



COB-2021-0985

STANDARDS CLASSIFIER: APPLICATION TO OPERATIONAL SUPERVISION OF DIESEL GENERATING UNITS

Rafael Novais Lacerda de Oliveira

Guilherme Gomes da Silva

Edgar Campos Furtado

Centre for Innovation, Research and Teaching in Mechatronic (NIPEM).

Telecommunications and Mechatronic Engineering Department (DETEM).

Federal University of São João del-Rei (UFSJ)

Rod. MG-443, km 7. Ouro Branco, MG, Brazil. CEP: 36420-000.

nlacendarafael@gmail.com

guilhermegomes@ufs.j.edu.br

edgar@ufs.j.edu.br

Carlos Eduardo Alves da Costa

Brentech Energia S.A.

R. 24 de outubro, SN, Jardim Ipanema. Aparecida de Goiânia, GO, Brazil. CEP: 74.984-290.

carlosalves.eng@gmail.com

Abstract. *This paper presents the development of an intelligent system for monitoring and operational supervision of Diesel-Generator Units (DGU). This system is based on machine learning algorithms, focused on the preventive and predictive maintenance of the DGU. In this context, such technologies are useful tools for learning, detecting, diagnosing, and classifying machine operation patterns, acting assertively in assisted supervision and decision making, both at the operational level and the managerial levels. Diesel generators are an alternative source of electrical energy used to diversify the Brazilian electrical grid, aiming to mitigate events of unavailability of electricity generation and supply throughout the national territory. A Diesel Generator Unit becomes a system where monitoring is complex, with an extensive range of mechanical and electrical parameters to be correlated and supervised, to ensure the lifespan of the equipment and safety operation. Furthermore, a failure or imprecision of supervision and operation can generate impacts on production that could affect all layers involved in the process of generation and consumption of energy, and also cause injuries to operators. The intelligent system developed in this research applies machine learning techniques to classify operational patterns based on real data from more than sixty DGUs connected to the Brazilian electrical grid. To enable the learning of standards, unsupervised learning techniques were investigated. As demonstrated in this article, the developed system proved to be assertive in detecting operating incidents by supporting the operating team in making decisions and adjusting critical parameters of the system.*

Keywords: Diesel-generator unit, intelligent system, machine learning, operational patterns, maintenance.

1. INTRODUCTION

Since 2001, electricity generation and transmission companies in Brazil have faced several challenges, such as lack of investments in infrastructure, expansion of the electricity grid, prolonged drought, little diversification of primary sources, among others. This situation was even more evident in the years 2009 and 2012 when problems in equipment in large generating parks left several Brazilian states without electricity for hours (TCU, 2013).

Recently, since September 2020, when started the most recent hydrological period in Brazil, the National Electrical System Operator (ONS) has been reaching successive historical minimum levels records of rainfalls over the reservoirs of the main hydroelectric plants in the country (ANEEL, 2021). The rainfall historical series is recorded since 1931.

After these events, the government and energy companies have sought to invest in several fronts to mitigate these faults, particularly the diversification of primary sources of electricity. In fact, the electric energy matrix in Brazil is concentrated in hydroelectric plants, with more than 66% of electric generation (EPE, 2019). These plants are high affected by prolonged droughts, a problem that is beyond the reach of companies and the government. These weather events cause a reduction in the level of reservoirs that, in some cases, reach critical levels, making the operation of such hydroelectric plants unfeasible for a considerable period of time.

Allied to the climate issue, there is a growing increase in demand for electricity in Brazil (Oliveira, 2019). According

to EPE (2017), it is estimated that the increase in energy demand between 2016 and 2021 will be 3.2% and between 2021 and 2026 of 3.9%.

An alternative for such diversification is the use of Diesel Generating Units (DGU). Electric generation plants based on DGU can benefit from the country's wide offer of Diesel oil and/or bio-Diesel, require low installation time, have good logistical flexibility, as well as can be installed in a wide variety of locations places, in general, adjacent to transmission lines.

Despite the installation conveniences, the operation and maintenance of such DGU plants presents a considerably number of unexplored technical-scientific challenges, which are of outstanding relevance for ensuring the operational safety of these units, as well as for ensuring the quality of energy generated. Among these challenges, the following stand out: operational monitoring, problem diagnosis and decision-making, all in an autonomous and integrated manner.

In fact, DGU can be divided into two coupled subsystems: a mechanical engine motor and a electromagnetic generator. The mechanical subsystem is primarily constituted by a Diesel combustion engine. The electromagnetic subsystem is formed by a synchronous electrical generator. Indeed, both subsystems have, in a isolated sense, a wide set of operational issues, that must be monitored to reach an efficient operation, a costs reduction and the maintenance of the equipments lifespan.

However, when considering the coupling subsystems, the dimension of the monitoring universe becomes much larger and more complex, mainly due to the inter correlation between phenomena. In fact, problems with fuel injection nozzles, oil filters, non-standard temperature, excessive vibrations, inconsistent power generation, frequency controllers faults, are some of the possible examples of situations that reduce lifespan, increase operation costs or even permanently damage the DGU (ABNT, 1997).

Thus, the identification of unwanted behavior patterns of the DGU, as well as the monitoring of signals in order to keep the operation at acceptable levels, are matters of great relevance. However, these issues may involve several factors combined, such as: intrinsic characteristics of the subsystems and/or knowledge of specialists in DGU.

The synthesis of this information in order to find response patterns in the DGU is an important step in increasing efficiency. From the mapping of patterns, it is possible, for example, to predict actions and/or prevent undesirable situations, that is, operational decisions can be made in relation to the DGU.

The design of an autonomous classifier to quantify the efficiency and guarantee the maintenance of the operational safety of the DGU constitutes an important contribution in innovation. From the classifier, it is possible to implement automatic decision making, such as changing the electrical power generated during the operation, signaling a possible problem, predicting maintenance actions or even performing the DGU shutdown, in critical cases.

The objective of this work is to propose an autonomous classifier to distinguish the DGU operation patterns, in order to identify potential abnormal operational states. From the previous knowledge of such operational abnormal patterns, it is possible to do predictive repair, automatic system interventions, or even manual intervention by the operating team, promoting prevention of faults, cost reduction, and increase of lifespan of the equipment.

In order to design the classifier, a set of operational information was collected from a group of sixty-two DGUs, in partnership with the company Brentech Energia SA, located in Aparecida de Goiânia, GO. The classifier is based on three process variables: the engine oil pressure, the average temperature of the cylinders, and the electric current flow. The database was collected on February 10, 2021, for 12 hours, with a sample rate of five seconds.

2. THE PROCESS

A DGU is a machine powered by internal combustion coupled to an electric generator. The machine uses Diesel fuel as the primary source of energy and it produces electrical energy. The energy produced is delivered to the national electrical system by the connection of the DGU plant to power line transmissions.

Some features of DGU operating can be sampled to verify the conditions of the machine and also to make the system work in adequate conditions. There are several parameters that can be used to monitor the DGU operation, such as, pressures, temperatures, flow rates, voltage, and electric current. Such parameters can be recorded by the DGU interface controller, which can be connected to a supervision system, by a CAN network. These data are saved in a database, by the supervision system, that also provides information to the operational team. The actual monitoring system architecture is presented in Figure 1.

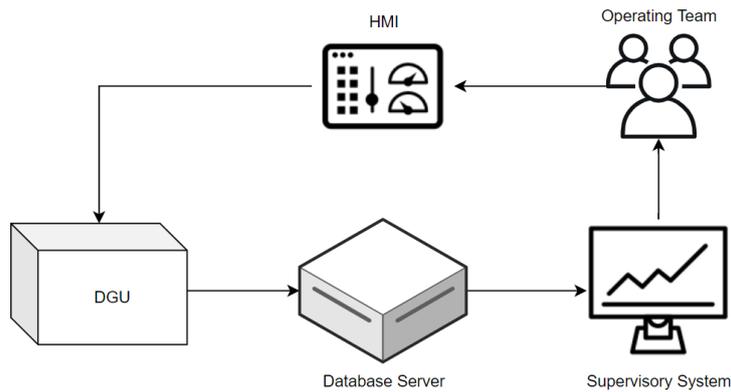


Figure 1. Actual monitoring system architecture.

The supervisory system makes a data treatment, aiming to prepare the data for better visualization by the user. The supervisory system is accessed directly by the operation team, which uses the information to analyze and control the machine operation. As can be seen, by Figure 1, the recognition of an abnormal DGU operation pattern is strongly dependent on the expertise of the operational team.

This work proposes a new architecture to the monitoring system operation, as presented in Figure 2.

In this new configuration, the data from the supervisory system is accessed directly by the operation team, however, the same data is used by the model to distinguish the conditions of the machine operation. The output of the model and the other features of the machine operation are shown to the operation team and they can be used to take action on the DGU operation. The main advantage of this configuration is the computational capacity for data processing, facilitating the analysis and the decision-making process of the operation team.

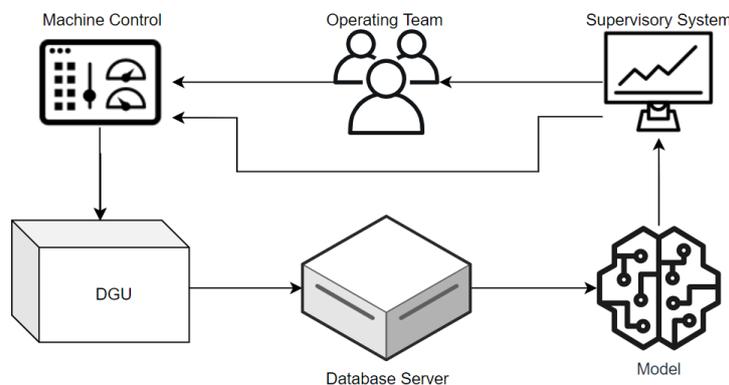


Figure 2. Proposed system architecture.

3. UNSUPERVISED LEARNING

The classifier proposed in this work is based on unsupervised learning theory. In such techniques, the models only receive input data from the system (Ghahramani, 2003; Brunton and Kutz, 2019). It means that there are neither target outputs for reference, nor environment rewards as feedback. Such models can find patterns in the data that can be used as in decision-making algorithms, predicting future states, etc. In general, with a properly adjusted model and a database of sufficient size, the algorithms can recognize patterns that would easily be mistaken as noise or non-relational data.

Additionally, unsupervised learning problems are focused on obtaining a probabilistic model of the data. Two classical questions in this context are clustering and dimensionality reduction (Ghahramani, 2003). Thus, this work aims to apply such techniques for clustering for abnormal patterns in the operation of the DGU.

3.1 Exploratory data analysis

As presented in Section 1, a real database, that includes data of sixty-two DGUs, was used in the work. The data was collected on February 10, 2021, for 12 hours, with a sample rate of five seconds. The database has, approximately, 535.000 samples.

The process of sampling through machine integrated instrumentation obtains features that can be classified by electric

nature or mechanical nature. Each sample contains twelve features of electric nature, such as electric power and electric current, and twenty-one features of mechanical nature, such as oil temperature and oil pressure.

An expressive size of the database is recommended to estimate an efficient model, however, in such cases, it is necessary to perform a treatment in the data set, in order to apply learning algorithms. This step is known as exploratory data analysis (Brunton and Kutz, 2019).

Some techniques can be used in order to optimize the training of a learning algorithm. Firstly, irrelevant information has to be removed from the data set, for instance, the existence of null and missing data. Then, it is necessary to verify outlier samples, which means, an observation that lies an abnormal distance from other values from a data population.

The final procedure in the exploratory data analysis is based on techniques of normalization and a combination of features. Database normalization is the process of structuring a data set in accordance with a common scale, without distorting differences in the ranges of values (Brunton and Kutz, 2019). The combination of features is also sometimes called feature optimization, the main goal of feature combination is to extract meaningful features from the primitive and original features.

3.2 K -means

One of the ways to classify data in an unsupervised manner is by the clusters method. For this, many associative, numerical, and statistical techniques can be applied. The results of these techniques are groups of data related one with others with hierarchical levels, *i.e.*, the data in the same group are more similar to each other than the data from the other groups. Among the clusters method, the K -means algorithm is one of the most popular and it is appropriated for a set of quantitative data (Hastie *et al.*, 2009). Therefore, this work uses the K -means algorithm.

The K -means method performs the grouping using the quadratic Euclidean distance as a dissimilarity measure. It receives as input the desired number k of clusters and, through an iterative process, minimizes the distance between the data. The method output is k centroids that represent the clusters (Vanem and Brandsæter, 2019).

One step to apply the K -means method is to define the number of centroids. Considering that there is no general optimal method, the work presented by Vanem and Brandsæter (2019) points out that for the definition of an adequate number of centroids, observe the ratio between the variance between the clusters and the total variance, applying the elbow rule, it is a viable strategy. In certain cases, knowledge of the application context and exploratory data analysis help in the decision process for this hyperparameter.

In this work, the number of clusters k is assumed to be fixed in k -means clustering and previously defined. Let the k centroids (w_1, \dots, w_k) be initialized to one of the n input data patterns (i_1, \dots, i_n) , *i.e.*,

$$w_j = i_j, j \in 1, \dots, k, l \in 1, \dots, n$$

Being the data subset C_j the j^{th} cluster whose value is a disjoint subset of input data patterns, the quality of the clustering is determined by the following error function:

$$E = \sum_{j=1}^k \sum_{i_l \in C_j} |i_l - w_j|^2$$

The k -means algorithm is presented in the Algorithm 1.

Algorithm 1 K -Means Algorithm

Require: Initialize the k centroids (w_1, \dots, w_k) such that $w_j = i_j, j \in 1, \dots, k, l \in 1, \dots, n$

repeat

for each input i_l from the data set, where $l \in 1, \dots, n$ **do**

 | Assign i_l to the cluster C_j with nearest centroid w_j , considering the Euclidean distance.

end

for each cluster C_j , where $j \in 1, \dots, k$ **do**

 | Update the centroid w_j to be the center of all currently samples in C_j

end

 Compute the Error: $E = \sum_{j=1}^k \sum_{i_l \in C_j} |i_l - w_j|^2$

until

E does not change significantly or cluster members no longer changes for some iterations.

▷;

The K -means algorithm has a number of iterations depending on the size of the dataset, the number of the clusters, and the input data distribution, *i.e.*, the necessary time to calculate the patterns may vary with these parameters. In the worst case for the Algorithm 1, the time required to the first for loop is $O(knd)$, the time required to update the centroids in the second for loop is $O(nd)$ and the time required for calculating the error E is $O(nd)$. (Alsabti *et al.*, 1997)

4. STANDARDS CLASSIFIER

The first step of this work was to identify the operating patterns of the DGUs. The graphical analysis provided by the equipment operating team was considered. For the present context, four operating states are considered, where the temporal behavior of the generator oil pressure is sufficient to evidence these states.

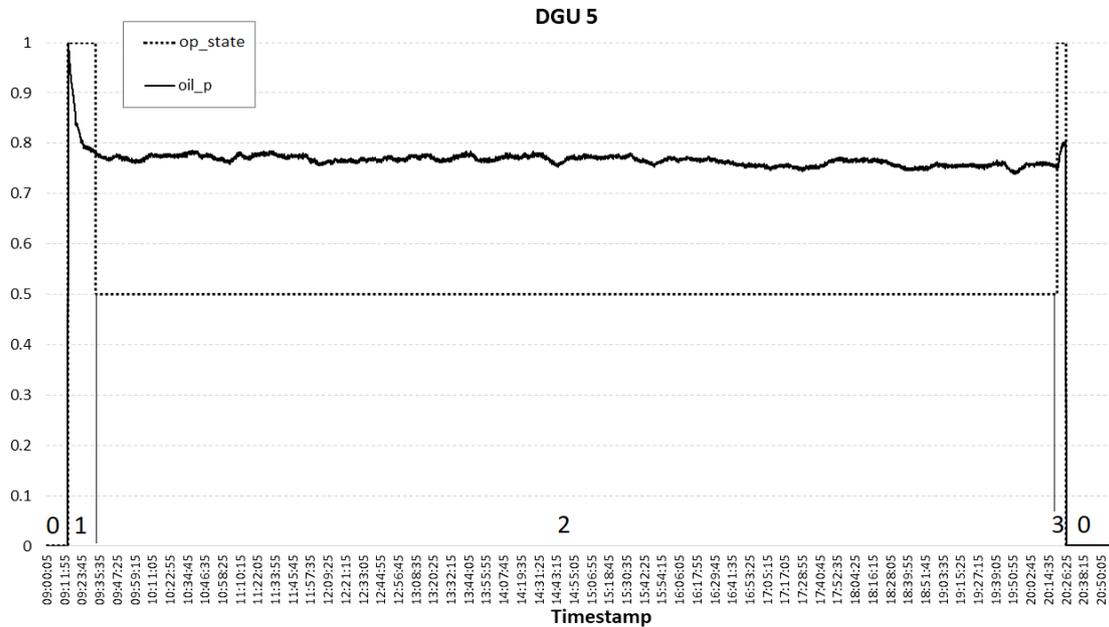


Figure 3. Operation states of the DGU.

Figure 3 presents a graph of the engine oil pressure, from normalized data, monitored over the 12-hour operating period on February 10, 2021. Numbered demarcations were inserted, representing the conditions of the operating state reported by the operating team. The state 0, 1, 2, and 3 represent, respectively, the shutdown state, setup state, operation state, and cooling state.

This work uses the K -means learning model aiming to detect automatically the operating states. The chosen input features in the method are the engine oil pressure and the average temperature of the engine cylinders. The number of clusters used in the method will be defined considering the number of machine states previously defined by graphical analysis, *i.e.*, each cluster will represent a state of operation.

4.1 Estimating the quantity of clusters

From the states indicated by the operation team and the analysis of results, it was defined the use of three clusters to configure the model initially. They represent the setup or cooling state, the operation state, and the shutdown state. The choice of these three groups is justified by the fact that the setup state and the cooling state present similar data, this similarity is readily identified by the model.

Figure 4 presents the model inputs, engine oil pressure (oil_p), average cylinder temperature (cylinder_avg) from the data set. This Figure also presents the model output, dotted line (op_mod). The output data is the operating mode of the electric generator classified by the algorithm. Three discrete values were assigned to the output, with zero being the shutdown state, one representing the operating under load mode, and finally two being the grouping of the setup and cooling modes.

Therefore, the developed model could identify the operating mode of the electric generator in all data from machines.

4.2 Defining patterns: normal and abnormal

The generated model was applied to the data of all DGUs. Analyzing the results from the K -means model, it is possible to see that, as scored by the responsible team, the standard operating characteristic of the equipment is given by the conventional sequence of shutdown state, setup state, operation under load state, cooling state, and shutdown state again. This pattern was identified in forty of the sixty-two machines analyzed, approximately sixty-five percent.

However, patterns with distinct characteristics were also found, where the alternation between the classified states does not follow the expected flow. One of the anomalous patterns found is represented in Figure 5. In this figure it is easy to verify that the machine began in the shutdown state, then it goes to the setup state and the operation state, respectively.

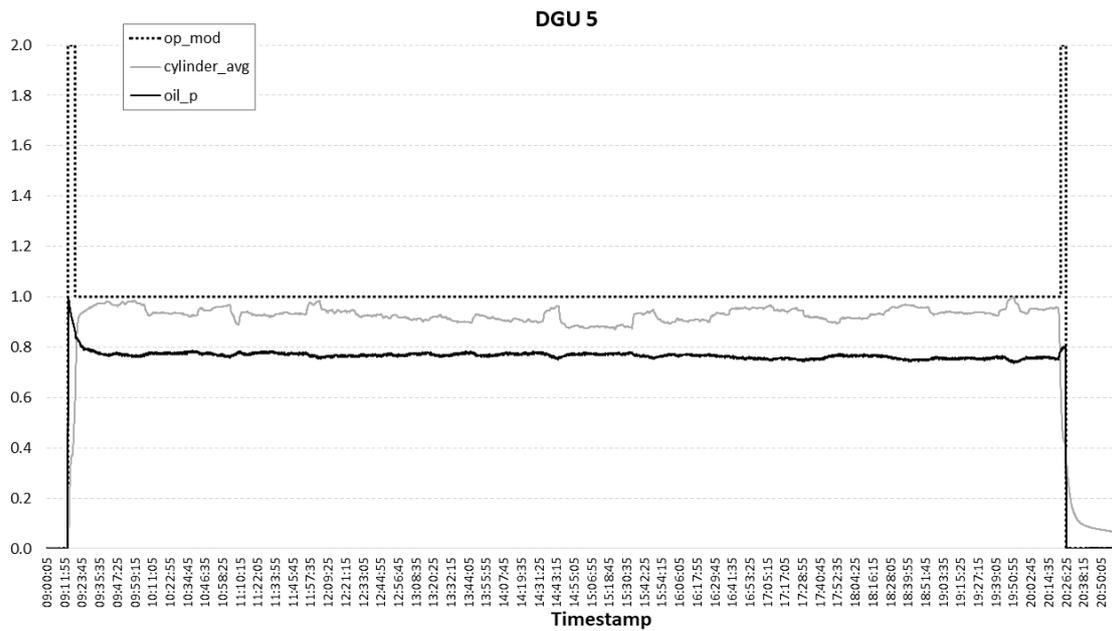


Figure 4. Classified operation states of the DGU.

Thus, during the operation state, the equipment turned off unexpectedly.

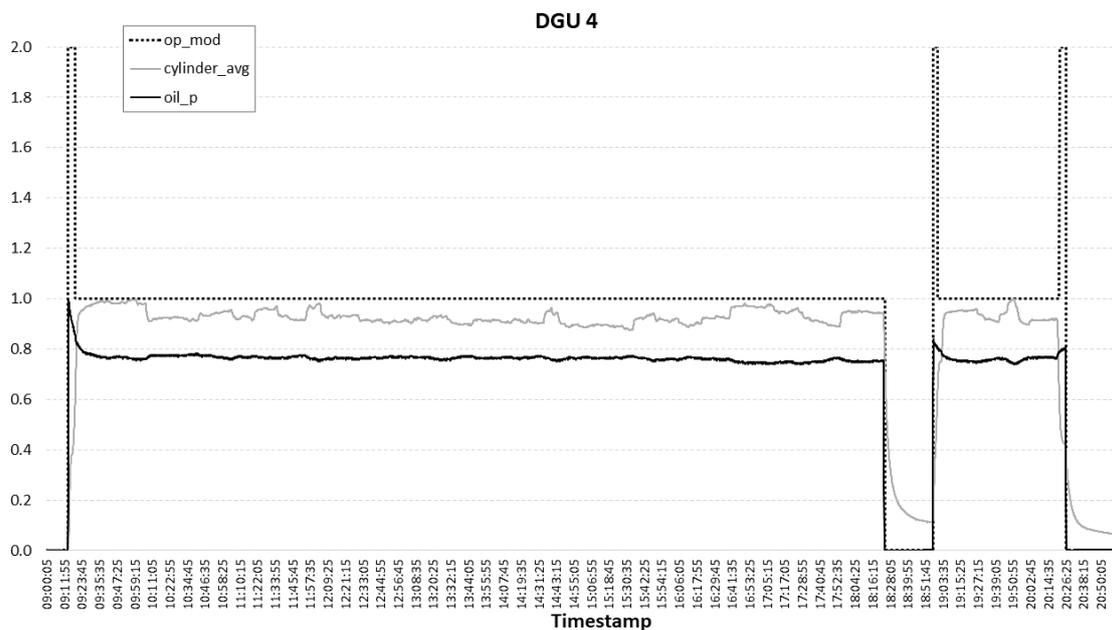


Figure 5. Anomaly detected by the model.

This behavior was signaled, after analysis, by the team responsible for the equipment as a sudden shutdown, where there is no proper cooling of the machine. This behavior can be dangerous to the DGU. Overheating, Diesel engine carbonization, presence of air in the supply system are some damages that can occur in DGUs by the unexpected shutdown. This anomalous behavior was identified in approximately twenty DGUs among the total machines analyzed.

Thus, the model of the *K*-means algorithm could also correctly identify the behavior pattern of all DGU analyzed.

4.3 Classifying the behavior pattern of the DGU

Based on the previous section, the *K*-means algorithm and the model developed can be used for classifying the behavior pattern of the DGU. It can identify DGUs with anomalous behavior and inform the technical operating team which machines present functionality imperfections.

As also stated before, the healthy behavior pattern of a DGU, on the analysis of the model, is a "0-2-1" (op_mod)

sequence, being the shutdown, setup state, and operation under load states, followed by a "1-2-0" sequence, being the operation under load, cooling, and shutdown states.

Thus, the first two behavior patterns can be defined as the normal startup being the op_mod "0-2-1" sequence and the normal shutdown being the op_mod "1-2-0" sequence.

Nonetheless, as presented in Figure 5 and in Figure 6 as well, the third and non-healthy pattern can be described by the op_mod "1-0" sequence. This pattern can be described as an emergency shutdown, where, under abnormal conditions, either the control module or the operator detects a dangerous situation and suddenly turn the machine off. Many conditions can lead to this not usual and undesired scenario, such as high fuel or coolant temperature, over current, low lubricating oil pressure, etc.

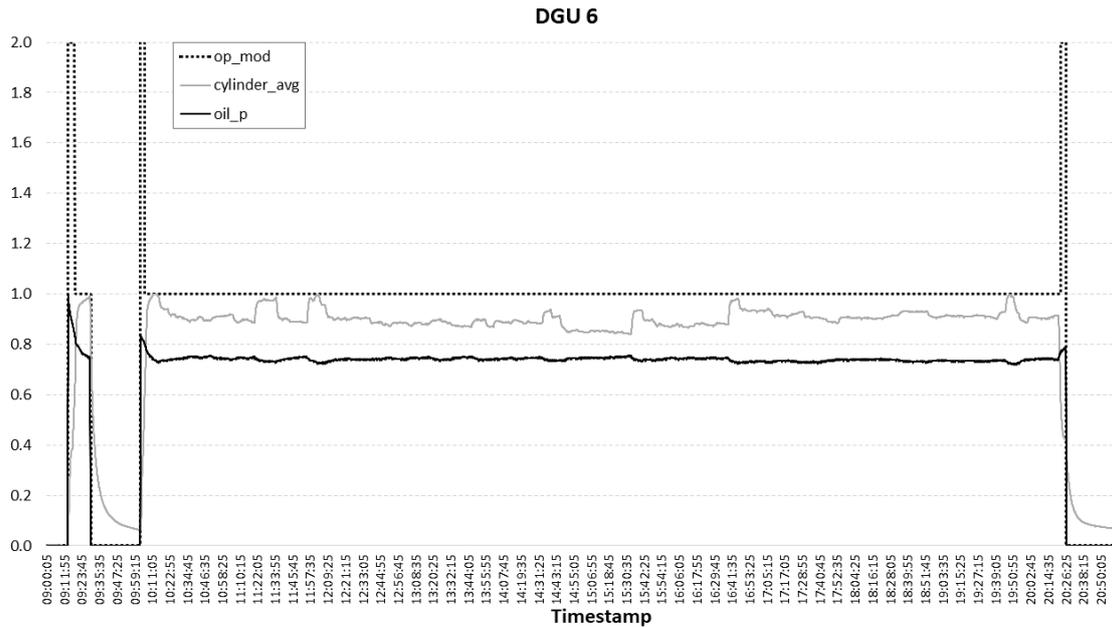


Figure 6. Emergency shutdown state.

The fourth pattern is represented by the op_mod "2-1-2" sequence and indicates the normal operation of a DGU. In this pattern, both the setup and cooling states are present and during the op_mod "1" the machine is in proper production. This pattern can be observed in Figure 4.

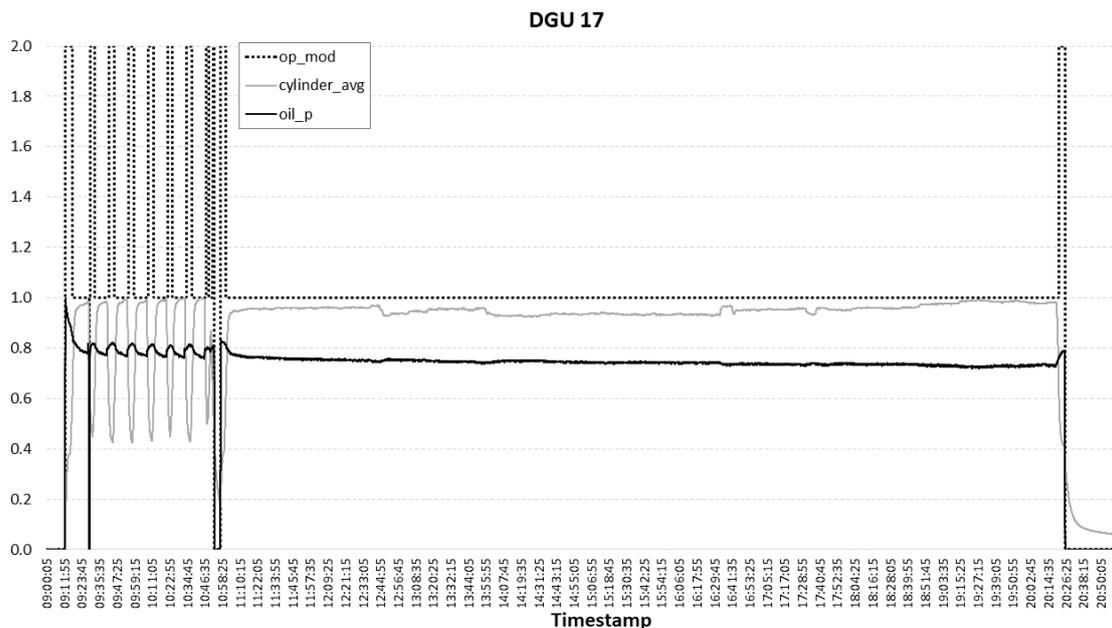


Figure 7. No-load operation pattern.

The fifth and also undesired scenario is established by the op_mod "1-2-1" sequence. According to the responsible

team of the DGUs, this pattern can be defined as a no-load operation. To dig into this conclusion, even though it was not applied as an input to the model, the engine current was taken into account, in order to assert that the machine was not turned off.

The no-load operation pattern can be observed several times in Figure 7. This pattern implies that the equipment is operating but it is not generating energy to the system.

This state can be manually activated by an operator or automatically set by the control module. In some cases, it is necessary both for heating or cooling the machine, although, as alerted by the operating team, more than ten minutes in this state leads to a problematic scenario with consequences such as high fuel consumption and equipment wear.

Another possible cause of this state is the loss of synchronism with the energy network, whether it being a voltage or frequency issue. In this case, the control module switches off the circuit breaker, and with the machine on tries to restore the synchronous condition.

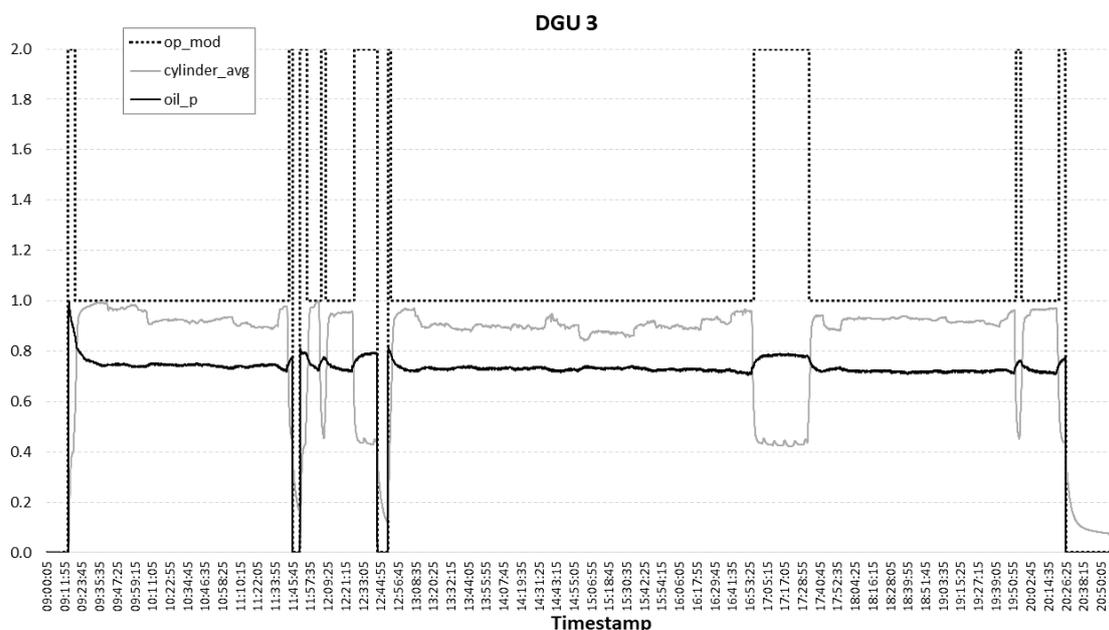


Figure 8. No-load operation undesired scenario.

Opposite from the machine of Figure 7, which presents short periods of no-load patterns, the DGU of Figure 8 is an evidence of a very unwelcome scenario. During hours 16 and 17, the machine stayed a long period of time operating under no load. Sometimes, some sort of maintenance is required for a proper machine recovery, thus rapid recognition of this unhealthy state is decisive for preventing failures and major wastage.

In this matter, such automated intelligent systems, as the one presented in this work, are handy tools for supporting the operation team on the supervision and maintenance of machines.

5. CONCLUSION

This paper has presented the development and results of a model to analyze and identify the operating states of Diesel Generating Units.

The model could successfully identify three operation states and, using these classifications, recognize five operation patterns, being three normal and two abnormal standards. This solution can assist the operating team in decision-making on predictive and corrective maintenance.

The results are relevant and justify the implementation of the proposed system architecture, in which the model sends information to a supervisory system, and this system is used to assist the operating team and supply the machine control with relevant data.

Thus, the decision-making process within the electric generation plant can be efficiently optimized. Reducing costs by efficient operation, decrease in unexpected stops, and increase of the lifespan of the machines are the main gains when applying such technologies.

6. REFERENCES

ABNT, 1997. *NBR 5477: NB 130: Presentation of the performance of internal combustion, reciprocating and compression-ignition engines. (in Portuguese).*

- Alsabti, K., Ranka, S. and Singh, V., 1997. “An efficient k-means clustering algorithm”. *Electrical Engeneering and Computer Science*, Vol. 43.
- ANEEL, 2021. “ANEEL define que bandeira tarifária de julho custará R\$9,492 a cada 100 kWh”. Disponível em: <https://bit.ly/3h3PDtN>. Disponível em: <https://bit.ly/3h3PDtN>.
- Brunton, S.L. and Kutz, J.N., 2019. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press. ISBN 978-1108422093.
- EPE, 2017. “Technical note DEA 001/17 - projection of electrical energy for the next 10 years (2017-2026) (in portuguese).” https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-245/topico-261/DEA%20001_2017%20-%20Projeções%20da%20Demanda%20de%20Energia%20Elétrica%202017-2026_VF%5B1%5D.pdf. Accessed 12 January 2021.
- EPE, E.D.P.E., 2019. “National energy balance: Summary report, base year 2018”. *Rio de Janeiro: Ministério de Minas e Energia*.
- Ghahramani, Z., 2003. “Unsupervised learning”. In *Summer School on Machine Learning*. Springer, pp. 72–112.
- Hastie, T., Tibshirani, R. and Friedman, J., 2009. “Unsupervised learning”. In *The elements of statistical learning*, Springer, pp. 485–585.
- Oliveira, N., 2019. “Energy consumption closes 2018 with increase of 1.1% (in portuguese).” <https://agenciabrasil.ebc.com.br/economia/noticia/2019-01/consumo-de-energia-fecha-2018-com-aumento-de-11>. Accessed 17 June 2021.
- TCU, 2013. “Survey with MME, ANEEL and ONS about regulatory, inspection and coercive measures adopted by the responsible agents of the sector to mitigate the occurrence of blackouts and their consequences. (in portuguese)”. <https://portal.tcu.gov.br/biblioteca-digital/levantamento-junto-ao-mme-aneel-e-ons-acerca-das-medidas-regulatorias-de-fiscalizacao-e-coercitivas-adotadas-pelos-agentes-responsaveis-do-setor-para-mitigar-a-ocorrencia-de-apagoes-e-suas-consequencias.htm>. Accessed 17 June 2021.
- Vanem, E. and Brandsæter, A., 2019. “Unsupervised anomaly detection based on clustering methods and sensor data on a marine diesel engine”. *Journal of Marine Engineering & Technology*, pp. 1–18.

7. RESPONSIBILITY NOTICE

The following text, properly adapted to the number of authors, must be included in the last section of the paper:
The author(s) is (are) solely responsible for the printed material included in this paper.