Leak Detection and Localization on Hydrocarbon Transportation Lines by Combining Real-time Transient Model and Multivariate Statistical Analysis

LUIS EDUARDO MUJICA, MAGDA RUIZ
and JUAN MANUEL MEJÍA

ABSTRACT

Safety and reliability of hydrocarbon transportation lines (pipelines) around the world represents a critical aspect for industry, operators and population. Lines failures caused by external agents, corrosion, inadequate designs, among others, generate impacts on population, environment, infrastructure and economy, besides it may be catastrophically. Therefore, it is essential to constantly monitor operating conditions and hydraulic lines to faults and thus to take measures to mitigate the failure.

Localization of leakage is more than comparison between simulated and measured flows, from the dynamic of these flows it can be inferred the localization of the leakage, and even its magnitude. One option is to develop an inverse Transient Model (TM) able to calculate parameters of the pipeline by using the measured flow. However, if the calculation of flows is computational expensive, the inverse calculation is even more. These phenomenological models reproduce as closely the response (flow and pressure) of the pipeline. The simulation contains information to optimize the pumping rate, the momentum and energy including a high number of inputs and constraints to consider that growing exponentially with the level of detail to get in the pipeline. Therefore, this method has a high computational cost. The other option is to simulate several scenarios by using TM and train some kind of classifier or predictor with the simulated measurements. The first phase of our complete proposed methodology under development is presented in this work. We have focused on carrying out simulations of pressure along a pipeline using TM and applying Principal Component Analysis (PCA) as a tool to recognize hided patterns which allow classify leakages in different locations and different magnitudes.

INTRODUCTION

Hydrocarbons started to play a prominent role in the global economy. Pipelines are usually employed for hydrocarbon transportation. The pipelines connect source and target units. These are pipes sequentially connected, buried on the terrain or over the surface.

L.E. Mujica and M. Ruiz, Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona. Consorci Escola Industrial de Barcelona. Universitat Politècnica de Catalunya. BARCELONATECH. Departament of Applied Mathematics III. luis.eduardo.mujica@upc.edu / magda.ruiz@upc.edu.
J.M. Mejia. Universidad Nacional de Colombia, Facultad de Minas. jmmejica@unal.edu.co.

Normally, the pipelines operate continuously or by batches. In these structures, it can be found a variety of faults due to the accumulation of sediments as wax or paraffin, leakage, rust and inappropriate designs due to the changing elevations along the line, among others. Hydrocarbons are volatile and flammable and any possible failure in these structures could be catastrophic for the population and environment, leading to severe businesses and structural losses. Therefore, it is essential to rely on fast and accuracy tools for detection and, more importantly, fault zone identification in the pipeline in order to proceed to control and mitigate the problem.

To establish a monitoring system feasible to transport the hydrocarbons through pipelines is not a simple task for any company. The problem is increasingly complex in order to optimize pumping and minimize costs associated with the operation while maximizing reliability. This type of study is relatively new (last decade), since the technology developed fifty years ago is still implemented. Nowadays, the market dynamics and environmental standards have demanded the application of more advanced techniques. These new techniques have proposed simulations of events with different work environments and common faults. Consequently, situations of high environmental and economic risk have begun to be quickly evaluated. However, currently there are no representation approaches that allow locate pipeline leakages efficiently and quickly [1].

Different leaks detection and localization technologies have been implemented in hydrocarbon transportation projects. These technologies can be based on [2]: (1) Regular or periodic monitoring of operational data, (2) Computational Pipeline monitoring, and (3) Data analysis method. For a given project, technology selection must be appraised on different criteria such as: cost, reliability, sensitivity, detection velocity, operational flexibility and ease of use and operation, among others. According to Wang et al. [3] no technology meets these criteria. However, real time transient modeling (RTTM) provides a good approach for the leak location identification [4]. RTTM is based on the inverse solution of the mass, momentum and energy transport equation of a fluid flowing through the pipeline. The numerical solution of the equations is compared to field data in order to evaluate the hydraulic condition. Many companies worldwide, such as Kuwait Oil Company, OMV, Adriatic-Wien Pipeline, Transalpine Pipeline, Saudi Aramco and Dow Chemical, among others, have implemented RTTM in their pipeline operation [5]. Since RTTM is based on the permanent numerical solution of a set of stiff, non-linear partial differential equations, the computational use is extremely high since the solution must satisfy the field data (inverse modeling) aiming to evaluate the hydraulic condition.

The computational limitation of RTTM can be bypassed using a hybrid approach presented in this paper. Here, we use a transient, phenomenological model that simulates different flow conditions of a given pipeline and the simulations results are used for the construction of a Multivariate Statistical Analysis based on Principal Component Analysis (PCA) [9]. RTTM describes the phenomenological mass transport, momentum and energy in pipelines. Then the flow, pressure, density and temperature along the pipeline can be obtained. PCA is a widely compress tool for feature extraction which maximize the variance and minimize the correlation among the variables. The goal of this work is to detect and localize leakages by means of simulations of the hydrocarbon flux in undamaged and damage conditions. To be more realistic, the mathematical model includes the dynamic behavior of the pressure in different locations for dead oil in horizontal topography.

TRANSIENT MODEL OF THE PIPELINE HYDRAULIC CONDITION

Leaks can be easily identified by monitoring pressure evolution between monitoring points. In a normal pipeline operation, if pressure decreases abnormally initially in one of the nodes then a leak is highly possible, and it is confirmed when pressure wave reaches the second node. The problem consists in identifying the position of the failure. RTTM uses the measured pressure wave evolution for matching a mathematical model of the pipeline with a leak. Then, an optimization algorithm is employed in order to match simulation results with field measurements. The procedure is time consuming because the numerical solution is complex and it must be systematically repeated until a decent matching is reached [6].

We propose the use of off-line simulations of the pipeline hydraulic behavior for different leaks sizes and locations, using a transient model. For a multiphase flow system, the one dimensional governing mass balance equation for the gas, liquid and drops are:

$$\frac{\partial}{\partial t}(\alpha_G \rho_G) = \frac{1}{A} \frac{\partial}{\partial x} [A \alpha_G \rho_G u_G] + \psi_G + G_G, \quad (1)$$

$$\frac{\partial}{\partial t}(\alpha_L \rho_L) = \frac{1}{A} \frac{\partial}{\partial x} [A \alpha_L \rho_L u_L] - \psi_G \frac{\alpha_L}{\alpha_L + \alpha_D} - \psi_e + \psi_d + G_L, \quad (2)$$

$$\frac{\partial}{\partial t}(\alpha_D \rho_D) = \frac{1}{A} \frac{\partial}{\partial x} [A \alpha_D \rho_D u_D] - \psi_G \frac{\alpha_L}{\alpha_L + \alpha_D} + \psi_e - \psi_d + G_D, \quad (3)$$

where x and t denotes x -direction and time, α_i , ρ_i , u_i , are the volume fraction, density and velocity of phase i^{th} , respectively [7]. Subscripts G , L , D and i denote gas, liquid, drops and interface, respectively. A is the pipe cross section, ψ_G is the mass transfer between phases, ψ_e is the drag rate, ψ_d is the drops deposition rate and G_i is the source/sink of phase i^{th} . Mass balance equations are coupled to the momentum equation for the gas/drops and liquid phases:

$$\begin{aligned} \frac{\partial}{\partial t}(\alpha_G \rho_G u_G + \alpha_D \rho_D u_D) = & -(\alpha_G + \alpha_D) \frac{\partial p}{\partial x} - \frac{1}{A} \frac{\partial}{\partial x} [A(\alpha_G \rho_G u_G^2 + \alpha_D \rho_D u_D^2)] \\ & - \lambda_G \frac{1}{2} \rho_G |u_G| u_G \frac{S_G}{4A} - \lambda_{in} \frac{1}{2} \rho_G |u_R| u_R \frac{S_{in}}{4A} \\ & + [\alpha_G \rho_G + \alpha_D \rho_D] g \cos \theta \\ & + \psi_G \frac{\alpha_L}{\alpha_L + \alpha_D} u_a + \psi_e u_{in} - \psi_d u_D \end{aligned} \quad (4)$$

$$\begin{aligned} \frac{\partial}{\partial t}(\alpha_L \rho_L u_L) = & -\alpha_L \frac{\partial p}{\partial x} - \frac{1}{A} \frac{\partial}{\partial x} [A \alpha_L \rho_L u_L^2] - \lambda_L \frac{1}{2} \rho_L |u_L| u_L \frac{S_L}{4A} \\ & - \lambda_{in} \frac{1}{2} \rho_G |u_R| u_R \frac{S_{in}}{4A} + \alpha_L \rho_L g \cos \theta + \psi_G \frac{\alpha_L}{\alpha_L + \alpha_D} u_a \\ & - \alpha_L d(\rho_L - \rho_G) g \frac{\partial \alpha_L}{\partial z} \sin \theta - \psi_e u_{in} + \psi_d u_D \end{aligned} \quad (5)$$

where p is the local pressure, g is the gravity constant, θ is the pipeline angle, S_L , S_G , S_{in} are wet perimeters for the liquid, gas and interface, respectively. u_a is the phase velocity u_R is the relative velocity between two phases and u_{in} is the interface velocity, close the liquid velocity [8].

The numerical solution of the Equations (1-5) provides the spatial distribution of phase pressure, velocity (and flow rate), hold-up and other relevant flow properties for a given fluid and pipeline configuration. In this work, four transient models of pipeline with 80Kms length are built: (i) on horizontal topography that transports natural gas (ii) on horizontal topography transporting heavy crude oil, (iii) on slopping hill transporting natural gas and, (iv) on slopping hill with heavy crude oil.

9 simulations are carried out by each model using undamaged condition. Models for natural gas employ around 3000 samples by each sensor and, models for heavy crude oil uses around 150 samples by each sensor. Those simulations belong to few days of operation where the measured variables are the pressure at 10 Km, 30 Km, 50 Km and 70 Km. Besides, 9 simulations are conducted by each model: three leakages (1 inch, 3 inches and 5 inches) in three different locations (20 Km, 40 Km and 60 Km), without considering the possible degradation of the pipeline for its use during the fault. In Table 1, it can be found a summary of the main parameters configured in the simulation.

TABLE I. PARAMETERS OF THE PIPELINE

Fluid	Heavy crude oil / Natural Gas
Pipe length (Km)	80
Pipe diameter (in)	22
Temperature (°F)	59
Environmetal pressure (psia)	14.7
Flow rate (MSTB/d)	60.4
Number of pressure gauges	4
Pressure gauges localization (Km)	10, 30, 50, 70
Leakages size (in)	1,3,5
Leakages localization (Km)	20, 40, 60
Undamaged simulations	9
Damage simulations	9

METHODOLOGY

A leak changes the hydraulics of the pipeline, and therefore changes the pressure or flow readings after some time. Local monitoring of pressure or flow at only one point can therefore provide simple leak detection. It is only useful in steady-state conditions, however, with the objective of classifying, locating or even identifying different kind of leaks, a uni-variate monitoring is not sufficient. The methodology that has been previously used by the authors for a multivariate analysis always include information related with the undamaged structure (baseline) to create a statistical model based on PCA. Afterwards, data collected by sensors when structure need to be assessed, are projected into the new space given by the PCA model. These projections provide information about how these new measurements are different to the baseline, therefore to know whether the structure still keep in pristine condition or not (damaged), how it has changed and, potentially to distinguish among different kind of damages [10-13]. In the current work, information of the "healthy" structure (no leaks) is given by simulation of the pressure in the different points previously mentioned when the system is operating in normal conditions considering all parameters included in the TM. This initial information is organized and arranged in the matrix $\mathbf{X}$ as shown in Equation 6 where the number of variables is given by the number of sensors ($m=4$).

The number of samples is given by the number of samples per simulation times the number of simulations (e.g. for one simulations of heavy crude oil $n = 152 \times 9 = 1368$). According to Equations 7 and 8, the statistical PCA model is calculated (Transformation matrix or loadings denoted by $\mathbf{P}$). Data acquired by simulations of the structure by each leak are organized in the same way, the new matrices $\mathbf{X}_{l1}$, $\mathbf{X}_{l2}$, ..., $\mathbf{X}_{l9}$, with dimension 162 samples $\times$ 4 sensors, they are projected into the PCA model previously calculated (Equation 9). Besides, statistical indices Q and T^2 -statistics are also determined (Equations 10 and 11). Finally, scatter plots of the first two scores and indices are depicted.

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{im} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nm} \end{pmatrix} = (\mathbf{v}_1 | \mathbf{v}_2 | \dots | \mathbf{v}_j | \dots | \mathbf{v}_m) \quad (6)$$

$$\mathbf{C}_X = \frac{1}{n-1} \mathbf{X}^T \mathbf{X} = \frac{1}{n-1} \begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \mathbf{v}_1^T \mathbf{v}_2 & \dots & \mathbf{v}_1^T \mathbf{v}_j & \dots & \mathbf{v}_1^T \mathbf{v}_m \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{v}_j^T \mathbf{v}_1 & \mathbf{v}_j^T \mathbf{v}_2 & \dots & \mathbf{v}_j^T \mathbf{v}_j & \dots & \mathbf{v}_j^T \mathbf{v}_m \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{v}_m^T \mathbf{v}_1 & \mathbf{v}_m^T \mathbf{v}_2 & \dots & \mathbf{v}_m^T \mathbf{v}_j & \dots & \mathbf{v}_m^T \mathbf{v}_m \end{pmatrix} \quad (7)$$

$$\mathbf{C}_X \mathbf{P} = \mathbf{P} \Lambda \quad (8)$$

$$\mathbf{T} = \mathbf{X} \mathbf{P} \quad (9)$$

$$Q_i = \tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i^T = \mathbf{x}_i (\mathbf{I} - \mathbf{P} \mathbf{P}^T) \mathbf{x}_i^T \quad (10)$$

$$T_i^2 = \mathbf{x}_i \mathbf{P} \Lambda^{-1} \mathbf{P}^T \mathbf{x}_i^T \quad (11)$$

where each row vector ($\mathbf{x}_i$) represents measurements from all the sensors at a specific time instant or experiment trial. In the same way, each column vector ($\mathbf{v}_j$) represents measurements from one sensor (one variable) in the whole set of experiment trials.

RESULTS

Firstly, the simulated measurements of the different sensors are analyzed to verify whether the leakage can be detected and identified. It can be seen from Figure 1 the measurements by each sensor before and after leaks start. Color and line belong to the location of the leaks (red dotted line to 20Km, blue dash dot line to 40Km and, green dashed line to 60Km). Shape belongs to the dimension of the leakage (square to 1in, circle to 3in and, diamond to 5in). First samples represent the transportation in normal operation. After that, the pressure is reduced according to the leakage. However, these profiles did not yield any relevant information about the location and dimension of the leakage. Only it can be observed the instant time when the leakages started.

After applying PCA to the baseline data (Pressure in the four sensors in normal operation) two principal components are selected and the confidence limit of 95% is determined. A scatter plot of the data simulated on the pristine pipeline and the different scenarios of leaks are depicted in Figures 2. The projection of each sample is represented by different shape and color as shown in Table 2.

Figure 1. Pressure in the four sensors before and after the leaks start. (a) Natural gas on horizontal topography, (b) Heavy crude oil on horizontal topography, (c) Natural gas on sloping hill topography and (d) Heavy crude oil on sloping hill topography.

TABLE II. MARKS USED FOR DIFFERENT PIPELINE CONDITIONS

	No leaks
◆	20Km 1in
■	20Km 3in
▲	20Km 5in
▼	40Km 1in
☆	40Km 3in
●	40Km 5in
◆	60Km 1in
■	60Km 3in
▲	60Km 5in

Figure 2. Projection of the measurements resulting from pipeline with different leaks into the the first two principal components. (a) Natural gas on horizontal topography, (b) Heavy crude oil on horizontal topography, (c) Natural gas on sloping hill topography and (d) Heavy crude oil on sloping hill topography.

When the pipeline is still operating without leaks, projections fall down into the limit of confidence (In Figure 2, these projections are located around the origin of coordinates). If we should be able to show how the projections are changing as the leaks are starting, we can see how the projections are moving away of the origin in a specific direction. Once the leak is stable, all the projections fall down in a specific region. Leaks in the same location take the same direction, but the stabilizing region is given by the magnitude (size of the crack). In this way it is clearly feasible a classification or even a localization of any leak considering simulation of some few leak scenarios.

This is clear, specifically for natural gas in both topographies. Since the complexity of the heavy crude oil behaviour, the number of samples in the simulations are lesser than the ones in natural gas (as was explained previously). Therefore, this effect is not visually obvious.

CONCLUSIONS

By projecting measurements into the PCA model, it can be seen that leaks in the same location take the same direction, but the stabilizing region is given by the magnitude (size of the leak). Additionally, it has been observed two possible errors in the simulation for natural gas with sloping hill topography: damage located at 20 Km with a leakage of 3 inches starts when leakage of 1 inch is finished for the same

location. Same behavior is observed for damage at 40 Km and a leakage of 3 inches. In conclusion, it is clearly feasible a classification or even a localization of any leak considering simulation of some leakage scenarios. The results have been successful and we can intuit a promising future by applying this methodology. Finally, we expect to validate the approach with data obtained from a real reservoir.

ACKNOWLEDGMENT

The authors thank to the “Escola Universitària d’Enginyeria Tècnica Industrial de Barcelona”, “Consorci Escola Industrial de Barcelona”, “Universidad Nacional de Colombia” and “Ministerio de Ciencia e Innovación” in Spain through the coordinated research project DPI2011-28033- C03-03

REFERENCES

1. Cafaro, D.C. 2009. *Programación óptima de operaciones en sistemas de transporte de combustibles múltiples a través de poliductos*. Universidad Nacional del Litoral, Facultad de Ingeniería Química, Santa Fe - Argentina, PhD. Thesis.
2. The U.S. Department of Transportation. 1997. Leak Detection Technology Study. H.R.5782.
3. Wang, X.J., Lambert, M., Simpson, A., Liggett, J. and Vitkovsky, J. (2005) ”Leak detection in pipelines using the damping of fluid transients”. *J Hydraul Eng*, 128 (7):697–711.
4. Hauge, E., Aamo, O.M. and Godhavn, J.M. (2009). “Computational Pipeline Monitoring”. *SPE projects, Facilities and Construction*.
5. http://www.psoilandgas.com/fileadmin/downloads/OIL_GAS/Publications/PSIpipelines_Brochure.pdf (26/11/2012)
6. Heskestad, K.L. (2005) “Field Data Analysis Using the Multiphase Simulation Tool OLGA 2000,” Norwegian University of Science and Technology.
7. Bendlk, K.H., Maine, D., Moe, R., Nuland, S. and Technology, E. (1991) “The Dynamic Two-Fluid Model OLGA : Theory and Application,” *SPE Prod. Eng.*, pp. 171–180, 1991.
8. Lakehal, D. “Advanced Simulation of Transient Multiphase Flow & Flow Assurance in the Oil & Gas Industry,” *Can. J. Chem. Eng.*, 9999: 1–14.
9. Jolliffe. I.T. (2002) *Principal Component Analysis*. Springer Series in Statistics. Springer-Verlag, second edition.
10. Mujica, L.E., Rodellar, J., Fernández, A. and Güemes, A. (2010) “Q-statistic and T2-statistic PCA-based measures for damage assessment in structures”. *Struc. Health Monitoring*, 10(5):539–553.
11. Mujica, L.E., Ruiz, M., Pozo, F., Rodellar, J. and Güemes, A. (2014) “A structural damage detection indicator based on principal component analysis and statistical hypothesis testing”. *Smart materials and structures*, 23(2):025014-1–025014-21.
12. 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*, 22:155–171.
13. 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):025028.