Lightweight Adaptive Data Rate Adjustment for Cost-Aware Industrial IoT Monitoring Systems

1st Aryan Morteza
School of CSEE, University of Essex
aryan.morteza@essex.ac.uk

2nd Mohammad Hossein Anisi
School of CSEE, University of Essex
m.anisi@essex.ac.uk

3rd Morteza Varasteh
School of CSEE, University of Essex
m.varasteh@essex.ac.uk

4th Faiyaz Doctor
School of CSEE, University of Essex
fdocto@essex.ac.uk

5th Mark Hadaway
Lubron UK Ltd. Colchester, UK
mark.hadaway@lubron.co.uk

6th Andy Dowell
Lubron UK Ltd. Colchester, UK
andy.dowell@lubron.co.uk

***Abstract*—The rapid proliferation of Industrial IoT (IIoT) systems has intensified the demand for intelligent data transmission strategies that balance monitoring fidelity with stringent energy and bandwidth constraints. This paper introduces a lightweight adaptive transmission framework designed for edge devices operating under strict cost and communication constraints. Our method predicts a utility-to-cost slope using machine learning and refines it through Bayesian fusion with historical priors, enabling anomaly-aware adjustments to multi-sensor data rates. The system extracts signal utility from frequency-domain and statistical features, including entropy and the Hurst exponent, while modeling cost through signal strength, anomaly likelihood, and SIM data usage. We deploy the framework on a 342-day water purification testbed and evaluate several machine learning models. The results demonstrate up to 60% data reduction with minimal impact on signal reconstruction fidelity, showcasing the framework's effectiveness for real-world, resource-constrained IIoT deployments.**

***Index Terms*—Industrial Internet of Things (IIoT), Adaptive Data Transmission, Edge Intelligence, Anomaly-Aware Sampling, Signal Utility Modeling, Bayesian Slope Fusion, Frequency-Domain Analysis, NB-IoT Communication**

I. Introduction

The rapid expansion of Internet of Things (IoT) deployments in smart cities, industrial automation, and environmental monitoring has led to a dramatic increase in the volume and frequency of sensor-generated data. These sensors often operate under stringent temporal and operational constraints: they are battery-powered, have limited communication bandwidth, and are frequently tied to cost-sensitive data plans, such as IoT SIMs with daily transmission limits [1], [2].

A broad range of methods has been explored to address these challenges. Data reduction and adaptive sampling approaches such as Kalman filtering and predictive modeling [2], compressed sensing [3], and spatial–temporal redundancy reduction [4] have demonstrated effectiveness in reducing redundant transmissions while preserving vital information. More recently, learning-based strategies, including reinforcement learning (RL) agents [5], [6] and parametric machine learning optimisers [7], have been introduced to dynamically adjust sampling or transmission decisions based on observed signal dynamics. However, these studies predominantly focus on instantaneous signal deviations or prediction errors, without incorporating temporal and operational constraints such as machine behavior trends under monitoring, daily data budgets, or cumulative usage over time.

Among these, RL-based methods typically frame transmission decisions as a two- or three-action space, such as sending, not sending, or classifying each sensor reading based on relevance or quality. and learn policies online through reward-based exploration. While powerful in principle, such approaches require continuous online training, are sample-inefficient, and introduce computational overhead unsuitable for IoT edge devices with strict memory, computational, and energy limitations. Furthermore, they involve inevitable hyper-parameter tuning, which reduces generalizability and increases the risk of unstable behavior.

On the other hand, heuristic or variance-based adaptive sampling methods [8] remain lightweight but are inherently rigid, using static parameters or thresholds and failing to generalise across heterogeneous sensor types or adapt to operational demands such as daily SIM data quotas or cloud ingestion caps.

Critically, current state-of-the-art strategies overlook several practical requirements fundamental to real-world IoT deployments. First, they rarely consider multi-sensor constraints, where end-nodes typically transmit all sensor readings as a single payload vector in one transmission rather than sending each reading separately. This architectural limitation, inherent to IoT communication protocols and common low-power IoT end-node designs, demands algorithms that manage the aggregate sending rate across all sensors jointly, not individually [3]. Second, existing methods fail to integrate daily budget awareness with continuous transmission rate adjustment in a lightweight manner suitable for deployment on constrained edge devices. Third, anomaly awareness is often neglected, where signal behavior changes indicative of faults or system events require increased sending frequency to maintain monitoring fidelity. These gaps highlight the need for mechanisms that can operate under real-world temporal and operational constraints, ensuring adaptability, efficiency, and robustness simultaneously.

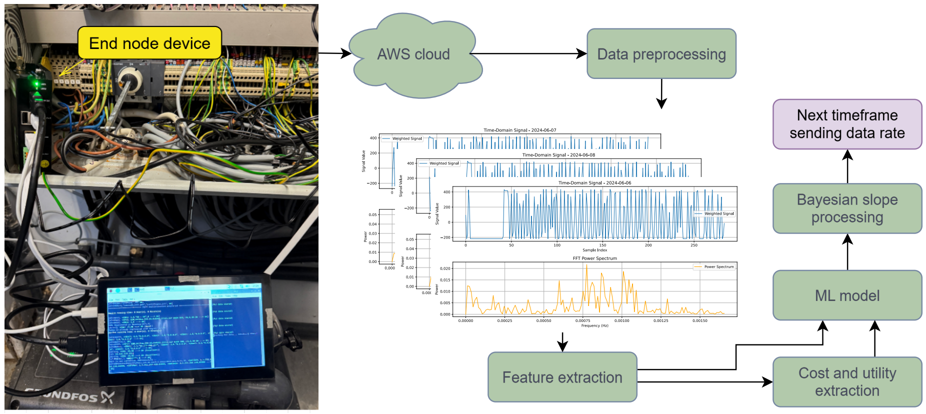

Fig. 1: Overview of the data flow from industrial machine to adaptive transmission framework: sensor signals are collected via PLC and transmitted to AWS cloud, where preprocessing includes trimming and FFT-based noise removal. Extracted features and costs are used to model and predict the next timeframe data sending rate.

In this paper, we propose a novel hybrid algorithm that adaptively controls data rates in constrained IoT environments while remaining sensitive to anomalies. Our approach begins with temporal and frequency-domain analysis of multi-sensor signals to derive lightweight yet informative features, including dominant frequency, signal entropy, and long-range dependencies. These features form the basis for a utility score, defined as the dominant period extracted from the fused multi-sensor signal via fast fourier transform (FFT), capturing the temporal informativeness and activity level of the system. We then integrate this utility signal into a cost-aware transmission control framework, where an adjustable timeframe transmission budget, referred to as cost, guides the sending rate. The total cost comprises three components: (i) signal strength, reflecting baseline energy consumption; (ii) deviation from median utility [9], indicating abnormal operational behavior; and (iii) total daily data volume in kilobytes, representing the communication burden. Our method leverages sensor-specific weighted norm modeling to account for the modality of sensors. It employs a ridge regressor, rather than other machine learning (ML) approaches, with a closed-form solution that offers computational simplicity, time efficiency, and minimal hyperparameter tuning, enabling fast, stable, and resource-efficient utility prediction for the next timeframe. To further enhance robustness, we introduce a Bayesian utility-cost slope fusion mechanism that integrates ML predictions with historical priors, yielding confidence-aware, adaptive transmission rates responsive to both normal and anomalous conditions.

We implement the proposed algorithm on a real-world end-node device within an industrial water purification system. Results show a 59.1% average reduction in daily transmissions while preserving anomaly detection and monitoring fidelity. This demonstrates the practical viability, computational efficiency, and deployability of our method on resource-constrained IoT edge devices under stringent temporal and operational limitations. Our main contributions lie in:

- **An innovative, multi-objective cost modeling** that integrates signal strength (for energy efficiency), deviation from median utility (as a proxy for anomaly likelihood), and SIM data usage (to capture communication budget constraints) into a unified, real-world transmission utility formulation. This modeling enables predictive and context-aware adaptation of sending rates, ensuring a balance between anomaly responsiveness and operational efficiency in constrained IoT deployments.
- **A novel Bayesian fusion mechanism for adaptive transmission control**, which combines machine-learned utility-cost slope estimates with historical similar-day priors. This confidence-aware strategy yields stable, uncertainty-calibrated transmission intervals, making it highly suitable for real-time deployment on resource-constrained edge devices in industrial environments.

The remainder of this paper is organized as follows: Section II presents the proposed model, including feature extraction, cost and utility, Bayesian interval selection technique. Section III provides evaluation results from a real-world water purification deployment. Section IV concludes the paper and discusses potential future enhancements.

II. Proposed Method

We consider an IoT end-node collecting data from N sensors, jointly transmitting them as a vector payload every T seconds. The primary objective is to dynamically adjust the sending interval T with respect to the costs while ensuring general anomaly detection capability. This approach aims to achieve two key goals: maintaining high signal fidelity through accurate reconstruction of transmitted data, and reducing the number of transmissions to remain within daily cost and bandwidth constraints. These objectives together ensure efficient, reliable, and anomaly-responsive monitoring under real-world operational limitations.

A. Testbed Setup

We deployed the proposed method on an IoT end-node in a water purification system, using a Raspberry Pi CM4 with SIM7070G NB-IoT module and Ethernet-connected PLC (Figure 1). Over a 342-day period, the system collected data from 12 sensors every 5 minutes, aggregated them into a JSON payload, and transmitted securely via MQTT over TLS.

B. Feature Extraction and System Modeling

To enable lightweight, real-time control of the IoT transmission rate, we extract a compact set of features that capture both temporal and frequency-domain characteristics of multi-sensor signals. Let $\mathbf{S}_t \in \mathbb{R}^{N \times M}$ denote the matrix of sensor readings over a timeframe t, where N is the number of samples (e.g., $N = 72$ for 6 hours of 5-minute intervals), and $M = 12$ is the number of sensors. To obtain a unified representation of system dynamics, we compute a single scalar signal by fusing all sensor modalities. Specifically, we normalize each sensor by its standard deviation σ_j over the window and compute a weight vector $w_j = \frac{\sigma_j}{\sum_k \sigma_k}$, where j indexes the current sensor and k spans all sensors in the window.

which emphasizes sensors with higher fluctuations. Each row $\mathbf{S}_{t,i}$ is scaled element-wise by $\mathbf{w}$, and a weighted Euclidean norm is applied. The resulting signal trace is defined as:

$$\text{signal}_t = \text{mean}\left(\|\mathbf{w} \odot \mathbf{S}_{t,i}\|_2\right) - \mu, \tag{1}$$

where μ is the average of all row-wise norms and $\odot$ denotes element-wise multiplication. This fluctuation-aware aggregation provides representative view of overall system behavior. The resulting signal_t is then used as the basis for extracting features such as dominant frequency, entropy, and long-term dependencies, as described below.

Nyquist-Derived Dominant Period (P_{dom}): We apply the discrete FFT on signal_t and compute its power spectrum. After filtering out low-frequency noise using the mean of the top 20 frequency components, the dominant frequency is selected as:

$$f_{dom} = \arg \max_{f_i > \bar{f}/2} |\hat{X}_w(f_i)|.$$

The corresponding period is computed as $P_{\text{dom}} = \frac{1}{2 f_{\text{dom}}}$, reflecting the half-cycle of the dominant oscillation [10]. Following the Nyquist criterion, this period defines an upper limit for the transmission interval without losing key information. We use this as the utility for each row sample, capturing temporal signal activity for adaptive scheduling.

Signal Entropy (H_t): To assess spectral complexity, we normalize the power spectrum $P(f_i)$ of signal_t into a distribution $p_i = \frac{P(f_i)}{\sum_j P(f_j)}$, where i and j index frequency bins. The entropy is then:

$$H_t = -\sum_i p_i \log p_i, \tag{2}$$

excluding zero-power bins [11]. Larger H_t denotes broadband activity; smaller values indicate periodicity.

Hurst Exponent (H_e): This metric characterizes the long-range dependency and persistence in the signal [12]. We estimate H_e via rescaled range (R/S) analysis by evaluating the relationship between lag l and the standard deviation $\tau(l)$ of lagged differences:

$$H_e = \text{polyfit}(\log l, \log \tau(l), 1)[0]. \tag{3}$$

The Hurst exponent feature is computationally efficient and well-suited for IoT end-node deployment, enabling real-time adjustment of transmission rates based on signal regularity, periodicity, and long-range dependency

TABLE I: Extracted Features and Corresponding Utility Periods (6 June 2024)

Time	Peak Freq.	Entropy	Hurst Exp.	Utility (min)
00:00	0.001080	3.3821	0.0831	7.72
06:00	0.000798	3.5123	-0.0105	10.44
12:00	0.000602	3.4204	-0.0367	13.85
18:00	0.000857	3.6931	-0.0320	9.72

The extracted features, dominant frequency, signal entropy, Hurst exponent, and the corresponding period, are computed from the unified scalar signal described above. Table I shows an example of these features and the resulting utility computed for a single 6-hour window. To address the trade-off between communication overhead and anomaly detection fidelity, we employ a compact cost model tailored to industrial sensor constraints. It jointly considers wireless energy consumption, data utility, and payload volume, capturing the system's demands on power, interpretability, and bandwidth.

Signal Strength Cost (C_{signal}): Represents the average received signal strength indicator (RSSI) across a day. Since IoT devices consume more energy when transmitting under weak signal conditions [1], this metric serves as a proxy for wireless energy expenditure:

$$C_{\text{signal}} = \frac{1}{N} \sum_{i=1}^{N} S_i, \tag{4}$$

where S_i is the RSSI at the i-th transmission, and N is the number of samples in the defined timeframe (e.g., 2 hours). *Lower RSSI values (e.g., near 2 dBm)* imply higher energy use and are undesirable. Thus, higher C_{signal} indicates better signal quality and lower transmission power [13].

Utility Deviation Cost (C_{utility}): Quantifies the deviation of the current signal utility from historical behavior, serving as a proxy for anomaly likelihood. Large deviations often correspond to significant events such as conductivity spikes or flow drops [2]:

$$C_{\text{utility}} = \left|U_{\text{today}} - \text{median}(U_{\text{hist}})\right|. \tag{5}$$

Here, higher C_{utility} indicates potential anomalies or dynamic behavior that may require increased sampling for accurate monitoring. Thus, high values are informative but may trigger cost increases.

Data Transmission Cost (C_{data}): Measures total data usage per day in kilobytes, reflecting SIM quota consumption. Assuming JSON-encoded packets from 12 sensors:

$$C_{\text{data}} = \frac{N_{\text{tx}} \times P_{\text{size}}}{1024}, \tag{6}$$

where N_{tx} is the number of transmissions and P_{size} is the payload size in bytes. Higher C_{data} indicates more frequent transmissions and thus higher operating cost, which is undesirable in constrained deployments.

The aggregated cost, computed as the sum of individual components, guides the transmission control. It serves as second input to mapping function that predicts the next sending interval T_{t+1} based on extracted features and current system conditions:

$$T_{t+1} = f\left(X_t, C_{total,t}\right). \tag{7}$$

The input feature vector X_t comprises the maximum peak frequency, signal entropy, Hurst exponent, and the corresponding period, extracted from the unified scalar signal signal_t. The scalar $C_{total,t}$ is the aggregated daily cost, combining signal quality, utility deviation, and payload size. The mapping function $f(\cdot)$ is learned via an ML model and integrated into the steps outlined in Algorithm 1, which summarizes the full adaptive transmission strategy.

Algorithm 1 Bayesian-Guided Adaptive Transmission Interval

1: **Input:** Sensor signal_t, historical data $\{(X_i, U_i, C_{\text{total},i})\}_{i=1}^{D}$
2: Extract features X_t: peak frequency, entropy, hurst exponent, and utility
3: Compute cost components: $C_{\text{signal},t}$, $C_{\text{utility},t}$, $C_{\text{data},t}$
4: Aggregate total cost: $C_{\text{total},t} = C_{\text{signal},t} + C_{\text{utility},t} + C_{\text{data},t}$
5: Train ridge regression on historical data to learn $\mathbf{w}, b$
6: Predict slope: $S_{\text{ml}} = \mathbf{w}^T X_t + b$
7: Compute σ_{util} as the standard deviation of training residuals
8: Retrieve similar-day slopes S_{hist} using 10-30 nearest neighbors in X space
9: Estimate variances:
10: $\quad \sigma^2_{\text{ml}} = (\sigma_{\text{util}}/C_{\text{total},t})^2, \quad \sigma^2_{\text{hist}} = (1.4826 \cdot \text{MAD}(S_{\text{hist}}))^2$
11: Fuse predictions via Bayesian update:
12: $\quad S_{\text{post}} = \frac{1/\sigma^2_{\text{ml}}}{1/\sigma^2_{\text{ml}} + 1/\sigma^2_{\text{hist}}} S_{\text{ml}} + \frac{1/\sigma^2_{\text{hist}}}{1/\sigma^2_{\text{ml}} + 1/\sigma^2_{\text{hist}}} S_{\text{hist}}$
13: Compute next transmission interval:
14: $\quad T_{t+1} = C_{\text{total},t} . S_{\text{post}}$
15: **Output:** Next timeframe's transmission interval T_{t+1}

Rather than directly estimating T_{t+1}, the model predicts the utility-to-cost slope S, defined as the ratio between the predicted utility at time $t+1$ and the known cost at time t. This slope reflects the expected return per unit of transmission cost. A higher slope indicates that the system is expected to yield more informative data for the same cost, suggesting stable behavior and allowing the node to transmit less frequently. Conversely, a lower slope implies a poorer return on cost, often due to anomalies or dynamic signal changes, prompting the system to transmit more frequently to preserve monitoring fidelity. In essence, the slope governs how aggressively or conservatively the system should invest its bandwidth, increasing data rates during critical periods and conserving them when the system is stable. We define the utility-to-cost slope as:

$$S = \frac{U_{t+1}}{C_{\text{total},t}}, \tag{8}$$

where U_{t+1} is the predicted utility of the next time window, and $C_{\text{total},t}$ is the current total cost. This slope quantifies the expected information gain per unit of transmission cost and serves as the core decision variable in our adaptive control strategy.

C. Bayesian Slope Fusion for Adaptive Transmission Control

As shown in Algorithm 1, the slope S is predicted using an ML model (line 6) based on features X_t and cost $C_{total,t}$ extracted from the current timeframe. To improve robustness and account for uncertainty, we retrieve slope values from historically similar timeframes in the feature space, identified using Euclidean distance, and denote them as S_{hist}. These values are then used to estimate the prior distribution.

The variance of this prior is computed using the Median Absolute Deviation (MAD), refer to Eq. 5, which is robust to outliers:

$$\sigma^2_{\text{hist}} = \left(1.4826 \cdot \text{MAD}(S_{\text{hist}})\right)^2. \tag{9}$$

We scale the ML prediction variance by the current total cost $C_{\text{total},t}$, incorporating sensitivity to budget:

$$\sigma^2_{\text{ml}} = \left(\frac{\sigma_{\text{util}}}{C_{\text{total},t}}\right)^2. \tag{10}$$

where σ_{util} is the standard deviation of modeling errors used in RMSE or MAE. Following Bayesian minimum variance estimation [14], the posterior slope estimate is computed as a precision-weighted combination of the ML prediction and historical prior:

$$S_{\text{post}} = \frac{\sigma^{-2}_{\text{ml}}}{\sigma^{-2}_{\text{ml}} + \sigma^{-2}_{\text{hist}}} \cdot S_{\text{ml}} + \frac{\sigma^{-2}_{\text{hist}}}{\sigma^{-2}_{\text{ml}} + \sigma^{-2}_{\text{hist}}} \cdot S_{\text{hist}}. \tag{11}$$

As in the final step of Algorithm 1, line 12, the next transmission interval is computed by dividing the current cost by the posterior slope:

$$U_{t+1} = T_{t+1} = C_{\text{total},t} \cdot S_{\text{post}} \tag{12}$$

This formulation integrates feature-driven learning with uncertainty-aware historical grounding, enabling stable, cost-adaptive transmission control suitable for industrial IoT deployments under constrained budgets.

III. Experimental Results

We evaluated the proposed adaptive transmission framework on an IoT end-node deployed in an industrial water purification system. Three strategies were tested: a fixed 5-minute baseline, a machine learning-based approach, and our uncertainty-aware hybrid method that integrates historical slope priors. The dataset spanned 342 days of uninterrupted operation under highly dynamic industrial load conditions, ensuring that both routine and anomalous behaviors were captured.

A rolling window evaluation was conducted across five interval configurations (4h, 6h, 8h, 12h, and 24h), assessing: (i) each model's ability to adjust transmission intervals based on signal dynamics, and (ii) reconstruction fidelity relative to the original 5-minute baseline. To ensure robustness and generalization, we employed cross-validation with 30% of the data reserved for testing. This setup closely mirrors real-world deployment scenarios in industrial IoT environments.

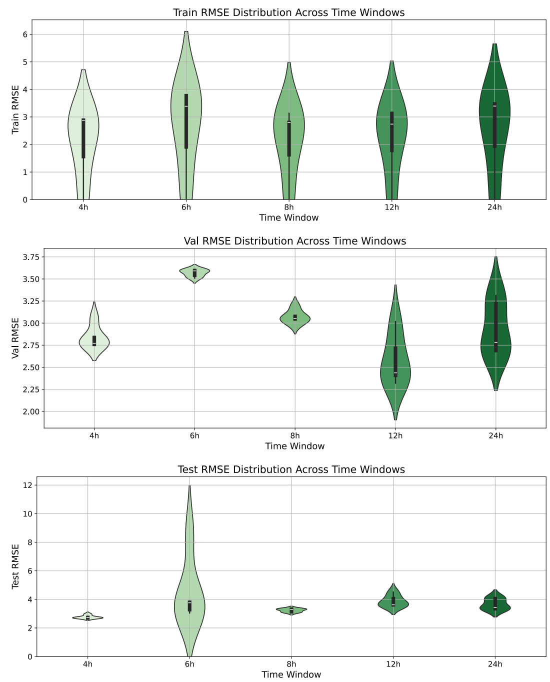

Fig. 2: Learning metrics across five ML models for utility prediction. Top: Train RMSE. Middle: Validate RMSE. Bottom: Test RMSE. Results reflect generalization and consistency across time windows.

Figures 2 compare the learning performance of five machine learning models across all time windows (4h to 24h). Train, validation, and test RMSEs capture model behavior during learning. Results show stable convergence and low test errors, confirming robust predictive capacity for adaptive transmission across varying temporal resolutions.

TABLE II: Impact of Bayesian Slope Fusion on reconstruction accuracy, transmission savings, and algorithm running time for ML-only vs. ML+Bayesian methods (4h interval).

Model	**ML Only**			**ML + Bayesian**		
	nRMSE	Save (%)	Time (s)	nRMSE	Save (%)	Time (s)
Random Forest	0.0405	45.48	49.39	0.0418	58.67	49.39
MLP	0.0417	46.82	44.64	0.0430	60.13	44.64
Linear Reg.	0.0406	46.10	0.41	0.0419	58.98	0.41
XGBoost	0.0402	44.57	27.71	0.0417	58.37	27.71
Ridge Reg.	0.0406	46.16	0.41	0.0420	58.95	0.41

Table II compares reconstruction accuracy and transmission savings across ML-only and ML+Bayesian configurations at a 4-hour interval. All models achieve high-fidelity reconstruction, with XGBoost performing best among ML-only methods (nRMSE: 0.0402, MAE: 0.0288). Bayesian fusion slightly increases error but substantially improves data reduction, up to 60.13% with MLP.

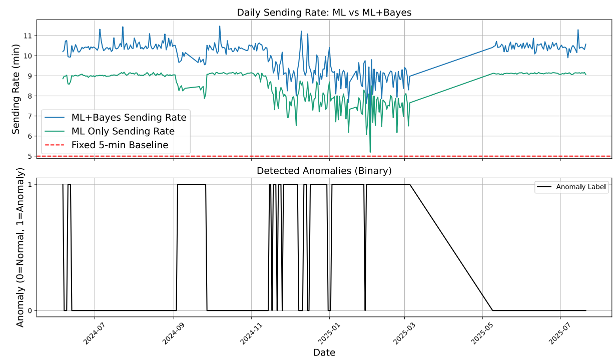

Fig. 3: Adaptive sending rates and anomaly detection using 5-min baseline, ML-only, and ML+Bayesian methods, showing system adaptability to dynamic conditions.

As shown in Figure 3, the ML+Bayesian method consistently selects longer transmission intervals than ML-only, particularly during stable periods, leading to substantial improvements in data efficiency. As presented in Table II, both configurations yield nearly identical nRMSE values, yet the Bayesian approach achieves significantly higher transmission savings. Both methods adaptively shorten the interval during anomaly-heavy periods (e.g., late 2024 to early 2025), reflecting responsiveness to dynamic system behavior. Moreover, the effectiveness of fusing historical slopes with the proposed learning strategy clearly demonstrates the strength of this approach, offering a lightweight and robust solution well-suited for deployment on resource-constrained IoT edge nodes.

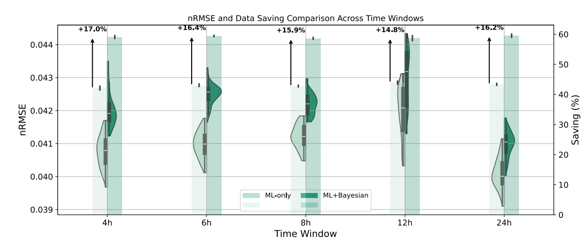

Fig. 4: Comparison of nRMSE (left y-axis) and transmission savings (%, right y-axis) across ML models. Bars represent savings; lines represent reconstruction error. Colors indicate model types (legend). ML+Bayesian offers savings with slight accuracy trade-off.

Figure 4 shows that the ML+Bayesian approach yields slightly higher reconstruction errors than ML-only. However, this increase is minor and justified by the substantial reduction in data transmissions, which lowers energy consumption at the node. By reducing sampling during low-utility periods, the proposed Bayesian slope fusion achieves consistent transmission savings across all models and timeframes. Despite fewer transmissions, signal fidelity is well preserved, demonstrating the method's effectiveness in adapting to signal dynamics without compromising reconstruction quality.

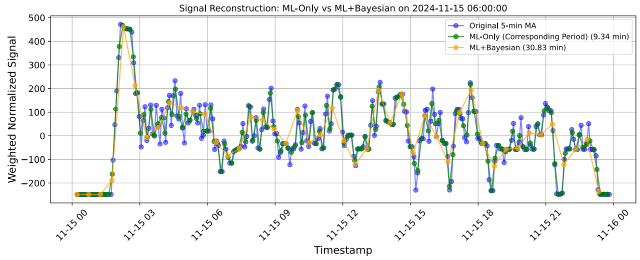

Fig. 5: Example of reconstructed vs. original signal on a random day using proposed technique with a daily window, illustrating temporal alignment and signal fidelity under adaptive transmission.

Figure 5 illustrates the performance of our proposed Ridge + Bayesian configuration on the mapped composite signal derived from 12 sensors. The reconstructed waveform closely follows the original signal, demonstrating that the model preserves structural fidelity while operating efficiently under resource constraints. This validates the suitability of the approach for real-time deployment on constrained IoT edge devices.

TABLE III: MAAPE comparison across time windows: ML-only vs Bayesian-guided models

Time Window	ML Only	Bayesian
4h	0.3575	0.3644
6h	0.3611	0.3711
8h	0.3571	0.3658
12h	0.3538	0.3674
24h	0.3499	0.3472

The Bayesian method yields competitive MAAPE scores across all time windows, balancing accuracy and efficiency. MAAPE's bounded, interpretable form suits low-magnitude sensor data. As shown in Table III, this is most evident with Ridge regression, where our approach achieves lower error and transmission cost than ML-only baselines. While the parametric adaptive sampling method in [7] reports a best-case MAAPE of 0.803 on environmental sensor data under a relaxed 0.5% error threshold, our approach achieves significantly lower values, ranging from 0.347 to 0.371 across all time windows. Although the datasets differ, our results are derived from multi-sensor industrial signals characterized by more complex dynamics and anomaly-aware conditions over 342 days.

IV. Conclusion

This paper presented a lightweight, deployable transmission framework for industrial IoT systems under strict communication and energy constraints. By extracting frequency-domain and statistical features from multi-sensor signals, the method estimates a utility score to guide adaptive transmission. A novel cost model, combining signal strength, anomaly likelihood, and data usage, is used alongside a Bayesian fusion mechanism that refines utility-to-cost slope predictions. This enables intelligent interval adjustment, achieving over 59.1% data savings while maintaining fidelity and anomaly responsiveness over 342 days of real-world deployment. Despite its effectiveness, the framework has certain limitations. During shorter timeframes, the model may struggle to adapt to sudden changes, as it relies on batch updates and assumes consistent cost-utility patterns. This limitation arises from its ML-based design, which requires sufficient data per window to ensure stable predictions. Future work will explore extending the framework to support heterogeneous sensor types, including high-dimensional modalities such as camera or video streams, enabling its use in a wider range of industrial monitoring scenarios. Adapting the system for both slow-changing environments and those sensitive to rapid signal spikes will further enhance its versatility.

V. Acknowledgment

This research was supported by a KTP project funded by Innovate UK (UKRI) under Partnership No. 10051698 (KTP Ref: 13392), in collaboration between LUBRON UK LTD and the University of Essex.

References

[1] Ferran Adelantado, Xavier Vilajosana, Pere Tuset-Peiro, Borja Martinez, Joan Melia-Segui, and Thomas Watteyne. Understanding the limits of lorawan. *IEEE Communications magazine*, 55(9):34–40, 2017.

[2] Handuo Zhang, Jun Na, and Bin Zhang. Autonomous internet of things (iot) data reduction based on adaptive threshold. *Sensors*, 23(23):9427, 2023.

[3] Huanzhuo Wu, Giang T Nguyen, Anil Kumar Chorppath, and F Fitzek. Network slicing for conditional monitoring in the industrial internet of things. *Transport*, 2018.

[4] Zaid Yemeni, Haibin Wang, Waleed M Ismael, Yanan Wang, and Zhengming Chen. Reliable spatial and temporal data redundancy reduction approach for wsn. *Computer Networks*, 185:107701, 2021.

[5] Muna Al-Hawawreh, Ibrahim Elgendi, and Kumudu Munasinghe. An online model to minimize energy consumption of iot sensors in smart cities. *IEEE Sensors Journal*, 22(20):19524–19532, 2022.

[6] Alaelddin F. Y. Mohammed, Salman Md Sultan, Joohyung Lee, and Sunhwan Lim. Deep-reinforcement-learning-based iot sensor data cleaning framework for enhanced data analytics. *Sensors*, 23(4), 2023.

[7] Hatem Algabroun and Lars Håkansson. Parametric machine learning-based adaptive sampling algorithm for efficient iot data collection in environmental monitoring. *Journal of Network and Systems Management*, 33(1):5, 2025.

[8] A. Rodríguez et al. An adaptive sampling period approach for management of iot energy consumption. *Sensors*, 22(19):1–19, 2022.

[9] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. Anomaly detection in time series: a comprehensive evaluation. *Proceedings of the VLDB Endowment*, 15(9):1779–1797, 2022.

[10] Fedaa et al. Ali. Feasibility study of signal processing techniques for vibration-based structural health monitoring in residential buildings. *Sensors*, 25(7), 2025.

[11] Leon Cohen. *Time-Frequency Analysis: Theory and Applications*. Prentice Hall Signal Processing Series, Pearson, 2nd edition, 2019.

[12] R. Bensaid, A. Ben Mnaouer, and H. Boujemaa. An evolving multivariate time series compression algorithm for iot tinyml applications. *Sensors*, 24(22):7273, 2024.

[13] Foivos Michelinakis, Anas Saeed Al-Selwi, Martina Capuzzo, Andrea Zanella, Kashif Mahmood, and Ahmed Elmokashfi. Dissecting energy consumption of nb-iot devices empirically. *IEEE Internet of Things Journal*, 8(2):1224–1242, 2021.

[14] Siliang Lu, Jingfeng Lu, Kang An, Xiaoxian Wang, and Qingbo He. Edge computing on iot for machine signal processing and fault diagnosis: A review. *IEEE Internet of Things Journal*, 10(13):11093–11116, 2023.