Evaluating the potential of short-term instrument deployment to improve distributed wind resource assessment

Lindsay M. Sheridan[1], Dmitry Duplyakin[2], Caleb Phillips[2], Heidi Tinnesand[2], Raj K. Rai[1], Julia E. Flaherty[1], and Larry K. Berg[1]

[1]Pacific Northwest National Laboratory, Richland, Washington, United States
[2]National Renewable Energy Laboratory, Golden, Colorado, United States

Correspondence to: Lindsay M. Sheridan (lindsay.sheridan@pnnl.gov)

Abstract. Distributed wind projects, which are connected at the distribution level of an electricity system or in off-grid applications to serve specific or local energy needs, often rely solely on wind resource models to establish wind speed and energy generation expectations. Historically, anemometer loan programs have provided an affordable avenue for more accurate onsite wind resource assessment, and the lowering cost of lidar systems have shown similar advantages for more recent assessments. While a full twelve months of onsite wind measurement is the standard for correcting model-based long-term wind speed estimates for utility-scale wind farms, the time and capital investment involved in gathering onsite measurements must be reconciled with the energy needs and funding opportunities that drive expedient deployment of distributed wind projects. Much literature exists to quantify the performance of correcting long-term wind speed estimates with one or more years of observational data, but few studies explore the impacts of correcting with months-long observational periods. This study aims to answer the question of how low can you go in terms of the observational time period needed to make impactful improvements to model-based long-term wind speed estimates.

Three algorithms, multivariable linear regression, adaptive regression splines, and regression trees, are evaluated for their skill at correcting long-term wind resource estimates from the European Centre for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5) using months-long periods of observational data from 66 locations across the United States. On average, correction with even one month of observations provides significant improvement over the baseline ERA5 wind speed estimates and produces median bias magnitudes and relative errors within 0.22~~1~~ m s^{-1} and four percentage points of the median bias magnitudes and relative errors achieved using the standard twelve months of data for correction. However, in cases when the shortest observational periods (one to two months) used for correction are misrepresentative of the longer-term trends in the wind resource, the resultant long-term wind speed errors are worse than those produced using ERA5 without correction. Summer months, which are characterized by weaker relative wind speeds and standard deviations for most of the evaluation sites, tend to produce the worst results for long-term correction using months-long observations. The three tested algorithms perform similarly for long-term wind speed bias; however, regression trees perform notably worse than multivariable linear regression and adaptive regression splines in terms of ~~mean absolute error and~~ correlation when using six months or less of observational data for correction.

Formatted: Left

1 Introduction

In the utility-scale wind energy industry, short-term (less than five years) wind measurements are widely used for correcting the long-term (decades-long) wind resource at a site of wind development interest in an expedient manner. Wind farm project planners install onsite wind measurement instruments (e.g., meteorological towers, lidars, sodars) to gather data, but cannot wait decades to establish long-term wind resource characterization based on such measurements alone (Lackner et al., 2008). Instead, wind analysts commonly simulate the long-term wind resource by correlating the short-term onsite wind measurements with long-term reference data, such as atmospheric reanalyses, high-resolution mesoscale models, or other nearby measurements using a measure-correlate-predict (MCP) approach.

Distributed wind projects, particularly those involving small wind turbines, are more subject to challenging financial, spatial, and time constraints than utility-scale wind. Onsite wind measurements are often not feasible or economically viable investments, leading to many distributed wind projects that relied solely on wind resource models to establish generation estimates, which are helpful but not entirely accurate. Between 2000 and 2011, the U.S. Department of Energy (DOE)'s Wind Powering America initiative sponsored an Anemometer Loan Program to provide a more affordable avenue for onsite wind resource assessment, which resulted in the installation of 128 anemometers across the U.S. and a template for around 20 state administered anemometer loan programs (Jimenez, 2013). More recently, researchers from Aurora College have begun deploying mobile lidar systems in northern Canada to study the viability of wind energy in communities that are reliant on diesel, which is expensive and difficult to transport to remote locations (Seto, 2022). In the U.S., communities that receive DOE technical assistance to transform their energy systems, such as through the Energy Transitions Initiative Partnership Project (DOE, 2023), continue to weigh the costs and benefits of onsite wind measurement. In addition to concerns regarding capital investment, communities must often reconcile the time investment involved for gathering onsite measurements with the energy needs and funding opportunities driving expedient deployment of wind turbines.

The vast majority of wind resource assessment literature supports collecting at least one year of onsite measurements to represent a full seasonal wind cycle, including the analyses of Dinler (2013), Liléo et al. (2013), Mifsud et al. (2018), Zakaria et al. (2018), Tang et al. (2019), and Chen et al. (2022). Miguel et al. (2019) found that uncertainty in long-term wind resource estimates reduced by 18%, 29%, 35%, and 40% when one, two, three, and four years of wind measurements, respectively, were added to the monitoring campaign. Additionally, private companies that specialize in providing resource assessment for wind projects also tend to require at least one year of onsite measurements to characterize the wind. For example, ArcVera performs high-resolution modelling (200 m) using WRF with boundary conditions from reanalyses including ERA5. They bias correct their model results using at least one meteorological tower with at least one full year of observational data to achieve mean absolute deviations within 0.2 m s^{-1}, including within areas of complex terrain (ArcVera, 2023). In their discussion of wind resource assessment, wind measurement company NRG Systems (2023) states that measurements of meteorological parameters are typically taken over the course of several years at a potential wind farm site using a combination of meteorological towers and lidars.

Formatted: Font: (Default) +Body (Times New Roman)

A small number of studies, however, explore adjusting long-term predictions using observational data with less than one year of temporal coverage. Taylor et al. (2004 as reported via Carta et al., 2013), Weekes and Tomlin (2014), and Basse et al. (2021) reported that wind speed errors were within 4%, 4.8%, and 4%, respectively, using three months of onsite measurements. The performance of long-term predictions using less than one year of on-site measurements varied according to the season(s) during which the onsite measurements were taken. For their experiment in the United Kingdom, Weekes and Tomlin (2014) found that the smallest errors occurred when using measurements taken in early spring or fall. Basse et al. (2021) performed MCP tests in Germany using linear regression and variance ratio algorithms and found that using variance ratio produced overestimates when using summer measurements and underestimations when using winter measurements, with the opposite trend noted for linear regression.

The present study expands previous analyses of long-term wind resource performance based on months-long observations to diverse locations across the United States, with a focus on measurement heights relevant to distributed wind installations (20 m – 100 m). We explore multiple MCP algorithm options and highlight the best performers for generating accurate long-term wind speed estimates for a variety of error metrics relevant to the wind energy industry. Section 2 describes the wind measurements employed as 1) the months-long training input for MCP-based long-term wind estimates and 2) the long-term datasets to validate the MCP results. The reference dataset, ERA5, is also discussed in Section 2, along with three MCP algorithms selected for evaluation and the error metrics employed to test their performance. Section 3 provides the performance analysis of MCP-based long-term wind resource estimation with less than one year of onsite measurements with a focus on establishing the minimum number of months needed for certain levels of accuracy. Additionally, Section 3 relates the performance results to a variety of influences, including geographical location and~~,~~ the time of year that measurements were gathered~~d, and interannual variability in the wind resource~~. Section 4 summarizes the performance analysis and relates the wind speed error metrics to impacts on wind energy generation expectations.

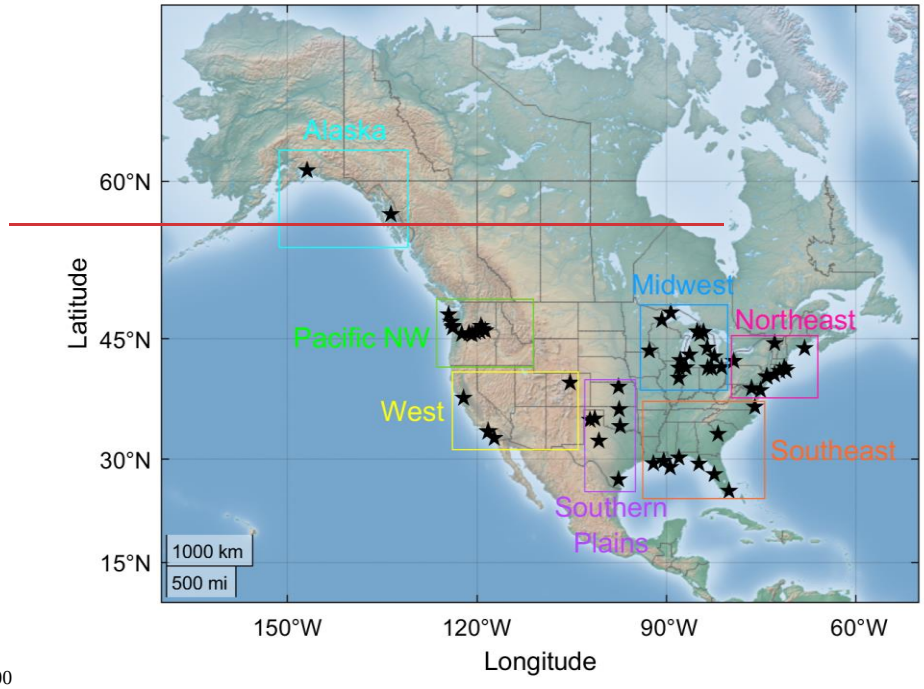
2 Data discussion and methodology

The MCP model for long-term wind resource assessment begins with establishing a short-term relationship between observational wind data measured at a target site and concurrent data at a nearby reference site. The reference data can be observations or model data and typically includes wind variables (speed and direction) and potentially additional relevant meteorological (temperature, pressure) or temporal (hour of day, month of year) variables. The resultant short-term relationship is subsequently applied to long-term data from the reference site to predict the long-term wind resource at the target site (Rogers et al., 2005).

2.1 Wind observations

The 66 wind speed observational datasets in this analysis are sourced from U.S. Department of Energy national laboratories, facilities, and projects, the National Data Buoy Center, the Bonneville Power Administration, and the Forest Ecosystem

Monitoring Cooperative. One observational dataset was collected using a lidar and the remaining measurements were gathered from anemometers on meteorological towers. Most of the observational collection is publicly available (57 sites), while a small number of datasets are subject to non-disclosure agreements (9 sites) as outlined in the data availability statement.



Formatted: Centered

100

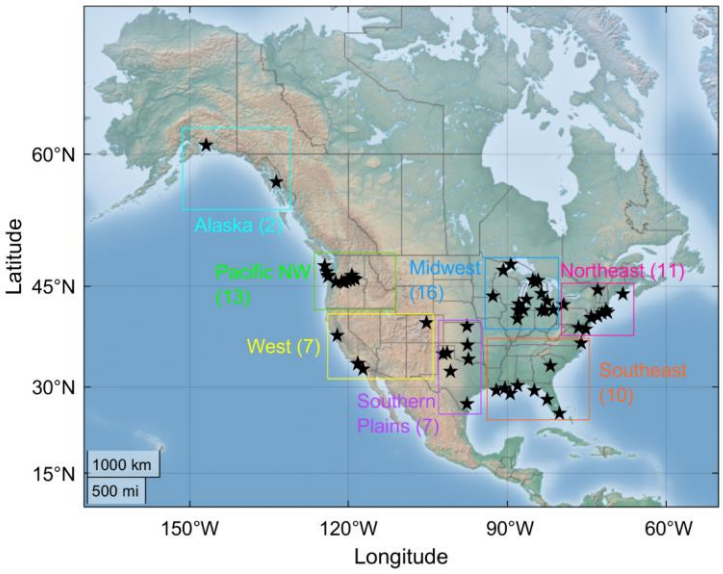


Figure 1. Locations of wind measurements assessed for establishing long-term performance based on months-long observations used in this study. The number of observational sites per region is included in parentheses.

The measurement sites utilised in this work span 28 states and are diverse in terrain complexity and land cover (Figure 1). The measurement heights are similarly diverse, from 20 m to 100 m (Figure 2a), representing a wide range of the distributed wind hub heights that are reported in the United States (PNNL, 2024). Many of the lowest observations, which align with small distributed wind turbine hub heights (between 20 m and 430 m), source from the National Data Buoy Center and are located along coastlines. The highest observations, which align with large distributed wind turbine hub heights (between 80 m and 100 m), are in Long Island, New York (85 m) and the San Francisco Bay Area, California (100 m).

To establish reliable datasets for MCP training and validation, the wind speed observations were quality-controlled by removing instances or periods of atypical or unphysical reported wind speeds (less than 0 m s^{-1}, greater than 50 m s^{-1}, or nonvarying periods of time greater than four hours) that might indicate instrument error due to an outage or weather impacts such as icing. The temporal coverage of the wind speed observations at the measurement sites ranged from 3.5 – 23 years, with an average of 11 years (Figure 2b). The long-term measurement wind speeds ranged from 2.0 m s^{-1} to 9.1 m s^{-1} with an average across all measurement sites of 5.6 m s^{-1}.

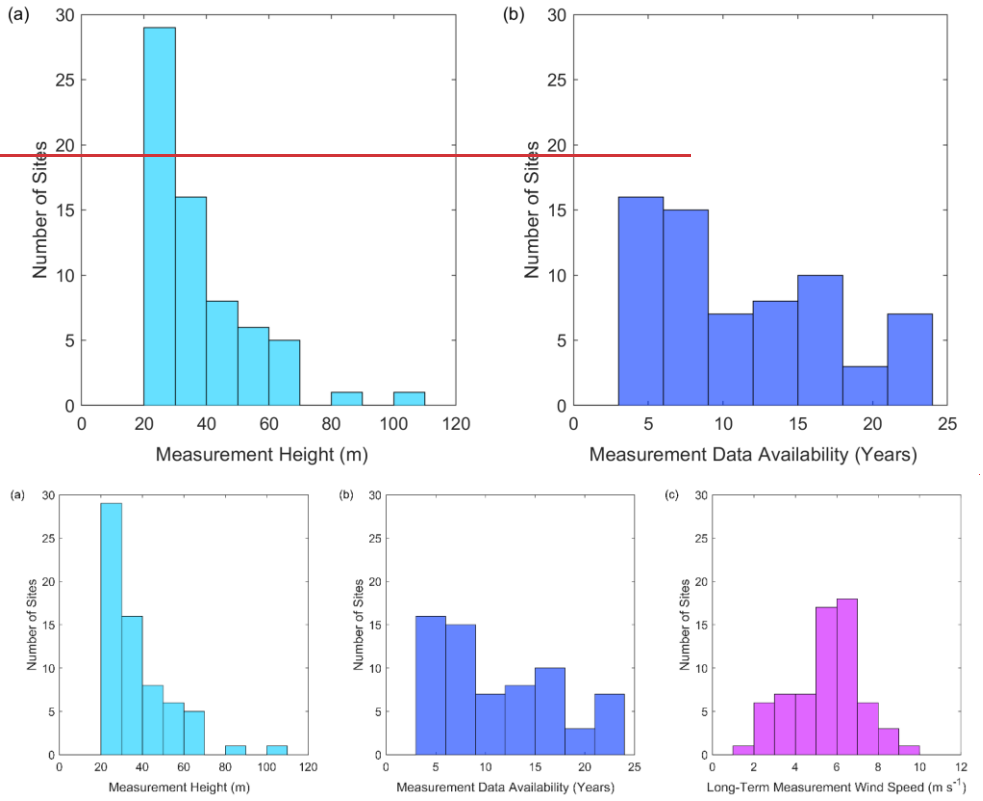


Figure 2. (a) Measurement heights, ~~and~~ (b) long-term measurement availability, and (c) long-term measurement wind speeds for the sites evaluated for long-term performance based on months-long observations.

2.2 Reanalysis model for long-term correction

ERA5 is a popular global reanalysis model (Hersbach et al., 2020) utilised for wind energy resource assessments in a variety of ways, including as a standalone product, as input boundary conditions to higher-resolution model runs, and as a reference dataset for MCP with local observations (Olauson, 2018; Soares et al., 2020; Hayes et al., 2021; de Assis Tavares et al., 2022). ERA5 is developed by the European Centre for Medium-Range Weather Forecasts and provides meteorological data from 1950-present at 1-hour temporal resolution with a horizontal grid spacing of 0.25° (Hersbach et al., 2020). At many locations where validation has been performed, ERA5 tends to produce high Pearson correlation coefficients (Eq. 1) and

negative biases (indicating underestimation) (Eq. 2) between the simulated (u_{sim}) and observed (u_{obs}) wind speeds. Ramon et al. (2019) utilised 77 meteorological towers around the globe with measurement heights ranging from 10 m to 122 m and found median ERA5 seasonal wind speed biases between 0 and -1 m s^{-1}, though ERA5 had the best correlation with observations among five reanalyses. Across 62 sites in the continental US, Sheridan et al. (2022) found that ERA5 underestimated the observed wind speeds by an average of 0.5 m s^{-1} but had higher correlations (average of 0.77) than two alternate reanalyses and wind models. Using measurements from more than 100 onshore and offshore lidars, sodars, and meteorological towers across the United States, Wilczak et al. (2024) determined that ERA5-derived wind power estimates were biased low by 20%. At locations across Europe, Murcia et al. (2022) determined that ERA5 slightly underestimated the observed wind speeds (average bias = -0.06 m s^{-1}) and provided a high degree of correlation (average of 0.92).

2.3 Metrics for performance evaluation

Formatted: Heading 2

This study aims to reduce the ERA5-based wind speed bias using months-long onsite observations while not degrading other metrics of error, such as~~including~~ correlation ~~and mean absolute error (MAE)~~.

~~—~~The Pearson correlation coefficient explains the degree to which the simulated and observed wind speeds are linearly related, with values close to one indicating a high degree of correlation (Eq. 1). The wind speed bias, i.e., the average difference between the simulated ($U\text{~~}u_{sim}$) and observed ($U\text{~~}u_{obs}$) wind speeds over a time series of length N, informs on whether a simulation tends to overestimate (positive bias), underestimate (negative bias), or accurately represent (zero bias) the observed wind resource (Eq. 2). This work also considers the bias magnitude (the absolute value of bias (Eq. 2)) when comparing multiple sites, as combinations of positive and negative biases can obscure the degree of error. The ~~MAE~~ relative error in the long-term wind speed simulation is the absolute difference between the simulated and observed wind speeds normalized by the observed wind speed ~~on a timestamp-by-timestamp basis~~, providing insight on the magnitude of error in a simulation (Eq. 3).

$$correlation = \frac{\sum_{i=1}^{N}\left(U\text{~~}u_{sim,i} - \overline{U\text{~~}u_{sim}}\right)\left(U\text{~~}u_{obs,i} - \overline{U\text{~~}u_{obs}}\right)}{\sqrt{\sum_{i=1}^{N}\left(U\text{~~}u_{sim,i} - \overline{U\text{~~}u_{sim}}\right)^2}\sqrt{\sum_{i=1}^{N}\left(U\text{~~}u_{obs,i} - \overline{U\text{~~}u_{obs}}\right)^2}} \quad (1)$$

$$bias = \frac{1}{N}\sum_{i=1}^{N}\left(U\text{~~}u_{sim,i} - U\text{~~}u_{obs,i}\right) \quad (2)$$

$$relative\ error\text{~~}MAE = 100\% * \frac{\left[\overline{U_{sim}} - \overline{U_{obs}}\right]}{\overline{U_{obs}}}\frac{1}{N}\sum_{i=1}^{N}\left|u_{sim,i} - u_{obs,i}\right| \quad (3)$$

To set a baseline of ERA5 performance for the suite of observations utilised in this work, we adjust the ERA5 wind speeds to each measurement height z (Figure 2a) using the power law (Eq. 4) with the shear exponent α calculated at each

timestamp using the ERA5 wind speeds at 10 m and 100 m (Eq. 5). Horizontally, we adjust the ERA5 wind speeds to the observational location using inverse distance-weighted interpolation.

$$U\sout{u}_{era5\_z} = U\sout{u}_{era5\_100m}\left(\frac{z}{100\ \mathrm{m}}\right)^{\alpha} \quad (4)$$

$$\alpha = \frac{\ln\left(U\sout{u}_{era5\_100m}/U\sout{u}_{era5\_10m}\right)}{\ln(100\ \mathrm{m}/10\ \mathrm{m})} \quad (5)$$

In all regions except for the Southeast, the ERA5 wind speeds tend to be lower than the observed wind speeds (Figure 3a). The underestimation is most pronounced in Alaska (median bias = -1.61 m s^{-1}), the Pacific Northwest (-1.18 m s^{-1}), and the Southern Plains (-1.74 m s^{-1}). The same three regions have the highest relative error ~~MAE~~s (median ~~MAE~~s = 2~~1~~8%~~.39 m s^{-1}~~, ~~1.86 m s^{-1}~~25%, and ~~2.09 m s^{-1}~~20%, respectively) (Figure 3b). The median correlations exceed 0.75 for all regions except Alaska (0.69) and the West (0.67), where the coarse ERA5 is challenged by mountainous and coastal terrain (Figure 3c). 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. Across all sites, the median statistics are: bias = -0.50 m s^{-1}, bias magnitude = 0.67 m s^{-1}, relative error ~~MAE~~ = 13%~~1.66 m s^{-1}~~, and correlation = 0.78. 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.

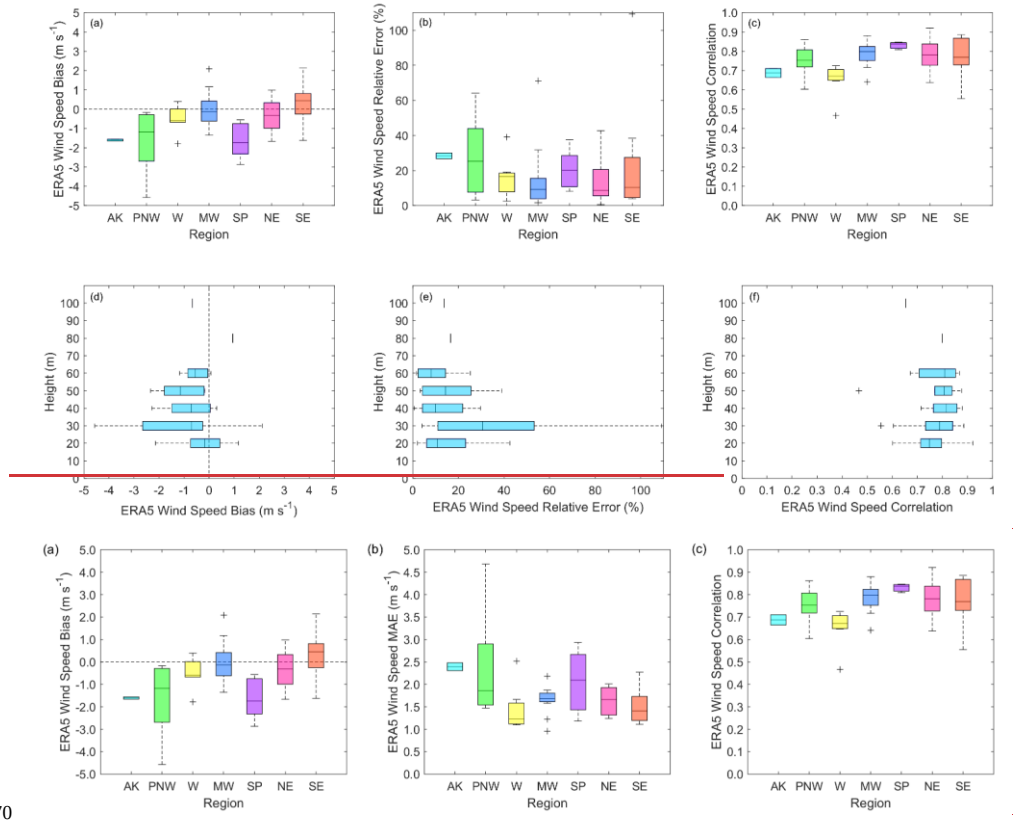


Figure 3. Long-term ERA5 wind speed (a), (d) bias (b), (e) relative error~~MAE~~, and (c), (f) correlation across 66 measurement sites in the United States, grouped by region (top) and measurement height (bottom). AK = Alaska, PNW = Pacific Northwest, W = West, MW = Midwest, SP = Southern Plains, NE = Northeast, and SE = Southeast.

2.~~3~~4 MCP methodologies

One of the advantages of utilising MCP for long-term wind resource estimation is the variety of algorithm choices, which range from simplistic linear regression to machine learning techniques that can be applied to link the short-term and long-term wind speeds. This analysis evaluates three algorithms for their skill at creating long-term wind resource estimates based on varying temporal lengths of observational data: (1) multiple linear regression (MLR), (2) adaptive regression splines (ARS), and (3) regression trees (RT). These algorithms were selected because they are broadly available and represent

diversity in complexity and approach. Multiple linear regression estimates the linear relationship between a target variable (onsite wind measurements in this analysis) and more than one reference variable (sourced from ERA5 in this analysis). Adaptive regression splines involve the construction of piecewise-cubic regression models based on the short-term target and reference datasets (Jekabsons, 2016). In this analysis, we utilise the default parameter configurations of Jekabsons (2016). The maximum number of basis functions follows the formula of Milborrow (2016): min(200, max(20, 2*the number of input variables)) + 1. The maximum degree of interactions between input variables is set to 1 for additive modelling, therefore the generalized cross-validation penalty per knot is set to 2 following the recommendation of Friedman (1991). Regression trees recursively partition and evaluate the concurrent short-term target and reference datasets into unique segments, which are subsequently used to predict long-term target behaviour. In this analysis, the ensemble aggregation method used is least-squares boosting with 100 learning cycles. Numerous additional algorithms have been developed and tested for their ability to improve simulation accuracy, and it is important to note that each feature different approaches, computational investments, complexities, skills, and limitations. For example, Rogers et al. (2005) note that linear regression techniques are easily implemented and well suited for performing bias correction but have a tendency to create a bias in the variance that variance-conserving MCP techniques are better suited to resolve.

To narrow down an effective training approach, different combinations of reference variables are explored for their impact on long-term wind resource assessment error metrics. The analysis of Phillips et al. (2022) identified reanalysis wind speed, reanalysis wind direction, and time of day as the most important variables for wind speed bias correction using a variety of techniques including multivariable linear regressions and regression trees. Therefore, we explore progressively increasing variable combinations of ERA5 wind speeds at the provided output heights of 10 m and 100 m (U~~H~~$_{era5\_10m}$ and U~~H~~$_{era5\_100m}$), power law-based wind speed estimates at the measurement height z (U~~H~~$_{era5\_z}$) (Eq. 4) using the shear exponent α based on the ERA5 wind speeds at 10 m and 100 m (Eq. 5), ERA5 u and v wind components~~wind directions~~ at 10 m and 100 m (u~~∂~~$_{era5\_10m}$, v_{era5_10m}, u_{era5_100m} and v~~∂~~$_{era5\_100m}$), and the hour of the day.

As an initial test of the performance of the algorithms and reference variable combinations, we develop ensembles of MCP-based long-term wind speed estimates at each measurement site using consecutive 12-month training periods when at least 75% of the observations each month (i.e., at least three weeks) are available after applying the quality control checks discussed in Section 2.1. For example, if a site has a 10-year long-term period ranging from January 2008 to December 2017, an ensemble of MCP-based 10-year wind speed estimates will be established based on training periods of January 2008 – December 2008, February 2008 – January 2009, … December 2016 – November 2017, January 2017 – December 2017 as long as each month in the training period meets the data recovery and quality threshold of 75%. All 66 observations utilised in this work have average and median monthly data recovery and quality rates exceeding the 75% threshold. The MCP-based long-term wind speed estimates cover the same temporal period that is available at each site (3.5-23 years) (Figure 2b). For each of the 66 sites, we compute the ensemble average error metrics (bias magnitude~~, MAE,~~ and correlation), with the averages and medians across all the sites presented in Table 1.

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Table 1. Long-term MCP-based wind speed error metrics averaged (median) across 66 sites using different combinations of algorithms and training variables. Values in bold indicate the optimal combination of training variables for each error metric and algorithm.

Error Metric	Algorithm	Training Variables			
		U~~u~~$_{era5\_10m}$, U~~u~~$_{era5\_100m}$	~~u~~U_{era5_10m}, U~~u~~$_{era5\_100m}$ U~~u~~$_{era5\_z}$	U~~u~~$_{era5\_10m}$, U~~u~~$_{era5\_100m}$ U~~u~~$_{era5\_z}$ u~~δ~~$_{era5\_10m}$, v~~δ~~$_{era5\_10}$~~θ~~$_{m}$ u_{era5_100m}, v_{era5_100m}	U~~u~~$_{era5\_10m}$, U~~u~~$_{era5\_100m}$ U~~u~~$_{era5\_z}$ u_{era5_10m}, v_{era5_10m} u_{era5_100m}, v_{era5_100m}~~δ_{era5_10m}, δ_{era5_100m}~~ *hour of day*
Bias Magnitude ($m\ s^{-1}$)	MLR	0.09 (0.08)	0.09 (0.08)	**0.0~~8~~9 (0.07)**	**0.0~~8~~9 (0.07)**
	ARS	0.09 (0.08)	0.09 (0.08)	**0.08 (0.07)**	**0.08 (0.07)**
	RT	0.09 (0.08)	0.09 (0.08)	0.08 (0.07)	**0.08 (0.06)**
Correlation	MLR	0.7~~8~~5 (0.~~8~~075)	0.78 (0.80)	0.~~8~~079 (0.8~~1~~0)	**0.~~8~~179 (0.8~~2~~0)**
	ARS	0.7~~8~~5 (0.~~8~~075)	0.78 (0.80)	0.8~~2~~1 (0.83)	**0.82 (0.8~~4~~3)**
	RT	0.7~~4~~1 (0.7~~6~~2)	0.74 (0.76)	0.78 (0.80)	**0.79 (0.8~~0~~1)**

Formatted Table

Formatted: Font: Not Bold

All MCP combinations of algorithms and training variables provide substantially improved long-term bias magnitudes compared to the ERA5 (i.e., U~~u~~$_{era5\_z}$) average (median) bias magnitude of 1.01 m s^{-1} (0.67 m s^{-1}) (Table 1). With 12 months of training time, the MCP average and median bias magnitudes vary by at most 0.02 m s^{-1} according to algorithm and training variables. More variability is seen according to the various combinations of MCP algorithms and training variables for ~~MAE. All combinations improve on the ERA5 average (median) MAE of 1.82 m s^{-1} (1.65 m s^{-1}), with the incorporation of all training variables producing the lowest MAE values~~correlation (Table 1). Using just U~~u~~$_{era5\_10m}$ and U~~u~~$_{era5\_100m}$ as training variables, MLR and ARS ~~all three MCP algorithms produce worse correlations~~improve on ~~than~~ the ERA5 average (median) correlation of 0.76 (0.77)~~,~~ while RT produces a lower correlation. ~~Adding the ERA5 wind speed adjusted to the measurement height, u_{era5_z}, to the training variables produces the largest relative improvement to the MCP correlations (average and median correlations increase by 0.03 – 0.05) but u~~Utilising all training variables generates the highest overall correlations (Table 1). Given the optimal ~~MAE and~~ correlation results found when training for 12 months with MCP using the complete variable set of ERA5 wind speeds, ERA5 u and v components~~ wind directions~~, and hour of the day, these variables are selected for evaluating long-term MCP wind speed estimates using months-long training periods. Using 12 months of observations, the three algorithms perform similarly for bias and ARS is the best performing algorithm for correlation~~all error metrics. and T~~the following sections explore whether that status holds when using months-long training periods.

The months-long analysis follows the same ensemble formula as the 12-month exercise, just with shorter consecutive training periods. For example, for the site that has a 10-year long-term period ranging from January 2008 to December 2017, the ensemble based on one month of observations would consist of MCP-based 10-year wind speed estimates using training periods of January 2008, February 2008, … November 2017, December 2017, provided each month in the training period meets the data recovery and quality threshold of 75%. The ensemble based on two months of observations would consist of MCP-based 10-year wind speed estimates using training periods of January 2008 – February 2008, February 2008 – March 2008, … October 2017 – November 2017, November 2017 – December 2017 as long as both months in each training period meet the data recovery and quality threshold of 75%. If every month in the 2008 observations for a site had ≥75% data recovery and quality except March, long-term MCP-based wind speed estimates based on January 2008 – February 2008 and April 2008 – May 2008 would be included in the ensemble, but not February 2008 – March 2008 or March 2008 – April 2008. Across the measurement sites, calendar months in the spring and fall had the most single instances of ≥75% data recovery and quality, followed by summer, and lastly winter (Figure 4). Median measurement data recovery and quality percentages according to calendar month ranged from 99.2% (December) to 99.7% (May) (Figure 4).

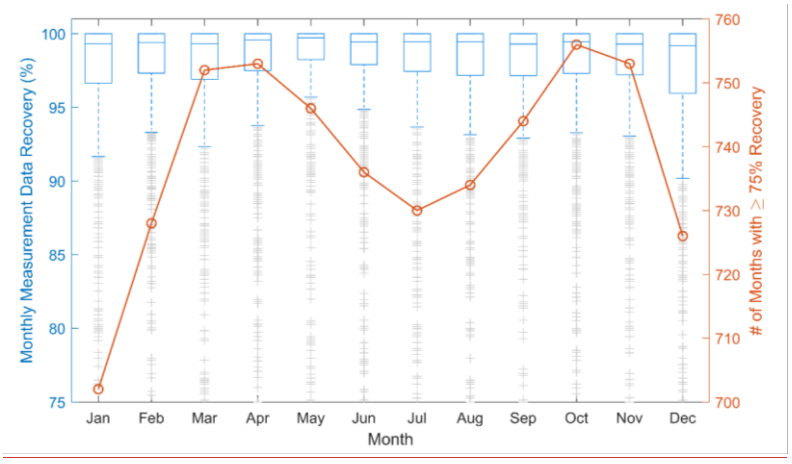


Figure 4. Monthly measurement data recovery and quality (boxplot) and number of months with ≥ 75% data recovery.

Formatted: Centered, Keep with next

Formatted: Font: (Default) +Body (Times New Roman), Not Bold

Formatted: Caption

3 Results and Discussion

3.1 Long-term wind estimation performance according to length of training period

On average, even one month of observations combined with any of the three assessed MCP algorithms provides substantial improvement over using ERA5 wind speeds alone for long-term wind speed estimation. Figure 5~~Figure 4~~a displays the

ensemble average bias magnitude at each site using increasing numbers of training months. The median bias magnitudes across all sites using one month of training are 0.288 m s^{-1} (MLR) and, 0.297 m s^{-1} (ARS, RT), and 0.29 m s^{-1} (RT), as compared with 0.67 m s^{-1} using ERA5 (Table 2, Figure 5a). The median bias magnitudes drop to ranges between 0.187 m s^{-1} (RT) and 0.19 m s^{-1} (MLR) using three months of training and 0.12 m s^{-1} (RT) and 0.124 m s^{-1} (MLR) using six months of training. For all training durations of at least two months, RT is on average the best performing, the MCP algorithms perform within 0.01 m s^{-1} of each other for medianreducing bias magnitude. Similarly, the relative errors in the long-term wind speed estimates decrease substantially from those using ERA5 (median = 13%) (Figure 5b). One month of training produces median relative errors of 6% across the algorithms, which decreases to 4% with three months of training and 2-3% with six months of training (Table 2). The standard deviations of bias magnitude provide an indication of the uncertainty of the MCP-based long-term wind speed estimates, ranging from 0.24 m s^{-1} (MLR, RT) to 0.28 m s^{-1} (ARS) using one month of training, 0.14 m s^{-1} (RT) to 0.15 m s^{-1} (MLR, ARS) using three months of training, and 0.08 m s^{-1} (ARS, RT) to 0.09 m s^{-1} (MLR) using six months of training (Figure 5c, Table 2).

Table 2. Median biases, bias magnitudes, standard deviations of bias magnitudes, relative errors, and correlations according to algorithm and number of training months.

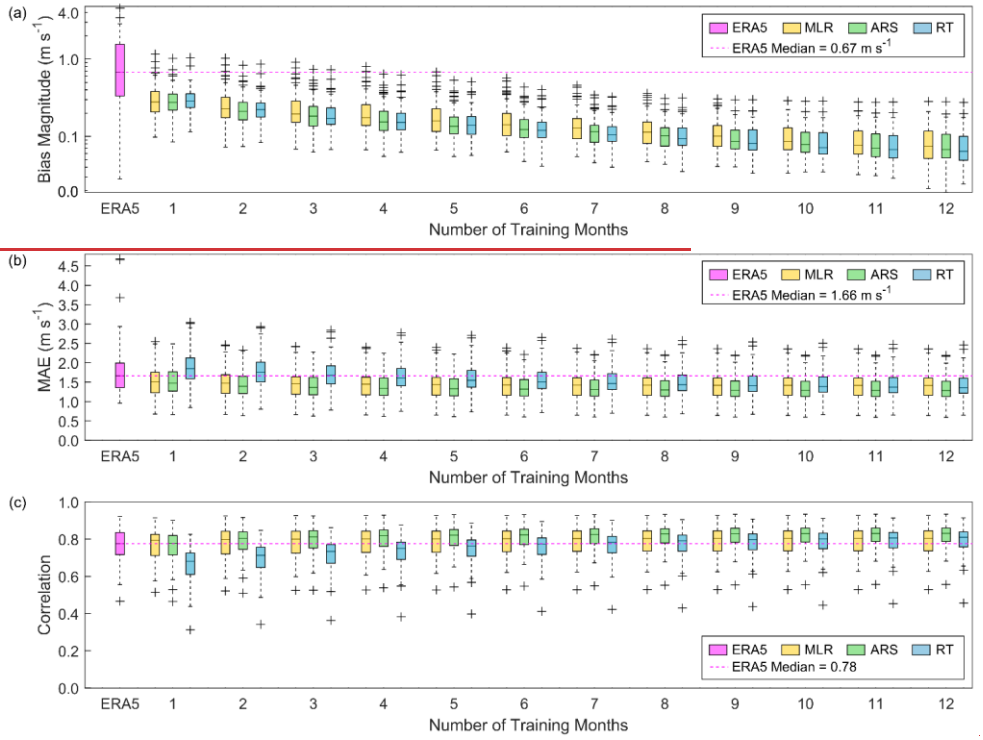
		Number of Training Months											
Error Metric	Algorithm	1	2	3	4	5	6	7	8	9	10	11	12
Bias (m s^{-1})	MLR	-0.03	-0.02	-0.02	-0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01
	ARS	-0.04	-0.02	-0.02	-0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01
	RT	-0.08	-0.04	-0.03	-0.01	-0.01	0.00	0.00	0.01	0.00	0.00	0.00	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

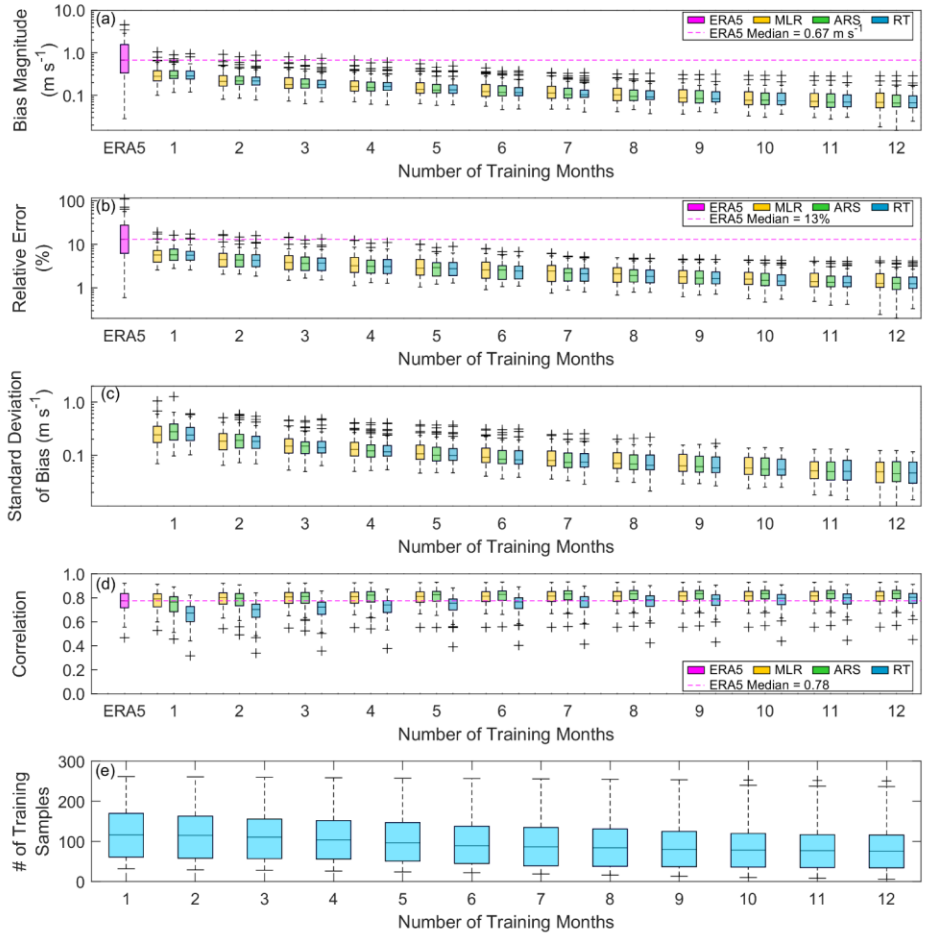
Considering the sign of the bias, we recall that ERA5 tends to underestimate the observed wind speeds with a median bias of -0.50 m s^{-1} across the sites. On average, incorporating months-long observations into the long-term estimations moves the bias substantially closer to zero. Applying just one month of observations results in median biases between -0.086 m s^{-1} (RT) and -0.031 m s^{-1} (MLRARS). The median biases using any MCP algorithm remain slightly negative with increasing training

period lengths until the ARS biases switch to slightly positive after five months of training data and the MLR and RT bias switch to slightly positive after seven months of training data. For all training period lengths of four months or greater, the median MCP-based biases are within ±0.01 m s^{-1}, regardless of algorithm (Table 2).

The MCP algorithms diverge in their performance for long-term wind speed MAE and correlation, especially when using four or fewer training months. On average, using one month of training improves the long-term MAE results when applying MLR (median MAE = 1.51 m s^{-1}) or ARS (median MAE = 1.47 m s^{-1}) relative to using ERA5 wind speeds (median MAE = 1.66 m s^{-1}) (Figure 4b). However, using one month of training with RT produces a median MAE across the sites of 1.85 m s^{-1}, showing worse performance than simply using the ERA5 wind speeds to establish long-term estimates. For MAE, RT begins improving over ERA5 when the training period length reaches four months. For all training durations, ARS is on average the best performing MCP algorithm for reducing MAE.

USimilar to the MAE results, using RT with a limited number of training months does not improve correlation relative to ERA5 and the other MCP algorithms. Using one month of training, MLR provides higher correlations (median = 0.79) than ERA5 (median = 0.78) and ARS produces similar correlations (medians = 0.79 and 0.778, respectively) to ERA5 (median = 0.78), while the RT correlations are quite a bit worse (median = 0.678) (Table 2, Figure 5Figure 4de). Only when the training period is at least seven months do the median RT correlations matchsurpass ERA5. An interesting correlation comparison is noted for the National Renewable Energy Laboratory's National Wind Technology Center, located in an extremely windy corridor of the complex terrain along Colorado's Front Range. All of the lowest correlation outliers in Figure 5Figure 4de are from this site and, while MLR and ARS match or improve upon the ERA5 correlation for the National Wind Technology Center (0.47) using any training period length of at least two months, RT never achieves the ERA5 correlation even with a full twelve months of training. For all training durations of at least four two months, ARS is on average the best performing MCP algorithm for correlation.


(a)
Bias Magnitude (m s-1)
4.0
1.0
0.1
0.0
ERA5
MLR
ARS
RT
ERA5 Median = 0.67 m s-1
ERA5 1 2 3 4 5 6 7 8 9 10 11 12
Number of Training Months
(b)
MAE (m s-1)
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
ERA5
MLR
ARS
RT
ERA5 Median = 1.66 m s-1
ERA5 1 2 3 4 5 6 7 8 9 10 11 12
Number of Training Months
(c)
Correlation
1.0
0.8
0.6
0.4
0.2
0.0
ERA5
MLR
ARS
RT
ERA5 Median = 0.78
ERA5 1 2 3 4 5 6 7 8 9 10 11 12
Number of Training Months



295

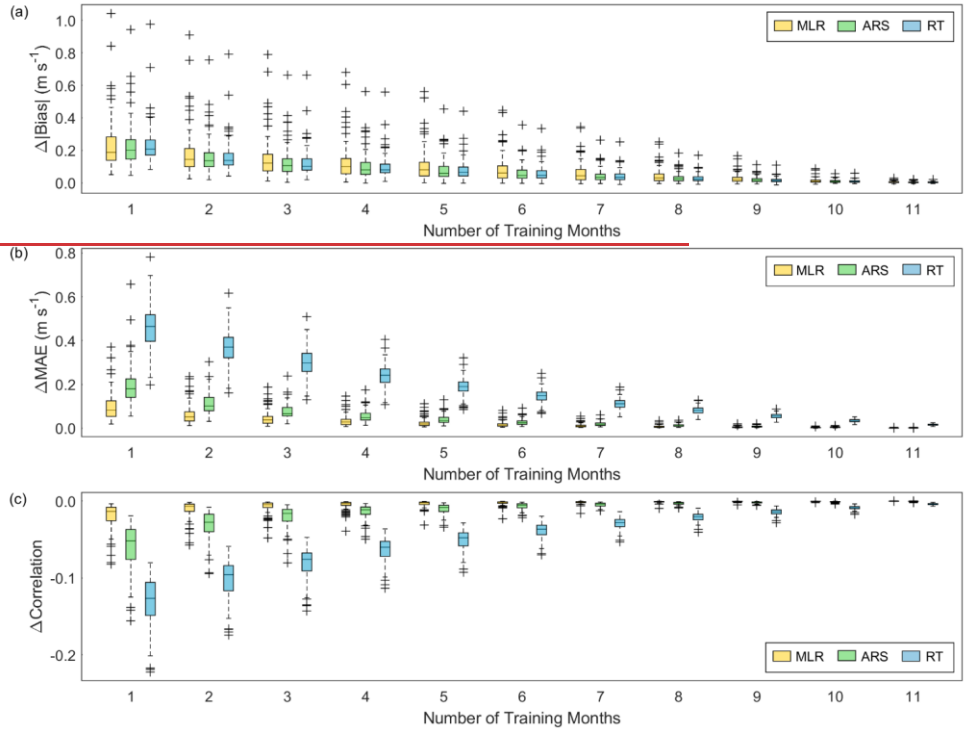
Figure 54. Average long-term (a) bias magnitude, (b) relative error ~~MAE, and~~ (c) standard deviation of bias magnitude, and (d) correlation~~correlation~~ for 66 sites comparing observations with ERA5 and MCP techniques using varying training period lengths, along with (e) the number of training samples per site and per number of training months.~~.~~

Given that a twelve-month training period is most commonly employed for long-term wind speed estimation using MCP,

300 it is beneficial to consider MCP performance using less than one year of observations relative to the twelve-month standard.

Using a one-month training period, the median long-term bias magnitudes (relative errors) for 66 sites are within 0.221 m s^{-1} (4 percentage points) of the median long-term bias magnitudes (relative errors) determined using twelve months regardless of MCP algorithm (Figure 6Figure 5a,b). However, at one coastal Alaskan site the difference in bias magnitude (relative error) when using one versus twelve months of observations can reachexceed 1 m s^{-1} (18%). Training with three months of observations places the median bias magnitudes (relative errors) within 0.12 m s^{-1} (3 percentage points) of the median bias magnitudes (relative errors) established when using twelve months, with outlier differences within 0.78 m s^{-1} (13%). Using six-month training periods results in median bias magnitudes (relative errors) within 0.057 m s^{-1} (1 percentage point) of the twelve-month median bias magnitudes (relative errors), with outlier differences within 0.45 m s^{-1} (6%).

Great disparity is noted among the MCP algorithms when comparing their MAE and correlation performance using months-long training periods versus training periods of one-year duration (Figure 6Figure 5b,c). The median MAEs using one month of observations are 0.08 m s^{-1} (MLR), 0.18 m s^{-1} (ARS), and 0.46 m s^{-1} (RT) different from their respective median MAEs when utilising twelve months of observations. The RT MAEs using less than one year of observations remain noticeably different from the RT MAEs using twelve months of observations, in contrast to MLR and ARS whose months-long MAEs converge on the year-long MAEs more rapidly. For example, using a six-month training period, the median MLR and ARS MAEs are within 0.02 m s^{-1} of their twelve-month median MAEs, while the six-month median RT MAE is 0.15 m s^{-1} different from the median RT twelve-month MAE. For correlation, the differences between the median months-long results and those using a full twelve months are within 0.01 regardless of training duration when using MLR. Using one month of training versus twelve months produces median correlation differences of 0.03 for MLR, 0.075 for ARS, and 0.13 for RT. With six-month training periods, MLR and ARS produces a median correlation less than 0.01 different than using twelve months, while the six-month median RT correlation is 0.04 different from using twelve months.


(a)
MLR
ARS
RT
Δ|Bias| (m s-1)
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
11
Number of Training Months
(b)
MLR
ARS
RT
ΔMAE (m s-1)
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
11
Number of Training Months
(c)
ΔCorrelation
0.0
-0.1
-0.2
MLR
ARS
RT
1
2
3
4
5
6
7
8
9
10
11
Number of Training Months

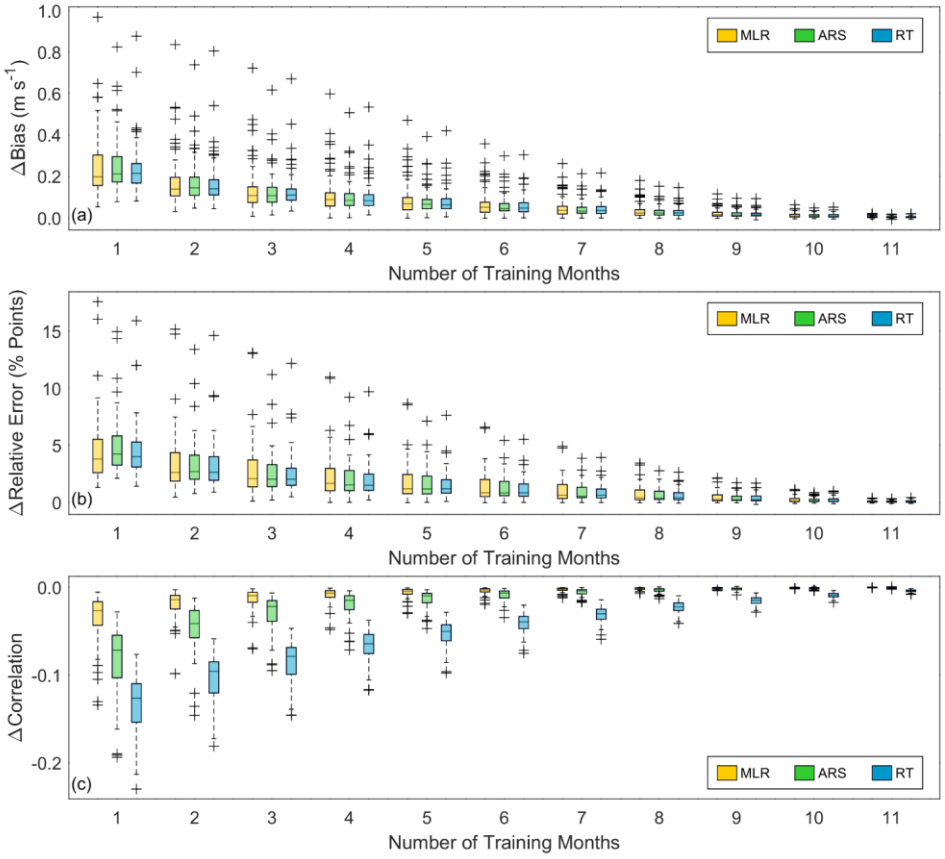


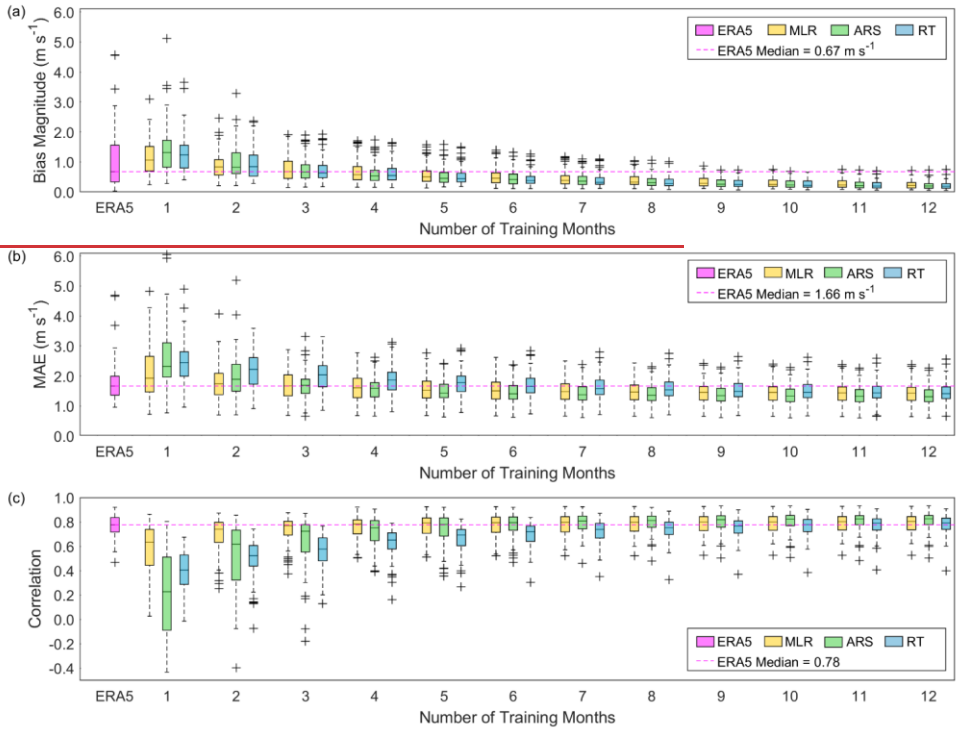
Figure ~~6~~5. Difference in average long-term (a) bias magnitude, (b) relative error~~MAE~~, and (c) correlation for 66 sites between MCP training periods of one to eleven months and MCP training periods of twelve months.

The results in Figure 5~~Figure 4~~ and Figure 6~~Figure 5~~ explain the degree of error expected on average when using MCP with observations less than one year in duration, highlighting similar performance across the MCP algorithms for long-term wind speed bias magnitude, but less successful performance when using RT with months-long observations for ~~MAE and~~ correlation. However, since wind measurement timelines tend to be more dependent on funding availability, project deadlines, and fair-weather deployment windows than on identifying the most representative period for long-term wind

representation, it is imperative to consider the worst-case scenario errors (Figure 7~~Figure 6~~). In this scenario, we identify the worst-performing ensemble member for each error metric (largest bias magnitude, largest MAE, smallest correlation) and each algorithm (MLR, ARS, RT) according to length of training period for each of the 66 sites. 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.

In the worst-case scenario, using a single month for MCP training will produce long-term errors significantly worse than simply using ERA5 to produce long-term wind speed estimates (Figure 7~~Figure 6~~). Despite its performance success on average (Figure 5~~Figure 4~~), ARS produces the largest bias magnitudes (median = 1.3~~1~~ m s^{-1}), the largest relative errors (median = 27%), and the smallest correlations (median = 0.2~~3~~) in the worst-case scenario. MLR and RT perform~~s~~ similarly for bias~~the best~~ using a one-month training period in the worst-case scenario (median bias magnitude~~s~~ = 1.23~~06~~ m s^{-1} and 1.18 m s^{-1}, respectively; median relative errors ~~MAE~~ = 22% for both algorithms~~1.93 m s^{-1}, median correlation = 0.63~~), while MLR performs best in terms of correlation (median = 0.63).

A training period of four months provides improvement in bias magnitude and relative error over simply using ERA5 (median~~s~~ = 0.67 m s^{-1} and 13%, respectively) for long-term wind resource estimation, even in the worst-case scenario, regardless of MCP algorithm choice (medians = 0.5~~3~~4 m s^{-1} – 0.5~~9~~8 m s^{-1} and 13%, respectively) (Figure 7~~Figure 6~~). ~~MLR and ARS show improvement in MAE over ERA5 (median = 1.66 m s^{-1}) at training durations of four months as well (medians = 1.60 m s^{-1}, 1.59 m s^{-1}, respectively), however the RT MAEs do not show improvement over ERA5 until a training duration of six months (median = 1.65 m s^{-1}) is established in the worst-case scenario.~~ For correlation, MLR exceeds ERA5 (median = 0.78) at five months, ARS at six months, and RT at twelve~~eleven~~ months (medians = 0.79). Though all three algorithms ~~RT~~ provide similar~~o~~ improvement in relative errors~~vides the most beneficial bias magnitudes~~ and MLR and ARS provide the most beneficial ~~MAEs and~~ correlations on average when using MCP with months-long observations (Figure 5~~Figure 4~~, Table 2), MLR is the least risky approach given the possibility of measurement during unfavourable wind conditions for MCP.


(a)
Bias Magnitude (m s-1)
ERA5
MLR
ARS
RT
ERA5 Median = 0.67 m s-1
Number of Training Months
(b)
MAE (m s-1)
ERA5 Median = 1.66 m s-1
Number of Training Months
(c)
Correlation
ERA5 Median = 0.78
Number of Training Months

355

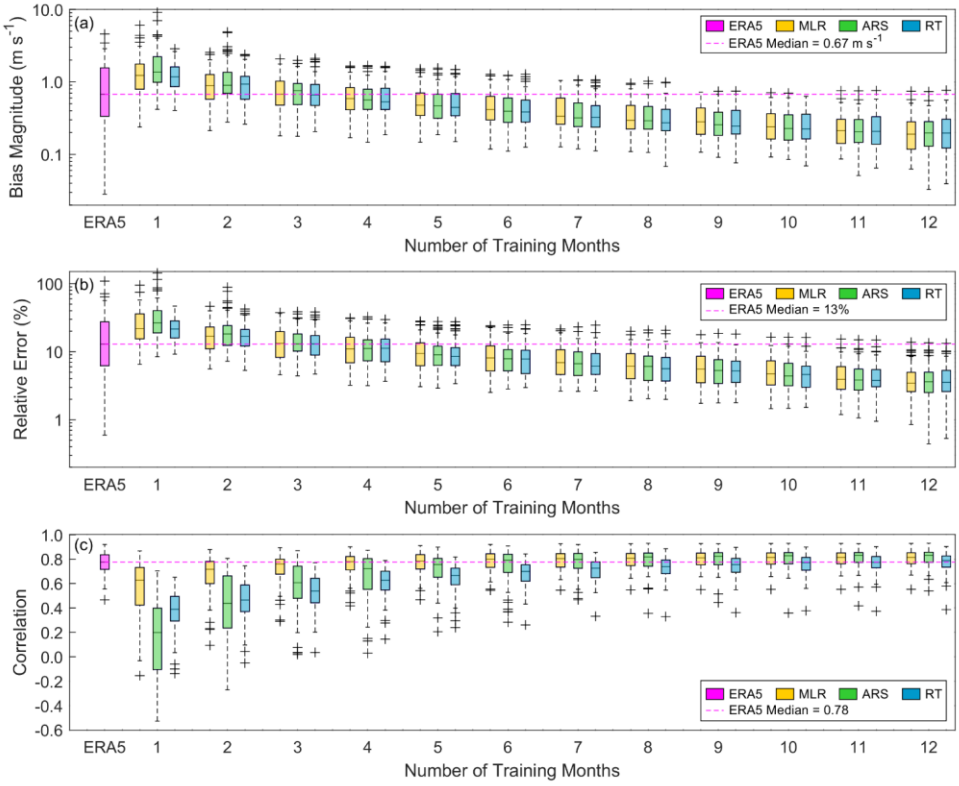


Figure 7~~6~~. Worst case long-term (a) bias magnitude, (b) relative error~~MAE~~, and (c) correlation for 66 sites comparing observations with ERA5 and MCP techniques using varying training period lengths. ~~Outliers not displayed: (a) 9.64 m s^{-1} bias magnitude for one training month with ARS, (b) 10.46 m s^{-1} MAE for one training month with ARS.~~

To help us understand what those unfavourable conditions are that might lead to a worst-case scenario, we investigate the characteristics of the wind speed observations during the worst-case scenario periods relative to the average characteristics of the wind speed observations across all training periods with the same duration (Figure 8~~Figure 7~~). It is interesting to note that the dates of the training time periods with the worst-case scenarios are variable according to MCP algorithm and error metric, particularly for the shortest duration training periods. We find that both the mean and standard deviation of the observed wind speeds during the worst-case scenario month or consecutive months tend to be lower than the mean and standard deviation of the observed wind speeds across all periods with the same durations, particularly when using ARS and

RT (Figure 8~~Figure 7~~). In other words, low wind speed time periods with correspondingly low variations in the wind speeds tend to provide the biggest challenges for long-term MCP accuracy. For example, the median observed wind speed mean (standard deviation) across the 66 sites for all one-month duration periods is 5.80 m s^{-1} (3.12 m s^{-1}), while the one-month worst-case scenario for ARS based on relative error~~bias magnitude~~ corresponds to an observed wind speed mean (standard deviation) of 4.61~~42~~ m s^{-1} (2.30~~25~~ m s^{-1}) (Figure 8~~Figure 7~~a,b). Increasing numbers of training months correspond with higher mean wind speeds and standard deviations during the worst-case periods, along with convergence of the worst-case scenario mean wind speeds and standard deviations to mean wind speeds and standard deviations across all training periods of the same durations.

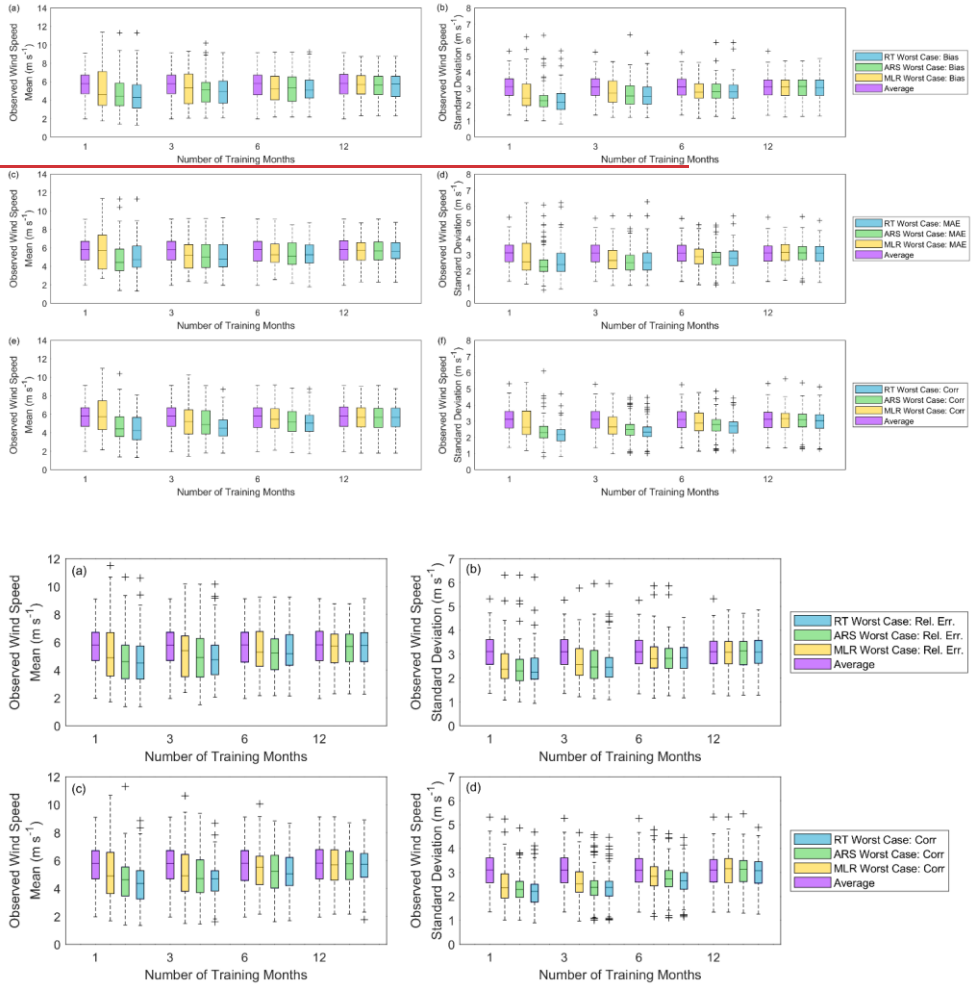


Figure 8~~7~~. (a), (c)~~, (e)~~ Mean wind speed and (b), (d)~~, (f)~~ wind speed standard deviation across 66 sites for MLR, ARS, and RT worst-case scenarios for (a), (b) relative error~~bias magnitude, (c), (d) MAE,~~ and (c~~e~~), (d~~f~~) correlation, and across all training periods for each site of the same durations. Only the results using one, three, six, and twelve months of training are presented for the sake of brevity. Increasing numbers of training months correspond with convergence of the worst-case scenario mean wind speeds and standard deviations to mean wind speeds and standard deviations across all training periods of the same duration.

3.2 Regional long-term wind estimation performance

On average, using one month of observations produces regional median relative errorsbias magnitudes within 8%0.35 m s^{-1} using any of the three MCP algorithms, except for the two Alaskan sitess (Figure 9Figure 8a-g). The Alaskan median relative errorsbias magnitudes are substantially higher (16%0.97 m s-1 using MLR, 14%0.88 m s^{-1} using ARS, and 0.93 m s^{-1}15% using RT), but still significantly lower than the ERA5 median relative errorbias magnitude for these two sites (1.61 m s^{-1}28%) (Figure 3a). The most impressive reduction in bias magnitude occurs for the Southern Plains, where the ERA5 produces a median relative errorbias magnitude of 20%1.74 m s^{-1} (Figure 3a) that decreases to 0.25 m s^{-1} 3% (MLR), 0.26 m s^{-1} (ARS), and 0.28 m s^{-1} (RT)) or 4% (ARS, RT) with the addition of just one month of observations (Figure 9Figure 8e). Using three and six months of observations, all median regional bias magnitudesrelative errors are within 0.26 m s^{-1}5% and 0.17 m s^{-1}4%, respectively, except for Alaska where the median relative errorsbias magnitudes range from 0.64 m s^{-1}11% (ARS, RT) to 0.78 m s^{-1}12% (MLR) using three months and 0.36 m s^{-1}6% (ARS, RT) to 0.50 m s^{-1}7% (MLR) using six months of observations. Using twelve months of observations, all regional bias magnitudes are within 0.15 m s^{-1}3%, including Alaska (Figure 9Figure 8a-g).

The overall trend of RT showing the worst performance on average for MAE (Figure 4b) holds for each regional designation as well (Figure 8h-u). On average, ARS with one month of observations improves on the median ERA5 MAEs in all seven regions, while MLR with one month of observations improves on the median ERA5 MAEs in all regions except Alaska. Only in the Southern Plains does RT with one month of observations improve on the median ERA5 MAE.

On average, using RT with one month of observations degrades the ERA5 median correlations in all regions. Considering all sites, MLR tends to improve on the ERA5 correlations using one month of observations, while ARS performs similarly to ERA5 (Figure 5Figure 4c). However, when exploring correlation on a regional scale, MLR and ARS tend to improve onperforms similarly to ERA5 correlations using one month of observations in the Pacific NW, Northeast, and Southeastall regions except the West, while disimproving ERA5 correlations in Alaska, the West, Midwest, and Southern PlainsARS performs similarly to ERA5 in the Pacific Northwest, Midwest, and Northeast and disimproves relative to ERA5 at all other regions (Figure 9h-n).

Formatted: Centered

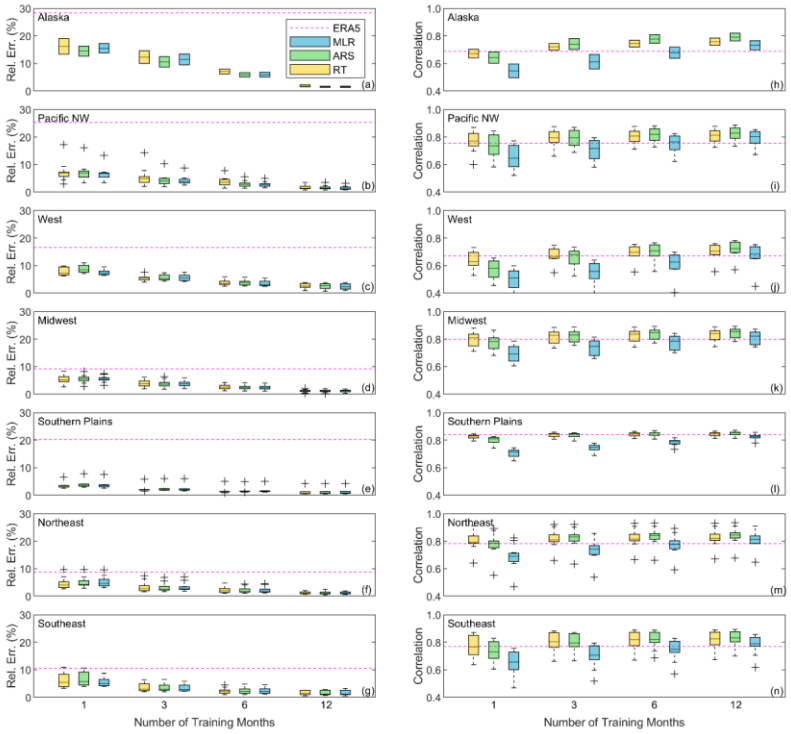
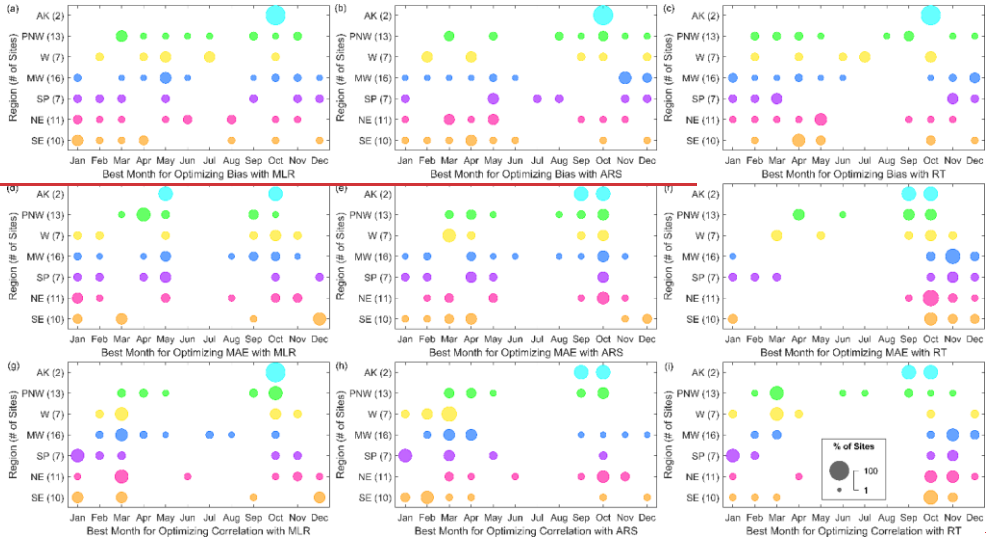


Figure 9~~8~~. Long-term wind speed relative error~~bias magnitude~~ (left)~~, MAE (centre),~~ and correlation (right) according to region for varying training period lengths.

3.3 Seasonal relationships between observations and long-term wind estimation performance

When considering using less than one year of wind speed measurements to correct long-term wind speed estimates, it is important to select the most optimal time of year to gather such measurements (and to understand what times of year to avoid). For each of the 66 sites, Figure 10~~Figure 9~~ and Figure 11~~Figure 10~~ share the best and worst single months, on average, for optimizing error metrics when applying MCP to ERA5. The best months for optimizing relative error~~bias~~ using all three MCP algorithms are quite variable across non-summer (June, July, August) months, with winter (December, January, February), spring (March, April, May), and fall (September, October, November) producing the most optimal relative errors at 21%-23%, 35%-41%, and 30%-33% of the 66 sites, respectively (Figure 10). For correlation, spring months produce the best results when using MLR and ARS for 55% and 48% of the sites, respectively. When using RT, fall months

produce the best correlations at 41% of the sites. ~~while spring (March, April, May) and fall (September, October, November) tend to produce the best MAE and correlation results, as shown by the larger circles (Figure 9). Specifically, using MLR, ARS, and RT, respectively, spring (fall) months produce the lowest MAEs at 24 (22), 26 (27), and 8 (45) of the 66 sites and the highest correlations at 26 (22), 29 (20), and 15 (33) of the 66 sites.~~ Summer months, particularly July and August, tend to produce the worst ~~biases, MAEs,~~relative errors and correlations (Figure 11~~Figure 10~~) and should be avoided if opting to create MCP-based wind speed estimates using a single season of observations. July and August produce the highest relative errors at 52%-56% of the 66 sites, depending on the selected algorithm, and the lowest correlations at 61%-73% of the sites.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
AK (2)
PNW (13)
W (7)
MW (16)
SP (7)
NE (11)
SE (10)
Region (# of Sites)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Best Month for Optimizing Bias with MLR
Best Month for Optimizing Bias with ARS
Best Month for Optimizing Bias with RT
Best Month for Optimizing MAE with MLR
Best Month for Optimizing MAE with ARS
Best Month for Optimizing MAE with RT
Best Month for Optimizing Correlation with MLR
Best Month for Optimizing Correlation with ARS
Best Month for Optimizing Correlation with RT
% of Sites
100
1

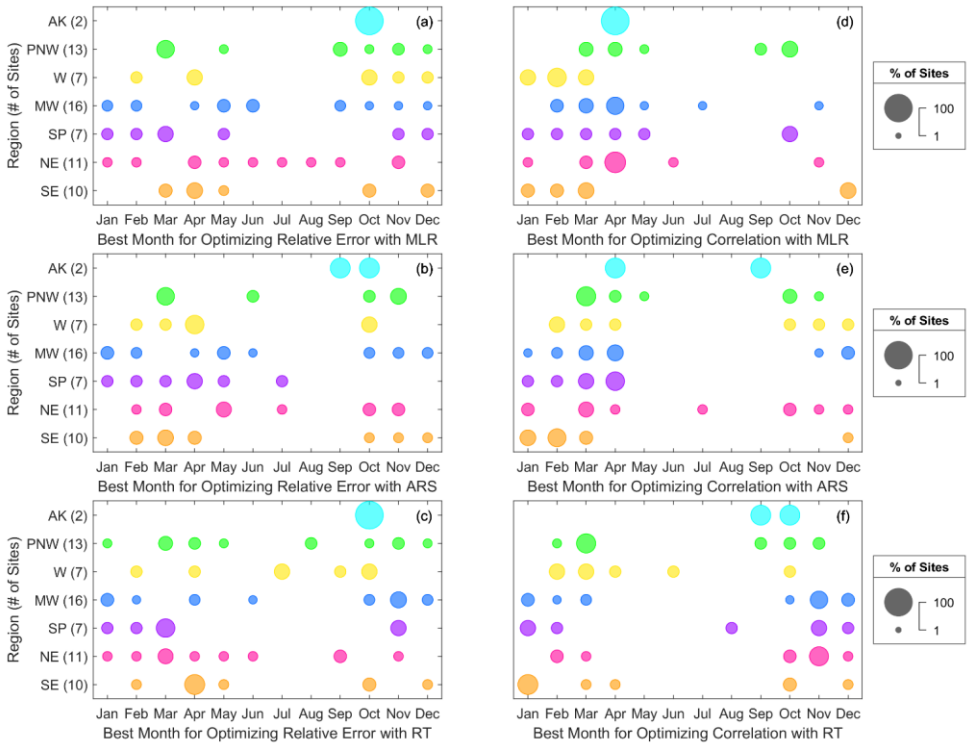
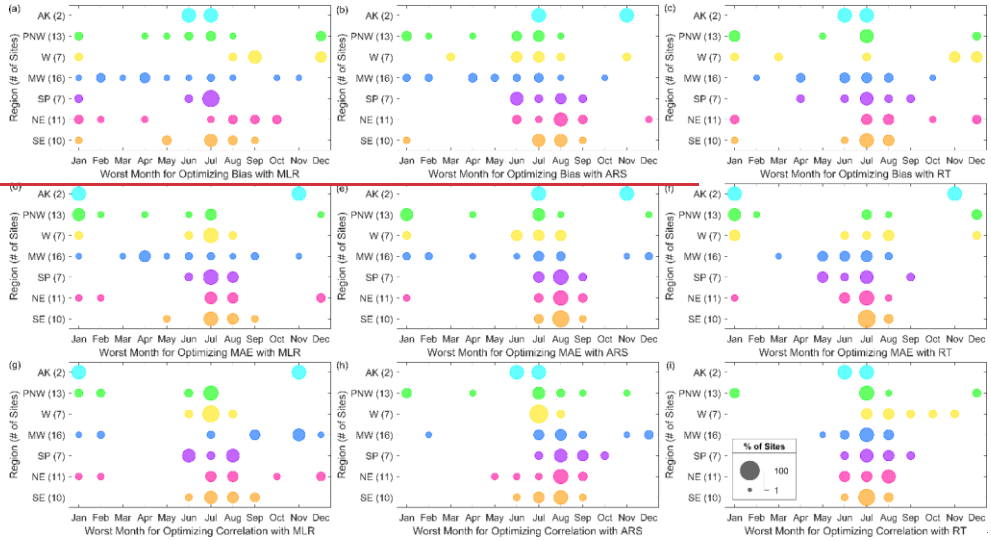


Figure 10. Best average single months for optimizing (a), (b), (c) relative error~~bias~~ and~~,~~ (d), (e), (f)~~ MAE, and (g), (h), (i)~~ correlation when creating MCP-based long-term wind speed estimates using (a), (d)~~, (g)~~ MLR, (b), (e)~~, (h),~~ ARS, and (c), (f)~~, (i)~~ RT according to US region. Larger circles indicate the best month for more sites in a given region.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
AK (2)
PNW (13)
W (7)
MW (16)
SP (7)
NE (11)
SE (10)
Region (# of Sites)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Worst Month for Optimizing Bias with MLR
Worst Month for Optimizing Bias with ARS
Worst Month for Optimizing Bias with RT
Worst Month for Optimizing MAE with MLR
Worst Month for Optimizing MAE with ARS
Worst Month for Optimizing MAE with RT
Worst Month for Optimizing Correlation with MLR
Worst Month for Optimizing Correlation with ARS
Worst Month for Optimizing Correlation with RT
% of Sites
100
1

435

Figure 11~~10~~. Worst average single months for optimizing (a), (b), (c) relative error and (d), (e), (f) correlation when creating MCP-based long-term wind speed estimates using (a), (d) MLR, (b), (e) ARS, and (c), (f) RT according to US region. Larger circles indicate the best month for more sites in a given region.~~average single months for optimizing (a), (b), (c) bias, (d), (e), (f) MAE, and (g), (h), (i) correlation when creating MCP-based long-term wind speed estimates using (a), (d), (g) MLR, (b), (e), (h), ARS, and (c), (f), (i) RT according to US region. Larger circles indicate a bad month for more sites.~~

~~3.4 Long-term wind estimation performance according to length of observations~~

~~Local wind speed trends evolve over decades due to climate evolution and land cover alterations. As this analysis involves a wide range of definitions of long-term period based on observational availability (3.5 – 23 years) (Figure 2b), it is valuable to assess the months-long MCP-based wind speed errors according to the length of the long-term period. Both on average and in the worst-case scenarios using six months of training, bias magnitudes and MAEs tend to increase with increasing length of long-term period (Figure 11a-d) as any six-month period becomes decreasingly representative of the long-term period. On average, median bias magnitudes range from 0.10 m s^{-1} to 0.11 m s^{-1} across the 8 sites with long-term periods of~~

3.5 – 5 years, 0.13 m s^{-1} across the 24 sites with long-term periods of 5 – 10 years, 0.10 m s^{-1} to 0.13 m s^{-1} across the 14 sites with long-term periods of 10–15 years, 0.12 m s^{-1} to 0.15 m s^{-1} across the 12 sites with long-term periods of 15–20 years, and 0.20 m s^{-1} to 0.25 m s^{-1} across the 8 sites with long-term periods exceeding 20 years. The performance in bias magnitude varies more significantly for the worst-case scenarios, with medians of 0.28 m s^{-1} for sites with long-term periods of 3.5 – 5 years and 0.69 m s^{-1} to 0.82 m s^{-1} for sites with long-term periods exceeding 20 years (Figure 11b).

MAEs from MCP wind speed estimated using six months of observations similarly increase with increasing length of long-term period. Median MAEs range on average from 1.16 m s^{-1} to 1.34 m s^{-1} for long-term periods of 3.5 – 5 years up to 1.53 m s^{-1} to 1.73 m s^{-1} for long-term periods exceeding 20 years (Figure 11c). For the worst-case scenarios, median MAEs range from 1.23 m s^{-1} to 1.42 m s^{-1} for long-term periods of 3.5 – 5 years up to 1.78 m s^{-1} to 1.95 m s^{-1} for long-term periods exceeding 20 years (Figure 11d). In contrast to the trends for bias magnitude and MAE, the average and worst-case correlations exhibit no such trend of degradation of performance with increasing long-term period length (Figure 11e,f). Rather, long-term periods of 3.5 – 5 years and exceeding 20 years yield the highest correlations, while long-term periods of 15 – 20 years produce the lowest correlations. Trends in MCP-based correlation are more influenced by region (Figure 8o-u) than the length of long-term period being represented.

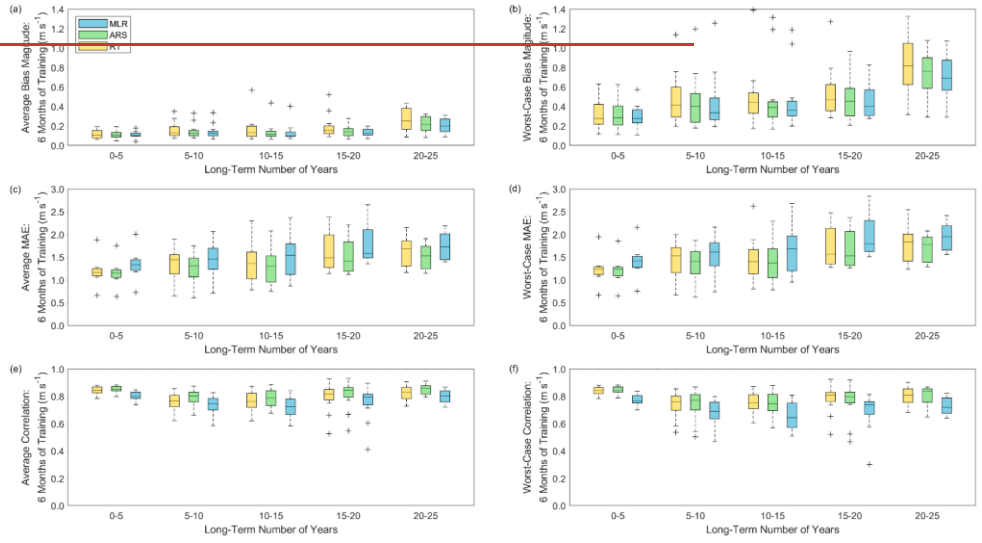


Figure 11. (a), (c), (e) Average and (b), (d), (f) worst-case scenario (a), (b) bias magnitude, (c), (d) MAE, and (e), (f) correlation based on six months of training according to length of long-term period.

3.44 Discussion

Formatted: Heading 2

On average, utilising months-long observations to correct reanalysis-based long-term wind speed estimates provides significant improvement in accuracy of predicted winds (Figure 5~~Figure 4~~). However, substantial risks to accuracy occur when the months-long observational period is misrepresentative of longer-term wind speed trends (Figure 7~~Figure 6~~). To evaluate the impacts of long-term correction using months-long observations on distributed wind energy production, we convert the long-term observed and simulated wind speeds to energy estimates using the National Renewable Energy Laboratory's 100-kW distributed wind reference power curve (NREL, 2019). Since variable lengths of long-term periods are utilised in this study, we opt to consider energy production in terms of the capacity factor, i.e., the energy estimate divided by the product of the total number of hours in the long-term period and the turbine rated capacity (100 kW in this example). The observed wind speeds in this analysis produce gross (i.e., no loss considered) capacity factors ranging 3% (inland Louisiana) to 66% (Texas panhandle), with an average (median) of 34% (36%). The low wind speeds at some of these sites are not suitable for wind energy deployment. Additionally, the power production at the low wind sites will be dominated by the tail end of the wind speed distribution, leading to potential significant differences between the skill of the MCP algorithms in reproducing the highest percentiles of wind speeds versus performing a bias correction, as in Sections 3.1-3.3. Therefore, sites with capacity factors based on observed wind speeds of at least 10% (58 sites in total) are considered for the following analysis.

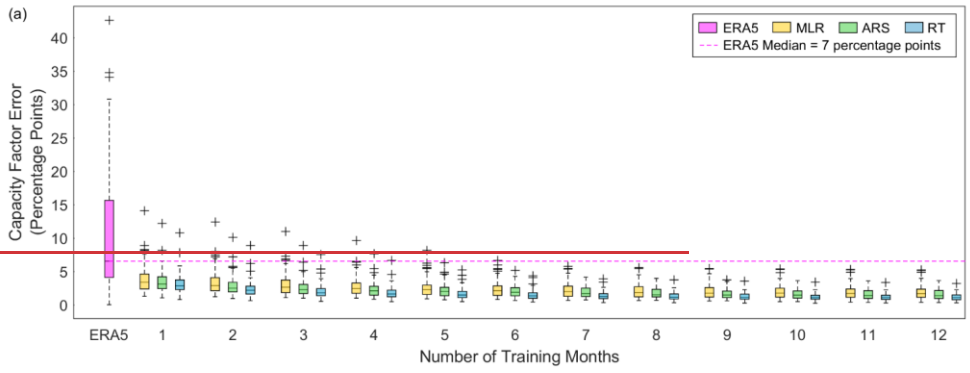
Across the 5866 observational sites, simply using ERA5 to produce long-term energy estimates results in capacity factor relative errors (|capacity factor based on simulations – capacity factor based on observed wind speeds| / capacity factor based on observed wind speeds) up to 894%~~3 percentage points~~, with a median of 24%~~7 percentage points~~. On average, just one month of observations reduces the capacity factor error range to within 51%~~14 percentage points~~ and the median capacity factor relative error to within 10%~~ 3 percentage points~~ (Figure 12~~Figure 12~~a, Table 3). The largest ERA5-based outlier (89%~~43 percentage points~~) is for ~~one of the~~ a complex terrain site in Oregon near the Columbia River with an observation-based capacity factor of 48%~~Bonneville Power Administration meteorological towers in Oregon~~, and MCP correction using one month of onsite observations reduces the capacity factor relative error at this site to 8%-11%~~4-8 percentage points~~ on average, depending on the algorithm. ~~Using three, six, and twelve months of training further reduces the capacity factor errors to within 11 (median ≤ 3), 7 (median ≤ 2), and 5 (median ≤ 2) percentage points, respectively.~~ As for the wind speed relative errors~~bias magnitudes~~ (Figure 5~~Figure 4~~b, Table 2a), RT is the best algorithm on average for reducing capacity factor error when using the 100-kW reference power curve.

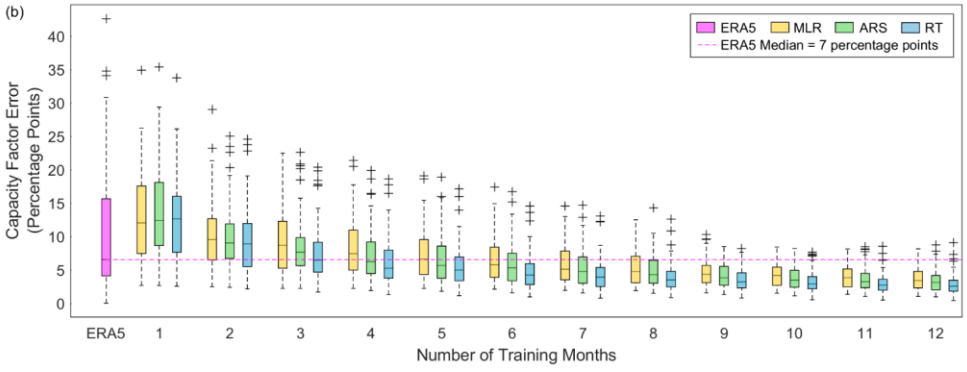
Like the wind speed evaluation, the potential for reduction of wind energy estimate error using months-long onsite observations comes with risk. With a one-month training period in the worst-case scenario, the ERA5 capacity factor error range (within ~~43 percentage point~~89%~~s~~) increases~~reduces~~ to within 94%~~35 percentage points~~, and~~but~~ the median capacity factor errors across the 58 sites increase from 24%~~7~~ to ~~12-13~~40%-47%~~ percentage points~~ (Figure 12~~Figure 12~~b). Using

three, six, and twelve months of training reduces the worst-case capacity factor errors to within 91%~~23~~ (median ≤ 26%~~9~~), 82%~~17~~ (median ≤ 16%~~6~~), and 37%~~9~~ (median ≤ 9%~~3~~) ~~percentage points~~, respectively.

Table 3. 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


(a)
ERA5
MLR
ARS
RT
ERA5 Median = 7 percentage points
Capacity Factor Error (Percentage Points)
0
5
10
15
20
25
30
35
40
ERA5
1
2
3
4
5
6
7
8
9
10
11
12
Number of Training Months


(b)
ERA5
MLR
ARS
RT
ERA5 Median = 7 percentage points
Capacity Factor Error (Percentage Points)
0
5
10
15
20
25
30
35
40
ERA5
1
2
3
4
5
6
7
8
9
10
11
12
Number of Training Months

505

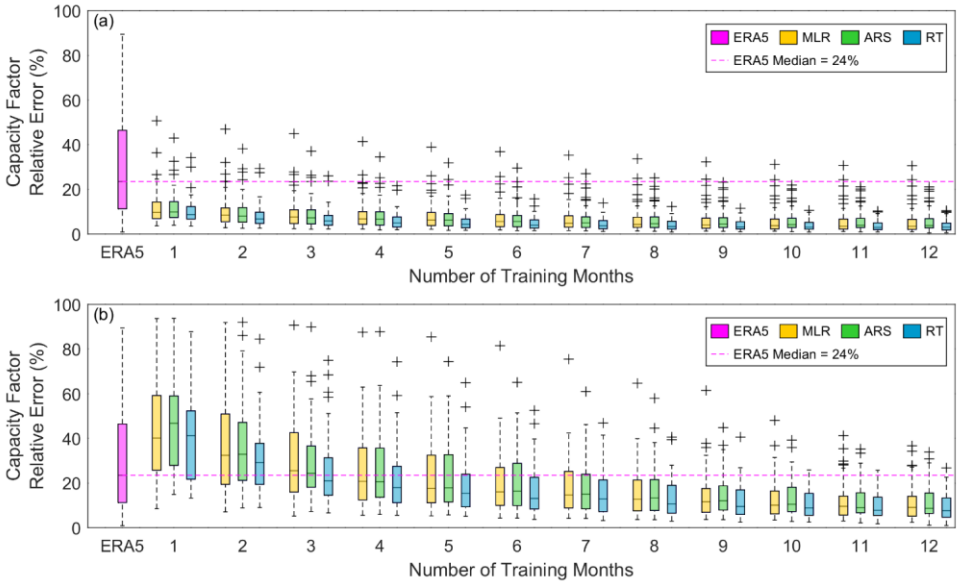


Figure 12~~12~~. (a) Average and (b) worst-case scenario capacity factor relative error (|ERA5/MCP capacity factor – capacity factor based on observed wind speeds|/capacity factor based on observed wind speeds) according to number of training months for 58 sites with observation-simulated capacity factors of at least 10%.

It is important to consider the performance of wind estimation methodologies from a variety of statistical standpoints, since behind-the-meter customers may care most about whether their long-term production meets their initial expectations (bias) and front-of-the-meter customers may be also concerned with distribution network integration with other energy technologies (bias, correlation). While the three MCP algorithms assessed in this study perform similarly in terms of bias magnitude (Figure 5~~Figure 4~~a), for the shortest training period lengths (one to three months) RT performs significantly worse on average for ~~MAE and~~ correlation than MLR and ARS (Figure 5~~Figure 4~~b,c).

Given ERA5's popularity in the wind energy community, along with its known challenges for wind speed bias and relative success in terms of wind speed correlation, this study provides a useful framework for evaluating the performance of MCP-based corrections to long-term wind speed estimates using months-long observations. However, the exercise would benefit in the future from expansion to include additional long-term models that provide wind resource information at different horizontal and vertical spatial resolutions, such as the new WIND Toolkit Long-Term Ensemble Dataset (Draxl et al., 2024). Additionally, the lessons learned in this work are being explored to quantify the geographic extent of their

potential application, with aims to support broader wind speed bias correction in distributed wind tools and to reduce the amount of onsite measurement sites needed to correct sites with similar wind profile characteristics.

Based on this analysis, we identify the following key recommendations for wind energy developers creating long-term wind resource estimates under the constraint of less than one year of onsite measurements which may occur for any number of reasons, including instrument outages or timing of funding opportunities:

- 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.
- Summer months, particularly July and August, should be avoided if opting to create MCP-based wind speed estimates using a single season of measurements in the U.S., as these months tend to be the least representative of long-term wind speed means and standard deviations.
- Since potential wind customers may care about long-term wind resource accuracy from multiple viewpoints, it is important to note that of the three MCP algorithms explored, RT produces the lowest wind speed and capacity factor relative errors~~bias magnitudes~~ and ARS the lowest wind speed ~~MAEs and~~ correlations. However, MLR is the least risky algorithm given the possibility of measurement during unfavourable wind conditions for MCP.

The results of this work highlight the potential for~~benefits of~~ anemometer or lidar loan programs to affordably assist future distributed wind energy customers with more accurate long-term wind resource estimates while maximizing the number of customers that can be served by reducing the measurement time needed.

Code and data availability

Many of the wind speed measurement datasets that support this study are publicly available. Measurements from U.S. DOE-sponsored laboratories, studies, and programs can be found at ANL (2022), ARM (2021), B2SHARE (2020), BNL (2020), LBNL (2020), NREL (2022). Measurements in the Pacific Northwest from the Bonneville Power Administration can be found at BPA (2022). Coastal measurements from the National Data Buoy Center are sourced from NDBC (2024). Measurements from the University of Vermont's Forest Ecosystem Monitoring Cooperative were formerly available at FEMC (2020). The remaining measurements are proprietary, subject to non-disclosure agreements, and have restricted access at DOE (2024).

The ERA5 reanalysis data is obtained from ECMWF (2024). The 100-kW reference distributed wind power curve utilised in this work is provided at NREL (2019). Data processing scripts are written in Matlab and are available from the author upon request.

Author contributions

Data management, software development, analysis, and manuscript preparation were performed by L. Sheridan. Team management was performed by H. Tinnesand and C. Phillips. All authors contributed to the research conceptualization, manuscript edits, and technical review.

Competing interests

The contact author has declared that none of the authors have any competing interests.

Acknowledgements

This work was authored by the Pacific Northwest National Laboratory, operated for the U.S. Department of Energy (DOE) by Battelle (contract no. DE-AC05-76RL01830). This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. DOE under contract no. DE-AC36-08GO28308. Funding was provided by the U.S. DOE Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. The authors would like to thank Patrick Gilman and Bret Barker at the U.S. DOE Wind Energy Technologies Office for funding this research.

References

ANL: Argonne 60 m Meteorological Tower, Argonne National Laboratory [data set], https://www.atmos.anl.gov/ANLMET/, last access: 14 February 2022.

ArcVera: Mesoscale and Microscale Modeling for Wind Energy Applications: Next-Generation Numerical Weather Prediction for Wind Flow Modeling. https://arcvera.com/mesoscale-and-microscale-modeling-for-wind-energy-applications-next-generation-numerical-weather-prediction-for-wind-flow-modeling/, accessed 24 August 2023.

ARM: Atmospheric Radiation Measurement User Facility, Southern Great Plains Site [data set], https://www.arm.gov/capabilities/instruments?location[0]=Southern%20Great%20Plains, last access: 7 June 2021.

B2SHARE: The Tall Tower Dataset, B2SHARE [data set], https://b2share.eudat.eu/records/159158152f4d4be79559e2f3f6b1a410, last access: 9 March 2020.

Basse, A., Callies, D., Grötzner, A., and Pauscher, L.: Seasonal effects in the long-term correction of short-term wind measurements using reanalysis data, Wind Energy Science, 6, 1473-1490, https://doi.org/10.5194/wes-6-1473-2021, 2021.

BNL: Current Observations, Brookhaven National Laboratory [data set], https://wx1.bnl.gov/, last access: 14 April 2020.

BPA: BPA Meteorological Information, Bonneville Power Administration [data set], https://transmission.bpa.gov/Business/Operations/Wind/MetData/default.aspx, last access: 17 March 2022.

Carta, J. A., Velázquez, S., and Cabrera, P.: A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site, Renewable and Sustainable Energy Reviews, 27, 362-400, https://doi.org/10.1016/j.rser.2013.07.004, 2013.

Chen, D., Zhou, Z., and Yang, X.: A measure-correlate-predict model based on neural networks and frozen flow hypothesis for wind resource assessment, Physics of Fluids, 34(4), 045107, https://doi.org/10.1063/5.0086354, 2022.
Dinler, A.: A new low-correlation MCP (measure-correlate-predict) method for wind energy forecasting, Energy, 63, 152-160, https://doi.org/10.1016/j.energy.2013.10.007, 2013.

de Assis Tavares, L. F., Shadman, M., de Freitas Assad, L. P., and Estefen, S. F.: Influence of the WRF model and atmospheric reanalysis on the offshore wind resource potential and cost estimation: A case study for Rio de Janeiro State, Energy, 240, 122767, https://doi.org/10.1016/j.energy.2021.122767, 2022.

DOE: U.S. Department of Energy, Atmosphere to Electrons [data set], a2e.energy.gov, last access: 13 February 2024.

Draxl, C., Wang, J., Sheridan, L., Aghili, C. T., Barton, R., Bodini, N., Buckhold, S., Edwards, P., Ferrier, J., Gilroy, N., Jung, C., Kotamarthi, R., Kumler, A., Mohammad, A., Olson, R., Phillips, C., Purkayastha, A., Young, E., Rosenlieb, E., Susmarski, P., and Tinnesand, H.: The WIND Toolkit Long-Term Ensemble Dataset, in preparation, 2024.

ECMWF: ERA5, ECMWF [data set], https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, last access: 10 January 2024.

FEMC: University of Vermont, Forest Ecosystem Monitoring Cooperative, Raw Proctor Maple Research Center Meteorological Data [data set], https://www.uvm.edu/femc/data/archive/project/proctor-maple-research-center-meteorological-monitoring/dataset/proctor-maple-research-center-air-quality-1/overview, last access: 15 April 2020.

Friedman, J. H.: Multivariate Adaptive Regression Splines (with discussion), The Annals of Statistics, Vol. 19, No. 1, 1991.

Hayes, L., Stocks, M., and Blakers, A.: Accurate long-term power generation model for offshore wind farms in Europe using ERA5 reanalysis, Energy, 229, 120603, https://doi.org/10.1016/j.energy.2021.120603, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Jekabsons, G.: Adaptive Regressions Splines toolbox for Matlab/Octave, version 1.13.0, http://www.cs.rtu.lv/jekabsons/Files/ARESLab.pdf, 2016.

Jimenez, T.: The Wind Powering America Anemometer Loan Program: A Retrospective, Tech. Rep., NREL/TP-7A30-57351, National Renewable Energy Laboratory, Golden, CO, USA, https://doi.org/10.2172/1081367, 2013.

Lackner, M. A., Rogers, A. L., and Manwell, J. F.: The round robin site assessment method: A new approach to wind energy site estimation, Renewable Energy, 33(9), 2019-2026, https://doi.org/10.1016/j.renene.2007.12.011, 2008.

LBNL: AmeriFlux, Lawrence Berkeley National Laboratory [data set], https://ameriflux.lbl.gov/sites/site-search/#filter-type=all&has-data=All&site_id=, last access: 16 April 2020.

Liléo, S., Berge, E. Undheim, O., Klinkert, R., and Bredesen, R. E.: Long-term correction of wind measurements, Tech. Rep., Elforsk Report 13:18, 2013.

Mifsud, M. D., Sant, T., and Farrugia, R. N.: A comparison of Measure-Correlate-Predict Methodologies using LiDAR as a candidate site measurement device for the Mediterranean Island of Malta, Renewable Energy, 127, 947-959, https://doi.org/10.1016/j.renene.2018.05.023, 2018.

Formatted: Font: (Default) +Body (Times New Roman)

Miguel, J. V. P., Fadigas, E. A., and Sauer, I. L.: The Influence of the Wind Measurement Campaign Duration on a Measure-Correlate-Predict (MCP)-Based Wind Resource Assessment, Energies, 12 (19), 3606, https://doi.org/10.3390/en12193606, 2019.

Milborrow, S.: Earth: Multivariate Adaptive Regression Spline Models [code] (derived from code by Hastie, T. and Tibshriani, R.), https://cran.r-project.org/web/packages/earth/index.html, 2016.

Murcia, J. P., Koivisto, M. J., Luzia, G., Olsen, B. T., Hahmann, A. N., Sørensen, P. E., and Als, M.: Validation of European-scale simulated wind speed and wind generation time series, Applied Energy, 305, 117794, https://doi.org/10.1016/j.apenergy.2021.117794, 2022.

NDBC: National Data Buoy Center [data set], https://www.ndbc.noaa.gov/, last access: 21 January 2024.

NREL: 2019COE_DW100_100kW_27.6, National Renewable Energy Laboratory [data set], https://nrel.github.io/turbine-models/2019COE_DW100_100kW_27.6.html, 2019.

NREL: Flatirons Campus Data, NREL [data set], https://www.nrel.gov/wind/nwtc/data.html, last access: 16 February 2022.

NRG Systems: Wind Resource Assessment. https://www.nrgsystems.com/products/applications/wind-resource-assessment/, accessed 30 August 2023.

Olauson, J.: ERA5: The new champion of wind power modelling? Renewable Energy, 126, 322-331, https://doi.org/10.1016/j.renene.2018.03.056, 2018.

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 Energy Science, 7(3), 1153-1169, https://doi.org/10.5194/wes-7-1153-2022, 2022.

PNNL: Distributed Wind Project Database. https://www.pnnl.gov/distributed-wind/market-report/datahttps://www.pnnl.gov/distributed-wind/data, accessed 12 January 2024.

Ramon, J., Lledó, L., Torralba, V., Soret, A., and Doblas-Reyes, F. J.: Which global reanalysis best represents near-surface winds?, Q. J. Roy. Meteorol. Soc., 145, 3236–3251, https://doi.org/10.1002/qj.3616, 2019.

Rogers, A. L., Rogers, J. W., and Manwell, J. F.: Comparison of the performance of four measure-correlate-predict algorithms, Journal of Wind Engineering and Industrial Aerodynamics, 93(3) 243-264, https://doi.org/10.1016/j.jweia.2004.12.002, 2005.

Seto, C.: Can wind power the North? Innovation, https://www.innovation.ca/projects-results/research-stories/can-wind-power-north, 30 August 2022.

Sheridan, L. M., Phillips, C., Orrell, A. C., Berg, L. K., Tinnesand, H., Rai, R. K., Duplyakin, D., and Flaherty, J. E.: Validation of wind resource and energy production simulations for small wind turbines in the United States, Wind Energy Science, 7, 659-676, https://doi.org/10.5194/wes-7-659-2022, 2022.

Soares, P. M. M., Lima, D. C. A., and Nogueira, M.: Global offshore wind energy resources using the new ERA-5 reanalysis, Environmental Research Letters, 15(10), 1040a2, 10.1088/1748-9326/abb10d, 2020.

Tang, X-Y., Zhao, S., Fan, B., Peinke, J., and Stoevesandt, B.: Micro-scale wind resource assessment in complex terrain based on CFD coupled measurement from multiple masts, Applied Energy, 238, 806-815, https://doi.org/10.1016/j.apenergy.2019.01.129, 2019.

Taylor, M., Mackiewicz, P., Brower, M. C., and Markus, M.: An analysis of wind resource uncertainty in energy production estimates, Proceedings of the European Wind Energy Conference and Exhibition, London, UK, 22-25 November 2004.

Weekes, S. M. and Tomlin, A. S.: Data efficient measure-correlate-predict approaches to wind resource assessment for small-scale wind energy, Renewable Energy, 63, 162-171, https://doi.org/10.1016/j.renene.2013.08.033, 2014.

Zakaria, A., Früh, W. G., and Ismail, F. B.: Wind resource forecasting using enhanced measure correlate predict (MCP), 6th International Conference on Production, Energy, and Reliability, Kuala Lumpur, Malaysia, https://doi.org/10.1063/1.5075569, 13-14 August 2018.