Response to Anonymous Referee Interactive Discussion

The authors would like to acknowledge the referee for the insightful and constructive comments regarding the manuscript. In the following, we answer to the comments and questions.

Scientific comments

1. In many machine-learning contexts (e.g., VAEs or standard autoencoders), the forward (decoder) and inverse (encoder) are trained jointly with a single objective. Clarify why two-step training is chosen over a single integrated approach, and discuss potential pros/cons.

In most autoencoders, the encoder and decoder must be found with a common goal. In contrast, herein, the decoder is dictated by a physical law and, therefore, it is fixed. We pre-train a Neural Network to approximate that physical law (decoder) for computational purposes since this enables access to its derivatives needed when training the encoder. Other than that, the decoder should follow the physical law and, therefore, it must be fixed. By having fewer unknowns (only those corresponding to the encoder) we minimize the problem difficulty; in particular, we decrease the number of local minima.

2. In the proposed method, the training of the forward operator is deterministic. Have you considered a probabilistic (or noisy) surrogate as well?

No, because the physical law governing the system is deterministic. The data employed to train the model incorporates noisy responses to account for measurement error. Since this work aims to explore how the uncertainty transfers from the measured observations to the estimated damage properties, we have neglected analyzing the uncertainty in the forward model.

3. Equation 9 needs to be clarified. What are the assumptions for the prior p(z)?

We assume uniform prior for the damage condition properties z1 and z2, constrained to the domain [a, b], as indicated in line 244 in the manuscript draft.

4. “Substituting Eq 9 in Eq 12 gives Eq 13”. This part needs more detailed explanation or derivation steps for better comprehension.

Thank you for the observation. We will provide a more detailed derivation of the expression.

5. What are the features used for the measurements? Statistics of the time series, properties of the PSD?

In section 3, we describe the specifications of the case study, indicating the selected features employed as the measurements. We employ five features, including time series statistics and frequency-domain features from acceleration signals: the mean, the standard deviation, two dominant peak frequencies, and the zero-th momentum. Exploring Cross Power Spectral Density (CPSD) features will be considered in future work.

6. What are the architecture and training parameters of the deterministic counterpart?

We have omitted the details related to the decoder (forward operator) as it was optimized in a previous work cited in this work. We will incorporate an appendix to summarize this information.

7. Does the provided uncertainty represent aleatory, epistemic components or mixed? In the latter case, how to decompose it?

This work focuses on exploring the aleatoric uncertainty, assuming that the available training data sufficiently covers all the potentially observed scenarios. As it is a proof of concept, we have selected frequently occurring damages and assume that no different damage may occur. However, we consider as future work handling this uncertainty.

8. The study focuses on two specific damage types within a single mooring line. This constraint simplifies the problem but may not represent the diversity of real-world conditions, where multiple damage types may occur at various locations.

We appreciate the comment and agree with the limitations of the proposed work. However, as a proof of concept, we consider it mandatory to constrain the problem. The lack of experimental data from damaged turbines requires generating computationally expensive synthetic measurements, forcing us to constrain the problem for practical reasons.

9. Damage data and measurement data will be needed to train the forward operator in a supervised manner. Similar to the above comment, it works in the case study because the same damage mode is simulated for both training and testing. But in reality, it wil not be the case. In this context, how would you address the following barriers for practical application:
    1. Damage data can be rarely collected from the real structure and therefore, simulations will be required to train the model. The difference between the simulated response and the real turbine response will affect the model robustness.

The authors agree that this issue is one of the key barriers to SHM applications. Incorporating a calibration task according to a set of available measurements enables the reduction of the discrepancy between the real and the simulated domains. Performing domain adaptation techniques to take advantage of limited raw experimental data and enhance synthetic measurements is a key challenge to overcome this limitation, which is a future research line for us. However, it depends upon the availability of real data.

    2. The simulated dataset will not cover all possible damage scenarios.

We agree that enlarging the dataset to sweep a wide range of possible damage scenarios is challenging (experimental data scarcity and computational cost). Thus, we must prioritize the most frequently occurring damage cases according to the knowledge from experts in the field and the experience of aging or long-term instrumented systems. Once we gain access to experimental data from operating systems, we will incorporate more cases prone to occur during service in future works.

Technical comments

Literature review should be more structured.

We will review the literature review organization.

Citation style should be proper and consistent throughout the manuscript. Cite in parentheses wherever is relevant.

We will review the citation style.

The manuscript still needs careful and thorough proof-reading.

We will review the manuscript writing to look for errors.

Thank you.

Sincerely,

A. Fernandez-Navamuel, N. Gorostidi, D. Pardo, and V. Nava.