The paper ‘About the suitability of different numerical methods to reproduce model wind turbine measurements in a wind tunnel with high blockage ratio’ is motivated by three main objectives: the provision of experimental data for validation of numerical wind turbine simulations, the validation of two specific numerical codes, and the assessment of the wind tunnel blockage effect. These three objectives have a significant scientific importance for the development and validation of reliable numerical codes and for wind tunnel studies with a high blockage ratio. The work in this paper is part of a larger wind energy research project about active flow control for wind turbines. Several papers have already been published as part of this project. The added value of this specific paper is the presentation of experimental data for the on-blade velocity, angle-of-attack and blade bending moment, as a function of azimuth, and the comparison with two numerical codes: a lifting line free vortex wake code called QBlade and a URANS flowsolver called FLOWer. The measurements and comparisons are done for three turbine yaw angles, and the simulations with FLOWer are done with and without wind tunnel walls.

Although the paper presents valuable experimental data and a comparison with two numerical codes, it requires in my opinion a major revision before it can be considered for publication. In general, I find the text not always clear, unstructured, and in many cases too vague. The lack of structure makes the text confusing to read, especially because the paper is investigation many different parameters, such as experiments, two different numerical codes, blockage effects and the effect of yaw misalignment. The new content of this paper is presented in chapter 3, with a comparison between experiments and simulations. Although the results are very interesting, the discussion is not thorough enough and the conclusions oversimplified. The differences between the experimental and numerical inflow conditions are neglected. However, the inflow conditions show a difference in the mean velocity of +- 10% compared to the simulation conditions, which seems not negligible. The authors should discuss the consequences of the different inflow conditions on the measurements in more detail. Furthermore, it would be interesting to verify and discuss specifically which physics are modeled correctly by the codes, and which not. For example, one could verify this by estimating the angle of attack, on-blade velocity, and bending moment with a simple BEM method, to verify the benefits of the Lifting Line Free Vortex Wake code. The comparisons for the angle-of-attack and on-blade velocity show a good agreement between simulations and experiments. However, the comparison for the bending moment, which is a result of the former two parameters, shows a surprisingly large difference. It would be interesting to discuss the possible causes for this difference in more detail. I have added a list of major and minor comments below.

Major comments:

- The introduction does not clearly motivate the research objectives with a literature overview. For instance on page 2, lines 13-15, it is mentioned that earlier studies already verified the effect of wind tunnel walls and blockage effects with simulations. What where the conclusions? From the introduction it is thus unclear what this paper will contribute to the study of wind tunnel blockage effects.
- When a comparison with measurements is done, it is important to consider the measurement uncertainty. Add an estimation of the measurement uncertainty.
- Section 2 about the ‘methodology and setups’ is badly organized and needs a significant improvement.
    - Each section and sub-section needs an introduction.

    - The section about the wind tunnel is too short and brief.
    - It is not necessary to mention the top-speed for the wind tunnel test section which is not used in this paper.
    - The measurements are performed in the settling chamber which has as purpose to condition the non-homogeneous and turbulent flow from the wind tunnel fan before it enters the test section. As figure 9 indicates, there is a significant mean shear over the cross section in the settling chamber, and the turbulence intensity is not negligible. It is important to provide a motivation for this configuration, provide a characterization of the inflow and turbulence properties, and discuss the effects it may have on the results.
    - Section 2.1 'Experimental setup' does not describe the velocity measurement setup.
    - Mention the specific acquisition devices, and not just the name of the manufacturer.
- P5 Figure 3: The actuators for the flaps and the 3-hole probes + air tubes on the smart blade look like they will influence the airflow around the blade. The presence and impact of this blockage should be discussed.
- P10 L3: An acquisition time of 16 seconds is short for velocity measurements. Please motivate, e.g. based on the integral time scale, that this is sufficiently long for good statistics.
- P15 L5: I don't agree with the statement that the error is small. The error in figure 12 is higher than 10% in a large part of the wake: shear region and center.
- Figure 18: The experiments show a significant dip around 90 degrees. This is not visible in the simulations. Is there a reason for this effect? Is this also due to the traverse? Explain the situation.
- The differences between the measured and simulated bending moments in figures 20 and 21 are significant. It is not ok to say that this is a good agreement. The experimental curves follow a different pattern, especially for CaseBase. Is there an explanation for this?
- The authors should be careful with copyrights. For instance figure 1a, figure 2, figure 3b and figure 8 can be found identically in the paper 'Reproducible inflow modifications for a wind tunnel mounted research HAWT' by Bartholomay S., et al. 2017.
- P24 L26: The main conclusion of this paper is too strong. The experiments have too many differences (e.g. vertical shear and turbulence) to make this statement. Furthermore, the agreement for the bending moment is not good at all. Instead make conclusions on what can be estimated correctly, what not, and which physics are modeled correctly.

Minor comments:

- Define abbreviations at first use in the main text. Don't define abbreviations in the abstract, and limit the use of abbreviations in the abstract. For example CFD, LLFVW, URANS, ..
- Throughout the text, several sentences are unnecessarily long, or have a structure where the subject is placed at the end, which can be confusing. Improving these sentences will benefit the clarity of the text. Some examples are:
    - P1, L1
    - P1, L2: This is a long sentence and not entirely clear. For instance 'methods of different fidelity' is too vague.

    - P1, L5: Is it relevant where the code was run?
    - P2, L3: long sentence.
- P2 L11: It is not clear what ‘a one third model’ is + this is a long sentence.
- P2 L17: The yaw angle is negative. Does this matter? The orientation is not mentioned.
- P2 L19: ‘the flow around the rotor’ is too vague.
- Define the term ‘far field conditions’.
- Units need to be formatted correctly.
- The first sentence of the introduction is too vague and unnecessary.
- P1 L17: You mention ‘simulations’ but the referenced paper presents experimental results.
- P2 L16: Instead of ‘three different states’, it would be more clear to mention ‘Three different yaw-misalignment cases’
- P2 L23: This sentence is very long, consider breaking it up in several more clear and well defined sentences.
- P3 L6: The text mentions ‘low Reynolds numbers’. Please describe the Reynolds numbers at which the experiments are run, and motivate if the experiments scale realistically.
- P3 L9: How is the boundary layer thickness estimated? Is it possible to indicate the tape on figure 3?
- P3 L15: It is confusing to mention at this point in the text the overall goal of the research project, as it is different from the objectives of this paper.
- P5 L2: What is meant with ‘trailing edge deployment’?
- Table 2: What are the units for the wake length?
- P6 L14: 21 panels are mentioned in the text, but in Table 2 15 panels are mentioned. Which one is correct?
- P7 l3: It is not entirely clear what is meant with ‘overlapped using the CHIMERA technique’, overlapping several grids?
    - P7 L7: Don’t use double brackets ‘(())’.
- Table 3: I suppose the units are millions of cells?
- P 8 L6-7: This is a repetition.
- Table 4 is an important table. Maybe it can be discussed earlier in the text.
- P10 L3: Which probe was taken as the reference then? How are the hot-wire probes calibrated?
- Figure 7 is unnecessary.
- P11 L15: How much are these corrections typically? Maybe indicate in figure 16.
- P12: The description of the strain gage setups should be done in the experimental setup section.
- P9 L17 The text mentions measurements at 1.05d , while P13 L3 doesn’t mention measurements at 1.05D. Be consistent, also with the unit of ‘D’.
- P13 L15 ‘More information about this topic can be found in..’ is too vague.
- P14 L6: ‘Some aspects’ is too vague.
- P14 L10-15: Conclusions on wake comparison are not clear, which of the two simulations is discussed?
- P15 L12: What does this mean for the measurement blade: isn’t 100% the maximal radial position?
- P17 L8-10 & p18 L10: ‘More information about..’ is too vague.
- P18 L13-L16: This should be mentioned in the introduction.