This study evaluates the performance of reanalysis, regional mesoscale simulations, and purpose-built wind datasets in a promising yet sparsely observed offshore environment near Hawaii. Using a year-long lidar buoy deployment off the eastern coast of Oahu, it characterizes the local wind resource and validates four datasets (ERA5, UH-WRF, NOW-23, GWA3) at different spatial resolutions. The region's persistent trade winds make it favorable for wind power, but limited hub-height observations underscore the need for accurate data. By comparing multiple products in this uniquely characterized setting, the work is relevant to the sector and appears suitable for publication pending the revisions below.

Major concerns

1. Attribution to PBL choice

The manuscript notes differences in PBL schemes across datasets but does not isolate the PBL contribution with, for instance, a sensitivity test. The study even presents several diagnostics informative of boundary-layer processes (e.g., stability-stratified errors, diurnal cycle, vertical shear), but the connection from these results to PBL-scheme attribution is not made explicit. If this is the case, please state explicitly how each diagnostic reflects expected PBL behavior, or soften the claim and note that other factors are in play.

2. Shifting focus from validation to campaign report

The buoy campaign is essential for the validation, but exhaustive operational/installation details distract from the paper's focus. Keep only what is needed for readers to understand, reproduce, and trust the results (site, period, completeness, QC, uncertainties, key post processing). Technical but relevant info can go in the Appendix, but non-essential logistics/engineering details should be moved to a separate deployment report.

Minor revisions

L89: Several times in the text, the word "trend(s)" could be replaced with "result(s)", "pattern(s)" or "behavior".

L92: Define MYNN and YSU on first use.

L93: Remove "using MYNN".

L97: "Nunalee and Basu (2014) noted …" — consider removing if it does not add to the argument in that paragraph.

L101: Is there a reference for the observational campaign (e.g., Krishnamurthy et al., 2023) and/or others; point to Appendix A for site-specific details.

L103: Replace "hitherto have not been explored in an observational manner" with "had not previously been observed."

L114: "…ERA5 will exhibit a low wind speed bias …" this entire sentence is ambiguous.

Section 2: Consider merging the opening of Section 2 with the the first paragraph of Section 2.1 to present a concise summary of the observational data (source, location, period, sampling/averaging, heights, instrument/method). The rest could be moved to a new subsection (e.g., Local Wind Characterization). You could briefly summarize the key points and refer to a single comprehensive figure that combines the already shown plots 2a and 2b, and adds (c) monthly and seasonal means and (d) the diurnal cycle, rather than exhaustively listing numbers. Any additional technical details should go in the Appendix.

Section 2.2: Clarify whether the ERA5 time series was horizontally interpolated or taken from the nearest sea or land/sea grid cell. The proximity to the coastline could influence results.

L179: Provide a reference for UH-WRF if available.

L205: Clarify that GWA3 wind data are provided as annual, monthly, and diurnal climatologies, not continuous hourly time series.

Table 1: Provide actual vertical levels used rather than “lowest N model heights.” For ERA5, “PBL scheme plus the effect of data assimilation”?

L238–239: It appears the manuscript uses “average/median bias” when referring to mean/median errors. It sounds a bit redundant as bias is already the mean error. Unless you compute multiple bias values (e.g., per height or per wind speed range) and then take their mean/median. Remove the word “average” in both lines if redundant and check the entire manuscript.

L241–243: “Figure 3a”.

Figure 3: There is no plot 3d.

Fig. 5: State normalized by what and its relevance in the text.

Section 3.2: Add one sentence explaining why testing different atmospheric stability conditions is relevant. Also, state explicitly that you use air-sea temperature differential as a proxy for stability and note the limitations.

L358: Include results/discussion on occurrences of LLJs, or remove earlier references if not analyzed.

L368: Reference figure 11a.

L383: Specify inconsistencies relative to what.

L433: Use “corroborate previous reports of” instead of “add another geographic data point to the trend of”.

L437–438: The winds may be consistent, but was 2023 a regular year? Maybe it would be helpful to include this information. Also, what is the distance from the buoy to the nearest coastline? If the ERA5 grid cell includes or lies close to the coast, this proximity could influence the results.

L453–455: Justify attributing UH-WRF's smaller bias primarily to the PBL scheme (YSU) rather than to other differences.

L487: There is no Shaw et al., 2020 in the references. Do you mean Gorton and Shaw, 2020?

Appendix A1: Keep only information that adds to the manuscript and is not already documented elsewhere: location, measurement method, campaign duration, data completeness, uncertainties, QC, and post-processing.

Appendix A3.3 and A3.4: Consider removing if not essential to understanding the manuscript.