

Review of the manuscript wes-2025-77, "Wind Tunnel Load Measurements of a Leading-Edge Inflatable Kite Rigid Scale Model" by J.A.W. Poland, J.M. Spronsen, M. Gaunaa, and R. Schmehl.

## Referee #1

Dear reviewer,

Thank you for providing detailed feedback on our manuscript. Following your global feedback, we have copied your comments from the annotated PDF and are answering those point-by-point.

Best regards,

Jelle Poland, on behalf of the authors

### Global feedback

This paper describes wind tunnel load measurements for a leading-edge inflatable kite.

- Scientific relevance: This is a relevant and important paper, providing a validation data set for numerical simulations.
- Scientific quality: The measurements are carried out carefully and they are well documents.
- Presentation quality: The paper is well-written and clear, logically structured and generally well-explained.

Some specific comments and questions can be found in the attached file.

**[Line 6.] This sentence is not clear to me. Does the "design geometry" mean the "deformed geometry"? If so, how do you know what the "deformed design geometry" is? If not, reword the sentence because it is not logical as it is.**

Good comment. We have changed the abstract accordingly.

**[Line 8.] I would quantify the ranges here**

The ranges have been included. We have changed the abstract accordingly.

**[Line 10.] Does that mean that the stitching was not present in the scale model, or why was the tape needed?**

Indeed, the stitching seam was not present on the scale model. The tape was applied to analyze the effect of a tripped boundary layer, which the stitching seam in reality may cause.

**[Line 13.] For the same geometry?**

Yes, for exactly the same geometry. We have changed the abstract to make this clearer.

**[Line 15.] Was anything learned about the aerodynamic performance of the kite?**

Yes, we altered the abstract to include more of this.

**[Line 24.] I'm not sure that I'd call the power output "continuous" given that the kite is being reeled in and out. Maybe "more continuous"? Or perhaps there's a better term.**

Excellent point. This is indeed confusing. We have changed the manuscript accordingly.

**[Line 25.] in the range of?** (context is: "wing pitched to a high angle of attack.")

This is very system-dependent, but for the V3 kite, it is roughly 8deg. As stated by

<https://wes.copernicus.org/preprints/wes-2024-182/> and cited later in the text on line 336.

**[Line 26.] of?** (context is: "maximum length")

This is also very system-dependent, even site-dependent -obstacles (trees, restricted airspace) may restrict certain ranges, and even wind condition dependent, i.e., it might be more favorable to fly higher/lower depending on the day. So it is hard to give this a number.

**[Line 33.] A close-up might be useful** (context is: The tubular frame of the wing consists of an inflatable leading edge tube and several connected inflatable strut tubes.)

The wing with leading-edge tube and connected inflatable struts is illustrated in Figures 1a and b, and an even closer view is provided in Figure 3, both figures clearly showing the leading-edge tube and struts.

**[Line 42.] Not sure exactly what you mean by that. Do you mean the design? Could you explain briefly?** (context: "aerodynamic system identification")

Aerodynamic system identification refers to figuring out what aerodynamic properties a given system has. In this context, once a kite has been built and is flown, the question often arises what the aerodynamic properties actually are, e.g., what is the lift-to-drag ratio of the deformed kite in flight?

**[Line 49.] How is it more scalable than field measurements? Because the industrial-scale kites mentioned above are still not "full" scale?**

The field experiments referenced above were conducted at full scale. However, experimental characterization of aerodynamic properties faces fundamental scalability constraints as system size increases—larger operational AWE systems require increasingly complex and costly flight tests due to expanding demands for measurement equipment, safety parameters, insurance, permitting, and so on. In contrast, once validated computational models are established, determining the aerodynamic characteristics of a 10MW system requires only marginally more computational resources than a 1MW system, while equivalent experimental characterization would demand orders of magnitude more

instrumentation and operational complexity. This computational advantage extends beyond cost to encompass time and feasibility, enabling evaluation of aerodynamic performance across different operating conditions and system configurations that would be prohibitively expensive or logistically impossible to measure experimentally.

**[Line 51.] Simulations are also limited by computational power, leading to simplifications such as RANS, as well as to computational stability. I would suggest adding this point to strengthen your argument for wind tunnel tests. This would allow you to argue that wind tunnel tests are important not ONLY for validation but also for controlled parametric testing.** (context: However, simulations necessitate validation, which is best achieved through wind tunnel testing that allows precise control of inflow conditions.)

Excellent point, thanks for the suggestion. We have changed the text and included the argument.

**[Line 58.] "downscaling" or "scaling down"**

We changed scaling to down-scaling.

**[Line 60.] and structural**

I added this to the sentence.

**[Line 66.] Define what this means** (context: anhedral angle)

We have changed the manuscript to include the explanation of the anhedral angle (opposite to dihedral), to include the following: *"..defined as the downward inclination of the wings relative to the horizontal plane when viewed from the front.."*

**[Line 67.] The results showed?** (context: Wind tunnel experiments using rigid kite models eliminate the aeroelastic scaling issues and provide aerodynamic data with a high degree of certainty on the inflow. Belloc (2015) presented wind tunnel measurements of a 1:8 scale paraglider model in which the anhedral angle follows an elliptical shape when viewed from the front, and it incorporates a spar made of a wood-carbon composite sandwich. During the tests, inflow velocities reached  $40 \text{ ms}^{-1}$ , corresponding to Reynolds numbers of  $9.2 \times 10^5$ . The experiments covered angles of attack ranging from  $-5$  to  $22^\circ$  and sideslip angles from  $-15$  to  $15^\circ$ )

We think you are asking about what the results of this wind tunnel study showed. Which is a good suggestion, to add the discussion. We have therefore altered the text and included the additional sentences to elaborate on what the results showed.

**[Line 80.] Stick to the past tense here.**

*We have changed the manuscript accordingly.*

**[Line 90.] first discusses**

We have changed the text accordingly.

**[Line 105.] why?** (context: This geometry differs)

We altered the manuscript to explain this in more detail.

**[Line 107.] Can this difference be quantified?**

Yes, it comes down to less than 1mm change in the global properties. We have altered the text to illustrate this.

**[Line 108 (figure caption).] This word should be at the end of the sentence**

We have altered the text accordingly.

**[Line 109.] How high was the blockage and why did you choose that amount?**

The blockage amount is detailed in Appendix A, in the discussion of the wind tunnel corrections. With the kite set at the maximum tested angle of attack of 24 degrees, the projected frontal area ( $A_f$ ) at  $\alpha = 24$  degrees is approximately 0.2m<sup>2</sup>. The octagonal wind tunnel opening has an area 7.47m<sup>2</sup>, resulting in a blockage factor of 3%.

We chose a 1:6.5 scaling as it was a compromise between having a large model, good for achieving higher  $Re$ , and having a small model, good for limiting blockage effects and low manufacturing costs.

**[Line 118.] alpha should be marked on Figure 4**

We added the angle alpha to this figure, in blue.

**[Line 125.] Should be marked on Figure 2**

We added the angle beta to Fig. 2

**[Line 128.] How were these particular values chosen?** (context is the measurement matrix in table 2)

We chose a range of values representative of kite operations when harvesting energy. For the angle of attack, different values were selected prior to the experiment. But the values have changed as the 6.3deg correction to the angle of attack was adjusted post-experiment upon re-re-re-

measuring the experimental setup. Furthermore, the reported angles of attack include the wind tunnel correction, which was also determined post experiment.

For the side slip (and in fact also for the angle of attack *range*), the angles chosen were first of all chosen to be representative of real flight. Secondly based on the available CFD data range, which is from 0 to 12deg, and thirdly it was decided to measure a larger range to provide validation data that could be used to analyze the extreme cases of up to 20deg.

The inflow velocity was determined to match Reynolds numbers encountered in flight. With the kite flying at approximately 30m/s during flight, and a chord length of 2.6m, we find a Reynolds number of approximately  $5e6$ . This is not achievable within the current wind tunnel setup, due to inflow velocity limitations (but also due to the structural vibration limitations). Therefore inflow velocities were selected that were as high as possible, to get as close as possible to this Reynolds number. The  $Re = 5e5$ , was furthermore selected as CFD data was available at that Reynolds number

**[Line 130.]  $c_{ref}$  should be defined and marked on a figure**

$c_{ref}$  is defined in Table 1 and the  $c_{ref}$ ,  $h$ , and  $w$  have now been added to Fig. 3 for extra clarity

**[Line 136.] Because?** (context: interference effect of kite and support-structure)

Because of their distance, and it not being measurable within the current campaign. One can only measure the support structure's own aerodynamics, and subtract this from the kite and support structure aerodynamics. But the effects of the interaction between the two cannot be captured using the presented setup.

**[Line 138.] And corrected for?** (context: sensor drift was analyzed)

The sensor drift is accounted for as a 'zero-run', a run without wind was conducted before every new measurement set. The sensor-drift present, and captured by measuring the static load, is in this way inherently counted for.

**[Line 142.] Is there a reason you write  $200 \times 10^5$  here rather than  $2 \times 10^7$ ?**

Yes, it was written to stay with the  $10^5$ . But we have changed it following your suggestion to  $2 \times 10^7$ .

**[Line 144.] I would put this at the end of the sentence**

Good idea, this has been altered in the text.

**[Line 147.] Specify that you mean the height threshold here (I was confused at first)**

Incorporated, we added 'height'.

**[Line 158 (in figure) ] mm & not minus** (context: no units are shown in Fig. 6)

The units were added correctly to the legend of Fig. 6. Thanks for the suggestion.

**[Line 171.] And was measured how, and how often?** (context: the density measurements)

We have changed the text to clarify this: “*the measurements were non-dimensionalised using the air density  $\rho$ , which was measured continuously and documented at each measurement point, ...*”

**[Line 173–186.] 3 suggestions to change “are” into “were”**

Yes indeed, this should be past tense. Thanks for the correction; it has been implemented.

**[Line 190.] Did you check how good this fit was?** (context: Missing data points were determined by interpolation, which was carried out by fitting two linear segments from the minimum to the mean and from the mean to the maximum, respectively)

This was done, yes. We checked the respective fit with either a single linear line, two linear lines or a parabola. A single linear line clearly did not capture the measured trends, e.g., lowest drag at a medium angle of attack and an increase with large negative and large positive. The parabola seemed to cause problems towards the ends of the measured range; hence, we decided upon using two linear lines.

**[Line 219.] Define & You define  $i = 1, 2, 3$ , etc. but then in Table 3,  $i$  seems to have values of L, D, and S. This needs clarifying.** (context: of the coefficient of variation)

This has been addressed by clarifying that  $\sigma$  represents the standard deviation, and the represented  $\overline{\sigma}$  represents the averaged standard deviation, and that  $\mu$  represents the mean. The  $i = 1, 2, 3, 4$ , etc. have been replaced with the respective force and moment coefficients.

**[Line 224.] Doesn't CV just decline with increasing  $Re$  because the mean increases??**

You are correct that a lower value of CV—the ratio between standard deviation and mean—does not necessarily indicate a lower standard deviation, i.e., a lower absolute uncertainty. Rather, it reflects a reduction in relative uncertainty. We realised this was not stated clearly in the original text and have adjusted the sentence accordingly.

**[Line 225.] Do you mean they generally exhibited lower relative uncertainty in this work, or they generally do? In which case, how well does that match in this work?**

This was not clearly stated. Thanks for the comment. We adjusted the text accordingly.

**[Line 227.] Are you saying that low CV = low uncertainty = high precision? I'm not sure that this is entirely true. What about constant offsets?**

You are making a good point. Low values of  $CV_i$  indicate high *precision* (i.e., low relative random uncertainty), but do not address measurement *accuracy*, which may still be affected by constant offsets or systematic errors. Our intent was to characterise the measurement precision; we have clarified this in the revised text.

*w*

**[Line 234.] Why did you choose this condition?** (context:  $\beta = 0$ ,  $\alpha = 9.4$ , for zigzag measurements)

The  $\beta = 0$  is the design condition and was chosen for that reason. The  $\alpha = 9.4$  was not chosen. Similar to the answer to the question on the measurement matrix (“**How were these particular values chosen?**”), it comes down to that the angles of attack were corrected post-experiment, both for the offset with steel supporting rods and the wind tunnel corrections. The intended design value was closer to  $8^\circ$ , the nominal operating condition during reel-out (power production) flight.

**[Line 243. (figure caption)] Which force coefficient is which?** (context subscript labels not visible)

This was a preprint publishing problem of WES and has been resolved.

**[Line 243.] Can you please explain how you came to this conclusion in more detail? It's not that easy to go through the plots and understand that when looking at them for the first time.**

The first sentence refers to Fig. 6. (not 7) Hence it is merely a restatement of conclusions made earlier based on the theoretical trip height calculation.

The second sentence was expanded by an additional explanation that we are talking about the effect of including a zigzag tape.

The third highlighted sentence is clear is fine as is.

To improve readability as requested, the whole part has been adjusted, adding a whole additional sentence.

**[Line 257.] I don't see this, please explain** (context: With increasing  $\text{Re}$ , the measurements show a converging trend and decreasing variation, consistent with the decreasing uncertainty observed in Table 3.)

It lies in the difference between the individual lines, representing different  $\text{Re}$ . These differences decrease with increasing  $\text{Re}$ , and the variation is at its maximum between the  $\text{Re} = 1.3 \times 10^5$  and  $\text{Re} = 2.5 \times 10^5$  cases and the smallest between (the highest)  $\text{Re} = 5 \times 10^5$  and  $\text{Re} = 6.1 \times 10^5$  cases.

The variation part is less clear; this is maybe what you are referring to. This part of the statement is hard to conclude from the figures; hence has been removed from the sentence.

**[Line 258.] Why not show  $CL/CD$ ?**

You are making an excellent point, discussing an increasing aerodynamic performance without clarifying what is meant by that specifically, and if it is CL/CD, then why not show this?

To resolve these issues, we have added CL/CD to the plot and adjusted statements about the aerodynamic performance to accompany the new insights. Lift increases with Reynolds, but drag as well, meaning that CL/CD does not increase with Re.

**[Line 260.] In what? You mean that it jumps around between Re and doesn't have a smooth trend? Please clarify.**

Good point. It would be better to use the term smoothness rather than variation. We have adjusted the text accordingly.

**[Line 261.] Not sure I can see that. You mean because it jumps around more? But does it?**

(context: The CL/alpha plot suggests that, compared to higher Re, there may already be stall development at lower angles.)

No, not because it jumps around more. But because the slope of the Cl-alpha curve starts changing earlier. Where they align nicely up to 8deg or so, beyond this point, the  $1.3 \times 10^5$  shows the first lower slope, indicating less lift generation with more angle of attack, hence the hypothesis of the development of stall cells. We have adjusted the text to make this point clearer.

**[Line 269.] You mean deviations between different values of beta?**

We meant asymmetric or anti-symmetric deviations. Which was not at all clear from the original sentence structure, so thanks for pointing this out. We changed the paragraph accordingly.

**[Line 275.] Define what you mean by performance. If you mean CL/CD, then why not show it?**

Similar to the comment before, we added CL/CD to the alpha sweep, but left it out for the beta sweep and changed the text accordingly to reflect that lift increases.

**[Line 283. (figure 8) ] Please colour these lines in a more intuitive way, i.e., getting darker or lighter as Re increases.**

Great suggestion, this has been done.

**[Line 307.] What does that mean exactly? How do you know that they are more accurate?**

(context: In previous studies, these polars were constructed using aerodynamic load correlations derived from a large set of CFD simulations (Breukels, 2011). In the present work, however, more accurate polars are employed, obtained from dedicated 2D RANS CFD simulations.)

Excellent point. A citation to another paper that is discussing these differences in great detail and introduces the RANS CFD simulations is missing here. We have added it and changed the text accordingly.

**[Line 311.] Which experimental data? And why not say this at the start of the present study in order to justify the chosen slow conditions?**

Good point, we adjusted the manuscript accordingly to include a reference to the experimental [data](#). It was stated in front of the comparison to numerical predictions section, as here it is most relevant to know in which range one wants the predictions to match the wind tunnel conditions.

**[Line 315.] I know the result is the same, but why do you call this L/D rather than CL/CD?**

No specific reason. We have therefore changed it to CL/CD.

**[Line 351–393.] The aim of this work was to provide a data set for numerical model validation. Can you comment on how and if the numerical studies were "validated" in this work? You made comparisons and suggested some reasons for differences, but are the codes now "validated"?**

Another excellent suggestion, we have added the following paragraph to the conclusion section to directly address this point.

## Referee #2

Dear reviewer,

Thank you for providing detailed feedback on our manuscript. Following your global feedback, we have copied your comments from the document, and answered those point-by-point, in the document below.

Best regards,

Jelle Poland, on behalf of the authors

### Global feedback

The study presents the wind tunnel measurements of the aerodynamic loads of a downscaled rigid model of a 25 square meter TU Delft V3 Leading-Edge Inflatable (LEI) kite. The experimental measurements are further compared with the results from numerical methods: particularly RANS results from the literature and VSM method. Authors claim to have observed similar trends and values

in the load data between the experimental data and the results from numerical methods under nominal operational conditions. Confidence intervals are presented for the measured data, and the discrepancies are attributed to the experimental setup in the wind tunnel. Correction of the experimental measurements are performed using standard methods considering blockage, downwash, and streamline curvature. The measured data and used codes are made publicly available through Zenodo.

Introduction along with the literature review is presented in the first part of the study, which is then followed by the description of the experimental setup and downscaled model. Mean and standard deviation of the measured aerodynamic loads as a function of angle of attack and side slip angle, considering uncertainty analysis, boundary layer transition and Reynolds number effects are presented.

The manuscript is well written and organized. The study provides novel experimental data sets on steady aerodynamic coefficients for the downscaled models of LEI kite. **However, the authors do not discuss/acknowledge about the unsteady loads or aerodynamic coefficients, which are responsible for the realistic aerodynamic performance of LEI kite.** In addition, the following comments should be addressed for improving the quality of manuscript.

### **Answer to global feedback**

The purpose of the experimental study was to provide reference aerodynamic data for a fixed wing geometry, mainly for validation of CFD analyses. It was not the purpose to aerodynamically characterize a real kite, that will always deform under load when in flight.

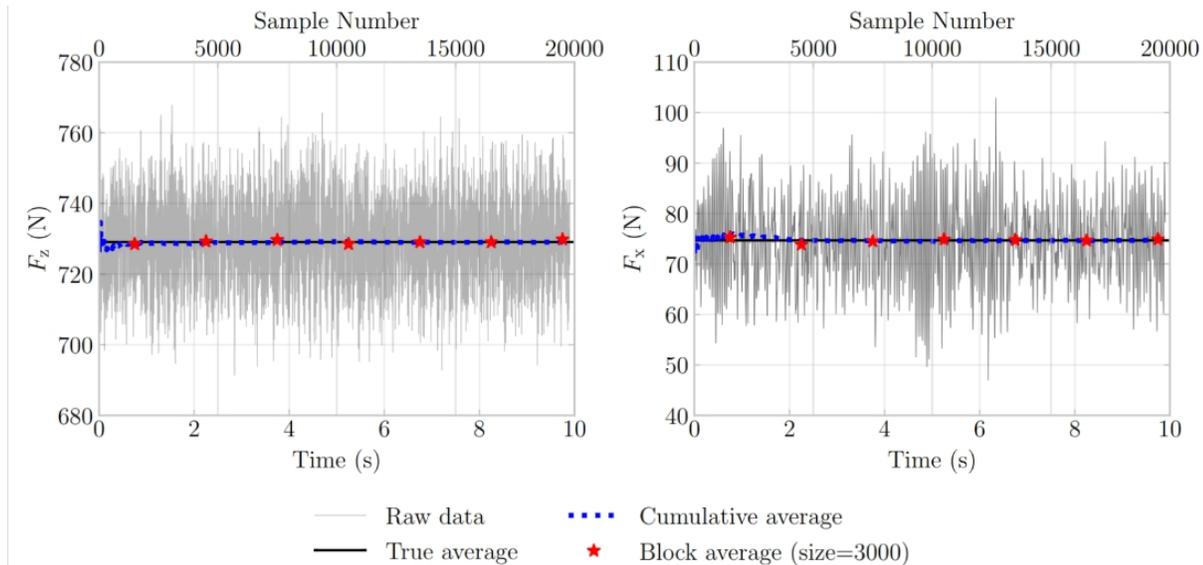
#### **1. Page 8, Line 134: A measurement period of 10 s is used. Is the time sufficient for the convergence of statistics? Discussions on the convergence analysis should be made.**

A 10-second measurement period was selected based on the characteristic aerodynamic time of the system, defined as the time required for a fluid element to traverse the kite's reference chord. For each tested condition, this resulted in 125 to 625 independent flow passages within the 10 s window (depending on the value of the free stream velocity), ensuring that statistical averages are based on many uncorrelated samples rather than being influenced by correlated fluctuations.

To validate statistical convergence, key measurement conditions (e.g.,  $\alpha = 5.7^\circ$  at  $U = 20\text{m/s}$  and  $\beta = -20, 0, 20^\circ$ ) were repeated three times. The close agreement in mean and fluctuating loads between these repeated trials demonstrates that a 10-second interval is sufficient to achieve converged statistics under the present test conditions. For extremely slow or rare dynamic processes, longer sampling periods might be required, but for the steady-state aerodynamic regime investigated here, the chosen duration is well justified.

We also added a figure here, which shows an analysis of a running average and a block analysis. The running average and block analyses show that the fluctuations in the computed statistics over the 10-second window are marginal, confirming that statistical convergence is achieved.

The figure below and subsequent analysis have been added to the manuscript, in Appendix A.



**2. Page 8, Figure 6: Please check the x-label of the figure,  $Re \times 10^{-5}$  or  $Re \times 10^5$ ?**

Thanks for the comment, it should indeed be  $Re \times 10^5$ , and has been corrected.

**3. Page 9, Line 155: where  $U_k$  is the local velocity at the roughness height, which may be approximated by  $U_e$ . Does the presence of a boundary layer near the leading edge affect the approximation of  $U_k$  as  $U_e$ ?**

Yes, the presence of a boundary layer near the leading edge does affect the approximation of  $U_k$  as  $U_e$ . Strictly speaking,  $U_k$  denotes the local velocity at the top of the roughness element, which resides within the boundary layer and is therefore somewhat lower than the local external velocity  $U_e$ . However, in the absence of detailed boundary-layer data and for the sake of a practical engineering estimate, this difference was assumed to be modest, and  $U_k$  was approximated by  $U_e$ . While this approximation simplifies the analysis, it is important to explicitly acknowledge that it constitutes an assumption, and the associated error remains unknown.

To this end, we have altered the part (around line 153) to include this assumption more explicitly.

**4. Page 10, Line 184:  $x_{cg}$  is repeated twice and needs to be corrected with  $z_{cg}$ . Does the change of angle of attack, side slip angle as well as deformation of the model at high test speed affect these values? If it does, how are they corrected?**

Thanks for the comment. This has been adjusted.

To answer the second part of the question:

The values are provided for the  $\alpha = 0$  case and indeed change with varying angle of attack. These changes are considered as the location of the reference point is computed as a function of the angle of attack.

With regards to the changes occurring due to the side slip angle, these are considered in a different fashion. Because the load balance rotates with the setup – i.e. is also placed on top of the rotary table – the whole reference frame rotates, and hence the distance does not actually change with changing side slip angle. This angular change, is later corrected in Step (4) (line 185), by a matrix multiplication with the rotation matrix  $R$ . To make the rotating part a clearer, a sentence was added to the manuscript.

The deformation of the structure is considered negligible. The rigidized carbon-fibre framework, reinforced with steel rods, was specifically designed to minimize deflection under the expected loads. Although absolute zero deformation is impossible, deflections were expected to be extremely small, which was also qualitatively confirmed by video analysis. Therefore, we assume that structural deformation does not affect the values of  $x_{cg}$  and  $z_{cg}$ .

**5. Page 12, Table 3: What is the angle of attack and side slip angle for these measurements?**

These are not for a specific angle of attack or side slip; these were instead computed over a large range of measurement points at the given Reynolds numbers. See preprint lines 220-224.

**6. Page 14, Line 269: Please discuss why  $\alpha = 7.4^\circ$  was selected.**

The measurements were made at  $\alpha = 7.4\text{deg}$  because the V3 kite is known to operate around  $8\text{deg}$  during the reel-out (power production) phase [as stated by <https://wes.copernicus.org/preprints/wes-2024-182/> and cited later in the text on line 336]. One might ask why  $8\text{deg}$  was then not selected, well the offset in angle arises from (a) wind tunnel corrections that were applied in hindsight and (b) the geometric offset from angle of attack to measured steel rods, which was remeasured multiple times post-experiment upon which a different offset angle was found, which changed the measured angles of attack. A sentence was added to the manuscript to introduce the rationale for selecting the  $\alpha=7.4\text{deg}$  case, as it closely approximates the average angle of attack experienced during reel-out.

**7. Page 16, Line 293: ‘.. we applied the area ratio of 3.7 as a correction factor to the Cs reported in Vire et al. What’s the reason for using this specific ratio? It should be clarified.**

The 3.7 ratio, is the ratio between the vertical area used by Vire et al. and the projected area used in this study. This 3.7 factor brings the coefficients from being non-dimensionalized by vertical area, to non-dimensionalized by projected area – consistent with CL and CD. This is explained in line 291, just prior to the is statement, as: “*The reported CS values differ from those reported in Viré et al. (2022) as we are using the platform area A, see Table 1, of the wing for the non-dimensionalization of the side force, as opposed to the projected side area that was used. For this reason, we applied the area ratio of 3.7 as a correction factor...*”

Which, was adjusted in the new manuscript to improve clarity.

**8. Page 16, Line 301: ‘.. contained a 1.02 degree offset in the angle of attack’. How has this been decided? Is it done for the better match of the experimental results with the results from numerical simulation?**

The 1.02deg offset in the angle of attack was not intentionally introduced to adjust agreement between experiments and simulations. Instead, it resulted from an unintentional feature in the original TUDELFT\_V3\_KITE CAD geometry: the vector from the mid-span leading edge to trailing edge was tilted by 1.02deg relative to the intended horizontal plane (see new Appendix F).

This geometric offset went unnoticed and was directly inherited by earlier RANS-CFD studies, such as Vire et al. (2020) and Vire et al. (2022). Consequently, the angles of attack reported in these publications do not strictly match the conventional definition, i.e., angle of attack = the angle between the incoming flow and the chord line.

The issue came to light during recent wind tunnel work. In follow-up discussions with G. Lebesque – whose MSc thesis underpinned the work of Vire et al. (2022)– he confirmed he had not been aware of the offset. This was further verified by inspecting cross-sectional plots: lines through the leading and trailing edges show a clear tilt compared to a true horizontal reference (see image/info graphic below).

To resolve this, we have corrected the geometry to remove the offset and have made the updated files openly available, see: [https://awegroup.github.io/TUDELFT\\_V3\\_KITE/docs/datasets.html](https://awegroup.github.io/TUDELFT_V3_KITE/docs/datasets.html) . The aim is to ensure consistency and accuracy in future experimental and numerical studies using this benchmark. The inconsistencies have also been reported to the international kite simulation community.

In summary: The offset was unintentional and not applied for tuning results. It was confirmed by geometric inspection and discussion with original contributors. Corrected files are now provided to align the CAD with the standard aerodynamic reference frame.

To make this clearer to the reader, an additional appendix F was created.

**9. Page 17, Line 324: In figure 10, the legend does not indicate with and without strut CFD studies.**

Excellent point. We have updated the legend; it is now plotted below and indicates which configurations do not have struts.

**10. Page 17, Line 329: 'All numerical models predict a higher maximum L/D...' Is this because of the difference in Reynolds number between the experimental setups and numerical simulations?**

No, we don't think so. There is only a difference in Reynolds number between one of the numerical CFD studies. But between the other numerical CFD studies and the numerical VSM study, the differences are very marginal, as they all round off to  $5 \times 10^5$ .

**11. Page 21, Line 377, 'the observed agreement suggests there is an aerodynamic potential within the presented numerical models...'. However, the experimental results show that there is a disagreement after a certain range of angle of attack and side slip angle. The statement should be justified with details.**

The statement is about the side force, where the disagreements arise mainly for side slip angles larger than  $\pm 10^\circ$ . This is stated in line 375, as a precursor to this statement, and hence is already justified with details.

**12. Page 28, Line 523: Comments on the sources of vibrations (resonance, vortex shedding) will make the manuscript more robust.**

Thanks for the suggestion. We added the following to make the manuscript more robust: "*Potential sources of the observed vibrations include structural resonance, wherein the natural frequencies of the experimental setup are excited by unsteady aerodynamic loads, and vortex shedding from the model or its mounting components, which can introduce periodic forcing. Both mechanisms are known to amplify dynamic responses in wind tunnel experiments, particularly at elevated angles of attack and higher wind speeds.*"