Dear Mr Goupee,
we are very thankful for taking your time to revise this extensive manuscript and for sharing your expertise and experience with us. We found your advice helpful, inspiring and in a positive spirit in all cases and were hopefully able to implement the vast majority of your comments.

We respond to your comments directly in the following:

This reviewer finds this manuscript by C.W. Schulz et al. to be of interest to the FOWT community. The article is also fairly well written and organized.

That said, this reviewer has several comments the authors may wish to consider addressing in a revised version of the manuscript. They are as follows:

1) There is surprisingly little quantitative information on the tested FOWT system (dimensions, mass properties, mooring geometry, etc.). This would appear to make the work difficult to reproduce. There is a mention of some information in Appendix B, but very little can be found there.

We agree on this. Unfortunately, the platform design is owned by a company. We are therefore not allowed to give more details here.

2) On a related topic, there is very little specific information on the numerical modeling inputs (e.g., aerodynamic properties, hull drag coefficients, mooring stiffnesses, etc.). Again, without providing this information, the work is difficult to reproduce. Also, adding detail in the modeling inputs may help the reader better understand discrepancies observed between the tests and simulations.

We added more details to appendix E.

3) The authors provide a fairly nice literature review, and do a good job of justifying their tank testing choices for the purposes of reducing uncertainty. However, no quantitative uncertainty information is provided for any of the measurements, nor is there any quantitative evidence provided that the approaches employed leader to less uncertainty than competing approaches. Based on the title of this manuscript, this reviewer thinks it is reasonable that a reader may expect this information to be present in the manuscript.

This is clearly a weakness of the submitted manuscript and we are thankful for this hint. We added a section (6) in incorporated the repeat tests into it. We hope that this provides more evidence for the improvement of the quality of the model tests. In addition, we added a more elaborate evaluation of the wind field quality and a comparison with other wind generators (see section 4.1). In this course, the evaluation of the non-uniformity of the wind generator was performed more precise so that even lower values resulted. Finally, a comparison with a wind tunnel measurement campaign was added under section 7.1 to prove evidence for the high quality of the wind field and sensing system.

4) Several choices are made to reduce uncertainty in the testing (wind machine design, wind turbine reduced size, focusing on regular waves, etc.). However, these choices deviate somewhat from the physical properties and design load case environments of usual interest in FOWT design. The authors may want to better defend the choice of reducing uncertainty for these ‘off-design’ scenarios as opposed to attempting more complicated, uncertain, but more realistic tests (closer to Froude-scale turbines, active turbine control, irregular wave environments, etc.). The former may reduce

uncertainty, the latter will provide data that can be used to exercise numerical models in areas more relevant for modern FOWT design. As an example, the environments considered in this work will not induce any second-order wave drift forces which can significantly impact certain FOWT dynamic responses. In addition, these are often the hardest to capture with numerical models, and as such, are of great interest in tank testing campaigns.

Thanks for this hint. We added a paragraph, which covers most of the points. However, we tried to keep it short in order to avoid lengthening the paper too much.

"… Therefore, although the Froude similarity of tower top motion velocity and wind speed is violated, the utility system is well suited for the validation of numerical methods. The special value of the proposed scaling approach is to enable a precisely known wind environment and wind turbine characteristic. This is achieved by a more elaborate wind generator and a much better covering of the (moving) rotor swept area compared to other tests due to the small size of the turbine. The ability to test the smaller rotor in a wind tunnel environment with sufficiently low blockage ratio and under highly controlled conditions (see section \ref{chap:windtunnel}) opens the possibility to validate the aerodynamic simulation model accurately and to identify measurement differences arising from the non-ideal wave tank environment. Both together leads to a well defined (and known) thrust force, which is applied to the tower top, with a comparatively low level of noise. This, in turn, yields a low contribution of the aerodynamic system to the random and systematic uncertainty of the platform motion, which is a major improvement in comparison to the above listed studies. In addition, the low level of noise enables the reliable compensation of gravitational and inertial loads from the tower top forces so that a direct validation of the simulated rotor thrust is possible during a motion cycle. Naturally, these improvements over existing test strategies are achieved in exchange with a less realistic behaviour of the model FOWT as discussed above. "

5) In Figure 4, the amplitude of the dynamic thrust force is approximately 1/3rd of the anticipated full scale amplitude. An explanation and derivation is provided, which is appreciated. However, is the uncertainty in this dynamic thrust force reduced by at least 1/3rd as well? If not, perhaps other approaches that can capture not only the mean thrust force, but the full-scale variation in the thrust force would be better to pursue as they better represent the desired physics and the uncertainty as a percentage of variation would be no worse than the proposed approach (consider reviewing some of the other works produced on the recent FOCAL test campaign).

This is an important note. We believe that we have to distinguish two things here: First, the application of a well defined and realistic rotor thrust to the tower top and second, the ability to measure this load during waves and compare it to simulations.

When considering the first point, it might be questionable if the introduced load by the rotor is three times more accurate. However, in the work of Mendoza (2022), it can be seen that the repetition error of the rotor thrust is in a range of 1% or a bit more, which is indeed more then our estimation of 0.6%. In addition, the ability of the numerical models to reproduce the rotor thrust over a certain range of TSRs shown in this work is much lower compared to our newly added figure the present publication. This itself is not an argument, because the numerical models might be erroneous, however, it gives a *hint that some physical effects like uncertainty in the wind field, the measurement system or the blade geometry also be present.*

From our point of view, the second point is more important: The aim of the testing is to identify modelling inaccuracies in particular part of the model. Therefore, the ability to measure the aerodynamic rotor thrust force while the platform undergoes motions is a key capability of this way of testing as this allows for a direct comparison with the aerodynamic simulations. We reviewed some of the the FOCAL-related works again and did not find a comparison between simulated in measured aerodynamic rotor thrust when the turbine undergoes wave motions. (And no quality check of a potential inertia and gravitational load compensation.) We are also not aware of any other successfully validated attempt to do so. Please let us know is you are aware of something like this. Considering our state of knowledge as given, a comparison of the uncertainty of rotor thrust measurements would not even possible because no other approach proofed to be working. Of course, this stands under the retention of the existence of other works. Please let us know if you know about such works. In this case, we would naturally compare the uncertainties to each other.

6) Several qualitative descriptions of the size of the wind machine relative to the wind turbine are provided. Consider precisely quantifying the room the rotor has to move in heave/sway while still remaining in the low spatial variation, low turbulence intensity portion of the wind machine jet.

We added a sentence:

"Deviations in oft the rotor position in heave and sway direction of approximately 0.2\,D are tolerable in this context. This limit is by far not exceeded in the present tests."

7) In Figure 6, it would be nice if the color bar variation was focused more on the rotor area; by including the pieces outside of the jet, it is hard to visually pick up on the turbulence intensity and spatial variation trends in the rotor plane.

We agree on this and changed the figure.

8) The last paragraph on page 14 discusses a modeling approach where an angle of attack offset is used as a viscous correction. Is this a standard approach? If so, can a reference be provided? This reviewer has not seen this method used before in a model correlation study.

This method is more or less unique because it is only relevant for panel methods and for the special situation, where we have a very narrow band of angles of attack at the rotor radial stations. The method is described a bit more detailed in

Schulz, C. W., Netzband, S., Özinan, U., Cheng, P. W., and Abdel-Maksoud, M.: Wind turbine rotors in surge motion: new insights into unsteady aerodynamics of floating offshore wind turbines (FOWTs) from experiments and simulations, Wind Energy Science, 9, 665–695, https://doi.org/10.5194/wes-9-665-2024, 2024

However, this is not relevant any more, because we decided to replace the panel method with our inhouse lifting-line method. This is due to the fact that we now have a LL model of the exact same rotor had to rerun all the cases when we incorporated the 'real' waves. The results are only slightly deviated by this change and the analysis is adjusted slightly.

9) No information is given on conditions to reach the ‘steady-state’ responses shown in the plots, nor how many cycles are included in the plots.

Thanks for this hint. We forgot to include this. The following sentence was added:

“For this and all following analyses, phase averaged data was computed on the basis of measurements, which started after aperiodic effects decayed, which took at least ten motion cycles after the approaching of the first wave. The data sets had a length of 6 - 12 motion cycles, depending on the quality of the generated waves and the occurrences of obstacles or noise in certain measurement channels”

10) This reviewer has seen other works that perform ‘hybrid’ simulations of mooring systems with better comparisons between experiment and simulation. The quality of the comparisons depends significantly on the numerical approach being employed in the mooring analysis (quasi-static, lumped mass, FEM, etc.) and mooring line hydrodynamic properties (e.g., transverse drag coefficient). The author M. Hall that has been referenced in this work has a good article that may be worth reviewing.

We are thankful for this hint. We did not know about this particular part of Hall’s work. This demonstrates nicely the idea that ‘hybrid simulations’ are of major value when identifying the source of particular mismatches between simulations and experiments. We added a reference to hist work. Indeed, our implementation of the mooring model is tightly oriented on the model Hall proposes in

“Validation of a lumped-mass mooring line model with DeepCwind semisubmersible model test data”.

However, we are aware of the sensitivity of the results on the choice of the mooring system parameters and could not find a suitable set of parameters, which leads to much better results. Therefore, we believe that our mooring system itself contains too complex effects (in comparison to real mooring systems) to be captured accurately. This view can be supported by considering the amount of high frequent oscillations and obstacles in the measured mooring loads. In addition, it has to be noted that Hall compares overall the mooring line tension, which is most likely dominated by the load in heave direction due to the angle of the mooring lines. In this direction, our simulations also show fair agreement, while the large differences occur in surge direction, which is not directly evaluated in Hall’s work.

11) For the full simulations, why aren’t the actual recorded waves used as inputs? Without doing so, it is hard to understand which discrepancies are due to modeling deficiencies/test uncertainties, and which are due to incorrect model inputs. This reviewer thinks it is common practice to use the as measured wave in model correlation studies, and am surprised that this is not done here.

We agree on this. Some software problems prevented us from doing so. However, these problems have been solved in the meantime and we repeated all simulations with the ‘real’ waves. All figures have been exchanged and the analysis was slightly reworked. However, the agreement between measurements and simulations did not improve significantly.

12) Is there a reason the results are not provided at full scale? Results are usually more intuitive when presented at full scale in this reviewer’s opinion.

Although it is common to present upscaled data, we decided to stick to real values. We did this for a number of reasons:

- *The simulations have been performed in model-scale as the viscous correction shall work in the correct Reynolds number regime. We do not want to suggest something else tot he reader.*
- *Model and sensing system uncertainties as well as estimations on repetition errors are given in absolute numbers. For us as researchers that deal with this kind of measurement equipment and data, absolute values are way more intuitive. In order to compare these uncertainties with the data, the scale should be the same.*
- *We decided to not normalise the rotor thrust force in terms of a Ct value as it should stay directly comparable to the mooring surge force.*
- *Due to the variety of turbine ratings used for FOWT analyses in literature (~2 – 22 MW), we do not find absolute values for heave, surge, rotor thrust and mooring loads at scale much more intuitive compared to the model scale data as those strongly depend on the full-scale turbine rating.*

13) There are some other minor issues that should be addressed: There are some widows/orphans, the paragraph indenting is inconsistent, and the figures are often not located very closely to their first mention in the text.

Thanks for the hint. We tried to fix the inconsistencies in indentation. As the typesetting will change strongly when the paper is converted to the journal layout, we will fix the remaining issues at this stage.

Best regards and many thanks for your extensive review,

Christian W. Schulz on behalf of the authors