# Review #1

**General comments**

The authors investigated the impact of the MCP correction on the long-term wind resource assessment. They focused on the accuracy of the MCP corrections using the months-long observations to investigate the possibility of reducing the measurement period used for the training data for the MCP. They found that one month of onsite wind speed measurements improves the longterm wind estimate on average, although four months of onsite measurements is a better option to mitigate the errors. It was also reported that the summer months should be avoided, as these months tend to be the least representative of long-term wind speed means and standard deviations. The study is well conducted, and the methods used are appropriate. The data are clearly presented. This study has shown quantitatively that the MCP, using a data period of less than one year, is effective in improving the performance of long-term wind resource assessment based on a large data set.

Thank you for your time and your thoughtful review. We are very grateful for your feedback and support!

These findings will be of interest to wind energy developers working on the distributed wind resource assessment, as well as to researchers in the field. However, I have following concerns for the manuscript:

• Page 7, Line 131–136: Correlation coefficient, bias, and MAE were used as error metrics. In the case of bias and MAE, they would be associated with the magnitude of the values. Accordingly, the use of relative values would be more appropriate when comparing these results for different wind climates.

We appreciate this suggestion and have updated the results section to include relative error per your recommendation.

• Also, the accuracy of ERA 5 would depend on the measurement height as well as the region. I'm not so sure that the combined results can show the true performance of the ERA 5 dataset near the surface. If the accuracy is strongly dependent on the measurement height, it would be better to narrow the range of observations used for the analysis.

In response to your helpful question, we have added the error metrics according to measurement height to Figure 3. We have also added the accompanying text to Lines 155-158: “No consistent trends in ERA5 performance are noted according to height above ground (Figure 3d, e, f). The wind speed relative errors are greatest for measurement heights between 30 m and 40 m (median = 31%), while the median relative errors for measurement heights between 1) 20 m and 30 m and 2) 40 m and 50 m are 11% and 10%, respectively.”

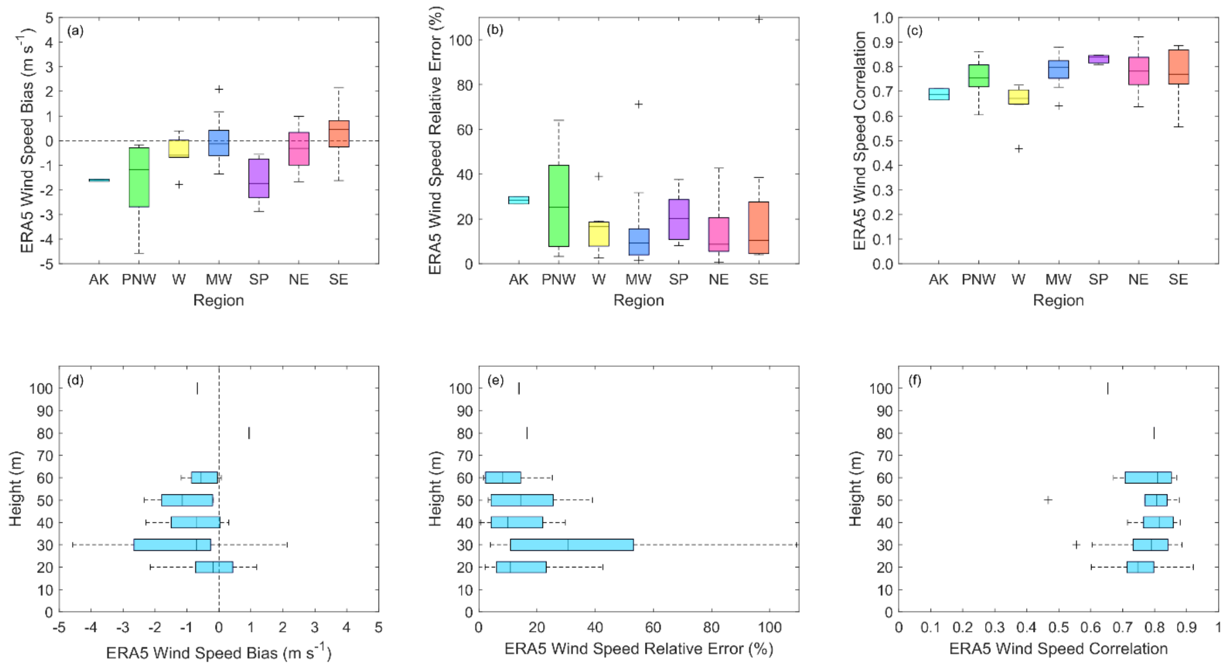


**Figure 1.** Long-term ERA5 wind speed (a), (d) bias (b), (e) relative error, and (c), (f) correlation across 66 measurement sites in the United States, grouped by region (top) and measurement height (bottom). AK = Alaska, PNW = Pacific Northwest, W = West, MW = Midwest, SP = Southern Plains, NE = Northeast, and SE = Southeast.

• Page 7, Lines 158–160: In addition to the three MCP algorithms used in this study, the other algorithms were also available. In fact, a commercial tool, such as WindPro provides methods using matrix and neutral network approaches. The reasons why these algorithms were chosen were briefly explained, but it is still unclear. Are there any reasons why they were chosen, e.g. because they gave better results than the other algorithms?

We agree that the three algorithms chosen are but a small subset of the techniques available for MCP exploration. Given the vast quantity of techniques available, we limited ourselves to three in order to optimize the effort we could spend analyzing their performance. In particular, we selected multiple linear regression and regression trees as they proved successful in our previous distributed wind-focused studies (Phillips et al., 2022) and added adaptive regression splines as a third method for comparison.

Phillips, C., Sheridan, L. M., Conry, P., Fytanidis, D. K., Duplyakin, D., Zisman, S., Duboc, N., Nelson, M., Kotamarthi, R., Linn, R., Broersma, M., Spijkerboer, T., and Tinnesand, H.: Evaluation of obstacle modelling approaches for resource assessment and small wind turbine siting: case study in the northern Netherlands, Wind Energ. Sci., 7, 1153–1169, https://doi.org/10.5194/wes-7-1153-2022, 2022.

• Page 11, Figure 4: The box plots for each month in the figures are based on the different numbers of samples. Is it possible to add the number of samples used for each box plot on the right axis? The authors would analyze a large dataset to derive the results. The information of the sample size would make it easier for readers to understand how much data was used in the analysis.

The box plots for each month are based on the average error metric for each site, so the boxes are composed of 66 values regardless of the number of months. But you are of course correct that different numbers of samples are going into the calculation of each of those averages according to number of training months. To help the readers understand how much data was used, we have added an additional box plot to Figure 5 that shows the number of samples per site incorporated according to the number of training months.

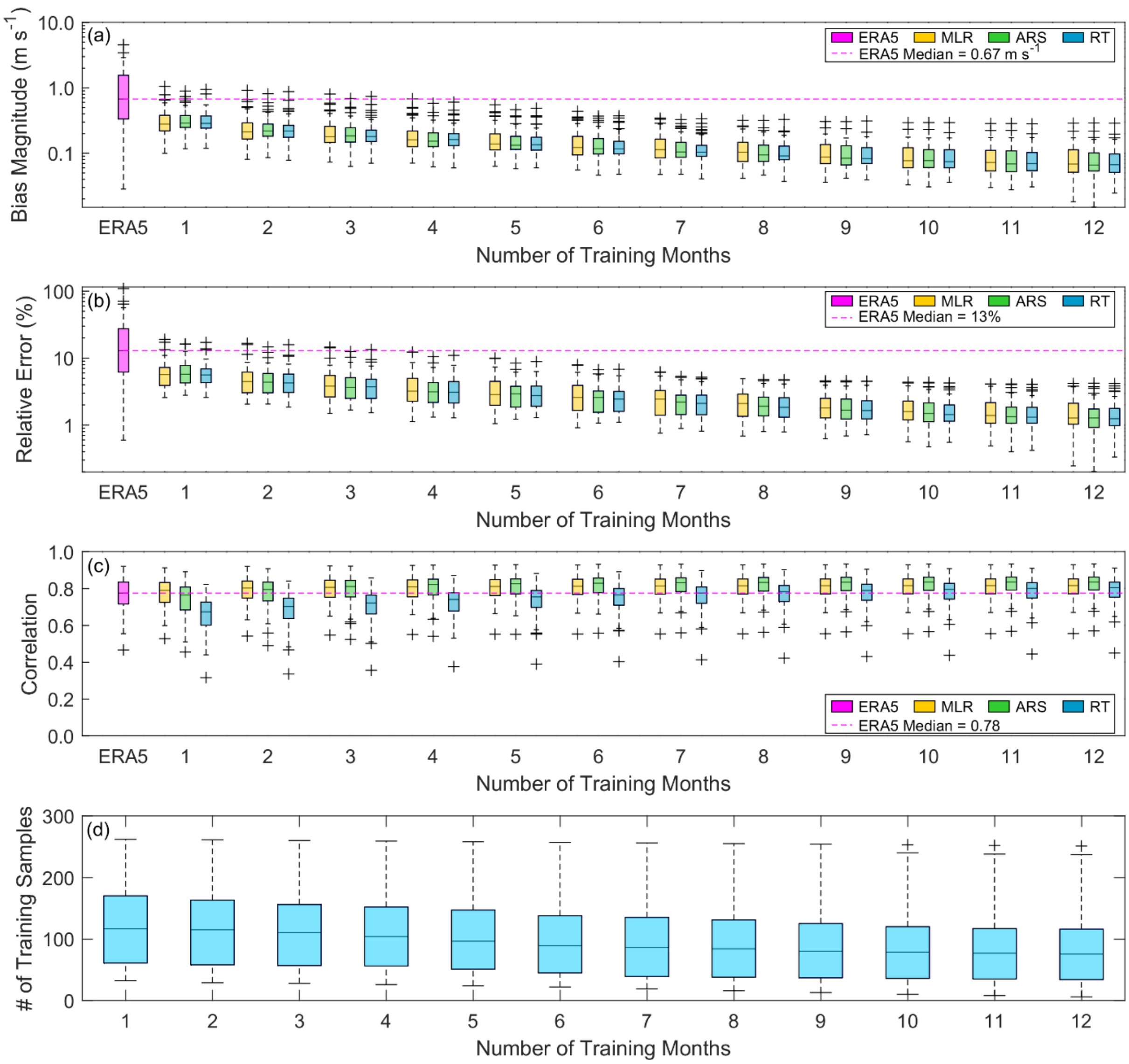


**Figure 2.** Average long-term (a) bias magnitude, (b) MAE, and (c) correlation for 66 sites comparing observations with ERA5 and MCP techniques using varying training period lengths, along with (d) the number of training samples per site and per number of training months.

• Page 22, Figure 12: Figure 12 (b) shows that the capacity factor errors appear to decrease when the training months reach four months. In the discussion, the authors concluded that four months is the preferred length of training months. If the aim of the investigation is to assess the capacity factor, is the MCP based the months-long observation an appropriate approach?

We stress in the discussion that four months is the minimum amount of time, as opposed to the preferred, that observations should be gathered in order to mitigate the errors that could occur if some of the wind speeds in the measurement period are misrepresentative of the longer-term trends (Lines 440-442): "While even one month of onsite wind speed measurements improves long-term wind speed estimates on average, incorporating at least four months of onsite measurements is a better option to mitigate the errors that could occur if some of the wind speeds in the measurement period are misrepresentative of the longer-term trends."

• The authors investigated the impact of observations using the error metrics with bias, MAE, and correlations. As shown in most of the figures in this study, the MCP methods would mainly affect the bias correction. Also, the improvement on the MAE scores would be due to the bias reduction, as discussed in Mattihas and Focken (2006). Is it necessary to use MAE and correlation coefficient for the KPIs as well as bias through the manuscript?

- Lange, Matthias, and Ulrich Focken. Physical approach to short-term wind power prediction. Vol. 208. Berlin: Springer, 2006.

We agree that some of our original choices of error metrics were redundant. We have removed MAE from the evaluation, but have kept correlation as we find it to be relevant for evaluating the performance of simulations in representing fluctuations in the wind, which is of interest when converting to power and assessing the implications of integration into a distribution network.

**Minor comments**

• Page 7, Figure 3: The error metrics were compared across seven regions. However, the number of sites used to drive the statistics would be different. One option would be to show the number of sites for each region in Figure 1.

We have updated Figure 1 with the number of observational sites in each region, per your helpful suggestion. Thank you!

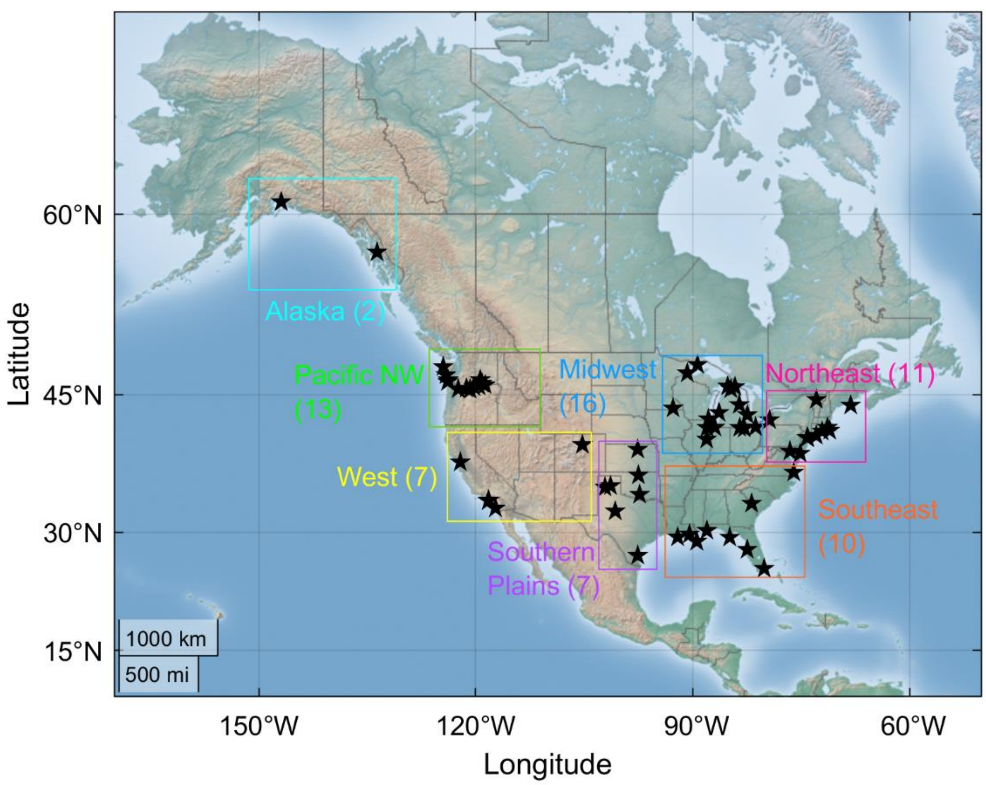


**Figure 3.** Locations of wind measurements assessed for establishing long-term performance based on months-long observations used in this study. The number of observational sites per region is included in parentheses.

• Page 7, Figure 3: The results of this study would be consistent with previous studies on the accuracy of the ERA5 dataset. The authors would be able to link the result to the previous study.

Thank you for this suggestion to link the ERA5 performance from our analysis to previous work. We have added the following on Lines 159-164: "The tendencies of ERA5 to underestimate the observed wind speeds in this analysis while exhibiting a relatively high degree of correlation with them aligns with the findings of Ramon et al. (2019), Murcia et al. (2022), Sheridan et al. (2022), and Wilczak et al. (2024) discussed in Section 2.2. The bias trends according to region (Figure 3a) also align with the findings of Wilczak et al. (2024) in that ERA5 underestimation is noted in the Pacific Northwest and Southern Plains, while a mix of overestimation and underestimated is noted for the Midwest."

• Page 8, Line 216–221: It's difficult to follow the numbers described in the main text. Please consider using more tables to show the numbers.

We agree and have added Table 2 and Table 3 to provide easier reference for the discussion.

**Table 1.** Median biases, bias magnitudes, relative errors, and correlations according to algorithm and number of training months.

| | | Number of Training Months | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Error Metric** | **Algorithm** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| **Bias (m $s^{-1}$)** | MLR | -0.03 | -0.02 | -0.02 | -0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | ARS | -0.04 | -0.02 | -0.02 | -0.01 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| | RT | -0.08 | -0.04 | -0.03 | -0.01 | -0.01 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Bias Magnitude (m $s^{-1}$)** | MLR | 0.28 | 0.21 | 0.18 | 0.16 | 0.14 | 0.12 | 0.11 | 0.10 | 0.09 | 0.08 | 0.07 | 0.07 |
| | ARS | 0.29 | 0.22 | 0.18 | 0.15 | 0.13 | 0.12 | 0.10 | 0.09 | 0.08 | 0.08 | 0.07 | 0.07 |
| | RT | 0.29 | 0.22 | 0.18 | 0.16 | 0.13 | 0.12 | 0.10 | 0.09 | 0.08 | 0.07 | 0.07 | 0.07 |
| **Standard Dev. of Bias Magnitude (m $s^{-1}$)** | MLR | 0.24 | 0.18 | 0.15 | 0.13 | 0.11 | 0.09 | 0.08 | 0.07 | 0.06 | 0.06 | 0.05 | 0.05 |
| | ARS | 0.28 | 0.19 | 0.15 | 0.12 | 0.10 | 0.08 | 0.07 | 0.07 | 0.06 | 0.05 | 0.05 | 0.05 |
| | RT | 0.24 | 0.18 | 0.14 | 0.12 | 0.10 | 0.08 | 0.07 | 0.06 | 0.06 | 0.05 | 0.05 | 0.05 |
| **Relative Error (%)** | MLR | 5.7 | 4.5 | 3.8 | 3.2 | 2.8 | 2.6 | 2.4 | 2.1 | 1.8 | 1.6 | 1.4 | 1.3 |
| | ARS | 5.7 | 4.4 | 3.6 | 3.1 | 2.9 | 2.6 | 2.2 | 1.9 | 1.7 | 1.5 | 1.3 | 1.3 |
| | RT | 5.6 | 4.3 | 3.7 | 3.1 | 2.8 | 2.4 | 2.1 | 1.8 | 1.6 | 1.4 | 1.3 | 1.2 |
| **Correlation** | MLR | 0.79 | 0.80 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 |
| | ARS | 0.77 | 0.80 | 0.81 | 0.82 | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 | 0.84 | 0.84 | 0.84 |
| | RT | 0.67 | 0.70 | 0.72 | 0.74 | 0.76 | 0.77 | 0.78 | 0.78 | 0.79 | 0.80 | 0.80 | 0.80 |

**Table 2.** Median site-average capacity factor relative errors according to algorithm and number of training months.

| | | Number of Training Months | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Error Metric** | **Algorithm** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| **CF Relative Error (%)** | MLR | 9.7 | 8.5 | 7.6 | 6.8 | 6.2 | 5.6 | 4.9 | 4.3 | 4.1 | 3.7 | 3.5 | 3.5 |
| | ARS | 9.9 | 8.0 | 7.3 | 6.7 | 6.2 | 5.6 | 5.0 | 4.6 | 4.4 | 4.2 | 3.9 | 3.8 |
| | RT | 8.6 | 6.8 | 5.9 | 5.0 | 4.3 | 4.0 | 3.7 | 3.4 | 3.2 | 3.0 | 3.0 | 3.1 |

• Page 15, Figure 6: Studying the worst-case scenario is certainly an interesting approach to investigating the risk of the MCP with the months-long observation. On the other hand, the large errors for the data with fewer training months would be due to the outliers. I assume that there is a possibility to improve such errors by applying robust regression algorithms that are insensitive to the outliers.

Thank you for pointing out this consideration. We have added the following text to Lines 303-306:

“It is important to keep in mind that the worst-case scenario error analysis is a conservative approach that is not analogous to assessing algorithm uncertainty. Additionally, more robust algorithms than those studied in this work could reduce the sensitivity to the outliers in the shortest training timeseries that drive error in the long-term estimates.”