Proceedings of the 7th European Conference of the Prognostics and Health Management Society 2022 - ISBN – 978-1-936263-36-3

Data-Driven Fault Detection for Transmitter in Logging-While-Drilling Tool

Karolina Sobczak-Oramus[1], Ahmed Mosallam[2], Caner Basci [3], and Jinlong Kang[4]

[1] *Schlumberger, Nowogrodzka Street 68, Warsaw, Mazowieckie, 02-014 Poland*
KSobczak@slb.com

[2,4] *Schlumberger, 1 Rue Henri Becquerel, 92140 Clamart, France*
AMosallam@slb.com
JKang5@slb.com

[3] *Schlumberger, 2-2-1 Fuchinobe, Chuo Ward, Sagamihara, Kanagawa 252-0206, Japan*
CBasci@slb.com

ABSTRACT

Logging tools widely used in the oil and gas industry are exposed to demanding environmental conditions that can lead to faster degradation and unexpected failures. These events can reduce productivity, delay deliverables, or even bring entire drilling operations to an end. However, such accidents can be avoided using a prognostics and health management approach. This paper presents a data-driven fault detection method for transmitter in logging-while-drilling tool adopting a support vector machine classifier. The health analyzer determines the component's physical condition in just a few minutes, demonstrating an exceptional value for both field and maintenance engineers. This work is part of a long-term project aimed at constructing a digital fleet management system for downhole testing tools.

1. INTRODUCTION

Prognostics and Health Management (PHM) combines the knowledge and experience from several disciplines such as engineering science, computer science, reliability engineering, and more, to assess a product's degradation and reliability. PHM has emerged recently as a momentous technology that makes an impact on maintenance practices for different industrial systems (Vachtsevanos, Lewis, Roemer, Hess, & Wu, 2006). Design, monitoring, and maintenance of complex systems such as aircraft, manufacturing, and industrial processes, and more have undergone a real transformation, being more data-driven thanks to the usage of PHM practices. PHM technologies are quickly evolving, the customer base for these technologies is expanding, and their potential application domains are increasing at a phenomenal rate (Vachtsevanos et al., 2006).

Prognostics and health management consist of the following main pillars: fault detection, fault diagnostics, fault prognostics, and decision support (Mosallam, Medjaher, & Zerhouni, 2016). Fault detection is used to determine that a problem has occurred within the monitored component. Fault diagnostics is the process of identifying faults and their causes. Fault prognostics helps estimate the remaining time left for a system or a component before it fails. Finally, decision support is used to select the proper maintenance actions based on the information gathered about the monitored system status.

SonicScope multipole sonic-while-drilling service tool (Figure 1) is a multi-function logging-while-drilling (LWD) tool developed for oil well drilling applications available in the following collar sizes:

- $4\frac{3}{4}$ in,
- $6\frac{3}{4}$ in,
- $8\frac{1}{4}$ in,
- 9 in.

It is commonly used in conjunction with other LWD equipment during the drilling phase of well construction.

Figure 1. SonicScope service tool in different collar sizes

LWD tools are often exposed to demanding environmental conditions such as shocks, vibrations, pressures, and elevated temperatures (Kirschbaum et al., 2020). Consequently, the degradation rate of the subsystems in the tools can increase over time resulting in tool failures. Hence, the information gained from the tools can be inaccurate, and this could compromise the operation. As a result, the deliverables can be delayed until the tool is fixed or, in a worst-case scenario, the entire operation may be canceled. Such situations lead to non-productive time and financial losses. To avoid such failures, after each run, field engineers are required to check the tool condition. Each tool consists of multiple different subsystems that contain many parts. To assess the overall tool condition, field engineers should analyze sensor signals for each part recorded during tool operation. Consequently, they need to decide if the tool and its subsystems are healthy and can be used again in the next run. Depending on the subsystem, they must decide if the tools should be repaired or junked. However, due to the large number of data channels generated at a record rate, which results in millions of data points from a single run, developing manually a solid analysis is extremely challenging (Mosallam, Laval, Youssef, Fulton, & Viassolo, 2018). A manual analysis of this data is time-consuming in an environment where time is critical, and the complexity of the signals limits the effectiveness of manual analysis.

Alternatively, the critical subsystems in the tool can be identified and a domain expert can select the channels that contain information about the tool condition and possible degradation of each subsystem. Statistical features that indicate the degradation of the system in time are extracted from the selected channels. These features can be used to build machine learning models that estimate the tool condition. Using the SonicScope service tool, a transmitter subsystem was identified as one of the most critical components by failure modes, effects, and criticality analysis (O'Connor & Kleyner, 2012). Thus, a fault detection algorithm was developed to help the field engineers identify whether the component behaved as expected or not.

In this paper, we present a data-driven fault detection method for transmitters in drilling tools. The transmitters are of distinct types and sizes depending on the version of the tool. As a result, the models and features extracted from the raw data differ between the types and sizes of the transmitters. The method is based on extracting relevant features that can identify healthy and faulty transmitters. A support vector machine model is trained on the features extracted from different runs labeled as healthy or faulty by a domain expert.

This paper is structured as follows. A literature review is presented in Section 2. Section 3 presents a description of the transmitter subsystem. The method and the results for the fault detection model are presented in Section 4. Finally, Section 5 concludes the paper.

2. Related Work

Transmitters are widely used in different types of equipment in different industries. Thus there are plenty of research works on transmitter fault detection. For instance, (Ganesh Kumar, Insozhan, & Parthasarathy, 2019) uses a fuzzy inference system to monitor transmitter circuit conditions in a wireless sensor network. (Tugova, Salov, & Bushuev, 2021) presented a fault diagnosis method of a pressure transmitter based on output signal noise characteristics. (S. Liu, Xu, Li, Zhao, & Li, 2018) proposed a hybrid fault diagnosis model for transmitters in water quality monitoring devices based on multiclass support vector machines in combination with rule-based decision trees. (C. Liu, Chen, Zhang, & Wang, 2018) introduced a fault diagnosis application of a short wave transmitter based on a stacked auto-encoder.

Generally, transmitters for different devices have different designs, which makes the fault detection method of the transmitter cannot be used universally. The device studied in this paper is a specific tool used for logging when drilling oil and gas wells. To the best of our knowledge, there is no published research work about LWD transmitter fault detection or fault diagnosis. Therefore, our study can give an idea of how to proceed with the specific case of transmitter's diagnostics.

3. Transmitter

SonicScope service combines high-quality monopole and quadrupole measurements to obtain compressional, shear, and Stoneley data in all formations and across a wide range of hole sizes from surface to true depth. Combined with a full-characterized tool design, simplified and automated operations, and advanced processing techniques, the SonicScope service delivers robust, accurate, and reliable acoustics measurements for many applications from petrophysics to cement evaluation.

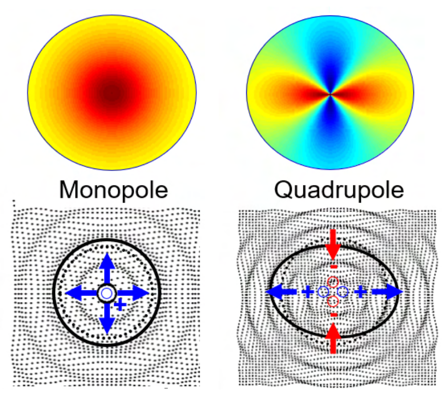

Figure 2. Monopole and quadrupole transmitters

SonicScope service carries two types of transmitters (Figure 2): monopole and quadrupole. Monopole transmitter is mainly used to measure the formation slowness for fast formations and quadrupole is used to measure the formation slowness for slow formations. Acoustic waves generated by the transmitters are captured by the array of receivers with a total quantity of 48. These receivers are located around all four sides of the tool. The transmitter consists of piezoceramic elements that generate acoustic waves when excited with high voltage. As a tool diagnostics, the tool records each transmitter's excitation voltage to interpret the health of the transmitter.

4. Proposed Method

We present a data-driven fault detection method to analyze the health of a transmitter. There are a couple of steps to conduct the data-driven PHM projects. The first phase of the fault detection approach is to conduct the data inventory, which allows collecting the data from different sources and services. The second step is to identify the scope, particularly perform an analysis of different failure modes and select the most critical ones. Several different problems can cause transmitter failures, such as low insulation resistance, different physical damages, and more. Subject matter experts (SMEs) select the critical problem in terms of priority, which in the case of the transmitter is insulation breakdown. Next, there are a few other phases like channel selection, data preprocessing and labeling, feature engineering, and modeling, which will be described in detail in the subsequent sections.

Cooperation with the SMEs is crucial to learn the details about the components' usage, health and failure patterns. The data scientists perform the research, and analysis based on the suggestions of SMEs and ensure that the data-driven approach is held. Moreover, the fact that the development of machine learning models requires specific knowledge and often quite a lot of experience leads to the fact that the models are developed by the data scientists with the support of SMEs, and not SMEs themselves.

The reason for developing the models in a data-driven approach, rather than a model-based approach, is that we can build models faster with less cost, using the historical data (Mosallam, 2014). Moreover, machine learning models provide efficiency and consistency, and are not dependent on the individual's mistakes. Furthermore, the model-based approach requires a deep understanding of the physical mechanism of the failure, extensive experimentation, expert knowledge, and model verification, which is highly time-consuming (Mosallam, 2014).

4.1. Data Description

After each run, tool data consists of a hundred number of data channels generated at a record rate, which results in millions of data points from a single run. Only some of the channels hold valuable information about the health of a transmitter. Removing the channels that do not contain useful information about the health of a component should help to reduce the noise and, consequently, increase the algorithm efficiency. SME domain knowledge determines the choice of channels that hold relevant information about the health status. Depending on the collar size of the tool, there can be a single type of transmitter (monopole & quadrupole) or separate monopole and quadrupole transmitters used during the run. For each transmitter, data is stored in two firing voltage channels with positive polarity and negative polarity, from which data can be collected from both or a single channel, depending on the operation mode. Based on those channels the features should be created to distinguish between a healthy and a faulty transmitter, and used to build the fault detection model. Taking into consideration that there are different types of transmitters (monopole & quadrupole) and that they differ between multiple collar sizes of the tool, in the subsequent sections, we are going to cover the details for each size and type of transmitter.

4.2. Data Preprocessing

Once the tool is initialized and put in the well, the onboard system records measurements of transmitters' voltage every 10 seconds. This data is available once the job is done and the tool is back to the surface. During the run, there can be situations when the transmitter is not firing, which affects in the measurement of 0 Volts. Such observations should be filtered out and not taken into consideration in the analysis. Figure 3 presents the time series of the transmitter's raw voltage and voltage after preprocessing.

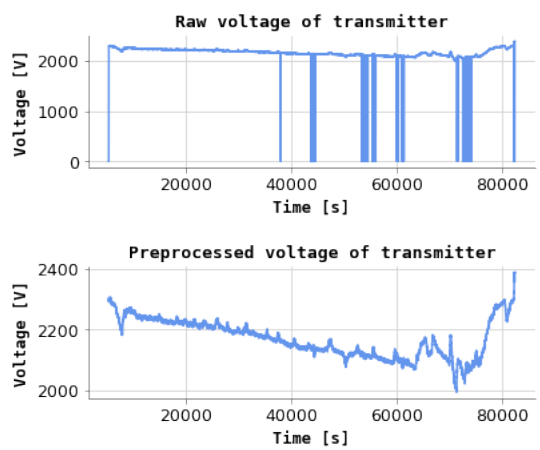

Figure 3. Transmitter's raw voltage and preprocessed voltage

4.3. Data Labeling

The step of labeling is crucial for the successive steps of feature engineering and modeling. This phase is performed by the SMEs who, using their domain knowledge and experience, and taking into consideration the available status of the run, label the transmitter as healthy or faulty.

4.4. Feature engineering

The goal of this step is to transform the original channels raw data into features that represent the transmitter's health after each run in a statistical way. Performing this phase requires a good understanding of the failure mode. It's necessary to know the symptoms, as well as how the whole subsystem works and how transmitters cooperate within the system. Based on the collar size of a tool and transmitter type, the features can be extracted in different ways, depending on how the subsystem works. For each transmitter, to extract the features for modeling, we use data from one voltage channel. The decision if we use positive or negative polarity is made after consultation with SMEs about the operation mode for a certain tool type.

Tool A: Monopole & Quadrupole Transmitters For this collar size of the tool, four single monopole & quadrupole transmitters are used during the same run. The behavior and firing of transmitters are similar throughout the run if each of them is healthy, which can be seen in Figure 4. Based on the analysis of the dataset, during the run, one or more transmitters can fail (but there is no run in the available data where all of them failed together). Figure 5 shows the transmitters voltages for the run when Transmitter 3 has failed. There is a visible drop-down in the voltage of that transmitter, while the other transmitters' voltages remain similar. The analysis showed that the drop-down in the failed transmitter's voltage differs between different failures. Therefore the features will be extracted, using the co-dependency of the transmitters, not simply taking into consideration a single transmitter's voltage.

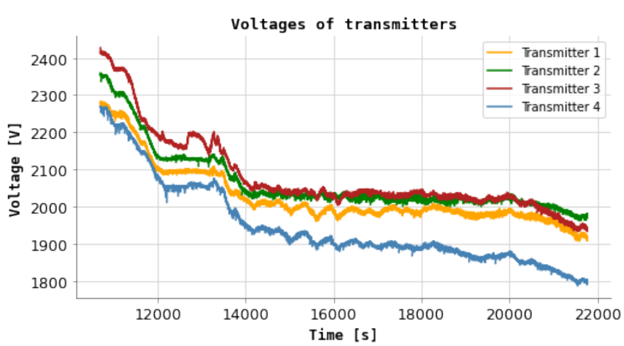

Figure 4. Tool A transmitters' voltages

Let T_i for $i \in \{1, 2, 3, 4\}$ denote a time series of voltage for transmitter used in a selected run. Let n denote the duration time (in seconds) of a run and $s \in \{1, ..., n\}$ denote the moment in the run. Consequently,

$$T_i = [t_1, ..., t_n] \tag{1}$$

is a vector of voltages for i-th transmitter within the selected

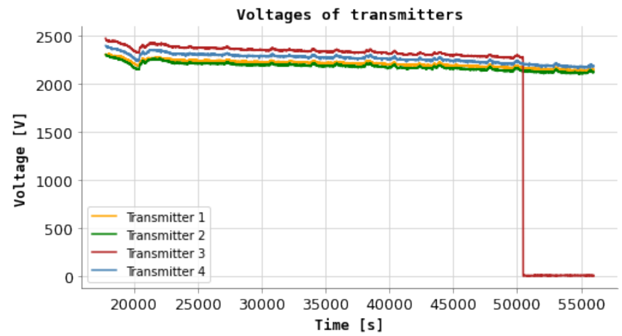

Figure 5. Tool A transmitters' voltages with one failed case

run. We calculate the features taking into consideration the maximum difference between the transmitters' voltages to differentiate the failed cases from healthy ones. For failed transmitters, depending on how many transmitters fail at once, at least one maximum difference will be high (as mentioned above, there were no runs for which we observed failures of all the transmitters). The following formula is used to calculate the feature x_k for $k \in \{1, 2, 3\}$, for the transmitter T_i:

$$x_k = \max_{s=1,...,n} |T_i - T_j|_s, \tag{2}$$

where $j \in \{1, 2, 3, 4\}$ and $j \neq i$. As a result, we get three dimensional space of features for each transmitter.

Tool B & C: Monopole & Quadrupole Transmitters In these tool types the monopole and quadrupole transmitters, are separate components. Four quadrupole and one monopole transmitters are used during the run. The reasons for creating features differently than for the transmitters of Tool A are the following:

- Like the transmitters described above, quadrupole behavior and firing are similar throughout the run if each is healthy, which can be seen in Figure 6. But we can observe that the firing mode is different from Tool A. It is noisier and the voltage can change drastically within 10 seconds (the spikes do not have to appear exactly at the same moment for all the transmitters).

- Monopole transmitters' voltage is similar to the voltage of transmitters of Tool A. Despite that, the fact that a single monopole transmitter is used within the tool requires different feature creation.

In the development phase, multiple statistical measures were checked such as a correlation between channels, standard deviation, minimum and maximum voltage, and so on, to highlight those that best separate healthy and faulty transmitters. Let T_i be defined by Equation (1). The following formula is

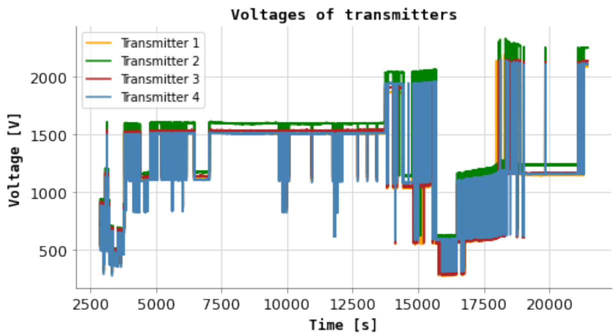

Figure 6. Tool B quadrupole transmitters' voltages

used to calculate the feature x for the transmitter T_i:

$$x = \min_{s=1,\ldots,n} t_s, \tag{3}$$

4.5. Modeling

This phase associate the features created for each transmitter with the labels assigned by SMEs to map the relation between them (x, Y) where:

- for a single monopole & quadrupole transmitter from Tool A:
$$x = [x_1, x_2, x_3], \tag{4}$$
where x_i for $i \in \{1, 2, 3\}$ is defined by the formula (2).
- for both monopole and quadrupole transmitters from Tool B and Tool C: x is defined by the formula (3)

and for all kinds of transmitters Y is given by the formula:

$$Y = \begin{cases} 1, & when\, transmitter\, is\, failed \\ 0, & when\, transmitter\, is\, healthy \end{cases} \tag{5}$$

Features and corresponding labels are used to develop the models. Models are trained separately for the following types of transmitters:

- Tool A monopole & quadrupole transmitters
- Tool B quadrupole transmitters
- Tool C quadrupole transmitters
- Tool B & C monopole transmitters

To determine if the transmitter is healthy or faulty, we used the classification model. For each type of transmitter support vector machine (SVM) model with linear kernel was trained. In SVM, a hyperplane is constructed in n-dimensional space, where n is the number of features used in the model, in a way that best separates healthy and faulty classes (Hastie, Tibshirani, & Friedman, 2009).

Due to imbalanced class distribution, the evaluation of the model is performed using both the Accuracy and F1-score (Fernández et al., 2018). The classification metrics are calculated as follows:

$$F1 = 2 * \frac{precision \cdot recall}{precision + recall} \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

where

$$precision = \frac{TP}{TP + FP} \tag{8}$$

$$recall = \frac{TP}{TP + FN} \tag{9}$$

and

- TP stands for $True Positive$ which is the number of the transmitters that model correctly classified as healthy,
- TN stands for $True\ Negative$ which is the number of the transmitters that model correctly classified as faulty,
- FP stands for $False\ Positive$ which is the number of the transmitters that are classified as healthy when they are faulty,
- FN stands for $False Negative$ that is the number of the transmitters classified as faulty when they are healthy.

The number of transmitters (with regards to healthy and faulty) used to develop the fault detection models is presented in the Table 1.

Table 1. Number of healthy and faulty transmitters per transmitter type

Transmitter Type	**Healthy**	**Faulty**
Tool A monopole & quadrupole transmitters	68	8
Tool B quadrupole transmitters	43	9
Tool C quadrupole transmitters	71	17
Tool B & C monopole transmitters	31	10

Figures 7, 8, 9, 10 present trained SVM models with corresponding hyperplanes (in 1-dimensional and 3-dimensional spaces depending on the features used for modeling). Healthy transmitters are indicated by green points and faulty by red ones. It is visible that all types of transmitters classes are well separated, which is a good premise for prediction.

Before the SVM classification models were developed, simple thresholds were used to differentiate between healthy and faulty transmitters for all transmitters types. One of the significant aspects of model deployment is constant model improvement and evaluation. Therefore, the models should be automatically retrained on a scheduled basis so that any new phenomenon in the data could be incorporated and the SVM hyperplane could be moved. Due to that, it was decided to deploy the SVM models instead of setting simple thresholds which would have to be manually reworked after some time.

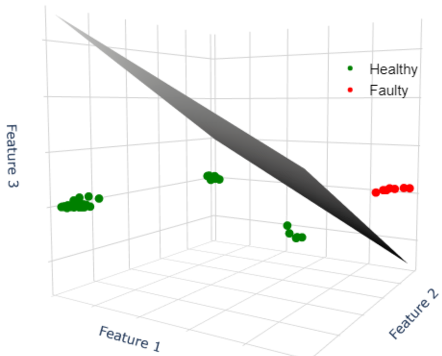

Figure 7. SVM hyperplane separating healthy and faulty monopole & quadrupole transmitters of Tool A

For Tool A monopole & quadrupole transmitters we observe three clusters of healthy transmitters. This phenomenon appears simply because we calculate the features in a certain order as it is mentioned in Equation (2). For example, for the Transmitter 1:

$$x_1 = \max_{s=1,\ldots,n} |T_1 - T_2|_s,$$

$$x_2 = \max_{s=1,\ldots,n} |T_1 - T_3|_s,$$

$$x_3 = \max_{s=1,\ldots,n} |T_1 - T_4|_s,$$

while for the Transmitter 2:

$$x_1 = \max_{s=1,\ldots,n} |T_2 - T_1|_s,$$

$$x_2 = \max_{s=1,\ldots,n} |T_2 - T_3|_s,$$

$$x_3 = \max_{s=1,\ldots,n} |T_2 - T_4|_s.$$

Therefore, the healthy cases are not grouped together. Looking at the three healthy clusters, we could conclude even more than the fact that a certain transmitter is healthy. We could use it to precisely say how many transmitters failed during the particular run (1, 2, or 3). However, for now, we use this model as a fault detection one, but there is a potential to be used in the future as a diagnostics model.

Model Performance Each machine learning model should be evaluated to check its performance and ability to generalize the learned pattern. Because we have highly limited amounts of data (see Table 1), a leave one out cross-validation method (LOOCV) was applied. This approach is a special case of K-fold cross-validation where the number of folds equals the number of transmitters that we have in the dataset.

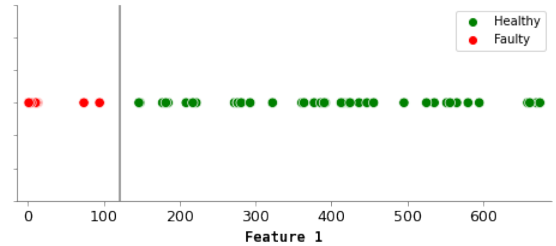

Figure 8. SVM hyperplane separating healthy and faulty quadrupole transmitters of Tool B

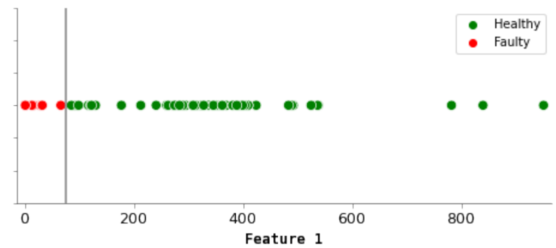

Figure 9. SVM hyperplane separating healthy and faulty quadrupole transmitters of Tool C

Thus, the algorithm is applied once for each transmitter, using the rest of the transmitters as a training dataset and adopting the selected one as a single test dataset. This method provides a low biased test accuracy and F1-score compared to using a single test dataset (Witten, Frank, & Hall, 2011). Table 2 presents the outcomes of the models' performance based on LOOCV.

The models show high confidence with zero misclassification for Tool A monopole & quadrupole transmitters and Tool B quadrupole transmitters and only one misclassification for Tool C quadrupole transmitters and Tool B&C monopole transmitters, and hence high accuracy and F1-Score for each one of them. It is worth mentioning that the performances of the models are very high due to good feature engineering, not only thanks to the SVM model. Before the SVM models were chosen to be deployed, they were compared to other algorithms, to choose the best performing and most beneficial one. Performance and benefits of using the SVM model were compared for example to the usage of logistics regression. The results were the same for the 1-dimensional space, but SVM outperformed the logistics regression for the 3-dimensional space. Therefore, it was decided to choose the SVM models to be implemented.

5. Conclusion

This paper presents a data-driven fault detection method for transmitters. Characteristics of the different types of transmitters were detected and analyzed in the exploratory data analysis phase. Each transmitter has two voltage channels that are

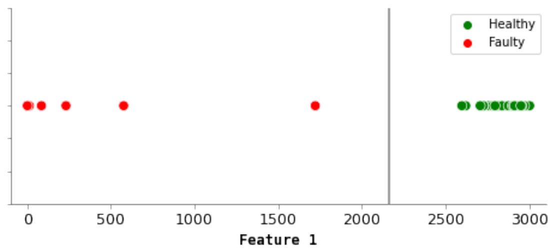

Figure 10. SVM hyperplane separating healthy and faulty monopole transmitters of Tool B&C

Table 2. Model performance metrics per transmitter type

Transmitter Type	Accuracy	F1-Score
Tool A monopole & quadrupole transmitters	100%	100%
Tool B quadrupole transmitters	100%	100%
Tool C quadrupole transmitters	99%	97%
Tool B & C monopole transmitters	98%	98%

used to construct representative features to identify the health status of a component. The classification model training is performed using these features to ensure the predictive power of the model, which is then used by the engineers and maintenance teams after each run to validate if the transmitter is healthy or not. The model performance validation resulted in a high F1-Score, which shows that the health of a transmitter can be identified correctly.

The proposed solution is deployed in the application that can be directly used by the field engineers and maintenance team to organize and plan their work more efficiently. The models and their performance are going to be constantly monitored and tested, and further improved if needed. Additional work is planned to build the diagnostics and prognostics models for the transmitters. However, several challenges such as the low historical data availability and uncertainty about the incipient failure mode of the components need to be considered and reworked.

REFERENCES

Fernández, A., García, S., Galar, M., Prati, R., Krawczyk, B., & Herrera, F. (2018). Performance measures. In *Learning from imbalanced data sets.* Springer, Cham.

Ganesh Kumar, D., Insozhan, N., & Parthasarathy, V. (2019). Recognition of faulty node detection using fuzzy logic in iot. *International Journal of Scientific and Technology Research*, *8*(12), 1112 – 1116.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Support vector machines and flexible discriminants. In *The elements of statistical learning.* Springer, New York, NY.

Kirschbaum, L., Roman, D., Singh, G., Bruns, J., Robu, V., & Flynn, D. (2020). AI-driven maintenance support for downhole tools and electronics operated in dynamic drilling environments. *IEEE Access*, *8*, 78683-78701. doi: 10.1109/ACCESS.2020.2990152

Liu, C., Chen, B., Zhang, H., & Wang, X. (2018). Fault diagnosis application of short wave transmitter based on stacked auto-encoder. In *2018 ieee 4th international conference on computer and communications (iccc)* (p. 119-123).

Liu, S., Xu, L., Li, Q., Zhao, X., & Li, D. (2018). Fault diagnosis of water quality monitoring devices based on multiclass support vector machines and rule-based decision trees. *IEEE Access*, *6*, 22184-22195. doi: 10.1109/ACCESS.2018.2800530

Mosallam, A. (2014). *Remaining useful life estimation of critical components based on bayesian approaches.* (PhD dissertation). Université de Franche-Comté.

Mosallam, A., Laval, L., Youssef, F. B., Fulton, J., & Viassolo, D. (2018). Data-driven fault detection for neutron generator subsystem in multifunction logging-while-drilling service. In *PHM society european conference.*

Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, *27*, 1037–1048. doi: 10.1007/s10845-014-0933-4

O'Connor, P., & Kleyner, A. (2012). *Practical reliability engineering, fifth edition.* John Wiley & Sons, Ltd.

Tugova, E. S., Salov, D. D., & Bushuev, O. Y. (2021). Diagnostics of a pressure transmitter based on output signal noise characteristics. In A. A. Radionov & V. R. Gasiyarov (Eds.), *Proceedings of the 6th international conference on industrial engineering (icie 2020)* (pp. 1298–1307). Springer International Publishing.

Vachtsevanos, G., Lewis, F. L., Roemer, M., Hess, A., & Wu, B. (2006). Intelligent fault diagnosis and prognosis for engineering systems. In *1st ed. hoboken.*

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

BIOGRAPHIES

Karolina Sobczak-Oramus is a data scientist at Schlumberger. She holds a Master of Science in Mathematics from the Jagiellonian University of Kraków, Poland. Her main research interests are in the fields of machine learning, artificial intelligence and data mining.

Ahmed Mosallam is an analytics manager at Schlumberger technology center in Clamart, France. He has his PhD degree in automatic control in the field of PHM from the University of Franche-Comté. His main research interests are signal processing, data mining, machine learning and PHM.

Caner Basci is a senior electrical sustaining engineer at Schlumberger. He holds an M.S. in the field of Very Large Integrated Circuits about Complementary Metal Oxide Semiconductor Image Sensors in Electrical Engineering Department at the University of Tokyo, Japan. His main interests are Power Electronics Circuits, Analog Circuit Design, Data Processing and Oil & Gas Exploration with Acoustics.

Jinlong Kang is a PhD student at University of Franche-Comté in Besançon and a data scientist at Schlumberger technology center in Clamart, France. He holds a B.S. in Industrial Engineering and an M.S. in Mechanical Engineering both from University of Electronic Science and Technology of China. His main research interests are Prognostic and Health Management, maintenance decision-making, data mining and machine learning.