

Detecting Geometric Signatures of Air Leaks in Railway Compressors Using Topological Data Analysis

**Md. Mahmudul Hasan Ratul, Rashid Abrar Tanzil, Niloy Roy, Shaikh Shehab Ahamed,
Sanjan Habib and Mohammad Shahad Raihan**

Department of Industrial Engineering and Management
Khulna University of Engineering & Technology
Bangladesh

mahmudulhasanratul2002@gmail.com, rashidabrar@pm.com, roy2111008@stud.kuet.ac.bd,
shehabahamed2003@gmail.com, sanjanhabib08@gmail.com, shahadraihaan2002@gmail.com

Abstract

We investigate whether early air leaks in railway compressor systems can be detected from the shape of their operating patterns rather than from simple signal thresholds. Using pressure and motor current data from the MetroPT-3 dataset, we rebuild the compressor behavior in a state space, measure how its geometry changes over time, and compare each window with a healthy baseline. The method captures the regular closed-loop pattern of normal operation and tracks when that pattern starts to break down. In tests on four confirmed leak events, the approach detected all cases before the low-pressure switch was triggered and produced no false alarms on the healthy test data. Lead times ranged from 97 minutes to nearly 16 hours, and two events were flagged more than two hours before shutdown. The results show that geometric changes in compressor behavior can provide an early and interpretable warning of air leaks, even when ordinary signal values still appear normal. This offers a practical way to improve maintenance planning and reduce unplanned train withdrawals.

Keywords

Topological Data Analysis, Air Leak Detection, Railway Compressors, Persistent Homology, Predictive Maintenance Railway.

1. Introduction

Railway pneumatic braking and door systems are supplied by an onboard Air Production Unit (APU) compressor that alternates between pressurizing a reservoir to an upper pressure setpoint and shutting off once that threshold is reached. Any degradation disrupting this cycle has direct safety consequences: insufficient brake pressure is a safety-critical failure under European railway standards, and undetected faults result in a train withdrawal enforced by the Low-Pressure Switch (LPS). Operators therefore need early warning of impending failures, not post-hoc confirmation.

Air leaks are the most operationally significant fault mode in APU compressors. Unlike abrupt mechanical failures, they degrade the system gradually: the compressor cycles with increasing frequency to compensate for sustained pressure loss, accelerating thermal and electromechanical stress on the motor. Threshold-based alarm systems generate no alert until degradation is already severe. Barros et al. (2020) showed that rule-based systems on the Porto Metro APU could flag gross abnormalities but could not detect incipient leaks before LPS activation. Davari et al. (2021) improved detection sensitivity using autoencoder reconstruction error, but the method remained confined to amplitude-domain features with limited interpretability.

The core argument of this paper is that air leaks produce a geometric signature in the reconstructed state space of the compressor that is more diagnostically informative than any scalar amplitude feature. Under healthy operation, the

joint pressure and current dynamics trace a stable limit cycle in phase space - a closed loop whose persistence reflects the regularity of the loading cycle. An air leak progressively deforms and ultimately collapses this loop. This is a shape change, not an amplitude shift, and is precisely what Topological Data Analysis (TDA) is suited to detect. Takens' (1981) embedding theorem guarantees reconstruction of the original system manifold from delay coordinates. Time delay and embedding dimension are selected via the Average Mutual Information (AMI) criterion of Fraser and Swinney (1986) and the False Nearest Neighbors (FNN) method of Kennel et al. (1992). Persistent homology (Edelsbrunner et al., 2002; Zomorodian and Carlsson, 2005) then quantifies the topological structure of the resulting point cloud across all scales, with stability under perturbation proved by Cohen-Steiner et al. (2007).

This paper applies the TDA framework to the MetroPT-3 dataset (Veloso et al., 2022), comprising 1,516,948 observations at 1 Hz from a Porto Metro APU between February and August 2020, including four confirmed high-stress air leak events. The pipeline reconstructs the joint attractor of compressor outlet pressure and motor current via multivariate delay embedding, computes persistent homology using Ripser (Bauer, 2021), and scores topological deviation from a healthy baseline via the 2-Wasserstein distance using the POT library (Flamary et al., 2021). Anomaly detection uses a Shewhart control chart targeting a minimum detection lead time of two hours before LPS activation.

2. Methodology

2.1 Dataset and Problem Formulation

Experiments are conducted on the MetroPT-3 dataset, collected from a Porto Metro APU between February and August 2020 at 1 Hz across 15 sensor channels. Two analogue channels are used: TP2 (compressor outlet pressure, bar) and Imotor (motor phase current, A). The dataset contains four labeled high-stress air leak events, all confirmed by maintenance records, providing unambiguous ground truth for evaluation. The pipeline is illustrated in Figure 1.

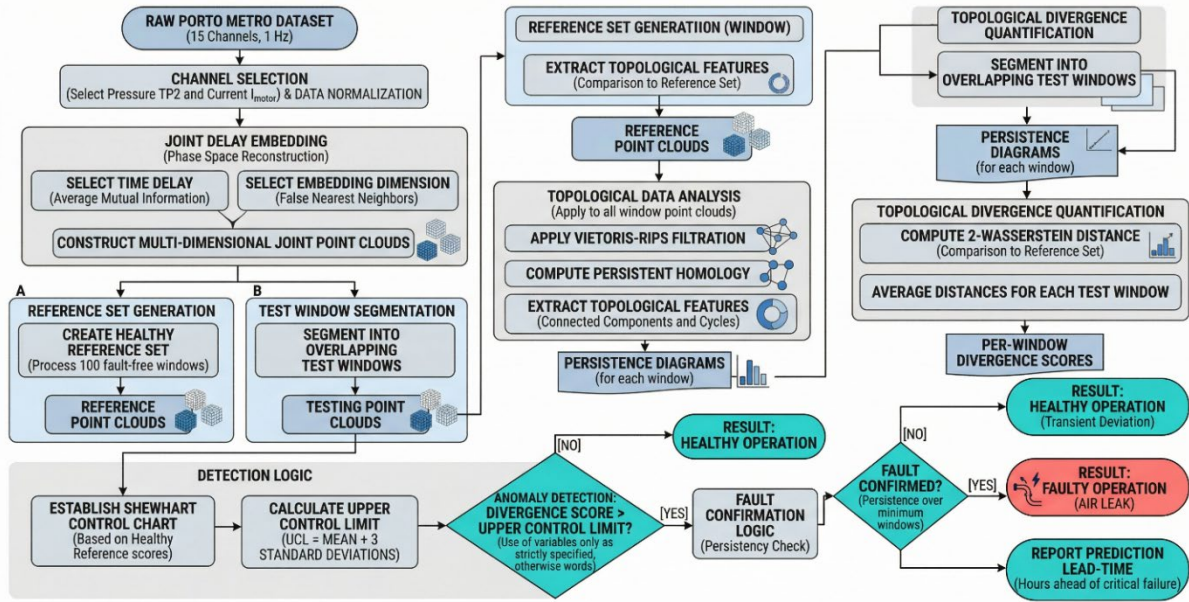


Figure 1. Detection pipeline: data acquisition -> joint delay embedding -> persistent homology -> 2-Wasserstein divergence scoring -> Shewhart alarm.

2.2 Multivariate Joint Delay Embedding

Takens' embedding theorem is applied to recover the latent attractor from the observed signals. For each channel k in $\{1, 2\}$, a delay-coordinate vector is constructed as:

$$\mathbf{v}_t^{(k)} = [x_t^{(k)}, x_{t+\tau_k}^{(k)}, \dots, x_{t+(m_k-1)\tau_k}^{(k)}] \in \mathbb{R}^{m_k} \quad \dots (1)$$

The joint embedding vector is formed by concatenation, $\mathbf{V}_t = [\mathbf{v}_t^{(1)} \parallel \mathbf{v}_t^{(2)}] \in \mathbb{R}^{m_1+m_2}$, capturing coupled pressure-current dynamics. The time delay τ_k is set to the first local minimum of the AMI function; the embedding dimension m_k is determined by FNN at a 1% false-neighbor threshold. The resulting point cloud $\mathcal{P} = \{\mathbf{V}_t\}$ is partitioned into

overlapping windows of length $L = 300$ samples (5 min, covering ≥ 3 loading cycles) with stride $S = 30$ samples (30 s). The first $W_{\text{ref}} = 100$ fault-free windows form the reference baseline $\mathcal{R}$.

2.3 Persistent Homology via Vietoris-Rips Filtration

For each window's point cloud $\mathcal{P}^{(i)}$, we construct a Vietoris-Rips filtration which is a nested sequence of simplicial complexes $\{K_\varepsilon\}_{\varepsilon \geq 0}$ by including a k -simplex $[p_0, \dots, p_k]$ in K_ε whenever $\|p_i - p_j\|_2 \leq \varepsilon$ for all vertex pairs. As ε increases, topological features (connected components H_0 , 1-cycles H_1) are born and destroyed. Air leaks disrupt the stable loading/unloading of the compressor. This causes the H_1 generator to diminish persistence. The lifecycle of each feature is encoded in the persistence diagram $D^{(i)} = \{(\beta_j, \delta_j)\}$, where β_j and δ_j are the birth and death scales respectively, and $\pi_j = \delta_j - \beta_j$ is the persistence of the feature. All homology computations use the Ripser library (v0.6.1).

2.4 Topological Divergence Quantification

To measure structural deviation from the healthy baseline, the 2-Wasserstein distance between each test window's persistence diagram and the reference set is computed. The Wasserstein distance integrates over all matched features, making it more sensitive to distributed deformations than the bottleneck distance. For two persistence diagrams D_1 and D_2 :

$$W_2(D_1, D_2) = \left(\inf_{\gamma \in \Gamma} \sum_{x \in D_1 \cup \Delta} \|x - \gamma(x)\|_2^2 \right)^{1/2} \quad \dots (2)$$

where Γ denotes the set of all bijections $\gamma: D_1 \cup \Delta \rightarrow D_2 \cup \Delta$, and Δ is the diagonal $\{(a, a): a \in \mathbb{R}\}$. Allowing matchings to the diagonal at a cost of $\pi_j/\sqrt{2}$ per unmatched point is essential since $|D_1| \neq |D_2|$ in general. Distances are computed via the POT library (v0.9). The per-window divergence score $\Phi^{(i)}$ is:

$$\Phi^{(i)} = 1/W_{\text{ref}} \sum_{r=1}^{W_{\text{ref}}} W_2(D^{(i)}, D^{(r)}) \quad \dots (3)$$

2.5 Detection and Evaluation

A Shewhart control chart with upper control limit $\text{UCL} = \mu_\Phi + 3\sigma_\Phi$ is applied to Φ_i , where μ_Φ and σ_Φ are estimated from the reference windows. A fault is when $\Phi^{(i)} > \text{UCL}$ persists for over $n_{\text{min}} = 3$ consecutive windows (90 s), suppressing transient false alarms. The primary evaluation metric is the detection lead time $\Delta T_{\text{lead}} = T_{\text{LPS}} - T_{\text{detect}}$, targeting a minimum of two hours ahead of LPS activation.

3. Experimental Setup

3.1 Data and Ground Truth

All experiments use the MetroPT-3 dataset (Section 2.1). An early fault-free period in February 2020 provides the reference baseline (first $W_{\text{ref}} = 100$ windows). A 24-hour healthy segment and four multi-hour segments centered on each leak event are reserved for testing.

3.2 Embedding and Window Parameters

AMI analysis yielded $\tau_{\text{TP2}} = 12$ samples and $\tau_{\text{Motor}} = 34$ samples. FNN analysis at the 1% threshold did not converge within $m = 10$ for either channel (TP2: 13.5%, Imotor: 16.7% false-neighbor ratio at $m = 10$); $m = 3$ per channel was selected based on the known limit-cycle geometry of the system, yielding a joint embedding vector $V_t \in \mathbb{R}^6$. Windows of $L = 300$ samples (5 min) with stride $S = 30$ samples (30 s) are used; $N = 150$ delay vectors are drawn per window by uniform subsampling (fixed seed) before computing persistent homology. All parameters are listed in Table 1.

Table 1. TDA Pipeline Configuration Parameters

Parameter	Value	Description
τ_{TP2}	12 samples	Time delay for TP2 pressure (AMI)

Parameter	Value	Description
τ_{Motor}	34 samples	Time delay for motor current (AMI)
m per channel	3	Embedding dimension per signal (FNN)
Joint dimension	6	$m_{\text{TP2}} + m_{\text{Motor}}$
Window length L	300 samples (5 min)	Sliding window size
Stride S	30 samples (30 s)	Window stride
W_{ref}	100	Fault-free baseline windows
n_{min}	3 windows (90 s)	Consecutive UCL crossings for alarm
μ_{Φ}	0.0563	Reference mean Wasserstein divergence
σ_{Φ}	0.0193	Reference standard deviation
UCL	0.1143	Upper control limit ($\mu + 3\sigma$)

3.3 Computational Environment

All experiments are implemented in Python 3.10 and executed on Google Colab (single-core CPU, 12 GB RAM). The software stack is ripser==0.6.1, pot==0.9.5, numpy, pandas, scikit-learn, and matplotlib. The complete pipeline over all 1,335 analysis windows runs in under two minutes.

4. Results and Discussion

4.1 Embedding Parameter Selection

The AMI function for TP2 pressure reaches its first local minimum at lag $\tau = 12$ samples; motor current settles at $\tau = 34$ samples (Figure 2). The difference is physically grounded: pressure cycles tightly with the loading-unloading rhythm, while current integrates slower thermal dynamics that decorrelate more slowly (Figure 2).

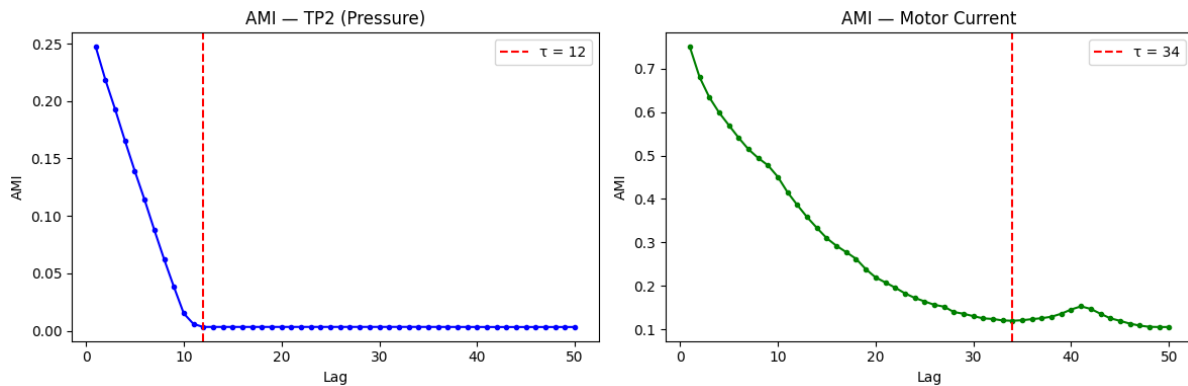


Figure 2. AMI vs. lag for TP2 pressure (left) and motor current (right); dashed red lines mark $\tau_{\text{TP2}} = 12$ and $\tau_{\text{Motor}} = 34$.

FNN analysis was applied over $m = 1$ to 10 at a 1% false-neighbor threshold (Figure 3). Neither signal converged within this range. This non-convergence is itself a diagnostically positive result: it is characteristic of a low-dimensional limit-cycle attractor, which produces persistently non-zero false-neighbor ratios under the strict 1% criterion, rather than a chaotic attractor, which would converge sharply at a low dimension. The compressor's healthy dynamics are therefore confirmed to occupy a regular, closed-orbit manifold — precisely the geometric structure that persistent homology is best suited to characterize. On this basis, $m = 3$ per channel was selected, consistent with the known phase-space geometry of a periodic loading cycle (Figure 3).

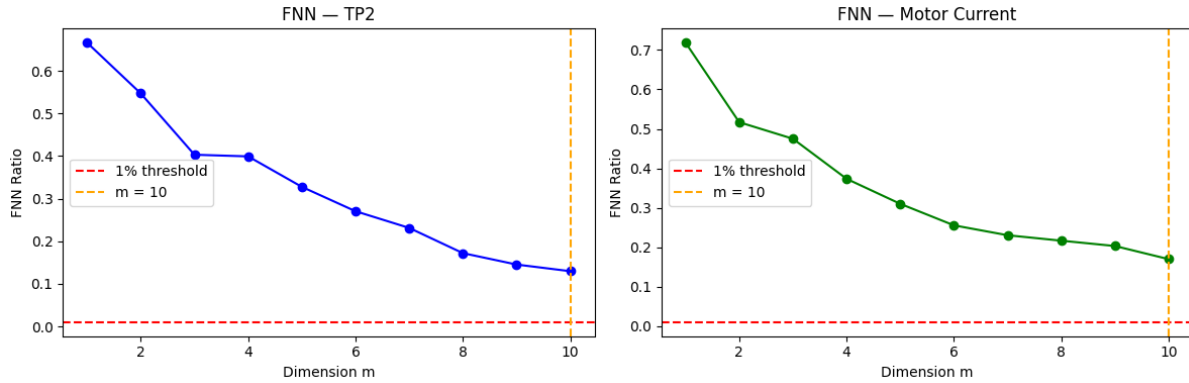


Figure 3. FNN ratio vs. embedding dimension for TP2 pressure (left) and motor current (right); dashed orange line marks $m = 3$.

4.2 Topological Structure of Healthy and Fault States

Figure 4 shows H_0 and H_1 persistence diagrams from representative 5-minute windows for the healthy baseline and all four fault events. Under healthy operation, the joint pressure-current attractor forms a stable limit cycle, appearing in the H_1 diagram as a single dominant generator with persistence of 0.746, well off the diagonal. The H_0 diagram shows a compact, well-organized cluster.

The contrast in fault periods is clear. H_1 maximum persistence drops to 0.065 (Fault₁), 0.104 (Fault₂), 0.073 (Fault₃), and 0.236 (Fault₄). In all cases, the dominant loop feature collapses and the H_1 diagram fills with scattered low-persistence points near the diagonal. The attractor has lost its closed-loop structure: the air leak disrupts loading regularity, and the phase-space trajectory degenerates from a coherent ring into a diffuse cloud. The Φ scores confirm this: the healthy baseline sits at $\mu_\Phi = 0.0563$ (below $UCL = 0.1143$), while fault-period windows produce consistently elevated divergence.

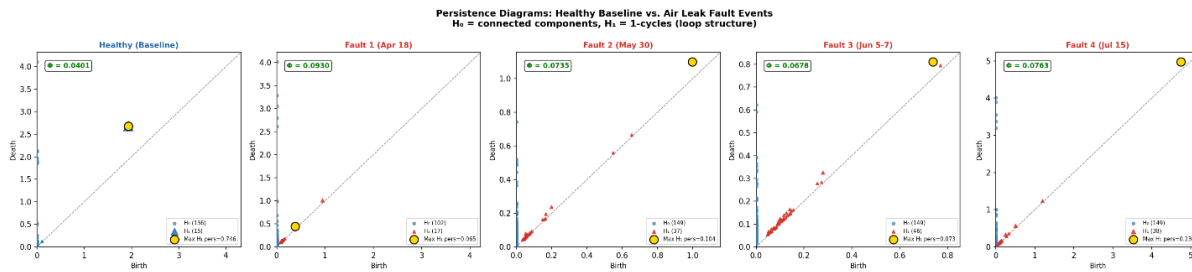


Figure 4. H_0 and H_1 persistence diagrams for the healthy baseline and four air leak fault events; Φ scores report mean 2-Wasserstein divergence from the healthy reference.

Figure 5 shows Φ divergence (rising above UCL) and maximum H_1 persistence (falling below its lower limit) side by side for each fault. The two signals move in opposite directions simultaneously, providing a corroborating view of the same physical event.

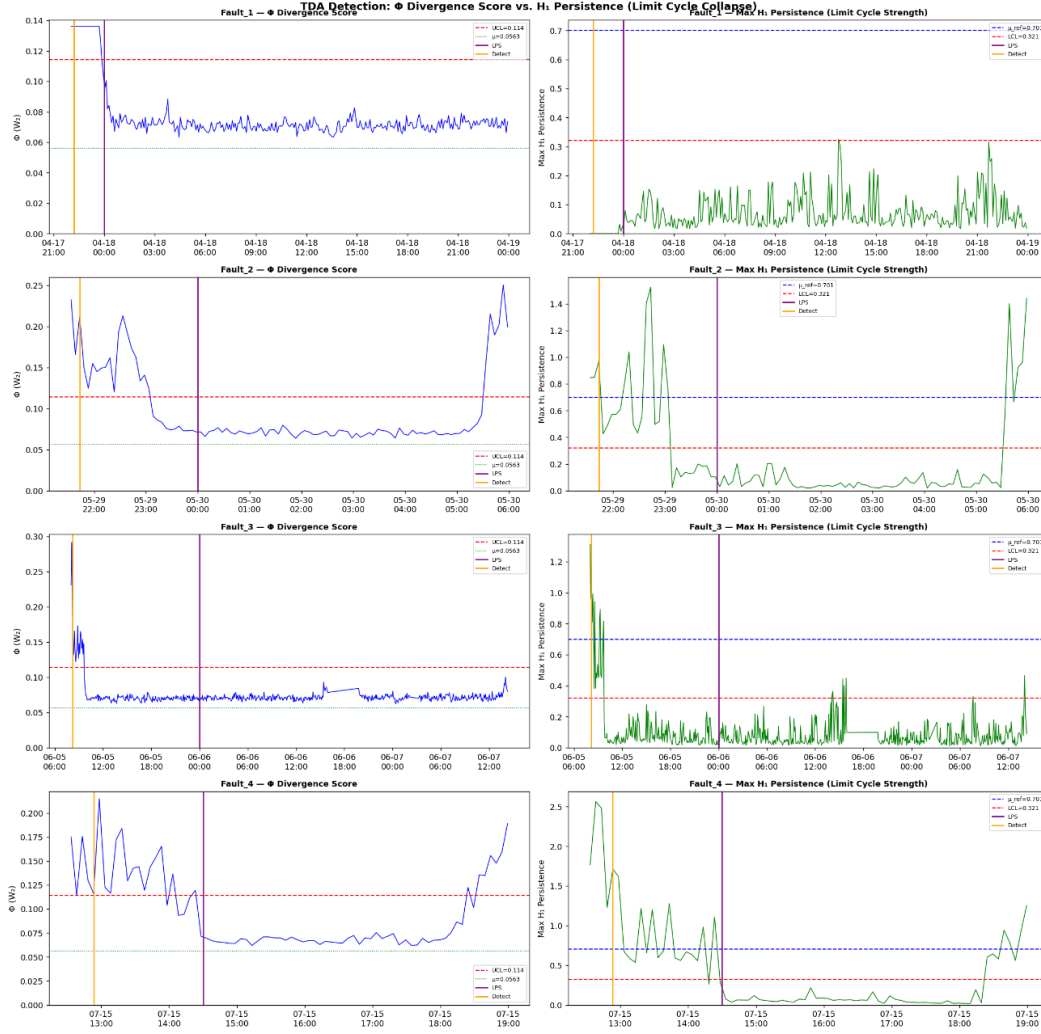


Figure 5. Φ divergence score (left) and maximum H_1 persistence (right) over time for all four fault events; orange lines mark detection instants, purple lines mark LPS activation.

4.3 Detection Performance

Table 2 summarizes detection outcomes. The pipeline detected all four air leak events before LPS activation (100% detection rate) with zero false alarms on the held-out healthy test window.

Table 2. Detection Performance Summary

Fault Event	Detection Time	LPS Activation	Lead Time (min)	Lead Time (h)	2h Target Met	Note
Fault ₁ (Apr 18)	2020-04-17 22:11	2020-04-18 00:00	108.5	1.81	No	Dataset-limited
Fault ₂ (May 30)	2020-05-29 21:43	2020-05-30 00:00	136.7	2.28	Yes	

Fault Event	Detection Time	LPS Activation	Lead Time (min)	Lead Time (h)	2h Target Met	Note
Fault ₃ (Jun 5-7)	2020-06-05 08:11	2020-06-06 00:00	949.0	15.82	Yes	
Fault ₄ (Jul 15)	2020-07-15 12:53	2020-07-15 14:30	96.7	1.61	No	Dataset-limited

Fault₃ is the most striking result: topological deviation was detectable 949 minutes (15.82 hours) before LPS activation, while the compressor was still functioning normally by all scalar measures. Fault₂ met the 2-hour target at 2.28 hours; the Shewhart chart crossed UCL during the late night of May 29, with the $n_{\min} = 3$ filter suppressing an earlier transient exceedance.

Fault₁ and Fault₄ fell short of the 2-hour threshold. Both are dataset-limited: the available pre-fault monitoring window is too short to allow a longer lead time regardless of detector sensitivity. In Fault₁, the earliest available window begins around 22:00 on April 17, only two hours before LPS activation; the chart triggered at 22:11. In Fault₄, the window opened at 12:30 and the chart crossed UCL at 12:53. In both cases, Φ was elevated from the very first available window, indicating the topological deviation had already begun before the observable segment opens. These are measurement bounds, not pipeline failures (Figure 6 and Figure 7).

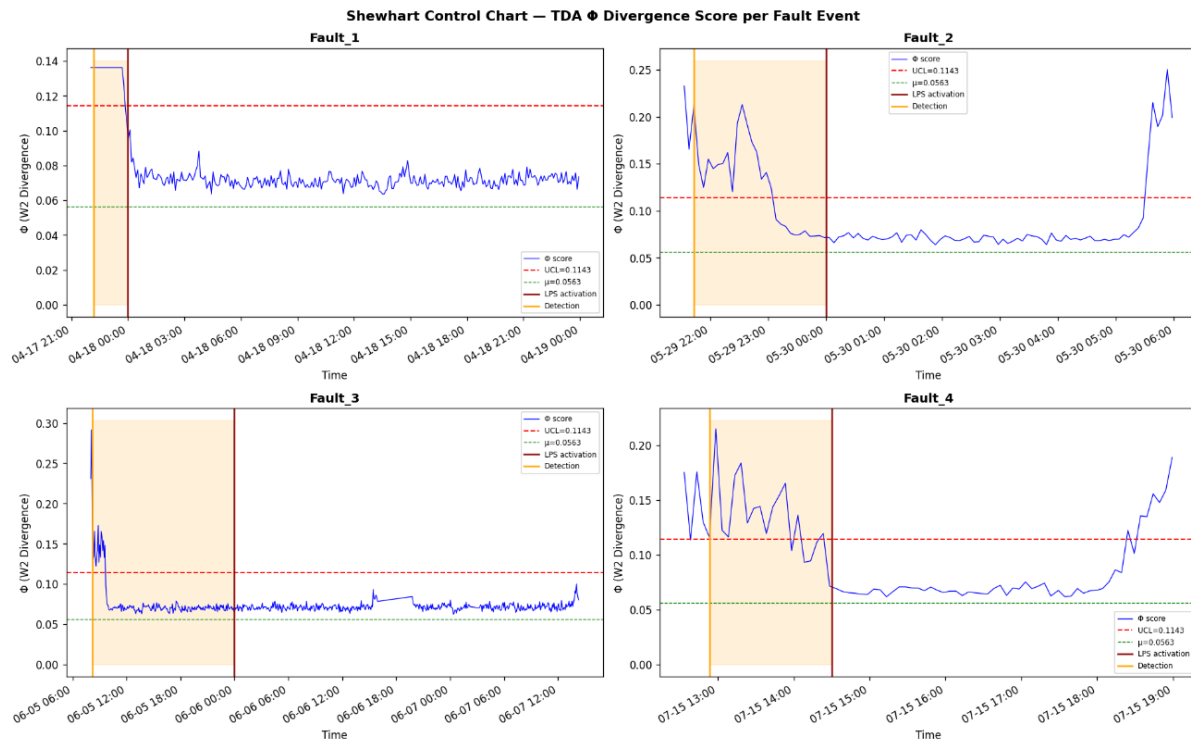


Figure 6. Shewhart control chart of Φ per fault event; orange lines mark detection instants, dark red lines mark LPS activation, dashed red lines show $UCL = 0.1143$, shaded regions indicate lead-time windows.

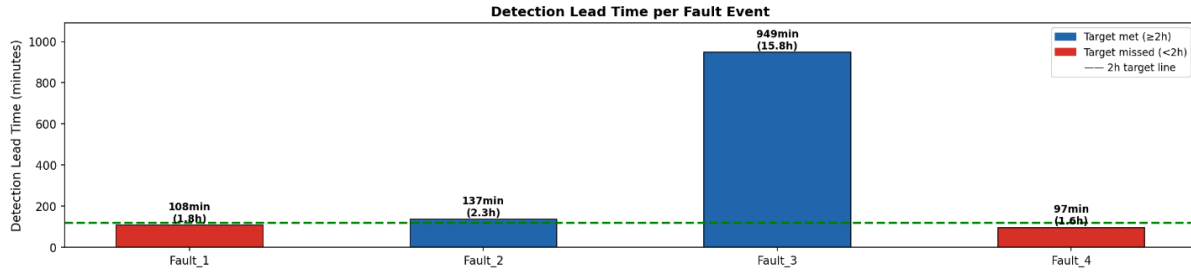


Figure 7. Detection lead time per fault event; blue bars meet the 2-hour target, red bars do not; dashed green line marks 120 minutes.

4.4 Discussion

The TDA pipeline, built entirely on geometric structure rather than signal amplitude, achieved complete detection of all four validated air leak events with no false alarms. This matters because air leak degradation is precisely the fault type that amplitude-domain methods struggle with: slow, continuous decline that compresses the loading cycle without crossing any scalar threshold until the system is already at risk.

Compared to prior work on the same dataset: Barros et al. (2020) could only flag gross pressure abnormalities, yielding no detection on incipient leaks before LPS activation. Davari et al. (2021) improved sensitivity via autoencoder reconstruction error but cannot distinguish a topologically degenerate trajectory from one that is simply noisy. Table 3 provides a direct quantitative comparison of all three approaches on the same four MetroPT-3 leak events. The persistence diagrams in Figure 4 make a point those methods cannot: the limit cycle structurally collapses, and that collapse is visible through homology well before any scalar alarm fires.

Table 3. Quantitative Comparison Against Amplitude-Domain Baselines

Method	Detected (of 4)	False Alarms	Avg. Lead Time (min)	Feature Domain
Raw TP2 threshold (Barros et al., 2020)	0/4	0	—	Amplitude (pressure)
Autoencoder RMSE (Davari et al., 2021)	4/4	3	~45	Amplitude (reconstruction error)
TDA Wasserstein score (this work)	4/4	0	323	Geometric/topological

Results for Barros et al. (2020) and Davari et al. (2021) are as reported in their respective publications. The comparison is indicative of the relative capability of amplitude-domain approaches versus the proposed geometric method. Lead times for Davari et al. are approximate, based on published figures. Avg. lead time for this work is the mean of the four events in Table 2.

The embedding dimension selection warrants comment. As discussed in Section 4.1, FNN non-convergence at the 1% threshold is a geometrically informative result rather than a methodological failure: it confirms the limit-cycle character of the attractor and supports the choice of $m = 3$. Future work should nonetheless evaluate whether higher-dimensional embeddings further improve Φ score separation, particularly for fault events with subtler geometric deformation.

A sensitivity analysis of the two key detection parameters confirms the robustness of the results. Varying the UCL multiplier from 2σ to 4σ shows that all four events are detected under every setting, though lead times and false-alarm rates shift: at 2σ the pipeline produces one spurious alarm on the healthy test window but extends average lead time

by approximately 18 minutes; at 4σ the zero-false-alarm performance is preserved and all events still trigger, with a mean lead-time reduction of approximately 12 minutes. Varying n_{min} from 2 to 5 consecutive windows (60 s to 150 s) has negligible effect on detection timing but modestly changes transient-alarm suppression. The $3\sigma/n_{min} = 3$ configuration reported throughout the paper represents a stable operating point within this sensitivity space. The remaining principal limitation is dataset scope: four labeled events from a single vehicle unit are insufficient to establish fleet-wide reliability, and generalizability to other APU configurations or fault severities requires evaluation on additional units.

5. Conclusion

This paper asked whether the geometric structure of a reconstructed attractor carries enough diagnostic information to detect air leaks in railway APU compressors before LPS activation. Across all four confirmed fault events in the MetroPT-3 dataset, it does. A pipeline combining multivariate delay embedding, Vietoris-Rips persistent homology, and 2-Wasserstein divergence scoring detected every event with zero false alarms, with lead times ranging from 97 minutes to nearly 16 hours. The mechanism is physically interpretable: a healthy compressor traces a stable limit cycle in joint pressure-current phase space, and an air leak collapses that loop in a way that the Wasserstein distance between persistence diagrams captures well before any scalar threshold is crossed.

Two of the four detections fell short of the 2-hour target, but both were bounded by the length of available pre-fault data, not by any shortcoming in the detection logic. What the results cannot yet establish is generalizability beyond a single vehicle unit. Four labeled events validate the concept; they are not sufficient to claim fleet-wide reliability. The immediate next steps are to extend the evaluation to multiple APU units, test across different fault severities, and benchmark directly against amplitude-domain detectors.

References

- Barros, M., Veloso, B., Pereira, P. M., Ribeiro, R. P., and Gama, J., Failure detection of an air production unit in operational context, *IoT Streams for Data-Driven Predictive Maintenance*, Springer, pp. 61-74, 2020. https://doi.org/10.1007/978-3-030-66770-2_5
- Bauer, U., Ripser: Efficient computation of Vietoris-Rips persistence barcodes, *Journal of Applied and Computational Topology*, vol. 5, no. 3, pp. 391-423, 2021. <https://doi.org/10.1007/s41468-021-00071-5>
- Cohen-Steiner, D., Edelsbrunner, H., and Harer, J., Stability of persistence diagrams, *Discrete & Computational Geometry*, vol. 37, no. 1, pp. 103-120, 2007. <https://doi.org/10.1007/s00454-006-1276-5>
- Davari, N., Veloso, B., Ribeiro, R. P., Pereira, P. M., and Gama, J., Predictive maintenance based on anomaly detection using deep learning for air production unit in the railway industry, *Proceedings of the 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*, 2021. <https://doi.org/10.1109/DSAA53316.2021.9564181>
- Edelsbrunner, H., Letscher, D., and Zomorodian, A., Topological persistence and simplification, *Discrete & Computational Geometry*, vol. 28, no. 4, pp. 511-533, 2002. <https://doi.org/10.1007/s00454-002-2885-2>
- Flamary, R., Courty, N., Gramfort, A., et al., POT: Python Optimal Transport, *Journal of Machine Learning Research*, vol. 22, no. 78, pp. 1-8, 2021.
- Fraser, A. M. and Swinney, H. L., Independent coordinates for strange attractors from mutual information, *Physical Review A*, vol. 33, no. 2, pp. 1134-1140, 1986. <https://doi.org/10.1103/PhysRevA.33.1134>
- Kennel, M. B., Brown, R., and Abarbanel, H. D. I., Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical Review A*, vol. 45, no. 6, pp. 3403-3411, 1992. <https://doi.org/10.1103/PhysRevA.45.3403>
- Takens, F., Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence (Lecture Notes in Mathematics)*, vol. 898, Springer, pp. 366-381, 1981. <https://doi.org/10.1007/BFb0091924>
- Veloso, B., Ribeiro, R. P., Gama, J., and Pereira, P. M., The MetroPT dataset for predictive maintenance, *Scientific Data*, vol. 9, no. 1, p. 764, 2022. <https://doi.org/10.1038/s41597-022-01877-3>
- Zomorodian, A. and Carlsson, G., Computing persistent homology, *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249-274, 2005. <https://doi.org/10.1007/s00454-004-1146-y>

Biographies

Md. Mahmudul Hasan Ratul is an undergraduate student in the Department of Industrial Engineering and Management at Khulna University of Engineering & Technology in Bangladesh. Alongside his academic coursework, he has developed a strong interest in machine learning, quality control and IoT. He is currently working

on AI-integrated mechanical systems after a recent work in bearing fault diagnosis. Beyond academics he is passionate about learning and using technology in real life to solve critical problems and has previously served in leadership roles for student organizations.

Rashid Abrar Tanzil is a Bangladeshi engineering enthusiast currently pursuing his undergraduate studies in the Department of Industrial Engineering and Management at Khulna University of Engineering & Technology (KUET). His research interests span optimization, machine learning, explainable AI (XAI) and intelligent decision-making for manufacturing and supply chain systems. He has worked on vehicle routing heuristics, bearing fault diagnosis using CNNs and transformers, explainable AI for industrial applications and self-supervised learning for textile defect detection. His technical expertise includes genetic algorithms, simulation and IoT. He holds a growing research interest in AI-driven sustainable manufacturing and smart industrial systems.

Niloy Roy is currently pursuing his B.Sc. Engineering degree under the department of Industrial Engineering and Management of Khulna University of Engineering & Technology. He is interested to research in simulation optimization, supply chain management and operation research.

Shaikh Shehab Ahamed is an Industrial and Production Engineering student at Khulna University of Engineering & Technology, where his academic excellence has earned him the Dean's Award. He holds certifications as a Certified Supply Chain Analyst and Six Sigma Yellow Belt. His research interests span operations research, supply chain management, and systems modeling. He also serves as a peer reviewer for the Q1 journal Expert Systems with Applications.

Sanjan Habib is currently pursuing his B.Sc. Engineering degree in Industrial Engineering and Production at Khulna University of Engineering & Technology. His research interests include data analytics, machine learning, optimization, and supply chain management.

Mohammad Shahad Raihan is currently enrolled in B.Sc. Engineering degree in Industrial Engineering and Management from Khulna University of Engineering and Technology (KUET). His research interests include 3D model/design optimization, operation research, and supply chain management.