47TH TURBOMACHINERY & 34TH PUMP SYMPOSIA
HOUSTON, TEXAS | SEPTEMBER 17-20, 2018
GEORGE R. BROWN CONVENTION CENTER

**Worldwide Deployment of Predictive Asset Management at Air Liquide**

**Cyril Defaye**
Rotating Equipment Expert
Air Liquide
Champigny, France

**Frederic Verpillat**
Predictive Maintenance Analyst
Air Liquide
Saint-Priest, France

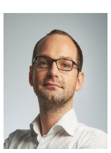
**James Huber**
Rotating Equipment Engineer
Air Liquide
Houston, TX, USA

**Ann Attaway**
Program Manager
Air Liquide
Houston, TX, USA

**Paul Gerke**
Machinery Support Engineer
Air Liquide
Houston, TX, USA

*Cyril Defaye is an Air Liquide Group International Expert in Rotating Equipment. He holds a PhD in mechanical engineering. Cyril started his career in 2002 as a R&D engineer at the French manufacturer of turbojet engines Snecma on R&D activities. In 2008, Cyril joined the engineering and construction division of Air Liquide as a rotating equipment specialist where he developed machine solutions and projects on the different cryogenic products (ASU, HyCO, CO2). Since 2014, he has worked in the Worldwide Industrial Management department of Air Liquide Large Industries where he develops strategic programs aimed at improving efficiency and reliability of rotating equipment. He's a member of the EIGA and CGA associations and he is leading the Air Liquide worldwide rotating equipment expert network. He has authored five papers in peer reviewed journals.*

*Ann Attaway is a mechanical engineering graduate of Rice University who joined Air Liquide in 1995. She has held a variety of positions within Air Liquide, including Facilities Manager, Plant Projects & Construction Director, Worldwide Technical Training Manager, and is now serving as the Program Manager for predictive analytics. She has spent most of her career figuring out how to make things happen in an organization, and is a certified Change Manager. Past experience includes Facilities Strategic Planning at Compaq Computer corporation during the years it was the fastest growing company in American business history, and as a plant engineering project manager for Gulf Oil Chemicals Corporation in a Polyethylene unit.*



*Frederic Verpillat is a predictive analyst for Europe area and data analyst expert. He graduated from Ecole Normale Supérieure and hold a PhD in Physics. He joined Air Liquide in 2013 in industrial IT, working on IoT and Big Data projects, then moved to Air Liquide Large Industries operations in 2017 to lead predictive maintenance deployment in South-West Europe plants.*

*Paul Gerke is a machinery support engineer assigned to the Air Liquide Center for Technical Expertise located in Houston, Texas. He holds a degree in Aeronautics from Saint Louis University and has been with Air Liquide since the beginning of 2002. He has held several roles supporting machinery repairs and reliability projects and has recently returned from Asia where he led the predictive analytic program for the Asia Pacific region of Air Liquide.*

*James Huber is rotating equipment engineer assigned to the Air Liquide Center for Technical Expertise located in Houston, Texas. He holds a degree in Mechanical Engineering from the University of Akron and started his career with Texaco. He has held various maintenance and engineering roles with various companies supporting machinery reliability and maintenance throughout his career. He has been with Air Liquide since the beginning of 2012. He has been active in supporting machinery design, repairs, and reliability.*

**ABSTRACT**

In this paper, the authors will describe how Air Liquide launched an international program to monitor and assess equipment asset health, resulting in a positive step-change in availability and reliability worldwide. Using predictive analytics, potential asset failures may be identified and appropriate intervention planned. Intervention prior to failure averts a possible reliability incident, adverse customer impact, and costly "emergency" maintenance activities.

## INTRODUCTION

Air Liquide is a worldwide organization producing industrial gases such as oxygen, nitrogen, argon, and hydrogen for industrial customers. Our business model results in many small facilities often located within our industrial customers sites. Subject matter experts may be located several hundred miles from the equipment in his or her scope of support. Having this "many mini" organization presents challenges with coordinating any worldwide initiative. The team is able to overcome these challenges by coordinating the people, processes, and tools required for success. Over 1500 models were deployed in over 30 countries in the first six months of the project. The purpose of this paper is to share successful practices that intelligently apply new digital tools to achieve acceptance in the organization.

Evaluating the root cause of unreliability in our network has underscored the potential for applying new analytical tools to asset health. Mechanical failure was clearly the primary contributor to the cost of unreliability, followed by instrumentation and electrical failure. Understanding the true value of a predictive tool helped the organization accept and embrace the changes necessary to take advantage of predictive data.

## PREDICTIVE ALGORITHMS

Predictive analysis (Nishchol 2012) is an advanced branch of data engineering which generally predicts some occurrence or probability. The process involves an analysis of historic data and based on that analysis to predict the future occurrences or events using Predictive Analytics modeling techniques. The form of these predictive models varies depending on the data they are using. In the following paragraphs, we detail the two main Predictive modeling approaches to detect abnormal events: Anomaly detection and Fault recognition.

### Anomaly detection

In anomaly detection (Chandola 2007), a model is built using 'good/normal' operating data that typically represents a wide range of operation. Each new point is evaluated against the model, and if the residuals are outside statistical limits, the point is considered as an outlier and the process could be seen as drifting outside its normal operating region. Anomaly detection is an unsupervised learning problem: it is the task of finding hidden patterns in unlabeled data. It determines that something unusual is occurring when conditions deviate from "normal" conditions of operation. Anomaly detection tasks are relevant when there are a large number of negative samples (normal operations) and a few positive samples (failure data). They work best when the failure is due to several factors, all of which cannot be modeled beforehand. For predictive maintenance of machines, anomaly detection tasks are the most relevant. Examples of mathematical concepts for unsupervised learning include PCA (Principal component analysis), SOM (Self organizing maps), Neural Networks, k-means clustering etc.

- Pros: This method can generally detect the potential (and unknown) failures
- Cons: The definition of the "normal" condition can be challenging (in particular for new processes). Moreover, when an anomaly is detected, you do not know what it is: equipment failure, new "normal" condition or a sensor fault?

### Fault recognition

Fault recognition captures the faint but precise sensor patterns from the very beginning of machine degradation and captures the stronger patterns as the condition develops and the machine operation deteriorates towards failure. Once the pattern of a fault signature pattern is captured, it can be used as a monitoring profile, and if the pattern ever emerges again, you will now know exactly what is happening, and an immediate warning can prompt action well before damage. Fault recognition is a supervised learning problem the task of inferring a function from labeled training data. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). If outputs/supervisors/KPIs cannot be naturally identified in the data, then an artificial one is created. Examples of artificial outputs are (i) time to failure or (ii) likelihood of failure in a given time period. Supervised learning tasks are meaningful when there are a large number of both positive and negative samples in the data (equal likelihood of good or bad data). They also work best when future positive samples (e.g. time to failure) are likely to have similar characteristics of the positive samples in the learning data. Such assumptions are unlikely for the monitoring tasks we are looking to accomplish, where all the faults cannot be expected to have the same root cause. Supervised learning methodologies include linear regression, decision trees, partial least squares, Bayesian networks, neural networks etc.

- Pros: Provide simultaneously Predictive and Prescriptive information
- Cons: A given failure can be detected only if it occurred in the past with a similar signature. Thus, this method performance may be limited when failures are rare or varied.

Fault detection algorithms may provides an automatic diagnostic and may be considered as more prescriptive compared to anomaly detection. However, the implementation of a fault diagnostic approach requires a large number of failure records. The most important rotating equipment, such as main air compressors, used in air separation units, are often tailor-made machines specific to each plant. For each type of machine, an exhaustive record of all pattern of failure is not available.

Several softwares based on anomaly detection algorithms are commercially available. The Air Liquide R&D department analysed several of them and rated them on mathematical, engineering and business criteria such as:

- Data Selection: What sources of realtime and historical data does the tool require? What engineering knowledge of the asset is imparted into processing the data? How much historical data does the tool require?
- Predictive Modeling: How accurate and robust are the mathematical techniques used to build a model of the desired operation?
- Statistical Analysis: How robust are the statistical techniques used to detect a deviation from desired operation?
- Preventive Action: Once an event is detected, what are the mechanisms to provide notification and alerts, from early warnings to recommendations for action?
- Maturity: Is the tool well referenced within the process industry? How is the vendor innovating? Is support available worldwide?

This benchmark showed that the mathematical core of the software is not a significant differentiating criteria for an industrial application. No significant differences were observed between the softwares on their sensitivity to detect a deviation. Differentiating factors include:

- Connectivity to the existing data historian and IT architecture
- Ease of data cleaning
- the possibility to deploy models from a template reference library
- the robustness of the associated expert system to provide pre-diagnostic, and
- the catch management system

These factors may be considered more or less important, depending on the needs and the organization of the user company.

Existing plant control systems supervise equipment performance based on instrumentation for single metrics such as pressure, temperature, vibration, etc. Low and high alarm setpoints are based on risk assessment and equipment design. However, daily operation is generally more restrictive than the entire allowable range. The predictive analytics tool is able to define a historical relationship between all the metrics describing asset health for each piece of equipment. When those relationships change, this may be an early precursor to failure, despite the fact that the alarm limits have not been breached. The Figure 1 explains in a simplified way the operating principle of the tool.

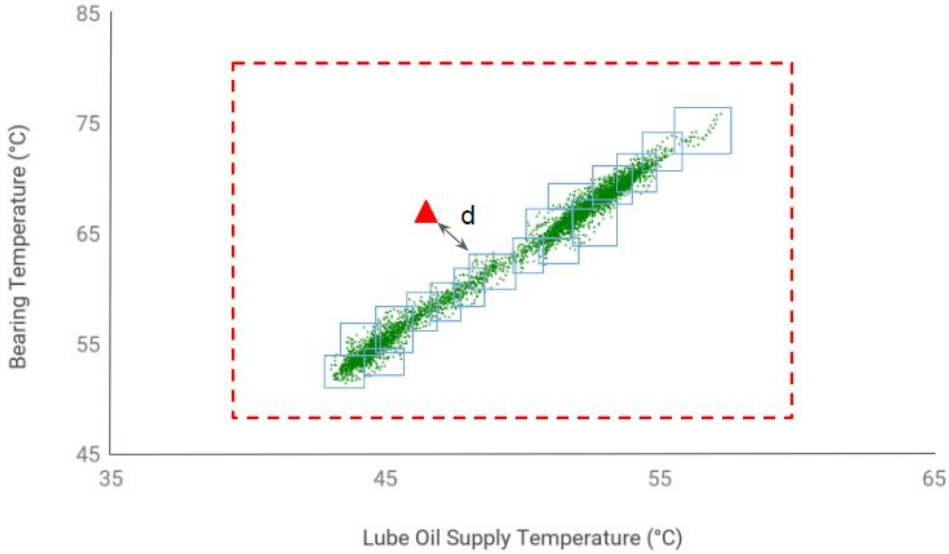


Figure 1. 2-Dimensional Correlation between Operating Parameters and Clustering of Historical Data

The correlation between two operating parameters of a compressor (oil temperature and bearing temperature) over one year of operation is shown with the green dots. In this example, the bearing temperature is increasing linearly with the oil temperature. For example, if the oil temperature is 47°C, the bearing temperature should be about 58°C based on historical data. The dashed red square in Figure 1 represents the typical low and high alarm setpoint implemented in the plant control system. Any operating point contained in this area is considered by the plant control system as a safe operating condition. An actual operating point, for example the red triangle in Figure 1 (oil temperature of 47°C and bearing temperature of 67°C), may be considered as a safe operating condition from the plant control system, despite the fact that it is clearly outside the pattern described by the historical data. The software is creating a model through a clusterization of the historical operating points. The clusters are schematically represented in blue in Figure 1. This model is stored in a database and is considered as the reference behavior of the machine. In real time, the tool will compare the actual operating point to the model prediction. When the actual operating condition falls inside a cluster, the machine behavior is considered as normal and no alarm is generated. When the actual operating condition falls out of a cluster, an alarm which is proportional to the distance *d* between the actual operating point and the nearest cluster is generated. This simplified example considers only two correlated parameters but the tool is actually able to create a model from several correlated parameters. Typically, the mechanical behavior of a air compressor rotor can be depicted by a model that includes the following parameters: vibrations, bearing temperatures, oil temperature, compressor load and discharge pressure. The dimensionless distance *d* calculated in n dimensions for a model with n parameters is the overhaul deviation of the actual operating conditions compared to the model built from historical operating conditions.

The selected predictive analytics software with pattern recognition compares current performance to historical data relationships. This is done on a set of metrics determined by subject matter experts for each class of machinery. It is important to create models that include parameters that are sufficiently correlated, otherwise there is a risk to decrease the sensitivity of the tool. For a air centrifugal compressor, it is recommended to create one model dedicated to the monitoring of the mechanical behavior and a second model dedicated to monitoring performance.. Matrix mathematics determines an overall deviation from known data clusters derived from the historical relationships between variables. When the deviation exceeds a set percentage on any metric, an alarm is registered and evaluated by the analyst.

## THE IMPORTANCE OF THE ORGANIZATION

Key roles in applying predictive analytics successfully include the Analyst, the Site Champion, and the Subject Matter Expert (SME). In our organization, these roles are in different departments but must closely coordinate as a team for maximum success:

- The Analyst deploys and monitors the software and reports through a centralized analytical organization to operations;
- The Site Champion assists the Analyst in deploying the model by ensuring the correct historian tags are selected for each metric and supplying local information on equipment use. The Site Champion is an operations employee generally reporting to the local plant manager.
- The SME resolves equipment issues by diagnosing issues and planning intervention before a reliability issue can occur. The Subject Matter Expert may report through a regional maintenance organization or a centralized team of experts.

Figure 2 depicts the different steps required to first create and deploy and then maintain the models. Before creating the models, it is important to prioritize the equipments of a fleet that has to be monitored. Several factors can be considered to define this strategy as

- criticity of the equipment for the production,
- availability of spare parts,
- ageing of the equipment,
- installed instruments,
- maintenance strategy…

Once equipment has been selected, the list of required models has to be established. For a compressor, depending on the instrumentation available and connected to the data historian, the models to be created can be:

- A mechanical model to monitor vibration, axial displacements and bearing temperatures
- A performance model to monitor either the global efficiency of the machine or the individual compression stage efficiency
- A lube oil system model to monitor the efficiency of the oil cooling and the oil pressure regulation...

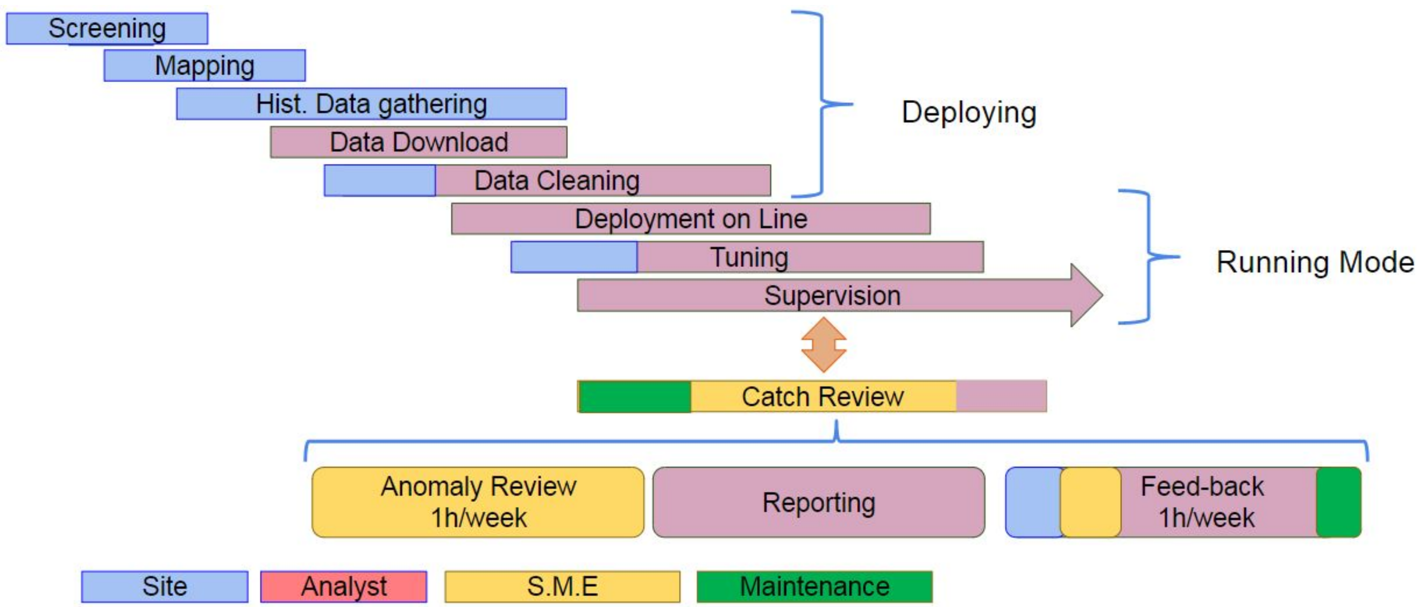



Figure 2. Deploying and Running a Predictive Analytics Tool

Each model requires a number of parameters called *metrics*. Determining which metrics to use for each class of machinery is a prerequisite to developing models. The input of the SME is required at this point to select the metrics for each model. To speed up the deployment and make it more robust, it is possible to create a library of generic models that can be easily used by the analyst to properly create the models without the support of the SME. These generic models are called a *template*. A template is a list of metrics such as temperature/pressure/vibration/load that describe the behavior of a specific part of the asset as explained above. The metrics used are defined for each type of asset and multiple templates may be used to monitor an asset. For instance, a air compressors requires four templates to monitor the lube oil system, mechanical systems, performance, and the motor. When launching the program, key subject matter experts were brought together to learn the software and develop these templates for major types of equipment. Note that the template comes also with alarm thresholds specific to each metrics, some pre diagnostic scenarios and filters to disable the model when the asset is in standby (for example). Each thresholds are associated with a persistence criteria. As the predictive analytics approach is focusing on the identification of early precursor signal, a deviation to the reference behavior can be confirmed during a given time before to generate an alarm. For models dedicated to the monitoring of the mechanical behavior of a compressor, a persistence of three hours is implemented. The deviation shall remain continuously above the given threshold during three hours before an alarm is effectively generated. The persistence is an effective criteria to reduce the number of false alarms generated by the tool. Performance models are usually less critical for the integrity of the machine and the persistence can be increased to a higher value to further reduce the number of false alarm.

The templates developed and tested by the SMEs were used by the Analysts and Site Champions to deploy the required models. Working together, they identify the tags available in the historian server for each metric, mapping tag names to the appropriate metric. Because this information is critical to program success, it was independently validated by both parties. The analyst then downloads historical data for all these tags to a server dedicated to the tool. The sampling frequency of the collected data depends on the process behavior, but one sample every fifteen minutes collected over one year appears to be a reasonable approach. Once downloaded, the data set has to be cleaned to remove non-standard operating conditions such as downtime, transient operating conditions during startup and shutdown of the equipment, invalid data recorded in the historian server, unexplained operating conditions, data before the last machine overhaul... The site champion has a key role in cleaning to explain specific operating conditions of the equipment to the analyst and to inform him or her about relevant maintenance history, such as the last overhaul on the machine. The sensitivity of the model is directly driven by the quality of the cleaning. Too much cleaning will lead to a high number of false alarms, while not enough cleaning will mask some deviations and will delay alarming. The cleaned dataset is checked for mathematical correlation between each metric in the template. Data is placed in clusters to create the model.

Once the profile is created, real-time monitoring begins. It means that on a regular basis, typically every 5 minutes, the tool will compare the actual value of each parameter to the value predicted by the model. An alarm will be generated when the overall deviation between actual and predicted values exceeds the threshold defined in the template. During the next two months, the profile is tested against daily operation. A high number of false alarms occur during this period as the model is field-tested and adjusted. The model may required additional data that are not yet present in the model such as extreme operating conditions, or possibly valid conditions that were deleted during cleaning of the data set. This period is a critical time for the Analyst to learn the equipment and the models and for the model to become robust enough to be credible with field personnel. A filter is added to cancel all alarms in the event the equipment is not in use and delay alarming immediately after startup until normal operation can be expected. The senior experts interact with the model to adjust the alarm threshold and the persistence required from a high deviation alert before an alarm is generated. Avoid giving too much information - some of which may be unfounded - to the site at this time. After the tuning period of the model, the number of false alarms is significantly reduced to an acceptable level that can be properly treated by the analyst, typically one false alarm per equipment per month.

After the model is deployed, if the present operation departs from historical norms, a deviation measure is calculated. If the deviation exceeds the set limitation and meets the persistence requirement, an alarm is generated. The analyst determines, based on historical trends and the information provided by the analytics, whether each alarm is most likely an instrument failure, a modeling issue with our software, or a true equipment anomaly. The tool allows to identify the main parameters contributing to an observed deviation and helps the Analyst to define a pre diagnostic explaining the deviation. The equipment anomalies are referred to the SME for further study and possible intervention. The instrumentation issues are referred to the site, and modelling issues are handled by the analyst. A flow chart demonstrating correct alarm management has been developed. One analyst is able to monitor approximately one hundred assets using this tool. During the supervision phase, a close collaboration between the analyst, the site champion, the SME and the maintenance team is mandatory to ensure the success of this approach. It is important to organize a weekly or bi-weekly call between these stakeholders to review all the alarms, classify them (false alarm, instrument failure, mechanical catch, performance issue) and define appropriate actions. It may take weeks or month before to act and fix an identified issue. Thus it is important to track all cases that remain open over the weeks and ensure they are finally properly closed. Only the most relevant alarms should be communicated to the operations personnel in order to not jeopardize the credibility of the methodology. On the other hand, a lack of collaboration and exchange between the stakeholders will result in a high number of ignored alarms with the potential to miss real catches. Figure 3 shows the difference between the number of catch per month of an affiliate with a robust and mature organization (zone 1) and a second affiliate (zone 2) of similar size where a dedicated analyst has not been officially assigned from the beginning of the deployment. The analyst in the zone 2 was assigned in May 2017. One month after his assignment the number of catches significantly increased, close to the level of the mature organization in zone 1. This example shows how organization is important for the deployment of predictive analytics solutions.

Senior management support is, of course, crucial to project success. While the senior managers were impressed with forecasts of possible savings and return on investment, they became more convinced of predictive analytics capabilities once we tested our list of metrics by backtesting actual past mechanical failures and calculated the significant potential savings.

Historical data were to create the equipment "normal operation" definition some time before the failure occurred. Data for 6 months or more prior to the failure was not considered. The model could then "fast forward" to the event as if the software had been monitoring as the event unfolded. We noted both the date of warnings and alarms and the mathematical "root causes" of the high deviations calculated when the equipment started to fail. Actually, in some cases of analyzed past failures, the tool provided early warning up to 9 months before the failure occurred on site. The subject matter experts who did these analysis were immediately convinced that predictive science works, and the senior managers supported immediate rollout and implementation. In fact, we reached our three year commitment within the first nine months of implementation.

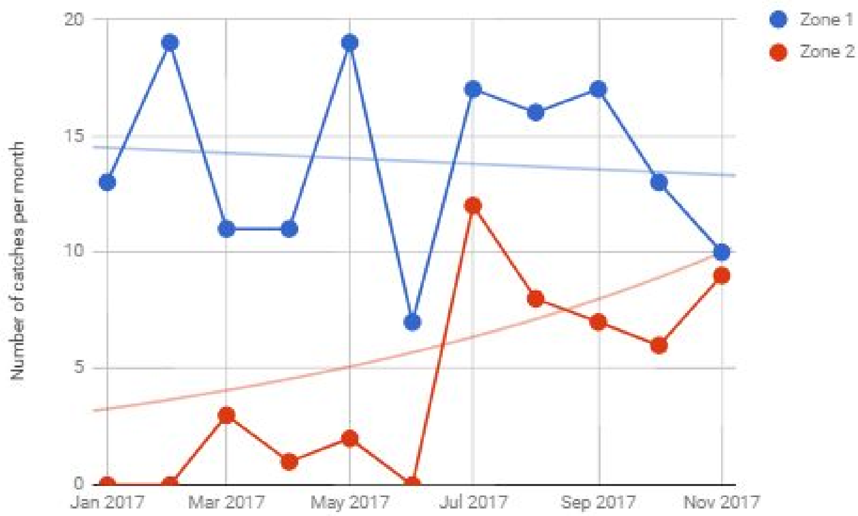



Figure 3. Influence of the Organization on the Number of Catch per Month

Software alone is not enough for successful implementation of digital tools. It has been demonstrated that the number of significant catches rises when - and only when - the right organisation is established and supported by the management team.

**RESULTS**

Warnings and alarms from three different types of "catches" are presented below. This includes a mechanical equipment catch, a malfunctioning instrumentation catch, and an equipment performance catch. Each catch demonstrates what was seen in the predictive analytics tool with accompanying photographs from the field corresponding to each event.

*Mechanical catch:*

A 7.4MW motor-driven main air compressor (Integrally geared type) of an air separation unit showed a sudden increase in the mechanical model overall deviation on February 26th, 2017. During the following days, the plant had an emergency shutdown and after startup, the overall deviation continued increasing. The deviation was mainly resulting of the following components: a 8°C decrease of the first stage bearing temperature, a 6µm p-p increase of the first stage vibration (Figure 4) and a 6°C increase of the fourth stage bearing temperature. An oil analysis was recommended by the SME and showed high varnish content in the oil. An external filtration system was implemented as temporary mitigation to reduce the short term risk of reaching the alarm level in the plant control system. After a spectral vibration analysis, it was decided by the technical committee to stop the machine for bearing inspection. Due to production obligations, the shutdown was postponed to November 2017. The inspection showed that one of the pads in the upper half of the 1st stage bearing was damaged with the white metal torn off. The damaged bearing was replaced. The bearings of other stages were found to be in good condition. They were thoroughly inspected, cleaned and reinstalled. Note that there was no alarm from the plant control system at the time of the maintenance.

In this case, the operations team received the first alarm eight months before the eventual shutdown and inspection of the compressor. The analysis of precursor signals with a predictive analytics software avoided extensive damage on the compressor shaft and significant production loss due to the unavailability of the compressor.

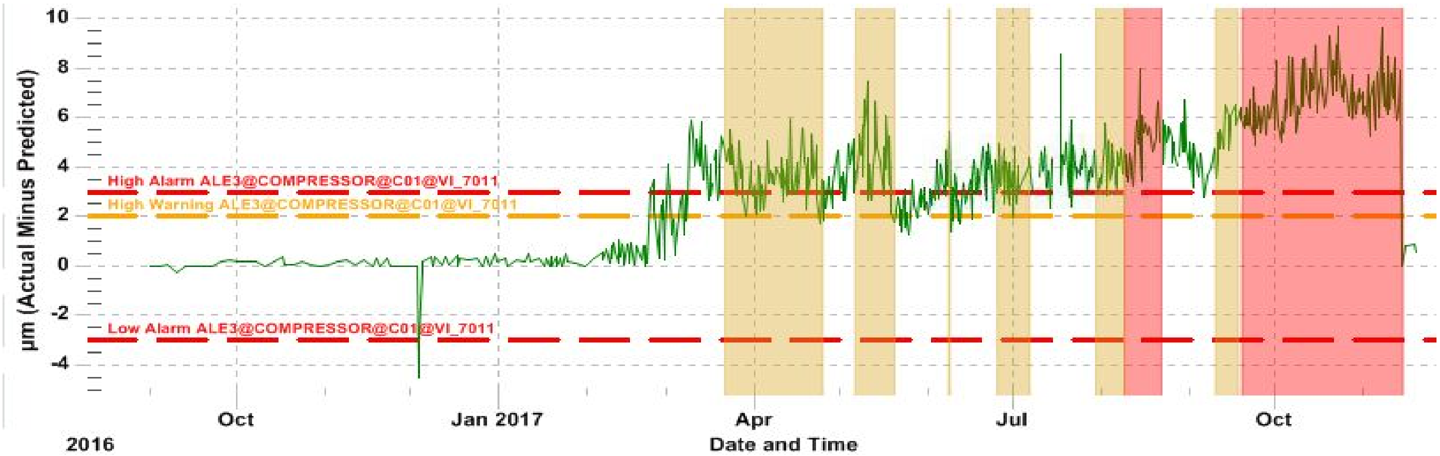



Figure 4. Evolution of the First Stage Vibration Deviation

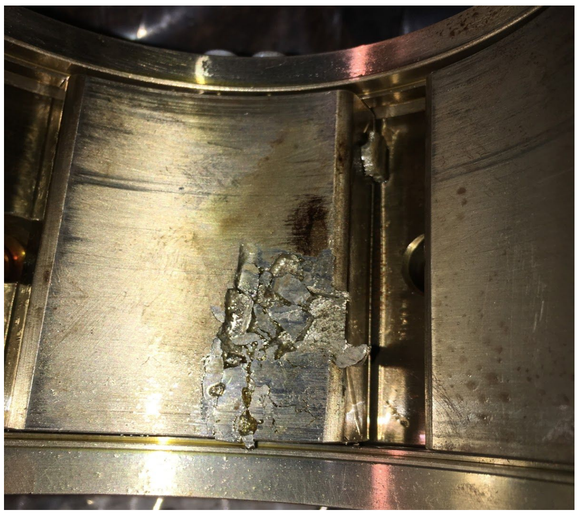


Figure 5. Observed Bearing Damages during Maintenance in November 2017

*Performance catch*

On June 8th, 2017 an alarm registered on the performance model of a 5.9MW booster air compressor model. The Analyst observed a lower actual flow compared to the predicted one as shown in the Figure 6. In this case, the flowmeter is installed downstream of the bypass line. It measures the flow effectively produced by the compressor to the plant. The observed average 5000Nm$^3$/h reduction in flow is corresponding to a leak or a recirculation internal to the the compressor system. Further investigation revealed a leak on the air instrument supply to the anti-surge valve, resulting in a partially opening the recycle valve. Following the repair of the leak on June 11th, the actual measured flow came back to a normal value, i.e. between boundaries of the predicted flow. A leak of 5000Nm$^3$/h corresponds to unnecessary power consumption of 400 kWh.

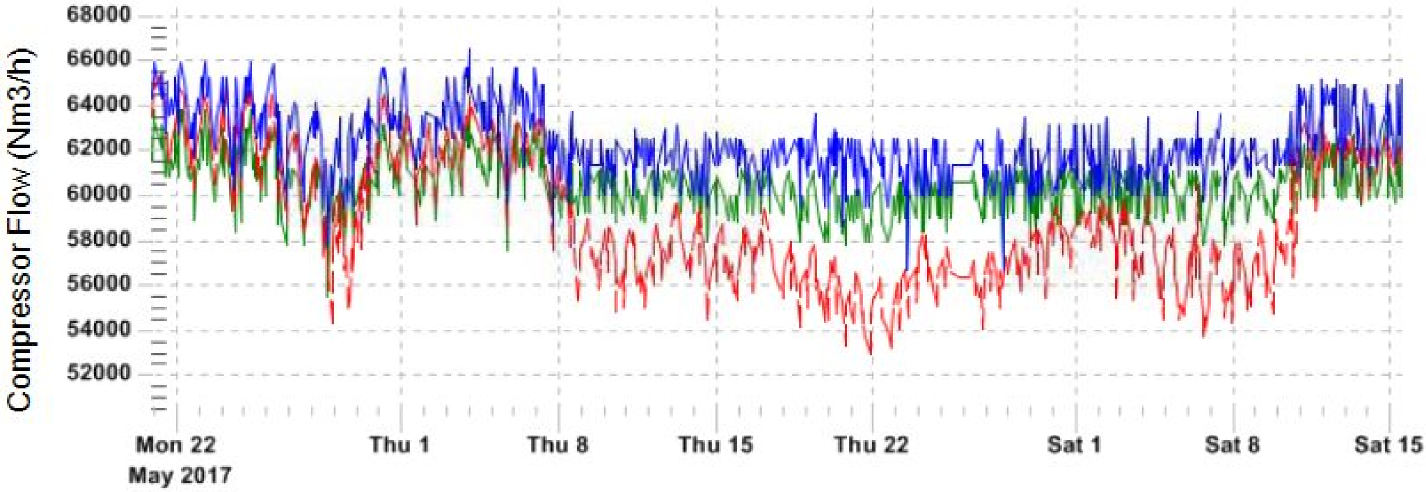



Figure 6. Actual Measured Compressor Flow (red) Compared to Lower (green) and Upper (blue) Boundaries of the Predicted Flow.

The use of predictive analytics can also enable significant savings in the maintenance budget and efficiency by confirming or completing the technical and efficiency assessment through more precise data analysis. Thank to a detailed analysis using the correlation between the parameters, it is possible to better understand the condition of an asset and determine maintenance requirements predictively. For example, this analysis has been done on a 16MW single shaft air compressor before a scheduled turn around. The analysis confirmed, along with the usual spectrum analysis, oil analysis, etc., that there was no significant wear on the mechanical parts of the machine. The analysis of the performance model showed a slight decrease of the efficiency due to fouling of both the rotor and heat exchangers. Opening the compressor to clean the rotor while the mechanical behavior is acceptable is a major maintenance task with high associated costs compared to the expected power savings. It was decided to focus on the cleaning of heat exchangers, as this task requires fewer resources with more results. Figure 7 shows the impact of heat exchanger cleaning on the compressor performance. For a given opening of inlet guide vanes, the flow capacity increased allowing about 100kW in savings on the power consumption for a minor investment. On the top of the power savings, the major overhaul of the compressor has been postponed allowing a significant saving on the maintenance budget of the current year.

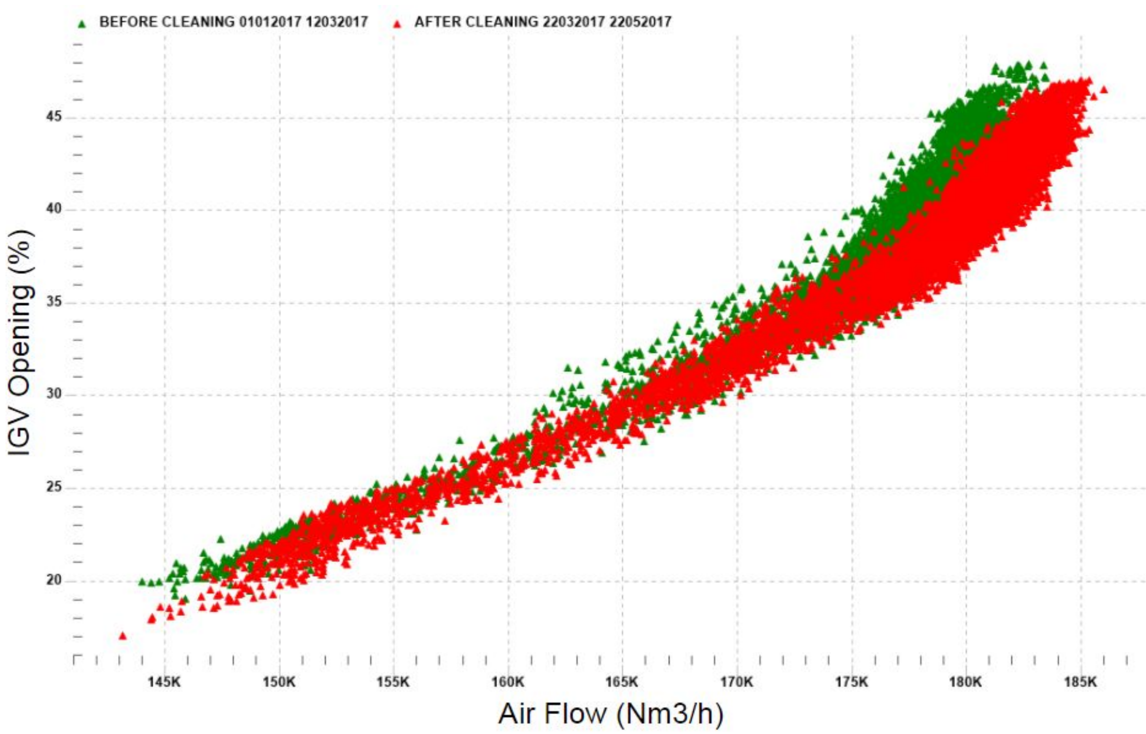



*Figure 7. Operating Data Before and After Cleaning of Heat Exchangers of the Compressor*

*Instrumentation catch:*

The present catch is related to an instrument failure detection that happened on an integral gear type air compressor. On December 8th, 2016 a discrepancy was observed between the predicted and actual values of the stage fourth bearing temperature, the actual measurement showing a temperature increase of 20°F compared to the prediction (Figure 8). Over the next couple of days this temperature rose to as high as 222 F° before stabilizing. Investigation discovered that the RTD was displaying a higher value in the control room than the RTD resistance value measured in the field. It was concluded that the card of the plant control system was at fault. The RTD card has been replaced. Further deviation of the measurement could have led at then end to a trip of the compressors and of the plant. The unavailability of the plant for few hours has been avoided.

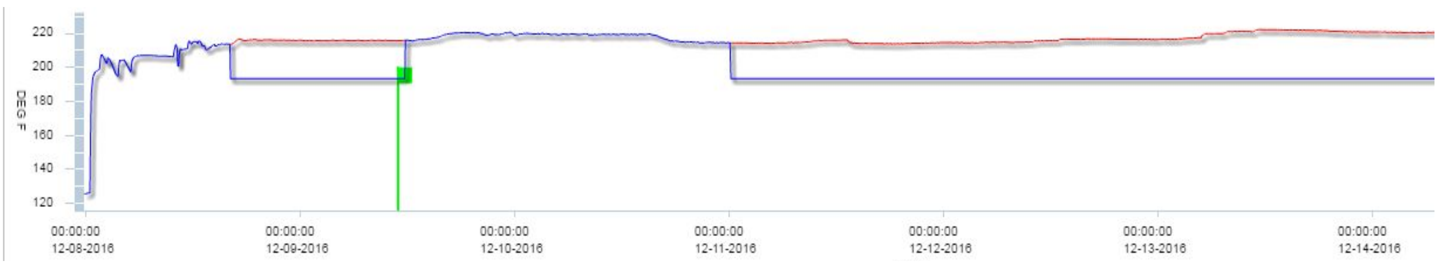



Figure 8. Discrepancy between Bearing Temperature Actual (red) and Predicted (blue) Values

**LESSONS LEARNED:**

We have been able to prove that the software works and adds significant value to our operation providing early warning weeks and even month before an alarm would have been generated in the plant control system. However, in this case, the ability to predict and prevent reliability incidents requires a cultural shift that is more complex than the matrix math that drives our predictions.

Change is always difficult for people and especially for large organizations with historical ways of working. Since “people support what they help create”, we were careful to involve key subject matter experts who garnered the respect of the operations organization. As they developed the templates for each type of asset and did the backtesting of real incidents, they realized they could have had months to predict and prevent some of the crisis situations we previously thought unavoidable. They immediately realized the impact on both their personal and professional lives and were eager to help us convince others to support the new methodology.

An important aspect of our success was to value each catch in terms of avoided cost of unreliability. This exercise is interesting from a return on investment perspective and has added motivation to speed adoption throughout our system. However, we do not want to allocate resources to calculate a precise value for each catch, as this adds little value to our customers. Nor do we want to create possible finger-pointing or useless friction arguing with operations on each nuance of each catch.

Instead, we calculated averages from large databases for each type of catch, taking into account equipment spares, reduction in time out of service, reduced expediting fees, and in some cases, less maintenance required overall as major incidents could often be avoided when wear parts were damaged and replaced, averting a possible catastrophic event. Of course, the cost of performing the required maintenance is subtracted from the calculation. Our catch values are conservative and it is generally considered that our total catch value is a very conservative estimate. At the end of the year, our catch value represented approximately 50% of the overall reduction in the reported cost of unreliability worldwide.

With more than 150 actionned catches in our system, we can observe (Figure 9) that 54% are instrumentation, 34% are mechanical and 12% are associated with equipment performance. However, the associated avoided cost are respectively 8%, 79%, and 33% of total savings.

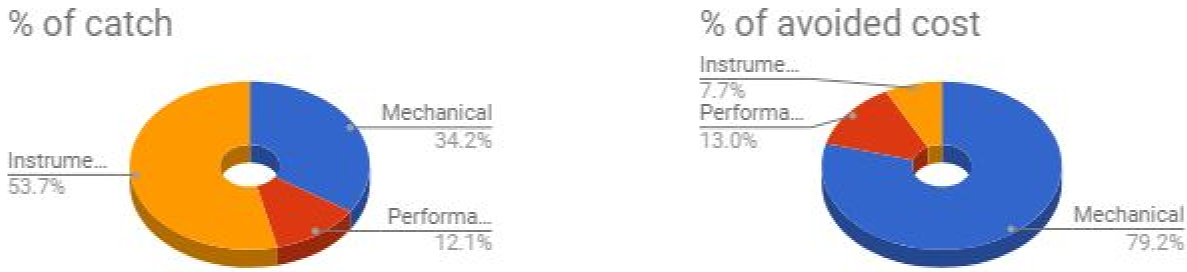



Figure 9. Distribution of Catch Categories in Term of Number of catch and Avoided Cost

Keeping executive support for this on-going program is critical. Keeping the analysts focused and motivated to find new applications and improve current practices is also a challenge. We've been able to do this several ways:

- Each geography keeps an alarm management spreadsheet. These sheets are hyperlinked to a web page showing the value of catches each year. Any change to catch value made on their tracking sheet is posted to the web within 5 minutes.
- Monthly "Flash" reports are issued for the "Catch of the Month" to highlight unusual applications or solid application of the predictive toolkit. The analysts, site champions, and involved subject matter experts are recognized by name for this global distribution.
- The annual report to executives showing the cost of unreliability showed a 50% step change that can only be explained by the use of predictive analytics

Creating an internal user community allows communication with the analysts, and encourages them to discuss amongst themselves. We had categories for software suggestions, discussion, new applications, and announcements. The analysts are able to coach each other with new ideas. This is also a great place to recognize achievements for the program and for individuals making significant contributions.

For instance, one of our senior analysts has added a new analytical tool for testing template viability. For a new application, he downloaded all the data available describing the asset status. As usual, he downloaded the historical data, removed non-standard operating conditions. Distribution of each sensors measurement is displayed on Figure 10.

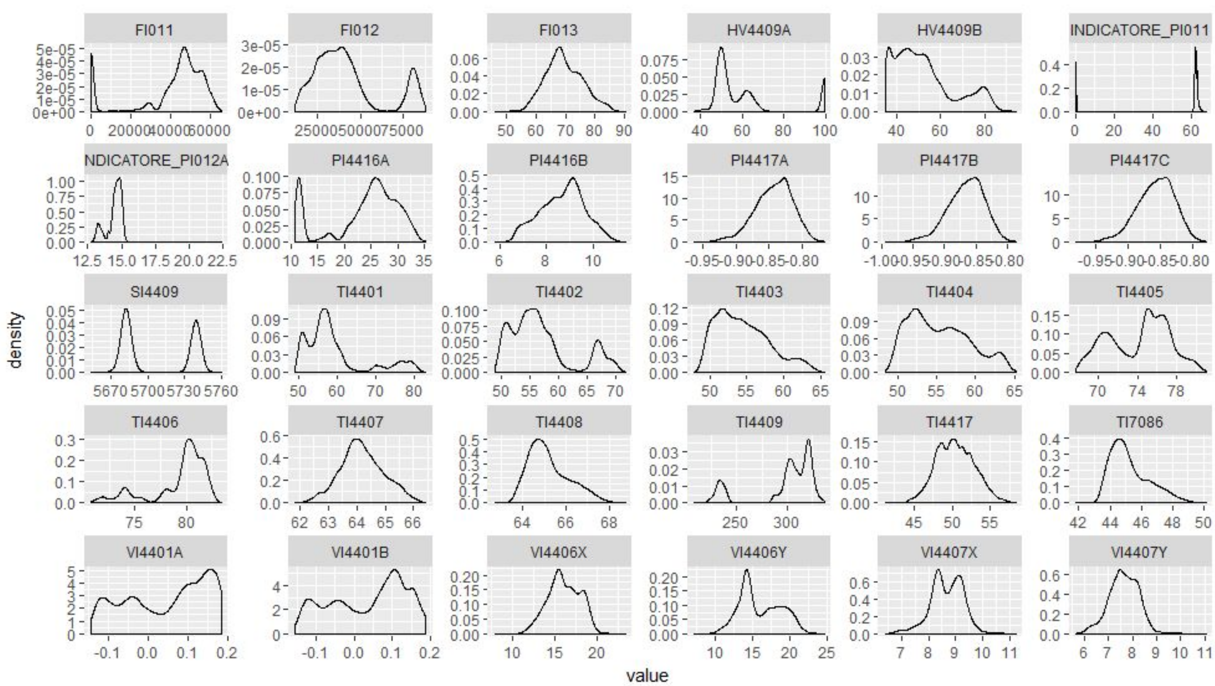



Figure 10. Distribution of each data points after cleaning

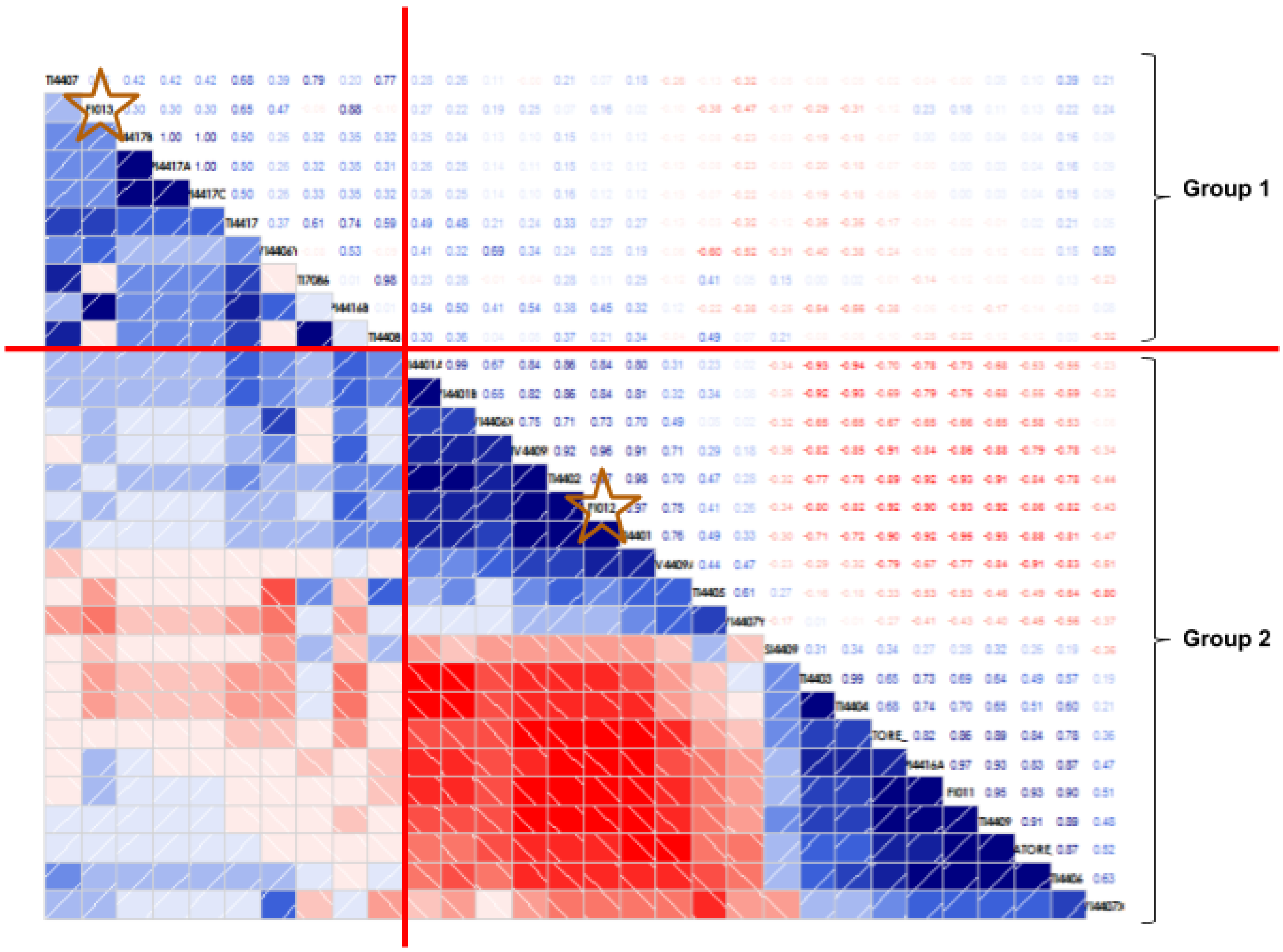



Figure 11. Correlation Matrix Showing Correlations Measurement between each Parameters (value between -1 and 1).

With the statistical computing language R, he calculated the correlation of each datapoint to the others and displayed the correlation matrix shown in  Figure 11. All the parameters used (pressure, flow, temperature, vibration, etc.) are listed on the diagonal. Above the diagonal, the correlation measurements between parameters are displayed  (value between -1 and 1). Below the diagonal, a visual representation of the correlation between the parameters is showed. In blue, parameters are correlated, in red they are anti-correlated. The intensity of the color represents the intensity of correlation. Datapoint are organised by cross-correlation. This visualisation is very powerful to identify groups of correlated data, which are two in this case :

- 1 : top left corner shows strong correlation coefficients, both positive and negative.
- 2 : bottom right corner : strong positive correlation coefficients

In this way, it was possible able to clearly draw the conclusion that for this equipment, there were two sets of correlated data. This analysis is also useful to identify the driving metric, marked with a star on the figure (ex: power measurement for a motor).  He was able to create two templates with the correlating metrics in each to improve the accuracy of the predictive tool.

**CONCLUSIONS**

Predictive Analytics is part of the digital revolution with proven applicability in reducing unplanned downtime in industrial applications. The Air Liquide experience shows that beyond the selecting software (which can be internally developed or commercially available), a wide industrial deployment of predictive analytics solutions requires a strong methodology, organization and communication to make it successful, providing the expected reduction of unreliability events. Areas of continuing research include water chemistry and cooling tower systems, static electrical equipment, electrical transformers, breakers, feeders, valve “stiction”, and backup system readiness.

The plant manager, our leadership and customers need even more information. They need to know there is an X% percent chance of failure in the Y coming days. The presented tool and methodology are not yet prognostic - the mathematics has to be developed within the software - but the authors believe a statistical approach utilizing a cross-linked database will be required to accomplish this goal.

## REFERENCES

Nishchol, M., and al, 2012, Predictive Analytics: A Survey, Trends, Applications, Opportunities & Challenges, International Journal of Computer Science and Information Technologies, Vol. 3 (3).

Chandola, V., Banerjee, A., Kumar, V., 2007, Anomaly detection : a survey.

## ACKNOWLEDGEMENTS

The authors acknowledge Jeff Arbogast and Athanasios Kontopoulos from the Air Liquide R&D center for their benchmark of commercially available predictive analytics softwares.