Full title: Monitoring of fuel oil process of Marine Diesel Engine

Authors: David Boullosa-Falces [a,*], Juan Luis Larrabe Barrena [a], Alberto Lopez-Arraiza [a], Jaime Menendez[b], Miguel Angel Gomez Solaetxe [a].

[a] Department of Nautical Sciences and Marine Systems Engineering, University of the Basque Country UPV/EHU, Maria Diaz de Haro 68, 48920 Portugalete, Spain.

[b] Ibaizabal Tankers Shipping Company.

* **Corresponding author.** Tel.: +34 946014850. E-mail address: david.boullosa@ehu.eus (David Boullosa-Falces).

Highlights

- Selection fuel oil process variables of a Marine Diesel engine.
- Process monitoring through Small Sudden Deviation Method (SSDM).
- Detection of small and sudden deviation of the process.
- Identification of the variables that caused deviations in the process.

Abstract

Hotelling´s T^2 control chart is very efficient for detecting sudden changes in a process; however, it loses sensitivity to detect small and progressive changes and its performance decreases when the number of variables monitored at the same time is high. Because of this, conventional methods for variable reduction such as PCA were used, but they have difficulties in detecting the variability of the process when the correlation between variables is poor.

We propose a Method for detection of Small and Sudden Deviations in the process (SSDM), applicable when the correlation between variables is low; which is typical in marine propulsion processes.

First, fuel oil process variables of a marine diesel engine running, poorly correlated between them, were reduced through the analysis of correlations.

Afterwards, the selected variables were monitored through Hotelling´s T^2 control charts and sudden, out-of-range changes were detected. The variable that generated the deviation in the process was identified and the predictive variables were monitored through Cusum charts; the origin of small and progressive changes in the process below the alarm threshold set by the manufacturer was identified.

The proposed method (SSDM), based on the combination of (Hotelling´s T^2 + Cusum), can be implemented in any type of process in marine propulsion in a satisfactory and economical way, helping in the identification of the origin of any type of deviation (small and sudden) in the process early enough to implement the right predictive actions.

1. Introduction

Nowadays, in the industry, due to technological advances and complexity of the processes, there are many situations in which the monitoring of two or more variables is required [1].

Monitoring of redundant variables unnecessarily increases the costs of measurement [2] and hinders the interpretation by the user when there is a high number of signals to be monitored.

Within the monitoring techniques, we can distinguish between computational and statistical. Computational techniques such as Artificial Neural Networks (ANN) have been used in the monitoring of different industrial processes [3]; i.e., energy efficiency with improved fuel consumption reduction on a marine diesel engine.

In the statistical process control (SPC), the contributions have been made through the control chart of Shewart [4]. This methods monitors variables through independent control charts, ignoring the possible correlation or interaction between them, so when there is a variation in the process, several of these charts detect it at the same time, being complex to detect the exact cause of the failure and sometimes giving rise to false alarms.

In an article published in January 2017 [5], in a 2-stroke marine diesel engine, for some specific working conditions, were monitored a small group of seven variables, corresponding to the cylinder lubrication process through Hotelling´s T^2 control charts in combination with the technique Mason, Young and Tracy (MYT).

In that work, the Hotelling´s T^2 control chart, monitors in a multivariate way and effectively deviations are detected regarding the optimal working condition of the process; nevertheless, small and progressive change wasn´t detected due to the difficulties that these types of control chart have to detect these types of behaviours.

Furthermore, the variables that generated the deviation in the process were identified through MYT decomposition; it facilitated the diagnosis of the change in the process.

The problem comes when the size of the monitored variables begins to be moderately large, thus complicating the interpretation of the variable which caused the deviation in the process.

The MYT decomposition of the T^2 statistic has been shown to be a great aid in the interpretation of signaling T^2 values, but when the number of variables is greater than 10, and the cause of signaling is not clear from the unique terms of decomposition, the possible combinations among them are increased exponentially and hide which variable is responsible for it. [6].

Although this problem has been noted by other authors e.g. [7], it fortunately has led to the development of computer programs that can rapidly produce the significant components of the decomposition for moderately large sets of variables.

However, the question on how these computational methods will work when there are hundreds of thousands of variables has yet to be answered.

Therefore, the efficient selection and reduction of the variables to monitor is a way to optimize the process, maximizing the efficiency and reducing the costs of measurement [8].

There are many factors influencing the capability of this procedure, and these include computer capacity, computer speed, the size of the data set, and the programming of an algorithm.

There are different statistics techniques reducing the variables to monitor such as the Principal Components Analysis (PCA). This technique is capable of reducing the variables space, generating uncorrelated principal components (PCs) [9]; however, monitoring through PCA, [10], has difficulties in detecting the variability of the process when the correlation between variables is low, like in main propulsion engines related processes.

Yifei Wang, Xiandong Ma et al. [11], proposed an optimal sensor selection method based on principal components analysis (PCA) for condition monitoring of a distributed generation (DG) system oriented to wind turbines. The aim was to identify a set of variables from a huge amount of measurement data which could reduce the number of physical sensors installed for condition monitoring, while maintaining sufficient information to assess the system´s conditions.

The results showed that under a faulty condition, the algorithm of selection reduced the dataset dimension and kept the vital functions associated with the fault in the retained dataset with a high accuracy.

In the marine industry [12], the condition of the ship through satellite using PCA was monitored. The software developed for transmission using PCA reduces the amount of data sent via satellite, reducing time and cost of communications in case of transmission of all signals together.

Vinicius Barroso Soares et al. [13], implemented a system of alarm management through the use of different correlation methods (Principal Component Analysis, Correlation Analysis and Cluster Analysis), in three natural gas processing plants, getting to replace groups of alarms correlated by a more meaningful one, provided that the processes were linear.

Other techniques for reduction of variables such as Partial Least Squares (PLS) were presented; José Carlos Vega-Vilca et al [14], compared the technical Principal Components Analysis (PCA) and Partial Least Squares (PLS) on a database of 252 cases, 17 predictor variables and 1 dependent variable, with the aim of reducing the dimensionality.

To select the best regression model, they used the predictive residual sum of squares (PRESS), determining that the best model for PCA was with 6 components and the regression PLS was with 7 components; For reasons of comparison, both models were estimated with 6 components, being the values PRESS for each of them 88.31 and 266.54 respectively. These results showed that the PLS regression exceeded those of the PCA.

These techniques of variables reduction can be combined with Hotelling´s T^2 control charts to reduce the limitation that they have when the number of variables is high; S. Joe Qin [15], analyzed the use of Hotelling´s T^2 control charts together with PCA and other methods of detection, identification and diagnosis of failures; Joyce M. F. Fonseca et al. [9], proposed a methodology based on the combination of PCA and Hotelling´s T^2 control charts, capable of dealing processes with multiple set points and non-stationary. The proposed methodology was implemented in a thermoelectric power plant, monitoring in real time to detect any changes in the operation conditions of critical units of the power plant, boiler and turbine-generator unit.

Finally, the Hotelling´s T^2 control charts and Principal Components Analysis (PCA), for monitoring and control of a multivariate normal process were proposed in metal industry [16].

The Hotelling´s T^2 control charts detected when the process had deviations regarding the normal operation conditions, but didn´t identify the variables which were out of range or possible trends that might be in the process variables; but the control chart of PCA detected when the process was out of range and also showed the trend that made the process to be in that situation.

As mentioned above, another feature of Hotelling´s T^2 control charts, similar to Shewhart charts for univariate process control, is that they lose sensitivity to small changes, below 1.5 σ, and progressive in the vector of averages of the process [17]. Thus, Aparisi F and Garcia JC [18] established a zone of attention as a way to increase the power of the chart to this type of behaviour; in low-speed machines, failures may develop slowly and they stay latent till some critical point of their development interval when it is too late to act preventively [19].

In these cases, alternative procedures such as Cumulative Sum charts (Cusum) are widely recommended. This type of charts represent the cumulative sum of deviations, which contains information of all the previous samples [20]; in this issue, during the process of elaboration of a piece for the automotive industry, Shewhart and Cusum control charts were compared, for a same magnitude in the process changes. While the Cusum charts detected changes successfully, Shewhart charts were not able to detect them, indicating that the process was in control.

These Cusum charts [21], also have been used to detect possible defects in the downwind main bearing; the method was fast and reliable, and offered an estimate on the development of the wear as a function of time.

Therefore, Hotelling´s T^2 control charts lose sensitivity to detect small and progressive changes in the process and they have difficulties in identifying the variable responsible for the change when the number of monitored variables is greater than 10.

On the other hand, revised reduction of variables methods such as PCA have difficulties to perform this task, when the correlation between variables is low.

Thus, we propose a new Method for detection of Small and Sudden Deviations in the process (SSDM), applicable when the correlation between variables is low; which is typical in marine propulsion processes.

First through the analysis of correlations between variables, we implemented a methodology to reduce the number of monitored variables, poorly correlated between them, of fuel process of a typical low-speed diesel engine running installed on a tanker ship and thus to improve the limitation that has MYT decomposition when the size of the monitored variables is large.

Afterwards, the selected variables were monitored through the Hotelling´s T^2 control chart, for some specific working conditions and through MYT decomposition; the variable that caused the out of range state with respect to the normal mode operation of the ship was identified.

In addition, the technique Hotelling´s T^2 was combined with univariate Cusum charts, to detect those variables which can generate small and progressive deviations in the process, typical in process where there are thermal exchanges, and cannot be detected through (Hotelling + MYT) control charts.

The main difference with current literature lies in the use of a new method called SSDM based on the combination of techniques (Hotelling´s T^2 + Cusum) in the main engine of a ship in seagoing conditions, offering reliable results and at the same time an economic and easy implementation.

2. Material and Methods

2.1 Machine study

The machine of study was the propulsion engine of a typical low-speed diesel engine, which is frequently installed in tanker ships and bulkarriers as the main engine. The basic technical details of the engine are listed in Table 1.

Manufacture	MAN B&W
Type	6S70ME-C8
Cycle	Low-speed 2 stroke
Nominal speed	91 r.p.m.
Rate power	19620 kW
Number of cylinders	6
Stroke	2800 mm
Bore	700 mm

Table 1. Technical details of engine studied.

The engine was installed on a Suez Max Crude Carrier with the characteristics listed in Table 2. This ship normally carries out regular voyages, between Western Africa, and Northern Europe where she is discharged.

Name of Ship	Confidential
Shipyard Build	Confidential
Year Built	2012
Type of Ship	Crude Oil Tanker
Class	Suez Max
Length Overall	274.20 m.
Extreme Breadth	48.04 m.
Draught	17 m.
Gross Tonnage	81.187
Net Tonnage	51.148

Table 2. Ship´s specifications.

In the fuel oil system [22], the fuel from the service tank is led to an electrically driven supply pump by means of which a pressure of approximately 4 bar can be maintained in the low pressure part of the fuel circulating system.

From here the fuel oil is led to an electrically-driven circulating pump, which pumps it through a heater and a full flow filter situated immediately before the inlet to the engine. This system is shown in Figure 1.

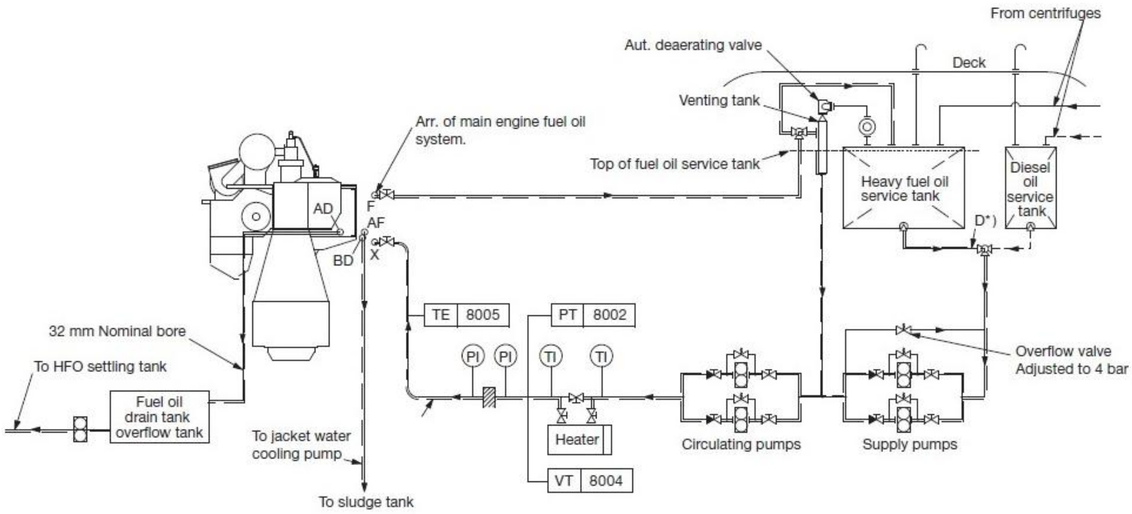

Figure 1 – Fuel Oil System

The fuel injection is performed by the electronically-controlled pressure booster located on the Hydraulic Cylinder Unit (HCU).

The Cylinder Control Unit (CCU) of the Engine Control System calculated the timing of the fuel injection and the exhaust valve activation, in accordance with the commands received from the Engine Control Unit (ECU).

To ensure ample filling of the HCU, the capacity of the electrically-driven circulating pump is higher than the amount of fuel consumed by the diesel engine. Surplus fuel oil is recirculated from the engine through the venting box.

2.2 Application of method

2.2.1. Step 1 – Data acquisition

The main engine has two monitoring systems and data acquisition: on one hand the CoCos EDS, a surveillance and diagnosis control system created by the engine manufacturer M.A.N.; and on the other hand the Integrated Automation System (IAS), where the thermodynamic process data are collected.

Our study only focused on the laden condition due to the high variability in the ballast condition, monitoring the behaviour of fuel oil process in the main engine during its voyage from Africa to Europe.

The fuel oil process of the main engine, was defined by p=11 variables: Engine Load, Fuel Index, Turbocharger speed Rpm (they were measured in the local control), Fuel Plunge Stroke (it is the average of the value of all the injectors), Scavenge air cooler air inlet temperature (it was measured from inlet of intercooler), Exhaust gas temperature at turbine inlet (it was measured from inlet of turbocharger), P (scav) (air pressure inlet combustion chamber), Estimate Effective Power (measured at the shaft), Compression Pressure (Pcom) and Maximum Pressure (Pmax) (they were the average of the value of all the cylinders, measured in the combustion chamber) and SFOC (fuel oil consumed by the engine) measured in a brake. For the selection of these, we have had the collaboration of the ship's engineers, and the manufacturer's data.

Data acquisition was performed under the following conditions: Speed over ground (SOG) between 12 and 14 knots with less than 18% slip, an average temperature of sea water of 20 ° C, average ambient temperature of 30 ° C and average temperature of the engine room of 37 ° C.

Four samples were taken daily, from all the selected variables, during 1 voyage which obtained a total of n=47 valid samples following the criteria previously mentioned.

The minimum, maximum, mean and standard deviations values of each are listed in Table 3. Each variable was identified with a correlative numbering.

No.	Variables	Unit	Min. Value	Max. Value	Means (µ)	Standard Deviations (σ)
1	Engine Load	%	54	61	56.85	1.546
2	Fuel Index	%	62.6	70.6	65.136	1.7644
3	Fuel plunger stroke	mA	2.58	2.77	2.662	0.0358
4	Scavenge air cooler air inlet temperature	°C	142	160	149.21	4.287
5	Exhaust gas temperature at turbine inlet	°C	354	407	372.63	15.306
6	Turbocharger speed	r.p.m.	10366	11202	10818.84	185.504
7	P (scav)	Bar	1.57	1.96	1.834	0.1093
8	Estimate Effective Power	kW	10318	10909	10540.11	130.761
9	Compression Pressure, Pcom	Bar	110.18	129.7	123.832	5.6346
10	Maximum Pressure, Pmax	Bar	138.43	142.47	140.739	1.0883
11	SFOC	g/kWh	154.28	164.01	158,511	2,042

Table 3. Means, standard deviations, maximum and minimum values.

There was a problem of lost data. The available data sampling period on board was too slow, therefore, it was necessary the use of interpolation technique to get the samples needed to implement the method; if these periods had been shorter, i.e, one sample, every half hour, the adjusted R^2 coefficients would had been higher, thereby increasing the reliability of the method.

Furthermore, to create the preliminary database, n=599 samples were generated of each variable through cubic spline interpolation [23].

With this, the number of samples needed to validate the study was achieved.The minimum sample size follows the equation (2) according to the number of variables, p [24]:

$$\text{Number of samples} = 2\text{p} + \frac{\text{p}\,(\text{p}-1)}{2} \qquad (2)$$

2.2.2. Step 2 Variable selection

With the samples generated in the preliminary database, a Pearson correlation analysis was performed [25], among the 11 variables in which fuel process was defined, listed in table 4.

		Variable identification number										
		1	2	3	4	5	6	7	8	9	10	11
Variable identification number	1	1.000	**0.880**	**0.750**	0.245	0.114	0.171	0.096	0.415	0.012	0.047	0.277
	2	**0.880**	1.000	**0.769**	0.373	0.125	0.282	0.156	0.358	0.046	-0.340	0.262
	3	**0.750**	**0.769**	1.000	0.181	0.101	0.119	0.067	**0.494**	-0.025	0.142	0.384
	4	0.245	0.373	0.181	1.000	0.160	**0.790**	0.414	0.409	0.235	-0.313	0.340
	5	0.114	0.125	0.101	0.160	1.000	-0.392	**-0.792**	-0.061	**-0.900**	**-0.693**	-0.377
	6	0.171	0.282	0.119	**0.790**	-0.392	1.000	**0.857**	**0.494**	**0.739**	0.115	**0.623**
	7	0.096	0.156	0.067	0.414	**-0.792**	**0.857**	1.000	0.404	**0.975**	0.457	**0.648**
	8	0.415	0.358	**0.494**	0.409	-0.061	**0.494**	0.404	1.000	0.286	0.318	**0.843**
	9	0.012	0.046	-0.025	0.235	**-0.900**	**0.739**	**0.975**	0.286	1.000	**0.563**	0.570
	10	0.047	-0.034	0.142	-0.313	**-0.693**	0.115	0.457	0.318	**0.563**	1.000	0.439
	11	0.277	0.262	0.384	0.340	-0.377	**0.623**	**0.648**	**0.843**	**0.570**	0.439	1.000

Table 4. Correlation between variables.

The process was monitored by the minimum number of variables, (p=3 variables: Fuel Index, Exhaust gas temperature at turbine inlet and Turbocharger speed), following the criteria: the selected variables have one correlation between them less than 0.49 and the selected variables had a correlation with at least one of the unselected variables equal to or higher than 0.49.

Finally, through SPSS software, it was found the adjustment of models among the three selected variables and their predictive variables using a multivarible regression analysis, obtaining the following coefficients of determination R^2 adjusted, 0.8, 0.95 and 0.96 for each model respectively.

Conventional methods for variable reduction such as PCA were not efficient; with two principal components represented only the 81% of the process. Five principal components were required to represent 96% of the process.

2.2.3 Step 3 – Purging process

To create the historical data set (HDS), n=599 samples from the preliminary database were monitored with mean μ and standard deviation σ, following a normal distribution, N_p (μ, σ) estimated for the multivariate process, through the Hotelling´s T^2 control chart [26], following the equation (3).

$$T^2 = (X_i - \overline{X})' \; S^{-1}(X_i - \overline{X}) \qquad (3)$$

Where:

$X_i = (X_{i1;} X_{i2}; \dots ; X_{iP})'$

Mean vector, $\overline{X}$ = (65.095 372.43 10821.55)

Inverse covariance matrix:

$$S^{-1} = \begin{pmatrix} 0.4515 & -0.0135 & -0.0016 \\ -0.0135 & 0.0057 & 0.0002 \\ -0.0016 & 0.0002 & 4.479\mathrm{E}-05 \end{pmatrix}$$

Depending on the circumstances, the T^2 statistic can be described by three different probability functions: the Beta, the *F* and the chi-square distributions.

When (μ, σ) are estimated, the Beta distribution is used in the purging process of a Phase I operation, whereas the *F* distribution is used in the development of the control process in a Phase II operation. When (μ, σ) are known, the chi-square has applications in both Phase I and Phase II operations [6].

During the purging process, the atypical observations of the process, obtained in the generation the preliminary database, were detected and eliminated in order to avoid possible errors in results.

For the calculation of the UCL (Upper Control Limit), we used the β distribution with α=0.05, for type I errors [27].

The UCL was determined by the following equation (4):

$$\mathrm{UCL} = \left\{\frac{(n-1)^2}{n}\right\} \beta_{\left\{\alpha; \frac{p}{2}; (n-p-1)/2\right\}} \qquad (4)$$

Where:

n: Size of data set,

p: Number of variables.

$\beta\{\alpha; p/2; (n-p-1)/2\}$, is the upper αth, quantile of the beta distribution, $\beta\{p/2; (n-p-1)/2\}$

If the value of T^2, which was monitored for an observation, exceeded the UCL, observation was purged from the preliminary data.

With the remaining observations, we calculated new estimates of the mean vector and covariance matrix. Again, we removed all detected outliers and repeated the process until a homogeneous set of observations was obtained. The final set of data was the HDS, the normal operation of the process, formed by 307 samples. The algorithms necessary to carry out the process of purged, are developed by using Labview software.

In Table 5, the detected outliers are represented in each step until the HDS was obtained.

No. of Observations	UCL	No. outliers detected
599	7.783	24
575	7.782	31
544	7.78	48
496	7.777	13
483	7.776	7
476	7.775	8
468	7.774	5
463	7.774	5
458	7.774	2
456	7.773	7
449	7.773	12
437	7.772	12
425	7.77	24
401	7.768	17
384	7.766	14
370	7.764	14
356	7.762	11
345	7.76	14
331	7.758	13
318	7.755	8
310	7.754	2
308	7.753	1
307	7.753	0

Table 5. Steps to get the HDS.

2.2.4. Step 5 – Control Procedure

In this step, was tested to see if a new entry of data generated a signal, with respect to the historical data set (HDS).

Due to the amount of monitored variables being small, the statistic T^2 Hotelling was used. Out of the different proposed methods in the literature to determine what were the variables that caused the out of range state of the process, the MYT decomposition was selected [28].

The new entry of data, corresponded to samples acquired during one voyage in laden condition, obtaining a total of 13 samples, listed in table 6, after having analysed them according to the criteria of the normal condition of the operation.

Obs. No.	**Fuel index (%)**	**Exhaust gas temperature (°C)**	**Turbocharger speed (r.p.m.)**
1	64.70	373	10992
2	66.30	373	11052
3	64.50	372	10960
4	68.40	378	11260
5	66.30	375	11076
6	65.70	381	11159
7	69.00	369	10987
8	66.90	370	11048
9	64.70	369	10906
10	69.50	369	11139
11	65.40	368	11154
12	64.30	369	11070
13	65.90	352	10141

Table 6. New data entry

The values T^2, for new data entry, were calculated, according to the following equation (5).

$$T^2 = (X_i - \overline{X})' \ S^{-1}(X_i - \overline{X}) \tag{5}$$

Where:

$\overline{X}$ and S^{-1}, the mean vector and inverse covariance matrix are obtained from the HDS and X_i, new data entry.$X_i = \ (X_{i1;}X_{i2}; \dots ; X_{ip})'$

$X_i = (X_{i1;}X_{i2}; \dots ; X_{ip})'$

For the calculation of the UCL (Upper Control Limit), we used the F distribution with α=0.05, for type II errors [29].The level of α can be variable, making more or less strict the method. The chosen alpha level is normally used in industrial processes and it depends on conditions of operation of the ship.

The UCL was determined by the following equation (6):

$$\mathrm{UCL} = \left\{\frac{\mathrm{p(n+1)(n-1)}}{\mathrm{n(n-p)}}\right\} \mathrm{F}_{\{\alpha;\mathrm{p};(\mathrm{n-p})\}} \tag{6}$$

Where p, is the number of variables, n, is the size of the HDS and $\mathrm{F}\{\alpha; \mathrm{p}; (\mathrm{n}-\mathrm{p})\}$, is the αth, quantile of $\mathrm{F}\{\mathrm{p}; (\mathrm{n}-\mathrm{p})\}$.

The values of T^2 which exceeded the UCL were declared as signals and thus concluded that the observation was out of range with respect to the mode of normal operation of the process.

Once the T^2 statistical detected samples which were out of range in the process from normal operating conditions, the MYT decomposition was used[30, 31], to identify the variables with more weight, responsible for the out of range state for each sample.

The general decomposition for "p" variables of the Hotelling´s T^2 statistic, follow the equation:

$$\mathrm{T}^2 = T_1^2 + T_{2.1}^2 + T_{3.1,2}^2 + T_{4.1,2,3}^2 + \ldots + T_{P.1,\ldots,P-1}^2 = T_1^2 \sum_{j=1}^{P-1} T_{j+1.1,---,j}^2 \tag{7}$$

The final T^2 value,T_1^2, is Hotelling´s statistic for the first variable. It reduces to the square of the univariate t statistic for the initial variable:

$$T_1^2 = \frac{\left(X_1 - \bar{\mathrm{X}}_1\right)^2}{S_1^2} \tag{8}$$

Where, $\bar{\mathrm{X}}_1$ and S_1 is the mean and standard deviation of variable X_1.

The statistic $T_{P.1,\ldots,P-1}^2$ is the p^{th} component of the vector X_i adjusted by the estimates of the mean and standard deviation of the conditional distribution of X_P given $X_1, X_2, \ldots, X_{p-1}$. It is given by

$$T_{P.1,\ldots,P-1}^2 = \frac{\left(X_{ip} - \bar{\mathrm{X}}_{\mathrm{P}.1,\ldots,\mathrm{P}-1}\right)}{\mathrm{S}_{\mathrm{p}.1,\ldots,\mathrm{p}-1}} \tag{9}$$

Where:

$\bar{\mathrm{X}}_{\mathrm{P}.1,\ldots,\mathrm{P}-1} = \bar{\mathrm{X}}_{\mathrm{P}} + \mathrm{b_p}' \,(X_i^{(p-1)} - \mathrm{X}^{(\mathrm{p}-1)})$,

$\bar{\mathrm{X}}_{\mathrm{P}}$ is the sample mean of n observations on the p^{th} variable, $\mathrm{b_p} = S_{XX}^{-1} s_{xX}$ is a (p-1) – dimensional vector estimating the regression coefficients of the p^{th} variable regressed on the first p-1 variables,

$S_{p.1,\ldots,p-1}^2 = S_X^2 - S_{xX}' S_{XX}^{-1} s_{xX}$ and $\mathrm{S} = \begin{pmatrix} S_{XX} & S_{xX} \\ S_{xX}' & S_X^2 \end{pmatrix}$.

2.2.5. Step 5 – Control of predictive variables.

Finally the univariate Cusum technique was used [32], for each of the predictive variables of the variable detected using MYT, and so, to detect if any of them was responsible for the out of range state of the process.

The tabular Cusum, calculated the deviations of each value with respect to the target value μ_0, distinguishing between deviations positive C^+ and negative C^-. The statistics C^+ and C^- have a form,

$$C_i^+ = max[0, X_i - (\mu_0 + K) + C_{i-1}^+] \quad (10)$$

$$C_i^- = max[0, (\mu_0 - K) - X_i + C_{i-1}^-] \quad (11)$$

Where the starting values are $C_0^+ = C_0^- = 0$

The reference value K, is often chosen about halfway between the target μ_0 and the out of range value of the mean μ_1 (15 % about μ_0), that we are interested in detecting quickly.

$$K = \frac{|\mu_1 - \mu_0|}{2} \quad (12)$$

The decision interval H, was determined by the following equation (13):

$$H = h \cdot \sigma \quad (13)$$

Where σ is the standard deviation of the process in control and h=5. The chosen h level is normally used in industrial processes.

If C_i^+ or C_i^- exceeds the decision interval H, the process is considered to be out of range.

The Average Run Length (ARL) was calculated according to the ARL approximation given by Siegmund [33]. For a one-sided Cusum (C_i^+ or C_i^-) with parameters h and k, Siegmund´s approximation has a form:

$$ARL = \frac{\exp(-2\Delta b) + 2\Delta b - 1}{2\Delta^2} \quad (14)$$

For $\Delta \neq 0$, where $\Delta = \delta^* - k$ for the upper one-sided Cusum C_i^+, $\Delta = -\delta^* - k$ for the lower one-sided Cusum C_i^-, b = h + 1.166, and $\delta^* = (\mu_1 - \mu_0) / \sigma$

The quantity δ^*, represents the shift in the mean, in units of σ, for which the ARL is to be calculated. Therefore, if $\delta^* = 0$, we consider an ARL_0 of 48, corresponding to the samples obtained during one day; whereas if $\delta^* \neq 0$, we calculate the value of ARL_1 corresponding to a shift of size δ^*.

The ARL of the two one-sided statistics of Cusum, ARL^+ y ARL^-, was determined by the following equation (15):

$$\frac{1}{ARL} = \frac{1}{ARL^+} = \frac{1}{ARL^-} \quad (15)$$

3- Results

3.1 Application of Hotelling´s T^2 statistical.

In this step, T^2 values were calculated according to the Eq. (5), for each one of the 13 new observations, and they were monitored in a control chart, according to Figure 2. with a upper control limit previously calculated, according to the expression Eq. (6), valued in UCL= 7.9808.

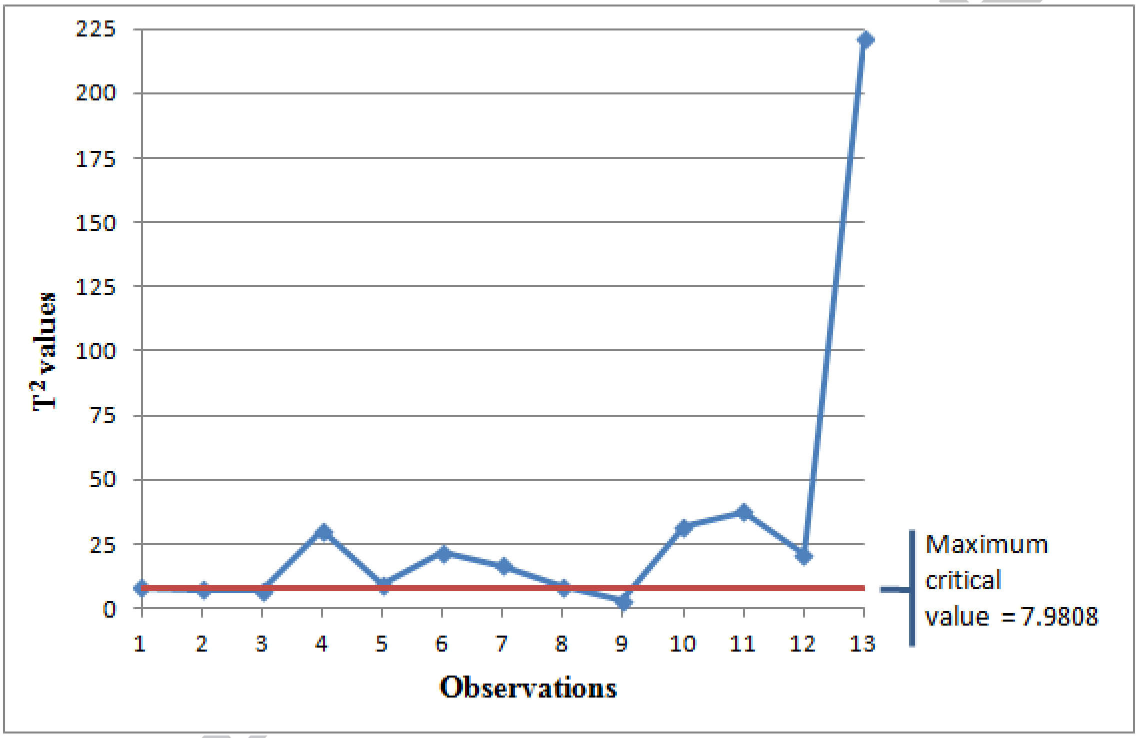

Figure 2. Control Chart

The control chart shows that there are values of T^2 above the UCL which indicates that, in that interval of time, the process had a deviation from its normal operation mode.

This situation does not mean that the engine is failing, but that the process has moved from normal operating conditions, and if this trend is repeated in time, corrective actions should be taken to prevent a malfunction in the process.

In the next stage, we identified which were the variables that had produced the out of range state for each observation.

3.2 Decomposition MYT

In this stage, using the MYT decomposition technique, each T^2 value was decomposed for each one of the signals to detect which was the variable which had contributed more strongly to the out of range state of the process. The unconditional terms were calculated following the Eq. (8), and the conditional terms were calculated following the Eq. (9), the decomposition is listed in table 7. This shows that the main variable that caused the deviation from its normal operating mode was the Turbocharger speed variable.

Observations	Variables
1	*Exhaust gas temperature*
4	*Turbocharger speed*
5	*Turbocharger speed*
6	*Exhaust gas temperature* and *Turbocharger speed.*
7	*Fuel Index*
8	*Turbocharger speed*
10	*Fuel Index*
11	*Turbocharger speed*
12	*Turbocharger speed*
13	*Turbocharger speed*

Table 7. Decomposition MYT

3.3 Application of Cumulative sum

In this stage, it was monitored the predictive variables of the Turbocharger speed variable, using the Cusum charts, to detect if any of them was responsible of the out of range state of the process.

The mean, standard deviation values of each predictive variable in control are listed in Table 8.

Variables	Unit	Means(μ)	Standard Deviations (σ)
Scavenge air cooler air inlet temperature	°C	147.35	2.24
P (scav)	Bar	1.84	0.1
Estimate Effective Power	kW	10542.36	103.63
Compression Pressure, Pcom	Bar	124.8	5.38
SFOC	g/kWh	158.58	1.7

Table 8. Mean and standard deviation of predictive variables.

61 observations for each of the variables were monitored; the first 48 observations corresponded to the ARL_0 and the following 13 were new input data.

Figures 3a, 3b, 3c, 3d, 3e, show Cusum charts for each of the variables. It was noted that the only variable that exceeded its decision interval was the SFOC variable, where at sample 50 is C_{50}^+ = 10.8. Since this is the first period at which C_i^+ > H=8.5, we would conclude that the variable was out of range in this point.

However, the tabular Cusum also indicates when the shift probably occurred. The first consecutive sample in which C_i^+ > 0 first exceed the value of H, was the period 49, C_{49}^+ = 5.36, thus indicating that the mismatch in the variable could have started in the sample 49.

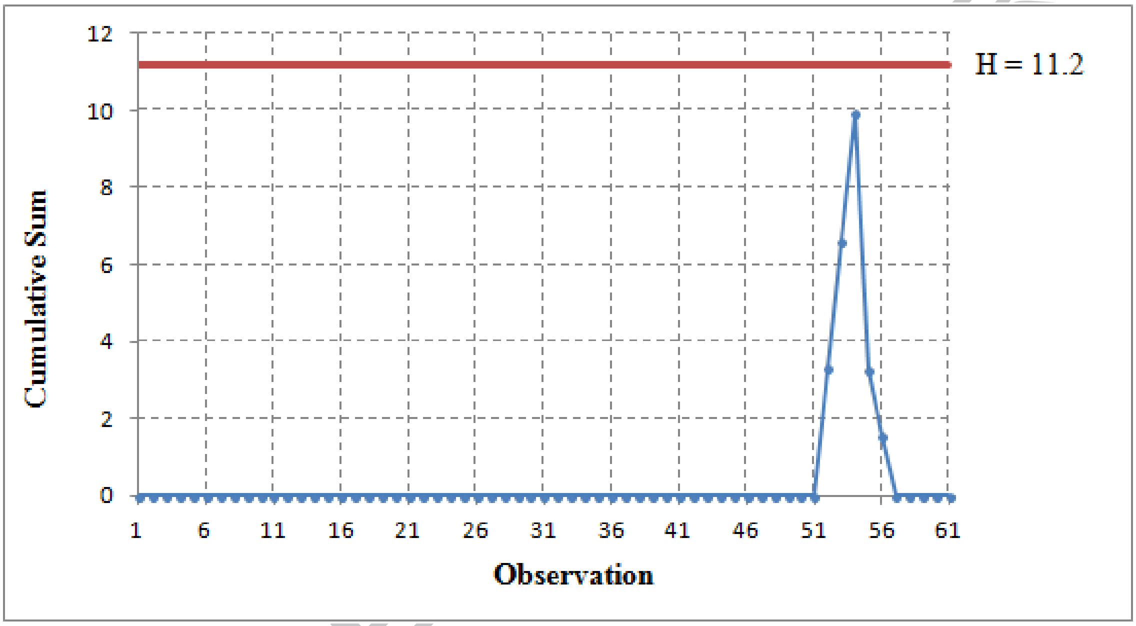

Figure 3a - Scavenge air cooler air inlet temperature (ARL_1=1.23).

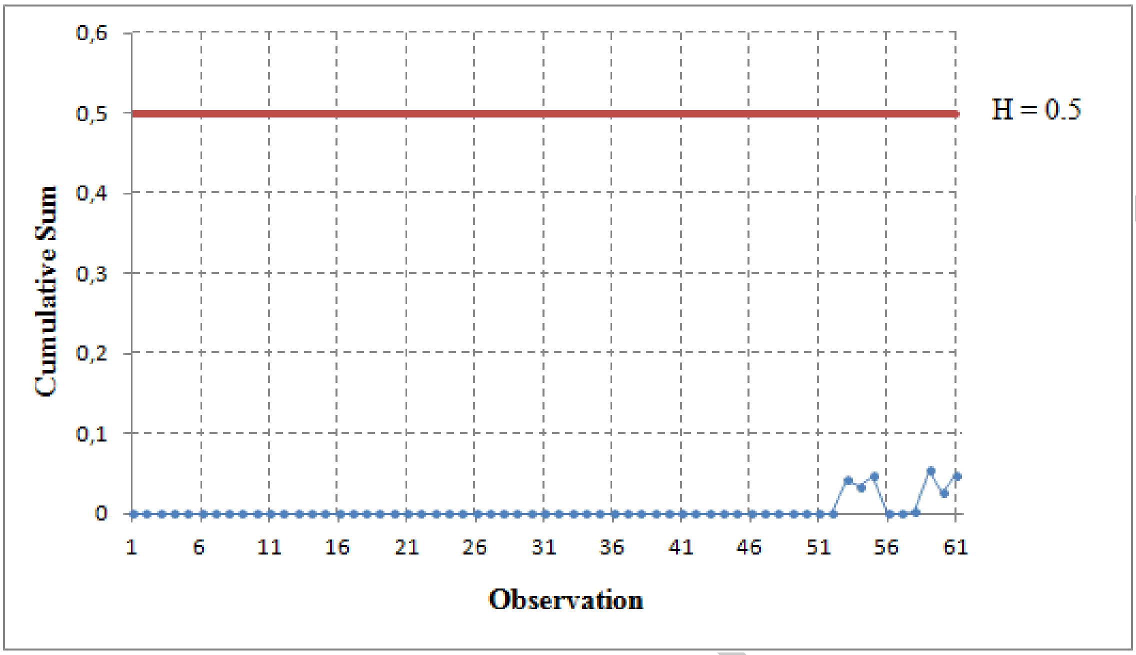

Figure 3b - P (scav) (ARL_1=4.29)

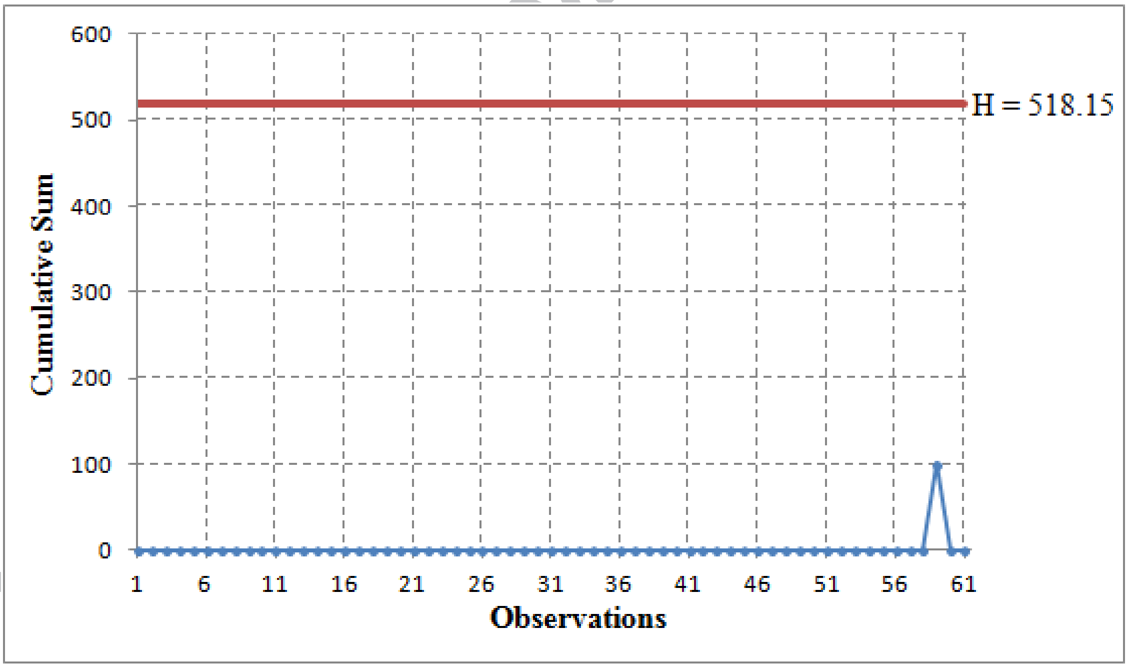

Figure 3c - Estimate Effective Power (ARL_1=0.79)

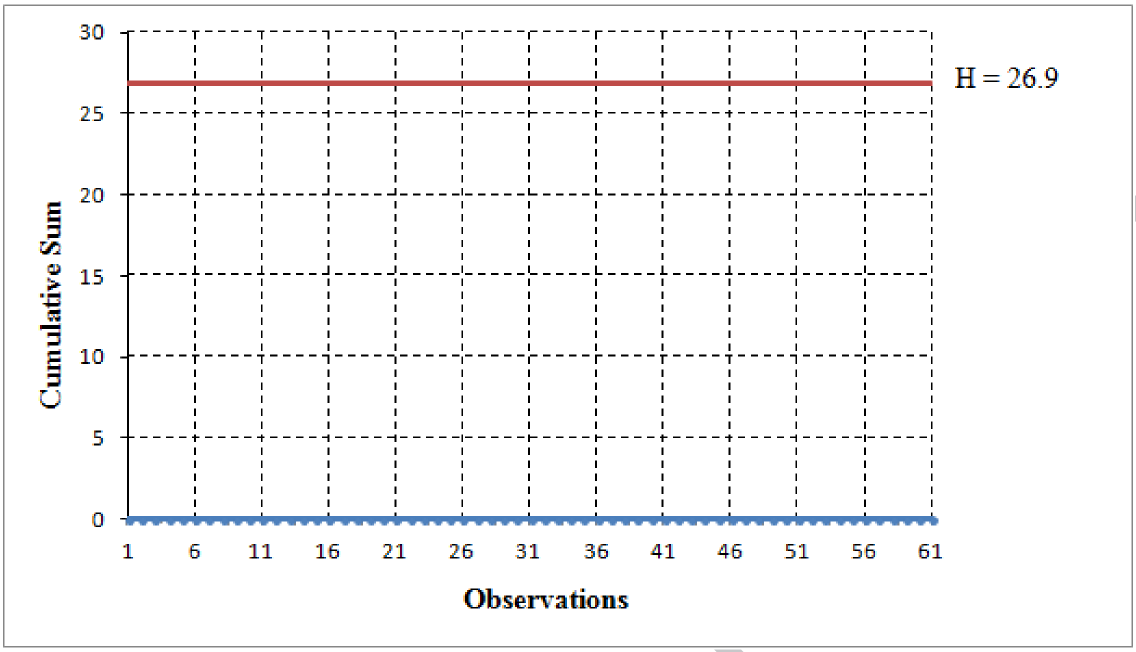

Figure 3d - Pcom (ARL_1=3.37)

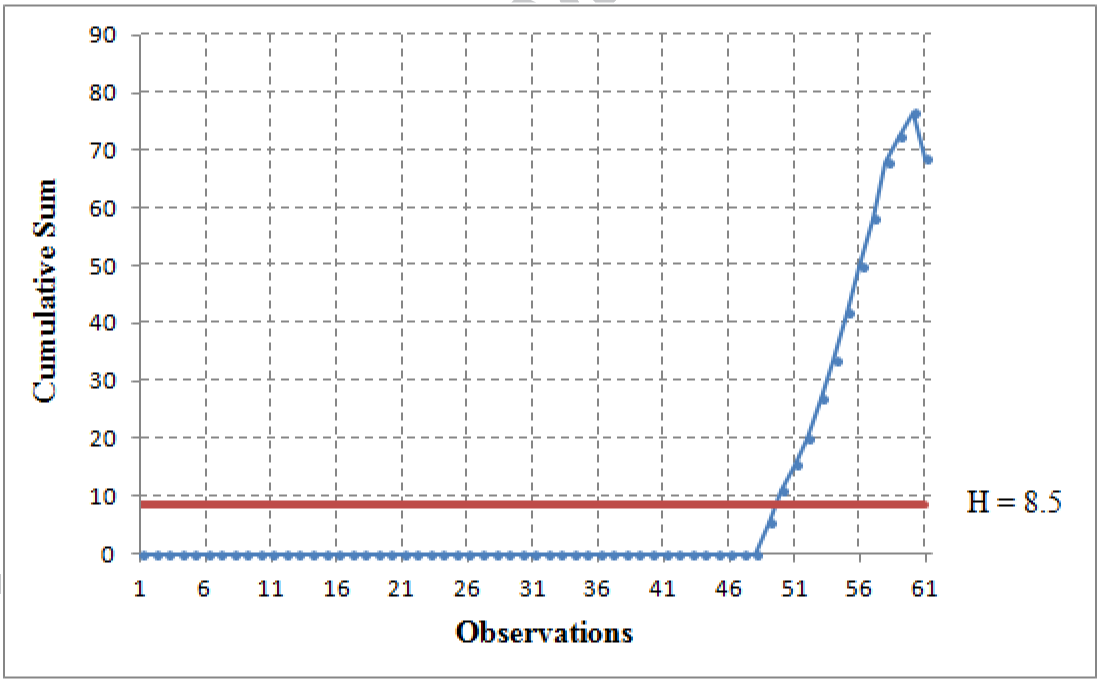

Figure 3e - SFOC (ARL_1=0.87)

4-Discussion

The fuel oil process of a 2-stroke marine diesel engine was monitored by only three variables with low correlation between them, through a combination of univariate and multivariate techniques (Hotelling´s T^2 + Cusum).

Hotelling´s T^2 control charts performance decreased as it increased the number of variables to be monitored. It was chosen the minimum number of variables to be monitored, p=3, from among the 11 variables representing the entire process through a multivariate regression analysis, ensuring fitting models between variables and their predictive variables, with coefficients of determination R^2 adjusted higher than 0.8.

Multivariate charts detected observations out of range with respect to the optimal conditions of the process; in the table 9, there is the chronology of the out of range observations, with its respective T^2 values.

Date	Observations	T^2
25/08/2016	1	8.042
28/08/2016	4	30.395
29/08/2016	5	9.294
30/08/2016	6	21.884
31/08/2016	7	16.745
01/09/2016	8	8.489
03/09/2016	10	31.785
04/09/2016	11	37.641
05/09/2016	12	21.212
06/09/2016	13	221.034

Table 9. Chronology of Observations.

Observations that were above the limit of control were decomposed, identifying the variable *turbocharger speed* as the main variable that originated the out of range state of the multivariate process.

Hotelling´s T^2 Technique has the advantage that effectively detects high and sudden changes in the process but can't detect small and progressive changes.

For this reason, the predictive variables in the variable *turbocharger speed* were monitored through Cusum charts, to try to detect small and progressive changes in the process that had not been detected by means of multivariate charts.

It was established a decision interval, less than the one marked by the manufacturer, 15% over the average value of each variable in optimal condition operation. The SFOC variable exceeded the threshold and was detected when it began to deviate from its normal condition before the established threshold.

The cleaning of the intercooler, for service reasons, only was made with chemical products, the last cleaning in depth had been 6 months ago; this situation generated a progressive fouling in the intercooler. In order to maintain the speed of the vessel a small deviation in the SFOC variable was caused.

5- Conclusions

The proposed methodology for reduction of variables, through the analysis of correlations between variables, was capable to reduce the number of variables, poorly correlated between them, of fuel process of a running marine diesel engine; conventional methods for variable reduction such as PCA was shown that were not efficient when the correlation between variables was poor.

Through proposed methodology of monitoring of variables SSDM based on the combination of (Hotelling T^2 + Cusum) charts, high and sudden and also small and progressive deviations in the process were detected.

The value of the differential pressure in the intercooler was not enough to overcome the threshold set by the manufacturer; a small deviation in the SFOC variable was generated. Without this identification, they would have had to wait for the value of the differential pressure was above the threshold set by the manufacturer, resulting in a higher fouling of the intercooler and an increase of the SFOC variable.

Many processes involved in the operation of a marine diesel engine have decay small and progressively in addition to suddenly, for this reason, this method has the advantage that is can be customized for any type of engines because it is capable of detecting any type of deviation (small and sudden) in the process; this can be performed in a simple and economical way, at the request of the shipowner, depending on the operational conditions of the ship.

In future work, the effectiveness of the method using Multivariate Cusum Charts (MCusum) in this type of process could be studied.

6- References

[1] Bersimis S, Psarakis S, Panaretos J. Multivariate statistical process control charts: an overview. Qual Reliab Eng Int 2007;23(5):517-543.

[2] González I, Sánchez I. Variable selection for multivariate statistical process control. Journal of Quality Technology 2010;42(3):242.

[3] Basurko OC, Uriondo Z. Condition-Based Maintenance for medium speed diesel engines used in vessels in operation. Appl Therm Eng 2015;80:404-412.

[4] Garcia-Alvarez D, Fuente M. Estudio comparativo de técnicas de detección de fallos basadas en el Análisis de Componentes Principales (PCA). Revista Iberoamericana de Automática e Informática Industrial RIAI 2011;8(3):182-195.

[5] Boullosa D, Larrabe JL, Lopez A, Gomez MA. Monitoring through T 2 Hotelling of cylinder lubrication process of marine diesel engine. Appl Therm Eng 2017;110:32-38.

[6] Mason RL, Young JC. Multivariate statistical process control with industrial applications. : Siam; 2002.

[7] Kourti T, MacGregor JF. Multivariate SPC methods for process and product monitoring. Journal of quality technology 1996;28(4):409-428.

[8] Control Estadístico Mega Variante para los procesos del siglo XXI. Dpto de Estadística e Investigación Operativa. UPV. Congreso Nacional de Estadística e Investigación Operativa. Valencia, España; 2003.

[9] Monitoring of a thermoelectric power plant based on multivariate statistical process control. Evolving and Adaptive Intelligent Systems (EAIS), 2016 IEEE Conference on: IEEE; 2016.

[10] Peña D. Análisis de datos multivariantes. : McGraw-Hill Madrid; 2002.

[11] Wang Y, Ma X, Joyce MJ. Reducing sensor complexity for monitoring wind turbine performance using principal component analysis. Renewable Energy 2016;97:444-456.

[12] Larrabe J, Gómez M, Sotés I, Alvarez F, Rey M, Mielgo V, et al. Vessel condition monitorization through satellite using principal component analysis. Journal of Maritime Research 2014;7(3):3-14.

[13] Soares VB, Pinto JC, de Souza MB. Alarm management practices in natural gas processing plants. Control Eng Pract 2016;55:185-196.

[14] Vega-Vilca JC, Guzmán J. Regresión PLS y PCA como solución al problema de multicolinealidad en regresión múltiple. Revista de Matemática Teoría y Aplicaciones 2011;18(1):09-20.

[15] Joe Qin S. Statistical process monitoring: basics and beyond. J Chemometrics 2003;17(8-9):480-502.

[16] Zamarrón AMC, Prado EM, Luis FZ. Monitoreo y control de un proceso normal multivariado. Conciencia Tecnológica 2012(43):29-35.

[17] Mason RL, Young JC. Improving the sensitivity of the T2 statistic in multivariate process control. Journal of Quality Technology 1999;31(2):155.

[18] Aparisi F, GARCIA JC. Aumento de la potencia del gráfico de control multivariante T 2 de Hotelling utilizando señales adicionales de falta de control. Estadística española 2001;43(148):171-188.

[19] Stamboliska Z, Rusiński E, Moczko P. Proactive condition monitoring of low-speed machines. Proactive Condition Monitoring of Low-Speed Machines: Springer; 2015. p. 53-68.

[20] Rodríguez LI, Rodríguez MI, Rodríguez MA. Análisis Comparativo de los Gráficos de Control Shewhart, CUSUM y EWMA, para Procesos en Condiciones de Control y Fuera de Control.

[21] Statistical fault diagnosis of wind turbine drivetrain applied to a 5MW floating wind turbine. Journal of Physics: Conference Series: IOP Publishing; 2016.

[22] MAN B&W S70ME-C8.5-TII – Project Guide. Electronically Controlled two-stroke engines. Copenhagen: MAN B&W Diesel, 2016. Available at: http://marine.man.eu/applications/projectguides/2stroke/content/printed/S70ME-C8_5.pdf. Accessed 12/08, 2016.

[23] McKinley S, Levine M. Cubic spline interpolation. College of the Redwoods 1998;45(1):1049-1060.

[24] Williams B. A sampler on sampling. : JSTOR; 1978.

[25] Pearson K. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London.Series A, containing papers of a mathematical or physical character 1896;187:253-318.

[26] Hotelling H. Multivariate quality control. In C. Eisenhart, M. W. Hastay and W.A. Wallis, eds.", Techniques of statistical analysis 1947. New York: McGraw Hill, pp 111-184,1947

[27] Mason RL, Chou Y, Sullivan JH, Stoumbos ZG, Young JC. Systematic patterns in T (2) charts. Journal of Quality Technology 2003;35(1):47.

[28] Maravelakis P. An investigation of some characteristics of univariate and multivariate control charts. Unpublished doctoral dissertation, Department of Statistics, Athenas University of Economics and Business.(Available from internet: http://www.stat-athens.aueb.gr/~jpan/diatrives/Maravelakis/Index.html) 2003.

[29] Mason RL, Tracy ND, Young JC. A practical approach for interpreting multivariate T2 control chart signals. Journal of Quality Technology 1997;29(4):396-406.

[30] Agog NS, Dikko HG, Asiribo OE. Decomposition hotelling's T2 statistic using four variables. Journal of Innovate Research in Science, Engineering and Technology 2014;3 (4).

[31] Mason RL, Tracy ND, Young JC. Decomposition of T2 for multivariate control chart interpretation. Journal of quality technology 1995;27(2):99-1108.

[32] Montgomery DC. Introduction to statistical quality control. John Wiley & Sons; 2009. 6th. ed.

[33] Siegmund DO. Sequential Analysis: Tests and Confidence Intervals Springer-Verlag. New York 1985.