Critical Review on Water Quality Analysis, Treatment and Monitoring Using Internet of Things and Artificial Intelligence for Drinking Water

Rajeshri U. Vaidya[1], Prof. Dr. Aashish A. Bardekar[2]

[1]Assistant Professor and PHD Research Scholar, Dept. of Computer Science and Engineering, Sipna College of Engineering and Technology, Amravati, Orchid id .https://orcid.org/0009-0005-0767-538X
[2]Professor and Ph. D. Research Supervisor, Dept. of Computer Science and Engineering, Sipna College of Engineering and Technology, Amravati, Maharashtra, India., Orchid https://orcid.org/0000-0001-8922-743

Abstract
The global challenge of ensuring safe drinking water access has intensified due to population growth, industrialization, and climate change. Conventional laboratory-based water quality assessment methods, while accurate, suffer from delayed analysis, limited sampling frequency, and high operational costs. This critical review synthesizes recent advancements in Internet of Things (IoT) and Artificial Intelligence (AI) technologies for drinking water quality analysis, treatment optimization, and real-time monitoring. IoT-enabled sensor networks facilitate continuous, high-frequency data acquisition on critical parameters—including pH (optimal 6.5–8.5), turbidity (<1–5 NTU), dissolved oxygen (>5 mg/L), and electrical conductivity (250–1500 μS/cm)—while machine learning algorithms including Random Forest, Gradient Boosting, Support Vector Machines, and Deep Neural Networks enable predictive forecasting, anomaly detection, and process optimization . The review examines the historical evolution from manual sampling to intelligent systems, critically evaluates the strong and weak points of current approaches, analyzes emerging trends such as digital twins and edge computing, and identifies persistent challenges including sensor drift, data interoperability, model interpretability, and regulatory acceptance. The findings indicate that while AI-IoT integration offers significant potential for transforming water management—demonstrating coagulant reductions of 10–22% in treatment plants and enabling early warning capabilities—substantial barriers remain for widespread operational deployment . Recommendations are provided for standardizing protocols, enhancing cross-site model generalization, and developing interpretable algorithms to bridge the gap between research innovation and practical implementation.

Keywords: Water quality monitoring, Internet of Things (IoT), Artificial Intelligence (AI), Machine Learning, Drinking water treatment, Predictive analytics, Smart water management, Real-time sensors, Anomaly detection, Digital twins

1. Introduction

Access to safe drinking water represents a fundamental human right and a cornerstone of public health. The World Health Organization reports that over 2 billion people worldwide lack access to safely managed drinking water services, with contamination from industrial effluents, agricultural runoff, urbanization, and inadequate sanitation infrastructure posing severe health risks . Traditional water quality monitoring relies on manual sample collection followed by laboratory analysis—a methodology that, despite its comprehensiveness, exhibits critical limitations: delayed information delivery, inability to sample entire distribution networks, high costs, and insufficient responsiveness to dynamic water quality fluctuations .

These limitations have catalyzed the search for innovative approaches capable of providing continuous, real-time, and cost-effective water quality assessment. The convergence of Internet of Things (IoT) and Artificial Intelligence (AI) technologies offers transformative potential for addressing these challenges. IoT enables dense sensor networks that continuously transmit water quality data from source to tap, while AI algorithms process these high-dimensional data streams to enable predictive forecasting, anomaly detection, and autonomous optimization of treatment processes .

The integration of these technologies has gained substantial research attention, with applications spanning source water monitoring, treatment plant optimization, distribution network surveillance, and household water quality assessment. However, despite significant progress in algorithmic development and sensor technologies, widespread operational deployment remains limited. This critical review aims to comprehensively examine the state-of-the-art in IoT-AI integration for drinking water quality management, critically evaluate the strengths and weaknesses of current approaches, analyze emerging trends, identify persistent challenges, and propose future research directions to bridge the gap between theoretical potential and practical implementation.

2. Definitions

Water Quality Analysis: The process of measuring and evaluating physical, chemical, and biological parameters of water to determine its suitability for drinking purposes. Key parameters include pH, turbidity, dissolved oxygen, electrical conductivity, total dissolved solids, temperature, and concentrations of specific contaminants such as heavy metals, nitrates, and microbial pathogens .

Internet of Things (IoT): A network of interconnected devices equipped with sensors, software, and communication technologies that collect, transmit, and exchange data. In water quality applications, IoT encompasses sensor nodes, wireless communication protocols (e.g., LoRaWAN, Zigbee, cellular networks), data storage infrastructure, and user interface applications .

Artificial Intelligence (AI): The simulation of human intelligence processes by computer systems. In water quality management, AI encompasses machine learning (ML), deep learning (DL), and reinforcement learning (RL) algorithms that learn from data to make predictions, detect patterns, and optimize decisions .

Machine Learning (ML): A subset of AI that enables systems to learn and improve from experience without explicit programming. Common ML algorithms applied to water quality include Random Forest, Support Vector Machines, Decision Trees, Gradient Boosting, and K-Nearest Neighbors .

Deep Learning (DL): An advanced subset of ML utilizing artificial neural networks with multiple layers to automatically extract hierarchical features from raw data. Popular architectures include Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Autoencoders, and Transformer models .

Water Quality Index (WQI): A numerical value that reflects overall water quality by integrating multiple parameters into a single score. Values below 50 indicate good quality, while values above 100 reflect poor conditions requiring treatment .

Predictive Analytics: The use of historical data and statistical algorithms to forecast future events. In water quality, this involves predicting pollutant levels, treatment chemical requirements, and contamination events .

Anomaly Detection: The identification of unusual patterns or outliers in data that deviate from expected behavior, indicating potential contamination, sensor malfunction, or system failure .

Digital Twin: A virtual representation of a physical water system that integrates real-time sensor data, historical information, and simulation models to enable predictive analysis, scenario testing, and operational optimization .

3. Need for IoT and AI Integration

The imperative for integrating IoT and AI into drinking water quality management arises from multiple converging factors:

Public Health Protection: Waterborne diseases remain a significant global health burden. Real-time monitoring and early warning systems enabled by IoT-AI integration can detect contamination events promptly, enabling rapid intervention and preventing disease outbreaks . Conventional laboratory analysis cannot provide the immediate notification necessary for timely public health responses.

Regulatory Compliance: Drinking water quality standards are becoming increasingly stringent worldwide. Water utilities must demonstrate continuous compliance through comprehensive monitoring. IoT-enabled systems provide the data density and audit trails necessary to meet regulatory requirements while reducing the administrative burden of manual sampling programs .

Operational Efficiency: Traditional water treatment relies heavily on operator experience and periodic jar tests for chemical dosing decisions. AI-driven optimization can reduce chemical consumption by 10–22% while maintaining or improving treatment efficacy, generating substantial cost savings and environmental benefits .

Resource Scarcity: Climate change and population growth are intensifying water scarcity. Efficient management requires precise understanding of water availability, quality, and consumption patterns. IoT networks provide the granular data needed for informed resource allocation decisions .

Emerging Contaminants: New pollutants, including pharmaceutical residues, microplastics, and per- and polyfluoroalkyl substances (PFAS), pose challenges for conventional treatment and monitoring. AI-enhanced detection and predictive models can improve identification and removal of these contaminants .

Infrastructure Resilience: Aging water infrastructure in many regions requires proactive maintenance and leakage detection. IoT sensors combined with AI analytics can identify anomalies indicative of pipe failures, enabling preventive maintenance and reducing water losses .

4. Aims and Objectives

4.1 Aim

This critical review aims to comprehensively evaluate the integration of Internet of Things and Artificial Intelligence technologies for drinking water quality analysis, treatment optimization, and monitoring, identifying strengths, limitations, current trends, and future directions to guide research and implementation efforts.

4.2 Objectives

1. To systematically review the historical evolution of water quality monitoring from conventional methods to IoT-AI enabled systems.
2. To critically evaluate the strong points of IoT-AI integration, including real-time monitoring, predictive capabilities, and operational efficiencies.
3. To analyze the weak points and persistent challenges, including sensor accuracy, data heterogeneity, model interpretability, and implementation barriers.
4. To examine current technological trends, including edge computing, digital twins, and multi-agent systems.
5. To synthesize evidence from peer-reviewed studies on quantitative improvements achieved through AI-driven water management.
6. To identify gaps in current research and propose future research directions.
7. To provide actionable recommendations for practitioners, policymakers, and researchers.

5. Hypothesis

Based on the synthesis of current literature, the following hypotheses are proposed:

1. IoT-enabled continuous monitoring systems provide significantly faster detection of water quality

deterioration compared to conventional laboratory-based sampling methods, enabling more timely public health interventions.
2. AI algorithms—particularly ensemble methods such as Random Forest and XGBoost—achieve superior predictive accuracy for water quality parameters compared to traditional statistical models, with documented R^2 values exceeding 0.95 in well-characterized systems .
3. The integration of AI-driven chemical dosing optimization in drinking water treatment plants reduces coagulant consumption by 10–22% while maintaining effluent quality standards, demonstrating measurable economic and environmental benefits .
4. Data quality issues, including sensor drift and calibration requirements, remain the most significant technical barrier to widespread IoT-AI deployment in water systems.
5. The gap between research-scale success and operational implementation is primarily attributable to challenges in model generalization across different water matrices, interpretability constraints, and integration with existing utility infrastructure.

6. Literature Search Strategy

A comprehensive literature search was conducted to identify peer-reviewed research, review articles, and conference proceedings relevant to IoT and AI applications in drinking water quality analysis, treatment, and monitoring.

Databases Searched: Web of Science, Google Scholar, ScienceDirect, Wiley Online Library, IEEE Xplore, and Scopus.

Search Keywords: The search employed a structured keyword framework integrating four complementary categories:

1. **IoT-related:** "Internet of Things," "wireless sensor networks," "IoT sensors," "real-time monitoring"
2. **AI/ML-related:** "artificial intelligence," "machine learning," "deep learning," "neural networks," "random forest," "support vector machines"
3. **Water quality-related:** "water quality monitoring," "drinking water," "water treatment," "chemical dosing," "water quality index"
4. **Application-related:** "anomaly detection," "predictive modeling," "process optimization," "early warning systems"

Time Frame: Publications from 2004 to 2025 were included to capture the evolution of AI applications while focusing on recent developments in the last five years .

Inclusion Criteria: Studies were included if they addressed (1) IoT sensor deployment for water quality parameters, (2) AI/ML algorithms for water quality prediction or treatment optimization, (3) integrated IoT-AI systems for drinking water applications, or (4) critical analysis of challenges and implementation barriers.

Exclusion Criteria: Studies focused exclusively on wastewater treatment, irrigation water quality, or marine environments were excluded unless they contained insights transferable to drinking water applications.

7. Research Methodology

7.1 Review Framework

This critical review follows a systematic approach to synthesizing and evaluating the literature, structured around four analytical dimensions:

Dimension 1—Unit Operation Categorization: Identification of water quality monitoring and treatment processes that require or benefit from IoT-AI integration across the drinking water supply chain, from source to tap .

Dimension 2—Algorithmic Capabilities: Comprehensive evaluation of AI methodologies—including ML, DL, and RL—for their capacity to optimize water quality analysis, treatment chemical dosing, and distribution network monitoring .

Dimension 3—Case-Based Validation: Critical benchmarking of reported performance metrics, including prediction accuracy (R^2, RMSE), chemical savings, and detection sensitivity across different applications .

Dimension 4—System Integration: Synthesis of findings into implementation frameworks and identification of barriers to practical deployment .

7.2 Data Extraction and Analysis

For each included study, the following information was extracted:

1. Application domain (source monitoring, treatment, distribution, or household)
2. IoT sensor configuration (parameters measured, communication protocol, power requirements)
3. AI/ML algorithms employed
4. Performance metrics and quantitative improvements reported
5. Identified challenges and limitations
6. Implementation context (research-scale, pilot, or operational)

7.3 Critical Evaluation Criteria

Studies were critically evaluated based on:

1. Methodological rigor and experimental design
2. Validation approaches (cross-validation, testing on independent datasets)
3. Comparability of performance metrics
4. Consideration of real-world operational constraints

5. Recognition of limitations and failure modes

8. History of Water Quality Monitoring and Treatment

8.1 Pre-20th Century: Sanitary Revolution

The recognition of water as a vector for disease transmission dates to the 19th century, with John Snow's 1854 cholera investigation in London establishing the link between contaminated water and disease. This period saw the emergence of rudimentary water quality assessment based on taste, odor, and visual clarity, along with the first chlorination trials in the late 19th century .

8.2 Early 20th Century: Establishment of Standards

The early 1900s witnessed the development of systematic water quality standards and laboratory testing methods. The US Public Health Service established the first drinking water standards in 1914, focusing on bacteriological quality. Chemical analysis methods evolved to detect inorganic contaminants, though testing remained labor-intensive and limited to centralized laboratories.

8.3 Mid-20th Century: Automation and Instrumentation

The post-World War II era brought significant advances in analytical instrumentation, including spectrophotometry, atomic absorption spectroscopy, and gas chromatography. These technologies enabled more sensitive and specific contaminant detection. However, sampling remained discrete, and analysis was conducted primarily in centralized laboratories, creating delays between sampling and results .

8.4 Late 20th Century: Electronic Monitoring

The advent of electronic sensors and telemetry in the 1970s–1990s enabled the first real-time water quality monitoring systems. Continuous analyzers for pH, turbidity, and residual chlorine became commercially available, though their high cost and maintenance requirements limited deployment to large treatment plants. This period also saw the emergence of water quality index methodologies to communicate overall quality status to stakeholders .

8.5 Early 21st Century: Digital Transition

The 2000s brought digitalization of water quality data management, with SCADA (Supervisory Control and Data Acquisition) systems becoming standard in water utilities. However, these systems remained largely reactive, with limited analytical capabilities. The focus was on data collection and visualization rather than predictive analytics or autonomous control.

8.6 2010s–Present: IoT and AI Revolution

The convergence of low-cost sensors, wireless communication, cloud computing, and AI algorithms has enabled a paradigm shift. IoT platforms now provide dense, real-time data from multiple points across the water supply chain, while AI algorithms extract actionable insights from these data streams. This period has seen the emergence of:

1. **IoT sensor networks** for distributed monitoring
2. **ML models** for WQI prediction and anomaly detection
3. **DL applications** for time series forecasting and pattern recognition
4. **AI-driven chemical dosing** for treatment optimization
5. **Digital twin development** for system simulation and scenario analysis

9. Strong Points of IoT and AI Integration

9.1 Real-Time Continuous Monitoring

The most significant advantage of IoT-enabled systems is the transition from discrete sampling to continuous monitoring. Traditional laboratory analysis provides instantaneous snapshots that may miss intermittent contamination events. IoT sensors can measure parameters such as pH, turbidity, temperature, and conductivity at frequencies ranging from seconds to minutes, enabling:

1. **Immediate detection of quality deterioration** through real-time alerts
2. **Capture of short-duration contamination events** that would be missed by periodic sampling
3. **Trend identification** through analysis of continuous time series data
4. **Early warning** for potential treatment issues or distribution system problems

Documented implementations demonstrate successful detection of turbidity increases and contamination agents under controlled conditions .

9.2 Predictive Analytics and Forecasting

AI algorithms excel at identifying patterns in complex, high-dimensional data. When applied to water quality time series, they enable:

1. **Short-horizon forecasting** of raw water quality changes, enabling proactive treatment adjustments
2. **Prediction of treatment chemical requirements** based on influent characteristics
3. **Water Quality Index (WQI) prediction** using multiple parameters, with ensemble models achieving R^2 values approaching 0.99
4. **Early warning of algal blooms** or other source water events

A study by Liu et al. demonstrated that AI-driven chemical dosing prediction enables proactive response to dynamic fluctuations in raw water quality, addressing the reactive limitations of conventional manual adjustments.

9.3 Treatment Process Optimization

AI integration in drinking water treatment plants has demonstrated substantial operational benefits:

1. **Chemical dosing optimization:** AI models predict optimal coagulant and oxidant doses based on real-time water quality data. Pilot studies report coagulant reductions of approximately 10–22% while maintaining settled turbidity below 1 NTU . This translates to significant cost savings and reduced environmental burden from chemical manufacturing and sludge disposal .
2. **Energy efficiency:** Optimization of pump scheduling and treatment processes through AI reduces energy consumption while maintaining water quality standards .
3. **Reduced chemical waste:** Precise dosing minimizes chemical overuse, reducing sludge generation and associated disposal costs.

9.4 Anomaly Detection and Contamination Early Warning

AI algorithms can detect subtle deviations from normal patterns that may indicate contamination, sensor drift, or system malfunction:

1. **Multiparameter anomaly detection** identifies events that may not trigger individual parameter exceedances
2. **Drift detection** identifies sensor calibration issues before they affect water quality assessment
3. **Pattern recognition** enables classification of anomaly types (e.g., biological contamination vs. equipment failure)

These capabilities are essential for: protecting public health, maintaining consumer trust, reducing response times to contamination events, and preventing costly waterborne disease outbreaks.

9.5 Scalability and Cost-Effectiveness

IoT platforms offer cost-effective monitoring solutions, particularly for resource-limited environments:

1. **Low-cost sensor nodes** can be deployed at multiple locations
2. **Low-power wireless communication** (e.g., LoRaWAN operating at ~29 W with telemetry ranges up to ~2 km) minimizes energy requirements
3. **Cloud-based data processing** eliminates the need for on-site computing infrastructure
4. **Open-source software platforms** reduce implementation costs

The documentation of IoT platforms powered by photovoltaic panels demonstrates the feasibility of self-sustaining monitoring nodes in rural or semi-rural environments .

9.6 Enhanced Distribution Network Management

Beyond treatment plant optimization, IoT-AI integration benefits distribution network management:

1. **Water quality monitoring** at multiple points in the distribution system, including households
2. **Leakage detection** through flow and pressure sensor analysis
3. **Water consumption monitoring** to identify unusual usage patterns indicating leaks or unauthorized use
4. **Residual chlorine management** to ensure disinfection throughout the network

The development of mobile applications providing water quality information to consumers represents an additional benefit, enhancing transparency and consumer engagement .

10. Weak Points and Persistent Challenges

10.1 Sensor Accuracy and Calibration Drift

Despite advances in sensor technology, significant limitations persist:

1. **Calibration drift** affects the accuracy of measurements over time, particularly for parameters such as free chlorine, dissolved oxygen, and nitrate
2. **Frequent maintenance and calibration requirements** make long-term unattended monitoring challenging
3. **Selectivity limitations** in low-cost sensors lead to cross-interference between parameters
4. **Lifetime limitations** of some sensors require frequent replacement, increasing operational costs

Free chlorine monitoring is identified as particularly challenging due to high sensitivity to changes in pH, temperature, flow, and pressure, while nitrate detection methods are either prohibitively expensive (UV spectrophotometric) or prone to malfunction (ion-selective electrodes) .

10.2 Data Quality and Heterogeneity

AI model performance depends critically on data quality:

1. **Manual sampling errors** in calibration and reference data affect model training
2. **Power supply interruptions** and network communication failures cause data gaps
3. **Severe weather events** and engineering maintenance cause sudden data loss
4. **Heterogeneous data standards** across different utility systems limit data sharing and model portability
5. **Inconsistent sampling frequencies** between parameters complicate temporal analysis

10.3 Model Interpretability and Regulatory Acceptance

The "black box" nature of many AI algorithms presents regulatory challenges:

1. **Limited interpretability** of deep learning decisions makes it difficult to explain model outputs to regulators and the public
2. **Regulatory acceptance** requires demonstrated understanding of model failure modes and uncertainty quantification
3. **Legal liability** concerns arise when autonomous systems make treatment decisions
4. **Trust barriers** among operators accustomed to empirical, experience-based control

10.4 Model Generalization Across Diverse Water Matrices

AI models trained on one water system often fail to generalize to others:

1. **Unique water chemistry** characteristics of each source limit cross-site applicability

2. **Seasonal variations** in source water quality affect model performance
3. **Climate variability** introduces patterns not captured in training data
4. **The "model adaptability" bottleneck** across different water matrices remains unresolved

10.5 Implementation and Integration Barriers

Technical success does not guarantee operational adoption:

1. **Legacy infrastructure** compatibility issues with modern IoT platforms
2. **Organizational resistance** from utility staff accustomed to traditional methods
3. **Insufficient domain expertise** in water utilities for developing and maintaining AI systems
4. **IT security vulnerabilities** associated with connected systems
5. **Cybersecurity threats** to water infrastructure are a growing concern

10.6 Limited Research-to-Operational Translation

A significant gap exists between research and operational deployment:

1. **Predominantly research-oriented** applications with limited operational validation
2. **Fragmented implementations** lacking coherent deployment frameworks
3. **Short-term pilot studies** without evidence of sustained performance
4. **Lack of benchmarking** and performance comparisons across different systems

10.7 Emerging Contaminants and Detection Limitations

AI algorithms can only detect what sensors measure:

1. **Emerging contaminants** (pharmaceuticals, microplastics, PFAS) lack practical real-time sensors
2. **Microbial pathogens** require specialized detection methods beyond standard physicochemical sensors
3. **Complex contaminant mixtures** create interactions not captured by individual parameter measurements

11. Current Trends

11.1 Digital Twins and Hybrid Modeling

Digital twin technology represents a significant advancement in water system management:

1. **Integration of real-time IoT data** with physics-based simulation models
2. **Scenario testing** capabilities for operational planning and emergency response
3. **Predictive maintenance** through simulation of infrastructure deterioration
4. **Operator training** using realistic virtual environments

Ensemble models and digital-twin workflows demonstrate high predictive performance for fouling evolution, flux decline, and WQI estimation . The integration of multi-agent systems and digital twins is identified as a future research hotspot and major challenge .

11.2 Edge Computing and On-Device AI

Edge computing addresses latency and bandwidth limitations:

1. **On-device AI processing** enables real-time analysis without cloud connectivity
2. **Reduced data transmission** minimizes communication costs and battery consumption
3. **Improved response times** for critical alerts
4. **Enhanced privacy and security** by keeping data local

Decision Tree Machine Learning algorithms running on central processing units (Raspberry Pi) demonstrate the feasibility of edge-based water quality management .

11.3 Deep Learning Advancements

DL techniques are increasingly applied to water quality challenges:

1. **LSTM networks** excel at time series forecasting of water quality parameters
2. **CNN architectures** enable image-based analysis of water quality (e.g., algal bloom detection)
3. **Transformer models** offer enhanced attention mechanisms for complex pattern recognition
4. **Autoencoders** enable unsupervised anomaly detection and data compression
5. **Deep Reinforcement Learning** shows potential for autonomous control optimization

DL applications are significantly more prevalent in water resources and distribution networks than in treatment and building supply systems, indicating areas for expanded research .

11.4 Multi-Parameter Sensor Integration

Modern IoT platforms integrate multiple sensor types:

1. **Physicochemical sensors**: pH, turbidity, temperature, conductivity, ORP, dissolved oxygen
2. **Flow and consumption sensors**: water usage patterns and network behavior
3. **Specialized sensors**: for specific contaminants (selective)
4. **Environmental sensors**: for contextual information (rainfall, temperature)

HydroScan Pro demonstrates integration of turbidity, pH, Total Dissolved Solids, and temperature sensors for comprehensive water quality assessment .

11.5 Explainable AI (XAI)

Addressing the interpretability challenge:

1. **Development of model-agnostic interpretation** techniques (e.g., SHAP, LIME)
2. **Feature importance analysis** to identify key water quality drivers
3. **Transparent model architectures** that enable understanding of decision-making
4. **Regulatory-compliant AI** with built-in interpretability

11.6 Integration of Water Quality Indices and AI

WQI methodologies are being enhanced through AI:

1. **Multiple WQI approaches** integrated with AI analysis (WAWQI, EWQI, etc.)
2. **WQI prediction** using ML models as a health risk communication tool
3. **Identification of critical parameters** for specific contamination scenarios
4. **Reduction of eclipsing effects** through advanced weighting approaches

11.7 Consumer-Facing Applications

Engaging consumers through water quality information:

1. **Mobile applications** providing real-time water quality information to consumers
2. **Consumer alerts** for contamination events
3. **Consumption tracking** enabling water conservation behaviors
4. **Trust building** through transparency

12. Discussion

12.1 Synthesis of Evidence

The literature consistently demonstrates the potential of IoT-AI integration to address critical limitations of conventional water quality monitoring. The documented quantitative improvements—including R^2 values of 0.93–0.99 for WQI prediction using XGBoost, coagulant reductions of 10–22%, and real-time detection of contamination events—support the hypothesis that these technologies can significantly enhance water management . However, the evidence base remains primarily research-oriented, with limited long-term operational validation.

The comparison of ML algorithms reveals consistent patterns: ensemble methods (XGBoost, Random Forest) generally outperform individual classifiers (Decision Tree, KNN) in prediction accuracy and generalization, while deep learning excels at pattern recognition in complex time series data . The selection of appropriate algorithms depends on application context, available data, and operational constraints—a finding that complicates the development of universal solutions.

12.2 Paradox of Progress

A fundamental paradox emerges from this review: despite significant technological advances, operational adoption remains limited. This discrepancy has multiple explanations:

First, the gap between controlled research conditions and operational reality is substantial. Studies reporting high prediction accuracy often use carefully curated datasets, validated against specific water sources with limited variability. Real-world water utilities must contend with seasonal changes, equipment failures, and the unpredictable challenges of actual operational environments.

Second, the regulatory and institutional framework has not kept pace with technological capability. Water utilities operate under conservative frameworks designed around traditional monitoring and control approaches. The absence of regulatory guidelines for AI-driven decisions creates liability concerns that deter implementation.

Third, the organizational and human dimensions are often overlooked. Water utility staff—predominantly civil and environmental engineers with limited computational training—are skeptical of systems they perceive as "black boxes." Success requires not just technical solutions but also organizational change management and training.

Fourth, the economic case for AI-IoT integration is not uniformly compelling. While chemical savings of 10–22% appear attractive, this must be balanced against sensor costs, ongoing maintenance requirements, and the potential consequences of system failures. The business case varies significantly with plant size, raw water variability, and regulatory stringency.

12.3 Contextual Considerations

The application of IoT-AI technologies for drinking water quality management is not a one-size-fits-all solution. The appropriateness and benefits vary significantly across contexts:

Developed vs. Developing Countries: Developed countries benefit from established infrastructure, funding, and technical expertise but face legacy system integration challenges. Developing countries may benefit more from low-cost, off-grid solutions but face significant institutional and capacity barriers.

Treatment Plant Scale: Large facilities can justify investment in sophisticated systems and dedicated data science expertise. Small facilities require simpler, more affordable solutions that may not require specialized staff.

Water Source Characteristics: Surface water sources with rapid quality fluctuations benefit more from real-time monitoring and predictive analytics than stable groundwater sources with minimal variability.

Regulatory Environment: Stringent regulatory regimes may facilitate or hinder adoption depending on whether regulations accommodate or prohibit AI-driven decisions.

12.4 Future Outlook

Despite persistent challenges, the trajectory toward IoT-AI integration in water management appears inevitable. The technology continues to improve, costs decrease, and institutional learning advances. The COVID-19 pandemic highlighted the value of early warning systems, reinforcing investment in monitoring technologies. Climate change-induced water quality variability increases the need for adaptive management approaches.

The transformation will likely follow a staged progression:

1. **Monitoring augmentation:** IoT sensors supplement conventional monitoring for specific high-risk applications
2. **Predictive analytics:** AI systems provide decision support for operators while maintaining human oversight

3. **Targeted automation:** Autonomous control for well-characterized, low-risk processes
4. **Systemic optimization:** Integrated AI-driven management of entire water supply systems

Each stage requires technical, institutional, and regulatory advancement, with Stage 1 and 2 representing the primary near-term opportunities.

13. Results

13.1 Quantitative Improvements Documented

Based on the synthesis of peer-reviewed studies, the following quantitative improvements have been documented:

Prediction Accuracy:

1. Multiple Linear Regression achieved $R^2 = 0.9992$ (RMSE = 0.338) for WQI prediction
2. XGBoost achieved $R^2 = 0.9998$ (training), 0.9726 (testing), and 0.9309 (cross-validation)
3. Ensemble models (RF, Extra Tree) demonstrate robust cross-validation performance

Treatment Optimization:

1. Coagulant reductions of approximately 10–22% while maintaining settled turbidity below 1 NTU
2. AI-driven chemical dosing optimization enables predictive, proactive response to water quality fluctuations

Monitoring Capabilities:

1. Real-time detection of turbidity increases and contamination agents under controlled conditions
2. Continuous measurement of pH (optimal 6.5–8.5), turbidity (<1–5 NTU), and electrical conductivity (250–1500 µS/cm)
3. Integration of flow monitoring with quality sensors for comprehensive water management

Energy Efficiency:

1. Low-power IoT sensing networks operating at ~29 W
2. Telemetry ranges up to ~2 km using LoRaWAN protocols

13.2 Algorithm Performance Comparison

Algorithm	Application	Performance Metric	Strength	Limitation
XGBoost	WQI Prediction	$R^2 = 0.93–0.99$	Excellent generalization	Complexity
Random Forest	WQI Prediction	$R^2 = 0.97–0.98$	Robust, handles non-linearity	Training time
MLR	WQI Prediction	$R^2 = 0.999$	Simple, interpretable	Linear assumptions
SVM	WQI Prediction	Strong generalization	Consistent performance	Tuning required
LSTM	Time series forecasting	Pattern recognition	Captures temporal dependencies	Data hungry
CNN	Image-based analysis	Feature extraction	Excellent for visual data	Compute intensive

13.3 Identified Research Gaps

1. **Cross-site model generalizability** across diverse water matrices
2. **Interpretable AI** algorithms suitable for regulatory acceptance
3. **Long-term operational validation** beyond pilot studies and controlled conditions
4. **Integration of multi-agent systems** and digital twins
5. **Standardized reporting protocols** for benchmarking AI performance
6. **Detection of emerging contaminants** through IoT-AI integration

14. Conclusion

This critical review has comprehensively examined the integration of Internet of Things and Artificial Intelligence technologies for drinking water quality analysis, treatment optimization, and monitoring. The convergence of these technologies represents a paradigm shift from reactive, discrete sampling toward predictive, continuous management, with documented improvements in monitoring capabilities, treatment efficiency, and early warning detection.

The strong points of IoT-AI integration are substantial and well-documented. Real-time continuous monitoring enables capture of quality fluctuations that conventional sampling cannot detect. Predictive analytics and machine learning algorithms—particularly ensemble methods—achieve high accuracy in forecasting water quality parameters and treatment chemical requirements. AI-driven optimization demonstrates quantifiable benefits including coagulant reductions of 10–22% and energy efficiency improvements. Anomaly detection algorithms provide early warning of contamination events, enhancing public health protection. These capabilities are increasingly accessible through low-cost sensor platforms and open-source software.

However, significant challenges persist. Sensor accuracy and calibration drift remain technical barriers to reliable long-term monitoring. Data quality issues, heterogeneity, and interoperability problems hinder model development and performance. Model interpretability and regulatory acceptance constrain deployment of autonomous AI systems. The gap between research-scale success and operational implementation is substantial, limited by inadequate model generalization, organizational resistance, and unclear business cases. Cybersecurity vulnerabilities of connected water infrastructure represent an emerging concern.

Current trends—including digital twins, edge computing, deep learning advances, and explainable AI—offer pathways to address some of these challenges. The integration of multi-agent systems and digital twins is identified as the most promising direction for next-generation smart water management. Standardized reporting protocols and cross-site benchmarking will be essential for validating the generalizability of AI performance claims.

The transformation of water quality management through IoT-AI integration is underway but remains in early stages. A staged progression from monitoring augmentation to systemic optimization is anticipated, with each stage requiring continued technical, institutional, and regulatory development. The ultimate success of this transformation will depend not only on technological advancement but also on organizational change, human capacity building, and the development of governance frameworks that balance innovation with protection of public health.

15. Suggestions and Recommendations

15.1 For Researchers

1. **Prioritize interpretable AI:** Develop explainable AI architectures that enable regulatory acceptance and operator trust. Model transparency is essential for operational deployment, not merely an academic consideration .
2. **Conduct long-term validation studies:** Extend beyond short-term pilot studies to evaluate performance over multiple seasons and under diverse conditions. Document failure modes and uncertainty bounds .
3. **Develop cross-site generalization methods:** Create transfer learning approaches that enable models trained on one water system to adapt to others. Establish common datasets for benchmarking .
4. **Address sensor limitations:** Research new sensor technologies for emerging contaminants (microplastics, PFAS, pharmaceuticals). Improve existing sensor selectivity, stability, and calibration requirements .
5. **Standardize reporting:** Adopt consistent metrics and reporting protocols for AI performance claims. Include both accuracy measures and operational constraints .
6. **Integrate multi-agent and digital twin approaches:** Research the integration of autonomous agents with virtual representations of physical systems .

15.2 For Practitioners

1. **Assess readiness and business case:** Evaluate specific needs, source water characteristics, and existing capabilities before pursuing AI-IoT integration. Large, variable surface water treatment plants offer the clearest early opportunities .
2. **Start with decision support:** Implement AI systems as decision support for operators rather than fully autonomous control. Build confidence through demonstrated reliability before automating critical decisions.
3. **Address data quality first:** Ensure reliable sensor data through careful selection, regular calibration, and robust data management. "Garbage in, garbage out" is a critical constraint .
4. **Invest in capacity building:** Develop staff skills in data analysis and AI interpretation. Success requires organizational change, not just technology acquisition.
5. **Plan for cybersecurity:** Implement robust security measures for connected water infrastructure. The integration of IoT creates new attack surfaces requiring protection .
6. **Engage stakeholders:** Build communication protocols to keep regulators and consumers informed. Transparency builds trust and facilitates adoption .

15.3 For Policymakers

1. **Develop regulatory guidelines:** Establish frameworks that accommodate AI-driven water quality decisions while maintaining accountability. Include standards for model validation, failure analysis, and human oversight .
2. **Create standards and protocols:** Develop technical standards for sensor accuracy, data formats, and AI performance reporting. Interoperability enables economies of scale .
3. **Support innovation funding:** Dedicate resources for demonstration projects and scaling of successful pilots. Public sector leadership can stimulate private sector development.
4. **Address rural and resource-limited contexts:** Target development of low-cost, off-grid solutions for underserved communities. Ensure that technological advancement benefits all populations .
5. **Facilitate data sharing:** Create mechanisms for secure, anonymized data sharing to enable model development without compromising privacy or security.
6. **Integrate water quality with other infrastructure:** Consider smart water as part of broader smart city initiatives. Integration with other urban systems can enhance benefits and share costs.

16. Future Scope

16.1 Short-Term (1–5 Years)

Multi-parameter Sensor Development: The development of practical, low-maintenance sensors for emerging contaminants including microplastics, pharmaceutical residues, and PFAS will enable AI models to address a broader range of contaminants. Priority should be given to parameters that are both toxicologically significant and technically feasible for real-time measurement.

AI-Enabled IoT Platforms: Integration of on-device AI processing at the sensor node level will reduce latency, enable detection even without cloud connectivity, and enhance privacy. Edge computing architectures will become standard for time-critical applications such as emergency contamination alerts.

Standardization and Interoperability: Industry-wide data standards will emerge, enabling model sharing and benchmarking across utilities. Water utilities will increasingly adopt common data formats, reducing implementation costs and enabling AI models trained on multiple datasets.

Decision Support Systems: Utilities will increasingly implement AI-based decision support tools for chemical dosing optimization, with the transition from "alerts" to "recommendations" to build operational confidence prior to autonomous control.

16.2 Medium-Term (5–10 Years)

Digital Twin Systems: Full implementation of digital twins representing entire water supply systems from source to tap. These will integrate real-time sensor data, historical information, and simulation models for comprehensive operational optimization and scenario testing .

Predictive Treatment Control: Widespread adoption of AI-driven chemical dosing control for routine operations, with human oversight for non-routine conditions and emergency management. Autonomous control will be accepted for well-understood, low-risk treatment steps with robust model validation.

Integrated Urban Water Management: IoT-AI integration will extend beyond drinking water to include stormwater, wastewater, and water reuse systems, enabling holistic urban water management and circular economy approaches.

Smart Water Infrastructure: Smart sensors embedded in water distribution networks will become standard infrastructure, enabling real-time leakage detection, water quality monitoring, and predictive maintenance of aging water assets.

16.3 Long-Term (10+ Years)

Autonomous Water Systems: Fully autonomous water management systems capable of self-optimization, self-healing, and learning from operational experience. Advanced multi-agent reinforcement learning will enable systems that adapt to changing conditions without human intervention.

Cyber-Physical Water Security: Integration of water quality monitoring with broader public health surveillance systems, enabling real-time tracking of waterborne disease outbreaks and environmental health risks.

Global Water Monitoring Networks: Interconnected monitoring systems enabling cross-border water quality sharing, early warning for transnational water quality events, and global databases for model training on diverse water chemistries.

Climate-Adaptive Water Management: AI systems that incorporate climate models to predict long-term changes in water quality and adapt treatment strategies accordingly. These will enable proactive rather than reactive adaptation to climate change impacts.

Universal Access: The deployment of low-cost, autonomous monitoring systems in underserved regions globally, enabling every community to access real-time drinking water quality information. The affordability and robustness of IoT platforms will be essential to achieving this goal .

17. References

[1] P. Dutta and S. Sarma, "Smart Water Management: Role of IoT, AI, and Machine Learning in Water Quality Monitoring," *Water Environment Research*, 2025.

[2] Y. Liu, J. Ren, Y. Liu, W. Yu, Z. Lu, X. Li, Y. Yang, and Z. Zhou, "Artificial intelligence in chemical dosing for drinking water treatment: A systematic review of algorithmic applications, implementation frameworks, and future challenges," *ScienceDirect*, 2025.

[3] "HydroScan Pro: Advanced IoT Water Level Monitoring and Smart Irrigation," *IEEE Xplore*, 2025.

[4] A. Das, "Evaluation and prediction of surface water quality status for drinking purposes using an integrated water quality indices, GIS approaches, and machine learning techniques," *Desalination and Water Treatment*, 2025.

[5] "Smart Water Management: Role of IoT, AI, and Machine Learning in Water Quality Monitoring," *Dimensions for DTIC*, 2025.

[6] "Artificial Intelligence Applications in Water Treatment Systems: A Mini Review," *CiNii Research*, 2025.

[7] "A wireless sensor network IoT platform for consumption and quality monitoring of drinking water," *Discover Applied Sciences*, Springer, 2024.

[8] "Surface water quality assessment and its evaluation of potential pollution risks for drinking purposes employing water quality indices and various machine learning techniques," *ScienceDirect*, 2025.

[9] "Smart materials and AI–IoT platforms for drinking-water treatment: Recent progress, operational efficiency, and risk-management perspectives," *Dimensions for DTIC*, 2025.

[10] Y. Bai, B. Geng, H. Yan, T. Tao, J. Hao, and K. Xin, "Building smart urban water systems from source to tap: Applications and challenges of deep learning," *ScienceDirect*, 2025.

[11] "Edge based Device using Machine Learning for Water Quality Management in a Smart Campus," *IEEE Xplore*, 2024.

[12] R. Sandhu, A. Arora, H. K. Channi, G. S. Cheema, and K. Lakhwani, "Analysis of water quality for sustainable resource management using machine learning techniques," *Taylor & Francis*, 2025.

[13] Ahmad, et al., "Occurrence, source identification and potential risk evaluation of heavy metals in sediments of the

Hunza River and its tributaries," *Environmental Technology & Innovation*, 2020.
[14] Belkhiri, et al., "Spatial analysis of groundwater quality using self-organizing maps," *Groundwater for Sustainable Development*, 2018.
[15] Das, A., "Multivariate statistical approach for the assessment of water quality of Mahanadi basin, Odisha," *Materials Today: Proceedings*, 2022.
[16] Das, A., "Surface water quality evaluation of Mahanadi and its Tributary Katha Jodi River," *Applied Water Science*, 2025.
[17] Das, A., "A data-driven approach utilizing machine learning and GIS-based time series analysis for water quality assessment in Mahanadi River Basin," *Discover Sustainability*, 2025.
[18] Frincu, et al., "Artificial intelligence in water quality monitoring: a review of water quality assessment applications," *Water Quality Research Journal*, 2025.
[19] Gandhimathi, et al., "Comprehensive river water quality monitoring using convolutional neural networks and gated recurrent units," *Journal of Environmental Management*, 2024.
[20] Gao, et al., "Hydrogeochemical characterization and quality assessment of groundwater based on integrated-weight water quality index," *Journal of Cleaner Production*, 2020.
[21] Liu, J., et al., "Artificial Intelligence in Water Quality Monitoring: A Review of Machine Learning Applications," *Water Research*, 2023.
World Health Organization, "Drinking-water quality guidelines," 4th ed., 2017.