



Bias Characterization, Vertical Interpolation, and Horizontal Interpolation for Distributed Wind Siting Using Mesoscale Wind Resource Estimates

Dmitry Duplyakin, Sagi Zisman, Caleb Phillips,
and Heidi Tinneland

National Renewable Energy Laboratory

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2C00-78412
January 2021



Bias Characterization, Vertical Interpolation, and Horizontal Interpolation for Distributed Wind Siting Using Mesoscale Wind Resource Estimates

Dmitry Duplyakin, Sagi Zisman, Caleb Phillips,
and Heidi Tinnesand

National Renewable Energy Laboratory

Suggested Citation

Duplyakin, Dmitry, Sagi Zisman, Caleb Phillips, and Heidi Tinnesand. 2021. *Bias Characterization, Vertical Interpolation, and Horizontal Interpolation for Distributed Wind Siting Using Mesoscale Wind Resource Estimates*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-2C00-78412. <https://www.nrel.gov/docs/fy21osti/78412.pdf>.

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2C00-78412
January 2021

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

List of Acronyms and Abbreviations

CONUS	contiguous United States
IDW	inverse distance weighting
MAE	mean absolute error
MSE	mean square error
NREL	National Renewable Energy Laboratory
REX	REsource eXtraction Tool
RMSE	root mean square error
TAP	Tools Assessing Performance
WIND	Wind Integration National Dataset
WRF	Weather Research and Forecasting (model)
WRMDB	Wind Resource Meteorological Database
WRM-WRM	See Section 3 (page 9).
WTK	WIND Toolkit
WTK-WRM	See Section 3 (page 9).

Executive Summary

Much like their counterparts in utility-scale wind energy, developers of industrial, small-scale, and distributed wind turbine deployments need to understand and accurately characterize the wind resource to properly assess the power generation and financial ramifications during siting and planning. The National Renewable Energy Laboratory's (NREL's) WIND (Wind Integration National Dataset) Toolkit (WTK) (National Renewable Energy Laboratory 2020b) provides a best-in-class wind resource data set generated using the Weather Research and Forecasting (WRF) model (National Center for Atmospheric Research 2020). This data set includes parameters such as the wind speed, wind direction, and temperature at various heights, as well as atmospheric stability near the surface. These data are available at 2-km spatial resolution and 5-minute temporal resolution across 7 years, from 2007 to 2013, through a publicly accessible application programming interface (Phillips, et al. 2018). The U.S. Department of Energy's Tools Assessing Performance (TAP) project seeks to extend this data set to allow long-term resource estimates and leverage it to better equip distributed wind equipment manufacturers, owner-operators, and installation professionals with better tools for practical siting applications.

To practically utilize mesoscale WRF model outputs for wind turbine siting, gridded data (at 2-km resolution in this case) must be bias-corrected to account for local deviance from the national model, spatially projected (i.e., interpolated to the point of interest), and vertically downscaled (i.e., interpolated to the height of interest). Figure ES-1 shows this process schematically as it is imagined in the TAP project. The work described in this report considers the joint problems of bias estimation and vertical downscaling in order to define a basis for bias correction, horizontal combining/interpolation, and eventual development of a complete methodology for leveraging mesoscale inputs for location specific siting. This work also supports the related effort in the TAP project to develop wind speed deficit models for obstacles, to further refine these point-estimates with the specific complexities of the local environment.

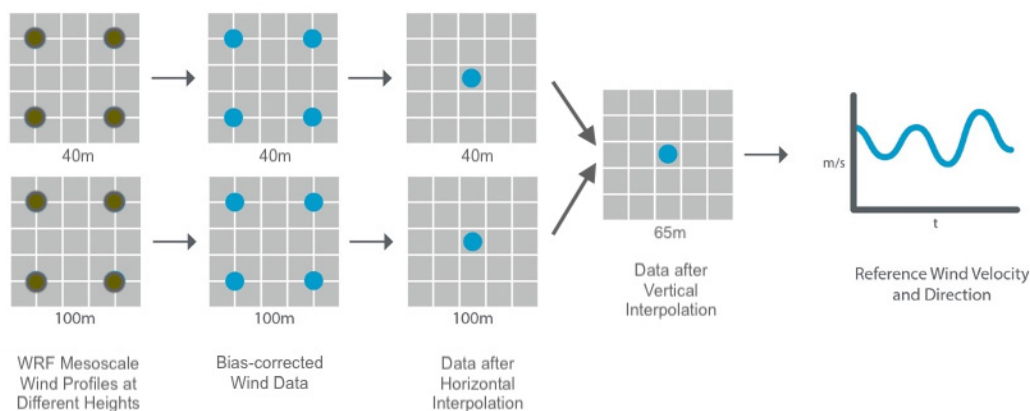


Figure ES-1. Schematic showing the process for developing accurate site-specific wind resource data from mesoscale model outputs

The study presented in this report uses NREL's Wind Resource Meteorological Database (WRMDB) (NREL 2020) as a basis for evaluating bias (error) arising when the WTK is applied to resource assessment for individual sites, as well as isolated error occurring from vertical downscaling or interpolation. From the part of this study with fixed horizontal interpolation

(nearest neighbor), we find that neutral log and power law methods minimize the error from vertical interpolation in this application. The nearest neighbor method for vertical interpolation also performs exceedingly well especially when considering its relative simplicity. This is likely because the WTK has good vertical resolution, by which the data for some heights can help construct accurate approximations for other heights' data. Overall estimation error in wind speed extrapolation from WTK to a specific WRMDB site can be characterized by a root mean square error (RMSE), which is in the range of 3.68–3.73 m/s. Further analysis of the possible causes of this error shows that the component related to the vertical interpolation may be responsible for 28% of this error. These errors do not appear to be strongly correlated with height, time of day, or season, and the methods evaluated appear appropriate for the ideal range of distributed wind heights: 20–60 m.

Based on the relative independence of WTK bias that is due to choosing a vertical interpolation strategy, and their strong performance in an isolated analysis, we recommend the use of neutral log law or neutral power law methods for performing vertical downscaling (interpolation) in a distributed wind siting context when turbulence and terrain roughness data are available. When these data are unavailable, nearest neighbor (i.e., taking data from the nearest height in the WTK model data) is a reasonable alternative. The high-level results of our vertical interpolation analysis—the distribution of errors and the method comparison—are shown in Figures ES-2 and ES-3.

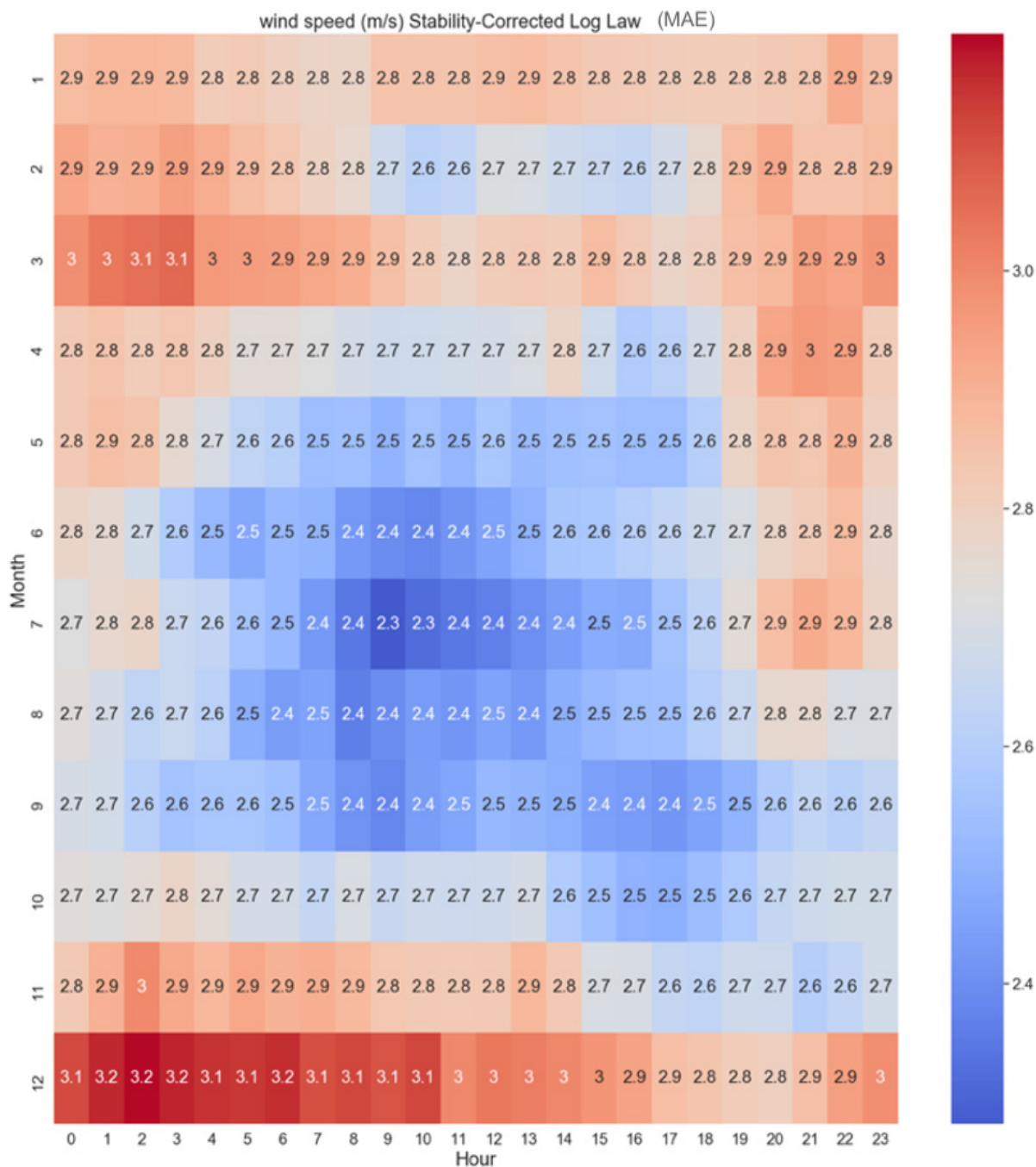


Figure ES-2. Mean absolute error for extrapolation of WTK data to observational points using stability corrected log law vertical interpolation method

MAE = mean absolute error

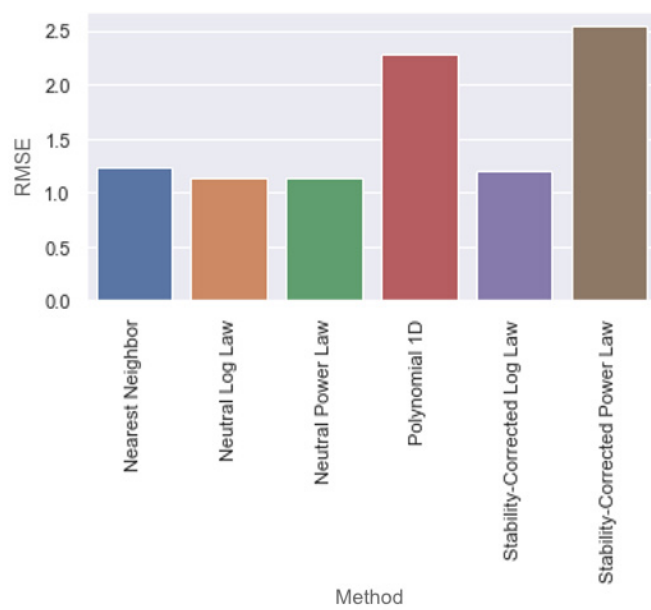


Figure ES-3. Average performance of isolated vertical interpolation methods (best-performing methods)

RMSE = root mean square error

In the part of our study focused on varying choices of both vertical and horizontal interpolation being used, we aim to determine the best performing method combinations for all validation sites, as well as for each such site considered individually. First, we evaluate 50 combinations (10 variants of horizontal interpolation and 5 selected best-performing vertical interpolation techniques) and we then compare the errors estimated against the WRMDB measurements. Based on this error analysis, we conclude all studied horizontal interpolation methods perform similarly, yet the inverse distance weighting method with 16 interpolation points outperforms the alternatives. Paired with Polynomial Degree 1 for vertical interpolation, the inverse distance weighting method creates the best method combination, which yields the lowest validation error in the set of the studied methods. Additionally, we determine that the estimates for the WRMDB sites in complex mountainous terrain are among the least accurate, which indicates the need for using more-complex, terrain-aware interpolation methods in these areas. Finally, we investigate the sensitivity of the validation error to the method selection for both horizontal and vertical interpolations. In this effort, we conclude the horizontal method sensitivity is *comparable to or larger than* the vertical method sensitivity for the set of methods being investigated. This conclusion is drawn based on our comparison of method sensitivity for the available data, yet a more direct and more conclusive summary can only be produced based on a larger set of validation measurements with denser spatial (horizontal) coverage.

Our interpolation error analysis reveals that though horizontal and vertical interpolation contribute to the overall estimation process, no individual method, or combination of the methods we study perfectly increases the estimate accuracy. In other words, regardless of your choice of interpolation strategy, you may expect to see the interpolation errors with the following characteristics:

- Mean absolute error (MAE) greater than 2.4

- Root mean square error (RMSE) above 3.3
- Mean error around or above 0.4 m/s.

We interpret these estimates as follows. The error distribution with the listed characteristics is the inherent and irreducible error distribution of the “virtual met mast” approach to which we adhere, at least for the spatiotemporal resolution and accuracy of this particular data set (and relative to the validation data we study). We believe this is an important result that helps establish common expectations for maximum fidelity of estimating wind resources (on an hourly basis at a given site) using the WIND Toolkit. This result also suggests more-complex approaches will be needed to reduce this error further, and the report provides details on where such estimate refinement would be the most needed.

Table of Contents

1	Introduction	1
1.1	Overview of Interpolation	2
1.2	Horizontal Spatial Interpolation Techniques	4
1.3	Vertical Interpolation Techniques	5
2	Data	7
2.1	WRMDB	7
2.2	WTK	8
3	Methods	9
4	Results	12
4.1	WTK-WRM Results	12
4.2	WRM-WRM Results	21
4.3	Learning from the WTK-WRM and WRM-WRM Results	28
4.4	Evaluating Combinations of Horizontal and Vertical Interpolation Techniques	29
4.4.1	Characterizing Mean Bias	30
4.4.2	Comparing MAE	31
4.4.3	Dissecting Spatial Patterns	32
4.4.4	Selecting Best Combinations	34
4.4.5	Conducting Sensitivity Analysis	35
5	Conclusions	37
	References	38

List of Figures

Figure ES-1. Schematic showing the process for developing accurate site-specific wind resource data from mesoscale model outputs.....	v
Figure ES-2. Mean absolute error for extrapolation of WTK data to observational points using stability corrected log law vertical interpolation method.....	vii
Figure ES-3. Average performance of isolated vertical interpolation methods (best-performing methods)	viii
Figure 1. Schematic showing the process for developing accurate site-specific wind resource data from mesoscale model outputs	3
Figure 2. Interpolated and observed wind speed timeseries shown for a selected WRMDB site.....	5
Figure 3. Map of WRMDB sites used in this study	7
Figure 4. Distribution of heights in WRMDB	10
Figure 5. Histogram showing distribution of observed wind speeds with kernel smoothed density function	12
Figure 6. Per-method distribution of interpolated wind speeds	13
Figure 7. Distribution of errors for each method (estimated minus observed)	14
Figure 8. 2D histogram showing WTK bias (error) with the best-performing vertical interpolation method (Polynomial Degree 1).....	16
Figure 9. Stability-corrected log law Monthly MAE for all sites and heights.....	17
Figure 10. Stability-corrected log law hourly MAE for all sites and heights	17
Figure 11. Heat map showing MAE for each month/hour.....	20
Figure 12. Relationship between error and height for well-performing vertical interpolation methods.....	20
Figure 13. Estimated wind speed distributions for each method for WRMDB	21
Figure 14. Distribution of error for each model.....	21
Figure 15. Comparison of RMSE for the best performing vertical interpolation methods.....	22
Figure 16. Errors for all methods by height.....	23
Figure 17. Overall MAE for each height grouping and method	24
Figure 18. Monthly MAE for all sites and heights by method.....	25
Figure 19. Hourly MAE for all sites and heights by method.....	26
Figure 20. Error according to height difference and extrapolated/interpolated	27
Figure 21. Error relationship between height difference between nearest input data height and target height.....	27
Figure 22. Degree of difference between method predictions	28
Figure 23. Violin plot comparing fraction of error due to vertical interpolation alone compared to combined error in extrapolating from WTK to WRMDB data.....	29
Figure 24. Comparison of mean bias estimates	31
Figure 25. Comparison of MAE estimates.....	32
Figure 26. Maps depicting WRMDB sites based on WTK-WRM errors: from small (top) to moderate (middle) and large (bottom)	33
Figure 27. Error variability caused by method selection	36

List of Tables

Table 1. Global Vertical Interpolation Error Measures for WTK Estimation	15
Table 2. Performance for the Hours with Minimum and Maximum Errors	19
Table 3. Performance for the Months with Minimum and Maximum Errors	19
Table 4. Isolated Vertical Interpolation Error for Those Methods Where Data are Sufficient to Draw Conclusions	22
Table 5. MAE Ratios for Locations and Heights Below 1.0	29
Table 6. Global Interpolation Error Measures for Estimates that Use IDW with 16 Points as Horizontal Interpolation and Various Techniques for Vertical Interpolation	32
Table 7. Top Three Methods in the Per-Site Selection of Horizontal Interpolation Method.....	34
Table 8. Top Three Methods in the Per-Site Selection of Vertical Interpolation Method	34
Table 9. Top Three Combinations of Interpolation Methods	35

1 Introduction

The U.S. Department of Energy's Tools Assessing Performance (TAP) project seeks to extend the National Renewable Energy Laboratory's (NREL's) Wind Integration National Dataset (WIND) to allow long-term resource estimates and leverage it to better equip distributed wind equipment manufacturers, owner-operators, and installation professionals with better tools for practical siting applications. To that end, TAP aims to provide access to newly available wind resource data and allow users to perform timely and accurate performance assessments. Additionally, TAP will develop several models for assessing the impact of obstacles within the project site. Unlike solar resource, the wind resource is highly sensitive to obstacles in the surrounding environment including buildings and vegetation. The project will then bring these two elements together with the most accurate approaches for horizontal and vertical interpolation to create a wind resource estimate that is specific to that project location. The overall goal of the TAP project is to address these resource assessment challenges that are hampering the widespread adoption of DW technology identified above and enable a market transformation. Successfully completing these tasks will lead to significant reductions in the uncertainty in pre-construction turbine performance estimates, thereby reducing the risk perceived by the financial community and consumers. As a result, better pre-construction estimates should help reduce financing costs and increase consumer adoption of DW technologies. This will ultimately lead to a reduction in the DW *levelized cost of energy* and a significant increase in the installed capacity of DW across the United States.

As we describe in this report, we plan to establish and utilize accurate interpolation procedures in order to facilitate the aforementioned TAP capabilities. In the context of interpolation efforts, we face several data-centered challenges. When assessing potential wind power generation, the site and height of a wind turbine must first be decided. One major challenge in real-world data collection is the discrete nature of the model and observational data in both space and time. WIND Toolkit (WTK) (National Renewable Energy Laboratory 2020b) is defined at a grid of 2 km X 2 km squares that spans the contiguous United States (CONUS), while at each grid point only nine heights are available. Thus, a tower at a particular location with a particular hub height will not (likely) be exactly estimated by the model. To ameliorate this problem, horizontal (x,y) as well as vertical (z) interpolation is performed on the wind speed model data for it to be used as an estimation of the point in question. Interpolation can include both numerical as well as physics-based techniques.

Because the time component is discrete in nature in both the WTK model and the observational data, an alignment in time or resampling of one or both of the time series might be required. A suboptimal choice for the alignment or resampling can be a source of error. Considering the discreteness of the time series, it should be noted that the errors of wind speed estimation from this type of analysis do not represent the true error in the wind speed. The wind speed variable is continuous, and thus the error measures and optimization are constrained by the time resolution and what the time series represents. For instance, if an hourly time series represents an averaging over the hour, any anomalous wind speeds such as short gusts are dampened or lost in the error calculations. We further discuss the temporal alignment and how it is integrated with the interpolation techniques later in this section, when we describe our overall wind speed estimation process.

Among the related literature, it is worth noting the recent study on adding “super resolution” to climatological models, including wind and solar data (Stengel, et al. 2020). The authors present their work on going from 100- to 2-km resolution for wind speed data—i.e., from the data obtained using NCAR Community Climate System Model (Taylor, Stouffer and Meehl 2012) to WTK data—using a novel adversarial deep learning approach. On one hand, their work can be viewed as comparable to our efforts at the high level because both studies seek to provide accurate estimates for arrays of locations for which high-resolution data is unavailable. On the other hand, they use compute- and data-intensive procedures to train their networks for acceptable interpolation, whereas we focus on simpler, much more lightweight estimation schemes that will likely provide sufficiently accurate interpolation results in most cases. Furthermore, we investigate vertical interpolation techniques, as well as the horizontal interpolation techniques, whereas their approach targets only the latter. They have a target spatial resolution, e.g., 2 km X 2 km—the resolution for which they train their deep learning networks—yet our work, in contrast, does not aim to resolve the wind patterns in the whole region with fixed resolution because we are primarily concerned with the accuracy of point estimates within TAP. Despite these differences, we can imagine combining these techniques (e.g., deep learning-based horizontal interpolation combined with domain-specific vertical interpolation); however, as we further discuss in Section 4.4.5 of this report, the spatial coverage of the observational data which are available to us and can be considered high-resolution data for the purposes of training is quite low. This limitation currently makes training of deep learning networks based on the available data infeasible, but the simpler methods we have identified and studied demonstrate satisfying results.

In the following parts of this section, we continue with the introduction of the selected interpolation methods and put them into context. The rest of the report is divided into four sections. In Section 2, we discuss the data we used, including preprocessing steps. The methods section covers the experimental design including the WTK bias and isolated vertical interpolation analyses in detail. The results section summarizes what we find through performing a sequence of four analyses using the available data and proposed techniques. In the conclusions section, we discuss key findings and outcomes.

1.1 Overview of Interpolation

Given a point (x,y,z) , we are interested in estimating the wind speed time series ideally at any arbitrary point in time. To do so, the main tool we use is the output of the WTK. As part of the estimation process, we have choices to optimize, as diagrammed in Figure 1 and which include:

- a) Aligning in time the WTK time series to observed time series
- b) Horizontally interpolating the WTK to an arbitrary (x,y) point
- c) Vertically interpolating the WTK to an arbitrary (z) point
- d) Bias correcting which includes all other modeling errors

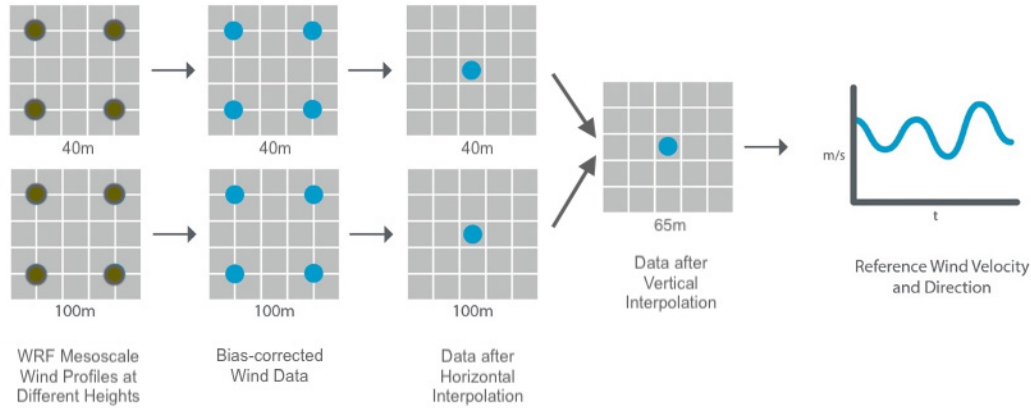


Figure 1. Schematic showing the process for developing accurate site-specific wind resource data from mesoscale model outputs

Assuming WTK time series are all aligned in time with each other, computations a), b), and c) above are commutative, but for now, we assume a composition of the functions in that order. The input is a set of WTK time series that can represent multiple (x, y, and z) neighboring wind speeds to our point-of-interest. Then after optimization, $c(b(a(\{wtk\})))$ represents the best estimation of the true wind speed minus bias correction. Bias correction is then an additional adjustment (function) that will get us a better estimate of the true wind speed by applying region or terrain-specific corrections to the national-scale WTK model outputs.

A closer look at the errors motivates the experimental approach we follow in this study. At a given spatiotemporal point, assuming an additive error model, we can write the true wind speed as $v_{true} = v_{estimate} + \epsilon$. The sources of error are due to suboptimal choices and irreducible error including:

- Temporal alignment
- Horizontal interpolation
- Vertical interpolation
- Bias correction
- Irreducible error.

We hold constant the temporal alignment and bias correction in the current study in order to investigate the impact of choosing horizontal and vertical interpolation techniques. When evaluating the error time series $Error = v_{true} - v_{estimate}$, we are evaluating a composite error of the five errors listed above. When we vary the choices of horizontal and vertical interpolations, we compare the composite errors and aim to choose the method that produces the lowest error. In cases where the error coming from the sources *other* than horizontal or vertical interpolations dominates the overall error, this error diminishes the effect of horizontal and vertical interpolation optimization. In fact, the results section (Section 4.1, page 12) for the WTK interpolation show exactly such a scenario.

The obvious question becomes: how do we choose the best technique possible? We need to evaluate the resulting errors of using such techniques using ground-truth measurements. The Wind Resource Meteorological Database (WRMDB) provides the basis for this ground-truth

assessment. It contains meteorological data from 205 sites, many at multiple heights,¹ during the same period as the WTK. Using WRMDB site data, we can assess bias (error) in the WTK and evaluate individual interpolation methods as a means for extrapolating from data on the WTK grid. The current study focuses on the evaluation of horizontal and vertical interpolation strategies.

1.2 Horizontal Spatial Interpolation Techniques

In the current study, we consider the following techniques and their variants:

- **Nearest Neighbor:** Given a set of nearby points, this method uses wind speed values from the geographic point that is closest to the point of interest. In contrast with the vertical interpolation (1D), this represents a multivariate (2D) approximation.
- **Linear Interpolation:** 2D spatial interpolation where polynomials of degree $n=1$ are fit to the wind speed data defined at (x,y) grid points. This method requires at least 2 points.
- **Cubic Interpolation:** 2D spatial interpolation where polynomials of degree $n=3$ are fit to the wind speed data defined at (x,y) grid points. This method is often used as an alternative to linear interpolation that provides smoother approximations. This method requires at least 4 points.
- **Inverse Distance Weighting (IDW):** 2D spatial interpolation that assigns values to unknown points that are calculated as *weighted averages* of the values available at the known grid points (GIS Geography 2020). This method is also known as *distance-based interpolation* and is formulated as follows:

$$\hat{v} = \frac{\sum_{i=1}^n \frac{1}{d_i} v_i}{\sum_{i=1}^n \frac{1}{d_i}},$$

where v_i – known values, d_i – distances from each known point to the point of interest, n – number of known points being considered, and the entire ratio gives an estimate of the unknown value at the point of interest.

We evaluate Linear Interpolation, Cubic Interpolation, and IDW methods with 4, 8, and 16 points used for interpolation to assess the role the amount of data used in interpolation plays in reducing, or increasing, the interpolation error. In contrast, Nearest Neighbor does not need a set of nearby points: it bypasses the computations by returning the win-speed values for the points closest to the points being investigated. Overall, this adds up to 10 variants of horizontal interpolation that we apply in our current investigation.

It is worth noting that the Linear Interpolation and Cubic Interpolation methods we use—both of which are implemented in the `interpolation.griddata` function inside the Python’s `scipy` package (SciPy.org 2020a)—perform the *data fitting* that uses all the points given to these methods. Unlike the Polynomial Degree 1 and Polynomial Degree 3 used in vertical interpolation (discussed in Section 4), these methods construct linear and cubic best-fitting

¹ The entire WRMDB data set covers 50 distinct heights at various sites, in the range of 15–180 m.

approximations from the given data utilizing all of it, when more than the minimum number of spatial points is used (two for Linear and four for Cubic Interpolation).

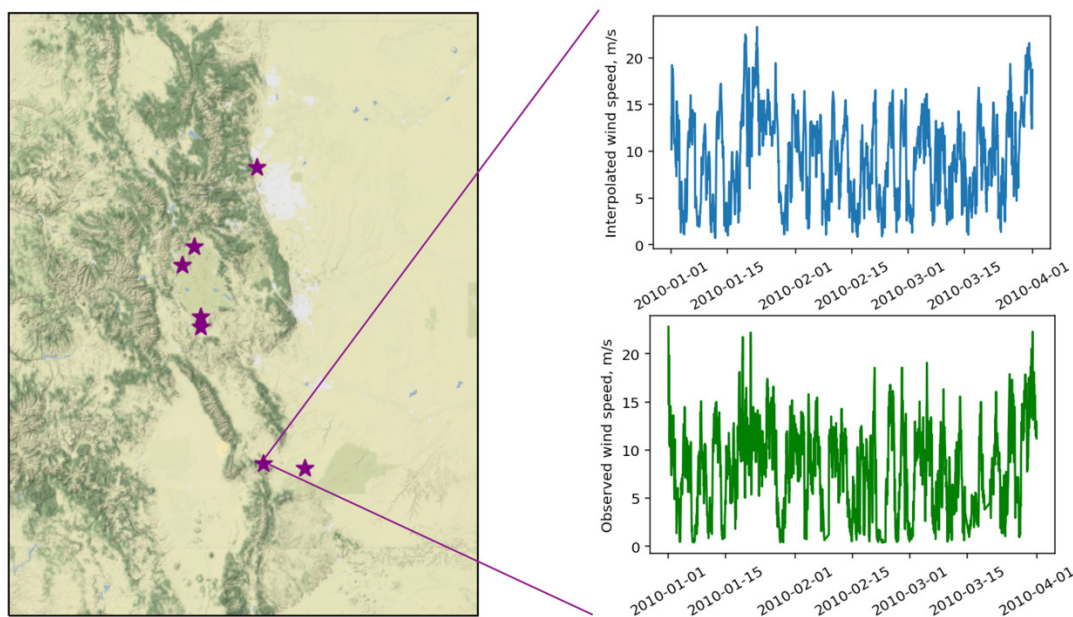


Figure 2. Interpolated and observed wind speed timeseries shown for a selected WRMDB site

Both timeseries depict data for 40 m hub height.

In Figure 2, we demonstrate the use of the Inverse Distance Weighting scheme with 16 interpolation points. Consistent with the overall windspeed estimation process shown in Figure 1, this visualization depicts a map of several WRMDB sites, one of which is characterized with two timeseries—horizontally interpolated and observed—for the first three months of 2010. While we can visually observe the similarities between the timeseries in this example, we need to quantify these differences across longer time intervals and for a number of available sites using different error metrics, which we return to and describe in detail in Section 4.

1.3 Vertical Interpolation Techniques

Broadly speaking, we evaluate two categories of interpolation: statistical/numerical and domain-specific. The numerical category comprises of techniques such as nearest neighbor and polynomial interpolation (including linear and quadratic). The domain-specific interpolation comprises the power/log law both in assumed neutral atmospheric stability regimes as well as in stability corrected techniques. Specifically, the vertical interpolation approaches considered are:

- **Nearest Neighbor:** Given an observation height, use the nearest neighboring height as the wind speed estimate.
- **Polynomial Interpolation:** Given n available WTK heights, we estimate the parameters of a polynomial that represents wind speed as a function of height in an exact fit. The wind speed of the observational height is evaluated. A different polynomial is fit at each point in time (t). When more data are available from other heights, higher degree polynomials are possible. In this study, we explore polynomials from degrees 1 to 8.

- **Neutral Log Law:** The neutral log law assumes neutral atmospheric stability conditions. A wind speed profile is characterized as wind speed proportional to the logarithm of the ratio of height and surface roughness. Our methods used known WTK wind speeds to solve for the proportionality constant and apply it to a desired height (Holtslag, Bierbooms and Van Bussel 2014).
- **Stability Corrected Log Law:** An extension of the neutral log law and a general form of log law that requires the Obukhov Length for determining stability condition. For the strongly stable regime, we apply the correction introduced in related literature in Equation 6 of (Holtslag, Bierbooms and Van Bussel 2014, Equation 6).
- **Neutral Power Law:** Use a relationship between the ratio of wind speeds at two heights to calculate an exponent of their height ratios. This exponent— α parameter—in this case is constant and taken to be 1/7, as described by Hsu (Hsu, Meindl and Gilhousen 1994).
- **Stability Influenced Power Law:** As an extension of Neutral Power Law, this method dynamically calculates the exponent α as a function of time using two known WTK wind speeds at different heights (Olauson and Bergkvist 2015).

2 Data

In this section, we describe the data that are available to us and which we use in TAP.

The analyses presented in this report rely on two data sets for estimation and validation:

1. WRMDB for estimation and validation
2. WTK for estimation

The supporting data that we use in the calculation of the stability-corrected log law interpolation includes:

1. National Land Cover Dataset for determining land cover classification of sites (Dewitz 2019)
2. Surface Roughness Constants as defined in the Continuum Open Source Repository (Walls 2020)

2.1 WRMDB

The WRMDB is a database of meteorological tower observations at various heights at sites in the CONUS. This database was curated from multiple prior observational studies and is available for use by NREL researchers on the NREL Eagle high-performance computing system. Thus, as part of quality control procedures, the coordinate verification and height verification were done for these sites before they were added to WRMDB. These sites are mostly geographically clustered in a few U.S. regions, including coastal Massachusetts, Colorado, Northern Missouri, and the Columbia River Gorge in Oregon. A map of site locations can be seen in Figure 3.

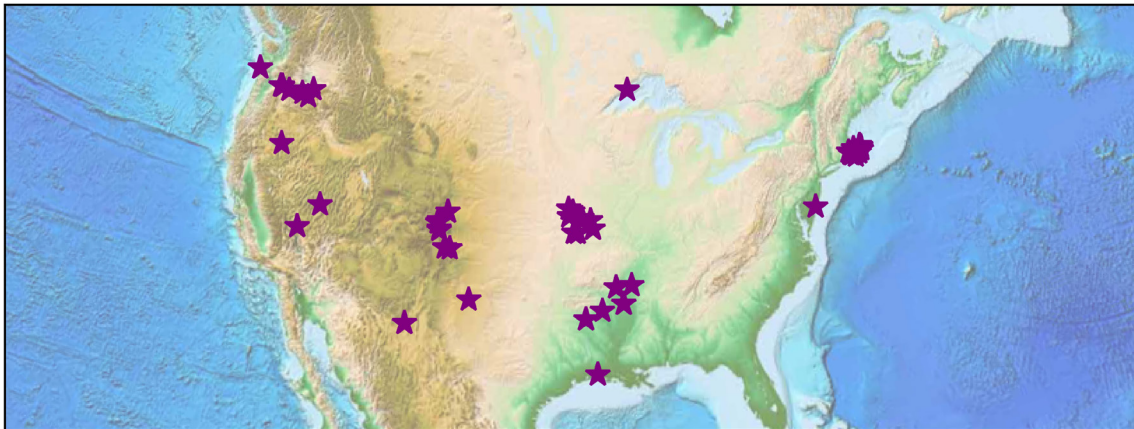


Figure 3. Map of WRMDB sites used in this study

The data are partitioned into channels (variables) that include wind speed, wind direction, temperature, and other variables. As mentioned previously, this study focuses on the wind speeds. The time of measurement of the wind speeds vary across locations and are mostly at the 10-minute time resolution rolled up to the *hourly resolution using an averaging scheme*. Some towers have measurements such as multiple wind speed channels at the same height because anemometers are designed to capture wind in different directions. As a preprocessing step, *the mean of these channels is taken to give one observational time series per site and height*.

As is common in industry, data points in the WRMDB for which the wind speed is less than or equal to 0.5 m/s are identified as instrument error and are excluded. We also exclude measurements in excess of 50 m/s, which are assumed to be similarly erroneous. One offshore site was also excluded, considering the relative inapplicability to the distributed wind context and our selection of vertical interpolation techniques, some of which assume terrestrial constants. Additionally, one WRMDB site was identified as providing unreliable measurements for its extremely high wind speeds, and it was removed from the vertical interpolation analysis, which relied on 64 sites for validation. After further per-site error investigation, we identified one site with abrupt and significant regime changes in the wind measurements and excluded it from the our study, which focused on horizontal evaluation. Thus, our horizontal analysis (discussed in Section 4.4) used the data from 63 WRMDB sites and 154 unique site-height combinations.

2.2 WTK

The WTK provides 2-km resolution of estimated atmospheric data variables for the CONUS based on the WRF model. The main variables of interest for our analysis are the wind speeds at a column of heights (10 m, 40 m, 60 m, 80 m, 100 m, 120 m, 140 m, 160m, 200 m) and the atmospheric stability variable called the *inversemoninobukohvlength_2m*. This analysis uses the *hourly resolution* of WTK.

3 Methods

This report covers the findings from four interrelated analyses we performed as part of our overall study:

1. We analyze WTK error and its relationship with the choice of vertical interpolation techniques using data from the WRMDB as a measure of “ground truth.” We call this approach the WTK-WRM analysis. We keep horizontal interpolation constant and use the nearest neighbor as the baseline.
2. We perform an analysis of vertical interpolation in isolation by evaluating the efficacy of each interpolation strategy at each WRMDB site. We call this approach WRM-WRM. No horizontal interpolation is involved in this analysis.
3. We compare and integrate the first two analyses.
4. Informed by the results from the first three analyses, we analyze the effect of varying choices of both horizontal and vertical interpolation techniques on the WTK-WRM errors. By doing this for the aggregate error measures and also on a tower-by-tower basis, we characterize how well each combination of choices performs and learn horizontal properties of the studied error distributions. As part of this analysis, we compare the levels of variability caused by selection of each vertical and special interpolation method, trying to answer the question about which method plays the biggest role in impacting the overall validation errors we observe.

For all these analyses, we align data to one-hour estimates of wind speed (m/s). In the WTK data, these estimates are instantaneous estimates, while for the WRMDB data, they are hourly averages.

For the WRM-WRM analysis, we follow a leave-one-out design, where all 58 WRMDB sites with more than one height are used. At each site, we “hold out” data from one height and attempt to interpolate these data from the remaining heights. Though generally this methodology is fair, it should be noted that the WRMDB does include potential skew in the distribution of heights (in addition to horizontal skew from the available locations and their terrain environments). Figure 4 provides a distribution of heights in the data set, showing that those heights in the target range for distributed wind applications (i.e., 20–60 m) are most well represented in the data.

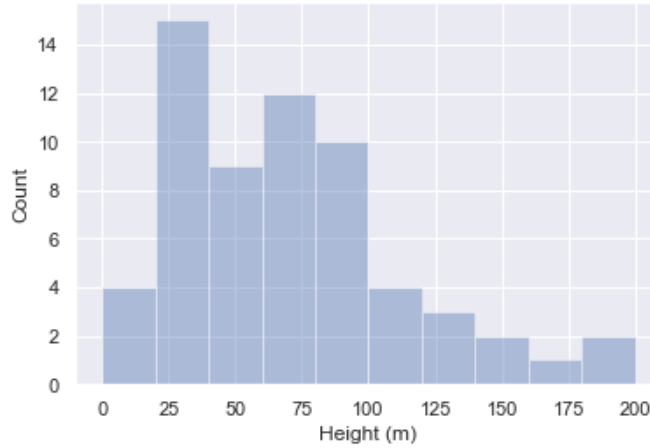


Figure 4. Distribution of heights in WRMDB

The plot shows 58 sites with more than one height.

In some scenarios, the vertical interpolation functions failed to make a reasonable estimate. In alignment with preprocessing, predicted wind speeds over 50 m/s and below 0.5 m/s are removed from the data, along with negative or infinite predictions. These data represent a very small fraction of the overall results, but they would significantly skew results and amount to a “failure to predict” by the vertical interpolation methods.

For each analysis, estimation and characterization of error is first evaluated visually, using observation/prediction scatterplots, histograms, and kernel-smoothed density plots. Relationships between error and time of day, season, and height are also explored. In addition, aggregate measures of error are considered, including:

- Statistical Bias in the Wind Speed Estimator: alternatively, we can refer to it as the mean error; though we use these terms interchangeably, they should not be confused with the word bias in *bias correction*.
 - $(1 / N) * \sum_{i=1}^N (v_{true} - \widehat{v_{est}})$
- Mean Absolute Error (MAE)
 - $(1 / N) * \sum_{i=1}^N |v_{true} - \widehat{v_{est}}|$
- Root Mean Squared Error (RMSE)
 - $(1 / N) * \sum_{i=1}^N (v_{true} - \widehat{v_{est}})^2$

Each of our four analyses starts with the spatial data in WTK structured as an array of points with (latitude, longitude) coordinates. Each point has a number of values associated with it for the variables (e.g., “winddirection_100m” or “windspeed_40m”). To transform these data into the state where we can calculate distances between points (which is a prerequisite for horizontal interpolation), we use the U.S. Geological Survey’s version of the Albers equal-area conic projection (spatialreference.org 2020). This projection is frequently used by the U.S. Geological Survey and the United States Census Bureau for the spatial data sets primarily within the CONUS. We order the results of these transformations, where the point coordinates are expressed in meters, based on the distances from each of the specific locations being

investigated and we then feed the corresponding wind speed WTK estimates into the horizontal interpolation methods for different heights and locations on the map.

To make this processing effort efficient, we do not search and process the data for *all* grid points in WTK, but rather we apply the aforementioned projection only to the grid points identified using NREL's REsource eXtraction Tool (REX) (National Renewable Energy Laboratory 2020a). This tool helps us find the points (1, 4, 8, or 16 maximum) that are closest to each of the WRMDB towers. REX, which is built on the NREL Highly Scalable Data Service (Phillips, et al. 2018) and utilizes the k-d tree-based technique (SciPy.org 2020b) for fast nearest-neighbor lookups, has proved to be efficient enough in our work.

4 Results

In this section, we describe the specifics of how we perform the analyses outlined in Section 3 and present our results. When possible, we discuss the instances where different findings relate to each other, aiming to better explain the observed patterns in studied wind estimates and their errors.

4.1 WTK-WRM Results

We start with the analysis of WTK errors and their relationship with the choice of vertical interpolation techniques using WRMDB data for validation. Figure 5 depicts the distribution of the wind speeds observed globally across all time for all sites and heights. Figure 6 shows similar distributions for each interpolation method. All studied interpolation techniques look similar at the level of the entire distribution and demonstrate good agreement with the observational data. In other words, all these wind speed distributions have peaks around 5 m/s, while the long tails of these distributions represent small fractions of speeds that are above 20 m/s.

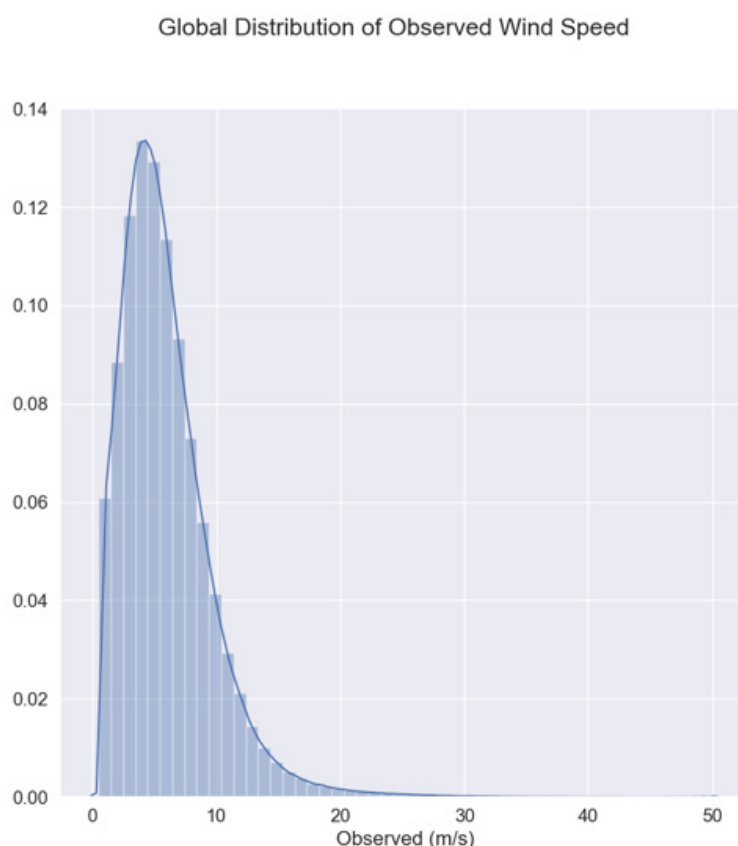


Figure 5. Histogram showing distribution of observed wind speeds with kernel smoothed density function

In Figure 7, we see the global distribution of wind speed errors—measured as *estimated minus observed* wind speeds. Any deviation of the mean of the error distributions around zero represents the statistical bias of the estimation, and the variability around zero is related to the RMSE. These errors appear normally distributed around the mean of approximately 0.4, irrespective of the

method being used, suggesting the overall errors are Gaussian; and generally, the overall bias in the WTK is positive, indicating that wind speeds are slightly and systematically high compared to observed.

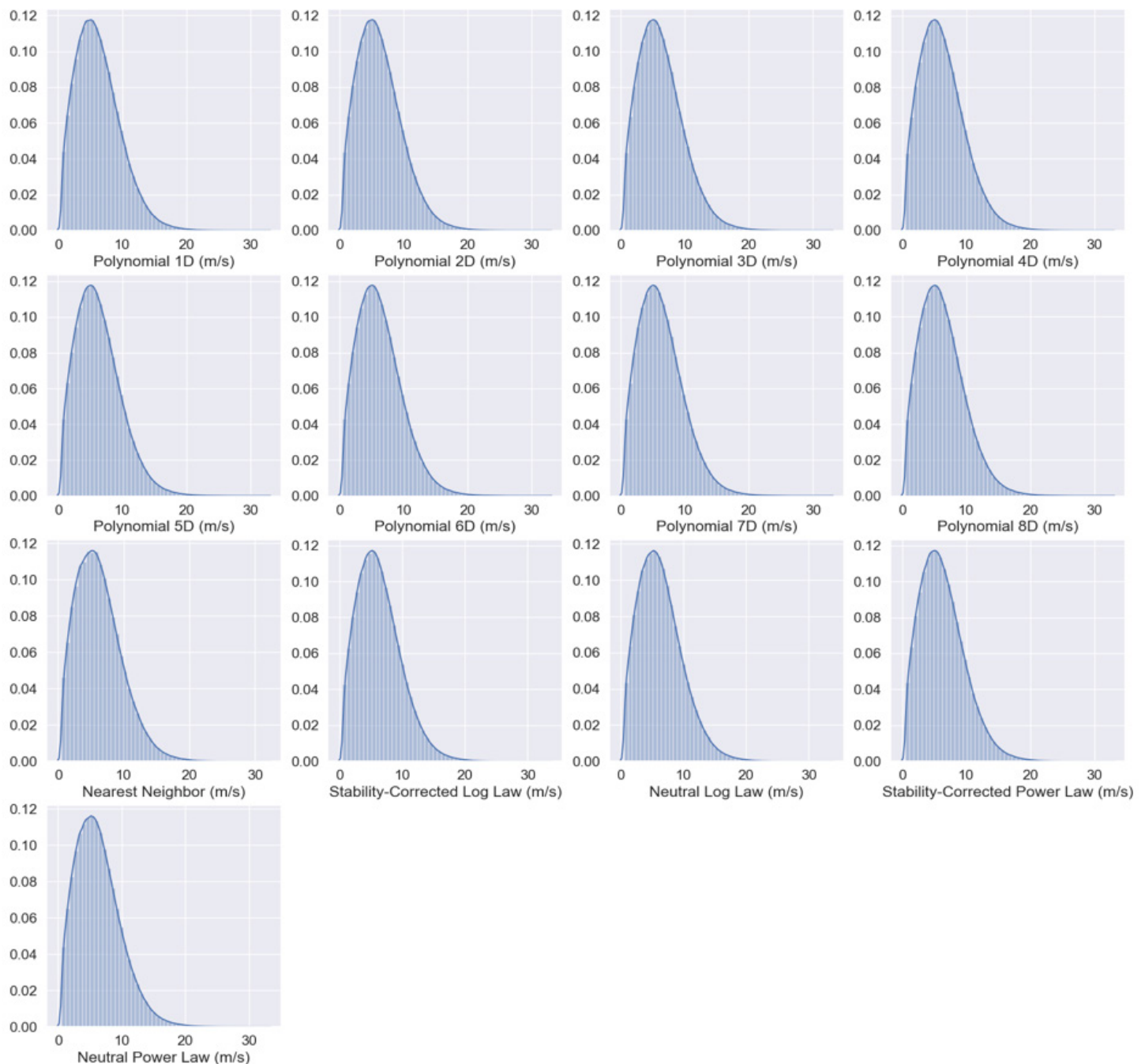


Figure 6. Per-method distribution of interpolated wind speeds

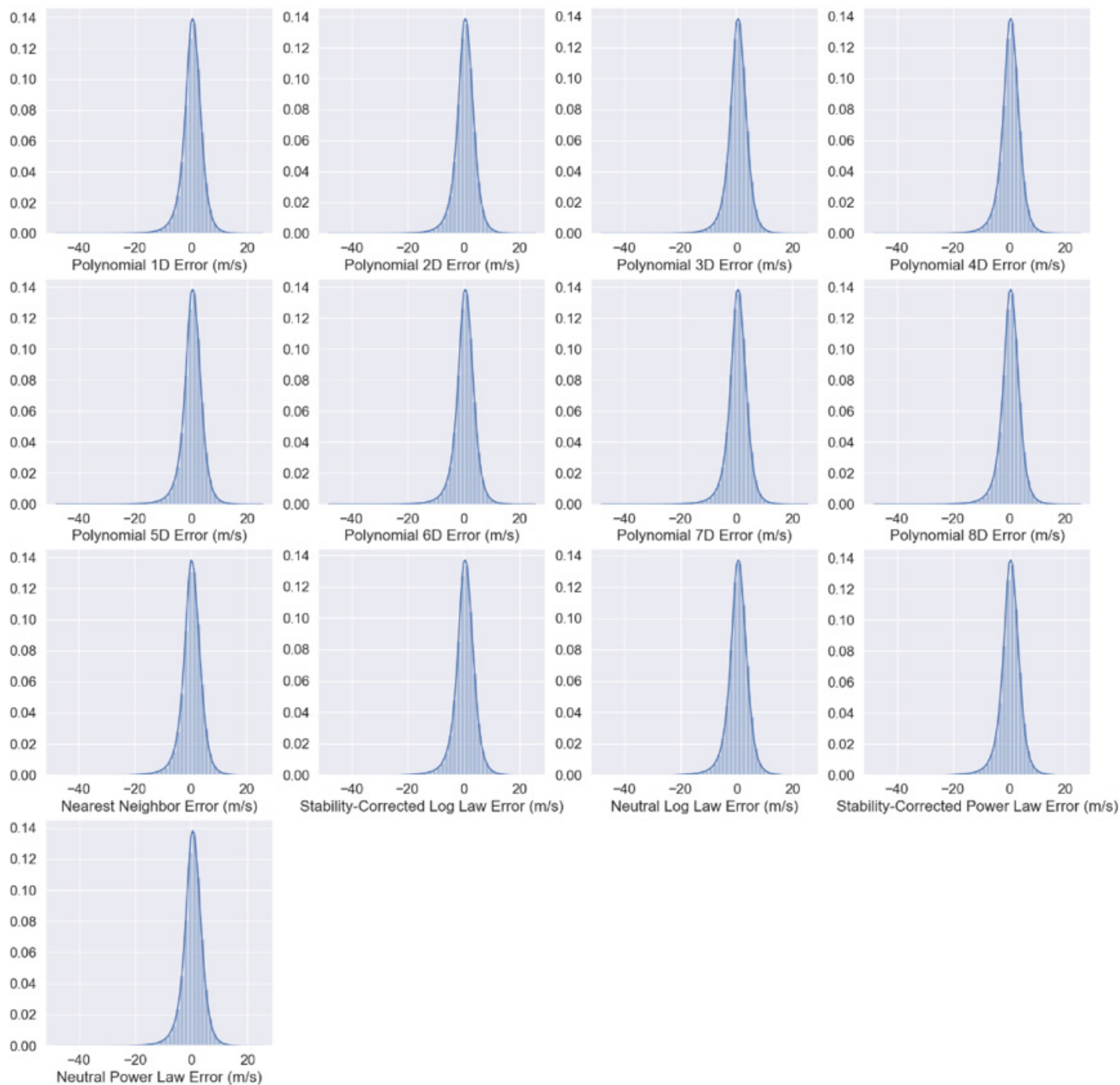


Figure 7. Distribution of errors for each method (estimated minus observed)

Aggregate performance errors are provided in Table 1. Based on this analysis, we can see that the Polynomial Degree 1 performs marginally better than the others in all error measurements. The Polynomial Degree 1 is the linear interpolation between the point above and below a given height to be estimated or a linear extrapolation for observational heights outside the range of WTK heights. Though it is tempting to assign a “best” interpolation technique, we can see that the variability among the techniques is quite small. Therefore, as mentioned earlier, it is very likely that the dominating feature in the errors is a different error.

Table 1. Global Vertical Interpolation Error Measures for WTK Estimation

Interpolation Technique	Mean Error (m/s)	MAE (m/s)	RMSE (m/s)
Polynomial Degree 1	0.362	2.68	3.68
Polynomial Degree 2	0.388	2.69	3.69
Polynomial Degree 3	0.398	2.70	3.69
Polynomial Degree 4	0.400	2.70	3.69
Polynomial Degree 5	0.401	2.70	3.70
Polynomial Degree 6	0.403	2.70	3.70
Polynomial Degree 7	0.406	2.70	3.70
Polynomial Degree 8	0.410	2.70	3.70
Nearest Neighbor	0.341	2.70	3.70
Stability Corrected Log Law	0.445	2.72	3.73
Neutral Log Law	0.433	2.72	3.73
Stability Corrected Power Law	0.402	2.70	3.70
Neutral Power Law	0.385	2.71	3.71

Figure 8 provides a kernel density estimate of the global hourly wind speeds for the Polynomial Degree 1 vertical interpolation technique. This represents another view of the mean error and error deviations. Other techniques yield rather similar plots and are omitted here. On the x-axis, we see the estimated wind speed, and on the y-axis, the observed wind speed. The red line represents the $y=x$ line, and any translational asymmetries of the bulk of the distribution around the line represent the bias. Indeed, we can see that the bulk of the distribution lies under the line which represents an *overestimation* bias. This matches the results in Table 1. The variance of the error measures, specifically the MAE and RMSE can clearly be seen in the kernel density estimate plot as the width of the distribution around the red line.

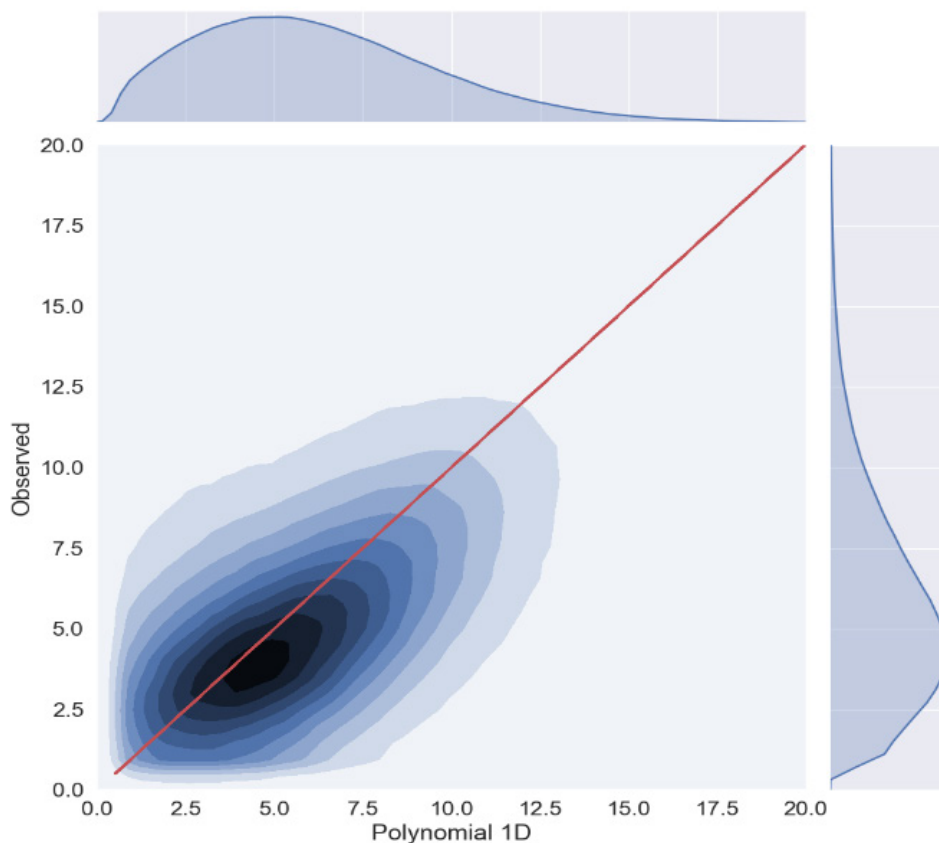


Figure 8. 2D histogram showing WTK bias (error) with the best-performing vertical interpolation method (Polynomial Degree 1)

Next, we consider an hourly (diurnal) partition, a monthly partition, and one by height. There are two main reasons to partition the input space and calculate separate error measures for each element of a partition. First, we do this to understand how each interpolation method performs in different scenarios and to understand the strengths and weaknesses of each. Second, there is no *a priori* reason to assume one vertical interpolation model has to be applied at each spatiotemporal point.

In the discussion that follows, the vertical interpolation method that we choose to illustrate the analysis is the stability corrected log law. Other methods show almost identical distributions and are omitted here. The implication is discussed at the end of the section.

In Figure 9 and Figure 10, we see MAE on a monthly and hourly basis, respectively, for the stability corrected log law vertical interpolation method. The plots for the other interpolation techniques are rather similar. We can see that the MAE of the stability corrected log law is lowest at the 11th hour of the day and highest during the 1st hour. At the monthly level, the MAE is lowest in September and highest in December. For the hourly partition, the minimum MAE occurs at various hours (mostly the 17th hour) but with insignificant differences in the minimum value and the maximum occurs at the 1st hour for each. For the monthly data, the minimum MAE occurs in September and the maximum MAE occurs in December for each vertical interpolation technique.

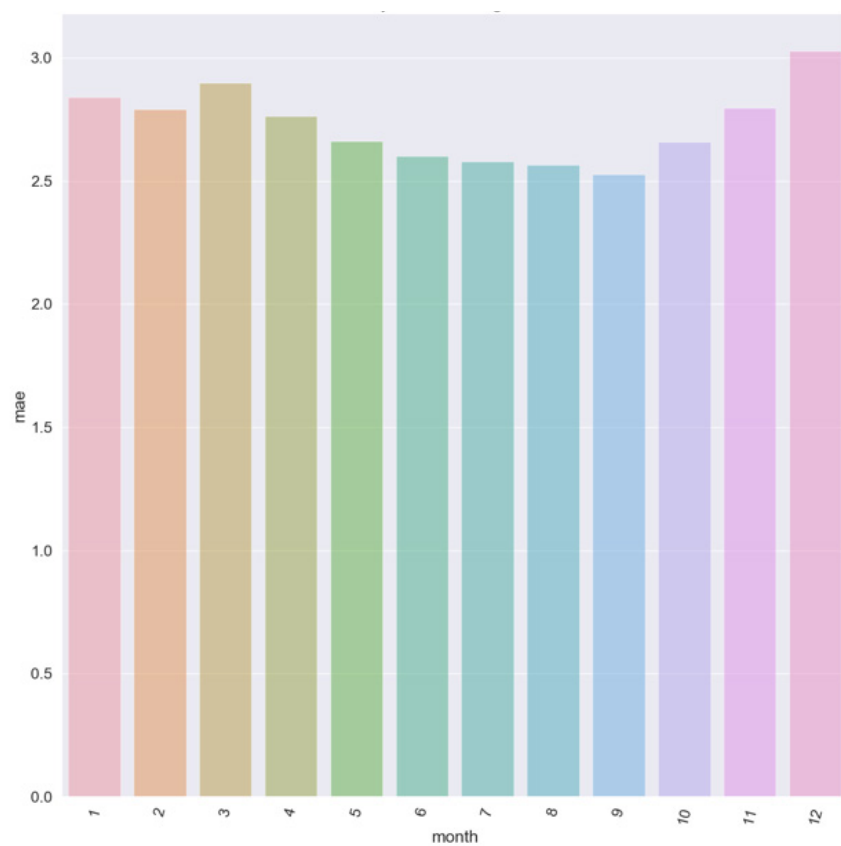


Figure 9. Stability-corrected log law Monthly MAE for all sites and heights

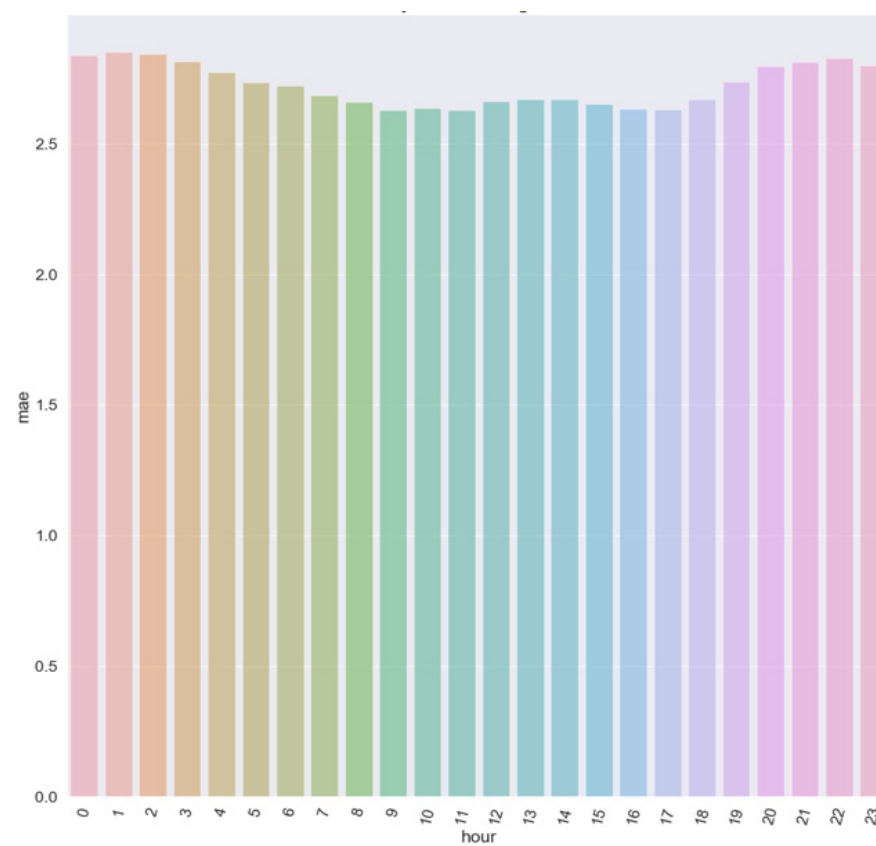


Figure 10. Stability-corrected log law hourly MAE for all sites and heights

We show the same error distributions in the tabular form in Table 2 and Table 3 (page 19). Table 2 shows the actual MAE for the hours corresponding to the minimum and the maximum error values. And Table 3 shows the MAE corresponding to the months with the minimum and the maximum estimation errors.

We can also combine the month and the hour into 2D heat maps as shown in Figure 11 (page 20). The lowest MAE times can be seen to be in the summer during the day and the highest MAE times during the winter largely at night. According to the previous related work (Holtslag, Bierbooms and Van Bussel 2014) (Newsom, et al. 2020), we explain these observations as follows. The daytime wind trends are defined by the unstable, well-mixed conditions that lead to little vertical wind shear, and the model estimates tend to handle this phenomenon well. The nighttime trends feature stable conditions with stronger vertical wind shear and the phenomena like low-level jets. The models tend to struggle more with capturing the degree of the observed wind shear. Also, the Monin-Obukhov similarity theory, which is incorporated into the physical interpolation schemes, tends to struggle with strongly stable conditions. At the same time, seasonally, slower wind speeds in the summer reduce the MAE estimates, as well as reduce their variability, while winter months generally have stronger winds and the larger MAE ranges.

Another partition can be performed by height categories. In Figure 12 (page 20), for a select number of methods, we plot the distribution of MAEs for a given height category for all locations. There appear to be higher absolute MAEs at the lower heights, but a deeper analysis needs to be performed to understand the exact dynamics at play. Moreover, there does not appear to be a strong need to select different methods at different heights. Other interpolation methods not plotted here show similar plots.

The results of the partitioned data for WTK estimation show that the WTK estimation accuracy varies in month, hour, and height. This result aligns with previous results from the TAP project. The vertical interpolation analysis shows that optimizing for a vertical interpolation technique for a given element of a data partition is unnecessary. Indeed, the results agree with the global analysis whereby minimal variability in error measures is observed.

Globally, the WTK estimation exhibits MAE of roughly 2.7 m/s and a mean error of 0.35–0.45 m/s. The choice of vertical interpolation strategy does not meaningfully impact the overall scale of error; in fact, this choice likely magnifies the impact of error correction efforts. It is important to remember that for each location and height, the WTK can potentially use any one of nine different heights. For an arbitrary height chosen, there likely exists a WTK height nearby. In other words, by using WTK data, there is less reliance on vertical interpolation. Such an observation might also help explain the low variability in performance among vertical interpolation techniques.

Interpolation Technique	Minimum MAE	Maximum MAE
Polynomial Degree 1	2.59	2.80
Polynomial Degree 2	2.60	2.81
Polynomial Degree 3	2.60	2.82
Polynomial Degree 4	2.60	2.82
Polynomial Degree 5	2.60	2.82
Polynomial Degree 6	2.60	2.82
Polynomial Degree 7	2.60	2.82
Polynomial Degree 8	2.61	2.83
Nearest Neighbor	2.60	2.81
Stability Corrected Log Law	2.63	2.85
Neutral Log Law	2.63	2.84
Stability Corrected Power Law	2.60	2.82
Neutral Power Law	2.61	2.82

Table 2. Performance for the *Hours* with Minimum and Maximum Errors

Interpolation Technique	Minimum MAE (9th month)	Maximum MAE (12th month)
Polynomial Degree 1	2.49	2.99
Polynomial Degree 2	2.50	3.00
Polynomial Degree 3	2.50	3.01
Polynomial Degree 4	2.50	3.01
Polynomial Degree 5	2.50	3.01
Polynomial Degree 6	2.50	3.01
Polynomial Degree 7	2.50	3.01
Polynomial Degree 8	2.51	3.01
Nearest Neighbor	2.50	3.02
Stability Corrected Log Law	2.53	3.03
Neutral Log Law	2.52	3.03
Stability Corrected Power Law	2.50	3.01
Neutral Power Law	2.51	3.03

Table 3. Performance for the *Months* with Minimum and Maximum Errors

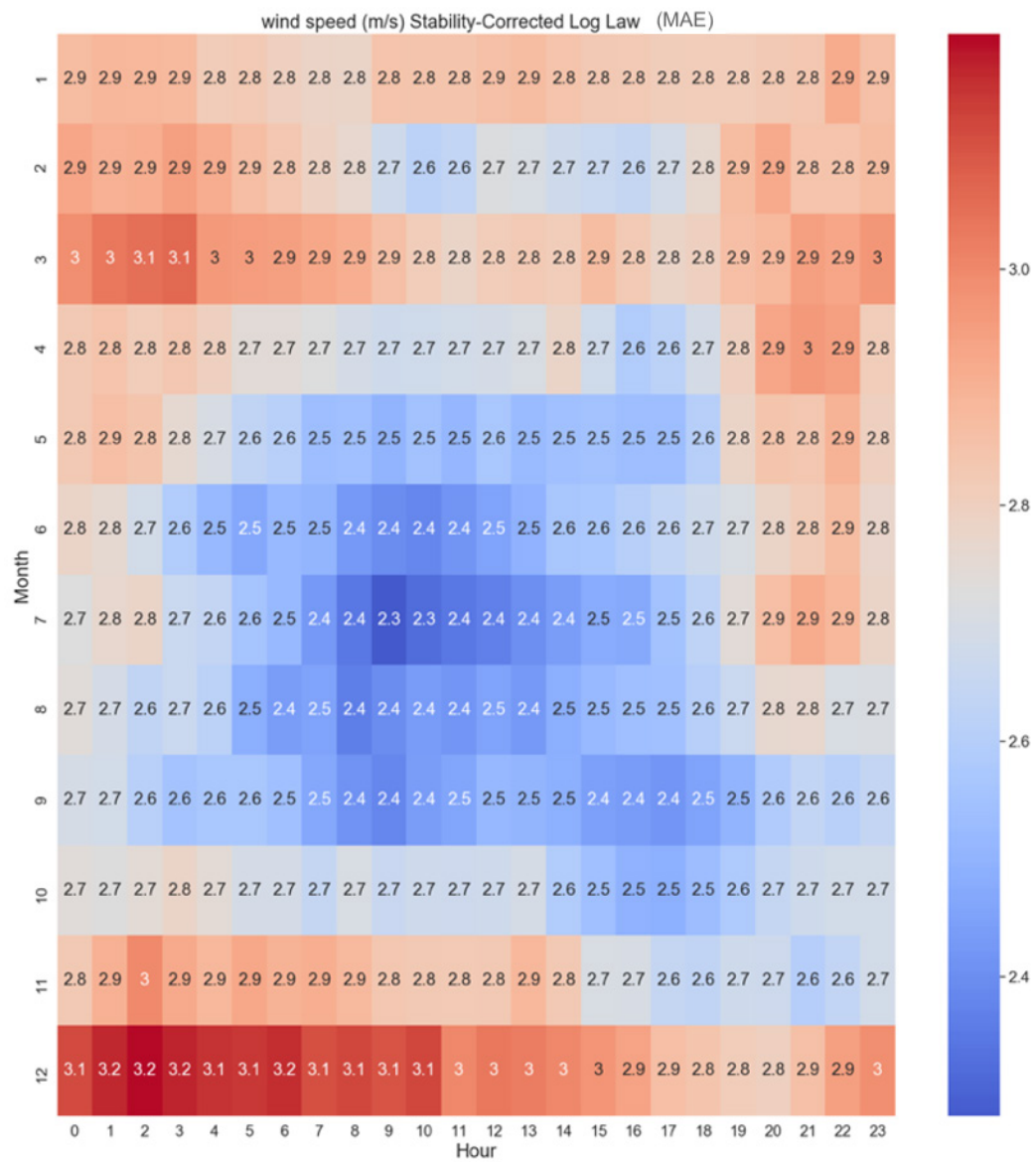


Figure 11. Heat map showing MAE for each month/hour

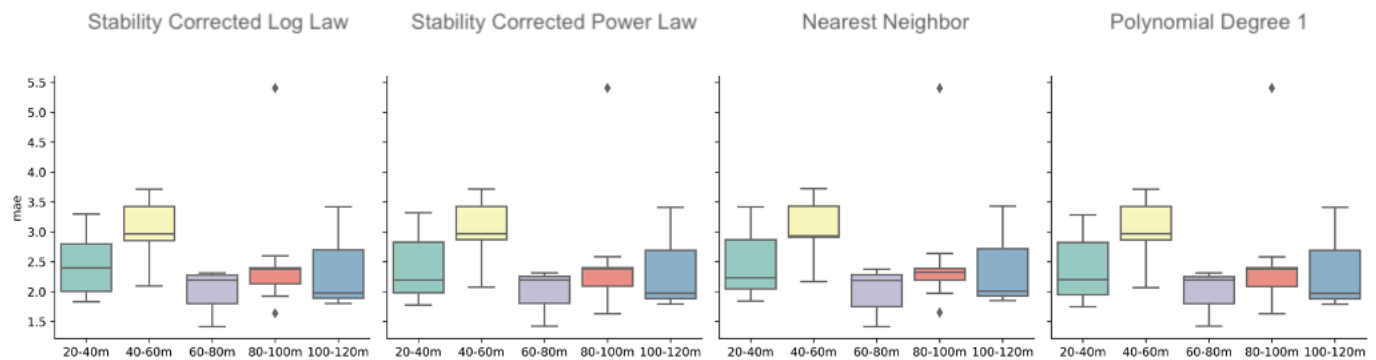


Figure 12. Relationship between error and height for well-performing vertical interpolation methods

4.2 WRM-WRM Results

We continue with the analysis of vertical interpolation at each WRMDB site. Figure 13 shows the global interpolated distributions of interpolated wind speeds in the WRMDB data. We see that these are similar to the WTK estimations for the same interpolation techniques.

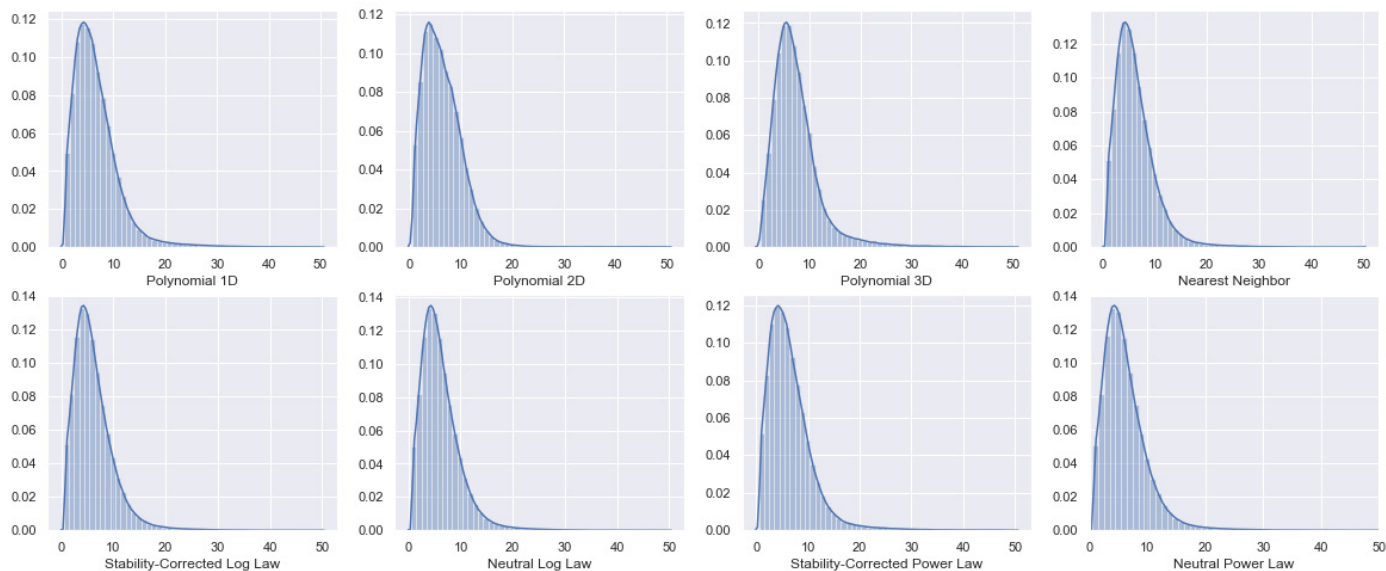


Figure 13. Estimated wind speed distributions for each method for WRMDB

The plot in Figure 14 shows the error distributions defined as observed wind speed minus estimated wind speed for different methods. A similar interpretation to the same plots for WTK estimation can be employed. We see that the peaks of these error distributions are around zero and the distributions are roughly symmetric around zero (i.e. positive and negative estimation errors have approximately equal probabilities) for all studied methods.

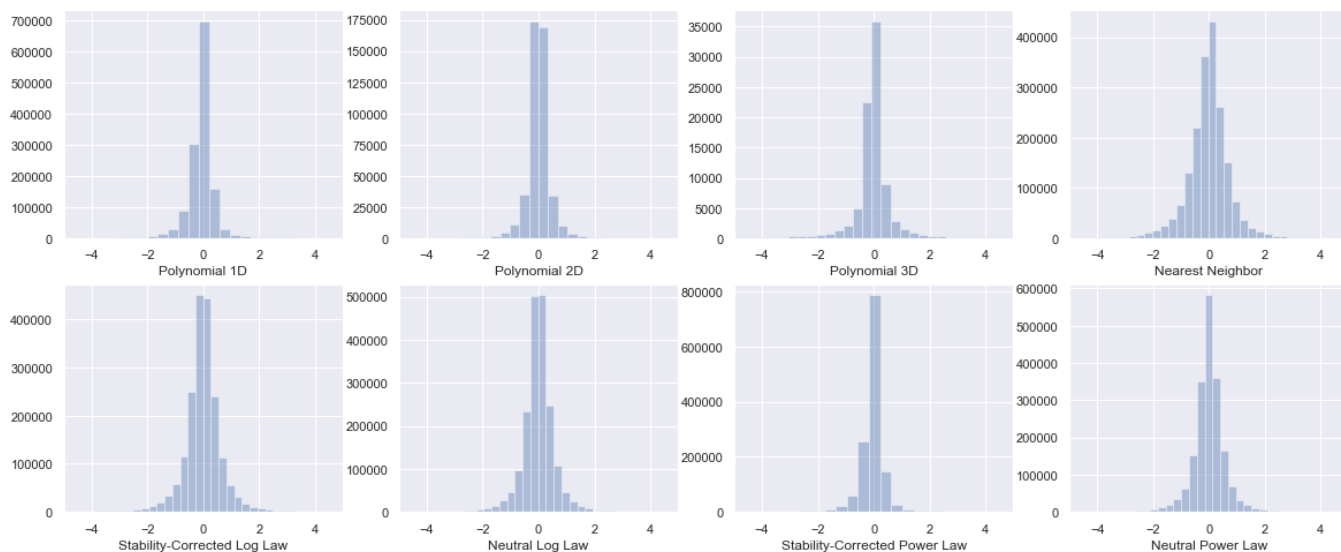


Figure 14. Distribution of error for each model

A global analysis of the vertical interpolation techniques is performed on all the sites and heights. This allows us to globally assess the error and choose the *best* interpolation technique. Figure 15 shows RMSE estimates for these techniques. We can immediately see that the Polynomial Degree 1 as well as the stability corrected power law perform the poorest of all techniques and the neutral log law or neutral power law performs the best. Table 4 summarizes the results for easy reference.

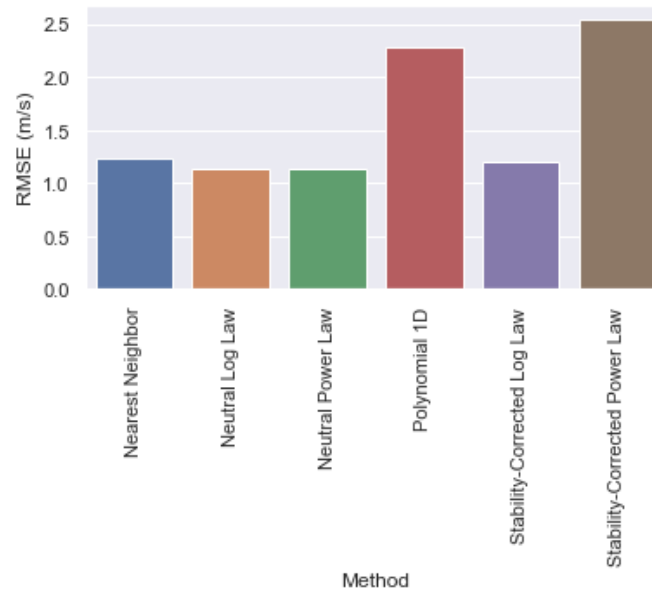


Figure 15. Comparison of RMSE for the best performing vertical interpolation methods

Table 4. Isolated Vertical Interpolation Error for Those Methods Where Data are Sufficient to Draw Conclusions

Interpolation Technique	Sites (n)	Heights (n)	Samples (n)	MAE (m/s)	RMSE (m/s)
Polynomial Degree 1	42	147	1,428,518	0.62	2.28
Nearest Neighbor	64	190	1,931,871	0.62	1.24
Stability Corrected Log Law	64	190	1,931,871	0.56	1.20
Neutral Log Law	64	190	1,931,871	0.49	1.14
Stability Corrected Power Law	42	146	1,408,403	0.61	2.54
Neutral Power Law	64	190	1,931,871	0.49	1.14

Similar to the WTK-WRM analysis, we can partition our data based on several variables. In Figure 18 (page 25), we compare the vertical interpolation methods by months, and in Figure 19 (page 26), we compare by hours. We can clearly see the variation of the errors in these variables. We see that—like the global data for the observational estimation and unlike the WTK estimation—there does exist variability among the methods. However, there does not appear to be a need to select a different method per month or hour. In both cases, the neutral log and power law minimize the error across each element of the partition.

Partitioning by height yields MAEs that are slightly lower at the lower heights (Figure 16). When looking at all the methods individually by height category (Figure 18), we can see that, surprisingly, the polynomial methods perform optimally at lower heights. However, before concluding this, we should note that the Polynomials Degree 2 and Polynomial Degree 3 techniques require 3 and 4 heights and the observational data set do not contain enough such locations. So, this result is likely an artifact of a dearth of data, but it is worth investigating.

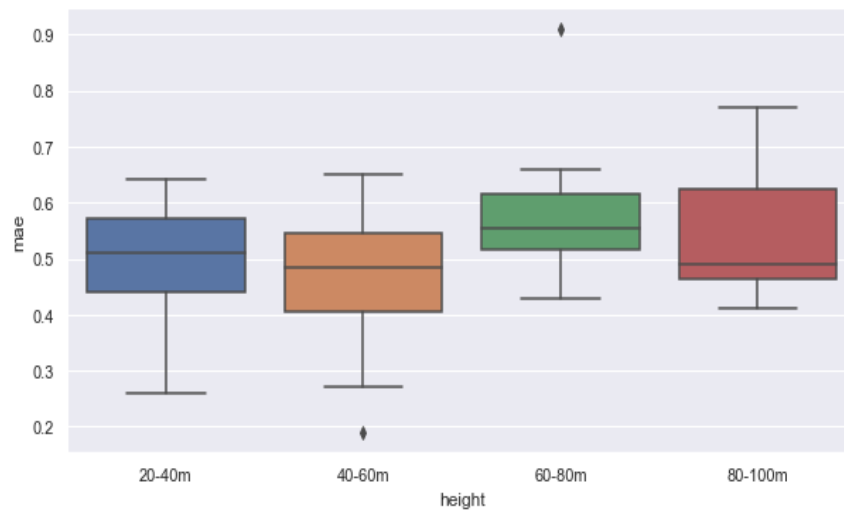


Figure 16. Errors for all methods by height

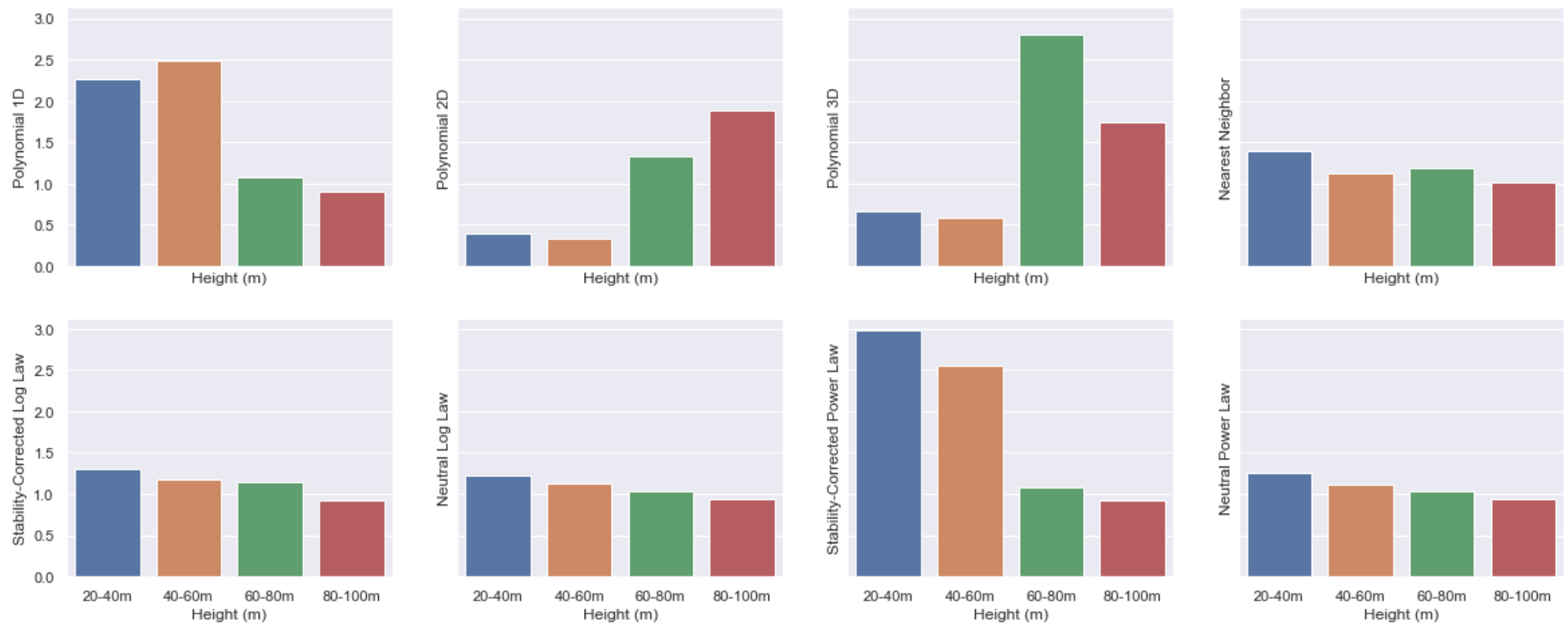


Figure 17. Overall MAE for each height grouping and method

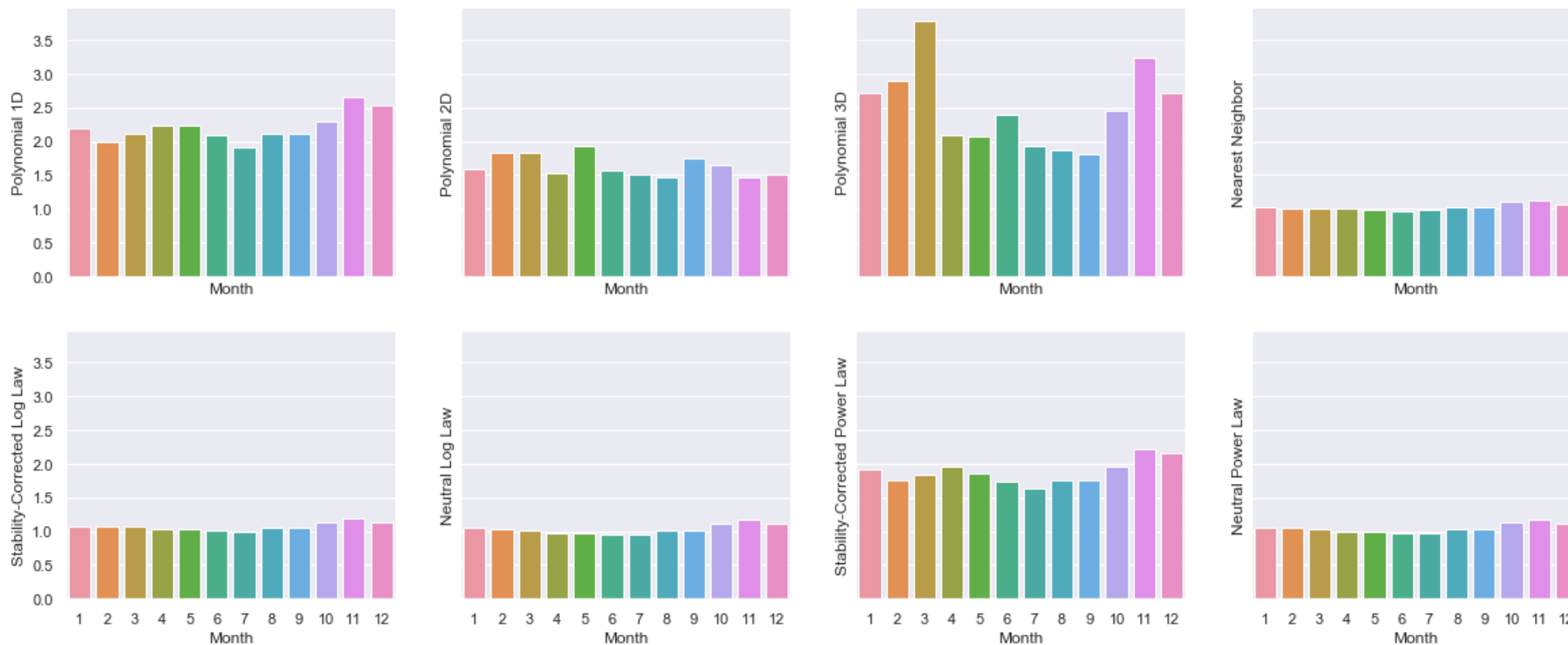


Figure 18. Monthly MAE for all sites and heights by method

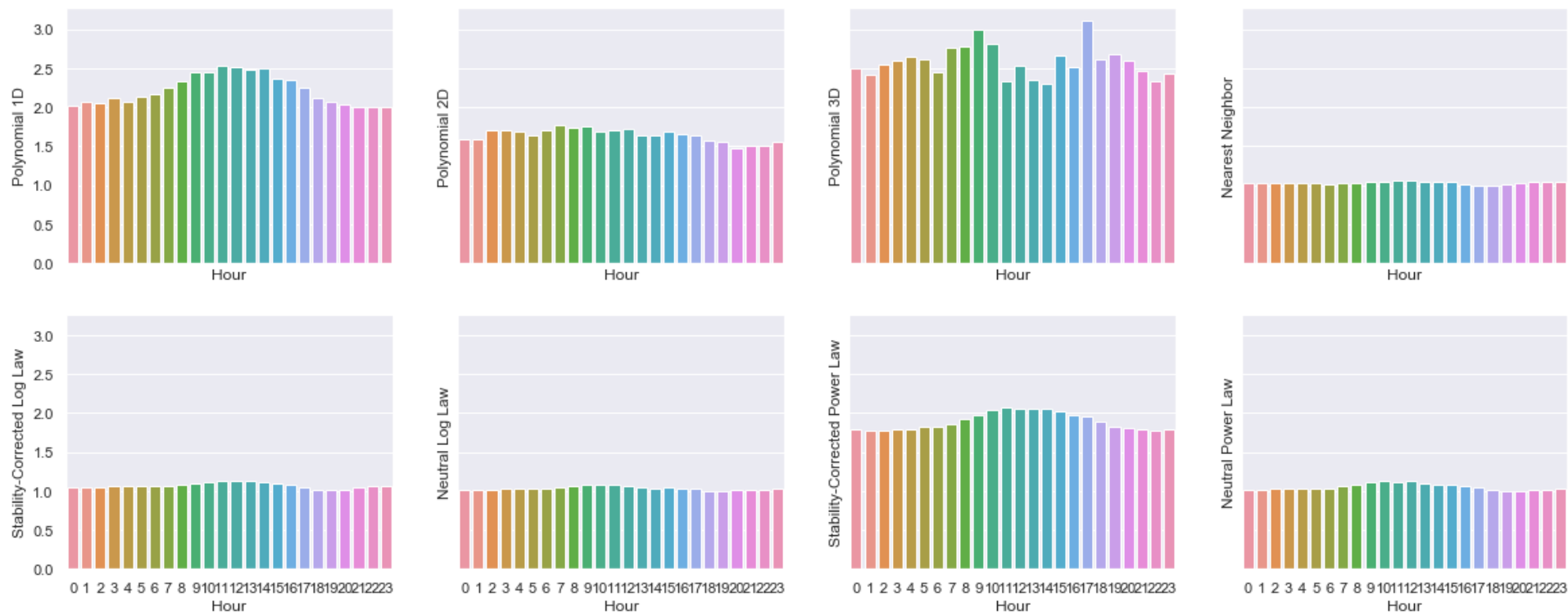


Figure 19. Hourly MAE for all sites and heights by method

To understand how the error may vary as a function of the vertical distance interpolated over, we group estimates according to this distance and analyze the resulting error statistics. We also partition the error by those estimates where an extrapolation is required (i.e., target height above or below all other heights). Figure 20 and Figure 21 show these results. According to these results, there does appear to be lower error overall with a smaller interpolation difference, and extrapolation (versus interpolation) might produce slightly higher errors on average. Combining results from all methods, we see an average MAE of 0.65 for interpolated points and 0.79 for extrapolated points. RMSE shows a similar difference: 0.92 for interpolated points and 1.1 for extrapolated points. The best performing method for extrapolated points is the neutral log law (MAE: 0.56, RMSE: 0.75), and the stability-corrected power law performs best for interpolated points (MAE: 0.48, RMSE: 0.72).

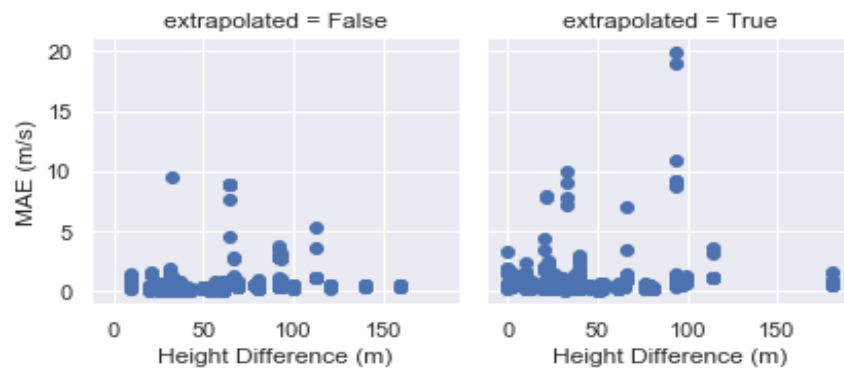


Figure 20. Error according to height difference and extrapolated/interpolated
(all models are shown).

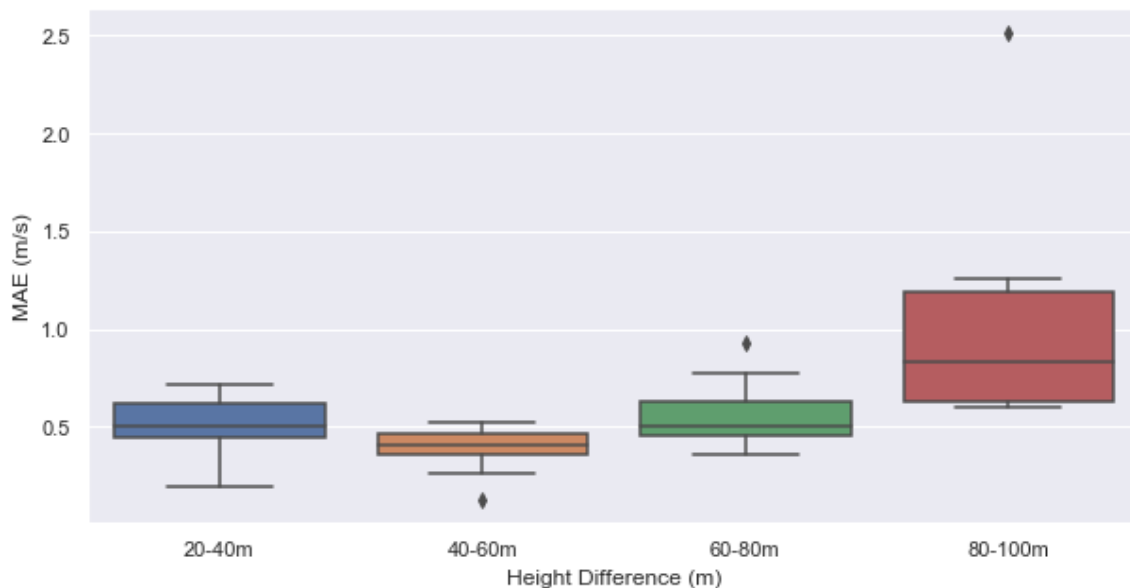


Figure 21. Error relationship between height difference between nearest input data height and target height

Because many of the methods we compare are similar mathematically, a final consideration we seek to address is whether there is a significant difference in the predictions they make. In other words, are the differences in predictions meaningful or due to random chance? To address this question, we perform a statistical analysis where 1,000 paired predictions are taken from all methods. A pairwise t-test is used to determine whether the mean difference is significantly different from zero. To account for potential false significance from multiple comparisons, we consider a difference significant if the p-value given by the t-test is smaller than $0.05/21 = 0.0024$, where 0.05 allows for 5% type 1 error and 21 is the number of comparisons. Figure 22 shows the result of this analysis. Overall, the Nearest Neighbor method is most different from the other methods. There appears to be significant similarity between the polynomial methods and stability-corrected methods. The neutral log law and neutral power law show some differences from other methods as well.

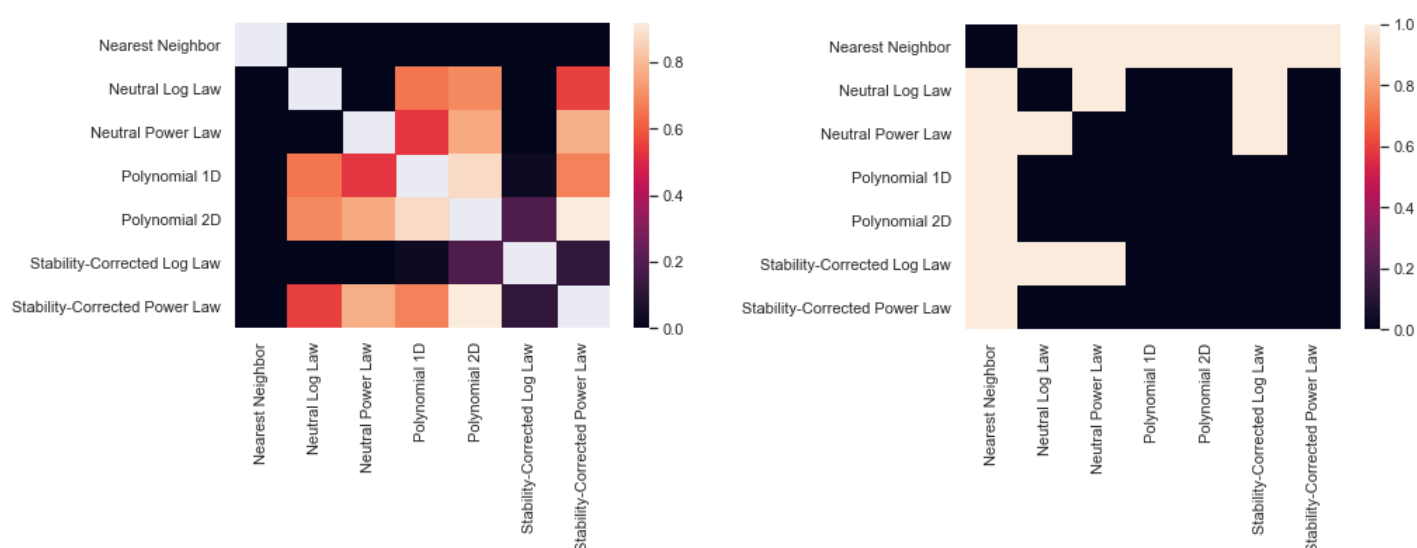


Figure 22. Degree of difference between method predictions

Left: The plot gives the p-value for each comparison.

Right: The plot highlights just those for which the p-value is less than the threshold set for significance.

4.3 Learning from the WTK-WRM and WRM-WRM Results

In this section, we discuss what we can learn from combining the insights drawn from WTK-WRM and WRM-WRM analyses.

Though the result from WTK-WRM and WRM-WRM analyses are not exactly mathematically comparable, we can get a sense of the contribution of the vertical interpolation error by calculating the MAE ratio of certain locations and heights from the WRM-WRM analysis and comparing those to errors for the same sites in the WTK-WRM analysis.

We perform an analysis of MAE ratios for 154 overlapping locations and heights. The calculated MAE ratios yield fractions between zero and one, as well as fractions above one. The range formed by these ratios is 0.02–6.68. Of 154 locations and heights, 145 exhibited MAE ratios below 1.0, while 9 exhibited ratios above one.

We suspect either that ratios above one represent bad readings from instrumentation or that a reliance on heights “far” apart from each other is required in the observational data. It is important to remember that for each location and height, the WTK can potentially use any one of nine different heights and the observational data varies from one through three different heights for most locations. In other words, WTK relies less on interpolation and the observational data relies more on interpolation.

The results for the ratios below 1.0 are shown in Table 5.

Table 5. MAE Ratios for Locations and Heights Below 1.0

Min	2%
Median	16%
Mean	28%
75th Percentile	28%

The violin plot in Figure 23 shows the distributions of the MAE ratios for all 145 locations and heights below a ratio of 1.0 broken by interpolation technique. Most methods show the bulk of ratios below 0.2. The Nearest Neighbor technique shows the highest median ratio of all techniques possibly because a reliance on one height is more likely to not be representative in the observational only data versus WTK.

Overall, we see that around 16%–28% of composite error in the WTK may be attributed to vertical interpolation depending on the interpolation method being used.

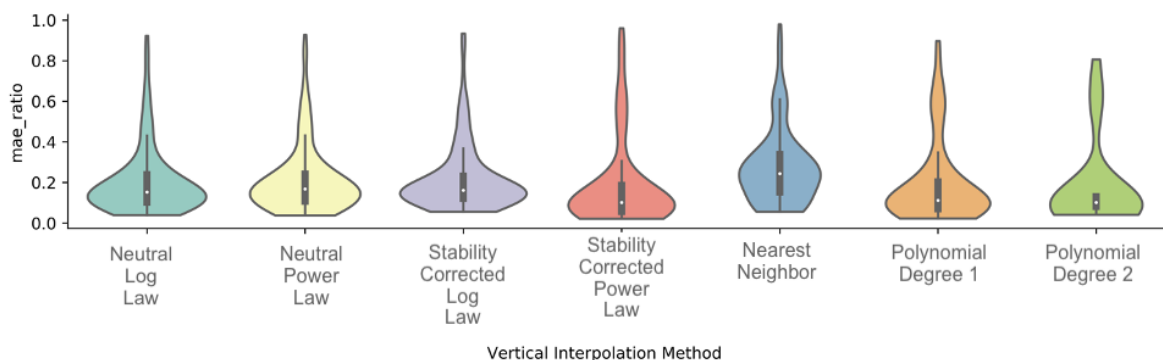


Figure 23. Violin plot comparing fraction of error due to vertical interpolation alone compared to combined error in extrapolating from WTK to WRMDB data

4.4 Evaluating Combinations of Horizontal and Vertical Interpolation Techniques

In this section, we describe how we analyze the horizontal and vertical dimensions simultaneously, using various choices for horizontal and vertical interpolation techniques. We explore this space via the following five analysis steps:

1. We assess four horizontal techniques, where all methods but Nearest Neighbor use only four neighboring points (considering these variants to be establishing the baselines for

others) and five vertical interpolation schemes (all these schemes were described in detail in Section 1). In this combination space, we show how these methods compare with each other with respect to the aggregate mean bias.

2. We expand the combination space by considering all 10 horizontal variants (where all 4-, 8-, and 16-point variants are considered), and we characterize the error distributions using the mean absolute error (MAE) metric. This investigation gives us a sense that IDW with 16 interpolation points yields the lowest errors when paired with Polynomial Degree 1 for vertical interpolation.
3. We study this best-performing combination further and look at how the errors are distributed spatially with respect to the individual sites in WRMDB.
4. We look at the per-site best methods and their combinations.
5. We compare how the estimation error responds to the method selection, for vertical and horizontal methods, isolating the two types of methods as best as we can using the available data.

4.4.1 Characterizing Mean Bias

We obtain wind speed estimates using different combinations of interpolation techniques and then perform the WTK-WRM error analysis as described in Section 4.1. Figure 24 shows the mean bias estimates we obtain for different combinations of techniques. Considering that this bias characterizes (estimated minus observed) wind speed errors, the positive cases, which represent the majority here, correspond to the overestimation, which we have encountered already in the context of the WTK-WRM results. The left plot of Figure 24 shows the set of results where we hold the vertical interpolation constant by using Nearest Neighbor. This set shows the lowest errors: medians for the studied four horizontal techniques fall within the range [0.50, 0.55]. The right plot shows the case with the highest errors; there, we use the Stability Corrected Power Law for vertical interpolation. In this set, the individual error medians are a bit larger: they are in the range [0.60, 0.62]. Numerically this difference may appear small, yet such difference showing up at this aggregate level—where we group the estimates for *all site and height combinations for all years* represented in our data set—indicates that at the instance-by-instance level we are likely to see the differences between the methods that are quite large (we present additional insights into the large method differences in Section 4.4.5). It is also worth noting here that based on these error measures, deciding which horizontal technique performs best is difficult, as the shown boxplots are nearly identical within each plot. We attempt to distill the differences between the methods in the following analysis that considers MAE.

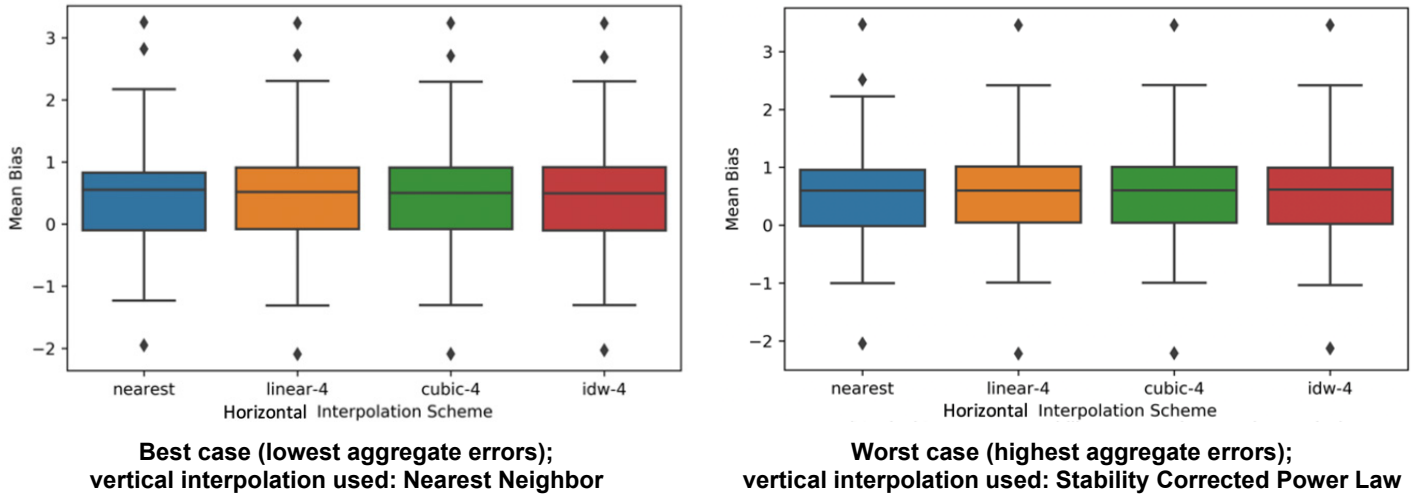


Figure 24. Comparison of mean bias estimates

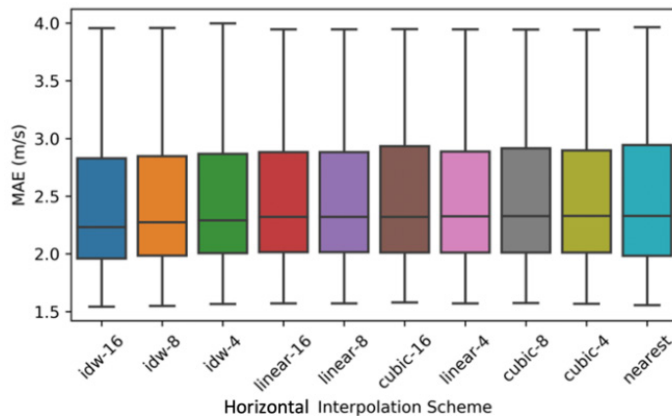
“-4” in the labels indicate the methods use four horizontal points for interpolation.

4.4.2 Comparing MAE

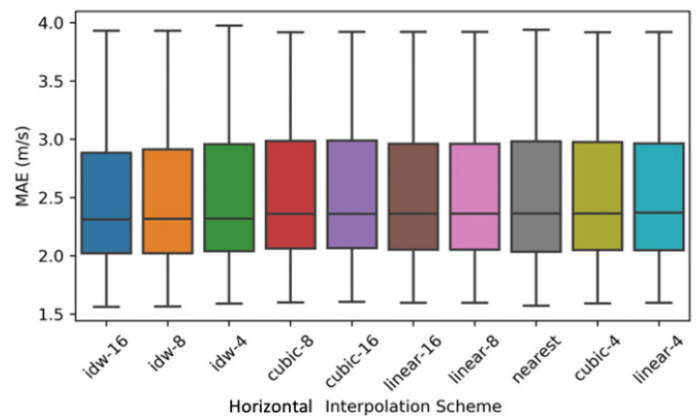
Next, we use MAE for the WTK-WRM error analysis instead of the mean bias. The former does not allow us to see the direction of the error (i.e., we cannot distinguish the overestimation and underestimation errors), whereas the latter can in principle “hide” the errors when positive and negative errors cancel each other out. Based on this, neither one method is perfect; therefore, we try to learn from the insights revealed by each of these error metrics.

With the expanded set of combinations of interpolation methods, we can observe subtle differences between the method choices, as shown in MAE plots in Figure 25 (page 32). Each of these plots on the x-axis has the horizontal methods ordered by the median error: from the lowest on the left to the largest on the right for each plot. The left plot shows Polynomial Degree 1: the best performing vertical interpolation method in the studied set with respect to this metric; in this set, the medians for different horizontal methods are in the range [2.23, 2.33]. Stability Corrected Power Law (not shown in these plots) is the second-best in this comparison, with the median range being very close to the best case: [2.24, 2.33]. In the worst case, shown in the right plot, the Neutral Log Law method yields the median errors in the range [2.31, 2.37].

What is noticeable in the plots in Figure 25 is that IDW outperforms other horizontal interpolation methods. Its variants with 16, 8, and 4 (in this order) can be considered the best choices based on these results, for both the best and the worst-case scenarios with respect to the vertical method selection. In contrast, Nearest Neighbor is near the end of the list in the shown sets. The Linear Interpolation and Cubic Linear variants are interleaved in the middle of these sorted sets.



**Best case (lowest aggregate errors);
vertical interpolation used: Polynomial Degree 1**



**Worst case (highest aggregate errors);
vertical interpolation used: Neutral Log Law**

Figure 25. Comparison of MAE estimates

Based on these results, we conclude that one specific combination of techniques—IDW with 16 points for horizontal and Polynomial Degree 1 for vertical interpolation—outperforms the rest of the combinations based on aggregate MAE. Table 6 shows how both mean bias and MAE estimates for this combination compare against other error estimates. We further analyze this combination by looking at the distribution of its errors, uncovering spatial trends discussed next in Section 4.4.3.

Table 6. Global Interpolation Error Measures for Estimates that Use IDW with 16 Points as Horizontal Interpolation and Various Techniques for Vertical Interpolation

Vertical Interpolation Technique	Mean Error (m/s)	MAE (m/s)	RMSE (m/s)
Polynomial Degree 1	0.437	2.481	3.340
Nearest Neighbor	0.405	2.499	3.362
Neutral Log Law	0.509	2.518	3.377
Stability Corrected Power Law	0.481	2.495	3.355
Neutral Power Law	0.457	2.501	3.363

4.4.3 Dissecting Spatial Patterns

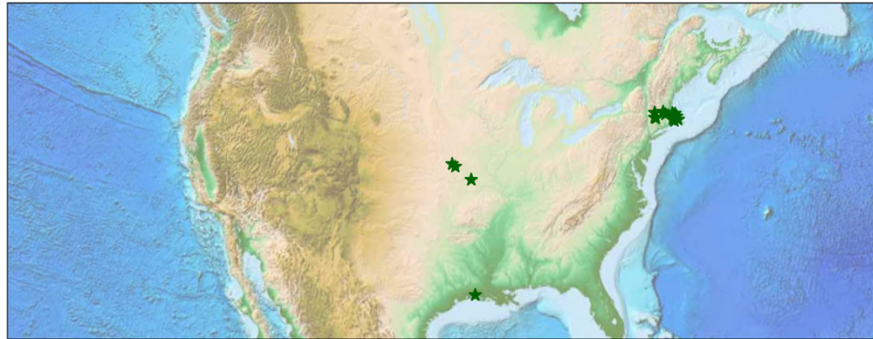
We study the best-performing method combinations further and analyze spatial patterns, as discussed in this section.

We dissect the results we obtained for the combination with *IDW with 16 points for horizontal and Polynomial Degree 1 for vertical interpolation* into three sets:

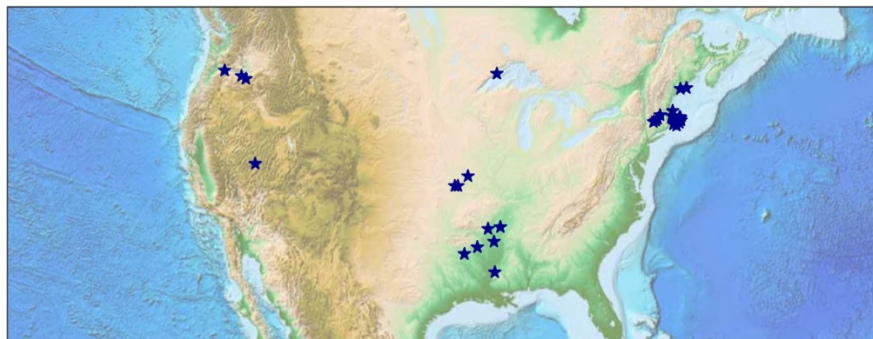
- **Small Errors:** errors in the first quartile of this distribution, with MAE of [1.54, 1.96]
- **Medium Errors:** errors in the second and third quartiles, with MAE of [1.96, 2.82]
- **Large Errors:** errors in the fourth quartile, with MAE of [2.82, 3.96].

Visualizations depicting the WRMDB sites corresponding to these error sets are shown in Figure 26.

Sites with **Small Errors**
(i.e., errors in the first
quartile; MAE < 1.96)
16 sites



Sites with **Moderate Errors**
(i.e., errors in the second
and third quartiles; MAE of
[1.96, 2.82])
31 sites



Sites with **Large Errors**
(i.e., errors in the fourth
quartile; MAE > 2.82)
16 sites



**Figure 26. Maps depicting WRMDB sites based on WTK-WRM errors:
from small (top) to moderate (middle) and large (bottom)**

We can draw several conclusions from these maps. The sites in coastal Massachusetts fall into all three sets, but primarily into the sets with small and moderate errors. It remains to be investigated why several of those sites are associated with larger errors even though they are closer to the rest of them. One way to investigate this further (which falls outside the scope of the current work) would be to study the raw time series for these sites and see whether these large errors are of a transient or more permanent nature and compare the time series with the ones for nearby sites with smaller errors. Another insight we gain here is that the sites in the Midwest and in the South can be estimated quite well using WTK: the errors that we observe there are small to moderate. As expected, the sites in the Rocky Mountain West, which are in areas of complex terrain, show errors that are larger than those of most other sites.

Evaluation of the terrain-aware spatial interpolation methods will be a part of future work. However, from the analysis we presented here, we can get initial insights into where (i.e., for which particular sites) applying the more sophisticated, terrain-aware techniques is most needed. Thus, the 16 sites showing large errors, most of which are located in the areas of complex, mountainous terrain in the Western region, as we can see in Figure 26, would be among the immediate candidates.

4.4.4 Selecting Best Combinations

If we take an approach different from “one method fits all” and allow each site to optimize its own choice of horizontal and vertical interpolations independently of other sites, we obtain the following results.

Table 7, Table 8, and Table 9 provide the details for the optimal choices: exact methods and their combinations that are at the top of the lists of the most selected choices based on per-site MAE as well as how many times each option was selected, in absolute and relative terms. Thus, for about 70% of the studied sites, IDW with 16 points is the best choice of the horizontal interpolation (which matches what we observed based on the aggregate MAE comparison earlier); it is followed by Nearest Neighbor and Cubic Interpolation with 16 points, but these two options are best in fewer cases (combined, around 17%). With respect to the vertical interpolation, our results suggest Nearest Neighbor and Polynomial Degree 1 are equally good (43% of cases each); we have seen the evidence of good performance for the latter in the WTK-WRM study (Section 4.1), and these results match that summary. Stability Corrected Power Law is the third choice (best in only 8% of cases). Finally, in the ordered list of combinations of the techniques (Table 9), we see nothing unexpected considering all the findings already mentioned: IDW with 16 points contributes to all three of the best three combinations, and the vertical interpolation list is: Polynomial Degree 1, Nearest Neighbor, and Stability Corrected Power Law (i.e., the same list as the top three vertical interpolation methods selected independently of the horizontal methods, as shown in Table 8).

Table 7. Top Three Methods in the Per-Site Selection of Horizontal Interpolation Method

(selected irrespective of the vertical method; pooled across all heights for each site)

Method	Number of Times Selected (%)
IDW with 16 points	44 out of 63 (70%)
Nearest Neighbor	7 out of 63 (11%)
Cubic Interpolation with 16 points	4 out of 63 (6%)

Table 8. Top Three Methods in the Per-Site Selection of Vertical Interpolation Method

(selected irrespective of the horizontal method; pooled across all heights for each site)

Method	Number of Times Selected (%)
Nearest Neighbor	27 out of 63 (43%)
Polynomial Degree 1	27 out of 63 (43%)
Stability Corrected Power Law	5 out of 63 (8%)

Table 9. Top Three Combinations of Interpolation Methods

(pooled across all heights for each site)

Method	Number of Times Selected (%)
Horizontal: IDW with 16 points Vertical: Polynomial Degree 1	18 out of 63 (29%)
Horizontal: IDW with 16 points Vertical: Nearest Neighbor	17 out of 63 (27%)
Horizontal: IDW with 16 points Vertical: Stability Corrected Power Law	5 out of 63 (8%)

One other dimension involved in the selection of the best methods is the amount of data used in the horizontal interpolation methods (specifically, for Linear Interpolation, Cubic Interpolation, and IDW). We experimented with the variants that use 4, 8, and 16 nearby points, and now we can reflect on these choices through the results of this per-site optimization process. Similar to the how we ranked the methods in Table 7, Table 8, and Table 9, we can rank 4-, 8-, and 16-point variants based on how many times they were selected as the best interpolation techniques (based on the lowest MAE estimates within the selected set of techniques). Of 56 cases (63 sites total minus 7 cases where Nearest Neighbor was selected, which does not have such variants), 16-point schemes were selected 50 times (89%), 8-point schemes 3 times (5%), and 4-point schemes 3 times (5%). These results speak to exploring how much further we can improve the quality of interpolation by adding more points and at what point the added benefit becomes negligible, which we can study in future work.

4.4.5 Conducting Sensitivity Analysis

In this section, we compare how the estimation error responds to the method selection based on the proposed per-site method sensitivity analysis.

The results of the WRM-WRM study (Section 4.2) allowed us to estimate the impact of the selection of the vertical interpolation method on the errors. That analysis was made possible by the presence of sites in WRMDB with sufficient data for multiple measurement heights. In the horizontal analysis, the equivalent would be if we had multiple towers in the same area, where the distances between the towers were small. They should be small enough to help us see the fine-resolution wind patterns that potentially impact the distributed wind siting process. Thus, in this context, we are trying to go beyond the 2-km resolution that is in WTK and explore finer-scale wind phenomena. As we examine the sites in WRMDB, we notice that the largest number of sites is in Massachusetts (which makes it a good candidate set for additional horizontal analysis); however, the distances between sites are not particularly small. The minimum distance between two sites in Massachusetts is nearly 6 km, the median in the set is 41 km, the average is 44 km, and the maximum is 94 km. With this set of measurements for the sites that are not close to each other, we are unable to extrapolate data from some sites to other sites with sufficient confidence. Therefore, we cannot draw the same type of conclusion we drew in Section 4.4.3 (about 16%–28% of composite error being attributed to the vertical interpolation).

Because the spatial coverage of the measurements in WRMDB is limiting the extent to which we can continue our horizontal error analysis, we propose using the available data to assess the per-

site (instead of global) relative impact of vertical and horizontal method selections. Specifically, we start with the method combinations we found to be the best-performing for each site (as discussed in Section 4.4.4), and then we vary for each site: (a) only the vertical method being used (and keep the horizontal from the best combination) and alternatively (b) vary only the horizontal (keeping the vertical from the best combination).

We calculate the differences between the best-combination MAE errors and the MAE values for the method combinations in sets (a) and (b) and express them in relative terms; for instance, a change of 10% corresponds to the MAE increase of 10% from the best combination (e.g., IDW with 16 points combined with Polynomial Degree 1) to the method combination being considered (e.g., IDW with 16 points combined with Stability Corrected Power Law). Then, we compare these differences in sets (a) and (b), which characterize how the distribution of errors across all sites is *sensitive* to the method selection along *each of these two dimensions*—vertical and horizontal. The distributions of these differences are shown in Figure 27. As we can see, the compared distributions are roughly the same, with the minimums near zero and the maximums around 30%–35%. In these long-tailed distributions, the median is 1.50 for set (b) (varying horizontal) and 0.86 for set (a) (varying vertical).

Solely based on this result, we can say that the *selection of the horizontal method has a larger effect on the errors* than the selection of the vertical method *within the studied sets of methods*. The comparison of the 75th percentiles in these distributions backs this outcome. However, the 95th percentiles, which characterize the tails for the distributions, and in other words, large “mistakes” caused by the selection of bad method combinations tell a different story: the inadequate selection of the vertical method yields larger error increases than the bad horizontal method choices, in the extreme cases. Overall, it appears that the *selection of the horizontal method has a comparable or larger effect on the errors than the vertical method selection in most cases* for the studied data and the selected methods.

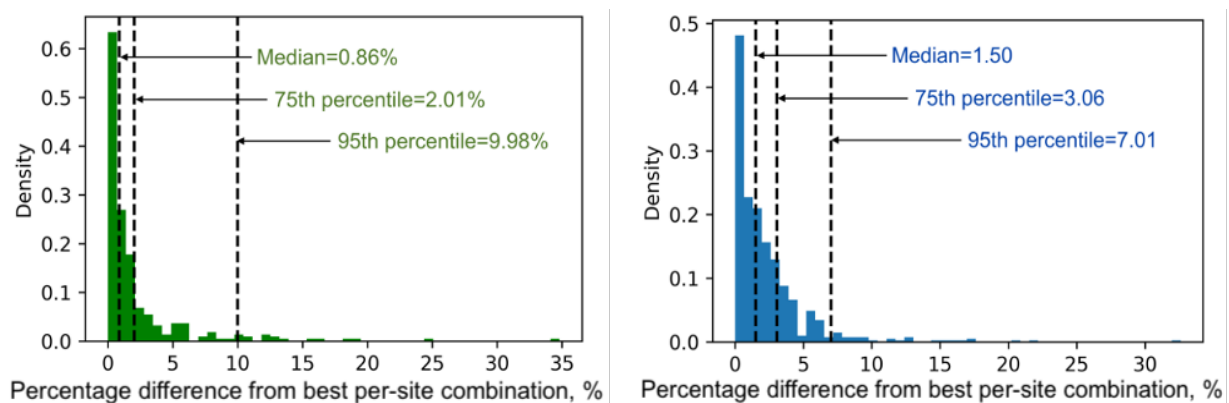


Figure 27. Error variability caused by method selection

Left: varying vertical interpolation method (horizontal minus fixed)
Right: varying horizontal interpolation method (vertical minus fixed)

5 Conclusions

The vertical and horizontal interpolation analyses in the current study were designed to optimize for, and choose, interpolation methods, as well as to assess the error inherent in the time series transformations.

The first two analyses we conducted show that though some vertical interpolation methods perform significantly better than others for the WRMDB data set (the WRM-WRM analysis), the differences are less apparent when used as a component of a methodology to predict observational data at a given point from WTK data (the WTK-WRM analysis). Nevertheless, the WRM-WRM analysis does show that more complex methods that utilize terrain and stability information generally perform best, when tested in isolation. The Nearest Neighbor method also performs remarkably well, given its simplicity. Based on this result, we recommend use of the Neutral Log or Power Law methods in distributed wind siting applications, when terrain and stability data are available, and the Nearest Neighbor method when they are not. This recommendation holds true if the horizontal interpolation is held constant and we rely on the Nearest Neighbor horizontal approximation. We remove this constraint when we perform the broader search for the best method combinations in one of the subsequent studies.

In our last analysis, we evaluated 50 method combinations: 10 variants of horizontal interpolation and 5 selected vertical interpolation methods. Based on the WTK-WRM error analysis (which we used to evaluate vertical interpolation with fixed horizontal), we conclude that all studied horizontal methods perform similarly. At the same time, the variant *t* implements Inverse Distance Weighting with 16 interpolation points can be considered the best method globally, if it is selected across all site and height combinations in this comparison. When paired with Polynomial Degree 1 for vertical interpolation, it creates the method combination with the lowest error globally, as judged by MAE estimates. Moreover, when we compare the impact of the vertical interpolation method selection with the selection of the horizontal method, we determine that the horizontal impact on the error is *comparable to, or larger than*, the vertical impact for the studied sets of method.

Our analysis allows us to determine the lower bounds for the WTK-WRM validation errors: the smallest MAE is around 2.4, the smallest RMSE is at 3.3, and the smallest mean error is approximately 0.4 m/s. These estimates are obtained for the best performing method combinations, and the other combinations yield the errors that are larger.

All these conclusions are necessarily limited by the data set being analyzed, and much care should be taken when extending them outside this context. Further work is needed, for instance, to understand the relationship between height and interpolation accuracy, particularly at the lowest and highest measurement heights, as the WRMDB is skewed toward sites with measurement data between 20 m and 60 m. For more comprehensive horizontal interpolation analysis with additional insights drawn for both global and site-specific error measures, a data set like WRMDB that is used for validation needs to include more measurements with high spatial density.

References

- Brutsaert, Wilfried. 1982. "The Surface Roughness Parameterization." *Evaporation into the Atmosphere* (Springer) 112-127.
- Dewitz, John. 2019. *National Land Cover Database (NLCD) 2016 Products: U.S. Geological Survey data release*, <https://doi.org/10.5066/P96HHBIE>.
- GIS Geography. 2020. *Inverse Distance Weighting (IDW) Interpolation*. 12 14. Accessed 12 22, 2020. <https://gisgeography.com/inverse-distance-weighting-idw-interpolation/>.
- Holtstag, M. C., W.A.A.M. Bierbooms, and G.J.W. Van Bussel. 2014. "Estimating Atmospheric Stability from Observations and Correcting Wind Shear Models Accordingly." *Journal of Physics: Conference Series* (IOP Publishing) 555 (1): 012052.
- Hsu, S.A., Eric A. Meindl, and David B. Gilhousen. 1994. "Determining the Power-Law Wind-Profile Exponent under Near-Neutral Stability Conditions at Sea." *Journal of Applied Meteorology* 33: 757-765.
- National Center for Atmospheric Research. 2020. *Weather Research and Forecasting (WRF) Model*. Accessed 12 22, 2020. <https://www.mmm.ucar.edu/weather-research-and-forecasting-model>.
- National Renewable Energy Laboratory. 2020a. *REsource eXtraction (rex) Tool*. Accessed 12 22, 2020. <https://nrel.github.io/rex/>.
- . 2020b. *Wind Intergration National Dataset Toolkit*. Accessed 12 22, 2020. <https://www.nrel.gov/grid/wind-toolkit.html>.
- Newsom, Rob K., Lindsay M. Sheridan, Brian J. Gaudet, Gabriel Garcia Medina, Zhaoqing Yang, Raghavendra Krishnamurthy, and William J. Shaw. 2020. *A Study on Modeled Wind Speed Errors Using the US Department of Energy Buoys*. Richland, Washington: Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830.
- NREL. 2020. "Wind Resource Meteorological Database (WRMDB)." Personal Communication.
- Olauson, John, and Mikael Bergkvist. 2015. "Modelling the Swedish Wind Power Production Using MERRA Reanalysis Data." *Renewable Energy* (76): 717-725.
- Phillips, Caleb, Caroline Draxl, John Readey, and Jordan Perr-Sauer. 2018. *Power from Wind: Open Data on AWS*. Amazon Web Services: Big Data. 03 20. <https://aws.amazon.com/blogs/big-data/power-from-wind-open-data-on-aws/>.
- SciPy.org. 2020a. *scipy.interpolated.griddata*. Accessed 12 22, 2020. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.griddata.html>.
- . 2020b. *scipy.spatial.cKDTree*. Accessed 12 22, 2020. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.cKDTree.html>.
- spatialreference.org. 2020. *SR-ORG:7480 USA Contiguous Albers Equal Area Conic USGS version*. Accessed 12 22, 2020. https://spatialreference.org/ref/sr-org/usa_contiguous_albers_equal_area_conic_usgs_version-2/.
- Stengel, Karen, Andrew Glaws, Dylan Hettinger, and Ryan N. King. 2020. "Adversarial Super-Resolution of Climatological Wind and Solar Data." *Proceedings of the National Academy of Sciences* 117 (no. 29 (2020)): 16805-16815.
- Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl. 2012. "An overview of CMIP5 and the experiment design." *Bulletin of the American Meteorological Society* 93 (4): 485-498.
- Walls, Liz. 2020. *oneenergy-software / Continuum*. Accessed 12 22, 2020. <https://github.com/oneenergy-software/Continuum>.