 ISSN:2302-8149
e-ISSN:2540-9719

AI-Enhanced Fault-Aware LPG Leak Monitoring with Environmental Compensation and Network-Independent Local Alarming

[1]Zainab marid alzamili*
[1]Education Directorate of Thi-Qar, Ministry of Education, Iraq
*e-mail: Zainab.alzamili@utq.edu.iq

(***received*:** 9 October 2026, ***revised*:** 2 November 2026, ***accepted*:** 5 November 2026)

Abstract

The use of a very short prototype demonstration is often reported for the low-cost liquefied petroleum gas (LPG) alarms without explicit assessment of environment cross-sensitivity, drift, device transfer, nuisance alarms and communication failure. This paper introduces a fault-aware reference architecture for the Arduino/MQ-6 device which performs local alarm decision and optional network notification in a separate path, and adds temperature and relative humidity into the detection pathway. The finished evaluation is a synthetic benchmark that consists of 12 development sessions and 36 held-out sessions, with 216 target events in 18 simulated monitoring hours over six virtual machines (two unseen). For the rule-based approach, the data from all environments were compensated, the resulting number of new alarms was timely, and the approach yielded 205/216 (94.91%) of new alarms when the session-bootstrap interval for the percentage of timely new alarms was calculated under the 95% rule for type I errors, with 4.19 nuisance alarms per valid background hour. B2 had a stronger overall benefit than filtered uncorrected detection (B1) with the number of timely new alarm it detected being 40.74 percentage points higher, while the 1.07 nuisance-alarm reduction per hour was inconclusive. The use of a raw-signal override (B3) did not provide an improvement in timely detection, and instead caused nuisance alarms. For future leakage safe comparison, four AI approaches including Random Forest, RBF-SVM, gradient-boosted trees, and a compact multilayer perceptron are given. Experimental performance of these approaches is not reported in this paper. These simulated results do not illustrate the accuracy of detecting physical LPGs, the accuracy of alarms, or the proven ability to withstand network failures.

Keywords: In this report the following usages of the devices are described: LPG leak monitoring, MQ-6, compensation for environmental influences, sensitivity by events, nuisance-alarm rate, synthetic simulation, alarming by error, sensor drift, and transfer of a device.

Introduction

LPG monitoring chain consists of physical exposure to sensing, analog conversion, calibration, decision logic, local warning, power integrity, and, if available, remote notification. A seemingly good lookalike prototype can be ruined by failure at any point. Low cost sensing, local alarms, wireless messaging, and automatic mitigation are all possible with published Arduino/IoT systems [1], [2], [6] but this does not guarantee that the measurements are valid or the systems are reliable.

Distribution shift is the main challenge. In addition to detecting a target gas, MOS sensors are sensitive to temperature, humidity, aging, contaminating species, manufacturing differences, airflow and interfering species. Recent studies, however, have been increasingly focusing on gas sensing as a calibration-transfer problem and a drift adaptation problem instead of a fixed threshold problem [3]–[5], [7]–[12]. This is done in the current study, where a clear rule-based safety path is implemented, alongside an optional AI decision path, allowing for the added complexity to be assessed, rather than taken as granted.

Four questions are posed in the study. RQ1: How is the sensitivity of the new alarm, nuisance alarms per valid background hour and detection delay in the synthetic benchmark compared to the uncorrected rules influenced by the T/H compensation? RQ2: What is the stability of these results throughout subsequent sessions, across environmental-stress scenarios, and virtual devices not used during development? RQ3: How to design an experiment to test whether compact AI classifiers can improve the quality of the alarms versus the rule-based baseline without causing unacceptable edge

 ISSN:2302-8149
e-ISSN:2540-9719

latency or memory usage? RQ4: What needs to be done to ensure alarms are not suppressed, cleared or delayed on the local alarm by the loss of remote connectivity? The contribution is a dedicated local state machine (with validity/explicit fault), separate asynchronous notification channel, session/device-aware synthetic evaluation and AI/physical validation protocol. RQ4 is only at the architectural and logical design level: there is no claimed physical experiment of communication-disconnection nor any measured timing of hardware; there is neither any asserted safety performance nor any certified safety performance.

1 Related Work and Research Positioning

There are four complementary literatures. First, LPG prototypes employing MQ-series sensors and microcontroller-based control are shown to be feasible, and typically use threshold triggers and short duration tests [1], [2], [6]. Second, sensor response in MOS compensation studies can be systematically affected by temperature and humidity, which can be mitigated by including explicit covariates and performing regression correction [3]. Third, longitudinal e-nose studies show that drift is a long-term domain-shift problem: In the one-year study of Wörner et al. [5] drift is a problem that can be evaluated over time, and more recently, studies in machine learning have applied ensemble learning, knowledge distillation, domain adaptation, and meta-learning to ensure accuracy across batches or devices [7]–[12]. Fourth, real gas-leak analytics are becoming more and more based on IoT time-series features and machine learning instead of single instantaneous thresholds [4].

The recent AI evidence is especially relevant, but cannot be directly applied to an MQ-6 LPG alarm. Dong and Han used iterative Random Forest (RF) correction and incremental domain adversarial network (IDN) for long-term MOS drift [7]. Lin and Zhan compared the performance of KD to DRCA with multiple partitioning of the UCI Gas-Sensor benchmark and achieved relative gains in accuracy and F1-score [8]. Schauer et al. put forward the concept of interpretable calibration transfer for complex mixtures and explicitly considered sensor-to-sensor and temporal domain shifts [9]. For few-shot drift adaptation, Yang et al. adopted the meta-learning with dual-branch feature extraction approach [10]. Transfer learning, contrastive learning, attention and real-world limitations of labeled data and edge computation are highlighted in a 2026 review by Li et al. [11]. These studies support adding AI “baselines” but the results of these studies are primarily for classification or regression of gas rather than for latched event alarming with false alarms per monitoring hour.

We believe the missing link here is the relationship between drift-aware learning and safety-related event detection. A good LPG alarm should be able to recognize new events, not be falsely triggered, be able to count missed/late events, be able to pass the local alarm through network failures, and show transfer between hardware units. The most relevant studies are analyzed in Table 1 with regard to the sensor/target, data, algorithm, environmental treatment, validation design, reported result, limitation, and implication for this work.

Table 1 Detailed Analytical Comparison of Recent LPG/MOS Sensing and AI Drift-Compensation Studies (2022–2026)

Study	Sensor / target	Data / system	Algorithm / logic	Env. / drift handling	Validation design	Reported outcome	Main limitation	Implication here
Abdullah et al. [3], 2022	MOX; ethanol	4 sensors + T/RH	3-D linear regression	Explicit T/RH correction	Multiple concentrations and environmental levels	Reduced response variability after correction	Not LPG; no event alarm	Use T/RH as covariates; test on held-out sessions
Jena et al. [6], 2023	MQ-series; LPG	Microcontroller + GSM	Threshold trigger	Not modeled	Prototype functional tests	Integrated detection and notification	Limited calibration/ transfer evidence	Integration is prior art; quantify reliability
Pamungkas et al. [1], 2024	MQ-6; LPG	Arduino + local actions	Threshold/ sequence logic	Not quantified	Prototype response tests	Demonstrated MQ-6 alarm actions	No independent truth or long monitoring	Use reference-labeled events and false alarms/h
Rivera-Martinez et al. [4], 2024	MOX; methane	Controlled releases	Reference-analyzer modeling	Field effects included	Independent controlled releases	Quantitative leak/source analysis	Methane; different endpoint	Independent exposure truth and synchronized timing

 ISSN:2302-8149
e-ISSN:2540-9719

Study	Sensor / target	Data / system	Algorithm / logic	Env. / drift handling	Validation design	Reported outcome	Main limitation	Implication here
Wörner et al. [5], 2025	62-MOX e-nose	One-year dataset	Drift benchmark	Temporal drift explicit	Repeated acquisition over 12 months	Public long-term drift resource	Not LPG/ safety alarm	Use later-session and device-transfer tests
Jiang et al. [12], 2025	E-nose arrays	Two drift datasets	Semi-supervised ensemble + center loss	Domain drift modeled	Long-/short-drift experiments	76.06%/82.07% average classification; R^2 0.804 regression	Classification/ regression; array-level compute	Include ensemble trees as compact AI baseline
Dong & Han [7], 2025	MOS arrays	Multi-channel temporal data	Iterative RF + incremental domain-adversarial network	Real-time error + long-term drift	Temporal drift evaluation	Improved correction and drift robustness vs conventional methods	Multi-sensor architecture; not LPG event alarm	RF offers interpretable edge baseline; deep adaptation is optional
Lin & Zhan [8], 2026	UCI e-nose	Gas Sensor Array Drift dataset	Knowledge distillation vs DRCA	Long-term domain shift	30 randomized test partitions; two transfer tasks	Up to 18% relative accuracy and 15% F1 improvement	Benchmark classification; no alarm persistence	Use repeated/ paired validation; do not rely on one split
Schauer et al. [9], 2026	MOS; 8 volatiles	Lab calibration, complex mixtures	Interpretable calibration transfer	Sensor-to-sensor + aging/ poisoning shifts	Cross-sensor/ domain transfer	Interpretable quantification under domain shift	Complex mixtures; laboratory system	Treat sensor replacement as domain shift, not same-device calibration
Yang et al. [10], 2026	E-nose	Drift + twin-array benchmarks	Meta-learning dual-branch network	Few-shot unseen drift/ device adaptation	Long/short drift + cross-device	95.53%, 95.71%; 97.89% cross-device (reported)	Higher complexity; benchmark classification	Use unseen-device evaluation; compare benefit against edge cost
Li et al. [11], 2026	MOS arrays	Review literature	Transfer, contrastive, attention, deep DA	Drift mechanisms synthesized	Comparative review	Identifies strong trend toward adaptive learning	Review; heterogeneous datasets	Motivates AI, but requires LPG-specific evidence
Present study	MQ-6 target LPG	Synthetic sessions + planned lab campaign	B0–B3 rules + RF/SVM/ boosting/ MLP protocol	T/RH, drift, device shift, faults	Session/ device holdout + later physical validation	Rule-based synthetic outcomes already executed; AI results not yet claimed	No physical LPG measurements yet	Combines safety event endpoints with AI transfer testing

2 System Architecture and AI-Enhanced Decision Layer

2.1 Hardware Scope and Local Independence

The reference design includes an analog channel for LPG-related response (MQ-6), an Arduino-compatible local controller, temperature and RH sensing, local buzzer and response indicators, a display, power supervision and an optional Wi-Fi interface. The MQ-6 is considered to be a “broadly responsive” MOS sensor and an alarm does not necessarily indicate the proof of chemical identity, but only that the measured response is similar to the target conditions that are calibrated. In Figure 1, you can see how the separation between sensing/decision logic and remote notification works. This separation is a reliability requirement and an example would be that if a network timeout occurs, it should be detected as degraded but not cause disruption to any sampling or clear a latched local alarm.

 ISSN:2302-8149
e-ISSN:2540-9719

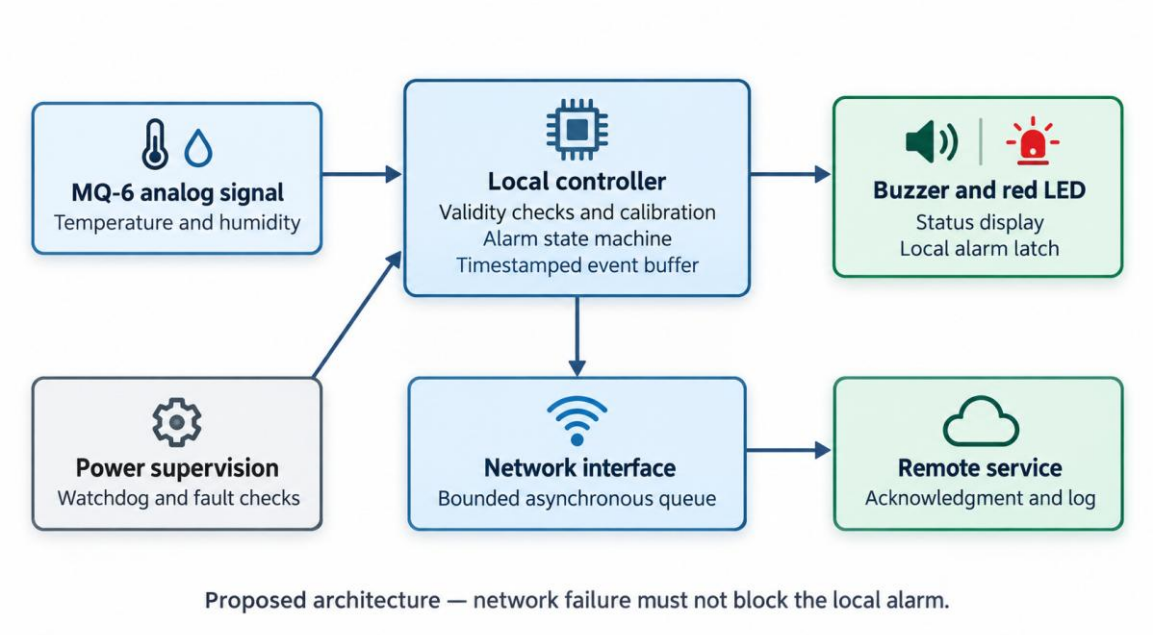


Figure 1 Proposed Monitoring Architecture with Local Alarm Ownership, Environmental Sensing, Fault Supervision, and Asynchronous Network Delivery

Table 2 summarizes the evidence needed and roles for implementation of each component. The table is designed for verification: it isn't enough to use a schematic for a hardware claim, but needs to be supported by a sensor/module identifier, calibration record, timing measurement or fault test.

Table 2 Proposed Hardware Roles and Evidence Required for Physical Validation

Element	Role	Evidence required
MQ-6 module	Target-responsive analog signal	Module/circuit identity, load resistance, heater/supply history
T/RH sensor	Environmental covariates	Exact model, calibration/uncertainty, placement
Local controller	Acquisition, validity checks, alarm and AI inference	Firmware version, ADC reference, loop timing, memory
Buzzer/LED/display	Independent local warning	Driver operation, startup/fault behavior, audibility/visibility checks
Wi-Fi interface	Supplementary event delivery	Module, supply integrity, timeout/retry and acknowledgment logs
Power/fault supervision	Detect invalid operating states	Thresholds, watchdog/reset logs and fault injection

2.2 Local State Machine

The local controller has six operating states (Figure 2), as specified: INITIALIZING, MONITORING, WARNING, ALARM, FAULT, RECOVERY. Local alarm latch is owned by the controller and will not reset because of a remote acknowledgment, loss of communication or a communication timeout. LATCH SELIT TO FAULT does not clear any previously latched ALARM Only local conditions, and an approved RESET policy, will clear and recover. The network interface is not an input to the sensing/alarm transition, but is simply an additional asynchronous notification channel. This is a design constraint and NOT a measurable time or statement of hardware failure. Physical tests should be performed to ensure all of the following is recorded and disconnected from the network: Local output, event timing, power state, and controller logs. The AI decision layer is an auxiliary element of deterministic sensor-validity checks and fault behavior, if implemented.

 ISSN:2302-8149
e-ISSN:2540-9719

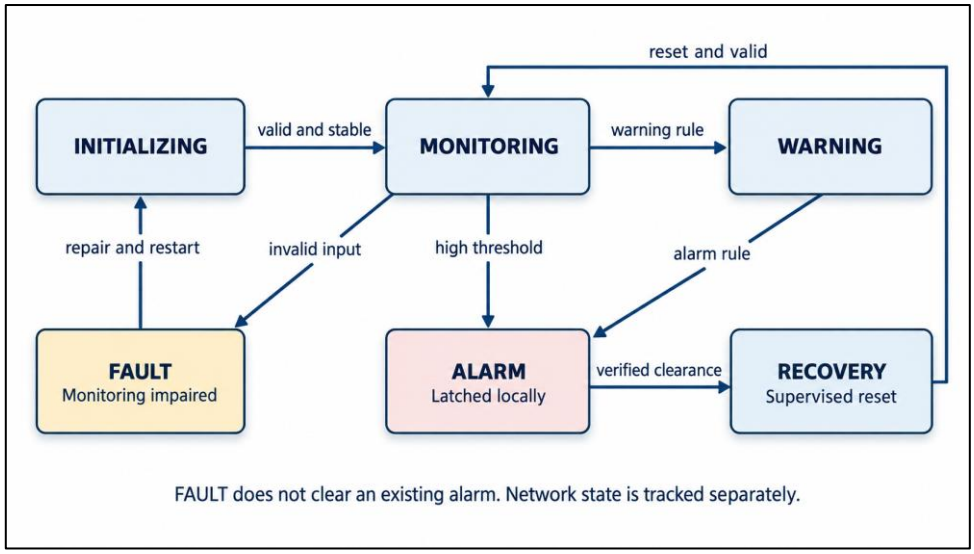


Figure 2 Local State Transitions. Network health is tracked independently, and FAULT does not clear a pre-existing ALARM latch

2.3 Signal Processing and Environmental Compensation

ADC code is converted to voltage and sensor resistance for a verified divider topology, based on the measured reference and supply voltage. A per device reference resistance R_0 is determined from documented calibration not from a clean air constant. Normalized log-response

$$x(t) = \log(R_s/R_0) \tag{1}$$

can be used to model against reference concentration and environmental deviations. A linear correction,

$$u(t) = x(t) - \beta_T \Delta T - \beta_H \Delta H, \tag{2}$$

is kept as the clear primary model; nonlinear terms are explained only if they enhance held-out performance. Then, a causal exponential filter

$$z(t) = \alpha u(t) + (1-\alpha) z(t-1), \tag{3}$$

consecutive-sample confirmation, hysteresis and controlled clearance produce transitions between warning/alarm. The parameters are hard-coded in before the test is evaluated.

2.4 Added Artificial-Intelligence Algorithms

The AI extension should not be used as a substitute for validity checks, local latching or conservative fault behavior. Interpretable tree ensembles, margin-based classification, nonlinear boosting, and a small neural network are selected for the four algorithms used—Random Forest (RF), radial-basis-function Support Vector Machine (RBF-SVM), gradient-boosted trees (XGBoost or an equivalent histogram gradient-boosting implementation), and a compact multilayer perceptron (MLP). Causally computed inputs are the current and recent samples of raw/normalized response, compensated response, first difference, short-window slope and variance, temperature, humidity, interaction T×RH, filter state, time since start-up, sensor-validity flags, and optionally the device calibration metadata. The target is a label based on event proximity or leak state that is realized and contained solely by independent reference truth in physical experiments.

Each algorithm listed in Table 3 has a specific role and is evaluated in a leakage-safe manner. Hyperparameters can only be optimized on development sessions. The scaling, imputation, feature selection and calibration are performed within the training partition. Preprocessing cannot be affected by any sample taken from a held-out device or held-out test session. If deployed in the embedded application, the final model is chosen, not only for the performance, but also for limited resources; a slightly more accurate model is not preferred if latency, memory or calibration instability will affect local operations.

 ISSN:2302-8149
e-ISSN:2540-9719

Table 3 AI Algorithms Added to the Study and Their Evaluation Role

Algorithm	Why included	Core settings to tune	Output / interpretation	Deployment note
Random Forest	Nonlinear interactions; robust tabular baseline; feature importance	Trees, depth, min leaf, class weights	Leak probability / vote; impurity or permutation importance	Good edge baseline; memory must be measured
RBF-SVM	Strong small-data nonlinear separator	C, γ, class weights; probability calibration	Margin or calibrated probability	Requires feature scaling; support-vector count affects runtime
Gradient-boosted trees	Captures weak nonlinear environmental effects	Trees/iterations, learning rate, depth, subsampling	Probability + SHAP/feature contribution where supported	Compact models possible; guard against overfit
Compact MLP	Tests whether learned nonlinear interactions add value	1–2 hidden layers, width, L2/dropout, early stopping	Sigmoid probability	Use only if cross-device gain exceeds complexity cost

3 Validation Design and Reproducible Benchmark

Physical validation is still required as all of the evidence is currently from synthetic signals. The logs for a real-world campaign include gas composition, identity and calibration of reference instrument, concentration and time of exposure, chamber geometry and airflow, ambient temperature and relative humidity, sensor position, sensor age and heating history, firmware and sampling clock, supply voltage and fault/network injection logs. The ground-truth leak events should be from an independent reference device and not from an MQ-6 threshold. Complete sessions and devices are assigned to separate partitions: All the preprocessing, calibration, feature selection, model tuning, and threshold selection should be done using development data only, not adjacent samples or overlapping windows. Three sensor units, independently identified, meeting the requirements of feasibility, is a starting point, but not evidence of a population-wide device generalization.

The anticipated factors are: Target gas level, Temporal profile, Approved non-target interferents, Environmental stress, Device replacement, Later sessions, Network availability and Controlled faults. Primary endpoints are the number of alarms triggered in the right time window (30 s prespecified) divided by the total number of events that are labeled as references, while false alarms or delayed alarms remain in the sensitivity denominator and missed or delayed events are not. Conditional detection delay is a summary of detection delay for only those events that are detected and is not a substitute for detection sensitivity. The number of events and the valid background duration should be recorded in the detector output for each held-out session, and the uncertainty in comparison should be obtained by resampling the events in each session. The simulated rule based outputs below were based on 2,000 paired session-bootstrap resamples and 95% percentile intervals. In the current study, no physical LPG samples were collected and no measurements of runtime, memory, calibration uncertainty, network-disconnection performance and accuracy of the AI model were taken.

The completed simulation compared four prespecified rule-based configurations on the same held-out synthetic sessions: B0, a raw threshold; B1, causal filtering, consecutive-sample confirmation, and hysteresis; B2, B1 plus environmental compensation; and B3, B2 plus a raw high-threshold override. Table 4 records the available simulation design details: six virtual units (0–3 represented in development; 4–5 held out), 12 development and 36 held-out sessions, six target events per held-out session, 1-s sampling, 216 target events, and 18 simulated hours. The generator includes target pulses/ramps, virtual-unit gain and offset, temperature/humidity interference, drift, and noise. Reported results apply only to this specified synthetic benchmark; full independent reproduction additionally requires release of generator code, all parameters and random-number streams, immutable session/device assignments, detector thresholds, event matching and alarm-latching code, background-duration definitions, and per-event predictions. The four AI models in Section 3.4 describe a future evaluation protocol, not models trained and tested in this reported simulation.

 ISSN:2302-8149
e-ISSN:2540-9719

Table 4 Synthetic Benchmark Design Used for the Existing Rule-Based Results

Item	Setting
Partition	12 development sessions; 36 held-out test sessions; 6 events/session
Virtual devices	Six; units 0–3 represented in development; units 4–5 held out
Test set	216 target events; 18 simulated hours; nominal and stress sessions
Generator	1-s sampling; target pulses/ramps; virtual device gain/offset; T/RH interference; drift/noise
Rule-based models	B0 raw threshold; B1 filtering/hysteresis; B2 + environmental correction; B3 + raw override
Uncertainty	2,000 paired session-bootstrap replicates; 95% percentile intervals
Evidence scope	Software behavior only; no physical sensor, gas concentration, communication, or safety validation

4 Illustrative Synthetic Results

This section presents all results that are synthetic or not measurements from a physical LPG sensor. Table 5 shows the results of each of the following distinct operations that should be examined together: timely new alarm sensitivity, false alarms per valid background hour, delay conditional on timely detection, and events not detected in time with a new alarm. B2 also recorded 205/216 (94.91%) timely new alarms and 4.19 nuisance alarms per hour against 117/216 (54.17%) and 5.26 nuisance alarms per hour respectively recorded by B1. The point estimates are therefore skewed towards B2 on both endpoints, but a decrease in nuisance alarms as compared to B1 is not considered to be proven since the paired uncertainty intervals contain zero. B0 had a significantly increased nuisance burden of 109.70/h. These two operational factors are visualised in figure 3, and examined not only because of the importance of the sensitivity as a safety criterion but also because of the other operational considerations.

Table 5 Held-Out Synthetic Rule-Based Outcomes; 95% intervals are session-bootstrap intervals.

Method	Timely new alarms	Sensitivity %	False alarms/h	Delay median / p90 (s)	No timely new alarm
B0	158/216	73.15 [67.59,78.72]	109.70 [97.25,122.71]	7 / 23	58
B1	117/216	54.17 [48.15,59.72]	5.26 [4.36,6.23]	13 / 27	99
B2	205/216	94.91 [91.20,97.69]	4.19 [3.30,5.10]	14 / 24	11
B3	204/216	94.44 [90.74,97.69]	4.91 [3.91,5.95]	14 / 24	12

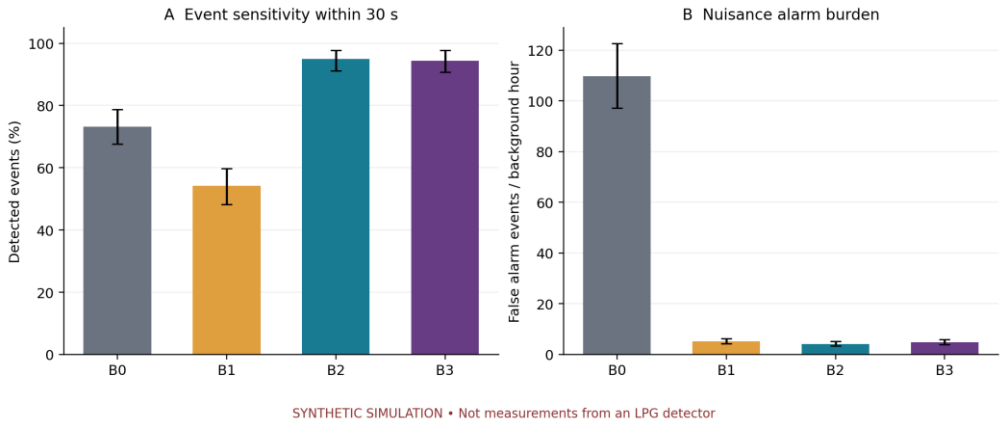


Figure 3 Synthetic Timely New-Alarm Sensitivity and Nuisance-Alarm Rates.

 ISSN:2302-8149
e-ISSN:2540-9719

The timing interpretation is extended in Figure 4. The cumulative curves include non-successful events in the denominator and the boxplots summarize the delays of only timely detections. These were 14-s median and 24-s p90 conditional delays for B2. The medians are descriptive, as they are different for each method, and are not a paired causal estimate of added delay.

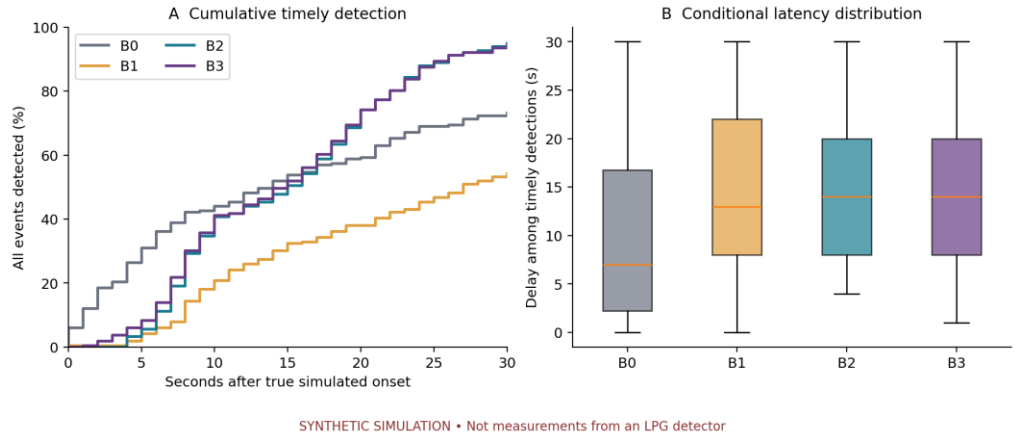


Figure 4 Synthetic Cumulative Detection and Conditional Latency Distributions; undetected events remain visible in the cumulative denominator.

Another key difference between a new alarm and a system that is already sounding at the start of an alarm is alarm persistence. Despite the high alarm-active coverage rates (90.28% for B1, 97.59% for B2), the rates of timely-new-alarm (TNA) were different: 79 events had an alarm already set for B1 at the moment of target onset, whereas only five events had an alarm already set for B2 at the target onset (see Table 6). B2 had 97.22% active coverage, but failed to report on 11/216 "alarms" in time. These are required as the nuisance alarm might be active for a long period of time, with little discriminative event information being provided. The denominator remains the same as in the existing record, which contains 216 simulated target events, but the previous reference to 149 target onsets was ambiguous and this has been eliminated.

Table 6 Secondary Synthetic Endpoints Describing Alarm Persistence at Target Onset

Method	Alarm active within 30 s	Active coverage %	Already active at onset
B0	195/216	90.28	45
B1	195/216	90.28	79
B2	210/216	97.22	5
B3	210/216	97.22	6

4.1 Time-Series Behavior, Transfer, and Component Effects

Figure 5 is not a cherry-picked example, it’s the first prespecified test session. The raw response captures both target and environmental motion, with part of the slow environmental displacement corrected in the corrected response, but the same remains true for the non-target pulse – nuisance transitions remain. The figure is useful for the mechanism inspection, but the above aggregate metrics are the key evidence.

 ISSN:2302-8149
e-ISSN:2540-9719

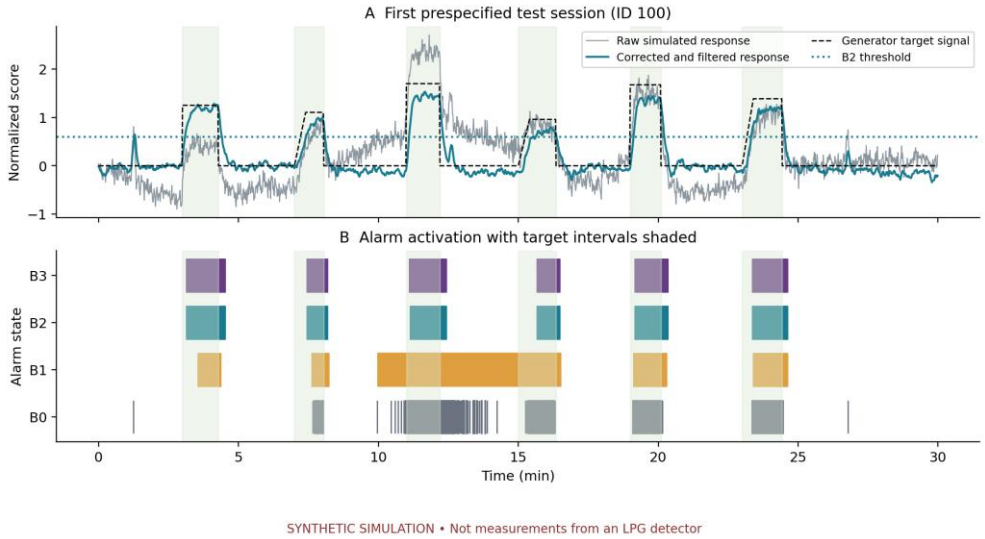


Figure 5 First Prespecified Synthetic Test Session: Generated Target Intervals, Raw/Corrected Response, and B0–B3 Alarm Traces

The heterogeneity by virtual device and stress condition is illustrated in Figure 6. New-alarm sensitivity for B2 was 95.37% in nominal, and 94.44% under stress, situations; 97.92% for development-seen, and 88.89% for unseen, units. This lower unseen-unit value is exactly why it is important to keep device holdout when incorporating the AI models as this would mask this transfer difficulty through random sample splitting.

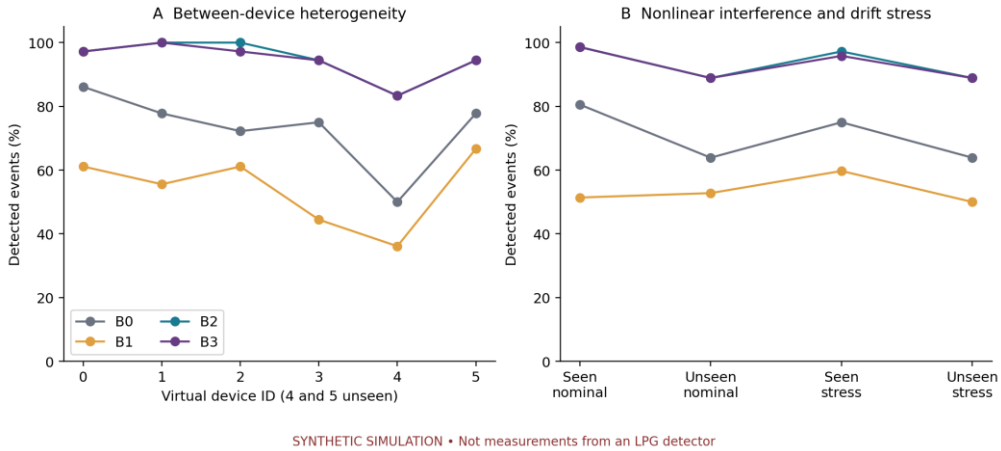


Figure 6 Synthetic Heterogeneity Across Virtual Devices and Environmental-Stress Strata, Including Held-Out Virtual Devices

The paired component estimates shown in Figure 7 alleviate the sensitivity/nuisance trade-off. There was a 40.74 percentage point (95% paired session-bootstrap interval +35.19 to +46.76) increase in timely new-alarm sensitivity when switching to B2 from B1. The reduction in false alarm rate was not statistically demonstrated, although the B2 point estimate was lower (false alarm rate changed by −1.07 per background hour (95% interval −2.33 to +0.18)). However, B3 had an increase in false alarms of +0.71/h (95% interval +0.18 to +1.42) with no improvement in timely new alarms (204 versus 205). Based on this simulated evidence, therefore, it is assumed that increased event sensitivity will be beneficial to the environment, and there is no unequivocal nuisance/alarm benefit or need for the additional raw override. The difference should not be extrapolated to physical LPG exposure, unless tested separately.

 ISSN:2302-8149
e-ISSN:2540-9719

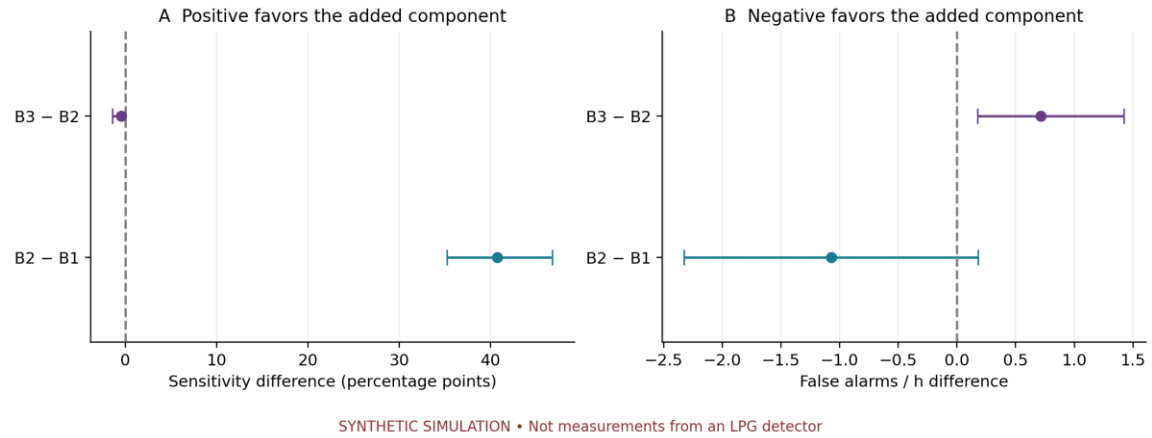


Figure 7 Paired Synthetic Component Effects with 95% Session-Bootstrap Intervals. Crossing zero indicates an inconclusive paired effect.

5. Discussion

The main simulated results are a significant improvement in new-alarm sensitivity when taking environmental correction into account (B2: 94.91%, B1: 54.17%). This gain is only practically significant when considered in conjunction with nuisance burden, detection time and alarm persistence. The rate of 4.19 false alarm events per valid background hour that B2 recorded is a rate that may be disruptive in practice and can lead to alarm fatigue. The nuisance point estimate for B2 is less than that of B1 (5.26/h), but the paired 95% interval for the difference between B2 and B1 (−2.33 to +0.18/h) does not reliably show a reduction. The raw override (B3) triggered nuisance alarms of +0.71/h compared to B2, and missed one event during the evaluation window, so is not justified by the reported simulation. The nuisance cost of B0 was very high at 109.70 false alarms/h, which is very high, especially when compared to B1. The results demonstrate the importance of considering a combination of these parameters – new alarms, existing alarm cover, false alarms per valid background hour, missed events, and event delay – for operational evaluation in addition to maximizing alarm percentage. Environmental interference is intentionally included in the generator, and its removal tests the expected mechanism and not to prove the superiority of MQ-6 in reality. Before claims can be made in practice, independent reference-labeled LPG trials, non-target gas challenges, longitudinal drift, device replacement and end-to-end fault/network tests are required.

The motivation for the AI extension is not due to the fact that machine learning has been shown to be an advantage in this particular experiment, but rather due to recent drift-compensation research. Non-linear interactions between the magnitude of the signals, the temperature/relative humidity, the slope of the signals, and the metadata of the device are represented by random forest and gradient boosted trees, while RBF-SVM is a nonlinear small data comparator and a compact MLP tests learned feature interactions. None of these 4 models have a measured score in this manuscript and thus, any comparative performance or edge-deployment benefit is still prospective. Future comparisons must be based on the same fixed held-out event labels, device/session partition, alarm rules and nuisance-hour denominators used in the baseline, and separate reporting of memory, latency and cross-device calibration. Only after the physical devices repeatedly shift due to repeated evidence should the more complex approaches to adaptation [7]–[12] be considered. The instability of the model after replacing the sensor should also be investigated as one of the signatures for detecting domain shift.

There is still a safety limit. This is a research prototype specification for use, not a certified combustible-gas detector. The identification of a gas and the universal selectivity of an MQ-6 channel cannot be substantiated by a single channel. A server acknowledgment is not a warning given to a human, but it is an extra to network delivery. Likewise, an invalid sensing should not be considered a “safe” state, replaced by the AI confidence. In the current learned model, some validity checks are deterministic (non-learned), such as some of the following: conservative processing of saturation, local latching, fault logging.

 ISSN:2302-8149
e-ISSN:2540-9719

5 Conclusion

This paper defines a fault-aware LPG monitoring architecture, which features independent local alarm ownership, environmental compensation, deterministic sensor-validity handling, and an optional asynchronous remote communication. For the completed synthetic benchmark, the number of alarms generated was 205/216 (94.91% [91.20–97.69%]) and the conditional delay was 14 s with a median of 4.19 false alarms per valid background hour. The sensitivity of the timely-new-alarm was improved by 40.74 percentage points compared to B1, while there was no conclusive improvement in nuisance-alarm reduction with the B3 raw-signal override, and this also added to the number of false alarms. These results provide only the behavior of the simulated rule-based decision logic given the selected generator and held-out sessions. Not an architectural specific property, but one that needs to be verified is local independence from network outages. A future evaluation plan is Random Forest, RBF-SVM, gradient-boosted trees, and compact MLP – there are no claims of AI performance. The next steps consist of independent reference labeled LPG and interfering exposures, longitudinal and cross-device tests, measured end-to-end response times, controlled network and power fault injection and resource profiling on an actual embedded controller. Until testing has been done and proper certification achieved, no guarantee of field detection accuracy, universal LPG selectivity, safety certification, or deployment readiness can be made.

References

[1] D. R. Pamungkas, A. Setiawan, and A. Widiyanto, "Design of Arduino-based LPG Gas Leak Detection Tool using MQ-6 Sensor," *Innovation, Technology, and Entrepreneurship Journal*, vol. 1, no. 1, pp. 21–27, 2024, doi: 10.31603/itej.10687.

[2] S. W. I. Made, A. A. N. Gunawan, P. I. Ketut, and S. I. Ketut, "Design and Manufacture of LPG Gas Leak Detection Based on Arduino Uno Using MQ-2 Sensor," *Asian Journal of Research and Reviews in Physics*, vol. 8, no. 1, pp. 27–32, 2024, doi: 10.9734/ajr2p/2024/v8i1155.

[3] A. N. Abdullah et al., "Correction Model for Metal Oxide Sensor Drift Caused by Ambient Temperature and Humidity," *Sensors*, vol. 22, no. 9, 3301, 2022, doi: 10.3390/s22093301.

[4] R. Rivera-Martinez et al., "Using metal oxide gas sensors to estimate the emission rates and locations of methane leaks in an industrial site: assessment with controlled methane releases," *Atmospheric Measurement Techniques*, vol. 17, pp. 4257–4290, 2024, doi: 10.5194/amt-17-4257-2024.

[5] J. Wörner, J. Eimler, and M. Pein-Hackelbusch, "Long-term drift behavior in metal oxide gas sensor arrays: a one-year dataset from an electronic nose," *Scientific Data*, vol. 12, 1628, 2025, doi: 10.1038/s41597-025-05993-8.

[6] B. Jena, S. K. Pradhan, R. Jha, S. Goel, and R. Sharma, "LPG gas leakage detection system using IoT," *Materials Today: Proceedings*, vol. 74, pt. 4, pp. 795–800, 2023, doi: 10.1016/j.matpr.2022.11.172.

[7] X. Dong and S. Han, "Real-Time Correction and Long-Term Drift Compensation in MOS Gas Sensor Arrays Using Iterative Random Forests and Incremental Domain-Adversarial Networks," *Micromachines*, vol. 16, no. 9, 991, 2025, doi: 10.3390/mi16090991.

[8] J. Lin and X. Zhan, "Sensor-Drift Compensation in Electronic-Nose-Based Gas Recognition Using Knowledge Distillation," *Informatics*, vol. 13, no. 1, 15, 2026, doi: 10.3390/informatics13010015.

[9] J. Schauer, J. Morsch, D. Arendes, A. Schütze, and C. Bur, "Interpretable Calibration Transfer and Drift Compensation for MOS Gas Sensors in Complex Gas Mixtures," *Sensors*, vol. 26, no. 14, 4595, 2026, doi: 10.3390/s26144595.

[10] Q. Yang et al., "MDFE-Net: A Meta-Learning Driven Dual-Branch Feature Extraction Network for E-Nose Sensor Drift Adaptation," *ACS Sensors*, vol. 11, no. 5, 2026, doi: 10.1021/acssensors.6c00538.

[11] R. Li, Z. Li, B. A. Kofi, J. Sun, Y. He, and M. Jiao, "A Comprehensive Review of Algorithms for Drift Compensation in Metal Oxide Semiconductor Gas Sensor Arrays," *Chemosensors*, vol. 14, no. 6, 143, 2026, doi: 10.3390/chemosensors14060143.

[12] K. Jiang et al., "Gas Sensor Drift Compensation Using Semi-Supervised Ensemble Classifiers with Multi-Level Features and Center Loss," *ACS Sensors*, vol. 10, no. 4, pp. 2906–2918, 2025,

ISSN:2302-8149
e-ISSN:2540-9719

doi: 10.1021/acssensors.4c03655.
[13] Zhengzhou Winsen Electronics Technology Co., Ltd., "MQ-6 MOS Flammable Gas Sensor," product documentation, accessed Oct. 1, 2026.
[14] Arduino, "Analog Input Pins," *Arduino Documentation*, accessed Oct. 1, 2026.
[15] Espressif Systems, "ESP8266 Hardware Design Guidelines," ver. 2.8, accessed Oct. 1, 2026.