

Many thanks for reviewing the revised article. A response to the reported issues is given below in blue.

The paper describes a comparison between 19 simulation tools and the Dan-Aero experiment. The codes are either BEM based, Free Vortex methods or full CFD. The test cases are very little shear, significant shear and a yaw case. The BEM and the free vortex codes require as input described airfoil data and the results can depend a lot on these as also shown in the paper. The test cases unfortunately all are at a relatively high induction indicating that the overall flow is close to the turbulent wake case. The BEM results are sensitive to what Glauert models are used for the empirical  $CT(a)$  relationships when  $a > 0.3$ . This should have been better described for the applied BEM methods.

Text has been added to clarify the need for an empirical relation in the turbulent wake state. For further details about each individual code the reader is referred to the code description section of the final report of this IEA Task, freely available to download.

Some of the CFD codes also have a transition model, so why only show the fully turbulent cases, since including this could have made a better agreement with the experiment.

To promote consistency between CFD results, turbulent boundary layer modeling was used. As these result already over-predict loads in comparison to the experiment, using a transition model will probably further worsen the agreement due to the corresponding lift increase at these angles of attack.

It is very weird that the geometries for the CFD based models are different.

Retrieving a sectional blade slice in a 3D pre-bended rotor geometry can result in differences due to small inconsistencies like definition of radial coordinate, angular orientation of the sectional plane and direction of the chord line. What adds to that is the fact that different software packages are being used and the corresponding 'human factor'. Hence it can be considered evident that differences appear, which is representative for what happens in real life CFD applications. This is further clarified in the text.

For the shear case shown in Figure 9 it is clear that the BEM and free vortex codes underestimate the loads for an azimuthal position of around 0, indicating that the 2-D airfoil data used stall too early as also shown in the CFD synthesized airfoil data in Figure 11f. It could have been interesting to investigate and compare with some of the classical papers showing how to correct airfoil data for rotational effects.

As described in the paper a 3D correction was applied to the default airfoil data set. Also a classical stall delay model has been used (Snel method), which showed similar results. This is now added to the text. Unfortunately these models apply a correction mainly to the inboard part of the blade, whereas it is observed that largest discrepancies can be found mid-board and also outboard. Therefore the need for a stall delay model that works along the whole blade span instead of only the inboard part is stressed in the paper.

I strongly suggest and recommend that the geometrical data and the applied blade and airfoil data for the Dan Aero experiment be made public allowing others to compare their different aerodynamic tools.

It is stressed that this data is freely shared within IEA Wind Task 47 (open for joining to all parties) after signing a 'light' NDA as required by the industrial participants of the DanAero project. This has already been indicated in the 'Code and data availability section'.