High-resolution offshore wind resource assessment at turbine hub height with Sentinel-1 SAR data and machine learning

Louis de Montera, Henrick Berger, Romain Husson, Pascal Appelghem, Laurent Guerlou, Mauricio Fragoso

REVIEW – round 2

GENERAL COMMENT:

The authors addressed most but not all of my previous comments. Despite the modifications to the approach, I feel there are still some fatal flaws (and several other minor issues), as I will detail in my comments below.
Please note that I have reviewed up to Section 3, as I really need to see my major comment on the methodology addressed before I can evaluate the results of the analysis, should the authors still want to pursue publication of this work.

MAJOR COMMENTS:

1. I still cannot see any practical application for the approach the authors proposed in the paper.
   In the first part of the approach (described in Section 3.1), you 1) extrapolate lidar winds down to 10 m and 2) apply a machine learning model where the SAR-derived wind speed is an input, and the target variable is the 10-m lidar wind speeds. You then use this SAR-corrected winds for the extrapolation to hub-height.
   Now, to me this approach has two fatal flows.
   The first fatal flaw I am seeing is that if one needs to have lidar data available (you use them to correct the SAR winds, Section 3.1), why would one need to extrapolate SAR data in the first place, since the lidar provides hub-height winds already?
   To respond to this concern, I could see an application of this approach when someone only has let's say a 10-m sonic anemometer (whose wind speed observations are used to correct the 10-m SAR winds), without any hub-height observations. But for this application to be possible, the authors would need to test the generalization of the approach they propose with a round-robin validation. In other words, the authors would need to answer the following question: "how accurate is this whole approach when applied at a site (i.e., in my example, where I only have near-surface wind speed observations) different from the one where it has been trained (i.e., in my example, where I have lidar observations which already give me hub-height data)?". Since the authors have multiple observational locations, they could do this exercise, but currently this validation is not done in the paper.
   However, even if this validation exercise were to be completed, the second fatal flow I am seeing here would still kick in. If one needs to have *any* 10-m observation

of wind speed to correct for the SAR data, why would one use the SAR data in the first place, instead of just extrapolating to hub-height the 10-m observations coming from the instruments needed to correct the SAR data?

SPECIFIC COMMENTS:

1. L. 33: other data sources can be used to estimate hub-height winds, for example reanalysis products.
2. L. 62: “Nevertheless, the analysis of Lidar data shows that, above 40 m, the power law is no longer accurate.” is a strong and very general statement. While it has some merit, it needs references.
3. Section 2.1 have multiple instances of weird spacing between words.
4. L. 135: why did you consider two lidars only to determine the exponents of the power law, which are then applied to all the lidars?
5. L.240: “The default hyperparameters were found to be the most appropriate ones”. How did you find this? Remember, your work should be replicable! A similar comment applies to line 286.
6. Figure 8: the y-axis label can simply be “SAR – lidar wind speed bias (%)”
7. Despite my previous comment, a data or code & data availability statement is still missing. You should add one even if your code cannot be shared – simply state it.