EAI Endorsed Transactions
on Transportation Systems and Ocean Engineering Research Article EAI.EU

Hybrid GRU-VAE and Spectral Residual Anomaly Detection for OWC Predictive Maintenance: A Technical and Economic Assessment at Chu Lai Port

Tran Pham Ngoc Minh[1], Nguyen Cat Quynh Nhu[1], Nguyen Le Bao Phuc[1], and Tien-Dat Tran[1*]

[1]AIT Group, University of Transport Ho Chi Minh City, No. 02 Vo Oanh Street, Thanh My Tay Ward, Ho Chi Minh City

Abstract

The central coast of Vietnam holds substantial wave energy potential, in which Oscillating Water Column (OWC) systems deployed in this setting require robust condition-monitoring and predictive maintenance strategies to sustain reliable operation. The study proposes a serial hybrid anomaly detector in which a GRU-VAE learns the normal wave-driven behavior of the SCADA channels and Spectral Residual (SR) analysis is applied to the GRU-VAE reconstruction residual to expose short-term frequency anomalies. On a simulated test set of 7,481 time steps, with all thresholds and weights fixed on a validation set, the hybrid reached, on average over 3 simulated datasets, a Precision of 0.845, Recall of 0.864 and F1-score of 0.854 (AUC-ROC 0.962), comparable to, though slightly below, a standalone GRU-VAE (mean F1 0.868) and an LSTM autoencoder (mean F1 0.867); applying SR to the residual instead of the raw signals raised the AUC-ROC of the SR branch from 0.881 to 0.955. Taking a Precision of 0.91 and a Recall of 0.89 as the target detection performance, a 15-year life-cycle cost (LCC) model with a 9% annual discount rate yields an NPV of +USD 79,955 and an IRR of 14.76% (+USD 69,719 and 14.1% at the measured performance). Monte Carlo simulation indicated a 75.9% probability of a positive NPV under the stated input distributions, while Tornado analysis identified the failure frequency, the emergency repair cost and the recall of the detector as the main drivers of NPV. These findings indicate the potential technical and economic feasibility of integrating machine-learning-based predictive maintenance into OWC systems in Vietnam, under literature-based cost assumptions.

Keywords: GRU-VAE, Spectral Residual, Predictive Maintenance, OWC, Techno-Economic Analysis

Received on 05 September 2026, accepted on 19 October 2026, published on 22 October 2026

Copyright © 2026 Tran Pham Minh *et al.*, licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/tsoe.14956

1. Introduction

The central coast of Vietnam possesses considerable wave energy potential. Tung and Chien [1] reported an annual average wave power of 7-9 kW/m at a water depth of 30m, with a peak value of approximately 29 kW/m during periods of strong northeast monsoon activity in December. The Chu Lai area, with a coastline directly exposed to the northeast, is located in a region with favorable conditions for wave energy. The Oscillating Water Column (OWC) is a wave energy conversion technology suitable for shallow waters and can be integrated into existing port infrastructure without requiring an independent foundation.

The continuous operation of OWC systems in marine environments requires a Supervisory Control and Data Acquisition (SCADA) system capable of detecting anomalies in a timely manner. Conventional static thresholding and rule-based methods fail to adequately exploit multivariate correlations and temporal structures inherent in operational data. Unsupervised machine learning methods, such as Isolation Forest [2] and One-Class Support Vector Machine [3], do not require labeled data but remain limited in handling time series with complex periodic structures. Hybrid structures combining Recurrent Neural Networks (RNN) with generative models – particularly Variational Autoencoder (VAE) and Gated Recurrent Units (GRU) [4, 5] – can learn the probability distribution of normal operating states to detect deviations without the need for labeled data;

however, they still lack a dedicated mechanism for short-term frequency anomalies.

This study proposes a GRU-VAE + SR method that combines the GRU-VAE architecture [4, 5] with the Spectral Residual (SR) algorithm [6] through a weighted linear combination of normalized scores, with the detection threshold determined using the P99.5 percentile on the validation set. Because no OWC system is installed at Chu Lai and no Mutriku SCADA data were available, the detectors are trained and evaluated on simulated SCADA data with eight channels and five injected fault types. The pre-training and fine-tuning workflow of the proposed framework, pre-training on Mutriku conditions and fine-tuning for Chu Lai, is carried out on simulated data for both sites (Section 3.5). The Mutriku breakwater plant, which has reached a cumulative production of about 3 GWh [7], serves as the reference OWC configuration; nominal operating values are assumptions of the authors, and wave-driven channels are scaled by the ratio of significant wave heights between Chu Lai (1.2 m) and Mutriku (1.8 m). The reported metrics are estimates of expected performance under simulated conditions, not field measurements. The technical performance of the proposed model is incorporated into a Life Cycle Cost (LCC) framework [8] to assess its economic feasibility, together with Monte Carlo simulation and Tornado analysis to quantify investment risks under uncertainties in electricity prices and the carbon credit market. The contributions of this study are threefold. First, it shows that the placement of Spectral Residual analysis matters for wave-driven SCADA data: applied to the raw signals, SR lowers the ranking quality of the fused score, whereas applied to the GRU-VAE reconstruction residual it becomes a usable detector on its own (mean AUC-ROC 0.955 vs. 0.881); within the fused detector, however, it does not improve on a standalone GRU-VAE, and this is reported explicitly. Second, it links the Precision and Recall of the detector to maintenance costs and cash flows through an explicit chain of equations, so that the economic case can be evaluated at any operating point, including both the target performance assumed for the base case and the performance measured here. Third, it implements the intended workflow of pre-training on Mutriku conditions and fine-tuning for Chu Lai, evaluated on simulated data for both sites.

2. Research methods

2.1 Structural overview

The design rationale stems from two distinct anomaly types observed in OWC SCADA data:

(i) long-term statistical deviation, in which the operating state gradually drifts away from its normal distribution over successive time steps, characteristic of gradual mechanical degradation;
(ii) short-term frequency anomalies, manifested as energy spikes within a specific frequency band, characteristic of sudden events such as resonance vibration or grid desynchronization.

We hypothesize that no single method handles both types well: GRU-VAE is sensitive to type (i), whereas SR targets type (ii) but has no model of normal behavior. This hypothesis is tested in Section 3.3.

The proposed system comprises two stages applied in series to the input sequence $X = \{x_1, x_2, \dots, x_T\} \in R^{T \times D}$, where T denotes the sliding-window length and D the number of SCADA channels. The GRU-VAE stage learns the normal wave-driven behavior and produces a reconstruction x̂t; the SR stage operates in the frequency domain on the reconstruction residual $r_t = x_t - \hat{x}_t$, so that the periodic wave component is removed before frequency anomalies are sought. Applying SR directly to the raw signals was found to be less effective (Section 3.3). The overall architecture is illustrated in Figure 1.

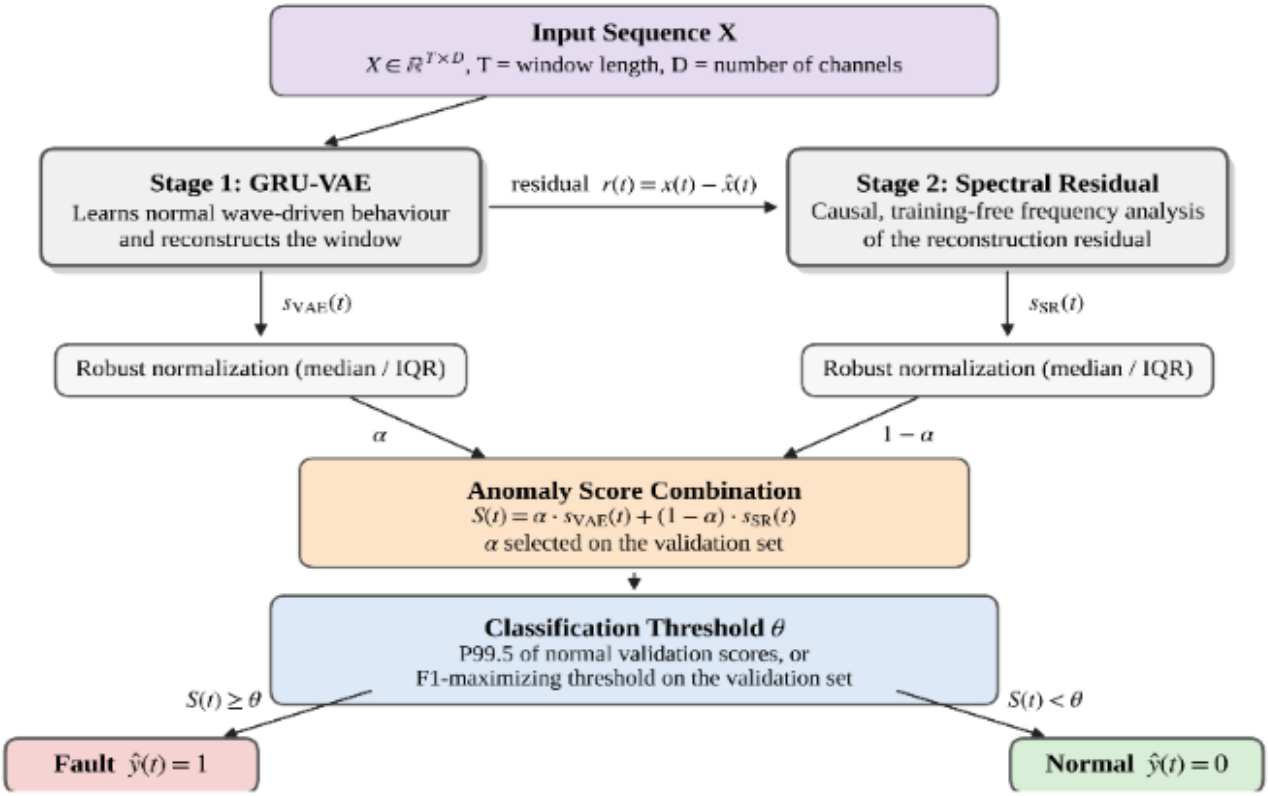


Figure 1. Overall architecture of the serial GRU-VAE + SR detector

2.2 Algorithm implementation

Score normalization and fusion. Because $s_{VAE(t)}$ and $s_{SR(t)}$ operate on different numerical scales, both scores are normalized by the median and inter-quartile range of their values on the normal validation windows, without clipping. The scores are fused by Eq. (1) [9]; the weight α is varied from 0 to 1 in steps of 0.1 and selected by Average Precision on the validation set (Table 3*)*:

$$S_{combined}(t) = \alpha \cdot S_{VAE}(t) + (1 - \alpha) \cdot S_{SR}(t) \quad (1)$$

Threshold determination. The classification threshold θ is calibrated exclusively on normal-condition samples from the validation set ($D_{val,normal}$) using the P99.5 percentile, meaning 99.5% of normal-condition scores fall below this cutoff. This non-parametric procedure makes no assumption about the underlying score distribution and fixes the false-alarm rate at 0.5% on the normal validation windows; the false-alarm rate realized on the test set is reported in Table 2. The P99.5 percentile was chosen over P99 or P95 because the design objective prioritizes high precision - a stricter cutoff lowers the false positive rate, consistent with the operational requirements of real-world OWC deployment - following Eq. (2) [10]:

$$\theta = Quantile_{P99.5}(\{S_{combined}(t) : t \in D_{val,normal}\}) \quad (2)$$

Final classification decision, per Eqs. (3) and (4):

$$\hat{y}(t) = 1\ (Fault)\ \text{ if } S_{combined}(t) \geq \theta \quad (3)$$

$$\hat{y}(t) = 0\ (Normal)\ if\ S_{combined}(t) < \theta \quad (4)$$

Two operating points are reported: (i) a label-free rule, with θ at the P99.5 percentile of the combined score on the normal validation windows; (ii) a label-assisted rule, with θ maximizing F1 on the labeled validation set. In both cases θ is fixed before the test set is used.

2.3 Simulated data and implementation

The simulated dataset has 50,000 time steps and $D = 8$ channels (chamber air pressure, turbine speed, generated power, generator temperature, air flow rate, shaft vibration, grid voltage and generator current), sampled every 0.25 s. Five fault types are injected: turbine stall, generator overheating, air-chamber seal leakage, a progressive power loss used as a proxy for biofouling, and a vibration growth used as a proxy for corrosion fatigue. Each fault segment lasts

50–200 steps (12.5–50 s), and a window takes the label of its last step. The data are split chronologically into training (65%), validation (20%) and test (15%) sets; the scaler is fitted on normal training samples only, and only normal windows are used to train the models. Each of the three simulated datasets (seeds 42, 1 and 2) has a test set of 7,481 windows, of which 559–938 are fault windows forming 4–7 contiguous fault events (17 events in total).

Table 1. Configuration of the data, model, training and evaluation

Parameter	Value
Data	50,000 time steps; 8 channels; sampling interval 0.25 s; 5 fault types; seeds 42, 1, 2
Sliding window	$T = 20$ steps (5 s); stride 1; label of the last step
Data split	65% / 20% / 15% in chronological order; scaler fitted on the normal samples of the training set
GRU-VAE encoder	GRU 128 → 64 → 32, dropout 0.15; latent dimension $z = 32$
GRU-VAE decoder	GRU 32 → 64 → 128; linear output layer with 8 channels
Loss function	reconstruction error $+ \beta \cdot KL, \beta = 0.4$
GRU-VAE score	window-averaged reconstruction error $+ \gamma \cdot KL, \gamma = 0.25$
Training	Adam, $l_r = 8 \times 10 - 4$ (ReduceLROnPlateau, patience 5, factor 0.5); gradient clipping at 1.0; batch size 128; 50 epochs; normal windows only
Residual for SR	rt = xt − x̂t at the last step of the window ending at t (causal), per channel
Spectral Residual	causal, over the most recent L steps; (L, q) selected by AP on the validation set, $L \in \{32, 64, 128\}, q \in \{3, 5, 9\}$; on the residual: $L = 64, q = 9$; on the raw signals: $L = 128, q = 5$; channel-wise z-score (statistics from normal training data), maximum over channels, averaged within the window
Score normalization	$(s - median)/I(QR)$, computed on the normal windows of the validation set; no clipping
Fusion	$S = \alpha \cdot S_{VAE}(t) + (1 - \alpha) \cdot S_{SR}(t)$; α selected by AP on the validation set, separately for each seed: serial $\alpha = 0.9/0.9 \ /0.5$ (seeds 42 / 1 / 2); parallel $\alpha = 1.0$ for all three seeds
Threshold	(i) P99.5 of the normal windows of the validation set; (ii) maximum F1 on the validation set; fixed before the test set is used
LSTM-AE (baseline)	64-unit LSTM in the encoder and decoder; MSE loss; same training configuration
Rule-based (baseline)	maximum channel-wise |z| within the window (statistics from normal training data)
Pre-training and fine-tuning (Section 3.5)	pre-training for 30 epochs on simulated data for Mutriku conditions (scaling factor 1.0; independent dataset); fine-tuning for 10 epochs, $lr = 4 \times 10^{-4}$, on 20% of the Chu Lai training data

3. Detection results

3.1 Comparison with baseline methods

Table 2 compares all methods on the test sets of the three simulated datasets, with every threshold fixed on the validation set. Figure 2 shows the confusion matrices of the serial hybrid under both threshold rules and the ROC curves of all methods for one of these datasets (seed 42).

Averaged over the three datasets, the serial hybrid reached Precision = 0.845, Recall = 0.864, F1 = 0.854 ± 0.012, AP = 0.842 ± 0.028 and AUC-ROC = 0.962 ± 0.017; the standalone GRU-VAE reached F1 = 0.868 ± 0.023, AP = 0.848 ± 0.040 and AUC-ROC = 0.963 ± 0.019, and the LSTM-AE reached F1 = 0.867 ± 0.007 and AP = 0.862 ± 0.029. The hybrid is therefore comparable to, though slightly below, the standalone GRU-VAE, and the differences lie within one standard deviation. Under the F1-maximizing rule, all methods, including the rule-based baseline (F1 = 0.794 ± 0.029), detect all 17 fault events. Under the label-free P99.5 rule, the serial hybrid keeps the false-positive rate on the test set at 0.6% but detects only 12 of the 17 fault events.

Table 2. Comparison of methods on the test set (thresholds fixed on the validation set; mean over three seeds: 42, 1, 2)

Method	Rule	Precision	Recall	F1	FPR	Events	AP	AUC-ROC
Rule-based	P99.5	0.794	0.234	0.346 ± 0.180	0.007	9/17	0.749 ± 0.009	0.949 ± 0.013
	F1max	0.811	0.779	0.794 ± 0.029	0.020	17/17		
LSTM-AE	P99.5	0.905	0.432	0.577 ± 0.104	0.005	11/17	0.862 ± 0.029	0.963 ± 0.016
	F1max	0.849	0.885	0.867 ± 0.007	0.018	17/17		
GRU-VAE	P99.5	0.892	0.420	0.567 ± 0.079	0.006	12/17	0.848 ± 0.040	0.963 ± 0.019
	F1max	0.848	0.889	0.868 ± 0.023	0.019	17/17		
SR on raw signals	P99.5	0.829	0.270	0.406 ± 0.069	0.006	12/17	0.580 ± 0.032	0.881 ± 0.020
	F1max	0.574	0.546	0.558 ± 0.056	0.045	17/17		
SR on GRU-VAE residual	P99.5	0.841	0.248	0.370 ± 0.150	0.006	15/17	0.702 ± 0.082	0.955 ± 0.012
	F1max	0.666	0.740	0.699 ± 0.028	0.041	17/17		
Parallel hybrid (α = 0.5; causal SR on raw signals)	P99.5	0.892	0.424	0.569 ± 0.085	0.006	12/17	0.837 ± 0.044	0.960 ± 0.022
	F1max	0.839	0.851	0.845 ± 0.033	0.019	17/17		
Serial hybrid (α selected on the validation set)	P99.5	0.892	0.419	0.566 ± 0.078	0.006	12/17	0.842 ± 0.028	0.962 ± 0.017
	F1max	0.845	0.864	0.854 ± 0.012	0.018	17/17		
Method	Rule	Precision	Recall	F1	FPR	Events	AP	AUC-ROC

Note: Precision, Recall and FPR are means over three seeds; F1, AP and AUC-ROC are means ± standard deviations. "Events": total number of fault events with at least one alarmed window, out of the 17 events of the three test sets (7, 6 and 4 events, respectively). The α of the serial hybrid is selected by AP on the validation set for each seed ($0.9-0.9-0.5$*); the parallel hybrid with α selected on the validation set (*$\alpha = 1.0$ *for all three seeds) coincides with the standalone GRU-VAE*

and is therefore not repeated. The configuration exactly as originally submitted (SR computed over the whole data segment, non-causal) is reported in the appendix.

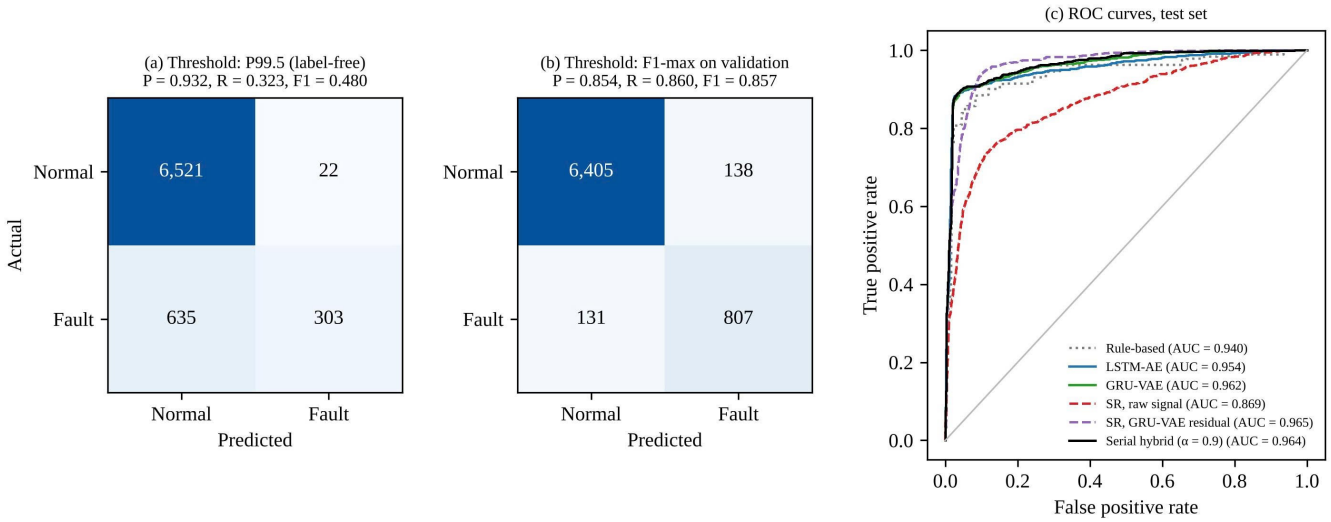


Figure 2. Serial GRU-VAE + SR detector on the test set of one simulated dataset (seed 42): confusion matrices with the label-free P99.5 threshold (a) and with the F1-maximizing threshold selected on the validation set (b); ROC curves of all compared methods (c). Mean values over the three simulated datasets are given in Table 2.

3.2 Threshold-independent and operating-point metrics

AP and AUC-ROC do not depend on the operating threshold, whereas Precision, Recall, F1 and FPR do, and are always reported with their threshold rule. Because the test set is imbalanced (938 fault vs. 6,543 normal windows), the Precision–Recall curve is shown alongside the ROC curve [11]. Applied to the raw signals, the SR score ranks windows less well than the GRU-VAE (mean AP 0.580 vs. 0.848 over the three simulated datasets), because the periodic wave component dominates the spectrum, and parallel fusion lowers AP monotonically as the SR weight increases. Applied to the GRU-VAE reconstruction residual, the SR score reaches AP 0.702 and AUC-ROC 0.955; with serial fusion, AP decreases more slowly as the SR weight increases, and AUC-ROC reaches 0.967 at $\alpha = 0.4$–0.5 (Table 3). For the serial hybrid, the label-free P99.5 rule keeps the test false-positive rate at 0.6% but detects only 12 of the 17 fault events, because many fault windows score below the upper tail of the normal scores.

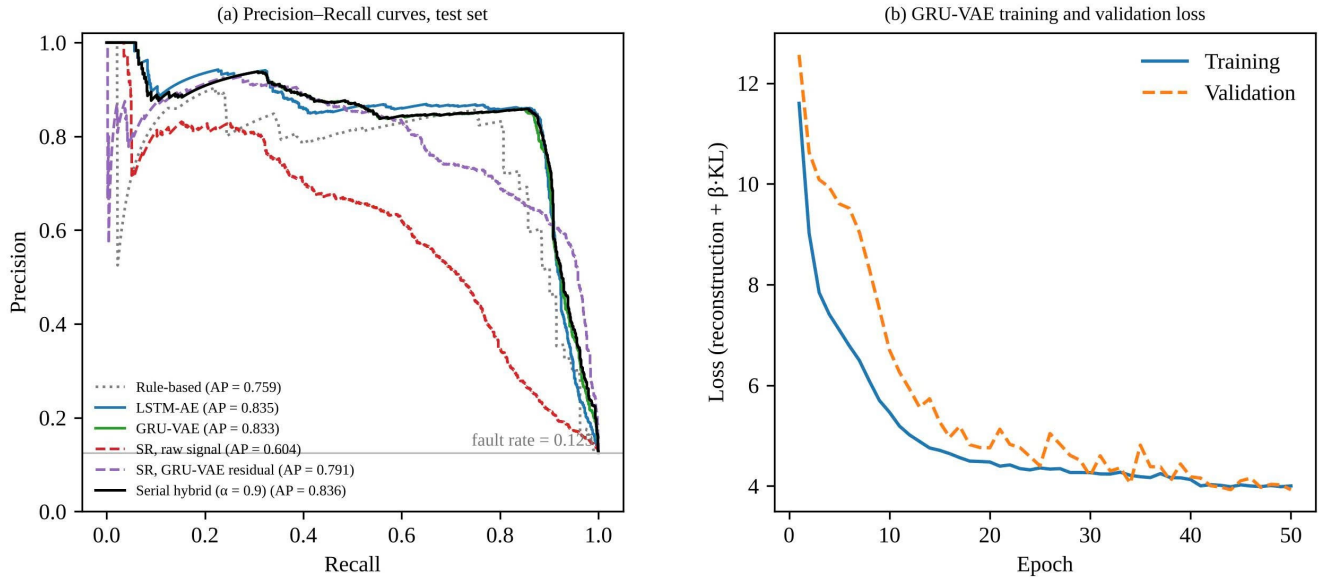


Figure 3. Precision–Recall curves of the compared methods on the same test set (a); training and validation loss of the GRU-VAE (b)

The training loss curve (Figure 3., right) confirms stable convergence with no evidence of overfitting across 50 epochs.

3.3 Fusion design and sensitivity to the fusion weight

Table 3 reports the sensitivity to the fusion weight α for both fusion designs, averaged over the three simulated datasets. With parallel fusion, validation AP peaks at α = 1.0 for every dataset, i.e. the SR branch does not contribute, and test AP decreases monotonically as the SR weight increases. With serial fusion, the selected α was 0.9, 0.9 and 0.5; test AP increases from 0.823 at α = 0.2 to 0.848 at α = 0.9–1.0, and test AUC-ROC reaches 0.967 at α = 0.4–0.5 (0.963 for the GRU-VAE alone), while no value of α gives a higher test F1 than the GRU-VAE alone (0.868). None of the three alternative fusion rules (Table 4) exceeded the GRU-VAE alone in F1 or AP. The benefit of the SR stage is therefore confined to a small gain in AUC-ROC.

Table 3. Sensitivity to α for the two fusion designs (threshold recomputed for each α; mean over three seeds)

α	Parallel: AP val	Parallel: F1 test	Parallel: AP test	Serial: AP val	Serial: F1 test	Serial: AP test	Serial: AUC test
0.0 (SR only)	0.595	0.558	0.580	0.737	0.699	0.702	0.955
0.1	0.771	0.764	0.779	0.818	0.760	0.810	0.964
0.2	0.793	0.792	0.806	0.824	0.800	0.823	0.965
0.3	0.806	0.812	0.821	0.825	0.805	0.829	0.966
0.4	0.814	0.827	0.831	0.827	0.817	0.833	0.967
0.5	0.820	0.845	0.837	0.829	0.824	0.837	0.967
0.6	0.825	0.854	0.842	0.831	0.839	0.840	0.966
0.7	0.828	0.860	0.845	0.833	0.845	0.843	0.966

0.8	0.831	0.867	0.847	0.834	0.853	0.846	0.965
0.9	0.833	0.866	0.848	0.835	0.864	0.848	0.964
1.0 (GRU-VAE only)	0.834	0.868	0.848	0.834	0.868	0.848	0.963

Note: the values are means over three seeds; "F1 test" uses the F1-maximizing threshold selected on the validation set. The bold cell in the "AP val" column is the largest mean value; the α actually used is selected separately for each seed (serial 0.9 / 0.9 / 0.5; parallel 1.0 for all three seeds; Table 1), so the α = 0.9 row does not coincide with the serial-hybrid row of Table 2. The bold cells in the "AUC test" column are the largest values.

Table 4. Fusion rules for the GRU-VAE and SR scores (all parameters selected on the validation set; test set; mean over three seeds)

Fusion rule	F1 val	Precision	Recall	F1	AP	Mean delay (s)
Linear, parallel (α = 1.0 for all three seeds)	0.859	0.848	0.889	0.868	0.848	3.72
Linear, serial (α = 0.9 / 0.9 / 0.5)	0.850	0.845	0.864	0.854	0.842	4.60
Maximum (parallel)	0.836	0.830	0.844	0.837	0.837	5.22
Rank average (parallel)	0.763	0.821	0.682	0.745	0.779	5.48
OR rule (separate threshold for each branch)	0.848	0.828	0.891	0.858	–	3.60

Note: for the OR rule, the threshold of each branch is selected by F1 on the validation set: GRU-VAE ≥ P98 or SR ≥ P99.5 (seed 42); GRU-VAE ≥ P99 with the SR branch switched off (seed 1); GRU-VAE ≥ P98 or SR ≥ P99.8 (seed 2). The OR rule does not produce a continuous score, so AP is not computed.

3.4 Event-level detection and detection delay

Table 5 reports the mean detection delay by fault type. The advantage of the deep-learning models appears mainly for gradually developing faults: for air-chamber seal leakage and power loss, the mean detection delay of the GRU-VAE is 6.0 s and 6.9 s, compared with 14.7 s and 12.3 s for the rule-based baseline. The serial hybrid does not detect any fault type earlier than the standalone GRU-VAE, and SR applied to the residual, when used alone, detects every fault type later than the GRU-VAE.

Table 5. Mean detection delay by fault type (seconds; F1-max rule; events pooled over three seeds)

Fault type	Events	Rule-based	LSTM-AE	GRU-VAE	SR on residual	Serial hybrid
Turbine stall	3	0.00	0.00	0.00	2.08	0.00

Generator overheating	2	3.63	4.63	2.63	6.13	4.38
Air-chamber seal leakage	3	14.67	5.58	6.00	9.17	7.67
Power loss (biofouling proxy)	6	12.29	6.13	6.88	12.25	7.33
Vibration growth (corrosion proxy)	3	0.33	0.42	0.42	1.33	0.58

Note: all methods in the table detect all 17 events; the delay is averaged over the events of the same type in the three test sets. Delays are measured from fault onset; no method raises an alarm before a fault occurs.

3.5 Pre-training on Mutriku conditions and fine-tuning for Chu Lai

To carry out the pre-training and fine-tuning workflow of the proposed framework, the GRU-VAE was pre-trained for 30 epochs on an independent simulated dataset representing Mutriku wave conditions (H_s = 1.8 m) and fine-tuned for 10 epochs (learning rate 4×10^{-4}) on 20% of the Chu Lai training data, representing the early operating stage when little site data are available. With only 20% of the Chu Lai data, pre-training raised the Average Precision from 0.803 ± 0.008 (training from scratch) to 0.828 ± 0.036, close to the 0.848 ± 0.040 obtained with the full Chu Lai training set, while F1 and AUC-ROC were almost unchanged (Table S1); with three datasets, this gain lies within one standard deviation and is therefore indicative rather than statistically significant. Table 2 thus represents the reference case in which sufficient site data are available, whereas this section represents the early operating stage, in which the detector is initialized from Mutriku conditions and fine-tuned on limited Chu Lai data. The Mutriku model without fine-tuning performed similarly (AP 0.838 ± 0.049), because the two simulated sites differ mainly in signal amplitude, which per-site normalization largely removes; these results therefore validate the pre-training and fine-tuning workflow but not the full site-to-site shift, which requires field data.

4. Techno-Economic Cost Analysis

4.1 Techno-Economic Assessment Framework

The IoT-PdM system comprises 12 multi-chamber OWC modules with 20 chambers per module and 240 chambers in total. The fault detection performance is integrated into a 15-year LCC model by comparing two scenarios: Scenario 1 – reactive maintenance without IoT-PdM, and Scenario 2 – predictive maintenance with IoT-PdM. Based on the differences in maintenance costs and operational benefits, the study calculates Net Cash Flow (NCF), Net Present Value (NPV), Internal Rate of Return (IRR), and Payback Period. Input variable uncertainty is quantified using Monte Carlo simulation with 10,000 scenarios and Tornado sensitivity analysis based on the One-at-a-Time (OAT) method. Table 6 summarizes the baseline assumptions.

Table 6. Baseline assumptions of the techno-economic model

Statistics	Symbol	Value	Basis/Assumption
Number of OWC failures per year	N_f	6 (failures/year)	Estimated from OWC literature [12]
Emergency repair cost	C_{repair}	~6,000 (USD/failure)	Estimated from OWC literature [12]
Periodic inspection cost	$C_{inspect}$	~800 (USD/inspection)	Estimated from OWC literature [12]
Number of periodic inspections	$N_{inspect}$	24 (inspections/year)	Two inspections/module/year
Number of AI alarms	N_{alert}	≈ 5.8681 (times/year)	$\frac{R \times N_f}{P}$

Emergency downtime	T_{down}	4 (hours/failure)	Assumption based on wave energy converter literature [13]
Planned downtime	T_{plan}	1.5 (hours/failure)	Estimated from OWC literature [12]
Downtime – Scenario 1	$T^1_{downtime}$	24 (hours/year)	$N_f \times T_{down}$
Downtime – Scenario 2	$T^2_{downtime}$	10.65 (hours/year)	$(1-R) \times N_f \times T_{down} + R \times N_f \times T_{plan}$
Revenue loss during emergency downtime	C_{down}	~160 (USD/hour)	Estimated from OWC literature [12]
Revenue loss during planned downtime	$C_{down_{plan}}$	144 (USD/hour)	$90\% \times C_{down}$ [14]
Planned repair cost	C_{plan}	1,500 (USD/failure)	Assumption based on wave energy converter literature [13]
Initial IoT-PdM investment	CAPEX	230,000 (USD)	
Mid-life replacement cost (Year 8)	–	23,000 (USD)	10% of CAPEX (Section 4.3)
Decommissioning cost (Year 15)	–	23,000 (USD)	10% of CAPEX (Section 4.3)
IoT system operating cost	OPEX	10,000 (USD/year)	
AI alarm Precision	P	0.91	Target detection performance used for the economic base case; measured on simulated data (mean over 3 datasets): P = 0.845, R = 0.864 (Section 3.1)
AI fault detection Recall	R	0.89	Target detection performance used for the economic base case; measured on simulated data (mean over 3 datasets): P = 0.845, R = 0.864 (Section 3.1)
Carbon credit price	P_{CO_2}	8 (USD/t_{CO_2})	Projected carbon credit price [15]
Grid emission factor	EF	0.6592 (t_{CO_2}/MWh)	Vietnam grid emission factor in 2023 [16]
Grid electricity purchase price	P_{grid}	90 (USD/MWh)	Estimated from electricity price literature [17]
Actual OWC power factor	CF_{OWC}	0.4	Estimated from OWC literature [18]
Total OWC capacity (12 modules)	P_{OWC}	4.4 (MW)	Estimated from OWC literature [19]
Discount rate	r	9 (%/year)	Average level reported in economic assessments of renewable energy projects [20][21]
Lifetime	n	15 (years)	

The number of false alerts and the avoided downtime follow from the recall and precision of the detector:

$$FP = N_{alert} - R \cdot N_f = R \cdot N_f \cdot (1 - P)/P \quad (5)$$

$$\Delta T = T_{downtime1} - T_{downtime2} = R \cdot N_f \cdot (T_{down} - T_{plan}) \quad (6)$$

In Equations (5)–(6), N_f is the number of OWC failures per year, and R and P are the recall and precision of the

classifier, respectively. The classifier detects a fraction R of the failures early, which are then handled by planned repair, and misses the remaining fraction $1 - R$, which are handled by emergency repair. N_{alert} (the number of alerts), $T_{downtime1}$ and $T_{downtime2}$ (the downtime in Scenarios 1 and 2) are defined in Table 6; FP is the number of false alerts per year, each of which leads to one inspection; and ΔT is the number of downtime hours avoided per year.

The annual costs of the two scenarios and the net cash flow are then:

$$C_1 = N_f \cdot C_{repair} + N_{inspect} \cdot C_{inspect} + N_f \cdot T_{down} \cdot C_{down} \quad (7)$$

$$C_2 = OPEX + (1 - R) \cdot N_f \cdot C_{repair} + R \cdot N_f \cdot C_{plan} + FP \cdot C_{inspect} \quad (8)$$

$$C_{down,2} = (1 - R) \cdot N_f \cdot T_{down} \cdot C_{down} + R \cdot N_f \cdot T_{plan} \cdot C_{down,plan} \quad (9)$$

$$B_{uptime} = \Delta T \cdot P_{OWC} \cdot P_{grid} \quad (10)$$

$$C_{2,total} = C_2 + C_{down,2} - B_{uptime} \quad (11)$$

$$Carbon_{rev} = \Delta T \cdot P_{OWC} \cdot EF \cdot P_{CO2} \quad (12)$$

$$- \quad NCF = C_1 - C_{2,total} + Carbon_{rev} \quad (13)$$

In Equations (7) – (13), C_1 is the annual maintenance cost of Scenario 1 (reactive maintenance), including the revenue lost during downtime; C_2 is the annual direct cost of Scenario 2 (IoT-PdM), comprising the OPEX, the emergency repair of missed faults, planned repairs and the inspections triggered by false alerts, before the revenue lost during downtime is added and before the uptime benefit is deducted; $C_{down,2}$ is the remaining revenue lost during downtime in Scenario 2, in which planned downtime is valued at $C_{down,plan}$ (Table 6); B_{uptime} is the uptime benefit of the avoided downtime hours; $C_{2,total}$ is the total annual cost of Scenario 2, including downtime and uptime, and is therefore directly comparable with C_1; $Carbon_{rev}$ is the carbon-credit revenue from the recovered generation, where EF is the grid emission factor and P_{CO2} is the carbon-credit price; and NCF is the annual net cash flow used for the indicators in Section 4.3.

The following assumptions are made: (i) each true alert leads to a planned repair instead of an emergency repair; (ii) each false alert leads to exactly one inspection; (iii) in Scenario 2, the 24 periodic inspections are replaced by alert-driven inspections, so that only the inspections triggered by false alerts remain; (iv) the revenue lost during planned downtime is 90% of that lost during emergency downtime (Table 6); (v) R and P remain constant over the project lifetime; (vi) NCF is held constant every year, with the replacement cost in Year 8 and the decommissioning cost in Year 15; (vii) the failure frequency, the repair, inspection and downtime costs, and the electricity price are taken from the literature and have not been measured at Chu Lai.

The uptime benefit (c) and the carbon credit revenue (d) are computed at the rated capacity $P_{OWC} = 4.4\ MW$, which is an upper bound on the energy recovered during avoided downtime.

4.2 Economic Performance

Based on the assumptions in Table 6, the total maintenance cost decreases from 59,040 USD/year (scenario 1) to 18,681.75 USD/year (scenario 2), representing a 68.4% reduction. This reduction is higher than the 20-35% range commonly reported in the literature [22] because the present study fully accounted for downtime costs and uptime benefits and is therefore not directly comparable with studies that consider only repair and inspection costs. As shown in Figure 4, the economic improvement in Scenario 2 is primarily driven by substantial reductions in technical inspection, emergency repair, and downtime costs, although additional planned repair costs and OPEX are incurred. The uptime benefit further contributes to the overall economic advantage of the IoT-PdM system.

Table 7 details the components of the annual maintenance cost of the two scenarios.

Table 7. Breakdown of the annual maintenance cost of the two scenarios (USD/year)

Item	Equation	Numerical substitution	Scenario 1	Scenario 2
Emergency repair	(7), (8)	6 × 6,000; 0.11 × 6 × 6,000	36,000	3,960
Planned repair	(8)	0.89 × 6 × 1,500	–	8,010
Inspection	(5), (7), (8)	24 × 800; 0.528 × 800	19,200	422.5
IoT system OPEX	(8)		–	10,000

Direct cost, excluding downtime and uptime (C_2 in Scenario 2)	(8)		55,200	22,392.5
Revenue lost during downtime	(7), (9)	6 × 4 × 160; 0.11 × 6 × 4 × 160 + 0.89 × 6 × 1.5 × 144	3,840	1,575.84
Uptime benefit (deducted from cost)	(10)	13.35 × 4.4 × 90	–	− 5,286.6
Total annual cost ($C_1;\ C_2, total$)	(7), (11)	22,392.5 + 1,575.84 − 5,286.6	59,040	18,681.75

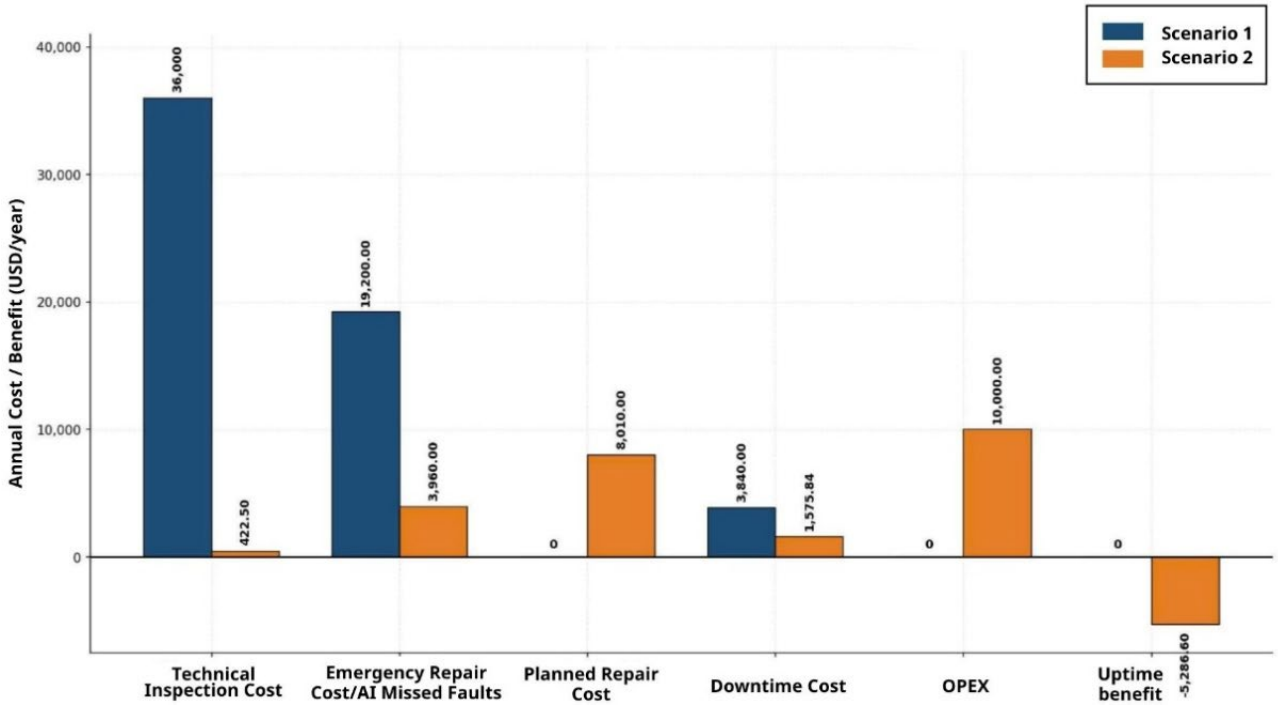


Figure 4. Comparison of cost components between the two scenarios

The annual benefit of the IoT-PdM system is generated from four sources: emergency repair cost savings, field inspection cost savings, uptime benefits, and carbon credit revenue. Figure 5 presents the contribution of each source, with a total gross annual benefit of 48,405 USD. The derivation of each component, and its reconciliation with the net annual cash flow used in Section 4.3, is given in Equations (14) – (17) and Table 8.

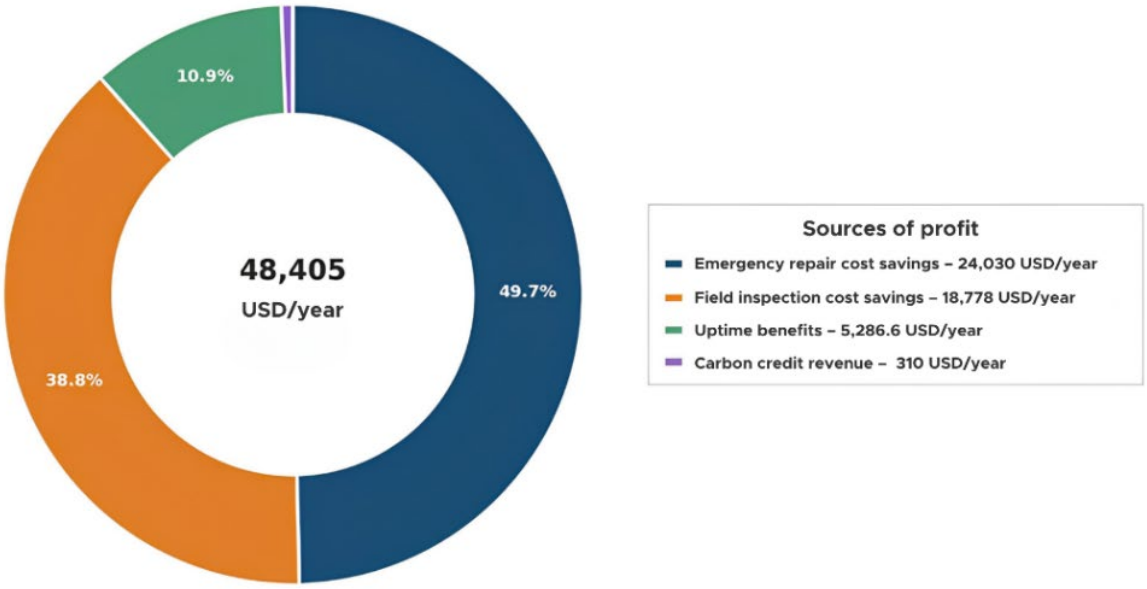


Figure 5. Composition of annual economic benefits generated by IoT-PdM

Figure 5 shows the gross benefit components (a)–(d); the net annual cash flow used for NPV is derived in Table 8 after deducting OPEX and adding the reduction in downtime revenue loss (e).

$$S_{repair} = R \cdot N_f \cdot \left(C_{repair} - C_{plan}\right) \quad (14)$$

$$S_{inspect} = N_{inspect} \cdot C_{inspect} - FP \cdot C_{inspect} \quad (15)$$

$$\Delta C_{down} = N_f \cdot T_{down} \cdot C_{down} - C_{down,2} \quad (16)$$

$$NCF = S_{repair} + S_{inspect} + B_{uptime} + Carbon_{rev} - OPEX + \Delta C_{down} \quad (17)$$

In Equations (14)–(17), $S_{repair}$is the saving in emergency repair, net of the cost of planned repair; $S_{inspect}$ is the saving in field inspections; ΔC_{down} is the reduction in revenue lost during downtime; and B_{uptime} and $Carbon_{rev}$ are given by Equations (10) and (12). The four terms $S_{repair}, S_{inspect}, B_{uptime}$ and $Carbon_{rev}$ are the four gross benefit sources in Figure 5; Equation (17) reconciles them with the NCF in Equation (13).

Table 8. Benefit components and annual net cash flow (USD/year)

Component	Equation	Numerical substitution	USD/year
(a) Emergency repair savings	(14)	0.89 × 6 × (6,000 − 1,500)	24,030
(b) Field inspection savings	(15)	24 × 800 − 0.528 × 800 = 19,200 − 422.5	18,778
(c) Uptime benefit	(10)	13.35 × 4.4 × 90	5,286.6
(d) Carbon credit revenue	(12)	13.35 × 4.4 × 0.6592 × 8	310
Total gross benefit (Figure 5)		(a) + (b) + (c) + (d)	48,405

IoT system OPEX			− 10,000
(e) Reduction in revenue lost during downtime	(16)	3,840 − 1,575.84	+ 2,264
Net cash flow, NCF	(17)		≈ 40,668 (≈ 40,000)

4.3 Financial Feasibility

The 15-year LCC analysis at a discount rate (r) = 9% per annum yields approximately 475,903 USD for LCC_1 (scenario 1) and 395,948 USD for LCC_2 (scenario 2 - including CAPEX and replacement/decommissioning costs incurred during and at the end of the system lifetime). The fact that $LCC_2 < LCC_1$ confirms that scenario 2 provides life-cycle cost savings even after accounting for initial CAPEX. The 15-year life-cycle cost (LCC) analysis compares the total present-value costs of the two scenarios at a discount rate of r = 9% per year. This rate lies within the range of hurdle rates, 7.5–10%, recommended for mature renewable energy technologies in Viet Nam [23] (see also [20, 21]); because wave energy is not yet a mature technology, this range is used only as a reference. A mid-life equipment replacement cost, estimated at 10% of CAPEX (23,000 USD), is incurred in Year 8. This reference value of 10% of CAPEX equals the decommissioning cost proposed for wave energy projects [24] and is applied as a conservative upper bound, because the partial replacement of electronic components and sensors involves less work than the removal of the entire device and its mooring system. It is also of the same order of magnitude as the operating costs reported for wave energy projects, whose annual OPEX ranges from 1.5% to 9% of CAPEX, with 5% of CAPEX proposed as a reference value [24]. In Year 15, the system incurs an additional decommissioning cost of 10% of CAPEX (23,000 USD), in line with the reference value for wave energy projects at early stages of development [24]. Based on the annual net cash flows, the project achieves a positive NPV and an IRR above the minimum discount rate, indicating financial feasibility under the stated assumptions (Table 9).

Têtu and Fernández Chozas [24] define the decommissioning cost as comprising the costs of removing the wave energy converter and its foundation or mooring system and consider 10% of CAPEX a reasonable estimate for this item, based on experience in the offshore wind industry [25] and on the LiftWEC cost database [26]. This percentage is applied here as a CAPEX-based reference value; no measured decommissioning or replacement costs are available for Chu Lai.

$$NCF_8 = NCF - 0.10 \cdot CAPEX = 40{,}668.02 - 23{,}000 = 17{,}668.02 \quad (18)$$

$$NCF_{15} = NCF - 0.10 \cdot CAPEX = 40{,}668.02 - 23{,}000 = 17{,}668.02 \quad (19)$$

In Equations (18) and (19), NCF = 40,668.02 USD is the annual operating cash flow; NCF8 and NCF15 are the net cash flows of Years 8 and 15 (USD) after deducting the mid-life equipment replacement cost and the decommissioning cost, respectively, each equal to 10% of CAPEX = 23,000 USD.

Table 9. Summary of financial indicators

Financial indicator	Value	Interpretation
NPV (r = 9%)	+79,955 (USD)	Positive NPV – profitable project
IRR	14.76 (%)	IRR > r = 9% - financially feasible under the stated assumptions
Simple Payback Period	5.66 (years)	Investment recovered before Year 6
Discounted Cash Flow Payback Period (DCF PP)	8.88 (years)	Discounted payback before Year 9
LCC_1 (scenario 1)	475,903 (USD)	Total LCC of Scenario 1
LCC_2 (scenario 2)	395,948 (USD)	Total LCC of Scenario 2
LCC savings	79,955 (USD)	Cost advantage of Scenario 2

At the detection performance measured on the simulated test set (mean P = 0.845, R = 0.864; Section 3.1), the same cash-flow model gives NCF = 39,398 USD per year, an NPV of +69,719 USD, an IRR of 14.1%, a simple payback period of 5.84 years and a discounted payback period of 9.32 years.

Figure 6 shows that the cumulative discounted cash flow of the IoT-PdM system starts at approximately -230,000 USD at the initial investment and reaches the break-even point in Year 8.88. After this point, the cumulative discounted cash

flow becomes positive and reaches a terminal NPV of +79,955 USD in Year 15. These results indicate that the system can recover its initial investment based on discounted cash flows within the expected operating lifetime.

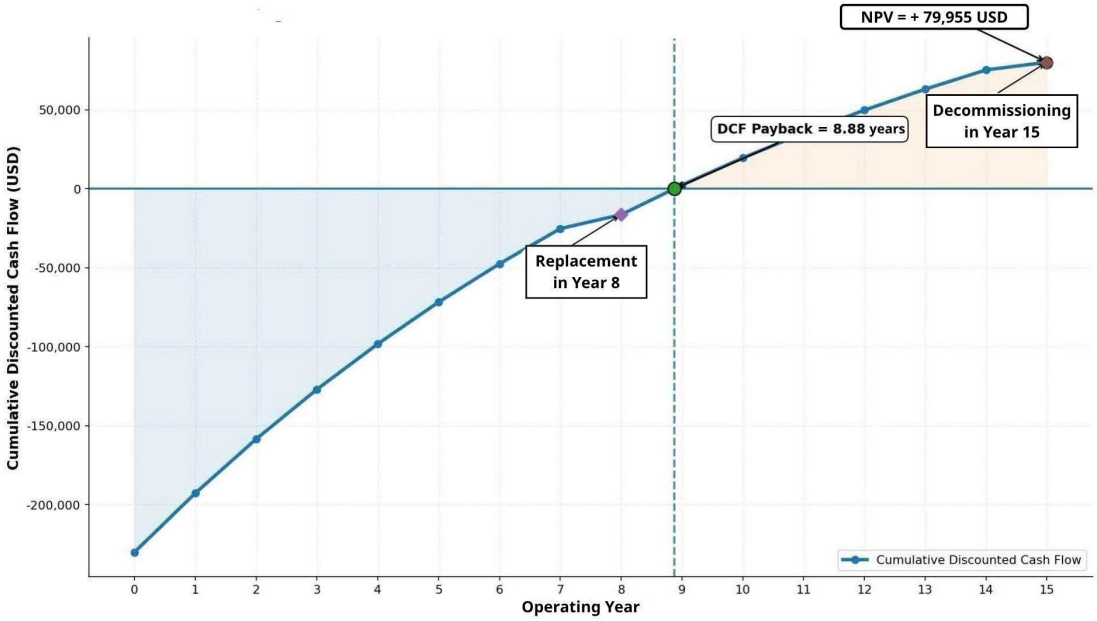


Figure 6. Cumulative discounted cash flow and discounted payback period over 15 years

The annual cash flows underlying Table 9 and Figure 6 are given in Table 10.

$$t_{SP} = 5 + \frac{26{,}660}{40{,}668} = 5.66\ years \quad (20)$$

In Equation (20), t_{SP} is the simple payback period, obtained by linear interpolation within Year 6, the year in which the cumulative undiscounted NCF changes from negative to positive: 26,660 USD is the part of CAPEX still to be recovered at the end of Year 5, and 40,668 USD is the NCF of Year 6. In Equation (21), t_{DP} is the discounted payback period, obtained by the same interpolation on the cumulative discounted NCF at $r = 9\%$: 16,453 USD is the amount still to be recovered at the end of Year 8, and 18,725 USD is the discounted NCF of Year 9.

$$t_{DP} = 8 + \frac{16{,}453}{18{,}725} = 8.88\ years \quad (21)$$

Table 10. Annual cash flows of the IoT-PdM system (USD; r = 9%)

Year	Operating NCF	One-off costs	NCF	Cumulative NCF	Discount factor	Discounted NCF	Cumulative DCF
0	–	CAPEX: −230,000	−230,000	−230,000	1.0000	−230,000	−230,000
1	40,668	–	40,668	−189,332	0.9174	37,310	−192,690
2	40,668	–	40,668	−148,664	0.8417	34,229	−158,460
3	40,668	–	40,668	−107,996	0.7722	31,403	−127,057
4	40,668	–	40,668	−67,328	0.7084	28,810	−98,247
5	40,668	–	40,668	−26,660	0.6499	26,431	−71,816

6	40,668	–	40,668	14,008	0.5963	24,249	−47,567
7	40,668	–	40,668	54,676	0.5470	22,247	−25,320
9	40,668	–	40,668	113,012	0.4604	18,725	2,272
10	40,668	–	40,668	153,680	0.4224	17,179	19,451
11	40,668	–	40,668	194,348	0.3875	15,760	35,211
12	40,668	–	40,668	235,016	0.3555	14,459	49,670
13	40,668	–	40,668	275,684	0.3262	13,265	62,935
14	40,668	–	40,668	316,352	0.2992	12,170	75,104
15	40,668	−23,000	17,668	334,020	0.2745	4,851	79,955

Note: Operating NCF = 40,668.02 USD per year (≈ 40,000 USD when rounded). Year 8: mid-life replacement cost; Year 15: decommissioning cost; each equal to 10% of CAPEX = 23,000 USD. From the table: NPV = 79,955 USD, IRR = 14.76%, simple payback period = 5.66 years and discounted payback period = 8.88 years, consistent with Table 9.

5. Uncertainty and Sensitivity Analysis

5.1 Monte Carlo Simulation

Executing 10,000 Monte Carlo simulation scenarios to quantify the uncertainty in economic evaluation, sampling input variables ($N_f, C_{repair}, P_{grid}, P_{CO2}$, R, P, OPEX, CF) based on distributing in accordance with each variable's characteristics (Poisson, Triangular, Normal, Uniform, Beta). The distribution family and parameters of each variable are given in Table 11. Table 12 summarizes the statistics of the resulting NPV distribution, while Figure 7 illustrates the overall distribution of the simulated NPV values and the corresponding percentile thresholds.

Table 11. Input distributions of the Monte Carlo simulation (N = 10,000 runs)

Variable	Base value	Distribution – parameters	P10 – P90	Rationale
Failures per year, N_f	6	Poisson, $\lambda = 6$	3 – 9	Independent random events per unit time
Emergency repair cost, C_{repair}	6,000 USD	Triangular (min = 4,000; mode = 6,000; max = 9,000)	5,000 – 7,775	Minimum, most likely and maximum values from the literature
Grid electricity price, P_{grid}	90 USD/MWh	Normal ($\mu = 90$; $\sigma = 15$)	70.8 – 109.2	Symmetric market fluctuation
Carbon credit price, P_{CO2}	8 USD/tCO2	Uniform $[4, 15]$	5.1 – 13.9	Immature Vietnamese carbon market; no preferred value
Detector recall, R	0.89	Beta ($\alpha = 8$; $\beta = 1$), mean 0.89	0.75 – 0.99	Bounded in $[0, 1]$; classification probability
Detector precision, P	0.91	Beta ($\alpha = 9$; $\beta = 1$), mean 0.90	0.77 – 0.99	Bounded in $[0, 1]$; classification probability
IoT OPEX	10,000 USD/year	Normal ($\mu = 10{,}000$; $\sigma = 2{,}000$)	7,437 – 12,563	Symmetric uncertainty around the base OPEX

OWC capacity factor, CF_{OWC}	0.40	Uniform $[0.30, 0.50]$	0.32 – 0.48	Range reported for OWC plants in the literature

All variables were sampled independently. The Tornado analysis (Section 5.2) uses the P10 and P90 of each distribution in Table 11, with all other variables held at their base values. The simulation used N = 10,000 runs. The distributions of R and P are centered on the target detection performance of the economic base case (Table 6); the economic result at the detection performance measured in Section 3.1 is reported separately in Section 4.3.

Table 12. Statistics of the resulting NPV distribution (N = 10,000)

Statistics	Value (USD)	Interpretation
Mean NPV, E[NPV]	+94,826	Expected value
Standard deviation (σ)	126,442	Dispersion of NPV
Median NPV (P50)	+81,196	50% of scenarios exceed this value
Optimistic NPV (P90)	+263,462	The best 10% of scenarios exceed this value
Pessimistic NPV (P10)	−52,409	The worst 10% of scenarios fall below this value
5th percentile (P5)	−83,406	5% of scenarios fall below this value
Probability of a positive NPV, P(NPV > 0)	75.9%	Share of scenarios with a positive NPV

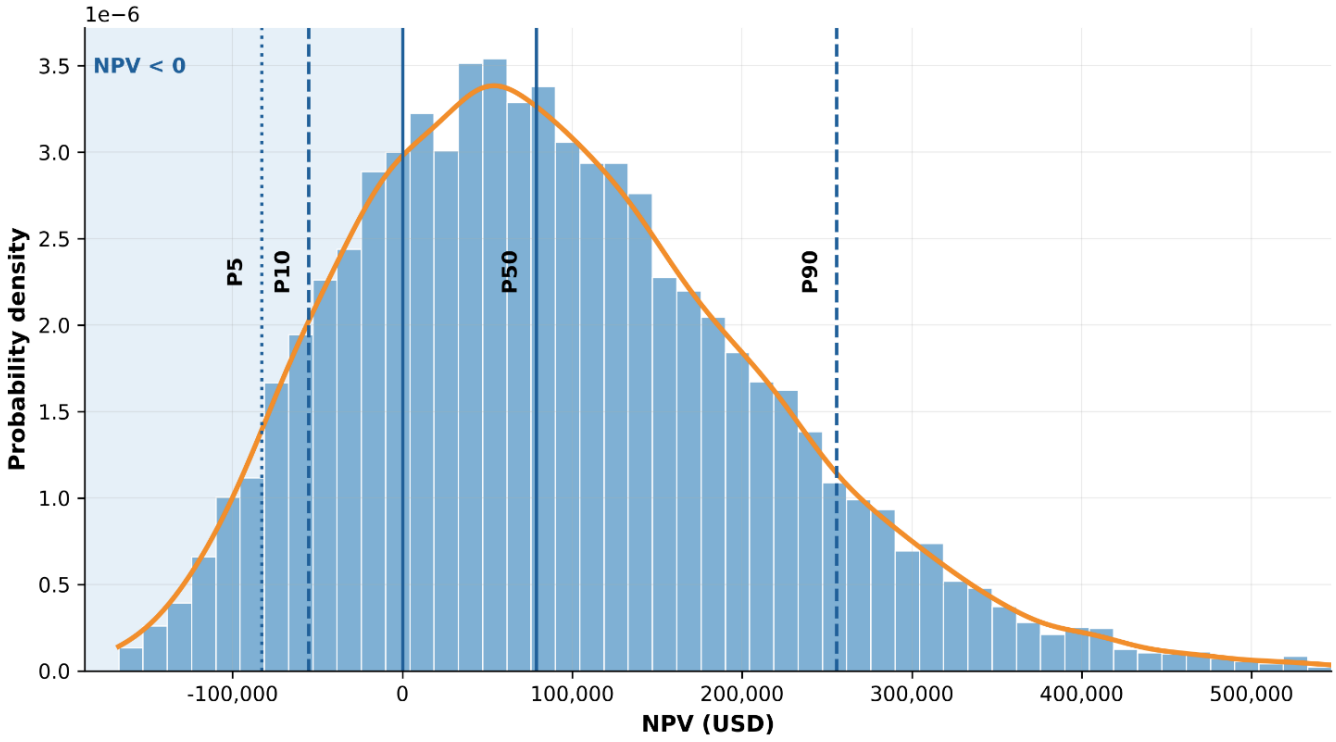


Figure 7. Distribution of NPV from 10,000 Monte Carlo runs with the input distributions of Table 11; vertical lines mark NPV = 0 and the 5th, 10th, 50th and 90th percentiles, whose values are given in Table 12

P(positive NPV) = 75.9% indicates that the project is profitable in about three quarters of the simulated scenarios under the stated input distributions. A negative NPV occurs mainly when the number of failures per year is low (4 or fewer), because the benefit of predictive maintenance scales with the number of avoided failures, and, to a lesser extent, when the recall of the detector is low (Figure 8). The 5th percentile of NPV is −83,406 USD.
The Monte Carlo and Tornado analyses were run from the same code and with the same input distributions.

5.2 Tornado Analysis

Tornado Chart quantifies the NPV Sensitivity with each input variable according to One-at-a-Time (OAT): each variable changes from P10 to P90 while the others maintain the base value (Figure 8).

As shown in Figure 8, the variables with the greatest impact on NPV are the number of failures per year Nf (253,654 USD), the emergency repair cost Crepair (119,458 USD) and the recall of the detector R (68,401 USD), followed by the IoT OPEX (41,321 USD), grid electricity price (26,000 USD), precision P (9,938 USD), OWC capacity factor (7,300 USD) and carbon credit price (2,747 USD); the NPV values at P10 and P90 of each variable are given in Table 13. Each bar in Figure 8 shows the NPV obtained when one input is set to its P10 (blue) or P90 (orange) value while all other inputs are held at their base values; the vertical line marks the base NPV of 79,955 USD, variables are ranked by the width of the NPV range, and the numbers at the bar ends are the input values at P10 and P90 (Table 11). For the IoT operating cost, the P90 input gives the lower NPV. NPV is most sensitive to the failure frequency, the emergency repair cost and the recall of the detector.

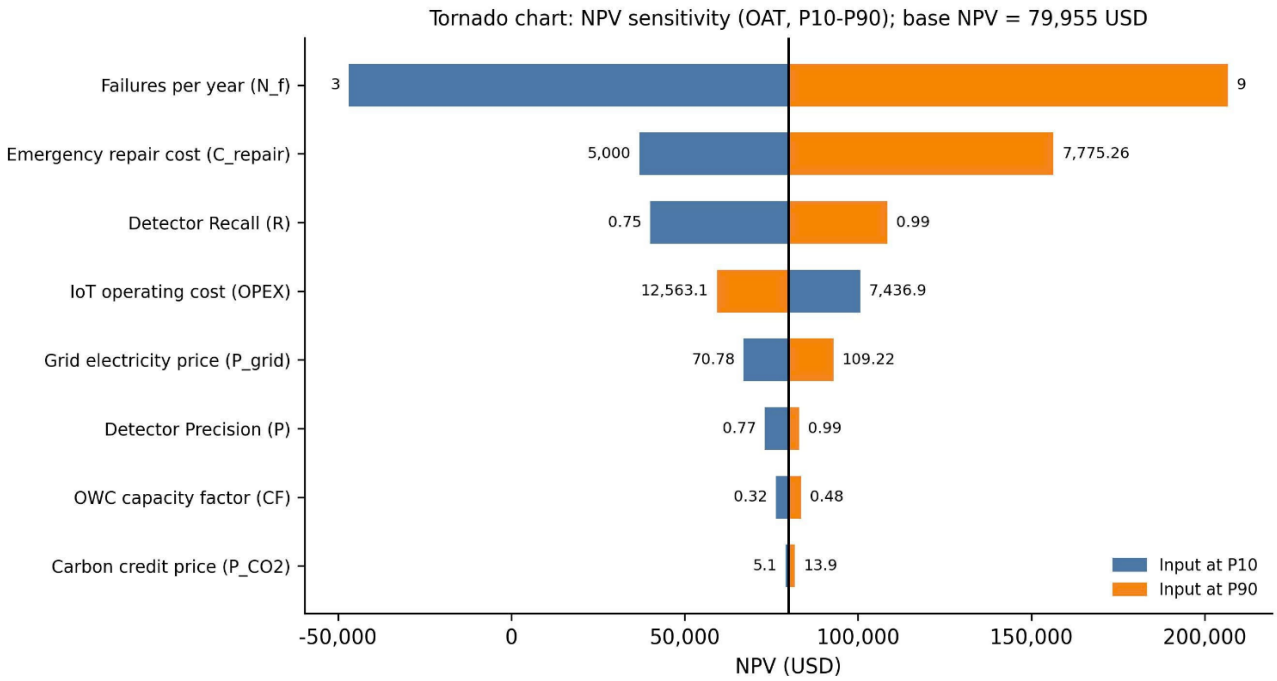


Figure 8. Tornado chart of NPV sensitivity (OAT, P10–P90; base NPV = 79,955 USD)

Table 13. Results of the Tornado analysis (OAT, P10–P90; base NPV = 79,955 USD)

Variable	P10	P90	NPV at P10 (USD)	NPV at P90 (USD)	ΔNPV (USD)
Failures per year, N_f	3	9	−46,872	206,782	253,654

Emergency repair cost, C_{repair}	5,000	7,775	36,911	156,369	119,458
Detector recall, R	0.75	0.99	40,054	108,455	68,401
IoT OPEX	7,437	12,563	100,615	59,295	41,321
Grid electricity price, P_{grid}	70.78	109.22	66,955	92,955	26,000
Detector precision, P	0.77	0.99	73,075	83,013	9,938
OWC capacity factor, CF	0.32	0.48	76,305	83,605	7,300
Carbon credit price, P_{CO2}	5.1	13.9	79,050	81,797	2,747

5.3 Integrated Analysis: Risks of reduced AI performance and market volatility

Integrating a stress-test scenario with two main risks - (1) a linear decline of Recall from 0.89 in Year 1 to 0.70 in Year 5, after which it stays at 0.70, and (2) a 20% decrease in the grid electricity price - results in $NPV_{stress} = +29{,}737\ USD$, with $P(NPV_{stress} > 0) = 62\%$ under the input distributions of Table 11. The project therefore remains profitable in about three in five stressed scenarios, highlighting the necessity of retraining the AI model periodically every 2-3 years to maintain Recall ≥ 0.85 and safeguard its economic advantages. The NPV remains positive as long as the recall of the detector exceeds about 0.61 at a precision of 0.91; the break-even recall is about 0.62 at a precision of 0.80 and about 0.64 at a precision of 0.70, with all other variables at their base values.

6. Conclusion

From a technical standpoint, the serial GRU-VAE + SR detector reached a mean F1-score of 0.854 and a mean AUC-ROC of 0.962 over 3 simulated datasets, comparable to, though slightly below, a standalone GRU-VAE (mean F1 0.868) and an LSTM autoencoder (mean F1 0.867). Spectral Residual was effective only when applied to the GRU-VAE reconstruction residual; applied to the raw signals, it lowered the ranking quality of the fused score. The SR stage did not shorten detection delays for any fault type. Pre-training on simulated Mutriku conditions and fine-tuning on 20% of the Chu Lai data increased the Average Precision obtained with limited site data (0.828 vs. 0.803; 0.848 with the full data), although this gain lies within one standard deviation over three simulated datasets.

From an economic standpoint, the improved fault-detection performance translates into concrete maintenance benefits: a 68.4% reduction in annual maintenance cost, an NPV of +USD 79,955, and an IRR of 14.76% over the 15-year project horizon. Monte Carlo simulation indicates a 75.9% probability of a positive NPV under the stated input distributions, while the Tornado sensitivity analysis shows that NPV is most sensitive to the failure frequency, the emergency repair cost and the recall of the detector, so that maintaining detection performance is essential to the economic benefit.

Several limitations should be acknowledged. First, all detection results are obtained on simulated data, because no OWC system is installed at Chu Lai; nominal operating values are assumptions, and wave-driven channels are scaled linearly by the wave-height ratio, although wave power scales approximately with the square of the wave height. Second, each test set contains only 4–7 contiguous fault events (17 over the three simulated datasets), so the reported detection metrics carry substantial statistical uncertainty. Third, faults are detected after onset (mean delay of a few seconds) rather than predicted in advance, and degradation processes such as biofouling and corrosion are represented by accelerated proxies lasting 12.5–50 s; the economic benefit of converting emergency repairs into planned repairs therefore assumes that detection occurs early enough for planned intervention. Fourth, the 0.25 s sampling interval is finer than typical minute-level SCADA logging, and the 5 s window is shorter than one wave period (7.1 s). Fifth, the maintenance cost parameters, the failure frequency ($N_f = 6$ per year), the downtime durations, the electricity price, and the mid-life replacement and decommissioning costs (10% of CAPEX each) were derived from literature-based estimates rather than field-measured operational data at Chu Lai; the reported NPV, IRR and payback periods should therefore be regarded as indicative. The results are most sensitive to the failure frequency, the emergency repair cost and the recall of the detector (Section 5.2), and the NPV becomes negative when the failure frequency falls to 4 per year or fewer. Sixth, the uptime benefit and the carbon credit revenue were computed at the rated capacity POWC = 4.4 MW, which is an upper

bound on the energy recovered during avoided downtime; with CFOWC = 0.4 these two terms would be about 2,115 and 124 USD per year, reducing the NPV by about 27,000 USD (to about 52,900 USD), which remains positive. Seventh, the uptime benefit (Eq. (10)) and the avoided loss of revenue during downtime (Eq. (16)) are both evaluated over the avoided downtime hours and may partly overlap; the reported NPV should therefore be read as an upper estimate. Finally, the model does not account for AI performance degradation as a continuous process over time. Future work should therefore validate these cost assumptions against on-site field data, extend the framework with an adaptive retraining mechanism, and assess its scalability to alternative OWC configurations.

Appendix

Table S1. Pre-training on Mutriku conditions and fine-tuning for Chu Lai (simulated data; Chu Lai test set; F1-max threshold selected on the validation set; mean ± standard deviation over three seeds)

Configuration	Chu Lai data	Precision	F1	AP	AUC-ROC
Training from scratch	100%	0.848 ± 0.025	0.868 ± 0.023	0.848 ± 0.040	0.963 ± 0.019
Training from scratch	20%	0.849 ± 0.021	0.867 ± 0.021	0.803 ± 0.008	0.962 ± 0.018
Mutriku model, no fine-tuning	0%	0.850 ± 0.022	0.868 ± 0.023	0.838 ± 0.049	0.963 ± 0.020
Mutriku pre-training + fine-tuning	20%	0.842 ± 0.032	0.864 ± 0.026	0.828 ± 0.036	0.962 ± 0.020

Table S2. Configuration exactly as originally submitted (Spectral Residual computed over the whole data segment, non-causal; parallel fusion with α = 0.5); test set; mean over three seeds

Method	Rule	Precision	Recall	F1	FPR	Events	AP	AUC-ROC
GRU-VAE + SR, full-series SR, $\alpha = 0.5$ (as submitted)	P99.5	0.881	0.368	0.516 ± 0.059	0.006	11/17	0.831 ± 0.039	0.963 ± 0.019
	F1max	0.851	0.852	0.851 ± 0.030	0.017	17/17		

Note: Precision, Recall and FPR are means over three seeds; F1, AP and AUC-ROC are means ± standard deviations; "Events" as in Table 2.

References

[1] T.T. Tung, N.Q. Chien. An assessment of potential wave power along a coastal province, Central Vietnam. Int. Marine Energy J. 2023; 6(1):27-35.

[2] F.T. Liu, K.M. Ting, Z.-H. Zhou. Isolation Forest. In: Giannotti F, Gunopulos D, Turini F, Zaniolo C, Ramakrishnan N, Wu X, editors. Proceedings of the Eighth IEEE International Conference on Data Mining. 15-19 December 2008; Pisa, Italy. Los Alamitos, CA: IEEE; 2008. p. 413-422.

[3] B. Schölkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, R.C. Williamson. Estimating the support of a high-dimensional distribution. Neural Comput. 2001; 13(7):1443-1471.

[4] D.P. Kingma, M. Welling. Auto-Encoding Variational Bayes. arXiv:1312.6114 [stat.ML]. 2013.

[5] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In: Moschitti A, Pang B, Daelemans W, editors. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). October 2014; Doha, Qatar. Stroudsburg, PA: Association for Computational Linguistics; 2014. p. 1724-1734.

[6] H. Ren, B. Xu, Y. Wang, C. Yi, C. Huang, X. Kou, T. Xing, M. Yang, J. Tong, Q. Zhang. Time-Series Anomaly Detection Service at Microsoft. In: Teredesai A, Kumar V, Li Y, Rosales R, Terzi E, Karypis G, editors. Proceedings of

the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 4-8 August 2019; Anchorage, AK, USA. New York, NY: Association for Computing Machinery; 2019. p. 3009-3017.

[7] Ente Vasco de la Energía (EVE). The Mutriku wave plant achieves cumulative electricity production of three million kilowatts per hour [press release]. 11 January 2024. Information on: https://www.eve.eus/en/news/news/the-mutriku-wave-plant-achieves-cumulative-electricity-production-of-three-million-kilowatts-per-hour

[8] M.H. Jahangir, S. Bahrizadeh. Life-Cycle Cost Assessment of Various Wave Energy Converters (Based on Energy and Carbon Intensity). In: Muthu SS, editor. Life Cycle Costing. Cham: Springer; 2023. p. 117-151.

[9] E. Trentini, T.C. da Silva, L. Melo Junior, J.F. de Macêdo. Model-centered Ensemble for Anomaly Detection in Time Series. In: Proceedings of the 12th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART. 2020. SciTePress; 2020. p. 700-707.

[10] D. Umsonst, J. Ruths, H. Sandberg. Finite Sample Guarantees for Quantile Estimation: An Application to Detector Threshold Tuning. IEEE Trans. Control Syst. Technol. 2023; 31(2):921-928.

[11] J. Davis, M. Goadrich. The Relationship between Precision-Recall and ROC Curves. In: Cohen WW, Moore A, editors. Proceedings of the 23rd International Conference on Machine Learning. 25-29 June 2006; Pittsburgh, PA, USA. New York, NY: Association for Computing Machinery; 2006. p. 233-240.

[12] J. Fernández Chozas, P. Johannesson, J. Cruz, P. Ruiz-Minguela, P. Etxaniz. Reliability and Cost Assessment of Critical Components: Electrical Generator Failure of IDOM Wave Energy Converter. In: Proceedings of the 15th European Wave and Tidal Energy Conference. 3-7 September 2023; Bilbao, Spain. EWTEC; 2023.

[13] S. Ambühl, L. Marquis, J.P. Kofoed, J.D. Sørensen. Operation and Maintenance Strategies for Wave Energy Converters. Proc. Inst. Mech. Eng. Part O: J. Risk Reliab. 2015;229(5):417-441.

[14] M. Borsotti, X. Jiang, R.R. Negenborn. Stochastic Optimization of Predictive Maintenance Scheduling for Offshore Wind Farms. Ocean Eng. 2026; 357:125424.

[15] Government of Vietnam. Decree No. 29/2026/ND-CP on the domestic carbon exchange. 2026.

[16] Department of Climate Change, Ministry of Natural Resources and Environment. Official Dispatch No. 1726/BĐKH-PTCBT on the grid emission factor of Vietnam in 2023. 2024..

[17] Ministry of Industry and Trade of Vietnam. Decision No. 1279/QD-BCT on electricity retail tariffs. 2025.

[18] L. Martinelli, P. Pezzutto, P. Ruol. Experimentally Based Model to Size the Geometry of a New OWC Device, with Reference to the Mediterranean Sea Wave Environment. Energies. 2013; 6(9):4696-4720.

[19] F. M'zoughi, J. Lekube, A.J. Garrido, M. De La Sen, I. Garrido. Machine Learning-Based Diagnosis in Wave Power Plants for Cost Reduction Using Real Measured Experimental Data: Mutriku Wave Power Plant. Ocean Eng. 2024; 293:116619.

[20] L.P. Truong, V.D. Nguyen, P.L. Le. Techno-Economic Analysis of Solar Power Plant Project in Binh Thuan, Vietnam. In: Proceedings of the 2018 4th International Conference on Green Technology and Sustainable Development (GTSD). 2018. p. 82-85.

[21] L. Castro-Santos, A. Filgueira-Vizoso, X. Costoya, B. Arguilé-Pérez, A.S. Ribeiro. Economic Viability of Floating Wave Power Farms Considering the Energy Generated in the Near Future. Renew. Energy. 2024; 222:119947.

[22] G.P. Sullivan, R. Pugh, A.P. Melendez, W.D. Hunt. Operations & Maintenance Best Practices: A Guide to Achieving Operational Efficiency. 3rd ed. Richland, WA: Pacific Northwest National Laboratory; 2010. Chapter 5, Types of Maintenance Programs. p. 5.1-5.6.

[23] B. Coleman. Assessment of a Social Discount Rate and Financial Hurdle Rates for Energy System Modelling in Viet Nam. OECD Environment Working Papers, No. 181. Paris: OECD Publishing; 2021.

[24] A. Têtu, J. Fernández Chozas. A Proposed Guidance for the Economic Assessment of Wave Energy Converters at Early Development Stages. Energies. 2021; 14(15):4699.

[25] M.J. Kaiser, B. Snyder. Modeling the decommissioning cost of offshore wind development on the U.S. Outer Continental Shelf. Mar. Policy. 2012; 36(1):153-164.

[26] A. Têtu, J. Fernández Chozas. Database of costs for wave energy projects [dataset]. LiftWEC project; 2021. Information on https://doi.org/10.5281/zenodo.4442080