Model-based Sensor-Fault Detection and Isolation in Natural-Gas Pipelines for Transient Flow

Khadija Shaheen*, Apoorva Chawla*, Ferdinand Evert Uilhoorn†, Pierluigi Salvo Rossi*,
*Dept. Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway
†Dept. Gas Engineering, Warsaw University of Technology, Warsaw, Poland
Email: shaheen.khadija@ntnu.no; apoorva.chawla@ntnu.no; ferdinand.uilhoorn@pw.edu.pl; salvorossi@ieee.org

Abstract—Sensor-based monitoring of natural gas pipelines is crucial for safety and detecting sensor faults is pivotal for reliable operations. This paper investigates the problem of fault diagnosis in a natural gas pipeline under transient flow, characterized by a system of hyperbolic partial differential equations (PDEs). A data-fusion approach based on an unscented Kalman filter (UKF) is employed to perform sensor-fault detection and isolation (SFDI). The architecture consists of a bank of local UKFs to provide the state estimates, which are further analyzed to identify sensor faults. The performance achieved by the proposed method is promising even in the case of a nonlinear flow model.

Index Terms—Data fusion, fault detection and isolation, partial differential equations, unscented Kalman filter, transient flow.

I. Introduction

Natural-gas pipelines are subject to failures because of deterioration, aging, and environmental factors. Unrecognized faults can cause accidents, property damage, and environmental contamination; thus sensor-based monitoring systems are commonly deployed to increase system safety and reduce the risk of catastrophic failures [1]. However, sensors are prone to errors and failures as well, thus sensor-fault detection and isolation (SFDI) is crucial for keeping the safety requirements.

SFDI methods are broadly categorized into model-based [2] and data-driven methods [3]. Although data-driven techniques have recently gained large interest [4], [5], they cannot guarantee reliable performance for dynamic/transient scenarios. For such cases, model-based approaches could exploit the availability of well-established accurate flow models describing the transient behavior of gases in pipelines. In this context, Bayesian filtering techniques were investigated for real-time estimation of gas-flow transients in pipelines [6], [7]. However, model-based SFDI are unexplored due to the highly nonlinear and complex nature of gas-flow models, characterized by a hyperbolic system of partial differential equations (PDEs).

This paper develops a model-based SFDI approach combining data fusion and unscented Kalman filter (UKF) [8] for a system of nonlinear hyperbolic PDEs, describing the transient flow of natural-gas in pipelines. The UKF-based architecture builds upon several local filters and exploits data fusion suitable for SFDI in large-scale systems. In comparison to centralized architectures [9], the distributed approach reduces the overall computational burden and improves decision support. The effectiveness of the proposed method is assessed for various types of sensor faults using simulated nonlinear spatial-temporal data. To the best of our knowledge, the proposed scheme is the first model-based approach that investigates SFDI for a highly nonlinear hyperbolic system of PDEs describing the transient gas-flow in pipelines.

The rest of the paper is organized as follows: Sec. II describes the system of PDEs, modeling the transient flow of natural gas in pipelines; the SFDI architecture is developed in Sec. III; Sec. IV presents and discusses the achieved performance; some final remarks are given in Sec. V.

II. System Model for Transient Flow

The mathematical model for natural-gas transients in pipelines can be modeled as a system of hyperbolic PDEs in terms of pressure, temperature, and flow [6] as

$$\frac{\partial \boldsymbol{w}}{\partial t} = -\boldsymbol{A}(\boldsymbol{w})\frac{\partial \boldsymbol{w}}{\partial x} - \boldsymbol{s}(\boldsymbol{w}) , \qquad (1)$$

where $(x,t) \in [0,L] \times [0,t_f]$ represents the system dynamics in space and time, respectively, with L being the pipeline length and t_f being the time span; and $\boldsymbol{w} = [p, \dot{m}, T]^{\mathrm{T}}$ is a vector collecting pressure, flow, and temperature. The coefficient matrix $\boldsymbol{A}(\boldsymbol{w}) \in \mathbb{R}^{3\times3}$ and the vector $\boldsymbol{s}(\boldsymbol{w}) \in \mathbb{R}^{3\times1}$ are defined in Eq. (2) at the top of the next page, where A and R denote the cross-sectional area and the ideal gas constant, respectively; z is the gas compressibility factor; C_p and ρ represent the specific heat at constant pressure and density[1]. The parameters α_1 and α_2 can be expressed as $\alpha_1 = 1 + \frac{T}{z}\left(\frac{\partial z}{\partial T}\right)_p$ and $\alpha_2 = 1 - \frac{p}{z}\left(\frac{\partial z}{\partial p}\right)_T$; α_s is the isentropic wave speed [6] and the frictional force w per unit length is given as $w = \frac{1}{8}f\rho v|v|\pi d$, where d is the diameter and v is the velocity. The friction factor f is computed using the Colebrook–White equation, $\frac{1}{\sqrt{f}} = -2\log\left(\frac{\epsilon}{3.7d} + \frac{2.51}{Re\sqrt{f}}\right)$, being ϵ the roughness and Re the Reynolds number. The heat transfer between the natural gas and its surroundings per unit length and time is $q = -\pi dU(T - T_s)$, where U is the overall heat transfer coefficient and T_s is the ambient temperature.

The numerical method of lines is an effective way to solve the system of PDEs. We used a 5-point, 4th-order finite difference method to spatially discretize the system of PDEs in Eq. (1) and convert into ordinary differential equations (ODEs)

---

[1]Thermodynamic properties C_p, z and ρ are taken from GERG2004 [10].

$$\boldsymbol{A}(\boldsymbol{w}) = \begin{bmatrix} -\frac{\dot{m}(\alpha_s^2\alpha_2 - RTz)}{Ap} & \frac{\alpha_s^2}{A} & \frac{\alpha_s^2\dot{m}\alpha_1}{AT} \\ A - \frac{\alpha_s^2\alpha_2^2 C_p \dot{m}^2 - R\alpha_s^2\alpha_1^2\alpha_2\dot{m}^2 z}{AC_p p^2} & \frac{\dot{m}(\alpha_2 C_p \alpha_s^2 - Rz\alpha_s^2\alpha_1^2 + RTC_p z)}{AC_p p} & \frac{\alpha_s^2\alpha_1\dot{m}^2(\alpha_2 C_p - R\alpha_1^2 z)}{ATC_p p} \\ -\frac{RT\alpha_s^2\alpha_1\alpha_2\dot{m}z}{AC_p p^2} & \frac{RT\alpha_s^2\alpha_1 z}{AC_p p} & \frac{R\dot{m}z(\alpha_s^2\alpha_1^2 + TC_p)}{AC_p p} \end{bmatrix}, \quad \boldsymbol{s}(\boldsymbol{w}) = \begin{bmatrix} -\frac{\alpha_s^2\alpha_1(Aqp + RT\dot{m}wz)}{A^2TC_p p} \\ w \\ -\frac{\alpha_s^2\alpha_2(Aqp + RT\dot{m}wz)}{A^2C_p p^2} \end{bmatrix} \quad (2)$$
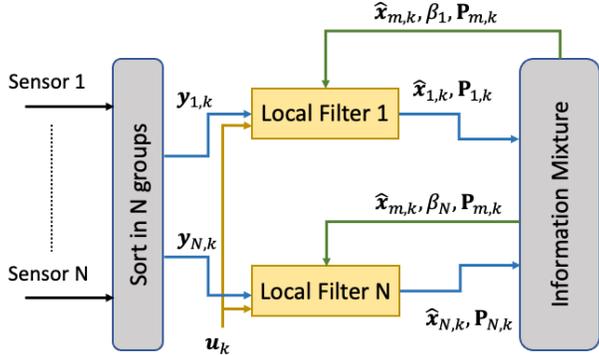
Fig. 1: UKF-based data-fusion architecture.

$$\frac{d\boldsymbol{w}(t)}{dt} = \boldsymbol{A}(\boldsymbol{w})\boldsymbol{D}\boldsymbol{w}(t) - \boldsymbol{s}(\boldsymbol{w}, t) , \quad (3)$$

where $\boldsymbol{w}(t) = [p_1(t), \ldots, p_n(t), \dot{m}_1(t), \ldots, \dot{m}_n(t), T_1(t), \ldots, T_n(t)]^T \in \mathbb{R}^{3n \times 1}$ is the state vector; $\boldsymbol{A}(\boldsymbol{w}) \in \mathbb{R}^{3n \times 3n}$ and $\boldsymbol{s}(\boldsymbol{w}, t) \in \mathbb{R}^{3n \times 1}$ are the assembled matrix and vector, $\boldsymbol{D}$ is the computational matrix given in [6]. The 4th-order Runge-Kutta method is used to solve the ODEs in Eq. (3). The discretized equations are modeled as a state-space model with solution advanced in time $\boldsymbol{w}(t + \Delta t) = \boldsymbol{w}(t) + \frac{1}{6}(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4)$, where $\boldsymbol{k}_1 = \Delta t \boldsymbol{f}(t, \boldsymbol{w}(t))$, $\boldsymbol{k}_2 = \Delta t \boldsymbol{f}\left(t + \frac{1}{2}\Delta t, \boldsymbol{w}(t) + \frac{1}{2}\boldsymbol{k}_1\right)$, $\boldsymbol{k}_3 = \Delta t \boldsymbol{f}\left(t + \frac{1}{2}\Delta t, \boldsymbol{w}(t) + \frac{1}{2}\boldsymbol{k}_2\right)$, and $\boldsymbol{k}_4 = \Delta t \boldsymbol{f}(t + \Delta t, \boldsymbol{w}(t) + \boldsymbol{k}_3)$.

III. SFDI ARCHITECTURE

The architecture, shown in Fig. 1, consists of 4 main stages: (i) sensors measurements are grouped into subsets; (ii) local filters produce local estimates of the state vector; (iii) global states are inferred from the information mixture built by fusing local estimates; (iv) local filters are updated with global estimates. The state-space model at the ith local filter is

$$\boldsymbol{x}_{i,k} = \boldsymbol{f}(\boldsymbol{x}_{i,k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{w}_{i,k} ,$$
$$\boldsymbol{y}_{i,k} = \boldsymbol{h}_i(\boldsymbol{x}_{i,k}, \boldsymbol{u}_k) + \boldsymbol{v}_{i,k} ,$$

where $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}_i(\cdot)$ are the nonlinear mappings. The state vector $\boldsymbol{x}_{i,k} \in \mathbb{R}^{n_x \times 1}$ is defined as $\boldsymbol{x}_{i,k} = [p_1(k), \ldots, p_n(k), \dot{m}_1(k), \ldots, \dot{m}_n(k), T_1(k), \ldots, T_n(k)]^T$; $\boldsymbol{u}_k$ and $\boldsymbol{y}_{i,k}$ are the system input and output, respectively; $\boldsymbol{w}_{i,k}$ and $\boldsymbol{v}_{i,k}$ denote the process noise and the measurement noise with covariance matrices $\boldsymbol{Q}_{i,k}$ and $\boldsymbol{R}_{i,k}$, respectively.

Data fusion exploits parallel UKFs as local filter to accurately estimate the nonlinear states in 5 steps.

Step 1: Initialization (depending on the use-case) of the state estimate $\hat{\boldsymbol{x}}_{i,0|0}$ and covariance matrix $\boldsymbol{P}_{i,0|0}$ at each local filter.

Step 2: Sigma points are computed for ith local filter with mean $\hat{\boldsymbol{x}}_{i,k-1|k-1}$ and covariance matrix $\boldsymbol{P}_{i,k-1|k-1}$, where

$$\boldsymbol{\chi}_{i,k|k} = [\hat{\boldsymbol{x}}_{i,k|k}, \hat{\boldsymbol{x}}_{i,k|k} + \boldsymbol{\xi}_{i,k|k}, \hat{\boldsymbol{x}}_{i,k|k} - \boldsymbol{\xi}_{i,k|k}],$$
$$\boldsymbol{\xi}_{i,k|k} = \sqrt{(n_x + \lambda)\boldsymbol{P}_{i,k|k}} , \quad \lambda = \alpha^2(n_x + k) - n_x ,$$

where n_x is the length of the state vector and λ specifies the sigma point spread.

Step 3: Time (resp. measurement) update of the ith local filter shown in Eq. (4) (resp. Eq. (5)) at the top of the next page.

Step 4: Global estimates via data fusion and information mixture as $\hat{\boldsymbol{x}}_{m,k} = \boldsymbol{P}_{m,k}\sum_{i=1}^{N}\boldsymbol{P}_{i,k|k}^{-1}\hat{\boldsymbol{x}}_{i,k|k}$, with $\boldsymbol{Q}_{m,k}^{-1} = \sum_{i=1}^{N}\boldsymbol{Q}_{i,k}^{-1}$ and $\boldsymbol{P}_{m,k}^{-1} = \sum_{i=1}^{N}\boldsymbol{P}_{i,k|k}^{-1}$.

Step 5: Local-filter update via global estimates, $\boldsymbol{Q}_{i,k} = \beta_i\boldsymbol{Q}_{m,k}$, $\boldsymbol{P}_{i,k} = \beta_i\boldsymbol{P}_{m,k}$, $\hat{\boldsymbol{x}}_{i,k|k} = \hat{\boldsymbol{x}}_{m,k}$, with $\sum_{i=1}^{N}\beta_i = 1$.

For SFDI, we consider a threshold rule on the state variance $\boldsymbol{V}_k \in \mathbb{R}^{n_x \times 1}$ and the state residual $r_{i,k}$ for each local filter, defined as $V_{j,k} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{x}_{i,k|k}^{(j)} - \frac{1}{N}\sum_{i=1}^{N}\hat{x}_{i,k|k}^{(j)}\right)$ and $r_{i,k}^2 = (\hat{\boldsymbol{x}}_{i,k} - \boldsymbol{x}_{m,k})^T(\hat{\boldsymbol{x}}_{i,k} - \boldsymbol{x}_{m,k})$. The state variance exhibits large values in case of sensor faults, thus performing fault detection. Isolation is performed based on specific combinations of local sensors: with N sensors, N locals are used with each local having $N-1$ measurements. For instance, if the jth sensor is faulty, then $N-1$ locals containing the jth sensor will be affected while only one local would be accurate; thus isolation is based on the maximum value of residuals by identifying the missing sensor from the accurate local.

IV. SIMULATION RESULTS AND DISCUSSION

We ran numerical simulations for a high-pressure natural-gas pipeline equipped with sensors measuring pressure, temperature, and flow rate, assuming the following specifications: $L = 150$ km, $d = 1.4$ m, $\epsilon = 0.016$ mm, $T_s = 5$ s and $U = 2.84$ Wm^{-2}K^{-1} and boundary conditions $p(0,t) = 8.4$ MPa, $T(0,t) = 303.15$ K, $\dot{m}(L,t) = f(t)$, for a time interval $t_f \in [0, 3600$ s$]$. The spatial-temporal evolution of the fault/noise-free state variables is shown in Fig. 2. The measurements are generated including zero-mean additive Gaussian noise with st.dev. 0.0004 MPa, 1.5 K and 2.5 kgs^{-1}, respectively, for pressure, temperature and flow rate.

The proposed algorithm (fusing UKF) was tested on: ***Scenario 1***, the SFDI architecture processes noisy measurements; ***Scenario 2***, faulty sensors are simulated by superimposing fault signals to the noisy measurements. We considered 21 sensor locations providing 63 signals (21 for pressure, 21 for temperature, and 21 for flow rate), and employed $N = 63$ local filters, each receiving 62 sensor measurements. The initialization was $\boldsymbol{P}_{i,0|0} = \boldsymbol{I}_{63}$, $\boldsymbol{Q}_{i,k} = \sigma_{w,k}^2\boldsymbol{I}_{63}$, with $\sigma_{w,k}^2 = 0.1 \sigma_{v,k}^2$, $\sigma_{w,k}^2$ and $\sigma_{v,k}^2$ being process and measurement noise variances.

Scenario 1 provides an assessment of noise impact on the estimation capabilities via root mean square error (RMSE). The

$$\boldsymbol{\chi}_{i,j,k|k-1} = \boldsymbol{f}(\boldsymbol{\chi}_{i,j,k-1|k-1}, \boldsymbol{u}_{k-1}) \,, \ \hat{\boldsymbol{x}}_{i,k|k-1} = \sum_{j=0}^{2n_x} W_j^{(m)} \boldsymbol{\chi}_{i,j,k|k-1} \,, \ \boldsymbol{\gamma}_{i,j,k|k-1} = \boldsymbol{h}_i(\boldsymbol{\chi}_{i,j,k|k-1}) \,, \ \hat{\boldsymbol{y}}_{i,k|k-1} = \sum_{j=0}^{2n_x} W_j^{(m)} \boldsymbol{\gamma}_{i,j,k|k-1} \,,$$

$$W_0^{(m)} = \frac{\lambda}{\lambda + n_x} \,, \ W_0^{(c)} = W_0^{(m)} + (1 - \alpha^2 + \beta) \,, \ W_j^{(m)} = W_j^{(c)} = \frac{1}{2(\lambda + n_x)} \,, \ j = 1, \ldots, 2n_x,$$

$$\boldsymbol{P}_{i,k|k-1} = \sum_{j=0}^{2n_x} W_j^{(c)} (\boldsymbol{\chi}_{i,j,k|k-1} - \hat{\boldsymbol{x}}_{i,k|k-1})(\boldsymbol{\chi}_{i,j,k|k-1} - \hat{\boldsymbol{x}}_{i,k|k-1})^T + \boldsymbol{Q}_{i,k} \,, \tag{4}$$

$$\boldsymbol{P}_{i,k|k-1}^y = \sum_{j=0}^{2n_x} W_j^{(c)} (\boldsymbol{\gamma}_{i,j,k|k-1} - \hat{\boldsymbol{y}}_{i,k|k-1})(\boldsymbol{\gamma}_{i,j,k|k-1} - \hat{\boldsymbol{y}}_{i,k|k-1})^T + \boldsymbol{R}_{i,k}, \ \boldsymbol{P}_{i,k|k-1}^{xy} = \sum_{j=0}^{2n_x} W_j^{(c)} (\boldsymbol{\chi}_{i,j,k|k-1} - \hat{\boldsymbol{x}}_{i,k|k-1})(\boldsymbol{\gamma}_{i,j,k|k-1} - \hat{\boldsymbol{y}}_{i,k|k-1})^T$$

$$\boldsymbol{K}_{i,k} = \boldsymbol{P}_{i,k|k-1}^{xy}(\boldsymbol{P}_{i,k|k-1}^y)^{-1} \,, \ \hat{\boldsymbol{x}}_{i,k|k} = \hat{\boldsymbol{x}}_{i,k|k-1} + \boldsymbol{K}_{i,k}(\boldsymbol{y}_{i,k} - \hat{\boldsymbol{y}}_{i,k|k-1}) \,, \ \boldsymbol{P}_{i,k|k} = \boldsymbol{P}_{i,k|k-1} - \boldsymbol{K}_{i,k} \boldsymbol{P}_{i,k|k-1}^y \boldsymbol{K}_{i,k}^T \,, \tag{5}$$

(a) Boundary condition (b) Pressure

(c) Temperature (d) Flow Rate
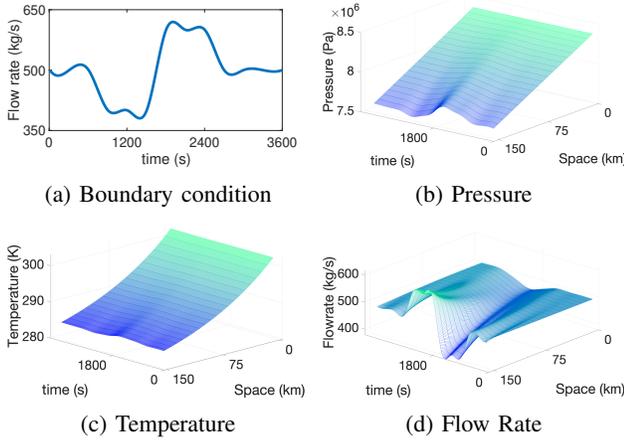
Fig. 2: Simulated data using ODEs.

	Pressure (10^{-3}MPa)	Temperature (K)	Flow Rate (kg/s)
Fusing UKF	**0.123**	**0.347**	**0.646**
Classic UKF	0.158	0.449	0.802
EKF	2.602	0.333	0.869
PF	1.546	0.164	1.002

TABLE I: RMSE in the presence of measurement noise.

proposed approach is compared in Tab. I with 3 benchmarks (classic UKF, extended Kalman Filter (EKF), and particle filter (PF)), achieving higher estimation accuracy. Scenario 2 shows the effectiveness of the proposed architecture in the presence of sensor faults. Bias and drift faults as in [5] are considered.

Fig. 3 shows the different behavior of SFDI based on fusing UKF and classic UKF: the former successfully detects and isolates the faulty sensors and provides a reliable state estimate in presence of faulty sensors; the latter fails to provide reliable detection, isolation, and estimation. More specifically, we used the probabilities of detection and false alarm to assess SFDI performance, and the receiver operating characteristic (ROC) curves for combinations of weak/strong bias/drift faults are shown in Fig. 4. ROC curves are computed separately for each physical quantity (pressure, temperature, and flow rate) and show that the proposed technique achieves high probability of detection and low probability of false alarm even in the presence of weak faults (usually more challenging to detect). Also, corresponding accuracy and RMSE are shown in Tab. II for fixed (heuristically-selected) threshold.
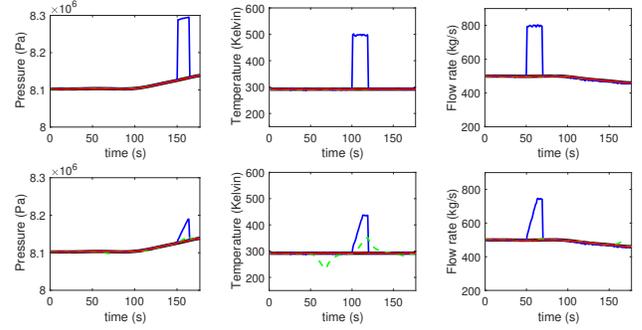
Fig. 3: State estimation in presence of faulty sensors with bias (top) and drift (bottom). Actual/faulty values in black/blue, estimate from classic/fusing UKF in green/red.
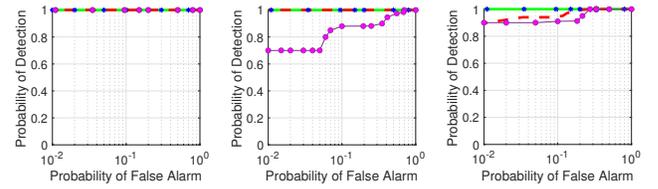
Fig. 4: ROC curves for the fusing UKF under strong/weak bias (green/red) and strong/weak drift (blue/magenta) faults.

Faulty Sensor	Metrics	High Bias	Low Bias	Strong Drift	Weak Drift
Pressure	Accuracy (%)	100	100	100	100
	RMSE(10^{-3}MPa)	0.48	0.453	0.476	0.431
Temperature	Accuracy (%)	93.33	93.33	93.33	86.67
	RMSE (K)	0.906	1.0308	1.0812	1.8048
Flow rate	Accuracy (%)	100	100	100	100
	RMSE (kg/s)	2.1312	2.1528	2.088	2.1876

TABLE II: Accuracy and RMSE of fusing UKF when operating with different types of faults.

V. CONCLUSIONS

We proposed a model-based SFDI method exploiting data fusion and UKFs for monitoring natural-gas pipelines whose transient behavior is characterized by nonlinear hyperbolic PDEs. Numerical simulations demonstrate the superior performance of the proposed technique (combining data from multiple sensors) with respect to popular benchmarks. The results show excellent performance even in the challenging case of weak faults. Performance validation of the proposed SFDI method with real-world data from industrial processes will be considered as future work.

REFERENCES

[1] M. A. Adegboye, W.-K. Fung, and A. Karnik, "Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches," *Sensors*, vol. 19, no. 11, p. 2548, 2019.

[2] J. Yang, F. Zhu, X. Wang, and X. Bu, "Robust sliding-mode observer-based sensor fault estimation, actuator fault detection and isolation for uncertain nonlinear systems," *Int. J. Control, Automation and Systems*, vol. 13, no. 5, pp. 1037–1046, 2015.

[3] Y. Feng and H.-X. Li, "Detection and spatial identification of fault for parabolic distributed parameter systems," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7300–7309, 2018.

[4] H. Darvishi, D. Ciuonzo, E. R. Eide, and P. Salvo Rossi, "Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4827–4838, 2021.

[5] H. Darvishi, D. Ciuonzo, and P. Salvo Rossi, "A machine-learning architecture for sensor fault detection, isolation, and accommodation in digital twins," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2522–2538, 2023.

[6] F. Uilhoorn, "Comparison of Bayesian estimation methods for modeling flow transients in gas pipelines," *J. Natural Gas Science and Engineering*, vol. 38, pp. 159–170, 2017.

[7] F. E. Uilhoorn, "A particle filter-based framework for real-time state estimation of a non-linear hyperbolic PDE system describing transient flows in CO_2 pipelines," *Computers & Mathematics with Applications*, vol. 68, no. 12, pp. 1991–2004, 2014.

[8] A. Vafamand, B. Moshiri, and N. Vafamand, "Fusing unscented Kalman filter to detect and isolate sensor faults in DC microgrids with CPLs," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–8, 2021.

[9] D. Mori, H. Sugiura, and Y. Hattori, "Adaptive sensor fault detection and isolation using unscented Kalman filter for vehicle positioning," in *IEEE Intelligent Transportation Systems Conf. (ITSC)*. IEEE, 2019, pp. 1298–1304.

[10] E. W. Lemmon and R. T. Jacobsen, "Viscosity and thermal conductivity equations for nitrogen, oxygen, argon, and air," *Int. J. Thermophysics*, vol. 25, no. 1, pp. 21–69, 2004.