Reviewer 2

The authors investigate the ability of satellites to provide roughness length and displacement heights for wind resource assessment. They consider 3 satellite platforms which provide tree height, land cover, and leaf area index. Then, they use 3 forest modules to convert these data to roughness and displacement. They consider these satellite-derived estimates, along with other traditional ways (global land-cover maps; aerial lidar scans; manual digitalization) of obtaining these quantities, and perform power predictions at various sites using WASP. They take each sensor on a met tower and use it to predict the wind speed and power at all other sensors on that tower (vertical predictions) and other towers at the same site (horizontal predictions). A total of 10 sites are considered. The main result is that WASP-derived power predictions obtained with satellite vs lidar land characteristics were comparable (similar error of ~10-11%), indicating that satellite data can be a more affordable alternative to costly aerial lidar campaigns over forested terrain.

Major comments

The writing quality is somewhat poor. Subsection 5.3 and Conclusion are the most well-written ones in the manuscript. The rest is certainly readable, but it takes extra effort from the reader to sort through ambiguities and incoherencies. I would expect to be able to focus on the science and results being presented while performing a review, but the writing made it difficult and I spent a substantial amount of time just trying to understand the content that was being presented. Each paragraph should have a clear message and a reason to exist. Figures should stand alone with well-labeled axes and captions, and always show the units. The reader shouldn’t have to read the manuscript in detail to understand what the figure is trying to show. Especially since there are so many acronyms for the datasets/models and foreign names for the site locations, it can become difficult to keep track of what’s what if the text is not easy to follow. The results presented implicitly include a validation of the WASP models themselves. I suggest that the analysis be reframed to focus exclusively on the effects of what the authors are actually testing: several ways of obtaining $z_0$ and $d$. Either that, or the narrative of the paper should be expanded to include the WASP validation that is being carried out (even if WASP is being employed in a less traditional way, using only one sensor, without the ability to fit to the entire mast profile). It feels like 90% of my time was spent reading the introduction, data,
and methodology (19 pages) and then not many results were presented (less than 3 pages). Maybe that’s because some results are shown in the methodology? Either way, please consider including some of the results that are not shown (as per manuscript text) and more examples that show spatial variations. It would help the interpretation of the results if the readers could see for themselves the spatial distribution of $z_0$ and $d$ at all sites, which could be accomplished by adding a single figure with subpanels. It might also be helpful to spatially see the progression of data from satellite-derived quantities (h, land cover, lai) to $z_0$ and $d$ for the best-performing and worst-performing sites. Figure 7 is also lacking discussion of some key points: e.g. I don’t understand why the hand-digitized map is so bad for one of the MX sites; I don’t understand why the proposed method performs worse than MODIS/GLOB/CORINE for Østerild and the two MX sites, and it’s not discussed. It seems like Figure 7 can be modified and milked for a much more valuable discussion (and shown more than once, highlighting different aspects of the results, like e.g. vertical vs horizontal predictions). A key, unanswered question is: in what types of sites should one go through the trouble of deriving $z_0$ and $d$ from satellites vs using the global datasets? I would like to see the authors using WASP as they say it’s usually used (page 25, lines 483-487) for the sites with two masts: Alaiz, Finland, Perdigao, Sweden, Østerild to get the best prediction possible and then evaluate the effect of different $z_0$ and $d$ estimates. By using WASP the way it’s usually used, won’t you get results that are more interesting to your audience, and be able to tease out the real sensitivity to these lower-boundary forcing parameters?

We first of all have largely reorganised the manuscript, moving parts of the methodology to the results section. All figures have been remade with more clearly labelled axes. To produce a lighter version of the methodology we removed the discussions around RIX and the Perdigao and Alaiz sites. As we stated in the first version of the manuscript, one should not use the linearized flow model at these sites. In our original submission we filtered out these sites for most of the plots, but now we removed them altogether from the manuscript. As suggested by the reviewer we added a figure with the roughness length around all the sites. We removed the figure with the tree heights and the databases, because they could be added to table 4 instead. Large parts of results that were presented in the methodology have been moved to the results (mostly former sections 2.3 and 3.6), which resulted in a more balanced paper regarding the methods and results sections.

The suggestion to do profile fitting in WAsP was intentionally avoided. We want to focus on the value of an automated model chain as input for the WAsP model. By fitting a profile ‘by eye’ we introduce a degree of subjectiveness in the paper. Furthermore, the wind shear is largely determined by stability as well, so a good fit of the profile can be caused by a wrong roughness length and stability description that are cancelling each other out. We added this in line 492: “However, such a fitting process can be deceiving, because a good fit is influenced by both $z_0$ and atmospheric stability conditions. It may thus me subject to compensating errors.” Here we avoided that by focusing solely on maps that were put into the software without human intervention.

We added more information why some sites are not having lower errors. See lines 397-402 in the new manuscript (“The former is a very complex site, that is characterized by forest in all directions. One possible explanation for the higher errors at this site can be a clearing in the forest that is shown to the southeast in the Sentinel maps, but which does not appear in reality. At Østerild the dominating wind directions is from the west, which is characterized by many wind breaks. These are not generally detected in the Sentinel based roughness maps, but contribute significantly to a higher $z_0$. Therefore the background roughness for grassland that was used in the Sentinel maps (see Table A5), might be too low at Østerild.”). In any model comparison one cannot expect all sites to perform better, which is why we look at aggregated model performance
metrics in this paper. However, we have added hypotheses about why the results were worse at the Østerild site; the poor results of the hand-digitized map at Mexican site was caused by a mislabelled roughness map in our validation script, which has now been corrected.

We added a paragraph in the discussion where we highlight at which sites one can expect the largest impact of using tree height maps (lines 444-447 in the new manuscript). “From Fig. 6 it is clear that using the Sentinel maps of tree height rather than standard land cover maps has the highest benefit at sites where masts or turbines are in the middle of the forest (Ryningsnas and Sweden). Taking the displacement height into account for such sites leads to significantly lower $\epsilon_p$. Because there are seven masts at the Swedish site, leading to a large number of cross-predictions, this site has a large impact on the aggregated results (Table 4).

**Minor comments**

A very large amount of minor comments is annotated directly onto the pdf.

We appreciate that the reviewer has taken the time to provide concrete suggestions for how the text can be improved. We have taken these suggestions into consideration and adopted most of them in the manuscript. We have also replied directly in the annotated PDF, which is attached in this reply.

Please also note the supplement to this comment: [https://wes.copernicus.org/preprints/wes-2021-28/wes-2021-28-AC2-supplement.pdf](https://wes.copernicus.org/preprints/wes-2021-28/wes-2021-28-AC2-supplement.pdf)