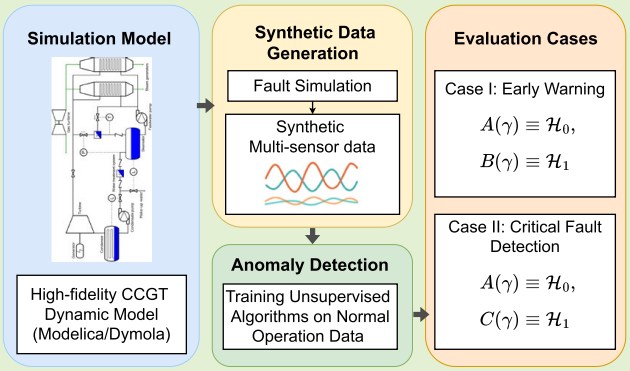
Unsupervised Leak Detection for Heat Recovery Steam Generators in Combined-Cycle Gas and Steam Turbine Power Plants

Mohammed Ayalew Belay, *Graduate Student Member, IEEE*, Lucas Ferreira Bernardino, Sindre Stenen Blakseth, Adil Rasheed, Rubén Mocholí Montañés, and Pierluigi Salvo Rossi, *Senior Member, IEEE*

***Abstract*—A predictable and reliable power supply is crucial for safe and efficient operation in the energy sector. Hence, digital twins with real-time monitoring, simulation, and optimization capabilities receive increasing attention. In particular, early detection of faults and other operational issues based on real-time processing of sensor data can greatly reduce downtime and maintenance costs. In this work, we compare the performance of conventional and state-of-the-art unsupervised anomaly detection methods to detect tube leaks in the steam generator that recovers heat from a combined-cycle steam and gas turbine (CCGT). Since real CCGT operational data with known faults is not available, the comparison is performed using data from a high-fidelity dynamic model based on a CCGT installed on an offshore oil and gas platform. Specifically, we evaluate local outlier factor (LOF), one-class support vector machine (OC-SVM), principal component analysis (PCA), low-rank and sparse (LRS) decomposition, and a transformer autoencoder (TAE) using receiver operating characteristic (ROC)–area under the ROC curve (AUC), geometric mean (G-Mean), and true positive rate (TPR) at fixed false positive rate (FPR). For small-leak detection using the full sensor set, ROC–AUCs are PCA = 0.77, LOF = 0.87, OC-SVM = 0.95, LRS = 0.98, and TAE = 0.99, while all methods approach near-perfect AUC under large leaks. Under reduced data availability (e.g., no real-time makeup water flow data), LRS remains the only method that sustains high detection performance. In general, LRS delivers consistently strong accuracy and robustness across leak sizes and noise levels, offering a practical accuracy–complexity trade-off for deployment.**



***Index Terms*— Anomaly detection, combined-cycle gas turbine (GT), dynamic model, leak detection, unsupervised training.**

Received 18 September 2025; accepted 3 November 2025. Date of publication 14 November 2025; date of current version 30 December 2025. This work was supported in part by the Research Council of Norway through the Project DIGITAL TWIN within the PETROMAKS2 Framework under Project 318899. The associate editor coordinating the review of this article and approving it for publication was Dr. Zhiliang Liu. *(Corresponding author: Mohammed Ayalew Belay.)*

Mohammed Ayalew Belay is with the Department of Electronic Systems, Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: mohammed.a.belay@ntnu.no).

Lucas Ferreira Bernardino, Sindre Stenen Blakseth, and Rubén Mocholí Montañés are with the Department of Gas Technology, SINTEF Energy Research, 7491 Trondheim, Norway (e-mail: lucas.bernardino@sintef.no; Sindre.Blakseth@sintef.no; ruben.mocholi.montanes@sintef.no).

Adil Rasheed is with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: adil.rasheed@ntnu.no).

Pierluigi Salvo Rossi is with the Department of Electronic Systems, Norwegian University of Science and Technology, 7034 Trondheim, Norway, and also with the Department of Gas Technology, SINTEF Energy Research, 7491 Trondheim, Norway (e-mail: salvorossi@ieee.org).

Digital Object Identifier 10.1109/JSEN.2025.3630886

I. Introduction

THE combined gas and steam turbine (ST) power plants have emerged as an efficient alternative in modern power generation. In a typical combined-cycle steam and gas turbine (CCGT) system, the high-temperature exhaust gases from a gas turbine (GT) are utilized to produce steam using a heat recovery steam generator (HRSG), which then drives an ST to generate additional electricity [1], [2], [3]. In onshore applications, CCGT achieves high thermal efficiencies exceeding 60%, making it a preferred choice for modern power generation [4]. Consequently, these plants are increasingly deployed onshore and considered promising for offshore oil and gas production platforms as a means to reduce emissions by utilizing recovered excess heat [4], [5]. Through installation of a steam cycle downstream to the GTs of an offshore platform, the power demand of the platform can be met by burning less natural gas, thereby reducing CO_2 emissions by up to 25%. However,

the complexity of CCGTs makes them susceptible to various operational issues that demand constant monitoring [3]. In previous units installed offshore, there has been evidence of cracking of tubes in the HRSG due to vibrational damage and fretting. Also, the casing of the waste heat recovery unit has been subject to failures [6]. Consequently, fault detection is essential for maintaining the reliability and operational efficiency of CCGT plants. Early and accurate fault detection is critical for preventing unplanned downtimes, reducing maintenance costs, and ensuring plant safety. Effective fault detection enables timely interventions, ultimately enhancing the operational life and efficiency of CCGTs, thus maintaining their contribution to the power generation industry onshore and emissions reductions in offshore applications [7]. Detecting faults such as leaks in the steam cycle of a CCGT is critical for both safety and reduction in operational cost. Tube leaks in the HRSG can allow high-pressure water or steam to escape, reducing power output and potentially introducing water into downstream machinery. Left unchecked, an incipient leak often grows to catastrophic failure, forcing an unplanned shutdown. Furthermore, chemical water leaks could pose safety and environmental risks.

A number of studies have been carried out to explore methods for handling faults in power generation systems. For example, Ajami and Daneshvar [8] explore the fault detection and identification in the turbine system of a thermal power plant. They find that independent component analysis is much better suited for fault identification than principal component analysis (PCA). More recently, Sarwar et al. [9] proposed a multisensor data fusion framework for fault detection and diagnosis in an industrial GT engine. Also focusing on GTs, Sampath et al. [7] propose a hybrid approach that combines real-world sensor data and information from simulation models. Fast and Palmé [10] applied the artificial neural networks to monitor the condition and diagnose faults in a combined heat and power plant. In the context of CCGT plants, Arranz et al. [11] presented a commercially deployed multiagent anomaly detection system. Each agent is a neural network that is trained to predict key performance parameters during normal operation, such that faults can be recognized by deviations between actual measurements and the predictions of one or more agents. A fault diagnosis system for CCGTs based on feedforward neural networks was proposed by Camporeale et al. [12]. Finally, Davallo et al. [13] proposed an extreme learning machine framework for the detection and identification of leaks in an onshore CCGT.

As evidenced by the literature cited above, both supervised and unsupervised fault detection methods have been explored for CCGTs and related systems. Supervised methods require extensive labels for diverse failure modes, which is both time-consuming and resource-intensive, if at all feasible [14], [15], [16]. Consequently, unsupervised fault detection methods, which do not depend on labeled data, provide considerable advantages for proactive diagnostics [17]. However, unsupervised detection methods are agnostic to the nature and number of possible faults that may arise, and can be continuously aligned with any gradual changes in the normal system behavior [18]. For these reasons, we consider unsupervised methods to be well-suited for fault detection in CCGTs.

In this study, we present a quantitative and qualitative comparison of unsupervised anomaly detection methods for leak detection in CCGTs. The study's main objective is to elucidate the methods' strengths and weaknesses, such that power system operators can identify and apply suitable detection methods on their installations. Secondarily, by demonstrating the feasibility of early leak detection and equipping operators with suitable detection methods, we seek to increase the reliability and prevalence of CCGTs, thereby reducing emissions from GT-powered installations. To the best of our knowledge, it is the first offshore CCGT leak detection study based on a validated high-fidelity dynamic model.

Our contributions can be summarized as follows.

1) We utilized a dynamic model of an existing offshore CCGT system on the Norwegian continental shelf to generate extensive, realistic datasets under transient operating conditions, simulating both normal operational dynamics and leak-induced anomalies.
2) We demonstrate an unsupervised multisensor leak detection framework that uses only fault-free data for (near) real-time detection.
3) We evaluate conventional and state-of-the-art detection methods using realistic operational scenarios, systematically analyzing their detection capabilities across different fault severity levels and data quality conditions.

The remainder of this article is structured as follows. Section II presents the offshore reference power plant and the employed dynamic simulation model. Section III introduces the proposed unsupervised fault detection methodologies. The simulation setup for the dynamic model, the generated datasets, evaluation frameworks, and implementation details are described in Section IV. Results, comprehensive performance analysis, and discussions are presented in Section V. Finally, Section VI summarizes the conclusions and outlines potential future research directions.

II. Reference System and Dynamic Simulation Models of Steam Bottoming Cycle Offshore

The data used in this work were generated using a previously developed high-fidelity dynamic process model that simulates the dynamics of the steam bottoming cycle in an offshore CCGT. The model was presented in the work by Montañés et al. [19]. Here, the model is used to simulate normal operation and operation with leaks in tube bundles within the cycle's once-through steam generator (OTSG) unit. In this section, we present the reference system and components in the combined cycle. The dynamic models are briefly described, and the reader is referred to the cited literature for further description on the model development and its validation [1], [5], [19], [20], [21], and particularly to [19] for the specific model and control system utilized and implemented in this work. In addition, we present an introduction to the relevance of tube failures and leaks and how they are implemented in the dynamic model.

A. Reference CCGT Thermal Power Plant System

The combined cycle on the Oseberg Field Center, located in the North Sea, consists of two GE LM2500+ GTs that each

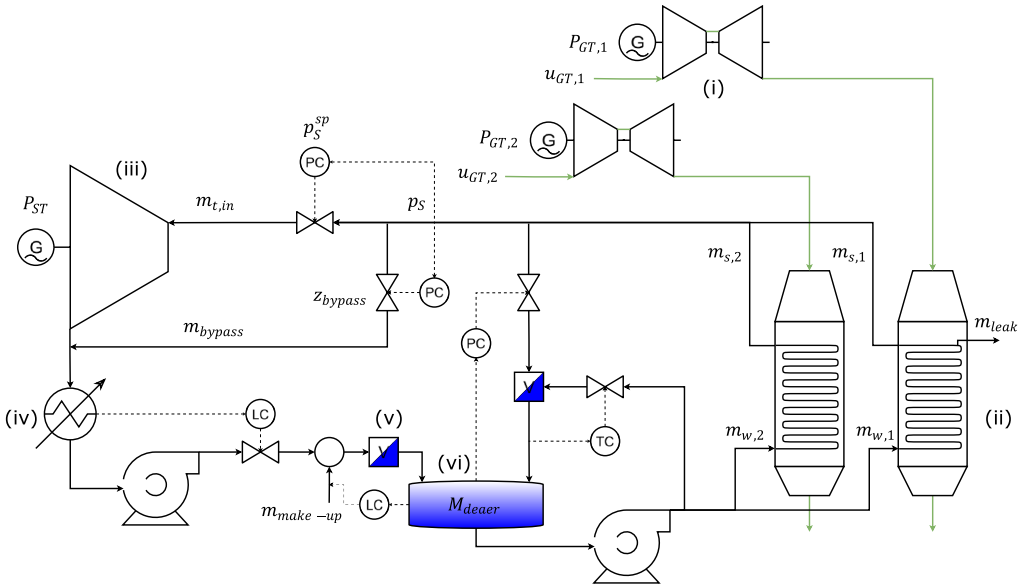


Fig. 1. Schematic of the steam cycle as implemented in the dynamic model, featuring (i) GTs, (ii) OTSGs, (iii) ST, (iv) water condenser, (v) water treatment system, and (vi) deareator. The variables utilized in the dynamic simulation (see Tables I and II) are highlighted in the corresponding component.

drive an export gas compressor. Downstream of each GT is a once-through HRSG. The GTs and OTSGs are located on the Oseberg D platform, whereas the ST, which is connected to an electric generator, is located on the Oseberg A platform [1]. Since the OTSGs and the ST are located on different platforms, there is a long steam supply pipe of about 400 m connecting them. The two OTSGs are designed for a live steam pressure of 16.5 bar(a) with a live steam temperature of 430 °C and a total steam mass flow rate of about 17.5 kg/s. A schematic of the proposed CCGT model is shown in Fig. 1. The steam bottoming cycle produces around 16 MW at nominal conditions [1].

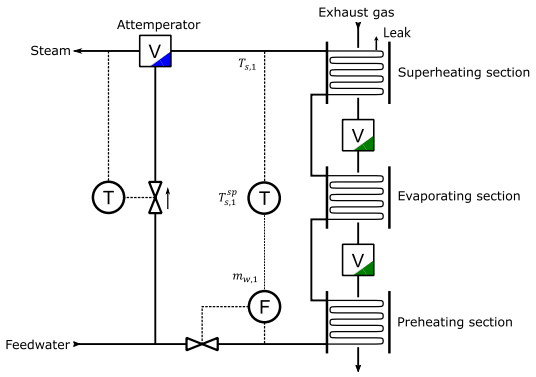


Fig. 2. Schematic of the OTSG dynamic submodel.

B. CCGT Components

Here, we describe the main features of the system components, as well as the main assumptions of the dynamic model used in this work. The reader is referred to [19] for details on the model itself. The dynamic model employs a modular structure that incorporates all main components of the steam bottoming cycle, including the GTs, the ST, the condenser, and the OTSGs. The system dynamic simulation models represent both quasi-static performance and transient responses under varying load conditions, with a modeling approach similar to that demonstrated in recent studies [1], [5]. Particularly, the dynamic model presented in [6] was built on the Dymola simulation environment (2020x) [22] using the Modelica language [23].

1) Gas Turbine: The GT model serves as the primary energy unit and simultaneously provides the exhaust heat that drives the steam generation process. GTs are typically modeled as quasi-static systems due to the relatively fast dynamics of GT operation compared to the steam cycle. That is, the transients of the GTs are modeled as a succession of off-design steady states, each state being in instantaneous equilibrium [1]. Using lookup tables and steady-state performance data from technology suppliers, the GT models produce outputs, such as exhaust gas temperature, mass flow rate, and power output. The computed exhaust parameters are fed as boundary conditions to the OTSG model, as the main source of variations that directly influence the dynamic behavior of the steam cycle [1].

2) Once-Through Steam Generator: The OTSG is one of the most critical components in the steam bottoming cycle, and it is used for converting high-temperature exhaust gases from the GTs into superheated steam. The key elements in an OTSG are the tube bundles, where the geometry is defined by key parameters, such as the number of tube rows, tube diameters, fin dimensions, and the number of passes, as detailed in [5]. The model predicts the OTSG's transient behavior by discretizing the heat exchanger into finite volume elements. This discretization approach captures the main mass, momentum, and energy balances and is crucial for predicting the time delays observed in outlet temperature responses following variations in GT operating conditions and load [5]. A schematic of the reference OTSG with three cores (three tube bundles) is shown in Fig. 2, adapted from [19].

3) Steam Turbine: The ST coupled to a generator is a key component of the bottoming cycle, where superheated steam from the OTSG expands to drive power generation. The model simulates the turbine's performance under varying steam process conditions [1], [19].

4) Condenser and Feedwater Systems: After expansion of superheated steam in the ST, the steam is condensed and recirculated in the feedwater system. The condenser is simulated as a tube-in-shell heat exchanger, employing established correlations for film condensation over tubes to accurately predict the condensation process and associated pressure drops, as reported in [5]. The feedwater system comprises deaerators and pumps, and it is modeled to ensure proper recirculation and preheating of water prior to its entry into the OTSG. Together, these subsystems complete the energy conversion process. Makeup water is added to the cycle via the condensate line upstream of the water treatment system and the deaerator.

5) Cooling System and Boundary Conditions: In offshore installations, seawater is commonly used as a coolant due to its availability and favorable thermophysical properties. The dynamic model incorporates cooling system models that simulate the condenser heat rejection process. The model incorporates boundary conditions, such as cooling seawater temperature, flow pressure, and flow rate. The boundary conditions establish the operational limits and ensure a realistic simulation of external factors impacting the efficiency of the steam cycle.

6) Control Structure: The system dynamic model integrates the different components in the combined cycle and includes a regulatory and supervisory control layer to maintain plant operation under varying load conditions. Such control strategies are critical for maintaining the desired operational setpoints and ensuring that transient disturbances are minimized. Regulatory and supervisory control layers are implemented using PI/PID controllers. These controllers control key process variables, including feedwater flow rate, turbine inlet pressure, and steam temperature. For example, feedforward control strategies based on an input transformation [20] are deployed to anticipate GT disturbances and adjust feedwater flow rate proactively to control the steam temperature. The actual control structures applied in this work, including the main manipulated variables and controlled variables, as well as the controller tuning method, are presented in [19]. The controllers are also illustrated in Figs. 1 and 2.

C. Leak Faults in OTSGs and Simulation Approach

Tube failure is one of the most frequent failures of HRSG and OTSG systems in combined cycles. For example, Tina et al. [24] reported a leaking tube that was discovered within the OTSG unit of a 110-MW cogeneration plant. The causes of tube leaks are plant-specific but typically related to water and steam quality, high variability in temperature and pressure conditions (thermo-mechanical fatigue), overheating (creep or creep-fatigue), and/or corrosion. Tube leaks result in unplanned shutdowns and corresponding operational loss; expensive maintenance and repair actions are normally required [25]. If leaks go undetected for long periods of time, their negative impact may increase [25]. Several works in the literature have reported tube leaks in HRSGs [24], [26], [27]. Normally, the failed segment of the tube is removed from the steam generator, and metallurgical examinations of the segment (or group of tube segments) are performed to identify the failure mode and cause, which differ from case to case [24].

TABLE I
DYNAMIC MODEL INPUT PARAMETERS

Variable	Description	Nominal value
$u_{GT,1}$	Load input for GT 1	80.76 %
$u_{GT,2}$	Load input for GT 2	80.76 %
$T^{sp}_{s,1}$	Outlet temperature set-point of OTSG 1	711.45 K
$T^{sp}_{s,2}$	Outlet temperature set-point of OTSG 2	711.45 K
p^{sp}_{s}	Steam pressure set-point	1.65×10^6 Pa
leak valve	Leak simulation input	0.0

In combined-cycle plants, water makeup assessments are performed at base load to identify unexpected leaks. However, during real-world operation, which includes operation at partial load and frequent load swings, opportunities to perform such precise leak tests are scarce [25]. However, the control strategies applied to the makeup water system often do not facilitate real-time leak detection during normal operation. Operators in the process control room may instead get indications of (severe) leaks by observing changes in pressure drops in water or steam lines and exhaust gas temperature instabilities. Visual inspections can be performed during shutdown or maintenance periods to identify and correct suspected leaks [27]. A real-time, online fault detection system could support leak detection and repair in a more timely and cost-efficient manner.

In this work, a leak simulation element is included in the dynamic simulation model to simulate leaks in the OTSG tube bundles, as shown in Fig. 2. The simulation is driven by including a leak trajectory over time as a disturbance to the model and is implemented by an orifice of variable size. The size (orifice opening) can be increased over time, resulting in an increasing mass flow rate of water/steam through the leak over time.

The dynamic simulation model of the reference power plant in this work integrates the models of GTs, an OTSG system, and the steam cycle (including the condenser and feedwater systems). The main input and output variables of the dynamic model relevant for this work are summarized in Tables I and II. The dynamic model is simulated by specifying time profiles for the variables listed in Table I, and the results from the time integration of the model for the variables listed in Table II are recorded. Refer to Section IV-A for further details on the actual simulations conducted in this work. With the exception of m_{leak}, the output variables have been selected since they can be measured by sensors typically implemented and available during normal operation of combined-cycle power plants. The variable m_{leak} represents the true value of the leak flow rate, and is therefore not used as an input to the leak detection algorithm.

III. UNSUPERVISED LEAK FAULT DETECTION

We present five unsupervised anomaly detection methods that are applicable to leak detection in CCGT systems. The

TABLE II
DYNAMIC MODEL OUTPUT PARAMETERS. WITH THE EXCEPTION OF THE LEAK FLOW RATE, THESE PARAMETERS CAN BE USED AS INPUT TO LEAK DETECTION ALGORITHMS

Variable	Description
$P_{GT,1}$	Shaft power production of GT 1 (MW)
$P_{GT,2}$	Shaft power production of GT 2 (MW)
P_{ST}	Power production of ST (MW)
$T_{s,1}$	Outlet temperature of OTSG 1 (K)
$T_{s,2}$	Outlet temperature of OTSG 2 (K)
$m_{w,1}$	Inlet water flow rate of OTSG 1 (kg/s)
$m_{w,2}$	Inlet water flow rate of OTSG 2 (kg/s)
$m_{s,1}$	Outlet steam flow rate of OTSG 1 (kg/s)
$m_{s,2}$	Outlet steam flow rate of OTSG 2 (kg/s)
p_s	Steam pressure (Pa)
M_{deaer}	Deaerator hold-up (kg)
$m_{\text{make-up}}$	Make-up water flow rate (kg/s)
$m_{t,\text{in}}$	Turbine inlet steam flow rate (kg/s)
z_{bypass}	Turbine bypass valve opening (–)
m_{bypass}	Turbine bypass steam flow rate (kg/s)
m_{leak}	Leak flow rate (kg/s)

dynamic simulation model of the CCGT plant generates synthetic datasets that reflect normal operating conditions as well as faulty conditions due to leaks. We consider unsupervised methods that use only data from normal operation to learn the behavioral characteristics of a CCGT without leaks. Leaks can then be detected as deviations from normal system behavior. This removes the requirement for labeled training data.

To detect leaks during operation, each method computes an anomaly score (s_n) at the generic nth time according to different assumptions. Then, the anomaly score is normalized using min–max normalization

$$\tilde{s}_n = \frac{s_n - s_{\min}}{s_{\max} - s_{\min}} \tag{1}$$

where $s_{\min}$ and $s_{\max}$ are the minimum and maximum scores as computed from the training data and compared with a predetermined threshold (τ)

$$\hat{y}_n = \begin{cases} 1, & \tilde{s}_n > \tau \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Here, $\hat{y}_n = 1$ (resp. $\hat{y}_n = 0$) denotes the system indicating the presence (resp. absence) of a leak. Typically, the threshold (τ) is set based on statistical considerations. The five methods we consider for computing anomaly scores are explained below.

A. PCA-Based Anomaly Detection

PCA is a widely used dimensionality reduction technique that projects high-dimensional data into a lower dimensional subspace while preserving maximal variance. PCA employs the inherent linear correlations among measurements in the CCGT system under normal operating conditions, and leaks manifest as deviations from the normal operational subspace.

Let $\boldsymbol{X} \in \mathbb{R}^{K\times N}$ be the multivariate sensor data under normal operation to be used for training, where K and N are the number of sensors and the number of observations, respectively. The mean vector $\bar{\boldsymbol{x}} = 1/N \sum_{n=1}^{N} \boldsymbol{x}[n]$, where $\boldsymbol{x}[n] = (\boldsymbol{X}_{1,n}, \ldots, \boldsymbol{X}_{K,n})^{\mathrm{T}}$ collects the measurements at time n, is removed from the original data as $\tilde{\boldsymbol{X}} = \boldsymbol{X} - \bar{\boldsymbol{x}}\mathbf{1}_N^{\mathrm{T}}$, where $\mathbf{1}_N$ is an N-dimensional vector of ones, and the covariance matrix is computed as $\boldsymbol{C} = 1/(N-1)\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^{\mathrm{T}}$.

The principal components are identified via eigen-decomposition of the covariance matrix

$$\boldsymbol{C}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \quad i = 1, \ldots, K \tag{3}$$

where λ_i and $\boldsymbol{u}_i$ represent the eigenvalues and eigenvectors of $\boldsymbol{C}$, respectively. Eigenvectors corresponding to the largest eigenvalues identify the most relevant directions; thus, the q most relevant eigenvectors are used to build the projection matrix $\boldsymbol{U}_q = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_q] \in \mathbb{R}^{K\times q}$, and the low-dimensional representation of $\boldsymbol{x}[n]$ is obtained as $\hat{\boldsymbol{x}}[n] = \boldsymbol{U}_q^{\mathrm{T}}(\boldsymbol{x}[n] - \bar{\boldsymbol{x}})$.

The reconstruction error is used as an anomaly score, $s_n = |\boldsymbol{x}[n] - \tilde{\boldsymbol{x}}[n]|_2^2$, where $\tilde{\boldsymbol{x}}[n] = \boldsymbol{U}_q\hat{\boldsymbol{x}}[n] + \bar{\boldsymbol{x}}$ is the reconstructed vector, as it quantifies how much a test vector ($\boldsymbol{x}[n]$) aligns with the learned normal subspace.

B. OC-SVM-Based Anomaly Detection

A one-class support vector machine (OC-SVM) identifies data points deviating significantly from a learned boundary around normal operating data. OC-SVM constructs a hyperplane in a high-dimensional feature space, effectively encapsulating normal operating conditions and isolating anomalies. OC-SVM maps input data ($\boldsymbol{x}[n]$) into a high-dimensional feature space using a nonlinear kernel function $\phi(\cdot)$, and aims to find the smallest hypersphere enclosing the data in feature space by solving

$$\min_{\boldsymbol{w},\xi_n,\rho} \frac{1}{2}\|\boldsymbol{w}\|^2 + \frac{1}{\nu N}\sum_{n=1}^{N}\xi_n - \rho, \ \text{s.t.} \begin{cases} \boldsymbol{w}^{\mathrm{T}}\phi\left(\boldsymbol{x}\left[n\right]\right) \geq \rho - \xi_n \\ \xi_n \geq 0 \end{cases} \tag{4}$$

where N is the number of samples available for training, $\boldsymbol{w}$ is the normal vector identifying the hyperplane, ρ is an offset defining the hyperplane, ξ_n's are the slack variables allowing some points to lie outside the boundary, and $\nu \in (0, 1]$ controls the trade-off between maximizing the margin and the fraction of outliers. For inference on the generic test vector ($\boldsymbol{x}$), OC-SVM relies on the following decision function:

$$\begin{aligned} f\left(\boldsymbol{x}\right) &= \text{sign}\left(\boldsymbol{w}^{\mathrm{T}}\phi\left(\boldsymbol{x}\right) - \rho\right) \\ &= \text{sign}\left(\sum_{n=1}^{N}\alpha_n k\left(\boldsymbol{x}\left[n\right], \boldsymbol{x}\right) - \rho\right) \end{aligned} \tag{5}$$

where α_n's are the Lagrange multipliers. The anomaly score is computed as the distance to the learned hyperplane

$$s_n = \rho - \sum_{m=1}^{N}\alpha_m k\left(\boldsymbol{x}\left[m\right], \boldsymbol{x}\left[n\right]\right). \tag{6}$$

In OC-SVM, to mitigate overfitting, a kernel and ν can be selected through cross-validation using temporally disjoint

subsets of the normal operating data. In addition, a regularization can be applied implicitly through the choice of ν, which constrains the fraction of support vectors and prevents the model from memorizing noise.

C. LOF-Based Anomaly Detection

Local outlier factor (LOF) is a density-based anomaly detection technique that identifies anomalous data points by assessing their local neighborhood density relative to their neighbors.

For each data point $\boldsymbol{x}[n]$, we define the k-distance $d_k(\boldsymbol{x}[n])$ as the Euclidean distance to its kth nearest neighbor, and the corresponding neighborhood includes all the points within this k-distance

$$N_k(\boldsymbol{x}[n]) = \{\boldsymbol{x}[m] \mid \|\boldsymbol{x}[n] - \boldsymbol{x}[m]\|_2 \le d_k(\boldsymbol{x}[n]),\ m \neq n\}. \quad (7)$$

Also, the reachability distance ($r_k(\cdot)$) from point $\boldsymbol{x}[n]$ to point $\boldsymbol{x}[m]$ is given as

$$r_k(\boldsymbol{x}[n], \boldsymbol{x}[m]) = \max\{d_k(\boldsymbol{x}[m]), \|\boldsymbol{x}[n] - \boldsymbol{x}[m]\|_2\} \quad (8)$$

and the local reachability density ($\ell_k(\cdot)$) at point $\boldsymbol{x}[n]$ is computed as the reciprocal of the average reachability distance to its k neighbors, i.e.,

$$\ell_k(\boldsymbol{x}[n]) = \left(\frac{\sum_{\boldsymbol{x}[m]\in N_k(\boldsymbol{x}[n])} r_k(\boldsymbol{x}[n], \boldsymbol{x}[m])}{|N_k(\boldsymbol{x}[n])|}\right)^{-1}. \quad (9)$$

The LOF of each point is considered as an anomaly score, and it is computed by comparing its local reachability density to the neighbors' average, i.e.,

$$s_n = \frac{\sum_{\boldsymbol{x}[m]\in N_k(\boldsymbol{x}[n])} \ell_k(\boldsymbol{x}[m])}{|N_k(\boldsymbol{x}[n])| \cdot \ell_k(\boldsymbol{x}[n])}. \quad (10)$$

A large LOF indicates lower local density compared to their neighbors, thus flagging a potential anomaly.

D. LRS-Based Anomaly Detection

Anomaly detection via low-rank and sparse (LRS) decomposition exploits the observation that, under normal conditions, the multivariate sensor data lie approximately in a low-dimensional subspace due to strong correlations among variables [28], [29]. LRS matrix decomposition separates the normal behavior from anomalies without requiring any labeled fault data. Let $\boldsymbol{X} \in \mathbb{R}^{K\times N}$ be the sensor data collected during normal operation, where K and N are the number of sensors and the number of observations, respectively. The core assumption of LRS decomposition is

$$\boldsymbol{X} = \boldsymbol{L} + \boldsymbol{S} \quad (11)$$

where $\boldsymbol{L}$ is a low-rank matrix representing the nominal (noise-free) behavior of the system, and $\boldsymbol{S}$ is a sparse matrix capturing faults, noise, and transients. The decomposition problem can be formulated as a minimization of rank and sparsity and solved via the augmented Lagrange multiplier method, i.e.,

$$\mathcal{L}(\boldsymbol{L}, \boldsymbol{S}, \boldsymbol{Y}, \rho) = \|\boldsymbol{L}\|_* + \lambda\|\boldsymbol{S}\|_1 + \langle \boldsymbol{Y}, \boldsymbol{X} - \boldsymbol{L} - \boldsymbol{S}\rangle + \frac{\rho}{2}\|\boldsymbol{X} - \boldsymbol{L} - \boldsymbol{S}\|_F^2 \quad (12)$$

where $\boldsymbol{Y}$ is the Lagrange multiplier matrix, $\rho > 0$ is a penalty parameter, and $\|\cdot\|_*$, $\|\cdot\|_1$, and $\|\cdot\|_F$ denote the nuclear norm, the ℓ_1-norm, and the Frobenius norm, respectively.

Anomaly detection relies on performing the SVD on the low-rank matrix $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{U} \in \mathbb{R}^{K\times K}$ contains the principal components. The generic test data vector ($\boldsymbol{x}[n]$)) is projected onto the low-rank subspace $\hat{\boldsymbol{x}}[n] = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{x}[n]$, and the projection distance is considered as an anomaly score as it quantifies the deviation of the data vector

$$s_n = \sum_{i=1}^{q} \frac{\hat{x}_i[n]^2}{\sigma_i^2} \quad (13)$$

where σ_i are the singular values from $\boldsymbol{\Sigma}$, and q is the number of principal components chosen for the analysis.

E. Anomaly Detection With TAE

This method employs a transformer-based autoencoder architecture tailored to capture complex spatio-temporal dependencies in high-dimensional sensor data. By learning a compact representation of normal operation sequences, the model detects anomalies via reconstruction error, under the assumption that abnormal patterns cannot be effectively reconstructed from the learned latent space [30], [31]. Let $\boldsymbol{X} \in \mathbb{R}^{K\times N}$ be the sensor data collected during normal operation, where K and N are the number of sensors and the number of observations, respectively. Data is segmented into overlapping windows of length L: $\boldsymbol{X}_n \in \mathbb{R}^{K\times L}$ and each window $\boldsymbol{X}_n$ is treated as a sequence input to the transformer encoder. The transformer encoder, denoted by $\mathcal{T}(\cdot)$, extracts latent representations by processing the input sequence enriched with positional encoding $\boldsymbol{P}$. The resulting embedding is $\boldsymbol{Z}_n = \mathcal{T}(\boldsymbol{X}_n + \boldsymbol{P})$. A convolutional decoder $\mathcal{C}(\cdot)$ composed of 1-D convolutional layers reconstructs the original input from the latent representation: $\hat{\boldsymbol{X}}_n = \mathcal{C}(\boldsymbol{Z}_n)$.

The model is trained to minimize the mean squared reconstruction error between the input sequence $\boldsymbol{X}n$ and its reconstruction $\hat{\boldsymbol{X}}n$

$$\mathcal{L}_{\mathrm{rec}} = \frac{1}{K\cdot L}\sum_{k=1}^{K}\sum_{\ell=0}^{L-1}\left(\hat{x}_k[n-\ell] - \hat{x}_k[n-\ell]\right)^2. \quad (14)$$

This loss encourages the network to learn the manifold of normal system behavior, thus enabling it to exhibit high reconstruction error on anomalous inputs.

At inference time, each test sequence $\boldsymbol{X}_n$ is passed through the transformer encoder and convolutional decoder to produce the reconstruction $\hat{\boldsymbol{X}}_n$. The anomaly score at a time n is defined as the ℓ_2 norm of the reconstruction error at the last step of the window

$$s_n = \left|\boldsymbol{x}[n] - \hat{\boldsymbol{x}}[n]\right|_2. \quad (15)$$

A higher anomaly score indicates deviations from normal operation patterns, allowing for effective unsupervised detection of various types of sensor anomalies.

IV. Experimental Setup

A. Dynamic Simulation Tools and Leak Simulation

The dynamic simulation model was built using Modelica, a nonproprietary, object-oriented, equation-based language for modeling complex physical systems [23]. The simulation is carried out in Dymola, which provides the advanced capabilities for simulating and analyzing Modelica models [22]. The model incorporates several standard and specialized libraries that including as follows.

1) *Modelon Thermal Power Library:* Providing a comprehensive set of components for modeling thermal processes, including GTs, two-phase flow, and condenser systems [32].
2) *Standard Modelica Blocks:* Providing standard blocks for control, signal processing, and system interconnections.

The dynamic model was presented in the work by Montañés et al. [19], and is also described in Section II in this article. The model is used to generate a synthetic dataset that includes both normal operation and operation with realistic development of leaks in the OTSG tubing. The synthetic dataset is used to evaluate the unsupervised leak detection techniques described in Section III.

Five time series corresponding to normal operation and leak development in the OTSG steam header were generated by simulating the dynamic model. In these time series, the steam cycle and OTSGs are subject to operating conditions that correspond to the normal variability of the connected GTs, with load oscillations around the nominal point and step changes corresponding to different operating nominal loads. The normal variability is based on historical data for mechanical drive GTs during one year of operation that was previously analyzed for a reference offshore platform in [5]. Each of the five time series generated in this work has a duration of one week, with a sampling resolution of 1 h. Three of the time series correspond to normal operation (without OTSG leaks), and the remaining two time series correspond to the gradual development of a leak in one of the OTSGs, as shown in Fig. 3. This is simulated by increasing the size of the leak over time (orifice opening), resulting in an increasing mass flow rate of water/steam through the leak over time. The main input and output variables of the dynamic model relevant for this work are summarized in Tables I and II.

B. Evaluation Metrics and Leak-Analysis Approach

For performance evaluations, we rely on binary classification metrics commonly used in anomaly detection. Let TP be the number of correctly detected leak events (true positives), FP be the number of false alarms (normal samples misclassified as leaks), TN be the correctly classified normal samples (true negatives), and FN be the missed leak events (false negatives). We define the true positive rate (TPR) and the false positive rate (FPR) as

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{16}$$

In real-world operational scenarios, the false alarms may lead to unnecessary maintenance and downtime, while missed detections can result in costly damage. Thus, the trade-off between TPR and FPR is crucial. To visualize this trade-off, we utilize the receiver operating characteristic (ROC) curve, which plots TPR against FPR across varying decision thresholds. We report the area under the ROC curve (AUC), which is a threshold-independent measure of separability between the leak and no-leak distributions. Also, we evaluate models under a fixed FPR constraint, focusing on the achieved TPR as the key indicator of detection performance across scenarios with different leak sizes and different measurement noise levels. However, we have utilized a geometric mean square (G-Mean = $((\text{TPR} \cdot (1 - \text{FPR})))^{1/2}$) to compare detection performance.

To evaluate the sensitivity and robustness of anomaly detection methods to varying levels of fault severity, we propose a structured testing framework. The data is partitioned into four segments, defined by a leak-size threshold (γ). The threshold γ can be interpreted as a normalized leak flow rate, such that $\gamma = 0$ corresponds to no leak and $\gamma = 1$ corresponds to the maximum leak flow rate (=0.335) observed in the dataset. The use of γ allows us to formulate binary classification tasks reflecting different fault detection scenarios, as illustrated in Fig. 4. For a given leak-size threshold (γ), we define four regions as follows.

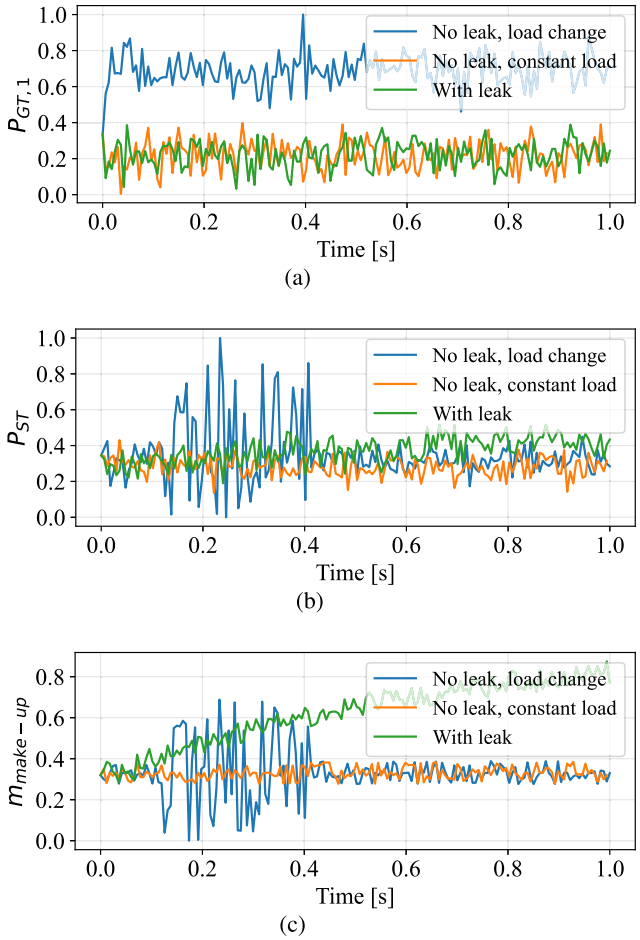


Fig. 3. Sample of normalized datasets. (a) GT 1 power production. (b) ST power production. (c) Makeup water flow rate.

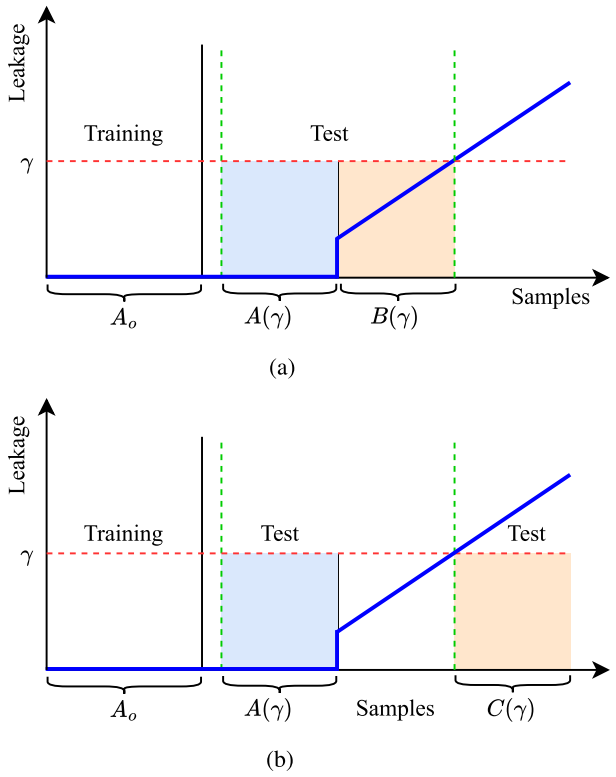


Fig. 4. Structured framework for leak analysis. (a) Low-severity leak. (b) High-severity leak.

1) A_o *(Normal):* Samples for which the leak size is exactly zero. These represent healthy, baseline system behavior and are used exclusively for *training*.
2) $A(\gamma)$ *(Normal):* Samples for which the leak size is exactly zero and are used for normal class *testing*.
3) $B(\gamma)$ *(Weak Faults):* Samples with leak sizes no larger than γ. These represent incipient or weakly observable anomalies and form the test set for early fault detection evaluation.
4) $C(\gamma)$ *(Strong Faults):* Samples with leak sizes larger than γ. These correspond to more pronounced and easier-to-detect anomalies, used for validating late-stage or critical fault detection.

Models are trained on data from A_o and tested on data from $A(\gamma)$, $B(\gamma)$, and $C(\gamma)$. During testing, each experiment proceeds by fixing a leak-size threshold (γ) and identifying the corresponding regions. More specifically, we define two binary classification tasks, where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the absence and presence of the leak, respectively, as follows.

1) *Case 1):* $A(\gamma) \equiv \mathcal{H}_0$, $B(\gamma) \equiv \mathcal{H}_1$ focuses on the capability of the models to detect *small leaks* from normal conditions (i.e., relevant for early warning).
2) *Case 2):* $A(\gamma) \equiv \mathcal{H}_0$, $C(\gamma) \equiv \mathcal{H}_1$ focuses on the capability of the models to detect *big leaks* from normal conditions (i.e., relevant for critical conditions).

C. Implementation Details and Tools

We conducted an evaluation of the proposed methods using baseline implementation in the PyOD Python library [33]. For training a transformer autoencoder (TAE) method, we utilized the PyTorch deep learning framework. For LRS, we optimize with maximum iterations of 100 and penalty parameter $\rho = 10^{-5}$. Following decomposition, PCA was applied on the low-rank component with 13 retained principal components. For TAE, we used input windows of five time steps with 14 features, model dimension $d_{\text{model}} = 14$, two attention heads, one encoder layer, feedforward and latent dimensions = 8, dropout = 0, convolutional decoder kernel size = 3, batch size = 4, and Adam optimizer (learning rate 10^{-3}) with mse loss, trained for 1200–1500 epochs. For the baseline methods, we used PCA with 13 components; OC-SVM with both polynomial and RBF kernels (degrees 2 and 5); LOF with contamination level 0.01 and $n_{\text{neighbors}} = 30$. In addition, we have documented experimental settings for evaluation, including leak thresholds ([0.03–0.1] and [0.1–0.3]), FPR levels ([0.05, 0.1] and [0.01–0.2]), and signal-to-noise ratio (SNR) levels (0–40).

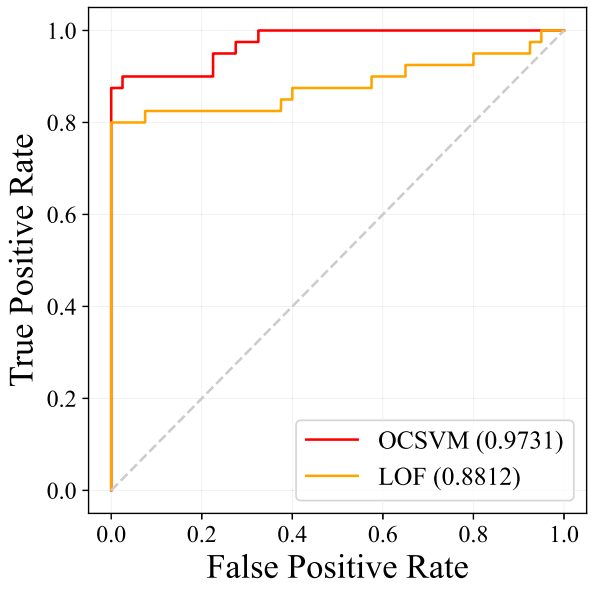


Fig. 5. ROC curves and corresponding AUC values when processing the condenser makeup water flow parameter only (small-size leak, $\gamma = 0.334$).

V. Results and Discussion

We begin our analysis by comparing the performance of two conventional methods (OC-SVM and LOF) for leak detection based on condenser makeup water flow rate data only. As shown in Fig. 3, the makeup water flow rate m_{makeup} is highly correlated to the presence of leaks. However, as discussed in Section II-C and shown in Fig. 3, m_{makeup} is also susceptible to large variations during operation with load changes. Due to this and control strategy-related limitations, m_{makeup} is not considered for real-time leak detection in current CCGT installations. The purpose of this initial analysis is thus to investigate the potential utility of m_{makeup}, should future cycles be designed with leak detection based on m_{makeup} in mind. Fig. 5 shows the performance of OC-SVM and LOF in terms of ROC curves and associated AUC values

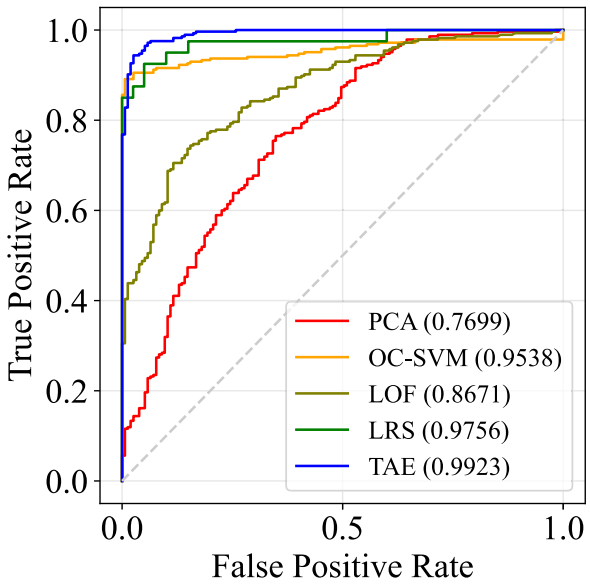


Fig. 6. ROC curves and corresponding AUC values when processing the complete feature set (small-size leak, $\gamma = 0.334$).

TABLE III
AVERAGE ($\pm$ STANDARD DEVIATION) TRAINING AND INFERENCE TIMES (SECONDS) OVER FIVE RUNS

Model	Training time (s)	Inference time (s)
PCA	0.0023 (0.0014)	0.0002 (0.0001)
OCSVM	0.0028 (0.0002)	0.0003 (0.0000)
LOF	0.0039 (0.0014)	0.0007 (0.0001)
LRS	0.0062 (0.0006)	0.0054 (0.0013)
TAE	34.9578 (5.5820)	0.0033 (0.0025)

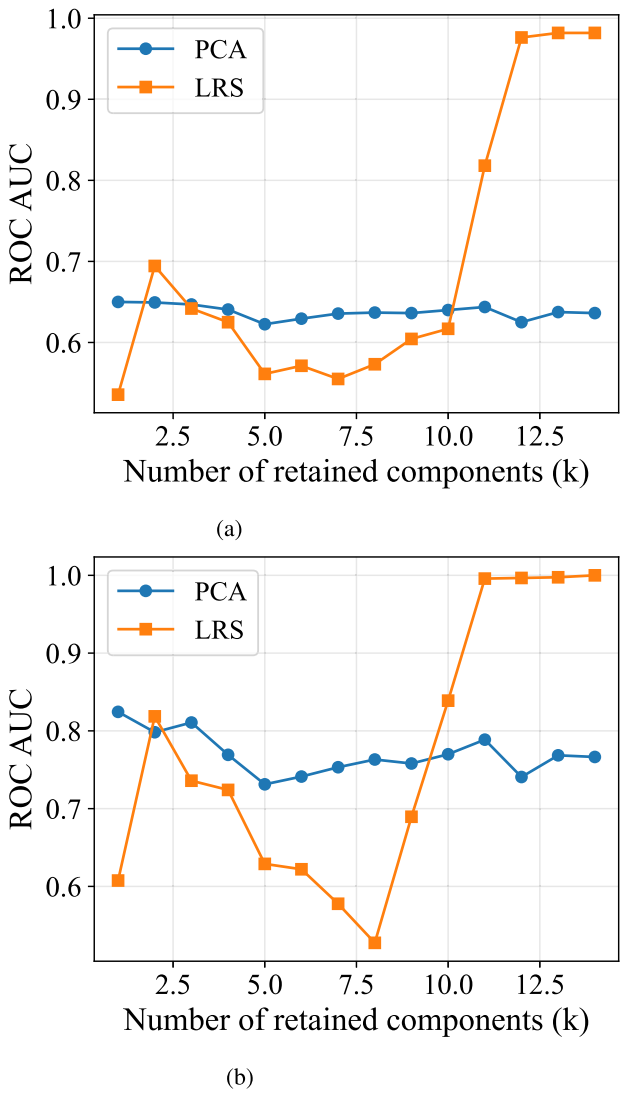


Fig. 7. ROC AUC versus retained components. (a) For small leak. (b) For big leak.

for Case 1—small-leak detection. Relying on m_{makeup} as the single input feature, the figure shows that the methods' performance is adequate to provide practical utility. The corresponding results for Case 2—big leak detection—are omitted because both methods performed almost flawlessly in this case.

If improved performance is desired in the small-leak case, a reasonable approach is to include the full set of available sensor data (see Table II) as data to the detection methods. Fig. 6 shows the performance of all five methods in the small-leak case when using the full set of available features. In this scenario, LRS and TAE achieve the best performance with AUC scores of 0.9756 and 0.9923, respectively. OC-SVM (AUC = 0.9538) and LOF (AUC = 0.8671) also perform reasonably well, but interestingly performed better in the case when m_{makeup} was the only input feature. PCA is the worst performer with an AUC of 0.7699. As in the previous scenario, all methods performed very well in the big leak case, so results from that case are omitted for brevity. In general, the results indicate that leak detection based on m_{makeup} alone can enable satisfactory leak detection if the water makeup control and monitoring system is set up appropriately.

Table III summarizes the average training and inference times for all methods, measured over five independent runs. Classical approaches, such as PCA, OC-SVM, LOF, and LRS, exhibit subsecond training and testing times, confirming their suitability for real-time monitoring scenarios. In contrast, the TAE requires substantially longer training time, reflecting the higher computational cost of deep sequence models, although its inference time remains comparable to that of LRS. These results highlight the trade-off between model complexity and computational efficiency: while advanced neural architectures may yield higher detection accuracy, lightweight linear and proximity-based methods offer clear advantages in deployment settings with strict real-time or resource-constrained requirements.

The G-mean balances sensitivity (TPR) with specificity, thereby improving detectors that achieve high TPRs without excessive false alarms. Table IV confirms the relative findings from the ROC analysis. PCA and LOF achieve only modest G-Mean values <0.5), indicating that their increased sensitivity under leak scenarios comes at the expense of frequent false positives. By contrast, LRS consistently achieves the highest G-Mean across both thresholds (≈ 0.9),

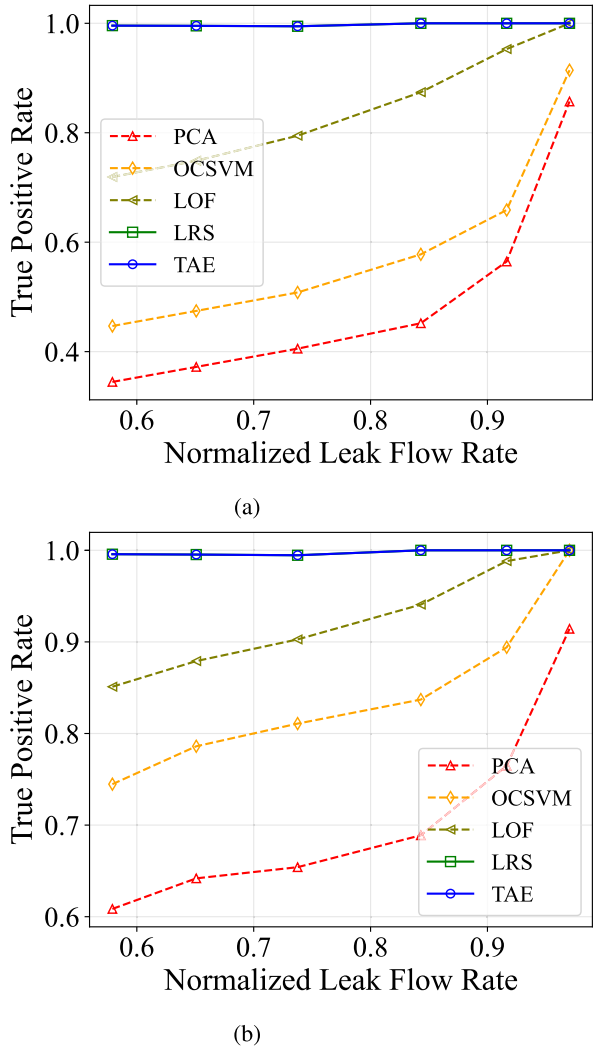


Fig. 8. TPR versus normalized leak flow rate threshold γ for fixed FPR (big-size leak). (a) FPR = 0.1. (b) FPR = 0.2.

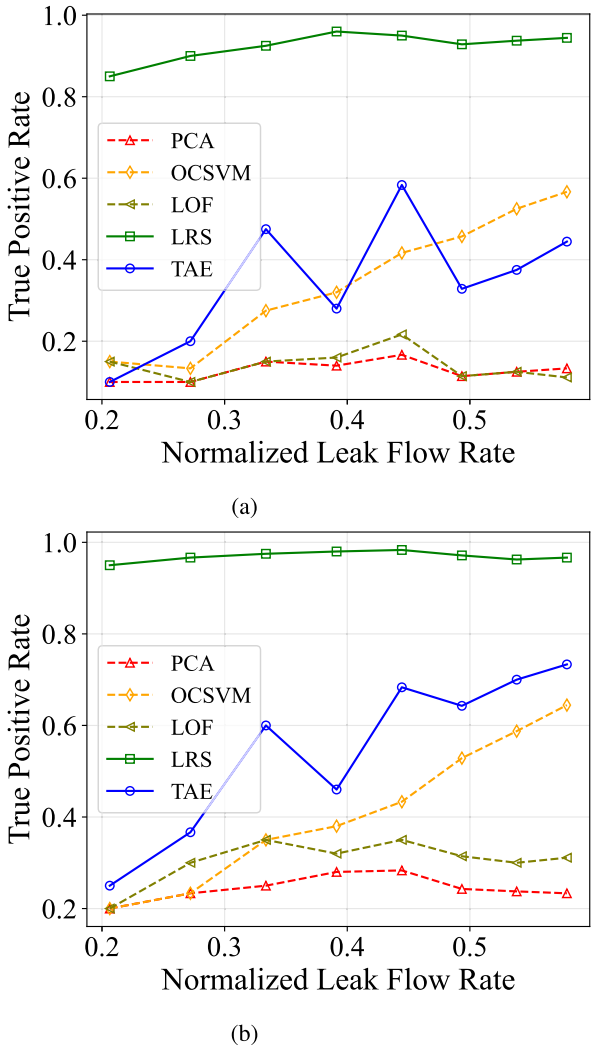


Fig. 9. TPR versus normalized leak flow rate threshold γ for fixed FPR (small-size leak). (a) FPR = 0.1. (b) FPR = 0.2.

TABLE IV
G-MEAN SCORE ACROSS LEAKAGE THRESHOLDS

Model	G-Mean@FPR=0.1	G-Mean@FPR=0.2
PCA	0.3390 (0.0292)	0.4422 (0.0228)
OCSVM	0.5504 (0.1303)	0.5694 (0.1072)
LOF	0.3535 (0.0425)	0.4930 (0.0380)
LRS	0.9120 (0.0164)	0.8806 (0.0045)
TAE	0.5366 (0.4215)	0.5060 (0.3974)

underscoring its robustness in maintaining strong leak detection while keeping false positives low. Interestingly, although TAE reaches very high AUC values (see Fig. 6), its G-Mean is considerably lower and more variable, which suggests a tendency toward overfitting, i.e., better separation in training but less stable performance under different leak thresholds. This highlights that LRS offers the most reliable trade-off between detection power and operational usability, making it particularly suitable for early-warning leak detection in practice.

We also analyze the effect of the retained component in PCA and LRS methods, as shown in Fig. 7. In both cases, LRS consistently outperforms standard PCA, achieving higher AUC values across nearly all component settings. This suggests that the LRS decomposition in LRS preserves more discriminative structure relevant to leak detection than PCA. The comparisons A versus B and A versus C further highlight that the separation between leak and no-leak is more challenging in the latter, reflected in overall lower AUC values, yet LRS still provides a clear advantage in capturing the underlying leak signatures.

In the remainder of this section, we focus on leak detection without including the condenser makeup water flow among the available features. This setup is in line with current CCGT installations and is more generally applicable. Fig. 8 presents a comprehensive comparison of all the methods with a big-leak perspective when operating at a fixed FPR. The consistent superiority of LRS and TAE is apparent, while other, more traditional methods are not capable of reliably detecting big leaks, with PCA being the one affected more severely. Fig. 9 presents an analogous comparison with a weak-leak

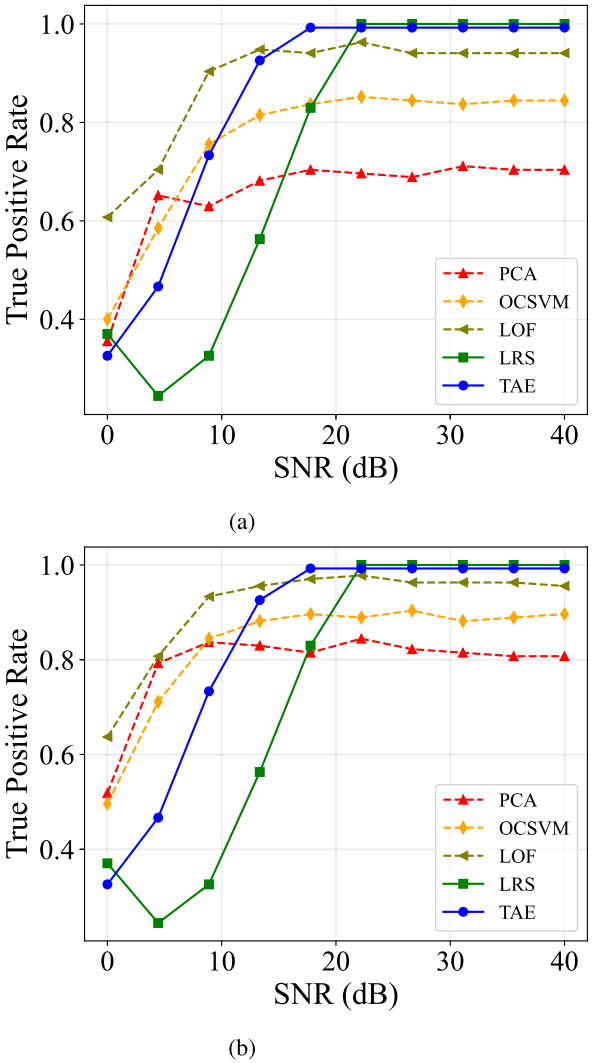


Fig. 10. TPR versus SNR for fixed FPR (big-size leak, $\gamma = 0.843$). (a) FPR = 0.2. (b) FPR = 0.3.

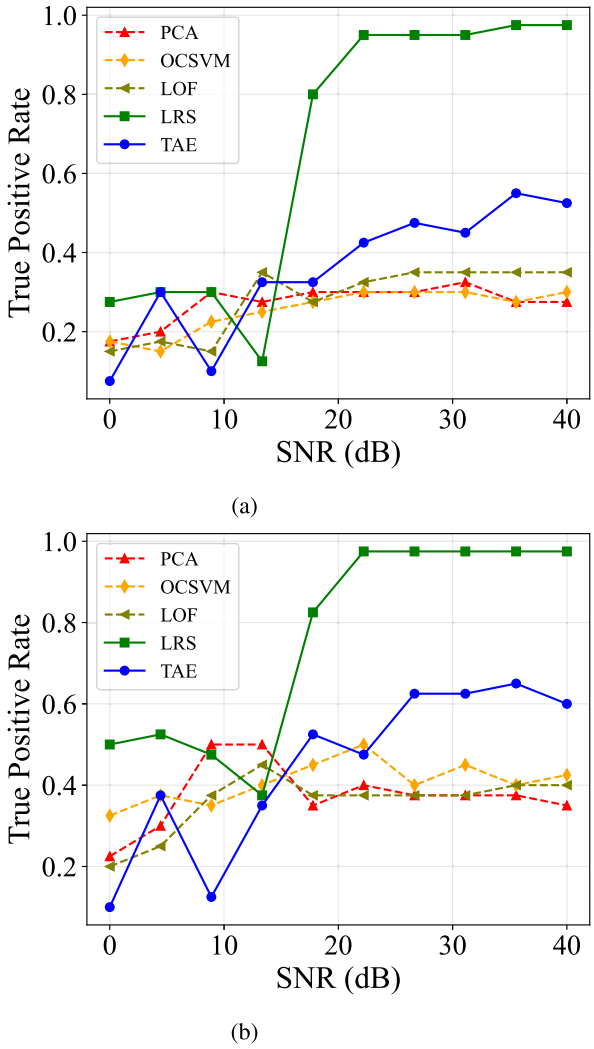


Fig. 11. TPR versus SNR for fixed FPR (small-size leak, $\gamma = 0.334$). (a) FPR = 0.2. (b) FPR = 0.3.

perspective. It is apparent and very informative to acknowledge that the LRS keeps achieving excellent performance, while all other methods are completely unreliable.

Finally, we consider the impact on the performance of data quality conditions, again when focusing on methods operating without including the condenser makeup water flow among the available features. More specifically, we corrupted the time series from the dynamic simulation model with additive white Gaussian noise to mimic realistic operating conditions. We superimpose zero-mean noise independently on each feature and assume the same SNR on each one. Fig. 10 presents a comprehensive comparison of all the methods and their robustness to noisy conditions with a big-leak perspective when operating at a fixed FPR. LRS and TAE achieve the best asymptotic performance for large SNR, while suffering a bit more at low SNR. Fig. 11 presents the analogous performance with a small-leak perspective, where LRS is confirmed to be the only reliable method at high SNR, while all methods are unreliable at low SNR.

VI. Conclusion and Future Works

This work presented a quantitative comparison of unsupervised algorithms for leak detection in CCGTs using synthetic data from a state-of-the-art high-fidelity dynamic model. If available under appropriate conditions, the condenser makeup water flow rate is, by itself, sufficient for adequate leak detection, even using simple methods. However, in the more common case that appropriate makeup water data is unavailable, we demonstrate that state-of-the-art methods nonetheless show excellent detection capabilities. In addition, detection sensitivity to leak size and robustness to noise were evaluated. Detection based on LRS decomposition was the only reliable algorithm in the most challenging case (small leaks and noisy measurements, while not relying on makeup water data). This study was performed exclusively on high-fidelity simulated data due to the unavailability of industrial leak-labeled datasets. This study was performed exclusively on high-fidelity simulated data due to the unavailability of industrial leak-labeled datasets. Should such data become available in the future, validation of the proposed detection framework using

that data would be a valuable continuation of the present work. This would help assess the practical performance of the algorithms under real-world conditions, including sensor faults, unmodeled dynamics, and operational interventions.

Acknowledgment

The authors acknowledge Magnus K. Windfeldt and Leif E. Andersson from SINTEF Energy Research for the initial discussions on how to implement faults into the dynamic model.

References

[1] L. O. Nord and R. M. Montañés, "Compact steam bottoming cycles: Model validation with plant data and evaluation of control strategies for fast load changes," *Appl. Thermal Eng.*, vol. 142, pp. 334–345, Sep. 2018.

[2] F. Alobaid, R. Postler, J. Ströhle, B. Epple, and H.-G. Kim, "Modeling and investigation start-up procedures of a combined cycle power plant," *Appl. Energy*, vol. 85, no. 12, pp. 1173–1189, Dec. 2008.

[3] S. E. Zitney, "Dynamic model-based digital twin, optimization, and control technologies for improving flexible power plant operations," in *Proc. Connected Plant Conf.*, 2019.

[4] E. Ersayin and L. Ozgener, "Performance analysis of combined cycle power plants: A case study," *Renew. Sustain. Energy Rev.*, vol. 43, pp. 832–842, Mar. 2015.

[5] R. M. Montañés, G. Skaugen, B. Hagen, and D. Rohde, "Compact steam bottoming cycles: Minimum weight design optimization and transient response of once-through steam generators," *Frontiers Energy Res.*, vol. 9, Jun. 2021, Art. no. 687248.

[6] R. M. Montañés et al., "Design optimization of compact gas turbine and steam combined cycles for combined heat and power production in a FPSO system—A case study," *Energy*, vol. 282, Nov. 2023, Art. no. 128401. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544223017954

[7] S. Sampath, A. Gulati, and R. Singh, "Fault diagnostics (don't short) using genetic algorithm for advanced cycle gas turbine," *Amer. Soc. Mech. Eng.*, vol. 2, pp. 19–27, Apr. 2002, doi: 10.1115/gt2002-30021.

[8] A. Ajami and M. Daneshvar, "Data driven approach for fault detection and diagnosis of turbine in thermal power plant using independent component analysis (ICA)," *Int. J. Electr. Power Energy Syst.*, vol. 43, no. 1, pp. 728–735, Dec. 2012.

[9] U. Sarwar, M. Muhammad, A. A. Mokhtar, R. Khan, P. Behrani, and S. Kaka, "Hybrid intelligence for enhanced fault detection and diagnosis for industrial gas turbine engine," *Results Eng.*, vol. 21, Mar. 2024, Art. no. 101841.

[10] M. Fast and T. Palmé, "Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant," *Energy*, vol. 35, no. 2, pp. 1114–1120, Feb. 2010.

[11] A. Arranz, A. Cruz, M. Sanzbobi, P. Ruiz, and J. Coutino, "DADICC: Intelligent system for anomaly detection in a combined cycle gas turbine plant," *Expert Syst. Appl.*, vol. 34, no. 4, pp. 2267–2277, May 2008.

[12] S. M. Camporeale, L. Dambrosio, A. Milella, M. Mastrovito, and B. Fortunato, "Fault diagnosis of combined cycle gas turbine components using feed forward neural networks," *Amer. Soc. Mech. Eng. Int. Gas Turbine Inst., Turbo Expo.*, vol. 1, pp. 549–561, May 2003, doi: 10.1115/gt2003-38742.

[13] H. E. Davallo, R. Bahrevar, and A. Chaibakhsh, "Fault diagnosis of combined cycle power plant using ELM," in *Proc. 7th Int. Conf. Robot. Mechatronics (ICRoM)*, Nov. 2019, pp. 40–45.

[14] M. A. Belay, S. S. Blakseth, A. Rasheed, and P. S. Rossi, "Unsupervised anomaly detection for IoT-based multivariate time series: Existing solutions, performance analysis and future directions," *Sensors*, vol. 23, no. 5, p. 2844, Mar. 2023.

[15] M. A. Belay, A. Rasheed, and P. S. Rossi, "Digital twin-based federated transfer learning for anomaly detection in industrial IoT," in *Proc. IEEE Symp. Comput. Intell. Eng./Cyber Phys. Syst. (CIES)*, Mar. 2025, pp. 1–7.

[16] M. A. Belay, A. Rasheed, and P. S. Rossi, "Digital twin knowledge distillation for federated semi-supervised industrial IoT DDoS detection," in *Proc. IEEE Symp. Comput. Intell. Secur., Defence Biometrics Companion (CISDB Companion)*, Mar. 2025, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/11010678/

[17] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.

[18] M. A. Belay, A. Rasheed, and P. S. Rossi, "Self-supervised modular architecture for multi-sensor anomaly detection and localization," in *Proc. IEEE Conf. Artif. Intell. (CAI)*, Jun. 2024, pp. 1278–1283.

[19] R. M. Montañés, M. Windfeldt, L. E. Andersson, and G. Skaugen, "A framework for physics based off-design and dynamic modelling and simulation of combined cycle power plants in weight and volume constraint environments," in *Proc. Heat Powered Cycles Conf.*, 2023, pp. 151–161. [Online]. Available: https://zenodo.org/records/10245219

[20] C. ZoticĂ, R. M. Montañés, A. Reyes-Lúa, and S. Skogestad, "Control of steam bottoming cycles using nonlinear input and output transformations for feedforward disturbance rejection," *IFAC-PapersOnLine*, vol. 55, no. 7, pp. 969–974, 2022.

[21] R. M. Montañés, C. Zotică, and A. Reyes-Lúa, "Operation and control of compact offshore combined cycles for power generation," *Energy*, vol. 290, Mar. 2024, Art. no. 130315. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544224000860

[22] *Dymola—System Dynamics Model | CATIA—Dassault Systèmes*. Accessed: Aug. 21, 2025. [Online]. Available: https://www.3ds.com/products/catia/dymola

[23] H. Elmqvist, S. E. Mattsson, and M. Otter, "Modelica—A language for physical system modeling, visualization and interaction," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, Apr. 1999, pp. 630–639.

[24] W. Tina, E. Donaldson, T. E. Dickinson, and W. H. Schmidt, "Failure analysis of once-through steam generator (OTSG) tube," *ASME Open J. Eng.*, vol. 2, pp. 969–974, Jan. 2023, doi: 10.1115/1.4062769.

[25] *Case Studies on HP Economizer Tube Leaks at Bouchain CCGT—Combined Cycle Journal*. Accessed: Aug. 12, 2025. [Online]. Available: https://www.ccj-online.com/case-studies-on-hp-economizer-tube-leaks-at-bouchain-ccgt/

[26] I. D. Wijayanti, H. C. Agustin, and A. Hariyadi, "Failure analysis of the leakage at boiler bottom wall tube of steam power plant," *AIP Conf. Proc.*, vol. 2384, no. 1, 2021, Art. no. 070006.

[27] P. Patil, B. Srinivasan, and R. Srinivasan, "Process fault detection in heat recovery steam generator using an artificial neural network simplification of a dynamic first principles model," *Comput. Aided Chem. Eng.*, vol. 44, pp. 2065–2070, Jan. 2018.

[28] M. A. Belay, A. Rasheed, and P. S. Rossi, "Multivariate time series anomaly detection via low-rank and sparse decomposition," *IEEE Sensors J.*, vol. 24, no. 21, pp. 34942–34952, Nov. 2024.

[29] M. A. Belay, A. Rasheed, and P. S. Rossi, "Sparse non-linear vector autoregressive networks for multivariate time series anomaly detection," *IEEE Signal Process. Lett.*, vol. 32, pp. 331–335, 2025.

[30] M. A. Belay, A. Rasheed, and P. S. Rossi, "MTAD: Multiobjective transformer network for unsupervised multisensor anomaly detection," *IEEE Sensors J.*, vol. 24, no. 12, pp. 20254–20265, Jun. 2024.

[31] M. A. Belay, A. Rasheed, and P. S. Rossi, "Autoregressive density estimation transformers for multivariate time series anomaly detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2025, pp. 1–5.

[32] *Thermal Power Library for Power Plant Operations | Modelon*. Accessed: Aug. 18, 2025. [Online]. Available: https://modelon.com/library/thermal-power-library/

[33] Y. Zhao, Z. Nasrullah, and Z. Li, "PyOD: A Python toolbox for scalable outlier detection," *J. Mach. Learn. Res.*, vol. 20, no. 96, pp. 1–7, 2022. [Online]. Available: http://jmlr.org/papers/v20/19-011.html

Mohammed Ayalew Belay (Graduate Student Member, IEEE) received the B.Sc. degree in physics from Addis Ababa University, Addis Ababa, Ethiopia, in 2011, the M.Sc. degree in physics from Bahir Dar University, Bahir Dar, Ethiopia, in 2013, and the M.Sc. degree in computer science from Hawassa University, Awasa, Ethiopia, in 2019. He is currently pursuing the Ph.D. degree with the Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

His research interests include deep learning, unsupervised learning, anomaly detection, time series analysis, and physics-informed neural networks.

Lucas Ferreira Bernardino received the Master of Science (M.Sc.) degree in chemical engineering from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 2019, and the Ph.D. degree in chemical engineering from Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2024, with a focus on a solid academic foundation in process systems engineering.

He is a Research Scientist with the Gas Technology Department, SINTEF Energy Research, Trondheim. His research interests center on process modeling, control, and optimization, with a significant emphasis on energy systems analysis and flexibility within industrial operations.

Sindre Stenen Blakseth received the Master of Science degree in applied physics from Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2021, where he is currently pursuing the Ph.D. degree in numerical mathematics.

He holds a part-time research position at the Gas Technology Unit, SINTEF Energy Research, Trondheim. His research expertise spans digital twin development, hybrid physics/data-driven modeling, surrogate modeling, machine learning, finite-element, and finite-volume methods, with applications in cryogenic storage (notably liquid hydrogen), gas turbine combined cycles, and industrial system digitalization.

Adil Rasheed is a Professor with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway. There, he works to advance the development of novel hybrid methods that combine big data, physics-driven modeling, and data-driven modeling in the context of real-time automation and control. In addition, he also holds a part-time position of a Senior Scientist with the Department of Mathematics and Cybernetics, SINTEF Digital, Trondheim, where he was previously the Leader of the Computational Sciences and Engineering Group from 2012 to 2018. His contributions in these roles have been the development and advancement of both the hybrid analysis and modeling and big data cybernetics concepts. Over the course of his career, he has been the driving force behind numerous projects focused on different aspects of digital twin technology, ranging from autonomous ships to wind energy, aquaculture, drones, business processes, and indoor farming. He is currently leading the digital twin and asset management-related work with the FME NorthWind Center, Trondheim.

Rubén Mocholí Montañés received the double Master of Science (M.Sc.) degree in mechanical engineering from the Faculty of Engineering, Lund University, Lund, Sweden, and in industrial engineering from the Polytechnic University of Valencia, Valencia, Spain, in 2012, and the Ph.D. degree in energy and process engineering from Norwegian University of Science and Technology Trondheim, Norway, in 2017.

He has been a Research Scientist at SINTEF Energy Research, Trondheim, since 2019. He conducted postdoctoral studies at the Energy Technology Division, Chalmers University of Technology, Gothenburg, Sweden. He currently leads and manages several large and small research and development and innovation projects in Norway and Europe. He is currently the Digital Twin KSP Project Leader. He has experience in experimental testing at research, pilot, and industrial facilities. His research interests include thermal process design optimization, dynamic process modeling and simulation, process control, operational flexibility of low-carbon thermal power plants, and CO_2 capture.

Pierluigi Salvo Rossi (Senior Member, IEEE) was born in Naples, Italy, in 1977. He received the Dr.Eng. (summa cum laude) degree in telecommunications engineering and the Ph.D. degree in computer engineering from the University of Naples Federico II, Naples, in 2002 and 2005, respectively.

He is currently a Full Professor and the Deputy Head of the Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. He is also a part-time Senior Research Scientist with the Department of Gas Technology, SINTEF Energy Research, Trondheim. Previously, he worked with the University of Naples Federico II; the Second University of Naples, Naples; NTNU; and Kongsberg Digital AS, Trondheim. He held visiting appointments with Drexel University, Philadelphia, PA, USA; Lund University, Lund, Sweden; NTNU; and Uppsala University, Uppsala, Sweden. His research interests fall within the areas of communication theory, data fusion, machine learning, and signal processing.

Prof. Salvo Rossi was awarded the Exemplary Senior Editor of IEEE COMMUNICATIONS LETTERS in 2018. He is/was with the Editorial Board of IEEE SENSORS JOURNAL, IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, IEEE COMMUNICATIONS LETTERS, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.