

Fault Detection using Principal Component Analysis and Mean Value Modeling in a 2 MW gas engine

Detección de fallas usando Análisis de Componentes Principales y Modelado de Valor Medio en un motor a gas natural de 2 MW

Yulineth Cardenas-Escorcia¹, Guillermo Valencia-Ochoa^{2*}, Juan Campos-Avella³

¹Magister en Gestión Energética, ydcardenas@mail.uniatlantico.edu.co, Orcid: 0000-0002-9841-701X, Universidad del Atlántico, Barranquilla, Colombia.

²Doctor en Ingeniería, guillermoevalencia@mail.uniatlantico.edu.co, Orcid: 0000-0001-5437-1964, Universidad del Atlántico, Barranquilla, Colombia.

³Doctor en Ciencias Técnicas, jcampos@e2energiaeficiente.com, Orcid: 0000-0002-9387-5302, Universidad del Atlántico, Barranquilla, Colombia.

How to cite: Y. Cardenas-Escorcia, G. Valencia-Ochoa, J. Campos-Avella “Fault Detection using Principal Component Analysis and Mean Value Modeling in a 2 MW gas engine”. *Respuestas*, vol. 25, no. 1, pp. 15-24, 2020.

Received on September 3, 2019; Approved on December 10, 2019

ABSTRACT

Keywords:

Fault detection;
Principal component analysis;
Mean value model;
Natural gas engine;
Thermodynamic model.

This paper describes the combination of statistical techniques and mathematical modeling in order to develop a fault detection system in a 2 MW natural gas engine under actual operation conditions. The Mixing chamber, turbochargers, intake and exhaust manifolds, cylinders, throttle and bypass valves, and the electric generator, which are the main components of the gas engine, were studied under a mean value engine to complement the statistical analysis. Objective: The main objective of this paper is to integrate two approaches in order to relate the faults with the changes of mean thermodynamic values of the system, helping to sustain the engine in optimal operating conditions in terms of reliability. The Principal Component Analysis (PCA), a multivariate statistical fault detection technique, was used to analyze the historical data from the gas engine to detect abnormal operation conditions, by means of statistical measures such as Square Prediction Error (SPE) and T2. These abnormal operation conditions are categorized using cluster techniques and contributions plots, to later examine its causes with the support of the results of a mean value mathematical model proposed for the system. The integration of the proposed methods allowed successfully identify which component or components of the engine might be malfunctioning. Once combined, these two methods were able to accurately predict and identify faults as well as shut downs of the gas engine during a month of operation. Statistical analysis was used to detect faults on a 2 MW industrial gas engine, also the result were compared with a mean value model in order to detect variations of the thermodynamic properties of the system at abnormal conditions.

*Corresponding author.

E-mail Address: guillermoevalencia@mail.uniatlantico.edu.co (Guillermo Valencia-Ochoa)

Peer review is the responsibility of the Universidad Francisco de Paula Santander.



This is an article under the license CC BY-NC-ND 4.0

RESUMEN**Palabras clave:**

Detección de fallas;
Análisis de componentes principales;
Modelado de valor medio;
Motor a gas natural,
Modelo termodinámico.

Este artículo describe la combinación de técnicas estadísticas y modelos matemáticos para desarrollar un sistema de detección de fallas en un motor de gas natural de 2 MW bajo condiciones reales de operación. La cámara de mezcla, los turbocompresores, los múltiples de admisión y escape, los cilindros, las válvulas throttle, bypass y el generador eléctrico, que son los principales componentes del motor de gas, fueron estudiados bajo un motor de valor medio para complementar el análisis estadístico. El objetivo principal de este trabajo es integrar dos enfoques para relacionar los fallos con los cambios de los valores medios termodinámicos del sistema, ayudando a mantener el motor en condiciones óptimas de funcionamiento en términos de confiabilidad. El Análisis de Componentes Principales (PCA), una técnica estadística multivariante de detección de fallas, se utilizó para analizar los datos históricos del motor a gas para detectar condiciones anormales de operación, por medio de indicadores estadísticos tales como el Error de Predicción Cuadrado (SPE) y T2. Estas condiciones anormales de operación se categorizan mediante técnicas de clústeres y diagramas de contribuciones, para luego examinar sus causas con el apoyo de los resultados de un modelo matemático de valor medio propuesto para el sistema. La integración de los métodos propuestos permitió identificar satisfactoriamente qué componente o componentes del motor podrían estar funcionando mal. Una vez combinados, estos dos métodos fueron capaces de predecir e identificar con precisión las fallas, así como las paradas del motor de gas durante un mes de funcionamiento. Se utilizó el análisis estadístico para detectar fallas en un motor de gas industrial de 2 MW, también se comparó el resultado con un modelo de valor medio para detectar variaciones de las propiedades termodinámicas del sistema en condiciones anormales.

Introduction

With the continuous technological progress and growth of society, the demand for energy has increased, more specifically the consumption of electrical energy; which has motivated the development of different renewable energy sources such as wind and solar energy, among others [1], [2]. In these processes, due to their complexity, they are increasingly exposed to a malfunction of some component of these processes, which could cause danger to operators, economic losses, among other external factors, so it must be ensured that these security systems must have excellent reliability.

To ensure that the operation of any process is correct, there must be a set of actions that must comply with three fundamental stages, failure detection, diagnosis and restoration of operating conditions as specified by the process, ie process monitoring and is applicable to any application [3], [4]. The industry has adopted methods and procedures that have allowed to automatically detect failures in generation and electric motors, extending their life cycle, improving their safety and providing financial savings [5], [6], these methods are based on the identification of highly probabilistic parameters, calculation of equations, estimation of state variables and multivalent statistical methods that help to determine

the causes of failure [7], [8] and other classical methods used is the rule of diffuse equations or neural network approaches [9], [10]. In general, the traditional methods of failure detection and diagnosis are based on dynamic and mathematical models [11], [12], having important aspects in failure detection as the expertise of the system operator and knowing the functions of the individual components and their respective connections, qualitative and/or quantitative in the process of nominal operation against abnormal operation, in relation to the presence of an abnormal event or called failure [13] in a machine, for example an engine, can influence its operation, efficiency and even the interruption of the main functions in the process [14].

Different methodologies [15], [16] have been developed to standardize the failure detection process, aiming at a repetitive cycle of Failure Detection-Identification-Diagnosis. This cycle corresponds to a global analysis of analysis and failure detection [17] in a system or process in order to standardize terms and procedures for failure diagnosis, which are usually based on models obtained with widely studied methodologies [18] - [20]. Some of these failure detection methods involve statistical methods for diagnosis, such as principal component analysis (PCA) [21], which is a powerful multivariate statistical method that has been successfully used to monitor and identify

abnormal dynamic operating conditions [22] - [24]. The above method is based on two statistical hypothesis tests: Hotelling's T2 and SPE, in order to determine when a failure occurs, but the causes are not known. Generally, contribution charts are used as a diagnostic tool in process monitoring [25]. Applications of the method have been made, where an algorithm of identification of the root cause of future failures based on PCA was proposed in order to help predictive maintenance [26], such as using the method for the reconstruction of variables between correlated sensors, proposing an indicator to determine the state of the sensors [27]. Studies have also been carried out to improve the failure detection technique by combining the PCA method with other methods [28]. The integration of PCA with experts has also given very good results [29], [30].

The main contribution of the research is the application of the principal component analysis (PCA) failure detection method to a generation engine, with the help of mathematical equations and a phenomenologically based semi-physical model, using the MATLAB simulation program to do so.

Materials and methods

PCA method

Consider a data set consisting of observation variables and observations for each of the variables, which is arranged in a matrix arrangement $X_0 \in \mathbb{R}^{n \times m}$ as shown below:

$$X_0 = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (1)$$

The X_0 matrix must be normalized or standardized in such a way as to allow analysis without a tendency to the units of each m . Normalizing consists in subtracting from each m or variable the corresponding mean by dividing it by its respective standard deviation if $d \in \mathbb{R}^m$ is defined as the "vector of means", a column vector containing the means of the variables, which can be calculated by means of equations 2 and 3.

$$X = \frac{x_{nm} - \mu}{\sigma} \quad (2)$$

$$d = \frac{1}{n} (X_0)^T I_n = [\bar{x}_1 \quad \bar{x}_2 \quad \dots \quad \bar{x}_m]^T, (I_n = [1 \ 1 \ \dots \ 1]^T \in \mathbb{R}^n) \quad (3)$$

And the standard deviations of the variables in a diagonal matrix as shown in equation 4.

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_m \end{bmatrix} \in \mathbb{R}^{m \times m} \quad (4)$$

Then data normalization can be carried out as follows:

$$X = (X_0 - I_n d^T) \Sigma^{-1} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \in \mathbb{R}^{n \times m} \quad (5)$$

The covariance matrix is now defined as follows:

$$S = \frac{1}{n-1} X^T X, S \in \mathbb{R}^{m \times m} \quad (6)$$

A spectral decomposition of S is then performed, so that

$$S = V \Lambda V^T \quad (7)$$

$$V = [v_1 \ v_2 \ \dots \ v_m], V \in \mathbb{R}^{m \times m} \quad (8)$$

where V corresponds to the goblet matrix $\Lambda \in \mathbb{R}^{m \times m}$ is the diagonal of the matrix containing the diagonal as recommended by Zwick and Velicer [31].

The projection $y = V^T x$ of an observation vector $x \in \mathbb{R}^m$ converts the observation space into a set of uncorrelated variables corresponding to the elements of y . The variance of the i -th y -element is equal to the i -th self value in the matrix Λ . Knowing that S is invertible (since it is a symmetrical matrix) and with the definition:

$$z = \Lambda^{-1/2} V^T x \quad (9)$$

the Hotelling statistic T^2 is given by the following conditions. Define a matrix $D \in \mathbb{R}^{m \times m}$, such that:

$$\Lambda = D^T D \quad (10)$$

this matrix coincides with the diagonal matrix corresponding to the decomposition into singular values of $\frac{1}{\sqrt{n-1}}X = UV^T$ X as shown in equation 11.

$$\frac{1}{\sqrt{n-1}}X = UTV^T \quad (11)$$

If D_a is now defined $D_a \in R^{m \times a}$ as the matrix containing the first rows and columns of D, then Hotelling's T^2 for the smallest space is shown in equation 12.

$$T^2 = X^T P D_a^{-2} P^T X \quad (12)$$

and the detection threshold δ_{T^2} for the T^2 statistic is defined by

$$\delta_{T^2} = \frac{\alpha(n-1)(n+1)}{n(n-\alpha)} * F_{\alpha}(a, n-a) \quad (13)$$

The observations corresponding to the (m-a) self values which can be monitored by the Q statistic also known as SPE (Squared Prediction Error) as shown in equation 14.

$$Q = r^T r \quad (14)$$

where r is calculated as shown in equation 15.

$$r = (I - PP^T)X \quad (15)$$

The detection threshold for the Q statistic is given by equation 16.

$$\delta_Q = \theta_1 \left[1 + \frac{\theta_2 h(h-1)}{\theta_1^2} + \frac{c_a \sqrt{2\theta h^2}}{\theta_1} \right]^{1/h} \quad (16)$$

where,

$$\theta_1 = \sum_{i=r+1}^n \lambda_i, \quad \theta_2 = \sum_{i=r+1}^n \lambda_i^2, \quad \theta_3 = \sum_{i=r+1}^n \lambda_i^3 \quad (17)$$

$$h = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \quad (18)$$

and c_a is the standard deviation for a given a of trust (1-a).

Engine description

Since the selected equipment is an engine, it must be defined as a set of process systems, where the operational behavior of the equipment can be simulated through mathematical equations, obtaining a phenomenologically based semi-physical model that can be used to perform experiments in simulations [32], [33]. With this, it is shown the process of a phenomenologically based semi-physical model of the air line of a 2 MW generation engine operating with natural gas, a compression ratio of 10.5, which has a cylinder capacity of 74.852 liters, has 12 V cylinders at 60°, the stroke length is 220 millimeters, the chamber diameter is 190 millimeters, the maximum torque it reaches is 60.66 kN*m, the maximum power it can reach is 1820 kW at a nominal speed of 1500 rpm. This engine has a three-phase electric generator at a frequency of 60 Hz, a power factor of 0.9% delivering a reactive power of 911 kvar, electric power 1975 kW and apparent power of 2177 kva, having an average voltage between lines of 13,264 V.

The beginning of the air line process starts in the mixer where, at a temperature between 30°C and 40°C, an atmospheric pressure of 1,007 bar and relative humidity of 75-85%, aspiration air and natural gas are mixed, shown in Figure 1. The above mentioned mixture is made at a line pressure between 1,152 and 1,211 bar and volumetric ratio between 110-140 L/s, with the purpose of an optimal flammable gas-air mixture and obtaining an air/fuel ratio between 1.4 and 1.8. The fuel composition is a methane with number 92.1 with Nitrogen (1.50), Carbon dioxide (0.16), with a calorific value of 2085.1 BTU/lbm.

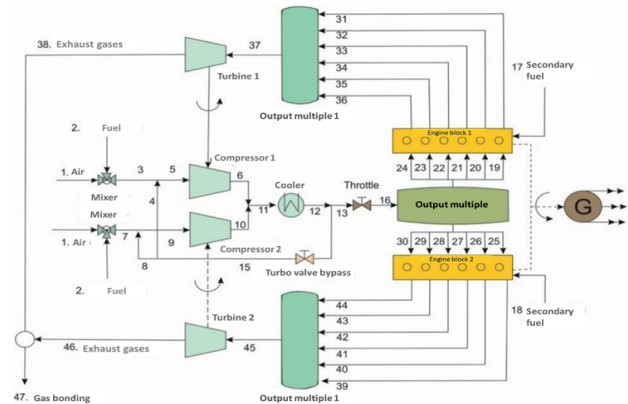


Figure 1. General flow chart for the engine generator system.

The gas-air mixture exits the mixer and then enters the compressor wheels belonging to the two turbo-compressors operating in parallel, sucking in the mixture, increasing the pressure by 3 to 5 bars and the temperature. After leaving the turbo-compressors, the gas-air mixture enters the engine suction chamber, passing through the mixture cooler, a mixture/water exchanger to reduce its temperature and then passing through the check valve to regulate the mixture's passage [34], according to its position, into the suction chamber of the crankcase, place in which there is a pre-chamber to make combustion and thus, to achieve homogeneity of the reaction, which is complemented with the combustion in case a new injection of a secondary fuel is required, generating exit gases in each cylinder, where the mixture is distributed to the 12 cylinders, with a mixture temperature between 60°C and 70°C and pressure load of 2.6 to 4.6 bar. With this, it is monitored and ready to send warnings when the mixture temperature is 71.1°C and failures are generated when it exceeds a temperature of 75°C.

The flow of the mixture can be regulated by the check valve, or throttle valve, and the turbo bypass valve, which recirculates the mixture leaving the mixer by regulating the engine power between 1000 kW to 1979 kW, each taking percentage openings depending on engine operating mode, which could be for the throttle valve 80% in mains independent operation or 98% in parallel operation with the mains and for the turbo bypass valve in turn, can take values between 15% to 50% independent of operating mode.

At the end of the process, the exhaust gases from the pre-combustion before the cylinders, which in each cylinder come out at temperatures between 580°C and 650°C and from the mixture that comes out in the rest of the cylinders, are mixed in a single outlet, having a concentration of gases that at the minimum load of 1000 kW with a compression ratio of 1.78 and maximum load of 1882 kW with a compression ratio of 1.97, it is understood that O₂ will have between 9.45% to 10.52%, 731-588 mg/m³ of CO, 461-468 mg/m³ of NO_x, 317-368 mg/m³ of NO₂ and 95 to 65 mg/m³ of NO.

Engine modeling

As part of the methodology, it is necessary to know the modeling of the engine made by a semi-physical model

of phenomenological basis, to evaluate by the method failure detection. Considering the objectives defined for the model, the mass and energy conservation balances have been applied to the study of the subsystems of interest in the engine, taking into account the following considerations:

The chamber and manifolds are considered as thermal energy reservoirs, with output properties equal to those of the control volume [35].

The two turbocharger systems and the cooler system are unified into a single system, which will be modeled as a single exhaust expansion stage and a single compression stage in the intake line.

One cylinder is studied and it is assumed that its behavior is the same for the others, where the exit properties of the gases will be equal to those of the control volume.

In the model proposed for the engine, each component is represented as a subsystem delimited by a fixed boundary as shown in Figure 2, where the mass and energy flows of each component are shown, where the mixing chamber, turbochargers, intake and exhaust manifolds, cylinders, throttle and bypass valves, and the electric generator stand out.

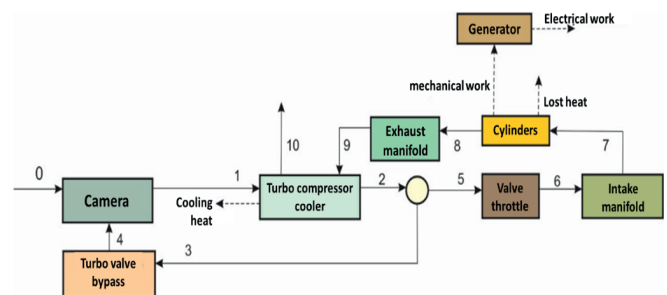


Figure 2. General flow chart for the engine generator system.

To find temperature 6, energy balances and replacements have been made suitable to express this in terms of temperature 1 (T_1) and the reference temperature (T_{ref}) expressed in equation (12), where η_{comp} is the compressor efficiency, $\lambda(t)$ AFR is the time dependent air-fuel ratio, $d_1=3.81451579330654$ 1, V_{gas0} is the volumetric flow rate of the gas, $d_2=49.9931450207536$, ρ_{aire0} air density, R_{gas0} is the particular constant of the gas at 0, and R_{aire0} is the particular constant of the air at 0.

$$T_6(t) = \frac{1-\varepsilon}{\eta_{comp}} \left[\eta_{comp} - 1 + \left(\frac{(\lambda(t)AFR+1)^2(d_1\dot{V}_{gas0}-d_2)}{(\lambda(t)AFR\rho_{aire0}+\rho_{gas0})(\lambda(t)AFRR_{aire0}+R_{gas0})T_1(t)} \right)^{\frac{k-1}{k}} \right] T_1(t) + \varepsilon T_{ref}(t) \quad (19)$$

where,

$$\varepsilon = a_1 T_{REF}(t) + a_2 (T_{ref}(t))^2 + a_3 \lambda(t) + a_4 (\lambda(t))^2 + a_5 (\lambda(t))^4 \quad (20)$$

and the constants are equal to

$$\begin{aligned} \alpha_1 &= 600.6 e_{exh} k_{ais} A_{ais} A_{amb} \\ \alpha_2 &= 28.6 k_{exh} k_{ais} A_{ais} A_{amb} \\ \alpha_3 &= 600.6 e_{ais} k_{exh} A_{exh} A_{amb} \\ \alpha_4 &= 21 k_{exh} k_{ais} A_{exh} A_{amb} \\ \alpha_5 &= 600.6 k_{exh} k_{ais} A_{exh} A_{amb} A_{ais} \end{aligned}$$

To find pressure 6, take the energy and mass balances and leave the values in terms of the flow densities 5 and 6 shown in equation 13, where ρ_{air6} is the air density at 6, ρ_{gas6} is the air density 6, ρ_{air5} is the air density at point 5, ρ_{gas5} is the gas density at point 5, A_5 is the area of step 5, A_6 is the area of step 6.

$$P_6(t) = \left(\frac{\lambda(t)AFR\rho_{aire6}+\rho_{gas6}}{\lambda(t)AFR+1} \right) \left(\frac{(d_1\dot{V}_{gas0}-d_2)(\lambda(t)AFR+1)}{(\lambda(t)AFR\rho_{aire5}+\rho_{gas5})} + \left(\frac{1-74.412 \exp(0.067\%AP_{TH})}{2} \right) \left(\frac{\rho_{gas0}\lambda(t)AFR(\lambda(t)AFR+1)}{A_5(\lambda(t)AFR\rho_{aire5}+\rho_{gas5})} \right)^2 (\dot{V}_{gas0}(t))^2 - \frac{(\lambda(t)AFR+1)^4 \rho_{gas0}^2 (\dot{V}_{gas0}(t))^2}{2A_6^2 (\lambda(t)AFR\rho_{aire6}+\rho_{gas6})^2} \right) \quad (21)$$

To know the values of the electrical power during time, we used the replacement that was made in equation 12 and found the electrical power in terms of temperature 1 (T_1) and pressure 6 (P_6), where $d_3=7.09033537747266$, $d_4=414.758117465621$ and $d_5=252.199816376828$, as shown in equation 22.

$$\begin{aligned} \dot{W}_{EL,EC}(t) &= d_3(1-\varepsilon) \left[1 - \frac{1}{\eta_{comp}} \right. \\ &+ \frac{1}{\eta_{comp}} \left(\frac{(\lambda(t)AFR+1)^2(d_1\dot{V}_{gas0}-d_2)}{(\lambda(t)AFR\rho_{aire0}+\rho_{gas0})(\lambda(t)AFRR_{aire0}+R_{gas0})T_1(t)} \right)^{\frac{k-1}{k}} \left. T_1(t) + \varepsilon d_3 T_{ref}(t) \right] \\ &+ \frac{d_4}{100} P_6(t) - d_5 \lambda(t) \end{aligned} \quad (22)$$

The volumetric efficiency was found by taking into account the pressure 7 (P_7) and temperature 7 (T_7) that vary with time, using the motor RPM, shown in equation 23.

$$\eta_V(t) = 0.24354 - 0.001371P_7(t) + 0.00035484T_7(t) + \frac{0.00013084RPM(t)}{60} + 5.1829\dot{V}_{gas0}(t) \quad (23)$$

Results and Discussion

More than 80,000 samples were taken for analysis, but the focus will be on the first thousand data, where, taking a minimum value where the engine performance is considered normal from the T2 statistics, an abnormality is presented from the 600 samples, where they exceed the safe limit and abnormality alarms are generated. When passing 1000 samples, a turning point is presented, where it goes down to the limit and takes T2 values of 100 and from 3000 to 6000 samples present constant abnormalities reaching maximum values of 150 and between 7000 and 8000 samples present an abnormality of more than 200, presenting the highest abnormality in this sample. With these values it is not really shown if there is a failure, breakdown or malfunction of the equipment, but it is shown that it is not working at its normal values. In the Q statistics, it shows the traces of the abnormalities presented in the graph of the T2 statistics, taking values above the Q limit, presenting abnormalities in the 1200 sample, and constantly manifesting between the 2000 and 3000 sample, which is repeated in the 5000 and 6000 sample and shows its highest peak between the 7000 and 8000 samples, demonstrating the highest abnormality, as shown in Figure 3.

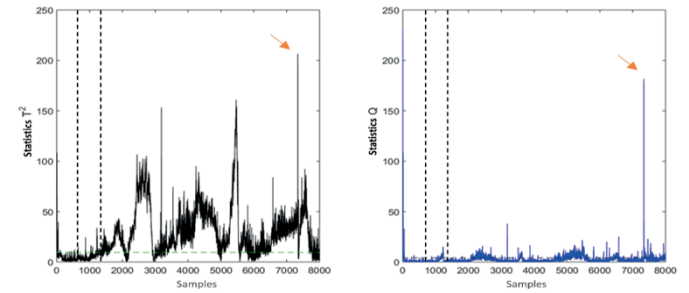


Figure 3. Response of standardized statistics

The sampling shown in figure 4, shows the behavior of the variables, showing in the variable 13, referring to the gas pressure, the one that presents abnormalities from the first 10 samples and generating critical values from sample 400.

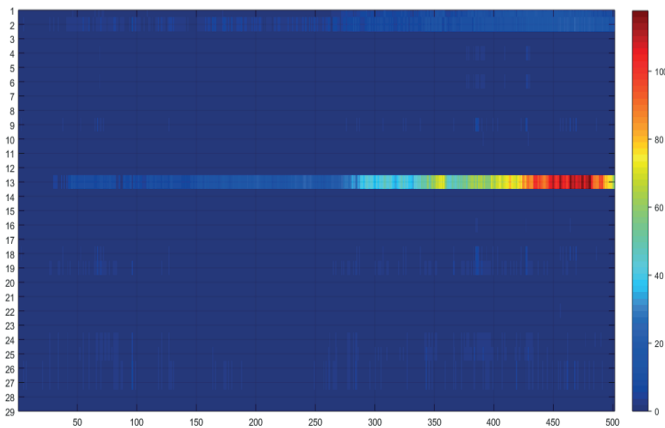


Figure 4. Behavior of variable 13

By graphing the eigenvalues with respect to the main components shown in Figure 5, it is shown that in the first 4 main components 95% of the dimensionless variability of each of the variables in their respective observation is represented. This means that, by analyzing these four main components, the real behavior of all the observations of all the variables is shown.

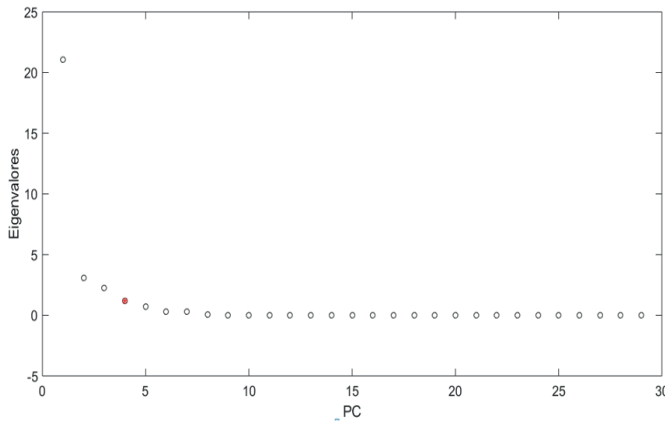


Figure 5. Behaviour of the eigenvalues of the main components

Taking the equations from the engine modeling, they set out to analyze the engine output variables such as temperature and pressure during the samples taken to perform them. Taking into account the abnormalities shown in Figure 3, there are abnormalities in the variables analyzed in Figure 6, temperature 6 and pressure 6 between samples 900 and 2600, where there are the raw abnormalities found. Temperature 6 presents a drop from 66.5 °C to 65.3 °C between samples 1000 and 1500, as well as the second abnormality between 1500 and 1700, where the temperature is reduced to a minimum of 65 °C. Pressure 6 is affected by the first abnormality, dropping drastically from 4.5 bar to 3.3 bar and reaching again

its value of 4.5 between samples 1200 and 1400, in the second abnormality where it drops drastically again to 3.4 bar and in the remaining samples it remains at values between 3.7 and 4.0 bar.

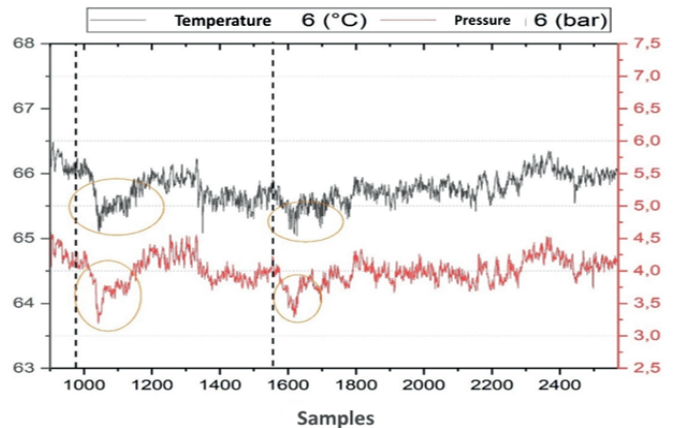


Figure 6. Motor output variable diagram.

In Figure 7, the 3 remaining output variables to be analyzed are electrical W, volumetric efficiency and effective efficiency. As shown in Figure 6, the analysis between samples 900 and 2600 is also shown, where the first abnormalities occur. The electric W, when the first abnormality occurs, goes down from 1820 KW to 1450 KW, where through the samples it is noticed that it recovers until the second abnormality where it goes down to 1460 KW and in the remaining samples it has a relatively stable behavior, having peaks between 1700 KW and 7160 KW. The effective efficiency in the first abnormality drops from a constant 0.393 to 0.385 and as the samples advance, there is no greater eventuality, with its lowest peak at 0.382. The volumetric efficiency in the first abnormality shows an increase from 0.618 to 0.623, where it remains intermittent between these values until the second abnormality where it shows an increase to its maximum peak of 0.625.

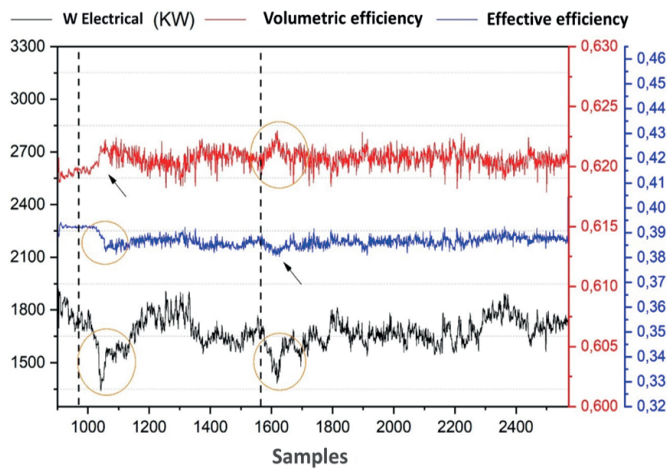


Figure 7. Remaining output variables

Conclusion

The statistical method of principal component analysis was used to detect failures in a generating engine with the help of dynamic modeling of the engine using a phenomenologically based semi-physical model to detect abnormalities in engine performance and display early alarms. Thanks to the combination of these two methods it was possible to detect failures and identify main variables where the process could have a malfunction and possible improvements to the components that perform the process. With the help of statistical analysis, it is possible to predict and accurately determine the failures, as well as the time in which the equipment should be stopped during the evaluated time. Further investigation is possible by performing better investigation of techniques such as clustering or discriminant analysis in order to show better failure diagnosis and classification when multiple failures occur over a period of time.

Financing

Scientific research article derived from the research project "Formulation of a model for diagnosis and failure detection (FDD) in spark plugs for natural gas generation engines", financed by "Universidad del Atlántico and the Kai energy efficiency management research group". Start year: 2019, end year: 2020.

Acknowledgements

The authors would like to thank the Vice-Rector of Research at the Universidad del Atlántico and the

company e2 Energía eficiente for the support received in this research.

References

- [1] G. Valencia-Ochoa, M. Vanegas-Chamorro and E. Villicaña-Ortiz, "Disponibilidad geográfica y temporal de la energía solar en la Costa Caribe colombiana", *1st ed: Universidad del Atlántico*, 2016.
- [2] G. Valencia-Ochoa, M. Vanegas-Chamorro and J. Polo-Jimenez, "Análisis estadístico de la velocidad y dirección del viento en la Costa Caribe Colombiana con énfasis en la Guajira", *1st ed. Universidad del Atlántico*, 2016.
- [3] L. Hurtado-Cortes, E. Villarreal-López and L. Villarreal-López, "Detección y diagnóstico de fallas mediante técnicas de inteligencia artificial, un estado del arte", *DYNA*, vol. 83, no. 199, 2016.
- [4] S. Qin, "Statistical process monitoring: basics and beyond," *Journal of Chemometrics*, vol. 17, no. 8-9, pp. 480-502, 2003.
- [5] R. Isermann, "Model-based fault-detection and diagnosis - Status and applications", *Annual Reviews in Control*, vol. 29, pp. 71-85, 2005.
- [6] A. Glowacz and Z. Glowacz, "Diagnosis of stator faults of the single-phase induction motor using acoustic signals", *Applied Acoustics*, vol. 117, pp. 20-27, 2017.
- [7] F. Kimmich, A. Schwarte and R. Isermann, "Fault detection for modern Diesel engines using signal- and process model-based methods", *Control Engineering Practice*, vol. 13, no. 2, pp. 189-203, 2005.
- [8] Y. Oguz and M. Dede, "Speed estimation of vector controlled squirrel cage asynchronous motor with artificial neural networks", *Energy*

Conversion and Management, vol. 52, no. 1, pp. 675-686, 2011.

- [9] J. Oliveira, K. Pontes, I. Sartori and M. Embiruçu, “Fault Detection and Diagnosis in dynamic systems using Weightless Neural Networks,” *Expert Systems with Applications*, vol. 84, pp. 200-219, 2017.
- [10] D. M. Himmelblau, “Use of Artificial Neural Networks to Monitor Faults and for Troubleshooting in the Process Industries”, *IFAC Proceedings volumes*, vol. 25, no. 4, pp. 201-206, 1992.
- [11] P. Frank, “Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results”, *Automatica*, vol. 26, no. 3, pp. 459-474, 1990.
- [12] M. Moore and L., Kramer, “Expert Systems in On-Line Process Control”, in *Third International Conference on Chemical Process Control*, Asilomar, pp. 839-867, 1986.
- [13] R. Isermann and P. Ballé, “Trends in the application of model-based fault detection and diagnosis of technical processes”, *Control Engineering Practice*, vol. 5, no. 5, pp. 709-719, 1997.
- [14] F. Farhani, A. Zaafour and A. Chaari, “Real time induction motor efficiency optimization”, *Journal of the Franklin Institute*, vol. 354, no. 8, pp. 3289-3304, 2017.
- [15] L. Desborough and T. Harris, “Performance assessment measures for univariate feedback control”, *The Canadian Journal of Chemical Engineering*, vol. 70, no. 6, pp. 1186-1197, 1992.
- [16] R. Dunia and S. Joe Qin, “Joint diagnosis of process and sensor faults using principal component analysis”, *Control Engineering Practice*, vol. 6, no. 4, pp. 457-469, 1998.
- [17] D. Godbole and R. Sengupta, “Tools for the design of fault management systems [automated highway systems]”, in *Proceedings of Conference on Intelligent Transportation Systems*, 1997, pp. 159-164.
- [18] H. Alvarez, R. Lamanna, P. Vega and S. Revollar, “Metodología para la Obtención de Modelos Semifísicos de Base Fenomenológica Aplicada a una Sulfitadora de Jugo de Caña de Azúcar,” *Revista Iberoamericana de Automática e Informática Industrial*, vol. 6, no. 3, pp. 10-20, 2009.
- [19] Á. A. Ruiz and H. Álvarez, “Escalamiento de Procesos Químicos y Bioquímicos basado en un Modelo Fenomenológico,” *Información tecnológica*, vol. 22, pp. 33-52, 2011.
- [20] C. A. Gómez, Y. A. Calderón and H. Álvarez, “Building phenomenological based semi-physical models: fermentation process case,” *Biotecnología en el Sector Agropecuario y Agroindustrial*, vol. 6, no. 2, pp. 28-39, 2008.
- [21] M. C. Verde, “Detección de Fallas usando Análisis de Componentes Principales”, *Congreso Anual la AMCA 2004*, vol. 1, pp. 431-436, 2004.
- [22] J. Kresta, J. Macgregor and T. Marlin, “Multivariate statistical monitoring of process operating performance”, *The Canadian Journal of Chemical Engineering*, vol. 69, no. 1, pp. 35-47, 1991.
- [23] J. F. MacGregor, “Multivariate statistical methods for monitoring large data sets from chemical processes”, in *Proceedings of AIChE Annual Meeting*, 1989.
- [24] B. M. Wise, “Adapting Multivariate Analysis for Monitoring and Modeling of Dynamic

- Processes.” *University of Washington*, 1991.
- [25] R. E. H. Miller and P. Swanson, “Contribution plots: a missing link in multivariate quality control,” *Applied Mathematics and Computer Science*, vol. 8, pp. 775, 1998.
- [26] D. R. Lewin, “Predictive maintenance using PCA”, *Control Engineering Practice*, vol. 3, no. 3, pp. 415-421, 1995.
- [27] R. Dunia, S. J. Qin, T. F. Edgar and T. J. McAvoy, “Identification of faulty sensors using principal component analysis”, *AIChE Journal*, vol. 42, no. 10, pp. 2797-2812, 1996.
- [28] S. Yoon and J. F. MacGregor, “Statistical and causal model-based approaches to fault detection and isolation”, *AIChE Journal*, vol. 46, no. 9, pp. 1813-1824, 2000.
- [29] A. Norvilas, A. Negiz, J. DeCicco and A. Çinar, “Intelligent process monitoring by interfacing knowledge-based systems and multivariate statistical monitoring”, *Journal of Process Control*, vol. 10, no. 4, pp. 341-350, 2000.
- [30] D. Leung and J. Romagnoli, “An integration mechanism for multivariate knowledge-based fault diagnosis”, *Journal of Process Control*, vol. 12, no. 1, pp. 15-26, 2002.
- [31] W. R. Zwick and W. Velicer, “Comparison of Five Rules of Determining the Number of Components to Retain,” *Psychological Bulletin*, vol. 99, pp. 432-442, 1986.
- [32] E. Hoyos, D. López and H. Alvarez, “A phenomenologically based material flow model for friction stir welding,” *Materials Design*, vol. 111, pp. 321-330, 2016.
- [33] J. Garcia Tirado, C. Zuluaga-Bedoya, V. Jaramillo and S. Jiménez Gómez, “Phenomenological-based Semi-Physical Model for a Pressure Control Plan”, 2017.
- [34] F. E. Moreno-García, J. J. Ramírez-Matheus, y O. D. Ortiz-Ramírez, “Sistema de supervisión y control para un banco experimental de refrigeración por compresión”, *Respuestas*, vol. 21, no. 1, pp. 97-107, 2016.
- [35] G. Valencia-Ochoa, C. Acevedo-Peñaloza y J.P. Rojas, “Thermoeconomic Modelling and Parametric Study of a Simple ORC for the Recovery of Waste Heat in a 2 MW Gas Engine under Different Working Fluids”, *Applied Sciences*, vol. 9, no. 21, pp. 4526, 2019.