

Received 11 August 2026, accepted 29 August 2026, date of publication 2 September 2026, date of current version 10 September 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3730232

## RESEARCH ARTICLE

# Performance Evaluation of Wave Energy Converters: A Data-Driven Case Study Using Supervised Learning and Stochastic Frontier Analysis

DANIEL COSTA<sup>id</sup>, (Student Member, IEEE), AND ALEXANDRE LUCAS<sup>id</sup>

INESC TEC, 4200-465 Porto, Portugal

Corresponding authors: Daniel Costa (up202207519@up.pt) and Alexandre Lucas (alexandre.lucas@inesctec.pt)

This work was supported by the Horizon Europe Research and Innovation Programme Federated Common European-scale Energy AI Testing and Experimentation Facilit (EnerTEF) under Grant 101172887.

**ABSTRACT** The commercial viability of Wave Energy Converters (WECs) depends heavily on reducing operation and maintenance costs by accurately detecting mechanical degradation, such as Power Take-Off (PTO) failures. A major challenge in the current state of the art is distinguishing true mechanical faults from false alarms triggered by a typical metocean conditions. This paper presents a test site case study proposing a three-phase data-driven methodology for the performance analysis of a WEC fleet. Due to early-stage development sensitivity, the study utilizes a semi-empirical dataset where both nominal metocean conditions and WEC power generation are extracted directly from real offshore telemetry, while environmental penalties and mechanical faults are introduced computationally to evaluate algorithmic robustness. The methodology integrates two key performance analysis models: a supervised learning approach and a parametric model. First, an Extreme Gradient Boosting (XGBoost) algorithm acts as a regressor to establish an absolute power generation baseline based on wave inputs. Second, a Cobb-Douglas Stochastic Frontier Analysis (SFA) evaluates the relative technical efficiency of the assets. The SFA decomposes the composite residual, isolating one-sided mechanical inefficiencies from symmetric environmental noise, such as spectral spreading. Finally, a deterministic decision engine merges the absolute power deficits with the relative SFA scores. Results demonstrate that this architecture effectively mitigates environmental false positives and identifies critical mechanical faults, providing a promising framework for offshore maintenance dispatch. To promote reproducibility and open science within the marine energy sector, the complete codebase developed for this study is made publicly available.

**INDEX TERMS** Condition monitoring, machine learning, predictive maintenance, stochastic frontier analysis, wave energy converters, XGBoost.

## I. INTRODUCTION

Ocean wave energy represents one of the most energy-dense and geographically predictable renewable resources available, with global estimates suggesting a theoretical potential of approximately 26,000 TWh per year [1]. Despite this abundance, wave energy technology remains at the pre-commercial threshold, largely because the Levelized Cost of Energy (LCOE) for current deployments cannot yet compete

with mature offshore wind or utility-scale solar. Among the primary cost drivers, Operation and Maintenance (O&M) expenditures are disproportionately large: in some reported WEC systems, maintenance activities alone account for up to 57% of total operational expenses [1]. Reducing this burden through intelligent, data-driven condition monitoring is therefore not merely an engineering optimization problem; it constitutes a commercial prerequisite for the transition of wave energy from prototype testing to grid-scale deployment.

Point Absorber WECs rely on mechanically complex, failure-prone Power Take-Off (PTO) systems to convert

The associate editor coordinating the review of this manuscript and approving it for publication was Dinesh Kumar.

wave motion into electrical energy. While this study utilizes a foundational case study inspired by CorPower Ocean's C4 technology [2], the proposed data-driven diagnostic methodology is strictly technology-agnostic. PTO degradation directly manifests as suppressed electrical power output. However, detecting these mechanical signatures early enough to trigger cost-effective maintenance is challenging because the wave-to-power relationship is highly non-deterministic. In harsh marine environments, power generation is heavily mediated by spectral shape, directional spreading, and resonance, a physical complexity that forms the core diagnostic problem addressed in this work.

### A. THE ENVIRONMENTAL FALSE-POSITIVE PROBLEM

The fundamental challenge in WEC condition monitoring is environmental false positives: alarms triggered by sub-optimal metocean conditions rather than mechanical degradation. Point absorbers maximize their Capture Width Ratio (CWR) through resonance with incident wave frequencies [3]. During atypical conditions, such as spectral spreading or confused seas, wave energy is distributed across broad frequency bands. Lacking a dominant resonant frequency, the conversion efficiency of the entire array drops symmetrically, aligning with empirical CWR boundaries [4]. Consequently, diagnostic systems evaluating assets in isolation misinterpret this fleet-wide environmental penalty as widespread mechanical failure, triggering prohibitively expensive vessel dispatches.

Left unfiltered, these environmentally induced false-positive rates render automated maintenance dispatch operationally unviable. Existing diagnostic approaches struggle precisely because they lack the spatial awareness required to determine whether a generation anomaly is isolated to a specific WEC or shared uniformly across the fleet.

### B. LIMITATIONS OF THE CURRENT STATE-OF-THE-ART

The bulk of offshore condition monitoring literature draws directly from established wind turbine diagnostic paradigms [1]. Consequently, normal behaviour models are widely implemented to raise anomaly flags whenever generation residuals exceed statistical thresholds [5].

However, while useful for rotary machinery, these absolute models carry critical structural limitations for WECs. By evaluating assets independently, they remain inherently blind to fleet-wide metocean deviations at inference time.

Furthermore, conventional methods treat the prediction residual as a single composite signal, conflating symmetric environmental perturbations with asymmetric mechanical losses. Applying static thresholds to this conflated signal inevitably triggers false positives under adverse seas or obscures early-stage faults. To the best of the authors' knowledge, no published WEC framework has formalized a real-time statistical decomposition of this composite residual, nor demonstrated explicit false-positive suppression for automated O&M dispatch.

### C. PROPOSED APPROACH AND CONTRIBUTIONS

To address these limitations, this paper proposes a three-phase data-driven architecture for WEC fleet condition monitoring, combining supervised machine learning with parametric econometric modelling. Evaluated on a semi-empirical array dataset anchored in real offshore telemetry, the proposed topology-agnostic architecture shows potential applicability to other fleets sharing a common metocean environment.

The primary contributions of this work are summarized as follows:

- 1) **Hybrid Diagnostic Architecture:** A multi-phase framework that orthogonally merges an Extreme Gradient Boosting (XGBoost) regressor, establishing an absolute generation baseline with  $R^2 = 0.9038$  on healthy fleet data, with a Cobb-Douglas Stochastic Frontier Analysis (SFA) to extract relative technical efficiency across the fleet in real time.
- 2) **Statistical Isolation of Environmental Noise:** The application of SFA to WEC condition monitoring. This approach mathematically decomposes the composite prediction residual into a symmetric meteorological noise component ( $v$ ) and a one-sided inefficiency component ( $u$ ), providing an estimate of inefficiency that is interpreted as potential PTO degradation, distinct from environmental volatility.
- 3) **Deterministic False-Positive Suppression:** A four-state decision matrix that filters alarms triggered by sub-optimal wave spectra. Integrating absolute power deficits, relative SFA scores, and a temporal threshold helps separate transient environmental noise from sustained mechanical degradation, guiding targeted maintenance.

The remainder of this paper is organized as follows. Section II reviews the relevant literature on data-driven WEC monitoring, frontier efficiency methods, and hybrid diagnostic architectures. Section III details the scope, physical constraints, and semi-empirical dataset processing. Section IV introduces the mathematical formulation of the three-phase methodology. Section V presents and discusses the operational results. Finally, Section VI provides concluding remarks and outlines future work.

## II. RELATED WORK

The literature on WEC condition monitoring [1] and operational performance optimization [6] spans three parallel streams: (i) machine learning for fault detection, (ii) frontier methods for efficiency benchmarking, and (iii) hybrid diagnostic architectures. This review situates our work at their intersection, motivating the proposed methodology.

### A. DATA-DRIVEN FAULT DETECTION AND POWER PREDICTION IN WECs

Proliferating SCADA telemetry has accelerated machine learning applications for WEC health monitoring. However, a comprehensive taxonomy by Mortazavizadeh et al. [1]

concludes that adapting wind energy techniques to WECs introduces unique diagnostic challenges, primarily due to the complex hydrodynamic coupling between wave spectra and PTO mechanics.

Recent literature confirms that simplified spectral representations systematically bias power predictions. For instance, assuming an idealized single-peak JONSWAP spectrum can overestimate annual production by 16.5% [7]. Furthermore, multi-directional seas dampen power fluctuations across point-absorbing arrays [8] and individual platforms [9] compared to unidirectional assumptions. Ignoring the disturbed wave field [10] and the macroscopic impact of energy extraction [11] further distorts array generation assessments. Methodologically, linear frequency-domain models overpredict power in irregular seas by missing non-linear viscous limits [12], necessitating second-order random wave formulations for accurate prediction [13].

Consequently, single-asset baselines failing to account for climate variability [14] inevitably misinterpret multi-modal seas as mechanical under-performance. Masoumi et al. [5] highlight this critical gap: most literature targets nominal power prediction, leaving fault isolation under environmental disturbances unaddressed. By operating in isolation, existing models forfeit the fleet-level context necessary to distinguish localized mechanical failures from fleet-wide meteorological events.

Within the supervised learning domain, gradient-boosted tree ensembles have emerged as the preferred baseline for tabular SCADA prediction tasks [15] due to their robust handling of complex non-linear feature interactions [16]. Consequently, XGBoost-anchored ensembles achieve state-of-the-art accuracy on multi-device WEC power prediction [17].

Deep learning alternatives, principally Long Short-Term Memory (LSTM) architectures [18], are frequently proposed for sequence-aware marine forecasting [5]. Additionally, convolutional networks are leveraged for wave power prediction [19] and spectral feature classification [20].

Nevertheless, the strict interpretability limitations of closed box algorithms [21] combined with their vulnerability to distributional shifts make them unsuitable for real-time condition monitoring frameworks [22].

## B. FRONTIER METHODS FOR OFFSHORE ENERGY EFFICIENCY ASSESSMENT

Complementing absolute power prediction, econometric research addresses relative technical efficiency using production frontier models. Stochastic Frontier Analysis (SFA), formulated by Aigner et al. [23], decomposes the gap between observed output and a theoretical frontier into two distinct components: symmetric random noise (capturing environmental volatility and measurement errors) and a non-negative inefficiency term (capturing mechanical waste). This property intrinsically separates spectral spreading penalties from physical PTO degradation in WECs.

Across the broader energy sector, SFA is widely utilized to benchmark technical efficiency, ranging from macro-economic assessments [24] and regional energy indices [25] to the operational evaluation of mining enterprises [26]. Within offshore renewables, applications remain strictly confined to wind energy. Iglesias et al. established a methodological precedent by extracting efficiency scores directly from wind farm SCADA data [27]. Subsequent research has utilized panel-level SFA to evaluate the technical efficiency of national wind industries [28] and demonstrated the robustness of parametric frontiers in assessing the long-term resilience of offshore wind farms [29].

Efficiency estimates remain highly sensitive to chosen functional forms and distributional assumptions [30]. Although semiparametric extensions like StoNED relaxes these constraints via flexible shaping [31], and hybrid models address complex non-linear dynamics [32], standard parametric SFA is consistently preferred over deterministic methods like DEA due to its superior resilience to measurement noise in harsh environments [33].

Despite these advancements, transferring frontier methods to wave diagnostics remains unexplored, a gap consistently noted in fault diagnosis literature [1] and ocean energy surveys [5]. While prior studies restrict SFA to industry-wide empirical assessments [28] or long-run offshore wind audits [29], this architecture uniquely exploits dynamic v-u decomposition as an instantaneous, real-time diagnostic signal for automated O&M dispatch.

## C. HYBRID DIAGNOSTIC ARCHITECTURES AND THE ENVIRONMENTAL FALSE-POSITIVE GAP

A principal limitation of existing approaches is their reliance on single modelling paradigms. Purely absolute models flag any statistical deficit, confounding mechanical wear with fleet-wide metocean events. Conversely, purely relative models (like standalone SFA) ignore absolute generation levels, potentially assigning nominal efficiency scores to a fleet degrading uniformly under an environmental shock.

In offshore renewables, particularly wind operations [34], predictive maintenance is transitioning toward AI-driven hybrid architectures [35]. This shift addresses the economic impact of false alarms; as May and McMillan originally noted [36], false-positive alerts trigger high-risk, unnecessary vessel deployments that erode monitoring benefits. This vulnerability is corroborated by surveys on AI-based mitigation [37] and broader reviews of offshore maintenance trends [38].

To mitigate this, recent literature integrates machine learning with statistical filters to cross-validate alarms. Specifically, k-Nearest Neighbour frameworks [39] and broader ensemble models [40] successfully identify substantial false positive rates in SCADA logs. Furthermore, normal behavioural boundaries are increasingly refined via semi-supervised deep learning [41] and statistical techniques tailored for highly imbalanced operating states [42].

However, analogous WEC monitoring approaches remain scarce because the false alarm problem is structurally distinct. While wind turbine false alarms often originate from sensor noise or internal transients [43], WEC false alarms are predominantly environmental. Fundamental extraction efficiency, governed by the Capture Width Ratio [3], can be suppressed fleet-wide without localized mechanical failure. These uniform power drops are typically driven by atypical conditions, such as broad spectral bandwidths restricting ideal energy capture [7], or short-crested seas dampening the array's collective response [8].

This systemic environmental vulnerability exposes a critical gap: single-asset diagnostics lack the spatial awareness required to identify fleet-wide efficiency drops. The proposed architecture addresses this by orthogonally combining an absolute XGBoost baseline with a relative Cobb-Douglas SFA frontier into a unified fleet-level pipeline. Fusing these complementary signals successfully differentiates uniform environmental penalties from isolated mechanical faults, aiding in overriding fleet-wide false alarms while triggering targeted maintenance dispatch. No prior WEC literature has combined supervised machine learning with SFA in this manner, nor explicitly demonstrated the quantitative suppression of spectral spreading false positives for automated O&M decisions.

### III. SCOPE AND CORE ASSUMPTIONS

The methodology relies on strictly defined physical and statistical boundaries to ensure diagnostic accuracy. Due to early-stage development sensitivity at the offshore test site, this study utilizes a semi-empirical dataset anchored in real metocean telemetry. Sourced directly from an operational Waverider buoy between March and June 2026, the data provides an authentic hydrodynamic foundation to model twelve C4-inspired point-absorber WECs deployed within a 6 km radius. This framework preserves genuine ocean temporal dynamics, establishing a representative environment to evaluate the diagnostic pipeline.

#### A. EXPERIMENTAL DESIGN AND TEMPORAL EPOCHS

To assess the condition monitoring architecture, the dataset is temporally segmented into three distinct operational epochs,

evaluating the 12 WECs over a high-frequency 30-minute resolution:

- **Epoch 1 (Nominal Operation / Golden Period):** Represents the baseline period where all WECs operate fault-free. This training set is strictly isolated using operational metadata, ensuring the absolute predictive models and relative econometric frontier are calibrated exclusively on confirmed healthy telemetry.
- **Epoch 2 (Environmental False Positive):** A testing period characterized by sub-optimal metocean conditions affecting the entire fleet uniformly. A 15% drop in energy extraction is simulated across all assets to evaluate the system's resilience against false alarms triggered by environmental factors.
- **Epoch 3 (Critical Mechanical Fault):** The final testing period where severe mechanical degradation (e.g., PTO failure) is isolated to a specific cluster of assets. Buoys 1 through 8 remain in healthy operation, while Buoys 9 through 12 suffer massive, localized efficiency drops. This evaluates the model's capacity to isolate broken assets within a mixed-health fleet.

#### B. SFA BOUNDARY CONSTRAINTS

The studied architecture monitors WECs with a nominal capacity limit of 350 kW [44]. The parametric formulation of the Cobb-Douglas Stochastic Frontier Analysis (SFA), being inherently log-linear, requires strictly positive input vectors to evaluate generation elasticity. To prevent mathematical divergence during Maximum Likelihood Estimation (MLE) caused by logarithmically undefined values, operational records containing non-positive Wave Power Flux or electrical generation were excluded. This pre-processing filter dropped 4,380 unfeasible rows, yielding a refined dataset of 65,892 observations across the 12 buoys.

Furthermore, because operational power generation within the monitored dataset peaked at 239.32 kW, the empirical response resides entirely within the dynamic elastic ramp-up region of the PTO. Since the 350 kW physical curtailment plateau was never triggered by the incident wave spectra, the dataset naturally avoids the horizontal saturation asymptotes that traditionally distort log-linear MLE fitting. For future deployments operating frequently at saturation limits,

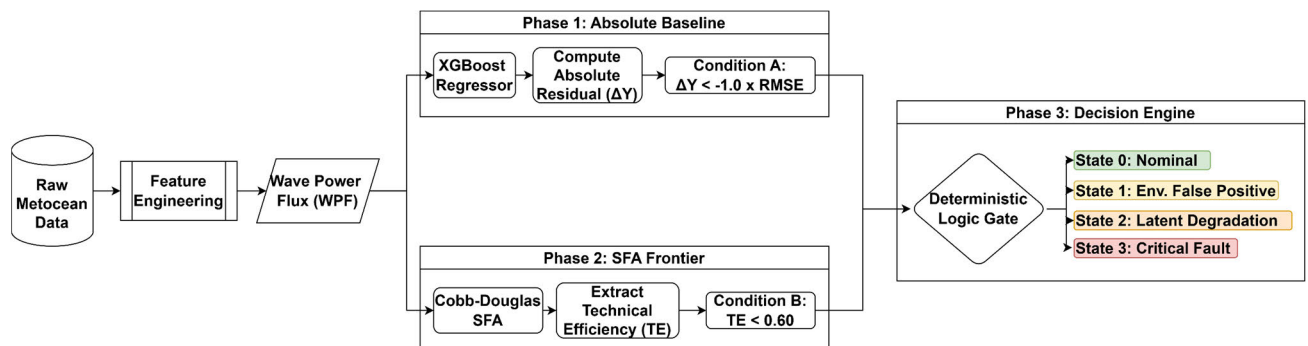


FIGURE 1. Architecture flow chart.

replacing the Cobb-Douglas curve with a Translog frontier could natively mitigate plateau distortions on inefficiency estimation [23].

### C. SPECTRAL SPREADING AS AN ENVIRONMENTAL FALSE-POSITIVE DRIVE

To test the decision matrix during Epoch 2, the dataset incorporates the hydrodynamic impacts of spectral spreading, representative of confused seas or multi-peaked spectra. Point-absorber WECs rely on mechanical resonance with dominant wave frequencies to maximize power extraction [3]. Under broad spectral bandwidths, the energy is distributed across multiple frequency bands rather than concentrated at the device's resonant frequency. Consequently, the Capture Width Ratio (CWR) experiences a pronounced drop despite the gross Wave Power Flux remaining high [4]. This environmentally induced penalty suppresses power generation uniformly across the entire farm. Embedding this physical phenomenon into Epoch 2 evaluates the architecture's capacity to recognize systemic, fleet-wide efficiency drops as benign environmental noise rather than misclassifying them as localized mechanical faults.

### D. DATA ACQUISITION AND FEATURE ENGINEERING

To ensure the data-driven models accurately capture the complex hydrodynamics of the WECs, rigorous data engineering was performed. The foundational ground, comprising high-frequency (30-minute) multi-spectral wave observations and the corresponding nominal WEC electrical power, was directly measured and extracted from operational offshore telemetry via API. To evaluate fleet-level condition monitoring, this single-asset baseline was computationally expanded into a simulated 12-WEC array using a spatial attenuation proxy. This proxy simulates the array's disturbed wave field and wake effects by iteratively applying a spatial scaling factor (ranging from 0.95 to 1.06) to the incident wave parameters and expected power output. Furthermore, an iterative temporal phase shift was applied to replicate wave propagation across the 6 km radius, alongside the injection of Gaussian micro-noise ( $\sigma=0.02$  for  $H_s$ ) to realistically decorrelate the environmental inputs across the assets.

Reflecting the harsh operational realities of offshore SCADA networks, the raw generated dataset of 70,272 observations contained severe sensor dropouts. Minor telemetry dropouts ( $\leq 2$  hours) were cured via piecewise cubic interpolation to preserve temporal continuity. Conversely, sustained structural data losses were flagged via operational metadata and subjected to a strict hard-drop policy, expelling 4,152 unfeasible observations. Furthermore, to satisfy the strict logarithmic requirements of the subsequent Stochastic Frontier Analysis (SFA), an additional 228 observations containing non-positive power or wave flux values were removed. This rigorous filtering cascade yielded a final, mathematically clean dataset of 65,892 observations prior to model training.

To prevent seasonal data leakage, explicit temporal variables (such as hour of day and month) were strictly excluded

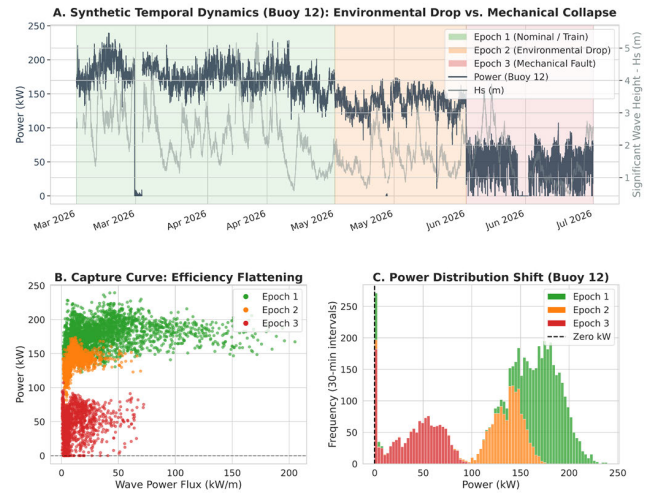


FIGURE 2. Temporal dynamics and power capture curve for Buoy 12.

from the feature space. By denying the model access to cyclical time anchors, the Extreme Gradient Boosting (XGBoost) algorithm is forced to map pure fluid mechanics rather than relying on temporal correlations. The expanded feature space utilizes 15 distinct variables, including macroscopic descriptors like Significant Wave Height ( $H_s$ ) and Energy Period ( $T_e$ ), alongside advanced statistical spectral data (e.g.,  $H_{max}$ ,  $H_{1/3}$ ,  $T_{avg}$ ).

To bridge the gap between pure statistical machine learning and physical oceanography, a domain-specific engineered feature was introduced: the Wave Power Flux (WPF). The WPF represents the raw hydrodynamic power available in the incident wave front per meter of wave crest (measured in kW/m) and is mathematically derived using the deep-water approximation:

$$\text{WPF} \approx 0.49 \times H_s^2 \times T_e \quad (1)$$

This physics-informed variable serves as the primary energy predictor, tightly linking the raw meteorological spectra directly to the mechanical absorption capabilities of the WEC fleet.

To rigorously evaluate the diagnostic framework, a chronological experimental design comprising three distinct operational epochs was embedded into this final 65,892-record dataset, as visualized in Fig. 2. Epoch 1 (32,880 observations) establishes the healthy baseline. Epoch 2 (17,472 observations) introduces a fleet-wide spectral spreading penalty, artificially suppressing the extraction efficiency uniformly across all 12 WECs (simulating a systemic environmental false positive). Finally, Epoch 3 (15,540 observations) simulates a severe PTO structural degradation exclusively within a target sub-cluster (Buoys 9–12), allowing the remaining healthy assets to recover. Fig. 2 illustrates this physical dynamic for a degraded asset (Buoy 12): the time-series demonstrates the temporary fleet-wide generation drop followed by a localized mechanical fault that significantly restricts its power extraction.

#### IV. PROPOSED THREE-PHASE METHODOLOGY

To accurately isolate mechanical faults from environmental noise, the proposed data-driven architecture is structured into three sequential phases: an absolute performance baseline, a relative efficiency evaluation, and a deterministic decision engine.

##### A. PHASE 1: ABSOLUTE ANALYSIS (DATA-DRIVEN BASELINE)

The first phase establishes an empirical baseline of healthy operations. Instead of relying on predefined theoretical power matrices, an Extreme Gradient Boosting (XGBoost) regressor is utilized as a data-driven predictive model. To prevent seasonal data leakage, temporal variables (e.g., month, hour) are strictly excluded from the feature space. Leveraging the physics-constrained feature space established during data acquisition, the model captures the non-linear mapping between macroscopic metocean conditions, primarily the Wave Power Flux (WPF), Significant Wave Height ( $H_s$ ), and Energy Period ( $T_e$ ), and the expected active power generation ( $Y$ ) of the WECs [17].

To optimize the bias-variance trade off and prevent overfitting on the highly volatile wave data, the XGBoost hyperparameters were strictly regularized. The model utilizes an ensemble of 500 decision trees ( $n\_estimators = 500$ ) with a maximum depth of 6 ( $max\_depth = 6$ ). To ensure robust generalization, stochastic gradient boosting techniques were applied, specifically row subsampling ( $subsample = 0.8$ ) and column subsampling by tree ( $colsample\_bytree = 0.8$ ), combined with a conservative learning rate ( $learning\_rate = 0.05$ ) and explicit regularization terms ( $reg\_alpha = 0.1$ ,  $reg\_lambda = 1.0$ ) to further penalize model complexity.

Rejecting blind chronological partitions, the training subset is strictly isolated using operational metadata from the Epoch 1 Golden Window. This guarantees that the algorithm learns exclusively from confirmed fault-free telemetry. The model's predictive accuracy is evaluated out-of-sample using three standard regression metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination ( $R^2$ ).

During inference, the absolute residual ( $\Delta Y_{i,t}$ ) for each asset  $i$  at timestamp  $t$  is computed. Instead of a hard-coded threshold, a dynamic confidence band is established using the historical RMSE extracted from the training calibration. In standalone operation, an absolute anomaly flag (Condition A) is triggered when the generation deficit strictly exceeds a  $-3\sigma$  error margin, acting as a robust Statistical Process Control (SPC) heuristic to isolate severe operational deviations:

$$\Delta Y_{i,t} < -3.0 \times RMSE_{train} \quad (2)$$

This chosen multiplier serves as an operational assumption for this proof-of-concept and requires calibration with actual operational data in future deployments.

##### B. PHASE 2: RELATIVE BENCHMARKING (SFA)

While Phase 1 effectively identifies deviations from historical expectations, it cannot mathematically distinguish whether a generation drop is caused by mechanical failure or unmeasured sub-optimal sea states. Phase 2 solves this by comparing the assets against each other using a Stochastic Frontier Analysis (SFA) [23].

The production frontier is defined using a log-linear Cobb-Douglas specification:

$$\ln(Y_{i,t}) = \beta_0 + \beta_1 \ln(WPF_{i,t}) + v_{i,t} - u_{i,t} \quad (3)$$

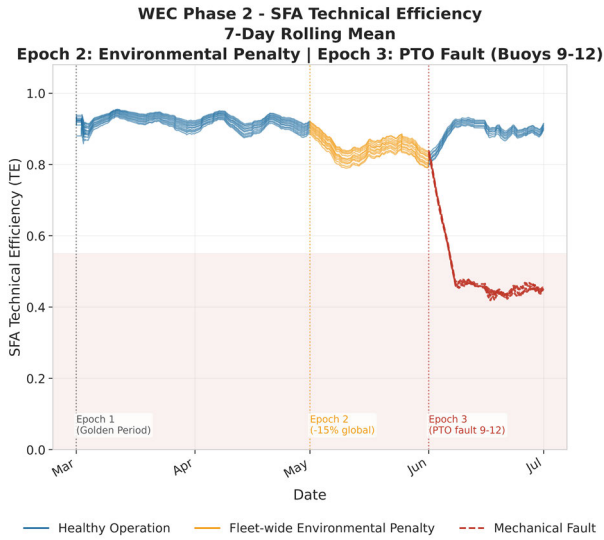
where  $\beta_0$  represents the intercept and  $\beta_1$  is the output elasticity with respect to the wave power flux. Methodologically, the log-linear Cobb-Douglas formulation is deliberately selected over flexible functional forms, such as the Translog frontier. While a Translog model captures second-order quadratic interactions, its unconstrained flexibility renders it highly susceptible to overfitting in harsh, highly volatile offshore environments, frequently violating strict monotonicity assumptions. The Cobb-Douglas model guarantees strict global monotonicity, ensuring that an increase in incident hydrodynamic energy consistently yields a non-decreasing power generation baseline, successfully preventing erratic frontiers amid wave noise.

The composite residual is decomposed into two distinct variables ( $\varepsilon = v - u$ ). Unlike deterministic models, this decomposition is the cornerstone of the diagnostic process. The symmetric noise component,  $v_{i,t} \sim N(0, \sigma^2_v)$ , captures natural sea-state variation and sensor noise. Because SFA measures relative efficiency against the fleet's empirical frontier, a fleet-wide environmental penalty (such as the spectral spreading in Epoch 2) shifts the global performance baseline uniformly. Since it does not increase individual relative inefficiency, this systemic variance is absorbed by  $v_{i,t}$ . Conversely, the technical inefficiency component,  $u_{i,t} \sim |N(0, \sigma^2_u)|$ , is a strictly non-negative term capturing unidirectional energy loss. While empirically serving as a proxy for mechanical PTO degradation in this framework, it also intrinsically absorbs unobserved asset heterogeneity and potential specification errors.

The model parameters are estimated via Maximum Likelihood Estimation (MLE). To ensure global convergence and avoid local minima, the L-BFGS-B optimization algorithm is employed with multiple randomized initial starting points.

Once the parameters are extracted from the Epoch 1 calibration, the relative Technical Efficiency (TE) of each WEC is computed utilizing the Battese and Coelli closed-form estimator [45]. This approach derives the efficiency from the conditional expectation of the inefficiency term  $u$  given the composite residual  $\varepsilon$ , assuming a truncated-normal distribution. The efficiency is formulated as:

$$\begin{aligned} TE_{i,t} &= E[\exp(-u_{i,t}) | \varepsilon_{i,t}] \\ TE_{i,t} &= \exp\left(-\mu_* + \sigma_*^2/2\right) \frac{\Phi(\mu_*/\sigma_* - \sigma_*)}{\Phi(\mu_*/\sigma_*)} \end{aligned} \quad (4)$$



**FIGURE 3.** Fleet-wide technical efficiency (TE) across three operational epochs.

where  $\Phi(\cdot)$  is the cumulative distribution function of the standard normal distribution, and the parameters  $\mu^*$  and  $\sigma^*$  represent the mean and standard deviation of the conditional distribution  $u_{i,t}|\varepsilon_{i,t}$ , respectively:

$$\mu_* = -\varepsilon_{i,t} \frac{\sigma_u^2}{\sigma^2} \text{ and } \sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma^2} \quad (5)$$

With  $\sigma^2 = \sigma^2 + \sigma^2$ . Under the composite error convention  $\varepsilon_{i,t} = v_{i,t} - u_{i,t}$ , a mechanical fault increases  $u > 0$ , which drives  $\varepsilon < 0$  and therefore  $\mu_* > 0$ , shifting the conditional distribution of  $u|\varepsilon$  toward larger inefficiency values and suppressing  $TE_{i,t}$  toward zero. This sign behaviour confirms that the estimator correctly maps PTO degradation into low efficiency scores.

To quantify the physical energy loss, the estimated frontier is back-transformed into the original linear space (kW). The expected instantaneous generation is defined as  $\hat{Y}_{kW} = \exp(\beta_0 + \beta_1 \ln WPF)$ . The absolute generation deficit attributable to mechanical waste is simply the difference between this theoretical optimum and the actual output.

A relative anomaly flag (Condition B) is triggered when an asset's instantaneous efficiency falls significantly below the empirical farm frontier, mathematically isolating internal mechanical waste:

$$TE_{i,t} < 0.60 \quad (6)$$

Rather than being a statistically tuned hyperparameter, this threshold is adopted strictly as an operational choice for this case study, not an economically validated figure. It represents an assumed boundary of acceptable relative efficiency loss (e.g.,  $\geq 40\%$  structural inefficiency) that would warrant an offshore dispatch, comfortably exceeding the natural variance of symmetric environmental noise ( $\sigma_v$ ). By requiring the efficiency to collapse below this strict boundary, the framework

ensures that minor sensor jitter or localized wave micro-noise is not misclassified as a mechanical fault.

(Note: While macroscopic degradation tracking employs a 7-day rolling mean for visualization, as illustrated in Fig. 3, the deterministic decision engine evaluates Condition B on the instantaneous high-frequency  $TE_{i,t}$  to preserve diagnostic granularity).

### C. PHASE 3: THE MERGE (DECISION ENGINE)

Phase 3 constitutes the final diagnostic stage, intersecting the absolute predictive capabilities of the data-driven baseline (Phase 1) with the relative Production Possibility Frontier (Phase 2). For each instantaneous observation, a deterministic logic gate evaluates Conditions A and B to classify the asset into one of four operational states.

Notably, the absolute error threshold multiplier (Condition A) is deliberately rendered more sensitive, shifting to  $-1.0 \times RMSE_{train}$  in this integrated phase, compared to the highly conservative  $-3.0$  factor utilized in the standalone Phase 1. This design choice is architecturally motivated. In isolation, the XGBoost baseline is entirely blind to complex multi-directional sea states and spectral spreading. Without a secondary filter, a standalone Phase 1 must enforce an extreme  $-3\sigma$  threshold to artificially suppress the massive wave of environmental false positives that naturally occur during confused seas. However, within the integrated decision engine, the relative SFA boundary (Condition B) assumes this suppression role, functioning as a statistically grounded safeguard against fleet-wide environmental noise. This orthogonal separation of responsibilities allows Condition A to operate with increased absolute sensitivity ( $-1\sigma$ ), capturing subtler mechanical generation deficits at a much earlier stage without inflating the false-positive rate.

- **State 0 (Nominal):** Both conditions FALSE. The WEC operates within its expected historical and relative margins.
- **State 1 (Environmental False Positive):** Condition A TRUE, Condition B FALSE. The asset underperforms its absolute historical baseline due to sub-optimal metocean factors, but maintains its relative position on the SFA farm frontier. The algorithm safely overrides the absolute alarm, classifying it as shared environmental noise.
- **State 2 (Relative Efficiency Anomaly):** Condition A FALSE, Condition B TRUE. The asset begins to deviate from its peers on the SFA frontier, indicating a relative performance drop, but the absolute deficit has not yet breached the relaxed  $-1.0 \times RMSE_{train}$  threshold.
- **State 3 (Critical Fault):** Both conditions TRUE. The asset falls below both its historical baseline and the instantaneous farm frontier, confirming an isolated mechanical PTO anomaly and triggering an immediate maintenance alarm.

By accumulating the total temporal duration of State 3 classifications across the operational epoch, the decision engine provides a filtered prioritization metric for O&M vessel

dispatch. Because the synthetic dataset provides a known ground truth of asset health, the Phase 3 diagnostic performance can be quantitatively assessed. An asset is flagged for physical dispatch if its accumulated State 3 duration strictly exceeds a predefined persistence threshold ( $\tau$ ). The system's robustness is evaluated using standard classification metrics (Precision, Recall, and F1-Score) adapted for operational time. True Positives (TP) represent hours of correct maintenance alarms, False Positives (FP) represent healthy hours erroneously flagged, and False Negatives (FN) quantify missed degradation. This temporal framing directly aligns statistical performance with real-world O&M costs.

## V. RESULTS AND DISCUSSION

The proposed three-phase architecture was executed and evaluated against the semi-empirical offshore conditions. To explicitly quantify the advantages of the proposed framework against traditional methods, the architecture is evaluated incrementally: Phase 1 acts as Baseline 1 (a purely absolute classification model) and Phase 2 acts as Baseline 2 (a purely relative classification model), culminating in Phase 3 as an integrated hybrid solution that establishes a reliable diagnostic baseline while rejecting environmental false positives.

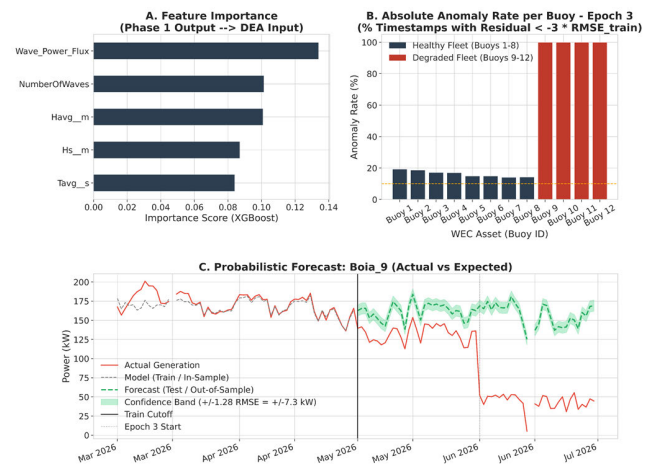
### A. PHASE 1: ABSOLUTE PERFORMANCE BASELINE

The XGBoost regressor was trained exclusively on the strictly isolated Golden Window partition (18,211 samples), which corresponds to the Epoch 1 operational window where all assets operate in a confirmed fault-free state. The model achieved strong in-sample convergence, capturing the complex, non-linear dynamics of wave energy conversion. As detailed in Table 1, the in-sample (Epoch 1) Root Mean Squared Error (RMSE) was restricted to 5.733 kW, achieving an  $R^2$  of 0.9038.

**TABLE 1. Phase 1 predictive metrics: Baseline calibration vs. out-of-sample inference.**

| Fleet Subset                                 | RMSE (kW) | MAE (kW) | $R^2$ Score |
|----------------------------------------------|-----------|----------|-------------|
| Epoch 1 (In-Sample/Train)                    | 5.733     | 4.494    | 0.9038      |
| Epoch 3 (Out-of-Sample: Healthy Buoys 1–8)   | 18.299    | 13.323   | 0.3130      |
| Epoch 3 (Out-of-Sample: Degraded Buoys 9–12) | 116.412   | 113.812  | -26.5745    |

Feature importance analysis (illustrated in Fig. 4A) revealed that the engineered Wave Power Flux (WPF) remained the primary predictive driver, accounting for 13.39% of the model's decision criteria. However, unlike idealized theoretical settings where a single WPF vector dominates entirely, the complex hydrodynamics of real-world ocean telemetry forced the algorithm to utilize a much broader distribution of spectral descriptors. Consequently, variables such as the Number of Waves (10.14%), Mean Wave Height ( $H_{avg}$  at 10.08%), and Significant Wave Height



**FIGURE 4. Phase 1 absolute analysis and XGBoost predictive outcomes.**

( $H_s$  at 8.70%) were heavily leveraged. This confirms that relying solely on simplified deep-water approximations is insufficient for accurate data-driven WEC monitoring in highly volatile seas.

During the out-of-sample inference phase, the absolute anomaly boundary was dynamically enforced at  $-3.0 \times RMSE_{train}$ . To explicitly evaluate the model's fault isolation capacity, Table 1 details the performance metrics extracted during the Epoch 3 testing window. The healthy fleet (Buoys 1–8) maintained a positive  $R^2$  of 0.3130, proving the model's capacity to establish a functional baseline without collapsing under out-of-sample sea states. Conversely, the algorithm correctly captured the structural failure of Buoys 9–12. This is corroborated by the absolute anomaly rate density isolating the degraded assets (Fig. 4B), and the probabilistic forecast showcasing the expected baseline boundary versus the actual generation collapse (Fig. 4C), where the Epoch 3  $R^2$  degraded massively to  $-26.5745$  (with the RMSE surging to 116.41 kW). This strongly negative coefficient confirms that the functional relationship between incident hydrodynamics and electrical output was severely disrupted by the PTO fault.

### B. PHASE 2: STOCHASTIC FRONTIER ANALYSIS (SFA)

While Phase 1 captured the absolute errors, Phase 2 successfully decomposed these residuals to isolate mechanical wear. The Maximum Likelihood Estimation (MLE) converged successfully during the Epoch 1 calibration (NLL =  $-26390.25$ ), yielding an output elasticity ( $\beta_1$ ) of  $+0.0440$ .

The structural variance parameters demonstrated a clear separation of noise, with the noise variance ( $\sigma_v^2$ ) minimized to 0.0076, while the inefficiency variance reached  $\sigma_u^2 = 0.0117$ . The derived signal-to-noise ratio  $\lambda = \sigma_u / \sigma_v \approx 1.24$  and variance ratio  $\gamma = \sigma_u^2 / \sigma_v^2 \approx 0.605$  confirm that 60.5% of the composite residual variance stems from inherent mechanical and thermodynamic system losses, establishing a baseline distinct from symmetric environmental noise. While idealized synthetic environments might force this

ratio closer to 100%, a 60.5% ratio directly reflects the inherent sensor jitter and metocean volatility of real-world telemetry. Crucially, this statistical separation indicates the model's capacity to decouple a robust inefficiency signal ( $u$ ) from benign sea-state fluctuations, aiding the isolation of subsequent mechanical degradation (Fig. 5). Empirical data introduces realistic complexities into these distributions. The healthy fleet's composite residual ( $\varepsilon$ ) distribution (blue) is slightly offset into the negative domain, reflecting the physical reality that structurally healthy WECs in harsh environments exhibit minor baseline inefficiencies relative to the absolute mathematical frontier. Conversely, the degraded fleet's distribution (red) shifts deeper into the negative domain with a visible leftward skew, confirming the severe energy penalty.

Furthermore, the decomposed components (Fig. 5, right panel) highlight the architecture's empirical robustness. While the environmental noise  $v$  (represented by Buoy 1) exhibits slight asymmetry, an expected empirical deviation from the strict theoretical Gaussian assumption when processing real-world sea states, this does not disrupt the decomposition. In stark contrast, the extracted technical inefficiency  $u$  (represented by Buoy 9) isolates in the positive domain with a one-sided, non-normal distribution, mathematically confirming the precise extraction of localized PTO degradation.

To comprehensively visualize the SFA mechanics, a triple frontier scatter analysis compares the fleet's response to environmental versus mechanical faults (Fig. 6). Panel A illustrates the full population distribution: during Epoch 2 (left), the entire fleet shifts uniformly downward due to spectral spreading, whereas in Epoch 3 (right), the degraded cluster (red) detaches massively from the healthy baseline (blue), which remains tightly bounded by the  $\pm 1\sigma_v$  stochastic noise band. Panel B isolates a single instantaneous common timestamp across all 12 assets to eliminate temporal variance. Under identical incident spectra, Epoch 2 exhibits a uniform generation deficit across the array. Conversely, the Epoch 3 snapshot reveals severe generation deficits directly tied to the mechanical state of Buoys 9 through 12. Finally, Panel C collapses each epoch into mean operating centroids per buoy, visually separating a systemic environmental penalty from a localized structural collapse.

In Epoch 3, the fault isolation capabilities of the Cobb-Douglas frontier were clearly demonstrated. Table 2 summarizes the asset criticality ranking derived from the back-transformed frontier distances. The SFA mathematically detached the degraded cluster, attributing massive mean generation deficits exceeding 123 kW per observation. Consequently, the Technical Efficiency (TE) for these assets collapsed to an average of  $\approx 45\%$ , falling far below the healthy fleet's nominal mean of  $\approx 90\%$ .

### C. PHASE 3: DETERMINISTIC DECISION ENGINE

The integration of absolute residuals and relative SFA efficiency scores in the Phase 3 decision matrix provided the

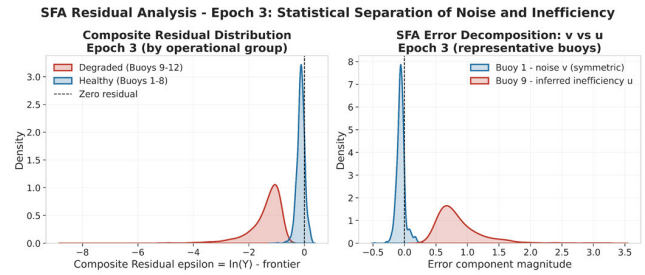


FIGURE 5. SFA residual decomposition in Epoch 3.

TABLE 2. Phase 2 SFA asset criticality ranking (epoch 3).

| WEC Asset | Mean TE | Mean Deficit (kW) | Status         |
|-----------|---------|-------------------|----------------|
| Buoy 9    | 0.4472  | 124.57            | Critical Fault |
| Buoy 10   | 0.4528  | 123.62            | Critical Fault |
| Buoy 11   | 0.4553  | 123.32            | Critical Fault |
| Buoy 12   | 0.4556  | 123.26            | Critical Fault |
| Buoy 8    | 0.9139  | 13.67             | Nominal        |
| Buoy 7    | 0.9119  | 15.14             | Nominal        |

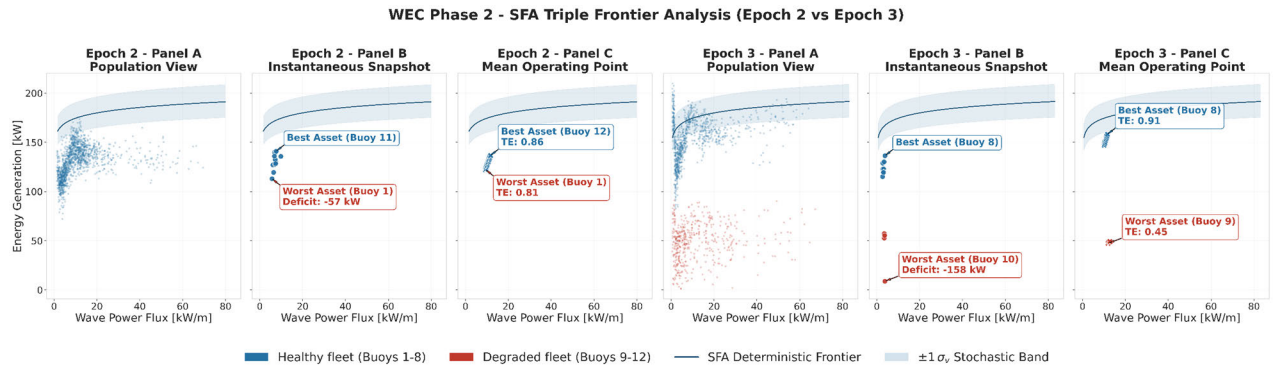
definitive O&M dispatch recommendation. To explicitly quantify the advantages of this proposed architecture, the integrated decision engine must be evaluated against its standalone components. If utilized in isolation, the absolute baseline would have successfully captured the vast majority of the mechanical faults. However, its fundamental blindness to fleet-wide spectral spreading would have triggered a total of 32,287 anomaly flags across the timeline (Condition A = TRUE). A detailed breakdown reveals that 27,259 of these were pure environmental False Positives (5,531 in Epoch 1; 17,225 in Epoch 2; and 4,503 on the healthy Epoch 3 cluster).

Conversely, relying exclusively on Phase 2 SFA introduces a structural vulnerability: lacking a historical generation anchor, a purely relative benchmark might erroneously assign nominal efficiency scores if systemic degradation suppresses the entire fleet's output simultaneously.

Integrating both models resolves these vulnerabilities. The hybrid intersection suppressed 99.7% of Phase 1's meteorological noise, collapsing environmental False Positives from 27,259 to 73 timestamps (36.5 hours). While this dual-condition is conservative, generating 304.5 hours of False Negatives by misclassifying premature degradation during calm seas, this trade-off is operationally vital to prevent costly false vessel deployments. Despite this filtering, the degraded cluster still accumulated massive, easily identifiable fault durations (e.g., Buoy 9 recorded 582.0 critical hours), as visually confirmed in Fig. 7.

To transparently evaluate the continuous diagnostic performance of the architecture, classification metrics were computed at the resolution of operational hours rather than simple asset counts, as detailed in Table 3.

The system correctly flagged 2,285.5 hours of true mechanical degradation (True Positives). The 304.5 hours



**FIGURE 6.** SFA triple frontier scatter comparing a fleet-wide environmental penalty (Epoch 2) with localized PTO faults (Epoch 3).

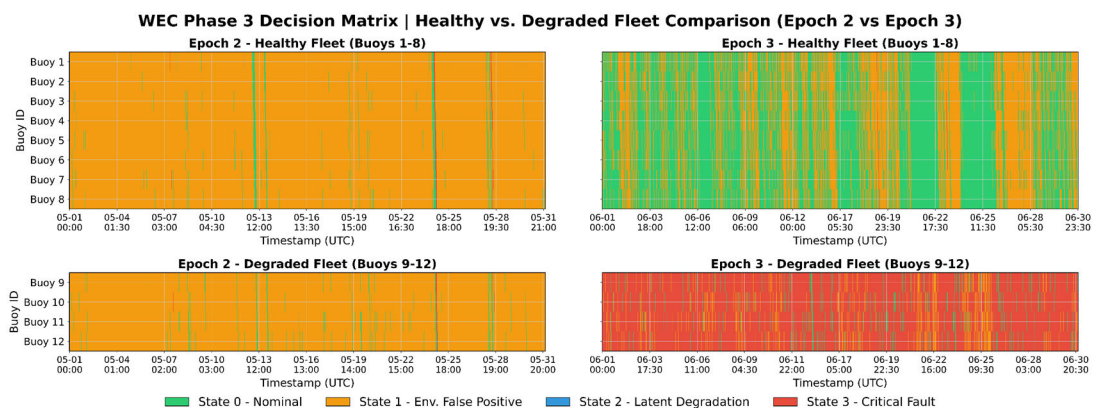
**TABLE 3.** Phase 3 diagnostic classification metrics (operational hours).

| Diagnostic Metric                                     | Value        |
|-------------------------------------------------------|--------------|
| True Positives (Degraded assets correctly dispatched) | 2,285.5h     |
| False Positives (Healthy assets falsely dispatched)   | 36.5h        |
| True Negatives (Healthy assets correctly ignored)     | 30,319.5h    |
| False Negatives (Degraded assets missed)              | 304.5h       |
| <b>Precision</b>                                      | <b>98.4%</b> |
| <b>Recall</b>                                         | <b>88.2%</b> |
| <b>F1-Score</b>                                       | <b>0.93</b>  |

of False Negatives represent a highly realistic operational phenomenon: during marginal degradation, incident wave energy is occasionally too low for the mechanical deficit to statistically overcome baseline environmental noise, causing intermittent algorithmic hesitation. The high diagnostic scores reflect the macroscopic nature of the injected PTO faults. While early-stage micro-anomalies might yield lower separability, the framework's primary objective, rejecting fleet-wide environmental false alarms during severe states, was robustly evaluated. Conversely, False Positives were reduced to a mere 36.5 hours out of over 30,356 healthy

operational hours, yielding a highly robust hour-level Precision of 98.4% and an F1-Score of 0.93.

To systematically filter the residual 1.6% false-positive rate, the decision engine requires an a priori temporal persistence threshold of  $\tau = 48$  accumulated (non-consecutive) State 3 hours before authorizing dispatch. Much like the relative efficiency limit, this was established strictly as an operational choice for this case study, not an economically validated figure. Representing two full days of critical fault telemetry, this cumulative approach ensures that actual degradation is not ignored due to intermittent fault visibility in calm seas. It must be explicitly noted that this 48-hour limit, the TE boundary and selected RMSE multipliers are operational assumptions for this proof-of-concept. They require rigorous calibration against actual operational and economic data in future real-world deployments. Applying this threshold yielded 100% asset-level dispatch accuracy: it successfully isolated all 4 degraded assets without triggering a single false deployment across the 8 healthy buoys. As the heatmap illustrates (Fig. 7, right panel), the persistent State 3 classification correctly isolates the structurally degraded assets (Buoys 9–12), supporting the architecture's potential viability for automated, high-precision offshore O&M dispatch.



**FIGURE 7.** Phase 3 decision matrix heatmaps.

## VI. CONCLUSION

Minimizing O&M expenditures is a critical prerequisite for the commercialization of Wave Energy Converters. This study introduced a simulation-based proof of concept proposing a three-phase data-driven methodology to address the challenge of environmental false alarms in offshore condition monitoring. By combining the predictive accuracy of an XGBoost regressor baseline with the relative econometrics of Stochastic Frontier Analysis (SFA), the proposed architecture aims to decouple mechanical PTO degradation from natural metocean variability.

The experimental results on a semi-empirical dataset, anchored in real offshore telemetry, demonstrated the system's robustness. The XGBoost model successfully mapped the baseline hydrodynamics ( $R^2 = 0.9038$ ), while the SFA effectively decomposed the prediction residuals. Crucially, the integrated decision engine successfully suppressed 27,259 environmental false positives caused by highly volatile sea states. The final deterministic framework successfully pinpointed the degraded cluster, identifying severe generation deficits ( $\approx 124$  kW) and triggering maintenance dispatches with high temporal precision: accumulating over 560 hours of correct critical fault flags for each degraded asset, compared to a mere 36.5 hours of transient statistical noise across the entire healthy fleet.

Future work will focus on validating this architecture against a fully empirical dataset containing documented real-world mechanical PTO failures. While the lightweight computational overhead of XGBoost and SFA inference highlights the framework's potential for continuous 30-minute evaluation on cloud infrastructures, the current methodology serves strictly as a semi-empirical proof of concept. Bridging the final gap to industrial deployment remains contingent on extensive validation against verified, physical fault data. Furthermore, if future arrays operate at power saturation, substituting Cobb-Douglas with a Translog frontier natively models horizontal curtailment. Additionally, integrating multi-modal telemetry (e.g., internal PTO pressures and hydraulic temperatures) into the Phase 1 feature space could further refine the latency of fault detection, accelerating the path towards fully autonomous offshore renewable energy management. To promote reproducibility within the marine energy sector, the complete codebase, data-processing scripts, and generated datasets used in this study are publicly available at: <https://github.com/ENERTEF/Performance-Evaluation-of-Energy-Forecasts-Service>

## ACKNOWLEDGMENT

The authors gratefully acknowledge project partner CorPower Ocean for providing the real offshore telemetry utilized in this study. The content of the article is the sole responsibility of them and does not necessarily reflect the views of European Commission. Additionally, they acknowledge the use of Google Gemini for text refinement and grammar correction throughout the manuscript and for code debugging support during the data analysis phase.

## REFERENCES

- [1] S. A. Mortazavizadeh, R. Yazdanpanah, D. C. Gaona, and O. Anaya-Lara, "Fault diagnosis and condition monitoring in wave energy converters: A review," *Energies*, vol. 16, no. 19, p. 6777, Sep. 2023, doi: [10.3390/en16196777](https://doi.org/10.3390/en16196777).
- [2] S. Pennock, M. Dickson, K. Rebenius, and J. H. Todalshaug, "Results and lessons learned from CorPower Ocean's full scale C4 wave energy converter first ocean deployment period," in *Proc. Eur. Wave Tidal Energy Conf.*, vol. 16, Sep. 2025, Art. no. 958, doi: [10.36688/ewtec-2025-958](https://doi.org/10.36688/ewtec-2025-958).
- [3] J. Falnes, *Ocean Waves and Oscillating Systems: Linear Interactions Including Wave-Energy Extraction*. Cambridge, U.K.: Cambridge Univ. Press, 2002.
- [4] A. Babarit, "A database of capture width ratio of wave energy converters," *Renew. Energy*, vol. 80, pp. 610–628, Aug. 2015, doi: [10.1016/j.renene.2015.02.049](https://doi.org/10.1016/j.renene.2015.02.049).
- [5] M. Masoumi, B. Estejab, and F. Henry, "Implementation of machine learning techniques for the analysis of wave energy conversion systems: A comprehensive review," *J. Ocean Eng. Mar. Energy*, vol. 10, no. 3, pp. 641–670, Aug. 2024, doi: [10.1007/s40722-024-00330-4](https://doi.org/10.1007/s40722-024-00330-4).
- [6] J. V. Ringwood, G. Bacelli, and F. Fusco, "Energy-maximizing control of wave-energy converters: The development of control system technology to optimize their operation," *IEEE Control Syst. Mag.*, vol. 34, no. 5, pp. 30–55, Oct. 2014, doi: [10.1109/MCS.2014.2333253](https://doi.org/10.1109/MCS.2014.2333253).
- [7] M. Alday, V. Raghavan, and G. Lavidas, "Effects of wave spectrum representation on power production estimations from point absorbers," *Appl. Ocean Res.*, vol. 161, Aug. 2025, Art. no. 104626, doi: [10.1016/j.apor.2025.104626](https://doi.org/10.1016/j.apor.2025.104626).
- [8] M. Göteman, C. McNatt, M. Giassi, J. Engström, and J. Isberg, "Arrays of point-absorbing wave energy converters in short-crested irregular waves," *Energies*, vol. 11, no. 4, p. 964, Apr. 2018, doi: [10.3390/en11040964](https://doi.org/10.3390/en11040964).
- [9] H. Wu and Z. Yuan, "Numerical investigation of the effects of short-crested irregular waves on the performance of a floating power capture platform," *Ocean Eng.*, vol. 313, Dec. 2024, Art. no. 119548, doi: [10.1016/j.oceaneng.2024.119548](https://doi.org/10.1016/j.oceaneng.2024.119548).
- [10] S. Zou, B. Robertson, A. Roach, T. Mundon, B. Rosenberg, and M. Penalba, "Wave energy converter arrays: A methodology to assess performance considering the disturbed wave field," *Renew. Energy*, vol. 229, Aug. 2024, Art. no. 120719, doi: [10.1016/j.renene.2024.120719](https://doi.org/10.1016/j.renene.2024.120719).
- [11] Z. Y. Tay and V. Venugopal, "The impact of energy extraction of wave energy converter arrays on wave climate under multi-directional seas," *J. Ocean Eng. Mar. Energy*, vol. 5, no. 1, pp. 51–72, Feb. 2019, doi: [10.1007/s40722-019-00127-w](https://doi.org/10.1007/s40722-019-00127-w).
- [12] Y. Wang, "Predicting absorbed power of a wave energy converter in a nonlinear mixed sea," *Renew. Energy*, vol. 153, pp. 362–374, Jun. 2020, doi: [10.1016/j.renene.2020.02.031](https://doi.org/10.1016/j.renene.2020.02.031).
- [13] Y. Wang, "A second order random wave model for predicting the power performances of a wave energy converter," *Acta Oceanologica Sinica*, vol. 40, no. 4, pp. 127–135, Apr. 2021, doi: [10.1007/s13131-021-1845-8](https://doi.org/10.1007/s13131-021-1845-8).
- [14] J. Morim, N. Cartwright, M. Hemer, A. Etamad-Shahidi, and D. Strauss, "Inter- and intra-annual variability of potential power production from wave energy converters," *Energy*, vol. 169, pp. 1224–1241, Feb. 2019, doi: [10.1016/j.energy.2018.12.080](https://doi.org/10.1016/j.energy.2018.12.080).
- [15] X. Guo, P. Zeng, X. Xiong, G. Wang, and Y. Cui, "Short-term wind power forecasting methods based on machine learning: A review and case study," *Energy Rep.*, vol. 14, pp. 3753–3782, Dec. 2025, doi: [10.1016/j.egyr.2025.10.040](https://doi.org/10.1016/j.egyr.2025.10.040).
- [16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [17] T. Muthamizhan, K. Karthick, S. K. Aruna, and P. Velmurugan, "AI-driven stacking ensemble for predicting total power output of wave energy converters: A data-driven approach to renewable energy processes," *Processes*, vol. 13, no. 4, p. 961, Mar. 2025, doi: [10.3390/pr13040961](https://doi.org/10.3390/pr13040961).
- [18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- [19] P. Bento, J. Pombo, M. D. R. Calado, and S. Mariano, "Ocean wave power forecasting using convolutional neural networks," *IET Renew. Power Gener.*, vol. 15, no. 14, pp. 3341–3353, Oct. 2021, doi: [10.1049/rpg2.12258](https://doi.org/10.1049/rpg2.12258).
- [20] C. Eadi Stringari, P. Veras Guimarães, J.-F. Filipot, F. Leckler, and R. Duarte, "Deep neural networks for active wave breaking classification," *Sci. Rep.*, vol. 11, no. 1, Feb. 2021, Art. no. 3604, doi: [10.1038/s41598-021-83188-y](https://doi.org/10.1038/s41598-021-83188-y).

- [21] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bénéttot, S. Tabik, A. Barbado, S. García, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: [10.1016/j.inffus.2019.12.012](#).
- [22] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007, doi: [10.1016/j.ymssp.2006.12.007](#).
- [23] D. Aigner, C. A. K. Lovell, and P. Schmidt, "Formulation and estimation of stochastic frontier production function models," *J. Econometrics*, vol. 6, no. 1, pp. 21–37, Jul. 1977, doi: [10.1016/0304-4076\(77\)90052-5](#).
- [24] P. Zhou, B. W. Ang, and D. Q. Zhou, "Measuring economy-wide energy efficiency performance: A parametric frontier approach," *Appl. Energy*, vol. 90, no. 1, pp. 196–200, Feb. 2012, doi: [10.1016/j.apenergy.2011.02.025](#).
- [25] W.-L. Hsiao, J.-L. Hu, C. Hsiao, and M.-C. Chang, "Energy efficiency of the Baltic sea countries: An application of stochastic frontier analysis," *Energies*, vol. 12, no. 1, p. 104, Dec. 2018, doi: [10.3390/en12010104](#).
- [26] U. Rzazade, S. Deryabin, I. Temkin, and A. Agabubaev, "Stochastic frontier-based analysis of energy efficiency in Russian open-pit mining enterprises," *Energies*, vol. 18, no. 13, p. 3257, Jun. 2025, doi: [10.3390/en18133257](#).
- [27] G. Iglesias, P. Castellanos, and A. Seijas, "Measurement of productive efficiency with frontier methods: A case study for wind farms," *Energy Econ.*, vol. 32, no. 5, pp. 1199–1208, Sep. 2010, doi: [10.1016/j.eneco.2010.03.004](#).
- [28] J. Zhang, Y. Wang, and L. Gao, "Empirical research on technical efficiency of wind power industry in China based on SFA method," *Environ., Develop. Sustainability*, vol. 26, no. 4, pp. 1–22, Mar. 2023, doi: [10.1007/s10668-023-03072-9](#).
- [29] G. Benini and G. Cattani, "Measuring the long run technical efficiency of offshore wind farms," *Appl. Energy*, vol. 308, Feb. 2022, Art. no. 118218, doi: [10.1016/j.apenergy.2021.118218](#).
- [30] Q. Van Nguyen, S. Pascoe, L. Cogan, and S. Nghiem, "The sensitivity of efficiency scores to input and other choices in stochastic frontier analysis: An empirical investigation," *J. Productiv. Anal.*, vol. 55, no. 1, pp. 1–10, Feb. 2021, doi: [10.1007/s11123-020-00592-8](#).
- [31] T. Kuosmanen and M. Kortelainen, "Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints," *J. Productiv. Anal.*, vol. 38, no. 1, pp. 11–28, Aug. 2012, doi: [10.1007/s11123-010-0201-3](#).
- [32] F. Liu, L. Li, B. Ye, and Q. Qin, "A novel stochastic semi-parametric frontier-based three-stage DEA window model to evaluate China's industrial green economic efficiency," *Energy Econ.*, vol. 119, Mar. 2023, Art. no. 106566, doi: [10.1016/j.eneco.2023.106566](#).
- [33] M. Molinos-Senante, A. Maziotis, and R. Sala-Garrido, "Benchmarking the economic and environmental performance of water utilities: A comparison of frontier techniques," *Benchmarking: Int. J.*, vol. 29, no. 10, pp. 3176–3193, Nov. 2022, doi: [10.1108/bij-08-2021-0481](#).
- [34] H. Fox, A. C. Pillai, D. Friedrich, M. Collu, T. Dawood, and L. Johanning, "A review of predictive and prescriptive offshore wind farm operation and maintenance," *Energies*, vol. 15, no. 2, p. 504, Jan. 2022, doi: [10.3390/en15020504](#).
- [35] O. Apata, J. L. Munda, and E. M. Migabo, "Artificial intelligence for predictive maintenance and performance optimization in renewable energy systems: A comprehensive review," *Energies*, vol. 19, no. 2, p. 536, Jan. 2026, doi: [10.3390/en19020536](#).
- [36] A. May and D. McMillan, "Condition based maintenance for offshore wind turbines: The effects of false alarms from condition monitoring systems," *IET Renew. Power Generat.*, pp. 783–789, 2013. Accessed: Jun. 7, 2026. [Online]. Available: [https://www.researchgate.net/publication/278673909\\_Condition\\_Based\\_Maintenance\\_for\\_Offshore\\_Wind\\_Turbines\\_The\\_Effects\\_of\\_False\\_Alarms\\_from\\_Condition\\_Monitoring\\_Systems](https://www.researchgate.net/publication/278673909_Condition_Based_Maintenance_for_Offshore_Wind_Turbines_The_Effects_of_False_Alarms_from_Condition_Monitoring_Systems)
- [37] A. M. Peco Chacón, I. Segovia Ramírez, and F. P. García Márquez, "State of the art of artificial intelligence applied for false alarms in wind turbines," *Arch. Comput. Methods Eng.*, vol. 29, no. 5, pp. 2659–2683, Aug. 2022, doi: [10.1007/s11831-021-09671-x](#).
- [38] G. Rinaldi, P. R. Thies, and L. Johanning, "Current status and future trends in the operation and maintenance of offshore wind turbines: A review," *Energies*, vol. 14, no. 9, p. 2484, Apr. 2021, doi: [10.3390/en14092484](#).
- [39] A. M. Peco Chacón, I. Segovia Ramírez, and F. P. García Márquez, "K-nearest neighbour and K-fold cross-validation used in wind turbines for false alarm detection," *Sustain. Futures*, vol. 6, Dec. 2023, Art. no. 100132, doi: [10.1016/j.sfr.2023.100132](#).
- [40] A. M. P. Chacón, I. S. Ramírez, and F. P. G. Márquez, "False alarm detection in wind turbine by classification models," *Adv. Eng. Softw.*, vol. 177, Mar. 2023, Art. no. 103409, doi: [10.1016/j.advengsoft.2023.103409](#).
- [41] V. F. Barnabei, T. C. M. Ancora, G. Delibra, A. Corsini, and F. Rispoli, "Semi-supervised deep learning framework for predictive maintenance in offshore wind turbines," *Int. J. Turbomachinery, Propuls. Power*, vol. 10, no. 3, p. 14, Jul. 2025, doi: [10.3390/ijtp10030014](#).
- [42] C. Velandia-Cardenas, Y. Vidal, and F. Pozo, "Wind turbine fault detection using highly imbalanced real SCADA data," *Energies*, vol. 14, no. 6, p. 1728, Mar. 2021, doi: [10.3390/en14061728](#).
- [43] A. Pliego Marugán and F. P. García Márquez, "Advanced analytics for detection and diagnosis of false alarms and faults: A real case study," *Wind Energy*, vol. 22, no. 11, pp. 1622–1635, Nov. 2019, doi: [10.1002/we.2393](#).
- [44] *Service 2-Performance Evaluation of Energy Forecasts for Offshore Energy Assets*, TEF RES Consortium, Brussels, Belgium, 2024.
- [45] G. E. Battese and T. J. Coelli, "Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data," *J. Econometrics*, vol. 38, no. 3, pp. 387–399, Jul. 1988, doi: [10.1016/0304-4076\(88\)90053-x](#).



**DANIEL COSTA** (Student Member, IEEE) is currently pursuing the master's degree in electrotechnical and computer engineering, specializing in energy with the Faculty of Engineering, University of Porto (FEUP), Portugal. He is also a Researcher at INESC TEC, where he develops anomaly detection frameworks for large-scale physical energy assets, including wind turbines, solar photovoltaics (PV), and wave energy converters. His recent work also encompasses electric vehicle (EV) resilience

strategies for transmission system operator (TSO) contracts. His research interests include machine learning, industrial AI, and predictive maintenance techniques to evaluate asset performance, energy systems modeling and optimization, data-driven decision support, explainable AI (XAI), and digitalization in the energy sector.



**ALEXANDRE LUCAS** received the degree in electrical engineering from the Instituto Superior de Engenharia de Lisboa (ISEL), Portugal, in 2007, the M.Sc. degree in business and industrial strategy from the Instituto Superior de Economia e Gestão (ISEG), in 2009, and the Ph.D. degree in sustainable energy systems from the Instituto Superior Técnico (IST), in 2013. In 2011 and 2012, he conducted research with Massachusetts Institute of Technology (MIT) and completed complementary

training in energy management at Harvard University. Following nine years of industry experience in power substations, he worked at the Joint Research Centre (JRC) of European Commission, from 2014 to 2020. He is currently a Researcher at INESC TEC, where he leads the Demand Side Energy Management Area. He has been actively involved in prominent European research and innovation projects, including Electra, Delta, Drimpac, Interconnect, OneNet, COSMIC, InterStore, Every1, and EnerTEF. His research interests include energy digitalization, energy management, demand response, and the integration of distributed energy resources.

...