An Innovative Machine Learning driven approach to detect anomalous behavior of Dry Gas Seal Heaters for Centrifugal Compressors

Meenali Sharma[1], Carmine Allegorico[2], Laura Nuti[3], Unnat Mankad[4], Gabriele Mordacci[5], Rajakumar Dharmaraj[6], Hasbullah, Aidil Fazlina[7]

[1,4]Baker Hughes, Bengaluru, 560037, India
meenali.sharma@bakerhughes.com
unnat.mankad@bakerhughes.com

[2,3]Baker Hughes, Florence, 50127, Italy
carmine.allegorico@bakerhughes.com
laura.Nuti@bakerhughes.com

[5,6,7]Baker Hughes, Kuala Lumpur, 50400, Malaysia
gabriele.mordacci@BakerHughes.com
d.rajakumar@bakerhughes.com
aidilfazlina.hasbullah@bakerhughes.com

ABSTRACT

Compressors are used in various industries to elevate the pressure of a gas to meet the process requirements. Dry Gas Sealing (DGS) system is used in compressors to contain the process gas within the casing with near zero leakage. Due to efficient sealing capabilities, Dry Gas Seals became key technology in industry to prevent the leakage to environment and to ensure compliance with health and safety regulations.

DGS is very sensitive to process gas conditions. To ensure health and integrity of DGS and to maintain required process gas conditions, Dry Gas Seal systems include various Auxiliary Equipment viz. Filters, Heaters, Control valves etc. Heaters maintain the Seal gas inlet temperature at a specified value to avoid presence of condensate, which can otherwise result in Seal failure. To maintain this temperature, Heaters operate in an on-off toggling pattern. Failure of these Heaters can lead to unit unavailability; Early detection of anomalous Heater operation can ensure timely action to avoid any possible negative impacts on Dry Gas Seal health which in turn can impact Compressor operation.

The paper makes a theoretical survey of existing pattern recognition algorithms for time series and examines their applicability in detecting anomalous operation of Seal Gas Heaters. Finding these methods not directly applicable, the paper presents a state-of-the-art physics plus data-driven approach. The method is developed by combining LSTM type Neural Network with Periodogram and Auto-correlation monitor to detect any deviations from normally expected operating behavior of Seal Gas Heaters. The LSTM learns an exhibited pattern and looks for the similar patterns in Heater signals. Auto-correlation monitor coupled with Periodogram helps in determining the dominant frequencies and window-size required for the LSTM component. The method is tuned to accommodate different operating modes of Heater based on Compressor running conditions. If Heater deviates from working in an established toggling mode, user is alerted before Seal gas temperature is impacted.

Applied in real-time, the method alerts engineers for any anomalies observed in Heater behavior, thus enabling swift action to prevent any harm to the Dry Gas Seals, caused by temperature upsets. The paper demonstrates the performance of this method when applied on 26 compressors, thus validating the applicability and accuracy in prognostic health

management of Dry Gas Seal systems of Compressors.

1. Introduction:

Dry gas seals are extensively utilized to seal the shaft ends of centrifugal compressors, effectively containing process gases and mitigating the risk of fugitive emissions in various industries. In modern practice, most centrifugal compressors are equipped with dry gas seals as a standard sealing solution.

According to Stahley, J.S. (2002) in Dry Gas Seal System Design Standard for Centrifugal Compressor Application (p. 145), dry gas seals are non-contacting, dry-running mechanical face seals comprising two primary components: a rotating mating ring and a stationary primary ring. The core functions of dry gas seals include containment of process gas, prevention of internal contamination, and preservation of compressor performance and efficiency.

To ensure a reliable supply of seal gas, dry gas seal systems typically require an auxiliary support system. This auxiliary system comprising of Seal Gas Heaters ensures to deliver seal gas that is clean, free from contaminants and condensates, and at adequate pressure and flow to maintain optimal seal performance.

As stipulated in API Standard 692 (2018), heaters are mandated as standard components within DGS systems to enhance reliability and performance by mitigating condensation and temperature-related instabilities. Figure 1 illustrates a typical configuration of a dry gas seal supply system.

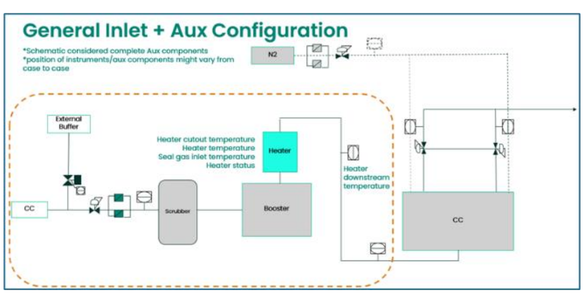


Figure 1: General configuration for Auxiliary system in DGS

Heater configurations are determined by process design requirements and may involve multiple heaters. In systems with more than one heater, the control logic becomes increasingly complex.

Historical data from the Baker Hughes' data repository has revealed several recurring anomalies related to DGS heaters, such as heater failure or underperformance, seal gas cooling below its dew point, and operational deviations. These issues have been linked to significant consequences, including moisture ingress into the seal chamber, contamination risks, and potential seal degradation. Additionally, heater malfunctions can compromise the system's ability to maintain optimal operating conditions, leading to increased energy consumption, reduced reliability, and, in severe cases, seal failure.

The Baker Hughes iCenter analytics platform enables diagnostic engineers to monitor and analyze heater-related anomalies in real time, providing timely alerts and preventive recommendations to customers. The platform delivers graphical trend visualizations, technical diagnostics, and actionable insights to facilitate root cause identification and streamline troubleshooting processes. To capture the anomalous operation of the heaters, a development of accurate analytical method is necessary which accurately captures the real anomalous behavior of heater operations. To address this, Baker Hughes conducted comprehensive studies, including requirement analysis, analytic model development, validation, and deployment on the platform.

Key monitored parameters include process variables, seal gas inlet temperatures, heater operational states (e.g., ON, OFF, CUTOFF, SELECTOR), and downstream temperature, pressure, and flow measurements. The analytic method should be able to detect a range of conditions, including abnormal temperature trends (slow degradation, fast degradation), heater malfunction against expected logic sequences (e.g., failure to cut out), and erratic toggling behavior.

2. DGS Auxiliary Heaters:

As outlined in the introduction, the dry gas seal heater is classified as an auxiliary or support subsystem within centrifugal compressor installations. Its primary function is to regulate the temperature of the seal gas to prevent moisture formation or condensation prior to entering the seal gas supply line.

Ensuring the gas is clean and dry is critical for protecting the seal faces and maintaining optimal system operability.

Typically, the seal gas heater regulates gas temperature to a predefined setpoint via an automatic thyristor-based control panel. In accordance with API Standard 692, a temperature transmitter must be integrated into the system to support safety interlocks and enable thyristor cut-out functionality in the event of abnormal temperature conditions. DGS Heaters can be classified based on their heating methodology into two main types: direct and indirect heating systems.

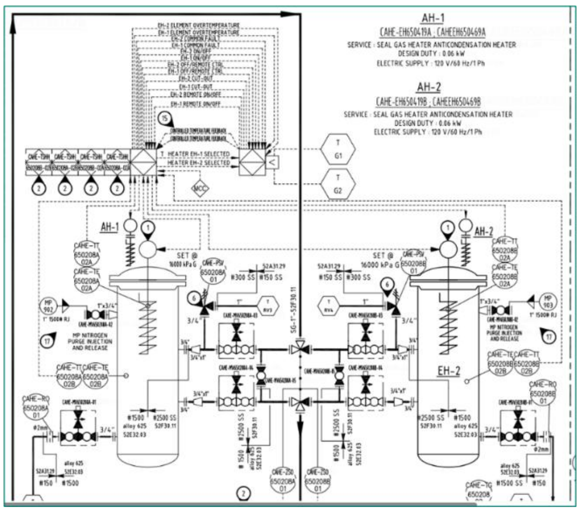


Figure 2: Dual heaters configuration

- Direct heaters utilize electric immersion elements that heat the seal gas directly through coils installed within the gas flow path. This method is typically employed for standard temperature applications where rapid and efficient heat transfer is required.
- Indirect heaters, on the other hand, are commonly used in applications involving higher seal gas temperatures or elevated dew points. These systems employ a heat transfer medium such as thermal oil or another fluid contained within a shell or jacket. A coil carrying the seal gas is submerged in or surrounded by the heated medium, allowing thermal energy to be transferred indirectly to the gas stream.

This classification allows for flexibility in heater selection based on process requirements, thermal performance, and safety considerations.

3. Pattern Recognition in Time Series:

Pattern recognition is the process of using machine learning algorithms to identify patterns and regularities in data. It finds use cases across domains- from stock market trends to recognizing visual patterns in images or repetitive sounds in audio signals. The data can be anything from text to images, audios or videos. In the context of time series data, Pattern recognition becomes especially important as it will allow us to uncover repeating trends or cycles.

Various types of Pattern recognition methods exist- from conventional statistical classifiers to ML based ones and advanced Deep Learning (DL) classifiers.

Besides the conventional use cases of identifying patterns, these methods can be used to identify anomalies in data by first detecting established patterns or expected behavior and then spotting the data points deviating from it.

Farahani et al. (2023) presented an extensive review of literature from 2002 to 2022 on Pattern recognition in time series data with special focus on applicability in Smart manufacturing systems. A summary of their work shows applications of DL/ANNs like FF Neural networks, RNNs to conventional techniques like SVMs, Decision trees and statistical classifiers like Naïve Bayes in previous research works.

Reger et al. (2018) applied pattern recognition using K means & support vector method of machine learning and compared with convolutional neural networks for manufacturing plants in both time and frequency domains.

Rewicki et al.(2023) compared six deep and classical methods for unsupervised detection in time series, applied on UCR Anomaly Archive data, which consists of 250 univariate time series from various fields including human medicine, biology, meteorology, and industry.

With development of more sophisticated models like Long Short-Term Memory, efforts have been made to apply these models on the time series data for identification of anomalous behavior.

Yang et al. (2018) applied LSTM model for Fault Identification in Turbomachines after converting Time Domain to frequency domain.

Homayouni et al. (2020) applied Autocorrelation based windowing technique on LSTM to accurately identify anomalies on time sensor data.

Meanwhile, Zhang et al. (2023) proposed Time Frequency Anomaly Detection model by decomposing time series for identifying specific types of anomalies.

4. Novel Methodology for pattern recognition:

The Heater toggling to On or Off operating condition occurs to control Seal gas supply temperature to the compressor at the setpoint. Figure 3 shows typical operations of heater for 2 different applications. State 0 shows Heater cut off and State -1 shows heater On.

Based on various parameters like incoming Seal gas temperature, process variations and heater degradation, it is observed that heater toggling following a nonstandard pattern are perfectly normal. Figure 3 highlights various toggling patterns occurring in a heater operation, which are non-standard, as observed by longer duration at which heater remains either On or Off for a specific period during a given cycle. However, they are considered as Normal operation. It to be further observed that a given pattern generally repeats multiple times.

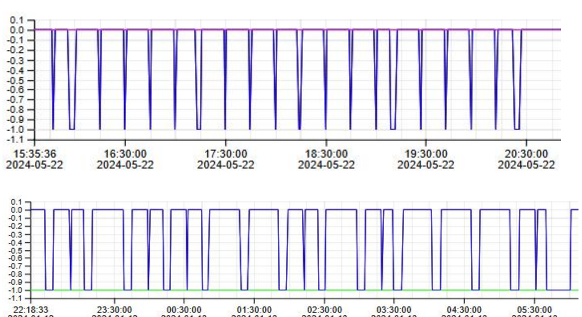


Figure 3: Typical operations of DGS Auxiliary Heater

The complex nature of DGS Auxiliary Heater renders application of any one method unhelpful to detect pattern change anomalies as the patterns are not distinctly repeating. Heaters can exhibit different patterns as their aim is to maintain seal gas supply temperature to compressor. In absence of a standard reference toggling pattern, it is not possible to directly apply conventional pattern recognition algorithms for identification of anomalous Heater operations. Hence, a hybrid approach was developed to first identify the established patterns and then detect deviations from these to spot anomalies. This paper presents a hybrid method to detect pattern anomalies in Time series data. The development is an in-series application of three analytical methods- Frequency Finder, Auto-correlation and Toggling Classifier.

Figure 4: Schematic of dataflow in pipeline

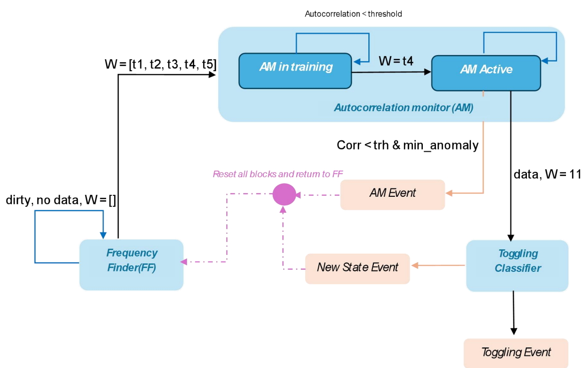
At any point of time, two out of three Heaters are required to operate to maintain the Skin and outlet temperature of Heater. Median value of all Heater running status tags is computed to arrive at one signal to process and identity pattern anomalies.

The solution follows a staggered approach for each of the components, which includes Training Phase, Active and Inactive phase. The three analytic components have an internal status which changes depending on the process sequence. Following is the high-level sequence of operation. With more detailed description of each of the components been discussed later.

1. At the beginning, Frequency Finder (FF) and Auto-correlation monitor (AM) are in training mode, learning on the incoming data, while the Toggling classifier (TC) remains inactive. In this training phase, FF and AM find most dominant window, which acts as a reference to analyze the upcoming data stream
2. In this step, FF becomes Inactive, while AM becomes Active & TC goes into training phase. TC receives the most dominant window size identified in the previous phase

and analyzes incoming hourly data to detect pattern changes in Heater operation.

3. AM and TC generate anomalies based on



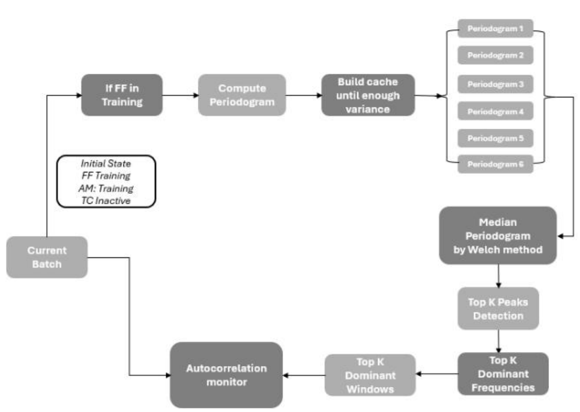
conditions elaborated in next section. Once an anomaly is generated by AM or TC blocks, entire pipeline is reset and returns to first sequence in the block i.e. FF to find new dominant Window size. An overall schematic of the pipeline is presented in Figure 5.

Figure 5: Event generation overview

A detailed breakdown of the pipeline is as below:

4.1. Frequency Finder:

As explained in figure 6, this component analyzes the signal to find dominant toggling periods by computing Periodogram of the signal in 1 hour batch and caches it to store Periodograms of 6 consecutive batches. Welch method, which is used for periodogram averaging, is then applied on these 6 cached Periodograms to compute one final Periodogram. The FF block then uses peak detection on the final periodogram to find top k dominant frequencies and corresponding time windows. The k dominant candidate windows are passed to the Auto-



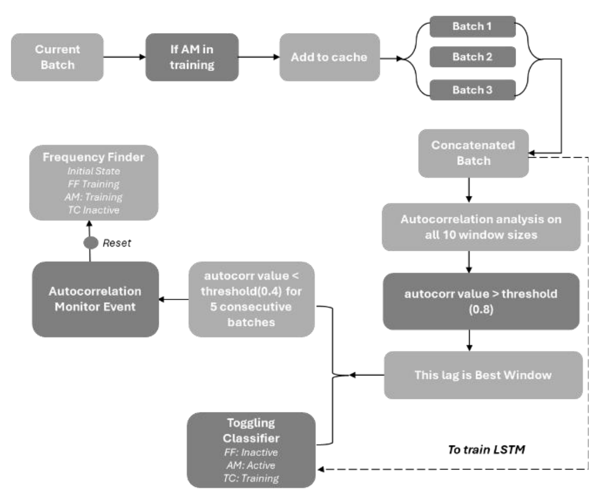
correlation monitor (AM) block.

Figure 6: Periodogram generation and dominant window identification

4.2. Autocorrelation Monitor:

Data for latest 3 hours is cached in Autocorrelation block. The top k Dominant candidate windows received from FF block are used to compute lagged autocorrelation to determine whether one period window, which gives autocorrelation value greater than the threshold, can be tracked as a reference, identified as W, for the Heater operating pattern. Autocorrelation block keeps computing lagged autocorrelations with this W for fresh batches of data.

As outlined in figure 7, an Autocorrelation anomaly is raised if for consecutive 5 batches, the value of autocorrelation is less than a minimum threshold. After the identification of anomalous operation, process of identifying the dominant frequency and window size via frequency finder and autocorrelation



is followed.

Figure 7: Autocorrelation analysis with dominant window size and event generation

4.3. Toggling Classifier:

If the autocorrelation anomaly is not raised in the previous stage, the dominant window size and the concatenated data is passed from AM to the TC block.

Heater toggling events which are more nuanced and couldn't be captured by AM block are detected with more advanced methodology in TC Block. The distinction between the two blocks doesn't lie in Heater's fault modes, but in their capabilities to detect pattern shifts based on subtlety and ease of identification. The Toggling Classifier block trains an LSTM on an augmented dataset composed of real and synthetic window sizes to identify pattern shifts from real window size.

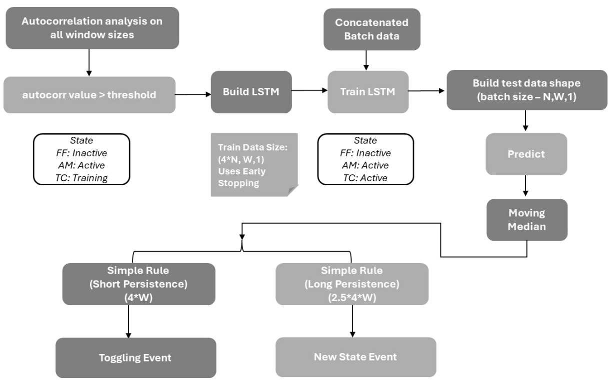
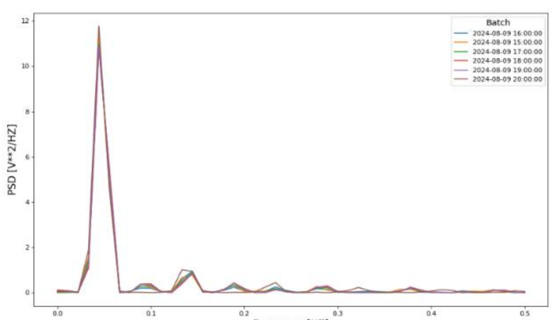


Figure 8: Toggling Classifier LSTM training and event generation for anomalous patterns

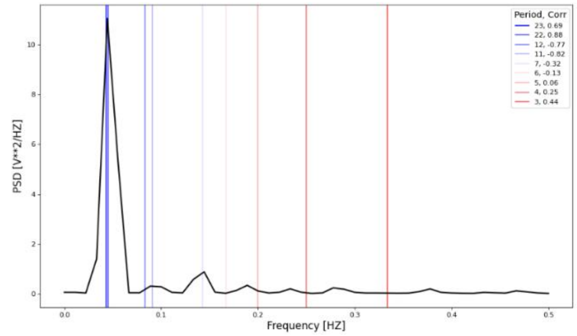
After the LSTM training, TC is set to Active, and the



network starts to predict in a rolling manner on incoming data streams. The Toggling Classifier tries to predict whether new Heater signal has same window patterns exhibited as the reference window. Otherwise, anomaly is raised.

5. RESULTS:

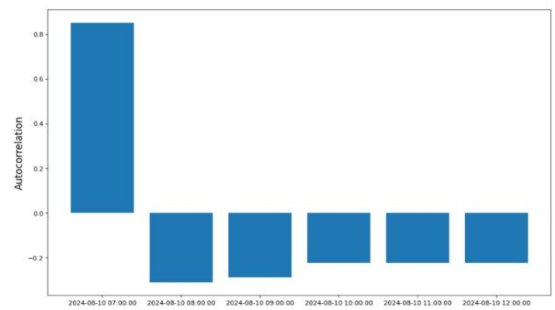
5.1 Demonstration of real Autocorrelation anomaly detection:



1. Six Periodograms are generated by six hourly batches as below:

Figure 9: Periodograms construction

2. Dominant frequencies are identified, and corresponding autocorrelations are computed



3. The frequency with autocorrelation value greater than the threshold of 0.8 is selected, time window corresponding to this frequency is taken as best lag.

Figure 10: Autocorrelation values at top window sizes identified

4. Anomaly is generated when autocorrelation value with this lag value is less than threshold of 0.4 for 5 consecutive batches.

Figure 11: Autocorrelation values before threshold of 0.4 for 5 consecutive batches

An example of Autocorrelation anomaly is as presented below where the Heater on and off pattern changes from high frequency to a low frequency mode.

Figure 12: Autocorrelation Anomaly Detection

Figure 13: Toggling Anomaly Detection

An example of detected Toggling anomaly where Heater on and off pattern shifts to a continuous Heater on mode.

5.2 Comprehensive result analysis on multiple compressors:

The algorithm was run on 26 centrifugal compressors for 1 year data sampled at 1 minute frequency. A detailed analysis of the results showed that out of total pattern change anomalies generated, ~45% were from Autocorrelation component while ~55% were from Toggling Classifier. To ensure, a robust solution which accurately captures real anomalous behavior of the heaters, a detailed verification of results was done by subject matter experts. The time series data is processed at one-minute frequency for pattern change identification. At this level of granularity, it is not feasible to manually and accurately identify Heater's non-standard toggling pattern changes. Hence, calculation of Recall and F1-score is limited. Therefore, model performance is evaluated using the Precision score. True and False positive assignment was done by experts for the algorithmically raised events, validating a Precision score of the method at more than 0.85.

Table 1. Result Statistics

Number of assets	26
Duration	12 months
Sampling frequency	1 minute
Autocorrelation component anomalies	~ 45%
LSTM component anomalies	~ 55%
Precision	> 0.85

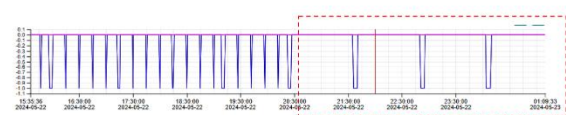
6. CONCLUSION:

This paper highlights the critical role of Auxiliary Heater for smooth operation of Dry Gas System of Centrifugal compressors. It establishes the importance of identifying anomalous operation of the heater to safeguard the system. The paper further demonstrates that due to complex nature of operation and interactions of systems, known methods of identifying pattern anomalies cannot be applied as is for the accurate detection of anomalous operation of real time industrial application. This paper then establishes a novel in series application of diverse techniques such as Frequency Analysis, Auto Correlation analysis and LSTM model to develop a robust solution which can accurately identify anomalous operation of the heaters. The method presented is capable of self-learning which is a critical aspect to enable a generalized application of the solution.

Lastly the paper talks of implementation of solution in Industrial IoT and Human in loop validation on Real time data coming from 26 centrifugal compressors running across the globe with observation period of 1 year. It calculates the precision score and further segregation on type of anomalies based on the inherent signature to provide critical insights to operation. The analytical solution has been deployed to detect Toggling pattern change anomalies in real time for continuous monitoring and is being used by engineering support teams for issue identification with Seal gas heaters thereby improving reliability and efficiency of Seal gas systems.

NOMENCLATURE

DGS: Dry Gas Seal
LSTM: Long Short-Term Memory
FF: Frequency Finder
AM: Autocorrelation Monitor
TC: Toggling Classifier



REFERENCES

Stahley, J. S. (2002). Dry Gas Seal System Design Standard for Centrifugal Compressor

Applications. Texas A&M University. Turbomachinery Laboratories

Yang, R., Huang, M., Lu, Q., & Zhong, M. (2018). Rotating Machinery Fault Diagnosis Using Long-Short-Term Memory Recurrent Neural Network. IFAC-PapersOnLine, 51(24), 228–232. doi:10.1016/j.ifacol.2018.09.392

Homayouni, H., Ghosh, S., Ray, I., Gondalia, S., Duggan, J., & Kahn, M. G. (2020). An autocorrelation-based LSTM-Autoencoder for anomaly detection on time-series data. 2020 IEEE International Conference on Big Data (Big Data), pp.5068–5077. doi:10.1109/BigData50022.2020.9378192

Rewicki, F., Denzler, J., & Niebling, J. (2023). Is it worth it? Comparing six deep and classical methods for unsupervised anomaly detection in time series. Applied Sciences, 13(3), Article 1778. doi:10.3390/app13031778 Rewicki, F., Denzler, J., & Niebling, J. (2023). Is it worth it? Comparing six deep and classical methods for unsupervised anomaly detection in time series. Applied Sciences, 13(3), Article 1778. doi:10.3390/app13031778

Reger, A., Dumler, J., Lobachev, O., Neuberger, J., & Steinhilper, R. (2018). Machine allocation via pattern recognition in harmonic waves of manufacturing plants. Procedia CIRP, 67, 69–74. doi:10.1016/j.procir.2017.12.187

Farahani, M. A., McCormick, M. R., Gianinny, R., Hudacheck, F., Harik, R., Liu, Z., & Wuest, T. (2023). Time-series pattern recognition in Smart Manufacturing Systems: A literature review and ontology. Journal of Manufacturing Systems, 69, 208–241. doi:10.1016/j.jmsy.2023.05.025

Zhang, C., Zhou, T., Wen, Q., & Sun, L. (2022). TFAD: A decomposition time series anomaly detection architecture with time-frequency analysis. Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22), 2497–2507. doi:10.1145/3511808.3557470

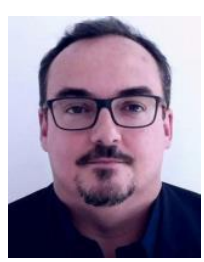
BIOGRAPHIES

Meenali Sharma is Staff Data Scientist at Baker Hughes, Bangalore, India. She received her Bachelor of Technology in Chemical Engineering from Nirma University and Post Graduate Diploma in Data Science from International Institute of Information Technology, Bangalore in India. In her current role, Meenali works on development of predictive analytics solutions for Baker Hughes' Turbomachinery for performance improvement.

Carmine Allegorico is Senior Principal engineer and experienced data scientist at Baker Hughes, Firenze, Italy. He received his master's degree in mechanical engineering from University of Napoli Federico II. In his current role, Carmine is a technical point of reference for the analytics discipline providing engineering guidance to other teams, helping to train new engineers and keeping abreast of industry trends and issues. He provides consulting during the development and implementation of advanced solutions for the on-line diagnostic and predictive maintenance, coordinates the creation of internal processes and support the adoption of new platforms and technologies.

Laura Nuti is Senior Data Scientist Manager at BH, energy technology Company. She has a master's degree in Statistics, with a math background. Laura plays a pivotal role in transforming the way in which organization use data to inform strategy and improve efficiency. In her current role, Laura leads engineering digital innovation and manage complex project to prevent issue in turbomachinery environment. With more than 13 years of experience in Industry and Energy sector, Laura contributes with expertise and leadership significantly to BHs' mission of advancing innovation in the energy market.

Unnat Mankad is a Principal Data Scientist at Baker Hughes, Bengaluru, India. He received his master's degree in mechanical engineering from Birla Institute of Technology and Science (BITS), Pilani. In his current role, Unnat develops complex analytics for on-line prognostics, diagnostics, and predictive maintenance of Baker Hughes' iCenter covered turbo machines. As expert of time series sensor data analysis, he also supports data science and machine learning driven algorithm developments focusing on emissions, reliability, and availability improvements.



Gabriele Mordacci is the iCenter engineering Manager at Baker Hughes, Kuala Lumpur, Malaysia. He received his master's degree in aerospace engineering from Universita degli studi (UNIPI), Pisa Italy, with a final thesis on the High-speed gearbox diagnostic with microphone dynamic analysis. He worked in RMD developing the first analytics based on the fleet statistical for different type of technology and translating the field experience in digital application. As OEM expertise supports diagnostic and data scientist to prepare and validate the new analytics algorithm and output.

Rajakumar Dharmaraj is an experienced Instrumentation and Electrical Specialist with deep expertise in Distributed Control System (DCS) design, system integration, and control logic implementation. He brings extensive field experience in commissioning, startup, and troubleshooting of turbomachinery—including gas turbines, turbo expanders, compressors, and associated auxiliaries—delivering operational excellence across the energy and industrial sectors. Currently, Rajakumar serves as a Senior Diagnostic Engineer at ICenter where he plays a pivotal role in diagnostics and remote monitoring of critical equipment. His responsibilities include analyzing data, identifying anomalies, and providing actionable insights to enhance reliability and reduce downtime. Leveraging his hands-on field expertise and systems knowledge to contribute to analytic development for continuous improvement initiatives.

Aidil Fazlina Binti Hasbullah is a Diagnostic Engineer at Baker Hughes, Kuala Lumpur, Malaysia. She received her former tertiary education from Universiti Teknologi PETRONAS (UTP) in Electrical & Electronic Engineering major in Instrumentation and Control. In her current role, Aidil Fazlina focuses on remote monitoring and diagnostics of turbo machines of global installed fleet. She also act as focal point for onshore LNG plants for Malaysia and India fleet. Prior joining Diagnostic team, she has experience of working with LNG plant as Instrument engineer, leading Factory Acceptance test for instrument control system. She also actively collaborates with Baker Hughes' Data Scientists to support advanced analytics development to improve the reliability and availability of the turbo machines.