Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Petroleum Science and Engineering 49 (2005) 223–238
Journal of Petroleum Science & Engineering
www.elsevier.com/locate/petrol

Leak detection in petroleum pipelines using a fuzzy system

Henrique V. da Silva [a], Celso K. Morooka [b,*], Ivan R. Guilherme [c], Tiago C. da Fonseca [b], José R.P. Mendes [b]

[a] Petrobras, Paulinia-SP, Brazil
[b] State University of Campinas, UNICAMP/FEM/DEP, C.P. - 6052, 13083-970, Campinas-SP, Brazil
[c] Paulista State University, DEMAC/IGCE. Rio Claro-SP, Brazil

Accepted 20 May 2005

Abstract

A methodology for pipeline leakage detection using a combination of clustering and classification tools for fault detection is presented here. 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 for 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 a modified data set including simulated leakages. The methodology is applied to a small-scale LPG 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, obtaining increased leakage detection with low computational costs.

Keywords: 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, flow rate 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 these approaches, data for the estimate of the mass balance are obtained from flow calculators installed in the pipelines. Usually, two equations are used for the mass balance: a more simplified one for the purpose of obtaining a rapid response to large leakages and another more detailed one to obtain slower responses (around 3 to 6 h) in the identification of small leakages (0.5% of the nominal flow). The method proposes the installation of devices at the extremities of the pipelines and occasionally at an intermediary point. This technique does not permit localization of the leakage point.

In the reviewed work, Ellul (1989) describes another technique also referred to as the deviation method. This method is based on the mathematical model of pipeline systems capable of deducing flows and pressures, which can in turn be compared with the measured values. Differences can indicate leakages. The system was capable of detecting leakages of 1.6% of the nominal flow

* Corresponding author. Tel.: +55 19 3788 3339; fax: +55 19 3289 4916.
E-mail address: morooka@dep.fem.unicamp.br (C.K. Morooka).

0920-4105/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.petrol.2005.05.004

rate capacity of a natural gas line in 40 min. For an ethylene pipeline, 4% of the leakage was detected between 1 and 13 min, depending on the location of the leakage. For other liquids, the system was capable of detecting leakages quickly, but without good performance in the localization, even in the case of relatively large leakages (4.7%).

Stouffs and Giot (1993) also presented a mass balance-base system, using a pipeline model for the purpose of evaluating changes in the pipeline inventory during flow transients. They highlighted the importance of the packing term, and concluded that the bottom limit for leak detection was in the order of 2% for the steady state flow and around 3% for transient flow.

Another model, proposed by Siebert (1981), was used to detect and locate leakage in the order of 0.2% of the nominal flow of a gasoline pipeline during operation. The data was collected every 1.7 s. The method is based on a statistical analysis of the data (crossed correlation), showing it to be capable of detecting, in a few hours, leakages in the order of 5% for gas lines and locate them with an error margin of approximately 20 km.

Wang et al. (1993) proposed a leak detection method on an autoregressive model. This method requires only four pressure devices and proves capable of detecting immediately a leakage in the order of 0.5% in a water pipeline of 120 m, using a sampling time of 2 min. However, this method requires the installation of various pressure devices along the length of the pipeline and synchronized transmission of the data. Costa et al. (2001) modeled a system, also based on the inverse analysis of hydraulic transients, in order to detect leakage in a water distribution system, with good performance.

Baptista et al. (2001) presented a study of pressure transients generated by a leakage in a pipeline of around 50 km in length, monitored at nine points along this length by three different fluids (diesel, gasoline and LPG). In this study, it was observed that the pressure drop in the pipeline, occurring in the presence of a leakage, in relation to the expected pressure in the steady state for the same point, is greater, the closer the leakage and the more viscous the fluid (in this case, the oil diesel proved to be more responsive). The tests presented in this work, demonstrate that soon after the beginning of the simulated leakage, the pressure conditions of the pipeline stabilize and another hydraulic profile establishes itself (a new regime condition). These details are important for the elaboration of the strategy presented here.

Parry et al. (1992) showed a performance obtained by a pipeline leakage detection system based on the balance of compensated volume. The system is interfaced with a data acquisition system in real time, in order to obtain real information direct from the field. Flow devices are installed at various points along the length of the pipeline. Leakages in the order of 2% in an LPG line can be detected between 46 min and 9.2 h, depending on the location of the leakage.

Zhang (1992) describes a statistical method for the dimensioning and location of pipeline leakages. This system uses flow and pressure devices at the extremities of a gas line extending to the length of 100 km. It was also tested in a liquid propylene pipeline of 37 km. Both the numeric simulations and the field tests showed that the system could detect leakages in the order of 1% with great precision in localization.

Jonsson and Larson (1992) studied the characteristics of wave propagation through a water pipeline of 5000 m, following the shutting down of a pump, with and without the occurrence of a leakage. The effects of a leakage were analyzed through the variations of the pressure signals of the system. The method succeeded in detecting leakages of up to 7% of the nominal flow rate.

Coelho and Medeiros (1999) presented another class of leakage detection system that is more promising, reliable and demands lower investment in its implementation. In this class, the quality of detection obtained depends on the models applied to the process, whether they are of a stochastic nature, generated by strategies based on the input and output variables (states), temporal series identification, or built according to phenomena principles and laws. In this same category we find the dynamic modeling in compressible outflow networks for application to real time leakage detection carried out by Neto (2002).

An Acoustic Emission Technique (Fantozzi, 2000) is also proposed. Initially, it is capable of rapidly detecting small leakages, but does not function very well in pipelines with a complex topology, or when background noise from valves, pumps or compressors is elevated (Ellul, 1989). Added to this, the spacing between the stations must be reduced to increase reliability.

Pressure waves generated by a leakage supply a potential method for the leakage detection by means of the measurement of pressure disturbances observed along the length of the pipeline, in accordance with a study by Silk and Carter (1996). However, this requires a complex system of real time data acquisition in order to allow the monitoring of temperature and pressure variations along the length of the pipeline (offsetting the speed of sound with any variable). Moreover, it is necessary to filter the normal pressure disturbances of the process.

Belsito et al. (1998) presented an important study comprising various methods of pipeline leakage detection, reproduced in part above, introducing the use of neural networks (ANN—Artificial Neural Network) for their application. In this work, they show that the neural networks possess certain attributes that make them an extremely adequate approach for the processing of data obtained in systems of pipeline transference, and could be used innovatively in methods of pipeline leakage detection with little requirement for a large sample frequency. A similar approach can be found in the work of Caputo and Pelagagge (2002).

The various approaches and proposed models for the solution of this problem, intensified in the last two decades, are indicative of the importance of this study, strongly motivated by the growing need for control of processes capable of causing damage to health and the environment.

The proposal presented here uses the model of compensated volume conservation, projected by Ellul (1989) and Stouffs and Giot (1993), with a modified approach, in which we use the measured transient level to evaluate the deviation observed in the mass balance application. Using this model and knowledge from a specialist, an intelligent system for leakage diagnosis is proposed. The projected system employs the concepts of fuzzy sets theory.

Real data collected every 10 s from a small LPG pipeline is used for the development of this system. The pipeline has an 8-in. diameter and a length of 2000 m with pressure, temperature and flow rate transmitters installed at its extremities. For the purpose of testing, the database was evaluated by an expert. After having been modified for abnormal situations through simulation, each stage of the transfer process and the in–out flow deviation was categorized.

Using the data from the LPG pipeline, an analysis using statistical techniques was carried out. The results showed a high correlation between the inlet–outlet flow rate deviation and the operational transients. This was the crucial factor applied in defining the fault detection strategy. Using the model and the statistical analysis it was possible to define the variables that would allow characterization of the leakage problems.

In the intelligent system proposed, two processes are performed. Firstly, the development of a classifier module that can identify the operational and process transients and determine the current stage of the transfer process. Secondly, a Fault Detection module that will evaluate the inlet–outlet flow rate deviation in order to detect a leak or abnormal situation, with a low level of false alarms, employs the output of this module.

Using a rule-base system developed from this database, a Fuzzy Inference System is employed to solve the presented problem. The system was evaluated by new data collected from the same process, and good results were obtained, with detection of increased leakage or abnormal situation. The low computational costs involved and low level of false alarms obtained are the most attractive aspects of the presented system.

2. Process description

The petroleum products produced by a refinery are spread to distribution companies by pipelines. The Measuring Station (EMED) composes the basic control system responsible by the custody transfers of petroleum products to the buying companies. The process variables measured in the EMED, such as pressure, temperature, flow and density, are available in real time. At the distribution companies, remote stations measure the flow, pressure and sometimes temperature.

The data available on the EMED computers and in the remote stations are sent to the Supervision System by serial links, which consolidates information from the extremities of the pipeline, on the same time base, for global supervision and leakage evaluation. Fig. 1 shows an available scheme for instrumentation and a transfer system.

In this simplified diagram, a single pump graphically represents the entire pumping station with its respective alignment of suction from the LPG drums. At the measuring station, only the instruments whose signals were used in the elaboration of the study are indicated; thus omitting the separator drums (used for separation and drainage of water), the filters (which eliminate residual particles), the provers (used for calibration of the measuring turbines) and the complementary instruments (pressure and temperature devices installed in the provers and separator drums, differential pressure devices installed in the filters, etc.). All of this infrastructure, as well as the operational and maintenance procedures, are developed in order to guarantee a level of uncertainty permitted in measuring systems destined for custody transference. The pressure and temperature devices (PT and TT) are used to calculate the respective volume factors associated with the compressibility and thermal expansion. The density devices (DT) are used to convert volumetric flow, calculated by the measuring turbine (FT), into mass flow.

Also in the distributing companies, represented in the diagram as ‘destination’, are installed some devices for balance closure and leakage detection, not having, therefore, the precision requirements necessary for the

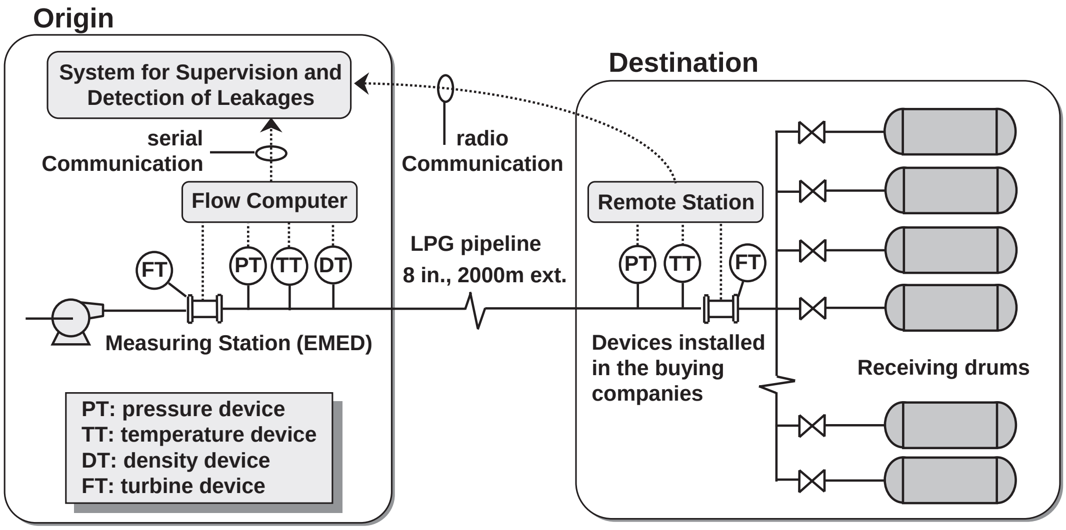

Fig. 1. Petroleum derivatives transference system and monitoring instrumentation.

system developed by the EMED. In this way, significant differences can be observed in the specification of the systems installed at each extremity, amongst these are: (a) the principle used by the flow devices installed in the distribution companies is ultra-sonic with an uncertainty of 0.5%, much greater than the uncertainty of the measuring turbine installed at the origin, in the order of 0.1%; (b) the temperature sensors installed at the destination are thermocouples with an uncertainty in the order of 1 °C, whilst those installed at the origin are thermo-resistant (PT500), with an uncertainty in the order of 0.3 °C; (c) the maintenance requirements (for example frequency of standardization), are also different.

In relation to data acquisition and processing, the equipment used at the EMED and at the destination also differs. At the origin, the measuring devices are connected to an instrument dedicated to this function—the flow computers. They have great precision and velocity during the sampling and digitalization of input data, with complex algorithms developed for the purpose of correcting the temperature and pressure effects by means of specific normalized spreadsheets (in the presented case, the API—American Petroleum Institute). At the destination, however, the data is processed by a general purpose Programmable Logical Controller (PLC), with less accuracy requirements and simplified correction algorithms (based on general equations, and not on specific spreadsheets).

Deviations in the calculation of the balance of mass between the origin and destination are expected, in virtue of the global uncertainty of the measuring system, as emphasized above, as well as the eventual lack of synchronization in the composition of data and the simplifications carried out in the model, strongly influenced during the transients that occur in the process.

The present paper focuses on the monitoring of an LPG (Liquefied Petroleum Gas) transference procedure. In this process, operational transients often arouse greater complexity. During this transference process, the pressure gradually rises whilst the LPG receiving drum is filled. When the LPG drum is completely full, the transference process is switched to a new drum. At that moment, a sudden expansion is observed and an increase in the flow rate occurs. During the drum filling process (steady flow state), there is only a small flow–balance deviation between the origin and destination. This deviation is expected according to mass balance model, and it is generated by the inherent uncertainties associated with the measuring process (Sattary, 1995). However, during the operational transient related to the receiving vessel switch procedure, the deviation observed here rises to significant values. This is motivated principally by the line packing effect, accounted for by the mass balance model, as a result of diverse responses from measuring devices and the eventual lack of synchronization in the data acquisition system.

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

3. System modeling and data analysis

In this section, the adopted modeling and the procedure for data analysis from the product-transfer process are presented.

3.1. System modeling

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 flow, as shown below:

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

where:

- Q_o Volumetric flow measured at the pipeline's origin;
- Q_d Volumetric flow measured at 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: ε is the uncertainty of flow measuring device.

The introduction of the term ε, further increases the difficulty of modeling by deterministic methods. Even considering that the uncertainty ε includes not only the uncertainty of measuring devices and entries to the data acquisition system, but also the eventual lack of synchronization in transmission of data, this variable presents distinct behavior in diverse operational conditions, since the effect of lack of synchronization is more relevant during the transients. Thus, for more adequate consideration of this term, the variable is broken down into two new terms, in Eq. (3).

$$\varepsilon = \varepsilon_{SS} + \varepsilon_{OT} \tag{3}$$

where:

- ε_{SS} Uncertainty of the measuring device in the state of steady flow.
- ε_{OT} Uncertainty of the measuring device during operational transients.

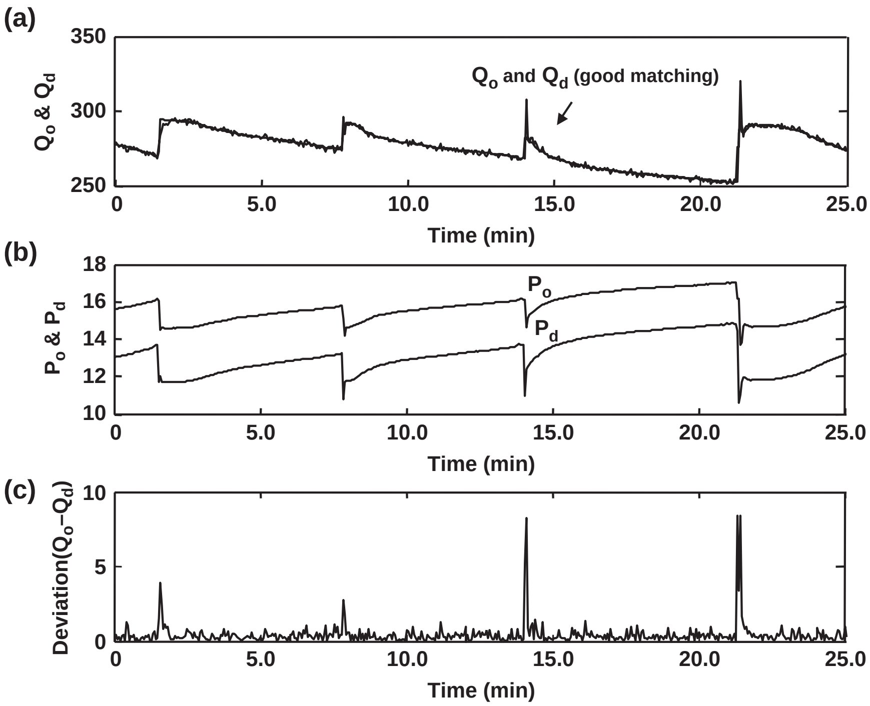

Fig. 2. Typical behavior of LPG transference in terms of (a) flow, (b) pressure and (c) deviation.

Eq. (2) can be rewritten accordingly:

$$(Q_o - Q_d) = -\frac{dLP}{dt} \pm \varepsilon_{SS} \pm \varepsilon_{OT}. \quad (4)$$

From Eq. (4) one can conclude, moving away from the hypothesis of the absence of leakage, that:

- In steady flow, the difference between flow at the origin and at the destination of the duct corresponds to the uncertainty of the measuring device ε_{SS}, as the term dLP is equal to zero for constant pressure and temperature profiles.
- During an operational transient, this difference is equal to the term dLP added to the uncertainty ε_{OT} of the measuring device during the transient.

The variable deviation is defined as:

$$\text{deviation}(t) = \frac{Q_o(t) - Q_d(t)}{Q_o(t)} \cdot 100 \ \forall \ Q_o \neq 0 \text{ and}$$
$$\text{deviation} = 0 \ \forall \ Q_o = 0. \quad (5)$$

Fig. 2 shows the typical behavior of different parameters in LPG transference, where (a) flow, (b) pressure and (c) deviation between origin and destination flow are depicted. Often, operational transients in this process occur during the receiving drum switch procedure, and increased deviation is calculated between the measured flow rates during these operations. This is emphasized in the presented study.

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

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 Eqs. (6)–(9).

Transient measured through average volumetric flow (Transqm):

$$\begin{aligned}\text{Transqm}(t) &= \text{abs}\left(\frac{Q_m(t) - Q_m(t-1)}{\Delta t}\right) \\ &= \text{abs}\left(\frac{Q_o(t) + Q_d(t) - Q_o(t-1) - Q_d(t-1)}{2\Delta t}\right)\end{aligned} \quad (6)$$

where: Q_m is the average instantaneous flow.

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

$$\text{Transdp}(t) = \text{abs}\left(\frac{(P_o(t) - P_d(t)) - (P_o(t-1) - P_d(t-1))}{\Delta t}\right) \quad (7)$$

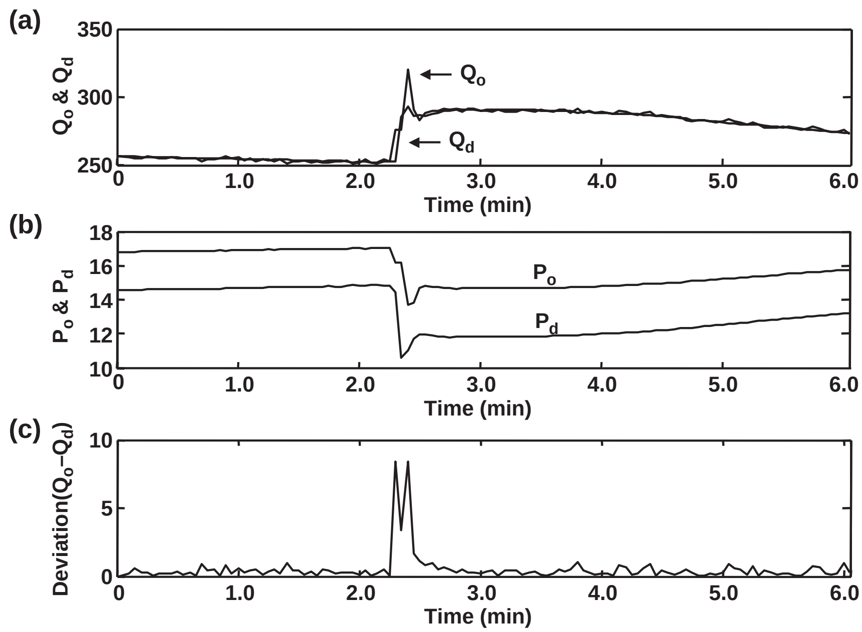

Fig. 3. Detailed behavior of (a) flow, (b) pressure and (c) deviation during the switch operation.

$$\mathrm{Transdpm}(t) = \mathrm{abs}\left(\frac{2\cdot\left(P_o^t - P_d^t\right) - \left(P_o^{t-\Delta t} - P_d^{t-\Delta t}\right) - \left(P_o^{t-2\Delta t} - P_d^{t-2\Delta t}\right)}{3\Delta t}\right) \tag{8}$$

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

$$\mathrm{Transcoef}(t) = \mathrm{abs}\left[\frac{\left(\dfrac{Q_o(t) - Q_d(t)}{\left(P_o(t) - P_d(t)\right)^2}\right) - \left(\dfrac{Q_o(t-1) - Q_d(t-1)}{\left(P_o(t-1) - P_d(t-1)\right)^2}\right)}{\Delta t}\right]. \tag{9}$$

Where Q and P are flow and pressure respectively, subscript o and d indicate the measuring point, respectively origin and destination. Superscript t, $t-\Delta t$ and $t-2\Delta t$ indicate the moment of measurement, respectively, the actual instant and the two previous samples.

From the variables defined above, the correlation between the temporal series (deviation × Transcoef; deviation × Transdp and deviation × Transqm) is found. The correlation is thus defined as in Eq. (10):

$$\mathrm{Corr}_{X,Y} = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} \tag{10}$$

where: $\sigma_X^2 = \frac{1}{n}\sum_1^n (X_i - \mu_X)^2$ and $\sigma_Y^2 = \frac{1}{n}\sum_1^n (Y_i - \mu_Y)^2 \mu_X$ and μ_Y represent the average of the values observed in the temporal series X and Y.

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

As the correlation is relatively high, around 0.8, the deviation can be associated with 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 flow state and operational transients. Fig. 5 shows a separate analysis of both transient and steady state regions. One can notice that in the transient region, the cor-

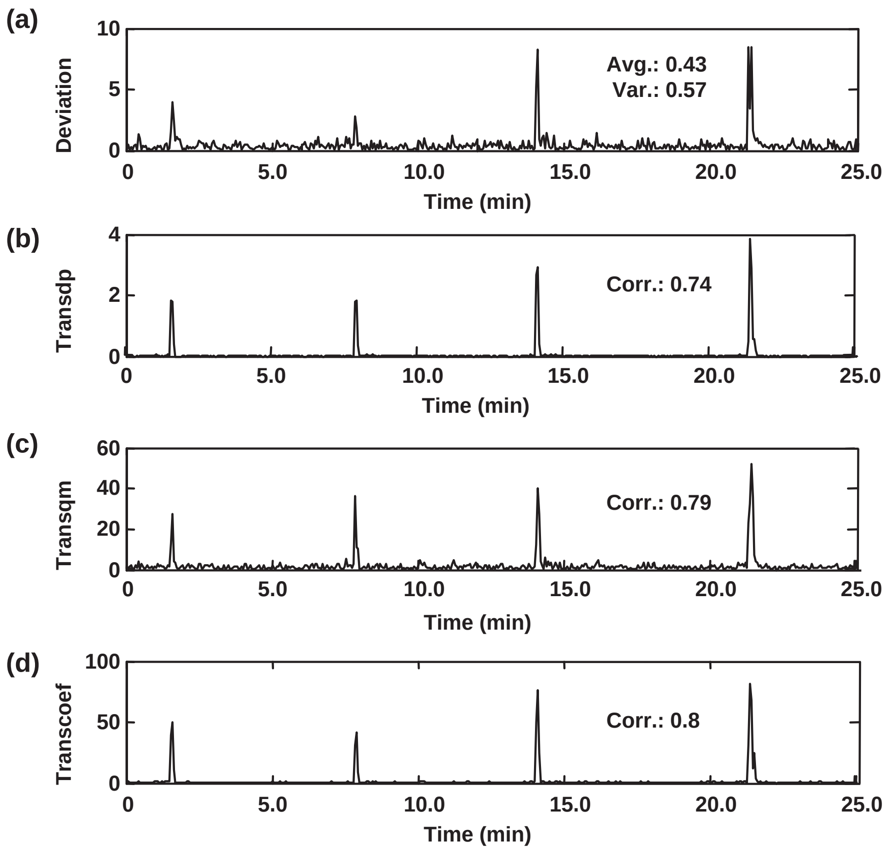

Fig. 4. (a) Deviation and variables capable of identifying the casual operational transients, along with the correlation between these variables and the deviation: (b) Transdq, (c) Transqm and (d) Transcoef.

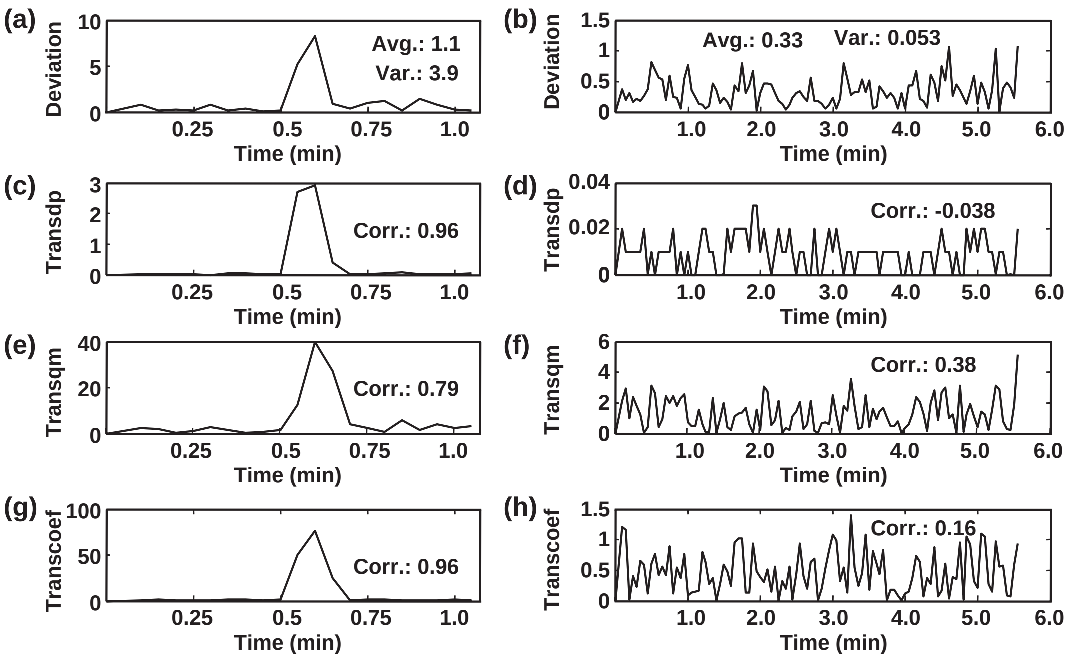

Fig. 5. Deviation in a (a) transient region and in a (b) Steady State region; Transdp and its correlation to the deviation in a (c) transient region and in a (d) Steady State region; Transqm and its correlation to the deviation in a (e) transient region and in a (f) Steady State region; and Transcoef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

relation degree is close to one and in the steady state region this correlation degree is close to zero.

This statistic leads to two main conclusions for the developed system:

1. In the steady state flow, or pseudo-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.

3.2. Data analysis

In the determination of the entry variable for the solution of the problem, essential for the development of the system, it is fundamental not only to possess knowledge of the correlation existing between the series, but also to carry out an analysis of the statistical distribution of the latter. In the previous item, instantaneous deviation was correlated with the level of operational transient and the series were analyzed separately under two operational conditions: pseudo-steady state flow and operational transient. It was observed that, when in pseudo-steady state flow, the temporal series presented a small variance. This statistical distribution is shown in Fig. 6. One can see that this distribution comes close to a normal distribution. This consistency is very important for the development of the system, guaranteeing a statistical predictability of this variable under the condition of pseudo-steady state flow.

The behavior of the mean deviation between the origin–destination flows was also studied. It was calculated in a time window of 10 min, in which only the percentage instantaneous deviations are considered, investigated under the steady state flow condition of this calculation. This mean percentage deviation was de-

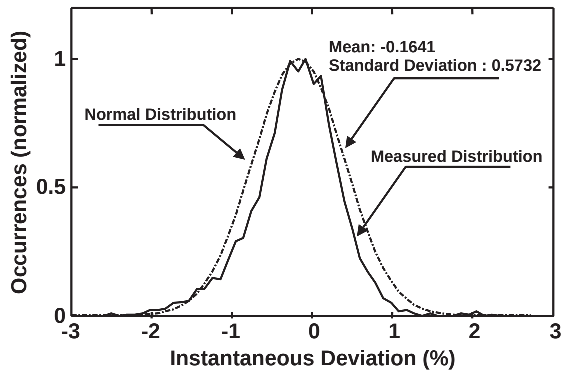

Fig. 6. Distribution of instantaneous deviation (%).

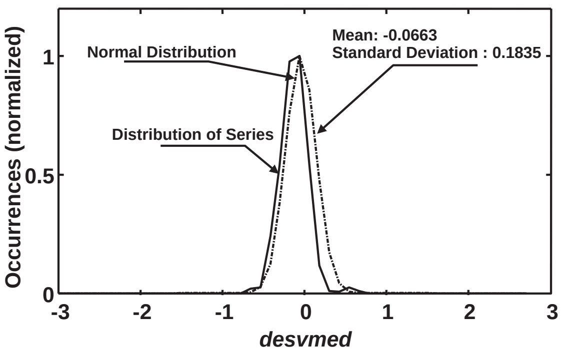

Fig. 7. Distribution of the mean deviation (%) measured in 10 min.

fined as devmean, in Eq. (11). In Fig. 7, we show the distribution of this variable.

$$\mathrm{devmean}(t) = \frac{\int_{t-\Delta t}^{t} |dev(t)|_{\mathrm{SS}}}{\Delta t} \quad \text{for } \Delta t = 10 \text{ min} \tag{11}$$

where: $|\mathrm{dev}(t)|_{\mathrm{SS}}$: series $\mathrm{dev}(t)$ measured under steady state flow condition.

As well as finding that the distribution of both variables can be associated with a normal distribution, one can also see that the standard deviation of the series devmean is around three times less than that observed in the series deviation. This finding allows us to conclude that the use of the variable devmean in the balance of the volume of average deviation allows for the detection of smaller leakages, although with a delay in their identification; whilst the use of the variable deviation allows for the immediate identification of larger leakages. A combination of these two system entries allows for optimization of the leakage identification process.

To continue with the statistical analysis of the variables relevant to the solution of the problem, in the following section, flow behavior is examined under two conditions: in steady state flow and during operational transient.

As expected, Figs. 8 and 9 confirm that the flow distribution cannot be represented by a normal distribution, given that this variable is not aleatory. It is, however, dependent upon the backpressure exercised by the receiver drum, which varies according to the characteristics and operational conditions (level and temperature) of each drum, as well as to the curve of the pump. However, the distribution of the flow values occurs within a band that corresponds to the observed extremes of contra-pressure for both of the studied conditions. During the transients, the maximum flow observed is greater, in line with the abrupt expansions that took place during the drum switching process. These limits,

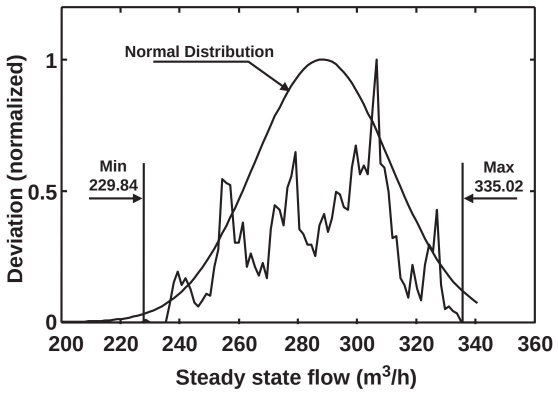

Fig. 8. Distribution of flow under steady state condition.

indicated in Figs. 8 and 9, are significant for the identification of the start-up and shutdown period of the transfer process, where flow levels beneath the lesser limit associated with elevated transients, are expected. Flow levels superior to the maximum limit can occur in the case of very low contra-pressure, characterizing an internal problem in the distributing company receiving the product, or even a fracture in the pipeline.

Variables that characterize the transient level observed in the process, such as transdp, transdpm, transqm or transcoef, defined in Eqs. (6)–(9), are fundamental to the solution of the problem, bearing in mind the direct correlation between these and the deviation. In the strategy for the detection of leakages developed over the course of the next section, these variables are used to classify outflow in regions and associate these with the expected deviation level, that is to say:

- during pseudo-steady state flow, characterized by a low transient level, low values are expected for the deviation, established in this work as a band located in the region between more or less 3 times the standard deviation observed in Fig. 6;

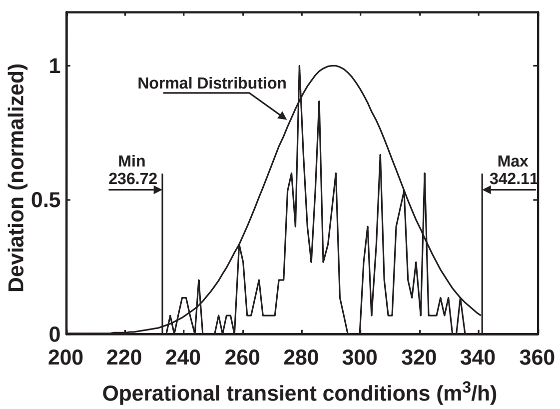

Fig. 9. Distribution of flow under steady state flow condition.

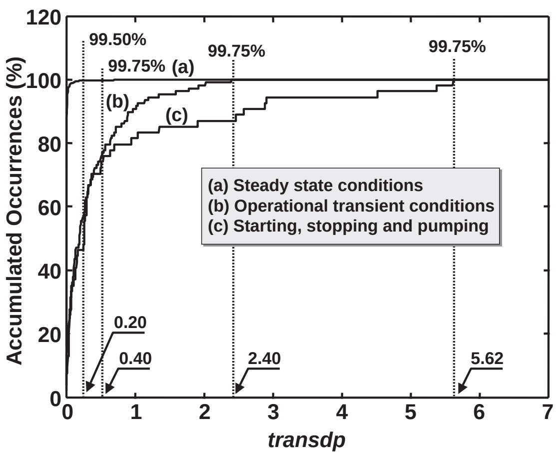

Fig. 10. Concentration of transdp values for each operational stage.

- during operational transients caused by valve operation, characterized by an average transient, medium values are expected for the deviation; and
- during strong transients caused by pump start-up or shutdown, large deviation values are expected.

For the purpose of the investigation, the study of the statistical distribution of the transients will be carried out separately in each region. The variable chosen for the study of transient was transdpm due to its independence in relation to the balance deviation data (transqm and transcoef use the same instruments for balance deviation determination) and its greater consistency when compared with the variable transdp (without the incorporation of any filter).

In Fig. 10, three graphs are presented; representing the concentration of the transdpm values under three conditions of outflow. In graph (a), pseudo-steady state flow, one can observe that 99.75% of the values of transdpm are below 0.40; and that 99.5% of the occurrences are concentrated below 0.20. In graph (b), in operational transient one notes that 99.75% of the occurrences of transdpm values are concentrated below 2.40. In graph (c), in start-up or shutdown, it is observed that in 99.75% of the occurrences of transdpm values are concentrated below 5.6281.

These boundaries are important for the later definition of linguistic terms associated with the variable transdp, which will be dealt with in the next section.

In the present section, after describing the transference process, we identified an existing correlation between the flow deviation and the operational transient, and analyzed some variables relevant to the solution of the problem. The information obtained in this section, will allow for not only the definition of the entry variables for the proposed system, but also the construction of a rule base, to be detailed in the next section.

4. Architecture of the system

In the intelligent system proposed, the Fuzzy Sets Theory (Pedrycz and Gomide, 1998) was employed to represent knowledge from a specialist used in the diagnosis. Three modules compose the intelligent system: 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. Fig. 11 shows the general architecture of the system. Each module will be depicted in further items, explaining the system architecture in more detail.

Initially the definition of a module for leakage detection, in which one of the entries would represent the transient level in existence at that moment, was proposed. However, it became evident that it is important to evaluate the transients in a separate module in order to characterize separately each operational state that the process went through—important information from an operational point-of-view. In addition to this, the evaluation of the variable devmean, defined in Eq. (6), depends upon prior classification of the operational state, bearing in mind that, as previously mentioned, this is a key variable in the identification of small leakages. The main operational states, which we shall detail at a later stage, are: blocked (B), start-up and shutdown (SuSd), operational transient (OT), steady state (SS) or unidentified operational problem (OP). The evaluation carried out by this module is a fundamental entry for the second module, specialized in the detection of leakages or other failures in the transference procedure.

Here, as previously mentioned, the applied methodology is presented and discussed through description and detailing of each module.

4.1. Fuzzy rules design

This module consists of a database implementation, based on real LPG transference data, classified by an expert. The database was analyzed using statistical tools, the results of the analysis leading to specific knowledge of the process. This knowledge is used to define the membership functions associated with fuzzy linguistic variables (Pedrycz and Gomide, 1998) used as input in the State Recognition and Deviation Eval-

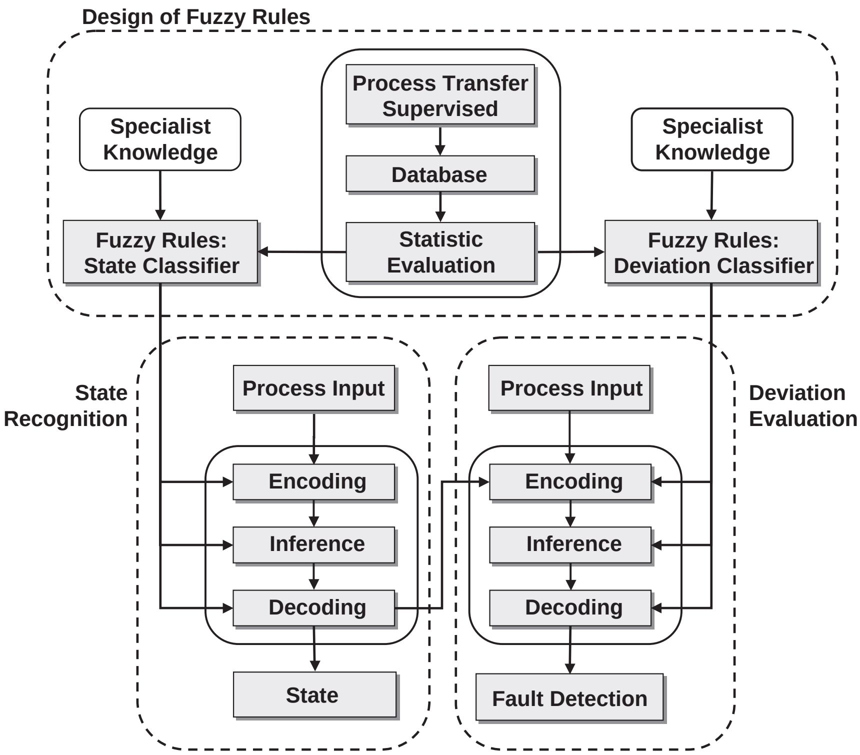

Fig. 11. General architecture of the system.

uation modules. To facilitate comprehension, variables raised by this module will be detailed during description of the next module.

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 inference technique proposed by Mamdani and Assilian (1975).

As previously shown in the System Modeling section, through the variables correlations, 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 selected respectively as the first and the second. This selection avoids common failure problems as they are taken from different measuring devices.

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

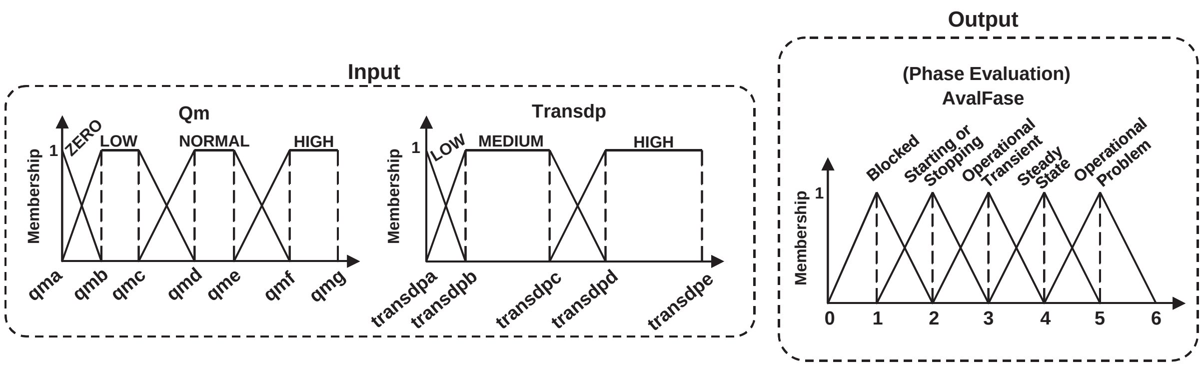

Fig. 12. Input and output variables associated to the state recognition.

4.2.1. Input variables and fuzzy linguistic terms

The average instantaneous flow measured between the origin and the destination (qm) and the transient evaluated by the variation of the pressure differentiation between origin and destination (transdpm), both previously defined, are the input variables used. The selection of the variable transdpm avoids common failure problems, as it is obtained from instruments independent of the other entry.

Fuzzy Linguistic terms were associated with each input variable. To each term, triangular and trapezoidal fuzzy functions were assigned, and are shown in Fig. 12 below:

Qm	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:

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

$$\text{qmc} = \overline{\text{qm}}|_{\text{SS}} - 3 \cdot \sigma_{qm|_{\text{SS}}};\ \text{qmd} = \min\left(\text{qm}|_{\text{SS}}\right);$$

$$\text{qme} = \max\left(\text{qm}|_{\text{SS}}\right);\ \text{qmf} = \overline{\text{qm}}|_{\text{SS}} + 3 \cdot \sigma_{\text{qm}|_{\text{SS}}};$$

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

$$\text{transdpb} = \overline{\text{transdp}}|_{\text{SS}} + 3 \cdot \sigma_{\text{transdp}|_{\text{SS}}};$$

$$\text{transdpc} = \overline{\text{transdp}}|_{\text{OT}} + 2 \cdot \text{r}_{\text{transdp}|_{\text{OT}}};$$

$$\text{transdpd} = \overline{\text{transdp}}|_{\text{OT}} + 3 \cdot \sigma_{\text{transdp}|_{\text{OT}}};$$

$$\text{transdpe} = 1,3 \cdot \max(\text{transdp})$$

where:

qm	Time series of the average flow rate measured between the origin and destination;
$\overline{\text{qm}}\|_{\text{SS}}$	qm series average in steady state conditions;
$\sigma_{\text{qm}\|_{\text{SS}}}$	qm series standard deviation in steady state conditions;
transdp	Time series of the origin–destination differential pressure transient;
$\overline{\text{transdp}}\|_{\text{SS}}$, $\overline{\text{transdp}}\|_{\text{OT}}$	transdp average in steady state and operational transient conditions;
$\sigma_{\text{transdp}\|_{\text{SS}}}$, $\sigma_{\text{transdp}\|_{\text{OT}}}$	transdp standard deviation in steady state and operational transient conditions.

In order to establish the parameters of each function defined, we used the study of statistical distribution presented in Section 3. The parameters established are listed below:

- qma and qmg are the universal limits for the variable qm. Its values were established respectively as: zero and 1.3 times the maximum value measured in the temporal series qm;
- qmb represents the limit around zero, in which the pump may cease despite some residual value being measured by the flow device. The established value was 3% of the maximum value observed;
- qmd and qme were defined by the flow limits observed during outflow in steady state flow. These values can be graphically visualized in Fig. 8;
- qmc and qmf were defined as the flow limits observed under the condition of outflow during operational transients, multiplied respectively by the factors 0.95 and 1.05. The limit values can be graphically visualized in Fig. 9;
- Transdpa and transdpe are the boundaries of the domain universe for the variable transdpm defined in Eq. (8). Their values were established respectively as: zero and 1.3 times the maximum value measured in the temporal series transdpm;
- Transdpb was defined as the value superior to the interval in which 99.5% of the occurrences of transient values measured under the condition of steady state flow were concentrated. This value can be graphically visualized in Fig. 10 (value of 0.20);
- Transdpc was defined as the value superior to the interval in which 99.75% of the occurrences of transient values measured under the condition of operational transient were concentrated. This value can be graphically visualized in Fig. 10 (value of 2.40);

- Transdpd was defined as the value superior to the interval in which 99.75% of the occurrences of transient values measured under the condition of pump start up and shut down were concentrated. This value can be graphically visualized in Fig. 10 (value of 5.62).

4.2.2. Output variable and fuzzy linguistic terms

The output variable is associated with the operational state of the transfer operation. This variable was denominated *Phase-Evaluation.* There are five linguistic terms associated with the system output: Blocked (B), Start-up and Shutdown (SuSd), Operational Transient (OT), Steady State (SS) and Operational Problem (OP). Each term corresponds to the operational state to be classified.

Phase-Evaluation (Blocked, Start-up and Shut-down, Operational Transient, Steady State, Operational Problem).
B Blocked (triangular function)
SuSd Start-up or Shut-down (triangular function)
OT Operational Transient (triangular function)
SS Steady State (triangular function)
OP Operational Problem (triangular function)

These functions are graphically represented in Fig. 12.

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 system's general knowledge. These fuzzy rules are summarized in Table 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.

In the next section, some important aspects considered in the elaboration of the table are detailed.

The state "Start-up and Shutdown" (SuSd) is characterized by high outflow transients in all flow conditions, or by medium value transients in a low flow condition. In the starting condition the transient is positive, whilst in the shutting down condition it is negative. In the definition of this transient its absolute value was used, as the differentiation of the two conditions can easily be inferred on the basis of the previous operational state. The measured value of flow during the starting up or shutting down operation remains low for most of the time. However, at the beginning of the stopping process this flow should be normal, or even high if combined with some other transient immediately before the command is given to shut the pump down.

The state "Blocked" (B) is characterized by a flow close to zero. Some noises generated by the ultra-sonic instrument can occur during measurement, despite the configuration of a *cut-off* (elimination of these noises in low flow states). Small flow levels can also occur due to pipe and fluid thermal expansion by sun exposure after the process of transference has been halted.

The state "Operational Transient" (OT) is characterized by medium transients within a normal band of flow, whilst the state "Steady State" is characterized by low transients. As we can see in Fig. 10, even during the operational transient we come across very low values of transdpm much less frequently, thus making the differentiation of this state in relation to steady state flow more difficult. Two complementary strategies could be used to combat this: (a) use of an additional rule to evaluate the permanence of the outflow in that state during a specific time, soon after its alteration; (b) combination of more than one entry to characterize the transient, for example, transdp, or transqm and transcoef, as well as transdpm itself. The first succeeds in eliminating some spurious points that might occur during permanence in one state. This seems sufficient given that the determination of the exact moment of migration from one condition to another is not very relevant. The second combination would increase the computational costs without any real benefits. It is worth highlighting that the main focus of the system is to determine leakages and other operational problems, the determination of the operational state being an important reference for this procedure. In this work, neither of the two proposed strategies was implemented; nevertheless, the output *defuzzified* by the CoG method guarantees the anticipated efficiency of the leakage detection system.

The state "Operational Problem" (OP) is observed in the stabilized condition (low transient) outside of the flow band or with medium transients in flow levels above the normal band. This condition could mean

Table 1
Rules of Module 2

		Transdp		
		L	M	H
qm	Z	B	B	B
	L	OP	SuSd	SuSd
	N	SS	OT	SuSd
	H	OP	OP	SuSd

partial blockage or problems with the pump (low flow levels) or leakages or measurement errors (in high flow levels).

4.3. Deviation evaluation

The objective of this module is to classify the deviation between the measured flow observed at the origin and at the destination as acceptable, thus giving evidence to show that the transference process is in regular operation; or if the results are above the acceptable level, to indicate a measuring device fault or a leakage. As described in Section 3, the deviation tolerance is related to the transient observed. It is thus expected that, in the case of 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 definition of parameters, follow.

4.3.1. Input variables and fuzzy linguistic terms

Based on the discussion above we defined as variables relevant to the solution of the problem, the deviations, deviation and devmean, labeled respectively in Eqs. (5) and (7), and the output obtained in the previous module. The use of the output from the previous module as opposed to some variable that directly indicates the transient, simplifies the system in terms of rule application.

DEV Deviation classification (zero, low, normal, high).

VN Very negative (trapezoidal function, parameters [deva deva devb devc])
N Negative (triangular function, parameters [devb devc devd])
Z Zero (triangular function, parameters [devc devd deve])
P Positive (triangular function, parameters [devd deve devf])
VP Very positive (trapezoidal function, parameters [deve devf devg devg])
EvalPhase Non-defuzzyficated output defined in the State Recognition module.

These functions are graphically represented in Fig. 13.

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

deva and devg are boundaries of the domain universe for the variable deviation defined in Eq. (5). Their values were established as −100 and 100, respectively. This definition is identical for the parameters devma and devmg associated with the variable devmean defined in Eq. (11).

devb, devc, devd, deve and devf were defined as follows, considering the normal distribution observed in the series, found in Section 4:

$$\text{deva} = -100;\quad \text{devb} = \min\left(\text{desv}|_{\text{OT}}\right);$$
$$\text{devc} = \min\left(\text{desv}|_{\text{SS}}\right);\quad \text{devd} = \overline{\text{dev}}|_{\text{SS}};$$

$$\text{deve} = \max\left(\text{desv}|_{\text{SS}}\right);\ \text{devf} = \max\left(\text{desv}|_{\text{OT}}\right);$$
$$\text{devg} = 100$$

where:

qm Time series of the deviation measured between the origin and destination;
$\overline{\text{dev}}|_{\text{SS}}$ dev series average in steady state conditions;

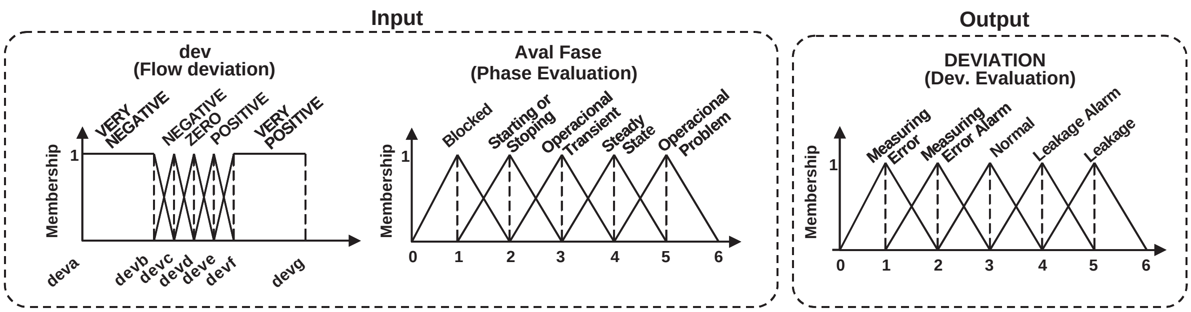

Fig. 13. Input and output variables associated with the deviation evaluation.

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

dev$|_{OT}$ dev series average in operational transient conditions.

4.3.2. Output variable and linguistic terms

The output variable is the fuzzy linguistic variable DEVIATION, to which are associated five linguistic terms, corresponding to each of the failure diagnoses: Measuring Error Alarm, Measuring Error, Normal, Leakage Alarm, Leakage. To each term are associated membership functions. The functions are presented as follows:

DEVIATION Deviation classification (measuring error alarm, measuring error, normal, leakage alarm, leakage)
ME Measuring error
MEA Measuring error alarm
N Normal
LA Leakage alarm
L Leakage.

These functions are graphically represented in Fig. 13.

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 Table 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.

5. Results

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

5.1. Results from the state recognition module: flow evaluation

Comparing the evaluation made by the system and by the specialist, a 93.67% correctness was obtained and most of the divergences were caused by the differ-

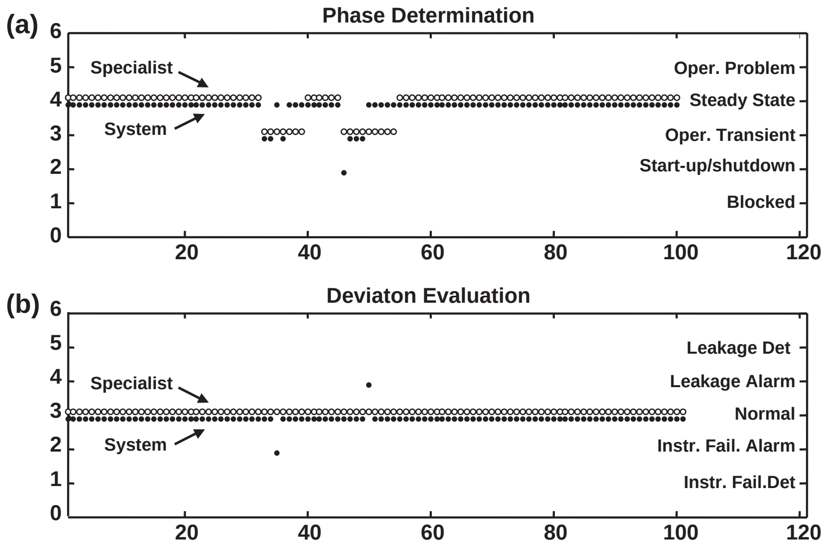

Fig. 14. Comparison of the system and the specialist's evaluation for the (a) Phase Determination and the (b) Deviation Evaluation.

ence between the system's and the specialist's conclusions during the transition of two close phases.

5.2. Results from the deviation evaluation module: deviation evaluation

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

In Fig. 14 we present a graph plotted from a specific section of the pumping operation, comparing the system and the specialist's evaluation.

6. Conclusions

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

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, working towards the elaboration of new, more robust and flexible systems, applicable to diverse transference processes.

References

Baptista, R., Roqueiro, N., Barañano, A., 2001. A study of pressure transients generated by leak in a multiproduct liquid pipeline, Brazilian Petroleum and Gas Institute—IBP, IBP04701, Nov.

Belsito, S., Lombardi, P., Andreussi, P., et al., 1998. Leak detection in liquefied gas pipelines by artificial neural networks. AIChE J. 44 (12), 2675–2688.

Caputo, A., Pelagagge, P., 2002. An inverse approach for piping networks monitoring. 4th International Conference on Inverse Problems in Engineering, Rio de Janeiro, Brazil.

Coelho, R.M.L., Medeiros, J.L., 1999. Detecção de Vazamento em Redes de Escoamento Incompressível via Reconciliação Não Linear. Anais do II ENPROMER—II Congresso de Engenharia de Processos do Mercosul, Florianópolis, Brasil, October.

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

Costa, D., Stoianov, I., Butler, D., Marksimovic, C., Graham, N.J.D., Ramos, H., 2001. Leakage detection in pipeline systems by inverse transient analysis: from Theory to Practice. International Conference on Computing and Control for the Water Industry. Leicester, U.K.

Ellul, I., 1989. Pipeline leak detection. Chem. Eng., 40–45 (June).

Fantozzi, M., 2000. Acoustic emission technique the optimum solution for leakage detection and location on water pipelines, captured in October 23, 2002. Online at the following address: http://www.ndt.net/article/wcndt00/papers/idn183/idn183.htm.

Jonsson, L., Larson, M., 1992. Leak Detections Through Hydraulic Transient Analysis, Pipeline Systems. Kluwer Academic Publishers, Dordrecht, Holland.

Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man-Mach. Stud. 7 (1), 1–13.

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

Neto, J.P.P., 2002. Modelagem dinâmica em redes de escoamento compressível para aplicações à detecção de vazamentos em tempo real. Petrobras Technical Bulletin, Rio de Janeiro, vol. 45(2) Apr/Jun (in Portuguese).

Parry, B., MacTaggart, R., Toerper, C. 1992. Compensated volume balance leak detection on a batched LPG Pipeline, OMAE, v. V–B, pp. 501–507, Pipeline Technology, ASME.

Pedrycz, W., Gomide, F., 1998. An Introduction to Fuzzy Sets. MIT Press, pp. 221–261.

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

Siebert, H., 1981. A simple method for detecting and locating small leaks in gas pipelines. Process Autom., 90–95.

Silk, M., Carter, P., 1996. A review of means of pipeline leak detection independent of flow, measurement, DEPIRE report, TEC-T031-01, AEA Technology, U.K.

Stouffs, P., Giot, M., 1993. Pipeline leak detection based on mass balance: importance of packing term. J. Loss Prev. Process Ind. 6 (5), 307–312.

Taillefond, N., Wolkenhauer, O., 2002. Fuzzy clustering and classification for automated leak detection systems. 15th Triennial World Congress, Barcelona, Spain.

Wang, G., Dong, D., Fang, C., 1993. Leak detection for transport pipelines based on autoregressive modeling. IEEE Trans. Instrum. Meas. 42 (1), 68–70 (Feb.).

Zhang, X.J., 1992. Statistical Method for Detection and Localization of Leaks in Pipeline. OMAE, v. V–B, Pipeline TechnologyASME, New York.