A Statistical Classification Approach to Valve Condition Monitoring Using Pressure Features

Jacob Chesnes[1], Jason Kolodziej[2], and Daniel Nelson[3]

[1,2]*Rochester Institute of Technology, Rochester, NY, 14623, USA*
jjc4939@rit.edu
jrkeme@rit.edu

[3]*Novity, Palo Alto, CA, 94304, USA*
danelson@parc.com

ABSTRACT

This paper provides effective features that can be used to estimate the level of valve leakage in reciprocating compressors. It is expected that the level of leakage will change the shape of the pressure-volume diagram. This method constructs a feature space using the polytropic exponent during the expansion and compression phases as well as the discharge and suction valve loss power. These features are extracted by measuring in-cylinder pressure, discharge pressure, suction pressure, and the crank angle. Linear and quadratic discriminant classifiers are used as machine learning approaches to classify the valve health of the compressor. This method is implemented on a single-stage double-acting industrial gas compressor operating on air. Faults are seeded by precisely machining valve poppets to simulate common valve leakage. The approach shows a high classification accuracy in determining the degree of leakage and shows promise for future work in prognostics.

1. INTRODUCTION

Reciprocating compressors are commonly used in the oil and gas industry because of their reliability with a variety of gases. While they have been improved on over the years, they still have maintenance costs that can be reduced. A large portion of repair and maintenance costs are associated with the valves (Shirmer et al. 2004). Condition monitoring approaches have been implemented to predict when valves should be serviced to combat these costs. Two common condition monitoring approaches for valves are done by analyzing the vibrations of the valve manifolds or by analyzing the P-V diagram.

For vibration analysis, accelerometers are strategically placed on valve manifolds to accurately represent the actions of the compressor such as valve opening and closing due to the pressure differential. This happens cyclically, so the data is often converted to the time-frequency domain for each cycle. Pichler et al. (2011) have created a method to classify valve health while the compressor is operating at a constant load. The method takes the cumulative difference between the time-frequency matrix of a healthy dataset and the time-frequency matrix of an unknown health dataset. If the cumulative difference is less than a threshold, the unknown dataset is classified as being healthy. An improvement on this method is to calculate more features that aren't affected by the load. Pichler et al. (2015) created a spectrogram which was the difference between the time-frequency representations for two cycles. Features were then extracted and a support vector machine (SVM) algorithm classified valve health with high accuracy. Another approach to valve health classification is to implement image processing on the time-frequency representation of non-invasive vibration data. The authors pursued both a Bayesian statistical-based approach (Kolodziej and Trout (2016)) and one based on a deep learning method (Chesnes and Kolodziej (2021)). The Bayesian classifier used image-based statistical features derived from a region of interest in the Wigner-Ville spectrum and was compared with a convolutional neural network implementation using the region of interest image as the direct input to the classifier. In both cases the algorithms had classification accuracies greater than 90%.

P-V diagrams are another common method in industry for health monitoring of valves. The P-V diagram is also cyclical and takes into account a combination of mechanical actions and thermodynamic principles. Operators observe various features of the diagram and can diagnose faults when those features deviate from expectations. Ideally, this approach can be automated, thereby reducing operator costs of

reciprocating compressors. Brüel and Kjæl Vibro (2015) suggest that the polytropic exponent can be estimated and used in health monitoring. They explore the reliability of sensor placement and conclude that the polytropic exponent will deviate with faults. Phillippi (2016) explores the thermodynamics of reciprocating compressors with explanations based on the P-V diagram. They go over inefficiencies associated with compressors that affect the thermodynamics in the cylinders. Some of these can be calculated through the P-V diagram such as valve loss power. Other factors such as the type of gas used and the speed of the compressor are explored and show how they affect these values as well. The valve loss power has been theoretically discussed as useful compressor performance metric (Phillippi 2016). However, to the best of the authors' knowledge, it has not been used as a feature in health monitoring algorithms.

The proposed method uses real-world data collected on a single-stage, dual-acting reciprocating compressor fitted with a variety of sensors. The pressure placements are as follows: in-cylinder pressure on the head and crank-side, suction pressure, and discharge tank pressure. The crank angle was measured by an encoder and is used to calculate the cylinder volumes. The features for this classifier are the polytropic exponent on the expansion and compression phases, discharge valve loss power (DVLP), and suction valve loss power (SVLP). The leaking poppets are seeded in the head-side discharge valve to varying degrees by increasing the number of leaky poppets. The features are then used to train a Bayesian classifier, with an 80-20 training to testing data split. The method is able to diagnose valve leakage to a high degree of using sensors commonly used in reciprocating gas compressors.

2. EXPERIMENTAL SETUP

Experimental data was collected on a single-stage, dual-acting reciprocating compressor while compressing air. The compressor operates at a nominal speed of 6.4Hz, while the speed is measured with a crank encoder. The cylinder dimensions are known so the volume of each cylinder is calculated for every data point. The data also includes the in-cylinder pressure for both cylinders. The suction pressure is included instead of assuming a constant inlet pressure to increase the accuracy of the features. An air filter covers the inlet manifold causing notable pressure fluctuations. The compressed air is discharged into an anti-pulsation tank that slowly releases air back into the environment through a back pressure valve, and the pressure inside the tank is recorded. All of the data is recorded at 25.6kHz. The load options for the compressor are 0%, 50%, and 100%, and when set to 50%, the crank-side of the compressor operates normally, while the head-side inlet valves are forced open. For the 100% load option, both sides of the compressor compress air producing the highest pressure output. Figure 1 shows the compressor setup with a front and side view.

a)

b)

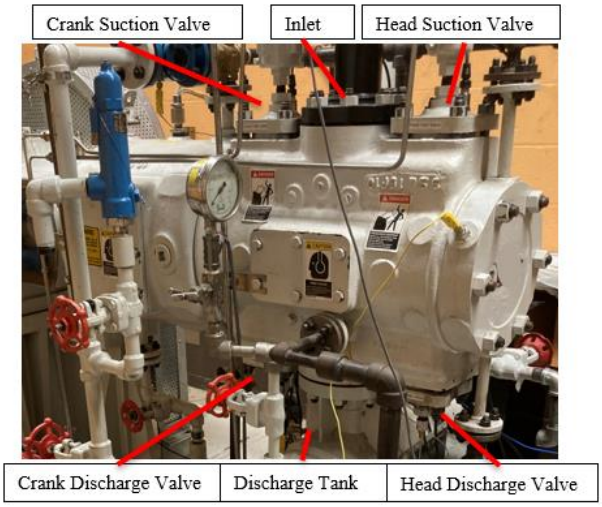

Figure 1. a) Front view of the reciprocating compressor. b) Labeled side view of the reciprocating compressor.

Faults are seeded into the head-side discharge valve which is on the right side in Figure 1 and data is collected while the compressor is operating at 100% load. There are 5 different levels of leakage with machined poppets simulating a chipped poppet. Figure 2 is showing the poppet array inside the valve assembly. Figure 3 is showing one of the machined poppets that is put inside the assembly. The first level of leakage is healthy data and the remaining 4 others have increasing levels of leakage. More poppets with a leak are added to the poppet array to increase the level of leakage representing multiple poppets being chipped.

Figure 2. The poppet array in the valve assembly.

Figure 3. A machined poppet.

3. FEATURE EXTRACTION

3.1. Polytropic Exponent

In an adiabatic compression or expansion process, the polytropic exponent is equal to the ratio of specific heats for a gas, or $n = c_p/c_V$ for an ideal gas, where c_p is the specific heat at constant pressure and c_v is the specific heat at constant volume. For a real process, the polytropic exponent can take on different values depending on the enthaply entering or leaving the working gas. Hence, estimating it's value should give some indication of the amount of leakage in the system. Equation (1) is used to calculate the polytropic exponent in this work. The polytropic exponent on the expansion phase is calculated when the volume of the cylinder is at 30 in^3 and 40 in^3, and the polytropic exponent for the compression phase is calculated when the volume of the cylinder is at 70 in^3 and 120 in^3. The specified volumes are also shown in Figure 5 in the zoomed-in plot of the compression phase for all the P-V diagrams. To get the points for the polytropic exponent, 10 data points are averaged around the volumes to reduce the impact of measurement noise.

$$n = \ln(p_2/p_1)/\ln(v_1/v_2) \qquad (1)$$

where;

$n = \ polytropic\ exponent$

$p_1 = pressure\ at\ point\ 1, psia\ (Pa)$

$p_2 = pressure\ at\ point\ 2, psia\ (Pa)$

$v_1 = volume\ at\ point\ 1, in^3\ (m^3)$

$v_2 = volume\ at\ point\ 2, in^3\ (m^3)$

The maximum pressure for the different leakage levels drops by about 3 psi with the higher levels of leakage as shown in Figure 4. Leakage can occur slowly over time and the pressure regulator can be adjusted to keep the discharge tank pressure as high as needed. However, the proposed method does not adjust the backpressure regulator with increased leakage. Figure 5 is showing a zoomed in portion of the P-V diagrams where the polytropic exponents were calculated.

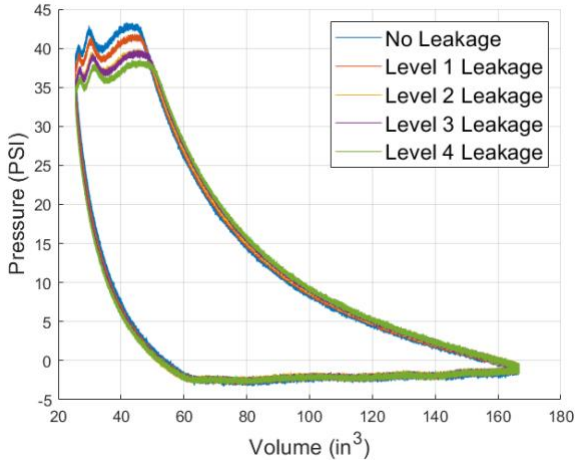

Figure 4. P-V diagram for all levels of leakage.

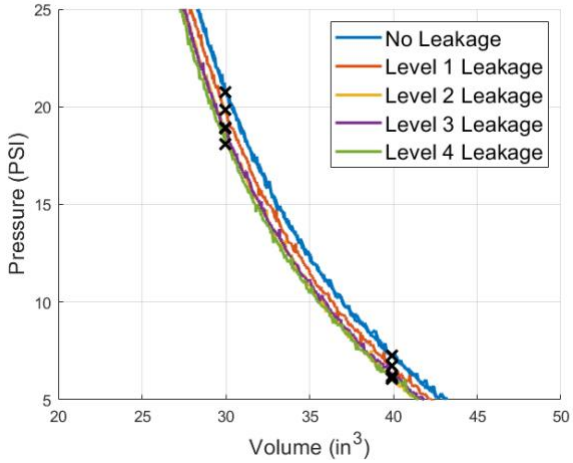

Figure 5. Compression phase of the P-V diagram with the markers for the polytropic exponent.

3.2. Valve Loss Power Estimation

The discharge tank pressure is needed to calculate the discharge valve loss power, and suction pressure is needed to calculate the suction valve loss power. Figure 6 shows a sample of data with the discharge tank pressure and suction

pressure. The suction valve loss power is calculated by numerically integrating the area between the suction pressure and the cylinder pressure over the cylinder volume and is labeled "A" in Figure 4. While the discharge valve loss power is the area between the cylinder pressure and discharge tank pressure and is labeled "B" in Figure 6.

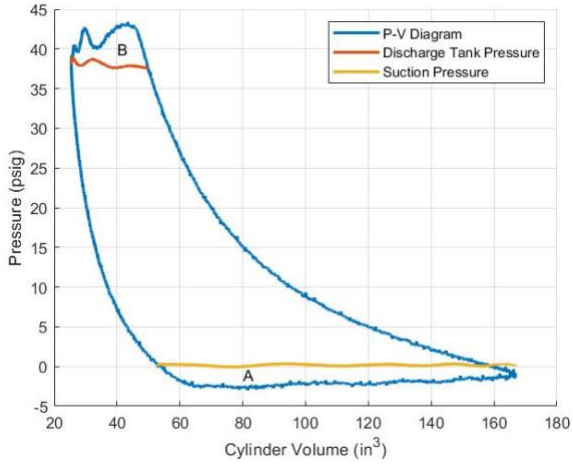

Figure 6. P-V Diagram with Discharge Tank Pressure and Suction Pressure

Three different feature combinations are used to train the classification algorithms. They are summarized in Table 1 where nE and nC are the polytropic exponents on the expansion and compression phases, respectively. Feature vector 2 is explored because the two features are directly affected by the seeded faults. The polytropic exponent on the compression phase is affected because air leaves the cylinder before the valves open. The last vector is explored because the data shows separability between the degradation levels.

Table 1. Caption of the table.

Scenario	Features
1	nE, nC, DVLP, SVLP
2	nC, DVLP
3	DVLP, SVLP

4. CLASSIFICATION

In practical applications, the selected features are expected to take on a continuous range of values as a function of leakage. Hence, such features may be appropriate for fault severity estimation and prognostics. To evaluate their effectiveness for fault estimation and prognostics, a quantitative measure of ground truth is needed – i.e. *how much leakage* is occurring. However, since the leakage is seeded in a discrete fashion, namely the number of affected poppets, a simple approach to evaluating the performance of the polytropic exponent and valve loss power in the context of fault detection is to formulate the problem as a classification problem. In a classification setting, features that result in high separation imply high sensitivity to the seeded faults. For this reason, the performance of the proposed features will be evaluated by their classification accuracy.

To formulate the present analysis as a classification problem, this work uses a linear discriminant classifier (LDC) and a quadratic discriminant classifier (QDC) to assess the health of the valves. These were selected because they are purely statistical and do not require hyperparameter tuning or iterations to converge to a solution like other classification algorithms. Bayesian inference is employed by QDC and LDC to classify which health condition is the most probable. The possible degradation levels are "0", "1", "2", "3", and "4" which correspond to number of machined poppets in the valve assembly. The objective function for these classifiers is in Equation (2), and Equation (3) is the *a posteriori* probability of an unknown feature vector being in a certain class, k.

$$J(\boldsymbol{x}) = argmin_{\omega}\{\sum_{k=0}^{4} C(\omega_i|\omega_k)p(\boldsymbol{x}|\omega_k)p(\omega_k)\} \quad (2)$$

$$p(\boldsymbol{x}|\omega_k) = \frac{1}{\sqrt{(2\pi)^d|\boldsymbol{\Sigma}_\mathrm{k}|}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_\mathrm{k})^T\boldsymbol{\Sigma}_\mathrm{k}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_\mathrm{k})\right) \quad (3)$$

where;

$\boldsymbol{x}$ = the feature vector of the unknown sample

ω_i = degradation level i; i = 0, 1, 2, 3, 4

d = the number of features (2, 4)

$C(\omega_i|\omega_k)$ = the cost of selecting class ω_k when ω_i is the true class.

$p(\boldsymbol{x}|\omega_k)$ = the a posteriori probability of the feature vector being with class ω_k

$p(\omega_k)$ = the a priori probability of the feature vector being with class ω_k

$\boldsymbol{\mu}_k$ = the mean vector of class k

$\boldsymbol{\Sigma}_k$ = the covariance matrix of class k

LDC and QDC both assume an *a posteriori* probability that is a multivariate Gaussian distribution, where the mean, $\boldsymbol{\mu}_\mathrm{k}$, and covariance matrix, $\boldsymbol{\Sigma}_\mathrm{k}$, are determined from the testing data. For LDC, the covariance matrices are assumed to be the same for all degradation levels. This is done by making the covariances for each degradation level the weighted sum of the measured covariances for all degradation levels, where the weights are the number of data points for each degradation level. The number of data points per degradation level is shown in Table 2. The covariance matrices for QDC are the measured covariance matrices for each degradation level. The number of data points for each degradation level is the same, so the *a priori* probability is 0.2 for all of the degradation levels.

Table 2. The number of cycles in the training and testing data points for each degradation level.

Degradation Levels	Training Data Points	Testing Data Points	Total Data Points
0, 1, 2, 3, 4	248	62	310
Sum	1240	310	1550

5. RESULTS

A sensitivity test for k-fold cross validation is performed to determine the effects of the k value on the accuracy. The test is completed in scenario 1. Figure 7 shows the minimum, average, and maximum classification accuracy for each k value. The maximum accuracy and average accuracy stay constant at 100% and 99.94% respectively. The minimum accuracy for each of the folds goes down proportionally with the number of folds. The constant average accuracy of 99.94% shows that 1 data point consistency gets misclassified, but the rest of the points have separability. The k value does not affect the overall performance of the algorithm, so a k value of 5 is chosen to be similar to the 80-20 testing to training split.

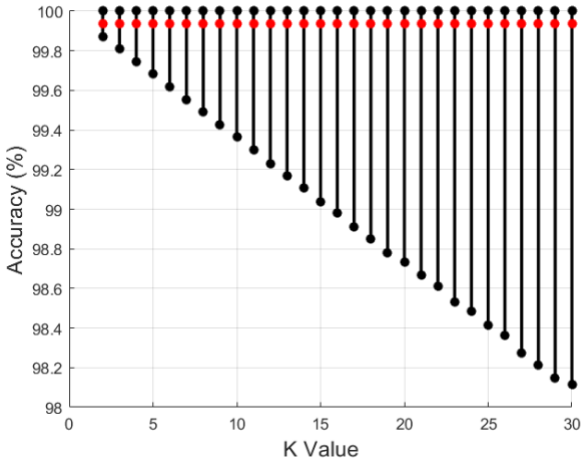

Figure 7. Sensitivity Test for K-Fold Cross Validation.

Table 3 shows the 5-fold cross validation accuracies for the 3 scenarios. There is not any significant difference between the QDC and LDC, implying that there is enough separability between the degradation levels and a more complex boundary is not needed. The LDC and QDC boundaries for scenarios 2 and 3 are shown in Figures 8 to 11. In the text by Webb and Copsey (2011), an implementation of principal component analysis (PCA) suggests that the coefficients in the eigenvector of a components can be used as an indicator of feature importance based on relative magnitude. The valve loss power features accounted for nearly 85% of the first principal component's eigenvector direction. This analysis is verified as there is no significant difference between scenarios 1 and 3 implying that the valve loss power features are indeed the most significant in separating the classes. Other feature selection methods have been applied including an exhaustive search based on classification accuracy corroborating the selection of these two features.

Table 3. The classification accuracy for 5-fold cross validation on all scenarios.

	LDC	QDC
Scenario	Accuracy (%)	Accuracy (%)
1	99.94	99.94
2	96.69	96.62
3	100	99.94

Scenario 2 does not have as high of a classification accuracy as the other scenarios, because it does not have a clear boundary between degradation levels 2 and 3 with a few data points on the wrong side of the boundaries. The feature space with DVLP and nC struggles to show complete separability with degradation levels 2 and 3, but DVLP with SVLP is enough to distinguish between the two leakage levels. As shown in Figures 8 to 11, the largest separation occurs between degradation levels 0 and 1 which correspond to the healthy dataset and the lowest amount of degradation. As the degradation level increases, the separation between the degradation levels decreases. The polytropic exponents do show a trend with leakage as shown in Figure 12 with the compression polytropic exponent positively correlating to the level of leakage while the expansion polytropic exponent negatively correlates to the level of leakage. The polytropic exponents do not have class separation, but the trends are promising for future work in prognostics.

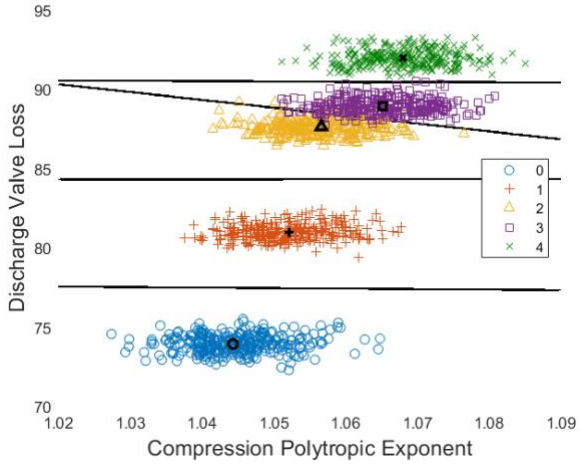

Figure 8. LDC classification boundaries for scenario 2.

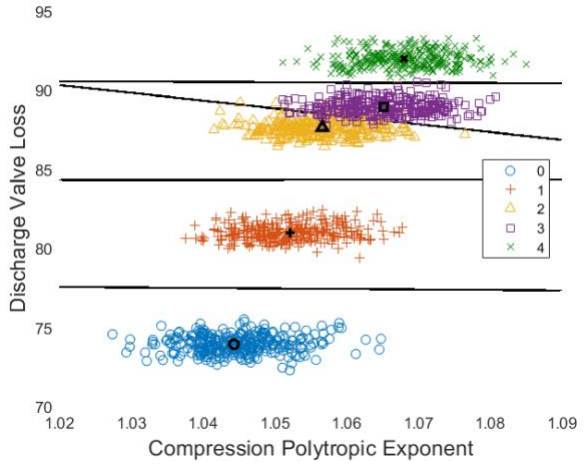

Figure 9. QDC classification boundaries for scenario 2.

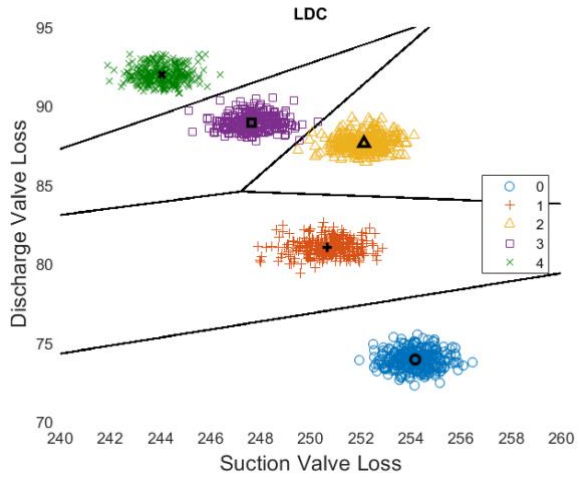

Figure 10. LDC classification boundaries for scenario 3.

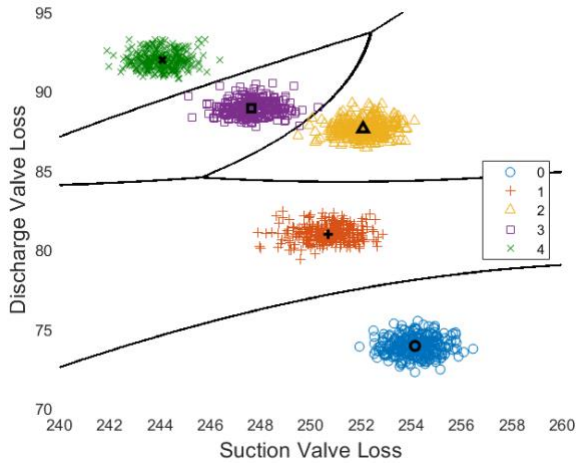

Figure 11. QDC classification boundaries for scenario 3.

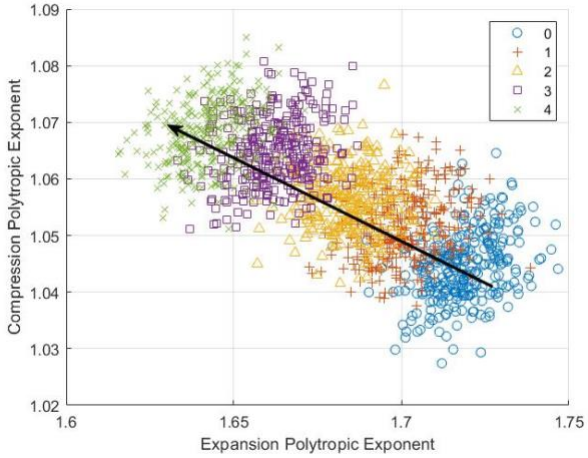

Figure 12. Expansion polytropic exponent vs. compression polytropic exponent.

6. CONCLUSIONS

The features extracted from the P-V diagram are sufficient for the classifiers to achieve a classification accuracy greater than 99%. The high level of accuracy stems from the discharge valve loss power and suction valve loss power. These two features achieve a high level of accuracy without the need for the polytropic exponent. The polytropic exponent does correlate to the level of leakage, but cannot be used to distinguish between small levels of leakage. Even though the leakages were seeded in the discharge valve and don't directly impact the suction valves, the suction valve loss power was still affected by the level of leakage. Future work is to explore how well this method works with seeded faults in other locations, mixed faults, and how adjusting the back pressure regulator to maintain discharge tank pressure impacts the classifiers.

ACKNOWLEDGEMENT

Thank you to PARC for supporting the current work on the compressor. Thank you to Dresser-Rand for donating the compressor and supporting previous projects to instrument the compressor.

REFERENCES

Brüel and Kjæl Vibro, (2015). Application Note | Monitoring strategy – Effective cylinder leak detection using polytropic exponent function

Kolodziej, J. R., and Trout, J. N., (2018). *An image-based pattern recognition approach to condition monitoring of reciprocating compressor valves*, Journal of Vibration and Control, 24(19), pp. 4433–4448.

Chesnes, J., & Kolodziej, J.K. (2021). An Application

Based Comparison of Statistical Versus Deep Learning Approaches to Reciprocating Compressor Valve Condition Monitoring. Annual Conference of the PHM Society, 13(1).

Phillippi, G., (2016). Basic thermodynamics of reciprocating compression, 45th Turbomachinery & 32nd Pump Symposia

Pichler, K., Schrems, A., Buchegger, T., Huschenbett, M., Pichler, M., (2011). *Fault detection in reciprocating compressor valves for steady-state load conditions*, Proceedings of the 11th IEEE International Symposium on Signal Processing and Information Theory, Bilbao, Spain, pp. 224–229

Pichler, K., Lughofer, E., Pichler, M., Buchegger, T., Klement, E. P., Huschenbett, M., (2016). *Fault detection in reciprocating compressor valves under varying load conditions*, Mechanical Systems and Signal Processing,70-71, pp. 104-119

Schirmer, A. G. F., Fernandes, N. F., Caux, J. E. D., (2004). On-line monitoring of reciprocating compressors, *Npra maintenance conference*

Webb, A. R. And Copsey, K. D., (2011). Statistical Pattern Recognition 3rd Edition, *Wiley*

BIOGRAPHIES

Jacob J. Chesnes is currently pursuing a Ph.D. in Mechanical and Industrial Engineering at the Rochester Institute of Technology (RIT) in Rochester, NY with expected completion in 2024. He received a B.S. in Mechanical Engineering and an M.E. in Mechanical Engineering from RIT in 2021. He has been a production support engineer co-op at G. W. Lisk tasked with decreasing the rate of solenoid valve failure. His most recent project was to design an inclusive retractable game net easily set up on standard cones for Inclusivity Inc.

Jason R. Kolodziej is an Associate Professor of Mechanical Engineering at the Rochester Institute of Technology (RIT) in Rochester, NY. He received his Ph.D. in mechanical engineering from the State University of New York at Buffalo in 2001 with a research focus in controls and nonlinear system identification. For eight years he worked in industry for General Motors Fuel Cell Activities as a Sr. Research Engineer with principle duties in hybrid electric-fuel cell vehicle powertrain controls and system architecture. To date, he has been granted 10 U.S. Patents. His present research focus is the study of fault detection, diagnosis, and prognostic health assessment of engineering systems. He currently has funded projects covering a wide range of industrial applications from electromechanical actuators in aircrafts to fuel cell automotive powertrains to large-scale compression equipment. He is a member of the ASME. In 2012, he was awarded RITs prestigious Eisenhart Provost Award for Excellence in Teaching.

Daniel Nelson has over 10 years of experience in high-fidelity modeling and numerical methods, with particular emphasis in fluid dynamics and tribo-mechanical systems. At Novity, he applies his expertise to the development of physics-based algorithms, including model-based reasoning, physics of failure and prognostics and diagnostics of industrial machinery. Prior to Novity, Dr. Nelson has worked for several OEMs in the oil and gas, chemical processing, and aerospace industries, where his focus was on multi-physics modeling of tribological elements and gas compression equipment. Dr. Nelson earned his Ph.D. at U.C. San Diego, where he studied computational fluid dynamics and dynamical systems, and his B.S. in Aerospace Engineering at San Diego State University.