as broker; a FastAPI ingestion service validates JSON Schema and writes immutable raw events to PostgreSQL/Timescale-compatible tables. Processed features and twin states are separate, versioned derivatives. Secrets are provisioned outside firmware source; key rotation and revocation are tested.

➢ *Frozen Edge Algorithms*

All confirmatory detector parameters are fixed here. Each continuous channel is processed in non-overlapping 10-s windows. A causal trailing Hampel screen uses the preceding 31 windows (310 s), median center, scale 1.4826×MAD, and rejection coefficient 3.0. An isolated candidate is provisionally replaced by the preceding median and flagged. At the third consecutive candidate in the same direction, all three raw values are restored, their spike flags are cleared, and the persistent regime shift is passed to CUSUM; this prevents a true sustained leak from being suppressed for an entire Hampel window. Features are standardized with training-normal median and 1.4826×MAD. For pressure and flow residuals, two-sided CUSUM states are $C^+_t=\max(0,C^+_{t-1}+z_t-k)$ and $C^-_t=\max(0,C^-_{t-1}-z_t-k)$, with $k=0.5$ and $h=5.0$. Their normalized channel scores are $cP=\min[1,\max(C^+P,C^-P)/h]$ and $cQ=\min[1,\max(C^+Q,C^-Q)/h]$.

The Isolation Forest is trained only on normal training windows with n_estimators=200, max_samples=256, max_features=1.0, bootstrap=False, contamination="auto", and random_state=42. To make score direction unambiguous, $s_t=-\text{decision\_function}(x_t)$, so larger values are more anomalous. Let q50 and q99 be the training-normal quantiles of s. The normalized score is $aIF=\text{clip}[(s-q50)/\max(q99-q50,10^{-8}),0,1]$. This epsilon rule prevents division by zero and is frozen. The twin-residual score is $rTwin=\text{clip}[0.5|eP|/10\text{ kPa} + 0.5|eQ|/1.0\text{ L·min}^{-1},0,1]$. The exact combined anomaly score is $A=0.40cP+0.25cQ+0.20aIF+0.15rTwin$. The primary alert threshold is $A\geq0.65$ for three consecutive 10-s windows. A separate safety alert is emitted when measured service pressure is <100 kPa for two consecutive windows; it does not count as a model detection in H2.

Leak localization is zone-level over exactly three candidates. For each event, the evaluation window is the first 12 complete 10-s windows after event onset (0–120 s). For candidate zone z and channel $X\in\{P,Q,H\}$, $\text{nRMSE}(X,z)=\text{sqrt}[(1/N_X)\Sigma_{t,s}(X_{obs,s,t}-X_{twin,s,t}(z))^2]/S_X$, where s ranges over all quality-valid sensors of that channel, N_X is the number of included sensor-window pairs, and S_X is the training-normal

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

interquartile range frozen separately for P, Q, and H. A channel with $<80\%$ valid pairs makes that event non-localizable but does not remove it from intention-to-monitor reporting. The distance is $D(z)=0.60\cdot nRMSE(P,z)+0.25\cdot nRMSE(Q,z)+0.15\cdot nRMSE(H,z)$. The predicted zone minimizes $D(z)$. Every zone satisfying $D(z)\leq D_{min}+0.02$ is reported as tied; top-1 evaluation counts a tie as correct only when the registered deterministic tie-breaker (lowest zone identifier) equals the true zone, while top-k evaluation reports the full tied set. No parameter is retuned after test access.

➢ *Availability Forecasting*

Service availability is defined numerically as pressure $P\geq 100$ kPa at the monitored service node for 30 consecutive 10-s windows (five minutes). Two targets are prespecified: (i) probability of an availability onset within the next 60 min and (ii) time-to-onset in minutes. Inputs are lagged pressure, flow, level, pump/valve state, hour-of-day and day-of-week harmonics, previous outage duration, and missingness indicators. Candidate models are seasonal persistence, regularized logistic/linear models, and gradient-boosted trees. The simplest model within one standard error of the best validation score is selected using validation data only. Probability calibration uses validation-only isotonic regression; test calibration is assessed once by Brier score and reliability bins.

➢ *Versioned Hydraulic Digital Twin*

The twin uses EPANET 2.2 pressure-dependent demand and WNTR-compatible orchestration [10–12]. The physical topology, pipe lengths/diameters, elevations, tanks, pumps, and valves are versioned. Initial roughness and demand multipliers are calibrated on the training period; validation selects regularization and update frequency. Data assimilation updates tank level and boundary conditions every 60 s when quality flags pass. The test period is held out for pressure/flow fidelity and leak-signature evaluation. A twin version cannot be overwritten; each simulation stores input snapshot, engine version, parameters, random seed where applicable, and output checksum.

➢ *Reference Physical Demonstrator Design*

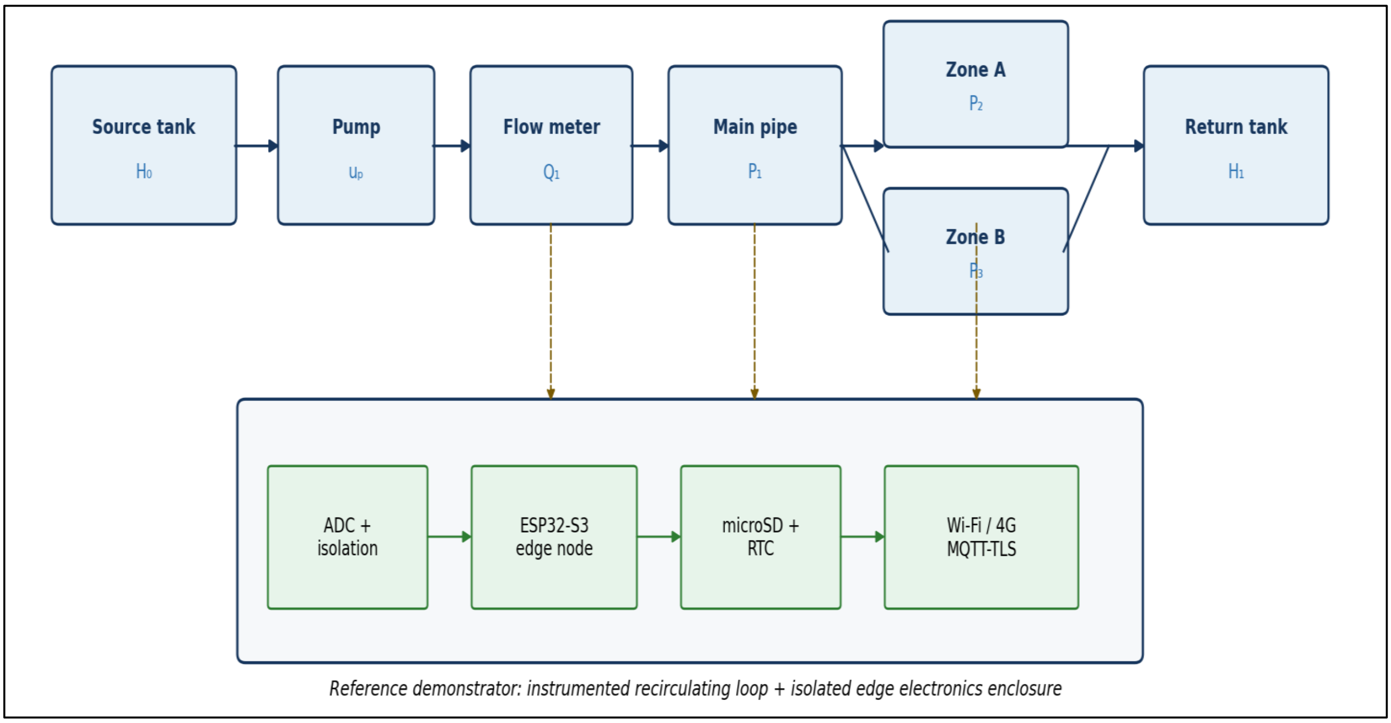

Fig 2 Reference Design for a Future Instrumented Recirculating Hydraulic Demonstrator and Isolated Edge-Electronics Enclosure. The Diagram is a Construction Specification, not a Photograph or Claim of Completed Assembly. Author-Generated Figure.

The reference physical architecture is a closed recirculating loop designed to reproduce supply start-up, normal demand changes, pressure decay, controlled zone leaks, sensor dropout, and communication interruption without wasting large volumes of water. A source tank feeds a pump and inline flow meter; pressure is observed upstream and in at least two distinguishable zones; a return tank closes the loop. Calibrated bypass/orifice branches provide repeatable leaks. The specified enclosure contains signal conditioning, an external ADC, an ESP32-S3-class controller, RTC, microSD storage, power monitoring, and Wi-Fi/4G communication. Physical emergency controls remain independent from software. Assembly evidence—BOM, dated scientific photographs, firmware hash, calibration sheets, logs, and real tests—must be added before the manuscript makes a physical-implementation claim.

Table 4 Reference Physical Demonstrator Design

Prototype interface	Implementation contract	Evidence retained
Sensor → edge	Timestamped 1-Hz acquisition, range check, calibration transform	Raw value, engineering unit, quality flag, calibration version
Edge → local decision	10-s robust features, CUSUM/IF score, persistence rule	Feature vector, model hash, threshold, decision time
Edge → broker	MQTT 5.0 QoS 1 over TLS; schema-versioned JSON	Sequence number, publish/ack time, retry count
Broker → twin	Validated telemetry updates boundary/state variables	Input snapshot, twin version, residuals, solver status
Twin → user	Zone hypothesis, forecast window, uncertainty and evidence	Alert payload, acknowledgement, operator action, audit record

➢ *Executable Reference Software Architecture*

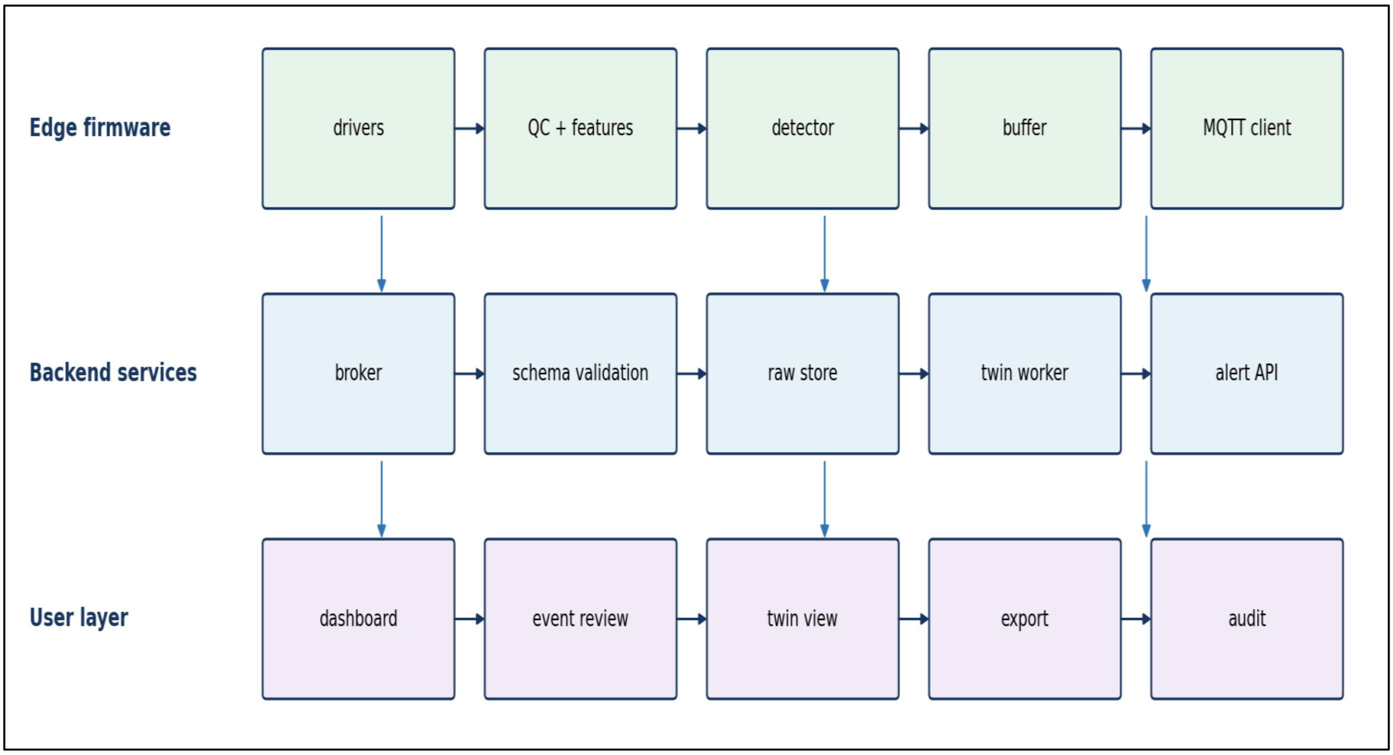

Fig 3 Executable Reference-Software Modules from Edge Emulation/Firmware Contract to Backend Services and User-Facing Decision Support. Author-Generated Figure.

The executable reference software follows replaceable modules with explicit contracts rather than a monolithic application. The edge layer separates sensor/replay drivers, quality control, feature extraction, anomaly inference, transactional buffering, and communication. Backend services separate broker reception, schema validation, immutable raw storage, feature/twin workers, alerts, and API delivery. The interface exposes current hydraulic state, time-series evidence, suspected zone, uncertainty, connectivity status, backlog state, and model/twin version. Every automated output remains traceable to the fixture and configuration that produced it.

The target server deployment is specified as containerized and version-pinned; it is an architectural requirement, not a claim that every backend service accompanies this review artifact. The supplied artifact executes the frozen detector plus a SQLite transaction/acknowledgement/de-duplication contract. Five unit tests verify the 60-s detection window, three-window persistence, anomaly-score direction, transactional acknowledgement/de-duplication, and required telemetry fields. MQTT broker integration, EPANET execution, dashboard delivery, and hardware-in-the-loop tests remain explicit future verification stages.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

IV. REPRODUCIBLE DEMONSTRATION RESULTS

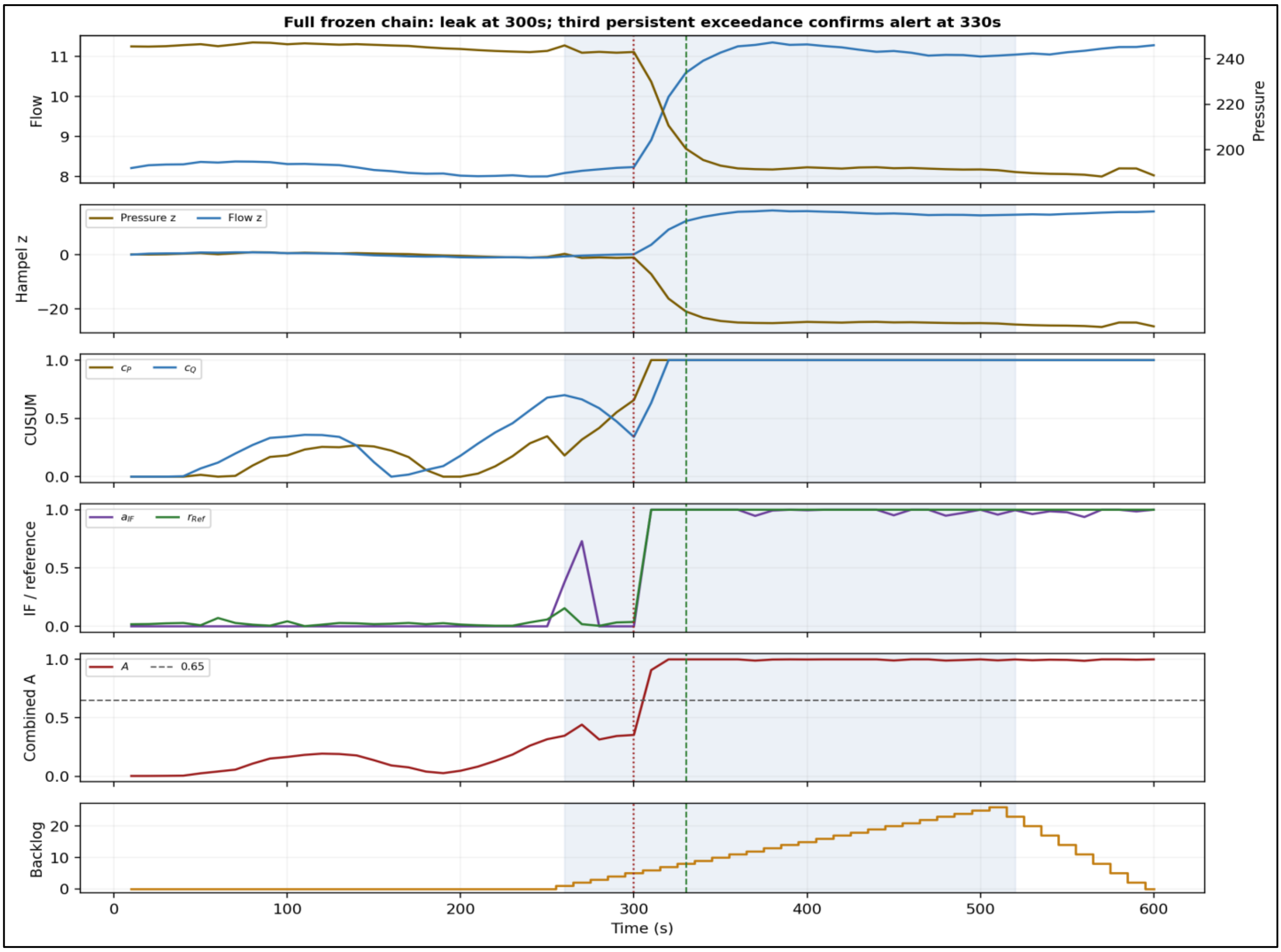

Fig 4 Full Deterministic Detector Chain Using a Declared Surrogate no-Leak Hydraulic Reference: Hampel-Standardized Signals, Pressure/flow CUSUM, Isolation Forest Anomaly Score, Surrogate-Reference Residual, Frozen Combined Score A, three-Window Persistence, Imposed Network Outage, and Backlog Trace. Seed = 42. Algorithmic Software-Execution Evidence, not EPANET, Prototype, or Field-Performance Evidence.

The reproducible demonstration executes one canonical scenario directly from the accompanying Figure 4 script. A zone leak is introduced at t=300 s and the third consecutive window with A≥0.65 confirms the alert at t=330 s, satisfying the prespecified 60-s event window. The plot exposes every frozen component—Hampel-standardized pressure/flow, cP, cQ, aIF, the declared surrogate-reference residual rRef, and A. The shaded interval is an imposed external-link outage: local inference continues and the backlog grows; a separate executable SQLite test verifies transaction commit, checksum-confirmed acknowledgement, ordered pending retrieval, and receiver de-duplication. This is intentionally an algorithmic and offline-contract demonstration. MQTT transmission, an executed EPANET/WNTR twin, dashboard integration, physical accuracy, energy autonomy, and field benefit remain confirmatory stages.

Table 5 Reproducible Demonstration Results

Demonstration checkpoint	Pass condition	Scientific interpretation
Acquisition	No timestamp reversal; sequence numbers monotonic	Sensor-to-edge chain is operational
Local inference	Alert is generated from frozen fixture and configuration	Algorithm executes reproducibly; accuracy is not yet inferred
Offline contract	SQLite commit, ordered pending retrieval and checksum-confirmed acknowledgement	Transactional mechanism executes in the supplied test
De-duplication	Repeated device_id + sequence_number is rejected	Receiver idempotency executes in the supplied test

Reference residual	Surrogate rRef values and combined A are exported	Detector fusion is reproducible; twin fidelity is not inferred
Artifact export	Figure, CSV trace, config and power result are generated	Reported demonstration can be independently inspected

This demonstration supports an executable algorithmic and offline-contract claim, not a complete-system, completed-hardware, EPANET-execution, or generalized-effectiveness claim. Detection accuracy, localization accuracy, energy autonomy, hydraulic fidelity, dashboard usability, and field benefit remain confirmatory outcomes governed by Sections 5–9.

➢ *Prespecified Experimental Protocol*

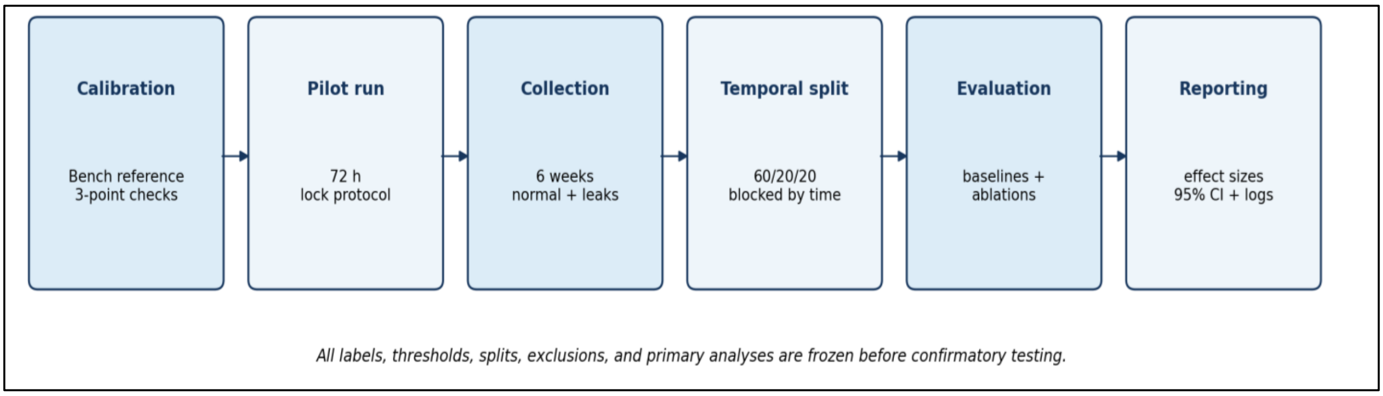

Fig 5 Prespecified Experimental Workflow and Data Lock. Author-Generated Figure.

- *Testbed, Duration, Sampling, and Observations*

The controlled testbed will comprise one source tank, one pump, exactly three hydraulically distinguishable zones, isolation valves, and instrumented upstream/downstream points. A separate 72-h pilot is used only for integration debugging; no pilot observation enters confirmatory analysis. The campaign continues until the fixed independent-unit counts below are completed, with a maximum calendar duration of eight weeks. The unit of inference is the independently initiated event, supply cycle, or impairment run—not the individual 1-Hz row.

Table 6 Testbed, Duration, Sampling, and Observations

Scenario family	Fixed levels	Fixed independent N	Ground truth
Normal intermittent cycles	Low/medium/high demand; three frozen supply schedules	120 cycles	Valve/pump log + reference gauges
Controlled leaks	5%, 10%, 20%, 30% of baseline flow × three zones	360 events: 30 per magnitude-zone cell	Calibrated leak orifice/bypass + timed volume
Non-leak disturbances	Demand step, pump start/stop, valve movement, sensor dropout	160 events: 40 per disturbance type	Operator script + video/time log
Connectivity impairment	Eight frozen profiles spanning outage, loss, latency, and bandwidth	128 paired runs: 16 per profile	Network impairment log
Forecast evaluation	All three supply schedules and demand levels	120 held-out supply cycles	Observed threshold-crossing time
Power limitation	Backup transition and controlled discharge	24 cycles	Power analyzer + event log

- *Calibration and Quality Assurance*

✓ Flow: at least five operating points across the planned range, three repeats each, using timed gravimetric or volumetric reference. Fit pulses-per-liter and report residuals, repeatability, hysteresis, and 95% uncertainty.
✓ Pressure: three-point ascending and descending checks at 20%, 50%, and 80% of span against a traceable reference gauge. Record zero offset, slope, hysteresis, and temperature.
✓ Level: empty/25/50/75/100% reference points with geometry-based conversion to volume; repeat after repositioning.
✓ Cross-sensor synchronization: common step event at all channels; permissible median offset ≤1 s after clock correction.
✓ Daily zero/range plausibility and weekly verification. Failure beyond acceptance limits quarantines subsequent records until recalibration.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

- *Missing Data and Exclusions*

Raw values are never imputed. For feature generation, gaps ≤2 s may be linearly interpolated only for continuous channels and are flagged; longer gaps produce missing features plus indicators. A 10-s window is excluded from the primary sensor-performance analysis when >20% of required samples are absent, during calibration, or when the reference instrument is invalid. Events remain in the intention-to-monitor analysis even if connectivity fails. All exclusions are enumerated in a CONSORT-like flow table, and sensitivity analyses use thresholds of 10% and 30%.

- *Temporal Partition and Anti-Contamination Controls*

Data are split chronologically by contiguous blocks: first 60% training, next 20% validation, final 20% locked test. Scenario repetitions are scheduled so every required class and leak zone occurs in each block, but no event crosses a boundary. Scaling, feature selection, calibration, threshold selection, and twin parameter estimation are fit on training/validation only. No random row-level cross-validation is allowed. Nested blocked resampling within training estimates model variability. Duplicate timestamps, near-identical windows from the same event, and future-derived variables are automatically checked. Test labels remain blinded until the prediction file and software hash are frozen.

- *Baselines and Ablation Study*

Table 7 Baselines and Ablation Study

Task	Baselines	Full MAYI-EDGE	Ablations
Leak detection	Fixed threshold; pressure-only CUSUM; Isolation Forest; twin residual only	CUSUM + IF + twin context	No pressure; no flow; no twin; no persistence; cloud-only
Zone localization	Largest pressure drop; twin signature only	Weighted sensor + twin signature	Single sensor; no state inputs
Availability onset	Last-observation/persistence; seasonal naïve; regularized model	Selected calibrated model	No level; no schedule features; no missingness flags
Continuity/latency	Cloud-only MQTT pipeline	Edge decision + store-and-forward	No local buffer; QoS 0; no batching

For H2, each baseline is fitted on training data and ranked on validation data by macro F1 computed across independent events, with false alerts per 24 h as the first tie-breaker and median detection delay as the second. The single highest-ranked baseline is then named in a signed/frozen analysis manifest before test labels are opened. H2 compares MAYI-EDGE only with that frozen validation-selected baseline. Test performance is never used to choose or reorder baselines.

- *Confirmatory Hypotheses and Decision Rules*

Table 8 Confirmatory Hypotheses and Decision Rules

Hypothesis	Directional claim	Primary outcome	Decision criterion
H1	Edge mode reduces anomaly-to-alert latency versus cloud-only under impaired connectivity.	Paired capped latency L* (s)	One-sided paired randomization test; Holm-adjusted p<.05; ΔL* and 95% CI
H2	Hybrid detection improves macro F1 over the validation-selected, frozen baseline.	Macro F1 across 520 independent events	One-sided paired label-randomization test; Holm-adjusted p<.05; ΔF1 and 95% CI
H3	Hybrid localization improves correct-zone accuracy over the frozen localization baseline.	Paired top-1 correctness	One-sided McNemar exact test; Holm-adjusted p<.05; Δaccuracy and 95% CI
H4	Forecasting reduces onset-time absolute error versus seasonal persistence.	Paired absolute-error difference	One-sided paired randomization test; Holm-adjusted p<.05; ΔMAE and 95% CI
H5 gate	The calibrated twin meets engineering fidelity bounds on held-out data.	Pressure nRMSE; floor-protected flow MAPE	Pass in every supply regime only if pressure nRMSE ≤10%, flow MAPE ≤15%, and valid coverage ≥90%

H1–H4 form the sole confirmatory statistical family; their four raw p-values are adjusted by Holm’s sequential procedure at family-wise α=.05. For H1, T_max=600 s is frozen: L*=min(T_alert−T_onset,600), and a mode producing no alert by 600 s receives L*=600 s. No-alert proportions are also reported separately, preventing undefined cloud-only latency during total outages. H5 is not included in Holm because it is a separate engineering validation gate, not a null-

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

hypothesis test. If H5 fails, no scenario-optimization or operational-twin claim is made even when H1–H4 support other system components. Secondary outcomes are explicitly exploratory.

For H5, flow MAPE is protected against near-zero denominators: $MAPE_Q=(100/N)\Sigma|Q_t-\hat{Q}_t|/\max(|Q_t|,0.5\ L\cdot min^{-1})$. Valid coverage is the fraction of scheduled held-out sensor-window pairs that pass calibration and quality rules and yield a converged twin solution. The gate must pass separately in each low-, medium-, and high-demand supply schedule; failure in any regime fails H5 globally. Zero-supply windows remain included through the 0.5 $L\cdot min^{-1}$ floor.

- *Outcome Definitions and Statistical Analysis Plan*

Table 9 Outcome Definitions and Statistical Analysis Plan

Domain	Metric and computation	Uncertainty / test
Detection	Event TP if alert overlaps event or occurs within 60 s of onset; precision, recall, F1, false alerts/24 h, detection delay	Cluster bootstrap by day/event, 10,000 resamples
Localization	Top-1 and top-2 zone accuracy; macro accuracy across zones	Wilson interval; clustered paired bootstrap
Forecasting	Onset MAE/RMSE; AUROC and AUPRC for 60-min probability; Brier score and calibration slope	Paired block bootstrap; DeLong only as sensitivity
Twin fidelity	Pressure RMSE/nRMSE; flow MAE/MAPE; empirical 90% interval coverage	Bootstrap by hydraulic cycle; Bland–Altman plots
Offline operation	Local decision availability; recovered-record fraction; duplicates; ordering errors; sync time	Median/IQR and 95% percentile CI by outage condition
Resources	Peak RAM/flash; CPU duty; energy per hour and per inference; backup autonomy	Repeated-measures estimates by scenario
Human factors	SUS score; task success; time; comprehension; alert appropriateness	Descriptive mean/median with 95% bootstrap CI; no interface-comparison test

All models are evaluated once on the locked test set. Binary outcomes use event-level confusion matrices. Repeated windows are nested within events and days; uncertainty is clustered at the highest relevant independent unit. The confirmatory tests are frozen rather than selected after assumption checks: H1, H2, and H4 use one-sided paired randomization procedures; H3 uses exact McNemar. Linear mixed models, cluster bootstrap, Wilcoxon tests, standardized effects, and rank-biserial correlations are sensitivity or descriptive analyses only and cannot replace the confirmatory tests. Model-selection comparisons on validation data are descriptive and do not consume confirmatory test alpha.

Sample sizes are fixed before data collection. The executable power scripts use seed 20260817 and the worst-case first Holm threshold $\alpha=.0125$. H1 assumes standardized paired capped-latency effect $d=0.45$ (128 paired impairment runs; power 0.998); H3 assumes discordant localization probabilities 0.12 versus 0.04 (360 leak events; exact-McNemar simulation power 0.941); and H4 assumes standardized paired absolute-error effect $d=0.35$ (120 cycles; power 0.944). H2 is not approximated by correctness probabilities: it simulates the actual paired macro-F1 estimand over all 520 detection events (360 leaks and 160 non-leak disturbances), with baseline sensitivity/specificity 0.78/0.85, MAYI-EDGE 0.86/0.90, and Gaussian-copula paired-error correlation $\rho=0.60$. Across 1,000 outer simulations and 999 label randomizations per simulation, estimated H2 power is 0.940 and median Δmacro-F1 is 0.0738. Assumptions and results are machine-readable in the accompanying package. Fixed counts cannot change after confirmatory collection begins; incomplete campaigns are reported without redefining the target sample.

- *Offline, Network, and Energy Validation*

Network impairment is injected with reproducible profiles for total outage, packet loss, bandwidth restriction, and variable round-trip time. Each profile is paired with the same hydraulic scenario in edge and cloud-only modes. Outcomes include sensing continuity, local inference continuity, local-alert latency, broker-delivery latency, backlog growth, checksum-valid recovery, duplicates, ordering errors, and time to synchronization. The primary integrity target is $\geq$99.5% recovery of valid buffered records with zero silently corrupted records.

Resource profiling records heap high-water mark, flash writes, CPU duty, telemetry bytes, and energy at idle, sensing, inference, publish, and reconnection. Autonomy is measured from a fully charged characterized backup source to the predefined safe-shutdown voltage. Energy claims are normalized by operating hour and number of inferences; laboratory supply readings are separated from field estimates.

- *Human Evaluation*

A formative prespecified usability study will recruit exactly 24 adults: 12 water-system operators/technicians and 12 household users. Recruitment stops only when both quotas are complete; shortfall is reported without replacement by another population. Eligibility, informed consent, demographic variables, prior digital experience, and role are recorded. Participants complete four fixed tasks: identify current network state, acknowledge an alert, locate the suspected zone, and interpret the next availability window. The study uses one interface only, a fixed training script and scenario order, screen recording with consent, task success,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26aug874

time, errors, the 10-item System Usability Scale (SUS) [13], and five comprehension/appropriateness items.

Participants do not evaluate leak-detection accuracy; they evaluate whether evidence and uncertainty are understandable and actionable. A facilitator who did not implement the tested interface administers the script when possible. Adverse comments, incomplete tasks, and missing questionnaire items are retained. The protocol seeks institutional ethics determination before recruitment; if only expert heuristic review is conducted, it will not be mislabeled as user satisfaction.

- *Reproducibility and Open-Science Plan*

- ✓ A versioned minimum reproducibility package accompanies the review manuscript as MAYI EDGE_Reproducibility_Package_v1.1.zip (SHA-256:
- ✓ 70f3f7691ec53569230ff3319a857cf8d285b28860439c721b49a76a2187c4c9). It contains the deterministic fixture, seed 42, Figure 4 script and trace, power simulations, SQLite offline contract, frozen detector configuration, MQTT/JSON schema, tests, README, license, citation metadata, and locked environments. A public immutable archive URL and DOI must replace this review-stage statement before submission; no repository DOI is claimed here.
- ✓ Before external journal submission, the identical package must be placed in a public repository, tagged v1.1.0, archived with a permanent DOI, and linked in both the manuscript and submission metadata. The final archive checksum is recorded in the Data and code availability statement.
- ✓ A private or embargoed extension may contain security-sensitive topology or physical data, but it cannot replace the public minimum package required to reproduce Figure 4.
- ✓ Machine-readable bill of materials; wiring diagram; topology; calibration sheets; MQTT topic/schema specification; configuration files; software and library versions.
- ✓ De-identified 1-Hz or 10-s data where governance permits, plus the supplied synthetic EPANET design topology and declared surrogate generator when physical topology cannot be disclosed.
- ✓ Fixed random seeds, deterministic flags where supported, serialized splits, frozen predictions, model cards, energy-test scripts, and an executable reproduction README.
- ✓ After the public package and fixed protocol exist, a permanent preregistration timestamp is created before test-label access; deviations are recorded with rationale, date, affected outcomes, and confirmatory/exploratory status.

- *Ethics, Security, and Data Governance*

Hydraulic telemetry can reveal household routines, infrastructure vulnerabilities, and sensitive operational conditions. The minimum necessary data are collected; household identifiers are pseudonymized; precise topology and credentials are access-controlled; retention periods and permitted uses are documented. Access follows least privilege, encryption is used in transit and at rest, and audit logs record administrative actions. Raw credentials never appear in released datasets or source code.

Human-participant recruitment begins only after ethics review or a documented exemption/determination. Consent covers study tasks, screen recording, withdrawal, and data retention. Operators can override automated recommendations. MAYI-EDGE is not a safety-certified autonomous controller: during the protocol, actuation recommendations require human confirmation, physical pressure relief remains independent, and network safety takes precedence over experimental completeness.

Governance is shared with the infrastructure owner and affected community. A data dictionary identifies controller, steward, permitted recipients, retention, deletion, breach response, and publication status for each data class. Security testing includes credential revocation, unauthorized topic access, replay, malformed payload, and denial-of-service resilience within safe test limits.

V. DISCUSSIONS, LIMITATIONS, AND FALSIFIABILITY

The protocol is expected to determine—not assume—whether edge processing provides operational benefit under outages, whether hybrid inference outperforms simpler baselines, and whether the twin is sufficiently faithful for scenario analysis. The architecture will be considered partially successful if monitoring continuity improves but predictive hypotheses fail. Null and negative results remain informative because component-level ablations identify where complexity does not add value.

The controlled network cannot represent every pipe material, elevation, illegal connection, demand pattern, or transient in Kinshasa. Low-cost sensors may drift, controlled leaks are easier to label than natural failures, and an eight-week controlled campaign cannot establish seasonal generalizability. Zone-level localization is intentionally narrower than pinpoint localization. Human evaluation measures usability, not long-term adoption or water savings. Any estimate of water saved requires a later counterfactual field design with comparable operating conditions and is not a primary claim of this protocol.

VI. CONCLUSION

MAYI-EDGE is presented as a rigorous reference architecture, reproducible algorithmic and offline-contract demonstration, and prespecified experimental protocol for intermittent water distribution—not as a completed physical implementation, executed EPANET twin, deployed end-to-end platform, or real-world validation. Its scientific value lies in the precise conjunction of intermittent supply, offline-first edge inference, transactional recovery, a versioned hydraulic-twin design, resource constraints, and falsifiable confirmatory evaluation. H1–H4 constitute the Holm-controlled statistical family; H5 is a separate engineering gate. After the public reproducibility package and permanent protocol timestamp exist, the study may accurately be described as preregistered. Until then, "prespecified" is the only valid designation.

DECLARATIONS

➢ *Ethics Status:*

No human participant has been recruited and no human-participant data have been collected for this architecture/protocol article. Recruitment is prohibited until the competent institutional body issues written approval or a formal exemption/determination; its name and reference number must be inserted before any report of human results.

➢ *Consent to Participate:*

Not applicable to the present software demonstration. Written informed consent is mandated for the future usability study.

➢ *Competing Interests:*

The authors declare no competing interests.

➢ *Funding:*

The present architecture and software-demonstration work received no externally reported project funding; every author must verify this statement before submission.

➢ *Data and Code Availability:*

The complete review package MAYI-EDGE_Reproducibility_Package_v1.1.zip accompanies this manuscript and is integrity-identified by SHA-256 70f3f7691ec53569230ff3319a857cf8d285b28860439c721b4 9a76a2187c4c9. A public immutable release URL/DOI is a mandatory author action before portal submission.

➢ *Author Contributions (CRediT):*

H.K.M.: conceptualization, methodology, supervision, project administration, and writing—original draft. G.G.D., D.K.M., E.I.M., K.E.M., K.M.L., J.-B.L.I., S.D.K., and F.F.S.: architecture development, executable-demonstration support, data curation planning, validation planning, and writing—review and editing. Physical-prototype implementation must not be assigned until contribution evidence exists. All authors must approve the final statement and manuscript before submission.

➢ *Study Type:*

Architecture-led system conception + reproducible algorithmic/offline-contract demonstration + prespecified confirmatory protocol.

REFERENCES

[1]. A. Pagano, D. Garlisi, I. Tinnirello, et al., "A survey on massive IoT for water distribution systems: Challenges, simulation tools, and guidelines for large-scale deployment," Ad Hoc Networks, vol. 168, 103714, 2025. doi:10.1016/j.adhoc.2024.103714.
[2]. R. J. L. Taloma, F. Cuomo, D. Comminiello, and P. Pisani, "Machine learning for smart water distribution systems: exploring applications, challenges and future perspectives," Artificial Intelligence Review, vol. 58, art. 120, 2025. doi:10.1007/s10462-024-11093-7.
[3]. P. Conejos Fuertes, F. Martínez Alzamora, M. Hervás Carot, and J. C. Alonso Campos, "Building and exploiting a Digital Twin for the management of drinking water distribution networks," Urban Water Journal, vol. 17, no. 8, pp. 704–713, 2020. doi:10.1080/1573062X.2020.1771382.
[4]. J. E. Pesantez, V. Alghamdi, S. Sabu, A. Mahinthakumar, and E. Berglund, "Using a digital twin to explore water infrastructure impacts during the COVID-19 pandemic," Sustainable Cities and Society, vol. 77, 103520, 2022. doi:10.1016/j.scs.2021.103520.
[5]. P. Ghorbani Bam et al., "Digital Twin Applications in the Water Sector: A Review," Water, vol. 17, no. 20, 2957, 2025. doi:10.3390/w17202957.
[6]. E. Q. Shahra et al., "Intelligent Edge–Cloud Framework for Water Quality Monitoring," Water, vol. 16, no. 2, 196, 2024. doi:10.3390/w16020196.
[7]. S. El-Zahab and T. Zayed, "Leak detection in water distribution networks: an introductory overview," Smart Water, vol. 4, art. 5, 2019. doi:10.1186/s40713-019-0017-x.
[8]. L. Romero-Ben, D. Alves, J. Blesa, M. G. Cembrano, V. Puig, and E. Duviella, "Leak detection and localization in water distribution networks: Review and perspective," Annual Reviews in Control, vol. 55, pp. 392–419, 2023. doi:10.1016/j.arcontrol.2023.03.012.
[9]. A. Pagano, D. Garlisi, F. Giuliano, T. Cattai, R. J. L. Taloma, and F. Cuomo, "Introducing and evaluating SWI-FEED: A smart water IoT framework designed for large-scale contexts," Computer Communications, vol. 237, 108146, 2025. doi:10.1016/j.comcom.2025.108146.
[10]. L. A. Rossman et al., EPANET 2.2 User Manual. Cincinnati, OH: U.S. Environmental Protection Agency, EPA/600/R-20/133, 2020.
[11]. U.S. Environmental Protection Agency, "EPANET: Application for Modeling Drinking Water Distribution Systems," 2025. https://www.epa.gov/water-research/epanet (accessed 17 Aug. 2026).
[12]. K. Klise, M. Bynum, D. Moriarty, and R. Murray, "A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study," Environmental Modelling & Software, vol. 95, pp. 420–431, 2017. doi:10.1016/j.envsoft.2017.06.022.
[13]. J. Brooke, "SUS: A 'quick and dirty' usability scale," in Usability Evaluation in Industry, P. W. Jordan et al., Eds. London: Taylor & Francis, 1996, pp. 189–194.
[14]. OASIS, MQTT Version 5.0, OASIS Standard, 2019. https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html.
[15]. A. M. Lieb, C. H. Rycroft, and J. Wilkening, "Optimizing intermittent water supply in urban pipe distribution networks," SIAM Journal on Applied Mathematics, vol. 76, no. 4, pp. 1492–1514, 2016. doi:10.1137/15M1038979.
[16]. ISO/IEC, ISO/IEC 27001:2022, Information security management systems—Requirements, Geneva, 2022.

[17]. M. Zuñiga-Uribe et al., “Artificial Intelligence in Water Distribution Networks: A Systematic Review of Models, Input Variables, Databases, and Output Strategies for Leak Detection,” Smart Cities, vol. 9, art. 45, 2026. doi:10.3390/smartcities9030045.

[18]. N. T. Mücke, P. Pandey, S. Jain, S. M. Bohté, and C. W. Oosterlee, “A probabilistic digital twin for leak localization in water distribution networks using generative deep learning,” Sensors, vol. 23, no. 13, 6179, 2023. doi:10.3390/s23136179.

[19]. V. Travaš, E. Gal, I. Lučin, and E. Žic, “Digital twin for a real-time leakage detection and localization in pressurized piping systems,” Journal of Hydroinformatics, vol. 27, no. 4, pp. 755–770, 2025. doi:10.2166/hydro.2025.304.

[20]. T. A. Syed et al., “Smart Water Management with Digital Twins and Multimodal Transformers: A Predictive Approach to Usage and Leakage Detection,” Water, vol. 16, no. 23, 3410, 2024. doi:10.3390/w16233410.

[21]. P. Locatelli, T. Cattai, P. Spadaccino, M. Casalbore, and F. Cuomo, “TwinAI: A digital twin and graph reinforcement learning framework for real-time management of water distribution networks,” Internet of Things, vol. 37, 101911, 2026. doi:10.1016/j.iot.2026.101911.

[22]. E. Steins and A. Cominola, “Addressing Practical Challenges of Stochastic Process Control for Leakage Detection in Water Distribution Networks: A Comparative Analysis,” Journal of Water Resources Planning and Management, 2025. doi:10.1061/JWRMD5.WRENG-6969.

[23]. H. Wang, “AI for detecting and localizing concurrent abrupt and incipient leaks in water distribution networks,” Water Research, vol. 295, 125538, 2026. doi:10.1016/j.watres.2026.125538.

[24]. K. V. Shah and H. M. Patel, “Source Contamination Detection and Identification of Affected Spots in the Intermittent Water Supply System Using Mixed Tool-Based and Data-Driven Machine Learning Techniques,” Journal of Water Resources Planning and Management, vol. 152, no. 3, 2026. doi:10.1061/JWRMD5.WRENG-7171.