Short-term wind forecasting using statistical models with a fully observable wind flow

J. Perr-Sauer, C. Tripp, M. Optis and J. King

National Renewable Energy Laboratory, 15013 Denver West Parkway Golden, CO 80401-3305

E-mail: Jordan.Perr-Sauer@nrel.gov

Abstract. The utility of model output data from the Weather Research and Forecasting mesoscale model is explored for very short-term forecasting (5-30 minutes horizon) of wind speed to be used in large scale simulations of an autonomous electric power grid. Using this synthetic data for the development and evaluation of short-term forecasting algorithms offer many unique advantages over observational data, such as the ability to observe the full wind flow field in the surrounding region. Several short-term forecasting algorithms are implemented and evaluated using the synthetic data at several different time horizons and for three different geographic locations. Comparison is made with observational data from one location. We find that short-term forecasts of the synthetic data considering wind flow from the surrounding region perform 26% better than persistence in terms of root mean square error at the 5-minute time horizon. This improvement is comparable to studies of observational data in the literature. These results provide motivation to use synthetic data for short term forecasting in grid simulations, and open the door to future algorithmic improvements.

1. Introduction

Accurate short-term wind speed forecasts are of significant value to the wind energy industry. These forecasts provide critical information that grid operators need to make economically important decisions, such as price and set point, at subhourly intervals [1, 2]. In recent years, the penetration level of variable, intermittent, and distributed energy resources, such as wind power plants, has increased on the electric grid. At the same time, smart controllers capable of regulating these resources at increasingly short timescales are reaching the marketplace.

The potential value of very short-term wind speed forecasts (5 to 30-minute horizon) when used by an electric grid replete with these autonomous, controllable, and distributed energy resources is not well understood. This problem is being studied through extensive computer simulations at the National Renewable Energy Laboratory (NREL) as part of the Autonomous Energy Systems project [3]. These simulations use atmospheric model output data from the NREL Wind Integration National Dataset (WIND) Toolkit [4] in-lieu of real atmospheric data. We call the model output data "synthetic data" in this context. To understand how these simulations will react to forecasting errors, short-term forecasts of wind speed from the WIND toolkit are being developed.

Using WIND toolkit model output data as synthetic input data for forecasting has many advantages over the use of observational data. In this paper, we explore these benefits and report on preliminary analysis of the performance of short term forecasting on the synthetic data. To do this, we apply well-known statistical models to the synthetic data at three locations

and compare their performance at multiple time horizons. Substantial performance improvement over persistence is shown by using regional wind data, with the magnitude of this improvement matching that demonstrated in the literature in some cases.

For an overview of the expected benefits and further motivation for using synthetic data, see Section 2. A short review of the literature is provided in Section 3. The WIND Toolkit data are described further in Section 4. The forecasting methods chosen for analysis are described in Section 5. As a case study in Section 6, we apply these methods to make short-term wind speed predictions from the synthetic data at the National Wind Technology Center (NWTC), in Boulder, CO. Evidence of accuracy improvement is shown. We compare the results of this case study to observational data, highlighting a limitation of the synthetic data. Section 7 provides a comparison of results between the NWTC and two other locations: Columbia River Gorge (CRG), and Brady, Texas (BDY). Section 8 concludes.

2. Potential Advantages Of Synthetic Wind Resource Data

Synthetic data from simulations of wind flow offer many advantages over observational data. Model output is, generally, fully observable; that is, any atmospheric variable can be extracted from any location within the simulation domain, at any time, with no measurement uncertainty. This allows for quick and cost-effective experimentation of sensor layout and sensor type. Observational data sets, in contrast, contain measurements only at specific locations that are determined by the sensor configuration. Additionally, synthetic data sets can be generated for large regions, and they can be created for new regions without expensive measurement campaigns. This provides almost limitless data across varied geography on which to train and test data-driven models.

Synthetic data sets have very few data quality issues, are easily accessible, and are available under permissive licenses, making them easy to use in analysis. Simulated data avoid many practical difficulties and sources of error present in real-world data used in forecasting studies while eliminating measurement uncertainty and opening new opportunities for analysis. This contrasts with observational data sets, which are often inaccessible because of proprietary restrictions and packaged in varying formats, sometimes lacking necessary metadata [2].

At the same time, the development of new remote sensing technologies continues to provide analysts with a more full and accurate picture of the wind flow field. As scanning lidar technology matures and becomes widely commercially available, it is increasingly feasible to use these high-resolution, high-rate wind speed measurements for forecasting. Modern lidar systems are capable of sensing wind speed at distances of several kilometers and at resolutions of fractions of a degree with low accumulation times and high scanning rates [2, 5, 6]; however, not many mature short-term prediction models exist to harness the predictive power of emerging lidar technologies for wind energy generation applications, especially in complex terrain. Work in the remote sensing community focuses on reconstructing the wind flow field from these measurements. Using synthetic data allows us to bypass this step and analyze the performance of statistical forecasting algorithms that are assumed to have access to this data. Along with the other advantages, synthetic data sets could provide an important avenue for developing, evaluating, and designing remote-sensing-based short-term wind forecasting systems.

3. Related Work

Statistical methods tend to dominate the literature on short-term forecasting. Auto-regression, artificial neural networks, and other machine learning techniques are commonly seen [7–9]. Persistence, which assumes that the wind speed will remain constant between measurements, has been identified as a well-performing benchmark that is increasingly difficult to beat at shorter prediction horizons. The use of persistence is justified at the 5 to 30-minute timescales studied in this work, and we use it here as a benchmark.

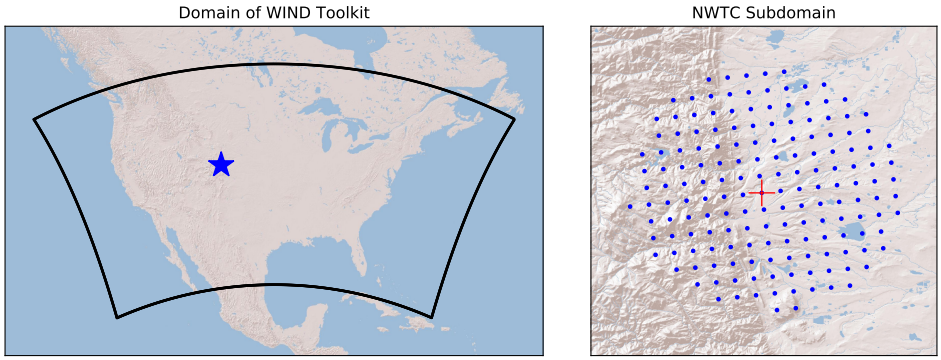

Figure 1. Left: Domain of the NREL WIND Toolkit with the location of the NWTC subdomain marked by a blue star. Right: The NWTC subdomain with each data point marked by a blue circle. The target site is marked by a red cross.

The spatiotemporal correlations of wind speed are well known and have been exploited to improve wind speed forecasting. In a study of three meteorological stations in the Aegean Sea, it was found that data from upwind locations could be leveraged by a dense neural network to improve their predictions by up to 19% over persistence at the 1 to 15-minute time horizon [10]. Another study of a wind power plant on the coast of Tasmania yielded similar results [11].

Much cutting-edge research on very short-term wind power forecasting has been oriented around new types of data made available by new sensors, such as scanning lidar. In one study, a prediction model that assumes linear advection of the incoming wind field in a coastal setting is able to achieve a 21%-38% performance improvement over a persistence model at the 5-minute time horizon [6]. Another recent study highlighted how forward-facing lidar sensors mounted on the nacelle of a wind turbine can capture spatiotemporal correlations, showing promise in the domains of wind ramp and gust detection [12].

Institutional efforts have been made to elucidate the field of wind power forecasting and to codify the best practices therein [13, 14]. A workshop was recently hosted by the International Energy Agency Task 32 and Task 36 to discuss the intersection between remote sensing and short-term forecasting [2]. One outcome of the workshop was a review paper that called for better, more standardized data sets and comparison methodologies among studies. The authors identified a lack of standardized data sets and forecast evaluation methodologies as a major issue facing the field.

A handful of studies have used synthetic Numerical Weather Prediction data to evaluate a forecasting algorithm. In [15], a mixed convolutional and recurrent neural network (which the author calls PDCNN) extracted information from the spatiotemporal correlations in NREL's WIND Toolkit data for one wind power plant in Indiana. PDCNN realized a 35% performance improvement on these data at the 5-minute time horizon. In [16], a 3-D convolutional neural network was trained to extract features from 30-minute resolution data from the National Center for Atmospheric Research's Weather Research and Forecasting (WRF) mesoscale model. Thus, there is precedent to use synthetic data for wind speed forecasting.

4. The NREL WIND Toolkit

The synthetic data set used in this paper is NREL's WIND Toolkit [4]. Based on the WRF mesoscale model [17], the WIND Toolkit provides meteorological data at a 2-km spatial resolution and up to 5-minute time resolution from the years 2007 through 2013. The domain, covering the continental United States, is shown in Figure 1. These attributes make the WIND

Toolkit the largest publicly available wind data to date for use in resource assessment, grid integration studies, and economic modeling [18].

The WIND Toolkit was developed using WRF Version 3.4.1 and nested 54-km, 18-km, 6 km, and 2-km domains. The model was initialized and forced at the boundaries with the European Centre for Medium-Range Weather Forecasts Interim Reanalysis. The WIND Toolkit used GTOPO30 data from the U.S. Geological Survey for terrain, roughness, and soil properties. Key physics parameterizations included the Noah land surface model, the Yonsei University boundary-layer scheme, and topographic wind enhancement [4].

A range of interactive online tools and download options have been made available for accessing and using WIND Toolkit data. [1] These options include techno-economic data at 120,000 points in the United States accessible through a website called the Wind Prospector as well as access to gridded time-series data through Amazon's Open Data Initiative.

5. Methods

5.1. Data Preparation

We extract regional data from the WIND Toolkit at three locations of interest for the years 2012–2013. For each location, we extract 80-m wind speed and wind direction at a 5-minute resolution from within a circular subdomain with 15-km radius using the haversine approximation. This radius was chosen so that weather fronts advecting at a relatively high wind speed would be present in the domain at the time of forecast. The coordinates for the NWTC, CRG, and BDY target sites are: (-105.23976, 39.90652), (45.63109, -120.64358), and (31.17370, -99.32876), respectively. Each subdomain contains 166, 163, and 177 measurement locations, which are located on a 2 km grid. We refer to an arbitrary location in a subdomain as x, and the location closest to the center of each subdomain as x_0, or the target site. The locations from the NWTC subdomain are plotted on a topographical map of the region and shown in Figure 1.

We consider forecast horizons of 5, 10, and 30 minutes. To achieve results for different time horizons, we first resample the source data to the desired forecast horizon using the mean. Predictions can then be made for one time step in the future. We refer to this quantity as Δ, or the timescale.

5.2. Baseline: Persistence Forecast

Persistence asserts that wind speed will remain constant at the forecast horizon. That is, $\hat{v}_{x_0,t+\Delta} = v_{x_0,t}$, where v is wind speed from the data set, $\hat{v}$ is a predicted wind speed, and t is an arbitrary time in the domain. Persistence is a surprisingly effective method of wind speed forecasting at the subhourly time scale, and its accuracy generally increases as the time horizon decreases. It is a commonly used baseline algorithm in the literature.

5.3. Second-Order Persistence Forecast

This method augments persistence by linearly extrapolating the change in wind speed between the previous two time steps. That is: $\hat{v}_{x_0,t+\Delta} = v_{x_0,t} + (v_{x_0,t} - v_{x_0,t-\Delta})$, following the same notation as the persistence method. Second-order persistence works well when the data are smooth and follow a clear linear trend.

5.4. Ridge Regression with Multiple Inputs

Ridge regression is an example of a linear regression model, and therefore it can be interpreted through inspection of its coefficients. It employs a regularization term in its objective function that is useful when strong colinearity is suspected between input variables, as is the case for spatially correlated wind field data. We applied ridge regression from the scikit-learn library

[1] https://www.nrel.gov/grid/wind-toolkit.html

[19]. Input data are first normalized by removing the mean and scaled by the standard deviation. We have:

$$\hat{v}_{x_0,t+\Delta} = v_{x_0,t} + \mu_{\Delta x_0} + \sigma_{\Delta x_0} \sum_{n=1}^{N} \frac{b_n}{\sigma_{\Delta x_n}} (v_{x_n,t} - v_{x_n,t-\Delta} - \mu_{\Delta x_n})$$

where N is the total number of locations, and $\mu_{\Delta x_n}$ and $\sigma_{\Delta x_n}$ are the mean and standard deviation of the differenced data $(v_{x_n,t} - v_{x_n,t-\Delta})$. Coefficients b_n were found through the minimization of the ridge regression objective function, $||y - Xb||_2^2 + \alpha * ||b||_2^2$, where X are the input data as described above, and α is the regularization parameter. This objective function uses the l^2-norm, thus penalizing large weights. We chose $\alpha = 10^4$ through a manual tuning process. The sensitivity of our results to this parameter was quite low, with a sensitivity of only 1% across all tested values.

5.5. Ridge Regression with Regime Switching

Based on some of our early, encouraging results with ridge regression, we implemented a variant that we call regime-switching ridge regression. In this model, the input data are separated into bins by the wind direction at the target site x_0. Separate ridge regression models, as described above, are trained for each wind direction bin. We found nine bins to be optimal through manual tuning. The sensitivity of our results to this parameter was quite high, with a spread of up to 10% for different bin sizes.

5.6. Multilayer Perceptron with Multiple Inputs

The multilayer perceptron (MLP) is a densely connected feed-forward neural network with one or more hidden layers. Gradient descent via back-propagation is used to train the model's weights. The MLP is an example of a nonlinear machine learning model. As such, we hoped this model would capture nonlinear spatial correlations in the wind field that ridge regression could not capture. We included it because of the success and proliferation of artificial neural networks in the literature on short-term wind speed forecasting. We used the MLP implementation from the scikit-learn library [19], using the same once-differenced and scaled training data as in the ridge regression model. Hyper-parameters were found manually by random sampling from a reasonable range, choosing a width and depth of 4, an activation function of $\tan(x)$, and learning rate $\alpha = 1$. We considered adding a regime switching MLP model, but chose not to do so since the ridge regression produces similar results and has more interpretable learned parameters.

5.7. Error Metric

We use the root mean square error (RMSE) as the error metric. RMSE is a commonly used error metric in the wind forecasting literature. It has the same units as the wind speed, ms^{-1}, and penalizes larger errors more severely than smaller ones. With T as the number of time steps, we have:

$$\mathrm{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\hat{v}_{x_0,t+\Delta} - v_{x_0,t+\Delta})}$$

6. Case Study at the National Wind Technology Center

We perform 24-fold cross validation using two years of WIND Toolkit data to assess the performance of each forecasting algorithm at each of the three geographic locations. For each fold, 23 months of data are placed in the training set and 1 month is left out for the test set. In this section, we report the results from the NWTC. This is a region of complex terrain that

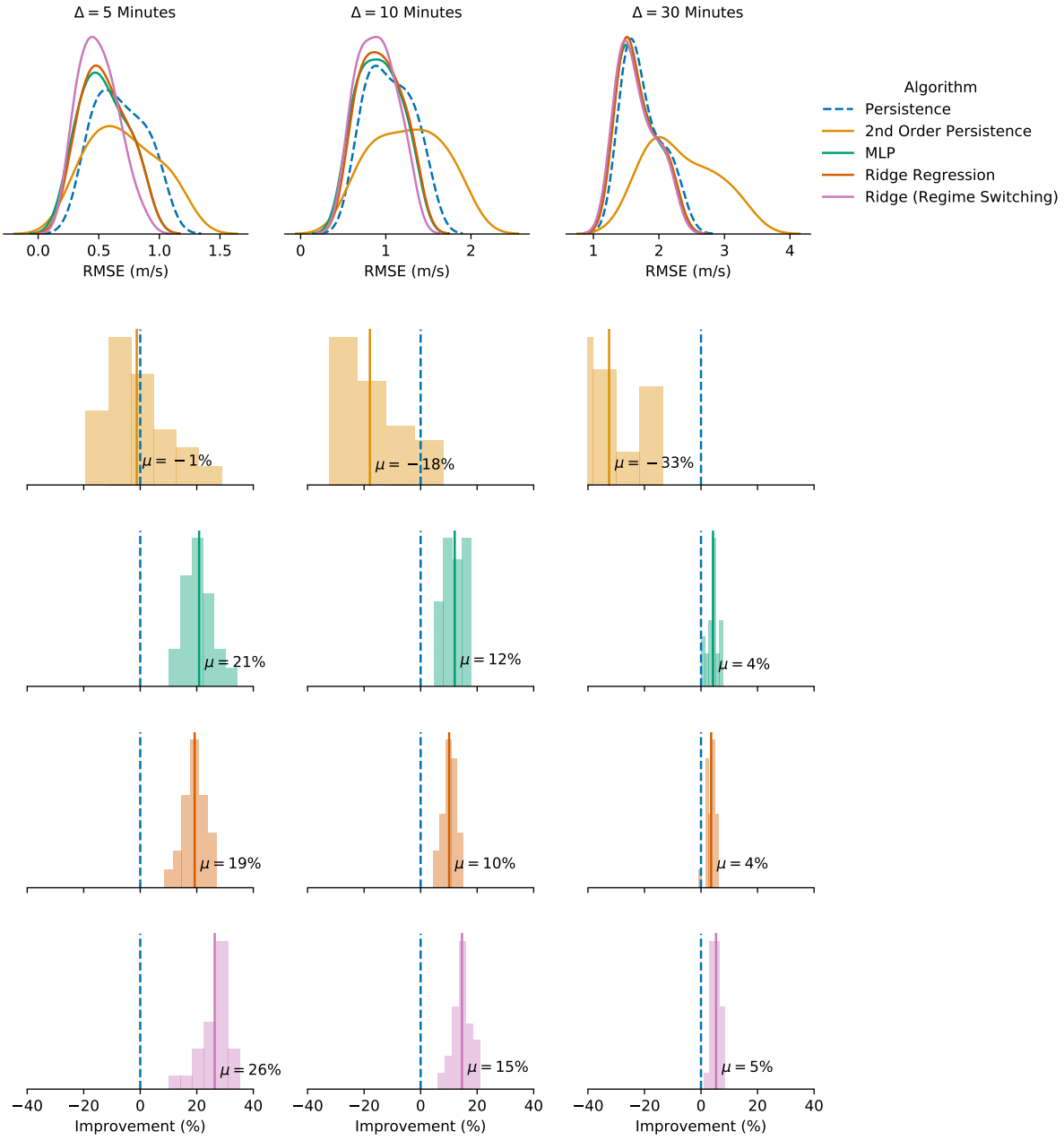

Figure 2. The results of running five different forecasting algorithms on the synthetic dataset for the NWTC region using 24-fold cross validation. The top row shows the distribution in RMSE for each forecasting algorithm and for three different timescales. The bottom four rows show the distribution of forecast improvement (with respect to persistence) for each algorithm, colored by the same legend. Columns show the timescale.

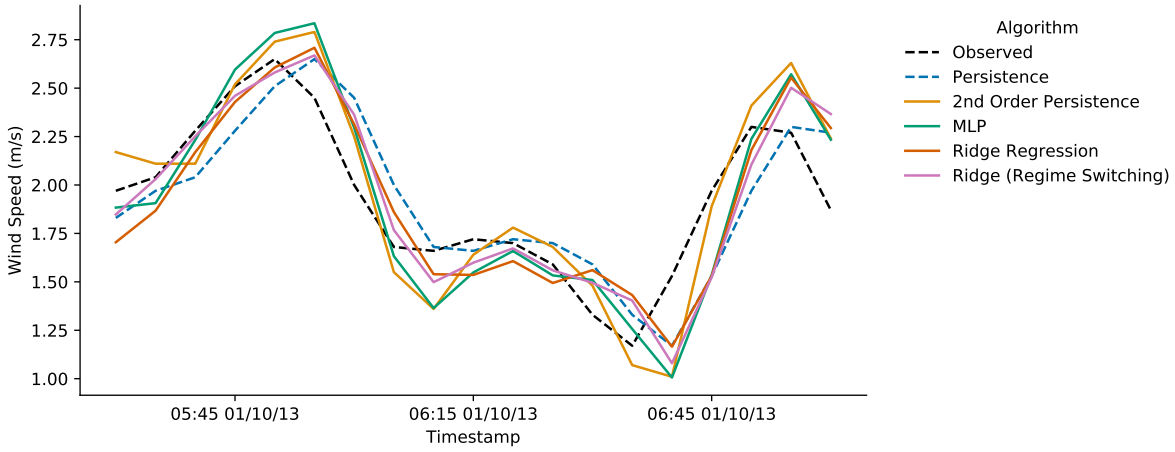

Figure 3. A snapshot of the forecasts produced by each method for a period of 1.5 hours on January 10$^{\text{th}}$, 2013. Persistence lags the observed data by the timescale (in this case, 5 minutes). The other forecasting algorithms somewhat correct for the lag, with MLP and ridge regression matching the observed data the closest.

sits adjacent to the foothills of the Rocky Mountains and features a meteorological tower with a cup anemometer from which single point observational data could be extracted.

Both ridge regression and MLP show improvement over persistence on the synthetic data. The only algorithm to perform worse than persistence was second-order persistence, which does not consider regional wind field data. The topmost row of histograms in Figure 2 shows the distribution in RMSE scores for each method over the timescales $\Delta \in \{5, 10, 30\}$ minutes. We can see that ridge regression and MLP outperform persistence on average. The improvement, however, is dwarfed by the variance of the RMSE scores. This can be seen by the relatively large width of the histograms. The mean RMSE for all methods also tends to increase with the timescale. For example, 1 ms^{-1} is a relatively high RMSE for the 5-minute timescale, but it is one of the lowest for the 30-minute timescale.

To better understand the quality of this forecast improvement, we show another set of histograms that report the total improvement percentage (with respect to persistence) for each fold in the cross validation. From the bottom three rows of Figure 2, it is clear that the ridge regression and MLP methods always outperform persistence. All algorithms show smaller percentage improvements as the timescale increases. This is partially, but not entirely, because of the increasing magnitude of RMSEs with the timescale. Another reason for this decline could be the choice of 15-km radius spatial domain, which may occlude some important regional weather data at the higher timescales. Additionally, this decline could be due to the effects of simplifying assumptions made in the WRF mesoscale model. We follow up on this in Section 6.2.

The best algorithm for the 5-minute timescale is regime-switching ridge regression. We see an average improvement of 26% in terms of RMSE over persistence at this timescale. Regime-switching ridge regression is able to maintain this advantage for all timescales tested, providing evidence that wind direction is an important feature. It is notable that ridge regression and MLP perform similarly, suggesting that the nonlinear MLP might not be extracting qualitatively more information from the input data than the linear method. A more thorough analysis of hyper-parameter choices and the nature of the source data are needed to support this conclusion.

The line graph in Figure 3 shows the actual predictions made at the 5-minute timescale

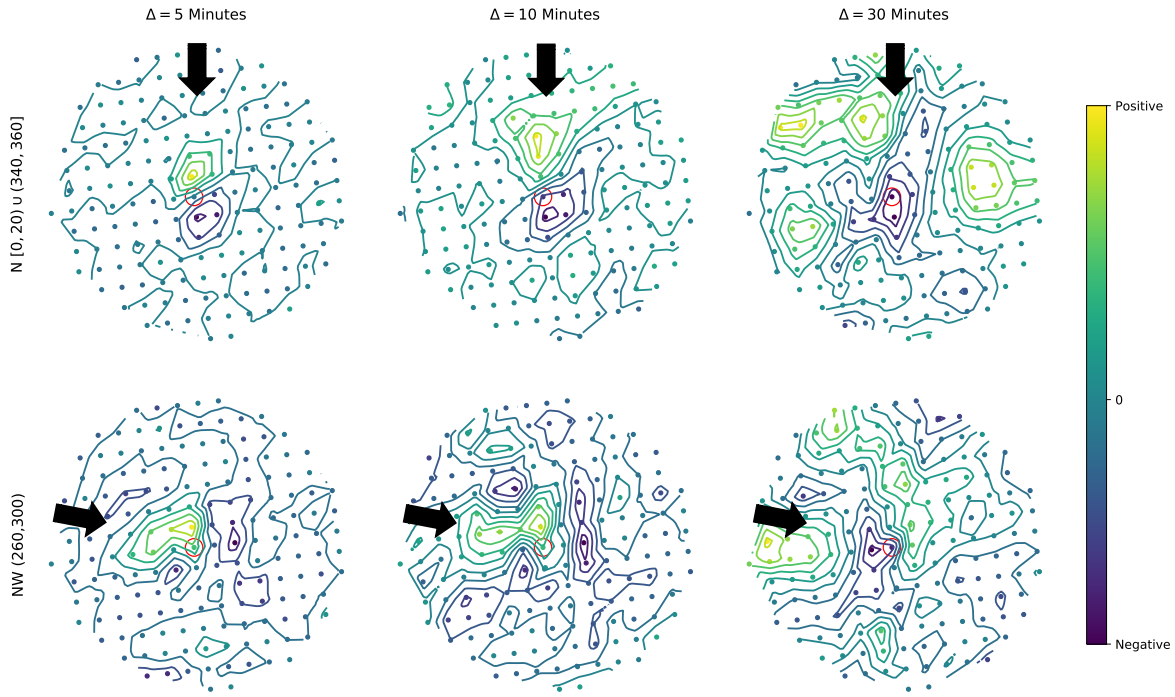

Figure 4. Coefficients for regime-switching ridge regression at the 5-, 10-, and 30-minute timescales, trained with two wind direction bins of 40° width. The center of each bin is annotated by a black arrow. The diagram shows that the model is most strongly weighting locations upwind of the target location and that the strongly weighted locations move farther from the target site and spread out as the timescale is increased.

for each algorithm during a particular 1.5 hour period. Persistence is easy to pick out for its consistent lag by the timescale. Qualitatively, the test methods are able to make predictions that do not consistently lag the observed data, more closely matching the observed wind speed. The models used for this figure were trained using only 2012 data, and evaluated using 2013 data. This train test split is used in all subsequent figures, and the RMSE values from these models are similar to those found in the 24-fold cross validation.

6.1. Visualization of Spatial Correlations using a Regime-Switching Method

One advantage of using a linear model is that the model can be interpreted through inspection of its coefficients. Figure 4 shows the magnitudes of the ridge regression coefficients b_n over the NWTC subdomain. Here, we use the regime-switching ridge regression, as described in Section 5, but we use only two manually specified wind direction bins, one for each dominant wind direction. These wind directions are $340 - 20°$ (N) and $260 - 300°$ (NW). This model was trained on all data from 2012 and tested on all data from 2013. Additionally, the regularization parameter was set to $\alpha = 1000$ to exaggerate the differences between coefficients.

This figure provides a snapshot of the spatial correlations in the data found at each timescale. We see that the closest upwind data points are most relevant to the prediction at the 5-minute timescale. As we increase the timescale to 10 and 30 minutes, the relevant physical locations move farther from the target site and are spread out over a larger number of physical locations. This matches the fluid dynamics principals of advection and diffusion seen in wind flows. Together, these results show that the most important physical locations to the ridge

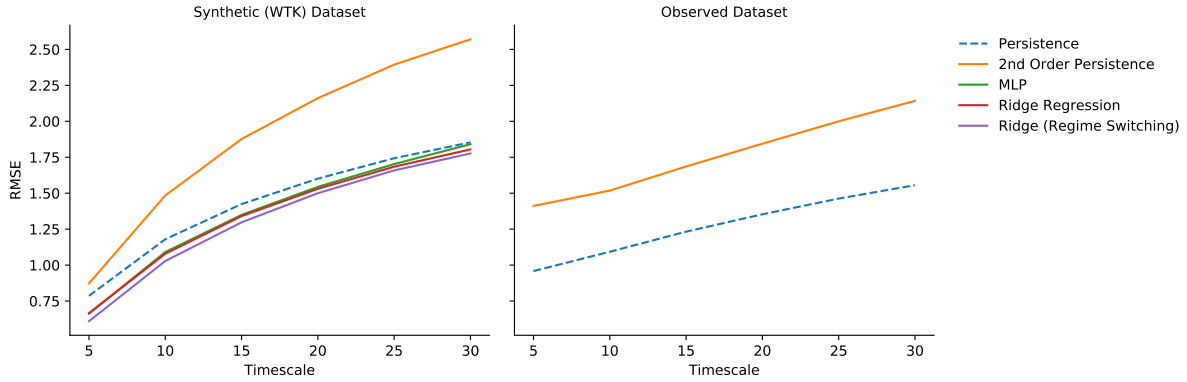

Figure 5. Prediction error (in terms of RMSE) for each algorithm considered in this study for several different timescales for the year 2013 at the NWTC. The leftmost graph shows this result on the Synthetic dataset, and rightmost graph shows the result for observational data from the same location.

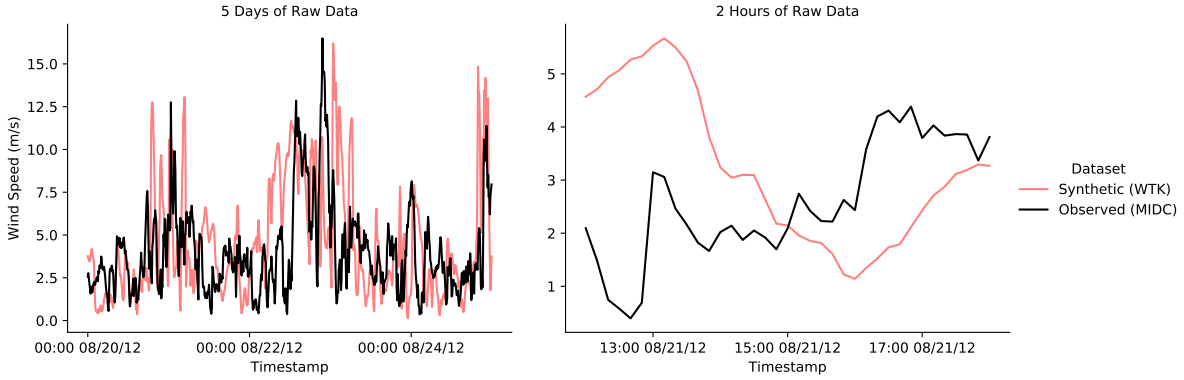

Figure 6. Raw time-series wind speed data from the WIND Toolkit (light red) against the real wind speed as measured by an anemometer sensor mounted on a meteorological tower (black) at the NWTC. The WIND Toolkit accurately captures variability at the hourly level, but it struggles to capture variability at the subhourly level.

regression model are those that are upwind some distance proportional to the wind speed at the target site. Directly modeling this advection would likely capture a significant portion of the improvement seen in this model.

6.2. Validation with Observed Data

The results in the previous section show that synthetic data from the WIND Toolkit can be used by short-term wind speed forecasting algorithms. We have not yet made any claims about how these algorithms perform on observational data, or if the performance is comparable. We begin to make that comparison by showing how persistence and second-order persistence perform on both the synthetic WIND Toolkit and real observational data from an anemometer at the NWTC campus. These observational data are publicly available from the Measurement and Instrumentation Data Center (MIDC) [20]. The data are provided at a 1-minute temporal resolution and time-averaged to match the forecast horizon.

Figure 5 presents a direct comparison of the the average RMSE for all methods on the

Table 1. Improvement over Persistence at the 5-minute Horizon for Three Locations.

Location	Second-Order Persistence	MLP	Ridge Regression	Ridge (Regime Switching)
BDY	-15% (15.50)	11% (5.36)	12% (4.79)	**19%** (6.82)
CRG	10% (12.74)	26% (5.20)	31% (7.88)	**35%** (5.68)
NWTC	-1% (11.75)	21% (5.22)	19% (4.18)	**26%** (5.65)

Mean percentage improvement (in terms of RMSE) over persistence for each forecasting method at the 5-minute time horizon for each region using 24-fold cross validation. Standard deviation of the percentage improvement reported in parenthesis.

synthetic data, and for persistence and second-order persistence on the observed data. We do not report scores for all forecasting algorithms on the observed data because of a lack of regional observational data at this site. General agreement is shown between the observed and synthetic datasets in terms of the ordering of the quality of forecasts. Persistence performs better than second-order persistence on both datasets. There is also a positive trend in RMSE as the timescale increases for both datasets.

It is clear from this figure that the WIND Toolkit is accurately capturing some components of the predictability of wind speed for persistence and second-order persistence for timescales greater than 20 minutes. At shorter timescales, however, the baseline methods produce smaller RMSE than on the observed data, and the RMSE seems to drop at an accelerated rate as the timescale is further reduced. At the 5-minute timescale, persistence and second-order persistence seem to converge, but this is not shown in the observed data set. This is an interesting finding because Section 6 shows that the largest percentage improvement for most algorithms is also at the 5-minute timescale.

To better understand the source of this discrepancy, we present raw wind speed data in Figure 6. Here, we see good agreement between synthetic and observed data at large timescales, but the observed data contain more variability than the synthetic data at the subhourly timescale. Thus, the mesoscale model is not accurately capturing the 5-minute variability for this particular location, which is to be expected. The mesoscale model's relatively coarse resolution does not allow it to resolve physics at such small timescales.

7. Results from the Columbia River Gorge and Brady, TX

We repeat the 24-fold cross-validation process with five prediction algorithms for each of the three locations, as described in Section 5. Table 1 shows a summary of mean percentage improvement over persistence for each algorithm.

It is clear that the performance improvement varies by region. In particular, second-order persistence underperforms persistence at BDY but outperforms persistence at CRG, though this result was within one standard deviation of the mean. Additionally, MLP and ridge regression outperform persistence by 19%–21% at the NWTC, but they show only a 11%–12% improvement at BDY. This cannot be explained by the relative magnitudes of the RMSEs. The RMSE of persistence at the NWTC and BDY are 0.67 and 0.35, respectively. From these magnitudes, we might expect the NWTC to have a lower percentage improvement, but this is not the case. This suggests that the reduction in the performance improvement at BDY is caused by another factor, such as the difference in the complexity of the terrain.

One common trend is that all methods that consider spatiotemporal data outperform persistence. We can also see that regime-switching ridge regression, our best method at the NWTC, outperforms all other methods at all locations. Not shown in the figure is that this result is actually robust across all timescales tested. Thus, we see evidence that regime-switching ridge regression provides a substantial and robust improvement over other methods.

8. Conclusion and Future Work

We see evidence that statistical short term forecasting algorithms can produce realistic forecasting errors on synthetic data in some cases. This provides evidence that short-term wind speed forecasts made using synthetic data could be useful for large-scale simulations of the electric grid. These algorithms will be integrated into the Autonomous Energy System grid simulation, where we hope to gain insight into their economic value to the electrical grid.

Further, we provide evidence that synthetic data sets have the potential to be useful in developing short-term forecasting methods for highly observable wind flow, such as those made available by lidar. We have shown that a subset of the 5-minute resolution synthetic wind data from the NREL WIND Toolkit, located over the NWTC campus and two other regions, can be exploited by multivariate statistical algorithms to produce an increase in forecast accuracy. Our best method, regime-switching ridge regression, achieves an average improvement in forecasting accuracy of 26% over persistence at the 5-minute timescale, at one location. This result is consistent with other studies that exploit similar spatiotemporal correlation in observational data [10, 11]. Additionally, we find evidence that locations upwind of the target site with distance proportional to the forecast timescale have the highest predictive power in the regime-switching ridge regression model.

We identified one potential drawback of using WIND Toolkit data for short-term forecasting. The timescale for which the prediction improvement results are strongest (5 minutes) is precisely the timescale for which the synthetic data look least like the observed data. This exposes a primary limitation of using simulated wind resource data to develop short-term forecasting methods. Simulated datasets are only as realistic as the assumptions and simplifications of their generating simulation. These assumptions and simplifications can have widely varying effects on the realism and utility of synthetic data. Common simplifications (often made to speed up the computation time) are simplified or nonexistent turbulence models, simplified handling of complex terrain, and coarse spatial and temporal resolutions. Other limitations of the current work include the small set of forecasting methods. We would like to add more and varied methods, such as physics-based and hybrid models, that consider an expanded set of atmospheric variables.

Finally, the general lack of available, standardized data sets for regional short term forecasting presents a real challenge in developing data driven algorithms, and hampered our efforts to validate those algorithms which require regional wind field data. Future work will include procuring high-fidelity simulated wind field data, such as large-eddy simulation for a particular region, as well as regional observational data from campaigns which feature advanced remote sensing devices. One region of interest is the Columbia River Gorge of Oregon, which is the location of the Weather Forecasting Improvement Project (WFIP2) study. High-fidelity synthetic data in this region with an overlapping period of record to the WFIP2 study could provide for a more complete validation.

References

[1] Wang Q, Martinez-Anido C B, Wu H, Florita A R and Hodge B M 2016 *IEEE Transactions on Sustainable Energy* **7** 1525–1537 ISSN 19493029 URL `https://doi.org/10.1109/TSTE.2016.2560628`

[2] Würth I, Valldecabres L, Simon E, Möhrlen C, Uzunoglu B, Gilbert C, Giebel G, Schlipf D and Kaifel A 2019 *Energies* **12** ISSN 19961073 URL `https://doi.org/10.3390/en12040712`

[3] Kroposki B, Dall'anese E, Bernstein A, Zhang Y and Hodge B M 2017

[4] Draxl C, Clifton A, Hodge B M and McCaa J 2015 *Applied Energy* **151** 355 – 366 ISSN 0306-2619 URL `https://doi.org/10.1016/j.apenergy.2015.03.121`

[5] Floors R, Peña A, Lea G, Vasiljević N, Simon E and Courtney M 2016 *Remote Sensing* **8** ISSN 2072-4292 URL https://doi.org/10.3390/rs8110884

[6] Valldecabres L, Peña A, Courtney M, von Bremen L and Kühn M 2018 *Wind Energy Science* **3** 313–327 URL https://doi.org/10.5194/wes-3-313-2018

[7] Soman S S, Zareipour H, Malik O and Mandal P 2010 *North American Power Symposium 2010, NAPS 2010* 1–8 URL https://doi.org/10.1109/NAPS.2010.5619586

[8] Colak I, Sagiroglu S and Yesilbudak M 2012 *Renewable Energy* **46** 241–247 ISSN 09601481 URL http://dx.doi.org/10.1016/j.renene.2012.02.015

[9] Jung J and Broadwater R P 2014 *Renewable and Sustainable Energy Reviews* **31** 762–777 ISSN 13640321 URL http://dx.doi.org/10.1016/j.rser.2013.12.054

[10] Alexiadis M C, Dokopoulos P S and Sahsamanoglou H S 1999 *IEEE Transactions on Energy Conversion* **14** 836–842 ISSN 08858969 URL https://doi.org/10.1109/60.790962

[11] Khalid M and Savkin A V 2012 *IEEE Transactions on Power Systems* **27** 579–586 ISSN 08858950 URL https://doi.org/10.1109/TPWRS.2011.2160295

[12] Simon E, Courtney M and Vasiljevic N 2018 *Wind Energy Science Discussions* **2018** 1–30 URL https://doi.org/10.5194/wes-2018-71

[13] Giebel G, Browsword R, Kariniotakis G N, Denhard M and Draxl C 2012 URL https://doi.org/10.13140/RG.2.1.2581.4485

[14] Möhrlen C, Collier C, Zack J and Lerner J 2017 *16th Wind Integration Workshop*

[15] Zhu Q, Chen J, Zhu L, Duan X and Liu Y 2018 *Energies* **11** 1–18 ISSN 19961073 URL https://doi.org/10.3390/en11040705

[16] Higashiyama K, Fujimoto Y and Hayashi Y 2018 *Energy Procedia* **155** 350–358 ISSN 18766102 URL https://doi.org/10.1016/j.egypro.2018.11.043

[17] Skamarock W, Klemp J, Dudhia J, Gill D, Barker D, Duda M, Huang X Y, Wang W and Powers J 2008 A description of the advanced research wrf version 3. NCAR Technical Notes (National Center for Environmental Research)

[18] Waite M and Modi V 2016 *Applied Energy* **183** 299 – 317 ISSN 0306-2619 URL https://doi.org/10.1016/j.apenergy.2016.08.078

[19] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M and Duchesnay E 2011 *Journal of Machine Learning Research* **12** 2825–2830 URL http://www.jmlr.org/papers/v12/pedregosa11a.html

[20] Jager D Andreas A 1996 URL http://dx.doi.org/10.5439/1052222

Acknowledgments
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 Autonomous Energy Systems project funded by the National Renewable Energy Laboratory's Laboratory Directed Research and Development program. 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.