Review of “Performance of wind assessment datasets in United States coastal areas”

https://doi.org/10.5194/wes-2024-115. Reviewer: R. Gandoin, C2Wind, Denmark.

Main text of the review

Thank you for a well written manuscript, the paper reads very well and provides useful comparisons and discussions. Still, I think that some additional needs to be provided regarding three main topics:

- Measurement datasets: please consider adding additional information on the datasets, the reader needs to be able to contextualise the datasets, and if needed go back to the original time series for doing their own analysis.
- The definitions of the 10m winds in ERA5 and NOW-23, see points 2) and 3) below.
- When classifying sites using the land/sea ratio around the site, please consider the frequency of occurrence of onshore or offshore flow cases. As discussed in my comment in the pdf (page 7), a coastal site may experience onshore winds most of the time, making it more alike a onshore site wrt to model bias.

Thank you very much for your review and constructive suggestions for our work! We are grateful for your time and assistance. We have updated the manuscript based on your valuable feedback regarding the three main topics that required attention as outlined below.

1) Could you please add additional information about the measurement datasets?

This should be, at the minimum, a table with a synthetic description of the measurement datasets with:

- coordinates
- explicit and univocal ID (so the user can easily find out which measurement dataset it is)
- data source with reference including link to the database

Ideally, this table would also include:

- a short description of the terrain (orography, roughness, obstacles nearby) and the measurement setup,
- a wind rose with a schematic coastline drawing/line
- an indication of whether the measurement location is in land or a sea model cell (or the value of the landmask, for ERA5)

Please consider as well providing:

- an energy-based metric about the type of most energetic flow case (onshore or offshore, see the main text of the review).
- the roughness length from the different modelling systems, at the measurement locations

Thank you for this great suggestion! We have added Appendix A and Table A.1 to describe the coordinates, measurement heights, site characteristics, station IDs, original data sources, and wind roses for each observational site. Additionally, we incorporated your recommendation of defining an energy-based flow metric for each site. For this we characterized the wind rose sectors with a definition of predominantly land cover or water cover using the Global Land Cover and Land Use Change 2000-2020

(Potapov et al., 2022). We applied a radius of 100 km from each observation location for this analysis to capture both onshore and offshore breezes based on the works of Gille et al. (2005) and Viner et al. (2021) and then weighted the amounts of land and water coverage by the distribution of wind across the directional sectors. This information was also added to Table A.1. The only observation-related recommendations we did not include were the land vs sea model cell information and the model roughness lengths as these were not universally available for all the datasets considered in this work.

New text was added to the main body of the article (Lines 158-162) to discuss the energy-based metric as follows:

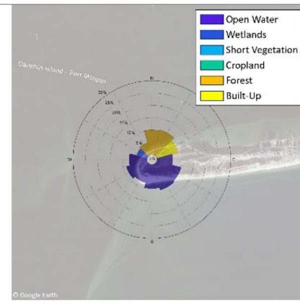
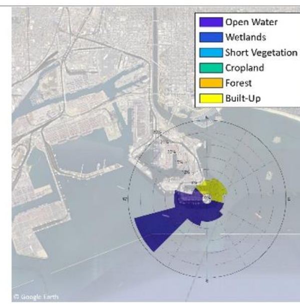
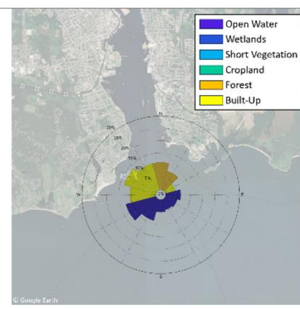
“When considering the distribution of flow direction within a 100 km radius to represent the extent of onshore and offshore breezes (Gille et al., 2005; Viner et al., 2021), the winds at 14 sites predominantly originate over land while the winds at 9 sites predominantly originate over water (Figure 2e, Appendix A) as determined by the Global Land Cover and Land Use Change 2000-2020 (Potapov et al., 2022) and the wind roses for each site.”

A more robust discussion of the observations and the energy-based metric is provided in Appendix A (Lines 475-482):

“Siting characteristics for the 23 coastal observations used for wind dataset validation are shared in Table A.1, including coordinates and measurement heights, wind roses, satellite imagery, general discussions of the land cover and nearby infrastructure, and citations to the original data sources. For each observational site, the 12 wind direction sectors used in this study are characterized as predominantly covered by land or water using the Global Land Cover and Land Use Change 2000-2020 (Potapov et al., 2022). For this analysis, an extent of 100 km from the observation location is utilised to capture both onshore and offshore breezes (Gille et al., 2005; Viner et al., 2021). Each 100 km long directional sector, with its designation of predominantly open water, wetland, short vegetation, cropland, trees, or built-up coverage, is then weighted by the frequency of observed winds occurring for that directional sector.”

Below is a capture of a portion of the new Table A.1:

Table A.1. Site characteristics of observations utilised in this study.

Observation	Coordinates/ Measurement Height	Site Characteristics and Land Cover	Station ID (Source)	Terrain/Wind Rose	% of Winds According to Land Cover
AL 1	30.228, -88.024 36 m	Long, low, narrow peninsula; few trees, large fort	FMOA1 (NDBC, 2024)		Open water 61% Wetlands 4% Short vegetation Cropland Forest 26% Built-up 9%
CA 1	33.733, -118.186 31 m	Shipping pier with high density infrastructure to the north and ocean to the south	PRJC1 (NDBC, 2024)		Open water 73% Wetlands Short vegetation Cropland Forest Built-up 27%
CT 1	41.306, -72.077 20 m	Lighthouse site with trees and infrastructure on mainland to the north and Long Island to the southwest; ~1 km to mainland	LDLC3 (NDBC, 2024)		Open water 42% Wetlands Short vegetation Cropland Forest 19% Built-up 39%

Per another of your helpful recommendations, we characterized the dataset errors according to the new energy flow metric instead of the simpler method we used earlier that only considered the land/water ratio and not the predominant wind directions:

Lines 289-294: "Of the 23 coastal sites in this analysis, 14 have wind flow distributions where most of the wind is arriving from land, while 9 have wind flow distributions where most of the wind is arriving from water (Figure 2e). Each region (Figure 1) is represented in both the water-dominant and land-dominant lists of sites according to flow. GWA3, NOW-23, WTK-LED Climate, and ERA5 perform notably better for the sites with water-dominant wind distributions, with median wind speed relative errors of 3.1%, 9.3%, 15.0%, and 8.6% respectively, than for the sites with land-dominant wind distributions, where the median relative errors are 15.8%, 25.5%, 18.3%, and 12.0% (Figure 6b)."

Lines 299-305: "All datasets follow the same trend of increasingly positive wind speed biases for land-dominant sites relative to the water-dominant sites. For GWA3, NOW-23, and WTK-LED Climate, the sites with land-dominant wind distributions experience a greater degree of dataset overestimation, with median wind speed biases of 0.81 m s^{-1}, 1.33 m s^{-1}, and 0.98 m s^{-1}, respectively, while the median wind speed biases for the water-dominant sites are 0.05 m s^{-1}, 0.36 m s^{-1}, and 0.71 m s^{-1} (Figure 1a). For ERA5, the degree of model underestimation is reduced for the land-dominant sites relative to the water-dominant sites, with median wind speed biases of -0.19 m s^{-1} and -0.54 m s^{-1}."

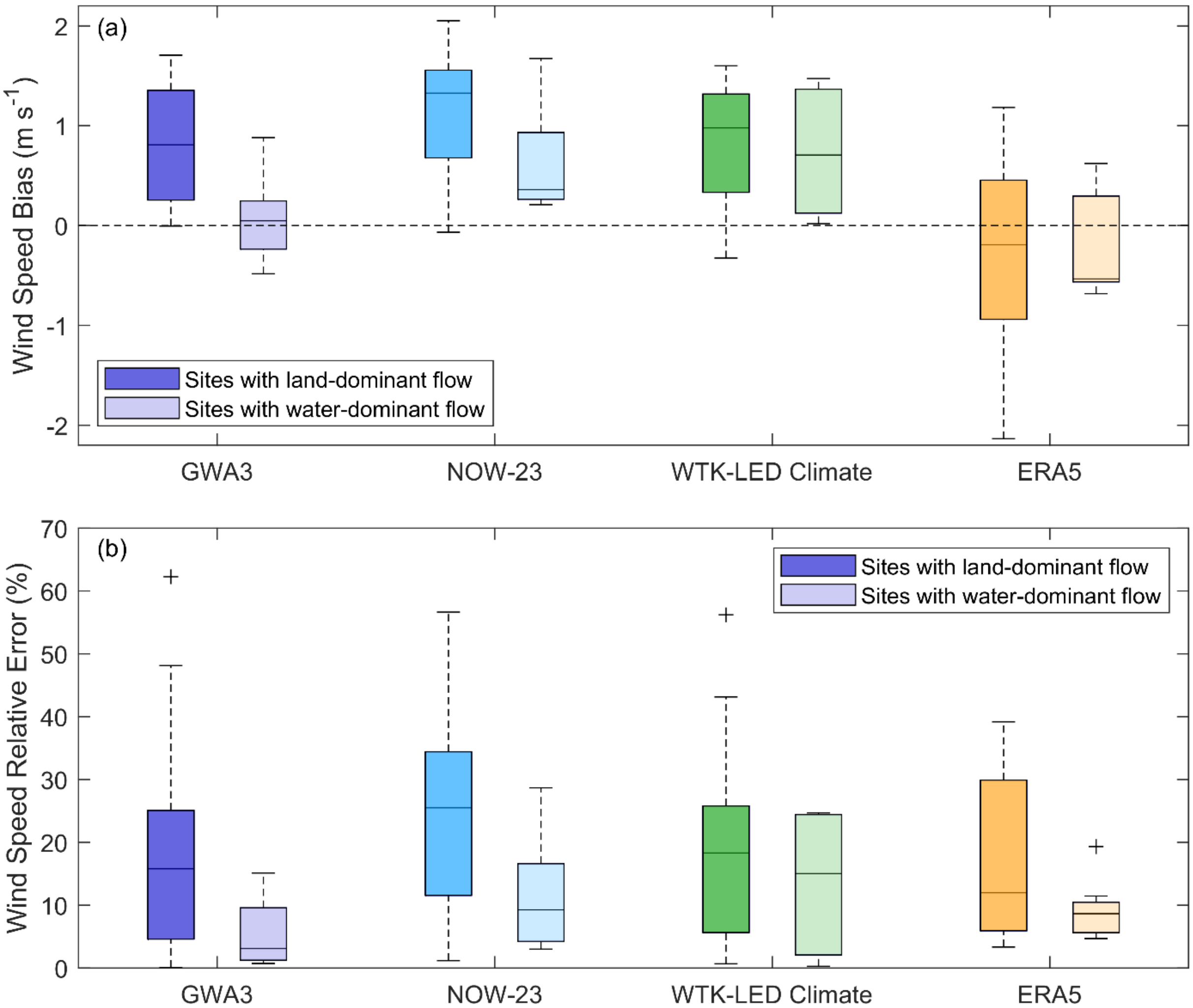


Figure 1. Annual average wind speed (a) biases and (b) relative errors across the 14 coastal sites with land-dominant wind flow and the 9 coastal sites with water-dominant wind flow.

Gille, S., Llewellyn Smith, S., Statom, N.: Global observations of the land breeze, Geophysical Research Letters, 32(5), https://doi.org/10.1029/2004GL022139, 2005.

Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., Zalles, V., Li, X., Khan, A., Stolle, F. and Harris, N.: The global 2000-2020 land cover and land use change dataset derived from the Landsat archive: first results. Front. Remote Sens. 3: 856903, https://doi.org/10.3389/frsen.2022.856903, 2022.

Viner, B., Noble, S., Qian, J-H., Werth, D., Gayes, P., Pietrafesa, L., and Bao, S.: Frequency and Characteristics of Inland Advecting Sea Breezes in the Southeast United States, Atmosphere, 12(8), 950, https://doi.org/10.3390/atmos12080950, 2021.

2) Could you please discuss the definition of the 10m wind in ERA5?

The single levels 10m wind in ERA5 is, for onshore areas, not the model 10m wind. It is a diagnostic "WMO 10m wind" corresponding to a roughness of 3cm, see Section 3.10.2 of https://www.ecmwf.int/en/elibrary/79697-ifs-documentation-cy41r2-part-iv-physicalprocesses.

I think it is important to discuss the influence this may have on the discussion you are providing on model results comparisons.

Thank you for the suggestion to add more information concerning ERA5 to this work, which we had neglected to do by instead focusing on the more recent datasets. First, we have improved Section 2.1 by including a discussion of ERA5 that incorporates your clarification on the 10 m wind along with the reference you kindly provided (Lines 129-133):

"ERA5 is a widely used global reanalysis model (Hersbach et al., 2020) in the wind energy community that began initial production in 2016. The single level ERA5 product outputs wind data at 10 m and 100 m above ground level (Table 1). The winds at the 10 m level are obtained via interpolation between the lowest model level and the surface and are corrected to align with open terrain observations. To adjust to the observations, the correction procedure for the ERA5 10 m winds involves an aerodynamic roughness length that is typical for open terrain with grassland (ECMWF, 2016)."

Second, we have added the roughness length consideration to the discussion of the dataset errors broken out according to land-dominant and water-dominant wind distributions on Lines 294-299:

"The significant decrease in dataset accuracy for land-dominant sites is likely due to a combination of challenges, including dataset representation of complex terrain (particularly for the western sites) and characterization of surface roughness length. Concerning the latter, the land-dominant sites tend to have wind flow distributions that favour cropland, forests, and built environments (Table A.1) which have greater roughness lengths than, for example, the open terrain grassland roughness length utilised for post-processing ERA5's 10 m single level output (ECMWF, 2016)."

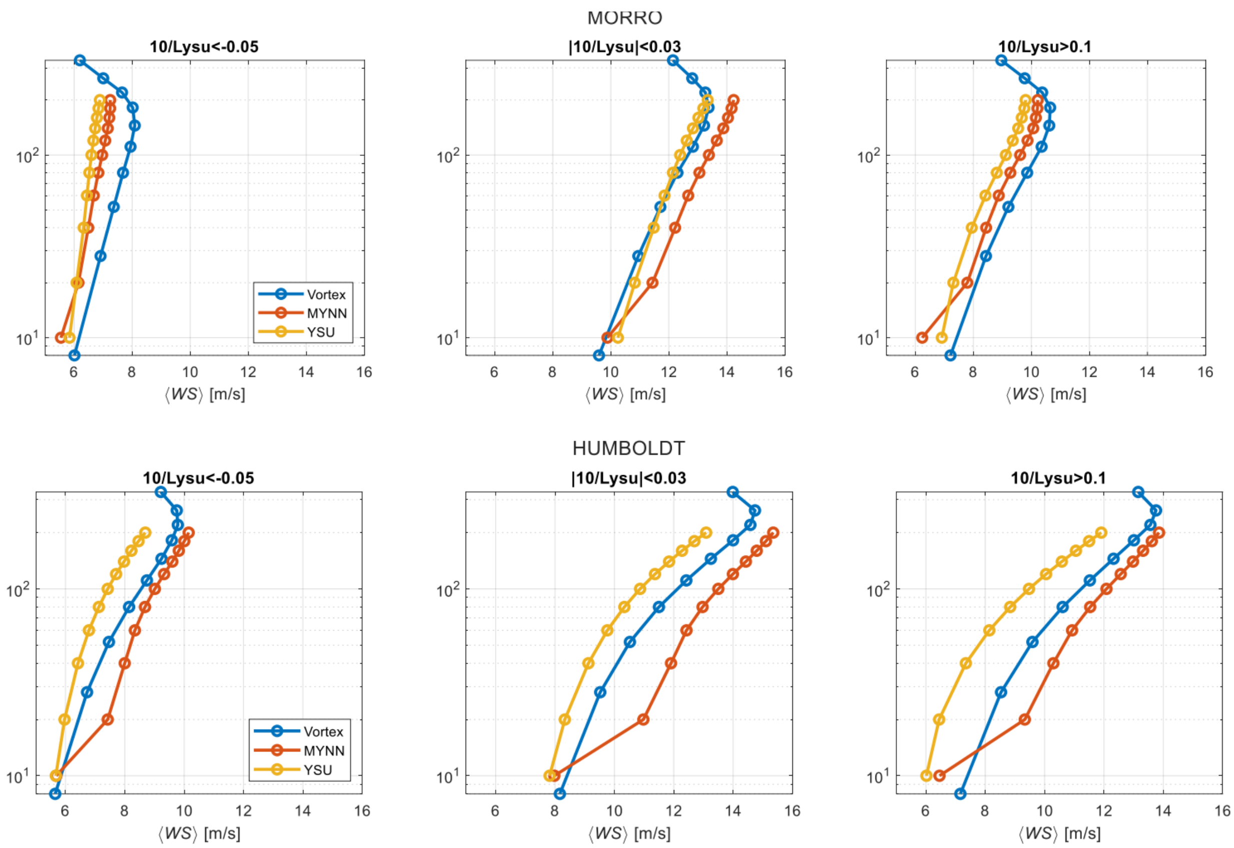
ECMWF: IFS Documentation CY41R2 - Part IV: Physical Processes, https://doi.org/10.21957/tr5rv27xu, 2016.

3) Could you please discuss the question of the 10m wind in the NOW-23 MYNN

According to Section 2.6.1 of "A Description of the MYNN Surface-Layer Scheme" https://repository.library.noaa.gov/view/noaa/30605 the 10m wind in the MYNN is a neutral wind (for WRF version > WRF–ARWv4.0) is some specific flag is set to 1. The NOW-23 has been produced w WRF 4.2.1 (Bodini 2023), so this explains maybe the offset in the 10m wind value compared with the rest of the elevations.

I have myself seen an offset when looking at the NOW-23 dataset at the two floating lidars on the Pacific coast, see below comparisons I did based on data I received from NREL (the Vortex time series is a free 6 months long WRF run as well, from https://interface.vortexfdc.com/). These are offshore locations, so I am unsure what MYNN does there wrt the z0 value (Charnock?), but I see the difference between the 10 m and the otherwise expected value grows with stability.

It is worth double-checking if there anything here with the 10m wind from MYNN datasets that requires attention for this paper



We really appreciate the profiles you shared and the ideas on why we mutually are seeing divergent behavior between 10 m and the rest of the wind profile for NOW-23 domains using MYNN. We are interested in further exploration; however, Reviewer 2 was strongly against our analysis of NOW-23 performance according to PBL scheme for the reason that each site has a unique, and therefore incomparable, PBL scheme. They found the discussion and the former Figure 8 (the NOW-23 relative wind speed errors according to PBL scheme) misleading, and we have accommodated their concern by removing the analysis.

We have, however, expanded the detail in our near surface shear exponent graphic, Figure 4, to depict the NOW-23 domains in hopes of encouraging further research into this interesting profile behavior. Additionally, we noted this discrepancy in our discussion section to encourage it as an area for future investigative research:

Lines 471-474: “Finally, it is hoped that the validations provided in this work identify areas of future research for dataset developers, such as accuracy improvements for locations dominated by land-based flow and understanding of the NOW-23 discrepancies between 10 m and the rest of the wind profile.”

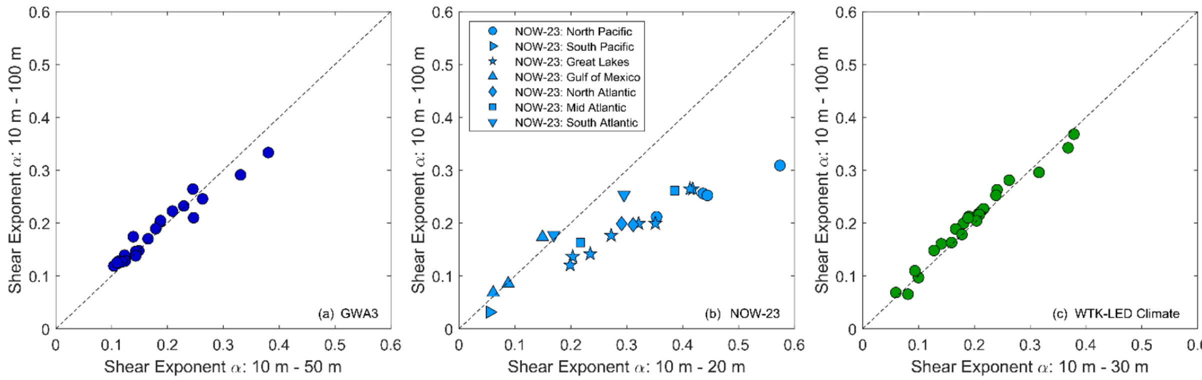


Figure 4. Shear exponents based on the lowest dataset output heights (x-axis) and 10 m and 100 m (y-axis) across 23 coastal sites from (a) GWA3, (b) NOW-23, and (c) WTK-LED Climate.

In addition to the main topics above, we are grateful for the notes you made throughout the body of the text and have addressed your suggestions as follows:

- Added the following to Lines 105-106: “Additionally, GWA3 provides Generalized Wind Climate files that include the wind speed and wind direction distributions for a number of roughness classes that a user can incorporate into WAsP.”
- Updated the start year of ERA5 from 1950 to 1940 in Table 1.
- Added the following footnote to the ERA5 spatial resolution in Table 1: “The ERA5 data have been converted from the native reduced Gaussian grid to a regular latitude-longitude grid at 0.25° (Hersbach et al., 2020).”
- Added “for the single levels product” to the ERA5 output heights in Table 1.