Leakage detection and location in gas pipelines through an LPV identification approach

P. Lopes dos Santos, T-P Azevedo-Perdicoúlis, G. Jank, J.A. Ramos, J.L. Martins de Carvalho

1. Introduction

Leak detection and location is one of the paramount concerns of pipeline operators all over the world. A timely evaluation and response to a leak, allows proper management of the consequences and an effective risk minimisation. Present methods for gas leakage detection range from manual inspection using trained dogs to advanced satellite imaging. They can be classified as acoustic monitoring, optical monitoring, gas sampling, soil monitoring, flow monitoring and model-based methods. Flow monitoring and model based methods are widely used in the gas industry. Both continuously measure the pressure and/or massflow signals at different sections of the pipeline (mostly only at the extremes). Leaks are detected from the massflow balance equations, which consist in balancing the flow at the boundaries plus the variation of linepack (LP), i.e. the amount of gas stored in the pipes. However, a considerable drawback is the LP model being strongly dependent on the noise of the pressure/temperature measurements. In [5], corrections to the LP model in the pipelines are used to obtain a more robust method. Although in [6] it is claimed that the Simone® simulator allows for calculating accurately the LP, even under extreme conditions, these methods cannot be considered completely reliable since a significant number of false alarm rates is registered everyday. This is mainly due to an integral term of the balance equation, which integrates the massflow...
difference at the boundaries. These flow measurements are corrupted by noise which will be also integrated, introducing a random walk term in the balance equation. This term is a non-stationary stochastic process with a variance proportional to the integration time. As a result, the balance equation is always corrupted by a significant amount of noise that can easily trigger false alarms. Some model based methods avoid this problem by using state observers with the fault parameters treated as augmented states [7,8]. Although this is an appealing approach, the models used so far were too complex, giving rise to estimators with too high computational costs.

In [9], the pipeline is modelled as an LPV system with the pressure as the scheduling signal. The model was identified from operational data using an algorithm described in [10,11]. The leakage is detected with a Kalman filter where the fault is treated as augmented states [7,8]. Although this is an appealing approach, the models used so far were too complex, giving rise to estimators with too high computational costs.

In [12], a similar approach was proposed. The pipeline was also modelled as an LPV system driven by the source node massflow but with the LP as the scheduling parameter. Given that the gas LP can be estimated from the massflow balance equation, a differential method is proposed to improve the leakage detector effectiveness. The proposed LPV Kalman filter based methods were compared with a standard mass balance method in a simulated 10% leakage detection scenario. The differential Kalman filter method proved to be highly efficient.

In a previous analysis of the gas dynamics within the pipeline, it was observed that a leakage triggers a pressure wave in both directions at sound speed [13,14]. Therefore, in order to locate the leakage, we need to detect at what time instants these pressure waves reach the pipeline ends. So, to include leak location in the approach described in [12], two detectors are needed, one at each end of the pipeline. For each detector, the linepack scheduling signal used in [12] is replaced by the pressure: the intake pressure for the intake detector and the offtake pressure for the other. The pressures are calculated using an approximation of the lumped transfer function model for high pressure natural gas pipelines derived in [15]; there, starting with a PDE model, a high order continuous state space linear model is obtained using a finite difference method. Next, from the SS representation an infinite order transfer function (TF) model is calculated. In the end, this TF is approximated by a compact non-rational function. This compact non-rational function may be further approximated by a simple integral model [13].

Through this paper, a case study was used to present the leak detection and location method. In this case study, a leakage was detected in a pipeline 36 km long, where the leak happens in an intermediate point. Since massflow are set at both ends of the pipeline, the leakage is detected using only pressure variations. In [12], a 36 km section of a much longer pipeline was considered and the leak occurred at the section output. Due to the length of the pipeline, the pressure variations in [12] were negligible during the leakage detection time interval. That is, the leakage was detected using only the massflow variation.

In this article, in Section 2 the model is identified and the output is simulated using a Kalman filter. In Section 3, the two leakage detectors are described. Each leakage detector is essentially a Kalman filter with an additional state variable mimicking a leakage modelled as a random walk. In Section 4, an interactive leakage locator is deduced by comparison of the two leakage detectors.

In Section 5, we withdraw some conclusions and point out some directions along which we would like the work to proceed.

2. Representation of the gas dynamics as an LPV model

The gas dynamics in a pipeline may be represented by the following hyperbolic partial differential equations (Nieplocha:1):

\[
\begin{align*}
\frac{\partial Q(s,t)}{\partial t} &= -S \frac{\partial P(s,t)}{\partial s} - \frac{\rho \rho_0}{\rho} \frac{\partial^2 Q(s,t)}{\partial s^2}, \\
\frac{\partial P(s,t)}{\partial t} &= -c^2 \frac{\partial Q(s,t)}{\partial s},
\end{align*}
\]

where \( s \) is space, \( t \) is time, \( P \) is edge pressure-drop, \( Q \) is massflow, \( S \) is the cross-sectional area, \( d \) is the pipe diameter, \( c \) is the isothermal speed of sound, and \( \rho \) is the friction factor. In the figure below, \( Q_i \) is the intake massflow and \( Q_o \) the offtake massflow and \( P_i \) is the intake pressure and \( P_o \) the offtake pressure.

The time pressure variations are assumed to have a unit correlation coefficient along the pipeline, and then are all proportional to a function \( \hat{p}(t) \), i.e., \( \hat{p}(s,t) = \hat{p}(t) \). Since the pressure varies slowly, this is a reasonable assumption for short length pipelines, i.e., ca. 35–50 km. Under this assumption, a discrete LPV model with affine parameter dependence, with \( T_s \) as the sampling period, was obtained [9]. Hence:

3. Gas pipeline LPV identification

In this section, an LPV model was identified from a mixture of measured and simulated data of a gas pipeline depicted in the Fig. 1 below:

Operational field measurements are all the intake/offtake massflows as well as the pressures along the entire network. These are available to Simone\textsuperscript{\textregistered}, a simulator installed at REN–Gasodutos headquarters, through a SCADA system. Intermediate flow rate measurements are not available and need to be simulated. These are calculated by Simone\textsuperscript{\textregistered} from the intake/offtake measured massflows. Simone\textsuperscript{\textregistered} also computes the pressures along the network, which are compared with the measured ones to assess the simulator performance [16,17]. In our example, we used measured values for the pressures and for the intake massflow (a source point) and simulated values for the offtake flow. To simulate a leakage, we calculate the pressure drop, and then subtracted it from the measured pressure values.

For the case study, we consider a cylindrical pipeline with a diameter of $d = 793$ mm, a length of $L = 36$ km, and a roughness factor of $k = 0.005$ mm. The TERMINAL\_A, JCT\_LEAK, and BV\_12400\_A are the pipeline intake node, the simulated leakage point, and the pipeline offtake node, respectively. The simulation reproduces a working gas day (March 2, 2009), in the closed interval $[0 \text{ h}, 24 \text{ h}]$, with no leakages, and at the constant temperature of $18.5 \degree C$. The data was collected with a sampling rate of 2 minutes.

Fig. 2 depicts a working day data, i.e., the profiles of the intake and offtake pressure and massflow. We can see that the pressure time pattern seems to be the same for both endpoints of the pipeline. In fact it presents a correlation coefficient value of 0.9998, which validates the pressure proportionality assumption.

For the first leakage detector the pipeline is modelled as the discrete LPV system, as in (2), using the time varying component (ac component) of the intake pressure as the scheduling signal. The second one uses the offtake pressure. The LPV

\[ x(k+1) = A_0 x(k) + A_p \bar{p}(k) x(k) + B_0 u(k) + B_p \bar{p}(k) u(k), \]
\[ y(k) = C_0 x(k) + C_p \bar{p}(k) x(k) + D_0 u(k) + D_p \bar{p}(k) u(k). \]
models were identified with the Successive Approximations Subspace Identification Algorithm in [10] with the intake/off-take pressure as the scheduling parameter. These pressures were computed from:

\[ P_{ic}(t) = \frac{K_G}{a} Q_i(t) + K_G \int_{-\infty}^{t} Q_i(\tau) d\tau - \frac{K_G}{a} Q_o(t) (t - T_L) - K_G \int_{-\infty}^{t} Q_o(\tau) d\tau, \]  

\[ P_{oc}(t) = \frac{K_G}{a} Q_i(t) - T_L) + K_G \int_{-\infty}^{t} Q_i(\tau - T_L) d\tau - \frac{K_G}{a} Q_o(t) (t - T_L) - K_G \int_{-\infty}^{t} Q_o(\tau) d\tau, \]

with \( T_L = \frac{1}{2} \) and parameters \( K_G, a \) being estimated from the data. See [13] for the derivation of the equations.

Fig. 3 compares the calculated with the measured values for both the intake and offtake pressures.

The LPV identification algorithm estimated two innovation models of the form:

\[
\begin{align*}
x(k + 1) &= A_0 x(k) + B_0 u(k) + A_p [\bar{p}(k) x(k)] + B_p [\bar{p}(k) u(k)] + K e(k), \\
y(k) &= C_0 x(k) + D_0 u(k) + C_p [\bar{p}(k) x(k)] + D_p [\bar{p}(k) u(k)] + e(k),
\end{align*}
\]

with \( u(k) = Q_i(k), y(k) = Q_o(k) \) and \( \bar{p}(k) = P_i(k) - \bar{P}, \) or \( \bar{p}(k) = P_o(k) - \bar{P} \), where \( \bar{P} \) means the average value along time. \( e(k) \) is the zero mean white noise.

With the calculated intake pressure as the scheduling signal the following parameters were obtained:

\[
\begin{align*}
A_0 &= 0.9661, & B_0 &= -0.1290 \times 10^{-1}, & C_0 &= -2.0120, \\
D_0 &= 0.2101, & B_p &= -0.5977 \times 10^{-3}, & A_p &= 0.3000 \times 10^{-2}, \\
C_p &= 5.370, & D_p &= -0.1340 \times 10^{-3}, & K_i &= -0.1540.
\end{align*}
\]

The left frame of Fig. 4 compares the simulated offtake massflow with its true value. The right frame compares the predicted offtake massflow with its true value.

When considering the calculated offtake pressure as the scheduling signal, the following parameters were obtained:

\[
\begin{align*}
K_{nm^3/h} &= 550, & Q_{i}(t), (\text{full line}) \text{ and } y_{i}(t), (\text{dotted line}) \text{ offtake massflow. Right: true } Q_{i}(t), (\text{full line}) \text{ and predicted } y_{i}(t), (\text{dotted line}) \text{ offtake massflow. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)}
\end{align*}
\]
Fig. 5. Left: true $Q_0(t)$, (dotted line) and simulated $y_{so}(t)$, (full line) offtake massflow. Right: true $Q_0(t)$, (dotted line) and predicted $y_0(t)$, (full line) offtake massflow. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

$$A_0 = 0.9528, \quad B_0 = -0.191 \times 10^{-1}, \quad C_0 = -2.1137,$$
$$D_0 = 0.1263, \quad B_p = -0.6013 \times 10^{-3}, \quad A_p = 0.2100 \times 10^{-2},$$
$$C_p = 6.550, \quad D_p = -0.393 \times 10^{-1}, \quad K_0 = -0.1570.$$

The left frame of Fig. 5 compares the simulated and the predicted values with the true ones.

4. Leakage detection

In this section, a leakage was simulated by subtracting to the offtake massflow 10% from its mean value. A method that uses a Kalman filter, built from an identified first order model, is described next. Thus, consider:

$$x_{\text{leak}}(k + 1) = x_{\text{leak}}(k) + e_{\text{leak}}(k),$$

$$x(k + 1) = A_0 x(k) + B_0 u(k) + A_p \hat{p}(k)x(k) + B_p \hat{p}(k)u(k) + Ke(k),$$

$$y(k) = x_{\text{leak}}(k) + C_0 x(k) + D_0 u(k) + C_p \hat{p}(k)x(k) + D_p \hat{p}(k)u(k) + e(k),$$

where $A_0, B_0, C_0, D_0, A_p, B_p, C_p, D_p$ are parameters of the identified model. $e_{\text{leak}}(k)$ is also a zero mean white noise term, not correlated with $e(k)$, whose variance is a design parameter. $x_{\text{leak}}(k)$ is the leakage detection signal and should be different from zero only in case of leakage.

This model has an additional state that is supposed to remain close to zero when there is no leakage. When a leakage occurs it should take the leakage value. From this idea, a differential Kalman filter based method was derived. This method consists in using two detectors and is identical to the one presented in [18], but the scheduling signals are now first the calculated pressure as the scheduling signal and the second uses the measured pressure. Since the first filter uses the calculated pressure it can never detect a leakage. Instead, it works as a reference signal. This filter leakage state is continuously compared with the corresponding state of the second filter, the one that uses the measured pressure as the scheduling signal. When there is no leakage, these states remain close to each other, but when a leakage occurs their difference takes the leakage value. In what follows, it can be seen, this method is very fast and accurate, and also well suited to detect small leakages.

We first describe the leakage detector whose scheduling signal is the intake pressure. The other leakage detector, i.e. the one that uses the offtake pressure as the scheduling signal is identical in every detail except for the chosen scheduling signal.

The calculated intake pressure, $P_{ic}$, is obtained from (3). Given that the leakage does not appear in this equation and the pressure solely depends on the intake/offtake massflows, the calculated scheduling signal always considers that there is no leakage in the pipeline. Thence, if we generate the scheduling parameter from this signal, a non leakage LPV model will always be identified and the Kalman filter should never detect a leakage. As a result, one should use this Kalman filter leakage estimate as a reference signal denoted by $x_{\text{leak}}^{\text{ref}}(k)$. In parallel, one must run another instance of this Kalman filter using a scheduling signal generated from the measured intake pressure, i.e., $P_{im}(k)$. This Kalman filter leakage estimate is denoted by $\hat{x}_{\text{leak}}(k)$. In the absence of leakage, both leakage estimates should be equal to zero. But, when a leakage occurs $x_{\text{leak}}^{\text{ref}}(k)$ remains zero and $\hat{x}_{\text{leak}}(k)$ takes the value of the leakage. As such, the signal

$$\hat{x}_{\text{leak}}(k) = x_{\text{leak}}^{\text{ref}}(k) - \hat{x}_{\text{leak}}(k).$$

is leak sensitive and will be used to detect leakages. This signal depicted in Fig. 6 represents a non faulty situation.
It has an expected value of $M_i = -0.1817 \text{K nm}^3/\text{h}$ and a standard deviation $\sigma_i = 0.5631 \text{K nm}^3/\text{h}$. From these values, upper and lower bounds of $T_{\text{upper}} = M_i + 3\sigma_i = 1.5075$ and $T_{\text{lower}} = M_i - 3\sigma_i = -1.8709$ were defined. A leakage is detected when this signal leaves the interval defined by these detection thresholds.

To illustrate this method, a leakage has been simulated at $t = 12 \text{ h}$ and at the distance of 12 km from the intake node. The left frame of Fig. 7 shows the leakage massflow (ca. 10% massflow) and the right frame of Fig. 7 shows the respective pressure reduction at the intake node of the pipeline. $x_{\text{leak}}(k)$ was estimated for this leakage scenario and then $\hat{x}_{\text{leak}}(k)$ was calculated. Fig. 8 shows the bounds just defined to be adequate for the detection of the leak. Recall that the leakage occurred at $t_{\text{leak}} = 12 \text{ h}$ with its influence being felt at $t_{\text{leak}} = 12\text{ h00}/20\text{h}$. Detection was done at $t_i = 12\text{ h12}$. So, it took 11‘40” to be detected.
To obtain the offtake pressure leakage detector, as we follow exactly the same procedure as for the intake pressure leakage detector, a leakage has been simulated in equal conditions. Its influence was felt at \( t_{\text{leako}} = 12 \, \text{h} 00 \). Detection was done at \( t_{\text{leak}} = 12 \, \text{h} 10 \). The detector took 9020 to find the leak (see Fig. 9).

The difference between the time instants, \( t_{\text{leaki}} \) and \( t_{\text{leako}} \), when the endpoints pressures begin to drop due to a leak, is a linear function of the leak location. However, this does not coincide with \( t_{\text{leak}} - t_{\text{leak}} \) due to factors such as noise and the different filters dynamics. As the leakage location is not a known function of the \( t_{\text{leak}} - t_{\text{leak}} \), we need more accurate measures of \( t_{\text{leaki}} - t_{\text{leako}} \) to locate a leakage. This will be done in the next section where an interactive methodology for leakage location is proposed.

5. Leakage location

Figs. 8 and 9 show that the signals \( \tilde{x}_{\text{leaki}}(k) \) and \( \tilde{x}_{\text{leako}}(k) \) are smooth before the leakage is detected and a sudden change occurs with the signal becoming more variable once the leakage is perceived. The idea is to use this variation to estimate \( t_{\text{leak}} \) and \( t_{\text{leak}} \). This sudden change causes a pulse at each differential signal:

\[
\delta \tilde{x}_{\text{leaki}}(k) = \tilde{x}_{\text{leaki}}(k) - \tilde{x}_{\text{leaki}}(k - 1),
\]

\[
\delta \tilde{x}_{\text{leako}}(k) = \tilde{x}_{\text{leako}}(k) - \tilde{x}_{\text{leako}}(k - 1).
\]

From the left frame of Fig. 10, we notice that this pulse is masked by the measurement noise. However the same pulse is perceptible in the right frame of Fig. 10, since this is a magnification of the left frame around the instance that the leakage occurs. Here, a jump in \( \delta \tilde{x}_{\text{leak}}(k) \) is clear.

To detect the leakage, we define a threshold:

\[
T_{\delta} = \delta M_i + 2\delta \sigma_i,
\]

where \( \delta M_i \) and \( 2\delta \sigma_i \) are, respectively, the expected value and the standard deviation of \( \delta \tilde{x}_{\text{leak}}(t) \) before the leak is detected. The following values were obtained for these parameters: \( \delta M_i = -0.0011 \), \( \delta \sigma_i = 0.1608 \) and \( T_{\delta} = 0.3206 \). From the observation

---

**Fig. 9.** Leak detection using the offtake node pressure as scheduling signal.

---

**Fig. 10.** Left: \( \tilde{x}_{\text{leak}}(t) \) and its \( \delta \tilde{x}_{\text{leak}}(t) \). Right: \( \tilde{x}_{\text{leak}}(t) \) and its \( \delta \tilde{x}_{\text{leak}}(t) \) near the the time instant where the leak occurs. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

---

of Fig. 10 one may expect a considerable number of false alarms whenever this threshold is adopted. This is confirmed by Fig. 11 were the leakage alarms are shown (the red).

One may also observe that the density of such points increases significantly immediately before $\dot{x}_{\text{leak}}(t)$ overtakes the leakage detection threshold. Consequently, we may select the first of these points as the first instant that the leakage is perceived by the intake node. That is $t_{\text{leak}} = 12\,\text{h04}$, and this is exactly the sampling instant after the leakage is perceived by the intake pressure.

An identical procedure has been adopted to find the first instant at which the leakage is perceived by the offtake pressure. The following values correspond to the expected value and standard deviation of $\dot{x}_{\text{leak}}(k)$, respectively: $\delta M_0 = -0.0019$, $\delta \sigma_1 = 0.2540$ and $T_0 = 0.3206$ and these values lead to the threshold $T_0 = 0.5061$.

Fig. 12 shows the leakage marks obtained from $\dot{x}_{\text{leak}}(k)$, which are also more dense immediately before the leakage is detected by the signal $x_{\text{leak}}(k)$. Four false alarms were registered before this sequence started at instant $t_{\text{leak}} = 12\,\text{h06}$, i.e., the instant immediately after the leakage is perceived by the offtake pressure.

Consider now $x$ the distance from the leakage to the intake node, $L$ the length of the pipe and $c$ the speed of the wave pressure. The wave pressure caused by the leakage takes $t_{\text{leak}} = \frac{x}{c}$ sec to be felt by the intake node and $t_{\text{leak}} = \frac{x}{c}$ to be felt by the offtake node.
As \( t_{\text{leak}} - t_{\text{tank}} = t_e - t_i = 2r = 120 \) s and \( c = 300 \) m/s then \( x = 0 \). This means that the leakage took place at the intake node. As a matter of fact, the leakage happens at \( x = 12 \) km! However as a sampling period of \( 120 \) s is the time it takes for a pressure wave to cross \( 36 \) km, the better available resolution is \( 18 \) Km. In order to achieve better resolutions, a finer sampling period is required. This is not viable due to technical/equipment limitations at REN-Gasodutos.

6. Conclusions

In this paper, an LPV differential Kalman filter leakage detector is proposed. Two identical detectors are run simultaneously. One considers the scheduling signal as the intake pressure and the other the offtake pressure. When these two detectors run simultaneously a location procedure becomes possible, since the leak location is a linear function of the difference between the time instants the leak is perceived at the pipe endpoints. Based on this fact, an interactive methodology for leakage location has also been presented.

The methodology has been tested with data supplied by REN-Gasodutos, however accuracy of the leakage locator was limited by the long sampling periods possible at REN. The application of the same methodology to more complex pipelines will be considered in the near future with higher sampling rates.

Acknowledgments

This work was supported in part by FCT, Fundação para a Ciência e Tecnologia.

References