Rebuttal of: "Flow Field Analysis of a Leading-Edge Inflatable Kite Rigid Scale Model Using Stereoscopic Particle Image Velocimetry"

Jelle Agatho Wilhelm Poland, Erik Fritz, and Roland Schmehl

1 General comments (RC1)

> **The manuscript presents a high-quality experimental study on a 1:6.5 scale LEI kite using stereoscopic PIV, providing a benchmark dataset for validating CFD and Vortex-Step Method models. The work is scientifically significant, as it elucidates local 3D aerodynamic phenomena, including strut-induced flow effects, and directly confirms stall onset at high angles of attack. Methodology and analysis are rigorous, and results are clearly presented. Minor technical revisions are suggested to improve clarity, including sentence refinements, explicit description of pseudo-2D assumptions in force calculations, and clarification of some PIV limitations. Overall, the manuscript represents a valuable contribution to the field of kite aerodynamics and is recommended for acceptance after technical corrections.**

Dear reviewer,

Thank you for your constructive assessment and for recommending the manuscript for publication after technical corrections. We appreciate your positive evaluation of the experimental methodology, the clarity of the presentation, and the value of the dataset for validating CFD and vortex-step-method models.

We have implemented the suggested revisions to improve clarity and transparency. Specifically, we refined a number of long sentences for readability, added an explicit description of the pseudo-two-dimensional assumptions underlying the sectional force estimates, and clarified the principal PIV limitations and their implications for interpreting gradient-based and drag-related quantities. Answers to the technical corrections are shown below in Sect. 1.1. All changes are documented in the attached tracked-changes manuscript.

We look forward to publishing this paper, which has greatly improved with your feedback.

Best regards,
J.A.W. Poland, on behalf of the authors

1.1 Technical corrections (RC1)

> **Split long sentences in Discussion and Conclusion for readability for example:**
>
> - **“Characterizing the aerodynamics of LEI kites with numerical prediction and experimental measurement poses several challenges, owing to the highly flexible nature, pronounced anhedral and sweep, and unconventional airfoil geometries.”**
> - **"The combination of increased downward, sideways, and upstream flow near the strut suggests the presence of a tilted or angled vortex structure.”**
> - **“Measurement quality could be further improved by employing a narrower laser light sheet to concentrate laser power, potentially reducing reflection intensity.”**

We agree, and we have addressed this throughout the manuscript. Several long sentences (including the three examples highlighted by the reviewer) were split into shorter, clearer sentences to improve readability and reduce syntactic complexity. These edits are best reviewed in the tracked-changes file.

> **"explicit description of pseudo-2D assumptions in force calculations"**

The manuscript has been revised to incorporate this description. In particular, we now state explicitly that the force estimates are *sectional, pseudo-two-dimensional loads per unit span* obtained from a single measurement plane, and that the momentum balance neglects out-of-plane fluxes. As a consequence, spanwise force transport and three-dimensional coupling effects are excluded by construction, reducing accuracy where out-of-plane components become significant (notably near the wing tip and strut junctions). This clarification has been added to Sect. 2.7 and is reflected consistently in the surrounding text.

> **"and clarification of some PIV limitations"**

We have clarified the PIV limitations more explicitly in the revised manuscript by (i) expanding the uncertainty analysis, (ii) identifying the dominant limitation mechanisms, and (iii) stating how these limitations constrain the interpretation of derived quantities.

First, the uncertainty section was expanded to distinguish Type-A (statistical) from Type-B (systematic and modelling-related) contributions following common PIV uncertainty practice (Sciacchitano and Wieneke, 2016). The reported uncertainties of the mean velocity are now presented strictly as Type-A convergence metrics, whereas systematic effects are discussed separately and are not conflated with statistical confidence intervals. This reduces the risk of using precision metrics as a proxy for overall measurement fidelity.

Second, the revised manuscript identifies the dominant experiment-specific limitations and their consequences for the resolved flow field. In particular, reflection-induced correlation loss and the progressive increase in measurement-plane misalignment towards the wing tip are stated explicitly as the primary constraints on interpretability. Their impact is quantified through plane-wise data-quality metrics (fraction of rejected and additionally masked vectors) and is linked directly to regions where near-wall information and velocity gradients become unreliable.

Third, the revised manuscript clarifies the implications for downstream analysis. At the start of the Discussion, a framing statement has been added noting that interpretation is primarily constrained by reflection-driven data loss and increasing plane–surface misalignment, which disproportionately degrade near-wall gradients and, consequently, gradient-sensitive and momentum-based sectional quantities. The detailed mechanisms, quantitative breakdowns, and their influence on circulation and pseudo-two-dimensional force estimates are then discussed in the dedicated uncertainty and Discussion subsections, with supporting material provided in the relevant appendices.

2 General comments (RC2)

> **The paper "Flow Field Analysis of a Leading-Edge Inflatable Kite Rigid Scale Model Using Stereoscopic Particle Image Velocimetry" describes an experimental study of a rigid, downscaled model of a flexible kite with an inflatable leading edge. A large amount of stereoscopic PIV measurements are conducted at various locations around the kite model and for two different inflow angles in a wind tunnel. The authors performed a flow field analysis and claim to have provided comprehensive validation data for CFD simulation results.**
> **The paper is well-written and clearly structured. However, while there is certainly a demand for experimental reference data for validation purposes of CFD simulations, especially considering the complex geometries that are relevant in the field of airborne wind energy, it is my assessment that the data provided and the analysis performed will be of no particular use for such efforts outside of the immediate circle of the authors. As the authors discuss themselves, the measurement data is seriously flawed and the insights generated from the analysis of the experimental data are very limited. A short, interesting analysis of CFD results at the strut location is provided in the paper but there is a limited connection with the experimental work that the paper is centred around.**
> **Therefore, it is my conclusion that I cannot recommend this paper for publication unless additional, higher quality experimental data is provided and the analysis is extended and improved.**

Dear reviewer,

Thank you for your comments and the time you put into this review. The reviewer's general concerns focus on four aspects: (i) applicability beyond the immediate circle of the authors, (ii) measurement quality, (iii) scientific insight, and (iv) the connection between the CFD strut analysis and the experimental investigation; each of these points is addressed explicitly below. The specific technical comments are addressed last and also point-by-point.

We look forward to publishing this paper, which has greatly improved with your feedback.

Best,
J.A.W. Poland, on behalf of the authors.

> **"However, while there is certainly a demand for experimental reference data for validation purposes of CFD simulations, especially considering the complex geometries that are relevant in the field of airborne wind energy, it is my assessment that the data provided and the analysis performed will be of no particular use for such efforts outside of the immediate circle of the authors."**

We respectfully disagree with the referee's assessment that the dataset is of no particular use beyond the authors' immediate circle. The present work provides the first *geometry-consistent, spatially resolved* flow-field dataset for a representative LEI-kite configuration, and thereby establishes a common reference case for method development and intercomparison.

This work is directly relevant to the broader airborne wind energy (AWE) community because it consolidates the TU Delft V3 kite as a *reference configuration*. A substantial body of openly documented modelling and experimental studies has already been developed around this geometry, spanning quasi-steady and dynamic models, in-situ measurements, and high-fidelity numerical simulations (Cayon et al., 2025; van der Vlugt et al., 2019; Oehler and Schmehl, 2019; Fechner et al., 2015; Bosch et al., 2014). Importantly, this body of work is not confined to a single institution. Recent independent contributions include a comprehensive review of AWE aerodynamics in which the TU Delft V3 kite is cited extensively, conducted at the Instituto Nacional de Técnica Aeroespacial (Spain) (Castro-Fernández et al., 2026); a Master's thesis at Ghent University (Belgium) that explicitly builds on the TU Delft V3 kite configuration (Van de Vondel, 2025); and large-eddy simulations of AWE wakes based on idealised straight and curved wings inspired by leading-edge inflatable kite geometries, performed at Université catholique de Louvain (Belgium) (Crismer et al., 2023).

By contributing spatially resolved flow-field measurements for this already widely adopted reference geometry, the present dataset fills a clear gap in the existing literature, where validation data have thus far been largely limited to in-flight measurements that do not resolved kite level details.

The relevance of this dataset beyond the authors' immediate circle is further demonstrated by the adoption of the TU Delft V3 kite as a common reference within the EU-funded Horizon Europe project MERIDIONAL (https://meridional.eu/). Within this framework, experimental and numerical datasets associated with the V3 kite are actively used by academic and industrial partners for model development, validation, and intercomparison.

The present PIV measurements, therefore, extend the existing benchmark by enabling validation against mean-flow topology, spanwise circulation trends, and pseudo-two-dimensional lift estimates. In this sense, the dataset complements the integral force measurements reported separately for the same configuration (Poland et al., 2025) and strengthens the role of the V3 kite as a genuinely community-level benchmark rather than an institution-specific case.

> **"As the authors discuss themselves, the measurement data is seriously flawed."**

We acknowledge that the stereoscopic PIV campaign on a strongly anhedral, double-curved LEI kite introduces optical and reconstruction challenges, most notably reflection-induced correlation loss and increasing measurement-plane misalignment towards the tip. We agree that, under the present experimental configuration and geometry, planar stereoscopic PIV does not achieve the near-wall completeness typically achieved in more conventional benchmarks (e.g. wake measurements, or wind-turbine airfoils); this motivates the plane-wise validity and uncertainty characterisation now provided in Table 4 and Table 5 of the revised manuscript. Accordingly, the dominant *systematic* (Type-B) limitations of the dataset are now stated explicitly (Sect. 3.1). However, we do not agree that these limitations render the dataset "seriously flawed" in a manner that precludes meaningful external use.

The impact of these limitations is -in the revised manuscript- quantified through a plane-wise breakdown of data rejection and additional masking (Table 4). This analysis shows that vector loss is concentrated in well-defined regions associated with reflections, shadowing, and out-of-sheet motion, while the retained vectors provide sufficient sample size for converged mean statistics (App. A) and low expanded uncertainty of the mean velocity (Table 5 and also addressed again later in this rebuttal).

The revised manuscript therefore provides quantitative validity information that allows external users to distinguish regions of reduced fidelity from regions that remain suitable for comparison.

The use of spanwise-defined, vertically aligned planes on a C-shaped wing reflects an explicit experimental trade-off rather than a design flaw. A robotic-arm concept that ensured local plane–surface perpendicularity was evaluated during the design phase but rejected in favour of a ground-based traverse system, which kept the measurement hardware outside the flow, avoided blockage and wind-induced vibrations, and provided superior positioning accuracy and repeatability.

Finally, the revised Discussion clarifies that the dataset is not intended as a full-field, near-wall benchmark. Instead, it is a first flow-field benchmark for a realistic LEI-kite geometry that transparently documents the attainable information content of the present experimental configuration and the resulting limitations of the measurements.

> **"and the insights generated from the analysis of the experimental data are very limited."**

We respectfully disagree.

The dataset yields several non-trivial insights that extend beyond integral loads: (i) it resolves spatially distributed flow topology, including separated-flow structure at high incidence, (ii) it quantifies spanwise circulation trends via direct velocity-field integration in planes with sufficient data coverage, and (iii) it provides pseudo-two-dimensional lift estimates in those same planes permitting physics-based comparison with sectional model predictions (and documents the corresponding two-dimensional NOCA derivation). These contributions are now highlighted and explicitly cross-referenced to the relevant figures and to the data-quality context (Sect. 3.1 and the revised Discussion).

In addition, the dataset provides transferable methodological evidence on how the present experimental configuration constrains planar stereoscopic PIV on strongly anhedral, double-curved LEI geometries. This is actionable for the community because it informs future campaign design and discourages the deployment of measurement strategies that cannot maintain approximate plane–surface perpendicularity across the span.

> **"A short, interesting analysis of CFD results at the strut location is provided in the paper but there is a limited connection with the experimental work that the paper is centred around."**

We agree that the original manuscript did not sufficiently articulate the connection between the strut-focused CFD analysis and the experimental evidence. This has been addressed in the revised version by making the comparison structure explicit: the CFD observations in the strut region are now directly linked to the deviations observed in the measurements in planes $Y3$ and $Y4$ in the vicinity of the strut. In addition, the revised text clarifies that the CFD analysis is not intended to replace the experiment, but to contextualise the experimentally observed flow features and their spatial localisation (see revised Discussion, Sect. 4.2–4.3).

> **"Therefore, it is my conclusion that I cannot recommend this paper for publication unless additional, higher quality experimental data is provided and the analysis is extended and improved."**

We have substantially strengthened the manuscript by (i) making the validity bounds quantitative (Table 4 and Table 5), (ii) addressing stitching uncertainty (iii) sharpening the claims to match what the data can support, and (iv) strengthening the CFD–experiment linkage, particularly in the strut region.

We respectfully submit that these revisions address the reviewer's substantive concern—namely, that external users must be able to interpret the dataset with clear knowledge of its limitations—without requiring additional experimental campaigns. The revised manuscript provides this transparency and therefore offers a usable reference dataset for the community: not as a "perfect airfoil-style benchmark", but as a first, well-documented flow-field benchmark on an LEI-kite geometry with explicitly reported constraints.

2.1 Specific and technical comments (RC2)

> **Line 62: It is not clear what the claimed novelty of the applied stereoscopic PIV measurements is. The technique has been developed more than 25 years ago. While the specific application to highly flexible LEI kites would be novel, this is mitigated by the fact that the authors chose to investigate a rigid model instead.**

We agree that stereoscopic PIV is a mature technique. The novelty claimed in the manuscript does not concern the measurement principle itself, but the first application of stereoscopic PIV to a leading-edge inflatable (LEI) kite geometry. In particular, the TU Delft V3 configuration combines strong anhedral curvature with extensive pressure-side recirculation and geometric features such as tubular leading edges and struts, which pose distinct optical and flow-physics challenges for stereoscopic PIV and have not been previously documented in the literature. This novelty framing and its implications for measurement constraints have been added explicitly to the revised manuscript.

We further clarify in the revised manuscript that the rigid scale model was a deliberate methodological choice rather than a limitation of novelty. For industrial-scale LEI kites, wind-tunnel testing necessarily requires geometric scaling, but aeroelastic similarity cannot be preserved in practice because the correct proportion between aerodynamic loading and structural response cannot be maintained across scale while simultaneously matching Reynolds number and the relevant structural parameters (Oehler et al., 2018; Poland et al., 2025). In particular, reproducing membrane pretension and inflatable-tube bending stiffness at model scale is constrained by manufacturing tolerances and material properties; for example, achieving the required bending stiffness would imply impractically high inflation pressures. This scaling rationale has been incorporated into the revised manuscript to motivate the rigid reference geometry.

Given that aeroelastic similarity cannot be preserved under geometric scaling for industrial-scale LEI kites, a flexible, scaled soft-wing configuration would not provide a sufficiently well-defined reference geometry, and describing its instantaneous shape would introduce an additional source of uncertainty. Any mismatch between the experiment and the aero-structural simulation would then be difficult to interpret, as discrepancies could arise from experimental uncertainty (including shape uncertainty), the aerodynamic model, the structural model, or the coupling strategy. By contrast, the rigid configuration provides a fixed, verifiable reference geometry and well-controlled inflow conditions, permitting direct experimental–numerical validation and a more unambiguous attribution of differences primarily to aerodynamic modelling assumptions. This advantage is strengthened further by the fact that the same V3 reference geometry has already been used in multiple numerical aerodynamic studies (Deaves, 2015; Viré et al., 2020, 2022; Cayon et al., 2023; Castro-Fernández et al., 2026), such that the present measurements augment an established community benchmark and can be compared consistently against an existing body of CFD and reduced-order predictions.

In conclusion, the novelty lies in the application: this study provides the first spatially resolved flow-field measurements on a representative LEI-kite geometry. A rigid wing was selected deliberately because geometric scaling of a flexible wing cannot preserve the required similarity ratios in practice, and a fixed, verifiable reference geometry is necessary to isolate the aerodynamic problem and permit unambiguous experimental–numerical validation.

> **Figure 2a: The model appears to be relatively large compared to the diameter of the open jet wind tunnel, which could result in undesirable effects. The discussion mentions that the model has been rotated and moved during the measurements. Has it been verified that the aerodynamics, in particular the tip flow, is not affected by the turbulent mixing layer that originates at the wind tunnel nozzle edges?**

The potential influence of open-jet shear-layer effects and blockage has been carefully assessed. The projected frontal blockage ratio of the model is approximately 3%, which lies well below commonly accepted guideline thresholds (Wickern, 2014; Barlow et al., 1999), and associated blockage-induced corrections are therefore expected to be small for the present configuration (Poland et al., 2025). In addition, shear-layer interference from the nozzle edges is unlikely to affect the present measurements, as the model remains well within the uniform core-flow region of the jet.

For the Open Jet Facility, the turbulent shear layers contract the jet core with a reported semi-angle of $4.75°$ near the nozzle exit (Lignarolo et al., 2014). With the model centred in the 2.85 m nozzle, the wing tips remained approximately 0.8 m from the nozzle edges in the nominal configuration. A linear extrapolation of the reported contraction rate implies that the shear layer would intersect the wing tips only several metres downstream of the measurement region, well beyond the PIV traverse range. While the apparent model size in Fig. 2a may appear large in isolation, the additional perspective provided in the figure below clarifies the substantial lateral clearance between the wing tips and the nozzle edges.

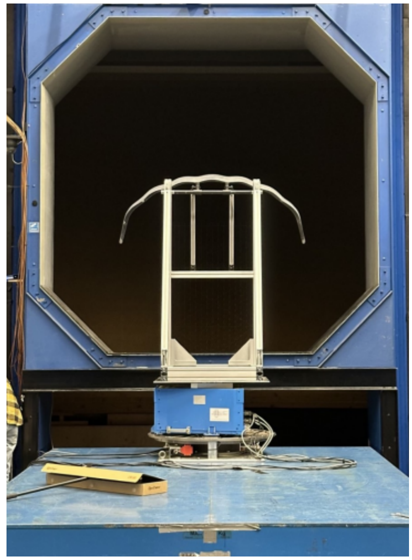

Figure 1. While the apparent model size in Fig. 2a in the manuscript can appear large in isolation, the additional perspective provided here clarifies the substantial lateral clearance between the wing tips and the nozzle edges.

During the measurement campaign, the cameras were repositioned only through controlled traverse motions required for PIV coverage, and the model was repositioned only through pitch rotations and small vertical adjustments to keep the wing-section inside the measurement plane. No rotations about the yaw axis were applied. Consequently, the lateral position of the wing tips relative to the nozzle edges remained effectively unchanged throughout all configurations. Consistently, the upstream outer-field velocities in the PIV measurement planes remained close to the free-stream value and did not exhibit pronounced transverse gradients indicative of shear-layer ingestion.

On this basis, open-jet shear-layer effects and blockage corrections are considered negligible for the present dataset. By contrast, streamline-curvature effects measurably modify the effective angle of attack; the corresponding correction values are reported in Table 1 (of the revised manuscript) with the underlying analysis documented in the companion paper (Poland et al., 2025). These corrections are applied to all reported α values.

> **Line 119: Aiming the laser directly at the curved surface is likely to create strong reflections in the PIV images that cannot be mitigated by black paint. A different measurement setup orientation should be selected to produce reliable measurements.**

We agree that surface reflections are a critical limitation for stereoscopic PIV on highly curved geometries. We interpret the reviewer's use of "directly" as referring to the laser light sheet intersecting the wing in a plane that is nearly aligned with the local surface, i.e. parallel to the wing surface. In that interpretation, we note that strong reflections are not solely due to this angle, as such reflections occur even at mid-span, where the measurement plane is approximately perpendicular to the local span.

In the present experiment, all measurements were performed in fixed vertical planes using a motorised traverse system, which permitted precise positioning and ensured consistent stereoscopic calibration across all spanwise locations. Re-orienting the measurement planes to follow the local wing surface would reduce plane–surface misalignment locally, but would require either manual repositioning with a full stereoscopic recalibration for each plane or an alternative measurement setup, e.g. the use of a robotic arm. The latter option was evaluated during the design phase but rejected as an explicit trade-off: a ground-based traverse system was selected instead because it kept the measurement hardware outside the flow, avoided upwind intrusion that could introduce blockage and wind-induced vibrations, and provided superior positioning accuracy and repeatability.

As discussed in Sect. 4.1, this experience highlights a limitation of the present experimental configuration when applying planar stereoscopic PIV to strongly anhedral, double-curved LEI geometries: surface reflections and increasing plane–surface misalignment progressively limit the reliability and attainability of near-wall measurements. These constraints motivate the use of volumetric approaches (e.g. tomographic or scanning PIV) in future studies. Importantly, the present work documents these limitations quantitatively and demonstrates which flow features remain robustly accessible despite them, thereby providing validation insight that cannot be obtained from integral load measurements alone.

> **Line 148: It appears that the measurement campaign was not finished, with important data points missing.**

We thank the reviewer for this observation. Measurements at $\alpha = 17°$ were limited to planes $Y1$–$Y4$. At this incidence, the onset of separated flow increased unsteadiness and substantially degraded the stereoscopic PIV signal quality, with reflection-driven data loss and elevated uncertainty already apparent in the inboard planes. Given the progressive increase in optical complexity and out-of-plane transport towards the tip, a further deterioration in data quality was expected for $Y5$–$Y7$. The experimental effort was focused on acquiring the most reliable high-incidence datasets inboard and on completing the spanwise-resolved measurements at $\alpha = 7°$, where robust quantitative comparisons were feasible. Despite the reduced spanwise coverage at $\alpha = 17°$, planes $Y1$–$Y4$ remain sufficient to document the separated-flow topology and to support comparison with stall behaviour inferred from the integral load measurements. The manuscript has been revised to incorporate this reasoning.

> **Line 160: The transitions between measurement regions were smoothed out but they are still clearly visible in Figure 6. Please report the difference in flow velocity at the same location across overlapping measurements as an indication of measurement error.**

We agree with the reviewer that, although the stitched transitions were blended using a smooth ramp, residual discontinuities remain visible in Fig. 6 and should be quantified. The revised manuscript therefore adds more details related to the stichting procedure, and a new Appendix with an explicit stitching-consistency assessment based on the velocity differences between independently acquired sub-planes in their overlap regions.

Specifically, for each measurement plane, the overlapping sub-plane pairs are compared at identical spatial locations after applying the same validity filtering as in the main analysis. Within each overlap region, component-wise difference fields are computed, and the discrepancy is summarised using the root-mean-square (RMS) velocity difference,

$$\mathrm{RMS} = \sqrt{\left\langle (u_A - u_B)^2 \right\rangle}, \quad (1)$$

where $\langle\cdot\rangle$ denotes averaging over all overlap points that remain valid in both sub-planes. The resulting RMS differences are reported for u_x, u_y, and u_z in the new appendix (App. D), together with a figure showing the sampling region (Fig. 15) and a summary table (Table E1).

Because missing vectors and locally degraded reconstruction quality limit overlap coverage near the wing, the comparison is evaluated on a restricted subset of overlap points sampled above a line in the upper part of the field. This restriction is stated explicitly, and the reported RMS values are interpreted as a location-specific indicator of stitching-related variability rather than a comprehensive PIV uncertainty metric. The manuscript now also cross-references this stitching uncertainty in the uncertainty analysis section to clarify that it represents an additional systematic contribution that is not captured by the Type-A uncertainty of the mean velocity reported in Table 5.

> **Line 179: There are no arguments provided why the formulation of Noca was used for the application of the conservation of momentum in integral form. On the contrary, the formulation of Noca was adapted to suit the particular test case. The classical formulation of the Navier-Stokes equations would be an obvious alternative here, which the authors do not consider. This formulation is often rejected in the literature when unsteady flows are analysed (due to the required volume integration and pressure computation), which is, however, not the case here. A significant drawback of Noca's method is the required computation of multiple velocity gradients, which are prone to amplifying measurement errors. I sincerely doubt that Noca's method is superior to the classical formulation in this case.**

We agree that the classical steady integral momentum balance represents an obvious alternative approach. The manuscript has been revised to acknowledge this explicitly and to clarify why it is not applicable in the present experiment, as explained below.

In the present experiment, however, the steady incompressible integral momentum balance is not applicable because it is not closed with planar PIV alone: (a) the static pressure on the control surface is not measured, and (b) near-wall velocity information is incomplete due to reflection-driven masking. Applying the balance would therefore require pressure reconstruction and modelling to fill masked regions, and the resulting forces would be strongly conditioned by those assumptions.

A pressure-free momentum-deficit formulation is only valid under special far-field conditions, where the control surface lies sufficiently downstream that static pressure has recovered to the free-stream value and viscous traction on the outer boundary is negligible. Under these assumptions, drag can be estimated from the streamwise momentum deficit in a downstream plane. Lift is more restrictive: recovering lift without pressure requires that the control surface captures the spatial extent of the induced velocity field (downwash and tip-vortex transport), and truncation of the lateral or vertical boundaries introduces non-negligible momentum- and pressure-flux contributions. These far-field conditions are not satisfied here because the available measurements are confined to the near field and necessarily intersect developing shear layers, separated regions, and the forming wake, where pressure recovery cannot be assumed.

The formulation of Noca et al. (1999) was therefore adopted not because it is universally superior, but because it provides a velocity-based boundary-integral closure that avoids explicit pressure evaluation on the control boundary and can be implemented with the measured velocity field on a prescribed contour. We fully agree with the reviewer that Noca's method involves multiple spatial derivatives and can amplify measurement noise. For this reason, the manuscript has been revised to clarify that we do not claim methodological superiority; rather, Noca's approach is used as a practical option under the specific constraints of this dataset, consistent with prior PIV-based force estimation studies where pressure is not directly available (Fritz et al., 2024a, b; LeBlanc and Ferreira, 2022).

Finally, we note explicitly that the present implementation applies Noca's inherently three-dimensional formulation in a pseudo-2D, planar setting; this is an approximation adopted to enable sectional estimates from the available data rather than a claim of strict force recovery.

> **Line 214: The performed PIV uncertainty analysis is too simple to understand the complex sources of error in the presented measurement data. Irrespective of that, Table 3 present uncertainties that go beyond commonly acceptable limits, particularly for measurement data that is intended to validate numerical simulations.**

The authors agree that a single statistical metric cannot capture the full spectrum of error sources in stereoscopic PIV on a strongly anhedral, double-curved LEI geometry. The revised manuscript therefore separates statistical (Type-A) uncertainty from systematic and modelling-related (Type-B) limitations, consistent with established PIV uncertainty practice (Sciacchitano and Wieneke, 2016).

To address the referee's reference to "Table 3" unambiguously, the revised manuscript has been restructured as follows. The quantities that quantify statistical uncertainty of the time-averaged velocity are now reported in Table 5, whereas Table 4 reports plane-wise data-quality metrics (data rejection and additional masking fractions). The referee's concern about "uncertainties" thus maps to two distinct items in the revised manuscript: (i) velocity uncertainty estimates (Table 5), and (ii) plane-wise data-quality limitations that constrain interpretability (Table 4).

Table 5 reports only the Type-A uncertainty of the time-averaged velocity due to the finite ensemble size ($N = 250$), expressed as two-sided 95% confidence intervals. These values quantify statistical convergence of the mean velocity field and, by definition, exclude systematic contributions such as stereo-calibration and mapping residuals, correlation-related bias (including peak locking), finite light-sheet thickness, and geometry-driven optical effects. When normalised by the free-stream velocity ($U_\infty = 15\ \mathrm{ms}^{-1}$ in the present study), the reported Type-A uncertainty levels correspond to approximately 0.2–0.7% at $\alpha = 7\ ^\circ$ and 0.7–1.7% at $\alpha = 17\ ^\circ$.

By contrast, Table 4 does not report velocity uncertainties. It reports plane-wise data-quality metrics: the fraction of vectors rejected by correlation and validation procedures (f_nan) and the fraction additionally masked by the out-of-plane velocity filter (f_{u_y}). Elevated values of f_nan and f_{u_y} indicate reduced data availability and increased post-processing intervention, rather than imprecision of the retained vectors. In the revised manuscript, these metrics are reported explicitly to allow readers to identify regions where reflection-induced correlation loss, shadowing, and progressive measurement-plane misalignment towards the wing tip limit the fidelity of derived quantities, and where validation use should be correspondingly cautious.

Study	Case	U_∞ ($\mathrm{m\,s^{-1}}$)	Uncertainty ($\mathrm{m\,s^{-1}}$)	Uncertainty / U_∞
Huang et al. (2023)	max (streamwise)	5.0	≤ 0.03	$\leq 0.6\%$
Huang et al. (2023)	max (in-plane)	5.0	≤ 0.02	$\leq 0.4\%$
Bensason et al. (2025)	maximum reported	2.7	0.07	2.6%
Present study	$\alpha = 7°$	15.0	0.03–0.10	0.2%–0.7%
Present study	$\alpha = 17°$	15.0	0.11–0.26	0.7%–1.7%

Overall, the data show that the present statistical uncertainty levels are of the same order as recent benchmarks when expressed as a fraction of U_∞, with the larger values occurring towards the wing tip and at $\alpha = 17°$, consistent with increased measurement-plane misalignment. We therefore respectfully disagree that the presented uncertainties go beyond 'commonly acceptable limits'.

> **Line 231: Despite already filtering the experimental data considerably for unreliable measurements in large regions, the discrepancies with the CFD results are still significant and lack physical explanations. This suggests that the PIV measurements should be repeated.**

We do not agree that the remaining discrepancies imply that the PIV measurements should be repeated. The dominant sources of disagreement between experiment and CFD are not attributable to random measurement error that could be reduced by repetition, but to structural limitations of planar stereoscopic PIV when applied to strongly anhedral, double-curved LEI geometries under the present experimental configuration.

Importantly, the present results provide value beyond pointwise numerical agreement. They demonstrate, in a controlled and quantified manner, which aspects of the flow field can be robustly accessed with planar stereoscopic PIV on LEI kites and which remain challenging to capture. In particular, the dataset supports qualitative validation of CFD-predicted flow topology and quantitative comparison of spanwise circulation and pseudo-two-dimensional lift in those regions where data quality is highest.

A further scientific contribution lies in transparently documenting why certain discrepancies persist despite filtering. By quantifying the effects of surface reflections, progressive measurement-plane misalignment towards the tip, and the breakdown of pseudo-two-dimensional force integration (including non-physical drag estimates), the study provides evidence-based guidance for future experimental design. This includes identifying where planar approaches become fundamentally limited and where alternative measurement concepts, such as volumetric techniques, would be more appropriate. As such, the value of the dataset lies not in claiming exhaustive agreement, but in pioneering flow-field resolved measurements of an LEI kite.

> **Line 268: The idea of comparing circulation directly between experiment and CFD is good because it circumvents the lift calculation. However, the selected approach of showing the results with 90% confidence intervals is misleading at least. One would expect the confidence intervals to capture the variable choice of integration contour shape (ellipse vs. rectangle). See also the following comment.**

We agree with the reviewer that the originally presented 90% confidence intervals were not an appropriate representation of uncertainty for the circulation comparison because they did not reflect the dominant source of variability, namely sensitivity to the integration-contour definition. The revised manuscript, therefore, fundamentally changes the uncertainty representation in Fig. 8.

Rather than statistical confidence intervals, Fig. 8 now reports contour-sensitivity intervals obtained from a systematic variation of the integration contour. For both CFD and PIV, the circulation is evaluated for two contour geometries (elliptical and rectangular) and, for each geometry, over an ensemble of 100 contour-size perturbations. The reported circulation value is taken as the mean of the shape-averaged results, while the plotted intervals span the full minimum–maximum range across the combined contour ensemble.

This revised representation explicitly captures the uncertainty associated with contour choice and placement, avoids conflating contour sensitivity with statistical sampling uncertainty, and provides a more physically meaningful basis for experiment–CFD circulation comparison.

> **Figure 8: The measurement point "PIV ellipse" near y = 0.4m is presented with a very small error bar and confidence interval, while it is clearly an erroneous data point (relative error on the order of -40% compared to CFD reference data). Representing incorrect data with low uncertainty indicated is a problematic decision, casting doubt on the entire data analysis performed.**

We agree with the reviewer that, in the original figure, this point was presented with an inappropriately small uncertainty. This occurred because the earlier version reported ellipse and rectangle contour results separately and the plotted intervals did not include contour-shape sensitivity; consequently, the elliptical result at $y \approx 0.4$ m appeared as a low-uncertainty outlier.

This has been corrected in the revised manuscript by adopting the contour-sensitivity intervals described above. Circulation is now averaged over both contour geometries, and the plotted bounds span the full minimum–maximum range across all contour perturbations for both shapes. Under this representation, the point no longer appears as a low-uncertainty outlier; instead, it correctly reflects the increased contour sensitivity of the PIV circulation estimate at this spanwise location.

> **Table 4: Negative drag values are not only reported for measurement data but also for CFD results based on Noca's method. This is in an indication related to my discussion point for Line 179 that the selected method is not well-suited for the analysis performed in this paper.**

Regarding the suitability of the selected method, as noted previously, the classical steady integral momentum balance was considered but is ill-conditioned for the present planar PIV dataset primarily due to closure limitations: static pressure is not measured and near-wall velocities are incomplete owing to reflection-driven masking. Pressure-free momentum-deficit approaches require restrictive far-field conditions, including pressure recovery and negligible outer-boundary tractions, which are not satisfied by the available near-field measurements. The boundary-integral formulation of Noca et al. (1999) was therefore adopted as a practical velocity-only closure under these constraints, without implying general methodological superiority, consistent with prior PIV-based force-estimation studies (Fritz et al., 2024a, b; LeBlanc and Ferreira, 2022).

We further acknowledge that the difficulty associated with drag estimation was not sufficiently emphasised in the original manuscript. The revised manuscript now explicitly notes that drag is widely recognised as a challenging aerodynamic quantity to measure and validate experimentally, particularly when derived from velocity-field measurements. Its magnitude is typically an order of magnitude smaller than lift, rendering it highly sensitive to measurement noise, spatial resolution, and systematic bias. Consequently, while PIV-based approaches often yield robust agreement for lift- or circulation-based quantities, drag-related components exhibit substantially larger uncertainty and variability (Fritz et al., 2024a, b). These difficulties are further exacerbated by wake truncation, masking, incomplete near-wall resolution, and three-dimensional transport effects, which disproportionately affect drag compared to lift (Huang et al., 2023; LeBlanc and Ferreira, 2022).

In the present experimental campaign, the primary objective was to resolve the flow field and not to measure drag accurately. Future studies aimed at accurate drag quantification would benefit from complementary measurement techniques, such as dedicated wake measurements.

> **Line 402: The basis of the conclusions are questionable. The fact stall occurs at alpha = 17 deg has been shown experimentally before and can be considered trivial with basis knowledge of airfoil aerodynamics. The stall onset angle and details on its development could be of interest but have not been investigated.**

The authors agree with the reviewer that the occurrence of stall at $\alpha = 17°$ is not, by itself, a novel finding and was already anticipated from prior integral load measurements and established airfoil aerodynamics. The original wording in the conclusions overstated this point by implying that the stall onset location was determined from the present measurements, which is not the case. This formulation was incorrect, and the manuscript has been revised accordingly.

Only two angles of attack were investigated, and the present study does not aim to determine the stall onset angle or to characterise the detailed development of stall. Instead, $\alpha = 17°$ was selected a priori as a representative high-angle-of-attack condition at which separated flow was expected. The role of the PIV measurements at this condition is therefore limited to qualitatively visualising the separated-flow topology and shear-layer structure, and to assessing the consistency of these features with CFD predictions and integral load measurements (Poland et al., 2025) in regions where optical access is reliable.

The revised conclusions explicitly reflect this scope. Stall is no longer presented as a newly identified or confirmed onset phenomenon, but as an anticipated operating regime used to enable a qualitative comparison of separated-flow behaviour between experiment and simulation.

> **Line 408: “good agreement with CFD” is a statement that I would tend to disagree with considering the data presented in the paper.**

We agree that the original phrasing overstated the level of agreement between PIV and CFD. The revised manuscript therefore softens and clarifies the relevant statements to reflect the actual scope and limitations of the comparison.

Specifically, the revised text states that circulation trends extracted from PIV follow the same overall behaviour as CFD and VSM predictions within the reported uncertainty bounds, and that sectional lift estimates are only broadly consistent with CFD at mid-span when elliptical integration boundaries are used, where measurement uncertainty is lowest. The revised conclusions further emphasise that agreement deteriorates towards the wing tip and that drag estimates are unreliable and interpreted only qualitatively.

We consider this revised wording to be a more accurate representation of what the data support and to avoid implying a degree of quantitative validation that is not justified by the present measurements.

> **Figure 2b: this figure is not discussed in the text. If the figure is not contributing to the discussion in the paper, it should be removed.**

Figure 2b was indeed insufficiently referenced in the original manuscript. Its purpose is to illustrate the smooth geometric transition between the circular leading-edge tube and the membrane surface of the LEI wing, which arises from practical manufacturing constraints and differs slightly from the idealised geometry used in the analysis. This clarification has now been explicitly incorporated into the manuscript text, ensuring that the figure's relevance and contribution to the geometric description are clear.

> **Figure 8: Data for Y7 is missing without explanation.**

The PIV data for plane $Y7$ were excluded from the quantitative circulation analysis because the boundary-curve evaluation required excessive local interpolation. Specifically, for plane $Y7$, the fraction of vectors requiring interpolation within the interpolation squares exceeded 1% for all 100 contour perturbations used in the sensitivity study (App. E). To avoid reporting circulation values dominated by interpolation artefacts rather than measured velocity information, plane $Y7$ was omitted from Fig. 8 and from the associated quantitative analysis. This rationale has now been incorporated explicitly into the revised manuscript.

References

Barlow, J. B., Rae, W. H., and Pope, A.: Low-Speed Wind Tunnel Testing, John Wiley & Sons, New York, 3rd edn., ISBN 978-0471557746, 1999.

Bensason, D., Sciacchitano, A., and Ferreira, C.: On the wake re-energization of the X-Rotor vertical-axis wind turbine via the vortex-generator strategy, Wind Energy Science Discussions, 2025, 1–37, https://doi.org/10.5194/wes-2025-3, 2025.

Bosch, A., Schmehl, R., Tiso, P., and Rixen, D.: Dynamic nonlinear aeroelastic model of a kite for power generation, Journal of Guidance, Control, and Dynamics, 37, 1426–1436, https://doi.org/10.2514/1.G000545, 2014.

Castro-Fernández, I., Sánchez-Arriaga, G., and García-Villalba, M.: A review of the aerodynamics of airborne wind energy systems, Progress in Aerospace Sciences, 161, 101 157, https://doi.org/10.1016/j.paerosci.2025.101157, 2026.

Cayon, O., Gaunaa, M., and Schmehl, R.: Fast Aero-Structural Model of a Leading-Edge Inflatable Kite, Energies, 16, 3061, https://doi.org/10.3390/en16073061, 2023.

Cayon, O., Watson, S., and Schmehl, R.: Kite as a Sensor: Wind and State Estimation in Tethered Flying Systems, Wind Energy Science Discussions, https://doi.org/10.5194/wes-2024-182, preprint, under review, 2025.

Crismer, J.-B., Trigaux, F., Duponcheel, M., and Winckelmans, G.: Large-Eddy Simulation of airborne wind energy systems wakes, Journal of Physics: Conference Series, 2505, 012 036, https://doi.org/10.1088/1742-6596/2505/1/012036, wake 2023 (Visby, Sweden), 2023.

Deaves, M.: An Investigation of the Non-Linear 3D Flow Effects Relevant for Leading Edge Inflatable Kites, Master's thesis, Delft University of Technology, https://resolver.tudelft.nl/uuid:ccb56154-0b70-4a41-8223-24b0f8d145c5, 2015.

Fechner, U., van der Vlugt, R., Schreuder, E., and Schmehl, R.: Dynamic Model of a Pumping Kite Power System, Renewable Energy, 83, 705–716, https://doi.org/10.1016/j.renene.2015.04.028, 2015.

Fritz, E., Boorsma, K., and Ferreira, C.: Experimental analysis of a horizontal-axis wind turbine with swept blades using PIV data, Wind Energy Science, 9, 1617–1629, https://doi.org/10.5194/wes-9-1617-2024, 2024a.

Fritz, E., Ribeiro, A., Boorsma, K., and Ferreira, C.: Aerodynamic characterisation of a thrust-scaled IEA 15 MW wind turbine model: experimental insights using PIV data, Wind Energy Science, 9, 1173–1187, https://doi.org/10.5194/wes-9-1173-2024, 2024b.

Huang, M., Sciacchitano, A., and Ferreira, C.: On the wake deflection of vertical axis wind turbines by pitched blades, Wind Energy, 26, 365–387, https://doi.org/10.1002/we.2803, 2023.

LeBlanc, B. and Ferreira, C.: Estimation of blade loads for a variable pitch vertical axis wind turbine from particle image velocimetry, Wind Energy, 25, 313–332, https://doi.org/10.1002/we.2674, 2022.

Lignarolo, L., Ragni, D., Krishnaswami, C., Chen, Q., Ferreira, C. S., and van Bussel, G.: Experimental Analysis of the Wake of a Horizontal-Axis Wind-Turbine Model, Renewable Energy, 70, 31–46, https://doi.org/10.1016/j.renene.2014.01.020, 2014.

Noca, F., Shiels, D., and Jeon, D.: A comparison of methods for evaluating time-dependant fluid dynamic forces on bodies, using only velocity fields and their derivatives, Journal of Fluids and Structures, 13, 551–578, https://doi.org/10.1006/jfls.1999.0219, 1999.

Oehler, J. and Schmehl, R.: Aerodynamic characterization of a soft kite by in situ flow measurement, Wind Energy Science, 4, 1–21, https://doi.org/10.5194/wes-4-1-2019, 2019.

Oehler, J., van Reijen, M., and Schmehl, R.: Experimental Investigation of Soft Kite Performance During Turning Maneuvers, Journal of Physics: Conference Series, 1037, 052 004, https://doi.org/10.1088/1742-6596/1037/5/052004, 2018.

Poland, J. A. W., van Spronsen, J. M., Gaunaa, M., and Schmehl, R.: Wind Tunnel Load Measurements of a Leading-Edge Inflatable Kite Rigid Scale Model, Wind Energy Science, https://doi.org/https://doi.org/10.5194/wes-2025-77, [under review, preprint available], 2025.

Sciacchitano, A. and Wieneke, B.: PIV uncertainty propagation, Measurement Science and Technology, 27, 084 006, https://doi.org/10.1088/0957-0233/27/8/084006, 2016.

Van de Vondel, N.: Multi-fidelity aero-elastic analysis of a soft kite for airborne wind energy applications, Master's thesis, Ghent University, Faculty of Engineering and Architecture, master's dissertation submitted in order to obtain the academic degree of Master of Science in Electromechanical Engineering – Mechanical Energy Engineering, 2025.

van der Vlugt, R., Bley, A., Noom, M., and Schmehl, R.: Quasi-steady model of a pumping kite power system, Renewable Energy, 131, 83–99, https://doi.org/10.1016/j.renene.2018.07.023, 2019.

Viré, A., Demkowicz, P., Folkersma, M., Roullier, A., and Schmehl, R.: Reynolds-averaged Navier-Stokes simulations of the flow past a leading edge inflatable wing for airborne wind energy applications, Journal of Physics: Conference Series, 1618, 032 007, https://doi.org/10.1088/1742-6596/1618/3/032007, 2020.

Viré, A., Lebesque, G., Folkersma, M., and Schmehl, R.: Effect of Chordwise Struts and Misaligned Flow on the Aerodynamic Performance of a Leading-Edge Inflatable Wing, Energies, 15, 1450, https://doi.org/10.3390/en15041450, 2022.

Wickern, G.: A Theoretical Approach towards the Self-Correcting Open Jet Wind Tunnel, Tech. Rep. 2014-01-0579, SAE Technical Paper, 2014.