

# Leaks Detection in a Pipeline Using Artificial Neural Networks

Ignacio Barradas\*, Luis E. Garza, Ruben Morales-Menendez,  
and Adriana Vargas-Martínez

Tecnológico de Monterrey, Campus Monterrey  
Ave. Garza Sada 2501, 64,849, Monterrey, NL, México  
{A00780799, legarza, rmm, A00777924}@itesm.mx

**Abstract.** A system based on Artificial Neural Networks (*ANN*) is proposed to detect and diagnose multiple leaks in a pipeline leaks by recognizing the pattern of the flow using only two measurements. A nonlinear mathematical model of the pipeline is exploited for training, testing and validating the *ANN*-based system. This system was trained with tapped delays in order to include the system dynamics. Early results demonstrate the effectiveness of the approach in the detection and diagnosis of simultaneous multiple faults.

**Keywords:** Leak detection, Fault detection, Diagnosis, Artificial Neural Network.

## 1 Introduction

Distribution of fluids in pipelines must occur under safe and trustable conditions, because damages may be caused by environmental and weather conditions, as well as aging or pressure changes. Pipelines are design to support impacts or internal over pressure, but occasionally pressure surges may lead to line breaks and leaks. In some cases, pipelines are underground or in the sea depths. And to complicate the scenario even more, normally the fluids in transport do not operate under steady state conditions, which makes more difficult to perform a fault inspection. Additionally, small leaks are harder to detect, because they are a consequence of corrosion and aging in the pipeline.

There are three approaches for leak isolation have been proposed (internal, external and hybrid). The first approach (internal) is based on physical models, such as, mass and volume balance, pressure analysis and real-time dynamic models. The second approach (external) is implemented using hardware, such as, sensors with impedance or capacitance changes, optic fiber, gas sniffing, acoustic sensors, ground analysis or infrared image. And the third approach is hybrid methods, which are a mixture between internal and external approaches, for example: acoustic and pressure analysis with mass and volume balance.

---

\* Any mail concerning this paper, should be send to the author's mail.

Leak isolation is still affected by expensive, noisy and vague instrumentation, uncertainties of the analytical model, and the relation between the operating point and leaks magnitude. A practical requirement for an automatic supervision system must be to detect the precise leak location as soon as possible and with a minimal amount of instrumentation.

In [1] several technologies to solve leak location applying automatic leak finders in pipelines where flow and pressure head instrumentation can be implemented only in their extremes is developed. In [3], a bank of observers with fixed leak positions satisfy the leak isolation and detection only if the pipeline is divided in three sections and two leaks are induced. The approach presented in [4] is the design of a parametric model in steady state which reduces the search interval. Although, these methods assure leak detection and diagnosis, they require intense mathematical formulation and a wide knowledge of the process. By this, there is a special interest in the application of Artificial Neural Networks (*ANN*) for solving fault diagnosis problems because of their classification and function approximation capabilities. *ANN* approach is convenient when an analytical model is difficult to obtain. In addition, *ANN* are highly robust to noisy inputs and to missing or new input data. Additionally, because of its parallel structure, *ANN*-based systems can be implemented for real time applications.

Recently, *ANN*-based approaches have taken special attention. In [2], two *ANN* cascade architecture were proposed, demonstrating that it is possible to detect leaks in pipeline. This work did not consider transient response when leaks occur. [6] demonstrated that using a neural-fuzzy system in a water distribution system makes possible to detect and classify faults in pipelines. A drawback of this approach is that multiple meters and gauges are needed in order to obtain the required information. Also, [5] used a fuzzy classifier to detect leaks in pipelines; this method used the transient response of the fluid in the pipeline, which requires very precise and continuous measurements. Finally, [10] compared several approaches for this application versus *ANN*-based systems.

This paper presents a method for detecting and isolating leaks in a pipeline. This method uses an *ANN*-based approach that recognizes the flow pattern using only two measurements. A mathematical model was proposed based on experimental data<sup>1</sup>.

The paper is organized as follows. In Section 2 the pipeline model is described. Section 3 presents the proposed scheme. In section 4 the testing procedure that validates the approach and results are shown. And finally, Section 5 concludes the paper.

## 2 Pipeline Model with Leaks

In [7],[8], [9] and [11] the following mathematical model was introduced. The model was also validated with experimental data. The dynamic of the fluid through the pipeline is given by:

---

<sup>1</sup> Thanks C. Verde because her support with experimental data.

$$\frac{\partial Q}{\partial t} + gA \frac{\partial H}{\partial z} + \mu|Q|Q = 0 \quad b^2 \frac{\partial Q}{\partial z} + gA \frac{\partial H}{\partial t} = 0 \tag{1}$$

where  $H$  is the pressure head ( $m$ ),  $Q$  is the flow ( $m^3/s$ ),  $z$  is the length coordinate ( $m$ ),  $t$  is the time coordinate ( $s$ ),  $g$  is the acceleration of the gravity ( $m^2/s$ ),  $A$  is the cross-section area ( $m^2$ ),  $D$  is the pipeline diameter ( $m$ ),  $b$  is the speed of sound ( $m/s$ ), and  $\mu = f/2DA$  where  $f$  is the *Darcy-Weissbach* friction coefficient.

A leak in point  $z_l$  will cause a discontinuity in equations (1)  $Q|_{z_l} = \lambda_i \sqrt{H|_{z_l}}$  where  $\lambda_i > 0$  is a function of the orifice area and discharge coefficient [12]. Because of this, a pipeline with  $n - 1$  leaks will be described by  $n$  pairs of differential equations, similar to equations (1) with a frontier condition between each pipeline segment given by:

$$Q^b|_{z_l} = Q^a|_{z_l} + Q|_{z_l} \tag{2}$$

where  $Q^b|_{z_l}$  and  $Q^a|_{z_l}$  are the flows before and after the leak. Having a pipeline of length  $L$  and assuming that the leaks are equally distributed along the space  $z$ , which can be divided in  $n$  segments of length  $\Delta_z = L/n$ . It is possible to approximate the partial derivatives of the pressure and flow with respect to the spatial variable  $z$  as follows:

$$\frac{\partial H}{\partial z} \cong \frac{H_{i+1} - H_i}{\Delta z} \quad \frac{\partial Q}{\partial z} \cong \frac{Q_i - Q_{i-1}}{\Delta z} \tag{3}$$

where, the index  $i$  is associated with the variables at the beginning of the section  $i$ , and the frontier condition for each section is described by:

$$Q_{i+1} = \lambda_i \sqrt{H_{i+1}} \tag{4}$$

Knowing that the frontier conditions are characterized by the pressure  $H_{r_i}$  and  $H_{r_o}$ , at the beginning and the end of the pipeline, and substituting equations (3) in equations (1), the model could be described as a set of  $n$  coupled nonlinear equations given by:

$$\frac{\partial Q_i}{\partial t} = a_1(H_i - H_{i+1}) - \mu|Q_i|Q_i \quad \frac{\partial H_i}{\partial t} = a_2(Q_{i-1} - Q_i) - (\lambda_{i-1} \sqrt{H_i})u_{t_i} \tag{5}$$

with  $H_1 = H_{r_i}$  and  $H_{n+1} = H_{r_o}$  as system inputs, and parametric constants  $a_1 = g\pi r^2 n/L$  and  $a_2 = b^2 L/g\pi r^2 n$  with  $n = 4$ ,  $u_{t_i} = u(t - t_i)$  is the unit step function associated with the occurrence time  $t_i$  of the leak  $i$ .

If leaks are not equally distributed,  $\Delta z$  is not constant and parameters  $a_1$  and  $a_2$  are function of the distance between leaks. For this study case,  $H_1$  and  $H_5$  are constant pressures heads at the input and output of the pipeline, while flows  $Q_1$ , and  $Q_4$  are the measurable flows at the extremes of the pipeline. For this study, only three leaks were considered each one at the frontier condition, Fig. 1.

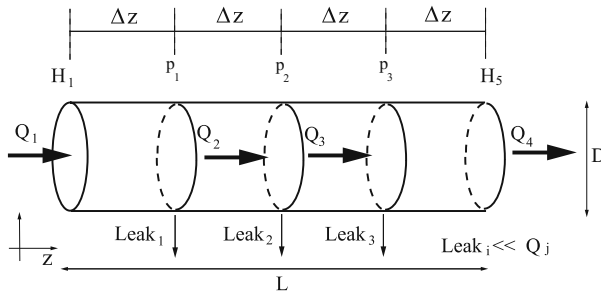


Fig. 1. Pipeline discrete model [10]

### 3 Proposed Scheme

Fig. 2 shows the implemented scheme. The scheme consists of an ANN that detects the leak and its location in the pipeline. Tapped delays signal from  $Q_1$  (flow measurement at the inlet) and  $Q_4$  (flow measurement at the outlet) were introduced as inputs of the ANN in order to include the system dynamics. The above resulted as an improvement of the ANN performance.

Many ANN configurations were tested. All of them were feed-forward multi-layer architecture. The classical back-propagation algorithm was used for the learning step. The basic differences in each ANN configuration were the number of neurons and layers; however, the number of delays in the input signals took the highest impact in the results. The input layer neurons use a tan-sigmoid activation function and the output layer neurons use a log-sigmoid function.

The leak detector is mainly based on the ANN performance. The ANN uses only the inlet/outlet flow measurements. The detector system identifies the possible pipeline operating states. And based on the state, the leak can be detected.

The ANN output will generate a leak signature according to Table 1, and it will be translated into an operating state by the state codifier. This codifier is based on simple logic rules, which will assign a state due to the outputs generated by the ANN.

The number of operating states that the codifier can estimate is given by the number of sections in which the pipeline is segmented. For this case, the pipeline was split in three segments; therefore, there are eight operating states, Table 1.

Table 1. Operating states of the pipeline

State	Activated Leaks	$f_1$	$f_2$	$f_3$	State	Activated Leaks	$f_1$	$f_2$	$f_3$
1	No leaks	0	0	0	5	1 and 2	1	1	0
2	1	1	0	0	6	1 and 3	1	0	1
3	2	0	1	0	7	2 and 3	0	1	1
4	3	0	0	1	8	1, 2 and 3	1	1	1

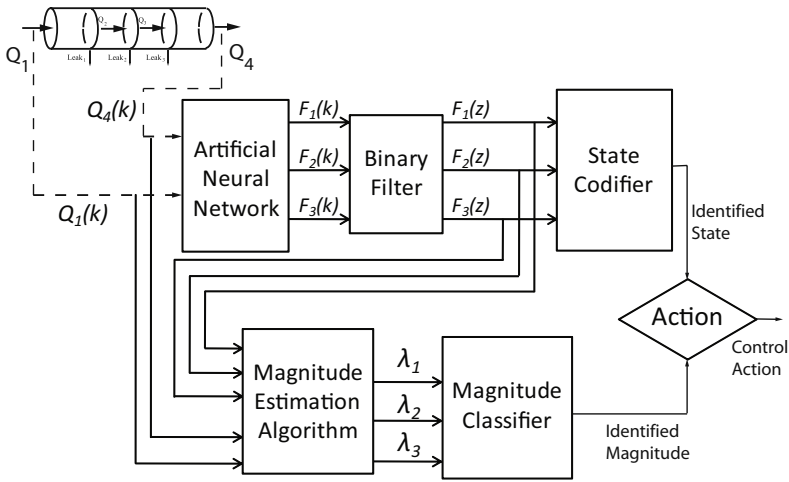


Fig. 2. Detection and diagnosis scheme

It is important to notice that the outputs of the ANN are activated by a hyperbolic tangent function; this means that the output can take a value between 0 and 1. A binary fault signature is needed, therefore a filter is included.

After estimating the operating state, the segment or segments of the pipeline in which the fault occurs are found. The accuracy of this approach will depend on the number of segments used in the pipeline.

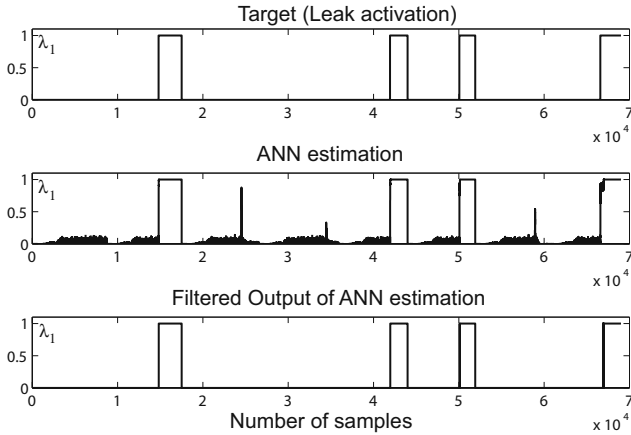
The ANN was trained with information about all possible leaks. All the possible state transitions are included in the training step based on a Markov chain simulation. The input data for the training step were the inlet/outlet flows and their delays; the output data was the leak signature. It is important to mention, that a better ANN training is possible if variations in the discharge coefficients are introduced in the generation of the training data set.

## 4 Results

The proposal approach was validated with 4 tests. *Test-1* was the introduction of never seen-before data to the ANN. *Test-2* consisted in adding noise ( $N$ ) to the flow signals. *Test-3* corresponds of two experiments for testing the robustness. First, the nominal pressure was changed from  $H_1 = 11$  m and  $H_5 = 5$  m to  $H_1 = 14$  m and  $H_5 = 8$  m; second, the input data was generated with variations in the value of the discharge coefficients ( $\lambda$ ). Finally, in *Test-4* the ANN was re-trained with a training data set that includes variations in the discharge coefficient.

The performance index corresponds to the error generated between the real states of the pipeline (multiple possible scenarios generated by the simulator) and the estimated states computed by the ANN.

Fig. 3 displays three plots, all of them represent the activation of leak 1. Top plot correspond to the real operating states, middle plot shows the estimated

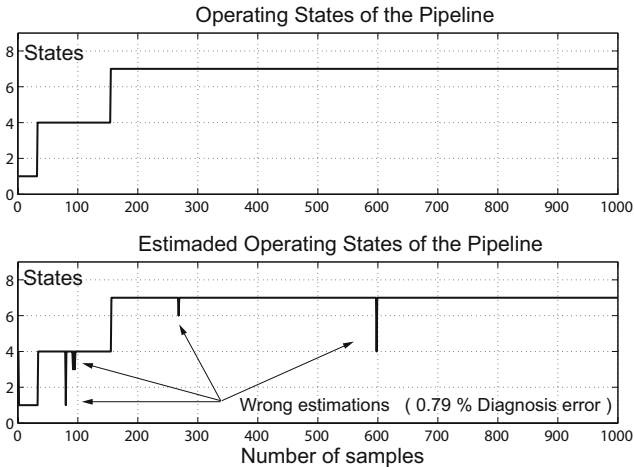


**Fig. 3.** Outputs from the ANN. Real state of pipeline; ANN output; filtered ANN output.

states, and the bottom plot represent the filtered states. This last signal plot is the one entering to the state codifier; and in combination with the other two leak signals the operating state of the pipeline is determined.

Fig. 4 shows the operating states estimations. This results proof that the proposal approach gives acceptable predictions of the real conditions in the pipeline.

Table 2 summaries the results. This table shows the error in the training process, and then the ANN configuration in each experiment depends on the flow signal delays. As it can be seen, in *Test-2* two different noise levels are added, 0.01% and 0.015%. As shown in the Table, the ANN that considers more delays



**Fig. 4.** State Estimation by the System. Real state of pipeline; ANN prediction.

**Table 2.** Summary of the results obtained after the experimentation

ANN	Error	Network Architecture	Test 1	Test 2		Test 3		Test 4
				$N = 0.01\%$	$N = 0.015\%$	Set point	$\lambda(\%)$	
$ANN_1$	0.00218	2-10-10-3	1.37%	1.37%	-	3.39%	27	4.53%
$ANN_2$	$4.9 \times 10^{-8}$	4-4-4-3	0.4%	3.8%	6.39%	3.79%	3.09	-
$ANN_3$	$4.4 \times 10^{-8}$	6-6-6-3	0.0001%	0.8%	2.4%	1.4%	12.7	6.13%

a better performance. As expected, having more information about the dynamic behaviour helps the ANN to recognize the pattern, even with a noisy signal. In *Test-3*, two validation experiments were conducted; first the head pressures were modified (set point), and second the size of the leaks were changed. These modifications impact directly the transient response of the flow (a new pattern is generated), as it can be seen in Table 2; even though, the performance of the ANN was acceptable.

The diagnosis error index (wrong estimated states/ total real states) was used as a main indicator of the performance of each ANN. It is important to notice that  $ANN_2$  and  $ANN_3$  used as inputs tapped delays.  $ANN_2$  used one delay, while  $ANN_3$  used 2 delays for each input signal.

[2] proposed a similar scheme for leak detection. In that research multiple sensors information was used in order to detect one and two leaks. In the present work similar results were obtained using only 2 measurements: inlet and outlet flow signals. The utilization of delayed signal allowed the algorithm to learn from the transient response, and therefore it was not necessary to use information of intermediate points in the pipeline.

It is important to mention that using this model allows the simulator to divide the pipeline in many segments. If there exist more segments, more data and operating states will be created as a consequence and therefore the accuracy of the leak location will be improved.

## 5 Conclusions

The main contribution of this work is that proves to be possible to estimate the location of the leak or leaks by only measuring the inlet and outlet flow, disregarding the pressure measurements and the size of the leak needed in [2], this can be observed in Fig. 3, and in Table 2, in which the ANN errors are 0.00218,  $4.9 \times 10^{-8}$  and  $4.4 \times 10^{-8}$  for the  $ANN_1$ ,  $ANN_2$  and  $ANN_3$ , respectively.

Using tapped delays as inputs of the network demonstrated to improve the ANN estimations, because the network is capable of learning the flow’s dynamic behavior. And this dynamic makes possible to identify the pattern and the different transitions between the operational states. In addition, it can be seen how the ANN using tapped delays has a smaller detection error, and also are more efficient in computer cost; require less training time and achieve a better training error value. Also, it is important to notice that the  $ANN_2$  and  $ANN_3$  were more sensitive to the discharge coefficients variation, due to the fact that they learn the transient response of the fluid dynamic (Table 2).

The training data set in which the network is trained is decisive. In order to generate a robust network, choosing a set that includes as many operational scenarios as possible is necessary. Therefore an acceptable detection scheme would be created. This can be observed in test 4, in which a better design set of training data were introduced to the networks.

## References

1. Ashton, S.A., Shields, D.N., Daley, S.: Fault Detection in Pipelines using Nonlinear Observers. In: UKACC Int. Conf. on Control IEE Conf., vol. 455, pp. 135–140 (1998)
2. Belsito, S., Lombardi, P., Andreussi, P., Banerjee, S.: Leak Detection in Liquefied Gas Pipelines by Artificial Neural Networks. *AIChE J.* 44(12), 2675–2688 (1998)
3. Billman, L., Issermann, R.: Leak Detection Methods for Pipelines. *Automática* 23(3), 381–385 (1987)
4. Blanke, M., Frei, Ch., Kraus, F., Patton, R.J., Staroswiecki, M.: What is Fault Tolerant Control? *SafeProcess* 35, 123–126 (2000)
5. Crowther, W.J., Edge, K.A., Burrows, C.R., Atkinson, R.M., Wollons, D.J.: Fault Diagnosis of a Hydraulic Actuator Circuit using Neural Networks an Output Vector Space Classification approach. *Proc. Inst. Mech. Eng. Part I: J. Syst. Control Eng.* 212(11), 57–68 (1998)
6. Izquierdo, J., López, P.A., Martínez, F.J., Pérez, R.: Fault Detection in Water Supply Systems using Hybrid Modelling. *Mathematical and Computer Modelling* 46, 341–350 (2007)
7. Verde, C.: Multi-leak Detection and Isolation in Fluid Pipelines. *Control Eng. Practice* 9, 673–682 (2001)
8. Verde, C.: Accommodation of Multi-leaks Positions in a Pipeline. *Control Eng. Practice* 13, 1071–1078 (2005)
9. Verde, C., Visairo, N., Gentil, S.: Two Leaks Isolation in a Pipeline by Transient Response. *Advances in Water Resources* 30, 1711–1721 (2007)
10. Verde, C., Morales-Menendez, R., Garza, L.E., De La Fuente, O., Vargas-Martínez, A., Velasquez, P., Aparicio, C., Rea, C.: Multi-Leak Diagnosis in Pipelines - A Comparison of Approaches. In: Special Session of the Mexican Int. Conf. on Artificial Intelligence, pp. 352–357 (2008)
11. Visairo: Detección y Localización de Fugas en un Ducto, PhD Thesis, SEP-CENIDET, México (2004)
12. Zhidkova, M.A.: Gas Transportation in Pipelines. In: Internal report written in Russian. *Naukov, Dumka, USSR* (1973)