Referee 1 - Anonymous

The authors thank the referee for the suggestions, comments and insights, which has led to improvement of the paper. Please find below the referee's comments (RC), the corresponding author's comments (AC) and the changes in the manuscript. PXLY refers to page X and line Y in the *revised* manuscript.

RC: The paper clearly shows that QuLAF can either under-predict or over-predict the results from FAST, and can sometimes match them perfectly through a favorable combination of discrepancies. For example, the authors show that for DLC1.6, a perfect match between the two models in tower base bending moments is obtained. However, this perfect march results from opposite discrepancies which cancel one another. In such a case, the reliability of the approach can be questionable as a good result is obtained for "bad reasons". Although the tool is of course intended for use in a pre-design phase, it would be useful if the authors could elaborate more on the reliability/repeatability of such results for different conditions and design types.

AC: Agree. The shortcomings of the QuLAF model observed for wave-dominated or wind-dominated situations are sometimes cancelling each other when combined wind and waves are applied (as in DLC1.6) for the present floater. We have updated the text to make it clearer and to emphasize this cancellation effect (see P12L10). The authors also point out that this "lucky" cancellation effect is no specific to this model only, but it can also show when e.g. comparing state-of-the-art numerical results to experimental measurements.

RC: Additionally, the QuLAF approach is restricted to 2D analyses with aligned wind and waves. It also models different physics than FAST (e.g. the mooring system in FAST introduces different sources of damping). More insights could be given on how these assumptions are likely to affect the accuracy and reliability of the results for different designs.

AC: The model is meant to complement existing state-of-the-art tools, giving a preliminary quick overview of the response and loads for a wide range of environmental conditions. After this preliminary screening, the time-domain model should be used to analyze in more detail specific load cases - e.g. cases with extreme loads or transient events (see P3L22-24 and P32L16-22). We have included a sentence on mooring in P4L20.

RC: P. 4 L. 27: An estimation of how much faster QuLAF is compared to FAST could be valuable

AC: We agree that this information would be valuable for the paper. We have added a comment on P4L4 regarding the computational times.

Please note that other minor changes have been introduced in the text to improve readability and fix a few typos.

Referee 2 – Maurizio Collu

The authors thank the referee for the suggestions, comments and insights, which has led to improvement of the paper. Please find below the referee's comments (RC), the corresponding author's comments (AC) and the changes in the manuscript. PXLY refers to page X and line Y in the *revised* manuscript.

RC: The only main comment I have is the following: precise quantitative differences between the results obtained in FAST and QuLAF are presented, but they are qualitatively classified as "good", "acceptable", and so on. It is not clear to me what is the criterion utilised to judge the goodness of the results, i.e. what would be the "unsuitable/acceptable/good/very good" thresholds (i.e. 30%/25%/10%/5% ? Different for different parameters?), based on state-of-the-art industry experience. I can appreciate that it is always difficult to have some precise numbers, but since this work has been carried out as part of the EU project Lifes50+ I wonder if the authors could add a discussion regarding this aspect, taking advantage of the close collaboration with some of the main FOWT support structure designers during the project.

AC: We agree with the referee that it was not very clear. Effort has now been put into streamlining the classifications of the results – utilizing: 0-5% (very good), 5-10% (good), 10-15% (fairly good) and 15-20% disagree. We have updated the results discussions in section 5. Our general comprehension is that in the pre-design phase you can accept lower accuracy just that the trends are right.

RC: Pag.6, line 19: "Six different wind and wave seeds were simulated for each environmental condition" and, later "a simulation time of 5400s with the same length of turbulent wind field was used for all the load cases including 1800s run-in-time to remove any transient response in the time-domain model". Does it mean that transient (1800) + 6 x 10 minutes simulations (each one with a different wind and wave seeds) have been adopted?

AC: We did not carry out any 10 min simulations. Each simulation of 90min (30min transient) is done for a specific peak period, wind/wave seed and mean wind speed, i.e. (7 wind speeds x 3 peak periods x 6 seeds) x 5400s. We have extended Table 3 on page 7 to include number of simulations per load.

RC: Pag.14, line 11: "The deviation levels in Table 8 are of the same magnitude and the reason for this is that only maximum values have been considered in the table. This might not be representative for this transient load case, where also the negative values have high influence, as can be seen in the left column of Figure6." Would it be possible to add a table relative to the max (in module) negative values, and discuss these as done for the max (in module) positive values?

AC: We agree and thank the referee for the comment. We have extended Table 8, page 15, the discussion and included Figure 7, page 16 with time series to clarify.

RC: Pag.2, line 7: "especially if they are carried out with time-domain numerical tools simulating at realtime CPU speed" Please clarify what it is meant by "at real-time CPU speed", indicating the simulated-tosimulation time ratio.

AC: Yes, by "real-time CPU" we mean a simulated-to-simulation time ratio of 1. We have updated the text, see P2L8.

RC: Pag.2, line 10: "when the concept design is more converged", please re-phrase, not very clear.

AC: We changed "converged" to "refined" (P2L11).

RC: Pag.10, caption of Table 5, please re-phrase expanding it (at the moment a bit difficult to understand).

AC: We agree. We changed all the captions of the result-tables to make it more clear (Table 5-9).

RC: Pag.15: "The simulations consists of 18 realizations (i.e. six seeds)" How long each simulation? Would it be possible to summarise the info below, adding them as additional columns on the right in Table 5? - Length of simulation - Timestep of integration - Number of seeds (and how many minutes for each seed).

AC: We thank the referee for the suggestion. We have adapted the idea into Table 5.

RC: Fig.9 The names of the load cases seem to be the names of the files used – i.e. not very clear. Furthermore, some of them are cut, and in general very small to be read. It is more important to highlight the fact that FAST and QuLAF agree or disagree on the load case ranking, than the specific name of the load case.

AC: The figure was indeed not very clear. We have changed the layout and labelling of the bars on Figure 10, page 20 so it is easier to read. As the reviewer writes, the reason for having this figure is to highlight the agreement or disagreement on the load case ranking.

Please note that other minor changes have been introduced in the text to improve readability and fix a few typos.

Referee 3 – Tor A. Nygaard

The authors thank the referee (and the additional contributors) for the suggestions, comments and insights, which has led to improvement of the paper. Please find below the referee's comments (RC), the corresponding author's comments (AC) and the changes in the manuscript. PXLY refers to page X and line Y in the *revised* manuscript.

RC: Although QuLAF is well described in the references, many of the readers working on floating wind turbines have most experience with time domain models, and I think the article would benefit from some clarifications.

AC: We agree that a more extended description of the model would be helpful. We chose the present quite short description to save on paper length and to avoid overlap with the original QuLAF paper which is also published in Wind Energy Science. The model is extensively described in the companion paper by Pegalajar-Jurado et al. (2018), where the frequency-domain solution is introduced and compared to the time-domain solution.

RC: In the left plots of figure 2, we have several results for each wind speed. The way I read the paper, for each wind speed, three sets of Hs and Tp are generated from the joint probability distribution. Each of these three realizations are computed with six different wind and wave seeds (also realizations). If indeed the use of several wind and wave seeds for one particular combination of Vm, Hs and Tp are used for the frequency domain model, please explain why this is done. Many frequency domain models work with distributions as input and output, directly giving the results for an infinite number of realizations. Here, however, does the input to QuLAF contains phase information for the particular realization at hand? Can the QuLAF results then be transformed back to the time domain, to be directly compared with the time domain FAST results, and post-processed with the same methods, such as rainflow counting?

AC: Yes, QuLAF contains phase information, since time-series of precomputed aerodynamic loads and freesurface elevation are input to the model. As a consequence, time-series of the results are available for comparison to time-domain models and for further analysis (note that fatigue damage-equivalent loads at the tower base are one of the metrics in this paper and in the previous paper by Pegalajar-Jurado et al. (2018)). We have added a sentence, see P4L23, to make this more clear.

RC: The ultimate nacelle accelerations are underpredicted in QuLAF, whereas the ultimate tower-base bending moments agree well. Often, accelerations are more sensitive to higher modes than ultimate bending moments. I did not find information on the number of tower modes used in FAST for this application. If it uses more than one tower mode, the following comment may be relevant: In addition to the underprediction of the wave excitation loads for strong sea states due to the omission of viscous hydrodynamic drag forcing, could the omission of the second tower mode in QuLAF also be part of the explanation? One way to examine this would be to turn off modes two and higher in FAST, or look at the response-spectra from FAST. Please include information on the number of tower modes used in FAST, and, under model limitations for QuLAF, mention that only first tower bending modes are used. Have any sensitivity studies on the number of tower modes ben carried out?

AC: The number of tower modes in FAST has been added now, see P3L8. Regarding the effect of the higher tower modes, we checked the response, where mode two and higher in FAST were turned off. This only

had very minor impact on the nacelle acceleration, thus we believe the under-prediction of the nacelle acceleration is due to the over-estimated damping of the tower mode.

RC: The aerodynamic damping model seems to be one area where changes could significantly improve the results. One possible improvement would be to perform the decay test in FAST with flexible blades, resulting in an eigen frequency closer to the coupled tower frequency in QuLAF, thereby reducing the overprediction of aerodynamic damping. It should also be possible to have an aerodynamic damping model in QuLAF model derived directly from a linearized BEM model.

AC: We thank the referee for the ideas of improvement. We are already exploring better ways to extract the aerodynamic damping. The results, however, are still not mature. Inclusion of flexible blades could also lead to an improvement, but introduces choices as to what specific mode one should choose (blades in phase (anti-phase etc.)). We have chosen to stay with the current simple approach and simply accept its limitations.

RC: I find it quite surprising, interesting and perhaps under-communicated that an emergency stop can be successfully computed with a frequency domain model. More details, such as direct comparison of the time series of tower base bending moments and nacelle accelerations would be very welcome.

AC: We agree, it is very interesting that QuLAF is able to reproduce a transient event and thank the referee for the suggestion of improvement to this comparison. We have added a figure showing the time series of a specific case in the discussion of DLC2.1, see Figure 7, page 16.

RC: Page 4, line 14: Did you check that there is no numerical damping in the decay test? One way to test this is to scale down the lift-and drag coefficients, or somehow provide an excitation of the tower top without rotor aerodynamics present.

AC: Regarding the numerical damping, we did a clamped pre-study of the model in 0 m/s and forced excitations of the tower top. The study showed that the response was undamped, i.e. no numerical damping present.

Please note that other minor changes have been introduced in the text to improve readability and fix a few typos.