Available online at www.eccomasproceedia.org
Eccomas Proceedia SMART (2023) 866-872

ECCOMAS Proceedia

X ECCOMAS Thematic Conference on Smart Structures and Materials
SMART 2023
D.A. Saravanos, A. Benjeddou, N. Chrysochoidis and T. Theodosiou (Eds)

ONLINE STRUCTURAL DAMAGE CLASSIFICATION METHODOLOGY FOR OFFSHORE WIND TURBINE FOUNDATIONS USING DATA STREAM ANALYSIS

JERSSON X. LEON-MEDINA[1*], NÚRIA PARÉS[2] AND FRANCESC POZO[1,3]

[1]Control, Data and Artificial Intelligence (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besòs (CDB), Universitat Politècnica de Catalunya (UPC), Eduard Maristany 16, 08019 Barcelona, Spain
e-mail: jersson.xavier.leon@upc.edu

[2] Laboratori de Càlcul Numèric (LaCàN), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besòs (CDB), Universitat Politècnica de Catalunya (UPC), Eduard Maristany 16, 08019 Barcelona, Spain
e-mail: nuria.pares@upc.edu

[3]Institute of Mathematics (IMTech), Universitat Politècnica de Catalunya (UPC), Pau Gargallo 14, 08028 Barcelona, Spain
e-mail: francesc.pozo@upc.edu

Abstract. Structural health monitoring (SHM) of wind turbines is crucial to improve maintenance and extend their lifespan. This study develops an online data analysis methodology using data stream analysis to classify damage in the links of an offshore wind turbine foundation. The methodology is validated using a laboratory-scaled jacket-type wind turbine foundation structure. 2460 measurements of the healthy structure were acquired, and a 5mm crack was applied to four different links to determine the four unhealthy classes. 820 measurements were taken for each of the unhealthy structures, resulting in a dataset with 5740 instances. As this is an imbalanced multiclass classification problem, a random sampler approach was used to treat the data. The only data obtained was from eight triaxial accelerometers distributed throughout the structure. Three different tree-based stream data classifiers were compared: Hoeffding Tree classifier, Extremely Fast Decision Tree classifier, and Hoeffding Adaptive Tree classifier. Each classification model underwent a tuning parameter procedure, and high values of the receiving operating characteristic area under the curve (ROC AUC) metric were achieved as a result. It is important to note that stream learning differs from batch learning.

Key words: stream learning; online data analysis; structural health monitoring; wind turbine foundation; machine learning; Hoeffding Tree classifier; data stream analysis; vibration analysis; condition monitoring

© 2023 The Authors. Published by Eccomas Proceedia.
Peer-review under responsibility of the organizing committee of SMART 2023.
doi: 10.7712/150123.9839.445442

1 INTRODUCTION

Wind energy is the leading source of renewable energy worldwide [1]. As of the end of 2021, the cumulative global installed wind capacity was approximately 837 GW [2]. To reduce the levelized cost of energy (LCOE) in the wind energy industry, various aspects must be evaluated, including operation and maintenance (O&M) costs. Advances in structural health monitoring (SHM) of wind turbines have improved their predictive maintenance. However, environmental and operational conditions in offshore wind turbines can create aggressive environments [3]. These conditions are related to the effects of wind and sea waves on structures. Recently, several techniques have been developed to carry out SHM of jacket-type wind turbine foundations, using only the vibration response of the structure with data acquired by accelerometers. These data have been used to train and test various machine and deep learning methodologies for performing structural damage classification in jacket-type wind turbine foundations.

Various machine and deep learning approaches have been used to process the data acquired from accelerometers and perform structural damage classification in jacket-type wind turbines, including Support Vector Machines [4], Extreme Gradient Boosting [5], Convolutional Neural Networks [6], Siamese Neural Networks [7], and Autoencoder Neural Networks [8]. Previous studies have typically used a batch setting approach to process the data. However, when continuous data streams are involved and models cannot be updated when new data arrives, stream data processing approaches become necessary [9]. This study presents a structural damage classification methodology based on stream data processing. Three different tree-based stream classifiers were compared to successfully solve an unbalanced classification problem in a jacket-type wind turbine foundation.

The remainder of this paper is organized as follows. Section 2 presents the materials and methods, describing the experimental setup and stream data processing algorithms. Section 3 presents the results obtained after performing a dimensionality reduction stage using the Principal Component Analysis (PCA) algorithm and the classification results using the Receiving Operating Characteristic Area Under the Curve (ROC AUC) metric. Finally, Section 4 describes the main conclusions.

2 MATERIALS AND METHODS

2.1 Experimental Setup

The data used in this study was obtained from a laboratory-scaled wind turbine. The wind turbine is 2.7 m high and consists of three parts: the jacket, the tower, and the nacelle. The structure is illustrated in Figure 1. To simulate the effects of marine waves and wind that offshore structures are subjected to, a white noise signal was applied to a shaker. This signal was then amplified in a function generator by factors of 0.5, 1, 2, and 3. Subsequently, the structure was excited, and its vibration response was measured using eight triaxial accelerometers located as shown in Figure 1.

Each experiment had a duration of 9.789 seconds, and a sampling frequency of 275 Hz was used. Thus, each measurement by each of the 24 accelerometers consisted of 2417 time instants.

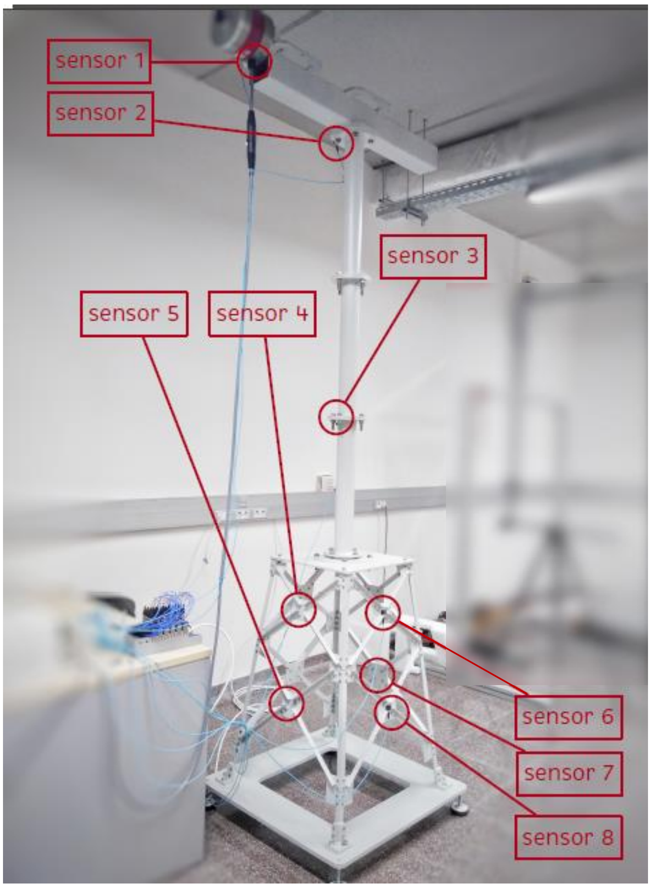

Figure 1: Placement of triaxial accelerometers on the jacket-type wind turbine foundation for vibration data acquisition.

After a data unfolding procedure, each signal was ordered one after the other such that the total size of a measurement was $2417 \times 24 = 58008$ data points per experiment. The structural damage introduced in the structure consisted of a 5mm crack in four different links, one at a time, as illustrated in Figure 2. There were five structural classes in total: the healthy structure and four different unhealthy structures. The total number of acquired samples of the healthy structure was 2460, while 820 samples were acquired for each damaged structure. Therefore, a total of 5740 samples were part of the final dataset obtained.

2.2 Hoeffding tree classifier

The Hoeffding tree [10, 11] is trained incrementally, using each sample in the stream only once. This means that the tree is the only information stored in memory [9]. The tree is updated by adding more leaves with its capability to grow. This tree can make predictions at any time; however, its performance improves incrementally as more data from the stream is processed.

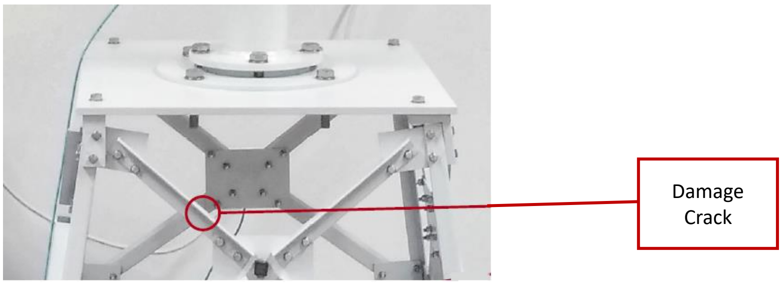

Figure 2: Visual representation of a damaged link with a 5 mm crack in the jacket-type wind turbine foundation.

2.3 Adaptive Hoeffding tree classifier

The Hoeffding Adaptive Tree method [12, 13] learns from streams that change their distribution over time. This method uses an adaptive sliding window size, based on defining instances of estimators of frequency statistics at every node [14]. The Adaptive Sliding Window (ADWIN) [15] is used as an estimator and drift detector. The Hoeffding Adaptive Tree variant measures drift in the branches in such a way that when accuracy decreases, the branches are replaced. In addition, bootstrap sampling is performed to improve its performance.

2.4 Extremely Hoeffding tree classifier

The Extremely Fast Decision Tree classifier (EFDT) [16, 17] has a variation in the way the algorithm performs splits. It does this by re-evaluating the performed split and determining if there is a better possible split, which is then used to replace the performed split. One important aspect of EFDT is that it maintains a steady tree size. The Extremely Fast Decision Tree classifier is also referred to as the Hoeffding AnyTime Tree (HATT) classifier.

3 RESULTS

The developed structural damage classification methodology was implemented in Python, using the *scikit-learn* [18] machine learning library and the *River* [19] dynamic data streams and continual learning library.

3.1 PCA dimensionality reduction

The Principal Component Analysis (PCA) method [20] was used to handle the high dimensionality of the two-dimensional unfolded matrix, which had 58008 columns. The new representation of the data is explained by the principal components. Eight principal components were used as input to the Hoeffding tree machine learning classifiers. These eight principal components represent a cumulative explained variance of 38%. Therefore, the reduced feature matrix input had a size of 5740×8. Figure 3 shows a bi-dimensional scatter plot of the second vs. third and first vs. second principal components, illustrating the formation of circular manifolds from the data.

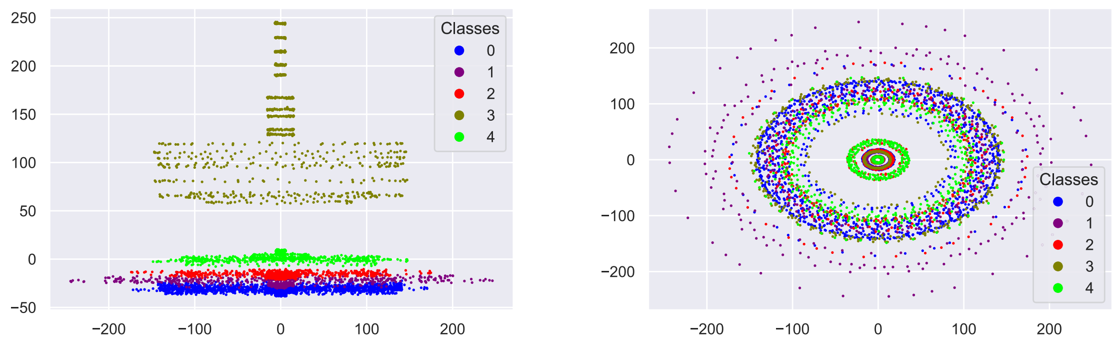

Figure 3: Scatter plot of the data in a reduced feature space obtained by the Principal Component Analysis (PCA) method. Left: second vs. third principal component. Right: first vs. second principal component.

3.2 Hoeffding tree stream classification

Different approaches were used to treat the imbalanced dataset used in this study. First, a baseline was constructed without modifying the samples. Second, a sampling method with a desired sample size and the random sampler method was defined using under-sampling and over-sampling according to a restriction imposed as a parameter. The size of the training data was also included as a parameter. Finally, third, a hybrid approach combining random under-sampling and importance weighting was used. For each of the three Hoeffding tree methods, the receiving operating characteristic area under the curve (ROC AUC) metric was calculated. The ROC AUC results are shown in Table 1. The results in the baseline setting show that the best method was the Hoeffding Adaptive Tree, reaching a ROC AUC value of 90.09%. Additionally, when the hybrid approach was used, an improvement in the ROC AUC value of the Extremely Fast Decision Hoeffding Tree was evident, reaching 92.48%.

Table 1: ROC AUC percentages obtained by the three different classifiers under three different settings: Baseline, Sampling with a Desired Sample Size, and Hybrid Approach.

	Hoeffding tree	Extremely Fast Decision Hoeffding Tree	Hoeffding Adaptive Tree
Baseline	72.73%	83.80%	90.09%
Sampling with a desired sample size	68.55%	64.23%	73.27%
Hybrid approach	49.15%	92.48%	85.09%

4 CONCLUSIONS

This study presented a structural damage classification methodology in a stream fashion using a Hoeffding tree. The methodology is composed of several steps, including data unfolding, data scaling, dimensionality reduction using the PCA method, stream learning classification using a Hoeffding tree, and the evaluation of model performance using a progressive validation score based on the calculation of the ROC AUC metric. The dataset used in this study was imbalanced and was treated using a random sampler technique. The hybrid approach combining random under-sampling with importance weighting yielded the best ROC AUC results. Furthermore, the Extremely Fast Decision Hoeffding Tree exhibited the best behavior, reaching a ROC AUC value of 92.48%. Further studies will focus on exploring different sampling techniques to treat the imbalance, as well as investigating the influence of the number of principal components as an input parameter in the classification process.

REFERENCES

[1] Z. Fan, S. Li, Z. Gao, L. Zhang, X. Zheng, W. Zhu, W. Shen, M. Sjöholm, T. K. Mikkelsen, T. Wang *et al.*, "On the importance of wind turbine wake boundary to wind energy and environmental impact," *Energy Conversion and Management*, vol. 277, p. 116664, 2023.

[2] G. W. E. Council, "Gwec global wind report 2022," *Global Wind Energy Council: Bonn, Germany*, 2022. [Online]. Available: https://gwec.net/global-wind-report-2022/

[3] X. Xu, F. Wang, O. Gaidai, A. Naess, Y. Xing, and J. Wang, "Bivariate statistics of floating offshore wind turbine dynamic response under operational conditions," *Ocean Engineering*, vol. 257, p. 111657, 2022.

[4] Y. Vidal, G. Aquino, F. Pozo, and J. E. M. Gutiérrez-Arias, "Structural health monitoring for jacket-type offshore wind turbines: Experimental proof of concept," *Sensors*, vol. 20, no. 7, p. 1835, 2020.

[5] J. X. Leon-Medina, M. Anaya, N. Parés, D. A. Tibaduiza, and F. Pozo, "Structural damage classification in a jacket-type wind-turbine foundation using principal component analysis and extreme gradient boosting," *Sensors*, vol. 21, no. 8, p. 2748, 2021.

[6] B. Puruncajas, Y. Vidal, and C. Tutivén, "Vibration-response-only structural health monitoring for offshore wind turbine jacket foundations via convolutional neural networks," *Sensors*, vol. 20, no. 12, p. 3429, 2020.

[7] J. Baquerizo, C. Tutivén, B. Puruncajas, Y. Vidal, and J. Sampietro, "Siamese neural networks for damage detection and diagnosis of jacket-type offshore wind turbine platforms," *Mathematics*, vol. 10, no. 7, p. 1131, 2022.

[8] Á. Encalada-Dávila, C. Tutivén Gálvez, B. Puruncajas Maza, and Y. Vidal Seguí, "Wind turbine multi-fault detection based on scada data via an autoencoder," *Renewable energy and power quality journal*, vol. 19, pp. 487–492, 2021.

[9] R. Kirkby, "Improving hoeffding trees," Ph.D. dissertation, University of Waikato, 2007.

[10] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2000, pp. 71–80.

[11] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, T. Abdessalem *et al.*, "Hoeffding tree classifier," 2023, https://riverml.xyz/dev/api/tree/HoeffdingTreeClassifier/.

[12] A. Bifet and R. Gavalda, "Adaptive learning from evolving data streams," in *Advances in Intelligent Data Analysis VIII: 8th International Symposium on Intelligent Data Analysis, IDA 2009, Lyon, France, August 31-September 2, 2009. Proceedings 8*. Springer, 2009, pp. 249–260.

[13] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, T. Abdessalem *et al.*, "Hoeffding adaptive tree classifier," 2023, https://riverml.xyz/dev/api/tree/HoeffdingAdaptiveTreeClassifier/.

[14] A. C. Bifet Figuerol and R. Gavaldà Mestre, "Adaptive parameter-free learning from evolving data streams," *Research report, Universitat Politècnica de Catalunya*, 2009. [Online]. Available: https://upcommons.upc.edu/handle/2117/87914

[15] A. Bifet and R. Gavalda, "Learning from time-changing data with adaptive windowing," in *Proceedings of the 2007 SIAM international conference on data mining*. SIAM, 2007, pp. 443–448.

[16] C. Manapragada, G. I. Webb, and M. Salehi, "Extremely fast decision tree," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1953–1962.

[17] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, T. Abdessalem *et al.*, "Extremely fast decision tree classifier," 2023, https://riverml.xyz/dev/api/tree/ExtremelyFastDecisionTreeClassifier/.

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[19] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier *et al.*, "River: machine learning for streaming data in python," *The Journal of Machine Learning Research*, vol. 22, no. 1, pp. 4945–4952, 2021.

[20] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.