

## ***Interactive comment on “Improving Lidar-Derived Turbulence Estimates for Wind Energy” by Jennifer F. Newman and Andrew Clifton***

### **Anonymous Referee #2**

Received and published: 29 July 2016

#### General comments:

The paper present the L-TERRA algorithm developed for correcting lidar measured turbulence intensity. Correcting here implies retrieving the same TI as that measured by a sonic anemometer. The algorithm contains two main parts:

1. Correction based on physical considerations: correction for the volume averaging effect and correction for the beam cross contamination effect. Those effects are corrected with methods previously developed in other pieces of work.
2. Correction of the remaining error after correction with the physical considerations, using machine learning models. This second part is to my understanding the main novelty of the piece of work presented in the paper.

C1

The first part of the paper is rather well organized and clear:

Section 2 provides a complete review of the work achieved on the topic of turbulence lidar measurement and suggested solutions for correction. Section 3 gives a high level overview of the different modules of the L-TERRA algorithm, although some part should include some more information. Section 4 gives an overview of the three datasets forming the basis of the analysis presented in the paper. Section 5 demonstrate the relevance of accurate TI measurement for wind turbine power performance estimation; a fortiori AEP estimate and resource assessment. An interesting point highlighted in this section is that most of the significant errors in power estimation occur for small errors in TI. The second part, presenting the actual work and results, would gain very much by following a more systematic approach.

The following questions should be answered in the paper:

1. How much of the correction is done by the physical modules and how much by the machine learning module?
2. What is the remaining error, after correction by the physical modules, due to? Is there any systematic pattern in the remaining error? If not (in extreme case, if the remaining error looks random) can a machine learning approach really make an improvement?
3. To which extent the machine learning module can improve the correction from the physical correction modules? In other words, can the improvement compared to previous work cited in section 3 be quantified and clearly demonstrated?
4. In the introduction, the L-TERRA algorithm is presented as a method easier to apply than other methods for lidar turbulence measurement correction proposed so far (p2, l25). However, in the end, the method does not look really much easier (and maybe even more complex) to implement since it re-uses or adapts the methods qualified as complex to the WC, combine them and add the machine learning module. The

C2

simplicity of the approach needs to be further demonstrated.

Furthermore, a more straight forward evaluation of the results, like a direct comparison of the lidar measured turbulence intensity to the sonic measurements before and after correction (as done in figure 11) would lead to clearer conclusions. The conversion of the TI/TI error into power/power error is very interesting to demonstrate the impact of the TI measurement error on power prediction (as done in section 5), but it tends to confuse the analysis objectives when it is used in the results (like it is now done in section 6.2.).

Detailed comments:

(p: page; l: line)

P2; l.27-32: this paragraph is misleading and/or misplaced. It gives the conclusions of the analysis. I suggest removing it.

P7, l5: is it 4s second scan or 5 sec scan?

P8, l7: "Two methods were evaluated..." The flow chart in fig. 2 shows 4 methods to take care of the effect of volume averaging (2 for each type of wind speed). Does that mean only 2 of them have been tested? Which ones: "spectral filtering 1" and "spectral filtering 2"?

P8, l11-12: could you please provide explanations on how you have applied the method from Krishnamurthy to DBS scans?

P8, l17-19: could you please provide explanations on how you have applied those techniques to a 5 beam configuration?

P8, l24: "there is still some error..." Why? What is the remaining error due to? Does it mean the physical correction models applied previously are not good enough? The assumptions are not verified?

P9, l1: for the paper to be comprehensive by itself, it would be good to have a short

C3

description of each of the 3 machine learning methods. P9, l21-22:

1. Some of these results are rather surprising to me: internal temperature and pitch. Could you please comment on those and provide some information regarding the range of each variable and its correlation with the TI?

2. TI and sigma\_w were not correlated?

3. Have you performed the same sensitivity analysis for the other flat site to see if you get the same final predictor variables?

P10, section 4 1. In the introduction you mention 2 flat sites and one semi-complex. Which one of the three sites described here is the semi-complex one? My guess is BAO, but needs to be clearly stated as it can influence the results of the lidar measurements. 2. In the semi complex site, were the lidar measurements corrected for the effect of complex terrain (e.g. FCR) or the same wind speed reconstruction algorithm was used at all sites?

P12, section 5 The main reason for the poor results at BAO is probably the effect of the terrain on the flow. In complex terrain, the assumption of horizontal homogeneity is not verified, then the reconstructed mean wind speed includes some error and therefore does not correlate with the sonic measurements as well as lidar measurements in flat terrain. If the mean wind speed comparison to the sonic is poor, the TI comparison is also expected to be poor.

P13, l19-20: Does this mean that the correction algorithm for the radial wind speed (right part of the flow chart in figure 2) has not been tested yet (or at least the results are not included in this paper)? This sounds in contradiction with p13, l22-23 (and p6, l29-30) stating that "all possible combinations... were evaluated". Could you please clarify which of the process presented in flow chart in fig 2 have actually been used for the results presented and discussed in this paper?

P13, l 30-31: Only the optimal combinations for each site are presented in table 3. 1.

C4

How much difference was there between the different combinations? Was it significant?  
2. Is one of the combinations more robust than the others? i.e. could you, based on this analysis, recommend one combination or is the idea to always try all of them and pick the smallest error? (This is maybe to be included in the discussion in section 6).

p.14, l13-28: 1. From this analysis, it sounds like the cross contamination effect is the most important. How much the correction for this effect change the lidar does measured TI? And how much does the machine learning correction part change the lidar measured TI?

2. Figure 11 shows that the slopes in the linear regression are usually getting closer to 1 after application of the L-TERRA correction, which mean the mean error is reduced. But the scatter is increased ( $R^2$  is lower), so the improvement is actually mitigated. Moreover, this shows that the method does not necessarily provide better estimate of every 10 minute value of TI measure by the lidar, whereas it was demonstrated in section 5 that is what was needed.

P15, section 6.2: only MARS and RF are discussed. Support Vector Regression was not tested?

---

Interactive comment on Wind Energ. Sci. Discuss., doi:10.5194/wes-2016-22, 2016.