Rebuttal of: “Computational aerodynamics for leading-edge inflatable kite design”

Jelle Agatho Wilhelm Poland, Kasper Raphaël G Masure, Oriol Cayon, and Roland Schmehl

1 General comments (RC1)

> **This paper brings together several elements related to the aerodynamic predictions for kites in airborne wind energy systems, using results and methods from multiple related publications.**

Dear reviewer,

Thank you for your thorough review and constructive comments. Below, we address each comment in detail. All changes are documented in the attached tracked-changes manuscript.

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

1.1 Specific comments (RC1)

> **The title is too broad, as the parametrization of the airfoil section seems to apply only/mostly to leading-edge inflatable kites, not to ram-air kites for instance, so the title should mention this limitation.**

We agree that the presented parametrisation applies specifically to leading-edge inflatable kites. The title has therefore been changed to “Computational aerodynamics for leading-edge inflatable kite design”, which states this limitation explicitly.

> **The title is somewhat misleading, as the “soft” aspect is not included in this paper, there is no aero-elasticity. This is also present in the manuscript and in the conclusion.**

We agree that the manuscript does not include aero-elastic coupling. The new title, “Computational aerodynamics for leading-edge inflatable kite design”, no longer suggests this.

> **The introduction with the step-wise build up and history of the lifting line method is interesting and suitable for a PhD thesis, but does not support the key message of this paper and is not required to understand what is being done in the presented work.**

By request, we have revised the introduction into a single continuous section without subsections. The former Sect. 1.1 (∼1000 words and one figure) has been condensed to a single paragraph of ∼200 words: the purely historical material (Munk, Glauert, Mutterperl, Sivells, Cone) and the former Fig. 1 have been deleted, and implementation-related citation content (wake alignment following Phillips, 2000; vortex-core and self-induction credits) has been relocated to Sect. 2, where it supports the method description.

We retained only the two developments the method is directly built on: the three-quarter-chord control point (Pistolesi; Weissinger) and the circulation-matching reformulation (Tani; Piszkin and Levinsky), as these are required to understand the hybrid viscous-inviscid character of the VSM and why it depends on representative sectional polars, which motivates the machine-learning surrogate central to this paper.

> **The machine learning model is now built separately for each Re, which is strange, as Re will constantly be varying during operation so also in realistic simulations. This should be changed by making a combined model, probably after gathering new data, or mentioned clearly as a significant limitation of the applicability.**

The intended usage is to interpolate the sectional polars between the bracketing per-Re models in $\log \mathrm{Re}$. This is well-founded within the model's own assumptions: all underlying simulations assume fully turbulent boundary layers (justified in Appendix C), which excludes the transition-driven regime changes responsible for abrupt Re-dependence. The remaining Re-dependence is weak, smooth, and monotonic, consistent with turbulent skin-friction scaling, and Viré et al. (2022) found negligible integral-force differences over a factor-2 change in Re for the same kite. This is also supported experimentally: the wind-tunnel campaign of Poland et al. (2026) measured the rigid V3 scale model across $\mathrm{Re} = 1.3 \times 10^5$–$6.1 \times 10^5$ (their Fig. 8) and found only gradual variation of the force coefficients with Re, with no abrupt regime changes; the earlier stall onset observed at the lowest Reynolds number reflects transitional effects below the fully turbulent range covered by the surrogate model, consistent with the validity limit now more clearly stated in the manuscript.

We verified the Re-dependence on the trained models directly: across the factor of 20 change in Re, the predicted C_l at $\alpha = 6°$ changes by only 1–8% and C_d decreases by 14–26% for the three representative test profiles of Table 7, with the three independently trained models remaining closely spaced across the operational band; where their ordering is not strictly monotonic, the differences lie within the regression uncertainty. These results are now quantified in Appendix D. The largest Re-sensitivity appears at stall onset, which the per-Re models resolve individually.

We therefore consider sticking to 3 separate per-Re models appropriate, and the interpolation (when done with $\log \mathrm{Re}$) not sufficiently limiting to motivate an increase in dimensionality (which would resolve from adding Reynolds as an additional input parameter).

> **The PCHIP interpolation for the angle of attack is a strong limitation, because that means a user or application should always extract the data from the model at the sampling points and then fit them before actually using the data. Smooth and physical behaviour should be built into the machine learning model by construction, e.g. by having coefficients of a polynomial for the data as output instead of the actual data.**

We would like to clarify a point that the original text obscured: the angle of attack is a continuous input feature of the ML model, alongside the six shape parameters. The model, therefore, returns coefficient values for any requested angle of attack directly; no extraction at the training sampling points is required to use it. We verified this empirically: querying the trained model over $\alpha \in [-10°, 25°]$ at $0.01°$ resolution for a representative profile, the raw predictions deviate from the PCHIP-fitted polar by less than 0.03 in C_l (RMS 0.006), well below the regression error itself. The PCHIP fit is thus a smoothness refinement, not a required decoding step: because tree ensembles are piecewise constant, the raw response varies in fine steps, and the spline is applied only where a C^1-continuous polar is needed, e.g., for derivative-based analyses. Within our framework, this step is automated and invisible to the user.

Regarding building smoothness in by construction via polynomial-coefficient outputs: we considered this, but deliberately moved away from it. The coefficient-based approach is precisely the Breukels (2011) formulation discussed in the introduction, whose fidelity is capped by the chosen basis. A low-order polynomial cannot represent the sharp stall break and post-stall plateau characteristic of LEI airfoils, while higher orders introduce spurious oscillations that shape-preserving PCHIP interpolation avoids by construction. Moreover, the non-smoothness originates from the tree architecture (piecewise constant in all inputs), so coefficient outputs would not resolve it; the architectural remedy would be a smooth surrogate such as a neural network. We agree that for applications requiring gradients through the model (e.g., gradient-based shape optimisation), such a smooth architecture would be preferable, and we consider this a valuable direction for future work; but for the polar-generation task at hand, the pointwise ERT model with optional PCHIP regularisation proved the most accurate option.

> **The actual contribution and goal of this paper are drowning in too much other information that is interesting as such, but not contributing to the message.**

We have streamlined the manuscript to bring the contribution forward more clearly. The historical introduction has been condensed to a single paragraph (see our response above), implementation-related material has been moved to the method description in Sect. 2, and additional supporting analyses have been relocated to the appendices.

> **Line 29: Give a range for the lift-to-weight ratios.**

The sentence has been rephrased: the qualifier "exceptionally" has been removed, and a quantitative range for the lift-to-weight ratios has been added.

> **Line 35: "But" is a coordinating conjunction and should only be used at the beginning of sentences to signal a shift in thought.**

The conjunction has been removed, and the sentence has been adjusted accordingly.

> **Line 69: Kutta (1902) ⇒ (Kutta, 1902)**

The citation has been changed to the parenthetical format (Kutta, 1902).

> **Line 182: Fritz (2024) ⇒ (Fritz, 2024)**

The citation has been changed to the parenthetical format (Fritz, 2024).

> **Line 226: Introduce abbreviation AIC here.**

The abbreviation AIC (aerodynamic influence coefficient) is now introduced at its first occurrence.

> **Line 275: Motivated ⇒ and which is motivated**

The phrasing has been corrected to “and which is motivated”, as suggested.

> **Line 283: w_i should be introduced after equation (19)**

The symbol w_i is now introduced directly after Eq. (19).

> **Line 328: Is this at 20% starting from the TE? Or 80% in the current formulation?**

The 20 % is measured starting from the second rear control point. The formulation has been adjusted to make this explicit.

> **Line 334: The sizes that were coming out of the sensitivity analysis should be summarized here.**

The cell sizes resulting from the sensitivity analysis are now summarised at this point in the text.

> **Line 338-9: It seems contradicting that one chord length is used in the z-direction and at the same time unit depth. Furthermore, it is not clear what the physical distance (in multiple of chord lengths) covered by the 201 cell layers is.**

There is no contradiction: the profile chord is normalised to unit length, so the unit-depth spanwise extrusion corresponds exactly to one chord length; the sentence has been reworded to state this once. Regarding the wall-normal extent: the 201 layers grow geometrically with ratio 1.1 from a first-layer height of $7.8 \times 10^{-6}\,c$, corresponding to $y^+ \approx 1$ at $\mathrm{Re} = 5 \times 10^6$, which places the circular far-field boundary at approximately 1.4×10^4 chord lengths from the profile. A domain convergence study confirming that the far-field placement does not influence the computed force coefficients is reported in Masure (2025). This information has been added to Sect. 3.2.

> **Line 346: Mention the remaining error.**

The remaining discretisation error has now been quantified in the manuscript. The production mesh ($575 \times 201 \approx 1.15 \times 10^5$ cells) was taken from the systematic mesh-resolution study of Masure (2025), in which the tangential resolution (275–775 nodes) and the wall-normal resolution (161–221 layers) of the structured O-grid were varied independently for representative conventional and highly cambered profiles at $\alpha = 0$, 10, and 15°. Relative to the next-finer resolution level, the adopted 575×201 grid changed the lift, drag, and moment coefficients by less than 2 % across the operational angle-of-attack range, with larger deviations occurring only at $\alpha = 0°$, where the extended pressure-side recirculation makes the steady-RANS solution least reliable. This remaining error of the order of 1 % in C_l and 2 % in C_d was considered acceptable, as it is small compared with the model-form differences between the numerical approaches and the wind-tunnel measurements (4–13 %).

> **Line 353-5: It is strange to mention documentation of 2025 for an older OpenFOAM version.**

The reference has been replaced by documentation matching the OpenFOAM version used, dated 2006.

Line 393: The importance of the variables could be quantified using Sobol indices.

We agree that Sobol indices would be the appropriate tool for quantifying parameter importance in general. In this specific case, however, the influence of ϕ cannot be quantified a posteriori: ϕ was fixed prior to dataset generation, so the training data contain no variation in this parameter. A variance-based sensitivity analysis of the trained surrogate would return a zero index for ϕ by construction (reflecting the sampling, not the physics), and evaluating the model at $\phi \neq 0.65$ would constitute extrapolation without supporting data. Quantifying the sensitivity to ϕ would therefore require a dedicated set of new CFD simulations varying ϕ, which we consider beyond the scope of this revision.

The decision to fix ϕ was instead based on a priori evidence: the CFD sensitivity study of Masure (2025), cited in Sect. 3.1, found that varying ϕ within its physically consistent range only slightly redistributes the pressure along the front spline and has a negligible effect on the integral coefficients, as the dominant flow features of LEI airfoils (the suction peak on the leading-edge tube and the pressure-side recirculation) are governed primarily by t, κ, and η.

Line 411: What did the flow field in the outliers look like? How come there is a strong variation in their results despite all being converged?

The full flow fields of the approximately 135,000 RANS simulations were not archived due to storage constraints; only integral coefficients and convergence metrics (iteration counts, final force-coefficient residuals, maximum y^+) were retained, so the outlier flow fields cannot be re-inspected directly. We have instead characterised the 397 removed outliers (0.93 %) of the $\mathrm{Re} = 5 \times 10^6$ dataset using these archived metrics. The outliers are genuinely converged: none reached the iteration limit, and their force residuals are of the same order as the retained cases. However, they are far from uniformly distributed: 84% occur at $\alpha = 6$–$12°$, i.e. around stall onset, and they are strongly biased towards the extremes of the geometric design space (largest thickness and camber, most forward camber position); 80% were flagged on C_d, the coefficient most sensitive to separation. Notably, no outliers are flagged at $\alpha \geq 14°$, where the entire population has stalled, and the statistical spread is correspondingly large.

This indicates that the strong variation among converged results reflects a change of flow state rather than numerical error: at stall-onset angles, geometries at the edges of the design space separate considerably earlier than the rest of the population, and the steady RANS solver converges to a massively separated solution whose force coefficients deviate strongly from the attached-flow mean at that angle of attack. Since steady RANS solutions in this bifurcation region are known to be unreliable (the flow is inherently unsteady and the converged state may depend on solver trajectory), we consider removing these isolated points (0.93 % of the data) a conservative choice that avoids contaminating the smooth surrogate with sparsely sampled, low-confidence data. A clarifying statement to this effect has been added to Sect. 4.1.

Line 464-5: This fits better with line 445.

We agree. In the revised manuscript, the PCHIP methodology has been consolidated in Sect. 4.3 (Model set-up) as part of the revision addressing the earlier comment on PCHIP interpolation. The sentence in Sect. 4.5 has been rewritten so that it no longer introduces the method but only describes how the comparison in Fig. 9 was constructed, with an explicit reference back to Sect. 4.3.

Line 482: The counterclockwise only applies on the pressure side (but that is also where the recirculation occurs here).

A specifier has been added to clarify that the counterclockwise rotation applies on the pressure side.

Line 487: encounters both three $\Rightarrow$ encounters three

The word "both" has been removed.

Line 580: $V \Rightarrow U_\infty$

The symbol has been corrected to U_∞.

Line 587: The last sentence is repetition from the line above.

The repeated sentence has been removed.

Line 639: Missing point at end.

The missing full stop at the end of the sentence has been added.

Line 711: $< 0 \Rightarrow > 0$

The sign has been corrected to > 0.

Figure 17: Indicate also unstable regions due to other criteria, e.g. in the right subplot with a second color.

The regions that are unstable due to the other criteria are now indicated in the each subplot using varying shades of grey.