Reviewer #1: This paper proposes a novel damage detection method for offshore wind turbines (OWT), which is used to consider the relative relationships between adjacent turbines as well as remove the environmental and operational variations. The manuscript is organized well and in line with the scope of the journal. However, after careful review, this manuscript cannot be recommended to be published. The specific comments are shown as follows:

We are grateful for your thoughtful comments and suggestions. Below are our responses to your comments.

1. Based on the existing introduction, the novelty and contribution of the proposed method cannot be illustrated well. It is suggested to modify this section to better state the novelty.

Thank you for your insightful comment. We agree that the introduction could better highlight the novelty and contributions of the proposed method. In response, we have revised the introduction to explicitly emphasize the novelty: innovative use of inter-turbine relationships. The method leverages the inherent similarity between adjacent turbines in a wind farm to eliminate the effects of environmental and operational variations (EOVs). This approach bypasses the need for additional sensors or EOV measurements, which are often costly or unavailable, transforming a significant challenge into a practical and scalable solution.

2. The proposed method is merely validated based on simulations. Also, authors have not given the details of the three-OWT wind farm, for example: How to simulate them? Which element do you choose for the OWT? Which method is used to obtain the structural responses? The images of the simulation?

Thank you for your valuable comment. We acknowledge that the proposed method has been validated solely through numerical simulations and does not include real-world data. The simulations were performed to demonstrate the feasibility and effectiveness of the method under controlled conditions, serving as a proof-of-concept. While this limits real-world validation, we agree that applying the method to real-world data is an important next step to evaluate its performance further. However, due to proprietary restrictions, access to real-world data for offshore wind turbines, particularly those experiencing scour damage, remains limited.

Regarding the simulation, we have ensured that the modelling techniques used are consistent with industry practices. Specifically, we used DNV Bladed software, which has been extensively validated through its consistent participation in the International Energy Agency (IEA) Wind Task 30 projects, including OC3 [1], OC4 [2], OC5 [3], and OC6 [4]. These international collaborations rigorously benchmarked simulation tools against experimental data and advanced models for offshore wind systems. Bladed’s active involvement in all phases demonstrates its reliability and accuracy in addressing complex dynamics of offshore wind turbines, solidifying its position as a trusted and well-validated tool in the industry.

Below, we provide additional details about the three-OWT wind farm simulation:

Simulation Setup. The simulated wind farm consists of three monopile-supported offshore wind turbines (OWTs) with identical structural properties but varying water depths to account for site-specific variations. This variation simulates realistic conditions of a small wind farm. Please note that the three OWTs are simulated separately in the software.

Finite Element Modelling. The turbines were modelled using Timoshenko beam elements for the tower and monopile, while the rotor-nacelle assembly was treated as a lumped mass. The soil-structure interaction was represented using linear springs. The whole system was linear.

Wind and Wave Field Generation. Independent turbulent wind fields were generated for each OWT using predefined seeds to ensure variability while maintaining realistic conditions. Irregular wave fields were also simulated for each turbine, with hydrodynamic forces calculated using Morison’s equation.

Structural Response Extraction. The tower top accelerations were simulated under the combined effects of wind and wave loading. These responses were then processed to extract the natural frequencies of the OWTs using the automated Operational Modal Analysis (OMA).

Visualization. Unfortunately, we no longer have access to Bladed after the first author's graduation. To facilitate understanding of the simulation setup, however, we have included an illustration (Fig. 1) of the model in Bladed.

A diagram of a wind turbine

Description automatically generated

Fig. 1. Illustration of the model in Bladed.

References:

[1] Jonkman, J., & Musial, W. (2010). *Offshore code comparison collaboration (OC3) for IEA Wind Task 23 offshore wind technology and deployment* (No. NREL/TP-5000-48191). National Renewable Energy Lab.(NREL), Golden, CO (United States).

[2] Popko, W., Vorpahl, F., Zuga, A., Kohlmeier, M., Jonkman, J., Robertson, A., ... & von Waaden, H. (2012, June). Offshore code comparison collaboration continuation (OC4), phase 1-results of coupled simulations of an offshore wind turbine with jacket support structure. In *ISOPE International Ocean and Polar Engineering Conference* (pp. ISOPE-I). ISOPE.

[3] Robertson, A. N., Wendt, F., Jonkman, J. M., Popko, W., Borg, M., Bredmose, H., ... & Guerinel, M. (2016). OC5 Project Phase Ib: validation of hydrodynamic loading on a fixed, flexible cylinder for offshore wind applications. Energy Procedia, 94, 82-101.

[4] Bergua, R., Robertson, A., Jonkman, J., Branlard, E., Fontanella, A., Belloli, M., ... & Goveas, A. (2023). OC6 project Phase III: validation of the aerodynamic loading on a wind turbine rotor undergoing large motion caused by a floating support structure. *Wind energy science*, *8*(4), 465-485.

3. In section 3.1, authors mentioned that this study adopted a clustering method to obtain the structural modal parameters because the p-LSCF method relies on users' expertise, but, in Section 5.1 and Fig.7, authors mentioned the modal parameters are obtained based on p-LSCF method, please explain it.

We have clarified this point in the manuscript. The clustering method is used to automate the selection of modal parameters after they are initially identified using the p-LSCF method. This automation reduces reliance on user expertise, ensuring consistency in the analysis process. A detailed explanation has been added to Section 2.1, which now reads as follows:

“However, the conventional p-LSCF method requires manual selection of poles, which relies on users’ expertise. This is not ideal for SHM, as SHM is intended to automatically monitor the state of structures. To address this limitation, several studies have proposed methods to automate the pole selection process [49, 50]. In this paper, a clustering method based on Devriendt et al. [1] is employed. The clustering method is designed to replicate the decisions made by an experienced modal analyst when examining a stabilization diagram.”

This clarification should resolve any confusion about the integration of the p-LSCF and clustering methods in the analysis. Thank you for pointing this out.

4. Which order of natural frequency has been used in Fig.9 for comparison? It is likely to be the 4th mode, please clarity this point. And it should be explained why authors adopted this mode rather than others.

Thank you for your comment. We confirm that the 4th mode (second side-side natural frequency, SS2) was used in Figure 9. This mode was chosen because it has higher sensitivity to scour damage compared to other modes, as reported in Devriendt et al. [1] and Weijtjens et al. [4]. We had addressed this point in the manuscript and explicitly stated the choice and justification for this mode in Section 5.1 - Automated Operational Modal Analysis in the original manuscript.

To improve readability, we have restructured Section 5.1 (now Section 4.1 in the revised manuscript) by moving part of the original first paragraph, which explains the selection of the SS2 mode, to the middle of the section. Additionally, we have revised the text to further emphasize this explanation, ensuring the rationale is clearer for the readers. Thank you for highlighting this and helping us enhance the clarity of the manuscript.

5. The manuscript about damage detection based on adjacent turbine lacks critical and scientific statement. It should be suggested to give more deep discussion on adjacent turbines, for example, what is the influence mechanism between adjacent turbines and why you can detect damage based on the natural frequencies of adjacent turbines.

Thank you for pointing this out. We agree that there is a lack of scientific statement to address the mechanism of the proposed method.

The underlying concept of the proposed method relies on the fact that adjacent turbines within the same wind farm experience similar environmental and operational variations (EOVs), such as wind and wave conditions. As a result, the natural frequencies of adjacent turbines are strongly correlated under undamaged conditions. When damage occurs in one turbine, such as scour or structural degradation, it alters the stiffness and dynamic properties of that turbine, leading to a deviation in its natural frequency. This deviation disrupts the correlation with the natural frequencies of its neighbouring turbines, enabling the detection of damage.

We have added a more detailed discussion of this mechanism in the end of Section 1, emphasizing how the method leverages these inter-turbine relationships to isolate damage-induced changes from EOV-induced variations. Additionally, we explain why this approach enhances robustness and reliability in damage detection compared to methods relying solely on single-turbine data.

------------------------------------------------------------------------------------------------------------  
Reviewer #2: The authors tackle an important issue in offshore wind farms by presenting a output-only approach for damage detection. While the paper approaches scour damage scenarions, the same philosophy can be extended for other components. Although I have not detected major flaws on the presented results, I do feel that the paper falls a little bit short in the depth of the analysis conducted and/or in the scenarios studied, while the presented methodology could allow for a much more extensive study. Below follow come comments that I believe could improve the final result.

We sincerely appreciate your thorough review and constructive feedback on our manuscript. They have helped us improve the quality of our manuscript. Below are our responses to your comments.

Major comments:  
- I do not think the description of the Problem justifies its own section. I suggest to merge Sections 2 and 3

Indeed, merging Sections 2 and 3 makes the manuscript more cohesive. Thank you for the suggestion. We have combined them in the manuscript.

- In page 5, line 55 it is written that the OWT can be considered a linear system when it is parked. Following the previous discussion starting in line 53, I do not see how a parked rotor eliminates the non-linear soil-pile-interaction. I do understand that using parked conditions may be reasonable for the purpose of this study (although it also reduces very significantly the number of available events for experimental applications), but I think this justification should be removed.

Regarding the non-linear soil-pile-interaction, we are specifically referring to the p-y curve (soil lateral resistance vs displacement), because this affects the natural frequencies the most, comparing to other soil curves (vertical or torsional responses). One characteristic of p-y curves is that the curves are almost linear for small responses. As an example, we took a screenshot of the Fig. 5 in the paper of Burd, Harvey J., et al. (2020), as shown below. It can be seen that the curves are generally linear for small responses.

When the wind speed is below the cut-in speed (normally 3m/s), the rotor will be parked. In this condition, the rotor blades are feathered and thus the wind force is small. Due to the small wind speed, the corresponding sea states are normally mild and thus small wave forces can be expected. Therefore, the total force applied to the OWT is small, and the soil-pile-interaction can be assumed as linear.

The requirement of parked conditions (due to the wind speed smaller than the cut-in speed) indeed significantly limits the applicability. It is still considered to be practical, as the wind speeds below 3m/s can normally take up 5%-10% in the wind speed distribution.

In the light of the above discussion, we inclined to keep the justification.

Meanwhile, we realised the justification was not well explained in the manuscript and we have revised it accordingly, as below.

“However, it can be seen as linear if the wind speed is below the cut-in speed and thus the OWT is parked. In this condition, both the wind and wave forces are small. As small responses are expected, the soil-pile-interaction can be regarded as linear (Burd, Harvey J., et al. (2020)).”

A group of graphs showing different types of growth

Description automatically generated with medium confidenceReference:

Burd, Harvey J., et al. "Application of the PISA design model to monopiles embedded in layered soils." Géotechnique 70.11 (2020): 1067-1082.

- In page 9, line 38 it is said that "nearly all of the sample points will fall between them [LCL and UCL]". While I agree that this is true, it contradicts the sentence that follows it that mentions "a point that plots outside of the control limits indicates that the process is out of control". Please rephrase.

The sentence has been rephrased for clarity. We now state it as below

“These control limits are chosen so that if the process is in control (e.g., the structure is undamaged), all the sample points will fall between them, while points outside the limits indicate potential damage and an alarm will be triggered.”

- In page 9, line 56, the monitored statistics are essentially defined as an average residual. Did the authors consider to weight the residues based on the predicted uncertainty level? Wouldn't this be a better metric?

Thank you for your insightful comment. We agree that weighting the residuals based on the predicted uncertainty level could provide a more balanced metric by accounting for varying confidence levels in the predictions. In the current study, the residuals were defined as simple averages to maintain a straightforward implementation and focus on the method's core principles.

Gaussian Process Regression (GPR) inherently produces uncertainty estimates along with its predictions, making this suggestion particularly promising. Incorporating these uncertainty estimates could emphasize predictions with higher confidence (lower uncertainty) and down-weighting those with lower confidence (higher uncertainty). This would potentially improve the method's sensitivity to anomalies and enhance its overall robustness.

We appreciate this suggestion and have included it as a potential enhancement for future work. Thank you for pointing out this valuable opportunity.

- In Eq. 17 are defined the LCL and UCL based on a 3sigma interval. Have the authors made any sensitivity analysis to tune this range?

Thank you for your question. No, we did not perform a sensitivity analysis to tune the range of the 3-sigma interval. The 3-sigma setting was selected as it is a widely used threshold in control chart applications and has demonstrated effectiveness in the current study. We acknowledge that adjusting this range based on specific application scenarios could further improve the method’s performance and consider this an interesting direction for future work.

- In Sec 4.1 the authors argue that it is equivalent to consider a varying tide level or different significant wave height and use this argument to study only the former. This is one of the critical points for me, since I do not agree at all with the claim: different significant wave heights may enhance or reveal non-linear dynamic effects that a different tide level simply doesn't. I am willing to accept that in this numerical study the tide level is the relevant parameter for frequency variations (mostly since the seabed stiffness matrix is constant), while I expect that waves energy content may be more relevant for the effective damping of the system, but I do not think the argument can be put as it is in the submitted manuscript.

Thank you for your insightful comment. We have revised this part to address the potential misunderstanding and clarify our reasoning. The updated text now reads as follows:

“In this study, a parked OWT operating under cut-in wind speed and calm sea conditions was assumed. Under these conditions, both tidal levels and wave heights primarily affect the natural frequency through added mass effects. To simplify the analysis, only the tide level was used to represent the combined influence of tidal levels and wave heights.”

We believe this revision avoids any misleading claims and better reflects the assumptions and scope of our study. Thank you for bringing this to our attention.

- In Eq. 19 the seabed stiffness matrix is assumed to be constant. Have the authors considered studying a non-linear one, or different scenarios with different excitation levels and different stiffness values to mimic a non-linear behaviour?

Incorporating non-linear soil stiffnesses will indeed be exciting to explore, as it broadens the potential application scenario of the method, extending beyond calm sea conditions. While we used a linear seabed stiffness matrix in this study, we agree that investigating non-linear behaviours could significantly enrich the analysis. We have included this as a recommendation for future research in Section 5.

- Were the simulations of OWT1, OWT2 and OWT3 conducted using independent wind and wave fields? If not, how are these correlated?

Yes, they are independent. This is achieved by setting distinct seeds for different OWTs when generating wind or wave fields. We acknowledge that, in reality, wind and wave fields are correlated to some extent; however, the wave effects can be ignored for parked OWTs. Despite this, the wind speed time series (or wave elevation time series) across OWTs should not be identical due to the spatial separation of the turbines, which introduces time lags in these series. For simplicity, we have used independent wind and wave fields, and we believe this approach demonstrates the robustness of the proposed method.

- The gaussian process regression defined for instance in Eq. 20 uses a directional link between the different OWT (for instance, f2 is predicted based solely on f1). Have the authors compared the results when f2 is also predicted from f3, or, more generally, predicted by f1 and f3?

“Have the authors compared the results when f2 is also predicted from f3?”

Thank you for your suggestion. In response, we have performed a study where f2 is predicted from f3, as illustrated in Fig. C1. Additionally, the other two prediction directions in Fig. C1 are also reversed compared to Fig. 6 in the manuscript. The corresponding results are shown in Fig. C2, similar to Fig. 11 in the manuscript.

From Fig. C2, it can be observed that subfigures (a) and (c) are out of control, whereas Fig. C2(b) is in control. Based on this result, OWT1 is identified as damaged. This case study has been added to the manuscript to further explore the effects of reversing the prediction directions (Section 4.4).

A diagram of a triangle with red and green circles and black text

Description automatically generated

Fig. C1. The reverted prediction directions between OWTs.

A diagram of a graph

Description automatically generated with medium confidence

Fig. C2. X-bar control charts of the residuals for: (a) OWT 1, (b) OWT 2 and (c) OWT 3 of the reverted prediction relationship. The horizontal lines represent the UCLs and LCLs.

“or, more generally, predicted by f1 and f3”?

Yes, this is an exciting direction and aligns with the authors’ vision for future development of this work. To start with, there are two potential approaches to incorporate information from both f1 and f3 when predicting f2.

1. Approach 1 – Combining Information in Prediction Relationships.

As shown in Fig. C3, f1 (OWT1) is predicted by f2 (OWT2) and f3 (OWT3). The residuals from these two predictions are then combined to monitor f1​. A straightforward method to achieve this is by adding the prediction residuals from f2​ and f3​, effectively summing the results from Fig. 11 in the manuscript and Fig. C2.

The results are displayed in Fig. C4. It can be seen that in Fig. C4(a), the performance for OWT1 improves, as the combination of two positive residuals enhances detectability. However, for OWT2 in Fig. C4(b) and OWT3 in Fig. C4(c), the residuals during the potentially damaged condition (31 May to 7 June) deviate from zero and are less consistent with the undamaged condition. While damage detection remains possible with adjusted control limits, the overall performance is not expected to exceed that of the original method.

A diagram of a diagram

Description automatically generated

Fig. C3. Illustration of incorporating multiple inputs – Approach 1.

A group of graphs with numbers

Description automatically generated

Fig. C4. Summed X-bar control charts of the residuals for: (a) OWT 1, (b) OWT 2 and (c) OWT 3, based on the residuals in Fig. 11 in the manuscript and Fig. C2. OWT1 is predicted by OWT2 and OWT3, as illustrated in Fig. C3. The same applies to OWT2 and OWT3.

1. Approach 2: Combining Inputs in the Gaussian Process Regression Model

In this approach, the monitored data (f1 and f3) are both used as inputs in the Gaussian Process Regression model to predict 𝑓2. This transforms the problem into a 2-dimensional input and 1-dimensional output scenario in GPR. We believe this approach holds significant potential for improving performance by leveraging additional information in the predictive model.

We appreciate this suggestion and have included it as a direction for future work in the manuscript.

Future Perspectives

Beyond using two OWTs (e.g., f1 and f3) for predictions, additional OWTs could be incorporated. This leads to a broader perspective where the wind farm is conceptualized as a network of interconnected units, such as 70 turbines. This opens exciting possibilities for linking this study with concepts from graph theory and social network analysis, enabling the development of more advanced methodologies for monitoring and damage detection in wind farms.

- The damage is introduced in one of the similar OWT (with the same water depth). Have the authors analysed the case where the damage was introduced in OWT3? Can these results be also included?

We did not analyse the case where damage was introduced in OWT3 in the original manuscript. OWT3 is unique among the three OWTs due to its water depth. Introducing damage to OWT3 would indeed be a meaningful case study, further demonstrating the robustness of the proposed method. We acknowledge the value of such an investigation and are happy to discuss its potential implications. However, the authors no longer have access to Bladed following the first author’s graduation. While it might be possible to reconnect with the Bladed team and apply for a new license, this process would involve significant uncertainties and potential delays.

To provide some perspective, we present an illustration (Fig. C4) that conceptually outlines our expectations for this case. While the figure is not based on new data, it represents a hypothetical scenario to explain how the proposed method might perform if damage were introduced to OWT3. In this scenario, we expect f3 to decrease due to stiffness loss caused by the damage. As shown in Fig. C4(b), the green line representing f3 from 31 May to 7 June drops compared to its undamaged state.

From this conceptual scenario:

1. The prediction from f2 to f3 (blue arrow in Fig. C4) would likely perform better in Fig. C4(b) than in Fig. C4(a) due to the increased vertical distance between the frequencies (i.e., the vertical distance of the blue arrow).
2. Similarly, the prediction from f3 to f1 (green arrow) would also improve under these conditions.
3. Predictions from f1 to f2 (red arrow) are expected to remain unaffected as this relationship is unchanged by the damage to OWT3.

Thus, based on these expectations, we infer that the overall performance in Fig. C4(b) would surpass that in Fig. C4(a), indicating that the proposed method would likely perform well when OWT3 is damaged.

A screenshot of a graph

Description automatically generated

Fig. C4. Hypothetical illustration for discussing damage detection performance. Dates from 16 May to 31 May represent the undamaged state, and dates from 31 May to 7 June represent the potentially damaged state. Prediction directions are the same as in Fig. 6 of the original manuscript. (a) OWT1 has damage, as in the original manuscript. During 31 May to 7 June, f1f\_1 decreases due to the damage. (b) OWT3 has damage. f3f\_3 decreases starting on 31 May when the damage is introduced.

Minor comments:  
- Some abbreviations are not defined in the text: BIC in page 7, line 36; SS2 in page 15, line 34 (although this one is defined later)

The abbreviations have now been defined upon their first use in the manuscript.

- In page 10, line 57, it is written "The wave was" where it should be "the wave field was" or the "the waves were". Please correct.

This has been corrected to "the waves were" in the manuscript.

- In page 20, line 50, it reads "greed" where it should be "green".

It has been corrected.

- Figures 4, 5, 6, 7, 10 and 11 should be moved further ahead in the text for the final version (I do understand that the final version layout may be different, but consider this).

Thank you for the suggestion. We will ensure the placement of these figures aligns with the content flow in the final version.

- After Eq. 18 Md is defined as the displaced water mass. I think it would be better defined as the mass of water with the same volume as the submerged body.

Your suggestion is indeed more accurate. We have updated the definition accordingly in the manuscript.

- In dates defined in Table 3 are used only as a reference to the tide levels in Fig 9, correct?

Yes, the dates are solely used to indicate the timestamps of the tide levels. We have replaced the header “Date” with “Tide level timestamps” for clarity.