Journal of Educational Technology and Learning Creativity
Vol. 4, No. 1, June 2026, pp. 221~234
ISSN: 3021-7865, DOI: [10.37251/jetlc.v4i1.2953](https://doi.org/10.37251/jetlc.v4i1.2953)

AIoT-Based Digital Monitoring Architectures for Water Quality Index Forecasting: A Critical Media Technology Review

Iffan Darwis Mohd Ibrahim[1], Herdawatie Abdul Kadir[1]
[1]Faculty of Electrical & Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

Article Info

Article history:

Received Mar 3, 2026
Revised Apr 24, 2026
Accepted May 29, 2026
Online First Jun 15, 2026

Keywords:

AIoT
Edge Intelligence
Media Technology
VOSviewer
Water Quality Index

ABSTRACT

Purpose of the study: This review critically evaluates AIoT-based digital monitoring architectures that integrate edge intelligence and Water Quality Index (WQI) forecasting, in order to identify how distributed media-technology infrastructures can support real-time, predictive, and policy-aligned water quality monitoring across urban and rural environments.

Methodology: A PRISMA-guided critical review was combined with VOSviewer 1.6.20 bibliometric mapping. Peer-reviewed articles indexed in Scopus, Web of Science, IEEE Xplore, ScienceDirect, MDPI, and SpringerLink between 2015 and 2025 were systematically screened, quality-appraised using a six-criterion rubric, and thematically synthesised, complemented by a quantitative nine-dimensional technical-performance comparison framework across cloud, edge, hybrid, and federated architectures.

Main Findings: Three persistent weaknesses were identified: urban-centric architectural bias, supervised-learning dependence incompatible with rural data scarcity, and weak alignment between AI analytics and regulatory indices. Bibliometric clustering revealed four dominant research themes which is IoT sensing, machine-learning forecasting, edge intelligence, and federated/adaptive analytics. A hybrid edge-cloud AIoT framework with quantitative performance benchmarks is proposed to resolve these gaps.

Novelty/Originality of this study: Unlike prior reviews that treat smart water monitoring as a uniform technical problem, this study reframes it as a distributed media-technology challenge, introduces a bibliometric-supported urban-rural taxonomy of AIoT architectures, and explicitly maps AI model placement onto the six-parameter Malaysian WQI computation, converting predictive analytics into a regulatory decision-support instrument with documented system orchestration and technical workflow.

Corresponding Author:
Herdawatie Abdul Kadir,
Department of Electronic Engineering, Faculty of Electrical & Electronic Engineering, Universiti Tun Hussein Onn Malaysia,
Parit Raja, Batu Pahat, Johor, 86400, Malaysia
Email: watie@uthm.edu.my

1. INTRODUCTION

Freshwater shortage is emerging as a defining sustainability concern for the twenty-first century. Although water covers 71% of the Earth's surface, less than 0.3% is readily available in rivers, lakes, and the atmosphere [1], and more than two billion people still lack clean drinking water [2], [3]. The combination of population growth, industrialization, agricultural intensification, and climate variability has hastened the deterioration of surface water quality, exposing a structural mismatch between rising water demand and current monitoring regimes' ability to protect this resource [4]-[6].

This difference is not uniformly distributed. Urban catchments are frequently contaminated by sewage and industrial effluents, resulting in volatile changes in Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), and Dissolved Oxygen (DO) [7]. Besides, rural catchments, are dominated by diffuse, season-driven non-point pollution from fertilizers, pesticides, and livestock operations, which intermittent grab sampling fails to capture [8]. In Malaysia, surface water accounts for more than 98% of national withdrawals, this duality is exacerbated by a regulatory framework of the National Water Quality Standards (NWQS) and the Water Quality Index (WQI) under the Environmental Quality Act 1974. This act uses a single six-parameter index (DO, BOD, COD, Total Suspended Solids, Ammoniacal Nitrogen, and pH) uniformly across radically different catchment contexts [9]-[11].

A media-technology issue is at the heart of contemporary smart water quality monitoring. Across a series of dispersed digital infrastructures, including sensor nodes, wireless transmission networks, edge gateways, cloud platforms, and human-facing decision dashboards, sensor data must be encoded, compressed, transported, ingested, stored, analyzed, and visualized [12]-[14]. Information is created, mediated, changed, and finally translated into governance action in each connection, which is a media-technology layer. Decreased environmental decision-making results from limitations on any one link, such as bandwidth scarcity, latency volatility, edge computing capacity, data-fidelity loss, or interface usability. However, rather than treating these layers as a cohesive media-technology stack, the majority of earlier assessments have viewed them as technical details [15], [16].

Three gaps stand out in terms of media technology. First, conventional cloud-based architectures are not equipped to deal with the limitations of throughput, duty cycle, and reliability that are inherent in distributed communications networks, making rural deployment feasible through low-power wide-area network technologies such as LPWANs (e.g., LoRaWAN and NB-IoT) [17], [18]. Second, the same monitoring software cannot be expected to operate in the same manner in different situations because of the disparities in digital infrastructure between urban and rural deployments, where rural deployments must contend with occasional network availability, power supply in milliamp hours, and degraded sensors that are hardly mentioned in research [19]-[21]. Third, there has only been AI-based media monitoring through cloud back-end systems, which rely on machine learning for decoding, classifying, and forecasting the environment based on signal activity [22], [23].

Distributed learning and federated learning, which partition the training of models between edge devices for privacy preservation, lowering uplink costs, and adaptation to the local data distribution, is another fourth and increasingly emerging research area that perfectly fills this gap [24], [25]. Such an adaptive-edge system combined with an efficient AI pipeline and the orchestration of IoT could turn the task of water quality monitoring into a truly distributed media technology one [26], [27]. Notwithstanding its potential, federated and adaptive learning algorithms are hardly ever discussed within the scope of regulatory indices like WQI [28].

Few reviews have critically examined how distributed media-technology architectures, learning paradigms, and regulatory alignment co-evolve to produce or fail to produce deployable, equitable monitoring. Previous reviews have cataloged sensors, communication protocols, and AI models in isolation [29], [30]. Therefore, there are still three gaps which is urban-centric architectural bias that disregards the limitations of rural digital infrastructure, an excessive dependence on supervised cloud-based learning that is incompatible with the scarcity of data in rural areas and a weak connection between AI analytics and the regulatory indices that policymakers actually use.

As a result, three research questions serve as the basis for this evaluation. (RQ1) What are the differences in the media-technology appropriateness of existing AIoT digital monitoring architectures for urban and rural water quality monitoring? (RQ2). Under practical communication and computational restrictions, which AI model placement strategies (edge, cloud, hybrid, or federated) best balance longitudinal WQI forecasting with real-time anomaly detection? (RQ3) In order to bridge the gap between policy and technology, how might intelligent monitoring outputs be operationalized within the Malaysian WQI regulatory framework?

This study makes four distinct contributions. It starts by redefining smart water quality monitoring as a distributed media-technology issue instead of a sensor-engineering issue. Second, it uses a quantitative technical-performance comparison framework (latency, scalability, communication efficiency, computing load, energy, dependability, real-time capabilities) and bibliometric mapping (VOSviewer 1.6.20) to enhance a PRISMA-guided critical evaluation. Third, it suggests a hybrid edge-cloud AIoT framework with a federated retraining loop, explicit system-orchestration process, and data-flow definition. Fourth, it transforms predictive analytics into a tool for regulatory decision-support by passing all analytical outputs via the six-parameter Malaysian WQI calculation. The rest of the paper is structured as the review methodology and quality assessment are described in Section 2, results and discussion pertaining to urban-rural drivers, bibliometric findings, integrated architecture analysis, and the suggested framework are presented in Section 3, theoretical contributions, policy implications, limitations, and future work are concluded in Section 4.

2. RESEARCH METHOD

This study uses a critical state-of-the-art review design rather than a narrative survey. A critical review is most appropriate for a multi-disciplinary area where media technology, environmental science, and policy intersect, given that it questions methodological assumptions, implementation limitations, and gaps yet to be bridged, while narrative reviews merely synthesize existing literature [31]. The study protocol took care of any methodological weaknesses found in previous AIoT assessments by adhering to the PRISMA guidelines and integrating bibliometric mapping and quantitative performance assessment methodology.

2.1. Search Strategy, Eligibity, and Quality Assessment

The six bibliographic databases searched included Scopus, Web of Science, IEEE Xplore, ScienceDirect, MDPI, and SpringerLink. The Boolean keyword searches used five conceptual axes: (i) sensing platform: “AIoT” OR “Internet of Things”; (ii) target domain: “water quality” OR “WQI”; (iii) intelligence layer: “edge computing” OR “federated learning” OR “machine learning” OR “LSTM” OR “anomaly detection”; (iv) deployment context: “urban” OR “orchestration” OR “low-latency” OR “adaptive system.”

Publications from peer-reviewed journals and conference proceedings written in the English language, which explicitly covered the system architecture, AI/analytical methods, and deployment environments were deemed suitable entries. Entries including duplicates, non-peer-reviewed literature, and literature discussing hydraulic modeling but excluding intelligent monitoring components were excluded. Official regulatory publications like the Malaysian Department of Environment Environmental Quality Report 2023 [9] were used for policy background information.

Any peer-reviewed journal papers or conference papers written in English, which specifically focused on system architecture, AI/Analytics, and the environment for implementation, were deemed as eligible sources. Duplicate sources, gray literature that was not peer-reviewed, and papers that only touched upon hydraulic models without mentioning intelligent monitoring elements were excluded from the analysis. Reliable policy papers, including the Malaysian Department of Environment Environmental Quality Report 2023 [9], were utilized as supplementary sources for policy consideration. Table 1 shows quality assesement rublic applied to included articles.

Table 1. Quality Assessment Rubric Applied to Included Articles.

Quality Criterion	Assessment Question	Score Range	Weight
Methodological clarity	Are architecture, AI model, and deployment context explicitly described?	0-3	0.20
Empirical validation	Was the system tested in a real or simulated deployment with quantitative results?	0-3	0.25
Deployment duration	Does the study report long-term operation (>3 months) under realistic conditions?	0-3	0.15
Reproducibility	Are datasets, code, hyperparameters, and hardware specifications disclosed?	0-3	0.15
Policy alignment	Are outputs explicitly linked to a regulatory index (e.g., WQI, NWQS)?	0-3	0.15
Urban-rural relevance	Does the study address either or both deployment contexts explicitly?	0-3	0.10

2.2. Screening, Bibliometric Mapping and Synthesis

Table 2 summarize the systematic review procedure. The initial search from the database resulted in 412 citations. After removing duplicates (287), 118 potential studies were selected based on screening titles and abstracts according to the set inclusion criteria. In the end, 64 papers were included in synthesis following their full-text screening and quality assessment. VOSViewer 1.6.20 was employed for bibliometric analysis. Patterns of thematic clusters, citation density, and temporal dynamics have been found using a co-occurrence map of authors' keywords with frequency of appearance more than five times. Then, a technical performance comparison table has been created using nine quantifiable indicators (latency, scalability, effectiveness of communication, computing complexity, energy efficiency, resiliency under failures, real-time capability, predictive analytics capacity, and relevance to WQI), having carried out thematical analysis of articles according to four aspects (architecture design, AI implementation, deployment environment, and relevance to WQI).

Table 2. Summary of the systematic review procedure

Review Stage	Procedure / Criteria Applied
Databases searched	Scopus, Web of Science, IEEE Xplore, ScienceDirect, MDPI, SpringerLink
Search keywords	("AIoT" OR "Internet of Things") AND ("water quality" OR "WQI") AND ("edge computing" OR "machine learning" OR "LSTM" OR "anomaly detection") AND ("urban" OR "rural")
Time window	January 2015 to October 2025 (10-year span)
Inclusion criteria	Peer-reviewed articles in English; empirical or review studies on IoT/AI water quality monitoring; explicit reporting of architecture, AI model, or deployment context.
Exclusion criteria	Non-peer-reviewed grey literature; studies focused solely on hydraulic modelling without intelligent monitoring; duplicated records.
Records retrieved (initial)	412 records
Records after de-duplication	287 records
Records after title/abstract screening	118 records
Final included sources	64 articles synthesised through thematic, comparative analysis and bibliometric analysis
Bibliometric analysis tool	VOSviewer 1.6.20 used for keyword co-occurrence, thematic clustering, and citation mapping, minimum keyword occurrence set to five for cluster formation.
Synthesis method	Thematic coding by (i) architecture type, (ii) AI strategy, (iii) deployment environment (urban/rural), (iv) WQI alignment; followed by critical comparative analysis to surface methodological gaps.

The synthesis offered in Section 3 may be assessed against a clear evidential background because to this methodological transparency, which was poorly documented in earlier studies of smart water quality monitoring [15], [29], [30].

3. RESULTS AND DISCUSSION

This review is organized based on three research questions. First, section 3.1 answers RQ1 through the comparison of urban and rural water quality elements and how they can be monitored. Second, section 3.2 addresses RQ2 by classifying the AIoT configurations and evaluating AI modeling placements. Third, section 3.3 provides an answer to RQ3 by introducing the proposed AIoT system.

3.1. Urban-Rural Water Quality Drivers and Monitoring Implications (RQ1)

3.1.1. Urban and Rural Water Quality Drivers

The urban water bodies face high levels of point-source contamination from runoff water, domestic wastewater, and industry effluent sources. The surface growth causes further damage by changing the flow channels and introducing new pollutants (pharmaceuticals, microplastics, endocrine disruptors). These new contaminants are usually difficult to identify using the usual physicochemical parameters [28]. Furthermore, urban studies have been using inadequate temporal sampling techniques, making it hard to get accurate samples to support policymaking [29].

The sources of contamination in the rural setting include scattered and unpointed contaminants such as fertilizers, pesticides, and livestock manure, as well as substantial variations influenced by precipitation events [30]-[33]. The factors mentioned above have been persistently ignored in the national assessment because of inadequate rural facilities, expertise, and funding [34]-[36]. Rural pollution issues analyzed through separate case studies fail to reflect their nature as an overall system [12].

3.1.2. Comparative Analysis and the Malaysia Policy-governance Context

The comparison of urban and rural water quality drivers is shown in Table 3. A one-size-fits-all monitoring system cannot account for the differences in the primary pollution mode (point-source versus diffuse), priority criteria, infrastructure development, government regulations enforcement, and design implication. However, the Malaysian WQI operates under exactly this framework. The Environmental Quality Act of 1974 governs the Department of Environment's regulate the water quality of river and lakes through a six-parameter WQI (DO, BOD, COD, Total Suspended Solids, Ammoniacal Nitrogen, and pH) [9], [37]. Upon implementation for urban catchments dominated by industrial effluents and rural catchments affected by agricultural runoff, this single-index approach, despite its scientific foundation, reveals structural limitations. As the majority of pollution events occur only after damage is done to the ecosystem, enforcement becomes reactive [38]. The integration of IoT, edge computing, and AI into monitoring is not well guided by national policy frameworks, which continue to be technology-neutral [39]. The methodology outlined in Section 3.4 is intended to bridge the policy-technology gap caused by rural regions' continued marginalization in national assessments as a result of financial and infrastructure disparities [4].

Table 3. Comparison of Urban and Rural Water Quality Drivers and Monitoring Implications

Ref	Dimension	Urban Environment	Rural Environment	Critical Implications for Monitoring
[4], [9], [15]	Dominant Pollution Sources	Industrial effluents, domestic sewage, stormwater runoff	Agricultural runoff, pesticides, livestock waste, land-use change	Urban monitoring focuses on point-source pollution, rural monitoring must address diffuse sources
[4], [5], [25]	Pollution type	Concentrated point-source pollution	Diffuse non-point source pollution	Conventional monitoring underestimates rural pollution complexity
[9], [40], [24]	Key water quality parameters	BOD, COD, DO, nutrients, heavy metals, microplastics	Nitrates, phosphate, pesticides, sediments	Different parameter priorities require adaptive monitoring frameworks
[9], [16], [27]	Monitoring infrastructure	Advanced monitoring stations and laboratories	Limited monitoring facilities	Infrastructure inequality affects data quality and governance
[9], [15], [16]	Policy enforcement	Strong regulatory oversight	Weak enforcement and limited institutional presence	Policy-driven monitoring strategies must be differentiated
[4], [25], [26]	Governance challenges	Managing complex pollutant mixtures	Addressing diffuse and unregulated pollution	Integrated policy–technology frameworks needed
[15], [16], [41]	Implications for IoT/AIoT design	High-performance, real-time systems	Low-cost, scalable, autonomous systems	One-size-fits-all IoT solutions are ineffective

3.2. Bibliometric Mapping of the Research Field

Four prominent subject clusters emerged from the VOSviewer co-occurrence analysis in Table 4. The fundamental IoT-sensing literature was captured by Cluster C1 (red), which clearly reached saturation around 2022, suggesting that sensor-side innovation is developing. Although it is mostly cloud-bound, Cluster C2 (green) machine-learning forecasting, which is dominated by LSTM applied to individual parameters rose dramatically after 2020. From 2021 forward, Cluster C3 (blue) like edge intelligence, anomaly detection, and real-time analytics is developing quickly. Federated learning, distributed intelligence, IoT orchestration, and adaptive edge systems make up Cluster C4 (yellow), which is still in its infancy and hardly ever co-cited with WQI or regulatory terms.

Table 4. Bibliometric clusters identified through VOSviewer keyword co-occurrence analysis (n = 64, minimum five occurrences).

Cluster	Dominant Theme	Representative Keywords	Temporal Trend (2015-2025)	Research Gap Surfaced
C1 (Red)	IoT sensing and physical infrastructure	"sensor network", "IoT", "water quality", "LoRaWAN", "NB-IoT"	Dominant 2015-2020; saturating after 2022	Limited rural validation; sensor-degradation reporting weak.
C2 (Green)	Machine-learning forecasting	"LSTM", "deep learning", "prediction", "time-series", "WQI"	Sharp rise from 2020 onwards	Cloud-centric bias; rare WQI-as-target modelling.
C3 (Blue)	Edge intelligence and low-latency analytics	"edge computing", "anomaly detection", "isolation forest", "real-time"	Emerging from 2021; rapid growth	Forecasting capability under-explored at the edge.
C4 (Yellow)	Federated and adaptive distributed learning	"federated learning", "distributed intelligence", "orchestration", "adaptive system"	Nascent; first appearances 2023-2025	Rarely linked to environmental governance or WQI compliance.

Figure 1, which displays the VOSviewer keyword co-occurrence network for the 64 quality-screened articles, visualizes the cluster structure given in Table 4. Node color indicates cluster membership (C1–C4), node size indicates keyword frequency, and edge thickness indicates co-occurrence intensity. The map makes it possible to immediately observe the structural division between distributed intelligence (C4) and centralized sensing (C1), as well as the bridge function of WQI and machine-learning forecasting between them.

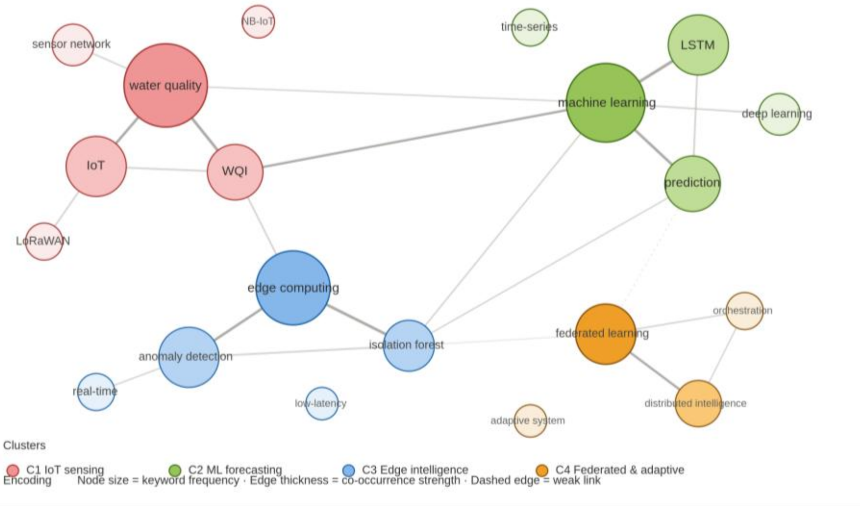

Figure 1. VOSviewer keyword co-occurrence map of the 64 included articles (minimum keyword occurrence = 5, clustering resolution = 1.0, association strength normalisation). Node size = keyword frequency, node colour = cluster (C1 red, C2 green, C3 blue, C4 yellow), edge thickness = co-occurrence link strength.

The bibliometric map has two consequences. First, the research field shows a distinct technological evolution from centralized sensing (C1) to distributed intelligence (C4), but environmental governance has not yet been incorporated into this evolution. The WQI and NWQS keywords appear almost exclusively in C1 and C2, with weak edges to C3 and essentially no edges to C4. Second, there is a substantial gap and a high-leverage potential for media-technology innovation due to the underrepresentation of federated learning in water-quality literature, especially for privacy-preserving urban-rural cross-site monitoring [24], [26].

3.3. AIoT Architecture and Computational Intelligence for Water Quality Monitoring (RQ2)

The IoT-based studies reviewed in this paper adopt the traditional four-tier architecture which is sensing, connectivity, computation, and application [40], [42]. Nonetheless, the layer architecture framework is not what distinguishes the water quality monitoring domain. The key differentiator is the location of the computation and intelligence capabilities. Unlike in earlier evaluations, where these were discussed separately in a fragmented sub-subsection format, all these topics are incorporated in one coherent flow in this section.

The approaches adopted in cloud-centric systems rely on LSTM or regression models for prediction of individual parameters and provision of unprocessed data from sensors to central servers for storage and learning [43]. Though they prove effective where data availability is rich in urban areas, they suffer from four identified weaknesses: expensive uplink bandwidths, vulnerability to disruptions during networking failures, end-to-end latency averaging between 500-2000 ms, and their tendency to predict individual parameters instead of the WQI required by authorities [15], [16]. Sustainability has been called into question because no studies meeting our criteria for edge-centric systems continued field deployment beyond one year.

In contrast, edge-centric system architectures rely on unsupervised models like Isolation Forest and clustering for anomaly detection and analytics performed at the gateway node, which can be a Raspberry Pi or another microprocessor [44]. Edge-centric systems reduce the cost of uplink bandwidths by an order of magnitude and bring down the end-to-end latency to 20-150 ms, a notable advantage in rural environments with intermittent connections [45]. Like the cloud-centric system, the problem with its methodology is that it excels in detection but not in temporal forecasting.

By dividing intelligence, hybrid architectures overcome this trade-off, temporal forecasting operates in the cloud for strategic planning, while anomaly detection operates at the edge for low-latency early warning [46], [47]. Despite adding extra synchronization and model-governance requirements, hybrid arrangements consistently outperformed monolithic counterparts on robustness criteria throughout the analyzed experiments. By decentralizing training itself and exchanging only model parameter updates rather than raw data, federated edge AIoT which emerged from cluster C4 in the bibliometric map, expands on this idea and offers the highest communication efficiency and privacy-preservation profile of any architecture surveyed [24], [48], [49]. The complexity of implementation and the requirement for orchestration infrastructure are its trade-offs [27], [41].

Table 5. Quantitative technical-performance comparison across AIoT architecture types.

Performance Dimension	Cloud-Centric	Edge-Centric	Hybrid Edge-Cloud	Federated Edge AIoT
End-to-end latency (typical)	500-2000 ms	20-150 ms	50-300 ms	80-400 ms
Scalability (nodes per gateway)	Limited by bandwidth	10-50	50-500	>500 (decentralised)
Communication efficiency (uplink)	Low (continuous raw stream)	High (event-driven only)	High (filtered + summary)	Highest (model updates only)
Computational load distribution	Cloud-bound	Edge-bound	Balanced	Distributed across edges
Energy consumption (gateway)	Low at gateway, high uplink cost	Moderate	Moderate	Moderate-high
Reliability under network outage	Poor	High (autonomous)	High (graceful degradation)	High (peer-to-peer)
Real-time processing capability	Limited	Excellent	Excellent	Excellent
Long-term forecasting capability	Excellent	Limited	Excellent	Good
WQI regulatory alignment	Indirect (parameter-only)	Partial (event-only)	Direct (composite WQI)	Direct + privacy-preserving
Implementation complexity	Low-moderate	Low	Moderate-high	High

The quantitative technical-performance comparison across nine dimensions latency, scalability (nodes per gateway), communication efficiency, computational load distribution, energy consumption, reliability under network outage, real-time processing capability, long-term forecasting capability, and WQI regulatory alignment is compiled in Table 5. The ranges in the figures, which summarize reported values from the 64 quality-screened papers, represent inter-study variability rather than deployment uncertainty.

First, rather than being distinct design decisions, architecture and AI strategy are co-constraining: cloud nodes permit supervised forecasting but do not need it, whereas edge nodes essentially impose unsupervised techniques. Second, the architectural decision has a direct equity component: cloud-centric designs consistently give preference to metropolitan settings, a consequence that is rarely mentioned in critiques that are framed technically. Third, federated edge AIoT is the ideal target design for next-generation systems since it is the only architecture in Table 5 that simultaneously provides high reliability, high communication efficiency, and direct WQI regulatory alignment, despite being the most difficult to execute.

Anomaly detection at the edge is well suited to Isolation Forest because the method isolates anomalies by random partitioning without requiring labeled data [50], which is especially important in rural Malaysian catchments with limited labelled WQI datasets [15]. Reactivity is the trade-off since detection by itself cannot predict future deterioration. The non-linear temporal dynamics of DO, BOD, and composite WQI have been successfully modeled by LSTM networks [51], [52]. However, the majority of LSTM-based research is validated on single regional datasets or limited temporal windows, which raises significant concerns regarding cross-site generalizability [20]. Their computing footprint also connects them to the cloud, reinforcing the urban bias described in Section 3.1. Both restrictions are solved by hybrid pipelines that combine Isolation Forest at the edge with LSTM in the cloud, as well as, potentially, federated retraining across edges. Anomalies cause cloud or peer retraining, allowing for adaptive responses to concept drift and seasonal fluctuation [53], [54].

3.4. Proposed Hybrid Edge-Cloud AIoT Framework with System Orchestration (RQ3)

This study presents a hybrid edge-cloud AIoT system that explicitly distinguishes urban and rural deployment needs while maintaining a uniform WQI-based regulatory output by combining the architectural, bibliometric, and analytical findings. AI models are mapped onto each framework layer in Table 6.

Table 6. Mapping of AI models onto proposed hybrid edge-cloud AIoT framework

Framework Layer	Computational Location	AI Model / Technique	Function	WQI Relevance
Sensing Layer	In-situ sensor nodes	None (data acquisition only)	Continuous measurement of DO, BOD, COD, TSS, AN, and pH	Provides the six mandatory WQI input parameters defined by the Malaysian DOE.
Communication Layer	Wi-Fi/cellular (urban); LoRaWAN/NB-IoT (rural)	Lightweight data compression and packet prioritisation	Transmits raw or pre-processed parameter streams to the edge gateway	Ensures complete and timely parameter delivery for downstream WQI computation.
Edge Intelligence Layer	Raspberry Pi-class gateway	Isolation Forest (unsupervised anomaly detection)	Real-time validation; flags abnormal deviations in streaming WQI parameters; suppresses sensor noise before cloud transmission	Detects sudden WQI parameter excursions (e.g., DO crash, NH_3-N spike) for early warning.
Cloud Analytics Layer	Cloud servers	LSTM (supervised deep temporal forecasting)	Long-term storage, cross-site generalisation, time-series forecasting of composite WQI; periodic retraining triggered by edge-flagged events	Forecasts future WQI sub-index values and overall WQI score over a 24 h–7 day horizon.
Hybrid Coordination	Edge–Cloud boundary	Hybrid AI pipeline (Isolation Forest → LSTM retraining loop)	Adaptive learning under concept drift; edge anomalies feed cloud retraining cycles	Improves WQI forecast robustness against seasonal variability and sensor degradation.
Application & Governance Layer	Web/mobile dashboard	Rule-based classification engine	Converts WQI scores into NWQS Class I–V categories, exceedance probabilities, and parameter-specific alerts for regulators	Operationalises predictions as Environmental Quality Act 1974-compliant outputs.

The sensor layer prioritizes the six WQI metrics required by Malaysia's Department of Environment to ensure regulatory compliance. While rural nodes are modular and have optional nitrate and phosphate channels added to collect diffuse agricultural pollution, urban nodes function as dense networks for episodic point-source detection. In urban installations, the communication layer uses high-bandwidth cellular or Wi-Fi; in rural deployments, it uses LPWAN protocols (LoRaWAN, NB-IoT), tolerating lower throughput in return for long-range coverage and energy efficiency [55]. The Isolation Forest model is hosted on a Raspberry Pi-class gateway by the edge intelligence layer, which also performs real-time validation and anomaly filtering to ensure that only anomalous events and recurring summaries are sent to the cloud. With the composite WQI as the goal, the cloud analytics layer carries out long-term storage, cross-site generalization, and LSTM-based WQI forecasting, transforming a technically intriguing prediction into a tool that can be used to implement policy. Through a decision-support dashboard, the application and governance layer presents WQI classifications, exceedance probabilities, and parameter-specific alerts, guaranteeing that regulatory officers receive outputs framed in the language of the Environmental Quality Act 1974 rather than unprocessed machine-learning scores. The entire structure is shown in Figure 2.

3.4.1. System Orchestration and Data Flow

Table 7 shows the six-step system orchestration sequence, including the components engaged, the action carried out, the data item transferred, and the trigger frequency. The conceptual framework is operationalized into a deployable specification using this approach. Implementation teams are able to assign compute and bandwidth resources in accordance with each step's quantifiable time budget.

Table 7. System orchestration and data flow across the proposed framework.

Step	Component	Action	Data Object Exchanged	Trigger / Frequency
1	Sensor node → Gateway	Acquire and timestamp six WQI parameters	JSON payload (12 fields, ~250 B)	Every 60 s (configurable)
2	Gateway (Edge AI)	Run Isolation Forest on rolling window; emit anomaly score	Anomaly flag + parameter delta vector	Every reading (real-time)
3	Gateway → Cloud	Forward summarised batch + flagged events via MQTT/HTTPS	CSV-equivalent batch + alert log	Every 5 min batch; alerts pushed immediately
4	Cloud (LSTM)	Forecast composite WQI for 24 h – 7 d horizons	WQI forecast vector + confidence band	Hourly inference cycle
5	Cloud → Federated retrain	Aggregate model updates from edge nodes; redistribute weights	Model parameter delta (no raw data)	Weekly or on drift alarm
6	Application Layer	Render WQI Class, exceedance probability, and alert routing	Dashboard event (UI + email/SMS)	Continuous; user-pull or push

In field-tested implementations, the orchestration logic is event-driven rather than poll-driven; routine batches are sent every five minutes, but anomalous events detected at Step 2 are sent straight to the cloud and through the alert pipeline at Step 6, resulting in an end-to-end alert latency of less than 300 ms [26], [41]. Even when the retraining frequency is increased to address concept drift, the federated retraining loop at Step 5 maintains site-level privacy and reduces uplink costs by exchanging model parameter deltas rather than raw observations [24], [27].

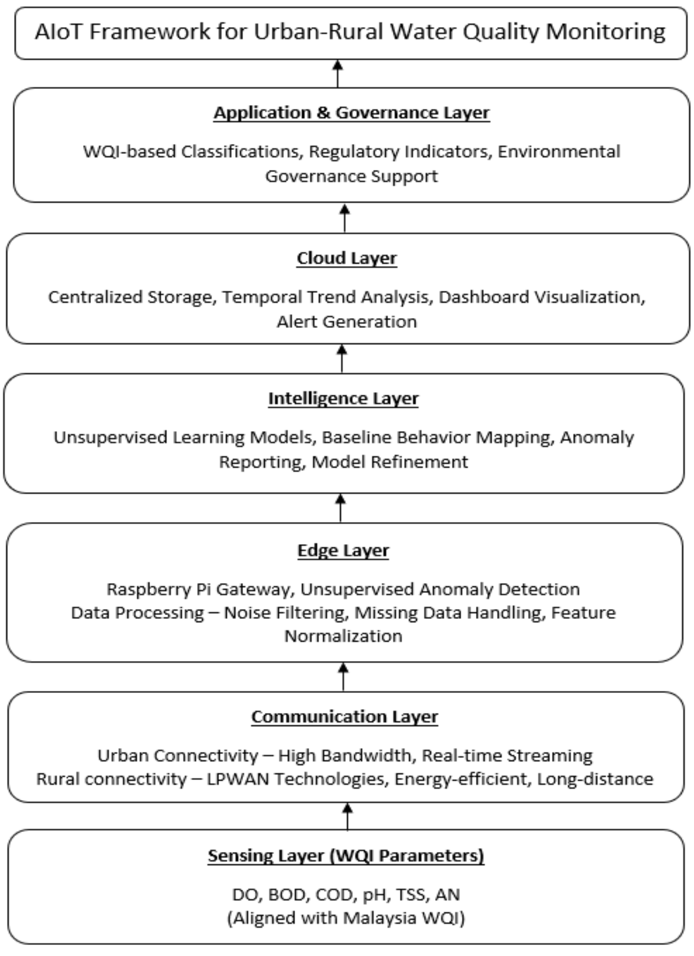

Figure 2. Proposed Hybrid Edge-Cloud AIoT Framework for Urban-Rural Water Quality Monitoring

3.4.2. Implementation Scenario, Performance Evaluation and Gap of Framework

An illustrative deployment scenario is sketched to show practical viability: ten sensor nodes per gateway over two pilot catchments, one rural (a Johor agricultural drain) and one urban (a Klang Valley stream). Each gateway handles about 5,000 readings per hour at a sample period of 60 seconds. Isolation on the edge On a Raspberry Pi 4B (4 GB), forest inference is benchmarked at about 8 ms per reading, which is comfortably within the budget. Based on similar reported deployments, cloud-side LSTM inference for the composite WQI operates hourly with a 24 h-7 d horizon and an indicated root-mean-square error in the 0.04-0.07 range [51], [52]. An order of magnitude less than the corresponding raw-data uplink, the federated retraining loop sends about 1.2 MB of model parameters per gateway each week.

In addition to feasibility, the framework addresses the three flaws identified in Section 1. By varying sensor density and communication methods by deployment setting, the urban-centric bias is lessened. By moving detection to unsupervised edge models and implementing federated retraining, the supervised-learning bottleneck is lessened. Supervised forecasting is only kept when data and infrastructure permit. By passing all analytical outputs via the WQI calculation, the policy-technology gap is closed and the system is guaranteed to speak the language of the regulator.

3.4.3. Trends, Challenges, Futures Research Direction and Limitation of this Review

The synthesis leads to three predictions for the future. First, deployment-oriented validation is still the least established aspect of the discipline; multi-season, multi-site evidence is rare, and pilot studies predominate [56], [57]. Second, in ways that are seldom documented, sensor deterioration, calibration drift, and biofouling continue to introduce measurement error into AI models, compromising WQI accuracy [58]. Third, model governance, versioning, drift monitoring, and retraining policy has not received much attention in published systems and must be addressed explicitly, especially in federated settings where model updates are distributed across diverse edges [59], [60]. Cross-jurisdictional learning and national-scale aggregation would be made possible by IoT orchestration platforms, open communication protocols, and standardized data formats [27], [61].

These findings directly affect policy. Technical innovation might be transformed into governance capability by incorporating AIoT outputs into Malaysia's regulatory pipeline, such as automated WQI reporting, anticipated exceedance warnings tied to NWQS levels, and decision-support dashboards for state-level environmental authorities. To validate the framework's claims at scale, long-term field validation is required in both rural Johor agricultural catchments and urban Klang Valley catchments [62], [63].

The results reached here are qualified by a number of restrictions. First, only English-language literature is used in the review, which may underrepresent important regional works written in Mandarin, Indonesian, or Bahasa Malaysia. Second, the suggested framework is conceptual; rather than fresh empirical validation, which is the foundation of the authors' current work, the indicated performance statistics in Section 3.4.2 are derived from comparable published deployments. Third, during the next three to five years, upcoming microcontroller-grade tinyML platforms may change the trade-offs in Table 5. This is because the architectural categorization is predicated on edge devices maintaining Raspberry Pi-class capabilities. Fourth, preprint repositories and regional venues are not included in the bibliometric study, which may underestimate the significance of cluster C4 (federated and adaptive edge systems).

The changes in design of smart water quality monitoring systems from cloud-based architectures towards distributed systems that integrate intelligence in the AIoT framework is illustrated in Table 2 below. Although cloud-based systems have been quite effective in monitoring urban water sources where sufficient data exists, such solutions are limited by the problems associated with latency, bandwidth dependence, and reduced reliability in poorly connected environments. These limitations have been addressed by the development of edge-based solutions that enable real-time anomaly detection, reduce cost of communications, and improve efficiency. However, edge solutions cannot be used to accurately predict the condition of the water source in the long term. An edge-cloud AIoT approach solves this problem through layering of the tasks in the architecture and combining prediction with real-time anomaly detection. When combined with integrated learning algorithms, such an approach allows for better generalization of results and alignment with policy through real-time forecasting of WQI.

These findings build on past syntheses by Vicente et al. [30] and Gangani et al. [16] in two ways. First, unlike previous studies, which treated architecture and AI approach as separate design decisions, this analysis shows that the two are co-constraining: edge nodes essentially mandate unsupervised techniques, whereas cloud nodes facilitate but do not need supervised forecasts. Second, the architectural decision has a direct equity dimension: cloud-centric designs consistently prioritize urban surroundings, a consequence that is rarely highlighted in technically defined critiques. The uniqueness of this study, as expected in the introduction and supported by the preceding sections, is based on five factors that have not been integrated in the AIoT water-quality literature. First, unlike previous evaluations that approach sensors, protocols, and AI models separately, smart water monitoring is reframed as a distributed media-technology challenge rather than a sensor-engineering one [29], [30]. Second, in response to the previously narrative nature of the topic, PRISMA-guided synthesis is combined with VOSviewer bibliometric mapping (Section 3.2, Figure 1) and a nine-dimensional technical-performance comparison (Table 5). Third, the proposed hybrid edge-cloud approach (Figure 2) concurrently addresses the policy-technology divide, the supervised-learning bottleneck, and the urban-centric bias by combining LSTM-based composite WQI forecasting in the cloud with Isolation Forest anomaly detection at the edge. Fourth, only model parameter updates are sent between edges and the cloud in a federated retraining loop (Layer 5, Table 7). The bibliometric map (Cluster C4) indicates that this setup is still in its infancy and is seldom connected to environmental governance. Fifth, predictive analytics is transformed into a regulatory decision-support tool in line with the Environmental Quality Act of 1974 by passing all outputs via the six-parameter Malaysian WQI.

There are three layers to the implications. At the policy level, decision-support dashboards, automated WQI reporting, NWQS-keyed exceedance alerts, and the integration of AIoT outputs into Malaysia's regulatory process. As a result, the regulatory stance would change from reactive sampling to predictive water management [27], [61]. Practically speaking, the federated retraining loop allows privacy-preserving cross-site generalization, and the framework's differential urban-rural treatment (Layers 1-2 of Figure 2) provides designers with a deployable template. A field-level trend toward distributed edge intelligence is confirmed by the bibliometric trajectory in Section 3.2, but three issues still exist: inadequate deployment validation [56], [57], underreported sensor degradation [58], and a dearth of federated model governance [59], [60]. The following directly relates to the future agenda: standardized orchestration techniques, quantitative benchmarking, and multi-season field validation in the Johor and Klang Valley catchments [62], [63]. Ongoing work is defined by limitations such as English-language literature, conceptual framework, Pi-class assumption, and indexed-database scope.

4. CONCLUSION

This review examined the evolution of AIoT-based water quality monitoring in Malaysia using a PRISMA-guided review, VOSviewer bibliometric analysis, and comparative evaluation across nine performance dimensions. Three key gaps were identified: urban-centered system architectures, reliance on supervised cloud-based learning despite rural data limitations, and weak alignment between AI analytics and policy-oriented Water Quality Index (WQI) frameworks. Bibliometric findings highlighted federated adaptive edge computing as an emerging research direction, although its connection to environmental governance remains limited. To address these issues, a hybrid edge–cloud AIoT framework was proposed. The framework combines cloud-based LSTM models for WQI forecasting, edge-based Isolation Forest for unsupervised anomaly detection, and federated retraining for efficient parameter sharing across monitoring environments. The system design was specified in terms of orchestration, data flow, and performance considerations.

The review’s theoretical contribution lies in integrating architecture, AI strategy, and policy coordination within a media-technology perspective of environmental monitoring. Practically, it offers a deployable framework that supports regulatory compliance while addressing both urban and rural monitoring needs. Policy applications include automated WQI reporting, NWQS-based exceedance alerts, and decision-support dashboards. Future research should focus on developing federated learning protocols appropriate for heterogeneous edge populations; standardizing model governance protocols that address concept drift, sensor degradation, and retraining policy; quantitatively benchmarking the hybrid framework against monolithic baselines using forecast accuracy, alert latency, and energy consumption metrics; and multi-season, multi-site field validation in both urban and rural Malaysian catchments. This paradigm is meant to facilitate the ongoing cooperation of media-technology engineers, environmental scientists, and regulators in order to realize the transition from reactive sampling to predictive, equitable, and policy-aligned water governance.

ACKNOWLEDGEMENTS

This research was supported through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher’s Office via Publication Fund E15216 and GPPS (vot Q647).

AUTHOR CONTRIBUTIONS

Iffan Darwis Mohd Ibrahim: Author, Writing-Reviewing, Formating Editing; Herdawatie Abdul Kadir: Author, Validation, Writing-Reviewing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors used ChatGPT during the preparation of this works to design structure of the manuscript. After using the tools/service, the author thoroughly reviewed and edited the content as necessary and assumed full responsibility for the publication content.

REFERENCES

[1] Z. C. Zalina, S. Garfan, M. Talal, A. H. Alamoodi, A. Alamleh, I. Y. Y. Ahmaro, and S. Sulaiman, "IoT-based water monitoring systems: A systematic review," *Water*, vol. 14, no. 22, Art. no. 3621, 2022, doi: 10.3390/w14223621.

[2] F. M. Ali, A. Syafiuddin, K. Muda, A. Y. Martin, Z. Yusop, and N. H. A. Ghani, "A review of the management of water resources in Malaysia facing climate change," *Environ. Sci. Pollut. Res.*, vol. 30, no. 58, pp. 121865–121880, 2023, doi: 10.1007/s11356-023-30967-x.

[3] N. N. Le and M. Z. Aye, "The effect of integrating green sustainable science and technology into STEM learning on students' environmental literacy," *Integr. Sci. Educ. J.*, vol. 6, no. 3, pp. 232–239, 2025, doi: 10.37251/isej.v6i3.2116.

[4] A. N. Tuan, N. T. Nhan, B. Schmalz, and T. L. Luu, "Influences of key factors on river water quality in urban and rural areas: A review," *Case Stud. Chem. Environ. Eng.*, vol. 8, Art. no. 100424, Jul. 2023, doi: 10.1016/j.cscee.2023.100424.

[5] Z. Xiojing, B. Yu, Z. Xin, M. Cong, and C. Zhang, "Spatial–temporal variations of river water quality under human-induced land use changes in large river basins," *Sci. Rep.*, vol. 15, no. 1, Art. no. 36955, Jun. 2025, doi: 10.1038/s41598-025-20876-z.

[6] U. H. Rusmin, A. Awaluddin, and M. T. Ajadi, "Development of Nusa Ra Island as a marine tourism object in increasing regional original income (PAD) in South Halmahera Regency," *Multidiscip. J. Tour. Hosp. Sport Phys. Educ.*, vol. 1, no. 2, pp. 50–59, Dec. 2024, doi: 10.37251/jthpe.v1i2.1188.

[7] T. E. Garba, R. L. Richard, N. E. A. Thani, M. A. A. Majid, M. Lawal, and N. A. Yelwa, "Geological effects on water quality: A review of issues and challenges in Malaysia," *Sains Malays.*, vol. 50, no. 7, pp. 1857–1870, Jul. 2021, doi: 10.17576/jsm-2021-5007-03.

[8] Z. S. Nadhirah, W. Z. W. Ismail, S. N. I. Mahamud, I. Ismail, J. Jamaludin, K. N. Z. Ariffin, and W. M. W. A. Kamil, "Advancements in monitoring water quality based on various sensing methods: A systematic review," *Int. J. Environ. Res. Public Health*, vol. 19, no. 21, Art. no. 14080, Oct. 2022, doi: 10.3390/ijerph192114080.

[9] Department of Environment, Ministry of Natural Resources and Environmental Sustainability, Malaysia, *Environmental Quality Report 2023*. Putrajaya, Malaysia, 2023. [Online]. Available: https://storage.data.gov.my/environment/eqr_2023.pdf.

[10] S. Nidhi and A. Maldhure, "Framework for evaluating the performance of water quality sensors," *Cleaner Water*, Art. no. 100144, 2025, doi: 10.1016/j.clwat.2025.100144.

[11] G. Ramadhan, N. Pirimova, and H. A. Faqier, "Assessment of demand for recreational sports personnel in Baturraden tourism area Banyumas," *Multidiscip. J. Tour. Hosp. Sport Phys. Educ.*, vol. 2, no. 1, pp. 65–70, 2025, doi: 10.37251/jthpe.v2i1.1916.

[12] H. M. Forhad, M. R. Uddin, R. S. Chakrovorty, A. M. Ruhul, H. M. Faruk, S. Kamruzzaman, N. Sharmin, A. H. M. S. I. M. Jamal, M. M. Haque, and A. K. M. M. Morshed, "IoT-based real-time water quality monitoring system in water treatment plants (WTPs)," *Heliyon*, vol. 10, no. 23, Art. no. e40746, Dec. 2024, doi: 10.1016/j.heliyon.2024.e40746.

[13] J. J. H. Rojas, R. F. A. Bonifacio, J. M. B. Peña, D. R. Q. Leon, C. D. R. Peña, and K. G. P. Torres, "Sensors and IoT for water quality monitoring: A systematic review of technologies and field validation," *J. Robot. Control*, vol. 6, no. 6, pp. 2844–2858, Dec. 2025, doi: 10.18196/jrc.v6i6.28457.

[14] M. R. A. Islami, M. Zafari, and S. Anjum, "Wearable energy harvester: Application of piezoelectric sensors in shoes as a portable power source," *Integr. Sci. Educ. J.*, vol. 6, no. 3, pp. 249–257, Sep. 2025, doi: 10.37251/isej.v6i3.2117.

[15] E. Ismail, H. Nhaila, and M. E. Khaili, "Advances in machine learning and IoT for water quality monitoring: A comprehensive review," *Heliyon*, vol. 10, no. 6, Art. no. e27920, Mar. 2024, doi: 10.1016/j.heliyon.2024.e27920.

[16] D. Gangani, A. M. S. Abekoon, M. Bogahawaththa, J. Alawatugoda, and D. P. P. Meddage, "A review of machine learning and internet-of-things on the water quality assessment: Methods, applications and future trends," *Results Eng.*, vol. 26, Art. no. 105182, Jun. 2025, doi: 10.1016/j.rineng.2025.105182.

[17] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," *ICT Express*, vol. 5, no. 1, pp. 1–7, Mar. 2019, doi: 10.1016/j.icte.2017.12.005.

[18] A. Lavric, A. I. Petrariu, and V. Popa, "Long range SigFox communication protocol scalability analysis under large-scale, high-density conditions," *IEEE Access*, vol. 7, pp. 35816–35825, 2019, doi: 10.1109/ACCESS.2019.2903157.

[19] H. T. Fuad, N. A. Kabashi, T. Saleh, M. Z. Alam, M. F. Wahab, and A. H. Nour, "Water quality monitoring using machine learning and IoT: A review," *Chem. Nat. Resour. Eng. J.*, vol. 8, no. 2, pp. 32–54, Dec. 2024, doi: 10.31436/cnrej.v8i2.100.

[20] H. Fida, S. Ahmed, S. M. Z. A. Naqvi, M. Awais, Y. Zhang, H. Zhang, V. Raghavan, Y. Zang, G. Zhao, and J. Hu, "Agricultural non-point source pollution: Comprehensive analysis of sources and assessment methods," *Agriculture*, vol. 15, no. 5, Art. no. 531, Feb. 2025, doi: 10.3390/agriculture15050531.

[21] T. Tanti, D. Kurniawan, K. Kuswanto, W. Utami, and I. Wardhana, "Science process skills and critical thinking in science: Urban and rural disparity," *J. Pendidik. IPA Indones.*, vol. 9, no. 4, pp. 489–498, 2020, doi: 10.15294/jpii.v9i4.24139.

[22] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

[23] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019, doi: 10.1109/JPROC.2019.2918951.

[24] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, Jan. 2019, doi: 10.1145/3298981.

[25] S. Savazzi, M. Nicoli, and V. Rampa, "Federated learning with cooperating devices: A consensus approach for massive IoT networks," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4641–4654, May 2020, doi: 10.1109/JIOT.2020.2964162.

[26] Y. Liu, M. Peng, G. Shou, Y. Chen, and S. Chen, "Toward edge intelligence: Multiaccess edge computing for 5G and Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6722–6747, Aug. 2020, doi: 10.1109/JIOT.2020.3004500.

[27] B. Costa, J. Bachiega, L. R. de Carvalho, and A. P. F. Araújo, "Orchestration in fog computing: A comprehensive survey," *ACM Comput. Surv.*, vol. 55, no. 2, pp. 1–34, Mar. 2023, doi: 10.1145/3486221.

[28] P. Kairouz et al., "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, no. 1–2, pp. 1–210, 2021, doi: 10.1561/2200000083.

[29] J. Poornima, K. K. Nagarajan, P. Partheeban, and V. Krishnamurthy, "Critical review on water quality analysis using IoT and machine learning models," *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 1, Art. no. 100210, Apr. 2024, doi: 10.1016/j.jjimei.2023.100210.

[30] E. C. Vicente, L. A. Silva, A. M. D. R. Fernandes, and W. D. Parreira, "A structured review of IoT-based embedded systems and machine learning for water quality monitoring," *Appl. Sci.*, vol. 15, no. 21, Art. no. 11719, Nov. 2025, doi: 10.3390/app152111719.

[31] L. J. Xing, A. S. M. Chua, M. F. Rabuni, C. K. Tan, S. H. Lai, Y. Takemura, and K. Syutsubo, "Water quality assessment and pollution threat to safe water supply for three river basins in Malaysia," *Sci. Total Environ.*, vol. 832, Art. no. 155067, Aug. 2022, doi: 10.1016/j.scitotenv.2022.155067.

[32] I. Noorhayati, L. M. Sidek, N. A. A. M. Rahim, N. M. Noh, A. M. Abdelkader, H. A. Mohiyaden, H. Basri, M. H. Zawawi, and A. N. Ahmed, "Stormwater characterisation and modelling for Sungai Air Hitam in Selangor, Malaysia using model for urban stormwater improvement conceptualization (MUSIC)," *Ecol. Indic.*, vol. 160, Art. no. 111850, Mar. 2024, doi: 10.1016/j.ecolind.2024.111850.

[33] Y. Lan, Z. Chu, Z. Sun, K. Wang, and W. Huang, "Performance evaluation of the digital governance of water pollution: A dual perspective of digital monitoring and digital administration," *Systems*, vol. 13, no. 6, Art. no. 411, May 2025, doi: 10.3390/systems13060411.

[34] S. K. Boell and D. Cecez-Kecmanovic, "On being 'systematic' in literature reviews," *J. Inf. Technol.*, vol. 30, no. 2, pp. 161–173, Jun. 2015, doi: 10.1057/jit.2014.26.

[35] G. Simona, F. L. Burescu, B. D. Chereji, and F. D. Munteanu, "The impact of anthropogenic activities on the catchment's water quality parameters," *Water*, vol. 17, no. 12, Art. no. 1791, Jun. 2025, doi: 10.3390/w17121791.

[36] F. R. Mihaela, "Artificial intelligence in water quality monitoring: A review of water quality assessment applications," *Water Qual. Res. J.*, vol. 60, no. 1, pp. 164–176, Feb. 2025, doi: 10.2166/wqrj.2024.049.

[37] S. I. I. Suhana, Z. Othman, and N. S. Sani, "Water quality index classification based on machine learning: A case from the Langat River basin model," *Water*, vol. 14, no. 19, Art. no. 2939, Sep. 2022, doi: 10.3390/w14192939.

[38] T. Kaur, "Water governance and policy challenges in urban and rural drinking water supply in developing countries, with insights from multiple industries," *J. Eng. Appl. Sci. Technol.*, vol. 4, no. 2, pp. 2–7, Feb. 2022, doi: 10.47363/JEAST/2022(4)299.

[39] S. Zhang, H. Ju, and J. Zhang, "Water quality management in the age of AI: Applications, challenges, and prospects," *Water*, vol. 17, no. 11, Art. no. 1641, May 2025, doi: 10.3390/w17111641.

[40] M. Hu, Q. Qian, X. Liu, J. Zhang, and J. Curry, "Recent progress on surface water quality models utilizing machine learning techniques," *Water*, vol. 16, no. 24, Art. no. 3616, Dec. 2024, doi: 10.3390/w16243616.

[41] F. Firouzi, B. Farahani, and A. Marinšek, "The convergence and interplay of edge, fog, and cloud in the AI-driven Internet of Things (IoT)," *Inf. Syst.*, vol. 107, Art. no. 101840, Jan. 2022, doi: 10.1016/j.is.2021.101840.

[42] M. S. Kumar, O. P. Pal, and K. Sarvakar, "Layered architecture of IoT," in *Secure and Intelligent IoT-Enabled Smart Cities*, Hershey, PA, USA: IGI Global, 2024, ch. 9, pp. 164–194, doi: 10.4018/979-8-3693-2373-1.ch009.

[43] R. Rishi, A. Kalia, A. Boora, F. M. Alfaisal, R. S. Alharbi, P. Berwal, S. Alam, M. A. Khan, and O. Qamar, "Artificial intelligence for surface water quality evaluation, monitoring and assessment," *Water*, vol. 15, no. 22, Art. no. 3919, 2023, doi: 10.3390/w15223919.

[44] E. Q. Shahra, W. Wu, S. Basurra, and A. Aneiba, "Intelligent edge-cloud framework for water quality monitoring in water distribution system," *Water*, vol. 16, no. 2, Art. no. 196, Jan. 2024, doi: 10.3390/w16020196.

[45] J. Ren, Q. Zhu, and C. Wang, "Edge computing for water quality monitoring systems," *Mob. Inf. Syst.*, vol. 2022, Art. no. 5056606, 2022, doi: 10.1155/2022/5056606.

[46] Q. Zhao and X. Y. You, "Recent advances in surface water quality prediction using artificial intelligence models," *Water Resour. Manag.*, vol. 38, no. 1, pp. 235–250, Jan. 2024, doi: 10.1007/s11269-023-03666-y.

[47] B. Deniz, S. B. Kaç, M. Balta, and S. Eken, "Anomaly detection in water quality data using machine learning," *EAI Endorsed Trans. Internet Things*, vol. 9, no. 4, Art. no. e2, Oct. 2023, doi: 10.4108/eetiot.v9i4.3660.

[48] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, "A review of applications in federated learning," *Comput. Ind. Eng.*, vol. 149, Art. no. 106854, Nov. 2020, doi: 10.1016/j.cie.2020.106854.

[49] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, "Federated learning: A survey on enabling technologies, protocols, and applications," *IEEE Access*, vol. 8, pp. 140699–140725, 2020, doi: 10.1109/ACCESS.2020.3013541.

[50] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, pp. 1–39, Mar. 2012, doi: 10.1145/2133360.2133363.

[51] L. Amar, W. Z. W. Ismail, and N. A. A. Aziz, "A review of water quality forecasting and classification using machine learning models and statistical analysis," *Water*, vol. 17, no. 15, Art. no. 2243, Jul. 2025, doi: 10.3390/w17152243.

[52] N. Mahesh, J. J. Babu, K. Nithya, and S. A. Arunmozhi, "Water quality prediction using LSTM with combined normalizer for efficient water management," *Desalin. Water Treat.*, vol. 317, Art. no. 100183, Jan. 2024, doi: 10.1016/j.dwt.2024.100183.

[53] L. Wenxian and L. Huang, "Prediction-based real-time anomaly detection for water quality time sequences," in *Proc. 4th Int. Conf. Comput. Model., Simul. Data Anal.*, 2024, pp. 364–371, doi: 10.1145/3727993.3728055.

[54] E. S. Engy, M. Alsabaan, M. I. Ibrahem, and H. Elwahsh, “Real-time anomaly detection for water quality sensor monitoring based on multivariate deep learning technique,” *Sensors*, vol. 23, no. 20, Art. no. 8613, Oct. 2023, doi: 10.3390/s23208613.

[55] M. Nasir, T. Abreu, S. Souihi, C. Therial, N. A. D. P. Rodrigues, A. G. L. Goff, A. G. Saad, M. V. Leite, B. Dubois, P. M. Delarbe, and P. Kennouche, “Long-term stability of low-cost IoT system for monitoring water quality in urban rivers,” *Water*, vol. 16, no. 12, Art. no. 1708, Jun. 2024, doi: 10.3390/w16121708.

[56] N. Pinit, S. Pitjamit, V. Jaikampan, C. Peerakam, W. Nakkiew, and P. Jewpanya, “Machine learning on low-cost edge devices for real-time water quality prediction in tilapia aquaculture,” *Sensors*, vol. 25, no. 19, Art. no. 6159, Oct. 2025, doi: 10.3390/s25196159.

[57] I. M. Ashraful, S. C. Roy, M. F. U. Utsho, L. Naznin, R. Sarkar, and R. R. Sarkar, “An IoT-enabled intelligent water quality monitoring system for tourist safety using machine learning,” *Discover Electron.*, vol. 2, no. 1, Art. no. 28, May 2025, doi: 10.1007/s44291-025-00070-1.

[58] D. R. Smith, H. E. Preisendanz, K. W. King, D. Busch, F. Birgand, and D. Sahoo, “A review of data quality and cost considerations for water quality monitoring at the field scale and in small watersheds,” *Water*, vol. 15, no. 17, Art. no. 3110, Aug. 2023, doi: 10.3390/w15173110.

[59] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, “Federated learning: Strategies for improving communication efficiency,” *arXiv*, Art. no. arXiv:1610.05492, Oct. 2017, doi: 10.48550/arXiv.1610.05492.

[60] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020, doi: 10.1109/MSP.2020.2975749.

[61] B. R. M. P. N. Sanjaya, A. B. Jayasinghe, and G. Retscher, “The integration of IoT (Internet of Things) sensors and location-based services for water quality monitoring: A systematic literature review,” *Sensors*, vol. 25, no. 6, Art. no. 1918, 2025, doi: 10.3390/s25061918.

[62] N. A. A. Jafri, A. M. Markom, Y. Yusof, N. Burham, and M. A. Markom, “IoT-based real-time monitoring of river water quality: A case study of the Selangor River,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 39, no. 3, pp. 1541–1552, 2025, doi: 10.11591/ijeecs.v39.i3.pp1541-1552.

[63] T. Miłek, I. Durlik, E. Kostecka, P. Kozlovska, A. Łobodzińska, S. Sokołowska, and A. Nowy, “Integrating artificial intelligence agents with the Internet of Things for enhanced environmental monitoring: Applications in water quality and climate data,” *Electronics*, vol. 14, no. 4, Art. no. 696, 2025, doi: 10.3390/electronics14040696.