Physics-Informed ML Diagnostic Engine: DXC'26 LiU-ICE Track Solution

Antoni Rogowski
Warsaw University of Technology, Poland

Anna Sztyber-Betley
Warsaw University of Technology, Poland

Abstract

The paper presents a solution to the DX'26 Competition LiU-ICE Track. The presented approach merges physics insights with data-driven residuals. The applied Grey-Box approach, involving the custom implementation of a recurrent LSTM network with state memory, is well-suited to environments with high inertia. The system does not directly estimate specific values of individual physical quantities, but focuses on determining their rate and direction of change, eliminating the problem of exploding residuals. Furthermore, this approach offers an advantage over classical LSTMs due to its physical representation of states. Structural analysis and the determination of the MSOs enabled the isolation of relationships between the observed variables, whilst simultaneously decoupling the training data from physical quantities unrelated to the variable under investigation. Data decoupling was leveraged to selectively incorporate faulty data into training. The solution implements four virtual sensors, each equipped with its own pair of neural networks. The proposed fault isolation approach is based on a fault diagnosis matrix, fixed alarm thresholds, and scalar product analysis. Additionally, adaptive covariate shift compensation mechanism was introduced to handle changes in external conditions.

The resulting algorithm proved to be highly resilient to measurement noise and capable of accurately analysing the process's variability, scoring True Detection Rate $TDR = 92.26\%$, False Alarm Rate $FAR = 7.86\%$, and True Isolation Rate $TIR = 79.62\%$ on hidden test data.

2012 ACM Subject Classification Computing methodologies → Causal reasoning and diagnostics

Keywords and phrases Diagnosis, Algorithms, Evaluation, LiU-ICE Benchmark, Gray-Box, Physics Informed Neural Networks

Digital Object Identifier 10.4230/OASIcs.DX.2026.15

Category DX Competition

Funding This research was partially funded by an internal grant titled "Qualitative modeling of a technological process based on process data and partial models," awarded in 2025 by the Scientific Council of the Discipline of Automation, Electronic, Electrical Engineering and Space Technologies at the Warsaw University of Technology.

Acknowledgements We extend our gratitude to Professor Janusz Piechna for his consultation and valuable insights on fluid dynamics and engine physics.

1 Introduction

The solution presented here was designed to address a problem that lies in the blind spot between two approaches. On the one hand, there are black-box models [3, 16, 20, 7, 18], which rely entirely on machine learning architectures and completely disregard the physical implementation of the underlying processes. On the other hand, there are PINNs (Physically Informed Neural Networks) architectures [17], which require detailed knowledge of the physical parameters of systems, such as their dimensions or the phenomena occurring within them, often described by complex physical equations, for example, the Navier-Stokes equations.

The framework implemented in this paper bridges these two extremes by constructing Grey-Box estimators inspired by physical knowledge. Specifically, rather than allowing the neural networks to develop uninterpretable latent representations, we explicitly constrain the hidden states to reflect tangible physical quantities, such as mass and energy accumulation within control volumes. Consequently, the dimensionality of the state space ceases to be a tunable, arbitrary hyperparameter and becomes strictly dictated by the thermodynamic complexity of the monitored subsystem. This methodology aligns with and advances emerging hybrid approaches [8].

DXC competition has a long tradition of inspiring diagnostic solutions [13, 15]. This paper proposes a solution to the LiU-ICE Track (see Section 2), where the task is to diagnose a turbocharged gasoline engine in dynamic conditions under limited data and system knowledge [9]. Earlier versions of the competition were already held at Safeprocess'24 [10] and DX'25 [15]. The paper builds upon earlier solutions [10, 2, 8], introducing *Physics-Informed ML Diagnostic Engine.*

The main contributions of this paper include:

- **Decoupled Grey-Box Architecture (Section 3.1):** Development of parallel neural network-based estimators that take into account the laws of physics, ensuring mass conservation and modelling other real-world constraints. This approach reduces the unpredictability of traditional black-box artificial intelligence without the need for complex, analytical fluid dynamics models.
- **Physics-Informed Feature Engineering (Section 3.2):** Integration of aerodynamic domain knowledge directly into the preprocessing pipeline by deriving analytical features– such as the square root of pressure differentials based on the orifice flow equation.
- **Structurally-Aware Training Strategy (Section 3.3.2):** A novel training paradigm that explicitly exploits the analytical independence of the derived MSO sets. In contrast to conventional residual estimators trained exclusively on fault-free data, this methodology deliberately incorporates selected faulty data into the training pipelines of the decoupled nodes. This forces the neural networks to actively ignore specific cross-coupled disturbances, enhancing fault isolation robustness.
- **Adaptive Covariate Shift Compensation (Section 3.4):** Design of a dynamic mechanism active during a predefined starting phase. This approach neutralizes thermal variations arising from environmental conditions, ensuring robust inference and minimizing False Alarm Rates.
- **Robust Residual Processing and Isolation (Section 3.5 and Section 3.6):** A combination of exponential moving average filtering with vector space isolation logic based on cosine similarity. This ensures effective noise suppression and enables rapid fault isolation.

While several of the above aspects were raised in the diagnosis literature, the paper merges them with dedicated physics analysis into the consistent solution to the LiU-ICE benchmark.

2 Competition Description and Problem Formulation

The DX'26 Competition LiU-ICE Track requires developing a robust Fault Detection and Isolation (FDI) system for the air path of a turbocharged internal combustion engine (Figure 1) under strict operational constraints [9]. Table 1 explains the notation and lists symbols used in the paper.

Table 1 Notation used throughout the paper.

Symbol	Description
Measured sensor signals (known variables z)	
y_{pic}	Intercooler pressure
y_{Tic}	Intercooler temperature
y_{pim}	Intake manifold pressure
y_{waf}	Mass flow through the air filter
y_{xpos}	Throttle actuator position
y_{ω}	Engine speed
y_{pamb}	Ambient pressure
y_{Tamb}	Ambient temperature
Control signals	
u_{mf}	Requested injected fuel mass
u_{wg}	Requested wastegate actuator position
Internal (unknown) model variables x in MVEM mean-value engine model	
p_{im}	Intake manifold pressure
p_{ic}	Intercooler pressure
p_{em}	Exhaust manifold pressure
T_{im}	Intake manifold temperature
T_{ic}	Intercooler temperature
T_{em}	Exhaust manifold temperature
V_{im}	Intake manifold volume (constant)
V_{ic}	Intercooler volume (constant)
V_{em}	Exhaust manifold volume (constant)
R	Gas constant of air (constant)
W_c	Compressor mass flow
W_{th}	Throttle mass flow
W_{cyl}	Cylinder mass flow
W_t	Turbine mass flow
W_{wg}	Wastegate mass flow
ω_{tc}	Turbocharger speed
P_c	Compressor power
P_t	Turbine power
J_{tc}	Turbocharger inertia (constant)
Grey-box architecture	
MSO	Minimally Structurally Overdetermined set
x_k	Hidden state vector at sample k
u_k	Input vector at sample k
x_{aug}	Augmented unmodeled state
N_x	Dimension of the hidden state of an MSO subsystem
$\mathbf{u}_{waf}$	Input feature vector for the WAF subsystem
r_k	Residual: $r_k = y_k - \hat{y}_k$
$g(x, u)$	State-transition subnetwork (estimated state derivative $\dot{x}$)
$g_{aug}(x_k, u_k)$	State-transition subnetwork for the augmented state x_{aug}
$h(x, u)$	Output subnetwork acting as virtual sensor
T_s	Euler integration step / sampling period ($T_s = 0.05\,\mathrm{s}$)

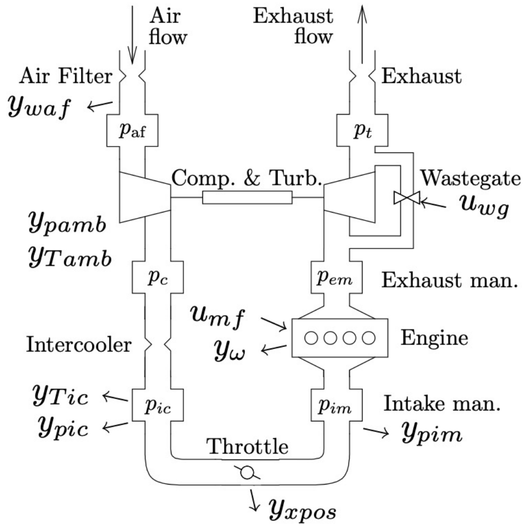


Figure 1 Schematic diagram of the internal combustion engine air path. Connections of the physical components and the placement of key sensors [9].

2.1 Physical and Operational Challenges

The diagnostic architecture must cope with severe nonlinearities, thermal inertia, and actuator lags. A major obstacle is closed-loop fault propagation. Every single fault affects all physical states across the entire air intake system. Furthermore, since obtaining fault data on physical engines is potentially destructive, anomalous training samples are scarce. This forces reliance on nominal data and physical system knowledge. Solutions must also respect strict execution timeouts, necessitating computationally lightweight, two-stage (detection and isolation) architectures suitable for real-time inference.

2.2 Diagnosis Targets

The system must monitor the air flow and, when it detects the fault, it should classify the operating condition as one of the following:

- Multiplicative biases on sensors: air mass flow (f_{waf}), intercooler pressure (f_{pic}), and intake manifold pressure (f_{pim}).
- Physical air leakage in the intake manifold (f_{iml}).
- Unknown fault (f_x)

2.3 Evaluation Criteria

Performance is evaluated using three metrics (normalized 0–1):

- **FAR (False Alarm Rate):** Percentage of false positive detections during nominal operation.
- **TDR (True Detection Rate):** Percentage of correctly detected faults.
- **FIA (Fault Isolation Accuracy):** Average probability assigned to the correct fault class upon detection.

The final ranking is determined by a composite objective function:

$$\text{Total Score} = \frac{(1 - FAR) + TDR + FIA}{3} \tag{1}$$

A naive system (e.g. one that never triggers alarms) achieves a baseline score of around 0.33.

3 Methods

3.1 Grey-Box State-Space Architecture

3.1.1 Hybrid modelling of system dynamics

To avoid the risks of overfitting and poor extrapolation, which are inherent in black-box models that cannot be interpreted from a physical perspective, we adopt a hybrid approach based on state space.

This work adopts the differential modelling framework proposed by [8], where the system dynamics are approximated using a discrete-time state-space Grey-Box model comprising two interconnected neural subnetworks. The state transition network, $g(x, u)$, estimates the system's gradient, and the hidden state is updated iteratively using Euler numerical integration (notation according to Table 1):

$$x_{k+1} = x_k + T_s \cdot g(x_k, u_k). \tag{2}$$

Subsequently, the output network, $h(x, u)$, acts as a virtual sensor, mapping the internal states and control inputs to the estimated measurement:

$$\hat{y}_k = h(x_k, u_k). \tag{3}$$

The raw diagnostic residual r_k, representing the deviation between the actual behavior of the physical system and the prediction of the model, is then calculated as:

$$r_k = y_k - \hat{y}_k. \tag{4}$$

Our architecture strictly adheres to the core structural principles of the Grey-Box methodology introduced by Jung [8]:

- **Structural Decentralization via MSO Sets:** Rather than utilizing a single global recurrent model, the architecture is strictly decentralized into four independent diagnostic tests (MSO sets, see Section 3.1.2).
- **Explicit Physical Constraining:** The hidden states are explicitly constrained to represent tangible physical quantities (e.g., mass and energy accumulation in control volumes). Consequently, the dimensionality of the state space is assigned to match the thermodynamic complexity of each specific MSO (ranging from 1 to 5 states) (details in Section 3.1.2).
- **Real-Time Hardware Synchronization:** The state update frequency (integration step T_s) is strictly dictated by the physical sensor sampling rate. Instead of using an arbitrary numerical step size, the integration process is closely tied to the frequency at which new data arrives from the sensor.

3.1.2 Structural analysis and selection of subsystems

3.1.2.1 Prerequisites

This solution uses structural analysis to find diagnostic tests. *Structural analysis* is a model-based diagnosis technique that abstracts a system of equations into an incidence matrix capturing only which variables appear in which equations, thereby enabling the identification of redundant subsets of equations and the derivation of diagnostic tests for large-scale systems without requiring their explicit analytical solution [1, 12, 6].

A Minimal Structurally Overdetermined (MSO) set, introduced in [12], is a concept important for deriving diagnostic tests. Consider a system described by a model $\mathcal{M}(z, x, f)$, where z denotes the vector of *known variables* (e.g., measured sensor outputs and controlled inputs), x denotes the vector of *unknown variables* (internal states and algebraic variables), and f is a vector of fault parameters. For any subset of equations $\psi \subseteq \mathcal{M}(z, x, f)$, the *structural redundancy* ρ_ψ is defined as the difference between the number of equations in ψ and the number of unknown variables appearing in them.

A subset ψ is said to be an MSO set if $\rho_\psi = 1$, i.e., it contains exactly one more equation than unknown variables, and no proper subset of ψ is itself overdetermined. An MSO set carries the minimum amount of analytical redundancy needed to perform a consistency check between the model and the measurements [12, 1].

3.1.2.2 Structural model of the engine

The engine MSOs were determined using the Python implementation of the Fault Diagnosis Toolbox [5]. While the original LiU-ICE benchmark provides a comprehensive Differential-Algebraic Equation (DAE) system of 94 relations [9], we utilized a condensed core version of the Mean-Value Engine Model (MVEM) [4]. By focusing strictly on air-path dynamics and turbocharger energy balances, and omitting auxiliary algebraic constraints, the core model was reduced to 21 physical relations, 18 unknown variables, and 10 measurable signals. With a redundancy degree of 3, this formulation yielded 27 Minimally Structurally Overdetermined (MSO) sets. The incidence matrix for MVEM is presented in Table 2.

MSO 0, MSO 1, and MSO 10 were selected to achieve a full fault isolability while minimizing the computational complexity. However, because purely structural MSO subsets struggled to isolate the air mass flow fault, an additional, dedicated residual (r_{waf}) was added for the detection of f_{waf}. The complete computation pipeline for all four diagnostic residuals is summarized in Figure 2.

Although the MSO sets dictate the input variables for each subsystem, feeding raw sensor data directly into the neural estimators often results in degraded predictive performance due to the highly non-linear nature of the engine's fluid dynamics. To address this, Physics-Informed Feature Engineering was employed prior to inference. For example, the input data for the WAF subsystem was non-linearly transformed using logarithmic mappings of engine rotational speed and throttle position (e.g., $x_1 = \log(\omega + \epsilon)\dot{m}_{air}$) (details in Figure 2).

3.1.3 Selected diagnostic tests

The MSO 0 subsystem ($N_x = 1$) models the local mass balance of the intake manifold. The fact that this subsystem is independent of the mass flow measurement (y_{waf}) and the turbine dynamics makes it robust to disturbances from the "cold side" of the system, allowing for a stable assessment of the manifold's tightness. The hidden state here follows the mass conservation equation:

$$\frac{dp_{im}}{dt} = \frac{R \cdot T_{im}}{V_{im}}(W_{th} - W_{cyl}) - f_{iml}. \tag{5}$$

The MSO 1 subsystem ($N_x = 5$) constitutes the most complex diagnostic node, providing a global thermodynamic consistency check. Notably, structural analysis dictates that this subsystem must remain strictly decoupled from the mass air flow (y_{waf}) and intercooler pressure (y_{pic}) sensors to ensure reliable fault isolation. Consequently, estimating the intake manifold pressure (y_{pim}) based solely on ambient conditions, injected fuel mass, and actuator positions requires the Grey-Box model to internally simulate the entire air-path dynamics. The five-dimensional state space perfectly reflects this physical propagation chain. It encompasses the four core differential variables of the Mean-Value Engine Model (intercooler pressure p_{ic},

Table 2 Incidence Matrix of the reduced Mean-Value Engine Model (MVEM). Rows represent structural relations (R), while columns correspond to unknown variables (X), known variables (Z), and faults (F). A bullet (•) denotes the incidence of a variable in a given relation.

	Unknown Variables (X)																		Known Variables (Z)										Faults (F)			
	W_{af}	W_c	W_{th}	W_{cyl}	W_t	W_{wg}	p_{ic}	$\dot{p}_{ic}$	p_{im}	$\dot{p}_{im}$	p_{em}	$\dot{p}_{em}$	T_{ic}	T_{em}	ω_{tc}	$\dot{\omega}_{tc}$	P_c	P_t	y_{waf}	y_{pic}	y_{pim}	y_{Tic}	y_{xpos}	y_{ω}	y_{pamb}	y_{Tamb}	u_{wg}	u_{mf}	f_{waf}	f_{pic}	f_{pim}	f_{iml}
R_0	•																		•										•			
R_1	•	•																														
R_2		•					•								•										•	•						
R_3		•	•					•					•																			
R_4							•	•																								
R_5							•													•										•		
R_6													•									•										
R_7			•				•		•				•										•									
R_8				•					•															•								
R_9			•	•						•																						•
R_{10}									•	•																						
R_{11}									•												•										•	
R_{12}				•	•	•						•		•																		
R_{13}											•	•																				
R_{14}				•										•														•				
R_{15}					•						•			•	•										•							
R_{16}					•						•			•	•			•							•							
R_{17}						•					•			•											•		•					
R_{18}		•					•								•		•								•	•						
R_{19}															•	•	•	•														
R_{20}															•	•																

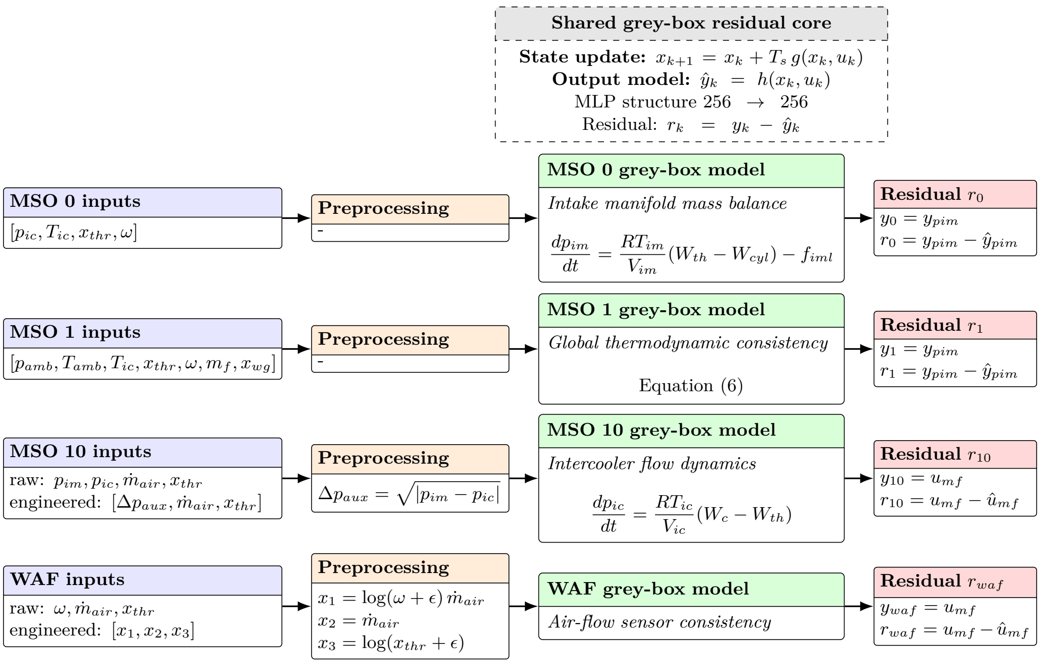


Figure 2 Four implemented Grey-box model residuals.

manifold pressure p_{im}, exhaust pressure p_{em}, and turbocharger speed ω_{tc}), supplemented by an augmented state (x_{aug}) that explicitly captures unmodeled inertias, such as exhaust thermal capacity and wastegate mechanical delays:

$$\frac{d}{dt}\begin{bmatrix} p_{ic} \\ p_{im} \\ p_{em} \\ \omega_{tc} \\ x_{aug} \end{bmatrix} = \begin{bmatrix} \frac{R \cdot T_{ic}}{V_{ic}}(W_c - W_{th}) \\ \frac{R \cdot T_{im}}{V_{im}}(W_{th} - W_{cyl}) \\ \frac{R \cdot T_{em}}{V_{em}}(W_{cyl} + u_{mf} - W_t - W_{wg}) \\ \frac{P_t - P_c}{J_{tc} \cdot \omega_{tc}} \\ g_{aug}(x_k, u_k) \end{bmatrix}. \tag{6}$$

The MSO 10 subsystem ($N_x = 1$) focuses on intercooler dynamics, enabling rapid verification of flow continuity before the throttle:

$$\frac{dp_{ic}}{dt} = \frac{R \cdot T_{ic}}{V_{ic}}(W_c - W_{th}). \tag{7}$$

Input data preprocessing was described in Section 3.2.1.

Integrated WAF Subsystem. ($N_x = 1$). This diagnostic node differs slightly from the previously described MSO subsystems. Initially, the diagnostic architecture was designed to rely exclusively on the derived MSO subsets for full fault isolability. However, preliminary model evaluations revealed that purely structural residuals struggled to consistently isolate mass air flow faults under highly transient operating conditions due to the severe nonlinearities. To address this fundamental limitation, we explicitly adopted the physics-informed feature engineering approach proposed by Corrini et al. [10]. The input vector $\mathbf{u}_{waf}$ directly utilizes the logarithmic transformations validated by the authors–specifically the interaction between the mass air flow and the logarithm of engine speed, derived from the state-space equations of a turbocharged engine, to linearize the flow dynamics (see also Section 3.2.2):

$$\mathbf{u}_{waf} = \begin{bmatrix} \log(y_\omega + \epsilon) \cdot y_{waf} \\ y_{waf} \\ \log(y_{xpos} + \epsilon) \end{bmatrix}, \quad \epsilon = 10^{-6}. \tag{8}$$

3.2 Physics Based Data Preprocessing

The effectiveness of machine learning algorithms in the diagnosis of non-linear systems depends critically on the appropriate representation of input features.

3.2.1 Implementation of non-linear flow relationships

In accordance with the structure of the MSO 10 subsystem, as determined during the structural analysis phase, the key element is the flow dynamics through the throttle. Instead of forcing the neural network to learn the non-linear relationship between flow and pressure, an auxiliary feature Δp_{aux} was introduced:

$$\Delta p_{aux} = \sqrt{|p_{ic} - p_{im}|} \tag{9}$$

This transformation directly reflects the simplified equation for flow through a constriction. This reduces the complexity of the network's optimisation task, allowing it to focus on approximating the variable flow coefficient rather than the square root non-linearity itself.

3.2.2 Feature engineering for the WAF subsystem

For the flow sensor fault detection module, pre-processing based on the analysis of state variable interactions was applied, strictly following the methodology proposed by Corrini et al. [10]. The input vector was constructed using a combination of raw sensor data and engineered synthetic features:

- The non-linear interaction between engine speed and air mass flow: $\log(y_\omega + \epsilon) \cdot y_{waf}$.
- The raw, untransformed mass air flow signal: y_{waf}.
- The linearized throttle position signal: $\log(y_{xpos} + \epsilon)$.

The use of logarithmic transformations allows for the decoupling of the highly non-linear exponential relationships prevailing in the turbocharging system. This explicit linearization significantly improves the convergence and learning stability within the Grey-Box neural network framework.

3.3 Truncated Backpropagation Through Time Training Scheme

Soft sensors training pipeline was implemented in PyTorch [14].

3.3.1 Optimisation and time horizon

The process of training the parameters of the g and h subnetworks was carried out using the backpropagation algorithm. Due to the recursive nature of Grey-Box models and the need to maintain gradient stability, the Truncated BPTT (TBPTT) [19] technique was employed. The training data was organised into batches of size 16, with each batch comprising 600 samples. At a sampling rate of 20 Hz, this corresponds to a fixed time horizon of 30 s.

The Adam optimiser [11] was used as the optimisation algorithm with an initial learning rate of $5 \cdot 10^{-4}$. To fine-tune the weights in the advanced stages of training, dynamic step reduction was implemented, which reduced the learning rate by 3% every 10 epochs ($\gamma = 0.97$). Empirical evaluation confirmed that this specific scheduling yielded the most stable convergence and the lowest validation error across the entire provided benchmark dataset, preserving a non-vanishing step size even after 1000 epochs. Furthermore, to ensure robust optimization against high-frequency sensor noise and prevent weight oscillations induced by large outlier residuals, the Mean Absolute Error (MAE) was chosen as the training loss function in preference to the standard Mean Squared Error (MSE).

To prevent overfitting, the training process, scheduled for a maximum of 2000 epochs, was coupled with an Early Stopping mechanism. Training was stopped if the loss function value did not improve (by at least 10^{-6}) over a defined patience period of 80 epochs. The networks typically achieved convergence between the 1200th and 1800th epochs. The final convergence epochs and the corresponding loss values for all generated diagnostic nodes are summarized in Table 3. Throughout the entire process, the network state corresponding to the lowest loss value obtained for the training dataset suited specific to each MSO was monitored and recorded each time, ensuring that models with the highest achievable result were obtained.

3.3.2 Decoupled Dataset Composition

A critical aspect of training the decentralized diagnostic nodes was the deliberate composition of the training datasets. While standard residual generators are typically trained exclusively on nominal, fault-free data, this architecture leverages the structural independence properties derived from the MSO sets to enhance isolation robustness.

Table 3 Training convergence statistics for the Grey-Box diagnostic subsystems. The training process utilized an Early Stopping mechanism with a patience of 80 epochs for all models. Note that the Final Loss is computed on normalized data uniformly scaled to the $[0, 1]$ range.

Subsystem	State Dim. (N_x)	Max Epochs	Stopping Epoch	Final Loss (MAE)
MSO 0	1	2000	1803	0.00685
MSO 1	5	2000	1485	0.00522
MSO 10	1	2000	1288	0.00940
WAF Subsystem	1	2000	1451	0.00697

If structural analysis dictates that a specific subsystem is perfectly decoupled from a particular sensor, feeding the network with datasets containing artificial faults from that sensor forces the Grey-Box model to actively learn to ignore those specific disturbances. Consequently, each sub-model was trained on a unique dataset combination, comprising both healthy operation data and selected faulty sequences. The exact composition of the training sets for each diagnostic node is detailed in Table 4.

Table 4 Composition of training datasets for individual subsystems. Models were intentionally exposed to specific fault scenarios (f) from which they are structurally decoupled.

Subsystem	Nominal Data	Injected Fault Data	Structural Justification
MSO 0	NF	f_{waf}	Manifold dynamics structurally decoupled from mass air flow sensor.
MSO 1	NF	f_{waf}, f_{pic}	Global thermodynamic model forced to ignore cold-side sensor anomalies.
MSO 10	NF	f_{waf}, f_{iml}	Intercooler dynamics decoupled from mass air flow and downstream manifold leakage.
WAF Sub.	NF	f_{pic}, f_{pim}	Targeted flow estimator explicitly decoupled from pressure sensor deviations.

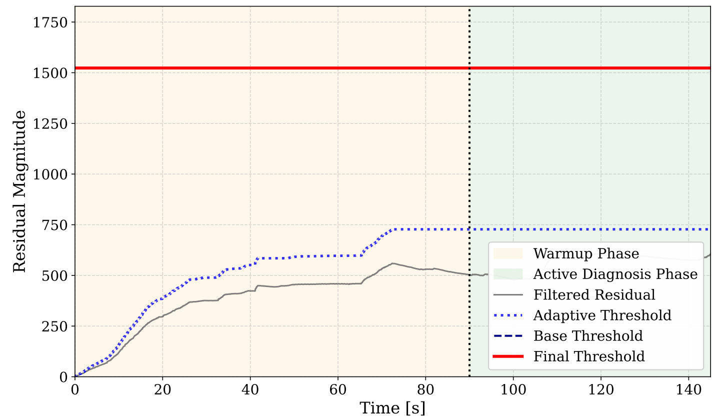
3.4 Environmental adaptation

Each incoming sample is mapped on-the-fly to the required input vectors and subjected to the physics-informed transformations defined during the pre-processing phase. The estimated state values are then compared with the actual system readings to generate the diagnostic residual vectors (see Figure 4).

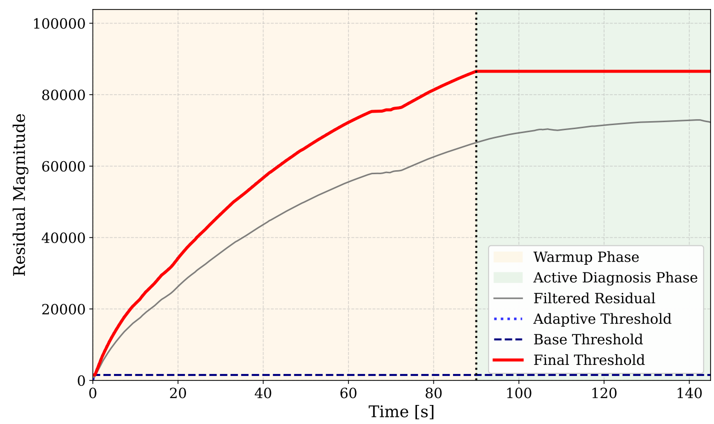
An innovative feature of this streaming workflow is the heuristic decision threshold adaptation mechanism applied specifically to the MSO 1 subsystem. In automotive air-path systems, ambient temperature (y_{Tamb}) and ambient pressure (y_{pamb}) act as absolute boundary conditions that fundamentally dictate intake air density. Consequently, they cannot be simply omitted, any fluctuation in these environmental factors cascades through the thermodynamic equations, significantly impacting baseline mass flows and pressure drops.

To compensate for the resulting Covariate Shift, the system suspends the generation of diagnostic alerts for the first 90 seconds (1800 samples) of the driving cycle. During this initialisation phase, it actively analyses the baseline prediction error induced by the current environmental conditions. The highest recorded deviation during this window establishes the new baseline for the adaptive alarm threshold, supplemented by a 30% safety margin. The adaptive thresholding mechanism was illustrated in Figure 3.

Crucially, this adaptive mechanism was deliberately omitted for the remaining subsystems (such as the MSO 0). Since MSO 0 relies heavily on highly dynamic, flow-restricting components like the throttle, calibrating its threshold during a relatively steady-state initialisation phase would lead to a severe underestimation of transient errors. This would inevitably result in a flood of false alarms during sudden throttle openings later in the cycle due to the problem of lack of enforcement.



(a) Training conditions (FTP-75).



(b) New conditions (Synthetic drift).

Figure 3 Comparison of the threshold adaptation logic: In standard training conditions (a), the system maintains the robust safety floor. Under significant covariate shifts (b), the adaptive candidate successfully supersedes the base floor, becoming the final active threshold to maintain diagnostic sensitivity without increasing false alarms. To simulate this covariate shift in ambient conditions, synthetic data featuring altered ambient temperature and pressure was generated, strictly mirroring realistic operational bounds.

3.5 Robust Detection Mechanism

To make the decision matrix independent of stochastic sensor noise and transients in fluid dynamics, a two-stage filtering architecture was implemented (Figure 4):

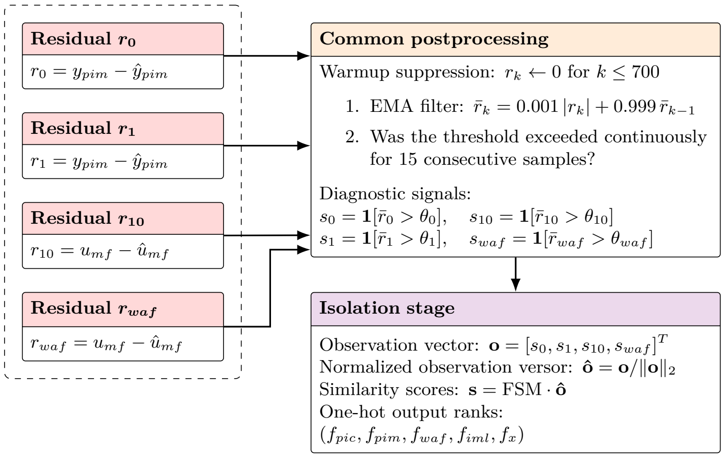


Figure 4 Fault detection and isolation mechanism. The diagnostic residuals are computed according to the architecture detailed in Figure 2.

1. EMA (Exponential Moving Average) filter: Raw residuals from the model outputs are smoothed in real time using a low-pass filter with a heuristically determined time constant ($\alpha = 0.001$), which mitigates the effect of high-frequency measurement noise (see Figure 5).

2. Inertial Filter (Debounce Filter): The signal indicating that the filtered residual has exceeded the alarm threshold is verified by an error persistence counter. The alarm is only confirmed if the threshold is exceeded continuously for 15 consecutive samples (0.75 s).

Such a rigorous configuration eliminates alarms generated by micro-jumps and the phenomenon of chattering at the decision threshold. Although the use of an inertial filter slightly increases the time to detect, this is an acceptable trade-off resulting from the prioritisation of reducing the false alarm rate.

3.6 Fault Isolation and Vectorized Signature Resolution

The fault isolation strategy constitutes one of the most critical stages in the proposed architecture, directly driving the overall performance in the DX Competition. The official evaluation metric strictly balances the False Alarm Rate (FAR), True Detection Rate (TDR), and Fault Isolation Accuracy (FIA), as defined in Equation (1). To maximize the FIA, the diagnostic logic must resolve the boolean threshold flags from the four independent subsystems $(s_0, s_1, s_{10}, s_{waf})$ into a single definitive fault classification.

3.6.1 Cosine Similarity Resolution

The isolation mechanism is triggered only after the active alarms pass a strict inertial debounce filter, which requires the threshold violation to persist continuously for at least 15 samples (0.75 s) to prevent decision chatter. Once confirmed, the active flags form an observation vector $\mathbf{o} = [s_0, s_1, s_{10}, s_{waf}]^T$.

To compare this empirically observed symptom pattern against the theoretically encoded fault dependencies, a cosine similarity approach is employed. The observation vector is first L^2-normalized into a unit versor ($\hat{\mathbf{o}} = \mathbf{o}/\|\mathbf{o}\|_2$) and subsequently projected onto the reference

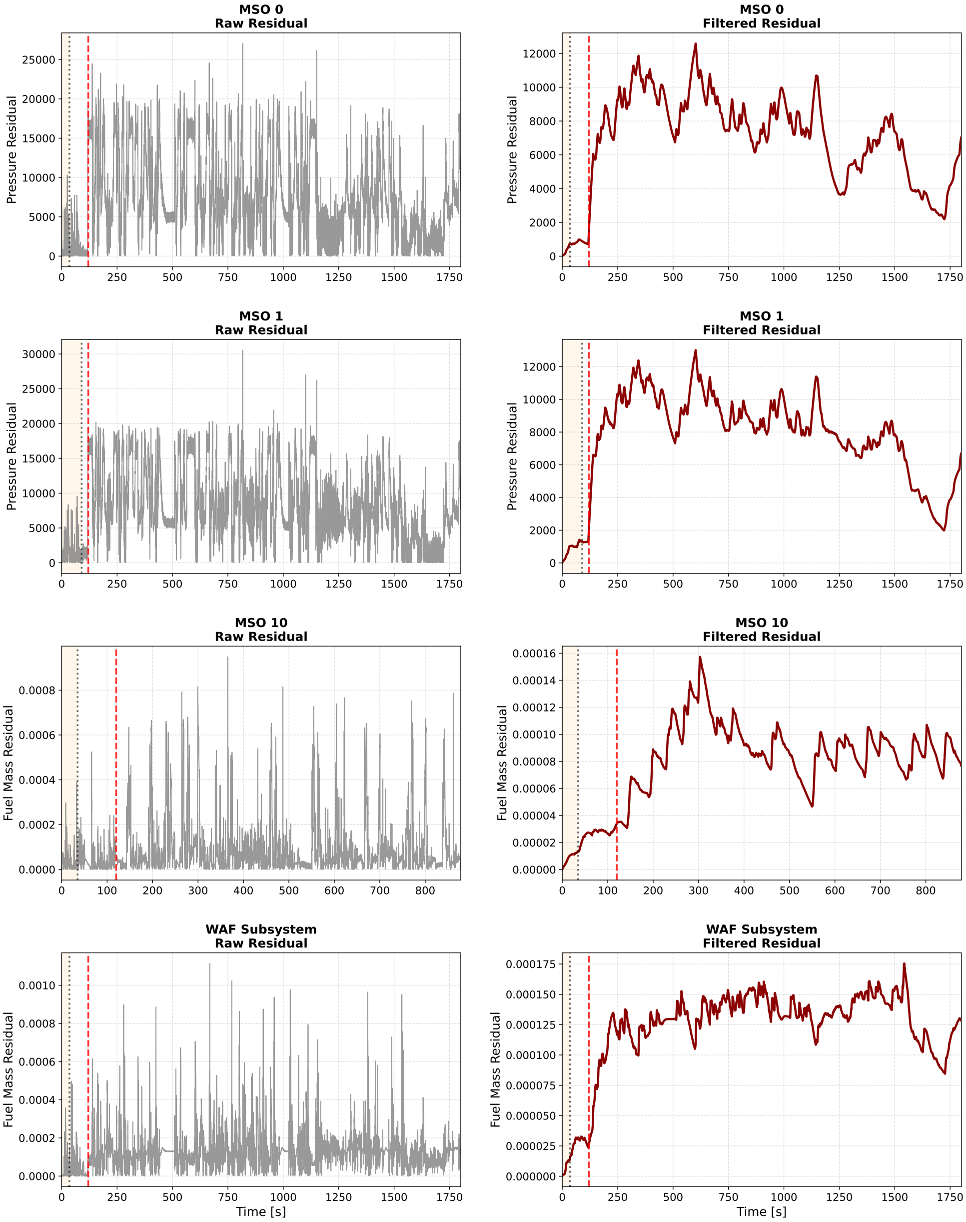


Figure 5 Comprehensive comparison of diagnostic residuals before (left column) and after (right column) Exponential Moving Average (EMA) filtration. The plots demonstrate the critical necessity of low-pass filtering in suppressing high-frequency sensor noise and transient spikes. MSO 0, MSO 1, and the WAF Subsystem are evaluated under an intake manifold leakage (f_{iml}) scenario, while MSO 10 highlights the response to an intake manifold pressure sensor bias (f_{pim}). A general initialization phase of 35 seconds (700 samples) is applied to all nodes to stabilize internal states and EMA filters, which is strictly necessary to prevent spurious false alarms during system startup. Furthermore, MSO 1 utilizes an extended 90-second environmental adaptation window to calibrate its dynamic threshold against ambient conditions. For each respective subsystem, the end of its required warmup phase is delineated by a vertical black dotted line. The red dashed line at $t = 120\,\mathrm{s}$ marks the onset of the injected fault.

Fault Signature Matrix (FSM) via the dot product:

$$\mathbf{s} = FSM \cdot \hat{\mathbf{o}} \tag{10}$$

where $\mathbf{s}$ represents the array of similarity scores across the known fault classes.

3.6.2 Hardmax Optimization and Troubleshooting Trade-off

The final isolation output is strictly binarized by applying an *argmax* function to the score vector $\mathbf{s}$, assigning a value of 1.0 to the most probable candidate and zeroing out all remaining possibilities. From a practical engineering perspective, outputting a continuous probabilistic distribution of suspected faults would be highly beneficial for real-world mechanical troubleshooting, as it preserves critical information regarding secondary fault suspicions. However, the implementation of this definitive "hardmax" algorithm was a deliberate design choice tailored specifically to maximize the competition's scoring function, which inherently rewards unambiguous isolation decisions.

3.6.3 Handling Unknown Faults

A crucial requirement of the benchmark is the ability to flag unknown anomalies (f_x). To prevent the system from forcibly assigning a novel anomaly to one of the four predefined classes, explicit structural safeguards were coded into the matrix logic. If the total similarity score evaluates to zero, or if a specifically unmodeled symptom pattern emerges (e.g., the simultaneous persistent activation of MSO 1, MSO 10, and the WAF subsystem without triggering MSO 0), the algorithm executes a case override. This explicitly routes the diagnosis to the unknown fault vector, ensuring safe and predictable behavior under unforeseen operating conditions.

Table 5 Fault Signature Matrix (FSM) defining the theoretical isolation logic. A bullet (•) indicates that the diagnostic node is sensitive to a specific fault.

Diagnostic Node	**Target Fault Scenarios (F)**			
	MAF (f_{waf})	**Leak** (f_{iml})	**IC Press** (f_{pic})	**IM Press** (f_{pim})
r_0 (MSO 0)		•	•	•
r_1 (MSO 1)		•		•
r_{10} (MSO 10)			•	•
r_{waf} (WAF Sub.)	•	•		

4 Results

4.1 Computational Complexity and Real-Time Feasibility

The evaluation script was executed in a standard desktop environment. The hardware configuration consisted of an Intel Core i7-10850H CPU (2.70 GHz) paired with 32 GB of RAM, operating on Windows 11 Pro. Notably, the inference was conducted exclusively on the CPU without reliance on a dedicated GPU accelerator.

While the final inference phase of the trained Grey-Box models is exceptionally fast–relying purely on lightweight matrix multiplications and basic numerical integration–the offline training process using Truncated BPTT is computationally intensive. Training each diagnostic subnetwork to convergence required approximately 6 to 10 hours of computation time on the specified CPU setup. However, once trained, empirical measurements during the validation phase revealed an average processing time of 0.82 ms per sample. Because $t \ll 100$ ms, this performance validates the architecture's compliance with the competition's timeout limits and its readiness for real-time industrial deployment. This time is as much restricted because the timeouts limit the maximum processing time per sample and not the average. Max from these files is Max: 49.5 ms which is about half of the timeout.

4.2 Evaluation on the shared datasets

The proposed Grey-Box diagnostic architecture was evaluated against the official benchmark datasets provided by the competition organizers. The validation set encompasses two driving cycles: the FTP-75 (Federal Test Procedure) and the WLTP (Worldwide Harmonised Light Vehicles Test Procedure). The datasets contain both nominal fault-free operations (NF) and various fault scenarios injected at different magnitudes (e.g., sensor biases ranging from 0.85 to 1.15, and manifold leaks of 4 mm and 6 mm).

Table 6 presents the aggregated evaluation metrics, averaged across all specific fault magnitudes within each category. The results demonstrate the high robustness of the architecture, an achievement attributed mainly to the decentralized MSO formulation, supported by physics-informed constraints. The system achieved a global Total Score of 92.29%, driven primarily by a low average False Alarm Rate (FAR) of 0.68% and a near-perfect Fault Isolation Accuracy (FIA) of 98.97%. Notably, for the majority of the sensor bias faults (f_{pic}, f_{pim}, f_{waf}), the isolation accuracy reached a perfect 100%, validating the effectiveness of the Cosine Similarity isolation logic combined with the structural decoupling strategy.

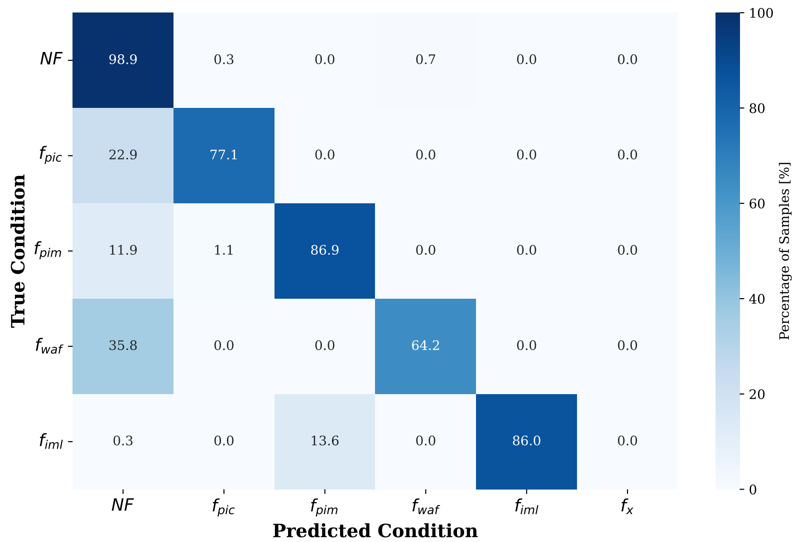


Figure 6 Confusion matrix.

Crucially, the proposed diagnostic architecture is entirely deterministic during the inference phase. The evaluation metrics (FAR, TDR, and FIA) remain perfectly consistent across multiple executions, demonstrating high reproducibility regardless of the underlying operating system or hardware architecture. This hardware-agnostic predictability is a significant advantage for potential real-world deployment. However, it should be noted that the final evaluation score exhibits a minor dependency on the Python runtime environment. As observed during testing, upgrading the interpreter from Python 3.11.8 distribution to version 3.11.12 yielded a slight improvement in the final score. This variation is strictly attributed to low-level differences in floating-point arithmetic handling and underlying library optimizations across Python releases.

Table 6 Averaged evaluation results across the FTP-75 and WLTP driving cycles. Metrics are aggregated by fault class, encompassing all tested fault magnitudes. (Note: TDR for NF indicates the correct classification of nominal state).

Driving Cycle	Fault Class	FAR [%]	TDR [%]	FIA [%]	Score [%]
FTP-75	NF	0.00	100.00	100.00	100.00
	f_{pic}	0.00	78.15	100.00	92.72
	f_{pim}	0.00	85.20	100.00	95.08
	f_{waf}	0.00	57.68	100.00	85.92
WLTP	NF	1.80	100.00	100.00	98.20
	f_{pic}	0.00	76.87	100.00	92.28
	f_{pim}	4.07	89.65	97.53	94.37
	f_{waf}	0.00	67.35	100.00	89.12
	f_{iml}	0.00	99.65	86.30	95.30
Global Average (41 files):		**0.68**	**78.75**	**98.97**	**92.29**

The performance drop observed in the True Detection Rate (TDR) for specific faults (e.g., the 0.95 and 1.05 bias magnitudes for f_{waf}) is a direct consequence of the rigorous 0.75 s inertial debounce filter, EMA filter and the deliberate tuning of the alarm thresholds. By sacrificing detection sensitivity for micro-faults that closely mimic natural sensor noise, the architecture successfully prevented error propagation and achieved the primary industrial requirement: avoiding false positives under highly transient conditions.

4.3 Evaluation on the Competition Datasets

To rigorously validate the generalisation capabilities of the proposed Decoupled Grey-Box Architecture, the final algorithmic solution was submitted to the competition organisers for an independent, blind evaluation. The system was tested against a set of datasets withheld during the training and validation phases. These hidden datasets span varying driving cycles (WLTP, FTP75 City, FTP75 Highway) and introduce different fault magnitudes, including categorised small and large intake manifold leakages (f_{iml}).

The official performance report returned by the evaluating committee is summarised in Table 7.

4.3.1 Performance Analysis

As evidenced by the official feedback, the architecture exhibits exceptional robustness and accuracy in diagnosing mass air flow (f_{waf}) faults and intake manifold leakages (f_{iml}). Notably, the system handles leakages – which are traditionally difficult to decouple from

Table 7 Official evaluation results provided by the competition organisers. High scores across most scenarios validate the robustness of the architecture, while specific degradation in two datasets highlights operational limits.

Driving Cycle	**Fault & Size**	**TDR [%]**	**1-FAR [%]**	**FIA [%]**	**Score**
FTP75 City 2	f_{iml} (large)	99.57	100.00	98.72	**99.43**
FTP75 City 2	f_{iml} (small)	97.13	100.00	99.33	**98.82**
FTP75 Highway	f_{iml} (large)	100.00	95.60	100.00	**98.53**
FTP75 Highway	f_{iml} (small)	96.66	100.00	98.87	**98.51**
FTP75 City 2	f_{waf} (-12%)	100.00	94.06	64.05	**86.04**
FTP75 City 2	f_{waf} (+12%)	96.83	100.00	100.00	**98.94**
FTP75 (Standard)	f_{pim} (-10%)	21.12	85.10	30.70	**45.64**
FTP75 (Standard)	f_{pim} (+10%)	99.93	100.00	91.25	**97.06**
FTP75 Highway	f_{pim} (-10%)	98.79	100.00	92.36	**97.05**
FTP75 Highway	f_{pim} (+10%)	99.22	100.00	100.00	**99.74**
WLTP	f_{pic} (-3%)	97.91	96.35	80.20	**91.48**
WLTP	f_{pic} (+8%)	100.00	34.61	0.00	**44.87**
	Overall Average:	**92.26**	**92.14**	**79.62**	**88.01**

rapid throttle transients – with near-perfect scores (averaging over 98.8%) across both highly dynamic city driving and steady-state highway conditions. The dynamic thresholding and covariate shift adaptation mechanisms successfully maintained a 100% non-false-alarm rate $(1 - \mathrm{FAR})$ in the majority of these independent test scenarios.

4.3.2 Boundary Limitations and Edge Cases

Despite the overall high performance, the blind evaluation exposed two specific boundary conditions where the diagnostic logic degrades. Firstly, a negative drift in the intake manifold pressure sensor (f_{pim} at -10%) under the standard FTP75 cycle resulted in a severely suppressed TDR (21.12%). Interestingly, the exact same fault magnitude was easily detected and isolated under Highway conditions (Score: 97.05). This suggests that the magnitude of a -10% pressure drop becomes indistinguishable from the model's natural uncertainty bounds during the frequent, low-load deceleration phases characteristic of urban FTP75 driving.

Secondly, a positive drift in the intercooler pressure sensor (f_{pic} at +8%) under the WLTP cycle caused a complete collapse of the isolation logic (FIA = 0.00%) alongside a high false alarm rate. A 100% TDR indicates that the residual generators correctly flagged an anomaly, but the resultant symptom vector fell outside the expected bounds of the Fault Signature Matrix. This independent finding highlights the extreme sensitivity of the air-path thermodynamics to positive pressure biases upstream of the throttle, identifying a clear direction for future structural improvements.

5 Discussion

Sensitivity to Mass Air Flow Faults

The most significant limitation observed during the evaluation was the relatively lower True Detection Rate (TDR) for the air mass flow sensor (f_{waf}). Despite incorporating specialized feature engineering and logarithmic transformations, the system struggled to consistently distinguish small fault magnitudes (e.g., 0.95 and 1.05) from the high-frequency measurement noise and model mismatches during steep transients. This suggests that while the integrated WAF module enhances isolability, its detection threshold must remain high to avoid false positives, ultimately sacrificing sensitivity to low-magnitude sensor drifts.

Detection Latency vs. Robustness

To achieve a near-zero False Alarm Rate (FAR), a rigorous 0.75 s (15 samples) inertial debounce filter and high inertia filter were implemented. While this successfully suppressed noise-induced spikes and transient-related residuals, it introduced a significant activation delay. Future work should focus on developing adaptive filtering techniques that can reduce this latency without compromising the system's robustness against measurement noise.

Intake Manifold Leakage Isolation

The isolation of physical air leakage (f_{iml}) showed room for improvement. Enhancing the IML isolation would require a more granular analysis of the leakage-induced pressure drop characteristics.

Strategic Data Utilization and Synthetic Scenarios

To overcome the inherent data scarcity of the benchmark, the proposed methodology leveraged cross-fault training. By feeding diagnostic nodes with data containing faults from which they are structurally decoupled (e.g., exposing MSO 0 to f_{waf} anomalies), the models were trained to actively ignore overlapping disturbances. Furthermore, there is significant potential in artificially synthesizing new, physically plausible fault scenarios.

The Role of Warmup Periods

The implementation of a 90-second (1800 samples) initialization and warmup period was essential for handling covariate shift caused by environmental factors. This period allows the system to calibrate its baseline residuals and adapt its thresholds to the current operating environment. However, this approach introduces a critical diagnostic risk which is fault masking. If a fault–particularly a subtle sensor bias–is present from the moment the engine starts, the adaptation logic might erroneously incorporate this anomaly into its healthy baseline calibration. Furthermore, during this 90-second window, the system's sensitivity is intentionally reduced, creating a temporal blind spot where critical failures could go undetected.

By identifying these bottlenecks, we provide a clear roadmap for future iterations of the system, focusing on the refinement of the feature-driven diagnostic nodes and the optimization of the isolation logic to handle even more subtle fault manifestations.

References

1 J-Ph Cassar and M Staroswiecki. A structural approach for the design of failure detection and identification systems. *IFAC Proceedings Volumes*, 30(6):841–846, 1997.

2 Austin Coursey, Abel Diaz-Gonzalez, Marcos Quinones-Grueiro, and Gautam Biswas. Data-Driven Fault Detection and Isolation Enhanced with System Structural Relationships. In Marcos Quinones-Grueiro, Gautam Biswas, and Ingo Pill, editors, *36th International Conference on Principles of Diagnosis and Resilient Systems (DX 2025)*, volume 136 of *Open Access Series in Informatics (OASIcs)*, pages 15:1–15:17, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/OASIcs.DX.2025.15.

3 Miguel Cubero, Luis Ignacio Jiménez, Daniel López, Belarmino Pulido, and Carlos Alonso-González. Towards Predictive Maintenance in an Aluminum Die-Casting Process Using Deep Learning Clustering and Dimensionality Reduction. In Marcos Quinones-Grueiro, Gautam Biswas, and Ingo Pill, editors, *36th International Conference on Principles of Diagnosis and Resilient Systems (DX 2025)*, volume 136 of *Open Access Series in Informatics (OASIcs)*, pages 6:1–6:16, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/OASIcs.DX.2025.6`.

4 Lars Eriksson. Modeling and control of turbocharged si and di engines. *Oil & Gas Science and Technology-revue de l'Institut Francais Du Petrole*, 62:523–538, 2007. URL: `https://api.semanticscholar.org/CorpusID:53476253`.

5 Erik Frisk, Mattias Krysander, and Daniel Jung. A toolbox for analysis and design of model based diagnosis systems for large scale models. *IFAC-PapersOnLine*, 50(1):3287–3293, 2017. 20th IFAC World Congress. `doi:10.1016/j.ifacol.2017.08.504`.

6 Maxence Glotin, Louise Travé-Massuyès, and Elodie Chanthery. MSO Sets and MTES for Dummies. In Ingo Pill, Avraham Natan, and Franz Wotawa, editors, *35th International Conference on Principles of Diagnosis and Resilient Systems (DX 2024)*, volume 125 of *Open Access Series in Informatics (OASIcs)*, pages 13:1–13:15, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/OASIcs.DX.2024.13`.

7 Yuxin Huang, Austin Coursey, Marcos Quinones-Grueiro, and Gautam Biswas. Time-series few shot anomaly detection for hvac systems. *IFAC-PapersOnLine*, 58(4):426–431, 2024. 12th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2024. `doi:10.1016/j.ifacol.2024.07.255`.

8 Daniel Jung. Automated design of grey-box recurrent neural networks for fault diagnosis using structural models and causal information. In Roya Firoozi, Negar Mehr, Esen Yel, Rika Antonova, Jeannette Bohg, Mac Schwager, and Mykel Kochenderfer, editors, *Proceedings of The 4th Annual Learning for Dynamics and Control Conference*, volume 168 of *Proceedings of Machine Learning Research*, pages 8–20. PMLR, 23–24 June 2022. URL: `https://proceedings.mlr.press/v168/jung22a.html`.

9 Daniel Jung, Erik Frisk, and Mattias Krysander. The liu-ice benchmark – an industrial fault diagnosis case study. *arXiv preprint arXiv:2408.13269*, 2024. `doi:10.48550/arXiv.2408.13269`.

10 Daniel Jung, Erik Frisk, Mattias Krysander, Anna Sztyber-Betley, Francesco Corrini, Andrea Arici, Nicolas Anselmi, Mirko Mazzoleni, Jiamin Xu, Siwen Mo, Zixuan Xu, Chongpan Yang, Zhile Du, Hossein Safaeipour, Mehdi Forouzanfar, Vahid Mirahi, Anna Pinnarelli, Vicenç Puig, Qiao Deng, Yufei Liu, Jiakun Liu, Haobin Ke, Wanting Zhu, Silke Merkelbach, Maryam Ahang, and Homayoun Najjaran. A fault diagnosis benchmark of technical systems with incomplete data — six solutions. *Control Engineering Practice*, 164:106427, 2025. `doi:10.1016/j.conengprac.2025.106427`.

11 Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015. `arXiv:1412.6980`.

12 Mattias Krysander, Jan Åslund, and Mattias Nyberg. An efficient algorithm for finding minimal overconstrained subsystems for model-based diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 38(1):197–206, 2008. `doi:10.1109/TSMCA.2007.909555`.

13 T. Kurtoglu, S. Narasimhan, S. Poll, D. Garcia, L. Kuhn, J. de Kleer, A. van Gemund, and A. Feldman. First international diagnosis competition-DXC'09. *Proc. Int. Workshop on Principles of Diagnosis DX*, 9:383–396, 2009.

14 Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: an imperative style, high-performance deep learning library. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 8024–8035, Red Hook, NY, USA, 2019. Curran Associates Inc.

15 Ingo Pill, Daniel Jung, Eldin Kurudzija, Anna Sztyber-Betley, Michał Syfert, Kai Dresia, Günther Waxenegger-Wilfing, and Johan de Kleer. The DX Competition 2025 and Its Benchmarks. In Marcos Quinones-Grueiro, Gautam Biswas, and Ingo Pill, editors, *36th International Conference on Principles of Diagnosis and Resilient Systems (DX 2025)*, volume 136 of *Open Access Series in Informatics (OASIcs)*, pages 14:1–14:19, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/OASIcs.DX.2025.14`.

16 Kélian Poujade, Louise Travé-Massuyès, Jérémy Pirard, and Laure Vieillevigne. Unsupervised Multimodal Learning for Fault Diagnosis and Prognosis - Application to Radiotherapy Systems. In Marcos Quinones-Grueiro, Gautam Biswas, and Ingo Pill, editors, *36th International Conference on Principles of Diagnosis and Resilient Systems (DX 2025)*, volume 136 of *Open Access Series in Informatics (OASIcs)*, pages 16:1–16:17, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/OASIcs.DX.2025.16`.

17 M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. `doi:10.1016/j.jcp.2018.10.045`.

18 Henrik Sebastian Steude, Lukas Moddemann, Alexander Diedrich, Jonas Ehrhardt, and Oliver Niggemann. Diagnosis driven anomaly detection for cyber-physical systems∗∗this research paper is funded by dtec.bw and the european union. *IFAC-PapersOnLine*, 58(4):13–18, 2024. 12th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2024. `doi:10.1016/j.ifacol.2024.07.186`.

19 Ronald J. Williams and Jing Peng. An efficient gradient-based algorithm for on-line training of recurrent network trajectories. *Neural Computation*, 2(4):490–501, 1990. `doi:10.1162/neco.1990.2.4.490`.

20 A. Youssef, H. Noura, A. El Amrani, E. El Adel, and M. Ouladsine. A survey on data-driven fault diagnostic techniques for marine diesel engines. *IFAC-PapersOnLine*, 58(4):55–60, 2024. 12th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2024. `doi:10.1016/j.ifacol.2024.07.193`.