DOI: 10.61782/fa.2025.0285

DETECTION OF OPERATIONAL REGIMERELATED FAULTS IN AIR COMPRESSOR SYSTEMS USING ACOUSTIC ANALYSIS AND MACHINE LEARNING

Andrej Hvastja **Jurij prezelj**
Faculty of Mechanical Engineering, University of Ljubljana, Slovenia

ABSTRACT

Compressed air systems are critical yet energy-intensive in industrial operations, making operational reliability essential for efficiency and safety. This study presents a novel approach using acoustic monitoring and machine learning to detect and classify operational regime-related faults in air compressor stations. Signal-processing and psychoacoustic features extracted from audio recordings enabled distinct regime identification through dimensionality reduction (autoencoders) and clustering (Gaussian Mixture Models). Regime duration analysis facilitated anomaly detection linked to operational inefficiencies. The proposed method effectively distinguished normal from abnormal patterns, highlighting regimes tied to compressed air consumption variability and equipment cycling, offering valuable improvements for predictive maintenance and reliability.

Keywords: *compressor station, acoustic monitoring, regime classification, anomaly detection*

1. INTRODUCTION

1.1 Compressed Air Systems in Industry

Compressed air is a versatile energy carrier widely used in various industrial sectors due to its applicability in numerous operational processes [1,2]. However, compressed air systems represent significant energy consumers within industrial facilities. According to [3], air compressors account for approximately 10% of industrial electricity consumption in the European Union and the USA, and up to 9.4% in China. According to [4], use of compressed air should be limited to scenarios involving safety improvements, substantial productivity improvements, or considerable labor reductions.

1.2 Economic and Operational Challenges

The operating costs of compressed air systems are predominantly due to their electricity consumption, which frequently exceeds the initial investment cost of the equipment [3, 4]. Furthermore, air compressor systems are prone to frequent faults and anomalies, making them expensive and inefficient, especially as the equipment ages [5]. Older compressors are notably susceptible to faults arising from wear and aging, highlighting the necessity for robust fault detection and diagnostic approaches [6]. In industries such as manufacturing, where machinery operates under extreme conditions and at high speeds, undetected faults can significantly disrupt production and compromise employee safety [7].

1.3 Fault Detection and Maintenance Strategies

Traditional preventive maintenance strategies, which rely on scheduled intervals, often result in premature component replacement, incurring unnecessary expenses [8].

Condition-Based Maintenance (CBM) addresses this inefficiency by monitoring the actual condition of the equipment to determine precise maintenance actions before failures occur. CBM involves tracking parameters such as vibration, acoustics, temperature, and pressure to detect early fault indicators [9].

1.4 Advantages of Acoustic Monitoring

Among various monitoring methods, acoustic sensors offer distinct advantages, including their non-contact nature, non-destructive implementation, and ease of integration into existing systems due to their compact size. The relatively low installation cost of microphones has encouraged widespread adoption for acoustic data collection in industrial environments. Despite these benefits, distinguishing fault-related acoustic signals from background noise remains a challenge due to noisy industrial conditions [10]. Machine learning (ML) has emerged as an effective tool to address this challenge, enabling efficient analysis and interpretation of acoustic data.

1.5 Operational Regime Classification

While conventional fault detection techniques primarily address abnormalities linked to fluctuating loads and environmental factors [10], much less attention has been paid to systematically classifying operational regimes. Yet industrial performance depends not only on mechanical reliability, but also on precise scheduling and evaluation of task durations. By comparing expected task timelines with actual operational performance, discrepancies become evident, revealing potential workflow anomalies or inefficiencies. This paper tackles that gap by emphasizing the necessity of classifying operational regimes, particularly those underscored by deviations between standard task durations and real-world operational times.

2. MATERIALS AND METHODS

2.1 Data Acquisition and Preprocessing

An operating compressor station, responsible for supplying compressed air to the Faculty of Mechanical Engineering in Ljubljana, was monitored using a microphone. The sound was recorded in WAV format over several days, including both weekdays and weekends. To accurately capture continuous operating conditions, a custom Python script segmented the audio into 2-second intervals with a 50% overlap.

2.2 Feature Extraction

For each sound segment the folowing signal-processing features were computed: Root Mean Square (RMS), Zero Crossing Rate (ZCR), Crest Factor, Short-Time Energy (STE), Variance, Skewness, Kurtosis and Arc length of the signal and its derivative, effectively capturing dynamic variations in the acoustic signals.

In addition to signal-processing features, the following psychoacoustic features were also computed on the sound signal: Sharpness, roughness, tonality, loudness, fluctuation strength, as they have been shown to be effective in machine fault detection [11].

2.3 Data filtering

To ensure data quality for the unsupervised clustering of operating regimes, outlier detection was performed using Isolation Forest on the aggregated feature set. Detected outliers, representing atypical data, were removed, preserving only normal operational data for subsequent analysis.

2.4 Dimensionality Reduction with Autoencoder

A deep autoencoder neural network was used for feature set dimensionality reduction [12]. The autoencoder architecture consisted of 4 dense layers with progressively reduced dimensions leading to a latent space dimension of 3 neurons. The model was trained with mean squared error (MSE) loss, using early stopping to avoid overfitting.

2.5 Operational Regime Classification

Latent features obtained from the autoencoder were clustered using a Gaussian Mixture Model (GMM). This unsupervised clustering approach grouped similar compressor operating states based on their latent representations, effectively classifying distinct operating regimes. Model selection criteria (AIC and BIC) were used to determine the optimal number of clusters. An illustration of the GMM clustering is provided in Figure 1 below, clustering latent features into 6 classes.

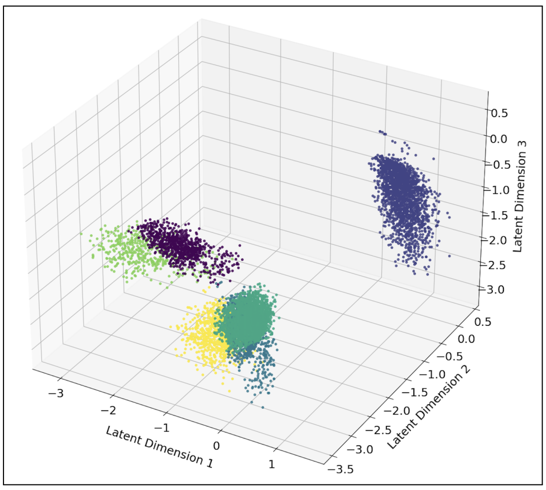

Figure 1. 6 clusters of operating regimes, based on GMM clustering.

2.6 Sequential Data Classification

The autoencoder and GMM were trained on randomly shuffled data to reduce overfitting and improve overall generality. In the second step, the same dataset was classified into identified operational regimes using the trained models. Whenever multiple consecutive sound recordings were assigned the same class, the event duration increased accordingly. However, to accurately evaluate the class or event duration, it is essential to process the recordings in their original sequence. This approach yields a feature that closely reflects the true duration of each event, and any deviation from this duration is flagged as an anomaly in the working regime.

3. RESULTS

This study examined a dataset of compressor station sound recordings spanning various operating regimes. Signal-processing features and psychoacoustic features, which capture human auditory perceptions and have been effective in machinery fault detection were derived from the sound recordings. The latent representation of the features was acquired using an autoencoder neural network and clustered using the Gaussian Mixture Model where the number of regime classes was selected using the AIC and BIC criteria.

Out of the six identified operating regimes, not all event durations are repeatable. As shown in Figure 2, the blue classes represent normal weekend operation and its typical duration range. Weekends serve as the baseline for “normal behavior” because of minimal compressed air usage. Through auditory inspection, Class 1—with its wider spread—was identified as the “working regime,” as its duration varies with compressed air consumption. In contrast, Classes 0, 2, 3, 4, and 5 remain similar across weekend (normal) and weekday (test) datasets, suggesting they correspond to more repeatable operations (e.g., turning compressors on/off or switching between them). Their durations should therefore remain relatively constant, given that these transient events (start-up, shutdown, etc.) are not driven by air consumption. Consequently, any deviation from expected durations in these transient regimes may indicate abnormal behavior and signal changes in compressor station operation.

Meanwhile, the purple set of regimes was recorded when a series of tests were conducted following compressor station maintenance. These regimes, labeled “overloading” and “rest,” demonstrated notably different durations, underscoring that the anomaly detection method successfully captured these unusual events.

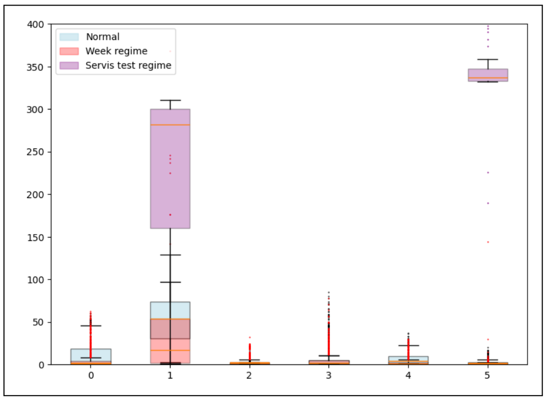

Figure 2. Event duration distribution of train (blue) and test (red) dataset.

The criteria for marking an event as an anomaly required its duration to fall below the 0.01 percentile or exceed the 0.99 percentile of the test dataset’s event durations. This approach proved most effective, as it flags events that are significantly shorter or longer than expected.

Figure 3 highlights these anomalous events using golden rectangles with thin red outlines. Several anomalies occurred in Class 1, which corresponds to compressor operation tied directly to compressed air consumption, and is therefore naturally more variable. Additional anomalies appeared in Class 2, typically representing regimes expected to have consistent durations. Notably, anomalies in Class 1 diminished over time as the operational durations became shorter and returned to expected ranges.

As expected, Figure 4 illustrates that all events from the unusual “overloading” testing regime following compressor maintenance were marked as anomalies. This confirms the method’s sensitivity and effectiveness in accurately detecting operational deviations.

4. CONCLUSION

This study demonstrated the effectiveness of acoustic monitoring combined with machine learning for detecting operational regime-related faults in air compressor systems. By extracting and analyzing signal-processing and psychoacoustic features from compressor sound recordings, distinct operational regimes were successfully iden-

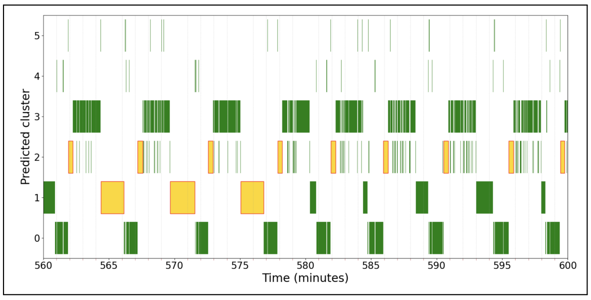

Figure 3. Work regime classes with corresponding durations.

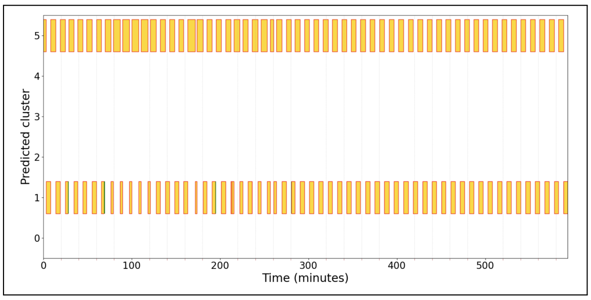

Figure 4. Overloading testing regime classes with corresponding durations.

tified using dimensionality reduction with autoencoders and clustering via Gaussian Mixture Models (GMM). The approach enabled the detection of anomalies based on deviations in regime durations, which were shown to correspond closely to actual operational events. Particularly, it was found that regimes associated with variable compressed air consumption exhibited significant variability in duration, whereas regimes related to equipment cycling and switching were highly consistent. Future work may focus on integrating this method into real-time monitoring systems to further enhance predictive maintenance and operational reliability in industrial compressor stations.

5. REFERENCES

[1] C. Y. Yuan, T. Zhang, A. Rangarajan, D. Dornfeld, B. Ziemba, and R. Whitbeck, "A decision-based analysis of compressed air usage patterns in automotive manufacturing," *Journal of Manufacturing Systems*, vol. 25, no. 4, pp. 293–300, 2006.

[2] J. J. Cabello Eras, A. Sagastume Gutiérrez, V. Sousa Santos, and M. J. Cabello Ulloa, "Energy management of compressed air systems. assessing the production and use of compressed air in industry," *Energy*, vol. 213, p. 118662, 2020.

[3] R. Saidur, N. Rahim, and M. Hasanuzzaman, "A review on compressed-air energy use and energy savings," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 4, pp. 1135–1153, 2010.

[4] D. Kaya, P. Phelan, D. Chau, and H. saraç, "Energy conservation in compressed-air systems," *International Journal of Energy Research*, vol. 26, pp. 837 – 849, 07 2002.

[5] B. Elliott, *Compressed air operations manual*. McGraw Hill Professional, 2010.

[6] J. Ma and J. Jiang, "Applications of fault detection and diagnosis methods in nuclear power plants: A review," *Progress in Nuclear Energy - PROG NUCL ENERGY*, vol. 53, pp. 255–266, 04 2011.

[7] B. G. Joung, C. Nath, Z. Li, and J. Sutherland, "Bearing anomaly detection in an air compressor using an lstm and rnn-based machine learning model," *The International Journal of Advanced Manufacturing Technology*, vol. 134, pp. 3519–3530, 09 2024.

[8] J. J. Montero Jimenez, S. Schwartz, R. Vingerhoeds, B. Grabot, and M. Salaün, "Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics," *Journal of Manufacturing Systems*, vol. 56, pp. 539–557, 2020.

[9] N. K. Verma, R. K. Sevakula, S. Dixit, and A. Salour, "Intelligent condition based monitoring using acoustic signals for air compressors," *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 291–309, 2016.

[10] Y. Tagawa, R. Maskeliunas, and R. Damaševičius, "Acoustic anomaly detection of mechanical failures in noisy real-life factory environments," *Electronics*, vol. 10, p. 2329, 09 2021.

[11] J. Murovec, L. Čurović, T. Novaković, and J. Prezelj, "Psychoacoustic approach for cavitation detection in centrifugal pumps," *Applied Acoustics*, vol. 165, p. 107323, 2020.

[12] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, 2016. RoLoD: Robust Local Descriptors for Computer Vision 2014.