Sections 2.1.1 and 2.1.2. I don't think it's necessary for the authors to explain the turbulence models used: these are well known to readers experienced in CFD and can be researched by the interested reader in provided references if necessary. It would be preferable to briefly go over advantages and disadvantages of each model and to elaborate on why they were used in this specific case. Also the choice of using two turbulence models in the same study seems odd, so further explanation on this aspect should be provided.

Figure 2: The scanned data points shown in figure 2 for the two ice-profiles appear to be quite far apart. From what is stated in section 3.2 the scanned profile is used to calculate the average sand grain roughness. Please justify the expected impact of the number of scanned points on the modelling.

Section 3.2: How can one differentiate between roughness and geometry? In other words what part of the geometry can be justifiably smoothed and included in the rougness wall functions and what part must be retained? A robust criterion is not provided here. See final comments for further remarks on this.

Section 3.2: Assuming roughness elements with a conical shape seems quite semplicistic. Please justify this choice

Section 3: Are rough wall functions used on the entire airfoil or just on the iced part? If it's the first, how does using rough wall functions on the entire airfoil influence the results? Perhaps it would be useful to check that all the wall functions proposed give reliable and consistent results for an airfoil with no icing.

Figures 4 and 5: Please explain these figures better. If the "fine" and "coarse" grid refer to cases where wall functions are used vs cases where they are not the authors must explain the difference in the size of the elements surrounding the airfoil. The coarse grid differs not only for the boundary layer but also in the flowfield. In my opinion this can significantly influence the results.

Figures 4 and 5: The meshes appear different in the fine and coars cases. "Fine" seems to use Cartesian cut-cell meshing, coarse uses polygonal cells. Please explain the differences.

Section 3.3: Was a mesh independence study performed?

Table1 : Why different Reynolds numbers? I suggest including the total number of elements of the mesh aswell.

Figures 7,9 & 11: The "Exp." Dataset appears to have duplicate data points for some AoAs. Fpr instance Fig. 9 AoA=12, 2 "Exp." Points are clearly visible. Please explain. Also, why are this figures relative to maximum lift? This could accentuate or diminish the difference between models depending on the value of maximum lift.

Section 4.4: As stated previously it is hard to judge differences in lift given the relative nature of figure 11. However, as it stands the differences seem pretty substantial even in the linear region.

Figure 12: here seems to be a considerable difference in figure 12b between momentum model and the others. This is not reflected in figure 11 at AoA=4°, why?

Line 240: This phrase is not clear, please revise it.

Section 5.1: Authors state that many icing profiles should be measured and averaged to get a good reference shape of the airfoils. What about the roughness values? How can those be estimated and used in computations?

Line 260: I suggest to edit figure 12 to include all of the tested roughness models. This would allow to compare if and where separation is predicted. Also, separation in this case seems like it is a consequence of the shape of the ice and not relative to roughness. If a model cannot correctly predict this separation can this be caused by excessive geometry smoothing? In other words, could agreement be improved is a different smoothing strategy was adopted? Or if no smoothing was adopted and roughness height measured differently?

Line 263: This is a good point. What kind of wall function is used in the non-iced part of the airfoils?