Int. J. Appl. Math. Comput. Sci., 2024, Vol. 34, No. 3, 391–407
DOI: 10.61822/amcs-2024-0028

APPROXIMATE AND ANALYTIC FLOW MODELS FOR LEAK DETECTION AND IDENTIFICATION

MAREK S. TATARA [a,*], ZDZISŁAW KOWALCZUK [a]

[a] Department of Robotics and Decision Systems
Gdańsk University of Technology
ul. Narutowicza 11/12, 80-233 Gdańsk, Poland
e-mail: {marek.tatara,kova}@pg.edu.pl

The article presents a comprehensive quantitative comparison of four analytical models that, in different ways, describe the flow process in transmission pipelines necessary in the task of detecting and isolating leaks. First, the analyzed models are briefly presented. Then, a novel model comparison framework is introduced along with a methodology for generating data and assessing diagnostic effectiveness. The study presents basic assumptions, experimental conditions and scenarios considered. Finally, the quality of the model-based diagnostic estimators is assessed, focusing on their bias, standard deviation, and computational complexity. Here, several optimality criteria are used as detailed indicators of the quality and performance of the estimators in a multi-criteria Pareto optimality assessment.

Keywords: leak detection and isolation, transmission pipelines, fault diagnosis, mathematical modeling.

1. Introduction

One of the most effective ways of transferring a fluid medium over long distances is through the use of transmission pipelines. Environmental pollution, millions of kilometers of pipelines transporting valuable goods susceptible to theft, and chemical fluids posing potential hazards and risks of contamination require meticulous monitoring and rapid detection of leaks. This is crucial not only from a business point of view, where a leak means a significant economic loss, but also from an ecological point of view, where leaks pose a serious threat to the environment. Therefore, a comprehensive approach to monitoring and solving problems related to pipeline safety and ecological pollution is necessary, along with the implementation of leak detection and isolation (LDI) systems to detect potential leaks and identify their parameters. For this purpose, we need the simplest possible models supporting the LDI methodology.

There are few mathematical models that are sufficiently reliable and accurate and can be applied in online practice for pipeline installation and operation tasks. This work is a comparative study of several different models suitable for such tasks.

The models describing the relationship between pressure and a flow rate date back to the mid-nineteenth century. One of the first models was proposed by Hagen and Poiseuille as well as Darcy and Weisbach (Brown, 2003). These models, due to their simplicity, presented only generalized relationships and could not express the full dynamics of the process as well as the effects of the wide variety of experimental settings occurring during the pipeline flow. To solve this problem, numerous domain-specific dynamic models have been proposed, tailored to the specificity of the LDI task, as shown, for example, by Billmann and Isermann (1987). The models have evolved to cover gas (Liu *et al.*, 2005 2017; Rui *et al.*, 2017; Wii *et al.*, 2020) or liquid (Abhulimen and Susu, 2004) pipelines, or a general case of fluid transmission that can be used for both (Gunawickrama, 2001; Kowalczuk and Tatara, 2020; Torres *et al.*, 2021). Note that there are models that also focus on a more complex flow, as, for example, the water-glycerol mixture flow model (Noguera *et al.*, 2019).

There are pipeline diagnostic systems built on a data-driven approach (see, e.g., Nowicki *et al.*, 2012; Quiñones-Grueiro *et al.*, 2018). Nevertheless, the mathematical model is an effective foundational element of many leak detection concepts. There are a number

*Corresponding author

Table 1. Comparison of selected state-of-the-art models for leak detection.

Modeling approach	Variables	Multiple leaks	Branched	Source item
Static flow model	Pressure and volumetric flow	Yes (two)	No	Ostapkowicz and Bratek, 2023
Quasi steady-state flow model	Pressure head and volumetric flow	Yes	Yes	Torres *et al.*, 2021
Pressure drop model for WDN	Pressure and fluid velocity	Yes	Yes	Fererdooni *et al.*, 2021
Artificial neural network	Pressure head and volumetric flow	No	No	Pérez-Pérez *et al.*, 2021
Transient with leak modeled + EKF	Pressure and mass-flow rate	No	No	Doshmanziari *et al.*, 2020

of works covering all or some aspects of this problem (Kowalczuk and Tatara 2020; 2021; Torres *et al.*, 2021; 2015; Verde and Torres, 2015). These models are also useful as the basis for simulation and emulation systems in which we replicate physical processes that occur in the real world. They can also serve as a benchmark against which the actual process can be compared, thereby gaining insight into potential leakage phenomena. Table 1 summarizes the selected models used for the pipeline monitoring task, taking into account the underlying modeling approach, the main variables monitored, whether the approach is suitable for multiple leaks, and whether it is applicable to branch pipelines.

Since such models are not the only element of the fault/leak detection and estimation (FDI/LDI) system that affects the overall quality of diagnosis, they must be subject to various conditions and limitations in order to be objectively assessed in terms of their effectiveness.

Generally speaking, detection tells us whether a fault has occurred, isolation tells us the type and/or location of the fault/damage, and identification tells us the extent of the fault (Kościelny, 1993; Korbicz *et al.*, 1999; 2004; Gunawickrama, 2001; Kościelny *et al.*, 2016). Therefore, from the FDI perspective, estimating the location of a leak means isolating the fault, and assessing the extent of the leak means identifying the fault.

The novelties introduced in this article are as follows:

(i) comparison of four instrumental analytical and approximate flow models useful in the task of detecting, isolating and identifying leakages,

(ii) proposing a framework for the comparative testing of different models,

(iii) analysis of the pipeline process model carried out in various experimental conditions, taking into account the variable impact of noise, and

(iv) comparison of the performance of models in terms of multi-criteria Pareto optimality.

Therefore, the aim of this paper is to demonstrate and compare alternative pipeline flow models in terms of computational complexity and performance for use in a range of leak detection and isolation tasks. The proposed framework can be used to compare models under uncertainty and to demonstrate a quantitative measure of the model's noncompliance with reality. The proposed models are a proven and effective alternative to the well-known base model, because with comparable results and operational efficiency they offer lower computational complexity, which is especially important in the context of online diagnostic applications.

Our algorithmic approach makes great practical sense for implementing FDI/LDI methods based on analytical models, generally called model-based approaches. This is a typical modern computer-based method of dealing with signal processing, used in the construction of controllers, identifiers, estimators, detectors, etc. The work presented is strongly related to industrial systems processes, including chemical sciences and pipeline fluid transport systems.

The article is organized as follows. First, we will introduce the modeling methodology used for residual generation. Next, we will bring up ideas related to leak detection, isolation and identification. It is important that, with industrial implementations in mind, the methods considered in this work are based on fundamental dependencies and are shaped in such a way that they are characterized by the greatest simplicity.

2. Methodology

This work aims to computationally investigate how the different instrumental models used in the standard structure of fault detection and isolation (FDI) projects affect the resulting system performance. In order to make the presented methodology comprehensible, the general LDI computing framework used to conduct our FDI experiments is first presented. Note that this document is not intended to describe branched pipelines or multiple leak scenarios. It focuses on comparing models under uniform conditions, excluding such complexities. Additionally, to facilitate comprehensive inclusion of branches in the analysis, effort must first be put into developing a good T-junction (three-way branch) model.

The LDI computing framework shown in Fig. 1 consists of a data generation block that provides data to each of the instrumental models at the same time. After each step of the simulation process, the same residual generation and leak assessment method is applied to all

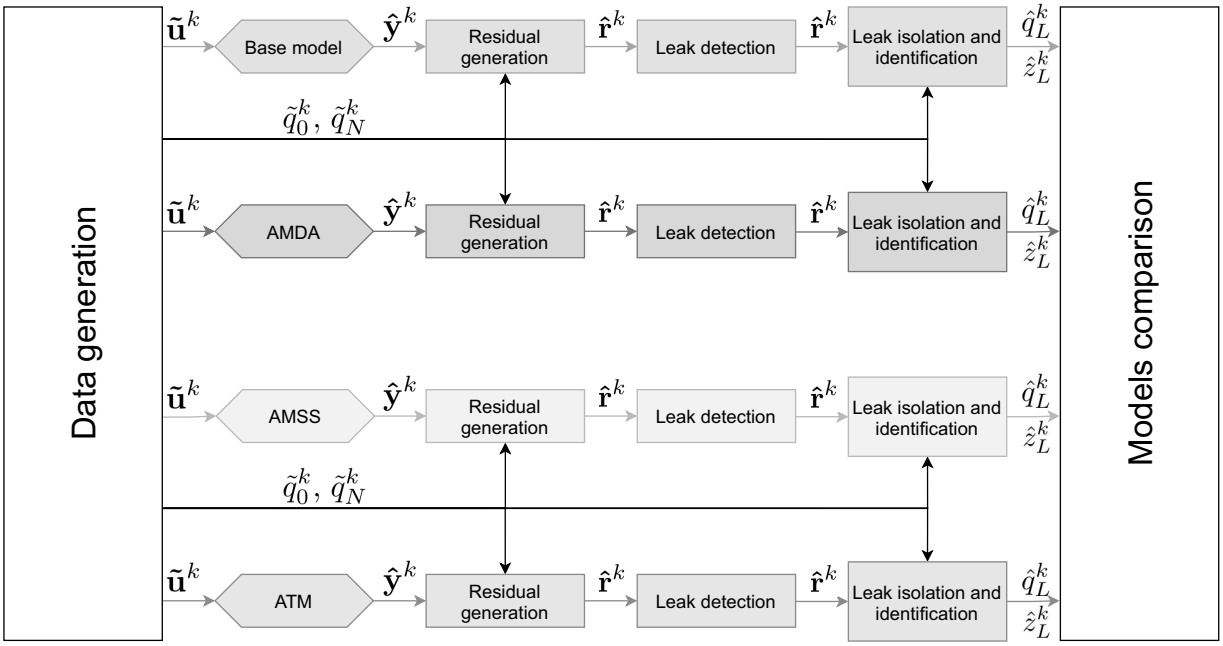


Fig. 1. Employed LDI computing framework with the same input supplied to each (parallel) path using a different instrumental flow-process model and a separate processing path consisting of the same processing blocks. The arrows indicate the data flow and each arrow shade indicates data related to a specific model. Note that the same input values ($\tilde{\mathbf{u}}^k$) containing pressures and inlet/outlet mass-flow rates ($\tilde{q}_0^k$ and $\tilde{q}_N^k$) are fed to each model in the same way.

instrumental models, each of which predicts the pipeline output differently, to obtain diagnostic results. Such outcomes, associated with each model, are then analyzed and compared in the model comparison block. All the models considered work under identical conditions (determined by the physical flow parameters and the discretization grid).

2.1. Base model. In order to compare the instrumental models, the one-dimensional (in the spatial sense) base model was adopted as the starting and reference point. Based on this, it can be conclusively stated what other instrumental models offer in relation to the method reliably established and already validated in the literature. As the base model is also the starting point for deriving new models, it is referenced before discussing the proposed LDI computing framework in detail.

The base model is derived from the following set of partial differential equations (PDEs) describing the flow in the pipeline under consideration (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 [$\frac{\mathrm{m}}{\mathrm{s}^2}$]. Properly used physical parameters concern not only the geometric characteristics of the pipeline, but can also reflect the type of fluid transported in it.

It is worth noting here that the friction coefficient used here (Darcy friction factor) serves as an aggregator of many friction-related influences, including, but not limited to, factors such as viscosity and roughness, which allows it to be treated as a generalized instrumental coefficient. A more specialized formula for calculating the friction coefficient can also be used, such as the Hazen–Williams equation (for water as a medium) or the Colebrook–White equation (White, 2011). However, due to the high level of complexity involved in modeling and estimating friction, but also taking into account the diversity of fluids, we do not delve into this area in this paper. We refer the interested reader to the considerations contained in the works of Dulhoste *et al*. (2011), Jiménez *et al*. (2017), Santos-Ruiz *et al*. (2021) and Pahlavanzadeh *et al*. (2024) for further exploration of this topic.

To obtain a discrete-time model, the following finite difference schemes are applied to the continuous-time

model of (1) and (2): using the second order backward time scheme (de Vahl Davis, 1986)

$$\frac{\partial x}{\partial t} \doteq \frac{3x_d^{k+1} - 4x_d^k + x_d^{k-1}}{2\Delta t} \tag{3}$$

and the second-order scheme in the space domain (time-averaged over two consecutive time steps) (Billmann and Isermann, 1987)

$$\frac{\partial x}{\partial z} \doteq \frac{x_{d+1}^{k+1} - x_{d-1}^{k+1} + x_{d+1}^k - x_{d-1}^k}{4\Delta z}, \tag{4}$$

where Δz and Δt are, respectively, the time and spatial steps, which are referred to as the discretization parameters.

Leading to relatively low-order systems, the above schemes provide smaller approximation errors than the first-order approach (de Vahl Davis, 1986). For brevity, it is assumed that the subscripts indicate the pipe segment index and the superscripts denote the step number in discrete time. After discretization, we obtain the following nonlinear transition equation (singular due to the $\mathbf{A}$ matrix) in the state space:

$$\mathbf{A}\hat{\mathbf{x}}^k = \mathbf{B}\hat{\mathbf{x}}^{k-2} + \mathbf{C}\left(\hat{\mathbf{x}}^{k-1}\right)\hat{\mathbf{x}}^{k-1} + \mathbf{D}\tilde{\mathbf{u}}^{k-1} + \mathbf{E}\tilde{\mathbf{u}}^k, \tag{5}$$

where the state and input vectors are defined as

$$\hat{\mathbf{x}}^k = \begin{bmatrix} \hat{q}_0^k & \hat{q}_2^k & \hat{q}_4^k & \cdots & \hat{q}_N^k \\ & & \hat{p}_1^k & \hat{p}_3^k & \hat{p}_5^k & \cdots & \hat{p}_{N-1}^k \end{bmatrix}^{\mathrm{T}},$$

$$\tilde{\mathbf{u}}^k = \begin{bmatrix} \tilde{p}_0^k & \tilde{p}_N^k \end{bmatrix}^{\mathrm{T}},$$

where p is the pressure and q is the mass-flow rate (Billmann and Isermann, 1987).

To limit the size of the paper, for a detailed description of the above matrices, we refer the reader to the relevant article (Kowalczuk and Tatara, 2021). The hat symbol ($\hat{}$) indicates the estimated measures, while the tilde ($\tilde{}$) denotes the measured values. As above, subscripts are spatial coordinates and superscripts are discrete time moments.

The parameters appearing in the state equations are directly related to the data resulting from the geometry, flow characteristics, and discretization constants. The detailed tuning aspects of these models are thoroughly covered in our previous work (Kowalczuk and Tatara, 2020), showing, among other things, how to select a Courant number based on these data. It should be noted that in flow models there is a friction coefficient, which is crucial for the reliability of the model and is an important tool in its adaptability (Kowalczuk and Gunawickrama, 2004; Kowalczuk and Tatara, 2021).

2.2. Data generation. To reliably compare the performance of the instrumental pipeline-flow models under different operating conditions, we need input and reference data. The process of collecting a representative set of data from real pipelines is both costly and difficult to coordinate, especially for long transmission pipelines.

On the other hand, extensive and representative data can be obtained by simulation means. This approach can provide data from a fully validated source, and at the same time opens the possibility to tune the flow parameters in any wide range (allowed for the implemented model). Certainly, the simulation has a numerical nature and the gained data will never match perfectly those obtained from the relevant real pipeline. Nevertheless, due to the uncertainty of parameters, bias and measurement noise, also the field signals collected in the real installation are characterized by high uncertainty and stochasticity. Therefore, we generally assume that for a sufficiently high order of the simulation algorithm (relating to the number of pipeline sections simulated) the data can be sufficiently accurate, which allows us to objectively compare the instrumental pipeline-flow models considered.

For the data generation task, a simulator developed and validated against experimental industrial data by Gunawickrama (2001) was used. This simulator is suited for isothermic and incompressible flow and consists of an adequately modified base model to generate the inlet and outlet pressures and flow rates. In the PDE model describing the flow process given by (1) and (2) used in this simulator, the leakage (q_L [$\frac{\mathrm{kg}}{\mathrm{s}}$]) can be easily included in the mass-balance equation as follows:

$$\frac{S}{\nu^2}\frac{\partial p}{\partial t} + \frac{\partial q}{\partial z} + \frac{\partial q_L}{\partial z} = 0. \tag{6}$$

After discretization, the above operation implies new elements in the corresponding difference equation, as a result of which the state space model (5) is modified by the expression describing the outgoing mass-flow rate:

$$\begin{aligned}\mathbf{A}\hat{\mathbf{x}}^k = {} & \mathbf{B}\hat{\mathbf{x}}^{k-2} + \mathbf{C}\left(\hat{\mathbf{x}}^{k-1}\right)\hat{\mathbf{x}}^{k-1} \\ & + \mathbf{D}\tilde{\mathbf{u}}^{k-1} + \mathbf{E}\tilde{\mathbf{u}}^k - \mathbf{Q}\left[\mathbf{v}^k + \mathbf{v}^{k-1}\right],\end{aligned} \tag{7}$$

with the matrix $\mathbf{Q} \in \mathbb{R}^{(N+1)\times(N+1)}$ defined as

$$\mathbf{Q} = \left[\begin{array}{c|c} \mathbf{0}_{\left(\frac{N}{2}+1\right)\times(N+1)} & \\ \hline \begin{array}{cccccc} b & b & 0 & 0 & \cdots & 0 \\ 0 & b & b & 0 & & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & & 0 & b & b & 0 \\ 0 & \cdots & 0 & 0 & b & b \end{array} & \mathbf{0}_{\left(\frac{N}{2}\right)\times\left(\frac{N}{2}\right)} \end{array}\right], \tag{8}$$

where $b = \frac{1}{4\Delta z}$ and the vector $\mathbf{v}^k \in \mathbb{R}^{N+1}$ is given by

$$\mathbf{v}^k = \begin{bmatrix} \mathbf{0}_{1\times\frac{d_L-1}{2}} & q_{L_{(d_L-1)}}^k & q_{L_{(d_L+1)}}^k & \mathbf{0}_{1\times N-\frac{d_L+1}{2}} \end{bmatrix}^{\mathsf{T}}, \tag{9}$$

where the symbol d_L (odd number) represents a pair of consecutive segments (therefore, odd and even) within which a leak occurs, the total number of such pairs being $N/2-1$. This pair symbol is sufficient to capture both relevant variables (pressure and flow rate) evaluated alternately in consecutive (even and odd) segments.

Consequently (except for the first segment), all odd segments $(3,\ldots,N-1)$ in (5) and (9) are represented by their preceding (even) segment marked with the index $d_L = 2,4,\ldots,N-2$ (note that flow rate $\hat{q}_0^k$ and pressure $\hat{p}_1^k$ are excluded from this leak consideration). In the simulation, we therefore have a total of $\frac{N}{2}-1$ points over which the leakage can be distributed. In practice, $q^k_{L_{(d_L-1)}}$ and $q^k_{L_{(d_L+1)}}$ placed in odd-numbered segments in the immediate 'vicinity' of the leaky pair of segments (understood as this segment and the preceding one) are calculated as follows:

$$q^k_{L_{(d_L+1)}} = q^k_L \frac{l_1}{l_1+l_2}, \tag{10}$$

$$q^k_{L_{(d_L-1)}} = q^k_L \frac{l_2}{l_1+l_2}, \tag{11}$$

where l_1 and l_2 are the leakage distances from the left and right ends of the leaking pair of segments, respectively, as shown in Fig. 2.

The above formulas can be used both in simulation and in estimation. It is worth emphasizing here that the position of the estimated leak does not have to be limited to the spatial grid used (along the z variable) because, as shown in the above two equations, any (exact) leak position in the range [0, L] can be calculated from leak estimates positioned in the grid.

In the scope of this research, we assume the existence of two distinct categories of leaks: those that manifest abruptly and those that develop gradually over time. These leaks may be conceptualized as sudden or slow-developing degradation, respectively. While the first type can be simulated as a step change, the second one can be modeled as follows:

$$\hat{q}_L^k = \begin{cases} 0 & \text{for} \quad k\Delta t < t_L, \\ q_L\left(1-\exp\left(-\frac{k\Delta t}{t_d}\right)\right) & \text{for} \quad k\Delta t \geq t_L, \end{cases} \tag{12}$$

where t_L is the leak occurrence time, and t_d is the time constant of leak development understood as the time required to achieve a 63% of the modeled leak size q_L.

To obtain experimental data, the state space model from (7) with a slowly developing leak (12) is used. The simulation model used to generate the data has already been verified in numerous experiments and in the literature (Gunawickrama, 2001; Korbicz *et al.*, 2004).

2.3. Residual generation and leak detection. As the basis for the leak detection performed, the residual signals are calculated by comparing the results of the instrumental models with the generated measurements. In order to obtain the appropriate conditions for the intended residual generation, the parameters and state of the LDI system must be properly initialized (including steady state setting), otherwise false detection symptoms can occur. A detailed description of the leak detection methodology can be found in the literature (Gunawickrama, 2001; Korbicz *et al.*, 2004; Isermann, 2011).

2.3.1. Residual generation. After the initialization phase, the residual signals are calculated in each iteration of the diagnostic algorithm:

$$\hat{\mathbf{r}}^k = \begin{bmatrix} \hat{r}_i^k \\ \hat{r}_o^k \end{bmatrix} = \begin{bmatrix} \tilde{q}_i^k - \hat{q}_i^k \\ \tilde{q}_o^k - \hat{q}_o^k \end{bmatrix}, \tag{13}$$

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 the estimates of the inlet and outlet flow rates obtained from the instrumental model under test, respectively. Note that the inlet and outlet flow rates are included in the state vectors of the model (5) as $\hat{q}_i^k \equiv \hat{q}_N^k$ and $\hat{q}_o^k \equiv \hat{q}_0^k$.

2.3.2. Leak detection. As noted, the leak detection task is performed on the basis of the residual signals. In the simplest case, the decision on whether or not to trigger an alarm can be made on the basis of the raw residuals (using only low-pass filtration). In the event of a fault or leak, such values simply deviate from zero. Nevertheless, due to the presence of different noise signals, the residual values usually fluctuate and are never really zero. On the other hand, the results are more reliable when we compute the cross-correlation of residuals using low-pass filtering with a forgetting factor β_c (Billmann and Isermann, 1987; Gunawickrama, 2001):

$$\Phi_{i,o}^k(\tau) = \beta_c \Phi_{i,o}^{k-1}(\tau) + (1-\beta_c)\,\hat{r}_i^{k-\tau}\hat{r}_o^k. \tag{14}$$

Next, the above is summed up for all analysed time shifts $\tau = 1,2,\ldots,\tau_{\max}$:

$$\Phi_\Sigma^k = \sum_{\tau=1}^{\tau_{\max}} \Phi_{i,o}^k(\tau). \tag{15}$$

With a leakage, the value of the aggregated cross-correlation index decreases. Then, to detect leaks, this indicator must be compared with some threshold Φ_{th}:

$$\Phi_\Sigma^k < \Phi_{\text{th}}. \tag{16}$$

Consequently, if the above condition is met, an alarm is triggered (a leak detection decision is made), followed by the diagnosis phase.

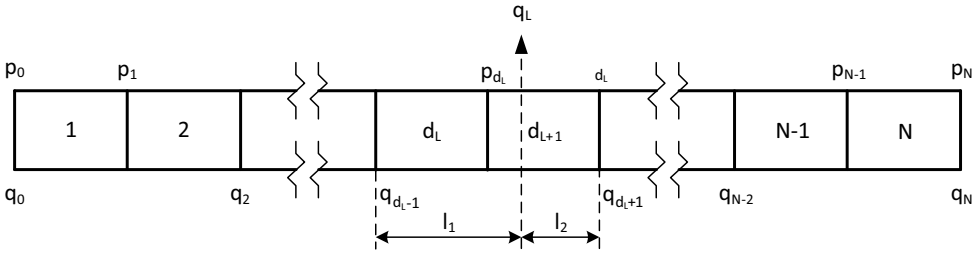


Fig. 2. Simulated pipeline with leak size q_L distributed in the immediate vicinity, where l_1 and l_2 are the distances from the right ends of two consecutive odd segments that are closest to the leak location (Tatara, 2014).

2.4. Leak diagnosis. When a leak is detected, a phase is performed to isolate and identify the leak. The purpose of this part of the diagnosis is to determine the location and size of an existing leak, and to monitor it for further changes.

2.4.1. Leak location. The location of an existing leak can be determined by the specific shape of the pressure distribution along the pipe. According to the established knowledge (Gunawickrama, 2001), the leak location z_L can be specified as

$$\hat{z}_L^k = L\left(1 - \frac{\left(\tilde{q}_i^k\right)^2 - \left(\hat{q}_i^k\right)^2}{\left(\tilde{q}_o^k\right)^2 - \left(\hat{q}_o^k\right)^2}\right)^{-1}. \tag{17}$$

The above can be rearranged to highlight the residuals

$$\begin{aligned}\hat{z}_L^k &= L\left(1 - \frac{\left(\tilde{q}_i^k - \hat{q}_i^k\right)\left(\tilde{q}_i^k + \hat{q}_i^k\right)}{\left(\tilde{q}_o^k - \hat{q}_o^k\right)\left(\tilde{q}_o^k + \hat{q}_o^k\right)}\right)^{-1} \\ &= L\left(1 - \frac{\hat{r}_i^k\left(\tilde{q}_i^k + \hat{q}_i^k\right)}{\hat{r}_o^k\left(\tilde{q}_o^k + \hat{q}_o^k\right)}\right)^{-1}.\end{aligned} \tag{18}$$

This formula can be further simplified assuming that the leak in question is small. For more information on such a simplification, the reader is referred to the works of Billmann and Isermann (1987), Gunawickrama (2001) or Isermann (2011). Nevertheless, our application assumes that the above leak location estimate is both feasible and sufficient.

To further reduce the impact of the measurement noise, another low-pass filter is implemented in an iterative form with the forgetting factor β_z:

$$\hat{w}_z^k = \beta_z \hat{w}_z^{k-1} + (1 - \beta_z)\,\hat{z}_L^k, \tag{19}$$

where $\hat{w}_z^k$ is the filtered estimate of the leak location variable z_L^k to be determined. The above filter, given by (19), will be called the leak location effective estimator.

2.4.2. Leak size. When a leak occurs, the inlet mass flow rate increases and the outlet mass-flow rate decreases, and the size of the leak can be easily determined as the difference between inlet and outlet flow rate measurements:

$$\tilde{q}_L^k = \tilde{q}_i^k - \tilde{q}_o^k. \tag{20}$$

To reduce the variance of this estimate, the improved balance sheet approach (Gunawickrama, 2001) can be used instead:

$$\hat{q}_L^k = \mathrm{E}\{\hat{r}_i^k - \hat{r}_o^k\}, \tag{21}$$

where $\mathrm{E}\{\cdot\}$ is the expected value operator. This approach involves the use of certain reference values (in particular, the output of the instrumental model contained in residual signals), which can also compensate for the calibration errors of sensors (Gunawickrama, 2001).

The expectation operator can also be mechanized using a low-pass filter with a forgetting rate β_q:

$$\hat{w}_q^k = \beta_q \hat{w}_q^{k-1} + (1 - \beta_q)\,\hat{q}_L^k, \tag{22}$$

which $(\hat{w}_q^k)$ will be referred to as the leak size estimator.

2.5. Estimation of the friction factor. When dealing with real pipelines, the value of the coefficient of friction is usually only roughly known. This is due to the complex (nonstationary nonlinear) nature of this quantity, which is affected by changes over time, including the aging of the pipe and its instrumentation. This phenomenon can be attributed to corrosion, roughness changes, fluid deposition, tuberculation, and other factors (Larson, 1960; Vítkovský *et al*., 2000). To obtain reliable LDI results, the friction coefficient assessment should be done online or at least performed at the start of the LDI system operation. This article uses the methodology described in a publicly available report (Kowalczuk and Tatara, 2020). Other approaches that can be used here are discussed by Dulhoste *et al*. (2011), Jiménez *et al*. (2017), Santos-Ruiz *et al*. (2021) and Pahlavanzadeh *et al*. (2024), while the problem of simultaneous evaluation of leakage parameters and friction coefficient is considered by Doshmanziari *et al*. (2020) and Pérez-Pérez *et al*. (2021).

The following is the equation for calculating the friction coefficient, derived from the Colebrook equation (Kowalczuk and Tatara, 2020):

$$\lambda(\mathbf{p}) = \left(-2\log_{10}\left[\frac{\epsilon}{3.7D} + \frac{2.51\mu\sqrt{L\kappa(\mathbf{p})}}{\sqrt{2\rho D^3(p_i - p_o)}} \right] \right)^{-2}, \tag{23}$$

where ϵ is the roughness height $[m]$ and $\kappa(\mathbf{p})$ is the correction factor resulting from the approximate nature of the Darcy–Weisbach equation given as (Kowalczuk and Tatara, 2020)

$$\kappa(\mathbf{p}) = \frac{4}{3}\left(1 - \frac{p_i p_o}{p_i^2 + 2p_i p_o + p_o^2} \right). \tag{24}$$

2.6. Summary of the methodology used. The experimental data were generated using a verified pipeline simulator. The resulting pressure measurement samples ($\tilde{\mathbf{u}}^k$) go to four instrumental models, as shown in Fig. 1. The vector

$$\hat{\mathbf{y}}^k = \left[\begin{array}{cc} \hat{q}_i^k & \hat{q}_o^k \end{array} \right]^{\mathrm{T}}$$

is then constructed from the estimates of the mass-flow rate at both the ends of the pipe. Ultimately, the mass-flow measurements and their estimated values are used to calculate the residual vector of (13).

When the collective cross-correlation computed by (15) is below a predetermined threshold, an alarm is triggered and the system proceeds to the next diagnostic phase in which the practical values of leak location ($\hat{z}_L^k$) and leak rate/size ($\hat{q}_L^k$) are computed using the diagnostic estimators described by (19) and (22), respectively.

3. Analyzed models

Starting with the base model as the most popular reference, we have developed three other models that describe the process of isothermal flow of an incompressible fluid through a pipeline. All these instrumental models will be validated in the LDI task. To conduct a comparative study, the LDI computational framework from Section 1 (Fig. 1) was used, which allows the analysis of the operation of each of the described models and evaluation in terms of diagnostic effectiveness. Below, rather than a full explanation, we only provide a brief overview of the main ideas behind each modeling approach, and refer the reader to other publications for more detailed information.

3.1. Analytic model of diagonal approximation. The second model is derived on the basis of the assumption that (5) can be shown in the following (nonsingular) form:

$$\begin{aligned} \hat{\mathbf{x}}^k = \mathbf{A}^{-1}\big(& \mathbf{B}\hat{\mathbf{x}}^{k-2} + \mathbf{C}\left(\hat{\mathbf{x}}^{k-1}\right)\hat{\mathbf{x}}^{k-1} \\ & + \mathbf{D}\tilde{\mathbf{u}}^{k-1} + \mathbf{E}\tilde{\mathbf{u}}^{k}\big), \end{aligned} \tag{25}$$

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

For the established, specific structure of the matrix $\mathbf{A}$, one can derive an analytical approximation of its inverse (by replacing the tridiagonal matrices with their diagonal equivalents) (Kowalczuk and Tatara, 2016). The model related to the approximate inverse of the matrix $\mathbf{A}$ is called the analytical model of diagonal approximation (AMDA). For a complete explanation of this instrumental model, the reader is referred to our previous work (Kowalczuk and Tatara, 2016).

3.2. Analytic Thomas model. To reduce the computational complexity, the sparsity of the matrices in the state space model given by (5) allows the re-arrangement of this model to the form corresponding to the Thomas algorithm (Conte and de Boor, 1980).

The related procedure in the work of Kowalczuk *et al.* (2018) takes into account the specific structure of the submatrices in the recombination matrix $\mathbf{A}$. Two of them are square and diagonal, whereas the other two are nonsquare band matrices with the band width equal to 2. As a result, the model of (5) can be shown as

$$\begin{aligned} \mathbf{A}\hat{\mathbf{x}}^k &\equiv \left[\begin{array}{c:c} \mathbf{A}_1 & \mathbf{A}_2 \\ \hdashline \mathbf{A}_3 & \mathbf{A}_4 \end{array} \right] \hat{\mathbf{x}}^k \\ &= \mathbf{w}\left(\hat{\mathbf{x}}^{k-1}, \hat{\mathbf{x}}^{k-2}, \tilde{\mathbf{u}}^k, \tilde{\mathbf{u}}^{k-1} \right), \end{aligned} \tag{26}$$

with

$$\hat{\mathbf{x}}^k = \left[\begin{array}{c} \hat{\mathbf{q}}^k \\ \hat{\mathbf{p}}^k \end{array} \right], \quad \mathbf{w} = \left[\begin{array}{c} \mathbf{g} \\ \mathbf{h} \end{array} \right],$$

where $\hat{\mathbf{q}}^k \in \mathbb{R}^{\frac{N}{2}+1}$ and $\hat{\mathbf{p}}^k \in \mathbb{R}^{\frac{N}{2}}$ being flow rates and pressures in subsequent segments in the k-th computational moment, while $\mathbf{g} \in \mathbb{R}^{\frac{N}{2}+1}$ and $\mathbf{h} \in \mathbb{R}^{\frac{N}{2}}$ are collective functions of the elements of the right side of (5).

By defining

$$\breve{\mathbf{g}} = \frac{\mathbf{g}}{c}, \qquad \breve{\mathbf{A}}_2 = \frac{\mathbf{A}_2}{c},$$
$$\breve{\mathbf{h}} = \frac{\mathbf{h}}{a}, \qquad \breve{\mathbf{A}}_3 = \frac{\mathbf{A}_3}{a},$$

where

$$a = \frac{3S}{2\nu^2\Delta t}, \quad c = \frac{3}{2S\Delta t},$$

we eventually get

$$\left(\mathbf{I} - \breve{\mathbf{A}}_2\breve{\mathbf{A}}_3 \right)\hat{\mathbf{q}}^k = \breve{\mathbf{g}} - \breve{\mathbf{A}}_2\breve{\mathbf{h}}, \tag{27}$$

$$\left(\mathbf{I} - \breve{\mathbf{A}}_3\breve{\mathbf{A}}_2 \right)\hat{\mathbf{p}}^k = \breve{\mathbf{h}} - \breve{\mathbf{A}}_3\breve{\mathbf{g}}. \tag{28}$$

The results of the matrix multiplications $\breve{\mathbf{A}}_3\breve{\mathbf{A}}_2$ and $\breve{\mathbf{A}}_2\breve{\mathbf{A}}_3$ are tridiagonal, and therefore the expressions $\mathbf{I} - \breve{\mathbf{A}}_3\breve{\mathbf{A}}_2$ and $\mathbf{I} - \breve{\mathbf{A}}_2\breve{\mathbf{A}}_3$ are also tridiagonal

matrices. Hence, the Thomas algorithm (Thomas, 1949) can be applied assuring the computational complexity of order $\mathcal{O}(N)$ (Conte and de Boor, 1980). The derived instrumental model, allowing the equivalent determination of successive state vectors, will be called the analytic Thomas model (ATM). The full derivation of this model can be found in our previous work (Kowalczuk and Tatara, 2017).

3.3. Model of steady state. Suppose the LDI system is installed in a constant flow pipeline when both pressure and flow rate remain the same over time (Kowalczuk and Tatara, 2018). Thus we can consider $\partial q/\partial t \to 0$ and $\partial p/\partial t \to 0$ in (1) and (2) to be satisfied approximately for such a constant flow. In such cases, the partial differential equations are reduced to ordinary differential equations,

$$\frac{\mathrm{d}q}{\mathrm{d}z} = 0, \tag{29}$$

$$\frac{\mathrm{d}p}{\mathrm{d}z} = -C_3 \frac{q|q|}{p}, \tag{30}$$

which can be solved for a constant mass-flow rate and pressure distribution in a pipeline. The resulting solution can be provided separately for the zero and nonzero inclination angle cases, as shown below, where the tilde does not appear over p_i and p_o, highlighting both the practical (computational and measurement) and theoretical (physical) aspects of the mathematical formulas presented below.

3.3.1. Zero inclination angle, $\alpha = 0$. The constant mass-flow rate is

$$q = \operatorname{sign}\left(p_i^2 - p_o^2\right) \sqrt{\left| \frac{DA^2}{\lambda \nu^2} \frac{p_i^2 - p_o^2}{L} \right|}, \tag{31}$$

where $\operatorname{sign}(x)$ is 1 for $x \geq 0$, and -1 otherwise, and the pressure distribution along the pipeline is

$$p = \sqrt{p_i^2 - \frac{p_i^2 - p_o^2}{L} z}. \tag{32}$$

3.3.2. Nonzero inclination angle, $\alpha \neq 0$. The mass-flow rate at a nonzero angle of inclination can be calculated in accordance with (33), and the corresponding pressure distribution along the spatial coordinate z is given by (34).

Note that we have a singularity above for $\alpha \to 0$. In this case, however, it is recommended to use the other model suitable for the zero inclination angle ($\alpha = 0$).

The above dichotomous models describing steady state flow for zero and nonzero slope angle, respectively, will be collectively referred to as the analytical model of steady state (AMSS), whose detailed derivation can be found in our previous work (Kowalczuk and Tatara, 2018).

4. Validation of the diagnostic models and estimators

In order to properly evaluate all of the above instrumental models of the pipeline process for their application to LDI systems, the influence of system (hyper)parameters on the desired diagnosis will be investigated. As the system output variables are strongly influenced by measurement noise, leak size/rate and leak location, the problem is multi-folded. Therefore, the practical experimental settings considered must be adequately reflected by rationally selected system parameters.

4.1. Statistical characteristics of estimates. Suppose that pressure and flow rate measurements are subject to noise, modeled as an additive Gaussian process with zero mean and variance σ_m^2. Additionally, let us assume that the noise (in percent) for the flow measurement is an order of magnitude greater than for the pressure measurement (here you can refer to data provided by manufacturers (e.g., Omega, 2019; Sensirion, 2019; JUMO, 2019; VEGA, 2019).

Of course, the assumption of Gaussian noise is a significant simplification of the problem (e.g., Lam *et al*., 2019), where the noise is modeled as a Rayleigh distribution). Nevertheless, the Gaussian model is widely used in the literature (Pérez-Pérez *et al*., 2021) to make modeling more convenient and to allow researchers to focus on the fundamental problems that need to be solved.

Diagnostic estimates of the size and location of a leak are evaluated for the expected bias and standard deviation of the error obtained. The corresponding systematic errors are determined as the deviation of the mean value (taking into account the sign):

$$\hat{\bar{w}}_q = \frac{1}{n} \sum_{j=1}^{n} \hat{w}_{q_j} - q_L \tag{35}$$

for the estimate of the leak size and

$$\hat{\bar{w}}_z = \frac{1}{n} \sum_{j=1}^{n} \hat{w}_{z_j} - z_L \tag{36}$$

for the leak location estimator, where n represents the number of experiment runs performed.

The standard deviation (STD) of estimation errors can be determined for the leak size estimator as

$$\hat{\sigma}_q = \sqrt{\frac{1}{n-1} \sum_{j=1}^{n} \left(\hat{w}_{q_j} - \bar{w}_q\right)^2} \tag{37}$$

and for the leak location estimator as

$$\hat{\sigma}_z = \sqrt{\frac{1}{n-1} \sum_{j=1}^{n} \left(\hat{w}_{z_j} - \bar{w}_z\right)^2}. \tag{38}$$

$$q = \sqrt{\left| \frac{2DS^2}{\lambda\nu^2} \frac{g\sin\alpha}{\nu^2} \left(\frac{p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L}}{e^{2\frac{g\sin\alpha}{\nu^2}L} - 1} \right) \right| \operatorname{sign}\left(p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L} \right)}, \tag{33}$$

$$p = \sqrt{e^{-2\frac{g\sin\alpha}{\nu^2}z} p_i^2 + \left(\frac{p_i^2 - p_o^2 e^{2\frac{g\sin\alpha}{\nu^2}L}}{e^{2\frac{g\sin\alpha}{\nu^2}L} - 1} \right) \left(e^{-2\frac{g\sin\alpha}{\nu^2}z} - 1 \right)}. \tag{34}$$

Table 2. Reference physical parameters of the pipeline flow process used in the simulation.

Parameter	Value
Length of the pipe L	$100\,\mathrm{km}$
Diameter of the pipe D	$0.4\,\mathrm{m}$
Speed of sound ν	$350\,\frac{\mathrm{m}}{\mathrm{s}}$
Friction factor λ	0.02
Number of segments N	10
Inclination angle α	0°
Inlet pressure p_i	$112.28\,\mathrm{bar}$
Outlet pressure p_o	$80\,\mathrm{bar}$
Leak location z_L	$\{0.1; 4; 13.87; 40; 73.75; 98.2\}$ [km]
Leak size q_L	$\{0; 0.04; 0.1; 0.4; 0.8; 1.6; 2; 4; 8; 16\}$ $[\frac{\mathrm{kg}}{\mathrm{s}}]$
Leak occurrence time t_L	$105.5\,\mathrm{min}$
Leak development time t_d	$17.5\,\mathrm{min}$

In the summary of LDI: we use two types of diagnostic estimators to determine the size and location of the leak. They will be assessed in terms of bias and standard deviation, as universal measures of estimation quality.

4.2. Deterministic and stochastic experimentation. Reference physical parameters of the flow process are given in Table 2. Using these parameters, simulation data was generated consisting of pressures and mass-flow rates at the inlet and outlet of pipelines for various experimental settings. The number N_s of segments implemented in the simulator was set to 100. The inlet pressure was selected so that the resulting flow under leak-free conditions was approximately $40\,\frac{\mathrm{kg}}{\mathrm{s}}$.

Another set of system parameters concerning the computation and simulation process is shown in Table 3. Parameters representing the modeled percentage of measurement noise and its standard deviation in the inlet and outlet measurements are presented in Table 4.

In the diagram of Fig. 1, pressures and mass-flow rates are determined in the Data Generation block (with the simulation environment) and introduced into the designed LDI systems directly and through appropriate instrumental models. However, the instrumental models were first initiated to reach a steady state and estimate the coefficient of friction, and only then were the models fed with the generated data. The residual signals were next calculated. After the leak was detected, appropriate estimates of the size and location of the leak were made.

Additionally, to visualize the stochastic properties of the estimates, the error probability distribution of this estimation was approximated by a Gaussian fit, with the number n of the runs (exceptionally) set to 1000. The distribution functions obtained are shown in Fig. 3 for the leak size estimates, along with their characteristic values of the systematic error and STD. Taking into account the shape of the probability distribution functions (PDFs) in question, the effectiveness of the leak location estimators was similar; therefore, they are not shown here.

As can be seen, the Gaussian curves fit the data satisfactorily, so the use of Gaussian PDFs is fully justified in modeling these quantities, and the resulting bias and standard deviation are sufficient characteristics of these probability distribution functions.

The issue of selecting the tool parameters of the analyzed model and research procedures is very complex. The reader interested in a deeper analysis of this problem is referred to our previous work (Kowalczuk and Tatara, 2021). Here we only provide a shortened commentary showing the general context that is important for correct implementation of the LDI system tuning process. In this process, the basic physical and geometric parameters of the flow process are taken into account, and the remaining variables are adaptively adjusted. In particular, the fundamental issue of selecting sampling periods and sensor locations, determining the granularity of the temporal and spatial domains (discretization mesh), can be solved within the Courant–Friedrichs–Lewy condition (Kowalczuk and Tatara, 2021). Further research on this topic is ongoing and focuses on finding new conditions for sampling variables describing the current flow that guarantee the stability of the discrete flow model. Our meticulous approach allows us to reconcile the number of observations with the complexity of the simulated process, ensuring the stability of the model and the appropriate

Table 3. Computational parameters of the LDI system used in the simulation study.

Parameter	Value
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
Length of segments Δz	$10\,\mathrm{km}$
Sampling time/numerical procedure step Δt	$4.76\,\mathrm{s}$

Table 4. Statistical indicators of the stochastic experimental system setup.

Indicator	Value
Noise percentage w.r.t. pressure readings v_{rdg_p}	0.1% of readings
Noise percentage w.r.t. mass-flow rate readings v_{rdg_q}	1% of readings
Standard deviation of inlet pressure noise $\sigma_{m_{pi}}$	$0.056\,\mathrm{bar}$
Standard deviation of outlet pressure noise $\sigma_{m_{po}}$	$0.04\,\mathrm{bar}$
Standard dev. of inlet mass-flow rate noise $\sigma_{m_{qi}}$	$0.2\,\frac{\mathrm{kg}}{\mathrm{s}}$
Standard dev. of outlet mass-flow rate noise $\sigma_{m_{qo}}$	$0.2\,\frac{\mathrm{kg}}{\mathrm{s}}$

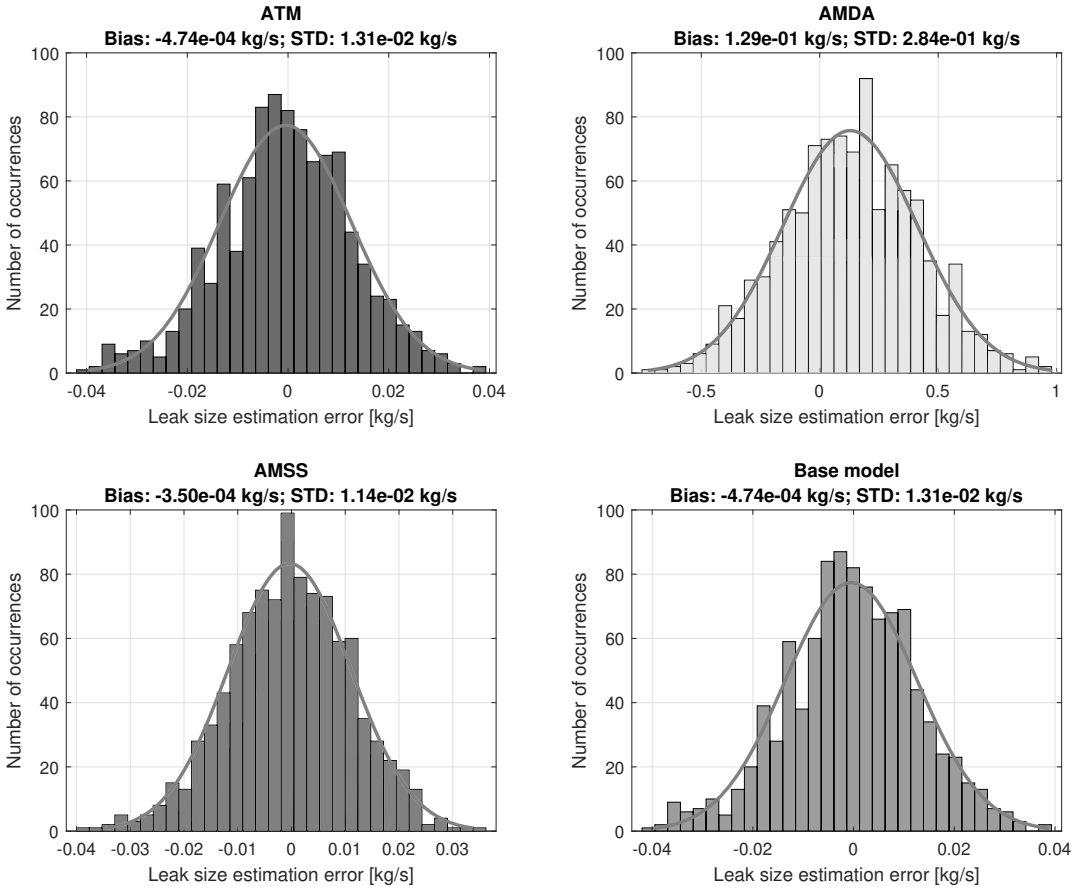


Fig. 3. Gaussian fitting of the estimation errors obtained in the leak size estimators considered based on the four instrumental flow models, for experimental data from 1000 observations (the values above the graphs are written in *E-notation*, a form of scientific notation in which the value before the letter 'e' is multiplied by 10 to a power equal to the value after the symbol 'e').

accuracy and usefulness of the research results.

4.3. Assessment of the quality of estimators. The quality of the estimators is assessed in terms of their accuracy, reflected in their bias, which tells how the obtained estimates relate to the actual value they estimate. Thus, the lower the bias, i.e., the difference between the actual value and its estimate, the more accurate the system. In turn, the precision of the estimator is reflected in its standard deviation, which the smaller it is, the more accurate the estimator (Bos, 2007).

In our LDI simulation study, the evaluation was performed for 60 experimental settings: 10 values of leak size $q_L \in \{0,\ 0.04,\ 0.1,\ 0.4,\ 0.8,\ 1.6,\ 2,\ 4,\ 8, 16\}\,[\frac{\mathrm{kg}}{\mathrm{s}}]$, and six of leak locations $z_L \in \{0.1,\ 4,\ 13.87,\ 40,\ 73.75, 98.2\}$ in km. In addition, seven different percentages $v_{rdg_p} \in \{0,\ 0.01,\ 0.05,\ 0.1,\ 0.5,\ 1, 5\}\,[\%]$ of noise introduced into the pressure readings (but only) were included for the leak events $q_L = 4\,\frac{\mathrm{kg}}{\mathrm{s}}$ and $z_L = 40\,\mathrm{km}$.

All the above-mentioned experiments were used to evaluate the analyzed diagnostic estimators for their bias and standard deviation. It is certainly difficult to clearly identify the best pair (model, estimator), because its diagnostic quality consists of many factors. Therefore, the problem of choosing the appropriate comparative criterion arises. One of the most common and elegant approaches to multi-criteria problems is optimization in the sense of Pareto, which introduces a practical ranking of solutions (Kowalczuk *et al*., 1999; Deb *et al*., 2003; Deb and Gupta, 2005).

The computational complexity of the analyzed instrumental models of flow processes may be related to both the computational time (t_i) necessary for a single iteration and the steady-state settling time (t_s).

The performance measures obtained in this part of the study are collected in Table 5, taking into account the parameters of the computational system and the experimental setup from Tables 2 and 3, respectively, and taking into account the isolation and identification tasks.

In order for the above measures to be comparable, they should be normalized with respect to their individual value ranges. As can be seen in Table 5, systematic errors in leak size and location can be both positive and negative, so they only need to be converted to positive (profit) or negative (loss) values before further processing and final use in the appropriate maximization or minimization task. This can be done by appropriately (as minimally as possible) shifting all measure values up or down (and applying negation if necessary).

Therefore, in our analysis, in the first step we normalize the obtained values to ensure that they are strictly positive or zero. This typical normalization procedure involves subtracting the minimum value within the criterion and dividing it by the range of the maximum and minimum values. As a result, all values for a given criterion will then fall within the specified range $[0, 1]$.

Assuming that the maximum value of the criterion is K and the minimum value is M, the conversion of the original value v_m (the criterion value for the instrumental model) to its normalized value v_n is expressed as follows:

$$v_n = \frac{v_m - M}{K - M}. \tag{39}$$

Consequently, this establishes a standardized minimization problem. To transform it into a convenient maximization scheme, each resulting value is subtracted from 1, keeping the range within $[0, 1]$. As a result of this maximization-adjusted data transformation, we obtain normalized performance metrics for all instrumental models presented in Table 6.

The sound speed considered, as indicated in Table 2, is specific to gases. To ensure the generalizability of the results to liquids, additional experiments were performed. These included scenarios with I: the speed of sound ($450\,\frac{\mathrm{m}}{\mathrm{s}}$), being the reference experiment, II: a higher speed of sound ($1482\,\frac{\mathrm{m}}{\mathrm{s}}$), III: this increased speed of sound combined with a higher coefficient of friction (0.06), and IV: this increased speed of sound at a leakage level of 10% of the flow rate. Each of experiments II–IV was compared with a gas-specific scenario I.

In all experiments except IV, a simulated leakage rate of $4\,\frac{\mathrm{kg}}{\mathrm{s}}$ was maintained at a distance of $40\,\mathrm{km}$. In experiment IV, the leakage rate was set at $1\,\frac{\mathrm{kg}}{\mathrm{s}}$.

The time indicators in all experiments presented similar values and were therefore excluded from further analysis. Furthermore, as shown in Appendix (Section A.1) with Table A1, in view of the data from experiments II–IV, the standard deviations and systematic errors (bias) in the leak size estimator were at least three times smaller than the errors observed in the reference experiment I. However, the bias in estimating the leak location is larger in the case of liquid (II–IV) than of gas (I). In all cases, the relative dependencies between the indicators assigned to individual models were maintained. Generally speaking, the gas scenario is more demanding for a leak detection system. Therefore, in the remainder of the article we will focus on experiments with a lower sound speed of $450\,\frac{\mathrm{m}}{\mathrm{s}}$, corresponding to the gas transport process. It is also worth paying attention to experiment III (with an increased friction factor), which shows optimality due to the errors in estimating the leakage size (lowest LSE bias and STD). Basically, we observe here the impact of reducing these estimation errors as a result of increasing the speed of sound or the coefficient of friction. This means that the increase in factors responsible for slowing down the flow rate (31) also contributes to the increase in precision and accuracy of estimating the leak size.

Table 5. Averaged quality measures of the estimators of $\bar{z}_L$ (isolation) and $\bar{q}_L$ (identification) and the computation times t_i and t_s used to compare the discussed models for a minimization task.

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]	$1.45 \cdot 10^{-2}$	$3.53 \cdot 10^{0}$	$2.73 \cdot 10^{-5}$	$2.24 \cdot 10^{-1}$

Table 6. Normalized local optimality levels for different performance indicators in a maximization task (*) used for further comparison of the models.

Performance indicator	ATM	AMDA	AMSS	Base model
* Bias of leak size estimator $\hat{\bar{w}}_q$	0.008	1	0	0.008
* STD of leak size estimator $\hat{\sigma}_q$	0.997	0	1	0.997
* Bias of leak location estimator $\hat{\bar{w}}_z$	0.787	0.437	0	1
* STD of leak location estimator $\hat{\sigma}_z$	0	0.842	1	0.663
* Single iteration computation time t_i	0.988	0.743	1	0
* Steady state computation time t_s	0.996	0	1	0.936

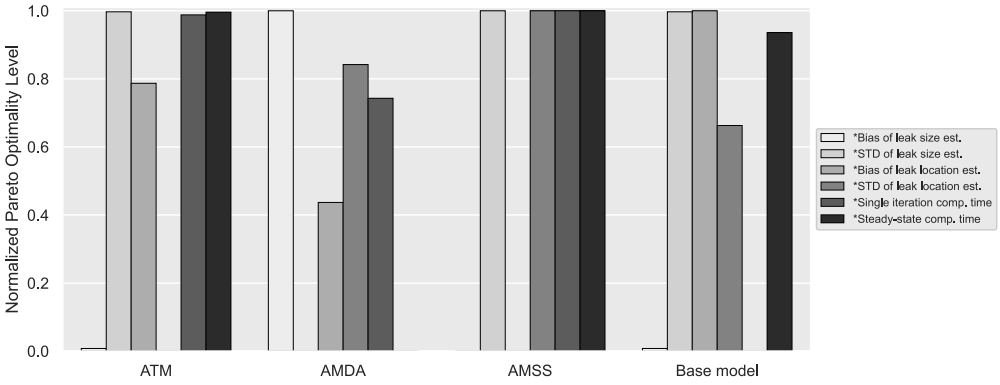


Fig. 4. Normalized local optimality level indicators for the analyzed models in the maximization task (*).

The analysis of the results collected in Table 6 leads to the conclusion that, in the Pareto sense, the three models (basic, AMDA and AMSS) are equivalent, because each of them is the best in at least one criterion (the ATM not much, but different from all). Thus, it is possible to indicate the optimal diagnostic estimator for the selected quality measure of the estimation of the leakage size, leak location or processing time.

However, in a practical case, all these aspects may be relevant. Then we can use a measure called the global optimality level (GOL) (Kowalczuk and Białaszewski, 2017). This criterion is usually able to accurately pinpoint the estimator that is best in terms of the worst case. In other words (in the maximization task), the GOL aims to maximize the lowest quality index for each case under consideration (here: the diagnostic estimator). This is equivalent to first (i) finding the lowest value for each column in Table 6, and then (ii) selecting the estimator by maximum among the rows assessed in this raw way.

However, in the given diagnostic case with normalized objectives, the GOL is equal to 0 for each of the instrumental models. This simply means that each model is worst in at least one criterion, and from the GOL point of view, the instrumental models analyzed cannot be distinguished. Therefore, another measure of global optimality should be used here.

In the simplest approach, you can calculate the average optimality index (AOI) for each diagnostic system by calculating the mean value for each column (describing the instrumental model and the estimator). The resulting AOIs are shown in Table 7, where you can see that the AMSS turns out to be the best (because its AOI is the

Table 7. Averaged optimality index for the diagnostic systems analyzed for global maximization (*).

	ATM	AMDA	AMSS	Base model
AOI	0.630	0.504	**0.667**	0.601

Table 8. Assessment of the diagnostic estimators, using the distance or demand-reference approach (DRA).

	ATM	AMDA	AMSS	Base model
DRA	1.425	1.552	**1.414**	1.450

Table 9. General assessment of instrumental models and appropriate diagnostic estimators, characterized only by the criteria of precision (without time), using various scalar measures (GOL, AOI, DRA).

	ATM	AMDA	AMSS	Base model
GOL	0	0	0	**0.008**
AOI	0.448	0.570	0.500	**0.667**
DRA	1.425	1.158	1.414	**1.048**

highest). As could be predicted according to Table 6, the ATM can be considered suboptimal (with the second highest score), and the baseline model has the third level of effectiveness. In this way, a ranking of models is obtained, measured by their average performance, taking into account the six criteria considered.

One of the popular approaches used in multi-objective optimization is the distance measure (e.g., Osyczka and Kundu, 1996). In particular, we can compare our diagnostic estimators using the demand-reference approach (DRA), which measures the Euclidean distance of the characteristics of the analyzed estimator from the reference point that models the ideal estimator (point $[1, 1, 1, 1, 1, 1]^{\mathsf{T}}$ in this case).

In this approach, the optimization problem naturally converts from expressive maximization to natural minimization of the distance from the ideal (reference) point. The values collected in Table 6 are first subtracted from 1 and then the second norm is calculated, resulting in a DRA measure that represents the distance from the ideal. The results are shown in Table 8.

Let us emphasize that the most effective estimator is AMSS, followed by the ATM and thirdly by the base model. In terms of the AOI and DRA, both estimators maintain the same ranking position.

It is worth noting that the ATM model does not achieve an optimal performance in any of the six criteria, although in most cases it is close to the best performance (Table 6). Nevertheless, both in terms of the (maximized) average estimate (AOI) and the (minimized) distance to demand (DRA), it clearly remains second in place.

Unlike other instrumental models, the best one, the AMSS, is particularly suitable for calculations in a steady state (its strength is the time of calculation). That is why it is worth conducting a similar analysis (in the context of the GOL, AOI and DRA), ignoring the time criteria (shown in the last two lines of Table 6). The results obtained are presented in Table 9.

Interestingly, from the point of view of only the set of analyzed global precision criteria (GOL, AOI and DRA), the base model seems the best. It is also not the worst estimator in terms of any of these criteria (there is no zero in Table 5). In the sense of the subglobal criterion GOL, the other estimators are equivalent. The AOI puts the AMDA model in second place, while the DRA favors the AMSS description here. The ATM algorithm is the worst in all indicators, which means that in the full ranking (more “global") time criteria significantly improve the assessment of this model.

The employed accuracy and precision criteria, systematic error and standard deviation were statistically estimated on the basis of 200 observations in each experiment. However, considering the entirety of the experiments conducted, with a total number of over 10,000 observations, the aggregate results given can be considered credible.

It is worth emphasizing that in the adopted working conditions and testing in the computational framework for LDI, new models are better than the popular base model mainly in terms of the computational complexity.

The presented environment for analyzing models in LDI tasks was used as a computational framework to compare the selected four models. It can also be applied to test other models and quantitatively assess their performance quality using the developed optimization indicators and criteria. While the synthesized flow models can be highly diverse and use different input variables and parameters, or even black box models, the end result can always be quantified using the proposed universal performance metrics (indicators).

5. Summary

The article presented the results of research on various instrumental models in terms of their diagnostic effectiveness. First, the computing framework for the LDI problem was presented and described in detail. Each block of the LDI framework was described using appropriate mathematical formulas. Then the analyzed models were briefly presented along with the basic principles of their synthesis.

We emphasize the importance of the simulation methodology and system complexity associated with the data acquisition process and its diversity. An example of such functionality is the ability to explore wide ranges of modeling parameters and implement complex

fluid flow process scenarios and diagnostic algorithms. Simulation also provides flexibility in the selection of some technological parameters, which is often impractical or even impossible in real installations.

In this work, we proposed a baseline assessment of diagnostic estimators in terms of the bias and the standard deviation of basic diagnostic estimates (leak location and size and time indicators). Various diagnostic situations (leakage and transport parameters) and measurement noise levels were taken into account. The proposed methodology can serve as a basis for comparing different flow models and diagnostic observers in terms of their diagnostic performance and computational efficiency.

Multivariable estimators resulting from the employed instrumental models of fluid transport processes were compared in terms of their global optimality level (GOL) and averaged optimality index (AOI) and demand-reference approach (DRA).

The analytical model of the steady state (AMSS) was the best in most cases in terms of the standard deviation, bias and computation time. The estimator based on the analytical model of diagonal approximation (AMDA) seems to be suitable only for leak locations. The analytical Thomas model (ATM) estimator is comparable to the base model in most cases, while having lower computational complexity than the AMDA and the base model. Considering only the global precision criteria (in the AOI and DRA versions), the basic model is the best, while the AMDA or AMSS are in the second place depending on the choice of the type of the joint/global measure (AOI or DRA).

The analysis performed excluding time criteria demonstrated the qualitative power of the base model in diagnostic use. However, it is important to remember that there is rarely a case where computational complexity is not important. Therefore, the proposed new models are a proven and effective alternative to the well-known base model because, with comparable results and operational efficiency, they offer lower computational complexity, which is particularly important in the context of online diagnostic applications.

References

Abhulimen, K.E. and Susu, A.A. (2004). Liquid pipeline leak detection system: Model development and numerical simulation, *Chemical Engineering Journal* **97**(1): 47–67.

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

Bos, A.V.D. (2007). *Parameter Estimation for Scientists and Engineers*, Wiley, Hoboken.

Brogan, W. (1991). *Modern Control Theory, 3rd Edition*, Prentice Hall, Englewood Cliffs.

Brown, G. (2003). The history of the Darcy–Weisbach equation for pipe flow resistance, *in* J.R. Rogers and A.J. Fredrich (Eds), *Environmental and Water Resources History*, American Society of Civil Engineers, Washington, pp. 34–43, DOI: abs/10.1061/40650(2003)4.

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

de Vahl Davis, G. (1986). *Numerical Methods in Engineering & Science*, Springer, Dordrecht.

Deb, K. and Gupta, H. (2005). Searching for robust Pareto-optimal solutions in multi-objective optimization, *in* C.A. Coello Coello *et al.* (Eds), *Evolutionary Multi-Criterion Optimization*, Springer, Berlin/Heidelberg, pp. 150–164.

Deb, K., Zope, P. and Jain, A. (2003). Distributed computing of Pareto-optimal solutions with evolutionary algorithms, *in* C.M. Fonseca *et al.* (Eds), *Evolutionary Multi-Criterion Optimization*, Springer, Berlin/Heidelberg, pp. 534–549.

Doshmanziari, R., Khaloozadeh, H. and Nikoofard, A. (2020). Gas pipeline leakage detection based on sensor fusion under model-based fault detection framework, *Journal of Petroleum Science and Engineering* **184**: 106581.

Dulhoste, J.-F., Besancon, G., Torres, L., Begovich, O. and Navarro, A. (2011). About friction modeling for observer-based leak estimation in pipelines, *50th IEEE Conference on Decision and Control/European Control Conference, Orlando, USA*, pp. 4413–4418.

Fereidooni, Z., Tahayori, H. and Bahadori-Jahromi, A. (2021). A hybrid model-based method for leak detection in large scale water distribution networks, *Journal of Ambient Intelligence and Humanized Computing* **12**(2): 1613–1629.

Gunawickrama, K. (2001). *Leak Detection Methods for Transmission Pipelines*, PhD thesis, Gdańsk University of Technology, Gdańsk.

Isermann, R. (2011). Leak detection of pipelines, *in* R. Isermann (Ed.), *Fault-Diagnosis Applications*, Springer, Berlin/Heidelberg, pp. 181–204.

Jiménez, J., Torres, L., Verde, C. and Sanjuán, M. (2017). Friction estimation of pipelines with extractions by using state observers, *IFAC-PapersOnLine* **50**(1): 5361–5366.

JUMO. (2019). Pressure transmitter dTRANS p32, *Technical specifications*, JUMO GmbH, Fulda, `https://www.jumo.net/attachments/JUMO/attachmentdownload?id=3018`.

Korbicz, J., Kościelny, J.M., Kowalczuk, Z. and Cholewa, W. (2004). *Fault Diagnosis: Models, Artificial Intelligence, Applications*, Springer, Berlin/Heidelberg.

Korbicz, J., Patan, K. and Obuchowicz, A. (1999). Dynamic neural networks for process modelling in fault detection and isolation systems, *International Journal of Applied Mathematics and Computer Science* **9**(3): 519–546.

Kościelny, J.M. (1993). Recognition of faults in the diagnosing process, *Applied Mathematics and Computer Science* **3**(3): 559–572.

Kościelny, J.M., Syfert, M., Rostek, K. and Sztyber, A. (2016). Fault isolability with different forms of the faults–symptoms relation, *International Journal of Applied*

Mathematics and Computer Science **26**(4): 815–826, DOI: 10.1515/amcs-2016-0058.

Kowalczuk, Z. and Białaszewski, T. (2017). Gender approaches to evolutionary multi-objective optimization using pre-selection of criteria, *Engineering Optimization* **50**(1): 120–144.

Kowalczuk, Z. and Gunawickrama, K. (2004). Detecting and locating leaks in transmission pipelines, *in* J. Korbicz *et al.* (Eds), *Fault Diagnosis*, Springer, Berlin/Heidelberg, pp. 821–864.

Kowalczuk, Z., Suchomski, P. and Białaszewski, T. (1999). Evolutionary multi-objective Pareto optimisation of diagnostic state observers, *International Journal of Applied Mathematics and Computer Science* **9**(3): 689–709.

Kowalczuk, Z. and Tatara, M. (2018). Analytical steady-state model of the pipeline flow process, *23rd International Conference on Methods and Models in Automation and Robotics, Międzyzdroje, Poland*, pp. 1–16.

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

Kowalczuk, Z. and Tatara, M.S. (2017). Numerical issues and approximated models for the diagnosis of transmission pipelines, *in* C. Verde and L. Torres (Eds), *Modeling and Monitoring of Pipelines and Networks*, Springer, Cham, pp. 39–62.

Kowalczuk, Z. and Tatara, M.S. (2020). Improved model of isothermal and incompressible fluid flow in pipelines versus the Darcy–Weisbach equation and the issue of friction factor, *Journal of Fluid Mechanics* **891**(A5): 1–26.

Kowalczuk, Z. and Tatara, M.S. (2021). Analytical 'steady-state'-based derivation and clarification of the Courant–Friedrichs–Lewy condition for pipe flow, *Journal of Natural Gas Science and Engineering* **91**(103953): 1–16.

Kowalczuk, Z., Tatara, M.S. and Stefański, T. (2018). Reduction of computational complexity in simulations of the flow process in transmission pipelines, *in* J.M. Kościelny *et al.* (Eds), *Advanced Solutions in Diagnostics and Fault Tolerant Control*, Springer, Cham, pp. 241–252.

Lam, M., Corredor, D., Camino, G. and Ghidaoui, M. (2019). Background pressure characterization in pipeline systems and the Rayleigh distribution, *Proceedings of the 38th IAHR World Congress, Panama City, Panama*, pp. 1–6.

Larson, T. (1960). Loss in pipeline carrying capacity due to corrosion and tuberculation, *Journal of the American Water Works Association* **52**(10): 1263–1270.

Liu, C., Li, Y., Fang, L., Han, J. and Xu, M. (2017). Leakage monitoring research and design for natural gas pipelines based on dynamic pressure waves, *Journal of Process Control* **50**: 66–76.

Liu, M., Zang, S. and Zhou, D. (2005). Fast leak detection and location of gas pipelines based on an adaptive particle filter, *International Journal of Applied Mathematics and Computer Science* **15**(4): 541–550.

Noguera, J.F., Torres, L., Verde, C., Guzmán, E. and Sanjuan, M. (2019). Model for the flow of a water-glycerol mixture in horizontal pipelines, *2019 4th Conference on Control and Fault Tolerant Systems (SysTol), Casablanca, Morocco*, pp. 117–122.

Nowicki, A., Grochowski, M. and Duzinkiewicz, K. (2012). Data-driven models for fault detection using kernel PCA: A water distribution system case study, *International Journal of Applied Mathematics and Computer Science* **22**(4): 939–949, DOI: 10.2478/v10006-012-0070-1.

Omega (2019). Coriolis mass-flow meter FMC-5000 series, *Technical specifications*, Omega Engineering, Norwalk, `https://assets.omega.com/spec/FMC-5000.pdf`.

Ostapkowicz, P. and Bratek, A. (2023). Two-leak case diagnosis based on static flow model for liquid transmission pipelines, *Sensors* **23**(18): 7751.

Osyczka, A. and Kundu, S. (1996). A modified distance method for multicriteria optimization using genetic algorithms, *Computers & Industrial Engineering* **30**(4): 871–882.

Pahlavanzadeh, F., Khaloozadeh, H. and Forouzanfar, M. (2024). Online fault detection and localization of multiple oil pipeline leaks using model-based residual generation and friction identification, *International Journal of Dynamics and Control* **12**(8): 2615–2628.

Pérez-Pérez, E., López-Estrada, F., Valencia-Palomo, G., Torres, L., Puig, V. and Mina-Antonio, J. (2021). Leak diagnosis in pipelines using a combined artificial neural network approach, *Control Engineering Practice* **107**: 104677.

Quiñones-Grueiro, M., Verde, C., Prieto-Moreno, A. and Llanes-Santiago, O. (2018). An unsupervised approach to leak detection and location in water distribution networks, *International Journal of Applied Mathematics and Computer Science* **28**(2): 283–295, DOI: 10.2478/amcs-2018-0020.

Rui, Z., Han, G., Zhang, H., Wang, S., Pu, H. and Ling, K. (2017). A new model to evaluate two leak points in a gas pipeline, *Journal of Natural Gas Science and Engineering* **46**: 491–497.

Santos-Ruiz, I., López-Estrada, F.-R., Puig, V., Torres, L., Valencia-Palomo, G. and Gómez-Peñate, S. (2021). Optimal estimation of the roughness coefficient and friction factor of a pipeline, *Journal of Fluids Engineering* **143**(5): 051304.

Sensirion (2019). Mass-flow meter SFM3300-D, *Technical specifications*, Sensirion AG, Stäfa, `https://sensirion.com/resource/datasheet/sfm3300`.

Thomas, L. (1949). Elliptic problems in linear difference equations over a network, *Watson Scientific Computing Laboratory Report*, Columbia University, New York.

Torres, L., Besançon, G. and Verde, C. (2015). Liénard type model of fluid flow in pipelines: Application to estimation, *12th IC on Electrical Engineering, Computing Science and Automatic Control, Mexico City, Mexico*, pp. 1–6.

Torres, L., Verde, C. and Molina, L. (2021). Leak diagnosis for pipelines with multiple branches based on model similarity, *Journal of Process Control* **99**: 41–53.

VEGA (2019). VEGABAR 82 pressure transmitter, *Technical specifications*, VEGA Americas, Inc., Mason, https://www.vega.com/en-us/products/product-catalog/pressure/hydrostatic/vegabar-82.

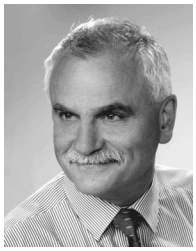
Verde, C. and Torres, L. (2015). Referenced model-based observers for locating leaks in a branched pipeline, *IFAC-PapersOnLine /Safeprocess* **48**(21): 1066–1071.

Vítkovský, J.P., Simpson, A.R. and Lambert, M.F. (2000). Leak detection and calibration using transients and genetic algorithms, *Journal of Water Resources Planning and Management* **126**(4): 262–265.

White, F. (2011). *Fluid Mechanics*, McGraw Hill, New York.

Wiid, A., le Roux, J. and Craig, I. (2020). Modelling of methane-rich gas pipeline networks for simulation and control, *Journal of Process Control* **92**: 234–245.

Marek S. Tatara was born in 1991 in Olsztyn, Poland. He obtained his PhD in 2019 from the Gdańsk University of Technology in the field of automatic control, electronics and electrical engineering. His scientific interests concern mathematical modeling of physical processes, diagnostics, signal processing applied to industrial processes, and evolutionary music composition. He is a member of the Polish Society for Measurement, Automatic Control and Robotics, and a vice-chair of the Technical Committee TC 7.5. Intelligent Autonomous Vehicles on Social Media of the International Federation of Automatic Control. He works as an assistant professor at the Gdańsk University of Technology and a chief scientific officer at DAC.digital.



Zdzisław Kowalczuk, Prof., DSc, PhD, MScEE (2003, 1993, 1986, 1978). Since 1978 he has been with the Faculty of Electronics, Telecommunications and Informatics at the Gdańsk University of Technology, where he is a full professor of automatic control and robotics, and the head of the Department of Robotics and Decision Systems. His scientific interests include robotics, automatic control, diagnostics, signal processing, modeling, identification, estimation, and artificial intelligence. He has (co)authored over 30 books, over 120 journal papers (50+ in JCR) and more than 350 conference publications and book chapters (with a citation index in Google Scholar exceeding 3500 and a Hirsch index of 22). He is the president of the Polish Consultants Society and the Polish Society for Measurements, Automatic Control and Robotics (IFAC), and a member of the Committee on Automatic Control and Robotics of the Polish Academy of Sciences.

Appendix

A1. Complementary, liquid experiments

One of the important systemic problems is the comparison of methodologies (in terms of errors in estimating the size and location of the leak) for various fluids, gases or liquids, which in the presented models are characterized by a specific speed of sound and a friction coefficient. Therefore, to illustrate and document this point, a set of additional four comprehensive experiments were performed, described as follows:

I. Leak size $q_L = 4$, leak location $z_L = 40$, friction factor $\lambda = 0.02$, speed of sound $\nu = 450\,\frac{\mathrm{m}}{\mathrm{s}}$.

II. Leak size $q_L = 4$, leak location $z_L = 40$, friction factor $\lambda = 0.02$, speed of sound $\nu = 1482\,\frac{\mathrm{m}}{\mathrm{s}}$.

III. Leak size $q_L = 4$, leak location $z_L = 40$, friction factor $\lambda = 0.06$, speed of sound $\nu = 1482\,\frac{\mathrm{m}}{\mathrm{s}}$.

IV. Leak size $q_L = 1$, leak location $z_L = 40$, friction factor $\lambda = 0.02$, speed of sound $\nu = 1482\,\frac{\mathrm{m}}{\mathrm{s}}$.

Experiment I concerns the comparative case of gas, which is the main subject of this work. Experiments II–IV are intended to collect data on liquid flows in various technological conditions (while maintaining possibly the same parameters as in the case of gas). In particular, in experiment II we have the higher speed of sound (1482 m/s), in experiment III this higher speed of sound and higher friction value (0.06), and in experiment IV this higher speed of sound and a leakage of 10% of the current mass flow. The results obtained are collected in Table A1.

The table shows that the standard deviations (STD) of both target estimators (LSE for size and LLE for location) are smaller for experiments II–IV compared with experiment I, while in experiment IV the standard deviation of leak location is comparable to experiment I. The biases of the LSE estimator in experiments II–IV are also usually smaller than in experiment I, with the exception of the AMDA model, especially in experiment IV (which may be due to the relatively small leakage). The leak location bias in experiments II–IV is larger than in experiment I (i.e., the increased sound speed leads to an increase in the location error).

A2. Additional analysis without the AMSS

As an extension of the presented study or an in-depth comparative analysis of models, the AMSS can be excluded (as the most optimal in the global/aggregate sense) and only the remaining instrumental models of flow processes (incorporated into the appropriate estimator) should be taken into account. In this situation, new, normalized quality indicators are obtained, which are presented in Table A2 and Fig. A1.

The GOL measure for these models remains the same (zero), and the calculated scalar global indices AOI and DRA for the discussed models in this case are given in Table A3. This time the base model is the best, followed by the ATM second in place and the AMDA third in place. It should be noted here that the score for the ATM is close to the result of the base model for both the criteria (they are therefore equivalent in practice, after excluding the AMSS).

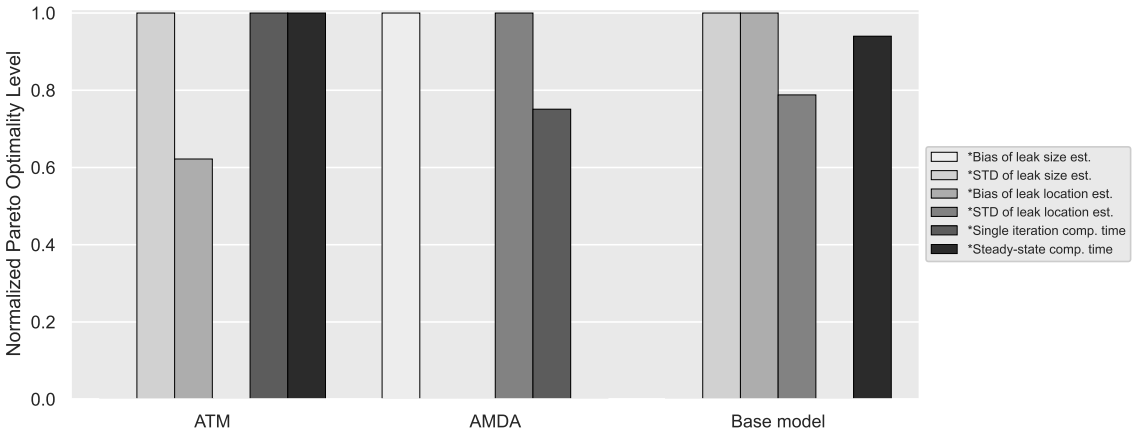


Fig. A1. Normalized local optimality levels for the inferior models in the maximization task (*).

Table A1. Data collected for additional experiments comparing the effects of lower ($\nu_1 = 450\ \frac{\mathrm{m}}{\mathrm{s}}$) and greater ($\nu_2 = 1482\ \frac{\mathrm{m}}{\mathrm{s}}$) speed of sound under different conditions. The parameters of the four experiments are as follows: (I) leak size $q_L = 4$, leak location $z_L = 40$, friction factor $\lambda = 0.02$ and speed of sound ν_1; (II) $q_L = 4$, $z_L = 40$, $\lambda = 0.02$ and ν_2; (III) $q_L = 4$, $z_L = 40$, $\lambda = 0.06$ and ν_2; and (IV) $q_L = 1$ (10% flow), $z_L = 40$, $\lambda = 0.02$ and ν_2. Leak size (LSE) biases and standard deviations (STD) are given in $\frac{\mathrm{kg}}{\mathrm{s}}$, while leak location (LLE) biases and standard deviations are given in km.

	Experiment I				Experiment II			
	ATM	AMDA	AMSS	BM	ATM	AMDA	AMSS	BM
LSE bias ($\times 10^{-3}$)	0.12	128	0.28	0.12	0.04	129	0.05	0.04
LSE STD ($\times 10^{-3}$)	13.0	289	11.3	13.0	3.04	69.0	2.63	3.04
LLE bias	0.28	0.46	-0.07	0.19	0.93	1.66	1.01	0.93
LLE STD	3.75	2.95	0.93	3.76	0.49	0.85	0.21	0.52
	Experiment III				Experiment IV			
	ATM	AMDA	AMSS	BM	ATM	AMDA	AMSS	BM
LSE bias ($\times 10^{-3}$)	0.02	131	-0.01	0.02	0.04	312	0.05	0.04
LSE STD ($\times 10^{-3}$)	2.00	39.4	1.51	2.00	3.04	69.0	2.63	3.04
LLE bias	0.98	1.62	0.99	0.98	0.92	1.69	0.88	0.92
LLE STD	0.24	0.46	0.12	0.26	1.79	3.24	0.83	1.78

Table A2. Local optimality of diagnostic estimators (with exclusion of the AMSS) in terms of normalized indicators in the maximization task (*).

Performance index	ATM	AMDA	Base
* Bias of leak size estimator $\hat{\bar{w}}_q$	0	1	0
* STD of leak size estimator $\hat{\sigma}_q$	1	0	1
* Bias of leak location estimator $\hat{\bar{w}}_z$	0.622	0	1
* STD of leak location estimator $\hat{\sigma}_z$	0	1	0.788
* Single iteration computation time t_i [s]	1	0.751	0
* Steady state computation time t_s [s]	1	0	0.940

Table A3. AOI and DRA scores for a limited set of diagnostic estimators (excluding the optimal model, the AMSS).

	ATM	AMDA	Base model
AOI	0.640	0.508	**0.656**
DRA	1.464	1.750	**1.431**

Received: 7 December 2023
Revised: 14 May 2024
Accepted: 15 June 2024