

PAPER • OPEN ACCESS

Magnetic Flux Leakage and Principal Component Analysis for metal loss approximation in a pipeline

To cite this article: M Ruiz *et al* 2015 *J. Phys.: Conf. Ser.* **628** 012027

View the [article online](#) for updates and enhancements.

You may also like

- [Quantifying predictive uncertainty in damage classification for nondestructive evaluation using Bayesian approximation and deep learning](#)
Zi Li and Yiming Deng
- [Accurate 3D reconstruction of complex defects based on combined method of MFL and MFDs](#)
Hengtao Li, Xiucheng Liu, Bin Wu et al.
- [Universal entanglement signatures of quantum liquids as a guide to fermionic criticality](#)
Siddhartha Patra, Anirban Mukherjee and Siddhartha Lal



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Magnetic Flux Leakage and Principal Component Analysis for metal loss approximation in a pipeline

M. Ruiz¹, L.E. Mujica¹, M. Quintero², J. Florez² and S. Quintero²

¹Departament of Applied Mathematics III. Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona. Consorci Escola Industrial de Barcelona. Universitat Politècnica de Catalunya. BARCELONATECH. Spain

²Department of Electronics. Research Institute of Corrosion CIC. Colombia. Km 2 via Refugio Sede UIS Guatiguará, Piedecuesta, Colombia

E-mail: magda.ruiz@upc.edu / luis.eduardo.mujica@upc.edu /
mquintero@corrosion.uis.edu.co / jflorez@corrosion.uis.edu.co /
squintero@corrosion.uis.edu.co

Abstract.

Safety and reliability of hydrocarbon transportation pipelines represent a critical aspect for the Oil and Gas industry. Pipeline failures caused by corrosion, external agents, among others, can develop leaks or even rupture, which can negatively impact on population, natural environment, infrastructure and economy. It is imperative to have accurate inspection tools traveling through the pipeline to diagnose the integrity. In this way, over the last few years, different techniques under the concept of structural health monitoring (SHM) have continuously been in development.

This work is based on a hybrid methodology that combines the Magnetic Flux Leakage (MFL) and Principal Components Analysis (PCA) approaches. The MFL technique induces a magnetic field in the pipeline's walls. The data are recorded by sensors measuring leakage magnetic field in segments with loss of metal, such as cracking, corrosion, among others. The data provide information of a pipeline with 15 years of operation approximately, which transports gas, has a diameter of 20 inches and a total length of 110 km (with several changes in the topography). On the other hand, PCA is a well-known technique that compresses the information and extracts the most relevant information facilitating the detection of damage in several structures. At this point, the goal of this work is to detect and localize critical loss of metal of a pipeline that are currently working.

1. INTRODUCTION

In Colombia, the petrochemical industry facilities have structures with more than 30 years in service. The ferrous pipe structures of oil and gas production and, the transmission pipelines are, in majority, buried. Nowadays, phenomena like corrosion, mechanical stress, soil erosion, worker mistakes and damages caused by third parts (for instance, excavation machinery that can strike the pipe causing scratches or dents) have generated several problems on pipelines.

The need to manage and maintain pipeline system has become increasingly in an important priority to operators. Thus, major investment on integrity programs with In-Line Inspection Tools - smart pigs - have been improved in order to examine the pipelines and avoid environmental, financial and social disasters. At the same time, international regulations have raised levels of requirements on reliable hydrocarbon and gas transmission.



The Colombian government is developing a regulation which contains the integrity concept of oil and gas transport in accordance with the international standards. These rules also should fulfill the national requirements. As a result of this process, operators will require the establishment of the physical condition of oil and gas pipelines, as well as the generation of correction and prevention strategies to integrity management assurances hold in the pipeline. Smart inspection is an effective technique to consider perform detailed inspections on pipelines.

Internal inspection of pipelines are currently carried out in Colombia by foreigner companies. However, the cost of the service is excessively high without negotiation opportunity, limitation in the analysis of results and without monitoring of action plans. All these disadvantages restrict the regular use of the In-Line Inspection tools.

Recently, Research Institute of Corrosion - CIC (Corporación para la Investigación de la Corrosión) run their own smart pig ILI tool in pipelines. This is the first device for this purpose developed completely in Colombia. The inspection technology is based on inertial and operational trends, ITION (Inertial Technology Inspection and Operational Trends). Up to date, the technology has been tested several times inside of pipelines providing valuable information along of thousand kilometers. These records contain a huge amount of data. An univariate statistical method can be used to determine the thresholds for each observation variable. However, it does not analyze the correlated information between variables. In this way, the main contribution of Principal Component Analysis (PCA) in this work is to monitor the structure by using the whole available variables to detect statistically significant events or damages since the information is compressed and a pattern recognition in the signal is performed for structural monitoring [1].

For this work, The CIC has provided the first measurement made with the ITION technology -smart pig ILI- pilot test. A brief summary of this technology is described in sections 2 and 3. PCA approach is explained in section 4. In section 5, the raw data analysis and PCA results are presented. Finally, discussion and conclusions are drawn.

2. ITION Technology Description

2.1. Generalities

One way to monitor pipelines is the In-Line Inspection pilot Tool - smart pig - which is a vehicle that travels inside the pipeline. The CIC has developed the Inertial Technology Inspection and Operational Trends -ITION-. This tool is composed by electronic system capable of acquiring, processing and storing data signals (Figure 1). The sensor system has mainly an Inertial Measurement Unit. It consists by accelerometers and gyroscopes with high precision and sensitivity. As the tool travels along the pipe, the inertial system records information on all components of accelerations and angular velocities associated with the movement dynamics of the device within the pipeline.

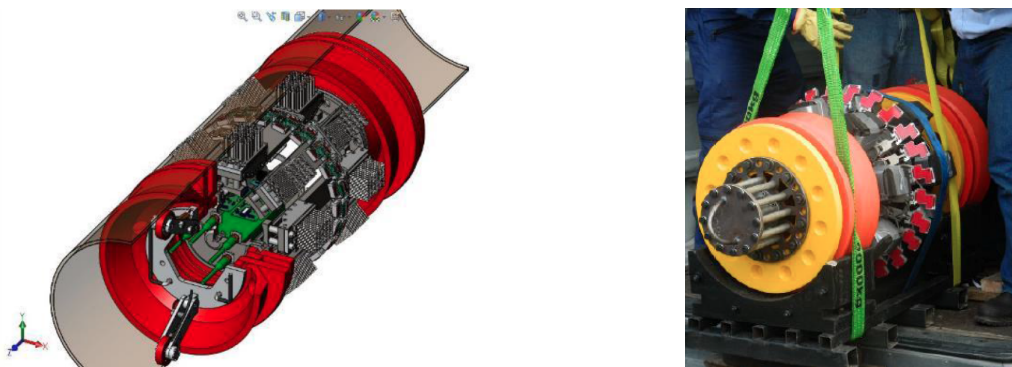


Figure 1: Inertial Technology Inspection and Operational Trends -ITION- developed by CIC

A second set of accelerometers is installed on the tool in order to filter and remove characteristic noise of inertial systems. In addition, the tool contains pressure and temperature sensors to study the phenomena present in the conveying pipeline. The system has been expanded to incorporate a prototype array of linear transducers that varies its output voltage in response to magnetic fields, providing a constant driving current to the sensors and amplifying the output signal. Odometers systems measures the distance traveled by the tool and allows the calculation of the instantaneous speed, information that is used to compensate the error associated with the nature of the inertial measurement ROA2012. All electronics are protected by a mechanical housing and designed to fit most of the conventional scraper routinely used in pipeline cleaning processes. The flexibility and adaptation of the ITION technology allow its application using high frequency with low cost [3]. In consequence, an effective monitoring and diagnosis of pipelines changes due to the passing of time can be achieved.

2.2. Signals description

After running the ITION, the following signals are recorded:

- Signal 1: Intensity of axial movement (the direction of flow of the transported product)
- Signal 2: Intensity rate of axial rotation
- Signal 3: Intensity resulting from the rotational sensors (three-axis)
- Signal 4: Intensity sensor remanent fields
- Signal 5: Propulsive force experienced by the tool
- Signal 6: Temperature of the transported product

Each signal has more than 7 millions of samples. These signals are depicted in Figure 2. For the purpose of this article, the signals are processed without emphasis on the characteristics of the phenomenon that describes or the behavior associated with any technique.

3. Magnetic flux leakage technique

Magnetic Flux Leakage technique - MFL - is used to detect metal loss defects and flaws in steel pipelines. It is implemented on In-Line inspection tools with powerful neodymium magnets, that requires a saturated field through the ferrous wall to achieve desired performance [4].

When the ITION goes through the pipe structure, a remanence effect is presented. The remanence effect is the magnetization left in the pipe wall by a pair of high energy permanent magnets employed in the inspection vehicle. Laboratory tests show that some kind of remanence are undesirable before running MFL inspection. Remanence is a magnetic distortion that affects the offset of the signal position in homogeneous fields on pipe walls without metal loss or corrosion. There is some array of magnetometers transducers installed inside and outside of the body of the tool; the external magnetometers register the residual field over the internal pipe wall surface.

The main objective is to understand the residual magnetization for improving the ability to reliably detect this kind of damage. In this way, identifying characteristics of the residual field signals is possible through the application of computational techniques [5]. In consequence, this work is focused on identifying pattern of defects, pipe accessories and reference grouped tags, which are important in the study and post-processing analysis of the data collected by ITION.

4. Principal Component Analysis

PCA is widely used in this kind of problems since it allows represent graphically as effectively as possible observations belongs to a general m -dimensional space in a small dimensional space (r) [6]. Besides, PCA allows transform original variables, usually correlated, to new uncorrelated variables, making easier its interpretation. The goal of PCA is to find a subspace with dimension

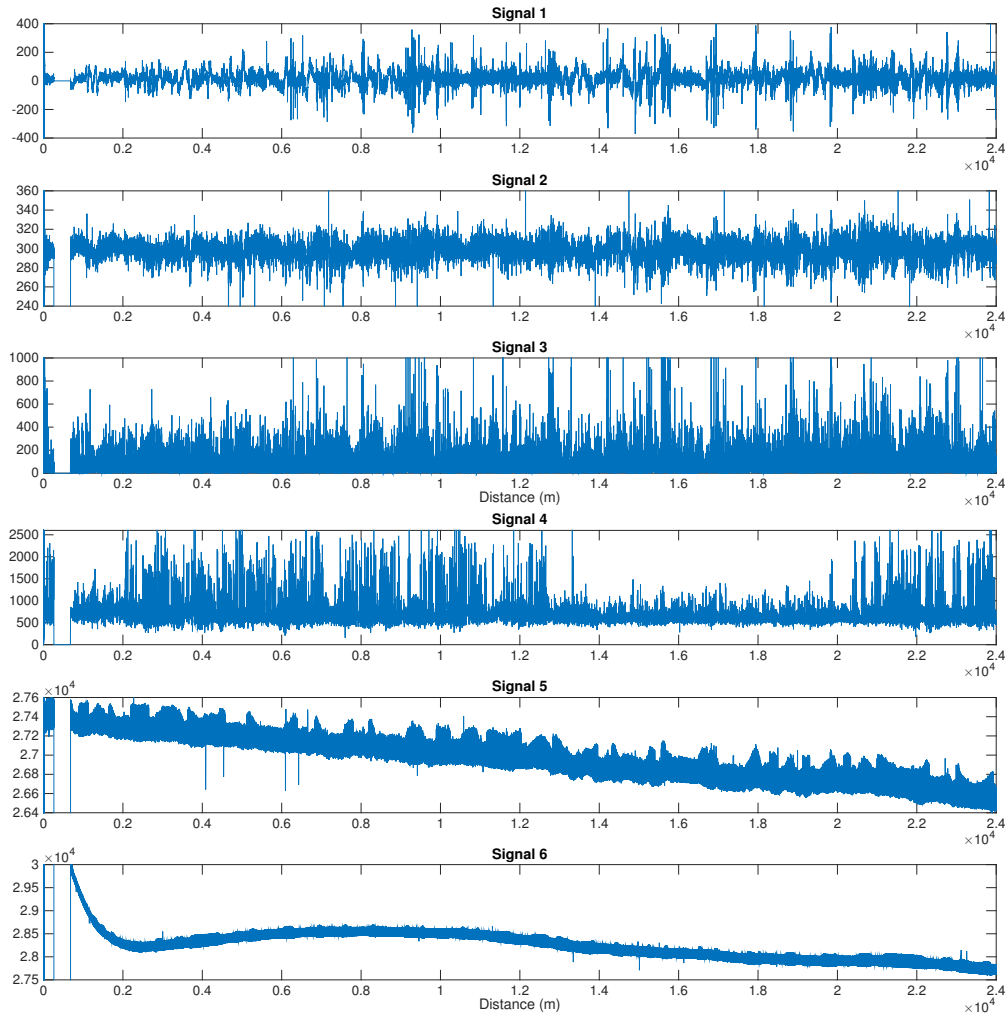


Figure 2: Original measurements detected by ITION

lesser than m such that projecting into it, the new variables keep its structure and minimize the distortion. In other words, a linear transformation orthogonal matrix \mathbf{P} , which is used to transform the original data matrix \mathbf{X} into the form

$$\mathbf{T} = \mathbf{X}\mathbf{P}. \quad (1)$$

In the literature, it can be found that the r -dimensional space ($r \leq m$) that represents better the original data is defined by the eigenvectors associated with the highest eigenvalues of the covariance matrix of the observations as follows:

$$\mathbf{C}_X = \frac{1}{n} \mathbf{X}^T \mathbf{X}, \quad (2)$$

$$\mathbf{C}_X \mathbf{P} = \mathbf{P} \mathbf{\Lambda}, \quad (3)$$

where \mathbf{C}_X is the covariance matrix of the original data \mathbf{X} , the eigenvectors of \mathbf{C}_X are the columns of \mathbf{P} , and the eigenvalues are the diagonal terms of $\mathbf{\Lambda}$ (the off-diagonal terms are zero). The eigenvectors p_j forming the transformation matrix \mathbf{P} (its columns) are sorted according to the eigenvalues by descending order, the eigenvector with the highest eigenvalue represents the

most important pattern in the data with the largest quantity of information. In Equation 1, Columns of \mathbf{P} are called the Principal Components (or loading matrix in other references) of the data set and \mathbf{T} the projected or transformed matrix to the principal component space (or score matrix in other articles).

In the full dimension case (using all the n principal components), this projection is invertible (since $\mathbf{P}\mathbf{P}^T = \mathbf{I}$) and the original data can be recovered as $\mathbf{X} = \mathbf{T}\mathbf{P}^T$. But, PCA also seeks to reduce the dimensionality of the data set \mathbf{X} by choosing only a reduced number r of principal components ($r < n$). Now, with \mathbf{T} given by the reduced matrix \mathbf{P} , it is not possible to fully recover \mathbf{X} , but \mathbf{T} can be projected back onto the original m -dimensional space and obtain another data matrix as follows:

$$\hat{\mathbf{X}} = \mathbf{T}\mathbf{P}^T. \quad (4)$$

Therefore, the original data matrix \mathbf{X} can be decomposed by the projected back data $\hat{\mathbf{X}}$ and the residual error matrix \mathbf{E} , which describes the variability not described by the model as follows:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}. \quad (5)$$

Two well-known statistics are commonly used to this aim: the Q -statistic (or SPE -statistic) and the Hotelling's T^2 -statistic (D -statistic). Q -statistic is based on analyzing the residual data matrix \mathbf{E} to represent the variability of the data projection in the residual subspace. It denotes the change of the events that are not explained by the model of principal components. The Q -statistic of the i -th sample or experiment (row vector x_i of data matrix \mathbf{X}) is defined as follows:

$$Q_i = e_i e_i^T = x_i (\mathbf{I} - \mathbf{P}\mathbf{P}^T) x_i^T. \quad (6)$$

where e_i is its projection into the residual subspace (row vector of residual data matrix \mathbf{X}).

T^2 -statistic is based in analyzing the score matrix \mathbf{T} to check the variability of the projected data in the new space of the principal components. The T^2 -statistic of the i -th sample (or experiment) is defined in the form:

$$T_i^2 = t_{si} \Lambda^{-1} t_{si}^T = x_i (\mathbf{P} \Lambda^{-1} \mathbf{P}^T) x_i^T. \quad (7)$$

where t_{si} is its projection into the new space (row vector of the score matrix \mathbf{T}) [7][9].

5. Results

5.1. Initial analysis and organization of the collected data

The six recorded signals, previously detailed in section 2.2 (See Figure 2), belong to a test in a buried pipeline of 20 inches diameter and 110 Km length. This Pipeline is made of sections of thin-walled steel tubing of 12 m length, which are welded together using a circumferential weld. It has been in service for more than 15 years and nowadays it transports gas. The owner of the structure provided the location of 58 tags across the pipeline (See Figure 3). These tags include elements of the pipeline (e.g. VA belongs to valves) and damages among others. Unfortunately no complete explanation can be given (due to confidence reasons) of what these events are. An initial analysis of the measurements is performed. The test is conducted in 24 Km of the structure, this means that the "smart pig" (ITION) traveled inside the pipeline 24 Km collecting data from 6 sensors. A total of 7'426.500 measurements (samples) were collected by each sensor. At the first 10 meters of the pipeline, during the tuning of the ITION, the sample frequency is higher than the rest of the structure (1'600.000 samples). On the other hand, in these profiles any event or damage can be directly observed. Since it is not possible to give some diagnosis of the structure observing directly the measurements, PCA is applied to carry out a multivariable

analysis, in other words, to analyze all measurements and its correlations as a whole [8]. In this way, possible patterns of strange events can be recognized (elements of the pipeline, damages, welds, etc). For this purpose, the original data are organized in a matrix. This $n \times m$ matrix contains information from m sensors and n experimental trials (samples). Consequently, each row vector represents measurements from all the sensors at a specific time instant or experiment trial. In the same way, each column vector represents measurements from one sensor in the whole set of experiment trials.

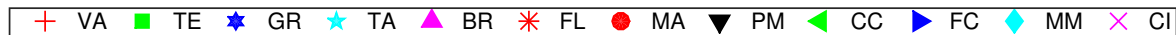


Figure 3: Tags provided by the owner of the pipeline

5.2. Multivariable analysis by using PCA

The original data ($7'426.500 \text{ samples} \times 6 \text{ sensors}$) is used to build the PCA model and compare with the tags provided by the pipeline owner. The data matrix is scaled, the loading and score matrices are calculated (\mathbf{P} and \mathbf{T} in Equation 1). Besides, statistical indices (Q and T^2) are calculated for each experiment (see Equations 6 and 7). However, the results for detecting tags are not entirely satisfactory. Going back to the initial analysis, the oversampling of the first 10 meters and, the profile of the signal 6 should be considered. The temperature of the transported product (signal 6) is irrelevant for the goal of the analysis: detection of tags. Therefore, signal 6 is removed from the original data matrix and three PCA models are built:

- Model 1 which uses all measurements ($7'426.500 \times 5$)
- Model 2 which uses the measurements of the first 10 meters of the pipeline ($1'600.000 \times 5$)
- Model 3 which uses the measurements of the rest of the pipeline ($5'826.500 \times 5$)

Each model is built with 4 principal components since around 90% of the cumulative variance is retained. Anew, scores and statistical indices (Q and T^2) are calculated and depicted for each model. Figure 4 shows scores 1 to 3 and the statistical indices for model 1. Apparently, the scores and statistical indices exhibit a local maximum (peak) in some tags. But a deeper validation confirmed that the best approximation is achieved by using Q and T^2 -statistics.

Model 2 describes the first 10 meters of the pipeline and the tuning of the ITION. 14 from the 58 tags (24%) are located in this section. In Figure 5, where the indices are depicted, it can be seen that some maximum match with some tags and others tags are nearby to be localized. 3 from the 9 tag named TA are well matched. VA, TE and GR are detected without any problem. In contrast, BR is not matched.

Model 3 describes the rest of the pipeline, 44 tags are localized in this section. According to Figure 6, a better tags localization is observed. First and third PM tag are detected, the second PM is detected only by score 1 (no shown). The set of tags around the same point are detected by at least one score or index: E.g. around the spot 0.2×10^4 meters (2 Km), tags are remarked by T^2 index; tags in 0.8×10^4 meters (8 Km) are detected by all scores and indices. In the interval from 14 Km to 18 Km it is observed some maximum values, but any tag is located there. Finally, in the last kilometers, all tags are correctly detected.

On the other hand, analyzing more in detail plots in Figure 6, it can be seen that indices increases every 12 meters (approximately). For instance, in Figure 7, that presents a zoomed section of Figure 6, a pattern can be identified every 12 meters. The reason of this "out of control" of these indices can be attributed to the weld that join the sections of the pipeline.

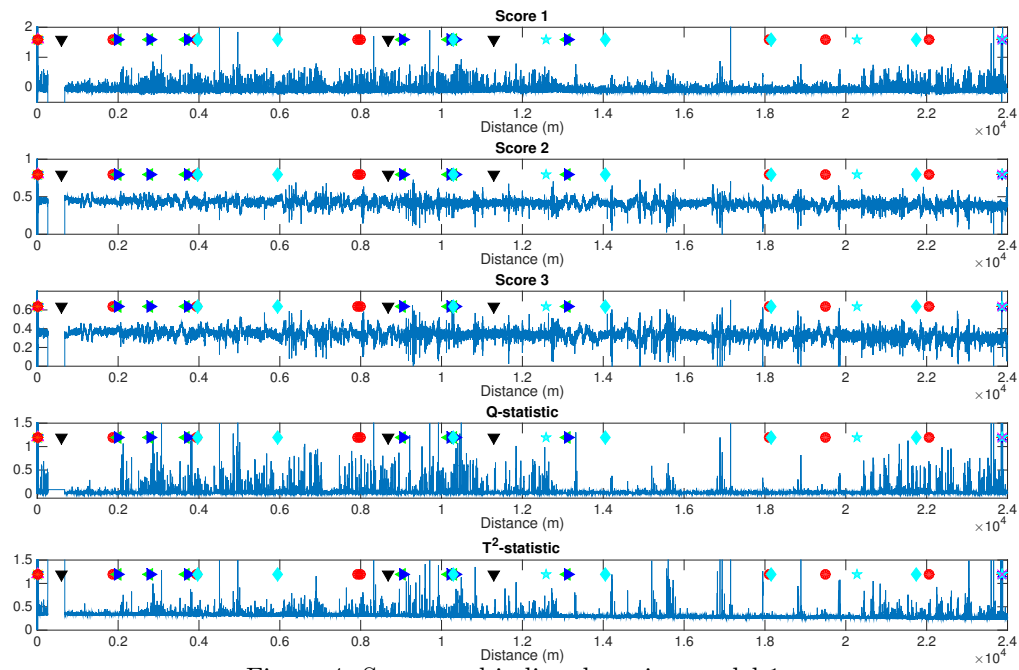
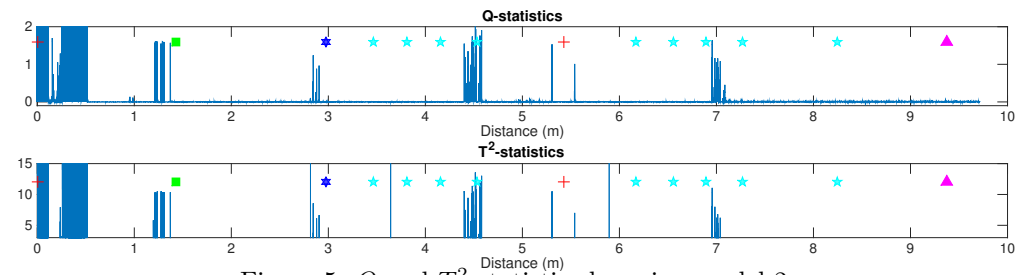
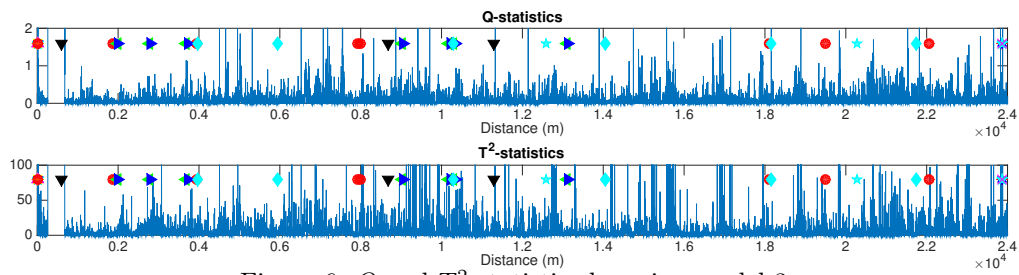
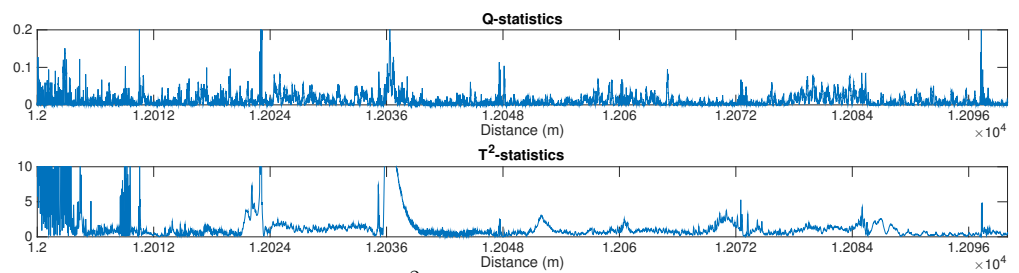


Figure 4: Scores and indices by using model 1

Figure 5: Q and T^2 statistics by using model 2Figure 6: Q and T^2 statistics by using model 3Figure 7: Q and T^2 statistics by using model 3 (Zoom)

6. Discussion and Conclusions

The main goal of this work is to localize critical loss of metal in a pipeline. The smart pig tool (ITION) and the methodology (based on PCA) have been validated in a pipeline (currently in service) made of tubing sections welded every 12 meters. The owner of the structure provided the location of 58 tags belong to different operational elements and damages of the pipeline, however any explanation of the meaning was given. A huge amount of measurements are gathered in the first run of the tool through 24 Km of the pipeline. This information is processed by means of PCA and some indices are calculated for every location of measurement. The validation of the methodology is carried out by comparing the location of the "alarms" or values out of control of the mentioned indices and, the location of the tags. From results, it is concluded that the localization of abnormal events (operational elements or damages) are improved considerably when data are arranged in two models (model 2 and 3). These results can be considered as successful (despite that it is a novelty detection real application) due to the pipeline complexity. Even though several possible false alarms are presented, it is inferred that the pipeline welds could be the responsibilities. In the near future, it is expected to run the second version of the smart pig (ITION) which includes 18 sensors and the sample frequency is higher. On the other hand, the methodology must be improved if the owner is not interested in detecting welds.

Acknowledgments

The authors would like to thank the support from the "Ministerio de Economía y Competitividad" in Spain through the Coordinated Project DPI2011-28033-C03-01, to the Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona, Consorci Escola Industrial de Barcelona and, the Colombian Research Institute of Corrosion - CIC-.

References

- [1] Russell E L, Chiang L H and Braatz R D 2000 *Data-driven techniques for fault detection and diagnosis in chemical processes "Advances in Industrial Control"*. London: Springer
- [2] Roa S, Parra J, Garnica H, Quintero S and Quintero M 2012 Design Optimization of a Magnetic Circuit and Sensor System for MFL Testing of Pipelines. *The 2012 Review of Progress in Quantitative Nondestructive Evaluation*, 2012.
- [3] Perez I and Dobmann G 2013 Surface Open Corrosive Wall Thinning Effects. *International Workshop on Smart Materials, Structures & SHM*
- [4] Jansen H J M, Van de Camp P B J and Geerdink M 1994 Magnetisation as a key parameter of magnetic flux leakage pig for pipeline inspection. *INSIGHT*, 36(9) pp 672-673
- [5] Zuoying H, Peiwen Q and Liang C 2006 3D FEM analysys in Magnetic Flux Leakage method. *NDT&E International*, 39 pp 61-66
- [6] Jolliffe I T 2002 *Principal Component Analysis*. Springer Series in Statistics. Springer-Verlag
- [7] Mujica L E, Rodellar J, Fernández A and Güemes A 2011 Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. *Struc. Health Monitoring* 10(5) pp 539-553
- [8] Mujica L E, Vehí J, Ruiz M, Verleysen M, Staszewski W J and Worden K 2008 Multivariate Statistics Process Control for Dimensionality Reduction in Structural Assessment. *Mechanical Systems and Signal Processing* 22 pp 155-171
- [9] Ruiz M, Mujica L E, Berjaga X and Rodellar J 2013 Partial least square projection to latent structures (PLS) regression to estimate impact localization in structures. *Smart materials and structures* 22(2) pp 025028