Dear referee 2, please find the answer to your comments below:

Review of manuscript "High-resolution offshore wind resource assessment at turbine hub height with Sentinel-1 SAR data machine learning" by Louis de Montera et al.

As I am providing the second review of the manuscript, and have had the change to go through the first reviewer's comments (referred to as RC1), I can start with pointing out that I fully agree with the raised criticism and recommended changes for an improvement of the manuscript.

Below I list my own major and minor comments, some of them being a repetition of those in RC1 but also including some additional input:

(comments in order of appearance in manuscript)

[l 11] When referring to "Lidar measurements", first time here in the abstract, please specify what kind of measurements you mean explicitly – e.g. Doppler wind lidar, somewhat ground-based, measurements of wind velocity profile in range relevant for wind energy applications, or similar.ll Please check the overall manuscript for a sufficiently specific terminology with this respect.

We could write: "Doppler wind Lidar measurements from commercial LiDAR units deployed for offshore Wind Energy site investigations purposes"

[ll 13-14] When stating a bias, you also need to mention the considered reference – please add this here.

We will say it is the bias compared to Lidar measurements.

[ll 19-20] When reading the sentence "The accuracy of the wind power…" it becomes not clear how you get to the numbers you compare. You should also refer to the process of deriving a power curve from (any) wind data, i.e. the involved derivation.

We will make it clear that we multiply the Weibull distribution by the power curve of a typical turbine in order to obtain the extractible power.

I am also wondering if you really need to consider this (as I understood later, very simplified power curved derivation) for your study, or instead could focus on a derivation of wind power density.

The problem with deriving a total wind power density is that we would have a much higher error in the end when comparing with Lidars. This would lead to conclude that SAR satellites are not a good way to estimate wind resources, whereas, in practice, when multiplying the Weibull pdf by a power curve, the error is much lower. This is why we chose to use a power curve (we can explain that in details in the revised paper).

More precisely, the error is higher when computing the total wind power density, because one has to consider very low wind speeds and very high winds speeds. In these ranges, wind speed is difficult to measure precisely with SAR satellites and the machine learning is failing to produce accurate estimates due to the lack of sufficient training data. Our opinion is that, in these ranges, the turbine is usually not functioning anyway, so that there is no need to consider them in practice when looking at the extractible power with a power curve. So, to sum up, we chose this approach because it is closer to industrial applications and because it is more appropriate to test the use of SAR satellites. Estimating the total wind power density with SAR satellites seems however possible, but with more time and research in the future.

[l 39] I believe you should refer here to ground- (or bottom-) based Lidars.

ok

[ll 43-44] The sentence seems incomplete – add "… of meteorological conditions [that may impact this extrapolation]", or similar.

ok

[l 59] Be more specific here: "found to give good results" for what explicitly?

The correct reference was put in the wrong sentence 3 lines after this sentence. The reference is Optis et al., 2021 in which the authors compare machine learning and theoretical approaches to extrapolate the surface wind to higher altitudes. They conclude that machine learning is giving the best accuracy.

[l 89] Here and at other places where you introduce already available models/methods, please add a reference – in this case, for WRF.

We will give this reference:
NCAR/TN–556+STR
NCAR TECHNICAL NOTE
March 2019
A Description of the Advanced Research WRF Model Version 4

[Figure 1] I suggest to add a larger map to help locating this cut-out. Please also introduce the used abbreviations.

ok

[Figure 1 and Table 1] Are you sure that all these datasets are from floating lidars? I have at least some doubt with respect to LEG and IJM. Please re-confirm.

After checking, we confirm LEG, EPL and IJM are not floating Lidars but actually on platforms.

[l 115] From own experiences and also in line with available recommended practices, I would not fully trust the lower height (as 4 m a.s.l.) measurements from floating lidar systems – mostly from in-situ sensors heavily influenced by the structure itself. Please have some thoughts on this, and possibly consider some added uncertainty.

Ok. Anyway, we had time to improve our approach. In the revised paper, we will avoid extrapolating Lidars measurements from 40m to 10m using these unprecise buoys. Instead, we will do the geometric correction of SAR measurements with the US NDBC buoy networks (around 40 selected buoys measuring wind speed at 4m asl).

This will also clarify why we present two different methods: the advantage of the NDBC dataset is that there is a wide variety of azimuth and incidence angles in the training dataset. This training can therefore be applied anywhere, contrary to the second method, which is more precise, but requires on-site Lidars.

[Figure 3] This figure is not very informative – please review the design and information included. Overall, I think you can and should reduce the number of figures in the manuscript, possibly combining some of them.

OK. Figure 3 will be removed because it is not useful.

[Figure 4 and Figure 5] In these plots it may be helpful to have the coastline (or something else) as reference.

These figures will be removed and included in Figure 1 that will represent only 3 years of Sentinel1 data

[section 2.4] As pointed out above, I think, the used power curve is too simplified. I would suggest to either apply a more realistic power curve, or instead consider another quantity as e.g. wind power density for this investigation.

We will use a more realistic power curve:

https://nrel.github.io/turbine-models/DTU_10MW_178_RWT_v1.html

As explained above, we prefer not to compute the wind power density instead since the SAR is less precise in measuring low and high wind speeds.

[Figure 6] I do not think that this figure is really needed, instead you could add more details in the text.

Ok

[section 2.5] I am confused by your mentioning of "one passage every two days" and "passage times are separated by 12 h" – please be more specific here.

Ok. Yes, this is confusing. The satellite passes every two days, and this can occur at specific times: or in the morning or in the evening. The two possible time have a difference of 12h.

[l 222] As already stated above, please add reference for the applied methods – here the "two types of machine learning regressor[s]".

Ok

[l 229] Please specify how "the relative importance" is defined and derived.

Ok. The relative importance is based on the mean decrease in impurity (over the trees). We will give the python function in the text (`feature_importances_` attributes of random forest regressors in scikit-learn package)

[l 239] Also the statement "Random forest was found to outperform neural networks" needs more explanation / details.

The accuracy of the outputs of neural networks was found to be lower and the training time was much higher. We attribute the success of random forest to its ability to model non-linear relations in an easier way than neural networks, because it uses thresholds rather that activation functions.

[Figure 9] Add a legend to the plots and the details of the red curve (fitting parameters).

Ok. We will indicate the number of degrees of the polynomial fit and add a legend. We will also say the color is the density.

[l 266] Again, "machine learning" needs more explanation and details.

We will add it is a random forest regressor and give the parameters.

[Figure 12 (and Figure 14)] It is not clear to me, why you have not used a smaller range for the vertical axis – please revise.

We could use a -5% to +5% axis.

[l 323] Sentence "In this case, …" is incomplete.

The word 'the distribution' is missing.

[Figure 18 and following] Please re-arrange these plots for better comparability – combine several plots in one figure, for instance.

We could make 2 figures, one for each zone, with four subplots each (model wind power, Sar wind power, difference in percentage and quality of the data)

We would like to thank you for these helpful comments and interesting insights.

Best regards,

The authors