Interactive comment on “Assessing Variability of Wind Speed: Comparison and Validation of 27 Methodologies” by Joseph C. Y. Lee et al.

Anonymous Referee #1

Received and published: 19 August 2018

Overall I find this work to be a substantial project and a valuable contribution to wind energy and wind climatology communities. I have only a few comments that I ask the authors to consider regarding this manuscript and perhaps their future work.

(1) Because the IAV statistical results are based on your filtered data (e.g., Section 2.2), more explanation/justification of your methods would be useful. For example, why use linear regression when power is a nonlinear function of wind speed? What exactly are your criteria for identifying "underproduction for reasons other than low wind speed" and "potentially erroneous overproduction" (lines 132-134), and how confident are you that these are legitimate outliers? What is the proportion of "derived energy data" included in each of the time series for the 204 stations that require such data (line 143)?
(2) Fig 1: Given the geographic distribution of retained sites, is there a need to consider geographically weighting the analysis results so that the central Plains results (for example) are not unduly influencing your interpretation of the statistics?

(3) Fig 2b, c: What would these figures look like if plotted with the R2- and r-filtered data?

(4) Fig 6b, c: What are the characteristics of those sites that parallel the "line" that goes through the TX site? What makes them not deviate so much on panels a, d? Are these the same sites that show this pattern in Fig A2 b, c?

(5) I am a fan of MAD-based statistics but not necessarily to the exclusion of other types of statistics. It would be helpful and interesting to include some discussion on why the different metrics give different results and how they may highlight different aspects of what the wind speeds are like at these stations (for example, in reference to the Oregon site in line 382). You do acknowledge the potential utility of different measures in the Discussion, lines 593-596, but the paper itself seems to be focused on identifying "the one" measure that should be used. Is that your explicit intention?