Reply to Reviewer #2:

The author would like to thank the reviewer for the helpful comments. Some revisions are made to the manuscript according to them. For some comments the author has different opinions on, explanations are given in this reply. The author hopes the revised manuscript is acceptable for the reviewer.

The author assume that the referenced power law adequately describes the impact of boundary-layer stability, whereas the authors of the power law point out that it applies only to near neutral conditions. Such an assumption will often be valid for strong winds (U10 > 15 ms^{-1}), however for wind speeds <7 ms^{-1} the departures from neutral conditions are likely to be substantial. Furthermore, buoys measure winds relative to the fixed Earth where as stress, which drives waves, is dependent on surface relative currents. For lower wind speed cases the impact of currents could be substantial. While this is mentioned later in the manuscript, it would be wiser to address it earlier and perhaps in the quality control of the input data.

The author understands that power law is only valid for the condition of neutral stability. The main reason for the author to use the simplest power-law profile is to make the inputs and the targets of the model consistent with Voermans et al. (2020) which is probably the first paper trying to establish a model to estimate wind information from wave spectra. With the same inputs and the targets, the performance of the two models can be directly compared. Besides, the author thinks that the physical meaning of wind speed indirectly estimated from the wave spectra (sea surface state) is probably closer to the equivalent neutral wind speeds derived from sea surface backscatter (e.g., space-borne scatterometers and altimeters). Therefore, the air-sea stability-dependent wind profile was not used here although it can give a more accurate extrapolation of the real 10-m wind speed. The DNN model in this study only perform well for 3-20 m/s, where the differences of extrapolated 10-m equivalent neutral wind between different methods (e.g., power law, log, LKB) are much smaller than the error of the wind estimation model itself (with respect to standard deviation). Therefore, different 10-m wind adjust method will also give similar results in this study.

Regarding the impacts of the ocean current, indeed, they are important for the cases of low wind speed. The impacts include not only the relative wind effect (the stress of wind is dependent on the relative movement between wind and currents), but also the “relative wave effect”, that is, the phase velocity of high-frequency waves and the current velocity are at the same order of magnitude during strong currents so that the dispersion relation will be distorted by the current. However, the current data are not available from the buoy data so that data with strong current cannot be discarded during the quality control. The impact of currents can only be regarded as the noise for wind estimation from buoy wave spectra in this case. Therefore, the author still feels it might be better to introduce this effect in the discussion of errors.

What (if any) quality control was applied to the data? Frankly, a paper should not be submitted without this information. If any quality control was applied, why was it applied and why is it likely to be sufficient? If it was not applied, then why is it not needed?

The data has been already quality controlled by National Data Buoy Center (NDBC) where the data is provided. The detailed information on NDBC data quality control can be seen from https://www.ndbc.noaa.gov/qc.shtml. In the description of the NDBC buoy data, the manuscript mentioned that "Many buoys from the National Data Buoy Center (NDBC) coastal-marine automated network can provide quality-controlled in-situ wave and wind measurements". Therefore, the author simply removed the data with bad-quality flags, and this has been clarified in the revised manuscript: "After removing the data records with bad-quality flags, more than 1.7 million records….were used in this study".

The Fourier characteristics of waves are poorly described and need to be much more clearly explained.

The author is a bit confused about the "Fourier characteristics of waves". Generally, the wave spectrum, which is the Fourier characteristics of ocean waves, is obtained by applying the Fast Fourier Transform to the time series of displacement, azimuth, pitch, and roll of buoys. However, this is almost common sense of the wind-wave community that needs not be explained in the manuscript. The author feels this is probably not what the reviewer is referred to. Therefore, it will be nice if the reviewer can explain a bit more on this comment.

The author tried to give a more detailed explanation on the five Fourier coefficients from the buoys in the revised manuscript, which now reads: "The buoy wave data includes five Fourier coefficients of waves for different frequencies in the range of 0.02-0.485 Hz (47 frequency bins) derived from the translational or pitch-roll information of buoys. The five Fourier coefficients are wave variance spectral densities (E) which describe the wave energy for each frequency, mean and principal wave directions for each frequency ($\alpha 1$ and $\alpha 2$), and first and second normalized polar coordinate of the Fourier coefficients ($r1$ and $r2$) which describe the directional spreading about the main direction the for each frequency. The five Fourier coefficients of different frequencies are the minimum requirement to reconstruct the directional wave spectrum." After this revision, the physical meaning of the five Fourier coefficients should be clear for the reader.

Does the lack of approximately uniform distribution over the parameter space impact the quality of the results, particularly for conditions that are poorly sampled? Normally there is a very large impact, with the results only applying to the conditions near the peak of the probability distribution.

Indeed, the lack of uniform distribution of wind speed will lead to large errors for the conditions that are poorly sampled. This will also lead to that the model performs the best near the peak of the probability distribution. These are also parts of the reasons for the author to have Figure 2c in the manuscript. Both of the two effects can be seen from Figure 2c of the manuscript where the error metrics were given as a function of wind speed. The error is the smallest (~1 m/s) for the wind speed of 2-10 m/s where we have the largest sample size and becomes large for extreme wind speed where very few samples are available. However, this is inevitable since it is difficult to a have large sample size in extreme wind cases, and the air-sea interaction becomes much more complicated during extreme wind (e.g., strong spray and surface wind-driven drifts). With Figure 2c, one can know not only the error property of the model but also the impact of the skewed distribution of wind speed. One can conclude from Figure 2c that the lack of approximately uniform distribution does not have a large impact on the performance of the model, at least for moderate wind speed between 2-17 m/s. Some revision has been made to the discussion on the error of wind speed DNN, and the above explanation has been included (L136-145 of the revised version).

Are the different Fourier components combined to produce a better result? I assume so, but the math suggests otherwise.

Yes, the combination of different Fourier components can produce a better result than only using only one set of Fourier components. The author fails to understand why the math suggests otherwise. It will be nice if the reviewer can explain a bit more on this point.

Can the one hour delay be better demonstrated with statistics and an appropriate graphic? It should be possible to show this result in a manner than much more clearly illustrates the width of the peak correlation or a time offset in the DNN.

This is a very good suggestion, and this figure is added to the manuscript as Figure 3 (Also shown below). The author also change the used wind speed data to the 10-minute resolution "continuous wind" from the buoy, according to the suggestion of Reviewer #3. Figure 3 indicates that the best correlation between DNN-estimated and direct-measured wind is under the condition of a time offset of 40-60 minutes.

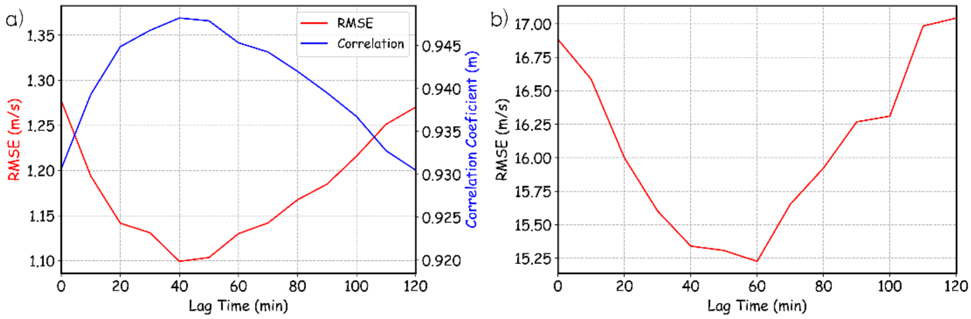

Figure R1. Figure 3 in the manuscript: (a) The RMSE and CC of the DNN-estimated wind speed as a function of lag time between wave and wind measurements (waves' end sampling time minus winds' end sampling time). (b) The RMSE of DNN-estimated wind direction as a function of lag time between wave and wind measurements for wind speed higher than 7 m/s.

Errors in the results are attributed to strong currents, but these errors are far larger than expected due to currents (at least in Figure 3a, and unlikely in 3b). Buoys don't survive long in such strong currents. Please consider alternative explanations or find evidence that the currents do exist.

The reviewer seems to misunderstand the condition in original Figure 3a and 3b. Actually, they have nothing to do with the ocean currents. The author used them to explain why the RMSEs for the two buoys shown in the Figure are large (the two buoys are two of the buoys with the largest overall RMSEs), and to show that this wind-estimation DNN model can serve as an additional quality control/monitoring method for wind and wave sensors on meteorological buoys. For example, in original Figure 3a, after 26-Jan-2014, the difference between the measured and estimated wind speed suddenly becomes very large. It is noted that the measured wind speed remains lower than 5 m/s for more than 15 days. This is unrealistic for ocean winds, therefore, there must be something wrong with the measured wind speed. However, these data are not screened out in the NDBC quality control procedure. A similar condition happens in original Figure 3b, where stable bias between the measured and estimated wind speed was suddenly observed. Because the DNN model is unbiased and time-independent, such a systematic underestimation or overestimation of U10 for a long period has to be attributed to the problem of either wind or wave sensor. Because the buoy data has been quality controlled by NDBC, such conditions of bad-quality data were only identified in the two cases in Figure 3. Even for the buoy the strongest impacts by currents, the error is not that large, as shown in the following example (Station 46087 at the Strait of Juan de Fuca where tidal currents are strong, and this is one of the buoys with the largest overall RMSE of DNN-estimated wind speed):

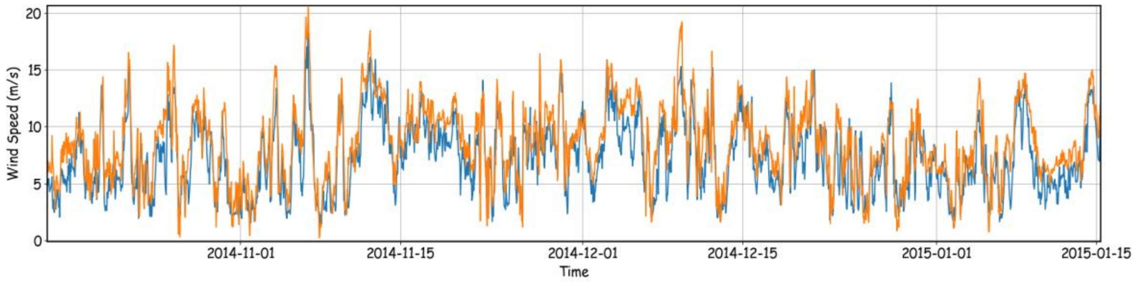

Figure R2. Time-series comparison of direct-measured (blue) and DNN-estimate (orange) wind speed for Station 46087 from 16-Oct-2014 to 15-Jan-2015.

In summary, the methodology needs to be greatly improved. The accuracy assessment should not be presented as an overall single value for the dataset, but rather as a function of wind speed. The explanation for the cause of large errors is highly unlikely to be correct, although I appreciate the authors efforts to provide an explanation. The lack of check the quality of the input data, the physics of the adjustment to a 10m wind, and poor assessment of the quality of data should be addressed.

The author believes that most of the points in this paragraph of the reviewer comment have been covered in the above responses except for the accuracy assessment. Of course, it is more reasonable to describe the error as a function of wind speed, and that is exactly what has been done in the manuscript. The error as a function of wind speed is shown in Figure 2c and presented in many places in the manuscript. However, many people in the community are also used to using a “typical” number to describe the error, maybe for simplicity. For instance, we often say that the error of wind speed, wind direction, and wave height for NDBC buoys are 1 m/s, 10°, and 0.2 m, respectively (e.g., https://www.ndbc.noaa.gov/rsa.shtml). The ocean remote sensing community also often says, for example, that the scatterometers have ~1 m/s error of wind speed (There are many papers, data product handbooks, and even textbooks, saying so). Therefore, the author thinks it is OK to present an overall single value of RMSE somewhere in the text.

Again, the author thanks the reviewer for his/her helpful comments.