Damage identification on a large-scale wind turbine rotor blade using sample-based deterministic model updating

Answers to the reviewers' comments

The authors would like to thank the editor and reviewers for their time and effort to review the article and for their constructive and relevant comments. We appreciate the chance to clarify the addressed points. This will undoubtedly help in improving the quality of the manuscript.
The article is revised according to the reviewers' remarks and queries, with detailed responses and explanations of the edits given below. Edits in the manuscript have been highlighted in red color for the first reviewer and in blue color for the second reviewer. We note that during the revision, the amount of lines has been changed. Therefore, in the proceeding paragraphs, we refer to the revised article.

Reviewer #1

In their manuscript submitted for publication in Wind Energy Science, the authors present a study of fatigue damage identification for a wind turbine rotor blade. The blade is a specimen of 31 m long which was tested in the lab and subjected to cyclic edgewise loading in order to generate (realistic) fatigue damage. During the test, the blade was instrumented with accelerometers to monitor changes in modal parameters resulting from the fatigue damage. Three states of the blade are considered in the damage identification which occurs through the updating of a finite element model of the blade. A beam as well as a shell model of the blade are used, where damage is represented as a reduction in stiffness in a zone of the blade, considering different damage parameterizations. As on objective function, a fit of the difference in modal properties between two (of the three) states is considered rather than tuning the model to each distinct state and subsequently checking the difference in parameters. Statistical uncertainty in the estimated modal characteristics is considered by repeating the updating for 53 sets of individually identified modal characteristics for different time records. A multi-objective optimization approach is used in the model updating of each set, considering a trade-off between the fit in the difference in natural frequencies between two states and the fit in the difference in eigenmode. It is concluded that the shell model provides the most accurate and reliable characterization of the evolution in damage from one state to another.

The work presented in the manuscript is valuable, in particular for which concerns the experimental data of the blade which are made publicly available on a repository of the institute of the first author. The damage identification presented in the manuscript is also of potential interest but I would like the authors to consider the following comments before it is given further consideration for publication:

1) A multi-objective optimization approach in model updating was previously presented by the group of Costas Papadimitriou at the University of Thessaly, see e.g. K. Christodoulou et al., CMAME, 2008. In their work, a Pareto front considering optimal solutions corresponding to trade-offs between fits in natural frequencies on the one hand and mode shapes on the other hand was presented as well. It also provides some interesting insight into the influence of model error on the trade-off. Please include this work in the state of the art and use it to situate the present work.

Thank you for this suggestion. We found the work very interesting and included it in our introduction, lines 45f, and as additional reference for multi-objective model updating in line 91. The added literature item is also marked in red color in the references.

2) The authors state that the key idea behind the SDMU approach "is to exclude the uncertainty from the design-variable dependent part of the objective function... Instead the uncertainty is incorporated by generating multiple discrete input samples". Please be more specific as to what type of uncertainty is considered here as "the uncertainty" in the present formulation suggests that uncertainty in general is considered. If I understand correctly, however, it is mainly the statistical uncertainty on the identified modal characteristics which is considered here and transferred to the model parameters (but this relates also to the next comment).

Thank you for this inquiry, we acknowledge that we did not specify the kind of considered uncertainty in enough detail. You are correct, we consider and propagate input uncertainty. In particular, in this work, the identification uncertainty of the modal parameters, including measurement uncertainty, is considered. Using the SDMU approach, this input uncertainty is propagated to the output, which is presented using CDFs that contain all (Pareto) optimal solutions (cf. next answer).
We revised our description of which uncertainty is particularly considered in this work (i.e., identification uncertainty) in Section 1, lines 93f and 107f. In addition, throughout the whole manuscript, we specified that, using the SDMU approach, input uncertainty is considered (cf. pages 1, 3, 8, 9, 10, 19; revisions are made in red color). In this context, we do not always refer to identification uncertainty in particular, which is specifically considered in this work, but rather revised the general term "uncertainty" to "input uncertainty". This is because the presented SDMU approach can also be utilized to consider additional input uncertainty and is not restricted to identification uncertainty.

3) It is not entirely clear how the CDF's of the model parameters are generated. It is clear that these CDF's cover the variability residing in the combination of the 53 sets for each two damage states, but it is not clear if these also include the variability in the Pareto optimal solutions of the multi-objective optimization framework for a specific set.

Thank you for this remark. Indeed, in this work, the CDFs include all Pareto-efficient solutions of the N_{sets}^2 optimization runs performed during one SDMU procedure. In response to your remark, we explained the determination of the empirical probability and the resulting CDF in more detail and provided the corresponding formulas in Section 2.2, lines 275f. Note that the original version of this paragraph was placed in Section 4 in the original manuscript and was now added to Section 2.2.

4) In my personal opinion, reporting results in terms of CDF's is less intuitive than histograms? This seems confirmed by the fact that the authors focus more on the slope of the CDF's rather than on the actual CDF's in the discussion, for example when it is stated "this probability reaches approximately 50%".

Thank you for this comment. We agree that, at first sight, histograms ore often more intuitive. However, histograms require the selection of bin sizes and bin locations, which introduces an additional level of arbitrariness and may influence the (visual) interpretation of the results. This is the reason why we chose to present CDFs, as they are independent of such choices and provide a complete and "foolproof" representation of the resulting distribution without discretization bias. We clarified this choice in more detail in Section 2.2, lines 288f.

5) The authors consider two finite element models of the structure, a beam model and a shell model, and this with different parameterizations of damage. First, I am wondering to what extent the simplification of the load shears to point masses in the beam model has an impact on the results. Wouldn't it be a more fair comparison between the beam and the shell model if the beam model also accounted for the rotational inertia of the load shears so that this part of the model is closer to the shell model. Second, one wonders how large the difference is between the results of both models. Can't the authors include a comparison of natural frequencies and mode shapes (e.g. MAC matrix for sensor positions) for the two models, using nominal parameter values in undamaged conditions.

Thank you for these suggestions. Regarding the first inquiry: This is an interesting question, which we did not explicitly consider in our previous analysis. The reason for this is that the model-updating procedure employed in this work is based exclusively on the first five eigenfrequencies and corresponding eigenmodes, which are predominantly bending-dominated. Consequently, torsional modes, for which the rotational inertias of the load shears may be more influential, are not explicitly targeted.

Nevertheless, we investigated your inquiry during the review process. We extracted the mass moments of inertia (i.e., rotational inertias) of the load shears from the shell model, where they are represented in full geometric detail, and added these as rotational inertias to the corresponding mass elements in the beam model using the "ROTARYI" element type in Abaqus. For the first five eigenfrequencies and the associated eigenmodes considered in the model-updating process, no significant differences were observed between the results obtained with and without implementing the additional rotational inertias. Hence, the simplification of the load shears only to point masses in the beam model did not affect the results relevant to this work. For your information, we listed the first five eigenfrequencies of the beam models with and without additional rotational inertias in the following Table 1.

Table 1: Comparison of the eigenfrequencies in Hz of the beam models with different load shear representation.

Eigenfrequency	Load shear representation	
No.	Mass	Mass and rotational inertia
1	1.020	1.019
2	1.870	1.868
3	2.791	2.782
4	4.741	4.730
5	5.692	5.666

For higher-order modes, we could observe a more pronounced influence, indicating that a more accurate representation of the load shears becomes increasingly relevant when a larger number of modes is taken into account. Therefore, we agree that this is an important aspect and added a short discussion in Section 3.2, lines 371f.
During the revisions made in response to this remark, a minor inconsistency in the representation of the first load shear was identified in the shell model. Subsequent verification confirmed that this discrepancy has no influence on the dynamic response of the FE model. Moreover, due to the relative formulation of the objective function, minor modeling inaccuracies would not affect the model-updating results in any case. Consequently, this minor discrepancy is considered negligible and is reported here solely for the sake of completeness.

Regarding the second inquiry: We added a comparison of the relevant modal parameters in Section 3.2, lines 382f. In particular, we added Table 5, listing the difference in the first five eigenfrequencies and the corresponding eigenmodes, quantified using the MAC value, between the two FE models in their healthy state and the modal parameters identified from the measurement data in rotor blade state I. Furthermore, we added Figure 8 to the same section, visualizing the first five eigenmodes of the beam and shell model in reference state in comparison to the eigenmodes identified using the measurement data of the rotor blade in state I.

6) One wonders at the end also what model would be preferred? In a Bayesian approach, one could apply a model selection approach which considers a trade-off between the complexity of the model and the fit of the data. Is a similar reasoning possible here?

Thank you, we acknowledge that we did not go into detail regarding this question before. We agree that, in principle, a Bayesian model selection approach could be applied to compare the different modeling fidelities using a quantified metric. However, we believe that a quantitative assessment of the trade-off between complexity and accuracy is more useful and informative when a larger set of competing models is available. In this particular case, only two FE models are considered, representing two clearly distinct levels of fidelity. In this situation, we think that the choice of which FE model is preferable is less a matter of statistical selection and more a question of the intended level of detail and computational efficiency required for the specific application. We already stated some of this in the conclusion, lines 675f, which we revised with regard to your comment, including a discussion concerning the possibility of applying a model selection approach.

7) In the damage identification, the authors use the natural frequencies and mode shapes of 5 modes. It would be useful to have an idea of the extent to which these data allow for an accurate identification of the data. For example, at the end the authors indicate some differences between the identified stiffness reduction and physical fatigue crack. I am wondering here, should the authors change the parameters of the stiffness reduction to bring it closer to the physical crack, what would be the difference in natural frequency shift and mode shape? If those shifts are insignificant, the identification errors are simply a consequence of the limited information contained in the modal data. Along the same line, it would also be useful to add a figure with the 6 identified mode shapes in subsection 3.1.

Thank you for this remark. Regarding the first part, we are actually currently already looking into the addition of a sixth design variable being the off-diagonal value of the covariance matrix. This extension of the design variable parameterization would allow for a possible inclination of the introduced stiffness alteration and, as a result, the ability to represent the occurred inclined crack correctly. If this is successful, we want to publish this further extension of the design variable parameterization and can hopefully answer your question with more insight.

Regarding the second part of your remark, and also in connection with your comment No. 5, we revised Section 3.2 (cf. lines 382f) and included Figure 8, which contains subfigures of the first five eigenmodes utilized in the model-updating procedure(s). Each subfigure shows a comparison between the eigenmodes of the two FE models in their reference state and the eigenmodes identified from the measurement data in rotor blade state I. In addition, we added Table 5 to the same section, listing the MAC values corresponding to the eigenmodes visualized in Figure 8 and the associated eigenfrequencies.

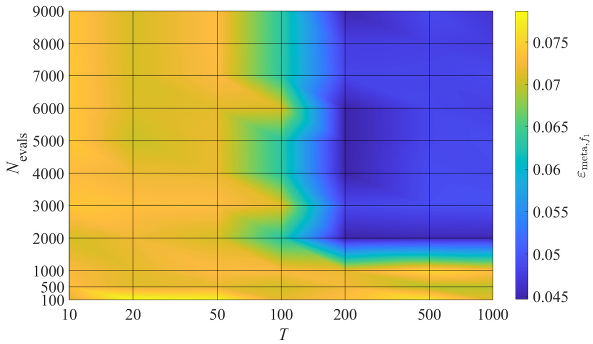
8) The authors parameterize the damage as a stiffness reduction, with a spatial distribution according to a (truncated) Gaussian distribution. This is fine, but at some points it seems that this spatial distribution is mixed up with a PDF of the damage which it is not. I would recommend not to use the term PDF when it is really a spatial distribution (for example when explaining equation (6)). Likewise, I find the following statement in subsection 4.1.2. misleading "As the one-dimensional damage distribution is formulated based on a Gaussian distribution function, $mu_L + sigma_L$ represents the range including 68% of the data". I think the statement is incorrect as mu_L and $sigma_L$ do not capture the variability in the crack location resulting from the statistical uncertainty in the data but describe the spatial extent of the stiffness reduction. Along a similar line, statements such as "The location . . . is identified at mu_L . . . , which overestimates the actual damage location" (also in subsection 4.1.2) are misleading. One can overestimate the intensity of damage or the extent of the damage zone but not the damage location.

Thank you for your detailed remark, we agree that our wording was misleading in some paragraphs. You are right, of course, the Gaussian distribution function is used only as the shape of the spatial distribution of the stiffness alteration and is not used to describe the probability of the design variables μ_L and σ_L. Accordingly, we revised parts of Section 2.1 to emphasize that we refer to a spatial distribution, cf. pages 5, 6 and 8; revisions are marked in red color. In addition, the description of the results in Section 4.1.2 was revised, cf. lines 486f (with regard to $\mu_L + \sigma_L$) and lines 519f (with regard to the wording 'overestimate the damage intensity'). Furthermore, some additional revisions with regard to your remark were made on pages 23, 24, 26, 27, 28 and 32, again marked in red color.

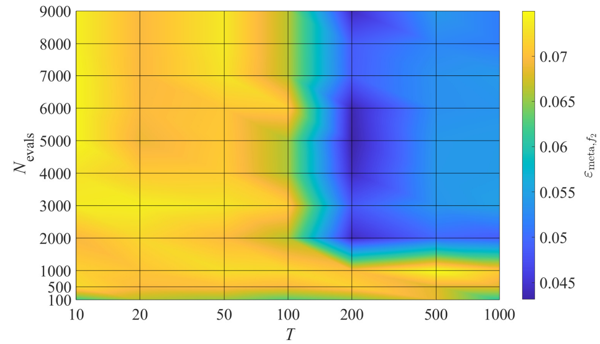
9) Can the authors give an indication of the accuracy of the meta-model relative to the range of the natural frequencies and mode shapes considered? For the meta-model to provide meaningful results for different sets of input data, the error should be small compared to the (small) range of modal data considered. A direct comparison of the error on individual natural frequencies and mode shapes would be more clear than the global error reported in figure 8.

Thank you for this comment. In response, we revised the error metric (Equation 18) in Section 4.1.1, lines 449f, and now evaluate the RMSE of the first five eigenfrequencies normalized by their respective value ranges. This allows for a direct assessment of the meta-model accuracy relative to the considered range of the individual eigenfrequencies. Accordingly, we also revised Figure 9 (page 21) and now present an example heat map of the relative error of the first eigenfrequency $\varepsilon_{\mathrm{meta},f_1}$. As the errors of the remaining eigenfrequencies are of similar magnitude, we did not show them individually in the manuscript. However, for your information, the results for all five eigenfrequencies are provided in the following Figure 1.
Regarding the accuracy of the meta-model for the eigenmodes, we evaluated the mean MAC value of the first five

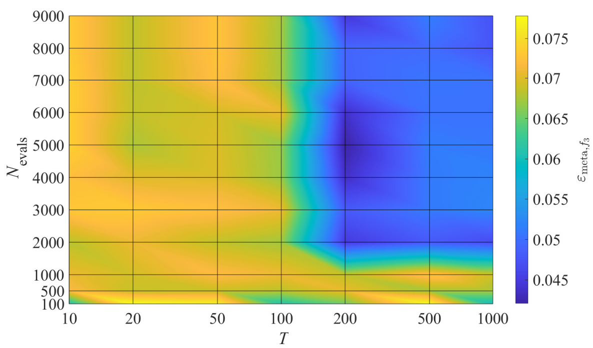
eigenmodes and the results showed that the variation of the hyperparameter settings for N_{evals} and T does not have a decisive influence on the meta-model accuracy. This is why, primarily, the evaluation of the eigenfrequency error metric was decisive for the choice of the hyperparameter settings of the considered optimization algorithm.



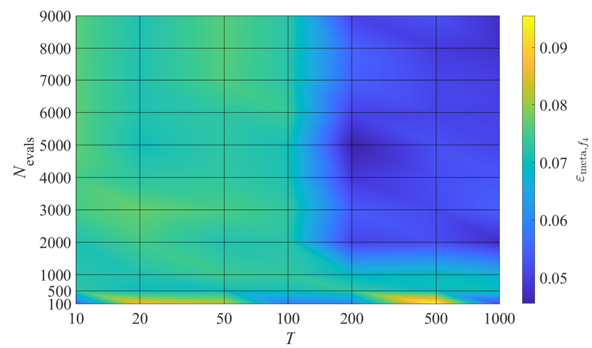
(a) $\varepsilon_{\text{meta},f_1}$



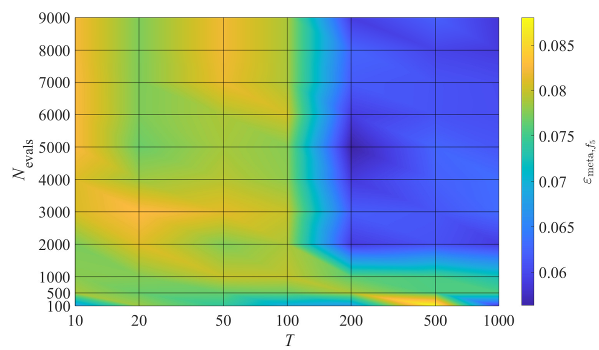
(b) $\varepsilon_{\text{meta},f_2}$



(c) $\varepsilon_{\text{meta},f_3}$



(d) $\varepsilon_{\text{meta},f_4}$



(e) $\varepsilon_{\text{meta},f_5}$

Figure 1: Evaluation of each meta-model set up using different combinations of N_{evals} and T. Surface colored according to the error $\varepsilon_{\text{meta},f_i}$.

10) It's a detail, but in section 4, the authors state "As demonstrated in Section 3.1, the measurement conditions were stationary, i.e. no significant variations in temperature, humidity, . . . ". I did not find any discussion of these parameters in subsection 3.1, however.

Thank you for this remark, we added a short description of the environmental conditions of the laboratory experiment in Section 3, lines 294f and 348f, and amended the sentence in Section 4, line 426.

11) It seems that the method adopted for the authors assumes that the stiffness reduction occurs in a single zone, so I was wondering what would happen if the procedure was applied for a case where it occurs in two distinct zones?

Thank you, you are correct, this is an underlying assumption and restriction of the presented damage distribution function, which we also mentioned in Section 2.1, lines 144f. Of course, this is an interesting question and we also think about and are currently working on extensions of the utilized damage distribution function to identify multiple damage locations. Finding two or more locations was successful to some degree but, so far, not consistently so. In this context, we struggle with the problematic need for a decision about the anticipated number of damage locations and the necessary higher number of design variables. If the number of damage locations is defined as part of the optimization problem, it becomes nonlinear and even more objective function evaluations would be required. Another approach, that we are currently looking into, is an approach where the amount of damage locations and their vicinity is anticipated by a specific sensitivity study, which determines the number of design variables and their upper and lower boundaries for the subsequent model-updating (i.e., optimization) procedure.
Nevertheless, in our view, the presented design variable parameterization so far still presents the best compromise between computational efficiency (i.e., few design variables) and meaningful output (i.e., the identification of one area where the most significant damage is present).

12) How did the authors consider the spatial variation of the stiffness in the mesh? Did they evaluate the stiffness distribution at the midpoint? How does $sigma_L$ compare to the element size?

Thank you for this detailed question. In the beam model, the smallest FE length is 5 mm and, in the shell model, the smallest FE length is 0.96 mm. As we wanted to apply the same lower (and upper) bounds for the design variables that are shared by the three different design variable configurations (i.e., μ_L, σ_L and D), we selected $x_{\text{lb}}(\sigma_L) = 0.001\,\text{m} = 1\,\text{mm}$ for both FE models (cf. Table 7, page 18). We added this information in Section 4, lines 402f.
Regarding the evaluation of the damage distribution function, it is indeed evaluated at the FE midpoint. For the beam model, this corresponds to the center of each beam element length in longitudinal direction and, for the shell model, this refers to the centroid of each shell element area. We added this information in Section 2.1, line 175 and lines 223f.

Reviewer #2

The paper "Damage identification on a large-scale wind turbine rotor blade using sample-based deterministic model updating" presents a very interesting, very important vibration-based damage identification study performed on a 31 m wind turbine blade under controlled fatigue loading. The authors apply a sample-based deterministic model-updating (SDMU) framework to both a beam and a high-fidelity shell model, using one-dimensional and two-dimensional Gaussian parameterisations of stiffness reduction to infer damage localization and severity. The work leverages a valuable full-scale dataset, and the systematic comparison of model fidelities and parameterizations offers useful insight for SHM practitioners.

In summary, the study is timely, relevant, and contains several strong elements, particularly the experimental dataset and the structured model-updating comparison. However, before reconsideration for full acceptance, the following issues should be addressed:

1) The Gaussian-shaped 1D and 2D damage distribution functions impose restrictive geometries on possible damage fields. There is no evidence that actual rotor-blade damage follows Gaussian distributions, nor is sensitivity to this assumption tested. This limitation seems to be indeed acknowledged, but deserves earlier and deeper discussion, ideally with a brief demonstration of its effect.

Thank you for this valid remark. A primary motivation for choosing a Gaussian distribution function as representation for suspected damage was to limit the number of design variables while keeping the optimization problem well-posed. For this purpose, a Gaussian-distributed parameterization is well-suited as it leads to a smooth damage field and, consequently, to a comparatively smooth objective function space. This, in turn, promotes faster and more stable convergence of the utilized optimization algorithm. Furthermore, we believe that a Gaussian-shaped damage representation can cover a wide range of damage scenarios, from highly localized defects affecting only few to one FEs to a smoothly distributed stiffness degradation over an extended region. In this context, we want to emphasize that the identified damage distribution function represents a spatial area with altered (in this case reduced) stiffness rather than an exact rendering of the physical defect geometry.
In an earlier publication, we compared the use of the presented Gaussian distribution function with the use of a continuous uniform damage distribution function for the damage identification via model updating of a laboratory steel cantilever beam with a reversible damage mechanism (cf. Wolniak et al. (2023)). The results showed no significant difference in the identification performance, which suggests that the results are not overly sensitive to the specific choice of the shape of the damage distribution function. Nevertheless, we acknowledge that this conclusion cannot be generalized to all structures and damage types and a systematic sensitivity study for the considered rotor blade remains an important topic for future work.
We explain the motivation for choosing the Gaussian-shaped damage representation in more detail in Section 2.1, lines 146f. In addition, we added a sentence in the introduction, lines 86f, and extended the outlook, lines 682f, with regard to this comment.

2) The approach cross-combines 53 $\times$ 53 BayOMA-derived modal data, implicitly assuming the independence and equal plausibility of all reference–analysis pairings. This statistical interpretation is not fully explained and may artificially inflate the uncertainty space. A clearer conceptual rationale, comparison with using full Bayesian posterior samples, or discussion of alternative sampling strategies would strengthen credibility.

Thank you for this useful remark. We tried to recognize this in the original manuscript in Section 4 by stating that the cross-combination strategy would no longer be sufficient for applications involving significant environmental and/or operational variations and an alternative sample provision approach would be required. However, we acknowledge that this was not sufficiently discussed.
The intent of using the cross-combination instead of a sampling method was actually to demonstrate that the proposed SDMU approach is applicable independently of prior assumptions about the distribution of the input data and, therefore, can be applied based on discrete input samples without the need for an explicit sampling procedure. However, you are of course correct that the cross-combination is based on the assumption that all BayOMA-derived modal parameters are statistically independent and equally probable samples. While this assumption enables a straightforward and computationally simple propagation of the variability, we acknowledge that it does not strictly correspond to the sampling from a Bayesian posterior distribution and may lead to a conservative estimation or even overestimation of the resulting variance. To further investigate this, we compared the mean of all $N_{\mathrm{sets}} = 53$ standard deviations per eigenfrequency, $\overline{\mathrm{std}(\boldsymbol{f}_{\mathrm{M},i})}$, with the standard deviation calculated from the $N_{\mathrm{sets}} = 53$ mean values per eigenfrequency, $\mathrm{std}(\overline{\boldsymbol{f}}_{\mathrm{M},i})$. Both calculation were made using the modal parameters (i.e., the eigenfrequency mean values and the corresponding standard deviations) identified using BayOMA in rotor blade state I. The results showed that there is only a marginal difference between the two values for the considered laboratory experiment. For example, for the first eigenfrequency, it was calculated that $\overline{\mathrm{std}(\boldsymbol{f}_{\mathrm{M},1})} = 0.0011$ and $\mathrm{std}(\overline{\boldsymbol{f}}_{\mathrm{M},i}) = 0.0012$.
To address this comment, we added a more detailed explanation of the motivation behind the use of cross-combination and a discussion of alternative sampling strategies in Section 4, lines 418f. Furthermore, we acknowledge that a comparison between the presented cross-combination results and results obtained using a sampling method poses an interesting outlook in lines 685f.

3) Localization performance can be checked against the crack position, but the inferred stiffness reduction magnitudes ($D_{\mathbf{1D}}$, $D_{\mathbf{2D}}$) seem not to be directly validated.

Thank you, this is a correct observation and we acknowledge that we did not discuss the interpretation or validation of the damage intensity D in detail. The design variable D does not actually have a direct physical interpretation in the same sense as the design variables μ (i.e., damage position) and σ (i.e., damage extent). Instead, it represents an alteration (in this case reduction) of the material stiffness within the utilized numerical model and can be understood as a phenomenological parameter rather than a direct measure of physical crack severity.
In general, it is very difficult to figure out the actual magnitude of a damage, especially in (fiber-reinforced) composite material. In the rotor blade fatigue test, the crack, that finally occurred at the leading edge, can be visually identified. However, even in this case, a quantification of the severity of this crack would require knowledge of geometric and mechanical characteristics such as the crack depth, crack tip opening displacement, and the specific damage mechanism (fiber breakage, matrix cracking, delamination, etc.). These parameters are typically not directly accessible and difficult to measure in a non-destructive manner. Moreover, the visible crack is likely not the only form of damage present in the blade. It is very plausible that progressive material stiffness degradation

occurred along the leading and trailing edges due to fatigue damage (cf. Figure 3) prior to the formation of the visible crack.
In summary, damage severity is in general not uniquely quantifiable, especially in composite rotor blades, and the design variable D, that we interpret as the "damage intensity", does not correspond to a distinct physical quantity that could be directly measured or validated. For this reason, an explicit validation of the value obtained for the design variable D is not really feasible for the rotor blade fatigue test considered in this work. We added a short discussion addressing this limitation in Section 4.1.2, lines 513f, and in Section 4.3, lines 650f.

4) Some modes included in the updating (e.g., flapwise modes) seem to exhibit minimal damage sensitivity. Using low-sensitivity modes in multi-objective optimisation may dilute the updating signal. A brief sensitivity analysis or justification for including all five modes would be appropriate.

Thank you for this comment. In the present case, you are correct that the flapwise eigenmodes are likely less sensitive to the observed damage as the fatigue load was applied in the edgewise direction and, consequently, the damage developed mainly in this direction. In addition, we are aware that performing a sensitivity analysis prior to the model-updating procedure can be beneficial and valuable especially when a large number of input parameters are available.
For the model-updating procedure(s) presented in this work, we included the first five eigenmodes and corresponding eigenfrequencies that could be reliably identified from the measurement data and matched both numerical models with good accuracy, which are fundamental prerequisites for a successful model-updating procedure. Given that "only" five eigenmodes and corresponding eigenfrequencies were used, we believe that a sensitivity study is not strictly necessary. Moreover, a separate sensitivity study would in principle be required for each combination of rotor blade states, resulting in six sensitivity studies (cf. Table 6). As these combinations include the three self-combinations, in which no damage is present at all, the sensitivity studies could suggest the utilization of different optimal mode sets for different combinations. However, for consistency and comparability, we prefer to utilize the same input (i.e., modal parameter set) for all model-updating procedures conducted in this work.
Nevertheless, you address an important point and we agree that investigating the influence of mode selection and damage sensitivity on the model-updating results would be valuable. We added a more detailed motivation for the inclusion of all five eigenfrequencies and corresponding eigenmodes in the objective function in Section 4, lines 407f.

5) Computational performance should be reported fully. Only modal analysis times seems to be explicitly reported; Readers may need the full cost of a complete SDMU run, especially for the shell model, to judge applicability in practice.

Thank you for this comment. The total computational time of the complete SDMU run(s) is highly dependent on the use of parallelization and computer or cluster (i.e., hardware) performance, which is why there is no definite computational time that we can specify. This is why, so far, we gave the time for one modal analysis using the respective FE models (cf. Table 4, page 16) and the number of FE evaluations necessary for the first deterministic model-updating run for the determination of the meta-model (cf. Table 8, page 21). Depending on the individual possibilities, readers can calculate the necessary computing time for this deterministic run with the given information.
The setup of the meta-model itself is, in our view, negligible, as it only takes a few seconds and, also, only has to be performed once per SDMU procedure, meaning once per rotor blade state combination. In this context, the computational time for the $N_{\text{sets}} \times N_{\text{sets}}$ model-updating procedures conducted using the meta-model is more

relevant but, again, highly dependent on parallelization possibilities and hardware performance. In addition, this part of the SDMU procedure is no longer dependent on the utilized FE model (as it is replaced by the respective meta-model) but dependent on the number of objective function evaluations. To give an indication, one optimization run with 5000 objective function evaluations using the meta-model set up for the beam model takes approximately 160 seconds if run on my laptop.
Due to the reasons mentioned above, we cannot report a definite computational time for the SDMU procedure with regard to your remark. However, we revised Table 8 (page 21) by listing the model evaluations in general, not only the FE model evaluations. Also, we added a more detailed description of the table in Section 4.1.1, lines 456f.

6) The state-of-the-art review is reliable but relatively limited. Please consider expanding it to include works such as https://doi.org/10.3390/s22041627.

Thank you for this suggestion, we included the work by Civera and Surace (2022) and additional, related work in the introduction, lines 21, 24f and 29f. The added literature items are also marked in blue color in the references.

7) The limitations section should explicitly mention the lack of environmental variability in the dataset.

Thank you, we added a sentence to the description of the experiment to further emphasize this limitation in Section 3, lines 297f (marked in red color due to a similar comment of reviewer #1). In addition, the lack of environmental variability is also mentioned in Section 3.1, lines 348f (marked in red color), in Section 4, line 427 (marked in red color), and was already stated in the conclusion, lines 687f.

8) It is not totally clear, in the text, if the Authors applied an already-available third-party version of BAYOMA, or if they implemented their own variant based on Siu-Kui Au's textbook Operational Modal Analysis Modeling, Bayesian Inference, Uncertainty Laws. This aspect should be clarified.

Thank you for this remark. We did not use an available third-party version of the BayOMA but implemented the BayOMA ourselves in Matlab. Our implementation strictly follows the methodology described in the mentioned references by Au (2013, 2017). We clarified this is Section 3.1, lines 332f.

9) Related to the above, for reproducibility, either a link to the used version (if available online) or a detailed pipeline (otherwise) should be provided.

As mentioned above, the implementation of the BayOMA strictly follows the methodology described in detail in Au (2013, 2017). The Matlab code was developed and continuously extended internally at our institute and is currently still under active development. For this reason, it has not yet been made publicly available.
To ensure correctness and reproducibility of the implementation, the developed code has been thoroughly tested and benchmarked against established operational modal analysis techniques, in particular against the covariance-driven stochastic subspace identification (SSI-COV), cf. Jonscher et al. (2023). A finalized version of the framework is intended to be published in the future. This is clarified in Section 3.1, lines 332f.

10) Please carefully revise the text for typos and mistakes. For instance, "Knebusch et al. (2020)" should be ("Knebusch et al., 2020)"

Thank you for this comment. We carefully reviewed the manuscript and corrected mistakes. Regarding the citation format, this is predefined by the WES journal style guidelines and, therefore, cannot be modified by the authors. In this case, we included the authors name in the sentence (cf. lines 56f).