Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 34 (2019) 876–884

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

47th SME North American Manufacturing Research Conference, Penn State Behrend Erie, Pennsylvania, 2019

# Vibration Analysis Utilizing Unsupervised Learning

Ethan Wescoat[a], Matthew Krugh[a], Andrew Henderson[b], Josh Goodnough[c], Laine Mears[a*]

[a]*Clemson University, Clemson, SC, USA*
[b]*Praemo, Inc Greenville, SC, USA*
[c]*BMW Manufacturing Co., Greer, SC, USA*

* Corresponding author. Tel.: +1-864-283-7229. *E-mail address:* mears@clemson.edu

**Abstract**

Many manufacturing environments have implemented methods of collecting data from their processes relating to vibration, temperature, or sound. With the data stored, manufacturers can run analytics to plan maintenance schedules and track machine health. However, in many cases, these maintenance schedules and health tracking are largely reactionary, largely implemented through experience rather than through predicting the onset of critical events and taking measures to prevent them. This paper describes a case of using time series data analytics of vibration from an automotive paint shop PVC dispensing pump (doser) attached to a robot using a novel combination of unsupervised learning and feature extraction. The goal is the determination of healthy versus unhealthy data and the implementation of predictive maintenance on the machine cell. Since the robot is a multi-axis robot, direct application of traditional health monitoring methods is lacking; instead a combination of methods suitable to the multidimensional nature of the robotic pumping process is employed. The goal of the first phase of the project is to build the tools to aid in this feature extraction using unsupervised learning and begin to establish a baseline of healthy data versus unhealthy data or fault data. The doser cell has been monitored for six months gathering data from seven sensor sets. Traditional methods of data analysis such as spectral analysis through Fast Fourier Transforms (FFTs) were used to establish the capability of reading vibration signals before moving to feature extraction of the time series data. For feature extraction, a Gaussian Mixture Model is utilized for the learning and the model building. These methods utilized not only determine the vibration of each specific component, but also help differentiate between the nozzle flow rate and angle. In extracting these features from the data, patterns can be traced from the variation of each production process and differentiation can take place based on what is healthy and unhealthy data. The goal of the continuing process phase is to inform the predictive maintenance function to improve equipment uptime.



*Keywords:* Data Analysis, Condition Based Monitoring, Unsupervised Learning

## 1. Introduction

Today's manufacturing industry has become quite data-heavy, leading to an expected 1 petabyte daily data generation in a smart factory by 2020 [1]. Companies are left questioning the proper tools and methods for making use of data. Manufacturing processes accumulate data quickly, collecting from people, machines, materials, and systems. The task is then to translate that data effectively to useful information for making decisions. Traditionally for sequential quantitative scales, analysis has been performed using either time series analytics, frequency analytics, or a combination of the two [2]. As data generation has exponentially expanded with the ease of sensing, the amount of available data vastly exceeds the capability of timely and traditional analysis after the initial collection. To address this issue, unsupervised learning has been introduced as an autonomous method in representing time series and frequency data, in order to analyze phenomena more effectively. Upon creating a fully functional model, unsupervised learning can rapidly detect trends and patterns to derive analysis useful to the operator or engineer of the process. More importantly, the key objective is to provide data to


10.1016/j.promfg.2019.06.160

operator and engineer when machine faults (degradation or impending catastrophic failure) are imminent. Unsupervised learning will continually teach itself and update as the model gathers more data. Sharp deviations can be treated as anomalous data and acted upon, thereby saving time and money.

Hardware to support this analysis approach is a coupled issue; cyber-physical systems and cloud computing have emerged in recent years to support monitoring and analyzing these large data sets as well as communicating data to both machines and people [3]. Cyber-Physical Systems (CPS) are hardware/software platforms for facilitating gathering, formatting, communicating, analyzing, and sharing information within the system. This integrated analysis and information sharing makes CPS useful to manufacturers.

Data collection as a first step has rapidly expanded using newly-available, low-cost, off-the-shelf sensors implementable on the shop floor. Using popular microcontrollers such as the Raspberry Pi or BeagleBone Black to collect data and communicate to a cloud database or a front facing visualization space is easily implementable [4]. Utilizing data protocols such as MTConnect or MQTT, the standardization of data from multiple sources allows the user quick access to make informed decisions [5]. With this end goal in mind, here we lay out details for the analytical phase of data.

The objective of this paper is to introduce a different approach from traditional methods for analyzing data from a manufacturing process. The paper focuses on vibration and health analysis of a high-use critical doser pump system with data gathered using condition-based monitoring. The manufacturing process being monitored is a PVC pump used for dispensing sealant to the joints and seams of an automotive body-in-white. This paper focuses on the time series analysis utilizing a method of *k*-means analysis, variance analysis, and peak-to-peak data [6]. The objective of this analysis is to establish trends in the cycle time, investigate the underlying signatures of data patterns, and use that information to determine when faults will occur to plan actions under a predictive maintenance program.

## 2. Literature Review

### *2.1. Condition Based Monitoring*

Condition based monitoring is not a new technique. It was formally introduced in the 1970s in the aerospace and petroleum industries specifically for gathering vibration data [2]. Condition-based monitoring is characterized by a predictive rather than preventive approach, specifically the idea of identifying component or system degradation that could lead to unexpected failure, and planning maintenance schedules around this knowledge. Preventive Maintenance (PM) has relied on the expected machine failure of equipment based on past uptime rather than active use over time. However, this method assumes the machine is back to “100%” work status. Rather, whenever repairs occur the machine does not go back to a perfect machine status or “as if new” [7]. Predictive maintenance however takes into account changes in equipment behavior over time, as characterized by signals from sensors, rather than assuming perfect functionality [8]. Predicting when a machine shall fail allows for managers and engineers to schedule downtime around an event to ensure the machine is back up and running in the smallest amount of time [8].

Figure 1 below highlights the shift of manufacturing through the past two centuries. Beginning with mechanical steam process first, as each successive industrial revolution occurs, more mechanization and automation has taken place. This leads to an increase in the need for sustainability of the work environment these machines are operating [9].

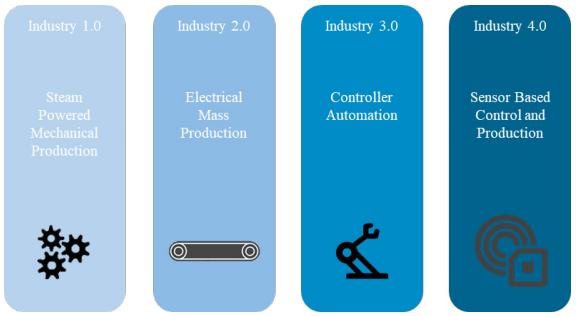



*Figure 1: The Evolution of Manufacturing through the years*

Condition-based monitoring has also given rise to an increased push in manufacturing for increased integration of sensing and data sharing (Industry 4.0 / Smart Manufacturing paradigm). Lee notes that cyber-physical systems are implemented in advancing Industry 4.0 referencing Figure 1 following a 5C principle of Configure, Cognition, Cyber, Conversion, and Connection [3]. Using these 5C devices for computation and communication increases the amount of computation power distributed throughout the factory. It is possible to perform all the analysis as well as communicate alerts to operators and maintenance staff when a problem has occurred. By differentiating between healthy data and faulty data, maintenance schedules can be determined based on the likelihood of a failure event [10]. Healthy data is initially recorded to determine when the machine is operating within normal parameters to determine a baseline. This baseline data allows subsequent data comparison in real time to determine if the machine is continually healthy, and the expected limits of that health in the measured dimension(s) [8]. Deviations of the data would signal that the machine, or a component has changed (assumedly for the worse). However, data from one sensor does not always readily indicate failure, especially for a complex process or machine. Sensor Fusion or other forms of multisensor analysis using multiple independent inputs allow more data to contribute to the baseline belief. This concept applies not just to multiple sensors, but also to multiple sources of information including the machine itself and people who interface with it.

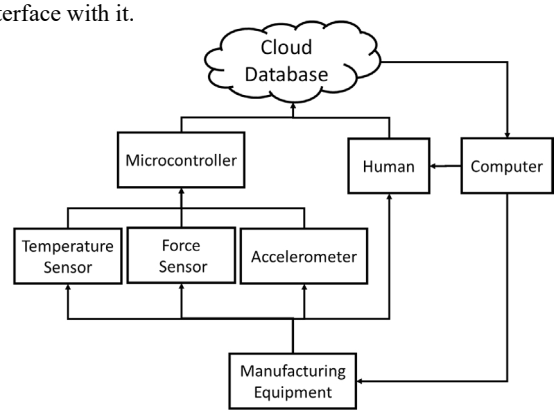


*Figure 2: Example data flow of a connected machine to a cloud database.*

Figure 2 shows an example of multiple sources of data collected from the machine and pathways for data communication. Multiple avenues of data allow for more complete analysis, resulting in a more accurate representation of the health of the machine and a means of communicating this information to the operator. Data flows bidirectionally to gather and then compare before it is stored and then communicated back to the user. Humans can also contribute data based on manual input of environmental factors. With the rise of new sensor technologies and lower costs, leading to increased collection of data, different methods can be utilized to store and process the data.

Cloud computing is a method of storing, analyzing location and communicating data across a wide network. Cloud computing, as defined by NIST, is characterized by on-demand self-service, broad network access, and resource pooling [11]. The cloud also provides a way to centralize large analytic tools needing high computational power. Commercial products such as Amazon Web Services and Microsoft Azure provide consumers and businesses with commercial cloud services; the main issues with such data security as well as managing energy expenditure [12]. Other techniques derived from cloud computing are edge and fog computing, which offer different location alternatives of where the data storage and analysis could occur.

Fog computing is a form of cloud computing that takes place closer to the physical processes on the network and locally spreads out the computational analysis of a data process [13]. The crucial difference between cloud and fog computing methodology is where the data is being processed and stored. The cloud serves as a "central hub" on the network, operating purely on the network whereas fog computing happens on multiple devices at the edge of the network, the idea being decentralized devices working on different type of computations, storing or sending the data to other parts of the network [14]. Fog computing suffers from limited storage capacity and coordination of all the devices working in conjunction with each other [14]. Utilizing the cost effectiveness of sensors and such computation methods, companies can handle the transference of data from machine to operator. Problems arise in terms of applications and the question of which method is better for a given situation.

### *2.2. Vibration Analysis and Feature Extraction*

Vibration analysis involves different methods of looking at acceleration data corresponding to machine shocks and feature extraction. For time series analysis, researchers look to the peak to peak output voltage of a transducer and determine variations against a prior model of a healthy machine, as well as the raw base level value, represented by the signal root mean square (RMS). RMS analysis is good for detecting systemic failures, but has trouble identifying failures with specific components. Two common measures of signal summary are the Kurtosis value and Crest Factor, measures which examine the distribution and threshold values of the signal, and the ratio of the expected and recorded value respectively; these are considered part of the time series analysis [15]. When performing this analysis, specific features are selected which contribute to certain instances in a machine's cycle. This is considered feature extraction and corresponds specific shapes and patterns in the data to determine if a machine failure. In the frequency domain models take shape using Fast Fourier Transform (FFT) and Wavelet Packet feature extraction and decomposition [15][16][17]. Feature Extraction also occurs in the time series data and has been used as a prediction analysis for others [18]. It congregates around similar scalar values relating to certain instances of this data. Pairing a form of feature extraction from time series analysis or frequency analysis with unsupervised learning would allow the model to update autonomously as more data is added, rather than having to reanalyze the data each time. This would allow for quicker fault detection and reduce the computational power required of the system.

### *2.3. Unsupervised Learning*

Traditional data analysis is performed typically over a finite data set [19]. The data must also be structured in a specific way for the analysis to take place. This usually comprises such methods as the Fast Fourier Transform and the Hilbert Huang Transform. There are other common methods as well for vibration analysis over multi point data [15]. Over smaller sets of data, these methods work well. However, these methods don't adapt or change over time, precluding their application to larger data sets and nonstationary systems. They are meant to run repetitively, performing analysis as each set of data comes in. However, in a complex process such a multi axis robot simply analyzing each individual frequency point is not a valid solution. When looking at the variation in the signal, there is no discernible trend looking at the frequency spectrum during an individual production process. Specifically looking at the robot doser pattern, there is too much variation regarding the spraying patterns of the robot and the angle that sealant is being applied to create a fixed model of the system behavior. Instead of breaking the production process into individual points, the entire process must be analyzed, and then specific features can be extracted.

Unsupervised Learning involves detecting patterns and extracting features among the measured signals through a method of learning that can either be autonomous or semi-autonomous. Unsupervised learning has been used to cluster word documents in large databases or organizing large amounts of big data [20]. Some popular uses of Unsupervised Learning involve self-organizing maps and k-means clustering, which make use of nodes that data clusters around that are similar to each other [20]. Self-Organizing Maps (SOM) allows as many reference vectors to be chosen irrespective of the data. These reference vectors then formulate the clusters that enable the unsupervised learning. Self-Organizing maps are used to take higher dimensional vectors onto a lower 2-D space. *k*-means assigns each observation/event to a cluster based on its proximity to that cluster's mean. The algorithm iteratively calculates a new cluster mean and an event's vector distance from that mean until the sum of squared errors (SSE) are

minimized. *k*-means also requires that the number of clusters (*k*) be defined upfront. As more clusters are generated for *k*-means, the slower the clustering occurs. For datasets with higher amounts, self-organizing maps are a better solution.

Gaussian Mixture Modeling (GMM) is utilized to perform the analysis below. GMM affords more with regards to flexibility. The data points can be considered gaussian distributed and a better cluster covariance [21]. The methodology follows like k-means, by first choosing several clusters. The probability that each point belongs to a cluster is calculated and then is assigned to the cluster that data point most matches. Applying that to manufacturing and looking for trends and patterns in the data could allow an operator to know the present cycle step from simply looking at the data flowing from the process.

Each of the seven tested sensors in this study can collect 20GB of data each week, for a total of 140 gigabytes of data a week in the test environment. Over the course of 6 months of testing, more than a terabyte of data has been passed. With regards to the scalability of the situation, the amount of data would vastly exceed any local storage or time sensitive analytics such as fault detection. This is the reason behind building the model with unsupervised learning and specifically using the Gaussian Mixture Model. More flexibility is afforded when utilizing this analysis.

For analyzing large sets of data, the need for parallel data analysis becomes apparent. Parallel data analysis distribute the analysis across many computer clusters or data platforms [22] With data that has been gathered from as many sources and then stored in those sources, utilizing parallel analysis speeds the up the overall analysis, sending to the model the key features gathered after the data has been analyzed locally. Rather than sending up the billions of data points, the completed analysis is sent for the overall model to make the decision. The need for dedicated computer clusters and parallel analysis therefore becomes a growing need to aid in this distributed analysis. In a shop or production facility where automated processes make up most of the production line, such level of computational hardware and software may become a necessity.

## 3. Methodologies

### *3.1. The Test Environment*

Data were collected from a sealant dosing pump carried by 6-axis robot arm in the paint shop of an automotive manufacturing facility. While in operation, this doser is subjected to vibrations from multi axis motions while dispensing PVC to the body of the car. Data show a tendency for this machine to break down from continuous vibration in key areas causing failure. While this is the most common failure, there are also failure cases such as material contamination. To gather this data, the vibration signal of the doser was captured using accelerometers mounted in a rigid frame to seven different areas of the pumping system. These seven areas are: the left ball screw case and the right ball screw case, the left and right inlet valves, the left and right outlet valves and the nozzle (see Figure 3 for a nozzle-mounted sensor system).

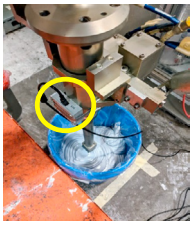


*Figure 3: The box circled in red is the sensor box containing accelerometers mounted to the nozzle of the robot.*

Like the nozzle mounting, 6 other boxes are also fixed to these locations similarly in the robot/pump system.

| Sensor # | Location | Side of the Robot |
|---|---|---|
| 1 | Nozzle | NA |
| 2 | Inlet Valve | Left |
| 3 | Inlet Valve | Right |
| 4 | Outlet Valve | Right |
| 6 | Bearing Case | Right |
| 7 | Outlet Valve | Left |
| 8 | Bearing Case | Left |

*Table 1: Device Number and Location*

Table 1 above highlights the naming convention while Figure 4 shows the robot in the production environment and highlights the location of all sensors. A typical production cell would have 4 robots.

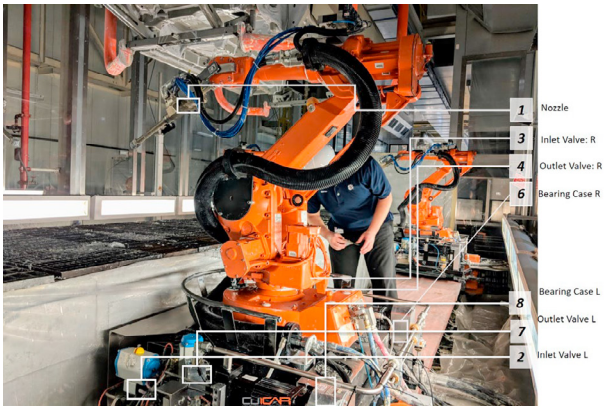



*Figure 4: Sensor Placement in Production Cell*

The boxes containing the sensors were constructed from aluminum to protect them from the harsh environment and ensure good vibration transmissibility from the underlying structure. The sensor kit, which met the compliance of the paint shop environmental, health and safety requirements, consists of a microcontroller, a real time clock and accelerometer (see Figure 5).

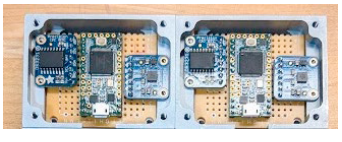


*Figure5: Sensor boxes used to gather data*

These components were chosen based on cost, off-the-shelf accessibility, hardware capability and software capability. The criteria for choosing the sensor involved sampling rate, cost, and the ability to sample in multiple axes. The ADXL335 met all of these design criteria. The real time clock utilized partnered well with the accelerometer used. The Teensy microcontroller was chosen based on both the size constraints as well as the ability to pass serial data. Having a separate microcontroller also allows for future sensor inputs to be added such as temperature and acoustic emission. The real time clock ensures the data are collected at specific time intervals and can be synchronized with each other. Data values corresponding to specific time instances can be related to events within the machine cycle. The sensor box samples at 500 Hz, as the dominant frequencies of operation were below 250Hz.

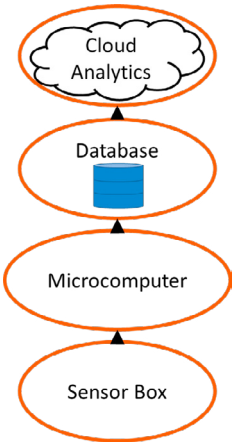



*Figure 5: Hierarchy of Devices in the test environment*

Figure 5 shows the hierarchy of the system from the data collection to the overall analysis level. After buffering, data are passed to a secondary collector microcontroller (Raspberry Pi) mounted outside the process through USB cable. To this end, the local collector microcontroller is used as the database and control for now but will soon be replaced by direct access to plant IT systems.

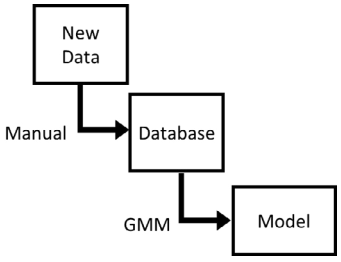



*Figure 6: Method of Data Collection and Analysis*

Figure 6 refers to the method of collecting data and then placing it into the database. The manual method refers to uploading the data to the higher database. The GMM refers to the method Gaussian Mixture Modeling used for building the model. The following Equation 1 is the base method for clustering Gaussian Mixture Models:

$$G_j(x) = \Phi_{\theta_j}(x), \theta_j = \left(\mu_j, \sigma_j^2\right) \quad (1)$$

In Equation 1, the *G(x)* refers to the gaussian density that will eventually be centered around the mean and variance given by $\theta_j$. This is what handles our clustering. The collected data (the new data box in Figure 6) are analyzed together with the historical data (the database box in Figure 6) to create a representative behavioral model which is updated through machine learning based on the new data. The analysis is performed at the top-level hierarchy in the cloud rather than the local level at either the secondary collector or the sensor box. Having the analysis occur in the cloud provides an "unlimited" computational power to access. At the local level, analysis is limited by the chip used for the microcontroller, due to the memory and graphics specifications. Cloud level analytics also allow for access to all data in a database, instead of limitation of local storage limits. The limit again is the amount of storage allocated for the project.

At the local level, a microcontroller is limited to the amount of storage available. Analysis can, however, occur at the local level with such embedded systems. Having the analysis distributed reduces the need for computational power to run the overall analytics at the cloud level. The microcontroller could also be utilized to run a data visualization module. This would provide the operator with the local analytics of a particular machine and a sensor area without having to draw in the entire model set from the cloud. Partnering this system on the cloud also allows access to the production data of the manufacturing system and correlation between good or bad parts can then be validated. The cloud analytics only then need to focus on the pattern analytics rather than the conversion of raw data to analyzed data. This reduces the need to have a huge database of unanalyzed data in addition to the analyzed data. This allows only relevant information that can provide actionable change to remain. Superfluous data can be discarded at the local level without having the cloud attempting the filtering. The model learning then updates with the new data added. Currently, the system is gathering data to build the model learning and beginning the event classification.

## 4. Results

Data were first collected from an offline test cell. This initial data gathering was done to simulate the signals expected from each of the key areas and to indicate the measurability of vibration readings from each area. Overall, 7 tests were initially run to simulate each flow pattern. Table 2 refers to each situation that was tested. The test was run over a purge cycle of the robot, during which the robot dispenses sealant inside of a bucket. The differentiator between each test is the nozzle angle and then whichever side is dispensing the sealant. The side column refers to which side is running. The event time for each test took around 100 seconds in dispensing sealant.

| Test # | Nozzle Angle | Side |
|---|---|---|
| 1 | 0 | Right |

| | | |
|---|---|---|
| 2 | 0 | Left |
| 3 | 45 | Right |
| 4 | 45 | Left |
| 5 | 90 | Right |
| 6 | 90 | Left |

*Table 2: Sealant Tests and* Description

Frequency analysis was performed on each set of test data in order to characterize the frequency content. A model is not being created at this point. Simply, there needs to be validation that vibration data can be taken from these key areas. In addition to moving in multiple dimensions, the sealant doser also applies sealant in different patterns and angles. The flow of sealant can vary greatly from one type of nozzle to another.

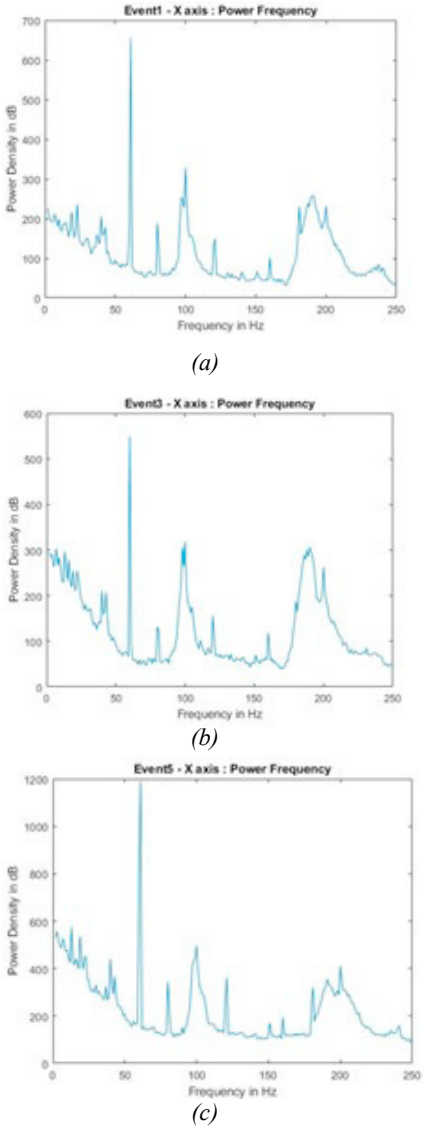



*Figure 7(a)-7(c): FFT results of each angle of sealant tests*

Figures 7(a) – 7(c) show the frequency content of three of the different tests on the inlet valve on the M1 side of the robot. They are tests 1,3 and 5 as referred in Table 2. In Figure 7(a) the robot was spraying at a 0-degree angle. Figure 7(b) was at a 45-degree angle, and Figure 7(c) shows spraying at a 90-degree angle. Each test shows a consistent peak of around 60 Hz. The variation in the signal depending on spraying pattern comes in the higher and lower frequency range based around peak. For the 0-degree angle and 45-degree angle are similar in the lower frequencies. The 0-degree angle has a higher power density in the higher frequency range. The 45-degree test and the 90-degree test have a similar higher frequency look, however the lower frequency power density for the 90-degree test is less than the 45-degree test.

Looking at the test results, the frequency analysis validated the ability to see changes in the system with regards to flow rate and nozzle angle. However, in terms of differentiating production process, the frequency analysis failed to help identify the different production processes. As a whole system health could be ascertained, however, variation is difficult to determine. Therefore, for the model learning, it was determined to utilize the variance analysis of the time series data to allow for each production process to be utilized.

This led the research team to utilize the time series analysis to examine the variance of the signal versus the frequency spectrum utilizing *k*-means and eventually a Gaussian Mixture Model. The goal will be models that update continuously in real-time and estimate the probability of a disruptive event. These disruptive events need to be identified or classified through the data. Once a data set with classified events is available, then various supervised training processes are widely available, such as support vector machines, for training the models. The time-consuming portion of the model building process (also the current status of the project) is collecting data and correctly classifying it. This is typically the work of skilled subject matter experts. Thus, the focus here is establishing techniques and tools to use for automating event classification. A secondary goal is providing the tools for consumers to repeat the learning on other equipment.

Raw signals from sensors, like accelerometers, do not have flags to indicate what an event is or when it starts and stops. This is where “traditional” forms of data analysis step in, such as the FFT and HHT. However, these processes cannot handle massive amounts of data continuously. A Discrete Fourier Transform, for example, requires a finite amount of time set to process and generate analysis. It does not function well to continuously add data and then run another test continuously. The data set cannot simply be added to the database and allow the learning to take place autonomously.

An example of the raw vibration data is shown in Figure 8 below. Signal processing tools are needed in order to extract data related to individual events. Since this is a valve, an event is defined as any time that the valve operates. Or, in other words, whenever the valve opens or closes. It is assumed that the accelerometer will sense vibrations that are greater in magnitude when the valve opens or closes than when it is resting in a static position. Over time and use, these peaks should become more pronounced with a higher amount of variability. This variation could occur with regards to either higher peaks or occur at different times over the event space.

Therefore, events are identified by peaks in the raw sensor signal and by increases in the data's variability over a short window of time.

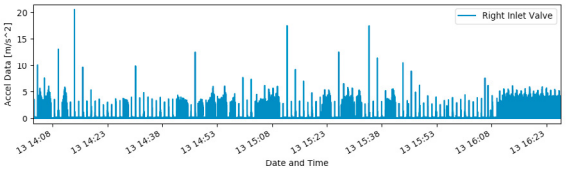



*Figure 8: Raw data from accelerometer on valve*

It can be seen from the data in Figure 8 that there are multiple types of these events, and they occur all throughout the spectrum of data and not in an organized manner or even a cyclic manner. Unsupervised learning is introduced to enhance the analysis, specifically through a learning technique known as *clustering*, used to identify different types of events. The vectors comprise descriptive features of the events.

Event duration and area under the curve were chosen as the features to describe these events for the subject application. Therefore, the duration and area under the curve were calculated for each event that was extracted from the raw signal and these features are used to establish clusters. A Gaussian Mixture Model, a variation of k-means, was used for this learning. Since $k$ needed to be defined up for the GMM to work an elbow plot was generated to determine optimal values for $k$. Multiple values of $k$ were used but ultimately improvements to the SSE were minimal beyond $k = 3$.

The points representing each event in feature space is shown in Figure 9, which also shows the cluster identifiers. Qualitatively, there is more variability within cluster 3 than the other two clusters. In Figures 10(a) – 10(c), each event is plotted based on the cluster it belongs too. The data in clusters 1 and 2 seem to follow a set path consistently with some variation at the peaks. This gives a consistent area under the curve which supports the clustering seen in Figure 9. For Cluster 3 the data seem to follow the first initial peak and then split off with peaks occurring at different instances from then out. The amount of variability also supports the clustering as seen in Figure 9. This seems to support that clustering was done successfully. Out of the 182 events clustered, only 1 event failed to fit into a given set. However even with this one event, a probability is obtained of the event and then used to estimate which cluster it should go to based on the other clusters.

Quantitatively, each cluster event follows a formula of acceleration with regards to time. The area under the curve is determined with regards to the following formula:

$$Velocity = \int_{t_0}^{t_f} a(t)dt \qquad (2)$$

The acceleration, $a(t)$, plotted is with regards to the vibration monitored during the production process. Each event is considered an individual production process of dispensing sealant. The event times, $T_0$ and $T_f$, also match up to the beginning and end of the production process of the doser robot for apply sealant. Therefore, they were chosen as the start and end times of each graph. The total event duration, then corresponds to the actual time duration of dispensing sealant to a car body. This was obtained by collecting data from the cell and taking a video of the production cell at the same time. Utilizing the timestamps, the event duration was able to be configured and matched to each cluster.

The event data from each cluster is color-coded and plotted on the raw timeseries data as shown in Figure 11. Cluster 3 represents the most frequent event in the data. Within the sample data shown, there are 142 events within Cluster 3 as compared to 26 events within Cluster 1 and 13 events within Cluster 2. The average duration for an event in Cluster 3 is less than 20 seconds, whereas the average duration of events in Clusters 1 and 2 are just over 60 seconds.

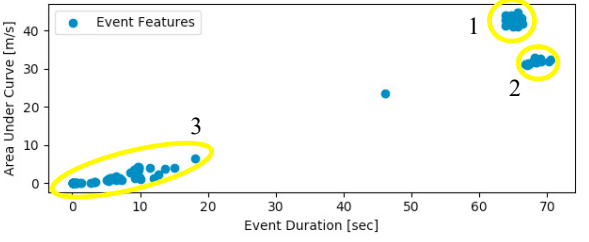
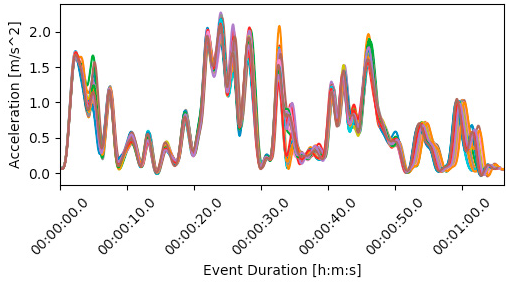
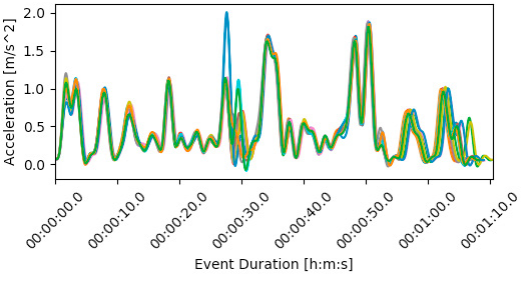


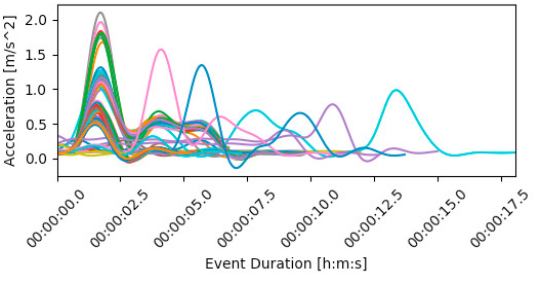

*Figure 9. Clusters within the feature space.*




*(a)*




*(b)*




*(c)*

*Figure 10. Data from multiple events: (a) cluster 1, (b) cluster 2, and (c)cluster 3*

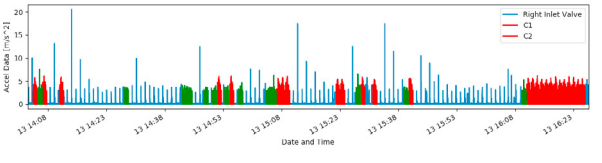


*Figure 11. Cluster overlay on raw data*

After observing the operation of this system, it was determined that Clusters 1 and 2 represent events when the entire mechanical system is moving and the vibrations due to the valve become masked by the vibrations of the system. While the doser is in motion, it will undoubtedly move in the same motion each time for the same operation accounting for the similarity of each point. The doser is also a multi-axis system, moving according to program. It is good to note that a current hypothesis of the team is that Cluster 1 and Cluster 2, are believed to be two different car bodies. This supports that the clustering can also help pick out different production processes and differentiate data points based around that.

However, events from Cluster 3 occur when the mechanical system is still, and the valve is the primary contributor to the sensed vibration, validating the increased variability. Therefore, Cluster 3 is expected to contain data that will reflect degradation of the valve over time. Qualitatively it is possible to observe variation in the dispensing time already from the differences in the shapes of the signals.

## 5. Conclusion

A review of condition-based monitoring has been presented with work done in vibration analysis and unsupervised learning in an application to a PVC dosing pump used in automotive painting. Key takeaways are an understanding of the selection of the materials used, the use of unsupervised learning to organize the data, in addition to the analysis taking place. Frequency spectrum tests were first utilized to characterize the signal and see if variation can be seen, before moving to unsupervised learning. With the unsupervised learning, it is a more robust method for model learning as well as variance analysis. From this point it is possible to determine the variance of signal and different production processes being performed on the machine. The sensors were readily accessible to the group and provided the versatility needed to capture the vibration signal. They met the compatibility of the sampling rate as well as the off-shelf accessibility. The sensor also paired well with the real time clock used to record timestamps and allow for matching the data to the production shift. The use of unsupervised learning sorts the data into a pattern from which a baseline and deviations are discerned. From the data collected, key events can be drawn to characterize the signal from when the valve on the doser is opening and closing, when the machine is inactive versus when the machine is active. Using this data as “healthy”, the model can start to be built to establish machine health and characterize more variability with the signal.

## 6. Further Work

Future work will involve implementing more on formalization of the predictive maintenance strategy; training the model to handle the predictive analysis rather than simply characterizing the data. This will also include modeling and testing against failed components and predicting instances during the doser dispensing cycle corresponding to specific times in the cycle time of the doser. More specific features for Cluster 3 will be generated and analyzed for long-term trends. With feedback regarding undesired valve behavior, a model will be trained to correlate feature trends with the undesired valve behavior for the prediction of that behavior. In addition, this will involve gathering more data, particularly fault data, and building up the infrastructure side. The infrastructure side involves implementing an off-the-shelf hardware solution as well as establishing automatic data transference and analysis on the industrial network. Production data will also be integrated into the system to allow for better matching of the clusters and data points to specific cars. On the IT side, computational analysis will occur using the techniques of fog and cloud computing discussed earlier. This will involve implementing a concrete database system to handle the device configuration as well as the analysis. Pipelines will be created to run the data from the database to the model. From the model, machine status will be displayed to the operator and engineers, providing real-time data.

A separate use case of vehicle track elevator Automated Storage and Retrieval System (ASRS) has been identified as a strong candidate for inclusion in a predictive maintenance program, as this system periodically experiences catastrophic bearing failure. While the ASRS and doser are functionally different, the practice of unsupervised learning and analysis is applicable to both. Upon the completion of the sensor box design, analysis strategy and system modeling, it is envisioned to deploy such a system to new equipment cases with new types of phenomena and failure modes.

### Acknowledgements

This work was done under the supervision and sponsorship of BMW research; the authors thank BMW for the support and time for this project as well as access to the factory and equipment under study.

## References

[1] J. Luse, “DATA DRIVES DESIGN – CONVERSATIONS IN IOT ARCHITECTURAL DESIGN,” 2018. [Online]. Available: https://blogs.intel.com/iot/2018/08/22/data-drives-design-conversations-in-iot-architectural-design/. [Accessed: 10-Feb-2019].

[2] E. P. Carden and P. Fanning, “Vibration based condition monitoring: A review,” *Struct. Heal. Monit.*, vol. 3, no. 4, pp. 355–377, 2004.

[3] J. Lee, B. Bagheri, and H. A. Kao, “A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems,” *Manuf. Lett.*, vol. 3, pp. 18–23, 2015.

[4] R. Lynn, E. Wescoat, D. Han, and T. Kurfess, “Embedded fog computing for high-frequency MTConnect data analytics,” *Manuf. Lett.*, vol. 15, pp. 135–138, 2018.

[5] R. Lynn, W. Louhichi, M. Parto, E. Wescoat, and T. Kurfess, “Rapidly Deployable MTConnect-Based Machine Tool Monitoring Systems.,” *Proc. 12th*

*ASME Manuf. Sci. Eng. Conf.*, 2017.

[6] D. Tom, R. Pintelon, J. Schoukens, and E. Van Gheem, "Variance Analysis of Frequency Response Function Measurements Using Periodic Excitations," *IEEE Trans. Intstrumentation Meas.*, vol. 54, no. 4, pp. 1452–1456, 2005.

[7] R. V. Canfield, "Cost Optimization of Periodic Preventive Maintenance," *IEEE Trans. Reliab.*, vol. 35, no. 1, pp. 78–81, 1986.

[8] X. Zhou, L. Xi, and J. Lee, "Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation," *Reliab. Eng. Syst. Saf.*, vol. 92, no. 4, pp. 530–534, 2007.

[9] A. D. Jayal, F. Badurdeen, O. W. D. Jr, and I. S. Jawahir, "Sustainable manufacturing : Modeling and optimization challenges at the product , process and system levels," *CIRP J. Manuf. Sci. Technol.*, vol. 2, pp. 144–152, 2010.

[10] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.

[11] P. Mell and T. Grance, "The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology," *NIST Spec. Publ. 800-145*, pp. 1–7, 2011.

[12] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: State-of-the-art and research challenges," *J. Internet Serv. Appl.*, vol. 1, no. 1, pp. 7–18, 2010.

[13] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog Computing and Its Role in the Internet of Things Characterization of Fog Computing," *Proc. first Ed. MCC Work. Mob. cloud Comput.*, pp. 13–15, 2012.

[14] L. M. Vaquero and L. Rodero-Merino, "Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 5, pp. 27–32, 2014.

[15] M. Lebold, K. Mcclintic, R. Campbell, C. Byington, and K. Maynard, "Review of vibration analysis methods for gearbox diagnostics and prognostics," *54th Meet. Soc. Mach. Fail. Prev. Technol.*, no. January 1985, pp. 623–634, 2000.

[16] G. Yen, Lin, and Kuo-Chung, "Wavelet Packet Feature Extraction for Vibration Monitoring," *IEEE Trans. Ind. Electron.*, vol. 47, no. 3, pp. 650–667, 2000.

[17] B. Chen, Z. Zhang, C. Sun, B. Li, Y. Zi, and Z. He, "Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors," *Mech. Syst. Signal Process.*, vol. 33, pp. 275–298, 2012.

[18] H. Sohn and C. R. Farrar, "Damage diagnosis using time series analysis of vibration signals," *Smart Mater. Struct.*, vol. 10, pp. 446–451, 2001.

[19] T. M. Romberg and A. G. Cassar, "A Comparison of Traditional Fourier and Maximum Entropy Spectral Methods for Vibration Analysis," *J. Vib. Acoust. Stree Reliab. Des.*, vol. 106, no. January 1984, pp. 36–39, 1984.

[20] T. Kohonen, "Exploration of very large databases by self-organizing maps," *Proc. Int. Conf. Neural Networks*, vol. 1, pp. PL1-PL6, 1997.

[21] G. Seif, "The 5 Clustering Algorithms Data Scientists Need to Know," *Towards Data Science*, 2018. [Online]. Available: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68. [Accessed: 10-Feb-2019].

[22] R. Pike, S. Dorward, R. Griesemer, and S. Quinlan, "Interpreting the Data: Parallel Analysis with Sawzall," *Sci. Program.*, vol. 13, no. 4, pp. 277–298, 2015.