Public Comment from Adam Theisen

Nice work analyzing this long-term dataset from the ENA radar wind profiler. The detailed calibration approach and evaluation of instrument performance over nearly a decade provide valuable insight into data quality and long-term stability. There are a few minor points that may be beneficial to review for possible updates.

- 2.2 KAZRARSCL –
  - L120: The KAZR operates in a precipitation mode as well in addition to GE and MD modes but is not included in KAZRARSCL at this time.
    - Correction: KAZRARSCL does use the PR mode.
  - L121: The KAZR is not calibrated using a corner reflector, which is generally used for the SACR. The KAZR data in KAZRARSCL has not been calibrated. The b1, calibrated and corrected, KAZR dataset for 2020-2024 should be released by ARM soon. https://arm.gov/publications/tech_reports/doe-sc-arm-tr-320.pdf

Thank you for the comment, this has been corrected in the manuscript.

- 2.4 Eddy correlation flux measurement system
  - L145: ARM's ECOR only has 1 gas sensor. It's deployed a methane sensor at NSA and OLI in the past but never at ENA.
  - It should be noted that the citation is for the ECOR data, Sullivan 1997 points to data from Oklahoma. I believe the DOI that's needed is http://dx.doi.org/10.5439/1025039
  - Another note is that the uncorrected fluxes are in the 30ECOR data product. It is recommended that the corrected fluxes in the QCECOR product be used until the ECORSF data product becomes available (2024-09-17) and provides corrected fluxes in the b-level datastream.

Thank you for these comments. We have clarified these details in text and will swap out the 30ECOR data for the recommended QCECOR data.

- Data Availability
  - Some of the DOI links go to blank pages in Data Discovery without data. Please verify the DOIs and reach out to ARM (Contact Us on arm.gov) if there are any issues.

Thank you for letting us know. We've gone through and corrected the faulty dataset citations.

RC1

This manuscript describes how the Department of Energy ARM program's 1290 MHz Radar Wind Profiler (RWP) deployed at the Eastern North Atlantic atmospheric observatory was calibrated. While the subject matter is worthy of publishing in the journal of Atmospheric Measurement Techniques, the manuscript contains many conceptual errors and does not use standard radar terminology that makes the manuscript hard to read and understand. Most of the errors this reviewer found should have been caught by co-authors and corrected before submission.

This manuscript found that due to degradation of the RWP hardware, 39 months (over 3 years) of the 10-year record are not suitable for atmospheric research (see Line 300). This is a significant finding. Yet this finding is not mentioned in the abstract.

While I would like to see a future version of this manuscript published, the current manuscript needs a lot of work to clarify the work that was done.

Note that this reviewer read the Abstract through Section 3 and did not read the last 10 pages.

We thank the reviewer for their insightful comments and suggestions for improving the manuscript. In-line responses below.

The following specific comments are in sequential order.

1. Line 31-44, Abstract. The abstract is misleading in that it does not state that the RWP's performance degraded over the 10-year period. The abstract does not mention the calibration constants changed over 30dB in the first 5 years, causing some radar parts to be replaced. One of the manuscript's objectives was to document the RWP performance, but this significant degradation of performance is not mentioned in the abstract.
   Great point. The abstract has been adjusted accordingly.

2. The range of values of '-84 to -44 dB, -15 to 9 dBZ, and 12 to 21 dBZ' are values taken from different operating modes and do not have much meaning when combined like this. It is more important to tell the reader that the calibration constant increased from A to B and the minimum detectable reflectivity increased from X to Y over the first 5 years of operation.
   Agreed, the abstract has been adjusted.

3. Line 92-94. In this introduction paragraph, inform the reader that all instruments were deployed by the US Department of Energy ARM program and all input data are available to the public.
   Added, thank you.

4. Line 96. Incomplete sentence.
   Fixed, thank you.

5. Line 96 and 97. The first sentence says the radar was operational from September 2014 through February 2025. The second sentence says there was a 14-month gap in operations. Both sentences can be true. Rewrite these sentences to reflect the 14-month gap in data collection.
   Thank you for the comment. This has been rewritten for clarity.

6. Line 99. If there was heavy corrosion to the antenna and phase shifters, was there any degradation to the data quality or height coverage of the radar. Inform the reader that there was decrease in height coverage of the radar and the main subject of this manuscript.
   Thank you for the comment. This sentence has been adjusted to acknowledge the decrease in data quality and height coverage.

7. Line 103-104. The manuscript needs to be consistent in describing the RWP 'modes' and the RWP 'beam directions'. It is incorrect to say that the "wind mode has three beams" and then refer to the beams as the "vertically pointing mode" and the "oblique mode". The manuscript needs to be clear that the 'mode' refers to the operating parameters as listed in Table 1, and the individual beams refer to specific beam pointing directions. This correction is needed in several places in the manuscript.
   Good point, thank you for pointing this out. Here and throughout the manuscript, we have corrected and standardized our terminology to reflect our analysis of the different radar beams.

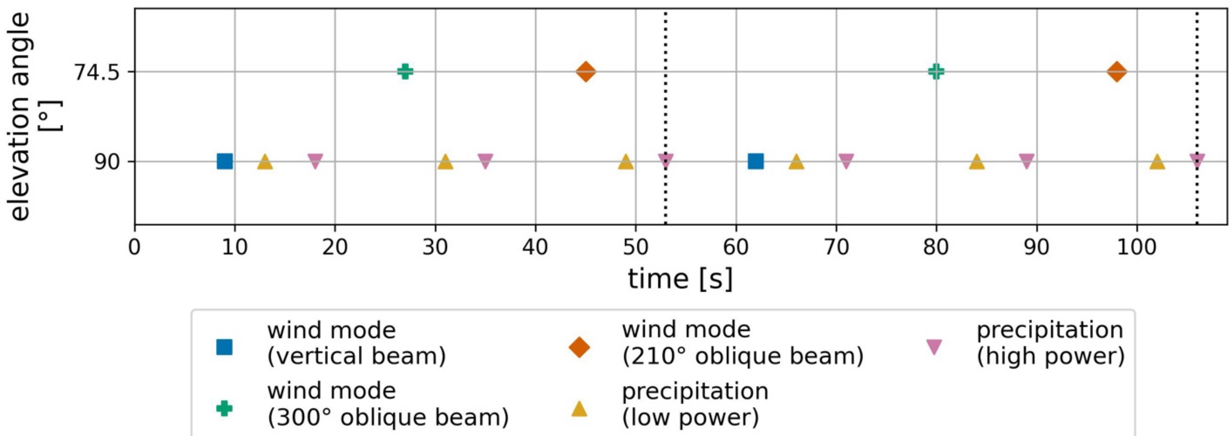
8. Line 104. There is no quantity with the name "Doppler spectra". The term Doppler refers to a velocity or a frequency. Also, the radar produces a power spectra. Thus, the RWP produces a Doppler velocity power spectrum at each range gate. Please change the manuscript so that the first reference to the RWP spectra is called 'Doppler velocity power spectra'. After this introduction, the manuscript can revert to a common name of spectra.
   Thank you for the comment. This section has been reworded to accurately convey the radar function and convention as you have described.

9. Line 104. Moments are not collected, they are calculated from the spectra.
   Thank you for the comment. Fixed.

10. Line 106-107. The sentence starting with "Between each wind mode …" is poorly written. Verbs do not match nouns. For example, "modes are operated" does not make sense because the modes do not operate. The RWP operates in different modes. Please re-write this sentence.
    Thank you for the comment. Vis-à-vis our response to #7, these lines have been rewritten.

11. Line 109. The full cycle time of approximately 53 seconds does not match the times listed in Table 1. In Table 1, the wind, long pulse, and short pulse modes require 27.6 seconds (3x9.2 seconds), 4.5 and 4 seconds, respectively. This adds up to less than 37 seconds. The text in the manuscript needs to match the values in the tables.
    This is because both precipitation beams are repeated between each wind beam. For one

period, the order is: wind (vertical), precip. short pulse, precip. long pulse, wind (300° oblique beam), precip. short pulse, precip. long pulse, wind (210° oblique beam), short pulse, precip. long pulse. These dwell times add up to 53 seconds for a full cycle. Because this is a bit confusing to read, we have added a new figure to the text visualizing this cycle, which is also included here:



12. Line 109. Again, the manuscript needs to be clear when discussing modes and beams. There are 3 modes described in Table 1.
    Thank you for the comment. Vis-à-vis our response to #7, these lines have been rewritten.

13. Table 1. The units for Inter Pulse Period (IPP) are wrong. They should be microseconds and not milliseconds.
    Good catch. Fixed.

14. Line 114. The RWP does not have a “vertically-pointing wind mode”. This inaccurate description of the RWP operating modes and the beam pointing directions is very confusing to the reader. This error occurs is several locals in the manuscript.
    Thank you for the comment. Vis-à-vis our response to #7, this line has been rewritten.

15. Line 117-124. The paragraph describing the KAZR radar and the ARSCL data product contains several conceptual errors, contains a misrepresentation, and has several grammar errors that this paragraph needs to be re-written. Specifically:
    1. Line 121. The KAZR radar is not calibrated with a corner reflector! This is a major error that raises questions about other errors that this reviewer is not catching.
       Thank you for catching this. Removed.

    2. Line 121-122. As written, the sentence referring to the work of Clothiaux et al. 2000, implies that the current work applied the merging and noise-filtering of Clothiaux et al. This is not true, and is a misrepresentation of the work not performed in this manuscript. The merging and noise-filtering described in

Clotheaux et al. was performed by ARM and is available to the public via the KAZR-ARSCL data product.
Thank you for your comment. We never meant to imply that we had done the labor of processing KAZR data according to the ARSCL algorithm and have altered the text to make it clear that this work was done by ARM as part of the VAP processing. To be absolutely clear, claiming that we did the work in processing KAZRARSCL data was not *our intention here.*

16. Line 137. Wow, the laser disdrometer is not very accurate if it underestimates drops with "mean hourly rain rates exceeding 2.5 mm h-1". Change the manuscript to have the correct units.
Thank you for the comment. While these units were correct (see Tokay et al., 2013, doi: 10.1175/JTECH-D-12-00163.1), the referenced study was characterizing the performance of the original Parsivel instrument, rather than the Parsivel2 (employed here). This is an important correction, as it does not suffer the same biases described here. We have adjusted the text accordingly.

17. Lines 150-160, first paragraph of section 3.1 Noise Filtering and binning. This paragraph does not describe the filtering process with enough clarity to be repeated by an independent researcher. The paragraph contains terms that are inconsistent with engineering principles. Examples of non-sensical statements include:
    1. Line 151-152. How are "daily profiles" of SNR binned to a "uniform time resolution"? This does not make sense to me.
    After cleaning the data, all data are binned to a 1-minute time resolution to match the disdrometer for calibration. This has been clarified in text.

    2. Line 152-153. How is the "noise floor" estimated from the SNR? The noise floor cannot be estimated from SNR measurements that contain atmospheric signals.
    3. Line 153. What are the visual inspection criteria used to determine the noise floor?
    (Responding points 2 and 3 here.) Thank you for calling attention to the vague description of this process. While interrogating this section of the manuscript, we decided to change this section of the methodology to instead rely on the in-built quality control flagging for removing pixels without significant signals. This is a more repeatable methodology for other scientists. As such, this section has been substantially altered and Figure 2 has been removed.

    4. Line 154. What is the "domain"? Does the domain span some or all range gates? Is there a temporal component to the domain?
    Thank you for the comment. Here, we refer to the temporal domain. In other words, we are referring to the ~24 hour period contained within a single file. This has been clarified in text.

    5. Line 157. What is the frequency response of the high-pass filter? Are the daily mean profiles being used or are the minute data being used in the high-pass filter?

6. Line 158. What is a ‘pixel filter’?
   Thank you for these comments. Per my response to points 2 and 3 above, we have replaced this aspect of the methodology.

7. Line 158. What does it mean to have “noise pixels’ when the measurements are SNR?
   Thank you for the comment. Here, we meant to describe a pixel that doesn’t contain a significant signal, or, as you accurately assess below, is below an estimate for minimum valid SNR.

18. Figure 2. What is the definition of “noise floor” when the measurements are signal-to-noise ratio? How is the noise power estimated from just the SNR? Is this figure showing some estimate of the minimum valid SNR level, but incorrectly referring to this estimate as a “noise floor”.
19. Figure 2. Is the “noise floor” estimated for every range gate? The minimum SNR will be different at low range gates than at far range gates. Include in the manuscript the range gate dependance of the “noise floor”.
    Thank you for the comments. Yes, this figure erroneously labeled “noise floor” on the Y-axis when we meant to describe an estimate of the minimum valid SNR level. As described above, we have changed the methodology and, accordingly, removed this figure from the manuscript and analysis.

20. Line 165-166. The sentence on line 165 states that ground-clutter pixels are removed. The sentence on line 166 states, “…the artifact can be suppressed without removing actual measurements.” How can these two sentences be consistent? The text implies that data is removed, yet, data is not actually removed. That is an interesting concept. Please correct the manuscript.
    Thank you for the comment. This has been corrected in text.

21. Line 167 and Table 2. What are the operating parameters during epoch A?
    The operating parameters are detailed in Table 1, we will correspond with the instrument mentor to see how/if the operating parameters changed at all between epoch A and B. Currently (to our knowledge), there is no published information regarding an operating parameter change.

22. Line 151 and 174. Line 151 says that the daily profiles were used to estimate the noise floor, yet line 174 says that the profiles were binned to 1-minute resolution. Please clarify the manuscript to remove this ambiguity.
    Thank you for the comment, we have clarified the manuscript.

23. Lines 109 and 176-177. Line 109 says that the two precipitation modes occur between the wind mode. Lines 176-177 says that the wind mode could occur once or twice in a minute, and the two precipitation modes occur 3 or 4 times in a minute. How do these sequencies match the 53 second cycle time listed in Line 109? Please make the changes in the manuscript.
    Thank you for the comment. This is because both precipitation beams are repeated

between each wind beam, and we didn't properly specify that it was the vertical *wind beam that can appear 1-2 times per minute. Clarified in text, see response to #11 for a more detailed discussion of this issue.*

24. Figure 3a and Figs. 4a and 4c. There is an increase in SNR (Fig. 3a) and reflectivity (Figs. 4a and 4c) at the lowest range gates. Is this increase real or an artifact?
Good catch. For Figure 3a, the high SNR at the lowest range gates is meant to be depicted, to demonstrate the artifact that gets removed after. However, for 3a and 3c (now 6a and 6c), these figures accidentally used an older version of the files which didn't properly clear the artifact. Fixed, new figure in the current version.

25. Line 184. Where the 915 MHz RWPs used in the Tridon et al. and Williams et al. work manufactured by the same company that built the 1290 MHz RWP deployed at ENA? The manuscript should mention the manufactures and discuss how the operating parameters impact the collected data.
Thank you for the comment. As you've alluded to, the radars used in Tridon et al. and Williams et al. were manufactured by a different company than the one described in this study. We have included additional commentary on this in text to clarify this difference.

26. The use of "minimum detectable signal" is another example of not using radar terminology correctly. The manuscript describes how the minimum detectable reflectivity is estimated, not the minimum detectable signal. Please change the manuscript.
Thank you for the comment. Fixed.

27. Figure 6. It is confusing to label the vertical axis as 'Sensitivity', when the value being plotted is the minimum detectable reflectivity. The figure would have more meaning to the reader if the vertical axis label was changed to minimum detectable reflectivity.
Thank you for the comment, this has been changed.

28. Line 300. Identifying 39 months of data to be 'bad' is a significant result. This should be stated in the abstract.
Thank you for the comment. This is reflected in the abstract.

RC2

This manuscript describes the characterization of radar wind profiler (RWP) measurements at the ARM Eastern North Atlantic observatory (ENA). The dataset is then used to derive parameters that represent the structure function of the refractive index and dissipation rates of turbulence kinetic energy. The authors then discuss the climatology of these parameters and relate the findings to previous studies of turbulence in the marine boundary layer.

The description of the RWP characterization was thorough and followed methodologies from previous studies but was sometimes confusing, for example, referring to terms that were not introduced until later in the manuscript. It also seemed that there were details discussed that were not correct with regards to the ARM datasets, which could have been found in the technical documentation available on the ARM website. Several grammar errors and inconsistencies in the figures suggest the need for thorough proof reading in the revision.

The abstract and introduction do not explain clearly what the goal of the study is from a scientific perspective though it became more apparent toward the end. It would be helpful if you mentioned more specific findings in the abstract and a description of what you were expecting to learn going into this study, and from the climatology.

A well characterized RWP dataset would be valuable for other researchers, and I hope that you will make it publicly available. For example, you were able to identify the "good and bad" days. Having this quality flag and corrected data (if possible) would be very valuable.

Overall, this manuscript provides a thorough analysis of RWP data and making these issues aware to other potential users is valuable and worthy of publication. As mentioned, there are some inconsistencies and areas where the analysis is confusing, but these can be easily clarified, and I have pointed out some specifics below. I recommend this manuscript for publication once these revisions are complete.

We thank the reviewer for their insightful comments and suggestions for improving the manuscript. Responses in line below.

Specific Comments

Fig 1b: It would be useful to show the location of the ENA site on the map.

Thank you for the comment. The ENA site is at the center of the windrose on the map, we will alter the figure to make this clearer.

Datasets: please state clearly the data level that you used for both the KAZR and RWP datasets and provide associated DOIs for data provenance and reproducibility purposes. For ARM data, a-level data is uncalibrated, b-level data is calibrated, and c-level is typically derived quantities using calibrated and corrected data. KAZR data at ENA was only recently calibrated to the b-level so it is most likely that the KAZR data in the ARSCL product that you are using is uncalibrated. Also, a corner reflector is not used for calibration. See the ARSCL technical report

for information about KAZR calibration. DOE Tech. Memo. ARM VAP-002.1
Thank you for the comment. We've clarified the properties of the data in the body of the main text as well as in the Data availability section at the end of the manuscript.

Figure 2 shows the signal to noise ratio (SNR) for the RWP and 4 data epochs but the discussion around this figure and the development of epochs is related to sensitivity that is not discussed until much later in the manuscript. I suggest reorganizing the manuscript to introduce and analysis the specific parameters before defining the epochs and describing the calibration techniques.

Thank you for the comment. Agreed. The manuscript has been significantly reorganized for clarity.

Section 3.2 Calibration technique: Suggest adding some additional details regarding calibration method. It seems all the elements are there, but the order of the description does not make it easy to follow. Suggest describing the calibration method in words first rather than jumping into equation (1). For example, state that you calibrate the RWP by comparing the surface reflectivity derived from surface disdrometer measurements described in Sec. 2.3 with RWP reflectivity during raining periods. Then describe the calculations of the reflectivity from the RWP and disdrometer. Several times in this section you refer to following the references from papers by Williams and Tridon but don't clearly describe the method in a logical way.

Thank you for the comment. We have significantly reorganized the paper including this section, and have moved the relevant figures and text around significantly

Lines 200-205: You state that 120 surface disdrometer measurements with >20 dBZ are needed on a given day – are these 120 continuous measurements?

They are not. This has been clarified in text.

Fig. 5: The 'x' symbols on the figure are very hard to see and distinguish from the larger, bolder 3-month average dots. Suggest making the 'x' symbols bolder in color.

Thank you for the comment. We're hesitant to make them too *bold, such that they make the bolder averages in the "foreground" more difficult to read, but agree that as-is they're practically invisible. To reflect this, the figure has been altered.*

Line 259-260: There are grammatical errors in this sentence.

Fixed, thank you.

Fig. 7: Line 290: You reference Figure 7b but it should be Figure 7d. Also in Fig 7 in the caption you refer to the KAZR ARSCL panel as '(e)' but should be '(d)'.

Fixed, thank you.

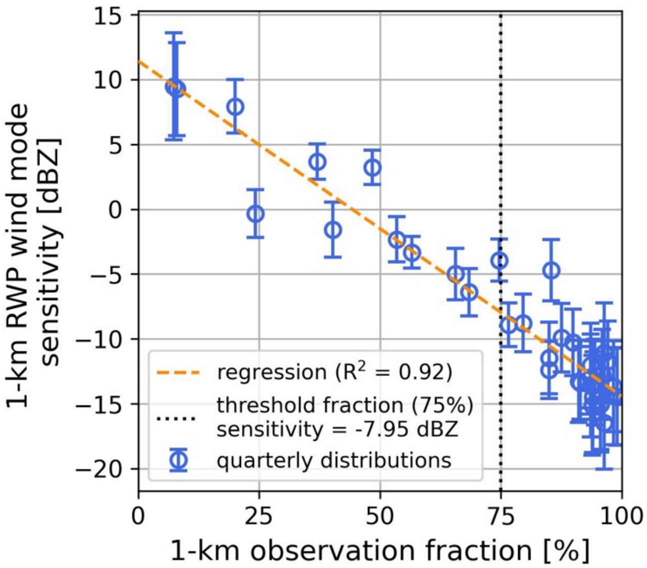
Lines 299-300: Explain how you determined this -7.5 dBZ threshold.

Thank you for the comment. We have rewritten this section, added a new figure, and refined this methodology. Previously, this was based on qualitative analysis of the profiles and matching where the radar observations started to look bad with the measured minimum detectable reflectivity (former "sensitivity"). Unfortunately, this is not exactly a well-described metric or replicable process! So, we altered this section to be more statistically rigorous.

Essentially, this all is motivated by the need for a "cut-off point" in the data since the sensitivity diminishes gradually during the time period. When a RWP is performing optimally it should retain – in contrast with hydrometeor-oriented radar (such as KAZR) – consistent valid observations at low altitudes throughout the day, as clear-air motions should be sensed more-or-less continuously. Therefore, we can infer that periods with consistently low observation counts at low altitudes are more likely to be the result of missing data from degraded sensitivity and are, consequently, not a robust representation of the day's weather. However, gating data purely on average observation frequency has its own issues. For example, one may imagine a period that recorded fewer average observations due to a higher-than-average frequency of low clouds, impeding observations above.

With this in mind, we regress our derived quarterly 1-km wind mode minimum detectable reflectivity (formerly "Sensitivity", from Figure 6 in the original submitted manuscript) on 1-km observation fraction. We define this new metric as the fraction of the day for which observations are recorded at 1 km, with 100% being equivalent to constant observations all day and 0% being equivalent to no valid observations all day. These results are shown in the figure below, now added to the manuscript:

With an R^2 of 0.92, we find that the averaged 1-km minimum detectable reflecitivity explains the overwhelming majority of the variance of missing data in this climatology. Based on this information, we set a threshold of RWP sensitivity at -7.95 dBZ (black dotted line in this figure), which corresponds to 75% observation fraction. Admittedly, this 75% threshold is an arbitrary



value based on the idea that it would be best to have at least 75% daily representation when deriving climatological statistics. Through this threshold, 39 months of data are removed.

Summary, lines 460-468: You mention comparing to ESMs, but this was only loosely mentioned in the summary, and it was unclear how you would compare the Cn^2 and ε to model output. Can you expand on this comment?

Good question. When writing this passage, we were thinking about interrogating and improving the performance of boundary layer parameterization schemes like CLUBB in E3SM. Meaningfully comparing ε directly would be impossible, but we can use measured ε to constrain relevant CLUBB tuning parameters (such as C1 or C2) in a perturbed parameter ensemble. Admittedly, the pathway towards utilizing C_n^2 in informing ESM improvement is less straightforward. As such, we will de-emphasize C_n^2 retrievals in this context and instead discuss the merits of having a large library C_n^2 retrievals for studying boundary layer processes in the context of LES and observational case studies. Thank you for this helpful comment.

References

Tokay, A., Petersen, W. A., Gatlin, P., & Wingo, M. (2013). Comparison of Raindrop Size Distribution Measurements by Collocated Disdrometers. *Journal of Atmospheric and Oceanic Technology*, *30*(8), 1672–1690. https://doi.org/10.1175/JTECH-D-12-00163.1