

LEAK DETECTION USING A FUZZY SYSTEM

Henrique Ventura da Silva
Petrobras/Refinaria de Paulínia
hventura@petrobras.com.br

Celso Kazuyuki Morooka
UNICAMP/FEM/DEP
C.P. 6052, Campinas, São Paulo, Brasil
morooka@dep.fem.unicamp.br

Ivan Rizzo Guilherme
UNESP/DEMAC/IGCE
Rio Claro, São Paulo, Brasil
ivan@rc.unesp.br

José Ricardo Pelaquim Mendes
UNICAMP/FEM/DEP
C.P. 6052, Campinas, São Paulo, Brasil
jricardo@dep.fem.unicamp.br

Tiago Cardoso da Fonseca
UNICAMP/FEM/DEP
C.P. 6052, Campinas, São Paulo, Brasil
tiago@dep.fem.unicamp.br

Abstract. *A methodology for pipeline leakage detection using a combination of clustering and classification tools for fault detection is here presented. A fuzzy system is used to classify the running mode and identify the operational and process transients. The relationship between these transients and the mass balance deviation are discussed. This strategy allows a better identification of the leakage because the thresholds are adjusted by the fuzzy system as a function of the running mode and the classified transient level. The fuzzy system is initially off-line trained with modified data set including simulated leakages. The methodology is applied to a small-scale LGP pipeline monitoring case where portability, robustness and reliability are amongst the most important criteria for the detection system. The results are very encouraging with relatively low levels of false alarms and obtaining an increased leakage detection with low computational costs.*

Keywords. *Classifiers, Pipeline leakage detection, Pattern recognition, Fuzzy Systems.*

1. Introduction

Pipeline is an efficient and economic transportation means for petroleum products. However, risks associated with accidental releases of transported product are still high (Costa, 2001). This issue has motivated the development of many methods for leak detection, mainly based on process variables, i.e., pressure, flowrate and temperature, such as the volume balance method (Ellul, 1989), or (Stouffs and Giot, 1993), where the importance of packing term in the transient flow is highlighted.

In the present paper, the high correlation between the inlet-outlet flowrate deviation and the operational transients is shown which is the important fact to ere applied define the fault detection strategy.

The applied strategy consists, at first, in the development of a classifier module that can identify the operational and process transients and determine the current stage of the transfer process. Then, the output of this module is used by a Fault Detection module that will evaluate the inlet-outlet flowrate deviation in order to detect a leak or an abnormal operation condition, with a low level of spurious alarms.

For this development real data collected at every 10 seconds from a small LGP pipeline is used. The pipeline has 8 inch diameter and 2,000 meters of extension with pressure, temperature and flowrate transmitters installed in its extremities. For tests, this database was evaluated by an expert. After having been modified for abnormal situations simulation, each stage of the transfer process and the in-out flow deviation was classified.

A Fuzzy Inference System is used to solve the present problem by using a rule-base system developed from this database. The system was evaluated by a new data collected from the same process, and good results has been obtained; with increased leakage or abnormal situation detected. The low computational costs involved and low level of spurious alarms obtained are the most attractive items in the present system.

2. Process Description

The petroleum products produced by a refinery are spread to distribution companies by pipelines. The Measuring Station (EMED), basically composes the control system that transfers petroleum derivatives to the buying companies. In general, main process variables arriving from the EMED, such as pressure, temperature, flow and density, are usually available in real time. In the destination, total flow, pressure and sometimes temperature are measured again. Figure (1) shows an scheme of transfer system and instrumentation available.

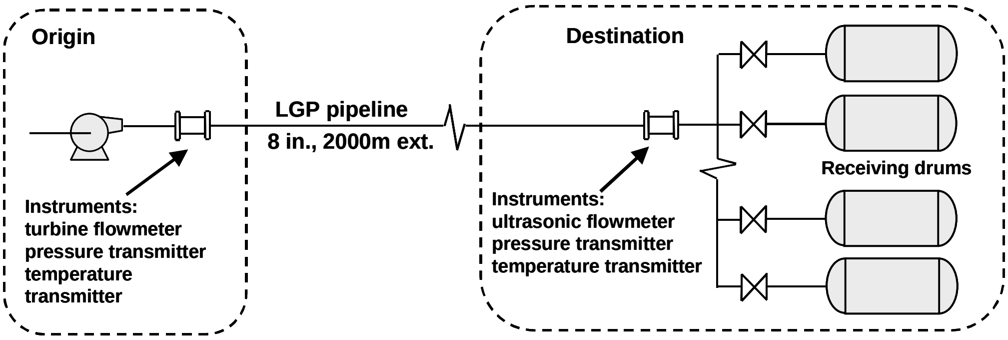

Figure 1. Petroleum derivatives transference system and monitoring instrumentation.

The present paper focus in the monitoring of a LPG (Liquefied Petroleum Gas) transference process, where often operational transients arouse larger complexity for the transference process. During this transference process, the pressure gradually rises while the LPG receiving drum is filled. When the LPG drum is completely full, then transference process is switched to a new drum. At that moment, a sudden expansion is observed and an increase in the flowrate happens. During the drum filling process (steady state flow), there is only a small deviation between the total flow measured in the origin and in the destination of the transference. The deviation is expected following mass balance model, and it is generated by the inherent uncertainties associated to the measuring process (Sattary, J.A., 1995). However, during the operational transient related to the receiving vessel switch procedure, the deviation here observed rises to significant values, which is mainly motivated by the line pack effect accounted by the mass balance model, due to diverse responses from measuring devices and by eventual lack of synchronism in the data acquisition system.

Modeling these transients through deterministic methods is a rather difficult task. The methods based on Fuzzy Logic are here highlighted in solving these problems (Taillefond, 2002). In the next sections, the system will be modeled and the correlations between data captured during distinct operational stages, which will support the Fuzzy System architecture and fault detection module development, will be analyzed.

3 Correlation Modeling and Evaluation

The mass conservation model states that any difference between the mass flowing in and out of a pipe, in a given time interval, must be analyzed as a function of the mass variation inside the pipe during this time interval. This mass variation is denominated line pack. If there is no leakage, the general equation might be presented as the function of the mass flows as shown in below:

$$(Q_o - Q_d)dt = dLP \tag{1}$$

where: $Q_o =$ Volumetric flow measured in the pipeline's origin;
$Q_d =$ Volumetric flow measured in the pipeline's destination;
$dLP =$ Line pack during one measuring cycle interval.

Adding the uncertainty of the measuring devices, it can be rewritten as follows:

$$(Q_o - Q_d) = \frac{dLP}{dt} + \varepsilon \tag{2}$$

where: $\varepsilon =$ flow measuring devices uncertainty.

Assuming no leakage, the following can be concluded from Eq. (2) above:

- in steady state flow, the difference between the origin and the destination flows is equal to the measuring devices' uncertainty, and;
- during operational transients, the line pack is added to the measuring devices' uncertainty.

Figure (2) shows the typical behavior of different parameters in a LPG transference, where (a) flow, (b) pressure and (c) deviation between origin and destination flow are depicture. Often operational transients in this process, occur during the receiving drum switch procedure, and increased deviation is measured between the measured flowrates during these operations. And, it is emphasized in the present study.

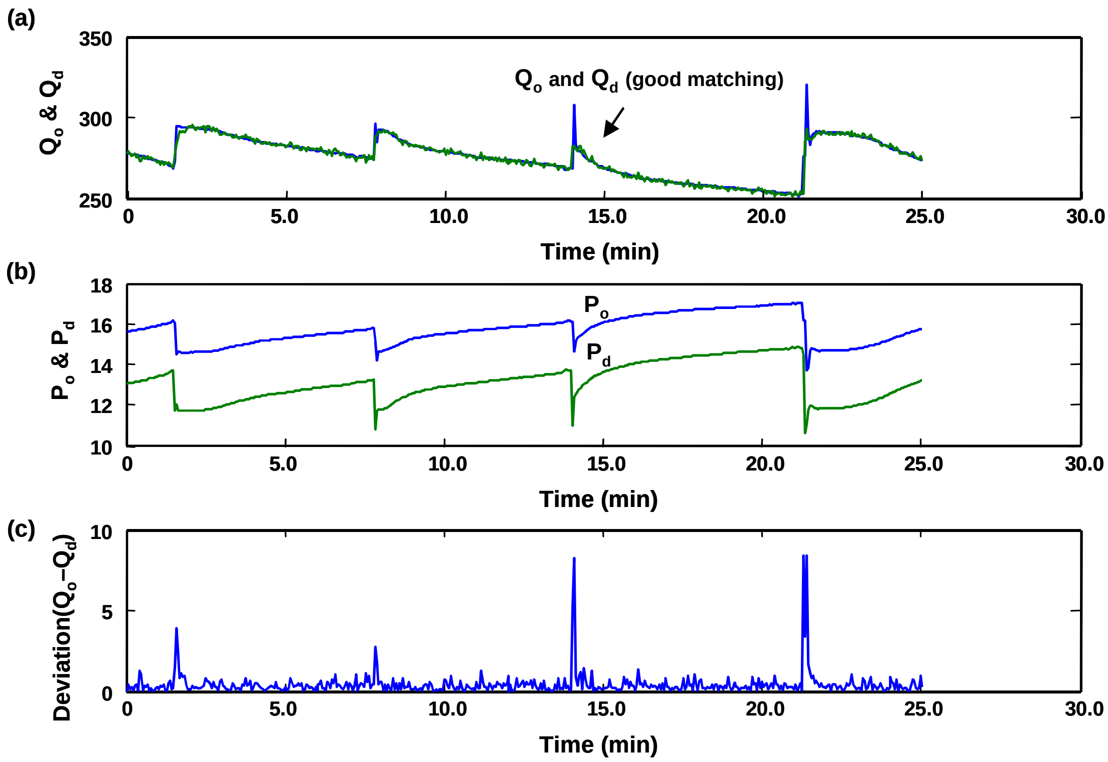

Figure 2. Typical behavior of an LPG transference in terms of (a) *flow*, (b) *pressure* and (c) *deviation*.

Figure (3) shows the detailed behavior of these variables during a drum switching operation. The hydraulic unbalancing and differences between the flow measuring devices' responses, in the origin and in the destination (turbine and ultra-sonic, respectively), are emphasized.

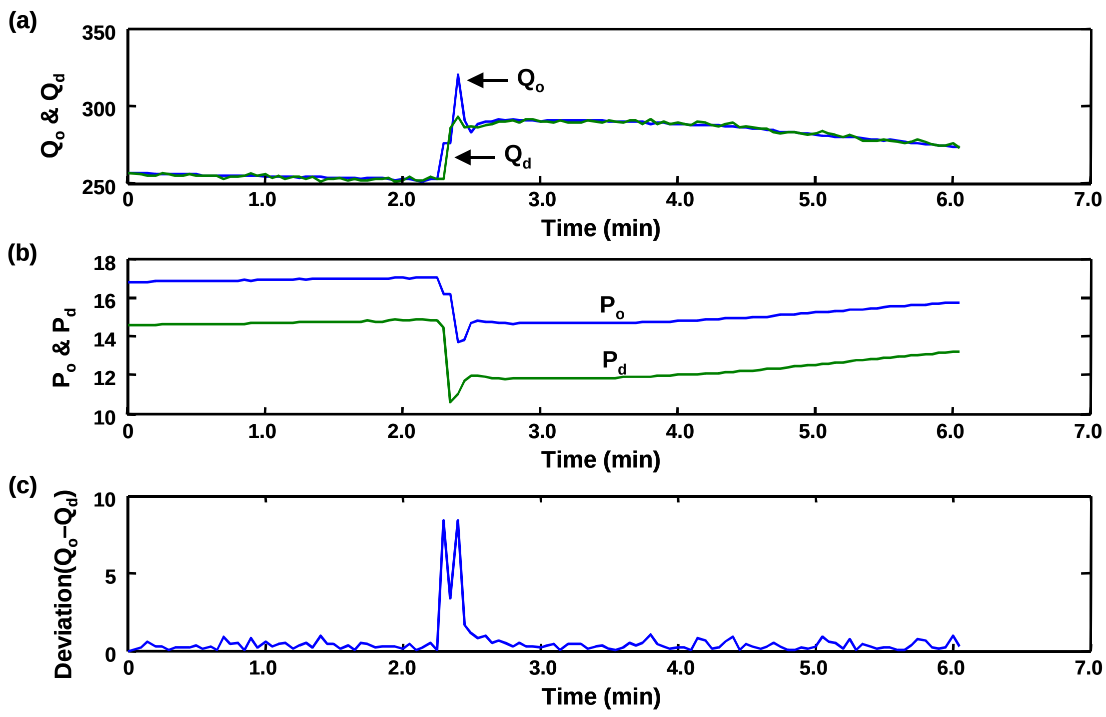

Figure 3. Detailed behavior of (a) *flow*, (b) *pressure* and (c) *deviation* during the switch operation.

In a conventional pipeline leakage detection system based on the mass balance model, if the above mentioned transient situation is not treated in an adequate manner, it usually generates a large number of false alarms (Moura, 2001). Due to this problem, some variables, capable of identifying the casual operational transients, can be redefined as presented in Eq. (3), (4) and (5).

Transient measured through average volumetric flow (*Transqm*):

$$Transqm(t) = abs\left(\frac{Q_m(t) - Q_m(t-1)}{\Delta t}\right) = abs\left(\frac{Q_o(t) + Q_d(t) - Q_o(t-1) - Q_d(t-1)}{2\Delta t}\right) \tag{3}$$

Transient measured through the origin-destination differential pressure variation (*Transdp*):

$$Transdp(t) = abs\left(\frac{(P_o(t) - P_d(t)) - (P_o(t-1) - P_d(t-1))}{\Delta t}\right) \tag{4}$$

Transient measured through the modified hydraulic coefficient variation (*Transcoef*):

$$Transcoef(t) = abs\left[\frac{\left(\frac{Q_o(t) - Q_d(t)}{(P_o(t) - P_d(t))^2}\right) - \left(\frac{Q_o(t-1) - Q_d(t-1)}{(P_o(t-1) - P_d(t-1))^2}\right)}{\Delta t}\right] \tag{5}$$

From the variables defined above, the correlation between the temporal series (*Deviation x Transcoef*; *Deviation x Transdp* and *Deviation x Transqm*) is found. The correlation is thus defined as in Eq. (6):

$$Corr_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X . \sigma_Y} \tag{6}$$

where: $\sigma_X^2 = \frac{1}{n}\sum_1^n (X_i - \mu_X)^2$ and $\sigma_Y^2 = \frac{1}{n}\sum_1^n (Y_i - \mu_Y)^2$

The result is shown in Fig. (4), using the same data as in Fig. (2).

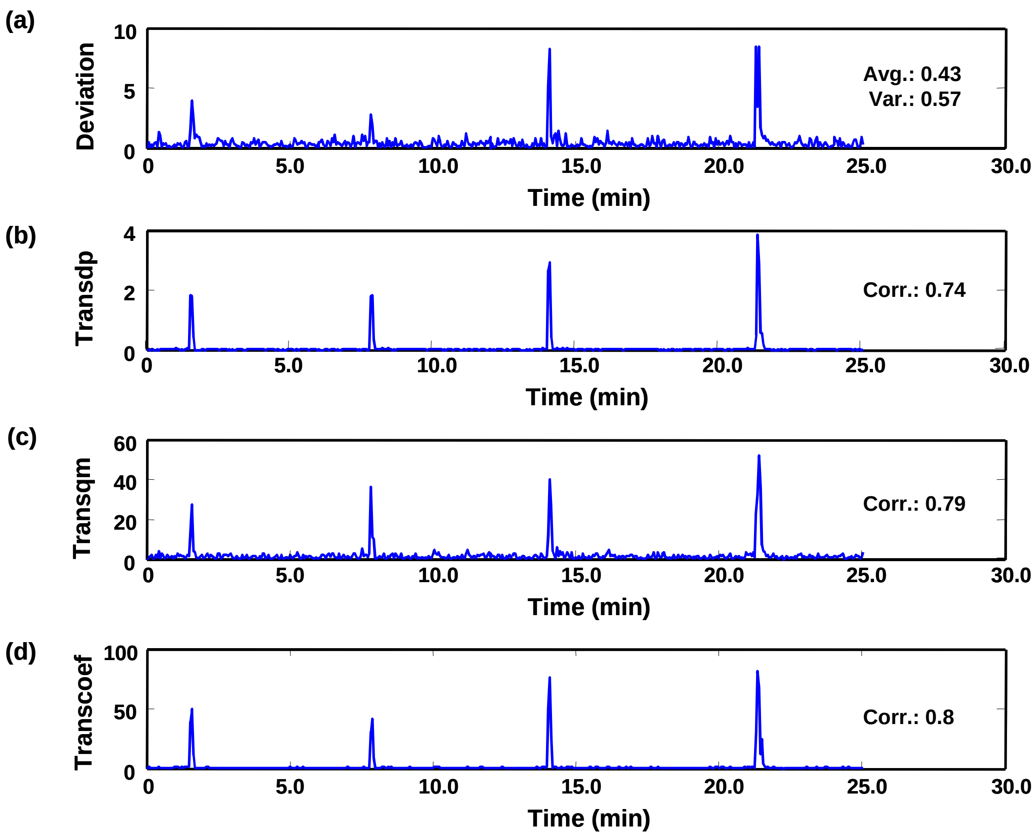

Figure 4. (a) *Deviation* and variables capable of identifying the casual operational transients, along with the correlation between these variables and the deviation: (b) *Transdq*, (c) *Transqm* and (d) *Transcoef*.

As the correlation is relatively high, around 0.8, the deviation can be associated to any variable that represents a process transient. It should be highlighted that the correlation is computed through the series in Fig. (4), which gathers the steady state flow and operational transients. Figure (5) shows a separate analysis of both transient and steady state regions. It can be noticed that in the transient region, the correlation degree is close to one and in the steady state region this correlation degree is close to zero.

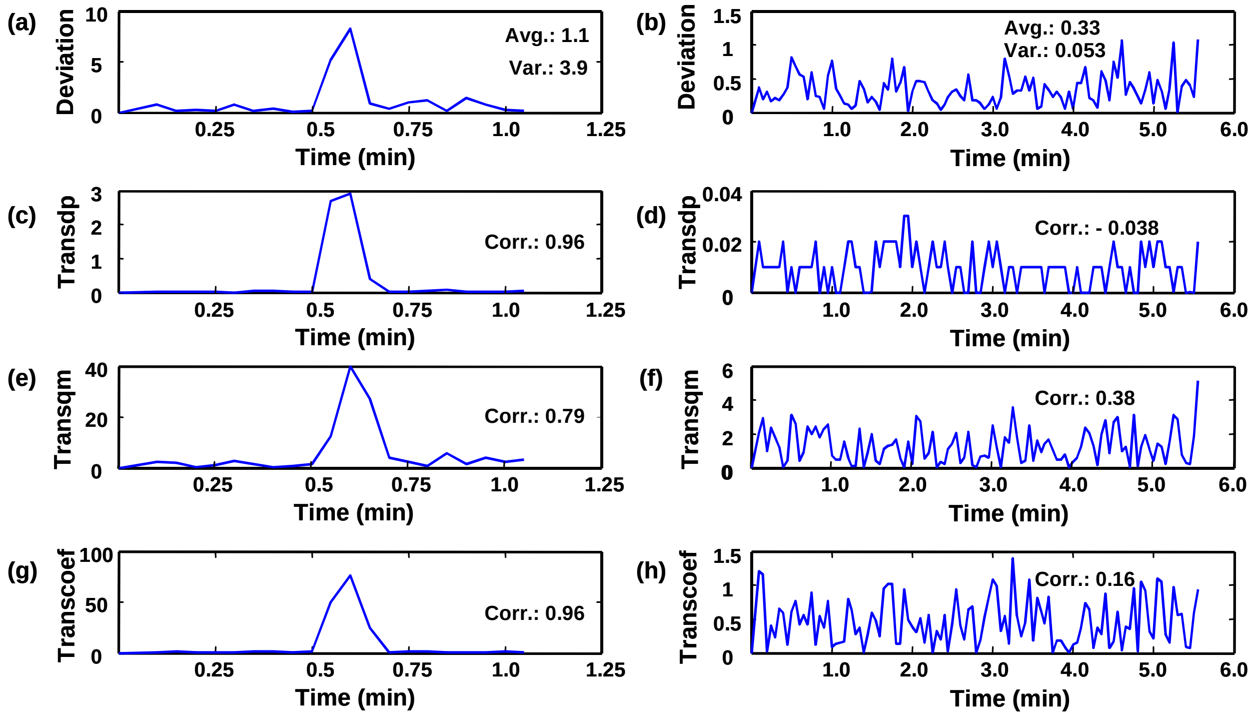

Figure 5. Deviation in a (a) *transient region* and in a (b) *Steady State region*; *Transdp* and its correlation to the deviation in a (c) *transient region* and in a (d) *Steady State region*; *Transqm* and its correlation to the deviation in a (e) *transient region* and in a (f) *Steady State region*; and *Transcoef* and its correlation to the deviation in a (g) *transient region* and in a (h) *Steady State region*;

This statistic allows two main conclusions for the developed system:

1. in the steady state flow, the correlation between the deviation and the transient is low and the deviation is statistically predictable, considering the low variance observed in the series, and;
2. during operational transients, the correlation between the deviation and the transient is high, allowing the "isolation" of this condition for a specific treatment.

4 Architecture of the System

The system is composed by three modules: Fuzzy Rules Design, State Recognition and Deviation Evaluation. In the Fuzzy Rules Design module, statistical tools are used to define the variables and the fuzzy membership functions. In the State Recognition and Deviation Evaluation modules, the rule based fuzzy systems used to classify the flow and identify the operational problems are implemented. Figure (6) shows the system's general architecture.

The applied methodology is here presented and discussed throughout description and detailing each module.

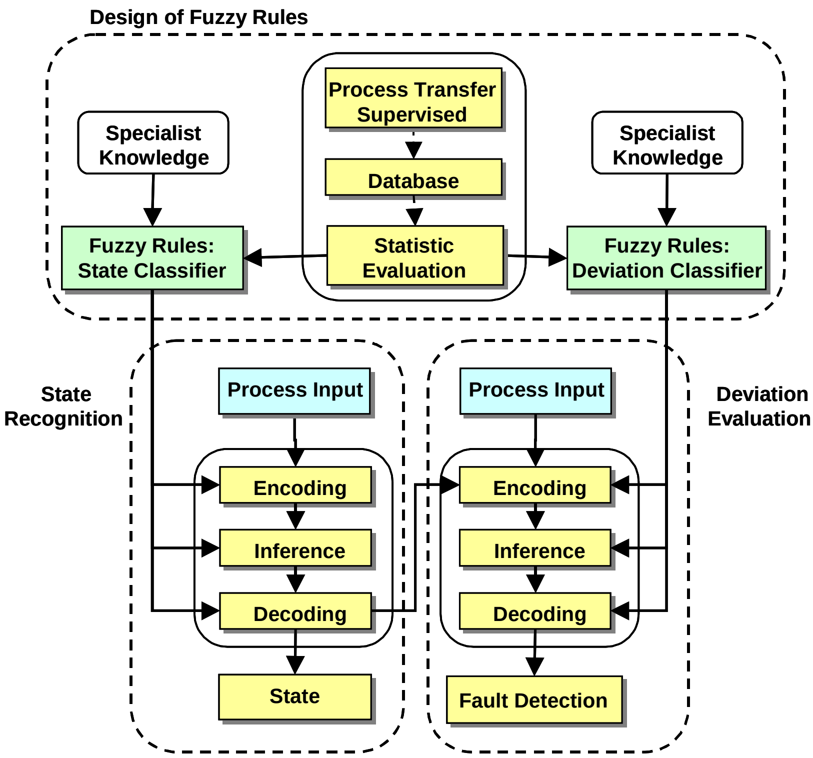

Figure 6. System's general architecture.

4.1 Fuzzy Rules Design

This module consists in a database generation, based on a real LPG transference data classified by an expert. This database was analyzed by using statistical tools and the results of this analysis leads to the specific knowledge of the process. This knowledge is used to define the membership functions associated to the fuzzy linguistic variables (Pedrydz and Gomide, 1998) used as input in the State Recognition and Deviation Evaluation modules. To facilitate comprehension, variables raised by this module will be detailed during the next modules's description.

4.2 State Recognition

This module consists of a Fuzzy Rules Based System composed by two inputs, one output and twelve fuzzy rules, using the centroid method proposed by Mamdani and Assilian (1975) as the defuzzyfication method.

As previously shown, at least two input variables are necessary to classify the flow, one characterizing the total flow level and the other the transient level. The average flow (qm) and the transient measured through the variation of the origin-destination pressure differential (transdp), both previously defined, were respectively selected as the first one and the second. This selection avoids common failure problems since they are taken from different measuring devices.

Linguistic variables associated to the input and output parameters and the definition of their characteristic functions follow.

4.2.1 Input Variables and Linguistic Terms

Linguistic terms were associated to each input variable. To each term, triangular and trapezoidal fuzzy functions were used, and they are shown in Fig. (7), as defined below:

Q_m – *Total Flow* (Zero, Low, Normal, High)
Z –Zero → Triangular Function, parameters [*qma qmb qmba*
L –Low → Trapezoidal Function, parameters [*qma qmb qmc qmd*]
N – Normal → Trapezoidal Function, parameters [*qmc qmd qme qmf*]
H –High → Trapezoidal Function, parameters [*qme qmf qmg qmg*]

Transdp - Transient measured through the origin-destination differential pressure variation (Low, Medium, High)
L – Low → Triangular Function, parameters [*transdpa transdpa transdpb*]
M – Medium → Trapezoidal Function, parameters [*transdpa transdpb transdpc transdpd*]
H – High → Trapezoidal Function, parameters [*transdpc transdpd transdpc transdpe*]

The parameters for the functions defined above are obtained from the database raised in the Fuzzy Rules Design module, according to the following definitions:

$qma = zero$; $qmb = 0{,}03.\max(qm)$;

$qmc = \overline{qm}\big|_{SS} - 3.\sigma_{qm|_{SS}}$; $qmd = \min(qm|_{SS})$;

$qme = \max(qm|_{SS})$; $qmf = \overline{qm}\big|_{SS} + 3.\sigma_{qm|_{SS}}$;

$qmb = 1{,}3.\max(qm)$; $transdpa = zero$;

$transdpb = \overline{transdp}\big|_{SS} + 3.\sigma_{transdp|_{SS}}$; $transdpc = \overline{transdp}\big|_{OT} + 2.\sigma_{transdp|_{OT}}$;

$transdpd = \overline{transdp}\big|_{OT} + 3.\sigma_{transdp|_{OT}}$; $transdpe = 1{,}3.\max(transdp)$.

where: qm : time series of the average flowrate measured between the origin and destination;

$\overline{qm}\big|_{SS}$: qm series average in steady state conditions;

$\sigma_{qm|_{SS}}$: qm series standard deviation in steady state conditions;

$transdp$: time series of the origin-destination differential pressure transient;

$\overline{transdp}\big|_{SS}$, $\overline{transdp}\big|_{OT}$: $transdp$ average in steady state and operational transient conditions;

$\sigma_{transdp}\big|_{SS}$, $\sigma_{transdp}\big|_{OT}$: $transdp$ standard deviation in steady state and operational transient conditions.

4.2.2 Output Variable and Linguistic Terms

DESVIATION – Deviation Classification (Blocked, Starting or Stopping, Operational Transient, Steady State, Operational Problem)

B – Blocked → Triangular Function, parameters [0 1 2]
SoS – Starting or Stopping → Triangular Function, parameters [1 2 3]
OT – Operational Transient → Triangular Function, parameters [2 3 4]
SS – Steady State → Triangular Function, parameters [3 4 5]
OP – Operational Problem → Triangular Function, parameters [4 5 6]

These functions are graphically represented in Fig. (7).

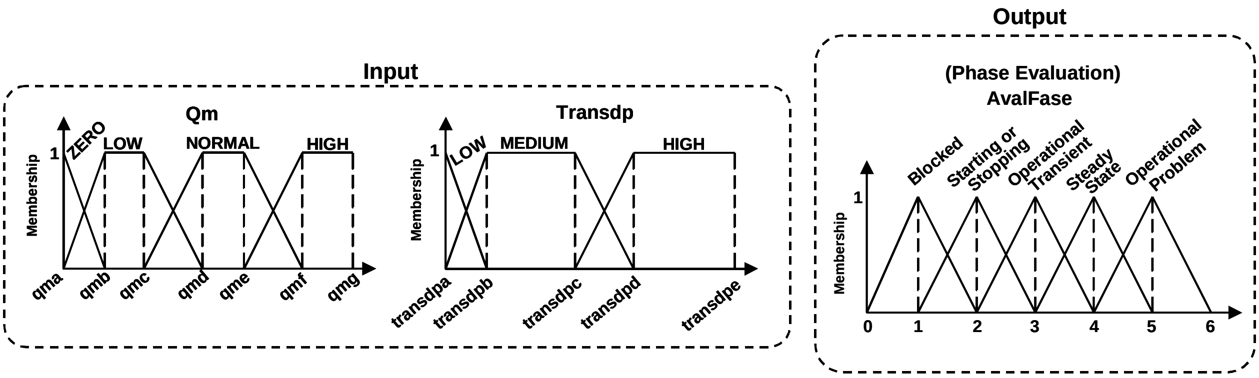

Figure 7. Input and Output variables associated to the State Recognition.

4.2.3 Fuzzy Rules and Inference Process

Using the analysis from the Fuzzy Rules Design module and the specialist's knowledge, fuzzy rules were generated based on the systems general knowledge. These fuzzy rules are summarized in Tab. (1). The inference process used is the Mamdani and Assilian (1975) method: where the logical operator AND was used as the minimum, the implication was used as the maximum operator and the centroid was used as the defuzzyfication method.

Table 1 - Rules of Module 2

		Transdp		
		L	**M**	**H**
Qm	**Z**	B	B	B
	L	OP	SoS	SoS
	N	SS	OT	SoS
	H	OP	OP	SoS

4.3 Deviation Evaluation

The objective of this module is to classify the deviation between the origin's and the destination's measured flow observed into acceptable, evidencing the transference process regular operation, or if they are above the acceptable level, characterizing a measuring device problem or a leakage. As described in section 3, the deviation tolerance is related to the transient observed. It is then expected that, during a small leakage, the system will initially identify an operational transient with acceptable levels, but as soon as the leakage is stable the system will detect it.

This module was also implemented through a Rule Based Fuzzy System as proposed by Mamdani and Assilian (1975), with two inputs, one output and twenty-five fuzzy rules, using the centroid as the defuzzyfication method. The deviation measured from input data and the flow classification from the State Recognition module are used as input variables.

The input and output variables, their characteristic functions and the parameters' definitions follow.

4.3.1 Input Variables and Linguistic Terms

DESVIATION – Deviation Classification (Zero, Low, Normal, High)

VN – Very Negative → Trapezoidal Function, parameters [*desva desva desvb desvc*]
N – Negative → Triangular Function, parameters [*desvb desvc desvd*]
Z –- Zero → Triangular Function, parameters [*desvc desvd desve*]
P – Positive → Triangular Function, parameters [*desvd desve desvf*]
VP – Very Positive → Trapezoidal Function, parameters [*desve desvf desvg desvg*]

AvalFase – Non-defuzzyficated output defined in the State Recognition module

The parameters of the *deviation* variable were obtained in the Fuzzy Rules Design module from the statistical base as defined in below:

$desva = -100$; $desvb = \min(desv|_{OT})$;

$desvc = \min(desv|_{SS})$; $desvd = \overline{desv}|_{SS}$;

$desve = \max(desv|_{SS})$; $desvf = \max(desv|_{OT})$.

$desva = 100$;

where: qm : time series of the deviation measured between the origin and destination;

$\overline{desv}|_{SS}$: *desv* series average in steady state conditions;

$desv|_{OT}$: *desv* series average in operational transient conditions.

4.3.2 Output Variable and Linguistic Terms

DESVIATION – Deviation Classification (Measuring Error Alarm, Measuring Error, Normal, Leakage Alarm, Leakage)

ME – Measuring Error → Triangular Function, parameters [0 1 2]
MEA – Measuring Error Alarm → Triangular Function, parameters [1 2 3]
N – Normal → Triangular Function, parameters [2 3 4]
LA – Leakage Alarm → Triangular Function, parameters [3 4 5]
L - Leakage → Triangular Function, parameters [4 5 6]

These functions are graphically represented in Fig. (8).

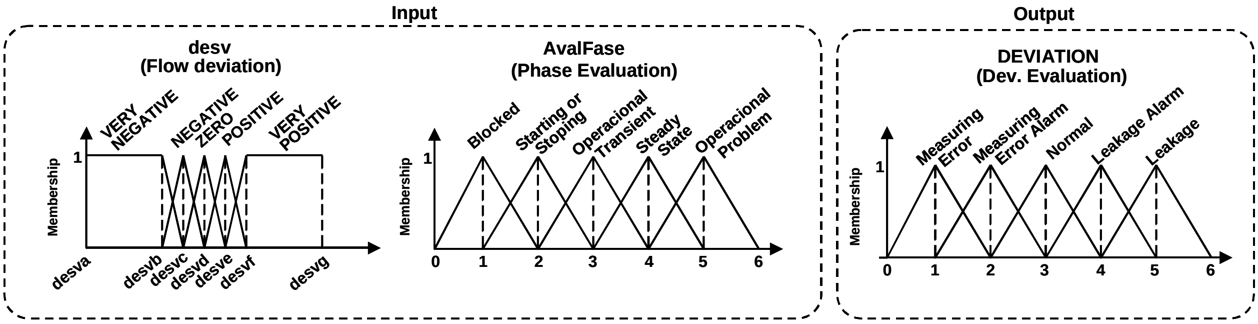

Figure 8. Input and Output variables associated to the Deviation Evaluation.

4.3.3 Fuzzy Rules and Inference Process

Using the Fuzzy Rules Design module analysis and the specialist's knowledge, fuzzy rules based on the system's general knowledge, were defined. These fuzzy rules are summarized in Tab. (2). The inference process used is the Mamdani and Assilian (1975) method: where the logical operator AND was used as the minimum, the implication was used as maximum operator and the centroid was used as the defuzzyfication method.

Table 2 - Rules of Module 3

		Deviation				
		VN	**N**	**Z**	**P**	**VP**
Phase	**B**	N	N	N	N	N
	SoS	N	N	N	N	N
	OT	MEA	N	N	N	LA
	SS	ME	MEA	N	LA	L
	OP	ME	MEA	N	LA	L

5 Results

A material transference data set was obtained from an oil refinery. From this data set, three pumping operations previously classified as classical by a specialist were used. Subgrouping was used to create the statistical base used in the Fuzzy Rules Design module and to determinate parameters for the fuzzy functions associated to the input and output parameters. Based on these parameters settings, two fuzzy modules were implemented. During the test phase, the system was used to evaluate another real pumping operation, also classified by the specialist. The following results were obtained:

Results from the State Recognition module: Flow Evaluation

Comparing the evaluation made by the system and by the specialist, a 93.67% rightness rate was obtained and most of the divergences were caused by the difference between the system's and the specialist conclusions during the transition of two close phases.

Results from the Deviation Evaluation module: Deviation Evaluation

Comparing both evaluations, a higher rightness rate was obtained: 98.03% for the pumping operation. It is good to notice that, among the divergences, only one detection failure was observed in universe of around three thousand failures. The other divergences were alerts, which went back to the original conditions right away.

In Fig. (9) we present a graphic plotted from a specific part of the pumping operation, comparing the system and the specialist's evaluation.

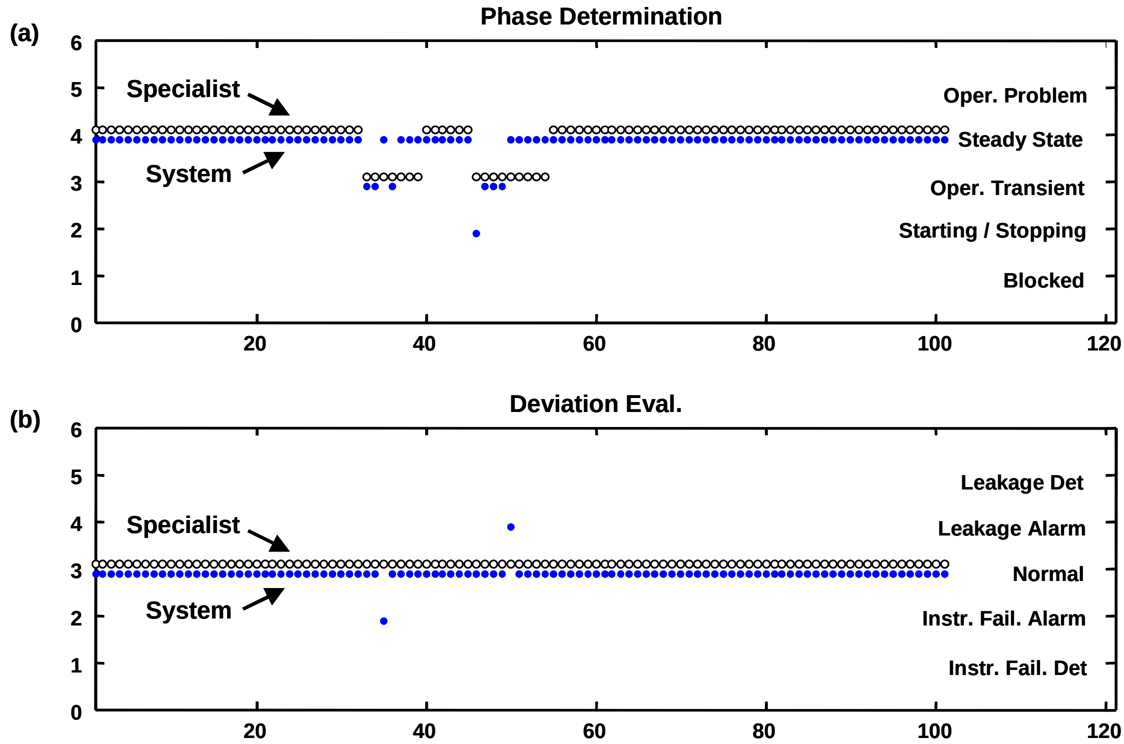

Figure 9. Comparison of the system and the specialist evaluation for the (a) *Phase Determination* and the (b) *Deviation Evaluation*.

6 Conclusions

The results obtained by the system are satisfactory, considering the low computational cost involved. It can be incorporated to the plant control and supervising system, with no need of a dedicated system. Establishing a new supervisory routine can eliminate the small variations' error through the process' continuous supervision.

The results obtained with the Rule Based Fuzzy System showed that the fuzzy logic used to evaluate a petroleum derivate transference process is a very adequate and promising tool. It may lead to other Artificial Intelligence techniques, such as neural networks built with the same fuzzy rules and input granularization criteria used herein, walking towards into the elaboration of new systems, more robust and flexible, to attend diverse transference systems.

6 References

Costa, A.L.H., "Leak Detection in a Pipeline", IBP, Brazilian Petroleum and Gas Institute, 2001.

Ellul, I., 1989, Pipeline leak detection, The chemical Engineer, pp 40-45, June.

Mamdani, E.H. and Assilian, S. (1975). "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller". International Journal of Man-Machine Studies, Vol. 7, N. 1, pp 1–13.

Moura, M.C.G., "Sistema de Medição Para Balanço de Massa em Dutos Críticos", IBP, Brazilian Petroleum and Gas Institute, 2001.

Pedrycz, W. and Gomide, F., An Introduction to Fuzzy Sets, MIT Press, (1998) 221-261.

Sattary, J.A., 1995, "Standardization and Evaluation of Uncertainty in Flow Measurement", 2nd. Brazilan Symposium on Flow Measurement.

Stouffs, P., Giot, M., 1993, Pipeline leak detection based on mass balance: Importance of packing term", J. Loss Prevention in Process Industry, vol. 6, no. 5, pp 307-312.

Taillefond, N. and Wolkenhauer, O. "Fuzzy Clustering and Classification for Automated Leak Detection Systems", 15th Triennial World Congress, Barcelona, Spain, 2002.