—- Main Positive Feedback —-

- The manuscript demonstrates strong scientific motivation and engages with a highly complex topic in a structured way. The authors succeed in breaking down the dynamic interplay of multiple driving factors into separate aspects and analyzing them in considerable detail.
- The separation into CFD-CSD/OV and CFD-CSD/TN configurations is particularly valuable, as it provides insights into the impact of the torsional degree of freedom that would otherwise remain hidden in real-world measurements.
- The figures and plots are generally clear, well-designed, and effective in illustrating the results. The effort to validate the structural model against five alternative models is commendable and shows attention to methodological robustness.
- The simulations also provide several interesting findings:
    - The 15 MW turbine exhibits slower wake recovery due to tower effects, in contrast to the behavior of the 5 MW turbine
    - The results highlight the major influence of the fidelity of both aerodynamic and aeroelastic solvers on the predictions. In particular, the study suggests that CFD-based approaches may be crucial for accurately capturing tower effects, blade dynamic response, and wake recovery.
    - The torsional degree of freedom has a clear impact on aerodynamic quantities, for example leading to a decrease in the power coefficient and affecting blade deformations.
      The presence of the tower produces peaks in CpC_pCp and CtC_tCt that could have important implications for fatigue loading.

—- Detailed Positive Feedback —-

- L. 530–532: *“The isolated low-frequency peaks found in BeamDyn and ElastoDyn suggest that these solvers tend to over-simplify the aerodynamic fluctuations associated with phenomena such as wind shear and tower shadowing.”* – This is a valuable observation and aligns well with what one would theoretically expect.
- L. 544–548: *“In agreement with previous studies, the results thus suggest that including the torsional degree of freedom in the structural solver is crucial for accurately describing the amplitude and dynamical behaviour of the aerodynamic quantities. Moreover, it is observed that duly taking into account the torsional degree of freedom reduces the value of Cp.”* – It is very good to see your work confirming both previous studies and qualitative trends that are expected on physical grounds.
- Figure 7 (page 13): The direct comparison between CFD-CSD/OV and CFD-CSD/T is highly informative in illustrating the impact of torsion on pitching moment and aerodynamic forces. Furthermore, the alignment between the corresponding CFD-CSD and OpenFAST variations (L. 540–542: *“Concerning the forces on the blade and the incidence angle, one can observe a rather good match between the CFD-CSD/OV solver and ElastoDyn, as well as between the CFD-CSD/T model and the BeamDyn solver”*) represents a strong qualitative finding.
- Figure 8 (page 14): This figure clearly highlights the oscillation at the tower position and is very effective in conveying the underlying physics.
- L. 568–570: *“Future work will explore the effect of turbulent fluctuations at the inlet to better investigate the impact of the atmospheric boundary layer on the aerodynamic*

forces, loads and deformations of the present turbine." – This is a logical and well-justified next step following the present analysis with laminar shear inflow.

—- Main Critique Points —-

Documenting the software and its validation status.

- As a reader, I found it a bit hard to trace how the different pieces of your in-house UTD-WF framework were developed across prior work, and what has been validated under which conditions. This is important to know that the in-house code you use - i.e., the foundation for all your work and insights - is credible and correct in the first place. Adding a short provenance/validation paragraph would strengthen the manuscript. For the protocol, this is what I understand from the literature:
    - Santoni et al. (2015) first introduced UTD-WF (they didn't call it "UTD-WF" yet, but rather incompressible LES + ALM, with IBM for tower/nacelle) and reported limited wind-tunnel validation at NTNU for a single model turbine (mean velocity & TKE), establishing basic fidelity of the ALM+IBM setup.
    - Santoni et al. (2017) then reproduced the NTNU "Blind Test" and compared simulations to Krogstad et al.'s measurements, quantifying the impact of tower and nacelle—further supporting the IBM+ALM approach.
    - At wind-farm scale, Santoni et al. (2020) used an actuator-disk representation within UTD-WF and a mesoscale–microscale coupling, considering both momentum-only and momentum+TKE variants.
    - Della Posta et al. (2022) introduced the two-way FSI coupling (ALM in UTD-WF + a modal structural model; ALM/IV/IVT options). While that paper mainly focused on methodology and inter-model comparisons, it did not include a dedicated experimental validation.
    - Finally, Della Posta et al. (2023) integrated a Beddoes–Leishman unsteady-aerodynamics model in the LES-FSI framework and examined uniform, laminar-shear, and turbulent ABL inflows; comparisons are discussed against reference datasets (including HAWC2-based results reported by Heinz, 2013).
- Clarifying which subsets constitute formal validation would help readers map prior evidence to your present setup. In that light, two clarifications would make your contribution easier to interpret:
    1. Scope of prior validation vs. present cases: Prior validations include uniform and ABL inflows for a 5 MW turbine; your study addresses the IEA-15 MW case for a sheared laminar inflow configuration. Since larger rotors can amplify aeroelastic effects, it would be helpful to note explicitly that the present application extends the validated setting and to discuss any implications or arising uncertainties.
    2. Citations for solver provenance: Where you cite the solver's origin, consider also referencing Santoni et al. (2015) as the first UTD-WF publication, prior to Santoni et al. (2020). Further mentioning the following papers and validations that were added to Santoni et al. (2015) would help readers that are unfamiliar with the code's history. Just so they know that the inhouse code you use is validated within a certain range.

Comparing BEM with CFD for aeroelastic simulations

- The conclusions drawn about the relative suitability of CFD-based versus BEM-based aeroelastic simulations are not entirely straightforward, because the setups being compared differ in more than one respect. In particular:
    - Case A (CFD-CSD) couples a finite-volume LES solver with ALM to a *linear* structural model (A.1: OV; A.2: TN).
    - Case B (OpenFAST with BeamDyn) uses BEM aerodynamics with a *nonlinear* structural solver.
    - Case C (OpenFAST with AeroDyn/ElastoDyn) combines BEM aerodynamics with a *linear* structural solver.
- Because both the aerodynamic *and* structural solvers vary simultaneously, it is difficult to isolate whether observed differences stem from the flow solver or from the aeroelastic solver (or from their specific implementations). Any causal conclusions should therefore be presented with appropriate caution.
- A related point concerns the comparison with Bernardi et al. (2023). That study focused on a different turbine (NREL 5 MW), under uniform inflow, and reports blade load distributions that do not match those of the present paper. Drawing direct inferences from those results to the IEA 15 MW case with sheared inflow is thus challenging, as multiple differences may drive the discrepancies. In addition, the current manuscript discusses the 5 MW results extensively without showing them, requiring the reader to switch between papers to follow the discussion.
- To strengthen this part, I would suggest either limiting the cross-paper comparison, or alternatively presenting a direct comparison in the present manuscript — for instance, by reproducing a nondimensionalized plot from Santoni et al. (2017):

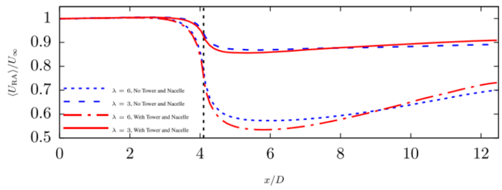

FIGURE 6 Streamwise velocity averaged in time and over the rotor disk, $\langle U_{RA} \rangle$: $\lambda = 3$ with (——) and without tower and nacelle (- - - -); $\lambda = 6$ with (— - —) and without (- - - - - - -) tower and nacelle. The vertical dashed lines denotes the position of the rotor [Colour figure can be viewed at wileyonlinelibrary.com]

alongside your own:

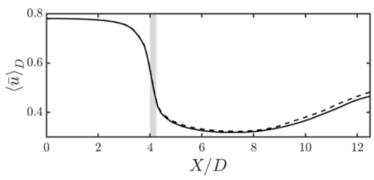

Figure 4: Rotor-averaged velocity along the streamwise direction normalized by the undisturbed velocity at the rotor height, namely, $U_\infty = 10\ m/s$. The grey region represents the area covered by the rotor. (RO - - - -, TN ——).

in a comparable format that allows readers to interpret the differences more clearly.

Convergence study

- The grid convergence section currently relies on only two simulations. While this gives a first impression, two points are generally not sufficient to establish a clear convergence trend. I would therefore suggest either adding at least one additional grid resolution to demonstrate a systematic trend, or explicitly referring to convergence studies from previous publications with a closely similar setup.

- In addition, the statement *“the results obtained using the coarse and fine grids are extremely close … the curves of the angle of attack are almost indistinguishable”* comes across as rather qualitative. It would strengthen the section to include a more quantitative measure of convergence (e.g. RMS difference, maximum deviation along the span, or an error norm), so readers can assess the actual magnitude of the differences. Further, the figure of the angle-of-attack comparison in the grid comparison study would be much more clear if it would be scaled to include a maximum angle of 15 degrees. The angles of attack on the cylindrical root section are not of interest to the reader.

Tip-loss models

- I also noticed that different formulations of the tip-loss model are used across the compared solvers. Specifically:
    - OpenFAST applies the classical Prandtl tip and root loss correction (l. 222).
    - The CFD-CSD solver employs the Shen et al. (2005) correction (l. 144).
- Since the Shen formulation is not equivalent to the standard Prandtl approach, the choice of model may itself contribute to differences in the results.
- For the sake of a fair comparison, you could report clearly which correction is applied in detail in each case and provide the coefficients used in the Shen model. This would make the comparison more transparent to the reader.

— Detailed Critique Points —

2 Methodologies

- 2.1 CFD-CSD Solver
    - L. 121: The subsection caption currently reads “2.1 CFD-CSD solver”. From what I know, subsection titles should follow a consistent style. Please either use title case throughout (e.g., “2.1 CFD-CSD Solver”) or sentence case consistently. At present, there is a mix (for example, “4 Results and Discussion” in title case versus “3 Flow and structural setup” in sentence case). Perhaps also check the WES journal guidelines, whether to use title or sentence case.
    - L. 123–125: The text states: *“The simulations of the flow around the wind turbine are carried out through Large-Eddy Simulations (LESs) of the incompressible, filtered, 3D Navier-Stokes equations, employing our in-house UTD-WF solver (Santoni et al., 2020).”* To my understanding, the first introduction of the UTD-WF code was actually in *Santoni et al. (2015), “Development of a high fidelity CFD code for wind farm control”*. The 2020 paper (*Santoni et al., 2020*) applies the solver at wind farm scale, but the code originates from the earlier work. Since UTD-WF is not as widely established as solvers like OpenFOAM, it would strengthen the methodology section if you also reference publications where the solver has been validated against either an established CFD solver of similar fidelity or experimental data.

    - L. 138: The manuscript currently cites *Troldborg (2009)* for the Actuator Line Method (ALM). A more appropriate primary reference would be *Sørensen and Shen (2002)*, who first introduced the ALM.
    - L. 144: You mention comparing the Shen tip-loss correction with the Prandtl tip-loss model. Could you please specify the chosen free parameters in the Shen model? It would also help to clarify how you ensured that both tip-loss approaches are consistent, so that observed differences in results are not simply due to discrepancies between the models themselves
    - L. 159–160: The structural model is attributed to *Della Posta et al. (2022, 2023)*. To provide readers with a clearer overview of how the full solver framework was assembled, it would be useful to briefly note the sequence of developments: *Santoni et al. (2015)*: introduction of the ALM–LES flow solver (UTD-WF), *Santoni et al. (2017)*: tower–nacelle representation using the immersed boundary method, *Della Posta et al. (2022)*: coupling of the aeroelastic solver, *Della Posta et al. (2023)*: implementation of the unsteady aerodynamics model (Leishman–Beddoes dynamic stall). You might phrase it along the lines of: *"The structural model used in the present study builds upon the UTD-WF framework progressively developed in Santoni et al. (2015, 2017) and further extended by Della Posta et al. (2022, 2023), where the aeroelastic solver and the Leishman–Beddoes dynamic stall model were implemented."*

3 Flow and Structural Setup

- L. 235: The subsection caption currently reads *"3 Flow and structural setup"*. For consistency with the other section titles, the first letter of nouns should be capitalized (e.g., *"3 Flow and Structural Setup"*). Please check the WES journal guidelines and apply the chosen style consistently throughout
- L. 248–249: The text states: *"The turbine location is 4 diameter units from the inlet and centered in the spanwise direction."* Could you please elaborate on whether this positioning is sufficient to ensure that there is no unintended influence between the turbine and neighboring boundaries? It would be helpful to explain how these dimensions were chosen—either by referencing a small domain-geometry convergence study (perhaps in the appendix) or by citing relevant literature that uses a similar domain setup.
- L. 250: The manuscript notes: *"Furthermore, we impose (...) a radiative outlet boundary condition."* This boundary condition type is relatively uncommon in CFD literature. A short explanation of its formulation would therefore be valuable for readers. For instance, *Santoni et al. (2017)* also use this condition and explicitly provide the form of the equation: ∂U~i∂t+C∂U~i∂n=0.\frac{\partial \tilde{U}_i}{\partial t} + C \frac{\partial \tilde{U}_i}{\partial n} = 0.∂t∂U~i+C∂n∂U~i=0. To my knowledge, this BC is commonly referred to as "convective boundary condition" If you wish to use the term "radiative boundary condition," it would be helpful to clarify this naming choice and cite *Santoni et al. (2017)* for context.

4 Results and Discussion

- All captions of the subsections, similar to the caption of chapter three, should have the first letter of the nouns in capital letters. So instead of "4.1 Flow analysis" -> "4.1 Flow Analysis" and the same for 4.2, 4.3, and 4.4

- 4.1 Flow Analysis
    - L. 280–283: The manuscript states: *"Contrary to what Santoni et al. (2017) observed in their work on the 5MW reference turbine invested by a uniform inflow, the rotor-averaged velocity for the TN configuration in the wake remains slightly lower than for the OR case, indicating that wake recovery is slightly hindered by the presence of the tower."* It is useful to compare with *Santoni et al. (2017)*, but in order for readers to interpret the comparison correctly, it would be important to also highlight the notable differences between the setups (e.g., *Santoni et al. (2017)* used no-slip boundary conditions at all walls — top, bottom, and lateral — and neglected aeroelastic effects). Furthermore, I recommend caution in how the observed results are phrased. Since this finding is based on a single CFD-based aeroelastic solver, it would be safer to state that wake recovery *appears* to be hindered by tower presence, but that further validation is required as the result does not fully align with expected aerodynamics. In the following sentence, the text currently shifts from observation to causality: *"The reason for this behavior can be found in the different aspect ratio of the tower for the present turbine."* Here, I would suggest a more cautious phrasing such as: *"One possible explanation for this unexpected trend could be differences in the tower-to-rotor aspect ratio."*
    - L. 291: The sentence *"Figure 5 represents the time-averaged TKE for both configurations on different planes."* refers to a figure that is placed two pages later. This is not a critical issue, but if the layout allows, placing the figure closer to its first mention would improve readability.
    - L. 296–297: The manuscript states: *"This suggests that the tower does not increase the kinetic energy entrainment but it rather has a slight shielding effect on wake recovery."* Similar to my earlier comment, I would recommend avoiding the phrasing *"this suggests"* here, as it conveys a level of certainty not yet supported by validation against other solvers or experimental data. A more neutral formulation (e.g., *"This result may indicate …"* or *"In this simulation, the tower effect appears to …"*) would be more appropriate.
- 4.2 Aerodynamic loads on the blade
    - Figure 7: It may improve clarity if the angle of attack (AoA) is shown in a zoomed-in range (e.g., [0°, 15°]), as otherwise it is difficult to interpret the details. The legend also appears to be missing for this plot, which makes comparison harder. In addition, the BeamDyn AoA and forces in the most outboard two blade sections seem to approach zero, which looks unexpected. It would be helpful to add a short explanation if there is a known reason for this, or otherwise acknowledge it as an open point.
    - L. 339–341: The sentence currently reads: *"Indeed, the blade-tower interaction leads to an oscillations of the aerodynamic forces and of the incidence angle around $\theta = 180^{\circ}$, i.e., when the blade is pointing down."* There is a grammar error here: it should read *"leads to oscillations"*.
    - Figure 8b: There appears to be a notable disagreement in F2 between CFD-CSD/OV and ElastoDyn, even though the two models agreed reasonably well on AoA. Moreover, given that CFD-CSD/OV predicts a lower AoA compared to ElastoDyn, one would expect a lower F2 compared to ElastoDyn, yet the opposite is observed. Could you please clarify this?

- L. 353–355: The manuscript states: *“On the other hand, when torsional feedback is included, CFD-CSD/T and BeamDyn solvers agree rather well for all the quantities considered, regardless of the linearity or non-linearity of the models.”* I would suggest reconsidering this phrasing. In Figure 8a there are visible discrepancies across all azimuthal angles, and in Figure 8c there are large differences at both low and high azimuthal angles. Instead of stating “agree rather well for all quantities,” it might be more precise to describe where the agreement is reasonable and where significant differences remain.
- Figure 9: To improve readability, it would be helpful to use a consistent color scheme across the colormaps, so that interpretation of the results is more straightforward.
- L. 374–377: The manuscript notes: *“Although some mild differences can be observed in their amplitudes and phases, the frequency of these oscillations appears consistent between the two solvers and comparable with the natural frequency of the first torsional mode.”* Here I would recommend two adjustments:
  1. The differences in amplitude, especially for α and F3, are not “mild” but rather notable. In some cases (e.g., BeamDyn-TN), the qualitative shape also differs, showing more oscillations than CFD-CSD/T.
  2. Instead of subjective wording like “mild” or “strong,” it would be clearer to describe the differences in quantitative or qualitative terms (e.g., “larger amplitude,” “different oscillation patterns”).

- 4.3 Power and Thrust coefficients
  - L. 388–390: The manuscript states: *“The results reflect the dependency of the power and thrust coefficients on the tangential aerodynamic force F2 and the normal aerodynamic force F3 at the 80% of the blade, respectively.”* For clarity and consistency, please use a single naming convention throughout. At present, F2/F3 are sometimes referred to as tangential/normal, and in other places as edgewise/flapwise. Consistency will make it easier for readers to follow.
  - L. 391–392: The text notes: *“Notice that, also here, we can observe that the drop in the Cp curve appears to be rather consistently predicted by BEM and CFD.”* I would recommend revising this statement. The qualitative shape of the drop differs between the two methods: the BEM prediction exhibits notable oscillations before and after the drop, whereas these are not present in the CFD results. Thus, describing the predictions as “consistently” aligned may be misleading.
  - L. 395–399: The manuscript explains: *“Indeed, the flow induced by a thinner tower (in diameter units), as in the case of the 15-MW wind turbine, might be better described by a potential flow solution compared to the one induced by a thicker tower, as in the case of the 5-MW wind turbine, and may thus lead to the observed improved agreement between BEM and CFD results.”* Since you make several direct comparisons with the results of 5MW turbine simulations, these comparisons become a central part of the discussion. To improve readability, I suggest including the relevant 5MW results directly in the figures, rather than expecting readers to switch between different papers.
  - L. 407–408: The sentence *“Overall, it can be said that the performance drop due to the passage in front of the tower is somewhat more limited for the*

15MW NREL turbine than for the 5MW counterpart (…)" could be made more precise. The repeated references to the 5MW turbine again suggest that the corresponding results should be explicitly shown and discussed here. Also, the phrases *"it can be said"* and *"somewhat"* are too subjective. A more precise wording might be: *"Results indicate that the performance drop relative to the corresponding average performance is smaller for the 15MW turbine than for the 5MW counterpart."* Ideally, a quantified comparison (e.g., percentage reduction or absolute values) would make this statement more rigorous. Further, the 5MW simulations were performed with uniform inflow, while the 15 MW simulations are performed using a sheared inflow. The lower wind speed in the lower part of the rotor plane means that the bottom half of the rotor plane, where the tower is located, will produce less than half of the total power. This will inevitably cause the performance drop due to the tower to be smaller relative to the total produced power. Therefore the observed difference is not only due to the change in turbine size, but also due to the change in inflow conditions.

  - L. 410–411: The manuscript states: *"Moreover, results seem to suggest that for very large rotors the presence of the tower may constitute a less critical issue for the blade deformations than for smaller rotors (…)".* This phrasing is somewhat misleading. It is not possible to draw conclusions about "issues" such as fatigue without dedicated simulations or experiments. What is actually shown are differences in blade deformation predicted by aeroelastic simulations, which have not yet been validated for this case. I would therefore recommend rephrasing along the lines of: *"Moreover, the present simulation results predict that for very large rotors the tower effect on blade deformations is less pronounced than for smaller rotors."* At the same time, it would be important to acknowledge that this conclusion is based only on one 15MW sheared inflow simulation, where the tower is located in the part of the rotor plane where deflections are generally smaller, and one 5MW uniform inflow simulation, so broader generalizations are not yet possible.

- 4.4 Structural response
  - L. 458–459: The manuscript states: *"The amplitude of the deformation is however consistent with that obtained by Trigaux et al. (2024) using LES."* The wording *"consistent"* is somewhat ambiguous here. It would strengthen the statement if you could clarify in which sense the amplitudes agree — e.g., whether they are similar in absolute magnitude, in relative deviation from a reference case, or primarily in qualitative trend.
  - L. 474–476: The text reads: *"However, the gap between the BEM and the CFD-CSD/T curves is quite large. This can be attributed to the different aerodynamic and structural model used in BEM and LES."*
    1. The expression *"quite large"* is subjective; it would be clearer to either quantify the gap (e.g., percentage difference, RMS error) or replace it with a more neutral term such as *"substantial"* or *"noticeable."*
    2. The phrase *"this can be attributed to"* presents speculation as a fact. Since both the aerodynamic solvers (BEM vs LES) and structural solvers (modal vs torsional models) differ, the discrepancies cannot be uniquely traced to one factor. A more balanced formulation would be: *"These differences likely arise*

from the combined effects of both aerodynamic and structural modeling approaches."

  - L. 491–493: *"the second and third flapwise natural frequencies are indeed recovered by all the numerical models"*. These peaks are visible, because they are at the 13th and 26th multiple of the rotational frequency (13p and 26p). They are not visible, because they are at the second and third flapwise mode, because flapwise modes have very large aerodynamic damping, which is also why the first flapwise mode is not visible.

5 Conclusions

- L. 522–527: The manuscript states: *"The entrainment of kinetic energy driven by the tower leads to higher turbulence levels in the near wake, but then result into a slightly decreased mixing behind the turbine, differently to what has been found for the NREL 5MW wind turbine, whose wake recovery was found to be promoted by the presence of the tower. This finding can have important implications for the aerodynamic loads on downstream turbines in wind farms and overall farm efficiency."*
 This is a counter-intuitive result, as prior studies (e.g., the NREL 5MW turbine case) suggested the opposite effect. Since the observation is based on a single solver without confirmation from experiments or other CFD-based aeroelastic simulations, I recommend formulating this more cautiously. For example, instead of *"this finding can have important implications"*, you could write that *"this result requires further examination, as it appears counter-intuitive and has not yet been confirmed by other studies."* This way, the uncertainty is acknowledged while still highlighting the potential significance.
- L. 551–553: The manuscript concludes that *"CFD-CSD can capture complex aerodynamic loading and turbulent effects better than BEM."* While this is likely true in general, the specific evidence presented here — larger amplitudes at lower frequencies — does not in itself constitute proof of that statement. It would strengthen the conclusion to either rephrase more cautiously (e.g., *"In this case, the CFD-CSD solver captured larger low-frequency fluctuations than BEM"*) or to provide additional evidence that directly supports the broader claim.
- L. 562–563: The differences between CFD-CSD/T and BeamDyn are described in a way that attributes them to solver fidelity. However, since the two approaches differ in both aerodynamic fidelity (LES vs. BEM) and structural fidelity (torsional vs. modal models), it is not possible to unambiguously assign the discrepancies to one component. I recommend rephrasing along the lines of: *"The observed differences likely stem from the combined effects of differences in aerodynamic and structural fidelity, and cannot be uniquely attributed to one component alone."*

Appendix A Grid Convergence Study for the LES Simulation

- L. 573: To demonstrate a clear trend in grid convergence, more than two grid resolutions are typically required. With only coarse and fine grids, it is difficult to establish whether the solution is truly converging. One possible approach would be to use the finest grid as a reference and compute a root-mean-square (RMS) error relative to it. Plotting the RMS across multiple grid resolutions would then allow you to illustrate the convergence trend more quantitatively.

- L. 579: The manuscript states: *"The comparison in figure A1 shows that the results obtained using the coarse and fine grids are extremely close each other along the entire blade span."* This phrasing is somewhat vague and subjective. To make the convergence assessment clearer, it would be better to quantify the agreement, for example by reporting RMS errors or percentage deviations. Also, there is a small grammar correction: it should read *"close to each other"* rather than *"close each other."*

Appendix B. Validation of the Structural model

- Caption: Please ensure consistent use of title case or sentence case across the entire manuscript. For example, the current caption *"Appendix B. Validation of the Structural model"* mixes styles. The WES journal guidelines specify which convention should be followed — it would be good to align with that.
- L. 599–600: The manuscript states: *"The computed values of the modal frequencies appear to be consistent with the other results, although some discrepancies in the higher-order modes are observed."* Since the 6th mode (corresponding to the first torsional mode) is likely the most relevant for the present study, it would strengthen the paper to explicitly address these discrepancies. Please clarify in which sense the results are "consistent" and provide more detail on how the 6th mode compares across solvers.
- Figure B2:
    - From my understanding, the 6th mode (first torsional mode) should be central to your analysis. I would expect clustering of results for solvers without torsion degrees of freedom (e.g., purple, green) as distinct from those including torsion (e.g., red, blue, grey, black). However, this clustering is not clearly visible. Is it possible that there has been a mix-up between the 6th and 7th modes in Table B1? For instance, in the 7th mode the aeroelastic solvers without torsion do not appear, whereas they do for the 6th mode. Even if this is not the case, a short discussion of why this expected separation is not observed would improve clarity.
    - There is also a large spread in the higher-order modes. It could be informative to explicitly separate the results into torsion and no-torsion degrees of freedom. For example, ElastoDyn does not include torsional degrees of freedom, so one would expect systematic differences relative to models that do. The same applies to H2-PTNT.

Sources

- The reference *"Hansen, M. (2015). Aerodynamics of wind turbines. Routledge."* should be adapted to match the citation style required by WES. At the moment, the publisher and format do not appear consistent with standard referencing. Please check the journal guidelines to ensure proper formatting. For example, depending on the required style, the same reference would look as follows:
    - APA Style: Hansen, M. O. L. (2015). *Aerodynamics of wind turbines* (3rd ed.). Earthscan.

    - Harvard Style: Hansen, Martin Otto Laver, 2015. *Aerodynamics of wind turbines*. 3rd edn, Earthscan, London, UK.
    - Chicago Style: Hansen, Martin O. L. 2015. *Aerodynamics of wind turbines*. 3rd ed. London, UK: Earthscan.
    - Vancouver Style: Hansen MOL. *Aerodynamics of wind turbines*. 3rd ed. London, UK: Earthscan; 2015.

— Conclusion and a Note —

Conclusion: The manuscript demonstrates a strong understanding of aeroelastic wind turbine simulations and clearly represents a substantial amount of work. At the same time, there are a number of methodological and presentation aspects that require improvement before the paper can be considered for publication. For this reason, I must recommend a major revision. To make the work more robust and easier for readers to follow, I suggest paying particular attention to:

- **Completeness:** Clearly indicate how the in-house solver has been validated in prior literature; include the 5MW results that are repeatedly discussed; note limitations and differences when comparing with *Bernardi et al. (2023, 2025)*; and provide justification for the chosen domain size and boundary conditions (either by citing literature or through a convergence study that also covers domain geometry). For the grid convergence study, additional grid levels are needed to demonstrate an actual trend.

- **Details:** Ensure correct and consistent citations, apply a uniform style for captions, use consistent terminology for forces (e.g., flapwise/edgewise vs. normal/tangential), and add legends where missing.

- **Scientific phrasing:** Avoid turning unvalidated observations (e.g., comparisons between 5MW and 15MW turbines with different inflow conditions) into firm statements. Whenever possible, support claims with quantitative or qualitative measures rather than subjective wording.

- **Critical reflection:** The reduced wake recovery attributed to tower presence is unexpected in light of prior literature. This result should either be further examined, discussed in more detail, or acknowledged as uncertain rather than used as the basis for speculative conclusions.

- **Conciseness:** At 31 pages, the paper is longer than necessary for the number of findings presented. For example, the introduction (2.5 pages) and theoretical background (4 pages) could be shortened, with common equations (e.g., Navier–Stokes or ALM) omitted. Tables and figures that add little to the analysis (e.g., Table 1, Figures 6, 9, 11, 13) could be streamlined or integrated into the text. A tighter discussion that focuses on observations and plausible explanations would also improve readability.

Personal note: I realize my comments are detailed and may sound strict, but they are not meant to be discouraging. On the contrary, they reflect the high potential I see in your work. Addressing these points will make the paper more rigorous, credible, and citable. You are very close to a strong and notable publication