The article describes a validation of two aeroservoelastic models using a unique dataset obtained in the field from a 4.3MW wind turbine where one blade is equipped with a trailing edge flap. The work is novel and very relevant. The article is fairly well written and I recommend the publication of the article. Before publication, I think the authors could improve the article further by following some suggestions:

1) Although slightly out of scope, it would be nice and relevant to read about the intended use of such flap system. Both the abstract and the introduction say that the ATEF system is promising. Why is that? Promising for what?

AG: example of potential benefits added in line 26 as follow "*For example, Ungurán and Kühn (2016) estimates a 10% reduction in the flapwise blade root bending moment and a 6% reduction in the tower side-side bending moment with an individual flap control strategy. These load reductions can be exploited to lower the components' cost or increase the energy production, as shown by Pettas et al. (2016) and Abbas et al. (2023), which estimated a potential reduction of the levelized cost of energy of 1.3%.*"

2) The authors use acronyms heavily. In my current job we have a communications department that oversees our manuscripts, and I've learnt that the use of acronyms should be minimized to improve readability AG: The use of acronyms has been significantly reduced. MBrM replaced by BMD.

3) The document relies heavily on the word "transient", which I found fairly confusing. What does a "maximum difference for the blade-to-blade MBrM transients below 1%" mean? Isn't enough to say "maximum difference for the blade-to-blade MBrM below 1%". We know MBrM varies in time/azimuth/wind speed. This is only one of the many uses of the word transient that I found confusing. AG: The use of transient word is improved and clarified in the whole document

4) you've split the validation in three steps: first the flap deflections, second the lift coefficients, third the full aeroelastic model. Although this is said multiple times, it doesn't always come out clearly. Maybe a scheme could help, or a clear numbered list in the intro? AG: The 3 phases has been numbered in the introduction (line 93). The Step number has been added also to the chapter name and in the figure 3

5) HAWC2 and BHAWC are similar models and indeed the results match very well between the two. Would it help to only report results from the former? AG: the purpose of the paper is also to show how close the results of the 2 codes are for the specific application, therefore results from both codes are included. It is clarified in the manuscript in line 85

6) Several paragraphs are very dense and not always clear. Maybe some schemes would help quick readers glance through the document. AG: Figure 3 now shows the flap model component involved in the 3 validation steps. Several improvements to the manuscript have been done to improve readability.

And some additional minor suggestions:

- Line 1: Why is “Wind Turbine” capitalized? AG: removed as the abstract is updated

- Line 25: it would be interesting to read more about the “potential benefits”. AG: See comment #1

- Line 107: parenthesis seems missing wrapping Bergami and Gaunaa, 2012. AG: parenthesis added

- Line 112: typo, initial vs indicial AG: Indicial is correct

- Lines 120-124: I don’t follow this paragraph. Can you please rephrase it? AG: Rephrased as “*This simplification is possible because the controller system controlled only the final pressure value and the pressure valve actuation time. The variation of the air pressure inside the hose after the valve actuation depended only on the layout of the pneumatic and flap systems after the valve (for example, the length, diameter, and material of the hose and the stiffness of the flap). Therefore, the air pressure and the corresponding flap deflection were expected to have a similar transient response for each pressure valve actuation.”*

- Line 131: mainly 4 million. Why mainly? AG: most of the measurements were run at 4 milions. Few at 3.5 or 3 milions. Rephrased with “The measurements, most of them run at a Reynolds number of 4 million,” in line 178

- Line 139: lift and drag coefficients don’t depend on the chord. Why were they adjusted to the chord? And how?
TB: The correction accounts for the flap chord percentage, so scaling the wind tunnel data (variation of the coefficients) for the actual percentage of the flap in full scale. The text has been re-phrased for clarity.

- Line 165-170: I find this paragraph somewhat hard to follow. AG: The paragraph has been rewritten as “*The BHawC flap model directly provides the instantaneous*

stationary 2D aerodynamic properties to the global wind turbine model. Instead, the HAWC2 ATEFlap model computes already the unsteady effects due to flow separation and the vorticity shed into the wake, providing the instantaneous dynamic aerodynamic properties to the global wind turbine model."

- Figure 4: I was surprised to see a transient of 3 seconds. Isn't the pitch actuator faster that that? Again a little out of scope, but the value of a "slow" ATEF becomes harder to justify. AG: future wind turbines will have longer blades and consequently slower rotor speeds. Therefore also a "slow" flap will be able to address 1P loads, reducing the use and consequent wear and tear of the pitch bearing and pitch actuator. Nevertheless, the VIAs project had the aim to develop the flap technology, building the knowledge for future and faster flap systems.

- Line 191: how can you target one wind speed experimentally? AG: Rephrased as "with an almost constant wind speed"

- Line 217: 60s+60s=120s (not 90)? AG: The "s" was a typo. Rephrased as *"completing a total of 90 cycles"*

- Line 235: this issue should be discussed further up or the previous paragraphs are confusing AG: The explanation is expanded here and removed from the previous chapters

- Line 237: if you average across azimuth, shouldn't the azimuth-variation be gone entirely? AG: the azimuth-variation is entirely gone only if the data is evenly distributed among symmetric azimuthal angles (minimum every 90 deg). Unfortunately, this is hardly achievable with measurements.

- Lines 248: I would put this paragraph first, and then the analysis. AG: Thank you for the suggestion. AG: We prefer to keep the current order: Azimuth variation of Cl, Simulations, Measurements, Comparison between measurements and simulations

- Figure 6: the legend covers the label: AG Figure updated

- Figures 12/13/14: include the grey band in the legend : AG Adding the error bands in the legend will make it significantly bigger and reduce the readability of the figure. The error band are anyway mentioned in the captions

- Line 480: I see a good match in Figure 15, what am I missing? AG: There is not a clear correlation between the BMD transient curves and the wind speed interval. However, the small differences between the averaged curves from different wind speed intervals have a magnitude comparable the validation error margin. This suggests the variation

of the external conditions can partly justify the differences observed between simulations and measurements.

- Line 494: example of excessive use of acronyms, what's o2o AG: acronym for one-to-one. Removed as it is rarely used in the paper

- Line 515: typo, affect AG: corrected

- Line 517: typo, improves AG: corrected

- Line 518: typo, overestimates AG: corrected

- Line 537: isn't a DT of 0.04 s excessively large? AG: 0.04 is the sampling time of the prototype data acquisition system. The simulation time step is actually 0.02s. The error has been corrected in the whole manuscript

- Line 565: thank you for the acknowledgment, even before reviews were in! To be seen if they improve the paper ;) – AG: You are welcome