Reply to reviewers (manuscript wes-2026-56)

The authors would like to thank the reviewers for their time and constructive feedback. All the comments have been taken into consideration and have contributed to improving the manuscript.

An annotated version of the revised manuscript (including all changes based on both reviewers' comments) has also been uploaded.

A list of point-by-point replies to the comments follows (reviewer comments in black, authors' response in blue).

Replies to RC1

Comments:

1. The manuscript is quite long. I suggest shortening a few parts, such as the detailed description and comparison of the machine learning models. Othwerwise, some parts could be moved to the Appendix, such as Fig.5 and Table 4.

We agree that the original manuscript was overly long in several places and have undertaken a manuscript-wide revision to improve conciseness. The main manuscript has been reduced from approximately 14,660 to 12,970 words. The main reductions include a more concise description of the flow-model calibration, substantial shortening of the surrogate-model section, consolidation of the operational-logic and aggregation methodology, removal of repeated explanations in the case-study description, and considerable trimming of the results discussion, particularly Section 3.2. Detailed flow-calibration information has been moved to Appendix A, while the neural-network implementation, selected architectures, validation figure, and complete surrogate-performance metrics have been moved to Appendix B.

The manuscript remains relatively extensive because the study integrates several disciplines and requires sufficient explanation of the ecological and regulatory context, field-data processing, wind-farm flow modeling, aeroelastic representation of normal and transient operating states, surrogate modeling, operational logic, and long-term energy and fatigue aggregation. We have therefore sought to balance conciseness with clarity and reproducibility.

2. The tuning procedure described in Section 2.1 is not very clear. Figure 2 facilitates comprehension, but it could be better supported by the text. For instance, mention the division into the two types of coefficients more clearly, the cost function, and the total number of coefficients that are tuned. If the length increases significantly, the section could be moved to the Appendix.

We have revised Section 2.1 to distinguish more clearly between the two groups of calibrated coefficients: the sector-dependent flow parameters, which represent terrain-induced speed-up factors, and the global wake-model parameters, which define the wake formulation. A more detailed description of the calibration procedure has been moved to Appendix A to maintain the focus and conciseness of the main manuscript. The appendix now states that 144 flow parameters and 10 global wake parameters, 154 parameters in total, are calibrated simultaneously using one year of operational wind-farm data. It also explains that the calibration minimizes a cost function based on the residuals between simulated and measured turbine power and uses singular-value decomposition to remove non-identifiable parameter combinations. The calibrated wake-parameter values and their association with the respective wake-model components have also been moved to Appendix A. Section 2.1 retains a concise overview of the procedure and refers readers to Appendix A and Braunbehrens et al. (2023), where the calibration methodology is described in detail.

3. In Section 4.7 (line 496), it is stated that bat activity occurs more frequently at lower wind speeds. However, I am not sure that this information can be easily obtained from Figure 9 (left) since the distribution is very similar to the typical Weibull distribution of wind speed occurrence. In other words, if the distribution of bat activity were the same as that of the wind speed occurrence, it would imply that bat activity does not depend on the wind speed. I suggest plotting these data without the bias of the wind speed occurrence to better convey your message.

We agree that the occurrence frequency of the environmental conditions must be considered when interpreting the distribution of recorded bat activity. The purpose of Figure 9 is not to estimate the conditional probability of bat activity at each wind speed, but to show the environmental regimes under which the observed contacts occurred, as these distributions form the basis for defining the thresholds used in static bat protection schedules.

We note that the bat-activity distribution in Figure 9 does not simply reproduce the ambient wind-speed distribution shown in Figure 8. The ambient distribution has peaks at higher wind speeds (around 5-7 m/s), whereas the recorded bat activity is concentrated more strongly in the lower wind-speed range (2-4 m/s) and decreases substantially toward approximately 8 m/s. This indicates a site-specific association with lower wind speeds, consistent with findings reported in the literature. Nevertheless, we agree that the original statement that activity was "strongly correlated" with wind speed and temperature was too strong. We have revised the text to describe the observed distributions more precisely and to clarify that no causal or generally transferable ecological relationship is inferred from this dataset.

4. Previous studies highlight the need to consider multiple sectors instead of only one rotor-averaged inflow estimate to predict the structural loads using surrogate models. This is mentioned in Section 4. However, the impact of this assumption could be tested with tools available in the literature, e.g. surrogate models that take sector-averaged inflow as input instead of rotor-averaged inflow. I recognize that this may be beyond the scope of the study but it would strengthen the results of this work.

We agree that sector-averaged or spatially resolved rotor inflow descriptors can improve structural-load predictions under wake/partial-wake conditions. We have clarified this limitation in Section 2.3 (l 323-333) and in the limitations discussion (l 695-700). The surrogate models used in the present study were trained on aeroelastic simulations with free-stream turbulent inflow and use rotor-averaged effective wind speed and turbulence intensity as inputs. Wake interactions are accounted for at the wind-farm level through the calibrated engineering wake model, which provides turbine-specific effective wind speed and turbulence intensity, but the detailed spatial structure of partial wakes is not represented in the aeroelastic inputs.

Testing the proposed sector-resolved formulation would require generating a new aeroelastic simulation database using spatially non-uniform, wake-affected inflow fields and retraining and validating the mode- and channel-specific surrogate models. It cannot be assessed by simply replacing the existing rotor-averaged inputs with sector-averaged values. We therefore consider such an analysis beyond the scope of the present study, as also acknowledged by the reviewer.

For the present application, the principal differences among the evaluated strategies arise from the frequency and duration of normal operation, idling, start-up, and shutdown states, particularly at the relatively low wind speeds relevant to bat protection curtailment. Moreover, the same inflow representation is applied consistently to all strategies, supporting their relative comparison within the adopted modeling framework. Nevertheless, we agree that sector-resolved inflow representations would strengthen the prediction of absolute fatigue loading and we consider it as an important direction for future work.

Minor comments:

1. In the Introduction (lines 45-50), the statement on the loads of the transients refers to floating wind farms. I suggest finding a reference for onshore farms or justifying why this also applies here.

We agree that the original ordering placed too much emphasis on evidence from floating wind turbines without first establishing its relevance to the present onshore application. We have therefore restructured the paragraph to begin with Ziegler et al. (2024), which reports measured start-up and shutdown loads from both an onshore and an offshore turbine. The offshore and floating studies are then retained as complementary evidence that transition-induced loading is relevant across different turbine and support-structure technologies, while clarifying that its magnitude and the most affected components are system-specific.

2. In Figure 2, the orange dots on the flow field are difficult to see. I suggest changing the color.

The figure has been updated as suggested.

3. In Figure 6, is it possible to add the DELs obtained from the aeroelastic solver? Keep it as possible suggestion.

We agree that showing the corresponding aeroelastic simulation results could provide a direct visual reference. However, Figure 6 is intended primarily to illustrate the mode- and inflow-dependent load trends represented by the selected surrogate models, rather than to serve as an additional validation figure. Each subplot already contains eight curves, corresponding to four operational modes and two TI levels, and overlaying the aeroelastic simulation points for all cases would substantially increase the visual density and reduce readability. The surrogate accuracy is assessed separately against independent aeroelastic test simulations in Figure 5 (moved to appendix in the revised manuscript) and is further quantified for all operational modes and load channels in Appendix B. We therefore retained Figure 6 without the additional simulation markers.

4. I recommend using the word ‘blanket’ also in the description of the methodology (Section 2.8) for consistency with the rest of the manuscript.

We have revised Section 2.8 to introduce the evaluated strategies as static (“blanket”) and dynamic (“event-triggered”) bat protection strategies. The term “blanket” has also been included when the three static cases are first described in detail. We retain “static” as the primary classification in the table and subsequent analysis for consistency with the case names and the terminology used throughout the manuscript.

5. In Figure 12, I suggest including the acronym (e.g. TBTOR) for each sublot to be consistent with the text, for instance, below each subtitle.

We have added the acronyms as suggested by the reviewer

6. In Figure 13, the plot on the right is difficult to read. The x-axis labels are a bit messy and the bar widths are very small. I suggest increasing the bin size of the TI.

We have increased the turbulence intensity bin width, thereby reducing the number of bars and improving readability. We have also enlarged the figure and increased the axis, tick-label, and legend font sizes.

7. For consistency with the other metrics, I suggest presenting the results of Table 6 as a histogram and including the information related to the inter-annual variability in Table 7 instead.

Thank you for this suggestion. We considered presenting the energy-production results as a histogram; however, we have retained the tabular format because the numerical differences between several strategies are small, often at the decimal level, and these differences are relevant when assessing energy-production impacts. On a common histogram scale spanning approximately −12% to 0%, mainly due to the larger loss for Static100, the remaining differences would be visually compressed and difficult to distinguish. We have also kept the year-specific and total values together in the same table, as this allows the interannual variation and the overall result to be compared directly without duplicating the same energy information in the following summary table.

Replies to RC2

Major comments:

1. The manuscript is very long. Part of this can be explained by the fact that the manuscript covers many different areas (describing the specific case study, the engineering wind farm model, the turbine load surrogate model, the bat protection strategy, and the effects in terms of both power and loads). However, the same information could in my opinion still be provided in a significantly shorter manuscript. The authors use a very verbose writing style, and many different paragraphs provide little to no new information with respect to provided figures and/or previous paragraphs. In the minor comments below I will make several suggestions for shortening the manuscript, but generally using a more concise writing style would be my number 1 recommendation.

We agree that the original manuscript was overly long in several places and have undertaken a manuscript-wide revision to improve conciseness. The main manuscript has been reduced from approximately 14,660 to 12,970 words. The main reductions include a more concise description of the flow-model calibration, substantial shortening of the surrogate-model section, consolidation of the operational-logic and aggregation methodology, removal of repeated explanations in the case-study description, and considerable trimming of the results discussion, particularly Section 3.2. Detailed flow-calibration information has been moved to Appendix A, while the neural-network implementation, selected architectures, validation figure, and complete surrogate-performance metrics have been moved to Appendix B.

The manuscript remains relatively extensive because the study integrates several disciplines and requires sufficient explanation of the ecological and regulatory context, field-data processing, wind-farm flow modeling, aeroelastic representation of normal and transient operating states, surrogate modeling, operational logic, and long-term energy and fatigue aggregation. We have therefore sought to balance conciseness with clarity and reproducibility.

2. At the end of section 2, on line 542, the authors say: “the dynamic cases are assumed to correspond to a 100% bat protection level”. I would argue that this assumption is in no way realistic in this case study. The wind farm uses only one bat sensor on the very edge of the farm, so it is highly unlikely that every bat activity within the entire wind farm is captured by the sensor. From Figs 9 and 10, it follows that bats have relatively predictable behavior, which is obviously why static protection algorithms are usually implemented. As a result, bat activity resulting in false-negative measurements by the bat sensor are likely still captured by these static algorithms, providing additional protection to bats that the dynamic algorithm does not. This effect is not mentioned anywhere in the manuscript, not here and only partly in Section 4. I understand that it is hard (perhaps impossible) to quantify this effect using the available dataset, but it should at the very least be mentioned more explicitly as a limitation of the dynamic protection algorithms.

Thank you for highlighting this important assumption. The original manuscript already stated in lines 531–542 that each recorded detection generates a farm-wide shutdown request and that the single monitoring signal is assumed to be representative of the wind farm while false-positive and false-negative detections are neglected, hence the 100% assumption. The limitations discussion in lines 730–744 also addressed sensing coverage, the possible need for multiple sensors and redundancy, and the effects of false-positive and false-negative detections.

We have now clarified the assumption, in more detail, where the dynamic cases are introduced. The nominal 100% protection level means that, under the adopted assumption of farm-wide representativeness and neglecting possible false negatives, the curtailment logic responds to 100% of the detections recorded by the monitoring system. It does not represent an independently verified measure of realized farm-wide ecological protection. We have also expanded the limitations discussion to acknowledge the reviewer’s point that the broader environmental and temporal coverage of static schedules may provide protection during some events missed by dynamic detection-based approaches. We identify direct comparisons between static and dynamic schedules under imperfect detection as an important topic for future work. This effect cannot be quantified with the available data because no independent farm-wide activity or mortality observations are available.

The figures in the manuscript do not match the quality of the manuscript itself. Figures use acronyms not explained anywhere in the surrounding text, some of the captions/textboxes in the figures are too small to read, and some of the data are too small to reasonably interpret. I suggest cleaning this up in the final manuscript. See minor comments below for some suggestions.

We have revised the figures throughout the manuscript to improve readability, terminology, labeling, and caption clarity. Specific changes are detailed in the responses to the related comments below.

3. Line 331-333: “ Instead, wake … of wakes”. Is it possible to quantify how well this method captures loads induced by wakes? If you have wind direction and load SCADA signals, you should also be able to see whether a turbine is waked. So, you could make a figure such as Figure 5 comparing waked and unwaked turbines. Such a figure would in my opinion be more valuable than the current Figure 5, as it is well-known that waked turbines see substantially increased loads and this situation therefore contributes significantly to the overall fatigue damage. If the

error of your model in waked situations is much higher than in unwaked situations, as I would expect, that is another unmentioned limitation of this study.

We agree that wake-induced loading, particularly under partial-wake conditions, can contribute significantly to fatigue accumulation. However, the proposed comparison between measured waked and freestream load responses cannot be performed with the available dataset, because the wind farm data do not include structural-load measurements. Furthermore, the OEM-specific aeroelastic model was unavailable, and the structural-response database was therefore generated using the IEA 3.4 MW reference turbine.

Wake effects are represented in the numerical framework through turbine-specific effective wind speed and turbulence intensity obtained from the calibrated engineering wake model. However, as already discussed in lines 322–338 of the original manuscript, the aeroelastic database uses free-stream turbulent inflow and rotor-averaged descriptors and therefore does not explicitly represent the spatial structure of partial wakes. The manuscript also cites studies using sector-averaged and spatially resolved inflow descriptions and identifies these approaches as a possible future extension.

The limitations section in lines 692–701 further acknowledges that this assumption may affect fatigue accumulation and clarifies that the quantitative results should primarily be interpreted as relative comparisons between strategies within the same modeling chain. In the present application, the main mechanism differentiating the strategies is the frequency and duration of curtailment-induced idling, start-up, and shutdown states at low wind speeds. Since the same wake and load-modeling assumptions are applied consistently to all cases, the framework remains suitable for investigating these relative effects.

4. Section 3: What I am missing in this section is an analysis of how well the different strategies really save bats. For example, how much of the measured bat activity in year 2 is missed by implementing the Static90, Static99, and Static100 strategies? Do we still achieve the promised percentage of bat protection? If not, how can this be improved (by using rolling average temperatures perhaps?). What if it were reversed and year 2 would have been the benchmark year?

This is a very interesting point. We agree that the transferability of a static curtailment schedule from one monitoring year to another is not guaranteed. Based on the practical experience of the industry coauthors, post-commissioning bat monitoring in France is commonly conducted over multiple activity seasons, and the curtailment schedule can be reviewed and adjusted when subsequent monitoring indicates materially different activity patterns. Quantifying year-specific protection efficacy, testing alternative benchmark years, and developing adaptive ecological thresholds would require a dedicated ecological analysis. However, this kind of discussion goes beyond the scope of this present paper, which focuses only on the effect of those bat protection strategies on energy production and structural loads, but it could definitely be the subject of a future study.

Minor comments:

Line 35: "In practice … wind farm". A source would be appreciated to support this statement

The statement reflects the practical experience of the industrial coauthors operating wind farms in several countries. In practice, the overarching regulatory framework is implemented through site-specific permitting and monitoring requirements defined on a case-by-case basis by the relevant regional or local authorities. The detailed project-level technical prescriptions are generally not consolidated in publicly accessible documents and therefore cannot be supported by a single general reference.

We have clarified the manuscript wording to make the intended scope explicit. The purpose of the statement is not to characterize the regulatory process itself, but to highlight that current assessments primarily focus on ecological protection indicators and approximate AEP losses, while structural loading, lifetime effects, and longer-term economic implications are generally not evaluated explicitly.

Figure 1: FLORIS, DELs, KPI are all undefined at this point in the text. The graphs are too small to make out and add no value.

We have revised Figure 1 to improve readability and make the conceptual workflow self-contained. The undefined acronyms have been replaced with descriptive terminology, including "engineering flow models," "fatigue-load metrics," and "energy, fatigue, and operational indicators." The operating states are now stated explicitly, and the text, spacing, and overall layout have been enlarged and simplified. The embedded graphics are intended only as schematic illustrations of the workflow; the specific flow model and fatigue metrics are introduced formally in the following subsections.

Figure 2: The cost function formulas add absolutely no value/clarity if they are not elaborated on/explained in the paper. I suggest removing the formulas from the figure.

We have retained the compact cost-function notation in Figure 2 because it identifies the objective used in the calibration workflow without introducing the full mathematical formulation. To support its interpretation, Appendix A now explains at a high level that the calibration minimizes the residuals between simulated and measured turbine power. The complete formulation and calibration methodology are provided in the referenced dedicated publication (Braunbehrens et al. (2023)).

Table 1: All the parameters listed in this table are not defined anywhere in the paper and therefore do not support the text in any way. If the authors insist on providing these values, I suggest moving them to the appendix (possibly together with the section describing the FLORIS tuning, as this is not the main contribution of the paper).

We agree that the detailed calibrated parameter values are not central to the main contribution of the paper. The table and the corresponding calibration details have therefore been moved to Appendix A. The table has also been reorganized to identify the wake-model component associated with each parameter, while the accompanying text explains the overall calibration structure. Detailed definitions and derivations of the individual parameters remain available in the referenced publication (Braunbehrens et al. (2023)).

Line 203: The case study wind farm and its turbines are mentioned here but are not defined yet. This caused confusion for me as I was reading it. I suggest moving section 2.6 up to before the current section 2.2.

Thank you for pointing out this potential source of confusion. The comment refers to the comparison between the IEA 3.4 MW reference turbine and the case-study turbines before the wind farm is formally described in Section 2.6. We have added a forward reference to Section 2.6 at the first mention of the case-study turbines. We have retained the current section order because Sections 2.1–2.5 present the modeling and evaluation framework, whereas Sections 2.6 and 2.7 jointly introduce the wind-farm, environmental, and bat-monitoring data used for the case study. Moving Section 2.6 before Section 2.2 would separate these closely related case-study inputs.

Table 2: Seed-avg points are not defined or mentioned anywhere else. I'm therefore not sure I understand what you mean by this.

We have revised both the paragraph and the table to clarify that six independent turbulence realizations are simulated for each wind-speed–TI combination and operational mode to account for seed-to-seed variability, and that the resulting load metrics are averaged to form one seed-averaged training point. The table headings have accordingly been changed to "Turbulence seeds" and "Seed-averaged training points."

Section 2.3: This section is VERY long. It describes and compares two different surrogate models, while this is as far as I see it not one of the main contribution of this paper, as both methods have been used and described in literature before. Given the length of the paper, I suggest considering to remove or to significantly trim this section.

We agree that the original Section 2.3 contained more detail than warranted by its role in the paper, since the comparison of spline and neural-network surrogates is not itself a main contribution. We have therefore substantially shortened the section. The main text now retains only the surrogate inputs and outputs, the mode- and channel-specific formulation, a concise description of the two benchmarked approaches, the independent-test-set evaluation, the rationale for selecting the spline models, and the main limitation associated with rotor-averaged inflow descriptors. The detailed neural-network implementation and hyperparameter search, the selected architectures, the surrogate-error distributions, and the complete performance metrics have been moved to Appendix B.

Line 262: "Li et al. (2018)" should be "(Li et al, 2018)".

Corrected

Line 270: How is the model accuracy measured? Do you use higher frequency SCADA data? This might be mentioned earlier in the text but it's not completely clear to me.

Thank you for pointing out this ambiguity. The accuracy discussed here refers specifically to the regression accuracy of the surrogate models, not to validation of the complete structural-load model against field measurements. The surrogate predictions are evaluated against DELs obtained from independent aeroelastic simulations at inflow conditions not used for training. The high-frequency SCADA data are used only to inform the simulated transition behavior and are not used in this accuracy assessment, since the available site data do not include structural-load measurements, and we use a reference wind turbine model for the simulations. We have revised the text to make this distinction explicit.

Figure 6: This figure looks very sloppy. The font size is too small, the use of "20k" is not very scientific, the legend uses inconsistent naming ("start", "shut"), and it is unclear to me why these specific channels are plotted while the third dimension channels and the tower-top channels are omitted.

We agree with the reviewer about the quality of figure 6. The original figure presented a subset of representative tower, blade-root, and drivetrain channels to limit the size and visual density of the figure. However, we agree that the rationale for this selection was not sufficiently clear and that omitting the torsional and tower-top responses was inconsistent with the accompanying discussion, which refers to several of these channels.

We have therefore expanded Figure 6 to include all 12 load channels considered in the study. The panels are now arranged in a 4x3 layout and grouped by component: tower bottom, tower top, blade root, and main shaft. We have also increased the figure and font sizes, replaced the abbreviated legend entries “start” and “shut” with “Start-up” and “Shutdown,” and revised the moment units to avoid informal axis notation such as “20k”. The surrounding text has also been adjusted accordingly.

Line 339: “response”, not “resposnse”

Typo corrected.

Section 2.4: This section is again quite long, while in my opinion providing limited new information with respect to literature. Consider trimming to only describe the most relevant findings.

We agree that the manuscript should avoid unnecessarily extensive discussion of established results. However, we have retained Section 2.4 because it presents a central methodological and contribution of the study rather than a general review of known turbine-load behavior.

To the best of our knowledge, the present work is the first to provide a systematic, mode-resolved aero-servo-elastic characterization of normal operation, idling, start-up, and shutdown across the wind-speed and turbulence-intensity ranges relevant to operational curtailment. A particular novelty is the proposed treatment of transient events, in which field-informed controller behavior is reproduced within the aeroelastic model to represent realistic start-up and shutdown maneuvers.

Retaining this section is also important because it explains how the different structural-load channels respond to idling and transient operation. These responses are strongly channel-dependent: some components experience substantial load relief during non-producing operation, whereas others can receive significant fatigue contributions from repeated start-up and shutdown maneuvers. This physical interpretation is necessary for understanding the cumulative fatigue results presented later and cannot be inferred from transition counts alone.

We therefore consider Section 2.4 useful beyond the present case study, as it provides both a transferable simulation approach for non-normal operating states and new insight into the load-channel-specific consequences of repeated operational transitions. For these reasons, we have retained the section in its current form.

Section 2.5: See above. This section could potentially be omitted completely or merged into one of the previous sections in my opinion.

We agree that the original section could be presented more concisely. However, we have retained it as a separate section because it defines how the curtailment requests are translated into chronological turbine-state sequences and how the resulting mode-specific power and fatigue responses are aggregated. These steps connect the instantaneous responses of the individual flow and load models to the long-term energy and fatigue results and are necessary for reproducibility.

We have substantially shortened the section by consolidating the descriptions of the static and dynamic control logic, removing repeated explanations, and presenting the state assignment, mode-specific response calculation, and metric aggregation more directly. We retained the state-machine treatment, the implications of the 10-minute discretization, the mode-dependent inflow and power assumptions, and the fatigue-damage formulation because these directly determine the reported results. Merging this material into the preceding sections would obscure the distinction between characterizing the individual operational modes and applying them sequentially over the evaluation period.

Line 466-469: “Missing timestamps … temporal coverage”. How reliable is this method? Would it not be better to simply omit the data that is unavailable, especially if it is only 0.4% of the dataset?

ERA5 data are used only to complete the 0.4% of timestamps for which the derived ambient SCADA time series is unavailable; it is not used as the primary source of the site conditions. We have retained these intervals to preserve a continuous and identical evaluation horizon for the baseline and all bat-protection strategies. Simply omitting them could exclude protection events occurring during those periods and could bias the normalized energy and fatigue changes by altering the cumulative quantities.

Line 498-499: “The seasonal …during May”. I assume this can be explained (at least partly) by the daily mean temperatures, i.e., year 1 having a warmer spring than year 2? Did you investigate that, and consider using that as a potential indicator of when to initiate any of the bat protection protocols?

A warmer spring could indeed be one possible explanation for the earlier activity observed in year 1, but establishing the ecological drivers of this difference is outside the scope and expertise of the present study. Bat activity can depend on several interacting, species-specific and site-specific factors, and we do not consider it appropriate to infer a causal explanation from the available two-year dataset.

Similarly, the selection of seasonal start dates or adaptive environmental triggers for bat-protection protocols is, in practice, undertaken by qualified bat ecologists and environmental specialists, based on the species present, site-specific monitoring, and a broader assessment of relevant environmental conditions. We have therefore avoided attributing the observed difference to temperature or recommending temperature as an additional initiation criterion.

Figure 9: This figure would be more useful in my opinion (at least the lefthand part) if it were shown relative to occurrence of each wind speed bin. Now it is hard to compare this figure with the lefthand side of figure 8 to truly assess when bats are most active. Furthermore, there’s a space missing between “Right:” and “bat”.

We agree that the occurrence frequency of the environmental conditions must be considered when interpreting the distribution of recorded bat activity. However, Figure 9 is not intended to estimate the conditional probability of bat activity within each wind-speed bin. Its purpose is to show the environmental regimes under which the observed contacts occurred, since the cumulative distribution of these contacts forms the basis for defining the thresholds used in static bat protection schedules.

We have therefore retained the activity distribution in its original form but clarified its interpretation in the manuscript. The distribution in Figure 9 does not simply reproduce the ambient wind-speed distribution shown in Figure 8. The ambient wind-speed occurrence peaks at approximately 5–7 m/s, whereas the recorded bat activity is concentrated mainly at approximately 2–4 m/s and decreases substantially toward 8 m/s. The activity distribution is therefore shifted toward lower wind speeds relative to the prevailing site conditions, indicating a site-specific association consistent with findings reported in the literature.

Nevertheless, we agree that the original wording describing bat activity as "strongly correlated" with wind speed and temperature was too strong. We have revised the text to describe the observed distributions more precisely and to clarify that the analysis is descriptive and does not infer a causal or generally transferable ecological relationship. The missing space has also been corrected.

Line 500-510: this whole section seems repetitive w.r.t. earlier parts of the manuscript.

We agree that this paragraph repeated information already provided in the íntroduction and in the subsequent case-study definition. We have therefore removed it.

Figure 10: I suggest plotting sunrise/sunset as a line instead of as dots, as it develops continuously throughout the year. I would also move the legend box, so it doesn't unnecessarily overlap with part of the plotted points, and cut off the figure at 0 and 24, as there is no 25th hour in a day. Similarly, I would suggest using 6-hour intervals for the y-label markers, which makes much more sense for time of the day.

The figure is now updated following all the suggestions, except for the strict cut for 0h and 24h on the y-axis, as the dots indicating the observations get larger with the number of activity, hence can extend further than 0 and 24 in the night.

Line 555: Why do you use a hashtag (year #1) here and throughout the manuscript instead of just "year 1" or even "2022" (or whatever year it was).

We have removed the symbol throughout the manuscript keeping the naming convention as year 1 and year 2.

Figure 11: I would be very interested to learn how the baseline downtime and start/stop events compare to the actual measured values. Please consider adding this to the manuscript.

Unfortunately, due to confidentiality reasons, we are not allowed to show the actual measured start and stop events. In practice, the static curtailment plan that is implemented at the farm is different from the 3 that are presented and discussed in the paper since for the same confidentiality reasons, we were not allowed to disclose its details. It would thus have been difficult to compare the predicted start/stop events from the 3 proposed curtailment plans with the actual events measured at the turbine since the conditions for starting/stopping the turbines are not the same.

Line 559 and after: you use decimal points to describe time, but you have a maximum accuracy of 10/60 minutes = 0.166 hour > 0.1. Please round up to full hours or use integer 10-minute intervals.

All mentions to durations have been rounded to full hours.

Section 3.2: Again very long, and a lot of the text adds no information that isn't provided by the figures.

We agree that the original section was overly detailed and repeated information already conveyed by the figures and by the mode-specific discussion in Section 2.4. We have therefore shortened Section 3.2 substantially. The revised text provides a more concise overview of the fatigue results, groups the load channels according to their response, and refers back to the previously discussed aeroelastic behavior rather than repeating it. We have also condensed the comparisons within the static and dynamic strategy families, the turbine-to-turbine spread, and the energy-production results, while retaining the interpretation needed to explain the cumulative fatigue response.

Figure 12: It would make a lot more sense to me if all the same elements (tower-base, blade-root, etc) are grouped together in the same row or column. Furthermore, because the Static100 case results in so much higher load variations compared to all other cases, the y-axis is stretched too much for all other signals, and subsequently, the value of the boxplot is diminished. Consider removing the Static100 case from this figure or plotting it separately.

We agree that grouping the load channels by turbine component improves the readability of the figure. Figure 12 has therefore been rearranged into four rows and three columns, with the tower-bottom, tower-top, blade-root, and main-shaft channels grouped by row. The standardized channel acronyms have also been added to the subplot titles, and the figure dimensions and text sizes have been adjusted.

We have, however, retained the Static100 case in the same panels as the other strategies. Static100 represents the most restrictive static strategy considered and is important for illustrating the full range of responses, including the pronounced sensitivity observed for blade-root flapwise loading. We recognize that its larger changes compress the visual differences among some of the remaining strategies. Nevertheless, these differences are generally only a few percentage points, and displaying them on a narrower or separate scale could visually overemphasize their magnitude. Retaining all strategies on directly comparable scales therefore better communicates both the generally small changes for most cases and the specific exceptions associated with Static100.

Line 599: period missing after Static100.

Period added.

Line 611-612: “This is illustrated in Figure 13”: In my opinion, this also follows logically from Table 5. I therefore see limited value in including Figure 13.

Thank you for this comment. We agree that the table provides the aggregate number of start-up and shutdown events and the associated downtime for each strategy. However, it does not show the wind speed and TI conditions under which the individual transitions occur. This distinction is important for interpreting the fatigue results, because the fatigue contribution of a transition depends not only on the total number of events but also on the inflow conditions at which they take place. Figure 13 therefore complements the aggregate information in the table by showing the distribution of the actual start-up and shutdown events over wind speed and turbulence intensity. In particular, it supports the interpretation of the increased blade-root flapwise fatigue observed for Static100, for which a larger proportion of transitions occurs at higher wind speeds. We have retained the figure and revised both the figure and surrounding text to make this purpose more explicit.

Figure 13: If you do keep this figure, define acronym WSP. Furthermore, the rhs of this figure is unreadable right now with the number of bins shown. Either increase the size of this figure or increase the bin size so the number of bars is reduced.

Thank you for these suggestions. We have replaced the abbreviations in the axis labels with the full names of the variables. We have also increased the TI bin width, enlarged the figure, and increased the font sizes.

Line 624-633: This paragraph in my opinion summarizes the true relevance of this section, combined with Figure 12. A lot of what is previously written in this section can honestly be removed.

We agree that the identified paragraph provides the main synthesis of the fatigue analysis, and we have retained it as such. However, the preceding discussion has not been removed entirely because it provides the evidence supporting that synthesis, including the channel-specific effects of idling and transient events, the differences within the static and dynamic strategy families, the influence of transition conditions, and the turbine-to-turbine variability. We have shortened this discussion, as explained in our reply to the reviewer's previous comment regarding section 3.2.