

# Condition Monitoring of a Reciprocating Compressor Using Wavelet Transformation and Support Vector Machines

Shawn Falzone<sup>1</sup> and Jason R. Kolodziej<sup>2</sup>

<sup>1</sup>*Rochester Institute of Technology, Graduate Student - Department of Electrical Engineering, Rochester, NY, 14623, USA  
sdf3445@rit.edu*

<sup>2</sup>*Rochester Institute of Technology, Associate Professor - Department of Mechanical Engineering, Rochester, NY, 14623, USA  
jrkeme@rit.edu*

## ABSTRACT

Condition monitoring techniques were applied to a reciprocating compressor in order to determine if faults were present in a system. Through the use of vibration based sensors, fault monitoring of the crank-side discharge valve springs was accomplished. Data was collected through a range of injected fault conditions and analyzed through the use of discrete wavelet transformations. The wavelet coefficients produced were transformed into a six-dimensional feature space through the use of first and second order statistics. By using a support vector machine classifier, the nominal and faulted condition data was used to train a fault monitoring classifier. This classifier was verified through the use of additional test data, and resulted in classification rates of 90% and above. This result is based on the trial of a multitude of different wavelets and support vector kernels in order to achieve the optimal performance for the dataset.

## 1. INTRODUCTION

Cost reducing, efficiency focused industrial processes have become more prevalent on the modern shop floor. With this push for initiatives like lean manufacturing and Six Sigma, machine downtime becomes a very important factor in the flow of a production line. Some machines can have an impact across the entire production floor, such as hydraulic pumps or air compressors. The reciprocating compressor is one such machine, sometimes supplying air pressure to a number of machines across the facility (Lin, Wu, & Wu, 2006). This compressor is widely used for its reliable pressure levels during operation and versatility with different gasses. Prior to the advent of condition monitoring techniques, two repair

Shawn Falzone et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

philosophies were used for reciprocating compressors, breakdown maintenance, which waited until the equipment failed, or preventative maintenance, which sets up a regular maintenance schedule (Bloch & Hoefner, 1996). Preventative maintenance, while more efficient than breakdown maintenance, can still result in more frequent and unnecessary downtime for the compressor. In order to further increase the efficiency of preventative maintenance techniques, condition monitoring can be implemented.

### 1.1. Condition Monitoring Techniques

The use of condition monitoring techniques as part of a maintenance program is known as Condition-Based Maintenance (CBM). Since its inception, CBM techniques have become prevalent in numerous different fields and industries. Through the use of CBM, the maximum reliable life of a component can be utilized, without the unexpected loss of the system due to a major failure (Prajapati, Bechtel, & Ganesan, 2012).

Previous condition monitoring techniques utilized Pressure-Volume, or P-V diagrams, in order to monitor the system. By comparing a real-time P-V diagram to a theoretically nominal diagram, deviations from normal operation can be found. (Trout & Kolodziej, 2016) When performing condition monitoring on reciprocating compressors, vibration analysis is another commonly used method. This method tracks the vibration signature of the compressor when running, and looks for deviations from the nominal signature caused by fault conditions. Vibration analysis is a great choice for reciprocating compressors as external sensors can be used and retrofitted to existing hardware with very little invasive modification. Due to the cyclical nature of the reciprocating compressor, the signals used for vibration analysis are also cyclical in nature. These signals are known as cyclostationary, as the majority of the signal information is

periodic, but with random elements that vary from one period to the next (Randall, 2010).

**2. SIGNAL PROCESSING**

In order to accurately extract the differences in these periodic signals, the various structural components of the cyclostationary signal must be broken out and quantified. This process can be completed using a variety of different methods. The most common methods for doing this involve transforming the signal in question from the time domain into the frequency domain, often done through the use of the Fourier Transformation. While this method is very commonly used to extract information regarding the power spectrum of a signal, it is a relatively general approach. A more modern approach to extract the vibration data from the signal is known as the wavelet transformation.

**2.1. Wavelet Transformation**

The wavelet transformation was developed by Morlet et al. in 1982 (Mallat, 1999). This original wavelet, now known as the Morlet Wavelet is given by,

$$\hat{\Psi}(\omega) = \sqrt{2\pi}\omega^2 \exp\left(-\frac{1}{2}\omega^2\right), \omega > 0 \quad (1)$$

where  $\hat{\Psi}(\omega)$  is the frequency domain representation of the signal, which turns out to be a modulated Gaussian function. This basic function can be turned into a family of functions through the addition of a scaling parameter. In general wavelets are defined by,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), a, b \in \mathbb{R}, a \neq 0 \quad (2)$$

where  $\psi$ , or the mother wavelet, is the original wavelet basis being used,  $b$  is the translation parameter which determines the time domain location of the wavelet basis and  $a$  is known as the scaling parameter which determines the amount of compression in the resolution of the signal. The scale of the wavelet is its duration in the time domain. By changing the scale factor, the length of the wavelet in time is either increased by a larger scaling factor, or shortened by a smaller factor. This is evidenced by the effect that  $a$ , the scaling factor, has on Eq. (2). Because of the longer length of the wavelet when the scaling factor is low, it is able to gain information about the lower frequency content of the signal under analysis. As this scaling is decreased, and the wavelet is made shorter, higher frequency content can be captured.

After Morlet’s implementation of his Morlet wavelet, a number of other mother wavelets were investigated for a variety of uses. A number of typical mother wavelets can be found in Figure 1.

The wavelet transformation is a multiresolution analysis technique which analyses the data through a number of resolutions in order to ensure that data is captured at each “scale” (Pan, 2009). This means that the information at a

wide range of frequencies can be captured, while still maintaining the information contained within the time domain for the signal. Because of this, wavelets are considered to be an improvement over the commonly used Short-Term Fourier Transform (STFT) for non-stationary processes. This phenomenon is captured effectively by investigating the wavelet transformation of a ‘chirp’ waveform as found in Figure 2.

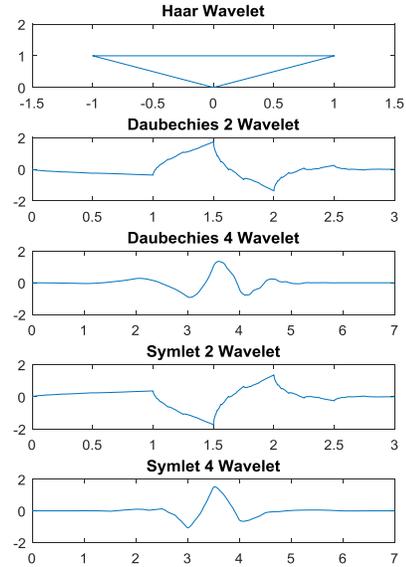


Figure 1. Examples of Various Wavelet Functions (The MathWorks Inc., 2016)

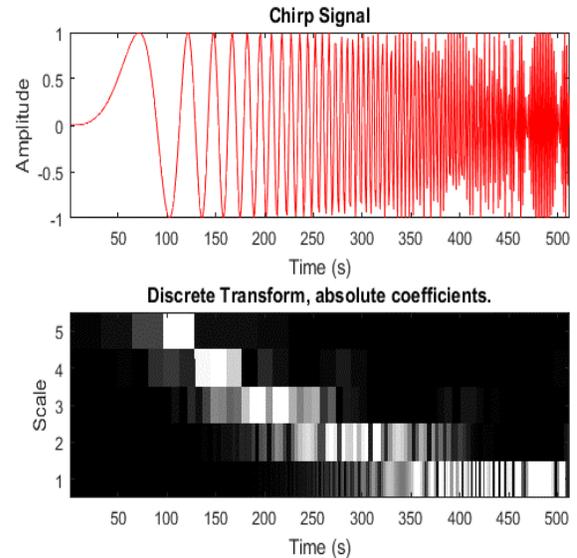


Figure 2. Wavelet Transform Plotted for a Chirp Function.

When analyzing this discrete transform, it becomes apparent that as the scale of the waveform is decreased, the frequency content that is captured increases. This creates a plot that shows the concentration of frequency content at a given scale over the time domain of the signal.

For the purpose of condition monitoring in this study, a selection of three mother wavelets were chosen for investigation in order to find the mother wavelet which supports the most accurate classification of nominal and damaged systems. The three mother wavelets chosen for investigation were Harr, Daubechies, and Symlet.

### 3. FEATURE EXTRACTION

When selecting the information that will be used as features in a machine learning algorithm for classification, the number of features must be limited appropriately. Increasing the number of features used to create a classifier means that the number of samples must also increase. Due to the limited number of samples in the data set, it would be very unrealistic to use each data point in the signal as a feature on its own. This means that a more concise feature vector must be used for classification purposes. In order to condense the information extracted from the wavelet transform of a signal, basic first and second order statistics of the resulting signal can be taken. This means that a more concise feature vector must be used for classification purposes.

When extracting the features for classification of the vibration data, first and second order statistics of the wavelet were used, referred to as Wavelet Statistical Features (WSF) (Jawahar, Babu, & Vani, 2014). For feature extraction from the wavelets used for condition monitoring, mean, standard deviation, minimum, maximum, skewness, and kurtosis were chosen, and formed into a feature vector for classification.

### 4. CLASSIFICATION

Classification algorithms learn from a set of feature vectors with known classes as the training data. The classifier can then predict the label of new unlabeled feature vectors, in some cases with very good certainty. Assumptions are made that the training data is normally distributed within each of the classes in order to maintain a simplified classifier model.

#### 4.1. Supervised Learning

Due to the seeded fault nature of the testing performed, there is readily available data on the current state of the system for each piece of training data. This means that a supervised learning classifier can be taken advantage of. Supervised classification can be performed more efficiently, allowing them to perform more complex and nonlinear categorization than their unsupervised counterparts (Izenman, 2008). One such supervised learning method, support vector machines (SVMs), has been widely used successfully for classification problems.

#### 4.2. Support Vector Machines

Support vector machines, a type of kernel machines, are a machine learning tool used primarily for binary classification. The algorithm was first proposed by Vapnik in 1963 (Vapnik & A., 1963). SVMs attempt to take a high dimensional input with some nonlinear dependency which results in an output. Because there is no prior knowledge of the interactions within the data, this classification must be done with only the information that is available in the training data. This is known as distribution-free learning. This makes SVMs opportune for datasets for which there is very little prior knowledge on the patterns within the data. In the case of the reciprocating compressor data, this is very helpful. Due to the frequency domain nature of the wavelet transformation, human intuition is not very useful in classification. This means that the distribution-free nature of SVM can be of large benefit when performing classification. The support vector machine methodology maps the input data vector into a Hilbert space, denoted as  $H$  in Eq. (3).

$$H_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (3)$$

This multidimensional space known as the feature space contains one orthogonal basis for each of the features being used for classification. The SVM algorithm then finds a multidimensional hyperplane which separates the training data, using this hyperplane to perform binary classification (Love, 2002). Due to the multidimensional nature of the SVM classifier being developed for these purposes, visualization of this hyperplane is not possible. For the purpose of explanation of the SVM algorithm, a model of just two dimensions will be used.

$$\sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T \mathbf{H} \alpha; \alpha_i \geq 0, \sum_{i=1}^L \alpha_i y_i = 0 \quad (4)$$

$$b = \frac{1}{N_s} \sum y_n - \sum \alpha_m y_m \mathbf{x}_m \cdot \mathbf{x}_s \quad (5)$$

The hyperplane is first optimized through the tuning of  $\alpha$  in Eq. (4) to reduce the classification error of the training data without overfitting the model by maximizing the value of the above equation. Because the data can be, and often is overlapping on the border between the two classes, a reasonable margin of error must be expected when optimizing the hyperplane.

Once the minimum classification error for the training data is found, a second optimization must occur. There are infinite hyperplanes which will reach the minimum classification error, and so the best of these hyperplanes must be chosen. In the case of SVM, the best possible hyperplane is defined as that which maximizes the margin between the hyperplane, and the data points within the dataset (Campbell & Ying, 2011). The data points which are closest to this hyperplane and define the maximum margin optimization are defined as the support vectors of the model, and are the data points which are most important for the definition of the model.

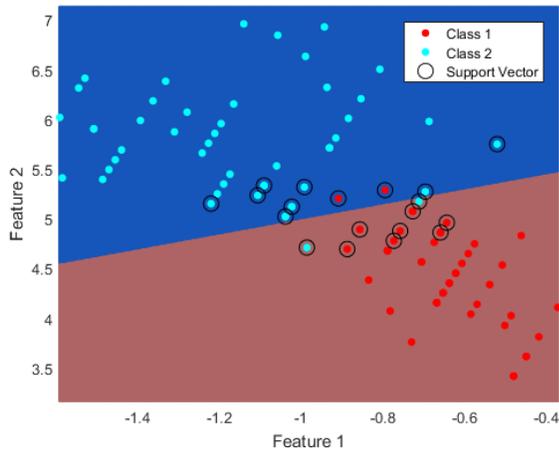


Figure 3. Support Vectors for a 2D SVM Classifier

The support vectors for a subset of the classifier used for the condition monitoring problem is shown in Figure 3.

This figure shows that the support vectors are those data points closest to the border between the two classes. Because this is just a subset of the feature space used to perform the condition monitoring, there is a great deal of overlap between the two classes. When SVM is extended to a multidimensional problem, this overlap becomes less prevalent, allowing for very accurate classification. Once an accurate model has been created for SVM classification, new data points can be classified through the prediction formula shown in Eq. (6).

$$y' = \text{sgn}(\mathbf{w} \cdot \mathbf{x}' + b) \quad (6)$$

This prediction takes the feature vector as an input, and produces a predicted value of the data points class. The interested reader is pointed toward numerous sources that exist on SVM theory, such as (Fletcher, 2008).

For the purpose of the condition monitoring of the reciprocating compressor, SVM has a few qualities which make it very well suited to the problem. SVM is resistant to local minima during optimization. This resistance comes from the nature of the optimization problem that SVM is attempting to solve. SVM presents a convex optimization problem, and so any local minimum is also a globally optimal solution (Kecman, 2005). This local minimum resistance makes SVM a very reliable and repeatable classification method to use. The secondary margin maximization also makes SVM a more robust algorithm to use. Also, because of SVMs use of kernels, the algorithm is able to model nonlinear relationships, which is very powerful, especially when dealing with overlapping datasets with rough boundaries.

### 5. DATA ACQUISITION

The data was taken from a series of seeded fault tests on a Dresser-Rand ESH-1 reciprocating compressor shown in Figure 4. This compressor was donated to Rochester Institute of Technology (RIT) and resides in RIT’s Compressor Test Cell. A series of accelerometers, rotational encoders, and pressure transducers were affixed to the compressor and connected to a National Instruments CompactDAQ. The data which was processed for classification purposes was taken from an accelerometer mounted to the crank-side discharge manifold of the compressor.

Seeded faults were added to the system to simulate deviations from nominal operating conditions. The fault condition which was monitored involved wear in the springs within the

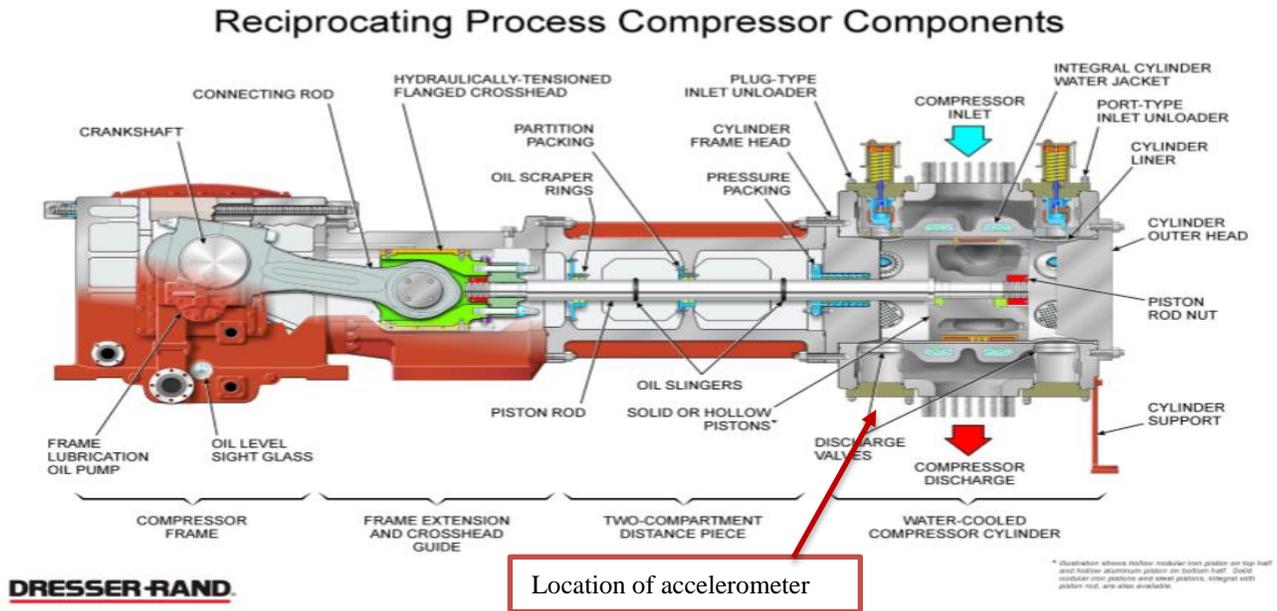


Figure 4. Reciprocating Compressor Cross-sectional Diagram

poppet valves of the compressor. These valves are used within the discharge valve assembly, and are pivotal for the proper operation of the reciprocating compressor. The poppet valve prevents compressed air from flowing backwards, out of the pressure tank, allowing the unit to continuously maintain pressure with every stroke of the piston. A cross sectional view of a typical reciprocating compressor can be found in Figure 4.

Spring fatigue was simulated by replacing the stock springs, considered nominal, with a softer spring to simulate a worn condition, and no spring at all, to simulate a completely broken spring. These springs were placed in a variety of configurations, as shown in Table 1 and Figure 5. These configurations placed springs in different conditions in each quadrant of the valve. Cases 1, 3, and 5 have every spring in the valve at the same level of wear, while cases 2 and 4 are blended conditions, with each half of the springs in a different condition.

Table 1. Seeded Spring Fault Cases

Case	Group 1	Group 2	Group 3	Group 4
1	Nominal	Nominal	Nominal	Nominal
2	Nominal	Degraded	Degraded	Nominal
3	Degraded	Degraded	Degraded	Degraded
4	Degraded	None	None	Degraded
5	None	None	None	None

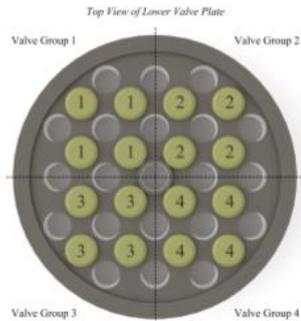


Figure 5. Valve Groupings

By seeding multiple cases for a binary classifier, a range of data is created. This ensures that there will be data points closer to the boundary line that SVM must determine. For the case of the classification problem, the cases shown in Table 1 are sorted into a binary problem, Nominal vs. Damaged. Nominal is only those points of data which were taken with all nominal spring sets, where the damaged datasets encompasses the rest of the points. A binary problem was chosen over multiclass SVM because the most critical

decision making point when performing condition monitoring is whether or not the system is performing within a nominal range.

6. RESULTS AND DISCUSSION

6.1. Collected Data

Due to the cyclic nature of the reciprocating compressor, data was broken into individual cycles of the compressor piston. The data was then taken from the time domain into the angular domain, where the data was plotted over the angle of the piston crankshaft. Figure 6 displays all of the data considered, plotted onto one figure. Both the pressure in the cylinders, and the accelerometer data used for classification are shown in the figure.

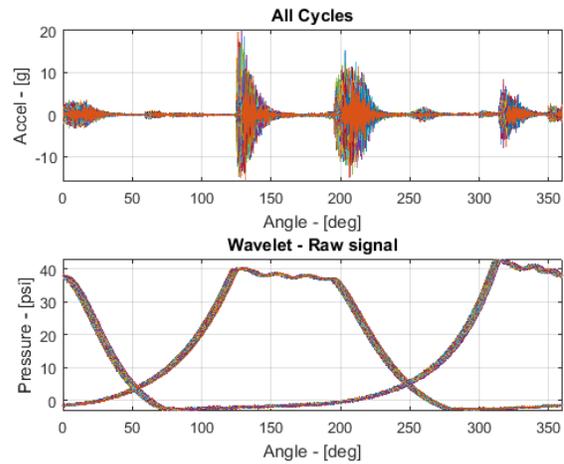


Figure 6. All Cycles after Angular Transformation

6.2. Wavelet Selection

In order to maximize the correct classification percentage, a variety of different mother wavelets must be considered and the best must be determined. Different wavelets are able to extract the information from a signal differently. In order to select the best wavelet for the reciprocating compressor vibration analysis, data was processed using a variety of different mother wavelets. This data was then classified using the same binary SVM classifier and kernel. The classification error of these various methods was compared in order to select the best performing wavelet. For the reciprocating compressor vibration analysis, the Haar wavelet, Symlets of order 2 and 4, and Debauches wavelets of order 2 and 4 were considered. Though all of the wavelets chosen for investigation performed well with above 90% classification accuracy, the Debauches wavelet of order 4 (db4) performed best, achieving a 91.8% classification accuracy. Therefore this wavelet was chosen for further investigation. An example of a multi-level wavelet transform performed on the periodic reciprocating compressor data is shown in Figure 7.

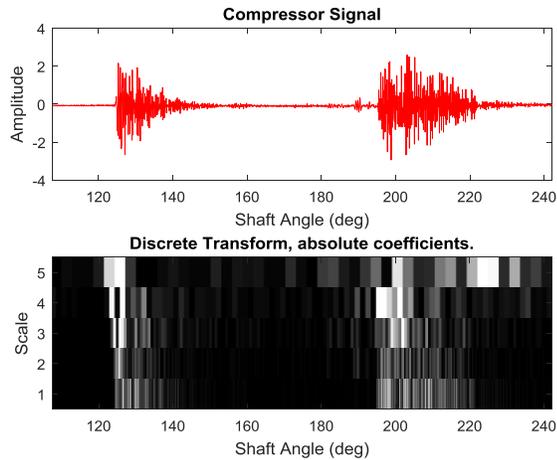


Figure 7. Wavelet Transformation of Compressor Vibration Signal

For the purpose of visualization, the data was zoomed to the range of interest. For this data, the plot is zoomed to the area ranging from 120 to 240 degrees. This range makes intuitive sense because it is the range during which the valve with the sensor affixed and seeded faults injected opens and closes. The lower axis of Figure 6 shows the pressure with respect to the shaft angle. As the center waveform peaks, the valves open. This creates the first accelerometer spike on the upper axis. This happens predictably just after 120 degrees on the shaft angle. The valves then close, with most of the vibrations dying out by 240 degrees on the shaft angle.

### 6.3. Feature Extraction

From the wavelet transformation of the data, the wavelet features were extracted. These features were placed into a vector which would function as the input to the classifier for each cycle of the shaft. Due to the dimensionality of the feature vector, visualization of the full feature set is not possible. A selection of features shown in Figure 8 shows that there is indeed separation between the different classes in the data.

### 6.4. Support Vector Machine Kernel Selection

In order to perform the binary classification problem, the training data was classified such that the nominal case (NNNN) was placed in the 'Nominal' class, while all other cases (SSSS, WSSW, WWWW, WNNW) were placed in the 'Damaged' class. This created two binary classes. Using the db4 wavelet, a number of different SVM kernels were tested, including linear, quadratic, cubic, and fine, medium and coarse Gaussian kernels. The accuracy of these different kernels was used to determine the strongest for this dataset. The investigation of the different SVM kernels led to the selection of the cubic kernel for further investigation.

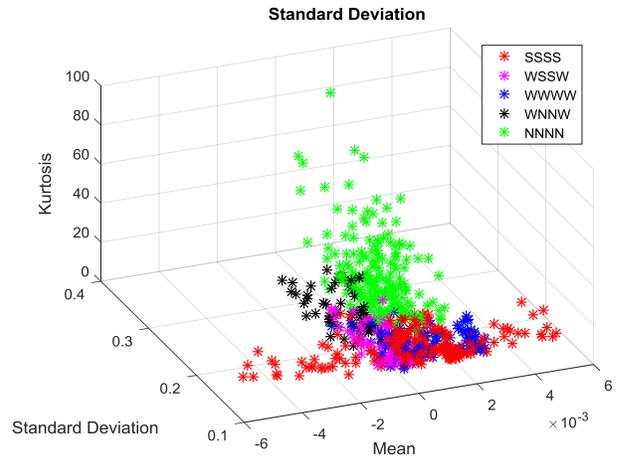


Figure 8. Feature Vector Comparison of Different Classes

Though all the kernels performed well with the dataset, achieving scores above 93%, the cubic kernel achieved an accuracy of 95.3% on the training data, making it the most accurate. The training set used was a randomly selected set containing 50% of the total data taken.

### 6.5. Support Vector Machine Kernel Selection

This classifier was then used to perform predictions on a new and randomly selected set of test data. The classifier performed very well achieving a prediction accuracy of 97.55% on the test set. The test set used was the remaining 50% of the data which had not been used for training the classifier, thus ensuring that the model was being tested only with fresh data. The confusion matrix for the test set of data is shown in Figure 9.

## 7. CONCLUSION

In order to create a condition monitoring system for a reciprocating compressor, a vibration analysis approach was developed and investigated. The use of wavelet transformations to process data, along with basic statistical analysis created a feature vector that could be used in conjunction with Support Vector Machines (SVM) with over 97% accuracy. Condition monitoring technique like those explored in this work, in conjunction with regular maintenance, can easily increase the efficiency of a production line by limiting maintenance to scheduled intervals.

Overall the approach developed was quite promising, with high accuracy even on relatively non-complex features. Further identification of important features within the wavelet data may serve to increase the accuracy of this method even further, resulting in nearly perfect classification.

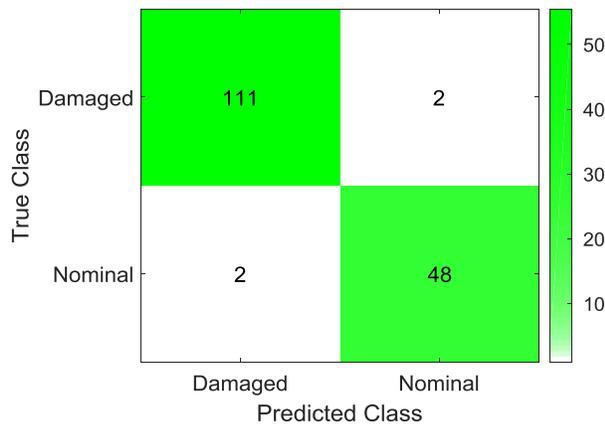


Figure 9. Confusion Matrix

#### ACKNOWLEDGEMENT

This work graciously acknowledges the continued financial support provided by Dresser-Rand and its partnership with RIT.

#### REFERENCES

- Bloch, H. P., & Hoefner, J. J. (1996). *Reciprocating Compressors - Operations and Maintenance*. Elsevier Inc. doi:<https://doi.org/10.1016/B978-088415525-6/50004-3>
- Campbell, C., & Ying, Y. (2011). Learning with Support Vector Machines. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 1-95.
- Feng, K., Jiang, Z., He, W., & Ma, B. (2011). A Recognition and Novelty Detection Approach Based on Curvelet Transform, Nonlinear PCA and SVM with Application to Indicator Diagram Diagnosis. *Expert Syst. Appl.*, 38(10), 12721-12729. doi:10.1016/j.eswa.2011.04.060
- Fletcher, T. (2008). *Support Vector Machines Explained*.
- Izenman, A. J. (2008). *Modern Multivariate Statistical Techniques - Regression, Classification, and Manifold Learning*. New York: Springer.
- Jawahar, M., Babu, N. K., & Vani, K. (2014). Leather texture classification using wavelet feature extraction technique. *IEEE International Conference on Computational Intelligence and Computing Research* (pp. 1-4). Coimbatore: IEEE.
- Kecman, V. (2005). *Support Vector Machines - An Introduction*. Berlin Heidelberg: Springer. doi:[http://dx.doi.org/10.1007/10984697\\_1](http://dx.doi.org/10.1007/10984697_1)
- Lin, Y.-H., Wu, H.-C., & Wu, C.-Y. (2006). Automated condition classification of a reciprocating compressor using time-frequency analysis and an artificial neural network. *Smart Materials and Structures*, 15(6), 1576-1584.
- Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, 829-835.
- Mallat, S. (1999). *A wavelet tour of signal processing, 2nd Edition*, . Academic Press.
- Pan, G. (2009). Intuitive Introduction to Wavelets. In G. Pan, *Wavelets in Electromagnetics and Device Modeling*, (pp. 15-29). Hoboken, N.J.: John Wiley & Sons, Inc.
- Prajapati, A., Bechtel, J., & Ganesan, S. (2012). Condition based maintenance: a survey. *Journal of Quality in Maintenance Engineering*, 18(4), 384-400. doi:<https://doi.org/10.1108/13552511211281552>
- Randall, R. B. (2010). *Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications*. Hoboken, N.J.: John Wiley & Sons, Inc. doi:<https://doi.org/10.1002/9780470977668>
- Trout, J. N., & Kolodziej, J. R. (2016). Condition monitoring and classification of reciprocating compressor valve faults using time-frequency analysis and image-based pattern recognition. *PHM Society*, (pp. 1-10). Denver, CO.

#### BIOGRAPHIES



**Shawn Falzone** is a systems engineer with the Aircraft Group at Moog Inc. in East Aurora, NY, specializing in flight control systems. He completed a dual B.S. and M.S. degree in electrical engineering with a focus on control systems at the Rochester Institute of Technology, in 2017. His current research interests include a wide range of systems modeling and test automation techniques.



**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 RIT's prestigious Eisenhart Provost Award for Excellence in Teaching.