International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 04 April-2026 | Impact Factor: 3.5

A Hybrid Air-Water Pollution Monitoring System Using Outlier Detection And Adaptive Feature Selection For Aqi-Based Environment Assessment

SRIKANTH M

NAVINRAJ S

GOKULRAJ S

KUMAR S

Department of Computer Science and Engineering

The Kavery Engineering College

Mecheri

Salem-636453.

(Affiliated to Anna University Chennai, Approved by AICTE, New Delhi)

Guided By

REVATHI R M.E

H.O.D: Dr. M. Balamurugan M.E., Ph. D

How to Cite this Article:

M, S., S, N., S, G., S, K. & Balamurugan, M. (2026). A Hybrid Air-Water Pollution Monitoring System Using Outlier Detection And Adaptive Feature Selection For Aqi-Based Environment Assessment. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04). https://doi.org/10.55041/ijcope.v2i4.596

https://doi.org/ 10.55041/ijcope.v2i4.596

ABSTRACT: Air and water quality are vital indicators of environmental health, significantly affecting ecosystem stability and human well-being. While deep learning (DL) has shown strong predictive capabilities, its dependence on large-scale datasets, high computational complexity, and limited interpretability restricts its adoption in regulatory decision-support scenarios. This study proposes an intelligent rule-based framework incorporating an Outlier Detection and Removal Algorithm (ODRA) to enhance data reliability and a Threshold-Aware Feature Selection Algorithm (TAFSA) to identify influential parameters. The framework provides an accurate, transparent, and explainable alternative to complex black-box models for environmental monitoring.

1.INTRODUCTION

Environmental pollution, driven by industrialization and urbanization, has created critical global challenges for public health and ecological balance. Traditional monitoring methods often fail to handle large volumes of data or provide timely assessments. While modern DL models offer accuracy, their "black-box" nature makes it difficult for policymakers to justify regulatory decisions based on their outputs.

Environmental pollution, driven by industrialization and urbanization, has created critical global challenges for public health and ecological balance. Traditional monitoring methods often fail to handle large volumes of data or provide timely assessments. While modern DL models offer accuracy, their "black-box" nature makes it difficult for policymakers to justify regulatory decisions based on their outputs.

2.LITERATURE REVIEW

The research builds upon existing studies in environmental informatics:

Sensor Reliability: Liu et al. (2018) highlighted the frequency of sensor malfunctions and proposed hybrid diagnostic frameworks combining SVM with decision trees.

Explainable AI (XAI): Kabir et al. (2020) and Zisad (2023) explored combining belief rule bases with machine learning to balance accuracy with transparency.

Fuzzy Reasoning: Aghaarabi et al. (2014) demonstrated that fuzzy logic-based systems effectively handle the inherent uncertainty in environmental data.

Feature Optimization: Elabd et al. (2026) utilized Particle Swarm and Grey Wolf Optimization to improve AQI classification, though noted high computational costs.

PROPOSED SYSTEM ARCHITECTURE:

The proposed framework addresses the limitations of existing systems, such as sensitivity to noisy data and the inclusion of irrelevant features.

3.1. OUTLIER DETECTION AND REMOVAL ALGORITHM (ODRA)

ODRA acts as a preprocessing filter. It identifies and removes abnormal readings caused by sensor faults or recording errors, ensuring the model is trained on consistent, high-quality data.

3.2. THRESHOLD-AWARE FEATURE SELECTION (TAFSA)

Instead of processing all available parameters, TAFSA evaluates each variable against predefined environmental thresholds (e.g., pH, PM levels). This reduces computational overhead while maintaining regulatory alignment.

3.3. LINGUISTIC VARIABLE-BASED CLASSIFICATION (LVBCA)

This algorithm maps numerical pollution values to linguistic categories:

- **Low:** Minimal pollution.
- **Moderate:** Acceptable levels but requiring monitoring.
- **High/Severe:** Dangerous levels requiring intervention.

4. SYSTEM REQUIREMENTS

- **Software:** Python (Spyder/VS Code IDE), Tkinter for GUI, NumPy for numerical operations, and CSV for structured data storage.
- **Hardware:** Intel Core i3/i5/i7 processor, 8GB RAM minimum, and 256GB SSD storage.

5. METHODOLOGY AND IMPLEMENTATION

The system follows a modular architecture:

1. Data Collection: Gathering air (CO, NO_2, NOx) and water (pH, Turbidity, Sulfate) parameters from monitoring sources.
2. Preprocessing: Applying ODRA to cleanse the dataset.
3. Optimization: Using TAFSA to select relevant indicators.
4. Inference: Using rule-based reasoning to determine the Air Quality Index (AQI) or Water Quality Index (WQI).

6. RESULTS AND DISCUSSION

The system was tested through a real-time GUI:

- **Air Quality Testing:** An input averaging 77.75 was classified as "Moderate" (range 51–100), with the interface dynamically updating to a yellow background.
- **Water Quality Testing:** An input index of 2488.14 exceeded the maximum safe thresholds and was classified as "Critical," triggering a red visual alert.

These tests confirm the system's ability to provide immediate, interpretable assessments suitable for regulatory use.

7. CONCLUSION AND FUTURE SCOPE

he proposed framework successfully demonstrates that rule-based models can offer high accuracy and superior interpretability compared to complex black-box approaches. By integrating ODRA and TAFSA, the system achieves computational efficiency and better alignment with environmental standards.

Future Work: Future developments will focus on integrating real-time IoT sensor networks for live streaming data and expanding the rule base to include emerging pollutants.

REFERENCES:

- Liu, S., et al. (2018). "Fault diagnosis of water quality monitoring devices." *IEEE Access*..
- Kabir, S., et al. (2020). "Integrated belief rule base and deep learning." *Sensors*..
- Aghaarabi, E., et al. (2014). "Comparative study of fuzzy rule-based approaches." *Stochastic Environmental Research*..
- Talagala, P. D., et al. (2019). "Feature-based procedure for detecting technical outliers." *Water Resources Research*..