The paper "Mid-fidelity simulations and comparisons of five techniques for axial induction control of a wind turbine" presents a numerical study of static derating axial induction control of a single wind turbine in view of increasing total power extraction in a virtual downstream turbine. The paper is original and relevant to the wind farm control community. The main highlight of the paper seems to be that only derating at increased thrust setpoints can potentially achieve overall power improvements, and that other techniques are incapable of concentrating energy gains at downstream rotor locations, which is the current state of literature (Annoni 2015). The paper provides some interesting observations, but I feel certain aspects should be further improved to attain scientific standards, especially related to the use of mid-fidelity CACTUS model. More detailed comments are listed below, which I believe could significantly enhance the quality of the paper.

Major comments

1. My main comment relates to the use of mid-fidelity tool CACTUS, and its capability to accurately describe the wake dynamics in response to different turbine setpoints. Based on its description in the paper, I find the overall adequacy of the CACTUS model for the current purpose questionable. More specifically,

a. p.4, L60: The authors mention using a mid-fidelity tool rather than high-fidelity LES for computational cost reasons. However, the computational cost of performing 16 simulations of a single actuator line model at different control setpoints in LES does not seem prohibitive to me. On the other hand, the amount of revolutions required for statistical convergence of the results could significantly increase the LES cost. Could the authors comments on the actual computational cost of CACTUS, and compare it to an
b. The authors comment around L150 that far-wake behavior is subject to 'numerical turbulence'. Is this similar to physical turbulence, or simply noise? If the latter, this would appear to limit its use for modeling wind-turbine interactions. Can this be linked to the spectral analysis in Figure 7 / Section 3.1?

c. Comparison with LES in a prior study showed a discrepancy in wake recovery rates (L170), which would again render CACTUS inappropriate for modeling turbine interactions for control purposes. The authors later mention (L175) that differences can 'likely' be viewed as bias errors, but I find no convincing justification for this claim. Please clarify.

d. The lack of dissipation is often mentioned as a limiting factor for interpreting CACTUS results (e.g. L157, L221), is this a fundamental limitation of the code? If no, why not include it?

e. The verification section is unconvincing. It consists of an energy balance within the domain, which is more a measure of numerical stability of the code rather than physical realism. The energy residuals of the maxCt cases are far from convergence, with residuals over 50% of total energy at 15 revolutions. The authors mention that additional revolutions will likely improve convergence. Given that the maxCt cases are the most performant cases regarding wake recovery, please explicitly show that this convergence can be attained. Furthermore, Figure 4 does not clearly illustrate positive and negative energy terms to be equal (plot them both as positive to allow visual comparison). Again, the imbalance for maxCt cases is worrying.

g. Considering all comments above, I find the accuracy of CACTUS for the given purpose comes across as relatively weak, and the attempt at verification in my opinion provides no further confidence in the model. Furthermore, literature shows that inlet turbulence strongly affects dynamics of wake recovery, so it is very doubtful that current conclusions can be transferred to turbulent ABL conditions. I thus believe that the authors should stress much more that current results should be interpreted with high caution (throughout the paper, including in the abstract), and that any true physical conclusions can only be formulated after validation in high-fidelity LES and/or wind-tunnel studies. The authors do so to a limited extent around line 175, but given the comments above, I feel this is a much more fundamental flaw of the paper and advise for caution should be formulated stronger.

As a suggestion, it might be feasible to include an LES of at least some cases (baseline, 10maxCt, 1 other 10% derate) in the current paper, as this would significantly strengthen it.
2. The discussion of Fig. 8, with the statement that wake recovery happens more slowly for higher derates, seems quite wishful thinking. The authors themselves point out the exception for maxCt cases, but I’d say that the differences do not seem significant for any of the cases actually.

3. One of the main conclusions is that the wake recovery in the maxCt cases is very different from the other cases. Please relate this to the recent study of wind turbines operating at high thrust coefficients in the following reference: Martínez-Tossas, Luis A., et al. "Numerical investigation of wind turbine wakes under high thrust coefficient." Wind Energy (2021).

4. In Figure 17, it is shown that power increase seem to vanish for larger streamwise distances. Can the figure be extended to include larger spacings? Typical turbine spacings in modern farms are well above the 6 rotor diameters shown in this figure.

5. The virtual rotor bending moment analysis in Figure 19 is somewhat misleading and adds little to the overall story. By normalizing with the baseline case, the maxCt appears to lead to significant load enhancement, whereas these high values might not correspond to problematic cases, as the baseline moment is simply very low due to the wake deficit. The load analysis for the bending moment hence simply comes down to a reduced wake deficit (which in itself if the target of the AIC...). The virtual rotor moment analysis can possibly be omitted.

Minor comments

- It would be good to comment on the distinction between static (focus of current paper) and dynamic axial induction control, which has been getting receiving increased attention in recent years, as well as add some references on the latter.

- L96: Reporting a default regularization value at 1e-7 in itself is meaningless, does it have a unit? Either report a unit, or some context for this value, or don't report the exact value at all.

- L122: dynamic viscosity should have a unit

- Figure 3: colormapping preferably in a perceptually uniform, b/w intelligible color (e.g. viridis/parula). If possible, it would be illustrative to explicitly include the contourlines for C_p at 10, 20, and 40% derating (in addition to the equispaced existing contourlines). The purple star appears to operate at higher C_P than the other purple symbols (should all be at 10% derate?).
- Many figures, e.g. 16, 17, non-dimensionalize distances with rotor radius $R$, whereas the accompanying discussion in the text refers to rotor diameters $D$. Please be consistent to allow a better direct interpretation between figures and manuscript text.

- Equations (12) and (13) seem to be functions on rotor radial location $r$ through the dependence of $C_t$ and $C_q$ thereon. However, Fig. 19 presents them as single scalar values. Should the small $r$ in (12) and (13) be a large $R$, or is there some implicit radial integration between Eqs. (12) and (13) and Fig. 19? Please clarify.

- Figure 2 is a nice visualization of the FVWM method, but it is not referred to in the text. Please do so in Section 2.