

УДК 004.03

DOI: 10.31653/1819-3293-2026-1-31-182-207

ARTICLE HISTORY

Received 11.03.2026

Accepted 06.04.2026

Vychuzhanin Volodymyr¹, Vychuzhanin Oleksii²
Odesa Polytechnic National University, Odessa, Ukraine
v.v.vychuzhanin@op.edu.ua¹, vychuzhanin.o.v@op.edu.ua²

Hybrid spatio-temporal neural network for predicting the risk of failures in ship power plant subsystems

Вичужанін Володимир Вікторович¹, Вичужанін Олексій
Володимирович²

Національний університет «Одеська політехніка», Одеса, Україна
v.v.vychuzhanin@op.edu.ua¹, vychuzhanin.o.v@op.edu.ua²

Гібридна просторово-часова нейромережа для прогнозування ризику відмов підсистем суднової енергетичної установки

Abstract. *This presents a hybrid spatio-temporal neural network for predicting failure risks in ship power plant subsystems within a digital twin framework. The proposed convolutional neural network and long short-term memory architecture integrates telemetry data, physics-informed constraints, and fuzzy diagnostic indicators to ensure consistent and reliable predictions. Experimental validation, including Hardware-in-the-Loop testing, confirms improved accuracy and stability under noisy conditions. The model supports real-time deployment and enhances interpretability through subsystem-level analysis and SHapley Additive exPlanations.*

Резюме. *Представлено гібридну просторово-часову нейронну мережу для прогнозування ризику відмов підсистем суднової енергетичної установки в межах цифрового двійника. Запропонована архітектура на основі згорткових нейронних мереж і мереж довгої короткочасної пам'яті інтегрує телеметричні дані, фізично обґрунтовані обмеження та нечіткі діагностичні ознаки. Експериментальна валідація підтверджує підвищену точність і стійкість моделі. Інтерпретованість забезпечується за допомогою методу SHapley Additive exPlanations.*

The increasing complexity of modern ship power plants (SPPs) and the nonlinear nature of their degradation complicate long-term risk prediction [1, 2]. Traditional analytical and fuzzy-based methods, while effective for local diagnostics, are limited in high-dimensional, time-varying environments [3, 4]. Neural network models trained on telemetry and digital twin (DT) data enable identification of latent dependencies between system states and failure probability, forming the basis of intelligent monitoring systems [5, 6]. Recent advances in prognostics and health management (PHM) highlight the effectiveness of deep learning, particularly hybrid CNN (Convolutional Neural Network)–LSTM (Long Short-Term Memory) architectures, in capturing spatial and temporal features [7]. However, purely data-driven approaches often lack physical interpretability and stability. Physics-Informed Neural Networks (PINNs) address this by embedding conservation laws into training, improving robustness and generalization under noisy or limited data [8], though real-time application remains challenging. Integration of DTs with machine learning enhances prediction accuracy by combining simulation and real-time data [9, 10]. Multi-source data fusion improves diagnostics [11], but consistency across heterogeneous domains remains an issue. Fuzzy logic supports uncertainty modeling [12], yet requires integration with deep learning in complex systems. Hybrid CNN–LSTM and multi-branch architectures improve feature extraction and interpretability [13, 14], while combining statistical reliability models with machine learning increases stability, especially for rare failures [15, 16]. Despite these advances, key gaps remain: limited integration of data-driven and physics-based models, lack of unified multi-source representation, and insufficient application in real-time maritime DT systems [17, 18]. The proposed approach addresses these gaps through a hybrid framework for SPP failure prediction, ensuring physical validity and cognitive consistency. It integrates telemetry, CSM (cognitive simulation model)-based indicators, and fuzzy diagnostics into a unified representation, and employs a multi-branch CNN–LSTM architecture for subsystem-specific modeling (fuel, cooling, electrical). A physics-informed loss function enforces conservation laws, improving consistency and generalization, while coordinated integration within a DT aligns neural, simulation, and fuzzy outputs.

The model is validated in a Hardware-in-the-Loop (HIL) environment, confirming robustness under realistic conditions. The study aims to develop an accurate and consistent predictive model, including data integration, architecture design, physics-informed training, and validation with robustness analysis under noise and ablation scenarios.

Within the DT architecture, the neural network subsystem operates at the top analytical layer as an intelligent integrator of prognostic risk estimates

for transmission to the Shore Control Centre (SCC). By fusing heterogeneous data into a unified cognitive model, it supports remote monitoring and informed decision-making. The subsystem interacts with three domains: the CSM (causal failure probabilities $P_i(t)$), the fuzzy module (aggregated risks $R_i(t)$), and the data acquisition layer (sensor and telemetry data). Its role within the control loop and data flows is formalized in the system architecture (Fig. 1), which describes the end-to-end transformation from raw signals to onshore management decisions.

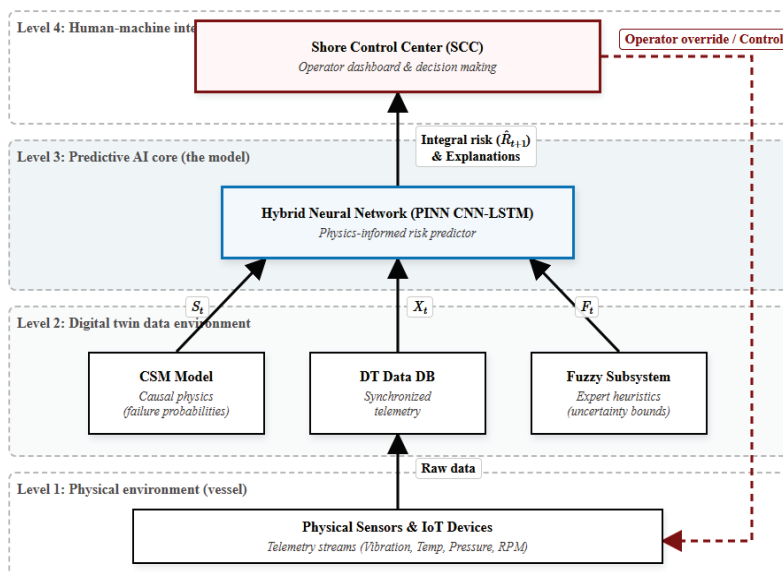


Fig. 1. Global system architecture of the cognitive DT: an echeloned data flow from physical sensors to the SCC

The end-to-end information flow (Fig. 1) is organized into four hierarchical levels: the physical layer (sensors and actuators), where IoT devices generate raw telemetry; the DT level, where data are enriched, while the CSM computes causal probabilities and the fuzzy subsystem formalizes expert knowledge; the artificial intelligence level, where a hybrid CNN-LSTM model integrates these inputs via a physics-informed loss function (L_{PDE}) to estimate the integral risk; and the decision-making level (SCC), where the validated forecast is transmitted via satellite for operational decisions.

The neural network acts as an intelligent bottleneck, filtering large volumes of raw data and delivering only interpretable, physically consistent risk estimates. The model performs one-step-ahead prediction of integral failure risk while accounting for multidimensional SPP dynamics.

$$\hat{R}_{t+1} = f_{\theta}(X_{t-k:t}) \in [0,1],$$

where f_{θ} is a parameterised neural network, $X_{t-k:t}$ is a time window of telemetry data with a depth of k steps, $\hat{R}_{t+1}$ is the predicted value of the integral failure risk at time step $t+1$, θ is a non-linear approximator (neural network) with a set of trainable weight coefficients θ , $(X_{t-k:t}) \in R^{k \times M}$ is the input feature matrix representing a time window of observations with a depth of k steps, k is the length of the observation window (a model hyperparameter), M is the number of input physical and diagnostic parameters.

The integration of the neural network with other DT modules, including the CSM and the fuzzy subsystem, ensures cognitive consistency of the forecast. The technical condition indicator S_t (from the CSM) and diagnostic features F_t (from the fuzzy subsystem) provide expert knowledge, enriching raw sensor data, accelerating convergence, and improving decision-making under uncertainty. The model captures time-varying dependencies, where the observation window k is a key hyperparameter. In this study, a window of $k = 60$ seconds was selected to balance short-term dynamics (e.g., vibration spikes) and computational efficiency. The input feature vector X_t integrates three data sources: OREDA databases (statistical failure parameters, e.g., Weibull, Lognormal), SPP DT data (pressure, temperature, vibration, rotation speed, consumption), and online telemetry from IoT sensors and diagnostic modules. The combined input ensures cognitive consistency and forms the basis for robust prediction. Figure 1 presents the three-level aggregation scheme used to construct X_t , with the overall vector defined for each subsystem accordingly:

$$x_t = [P_t, T_t, V_t, \omega_t, S_t, Q_t, F_t], \quad (1)$$

where: P_t , T_t is the pressure and temperature of working media at critical nodes of fuel, lubrication, and cooling circuits; V_t is the vibroacoustic parameters (RMS accelerations) of bearing units; ω_t is the rotational speed of the main engine shaft or turbocharger; S_t is a scalar indicator of technical condition from the CSM model; Q_t is the instantaneous fuel consumption characterising the current load regime; F_t is a fuzzy indicator of diagnostic uncertainty.

The detailed semantics and frequency characteristics of the vector components are presented in Table 1.

Table 1

Semantic structure of the input feature vector X_t

Component	Physical meaning of parameter	Data domain	Update frequency, Hz
P_t	Pressure of working media (oil, fuel, boost)	Telemetry	High (10–50)
T_t	Temperature gradients in power plant circuits	Telemetry	Medium (1–5)
V_t	Vibroacoustic activity (RMS acceleration)	Telemetry	High (100–1000)
ω_t	Rotational speed of shaft and turbochargers	Telemetry	Medium (10)
Q_t	Fuel / coolant flow rate (load condition)	Telemetry	Medium (1)
S_t	State indicator based on the CSM model ($S \in [0,1]$)	Causal (physics-based)	Low (event-driven)
F_t	Diagnostic feature of the fuzzy subsystem	Expert-based (logic)	Low (on demand)

This multidimensional feature vector structure ensures accurate modeling of complex SPP degradation processes. It correlates low-level stochastic sensor fluctuations with high-level structural health indicators, providing the SCC operator with a comprehensive and interpretable assessment of the asset's technical condition. To preserve the $k \times M$ matrix structure required for the operation of 1D-CNN layers, low-frequency expert features (S_t , F_t) are broadcast to all steps of the current observation window k . This ensures that the neural network takes into account a stable physical context when analysing high-frequency oscillations of pressure or vibration. The overall input vector X_t is a fusion of raw sensor data and synthesized diagnostic metrics, defined by Equation (1). For CNN–LSTM processing, its temporal structure is explicitly represented, with the combined input window $X_{t-k:t}$ forming a $k \times M$ matrix (e.g., 60×7), where $k=60$ s. High-frequency features $P_t, T_t, V_t, \omega_t, Q_t$ are updated at each time step Δt , whereas synthesized features are low-frequency. The technical condition indicator S_t , derived from the CSM, represents the subsystem's aggregated health (e.g., $S_t \in [0,1]$ or a binary state) and is updated at a coarser rate. Diagnostic features F_t , obtained from the fuzzy subsystem, quantify uncertainty and local risk and

are also updated infrequently. To preserve the required matrix structure for the 1D-CNN layer, S_t and F_t are broadcast across the entire observation window $X_{t-k:t}$, remaining constant for all k time steps and taken from the latest available low-frequency update prior to time t .

$$Feature(t') = \begin{cases} Sensor\ Reading(t') & \text{for } PTV, \omega, Q \\ S_{latest} & \text{for } S \text{ where } t - k \leq t' \leq t \\ F_{latest} & \text{for } F \text{ where } t - k \leq t' \leq t \end{cases},$$

where t' is the current step index within the time window; S_{latest} , F_{latest} is the latest available values of causal and fuzzy features up to time t .

This multidimensional structure enables mapping low-level stochastic sensor fluctuations to high-level structural health indicators (Fig. 2). The model's prognostic capability relies on cognitive consistency via integration of multi-domain features. The input vector X unifies process dynamics, causal physical knowledge, and expert uncertainty through a three-level aggregation scheme (Fig. 2).

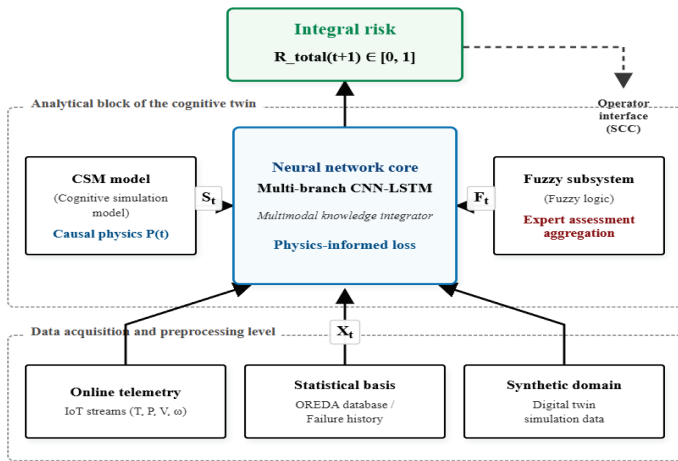


Fig. 2. Multi-level architecture of intelligent monitoring: integration of the neural network core with the functional domains of the DT

Fig. 2 shows that X_t concatenates features from dynamic (time-series telemetry), causal (CSM physics-based indicators), and fuzzy (uncertainty and local risks) domains. This links low-level sensor data with high-level health indicators, providing SCC operators with a comprehensive system view. The approach combines high-frequency temporal dynamics with low-

frequency expert knowledge. Preprocessing includes normalization, spline-based imputation of missing values, and median filtering to remove outliers. For each window of length $k=60$ s, a training tensor $X_{t-k:t}$ and R_t is formed. To capture complex spatio-temporal degradation, a hybrid CNN-LSTM architecture [13] is used, combining 1D-CNN for feature extraction and LSTM for long-term temporal dependencies. To better represent subsystem-specific degradation, the model is implemented as a multi-branch parallel structure (Fig. 3).

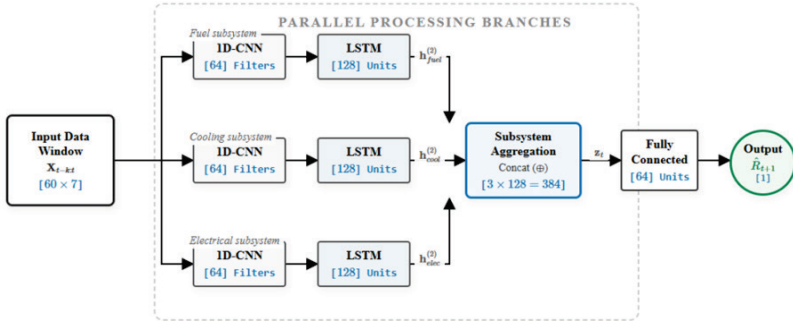


Fig. 3. Multi-branch CNN-LSTM architecture for SPP risk prediction: parallel feature extraction and subsystem aggregation

To prevent statistical interference between heterogeneous physical parameters, the predictive core employs a multi-branch CNN-LSTM topology (Fig. 3). The input telemetry window $X_{t-k:t}$ (60×7 matrix) is directly segmented into three domain-specific sub-matrices $X_{t,fuel}$, $X_{t,cool}$, and $X_{t,elec}$, enabling isolated spatio-temporal feature extraction for each major SPP subsystem.

Segmentation rule: $X_{t,fuel}$: includes parameters directly related to fuel system operation (e.g., fuel pressure P_{fuel} , fuel temperature T_{fuel} , and fuel flow rate Q_t); $X_{t,cool}$: includes parameters related to the cooling circuit (e.g., coolant temperature T_{cool} , coolant pressure P_{cool}); $X_{t,elec}$: includes parameters reflecting mechanical and electrical load/condition (e.g., vibration V_t , rotational speed ω_t , and auxiliary metrics). Each of these segmented sub-matrices is fed into its own independent, yet structurally identical CNN-LSTM branch (as defined by layers 1 and 2). This parallel processing ensures the dedicated extraction of local patterns and the robust modeling of long-term temporal dependencies specific to each physical domain. The resulting hidden state vectors ($h_{t,fuel}^{(2)}$, $h_{t,cool}^{(2)}$, $h_{t,elec}^{(2)}$) from the parallel LSTM blocks are subsequently concatenated in the subsystem aggregation layer to form

the unified system state vector z_t , as formally defined in Equation (2). This integrated approach guarantees the cognitive consistency of the forecast by synthesizing knowledge extracted from disparate physical domains of the complex technical system. The model includes four functional layers:

1. Feature extraction layer (1D-CNN).

The one-dimensional convolutional layer performs the extraction of local spatial patterns from the multi-domain input:

$$h_t^{(1)} = \text{ReLU}(W_c \cdot X_t + b_c),$$

where $h_t^{(1)}$ is the hidden state vector after the convolutional layer; W_c is the matrix of convolutional filters; X_t is the input observation vector; b_c is the bias term; ReLU is the activation function (rectified linear unit).

2. Temporal encoding layer (LSTM).

The recurrent layer based on Long Short-Term Memory (LSTM) blocks models the inertia of processes

$$h_t^{(2)} = \text{LSTM}(h_t^{(1)}, c_{t-1}).$$

where LSTM is the memory cell and forget gate mechanisms to account for the inertial properties of processes in the subsystems of the SPP (e.g., thermal and vibration effects); $h_t^{(2)}$ is the hidden state vector of the recurrent layer; c_{t-1} is the memory cell state at the previous step.

3. Subsystem aggregation layer.

The aggregation layer combines the data from parallel branches into a unified system state vector z_t :

$$z_t = \text{Concat}(h_{t,fuel}^{(2)}, h_{t,cool}^{(2)}, h_{t,elec}^{(2)}), \quad (2)$$

where $h_{t,fuel}^{(2)}, h_{t,cool}^{(2)}, h_{t,elec}^{(2)}$ - hidden state vectors obtained at the second level (LSTM layer) for different subsystems of the SPP: $h_{t,fuel}^{(2)}$ - fuel system; $h_{t,cool}^{(2)}$ - cooling system; $h_{t,elec}^{(2)}$ - power supply system.

This layer reflects the implicit interaction between subsystems by assuming that the concurrently calculated hidden states from each subsystem (fuel, cooling, electric) independently contribute to the overall system risk. This concatenation step is crucial as it ensures a comprehensive, cognitively consistent representation of the multi-domain system state before the final prediction.

4. Fully connected output layer.

Generates the forecast of the integral failure risk:

$$\hat{R}_{t+1} = \sigma(W_o \cdot z_t + b_o),$$

where $\sigma(\cdot)$ is the sigmoid activation function that normalizes the result within the range $[0,1]$; W_o, b_o is the weights and bias of the output fully connected layer.

The resulting value is $\hat{R}_{t+1}$ interpreted as the probability of the system transitioning into a degradation state at the next time step. The summary structure of the architecture is presented in Table 2.

Table 2

Structural specification of the layers of the hybrid neural network (Multi-Branch CNN-LSTM)

Layer type	Parameters / number of neurons	Activation function	Output	Output dimension
Input layer	Sliding window (k×M)	-	X_t	60×7
Convolutional (1D-CNN)	64 filters (3×1)	ReLU	$h_t^{(1)}$	64 (per branch)
LSTM memory cells	128 units	tanh / sigmoid	$h_t^{(2)}$	128 (per branch)
Dropout	rate = 0.2	-	-	128 (per branch)
Subsystem aggregation	Branch concatenation	-	z_t	3×128 = 384
Fully connected	64 neurons	ReLU	-	64
Output layer	1 neuron	Sigmoid	$\hat{R}_{t+1}$	1

The architecture (Table 2) is designed for real-time processing of streaming data. The initial 1D convolutional layer (64 filters) acts as a high-frequency spatial filter, extracting short-term mechanical anomalies directly from raw sensor inputs without prior preprocessing. The subsequent LSTM layer (128 units) captures long-term thermodynamic degradation trends. To reduce overfitting to stochastic marine noise, Dropout regularization (rate = 0.2) is applied in each branch. The subsystem aggregation layer then concatenates outputs from the fuel, cooling, and electrical branches into a unified 384-dimensional latent vector, enabling the fully connected network

to perform cross-domain feature fusion and compute the integral risk indicator.

To ensure a high level of confidence of the SCC operator in the generated predictions, model training is implemented within the paradigm of physics-informed learning. The final loss function is minimised iteratively: L_{data} is the standard prediction error based on observed data (for example, mean squared error (MSE) between the predicted and actual risk values); L_{phys} is the penalty for violation of physical dependencies derived from thermodynamic, hydraulic, and vibration models of the digital twin. The final loss function is expressed as [8]:

$$L_{total} = L_{data} + \lambda \cdot L_{phys},$$

where λ is the regularization coefficient that determines the degree of influence of physical constraints on the training process.

The regularization coefficient λ was selected as $\lambda = 0.5$ using the L-curve method and further tuned via grid search to balance prediction accuracy (L_{data}) and physical consistency (L_{phys}). Minimizing the total loss L_{total} enforces agreement with physical laws and prevents non-physical predictions. This integration enables training on both empirical data and structural dependencies within the system. The model is trained on a combined dataset including SPP digital twin time series (pressure, temperature, vibration, rotational speed, generator load, etc.), historical failure data (e.g., OREDA), and synthetic CSM data describing probabilistic-causal relationships between subsystems.

Such integration of data sources provides comprehensive model training - both from empirical operational data and from formalized structural dependencies among a complex technical system elements. The main component L_{data} is calculated as the MSE:

$$L_{data} = \frac{1}{N} \sum_{i=1}^N (R_i - R_i^{true})^2,$$

where N is the number of samples in the batch; R_i^{true} is the reference label of failure risk from the training dataset.

The Physics-Informed Loss (L_{phys}) regularizes training by enforcing SPP conservation laws and penalizing deviations from DT-based thermodynamic, hydraulic, and vibration constraints, including conservation of energy, mass, and momentum (heat balance, flow, pressure, and rotational dynamics).

Formal Definition of the Physical Loss Term

The physical loss term, L_{phys} is defined as the MSE of the residuals, R_i , of the governing physical equations. The residual R_i is the error that results

when the predicted parameters and their derived temporal derivatives (calculated from the neural network's input window) are substituted into the physical conservation equations (where the result should ideally be zero). This loss function quantifies the deviation from the conservation laws formalized within the DT's CSM. The general form of the L_{phys} is:

$$L_{phys} = \frac{1}{N \cdot K_{phys}} \sum_{i=1}^N \sum_{i=1}^{K_{phys}} R_i (X_{t-k:t}, \hat{R}_{t+1}, \theta)^2 ,$$

where N is the total number of time steps (windows) in the training batch; K_{phys} is the total number of distinct physical equations/constraints applied ($i=1 \dots K_{phys}$); $R_i(\dots)=0$ represents the i -th differential or algebraic conservation equation; $X_{t-k:t}$ is the input observation window; $\hat{R}_{t+1}$ is the predicted risk.

As an academic example, consider the simplified heat balance equation for the cooling circuit (Subsystem j), which is one component of the total L_{phys} :

$$R_{heat} = \frac{dE}{dt} - (\dot{\theta}_{in} - \dot{\theta}_{out}) = 0.$$

where dE/dt is the rate of change of internal energy (approximated from $\nabla_t \cdot T_j$ and P_j); $\dot{\theta}_{in} - \dot{\theta}_{out}$ is the net heat flow (derived from operational parameters θ_t and physical model φ).

The corresponding residual term R_j to be minimized by the neural network weights θ would thus be:

$$R_j \equiv (\nabla_t \cdot T_j) + \frac{1}{c} (f_j^{thermal}(P_j, \theta_t) - T_j) = 0.$$

The neural network is enalized if its prediction leads to an overall system state where the measured temperature change ($\nabla_t \cdot T_j$) significantly deviates from the value predicted by the physical thermal model $f_j^{thermal}(P_j, \theta_t)$ (which accounts for fluid flow, heat exchange, etc., and is derived from the DT's CSM).

Since the input data consists of discrete time series, the temporal derivative (∇_t) of the physical parameters (like pressure P_j or temperature T_j) is approximated numerically using the forward difference scheme between two consecutive time steps (Δ_t):

$$\nabla_t P_j = \frac{P_j(t+\Delta t) - P_j(t)}{\Delta t} ,$$

where $\nabla_t P_j$ is the estimate of the time derivative of pressure P with respect to time t ; Δt is the telemetry discretisation step.

$$T_j = \frac{T_t(t + \Delta t) - T_t(t)}{\Delta t}.$$

This numerical method provides the empirical dynamic behavior of the system, which is then enforced to align with the theoretical behavior f_j^{energy} during the model's training. The final total loss $L_{\text{total}} = L_{\text{data}} + \lambda \cdot L_{\text{phys}}$ is minimized iteratively to achieve a balance between empirical prediction accuracy (L_{data}) and physical consistency (L_{phys}). The L_{phys} component enforces dynamic consistency between observed physical changes and the modeled term f_j^{energy} , derived from governing equations (e.g., thermal and energy balance). The model's construction and training are formalized as a data flow diagram, with digital twin constraints integrated into the backpropagation loop (Fig. 4).

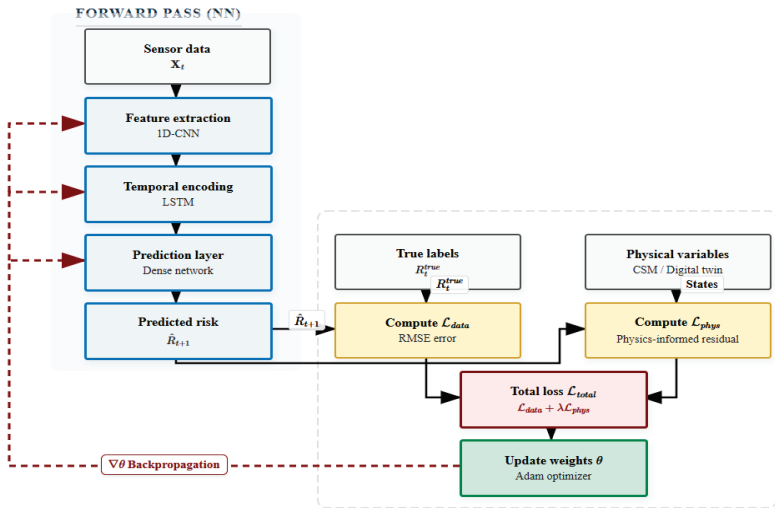


Fig. 4. Data flow diagram and backpropagation mechanism in the process of training the hybrid model

The scheme (Fig. 4) shows two coupled optimisation loops: the main forward flow from sensors to the predictive layer and an auxiliary branch of a priori constraints derived from the digital twin. Training iteratively minimises the total loss L_{total} , where L_{data} fits observations and L_{phys} serves as a strict regulariser. This feedback enforces agreement with noisy data while constraining predictions within engineering limits, preventing physically impossible results — critical for autonomous navigation safety.

Training of the PINN combines data-driven optimisation with physical constraints. Parameters φ and θ are updated adaptively as operating conditions change. The model is trained with mini-batches using Adam ($\alpha \approx 10^{-3}$), and its generalisation is assessed via 5-fold cross-validation. The physics-informed loss ensures consistency with conservation laws, improving reliability under incomplete data and robustness to sensor noise. Thus, the CNN–LSTM model functions as an intelligent predictor within the digital twin, uniting empirical patterns with physical laws. For validation under near-real conditions, a dedicated Hardware-in-the-Loop (HIL) test bench was developed (Fig. 5).

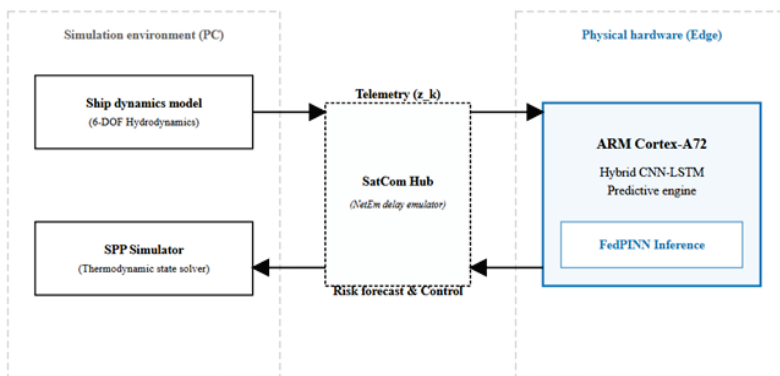


Fig. 5. Semi-physical (HIL) test ben diagram: integration of an ARM Cortex onboard computer with an SPP physical process simulator

Using a real ARM Cortex-A72 (4×1.5 GHz) in the loop enables objective latency assessment. With satellite communications (SatCom) degradation emulation, the model shows robustness to packet loss: inference time is 42 ± 5 ms per prediction, with peak RAM under 128 MB, meeting Hard Real-Time edge constraints. The PINN architecture is highly scalable within the SCC ecosystem: its multi-branch design enables vertical expansion without restructuring, supports cost-efficient horizontal scaling for fleet monitoring, and reduces SatCom traffic by transmitting only the aggregated risk vector (over 98 %).

A fixed observation window of 60 s (Table 3) was selected to balance capturing LSTM temporal dependencies with computational efficiency. The regularisation coefficient λ is a key parameter defining the trade-off between L_{data} and L_{phys} ; its optimal value was determined via grid search over $[0, 1]$ with 5-fold cross-validation.

Table 3

Technical parameters for deployment of the predictive model

Deployment parameter	Value	Comment
Response time (latency)	< 50 ms	Meets hard real-time requirements
Required memory (RAM)	≈ 120 MB	Acceptable for edge computing (ARM/Jetson)
Throughput	10 Hz	Synchronized with main sensors of the power plant
SatCom traffic savings	> 98 %	Due to onboard edge data processing

The hybrid CNN–LSTM model’s efficiency and stability depend on hyperparameters controlling learning, representation, and complexity. For reproducibility, the configuration (Table 4) was obtained using the same procedure; the regularization coefficient remains the most critical parameter and requires further analysis.

Table 4

Hybrid CNN-LSTM model hyperparameter configuration

Parameter	Value	Description
Observation window length (k)	60 s	Length of the time series segment used for prediction.
Input feature count (M)	7	The dimensionality of the input vector X_t per time step.
Batch size	32	Number of samples processed before updating the model weights.
Optimizer	Adam	Adaptive moment estimation algorithm.
Parameter	Value	Description
Learning rate (α)	10^{-3}	Adaptive learning rate for weight updates.
Epochs	100	Total number of full passes over the training dataset (with early stopping).
Regularization coefficient (λ)	0.5 (empirical)	Weighting factor balancing L_{data} and L_{phys}
Cross-validation scheme	5-fold	Scheme used to ensure generalization capability and prevent overfitting.

A deeper analysis is required to validate the critical trade-off parameters: the observation window length (k) and the regularization coefficient (λ). The selection methodology for the regularization coefficient is illustrated in Figure 6. As shown by the L-curve (Fig. 6), increasing λ from 0 to 0.5 sharply reduces L_{phys} while maintaining low RMSE; $\lambda = 0.5$ provides the optimal balance. For $\lambda > 0.5$, RMSE rises, indicating overemphasis on physical constraints at the expense of sensor-driven dynamics. The time window $k = 60$ s reflects SPP inertia and external disturbances: with load oscillations of 8–15 s, it captures 4–7 cycles, allowing the LSTM to separate short-term fluctuations from long-term degradation.

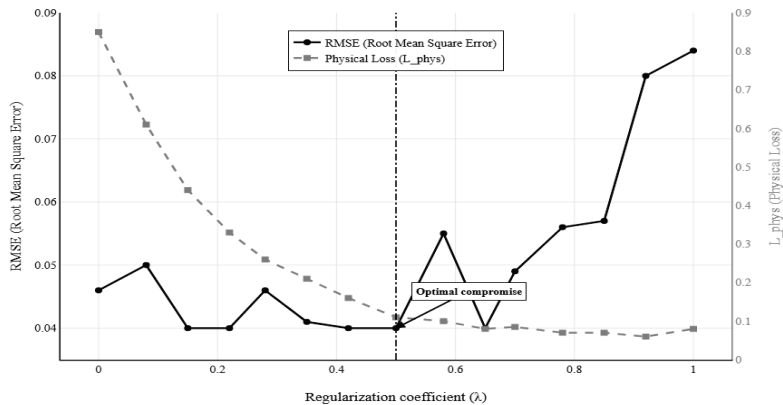


Fig. 6. Trade-off between empirical accuracy (RMSE) and physical consistency (L_{phys}) as a function of the regularization coefficient λ

The optimal regularisation coefficient $\lambda = 0.5$ corresponds to the L-curve “knee”: for $\lambda < 0.3$ the model overfits to sensor noise, whereas for $\lambda > 0.7$ physical constraints dominate, suppressing empirical features and increasing systematic error. To assess the model’s quality, standard regression metrics are applied: MAE; root mean square error (RMSE); coefficient of determination (R^2), as well as the k -fold cross-validation procedure, which guarantees model robustness against overfitting.

The choice $\lambda = 0.5$ follows the L-curve criterion, where gains in physical consistency no longer justify losses in empirical accuracy. Sensitivity of MAE to k and λ was also analysed (Fig. 7).

Analysis of Fig. 7 shows that reducing the window to 30 s increases error by 18 – 22 % due to aliasing of slow thermal dynamics, while extending it to 120 s yields negligible accuracy gains (<1.5 %) but sharply increases

random access memory (RAM) usage and introduces unacceptable real-time delays.

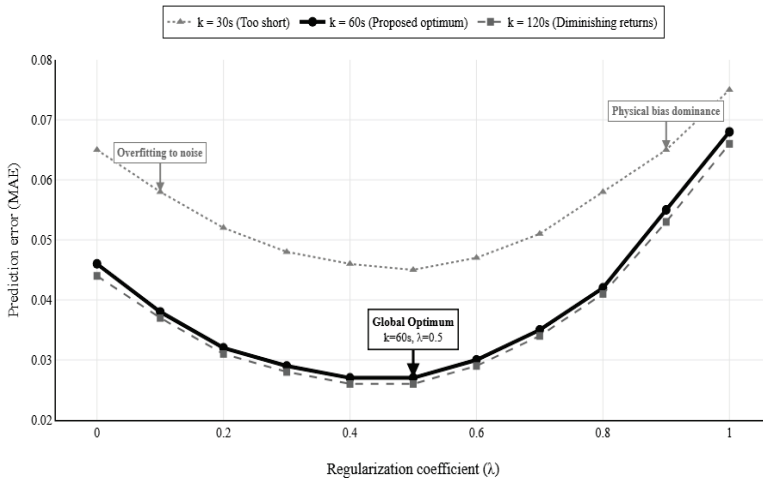


Fig. 7. Sensitivity analysis of the predictive model: the influence of the observation window length (k) and the regularisation coefficient (λ) on prediction accuracy (MAE)

Together, these elements form a trainable neural subsystem of the digital twin, providing adaptive prediction of failure probability and integral risk based on sensor data streams:

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{R}_i - R_i^{exp}|,$$

$$MRSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{R}_i - R_i^{exp})^2},$$

$$R^2 = 1 - \frac{\sum_i (R_i^{exp} - \hat{R}_i)^2}{\sum_i (R_i^{exp} - \bar{R})^2},$$

where R_i^{exp} is the actual (empirical) risk values; $\hat{R}_i$ is the predicted values produced by the model; $\bar{R}$ is the mean value of actual risks.

Before training, a statistical analysis of feature distributions using Kernel Density Estimation (KDE) was performed to verify data representativeness and reduce bias toward dominant classes. As shown in Fig. 8, data density is evaluated across four SPP subsystems: fuel, cooling, electric power, and control.

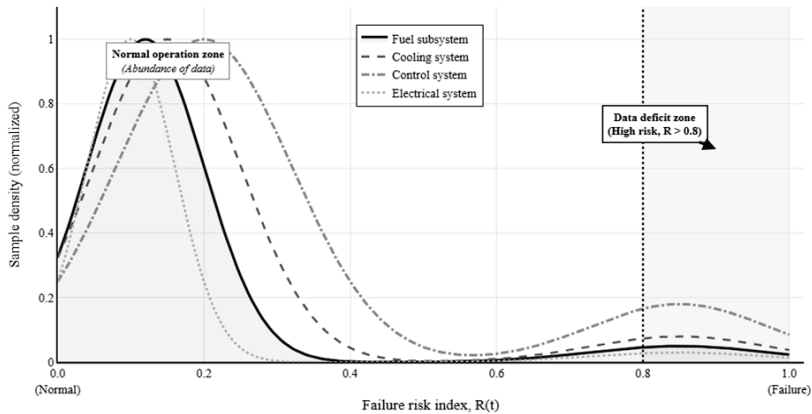


Fig. 8. Density distribution of the training dataset by integral risk levels for various SPP subsystems

Figure 8 shows data density across four subsystems (Fuel, Cool, Elec, Control). Most observations lie in the low-risk zone ($R < 0.5$), reflecting normal operation. Cooling and electrical systems dominate with medium-risk cases ($0.2 < R < 0.6$), while high-risk events ($R > 0.8$) account for only 5–7%, motivating data augmentation via CSM-generated scenarios. High-risk cases are mainly concentrated in the control subsystem, highlighting its critical role despite fewer parameters.

This distribution guides dataset balancing (e.g., oversampling, class weighting) to avoid bias toward dominant classes. Model adequacy was assessed by comparing predicted risks with DT-based reference values on an isolated test set containing rare emergency scenarios (fault detection and isolation attacks, cylinder–piston degradation), with results shown in Fig. 9. Analysis of the scatter diagram (Fig. 9) revealed a slight dispersion in the low-risk zone ($R < 0.2$), caused by hardware sensor noise, and a tendency towards slight overestimation in the ultra-high-risk zone ($R > 0.8$), which is acceptable from the standpoint of a conservative safety strategy (Safety First). Quantitative accuracy results for the key SPP subsystems are presented in Table 5.

Most predicted values lie close to the diagonal $y=x$, indicating strong agreement with reference data ($R^2 = 0.982$). Errors remain low ($\text{MAE} \leq 0.03$, $\text{RMSE} = 0.04$). Minor dispersion at low risk ($R < 0.2$) is due to sensor noise and weak degradation signals, while slight overestimation at high risk

($R > 0.7$) reflects limited training data for rare failures; this can be mitigated by denoising and data balancing (e.g., synthetic generation).

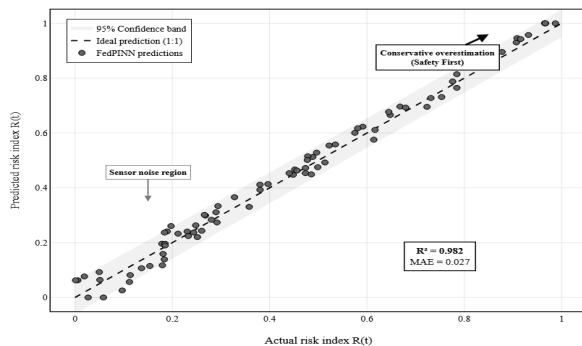


Fig. 9. Comparison of hybrid model predictions with reference DT data

Table 5

Accuracy metrics for failure risk prediction by subsystems

PPS	MAE	RMSE	Pearson correlation coefficient
Fuel system	0.024	0.038	0.985
Cooling circuit	0.031	0.042	0.978
Power unit (Generators)	0.028	0.035	0.981
Integral indicator	0.027	0.039	0.982

Benchmarking against Weibull, LSTM-only, and CNN-only models confirms the superiority of the hybrid PINN CNN–LSTM. Error reduction versus the best baseline (LSTM) is statistically significant ($p < 0.001$). The model matches or exceeds state-of-the-art results, benefiting from physics-informed constraints that improve accuracy under data scarcity. Unlike classical statistical models, it adapts to unit-specific degradation, while physical laws act as a filter that preserves accuracy in high-risk regimes where purely data-driven models degrade. A combined benchmark and ablation study across Weibull, LSTM-only, standard CNN–LSTM, and the proposed PINN CNN–LSTM confirms the contribution of each component. Summary results, including R^2 , are presented in Table 6.

Table 6
Comparative benchmark of predictive models on the test dataset

Predictive model architecture	MAE	RMSE	R ² Score
Weibull Model (Statistical baseline)	0.184	0.215	0.412
CNN-only (Spatial baseline)	0.078	0.092	0.815
LSTM-only (Temporal baseline)	0.062	0.075	0.894
Standard CNN-LSTM (Data-driven / No physics)	0.046	0.058	0.941
PINN CNN-LSTM (Proposed model)	0.027	0.039	0.982

Analysis of Table 6 confirms the theoretical hypotheses on architectural components, showing the hybrid PINN’s predictive reliability ($R^2 = 0.982$). The physics-informed approach is essential for Ship Power Plants (SPPs), which lack sufficient historical data on catastrophic failures. Purely data-driven networks (Standard CNN-LSTM) overfit to sensor noise, while PINN enforces thermodynamic constraints, reducing mean absolute error by 41% and preventing "mathematical hallucinations" under scarce data. Figure 10 illustrates these dynamics, comparing statistical baselines and the effect of the physical regularizer.

The graphical comparison (Fig. 10) shows that the proposed hybrid approach yields the lowest prediction error among all methods. The statistical Weibull model remains less accurate ($RMSE > 0.2$) due to its inability to capture environmental disturbances. Recurrent networks (LSTM) significantly reduce the error ($MAE = 0.062$), but remain sensitive to stochastic noise from SPP sensors. Ablation analysis (Fig. 10) indicates that removing the physical regularizer (i.e., transition from PINN to standard CNN-LSTM increases MAE by 41%. The combination of spatio-temporal filtering and physical inductive bias enables the PINN model to achieve the best performance ($RMSE = 0.039$), making it suitable as a core component of a digital twin.

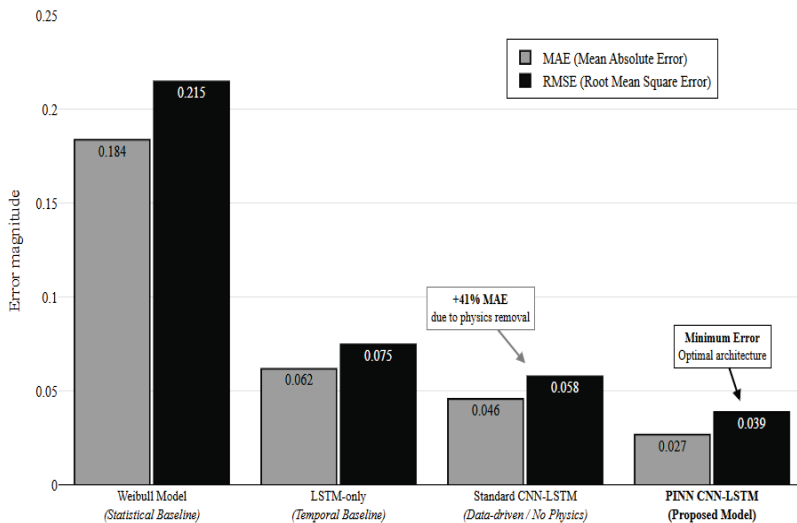


Fig. 10. Comparative benchmark and ablation analysis of predictive models: evaluation of accuracy (MAE and RMSE) and the influence of functional components

Beyond accuracy, the multi-branch architecture provides basic interpretability: separating feature flows allows not only failure probability estimation but also attribution of risk to specific subsystems (fuel, cooling, electrical) via contributions of hidden states. For deeper explainability, shapley additive explanations-based feature attribution can be applied. To illustrate the physics-informed core, convergence of the loss function was analyzed (Fig. 11).

In early training (epochs 0 – 20), the optimizer minimizes empirical loss L_{data} , while high volatility of L_{phys} reveals physically inconsistent intermediate states. This confirms that purely data-driven models may exploit statistical shortcuts without respecting conservation laws. In later stages (Pareto knee, Fig. 11B), competing gradients emerge: the physical regularizer acts as a structural constraint, penalizing infeasible representations and restricting the hypothesis space. Convergence at this point ensures consistency with both noisy sensor data and thermodynamic invariants, preventing statistical artifacts.

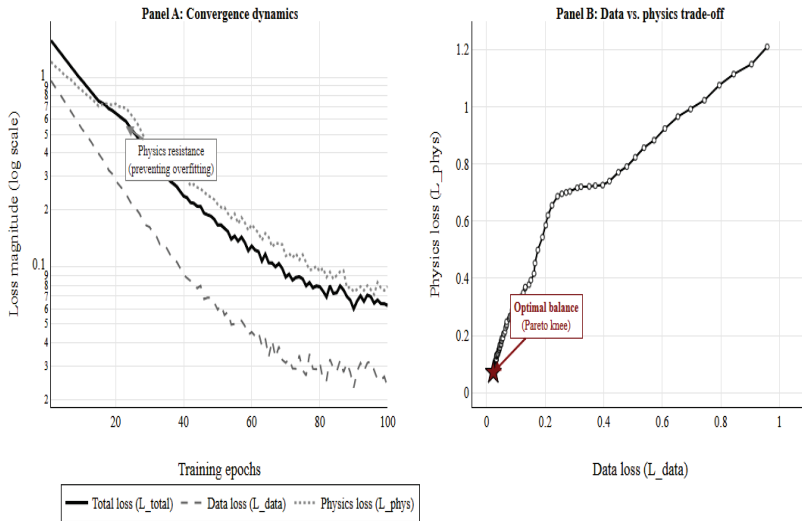


Fig. 11. Interaction of loss function components in the hybrid architecture: (A) dynamics of joint convergence; (B) analysis of the trade-off (Pareto trade-off) between data and physics

The study proposes a predictive architecture with several novel elements: a cognitive consistency method that integrates heterogeneous domains (CSM, fuzzy logic, sensor data), enabling the neural core to leverage prior physical knowledge for noise filtering; a multi-branch design for parallel feature extraction across key SPP subsystems (fuel, cooling, electromechanical), improving sensitivity to local faults; a real-time PINN with a modified loss enforcing mass and energy conservation to avoid physically inconsistent predictions; and Hardware-in-the-Loop validation on an ARM Cortex-A72, confirming robustness and real-time performance.

Limitations remain: despite ~ 42 ms inference, longer observation windows ($k > 120$ s) or added nonlinear constraints significantly increase memory use, potentially exceeding maritime programmable logic controller capabilities. The model depends on reliable boundary conditions, handling noise and partial drift but becoming uncertain under complete sensor failure. It is also focused on gradual degradation and cannot capture abrupt failures caused by external impacts.

Discussion

The results indicate that integrating data-driven models with physical constraints and multi-source data improves both accuracy and stability in failure risk prediction for complex systems. Hybrid CNN–LSTM models capture temporal dependencies and degradation trends [14, 21], but their purely data-driven nature limits robustness under changing conditions. The proposed PINN addresses this by embedding physics-informed loss functions, enforcing consistency with physical laws and improving generalization. While PINNs enhance stability and interpretability [22], they are typically used for continuous systems; this work extends them to discrete, multi-component marine power plants via aggregated conservation constraints. Digital twin approaches are widely used for monitoring and decision support [23, 24], yet often rely on either simulation or data alone. The proposed framework unifies telemetry, CSM simulation, and fuzzy inference into a consistent representation, improving accuracy and interpretability. Similarly, although multi-source fusion improves diagnostics [11], many methods remain purely statistical; here, structured fusion aligns heterogeneous inputs in a shared state space, reducing inconsistencies. Model robustness under noise, missing data, and imbalance remains challenging [25], which is mitigated here through physical constraints and multi-domain features as regularization. In maritime applications, machine learning-based diagnostics often overlook subsystem heterogeneity [26]; the multi-branch architecture addresses this via subsystem-specific feature extraction. Federated learning improves scalability and privacy [27], but rarely incorporates domain knowledge; the proposed approach combines distributed learning with physics-informed constraints. Unlike many studies based on simulation, this work includes Hardware-in-the-Loop validation, ensuring realistic performance evaluation [28].

Overall, the proposed method extends CNN–LSTM models with physical constraints, advances digital twin frameworks through coherent multi-source integration, improves robustness to noisy heterogeneous data, introduces subsystem-aware modeling, and is validated under realistic conditions. These results both confirm current research trends and contribute to the development of reliable SPP failure prediction.

Conclusion

This section describes the development and validation of a hybrid spatio-temporal neural network for predicting failure risk in marine power plant subsystems. The approach overcomes limitations of purely data-driven models by incorporating physical constraints within a digital twin framework. A multi-branch CNN–LSTM architecture is proposed for

subsystem-specific feature extraction from heterogeneous sources (sensor data, cognitive simulation, fuzzy diagnostics), capturing both short-term dynamics and long-term degradation. The inclusion of conservation-based constraints in the loss function ensures physically consistent predictions; their removal increases error by up to 41%, confirming their significance. The model achieves stable performance ($R^2 = 0.982$, $RMSE = 0.039$) under noise and incomplete data. Hardware-in-the-Loop validation on an ARM Cortex-A72 demonstrates near real-time applicability, with inference latency of 42 ± 5 ms and moderate resource demands, supporting onboard deployment. The architecture also improves interpretability through subsystem-oriented branches and physics-based regularization, forming a basis for explainable diagnostics. Overall, the proposed PINN CNN–LSTM model provides a consistent and practical solution for failure risk prediction and can be integrated into digital twin systems for advanced monitoring and decision support.

ЛІТЕРАТУРА REFERENCES

1. Вичужанин В. В. Информационное обеспечение мониторинга и диагностирования технического состояния судовых энергоустановок. // Вісник Одеського національного морського університету: зб. наук. праць. – 2012. – № 35. – С. 111 – 124.
2. Вичужанин В. В. Методы информационных технологий в диагностике состояния сложных технических систем : монография / В. В. Вичужанин, Н. Д. Рудниченко. – Одесса : Экология, 2019. – 178 с.
3. Вичужанин В. В., Рудниченко Н. Д. Метод управления рисками судовых сложных технических систем // Проблеми техніки. – 2014. – № 2. – С. 138 – 142.
4. Бойко В. Д., Вичужанин В. В. Модель оценки живучести судовых технических систем // Вестник Миколаївського кораблебудівного університету. – 2012. – № 3. – С. 62 – 67.
5. Vychuzhanin V., Vychuzhanin A. Model for determining similarity between cases in failure diagnosis of ship power plant equipment // Вісник Приазовського державного технічного університету. Серія: Технічні науки. – 2025. – № 50. – С. 9 – 17.
6. Vychuzhanin V., Vychuzhanin A. Integrated modeling of reliability and maintenance of ship's power plant equipment considering degradation and operational conditions // Системи контролю інформації та

інтелектуальні технології. Досягнення та застосування : монографія. - Lviv-Toruń : Liha-Pres, 2025. – С. 335 – 363.

7. Zhu Z., Lei Y., Qi G., Chai Y., Mazur N., An Y., Huang X. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery // *Measurement: Journal of the International Measurement Confederation*. — 2023. — Vol. 206. — P. 112346 – 112346.

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

9. Tao F., Qi Q., Liu A., Kusiak A. Data-driven smart manufacturing // *Journal of Manufacturing Systems*. – 2019. – Vol. 48. – P. 157 – 169.

10. Coraddu A., Oneto L., Ghio A., Savio S., Anguita D., Figari M. Machine learning approaches for improving condition-based maintenance of naval propulsion plants // *Ocean Engineering*. – 2020. – Vol. 209. – P. 1 – 24. – Art. 107379.

11. Lei Y., Yang B., Jiang X., Jia F., Li N., Nandi A. K. Applications of machine learning to machine fault diagnosis: A review and roadmap // *Mechanical Systems and Signal Processing*. – 2020. – Vol. 138. – P. 106587 – 106587.

12. Mittal K., Jain A., Vaisla K. S., Castillo O., Kacprzyk J. A comprehensive review on type-2 fuzzy logic applications: Past, present and future // *Engineering Applications of Artificial Intelligence*. – 2020. – Vol. 95. – P. 103916 – 103916.

13. Yang Y., Tu F., Huang S., Tu Y., Liu T. Research on CNN-LSTM DC power system fault diagnosis // *Frontiers in Energy Research*. – 2023. – Vol. 11. – P. 1258549 – 1258549.

14. Li X., Ding Q., Sun J. Q. Remaining useful life estimation in prognostics using deep convolution neural networks // *Reliability Engineering & System Safety*. – 2022. – Vol. 172. – P. 1 – 11.

15. *Offshore Reliability Data Handbook / OREDA Participants*. - 6th ed. – DNV GL, 2020. – 836 p.

16. Wang T., Yu J., Siegel D., Lee J. A similarity-based prognostics approach for remaining useful life estimation // *IEEE Transactions on Reliability*. – 2021. – Vol. 70, No. 1. – P. 246 – 261.

17. Vychuzhanin V., Vychuzhanin A. Stochastic models and methods for diagnostics, assessment, and prediction of the technical condition of complex critical systems : monograph. – Lviv-Toruń : Liha-Pres, 2025. – 176 p.

18. Vychuzhanin V., Vychuzhanin A. Adequacy and verification of an intelligent diagnostic model for ship power plants // *Informatics and*

Mathematical Methods in Simulation. – 2025. – Vol. 15. – No. 3. – P. 312 – 326.

19. Wang J., Cao H., Cui Z. et al. Intelligent fault diagnosis of marine diesel engines based on efficient channel attention-improved convolutional neural networks // *Processes*. – 2023. – Vol. 11. – No. 12. – P. 3360.

20. Lang X., Wu D., Mao W. Physics-informed machine learning models for ship speed prediction // *Expert Systems with Applications*. – 2024. – Vol. 238. – P. 121877 – 121877.

21. Babu G. S., Zhao P., Li X.-L. Deep convolutional neural network based regression approach for estimation of remaining useful life // *Proceedings of the Annual Conference of the Prognostics and Health Management Society (PHM)*. – 2016. – P. 214 – 228.

22. Karniadakis G. E. et al. Physics-informed machine learning // *Nature Reviews Physics*. – 2021. – Vol. 3. – P. 422 – 440.

23. Tao F., Zhang M., Liu Y., Nee A. Y. C. Digital twin in industry: State-of-the-art // *IEEE Transactions on Industrial Informatics*. – 2022. – Vol. 18. – No. 4. – P. 2405 – 2415.

24. Jones D., Snider C., Nassehi A., Yon J., Hicks B. Characterising the digital twin // *CIRP Annals*. – 2020. – Vol. 69. – No. 1. – P. 36 – 39.

25. Sambasivan N., Kapania S., Highfill H., Akrong D., Paritosh P., Aroyo L. “Everyone wants to do the model work, not the data work”: Data cascades in high-stakes AI // *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. – 2021. – P. 1 – 15.

26. Coraddu A., Oneto L., Ghio A., Savio S., Anguita D., Figari M. Machine learning approaches for improving condition-based maintenance of naval propulsion plants // *Proceedings of the Institution of Mechanical Engineers, Part M*. – 2016. – Vol. 230. – No. 1. – P. 136 – 153.

27. Nguyen D. C., Ding M., Pathirana P. N., Seneviratne A., Li J., Poor H. V., Niyato D. Federated learning for Internet of Things: A comprehensive survey // *IEEE Communications Surveys & Tutorials*. – 2021. – Vol. 23. – No. 3. – P. 1622 – 1658.

28. Fuller A. et al. Digital twin: Enabling technologies, challenges and open research // *IEEE Access*. – 2020. – Vol. 8. – P. 108952 – 108971.

Анотація. Розглянуто проблему підвищення точності прогнозування ризику відмов складних технічних систем на прикладі суднових енергетичних установок в умовах високої структурної складності, нелінійності процесів деградації та невизначеності експлуатаційних даних. Обґрунтовано обмеженість традиційних аналітичних, статистичних і нечітко-логічних підходів, які не забезпечують достатньої достовірності довгострокових прогнозів у динамічних

середовищах функціонування. Запропоновано гібридну просторово-часову нейронну модель, побудовану на основі інтеграції згорткових нейронних мереж та мереж довгої короткочасної пам'яті, яка функціонує в межах цифрового двійника. Особливістю підходу є поєднання трьох інформаційних доменів: високочастотних телеметричних даних, фізично обґрунтованих індикаторів стану, отриманих із когнітивної імітаційної моделі, а також діагностичних ознак, сформованих за допомогою нечіткої логіки. Це дозволяє забезпечити когнітивну узгодженість і фізичну валідність результатів прогнозування. У роботі реалізовано механізм фізично інформованого навчання, що передбачає включення законів збереження енергії та маси до функції втрат нейронної мережі. Такий підхід забезпечує обмеження простору гіпотез моделі та запобігає формуванню фізично неможливих прогнозів. Запропонована багатогілкова архітектура дозволяє виконувати незалежну обробку даних для окремих підсистем (паливної, системи охолодження та електротехнічної), що підвищує чутливість до локальних дефектів і деградаційних процесів. Експериментальна перевірка моделі проведена з використанням підходу Hardware-in-the-Loop, що забезпечує наближеність умов тестування до реальної експлуатації. Отримані результати демонструють підвищення точності прогнозування, стійкість до шумів і пропусків у даних, а також придатність моделі до використання в системах реального часу. Додатково реалізовано механізми інтерпретації результатів на основі аналізу внеску підсистем та методу SHapley Additive exPlanations, що підвищує довіру до прийнятих рішень. Отримані результати підтверджують доцільність використання гібридних нейромережеских моделей із фізичною регуляризацією для задач прогнозної діагностики та управління ризиками в складних технічних системах.