Machine Learning-based OWC Diagnosis Using Real Measured Data from Wave Power Plants

1st Fares M'zoughi
Automatic Control Group - ACG,
Department of Automatic Control and Systems Engineering,
Faculty of Engineering of Bilbao - EIB/BIE,
University of the Basque Country - UPV/EHU,
Rafael Moreno 3, 48013 Bilbao, Spain
ORCID: 0000-0003-2935-3830

2nd Jon Lekube
Biscay Marine Energy Platform, BiMEP,
Atalaia 2 bajo, 48620 Armintza, Spain
ORCID: 0000-0003-2611-2054

3rd Izaskun Garrido
Automatic Control Group - ACG,
Department of Automatic Control and Systems Engineering,
Faculty of Engineering of Bilbao - EIB/BIE,
University of the Basque Country - UPV/EHU,
Rafael Moreno 3, 48013 Bilbao, Spain
ORCID: 0000-0002-9801-4130

4th Aitor J. Garrido
Automatic Control Group - ACG,
Department of Automatic Control and Systems Engineering,
Faculty of Engineering of Bilbao - EIB/BIE,
University of the Basque Country - UPV/EHU,
Rafael Moreno 3, 48013 Bilbao, Spain
ORCID: 0000-0002-3016-4976

***Abstract*—This manuscript introduces a new classification-based power take-off diagnosis for wave energy converter farms. The suggested strategy has been tested on the Mutriku Multiple Oscillating Water Column-based wave power plant in order to reduce the Levelised Cost of Energy (LCoE) by implementing predictive maintenance strategies. This has been achieved by employing Linear Discriminant Analysis (LDA) to determine the furthermost relevant features from the measured data. Then the Support Vector Machine (SVM) has been implemented as a classification technique to classify the state of the OWC system.**

Keywords—Classification, Fault Diagnosis, LDA, machine learning, OWC, power take-off, SVM, wave energy.

I. Introduction

Levelized Cost of Energy (LCoE) within a wave farm may be reduced by increasing the power harnessing capability of Wave Energy Converter (WEC) technology. Enhancing availability, capacity factor, and Annual Energy Production (AEP) are further ways to reduce LCoE. To maintain ideal operational conditions while implementing these enhancements, effective monitoring and maintenance procedures are required. Indeed, maintenance has a substantial influence on the amount of downtime throughout the course of a plant's life, which helps to raise availability, power output, capacity factor, and AEP. Therefore, decreasing Operational and Maintenance expenses efficiently achieves LCoE mitigation [1].

Maintenance strategies are commonly categorized as reactive, proactive, and opportunistic based on task timing. The reactive approach, also known as corrective maintenance, involves addressing failures after they've transpired. This method proves effective when downtime-associated maintenance is minimal, making it suitable for highly reliable small farms [2]. Conversely, proactive maintenance entails pre-scheduled inspections and replacements to prevent minor issues from escalating into major failures. Various strategies, including preventive, condition-based, and predictive maintenance, fall within the proactive category [3]. Lastly, the opportunistic strategy amalgamates scheduled preventive and corrective maintenance tasks with unscheduled preventive operations aimed at addressing future component wear [4].

Given that both onshore and offshore power plants benefit from a proactive maintenance approach, data from both time-based and sensor-based sources are collected to create the best possible maintenance plan. Data processing is made more difficult by the sizeable volume of data that has been gathered and the large number of variables that have been measured. As a result, feature extraction is used to lessen dimensionality and eliminate redundancy—a widespread method across several areas [5,6,7]. Numerous feature extraction algorithms are available, with Linear Discriminant Analysis (LDA) ranking among the most widely utilized. LDA involves identifying a projection hyperplane that minimizes interclass variance while maximizing the separation between projected class means [8]. This objective is accomplished by addressing the eigenvalue problem, which yields the relevant eigenvector defining the pivotal hyperplane [9].

To employ the extracted data for plant health monitoring and failure detection, recognizing patterns of failure within the data is imperative. As a result, various studies have delved into the development of recognition or classification models [10,11,12]. These encompass the nonparametric *k*th-Nearest Neighbor (kNN) method, which employs "feature similarity" to predict the values of new data points based on their proximity to training set points [13]. Logistic models are also considered, elucidating the data and elucidating the correlation between a dependent binary

This work was supported in part by project IT1555-22 funded by the Basque Government and projects PID2021-123543OB-C21 and PID2021-123543OB-C22 funded by MCIN/AEI/10.13039/501100011033/FEDER, UE and through the Maria Zambrano grant MAZAM22/15 funded by UPV-EHU/MIU/Next Generation, EU.

variable and other nominal independent variables [14]. The decision tree (C4.5) is another technique that utilizes a recursive splitting method to divide a set of instances into disjoint subsets [15]. Additionally, Multivariate Discriminant Analysis (MDA) is a classification approach that constructs a discriminant function by maximizing the ratio of "between groups" variance to "within groups" variance [16].

The Support Vector Machine (SVM) is a well-established machine learning technique employed to address classification challenges within vast datasets [17]. Its applications are especially valuable in multi-domain scenarios within the context of big data [17]. However, it's important to note that SVM, despite its effectiveness, entails mathematical complexity and substantial computational requirements [18]. Yet, SVM shines in terms of its strong generalization capabilities, making it a reliable choice for achieving great classification precision, particularly in machine condition monitoring and fault diagnostics [19, 20].

The Basque Energy Agency (Ente Vasco de la Energa - EVE), which is located in north of Spain, formally inaugurated the Mutriku wave power plant in July 2011. This establishment, illustrated in Figure 1, is an onshore facility seamlessly built into the harbor breakwater of Mutriku. It holds 16 Oscillating Water Columns consisting of a Wells turbine coupled with a DFIG generator [21], [22], [23] and [24].

Fig. 1. Wave Power Plant of Mutriku in Spain.

II. MATERIALS AND METHODS

Following ten years of operation, the Mutriku Wave Power Plant (WPP) has documented various instances of degradation and failures. These specific occurrences are elaborated upon in the subsequent table.

All the potential damages that may affect any OWC system primarily correspond to three categories of malfunctions: bearing issues, resonance occurrences, and imbalance situations. As a result, these issues will directly result in elevated vibrations, as depicted in Figures 2, 3, and 4.

On the date 15/09/2021, the 24-hour vibration profiles of turbines T03, T06, and T07 in the Mutriku WPP are displayed in Figures 2, 3, and 4.

TABLE I. DAMAGES OCCURRED ON OWCS IN MUTRIKU WPP.

Component	**Cause**	**Damages**
Wells turbine	Exposure to saltwater and material fatigue from strong airflows	
Generator	Exposure to saltwater and/or strike by broken blades	
Bearing cover	Excessive axial force induced in the turbine shaft leads to bearings rubbing against the inside of the generator cover	
Cooling system	Salt accumulation	

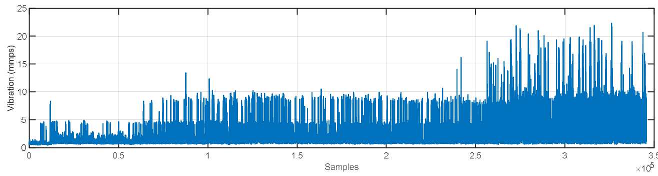

Fig. 2. Turbine T03's 24 hours measured vibrations with bearing problem.

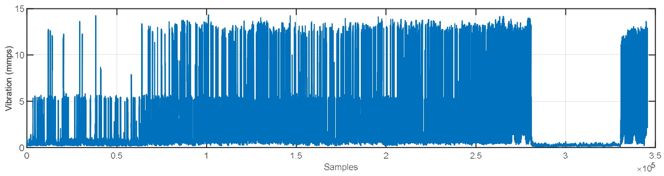

Fig. 3. Turbine T06's 24 hours measured vibrations with resonance problem.

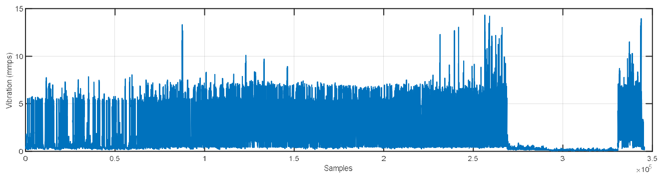

Fig. 4. Turbine T07's 24 hours measured vibrations with unbalance problem.

As depicted in the preceding figures, vibrations have the potential to surpass *20 mm/s*. If these undesirable vibrations are not addressed, they have the capacity to undermine the OWC's efficiency and potentially exacerbate its condition through degradation and potential component failure.

Enhancing the preventive maintenance of OWCs necessitates effective handling of the gathered data. By scrutinizing and analyzing the recorded vibrations across varying operational months, potential problems and faults can be identified and diagnosed. Fig. 5 presents the schematic representation of the proposed approach to identify the type of failures in OWC units, which, in turn, aids in planning upcoming maintenance activities to curtail OpEx expenses.

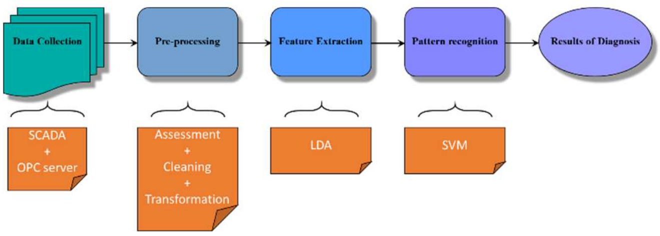

Fig. 5. ML-based power take-off diagnosis for Oscillating water columns.

Enhancing the proactive maintenance of the OWCs necessitates effective management of the collected data. The examination and assessment of recorded vibrations over various operational months are instrumental in recognizing and pinpointing potential problems and malfunctions [25], [26]. Fig. 5 outlines the schematic representation of the utilized approach to categorize the nature of failures within the OWC unit, which subsequently aids in planning forthcoming maintenance activities aimed at minimizing OpEx.

A. LDA-based Feature Extraction

A technique for preprocessing and reducing the computational complexity of a dataset is feature extraction. A classifier's training and classification stages might both experience significant computational and memory overheads as a result of increased feature dimensionality. A classification approach is used because finding patterns in high-dimensional data can be difficult.

The most often used classical linear technique to reduce dimensionality is Linear Discriminant Analysis (LDA). Within the feature-based projection space, LDA looks for a transformation matrix W that will optimally increase the ratio of the between-class disperse and decrease the within-class disperse matrix.

LDA search for a transformation matrix W, which will maximize the ratio of the between-class disperse and will minimize the within-class disperse matrix within the feature-based projection space [8]. LDA is an approach to obtain the linear sets of characteristics that best distinguishes between multiple classes of events or objects.

The within-class distribution matrix S_W is defined by [27,28]:

$$S_W = \sum_{i=1}^{c} \sum_{x \in Ci} (x - m_i)(x - m_i)^t \tag{1}$$

here c represents the classes' number while C_i represents the set of data in the i^{th} class, and m_i represents the mean of the i^{th} class. It's to be noted that the matrix of within-class distribution is a representation of the level of scattering inside classes as the sum of the covariance matrices of every class.

Another relevant parameter is the between-class scatter matrix, which maybe defined as [27,28]:

$$S_B = \sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^t \tag{2}$$

A criterion function is then defined using S_W matrix of the within-class scatter and S_B matrix of the between-class scatter to obtain the transformation matrix W described by [27,28]:

$$J(W) = \frac{|\widetilde{S}_B|}{|\widetilde{S}_W|} = \frac{|W^t S_B W|}{|W^t S_W W|} \tag{3}$$

The transformation matrix W is the one will maximize the criterion function $J(W)$. The generalized eigenvectors w_i in the columns of the optimum transformation matrix W correspond to the biggest eigenvalues in:

$$S_B W_i = \lambda_i S_W W_i \tag{4}$$

LDA seeks to identify a combination of features by effectively differentiating between various object classes. If S_W is full-rank, W may be calculated via the eigenvectors of $S_W^{-1} S_B$.

The LDA technique relies on linear adjustments to increase variation within a smaller dimension. LDA looks for linear discriminants to increase variation within different categories while at the same time lowering variance within each class.

B. SVM-based Power Take-Off Diagnosis

The goal is to develop and train a classifier utilizing the pre-processed data to distinguish between distinct OWC health statuses. The Support Vector Machine (SVM) approach is the classification methodology used in this investigation.

The classification-based OWC diagnosis technique described here aims to increase the effectiveness of maintenance scheduling by using predictive maintenance rather than relying solely on preventative maintenance.

Formally defined by a separating hyperplane, an SVM is a discriminative classifier. Classification and pattern recognition have seen much development and use of SVM [29,30]. SVM refers to a group of similar supervised learning techniques. A hyperplane classifier is essentially what SVM is. Finding a hyperplane that distinguishes the positive training samples from the negative training samples associated to the greatest margin serves as the decision surface for training an SVM classifier [31]. SVM's ability to handle nonlinearly separable data is one of the key factors contributing to its widespread use. For training samples pairs (x_i, y_i), with x_i represents the vector of the weighted feature while $y_i \in \{1, -1\}$ is the label.

For data that can be separated linearly, we may identify a hyperplane $f(x)=0$ that does the following:

$$f(x)=\sum_{i=1}^{n} w_i x_i + b = 0 \tag{5}$$

here w represents a n-dimensional vector and b as a constant. Both parameters w and b determine the position of the separating hyperplane. For every i either:

$$\begin{cases} w.x_i - b \geq 1 & \text{for } x_i \text{ of the first class.} \\ w.x_i - b \leq 1 & \text{for } x_i \text{ of the second class.} \end{cases} \tag{6}$$

The separating hyperplane is the hyperplane that produces the largest margin. By resolving the following issue with the consideration of the noise with slack variables ξ_i and error penalty C, the ideal hyperplane may be discovered:

$$\min_{w,b,\xi} P(w,b,\xi) = \frac{1}{2}\langle w.w \rangle + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \tag{7}$$

here ξ_i represents the gap from the margin to the sample x_i, which is positioned outside of it

$$\min_{w,b,\xi} P(w,b,\xi) = \frac{1}{2}\langle w.w \rangle + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \tag{8}$$

By transforming the Kuhn-Tucker conditional problem into an analogous Lagrange dual problem, the computations can be made easier:

$$V(\alpha)=\sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j K(x_i.x_j) \tag{9}$$

subject to

$$\sum_{i=1}^{l} y_i\alpha_i = 0, \qquad C \geq \alpha \geq 0, \qquad i=1,2,\ldots,l \tag{10}$$

A kernel function is a function that returns a dot product of feature space mappings of the original data points, $K(x_i.\ y_i)$. The quantity of training data equals the number of variables in the dual problem. The Karush-Kuhn-Tucker theorem states that the associated α must not be 0 in order for the equality criterion to apply to the training input-output pair $(x_i.\ y_i)$. Support vector (SV) is used as the training example x_i in this instance. SVM is extremely computationally efficient since the number of SVs is much smaller than the number of training samples. For the classification problem, SVM is a useful classifier.

III. Results and Discussion

To extract characteristics that indicate the information on plant health, the acquired data is processed. The traits that are most pertinent to our investigation are revealed by running an LDA on the OWC data. In contrast to the blue bars in Fig. 6, which reflect the proportion of variation described by each individual component (in%), the scree plot of Fig. 6 displays the cumulative variance explained by the additional Discriminant Component (DC).

According to Fig. 6, the first discriminant component accounts for 40.88% of the variance, whereas the second and third components each account for 26.43% and 24.78%. Therefore, it requires 3 components to explain 92.09% of the total variation. The top three DCs correspond to the vibration velocity, angular velocity and pressure features.

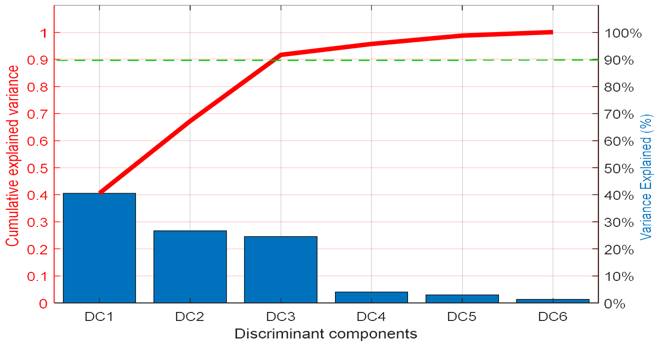

Fig. 6. Explained variance obtained from LDA

The scatter plots of the two most significant features are shown in Fig. 7 using data from a healthy Wells turbine and three distinct defective turbines, including imbalanced, bearing, and resonance issues, that were recorded at the Mutriku WWP facility on September 15, 2021.

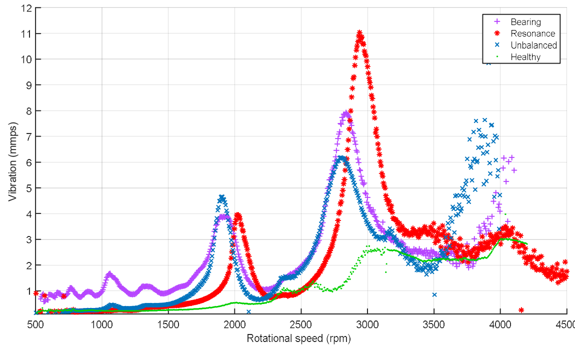

Fig. 7. Average vibrations speed vs. angular velocity in four turbines in Mutriku's OWCs.

As seen in Fig. 7, the average vibration changes depending on the rotor's angular frequency. In actuality, it is evident that each turbine registers two maxima at roughly *1900 rpm* and *2900 rpm*. An ideal turbine may vibrate up to *3 mm/s*, while ill turbines can vibrate up to *6 mm/s* at *2800 rpm* for imbalance issues, *8 mm/s* at *2838 rpm* for bearing issues, and *11 mm/s* at *2940 rpm* for resonance issues.

The scatter plot depicting the three most relevant features is presented in Fig. 8, derived out of the collected data on 15/09/2021 concerning the Wells turbine bearing issue. This 3D scatter plot distinctly reveals that the fluctuation in vibration is influenced not solely by the rotor's angular velocity but also by the OWC's pressure.

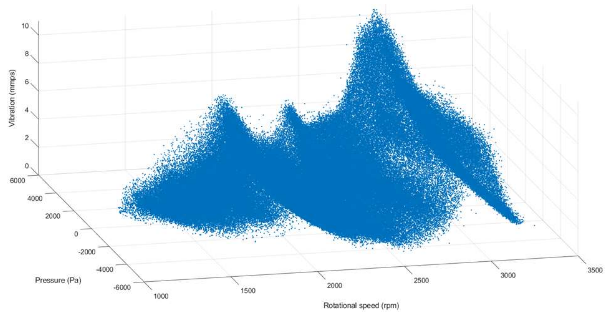

Fig. 8. Scatter plot of vibration vs angular velocity & pressure in turbine T03.

In Fig. 9, it's evident that there is minimal confusion between the classes. Precise predictions between classes stand at 89%, while inaccurate predictions are merely 11%.

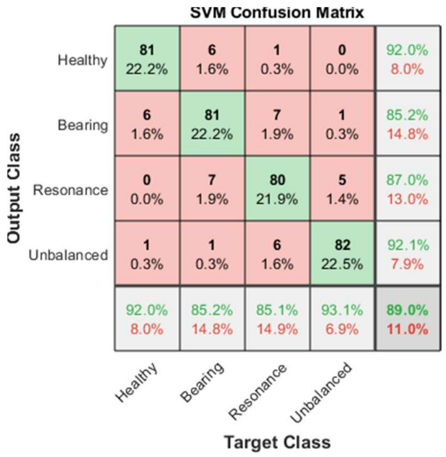

Fig. 9. Confusion plot of the trained SVM model.

The ROC depicted in Fig. 10 reveals that the SVM model exhibits an excellent performance, boasting an AUC of 0.81.

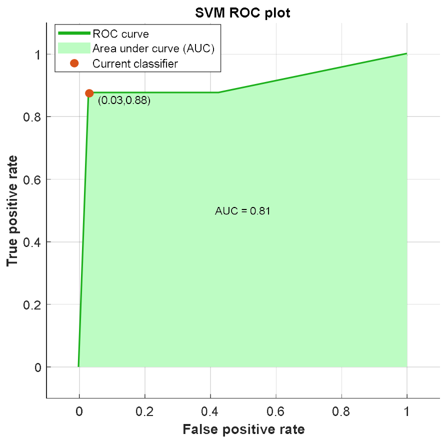

Fig. 10. Receiver Operating Characteristic of trained SVM model.

Fig. 11 illustrates simulations conducted to assess the precision of the trained SVM classifier concerning the count of features, utilizing the LDA components. A substantial enhancement is observed as additional features are incorporated, reaching a notable improvement up to three DCs. However, the accuracy gains become more marginal beyond four DCs.

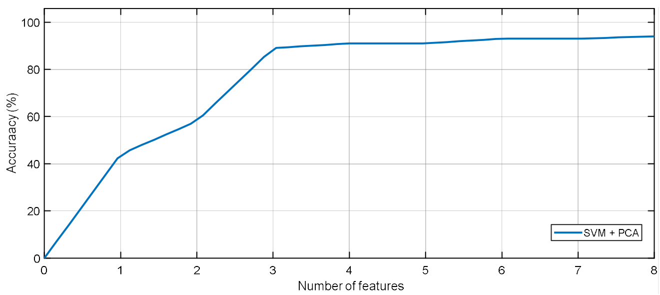

Fig. 11. Trained SVM's classification accuracy vs. features.

IV. Conclusions

This paper introduces a classification-oriented approach for diagnosing Power Take-Off systems in WECs to facilitate the implementation of predictive maintenance strategies. The proposed methodology utilizes actual plant data for training classifier models, enabling the prediction of the health status of individual WEC units. The applicability of this approach is assessed using the Mutriku wave power plant as a case study. Given the challenging environmental conditions at Mutriku, the PTO units of the Oscillating Water Column experience vibrations that can result in failures and breakages. These incidents contribute to significant downtime, consequently impacting the Levelized Cost of Energy (LCoE). In the context of this case study, the primary objective is to diagnose the underlying issues, mitigate vibration-related failures, and ultimately optimize the LCoE through the implementation of a robust predictive maintenance strategy.

The study utilized the feature extraction technique of Linear Discriminant Analysis (LDA) to identify the most pertinent features for OWC diagnosis. Through LDA, it was determined that three components are required to achieve a cumulative explained variance of 92.09%. Consequently, the initial three common components, encompassing vibration, rotational speed, and pressure features, were selected for the design and training of the classification model.

The classification technique employed in this study is the Support Vector Machine (SVM). The outcomes indicate that the SVM model, once trained, effectively attains a remarkable accuracy level of 89%.

Upcoming research endeavors will encompass the exploration and application of the classification-based PTO diagnosis for the integration of a predictive maintenance strategy. As indicated by existing research, this strategy holds the potential to yield an OpEx reduction of 18% and decrease plant downtime by 20%, thereby elevating plant availability to 81% and consequently reducing LCoE by as much as 23%.

Acknowledgment

The authors would like to thank the help and collaboration of the Basque Energy Agency (Ente Vasco de la Energía-EVE) and Biscay Marine Energy Platform-BiMEP for providing the real measured experimental data of the Mutriku wave power plant and the economic data.

References

[1] R. Zhengru, A. S. Verma, Y. Li, J. J. E. Teuwen, and Z. Jiang. "Offshore wind turbine operations and maintenance: A state-of-the-art review," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110886, 2021.

[2] A. Karyotakis, and R. Bucknall. "Planned intervention as a maintenance and repair strategy for offshore wind turbines," *Journal of marine engineering & technology*, vol. 9, no. 1, pp. 27-35, 2010.

[3] J. Peng. "Maintenance of wind turbine," *Electr. Equip.* vol. 28, mo. 6, pp.68-71, 2011.

[4] T. Édouard, É. Levrat, and B. Iung. "Overview on opportunistic maintenance," *IFAC Proceedings Volumes*, vol. 41, no. 3, pp. 245-250, 2008:

[5] J. Sklansky, "Image segmentation and feature extraction," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 4, pp. 237-247, 1978:

[6] P. Boonyakitanont, A. Lek-Uthai, K. Chomtho, and J. Songsiri. "A review of feature extraction and performance evaluation in epileptic seizure detection using EEG," *Biomedical Signal Processing and Control*, vol. 57, p. 101702, 2020.

[7] Z. Peng, F. Chu, and Y. He. "Vibration signal analysis and feature extraction based on reassigned wavelet scalogram," *Journal of Sound and Vibration*, vol. 253, no. 5, pp. 1087-1100, 2002.

[8] Izenman, Alan Julian. "Linear discriminant analysis." In *Modern multivariate statistical techniques*, pp. 237-280. Springer, New York, NY, 2013.

[9] J. Wen, X. Fang, J. Cui, L. Fei, K. Yan, Y. Chen, Y. Xu. "Robust sparse linear discriminant analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 2, pp. 390-403, 2018.

[10] M. Y. Kiang. "A comparative assessment of classification methods," *Decision support systems*, vol. 35, no. 4, pp. 441-454, 2003.

[11] J. Liu, J. Sun, S. Wang. "Pattern recognition: An overview," IJCSNS International Journal of Computer Science and Network Security, vol. 6, no. 6, pp. 57-61, 2006.

[12] G. K. Verma, U. S. Tiwary. "Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals," NeuroImage, vol. 102, pp. 162-172, 2014.

[13] M. A. Wong, T. Lane. "A *k*th nearest neighbour clustering procedure," Journal of the Royal Statistical Society: Series B (Methodological), vol. 45, no. 3, pp. 362-368, 1983.

[14] D. R. Cox. "Analysis of binary data," *Routledge*, 2018 Feb 19.

[15] J. R. Quinlan. "C4.5: programs for machine learning," *Elsevier*; 2014 Jun 28.

[16] R.A. Fisher. "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179-188, 1936.

[17] S. Suthaharan, S. Suthaharan. "Support vector machine," In Machine learning models and algorithms for big data classification: thinking with examples for effective learning. *Springer*, Boston, MA, pp. 207-235, 2016.

[18] D.A. Pisner, D. M. Schnyer. "Support vector machine," In Machine learning. *Academic Press*, pp. 101-121, 2020.

[19] A. Widodo, B.S. Yang. "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical systems and signal processing*, vol. 21, no. 6, pp. 2560-2574, 2007.

[20] S. Lee. "Monte Carlo simulation using support vector machine and kernel density for failure probability estimation," *Reliability Engineering & System Safety*, vol. 209, p. 107481, 2021.

[21] F. M'zoughi, S. Bouallègue, M. Ayadi, A.J. Garrido, I. Garrido. "Modelling and airflow control of an oscillating water column for wave power generation," In Proc. *4th International Conference on Control, Decision and Information Technologies (CoDIT)*, Barcelona, Spain, April 2017, pp. 938-943.

[22] F. M'zoughi, I. Garrido, A.J. Garrido, M. De La Sen. "Fuzzy gain scheduled-sliding mode rotational speed control of an oscillating water column," *IEEE Access*, vol. 8, pp. 45853-45873, 2020.

[23] F. M'zoughi, I. Garrido, A.J. Garrido, M. De La Sen. "Rotational speed control using ANN-based MPPT for OWC based on surface elevation measurements," *Applied Sciences*, vol. 10, no. 24, p. 8975, 2020.

[24] J. Lekube, A.J. Garrido, I. Garrido, E. Otaola, J. Maseda. "Flow control in wells turbines for harnessing maximum wave power," *Sensors* , vol. 18, no. 2, p. 535, 2018.

[25] P. Aboutalebi, F. M'zoughi, I. Martija, I. Garrido, A.J. Garrido. "Switching control strategy for oscillating water columns based on response amplitude operators for floating offshore wind turbines stabilization," *Applied Sciences*, vol. 11, no. 11, p. 5249, 2021.

[26] I. Ahmad, F. M'zoughi, P. Aboutalebi, I. Garrido, A.J. Garrido. "Fuzzy logic control of an artificial neural network-based floating offshore wind turbine model integrated with four oscillating water columns," *Ocean Engineering*, vol. 269, p. 113578, 2023.

[27] H. Yu, J. Yang. "A direct LDA algorithm for high-dimensional data—with application to face recognition," *Pattern recognition*, vol. 34, no. 10, pp. 2067-2070, 2001.

[28] Y.A. Ghassabeh, F. Rudzicz, H.A. Moghaddam. "Fast incremental LDA feature extraction," *Pattern Recognition*, vol. 48, no. 6, pp. 1999-2012, 2015.

[29] S. Suthaharan, "Support vector machine," In: Machine learning models and algorithms for big data classification, Integrated Series in Information Systems, vol. 36, pp. 207-235. Springer, Boston, MA, 2016.

[30] D. A. Pisner, and D. M. Schnyer. "Support vector machine." In: Machine learning, pp. 101-121. Academic Press, 2020.

[31] M. Awad, and R. Khanna. "Support vector machines for classification," In: *Efficient Learning Machines*, pp. 39-66. Apress, Berkeley, CA, 2015.