Diagnostic Models and Estimators for LDI in Transmission Pipelines

Zdzisław Kowalczuk * **Marek S. Tatara** **

Department of Robotics and Decision Systems, Faculty of ETI, Gdańsk University of Technology, Gdańsk, Poland
* *e-mail: kova@eti.pg.edu.pl*
** *e-mail: marek.tatara@pg.edu.pl*

Abstract: This article considers and compares four analytical models of the pipeline flow process for leak detection and location tasks. The synthesis of these models is briefly outlined. Next, the methodology for generating data and diagnosing pipes is described, as well as experimental settings, assumptions and implemented scenarios. Finally, the quality of model-based diagnostic estimators has been evaluated for their bias, standard deviations and computational complexity. The global level of optimality served as a general indicator of the quality and performance of multidimensional estimators.

Keywords: Leak Detection and Isolation, Diagnosis, Transmission Pipelines, Fault Diagnosis

1. INTRODUCTION

One of the most efficient ways to transfer a fluid medium over a long distance is to use transmission pipelines. During such transport, a leak may occur and cause financial losses, pollute the environment or endanger people. Therefore, it is necessary to implement a leak detection and isolation system (LDI), whose task is to detect potential leaks and identify their parameters. The following briefly introduces the methodology used with a view to using the simplest models possible.

2. ANALYZED MODELS

Four models will be analyzed and compared. Using a commonly accepted model as a reference, we develop three other models, which describe the process of isothermal flow of incompressible fluid through a pipeline.

2.1 Base Model

The first model, referred to as the Base Model, is derived from the following set of PDEs (Partial Differential Equations) (Billmann and Isermann, 1987):

$$\frac{S}{\nu^2}\frac{\partial p}{\partial t}+\frac{\partial q}{\partial z}=0 \tag{1}$$

$$\frac{1}{S}\frac{\partial q}{\partial t}+\frac{\partial p}{\partial z}=-\frac{\lambda\nu^2}{2DS^2}\frac{q|q|}{p}-\frac{g\sin\alpha}{\nu^2}p \tag{2}$$

where S is the cross-sectional area [m^2], ν is the surrogate velocity associated with the isothermal speed of sound in the fluid [$\frac{\mathrm{m}}{\mathrm{s}}$], D is the diameter of the pipe [m], q is the mass flow [$\frac{\mathrm{kg}}{\mathrm{s}}$], p is the pressure [Pa], t is the time [s], z is the spatial coordinate [m], λ is the generalized dimensionless friction factor, α is the angle of inclination [rad], and g is the gravitational acceleration [$\frac{\mathrm{m}}{\mathrm{s}^2}$].

The above model, after discretization, leads to the following nonlinear singular equation in the state space:

$$\mathbb{A}\hat{\mathbf{x}}^k=\mathbb{B}\hat{\mathbf{x}}^{k-2}+\mathbb{C}\left(\hat{\mathbf{x}}^{k-1}\right)\hat{\mathbf{x}}^{k-1}+\mathbb{D}\mathbf{u}^{k-1}+\mathbb{E}\mathbf{u}^k \tag{3}$$

where the state and input vectors are defined as

$$\hat{\mathbf{x}}^k=\left[\,\hat{q}_0^k\;\hat{q}_2^k\;\hat{q}_4^k\;\cdots\hat{q}_N^k\;\hat{p}_1^k\;\hat{p}_3^k\;\hat{p}_5^k\cdots\hat{p}_{N-1}^k\,\right]^{\mathrm{T}}$$

$$\mathbf{u}^k=\left[\,p_0^k\;p_N^k\,\right]^{\mathrm{T}}$$

where p is the pressure, and q is the mass-flow rate (Billmann and Isermann, 1987). The description of the matrices can be found in (Gunawickrama, 2001). Subscripts denote the spatial coordinate number, while superscripts denote time indices. Hat symbols indicate estimates.

2.2 Analytic Model of Diagonal Approximation

The second model is derived on the basis of the assumption that (3) can be shown in the following nonsingular form:

$$\hat{\mathbf{x}}^k=\mathbb{A}^{-1}\left(\mathbb{B}\hat{\mathbf{x}}^{k-2}+\mathbb{C}\left(\hat{\mathbf{x}}^{k-1}\right)\hat{\mathbf{x}}^{k-1}+\mathbb{D}\mathbf{u}^{k-1}+\mathbb{E}\mathbf{u}^k\right) \tag{4}$$

where the recombination (or descriptor) matrix $\mathbb{A}$ is inverted using matrix partitioning (Brogan, 1991).

For the specific structure of the matrix $\mathbb{A}$, it is possible to find an approximation of its inverse analytically - by substituting tridiagonal matrices by their diagonal counterparts (Kowalczuk and Tatara, 2016). The model associated with the inverse of approximate $\mathbb{A}$ is called the analytical diagonal approximation model (AMDA).

2.3 Analytic Thomas Model

To reduce the computational complexity of calculations, the sparsity of the matrices in the state-space model (3) allows us to rearrange this model in the form applicable to the Thomas algorithm (Conte and de Boor, 1980). Model (3) can be eventually shown as Kowalczuk et al. (2018).

By defining: $\tilde{\mathbf{g}} = \frac{\mathbf{g}}{c}$, $\tilde{\mathbb{A}}_2 = \frac{\mathbb{A}_2}{c}$, $\tilde{\mathbf{h}} = \frac{\mathbf{h}}{a}$ and $\tilde{\mathbb{A}}_3 = \frac{\mathbb{A}_3}{a}$ (where a and c are physical parameters), one eventually obtains

$$\left(\mathbb{I} - \tilde{\mathbb{A}}_2\tilde{\mathbb{A}}_3\right)\mathbf{q}^k = \tilde{\mathbf{g}} - \tilde{\mathbb{A}}_2\tilde{\mathbf{h}} \tag{5}$$

$$\left(\mathbb{I} - \tilde{\mathbb{A}}_3\tilde{\mathbb{A}}_2\right)\mathbf{p}^k = \tilde{\mathbf{h}} - \tilde{\mathbb{A}}_3\tilde{\mathbf{g}} \tag{6}$$

The results of the matrix multiplications $\tilde{\mathbb{A}}_3\tilde{\mathbb{A}}_2$ and $\tilde{\mathbb{A}}_2\tilde{\mathbb{A}}_3$ are tridiagonal, therefore the expressions $\left(\mathbb{I} - \tilde{\mathbb{A}}_3\tilde{\mathbb{A}}_2\right)$ and $\left(\mathbb{I} - \tilde{\mathbb{A}}_2\tilde{\mathbb{A}}_3\right)$ are also tridiagonal matrices and the Thomas algorithm (Thomas, 1949) can be applied. This model will be called the Analytic Thomas Model (ATM).

2.4 Model of Steady State

Let us assume that the LDI system works on a pipeline in steady flow. Thus, we consider $\frac{\partial q}{\partial t} \to 0$ and $\frac{\partial p}{\partial t} \to 0$ in (1) and (2) as (approximately) satisfied for steady flow. In such cases, the PDEs are reduced to ordinary differential equations that can be solved for constant mass-flow rate and pressure distribution along the pipe. The solution can be provided separately for zero inclination angle as constant mass-flow rate and pressure distribution:

$$q = \operatorname{sign}\left(p_i^2 - p_o^2\right)\sqrt{\left|\frac{DA^2}{\lambda\nu^2}\frac{p_i^2 - p_o^2}{L}\right|}, \quad p = \sqrt{p_i^2 - \frac{p_i^2 - p_o^2}{L}z} \tag{7}$$

where $\operatorname{sign}(x)$ is 1 for $x \geq 0$, and -1 otherwise; and for non-zero inclination angle with mass-flow rate calculated as

$$q = \sqrt{\left|\frac{2DS^2}{\lambda\nu^2}\frac{g\sin\alpha}{\nu^2}\left(\frac{p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L}}{e^{2\frac{g\sin\alpha}{\nu^2}L} - 1}\right)\right|} \cdot \operatorname{sign}\left(p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L}\right) \tag{8}$$

and pressure distribution along the spatial coordinate z:

$$p = \sqrt{e^{-2\frac{g\sin\alpha}{\nu^2}z}p_i^2 + \left(\frac{p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L}}{e^{2\frac{g\sin\alpha}{\nu^2}L} - 1}\right)\left(e^{-2\frac{g\sin\alpha}{\nu^2}z} - 1\right)} \tag{9}$$

This dichotomous model, describing the steady state flow, (for zero and non-zero angle of inclination), will be jointly called the analytical model of steady state (AMSS).

3. APPLIED METHODOLOGY

For the LDI task, the four considered models (the base model, ATM, AMDA, and AMSS) were implemented to emulate the flow process as a part of a model-based LDI system. The same input data was provided for each of the models, working under identical conditions (physical flow parameters and discretization grid).

The previously validated simulator Gunawickrama (2001) was used to generate the data. This simulator is suited for isothermal and incompressible flow and is based on an adapted version of the base model to accurately generate inlet and outlet pressures and flow rates. The simulator uses the fact that a leak of q_L $[\frac{\text{kg}}{\text{s}}]$ can be included in the PDEs (1) and (2) describing the flow process.

3.1 Residual Generation and Leak Detection

After initialization, residual signals are calculated in each iteration of the diagnostic algorithm:

$$\hat{\mathbf{r}}^k = \begin{bmatrix} \hat{r}_i^k \\ \hat{r}_o^k \end{bmatrix} = \begin{bmatrix} \tilde{q}_i^k - \hat{q}_i^k \\ \tilde{q}_o^k - \hat{q}_o^k \end{bmatrix} \tag{10}$$

where $\tilde{q}_i^k$ and $\tilde{q}_o^k$ are measurements of the inlet and outlet mass-flow rates, and $\hat{q}_i^k$ and $\hat{q}_o^k$ are estimates of inlet and outlet flow rates obtained from the model, respectively.

Leak Detection We calculate the cross-correlation of residuals using a low-pass filter implemented with the forgetting factor β_c (Billmann and Isermann, 1987) as

$$\Phi_{i,o}^k(\tau) = \beta_c\Phi_{i,o}^{k-1}(\tau) + (1 - \beta_c)\,\hat{r}_i^{k-\tau}\hat{r}_o^k \tag{11}$$

Moreover, the above is summed up over all analyzed time shifts $\tau = 1, 2, \ldots, \tau_{max}$ resulting in

$$\Phi_\Sigma^k = \sum_{\tau=1}^{\tau_{max}} \Phi_{i,o}^k(\tau) \tag{12}$$

In the event of a leak, the value of the above indicator decreases. We compare it to a threshold Φ_{th}:

$$\Phi_\Sigma^k < \Phi_{th} \tag{13}$$

If the above condition is met, an alarm is triggered and the diagnostic phase occurs.

Leak Location The location of the leak can be determined on the basis of the characteristic shape of the pressure distribution along the pipe in the event of leakage. The location of the leak (z_L) can be determined as

$$\hat{z}_L^k = L\left(1 - \frac{\hat{r}_i^k\left(\tilde{q}_i^k + \hat{q}_i^k\right)}{\hat{r}_o^k\left(\tilde{q}_o^k + \hat{q}_o^k\right)}\right)^{-1} \tag{14}$$

To reduce the influence of measurement noise, another low-pass filter is implemented.

Leak Size In a leaking case the inlet mass-flow rate is increasing, while the outlet rate is decreasing, and the size of the leak can be determined as a difference between the measurements of the inlet and outlet flow rates:

$$\hat{q}_L^k = \tilde{q}_i^k - \tilde{q}_o^k = \hat{r}_i^k - \hat{r}_o^k \tag{15}$$

However, to reduce the impact of noise, we use the following improved balancing method (Gunawickrama, 2001):

$$\hat{w}_q^k = \mathrm{E}\{\hat{q}_L^k\} \tag{16}$$

where $\hat{w}_q^k$ is referred to as the leak size estimator. This approach may require some reference values (Gunawickrama, 2001) that compensate for the sensors calibration errors. Calculation of the expected value can be realized by means of low-pass filtering.

3.2 Estimation of the Friction Factor

Usually for real pipelines, the coefficient of friction is only known roughly. This is due to the complex, variant/non-stationary and non-linear nature of this quantity.

For discrete-time models (Base Model, AMDA, ATM) the friction coefficient can be estimated assuming that the mean square error between the estimated flow rate and the measured one is minimized (Gunawickrama, 2001).

Similarly, for AMSS model we calculate λ that minimizes the mass-flow squared error as:

$$\frac{\partial \left(\tilde{q}_E^k - \hat{q}^k\right)^2}{\partial \lambda} = 0 \tag{17}$$

where $\tilde{q}_E^k$ is the measurement of flow velocity at one end of the pipeline, $E \in \{i, o\}$. The resulting friction factor be computed as

$$\hat{\lambda}^k = \left(\frac{C_z}{\tilde{q}_E^k}\right)^2 \tag{18}$$

where C_z for zero inclination angle is given by

$$C_z = \sqrt{\left|\frac{DS^2}{\nu^2}\frac{p_i^2 - p_o^2}{L}\right|}\,\text{sign}\left(p_i^2 - p_o^2\right) \tag{19}$$

whereas for the non-zero inclination angle case we have

$$C_z = \sqrt{\left|\frac{2DS^2}{\nu^2}C_4\left(\frac{p_i^2 - p_o^2 e^{2C_4L}}{e^{2C_4L} - 1}\right)\right|}\,\text{sign}\left(p_i^2 - p_o^2 e^{2C_4L}\right) \tag{20}$$

Due to measurements at the inlet and outlet of the pipe, the above (18) can be associated with the respective input and output estimates of $\hat{\lambda}^k$.

Again, the effective value of the friction factor λ can be estimated (Kowalczuk and Gunawickrama (2004)) using a simple recursive filtering with a forgetting factor.

4. EVALUATION OF THE DIAGNOSTIC MODELS AND ESTIMATORS

To evaluate the proposed models in terms of their applicability to LDI systems, the influence of various parameters on the diagnosis outcome was examined. Because the system output is influenced by measurement noise, leak size, and leak location, the problem is multidimensional, and the experimental settings have to be reflected by rationally selected parameters.

Diagnostic estimates of the size and location of the leak were assessed for the bias (expected estimate error) and standard deviation of that error. The respective biases are calculated using the mean signed difference:

$$\hat{\bar{w}}_q = \frac{1}{n}\sum_{j=1}^{n}\hat{w}_{q_j} - q_L, \quad \hat{\bar{w}}_z = \frac{1}{n}\sum_{j=1}^{n}\hat{w}_{z_j} - z_L \tag{21}$$

for the estimates of the leak size and location, where n is the number of executed experiment runs. The standard deviation (STD) of estimation errors can be determined for the leak size and location estimators as

$$\hat{\sigma}_q = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(\hat{w}_{q_j} - \bar{w}_q\right)^2}, \quad \hat{\sigma}_z = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(\hat{w}_{z_j} - \bar{w}_z\right)^2} \tag{22}$$

4.1 Experimental Setup

The physical parameters of the pipeline process are given in Table 1. Simulation data were generated using these parameters, with the number of N_s segments implemented for the data generation was set to 100 (while the LDI system itself uses $N = 10$).

Another reference set of parameters concerning the computation and simulation process are shown in Table 2.

4.2 Assessment of the Quality of the Estimators

The quality of estimators is evaluated in terms of their accuracy (reflected in respective bias) and precision - reflected in their standard deviations.

In the whole leak simulation study, the evaluation was performed for 60 experimental settings: 9 values of leak size $q_L \in \{0.04, 0.1, 0.4, 0.8, 1.6, 2, 4, 8, 16\}[\frac{\text{kg}}{\text{s}}]$, and 6 discrete leak locations (in km) $z_L \in \{0.1, 4, 13.87, 40, 73.75, 98.2\}$. Moreover, 7 different noise percentages introduced to pressure readings $v_{rdg_p} \in \{0, 0.01, 0.05, 0.1, 0.5, 1, 5\}[\%]$ were taken into account, but only for $q_L = 4\,\frac{\text{kg}}{\text{s}}$ and $z_L = 40\,\text{km}$.

It is difficult to clearly indicate a best model, because there are many factors determining their quality. The problem of comparative criterion therefore arises. One of the most recognizable approaches to multi-criteria problems is optimization in the sense of Pareto, which introduces a practical ranking of solutions (Deb et al., 2003).

For obvious reasons, the computational complexity of the developed and analyzed process models can be associated with both the computational time (t_i) necessary for a single iteration, and the steady-state settling time t_s.

All the performance parameters obtained in this part of the experimental study are collected in Table 3, considering the computational parameters of Tables 1 and 2.

To make the above indicators comparable, we go to the task of maximization with normalization relative to the highest value (for each indicator separately). So 1 means best score, while lower values refer to this best result. The normalized functionals obtained are plotted in Fig. 1.

In the P-sense, AMDA, AMSS and base models are equivalent (ATM is very close to them), and you can point a diagnostic estimator that is tailored for a specific task. In general, however, all of these aspects are important,

Table 1. Physical parameters of the flow.

Parameter	**Value**
Length of the pipe L	100 km
Diameter of the pipe D	0.4 m
Speed of sound ν	350 $\frac{\text{m}}{\text{s}}$
Friction factor λ	0.02
Number of segments N	10
Inclination angle α	0°
Inlet pressure p_i	112.28 bar
Outlet pressure p_o	80 bar
Leak location z_L	40 km
Leak size q_L	4 $\frac{\text{kg}}{\text{s}}$
Leak occurrence time t_L	105.5 min
Leak development time t_d	17.5 min

Table 2. Computational parameters.

Parameter	**Value**
Percentage of pressure reading v_{rdg_p}	0.1% of reading
Percentage of mass-flow rate reading v_{rdg_q}	1% of reading
Courant number μ	0.17
Detection threshold Φ_{th}	0.01
Forgetting factors $\beta_c, \beta_z, \beta_q, \beta_\lambda$	0.99
Maximum shift τ_{max} for cross-correlation	20
Number of runs n	200

Table 3. Quality measures of the estimators of $\bar{z}_L$ (isolation) and $\bar{q}_L$ (identification).

Performance indicator	ATM	AMDA	AMSS	Base model
Bias of leak size estimator $\hat{\bar{w}}_q$	$-2.65 \cdot 10^{-1}$	$-1.45 \cdot 10^{-1}$	$-2.66 \cdot 10^{-1}$	$-2.65 \cdot 10^{-1}$
STD of leak size estimator $\hat{\sigma}_q$	$9.75 \cdot 10^{-3}$	$3.37 \cdot 10^{-1}$	$8.78 \cdot 10^{-3}$	$9.75 \cdot 10^{-3}$
Bias of leak location estimator $\hat{\bar{w}}_z$	$-1.21 \cdot 10^{4}$	$-1.85 \cdot 10^{4}$	$-2.65 \cdot 10^{4}$	$-8.21 \cdot 10^{3}$
STD of leak location estimator $\hat{\sigma}_z$	$6.70 \cdot 10^{5}$	$2.18 \cdot 10^{5}$	$1.33 \cdot 10^{5}$	$3.14 \cdot 10^{5}$
Single iteration computation time t_i [s]	$3.53 \cdot 10^{-5}$	$2.05 \cdot 10^{-4}$	$2.73 \cdot 10^{-5}$	$7.18 \cdot 10^{-4}$
Steady state computation time t_s [s]	$1.45 \cdot 10^{-2}$	$3.53 \cdot 10^{0}$	$2.73 \cdot 10^{-5}$	$2.24 \cdot 10^{-1}$

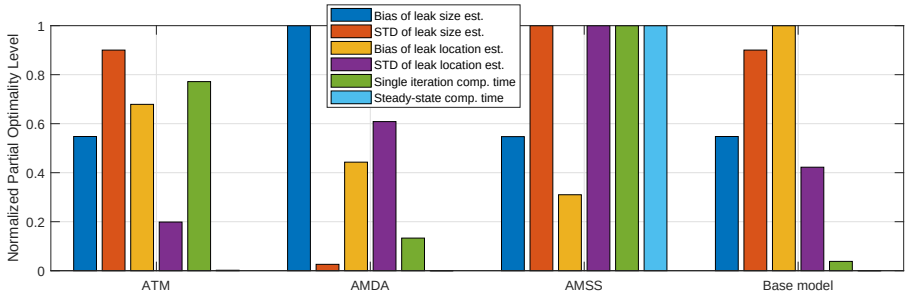

Figure 1. Normalized local optimality level indicators for the analyzed models.

then we can use a measure called Global Optimality Level (GOL) (Kowalczuk and Białaszewski, 2017). This criterion, aiming at maximizing the lowest quality indicator for each diagnostic estimator, is shown in Table 4.

Table 4. GOL for the diagnostic estimators.

	ATM	AMDA	AMSS	Base model
GOL:	$1.87 \cdot 10^{-3}$	$7.38 \cdot 10^{-6}$	0.31	$1.22 \cdot 10^{-4}$

We can see that the AMSS model is superior because its GOL is at least two orders of magnitude larger than GOL for other estimators. Nevertheless, the small values for the ATM and AMDA models and the base model result from the last quality indicator, i.e. the time required to reach steady state. The AMSS model is particularly suitable for this task (while other models are not). Thus, we get better GOL estimates after excluding this indicator, and the effect of this approach is shown in Table 5.

Table 5. Corrected GOL for the estimators.

	ATM	AMDA	AMSS	Base model
GOL:	0.20	0.03	0.31	0.04

It is now clear that the best diagnostic estimator is the one based on the AMSS. The second best is the ATM-based estimator. We also see that the AMDA model is slightly inferior to the base model.

It is important to clarify here that the quality criteria are measured with some uncertainty. Therefore, the above assessment is approximate, however, given the total number of runs (about 4000), the collective results seem justified.

5. SUMMARY

This paper has presented a comparative study analyzing different derivative models in terms of their diagnostic application. The evaluation of diagnostic estimators in terms of their bias and standard deviation has been proposed and evaluated taking into account different leak sizes, locations and measurement noise levels. Multivariate estimators resulting from the applied process models have been compared in terms of their Global Optimality Level.

The AMSS model turns out to be the best in most cases in terms of standard deviation, bias and calculation time, the AMDA estimator seems only suitable for leakage locations, and the ATM estimator, with lower computational complexity, is in most cases comparable with the basic model. In the GOL sense, the AMSS model is the best.

REFERENCES

Billmann, L. and Isermann, R. (1987). Leak detection methods for pipelines. *Automatica*, 23(3), 381–385.

Brogan, W. (1991). *Modern Control Theory. Third Edition.* Prentice Hall.

Conte, S. and de Boor, C. (1980). *Elementary Numerical Analysis: An Algorithmic Approach.* McGraw-Hill.

Deb, K., Zope, P., and Jain, A. (2003). Distributed computing of pareto-optimal solutions with evolutionary algorithms. In *Evolutionary Multi-Criterion Optimization*, 534–549. Springer, Berlin, Heidelberg.

Gunawickrama, K. (2001). *Leak Detection Methods for Transmission Pipelines.* Ph.D. thesis, GUT.

Kowalczuk, Z. and Białaszewski, T. (2017). Gender approaches to evolutionary multi-objective optimization using pre-selection of criteria. *Engineering Optimization*, Vol. 50, no 1, 120–144.

Kowalczuk, Z. and Gunawickrama, K. (2004). *Detecting and locating leaks in transmission pipelines*, 821–864. Springer, Berlin, Heidelberg.

Kowalczuk, Z. and Tatara, M. (2016). Approximate models and parameter analysis of the flow process in transmission pipelines. In *Advanced & Intelligent Comp. in Diagnosis and Control*, 239–252. Springer, Cham.

Kowalczuk, Z., Tatara, M., and Stefański, T. (2018). Reduction of computational complexity in simulations of the flow process in transmission pipelines. In *Advanced Solutions in Diagnostics and Fault Tolerant Control*, 241–252. Springer IP, Cham.

Thomas, L. (1949). Elliptic problems in linear difference equations over a network. *Watson Sci.Com.Lab. Report.*