
Revision for Wind Energy Science

Aeroelastic Validation of ExaWind for the Pazy Wing Wind Tunnel Experiment

---

May 31, 2026

Dear Profs. Aubrun and Uzol,

We would like to thank you for processing our manuscript so quickly and for obtaining two helpful and insightful reviews. We have addressed all comments, and have noted the resulting modifications to the manuscript. In the following pages, we provide inline responses to the comments from the reviewers.

If you or the reviewers believe the comments have not been satisfactorily addressed, we humbly request clarification and the opportunity to readdress them.

We believe the feedback from the reviewers and the resulting changes have made the manuscript stronger, and we are grateful for the time you and they have invested into this manuscript.

While performing the additional analyses requested by Reviewer 1, we uncovered some anomalous behavior from the structural model that we believe warrants further examination. If possible, we humbly request additional time to scrutinize the structural model and perform additional validation of said model for inclusion in this manuscript. We believe this additional work would dramatically strengthen the impact and appeal of this paper. Ideally, these additional analyses would be included before the second pass by the reviewers. However, we understand that this may not be possible, and if so, we hope that the attached responses sufficiently address the concerns of the reviewers.

Best regards,

Neil Matula nmatula@sandia.gov	Gopal Yalla gyalla@sandia.gov	Nathaniel deVelder ndeveld@sandia.gov
Lawrence Cheung lcheung@sandia.gov	Paul Crozier pscrozi@sandia.gov	

Sandia National Laboratories,
Albuquerque, NM

Bumseok Lee Bumseok.Lee@nlr.gov	Ganesh Vijayakumar ganesh.vijayakumar@nlr.gov	Michael Sprague michael.a.sprague@nlr.gov

National Laboratory of the Rockies,
Golden, CO

Reviewer 1

The paper is well written and addresses the relevant scientific questions of the accuracy of FSI style modeling of slender wings. The methods applied are not necessarily new, but the validation of the open source ExaWind software stack is relevant for the wind community and should have a broad international interest.

The paper is well structured, and the objective of documenting the capabilities of predicting stability of slender wings is clear. The scientific approach is generally well described, and only minor details are unclear.

We thank the reviewer for the kind words, and for taking the time to review our manuscript and provide helpful suggestions. The manuscript changes made for this reviewer are shown in blue.

In the present work, some approximation of the actual wind tunnel geometry and the wing setup is done. The present reviewer considers the approximations acceptable but is worried that neglecting the tip bar attached to the wing in the CFD setup might have some relevant influence on aerodynamics. It would be good with some discussion on this. One could make a very simple mesh around the tip bar and investigate the aero during a forced motion to evaluate the approximation.

The reviewer makes an important point: while the tip weight added in the experiment was likely intended to have a minimal aerodynamic impact, it is indeed not guaranteed that its absence will not affect the simulation results. However, the complex geometry would almost certainly require an unstructured tetrahedral mesh to maintain acceptable element quality, and at present, mesh motion is only available in the Kynema (formerly known as ExaWind) suite for hexahedral meshes. This is an important limitation of the software, and discussion was added to Section 3.2 to make this limitation explicit. However, we believe that the absence of the tip weight from the CFD mesh is a reasonable approximation, due to the following arguments:

- The tip weight has a slender "stinger-like" shape. While it is not explicitly said in the original paper by Avin et al., it was likely designed to have a minimal aerodynamic impact, instead serving simply as added mass.
- As far as the authors can tell, none of simulation results contributed to the workshop included the aerodynamic effect of the tip weight.

We realize, however, that these points do not guarantee a minimal aerodynamic impact. Because of this, we intend to investigate the aerodynamic impact of the tip weight as future work.

As both eigenfrequencies and twist angles are available in the experiment, the present reviewer finds it relevant to compare these, especially as modern wind turbines have considerable torsional degree of freedom. The comparisons with measurements are well discussed but should include the above discussion of twist and frequencies. I don't expect that including the twist information will change the clear conclusion that the FSI can reproduce the exp.
...
Mandatory changes:
Include comparison of eigenfrequencies (maybe a table) and twist angles.

This is a fair point. While the torsional deformation and frequency response were not quantities of interest for the workshop, they are indeed available from the original experiment. We have added comparisons and discussion in Section 3.

The main shortcoming of the paper is the references, where essentially only references withing the ExaWind community are given. It would be appropriate to position the work in context to other publications of blade resolved aeroelasticity in wind energy, and other codes used in the wind energy community, as this is the intended application.
...
Mandatory changes:
The work should be positioned in an international context; there are several international groups among others at Univ. Stuttgart, DLR and DTU Wind and Energy Systems who have worked intensively on FSI of blade resolved flows.

We agree that the previous work on this case is underrepresented in the manuscript. We have added a substantial number of references to the introduction in an attempt to correct this. If the reviewer believes that some important research groups are still underrepresented in this section, we humbly request additional clarification on the missing work and the opportunity to correct this.

It is mentioned in the text that the structural model is sub-stepping within each CFD time-step. It would be appropriate to mention how the loads are estimated at the intermediate time steps on the structural side.

We have added discussion on the sub-stepping scheme to section 2.4.

Fig 4, Page 11: In my pdf the time step is shown as 10^5 *not* 10^{-5} *as intended etc.*

We thank the reviewer for catching this. This appears to be a typo present in the preprint, and we will be certain that it is addressed moving forward.

Reviewer 2

The paper "Aeroelastic validation of ExaWind for the Pazy Wing wind tunnel experiment" describes in detail how the simulations have been carried out, which is more than sufficient for others to do the same thing including all the assumptions and simplifications that have been made.

We thank the reviewer for the kind words, and for taking the time to review our manuscript and provide helpful suggestions. The manuscript changes made for this reviewer will be shown in magenta.

This description takes 12 pages, but the final results (including the conclusions) need one page of written text and two plots. Here the paper lacks of discussion and interpretation of the results, while the conclusions read more like a summary than a conclusion part. I will try to explain what I think is missing.

We agree with the reviewer that the Conclusions section is indeed rather minimal. We have expanded it, in part using the feedback from both reviewers.

Figure 5 shows the results for the ExaWind (two versions) simulations, the experimental results and the results from a workshop. One 1/3 page the authors describe what is seen in that figure without trying to explain what can bee seen. Here the experimental results are shown as lines but in the text they mention that the differences between the two ExaWind simulations are small compared to the uncertainties of the measurements. If the uncertainties are so significant, it would be good to add them to the plot so the reader can also see them. If these uncertainties are so big, what is the goal of comparing the results to the mean values? Where are the uncertainties coming from. What determines a "good" result?

The reviewer makes a fair point: the authors' use of the word "uncertainty" was rather loose in this context, and we did not explain it sufficiently. We will attempt to explain below what our intent was in using this language, as well as update the manuscript to make this more clear.

In Figure 5 we show, for comparison, the results of two different wind tunnel experiments from Avin's original paper. The solid curve represents the experiment in which the root angle of attack was fixed, and the air velocity was varied across the shown range. The dashed curve represents the experiment in which the air velocity was fixed, and the root angle of attacked was varied. (Since these plots are for three individual AoAs, rather than a continuum of AoAs, we sampled the experimental results at the three AoAs of interest to add to this plot.) These two experimental setups are both practical ways of assessing the effect of AoA and wind speed on the wing deformation in a reasonable amount of time. However, they also both have the unfortunate side-effect of introducing temporal effects into the results, since as a sweep is performed through either AoAs or wind speed, the results at any state may be influenced by previous states. Hence, while the results should ideally only be functions of the AoA and wind speed, they are unfortunately also influenced by the choice of which parameter to vary, and how quickly that parameter is changed. To their credit, Avin et al. included results from both parameter sweeps. Since these two experimental setups are both imperfect ways of assessing the same quantity of interest (the deflection as a function of AoA and wind speed), we can use the differences between them as a proxy for the "uncertainty" in the measurement. Hence, when we say that the differences between our two sets of simulation results (with and without the transition model) are small compared to the uncertainties in the experiment, we mean small compared to the differences between the two parameter sweeps in the experiment.

We agree that this was not clear in the original manuscript, and we hope that the revised manuscript has addressed the reviewer's concern.

In the text the authors also mention, that the experimental results depend on the direction of the sweep – for which sweep direction are the experimental results, which are shown here? Why did the chose this direction of sweep for the comparison? What is the difference in terms of aerodynamics between the sweep in the experiments compared to the fixed angle fixed wind speed in the simulations? When changing the AoA dynamically, like in the AoA sweep, results depend on the rate of change of the AoA. This is not specified or even mentioned. The same might also be true for the velocity – the experimental conditions should be mentioned in more detail and also be considered in the discussion of the results.

We agree that this was not adequately explained in the original manuscript. We will attempt to explain more clearly below, as well as update the manuscript to make this more clear.

The first paper on the Pazy wing experiments, written by Avin et al., was primarily focused on static deflections. In this paper, the AoA sweeps were performed by starting at zero degrees, rotating first in the negative direction, then back to zero, then the positive direction, then back to zero. A small hysteresis effect is seen in these results; however, the digitized samples that we used for this paper represent the midpoint of the hysteresis loop. For the velocity sweeps, the values seen in Avin et al. are only in the positive direction (i.e. increasing velocity). However, the followup paper, Drachinsky et al., was more focused on the flutter boundaries, and this paper is where the flutter boundary comparison data used for the workshop originated. In that work, velocity sweeps were performed in both directions to assess the flutter onset and offset, and there is a substantial difference in the onset and offset speeds depending on the sweep direction. As the authors discuss there, prediction of offset is much more difficult than onset. The workshop used the data for flutter onset with increasing velocity as their benchmark data for the lower flutter boundary, and our comparison followed their example. While we were not able to find an exact reason in the workshop paper (Collaborative Pazy Wing Analyses for the Third Aeroelastic Prediction Workshop) for this choice, we suspect it was because a) it more closely fit the analytical comparison presented in the Drachinsky paper, and b) a real wing will generally encounter the increasing velocity instability before the decreasing velocity one. However, for our simulation results (and those presented in the workshop), no sweep through air velocity was performed. For our simulations, the AoA and velocity were fixed, and the initial transients either grew in magnitude or decayed.

We have updated the manuscript to more clearly communicate the conditions of the experiment and the choices of conditions for the simulations.

The results from the workshop are only shown as a grey area – how many different simulations contributed to these results? Are there others that perform maybe even better than the ExaWind ? If yes, what could be the reason for it. What are the known limitations that others should take into *account before using this method? Why not only showing the best results from the workshop and see how they compare to the ExaWind? What is the novelty with ExaWind that people should use that even though others might perform better? Right now, it looks to me that there is just another method that can be used and I don't think that this is the take home message.*

It is true that the workshop results are represented somewhat crudely by the grey band here. This is because it is difficult to separate the curves when digitizing the workshop data, due to the sheer number of curves present. We have recently requested the exact datasets presented in the workshop, and will update the plots with all the simulation curves if we are able.

It is also true that, among the lower-fidelity simulation results presented for the workshop, it is likely that some more closely match the reference data than Kynema/ExaWind. However, this is quite normal. Indeed, a surprising fact is that low-fidelity computations often more closely match experiments than high-fidelity ones. Part of the reason for this is that simpler models are tuned to match the data at hand. A common example of this would be any calculation which requires airfoil section coefficients, because those coefficients must be tuned to match either wind-tunnel data or more sophisticated aerodynamic calculations. Hence, the low-fidelity calculation is not really making a "blind prediction", since it leverages either experimental data or high-fidelity methods to make its predictions. Another reason low-fidelity methods often more closely match experiments is that those experiments are designed to be suitable for low-fidelity methods. For example, the Pazy wing experiments employed small root angles of attack, such that many low-fidelity methods (such as panel methods) could reasonably hope to capture the flap-wise forces. However, if highly separated flows were present, we would not expect good predictions from these methods. A third and final reason is that experimental data is often incomplete, and can contain small errors. To achieve the best possible predictions, the highest-fidelity calculations would often require more information about the conditions, and more accuracy in that information, than a real-world experiment can be expected to provide. Because of this, low-fidelity methods often better match experiments partly by chance.

For these reasons, we do not expect unprecedented agreement between the predictions made by Kynema/ExaWind, or any other high fidelity calculations, and the Pazy wing experiments. Indeed, if we had achieved a near-perfect match with the experimental results that was clearly superior to all competitors, we would have been forced to attribute that feat, at least partly, to good fortune. Instead, the goal of the paper is much less ambitious: to use the Pazy wing case to validate our predictive capability for FSI problems. We expect our results to be reasonably close to those of the experiments, and we expect them to be competitive with other methods available today. In showing the simulation results contributed to the workshop, we aim to show not necessarily that our results are superior, but only that they are comparable.

In figure 5 the x-axis is in dynamic pressure instead of wind speed in m/s – why?

The experimental static deflection results from Avin et al. were presented with dynamic pressure on the horizontal axis, and the workshop comparisons kept this convention, likely to make it easy for the reader to compare the publications. We decided it was best to retain this convention for this same reason.

The situation for figure 6 is somehow similar. Here the range is much wider compared to the *workshop and the stable condition for AoA of 3° ist missing. Was ist not possible to run another simulation after these?*

The range is indeed wider than the workshop data. However, this is due to the way flutter is being characterized in this work. Our understanding is that the simpler methods used in the workshop often permitted a kind of linearized analysis to determine the flutter onset boundary in an inexpensive fashion. In contrast, here we simply used individual simulations, which take a non-negligible time to run, and examined the long-term behavior of those simulations to see whether they demonstrated a growing instability.

Regarding the fact that no "unstable" (i.e. borderline) case was identified for 3 degrees, we could indeed run additional simulations to more precisely bracket the flutter onset boundary. However, the precise boundary found in those simulations will likely be sensitive to the design choices of the simulation, including the initial conditions, iterative convergence, and so on. Hence, a very precisely bracketed flutter boundary would somewhat overstate our claims about how robustly we can hope to predict this boundary with a high-fidelity time-dependent simulation. Instead, we opted for a much more modest claim: we were able to show that a flow velocity that is a few percent less than the flutter boundary is clearly stable in our simulations, and one that is a few percent higher is clearly displaying flutter. The borderline condition ("unstable") was simply intended to be another characterization of the flutter boundary, and not a precise prediction. The core result is simply the interval between the clearly stable and clearly fluttering wind speeds.

In figure 6 the abbreviations for the methods are different than before which makes it harder to understand whats going on. Why is the flutter region also added when only the onset of flutter is important? What about uncertainties in these results?

It is true that the labels for the near-flutter simulations are new to that section, and are not entirely intuitive, as they are specific to how we characterized the flutter boundary and are not used by other authors. We believe that, while imperfect, these labels are the best choice to represent their corresponding conditions. However, we would welcome suggestions for alternative labels.

Regarding the presentation of the flutter region in Figure 6, this was done for two reasons:

1. the inclusion of the flutter region allows the reader to easily orient themselves when comparing with Drachinsky et al. and the workshop results, and
2. the presented data points that are clearly fluttering were selected to be well inside the flutter region. Without the entire flutter region present on the plot, the reader could not be certain that the fluttering cases were beneath the experimental offset boundary (i.e. not too far to the right).

Regarding uncertainties in the results, the interval between our "stable" and "flutter" cases can be interpreted as a proxy for uncertainty in our results. We have updated the manuscript to reflect this. However, the uncertainty in the experimental values is more difficult to pin down. One could claim that the differences found in the onset and offset boundaries depending on the direction of sweep could serve as a proxy for the uncertainty. However, neither the authors of the experimental papers nor those of the workshop paper make any claims about the uncertainty in the onset boundary. Given that this clean presentation has been used in comparisons of simulation results in the past (i.e. the workshop results), we would prefer to retain it here rather than attempting to add experimental error bars based on the sweep direction. However, if the reviewer feels strongly that they should be included, we will be happy to do so.

I would highly recommend to extend the results part and add discussion.

We agree that the conclusions in the original manuscript are somewhat simple. We have added additional discussion there, and have also additional data to the results section with additional comparisons to the experiments.