Princeton University

CEE 588: Boundary Layer Meteorology

Characterization of an Autoregressive Model for Forecasting Wind Velocity

Joseph G. Tylka

josephgt@princeton.edu

Submitted: May 19th, 2018

Abstract

An autoregressive model for forecasting wind velocity is implemented and characterized. Autoregression is a well-established method for short-term (intraday) time-series forecasting of wind velocity, but the performance of the model depends on the temporal characteristics of the winds. Given a time-series dataset of past wind velocities, an autocorrelation is computed and used to determine relevant timescales over which past wind velocities might be useful predictors of future data. Subsequently, an autoregressive model is implemented and its performance is compared, in terms of a prediction error, to the performances of two simpler predictive models: persistence and a random sample. Additionally, the performance of the autoregressive model is characterized by examining the dependence of the incurred prediction errors on season and on time of day. Results of the timescale analysis suggest that the wind system will tend to "forget" previous wind conditions after approximately 5.5 days. Indeed, performance results show that, while the autoregressive model consistently outperforms the alternative models, it reaches a plateau in prediction error over approximately that timescale. The autoregressive model is also shown to yield significantly more accurate predictions in the Summer than in the Winter, and slightly more accurate predictions during the day than at night.

Contents

List of Figures

1 Introduction

In order to facilitate the integration of wind energy into the power grid, reliable methods of forecasting wind power are required. Existing forecasting methods generally take a statistical, physical, or combined approach [1]. For short-term (intraday) predictions, statistical approaches, such as direct time-series forecasting, are often preferred. Several such methods are reviewed by Giebel et al. [2, Chap. 2]. For longer-term predictions, physics-based numerical weather prediction (NWP) models are often preferred, or at least included in a combined approach, as such models have been shown to outperform time-series forecasts beyond about 6 hours ahead [2, Sec. 1.3].

It is well-established that modern weather forecasting models, which tend to employ a combination of statistical and physical models, can only accurately forecast about 5–7 days into the future [2, Sec. 1.2]. For time-series forecasting models such as an autoregression, whose predictions draw heavily on recent data, the upper limit on how far ahead such models can accurately predict is likely related to the timescale over which the wind system "forgets" about previous conditions. Consequently, the primary objective of the present work is to empirically determine this timescale and characterize the time-domain performance of an autoregressive model.

Brown et al. [3] first proposed using an autoregressive model to forecast wind speed and power, but found that their model only yields accurate predictions up to about 3 hours ahead. On a similar timescale, Landberg and Watson [4] found that persistence models, which model future wind conditions as identical to the most recent measurement, tend to outperform their NWP model up to about 6 hours ahead. Numerous other studies, many of which are reviewed by Giebel et al. [2, Sec. 2.1], have demonstrated that autoregressive (or, more generally, autoregressive moving-average) models can be optimized to outperform persistence over most, if not all, timescales. In view of these findings, here, we use persistence as a benchmark against which to compare the forecasting accuracy of the autoregressive model.

In order to determine the timescales over which an autoregressive model can be expected to perform well, we analyze the autocorrelation of a dataset of historical wind velocities. We then implement an autoregressive model, the parameters of which (i.e., the autoregression coefficients) are computed using these same historical wind data, and evaluate it using the same dataset but for a distinct (i.e., non-overlapping) time period. This evaluation consists of randomly selecting segments of measured wind data as an input to the model, and comparing the predictions of the model to the measured data immediately following the input segment. The performance of the autoregressive model is then compared, over various prediction timescales, to the performances of two simpler predictive models: persistence and a random sample. Additionally, the performance of the autoregressive model is further characterized by examining the dependence of the incurred prediction errors on season and on time of day.

The rest of this report is organized as follows. In Sec. 2, we review the mathematical theory used in this work as well as several forecasting models. Next, in Sec. 3, we describe the analyses conducted in order to determine relevant timescales and to characterize the performance of each forecasting model. In Sec. 4, we present and discuss the results of these analyses, and finally, in Sec. 5, we summarize this work and draw conclusions from the results.

2 Review of Mathematical Theory

In this section we first formulate the time-series forecasting problem. We then review four pairwise (i.e., "two-point") statistics functions: the cross- and autocorrelation and cross- and autocovariance functions. Finally, we review three forecasting models that we implement and evaluate in this work.

2.1 Problem Formulation

Consider the stream-wise and cross-stream wind velocities, $u(t)$ and $v(t)$, respectively, sampled at times t_n for all integers n with a sampling rate F_s, such that $t_n = n/F_s$. For simplicity, we consider $t_n < 0$ to be the "past" and $t_n \geq 0$ to be the "future." We write the wind velocity as a complex variable, $z(t) = u(t) + iv(t)$, where i is the imaginary unit, and denote the sampled wind data by $z_n = z(t_n)$. The goal of time-series forecasting of wind velocity is to use past wind data, say N_{in} previous samples, to predict N_{out} future samples. That is, we seek a set of forecasting functions f_n such that

$$z_n = f_n(z_{-1}, z_{-2}, \ldots, z_{-N_{\mathrm{in}}}), \text{ for all } n \in [0, N_{\mathrm{out}} - 1]. \quad (1)$$

2.2 Pairwise Statistics Functions

Consider two (possibly complex-valued) sequences X_n, Y_n, for all integers $n \in [0, N-1]$. An unbiased estimate of the cross-correlation of these sequences is given by[1]

$$R_{XY}(m) = \begin{cases} \dfrac{1}{N-m} \displaystyle\sum_{n=0}^{N-m-1} X_{n+m}\overline{Y_n}, & \text{for } m \geq 0, \\ \\ \overline{R_{YX}}(-m), & \text{for } m < 0, \end{cases} \tag{2}$$

for all integers $m \in [-(N-1), N-1]$, where $\overline{(\cdot)}$ denotes taking the complex conjugate of the argument. Furthermore, the autocorrelation of a single sequence, X_n, is given by $R_{XX}(m)$. Note that this definition differs from that given by Stull [5, Sec. 8.2.1], in which the mean of the sequence is removed (cf. Eq. (3) below) and a different normalization is used.

Similarly, an unbiased estimate of the cross-covariance of X_n and Y_n is given by[2]

$$C_{XY}(m) = \begin{cases} \dfrac{1}{N-m} \displaystyle\sum_{n=0}^{N-m-1} \left(X_{n+m} - \langle X \rangle\right)\left(\overline{Y_n} - \langle \overline{Y} \rangle\right), & \text{for } m \geq 0, \\ \\ \overline{C_{YX}}(-m), & \text{for } m < 0, \end{cases}, \tag{3}$$

where $\langle \cdot \rangle$ denotes taking the mean, i.e.,

$$\langle X \rangle = \frac{1}{N} \sum_{n=0}^{N-1} X_n. \tag{4}$$

Note that the cross-covariance is related to the cross-correlation by $C_{XY}(m) = R_{\delta_X \delta_Y}(m)$, where $\delta_{X_n} = X_n - \langle X \rangle$ and $\delta_{Y_n} = Y_n - \langle Y \rangle$ represent each sequence's deviation from its mean, and the autocovariance of X_n is given by $C_{XX}(m) = R_{\delta_X \delta_X}(m)$.

2.3 Forecasting Models

In this section we review three forecasting models used in this work. Here, we denote the measured wind velocity by z_n and the predicted wind velocity by z_n'.

[1] See: `https://www.mathworks.com/help/signal/ref/xcorr.html`

[2] See: `https://www.mathworks.com/help/signal/ref/xcov.html`

2.3.1 Persistence

One well-established yet inherently limited forecasting model is known as the "persistence" model [2, Sec. 1.5], wherein all future wind velocities are taken to be equal to the most recent wind velocity sample, i.e.,

$$z'_n = z_{-1}, \text{ for all } n \in [0, N_{\text{out}} - 1]. \tag{5}$$

Given this model's construction, we can expect (and indeed we will see in Sec. 4.2.1) that the prediction errors incurred by this model will be small in the very short-term, but will increase rapidly with increasing time.

2.3.2 Random Sample

Given a database of historical wind velocities, ζ_m, for all integers $m \in [0, M-1]$, we randomly select a segment of N_{out} consecutive samples as the predicted future data. That is, for some randomly selected integer $r \in [0, M - N_{\text{out}}]$, we take

$$z'_n = \zeta_{r+n}, \text{ for all } n \in [0, N_{\text{out}} - 1]. \tag{6}$$

For this model, we can expect (and again we will see in Sec. 4.2.1) that the prediction errors will be large and, on average, relatively constant with time, as the randomly selected segment is likely uncorrelated with the true future data.

2.3.3 Autoregression

An autoregressive model of order P uses a linear combination of the P most recent time samples to predict the following one. In general, this can be written as

$$z'_n = \sum_{p=1}^{P} \phi_p z_{n-p}, \tag{7}$$

where ϕ_p are the autoregression coefficients for all integers $p \in [1, P]$. In practice, this formula is applied successively to predict z'_0, then z'_1, then z'_2, and so forth. Note that, in contrast to much of the existing literature, this formulation simultaneously forecasts wind speed and direction (i.e., wind velocity), whereas often only the wind speed is considered [3, for example].

Here, we compute the autoregression coefficients using a database of historical wind data ζ_m for all integers $m \in [0, M-1]$. According to the Yule-Walker equations and provided that $P < M$, the autoregression coefficients are given by [6, Sec. 3.1.1]

$$\begin{bmatrix} C_{\zeta\zeta}(1) \\ C_{\zeta\zeta}(2) \\ \vdots \\ C_{\zeta\zeta}(P) \end{bmatrix} = \begin{bmatrix} C_{\zeta\zeta}(0) & C_{\zeta\zeta}(-1) & \cdots & C_{\zeta\zeta}(-P+1) \\ C_{\zeta\zeta}(1) & C_{\zeta\zeta}(0) & \cdots & C_{\zeta\zeta}(-P+2) \\ \vdots & \vdots & \ddots & \vdots \\ C_{\zeta\zeta}(P-1) & C_{\zeta\zeta}(P-2) & \cdots & C_{\zeta\zeta}(0) \end{bmatrix} \cdot \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_P \end{bmatrix}, \tag{8}$$

where $C_{\zeta\zeta}$ is computed using Eq. (3).

Similar to the persistence model, we can expect (and again we will see in Sec. 4.2.1) that the prediction errors incurred by this model will be small in the very short-term, but will then increase with increasing time as the wind system "forgets" about the previous wind conditions.

3 Methodology

In this work, we use a database of wind velocity measurements from the Cabauw meteorological tower,[3] which is a 213 m tall tower that has been recording, since 1986, the horizontal wind speed and direction at several vertical heights [7, Table I]. These wind speeds and directions are measured using a combination of cup and sonic anemometers and weather vanes [8, Sec. 2.1]. We perform a pairwise statistical analysis of this data to determine relevant timescales, and subsequently, we implement and characterize the three forecasting models reviewed in Sec. 2.3.

As our historical "training" dataset, ζ_m, we use the wind velocity data measured at 200 m from January 1$^{\text{st}}$, 2001, until (and including) December 31$^{\text{st}}$, 2010. These data are given at a sampling rate of $F_s = 6$ samples per hour. As a pre-processing step, we first determine a nominal "average" direction for the wind. To do this, we compute the wind direction $\theta_m = \arg(\zeta_m)$, where $\arg(\cdot)$ denotes taking the argument (i.e., phase angle) of a complex number, such that $\zeta_m = |\zeta_m| e^{i\theta_m}$. We then compute a histogram of these θ_m, with a bin width of $3°$, and subsequently find the peak of this histogram, which, for these data, occurs at $\hat{\theta} = -130.5°$. We then realign the training dataset

[3] Available for download here: `http://www.cesar-database.nl/Welcome.do`

and all other wind data used hereafter by computing

$$\hat{\zeta}_m = \zeta_m e^{-i\hat{\theta}} \qquad \text{and} \qquad \hat{z}_n = z_n e^{-i\hat{\theta}}. \tag{9}$$

3.1 Timescale Analysis

In order to determine the timescales over which an autoregressive model can be expected to perform well, we compute the autocorrelations of both U_m and V_m, where $\zeta_m = U_m + iV_m$. From these autocorrelation sequences, we define an integral timescale, T_X (where X can be replaced by either U or V), given by [9]

$$T_X = \sum_{m=0}^{M-1} R_{XX}(m)\Delta t, \tag{10}$$

where $\Delta t = 1/F_s$.[4] However, in practice, we may choose to sum only over $M_{\max}$ samples, where $M_{\max} < M$ and is chosen based on when R_{XX} reaches a sufficiently small value. Here, we choose $M_{\max}$ as the point at which R_{XX} reaches its "steady-state" value, $\tilde{R}_{XX}$.

Note that an autocorrelation sequence may have a non-zero steady-state value when, for example, the sequence X is sufficiently consistent such that its autocorrelation remains relatively constant over a wide range of m. In such cases, we use a modified form of Eq. (10), such that T_X is given by

$$T_X \approx \frac{1}{1 - \tilde{R}_{XX}} \sum_{m=0}^{M_{\max}-1} \left(R_{XX}(m) - \tilde{R}_{XX}\right)\Delta t. \tag{11}$$

Note that if $\tilde{R}_{XX} = 0$, then Eq. (11) reduces to the unmodified Eq. (10).

3.2 Performance Characterization

In order to characterize the performance of the autoregressive model, we first compare its performance to the performances of the other forecasting models reviewed in Sec. 2.3. Subsequently, we investigate the dependence of the prediction errors incurred by the autoregressive model on season and on time of day.

[4] Also see: https://en.wikipedia.org/wiki/Integral_length_scale

3.2.1 Comparison of Forecasting Models

For each of the forecasting models reviewed in Sec. 2.3, we compute an incurred prediction error. To do this, we randomly select test segments of measured wind data from an "evaluation" dataset, for which we use the Cabauw wind data from January 1$^{\text{st}}$, 2011, until (and including) December 31$^{\text{st}}$, 2017. Each test segment is $N_{\text{in}} + N_{\text{out}}$ samples long, where the first N_{in} samples are given as an input to each model, and the following N_{out} samples serve as a reference against which each model's prediction is compared. We then compute the prediction error, $\epsilon_n = z_n - z'_n$, for all integers $n \in [0, N_{\text{out}} - 1]$, where z_n and z'_n are the true (reference) and predicted wind velocities, respectively.

We perform this calculation for Q randomly selected test segments, and compute the root-mean-square (RMS) magnitude error over all test segments, given by

$$\bar{\epsilon}_n = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left|\epsilon_n^{(q)}\right|^2}, \tag{12}$$

where $\epsilon_n^{(q)}$ is the prediction error for the q^{th} test segment.

3.2.2 Dependence on Season and on Time of Day

To investigate the seasonal dependence of the prediction errors, we repeat a very similar analysis to that described above in Sec. 3.2.1, but restrict the randomly selected test segments to be from a given season. In particular, we ensure that the entire test segment is contained within the same season. Here, we use "Winter" to refer to December through February; "Spring" refers to March through May; "Summer" refers to June through August; and "Fall" refers to September through November. Similarly, to investigate the dependence on the time of day, we restrict the randomly selected test segments such that the reference wind velocity segments are no longer than 12 hours and are either "daytime" (starting at 8AM) or "nighttime" (starting at 8PM) segments.

4 Results and Discussion

In this section we present and discuss the results of the analyses described above.

4.1 Timescale Analysis

In Fig. 1, we plot the autocorrelation sequences $R_{UU}(m)$ and $R_{VV}(m)$ (for the stream-wise and cross-stream wind components, respectively), which have been normalized such that $R_{UU}(0) = R_{VV}(0) = 1$. By averaging over the widest "flat" portion of each sequence, the "steady-state" values of these sequences are found to be approximately $\tilde{R}_{UU} = 0.1$ and $\tilde{R}_{VV} = 0$. From the plot, we see that R_{UU} reaches its steady-state value after approximately 70 days, while R_{VV} reaches its steady-state value after only approximately 15 days. Choosing $M_{\max}$ as the first (i.e., smallest positive) value of m at which $R_{XX}(m) < \tilde{R}_{XX}$, we compute, using Eq. (11), the timescales T_U and T_V. For each curve in Fig. 1, the area between $\tilde{R}_{XX}$ and $R_{XX}(m)$ is shaded for $m < M_{\max}$, representing the value of the summation in Eq. (11).

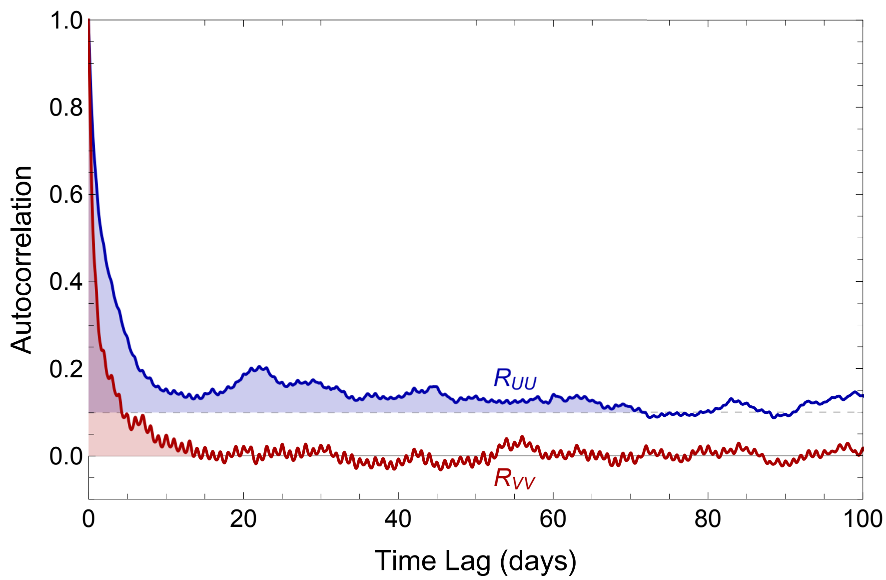

Figure 1: Normalized autocorrelation sequences R_{UU} (blue curve) and R_{VV} (red). The filled regions correspond to the summations computed in Eq. (11) to find the integral timescales T_U and T_V.

For these data, we find that the timescales computed via Eq. (11) are $T_U \approx 5.44$ days and $T_V \approx 1.71$ days. These values suggest that the stream-wise winds tend to remain correlated for ~ 5.5 days, while the cross-stream winds remain correlated for only ~ 2 days. Consequently, we can expect the errors incurred by the autoregressive and persistence models to attain relatively constant values after ~ 5.5 days, since, after that time, the wind conditions are essentially uncorrelated with the previous ones. This timescale corresponds well with the previously mentioned limitation of modern weather forecasting models, which are only able to accurately forecast about 5–7 days into the future [2, Sec. 1.2].

4.2 Performance Characterization

Here, we construct an autoregressive model, as described in Sec. 2.3.3, with an order P corresponding to 2 days, i.e., $P = 2 \times 24F_s$, where F_s is given in samples per hour. In each analysis below, we take $N_{\text{in}} = 2 \times 24F_s$ input samples and we compute, using Eq. (12), the RMS error over $Q = 100$ test segments.

4.2.1 Comparison of Forecasting Models

The RMS prediction errors incurred by each forecasting model are plotted in Fig. 2 for a forecast duration of 12 hours (i.e., $N_{\text{out}} = 12F_s$). From this plot, we see that the autoregressive model performs comparably to, if not slightly better than, persistence over almost the entire forecast duration. As expected, we see that the RMS errors incurred by the random sample model are approximately constant with time.

In Fig. 3, we again plot the RMS prediction errors incurred by each model, now for a forecast duration of 30 days (i.e., $N_{\text{out}} = 30 \times 24F_s$). From this plot, we see that, beyond about 1 day, the autoregressive model consistently outperforms persistence. Additionally, we see that the errors incurred by the persistence model are small at small t_n, but ultimately reach similar values to those incurred by the random sample model. We interpret this behavior as corresponding to the system "forgetting" about the previous wind conditions, and, in particular, we see that this happens when $t_n \approx T_U \sim 5.5$ days.

Similarly, we see that the autoregressive model reaches a plateau in RMS error beyond T_U. However, these errors are still smaller than those of the persistence and random sample models. This suggests that, even though the system has "forgotten" about the previous wind conditions, the autoregressive model is still able to achieve predictions that are, on average, slightly more

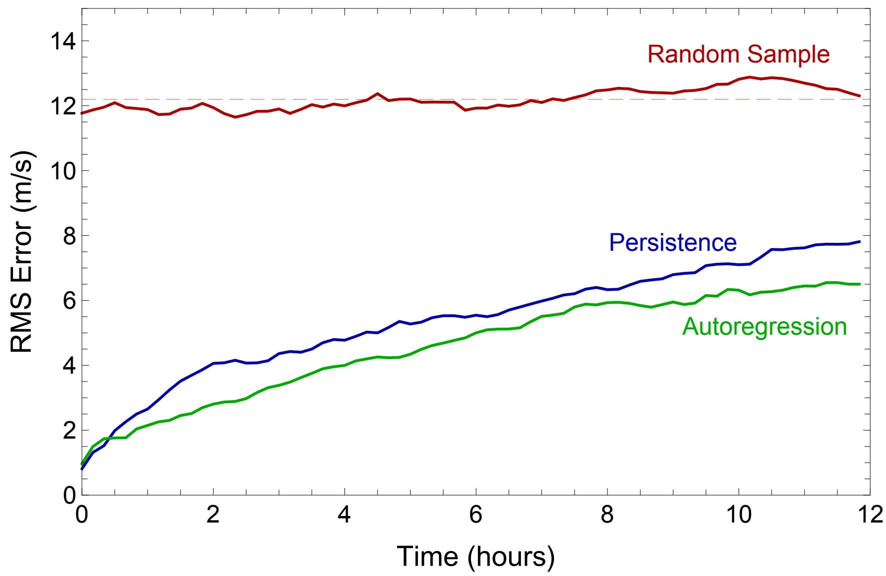

Figure 2: Comparison of RMS prediction errors incurred by the persistence (blue curve), random sample (red), and autoregressive (green) models for predictions up to 12 hours. The horizontal dashed red line indicates the average (mean) value of the RMS error for the random sample model (equal to approximately 12.5 m/s).

accurate than the totally independent predictions produced by the alternative models. As shown in Fig. 4, these plateaus remain constant even when forecasting up to 365 days into the future (i.e., $N_{\text{out}} = 365 \times 24F_s$).

4.2.2 Dependence on Season and on Time of Day

The RMS prediction errors incurred in each season by the autoregressive model are plotted in Fig. 5. From this plot, we see that the errors are smallest in the Summer and largest in the Winter. This suggests that wind conditions in Spring and Summer tend to be "more predictable" than those in Fall or Winter. Recall that the model coefficients used here were computed using 10 years of continuous data (see Sec. 3), which likely yielded "average" model coefficients that are generally suitable for all seasons, but not ideal for any season. Consequently, that there exists a significant seasonal dependence in the observed errors suggests that one approach to reduce these errors might be to compute and employ different model coefficients for each season.

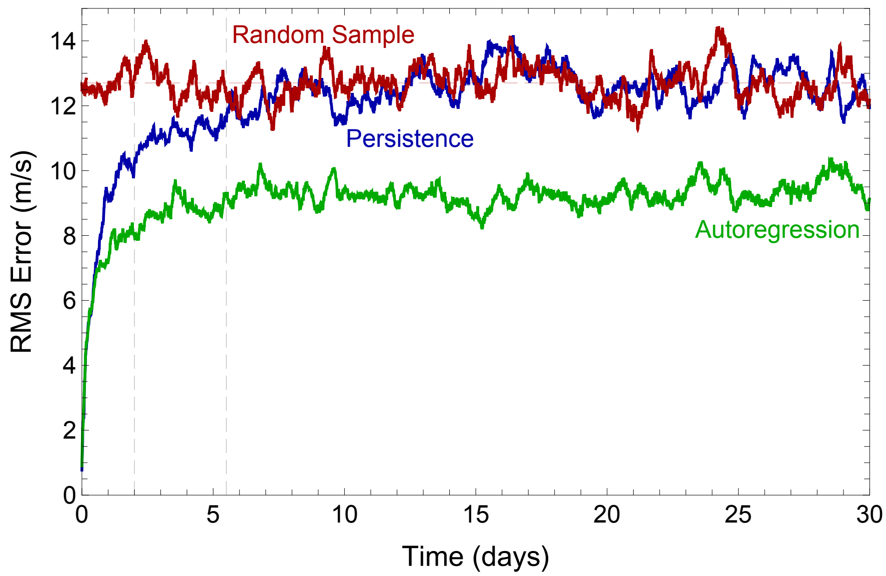

Figure 3: Comparison of RMS prediction errors incurred by each model for predictions up to 30 days. The vertical dashed lines indicate 2 and 5.5 days. See Fig. 2 for more details.

The RMS prediction errors incurred under daytime and nighttime conditions by the autoregressive model are plotted in Fig. 6. From this plot, we see that the prediction errors incurred during the day are usually slightly smaller than those incurred at night. This is somewhat surprising, as we might expect nighttime conditions to be "more predictable" due to the stability of the atmospheric boundary layer [5, Fig. 1.7]. However, these results suggest that the autoregressive model may be better able to predict daytime wind conditions given data from the previous night, than it is to predict nighttime wind conditions given data from the previous day.

5 Summary and Conclusions

In this work we implemented and characterized the performance of an autoregressive model for wind velocity forecasting. We first determined the relevant timescales over which an autoregressive model can be expected to perform well by analyzing the autocorrelation of a dataset of historical wind data. Subsequently, we compared, over various timescales and in terms of an RMS prediction error, the performance of the autoregressive model to the performances of two simpler predictive

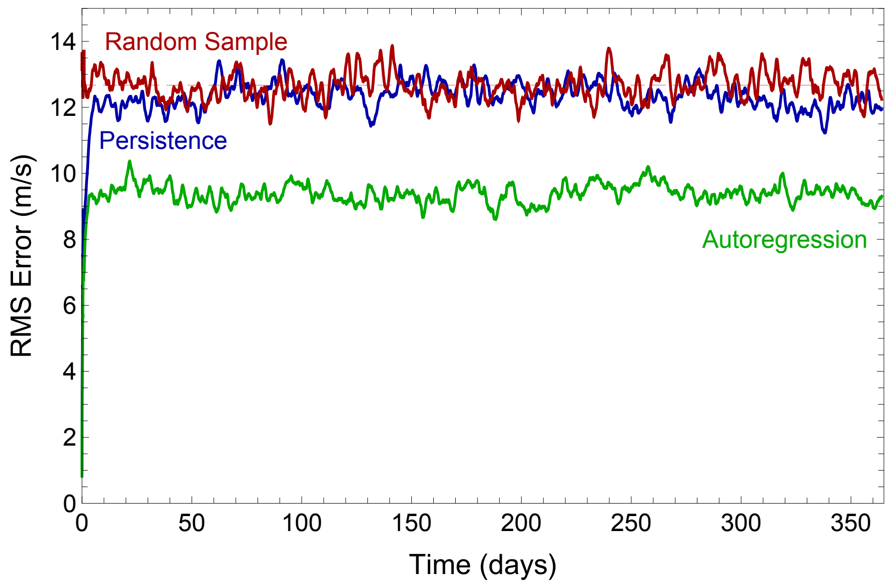

Figure 4: Comparison of RMS prediction errors incurred by each model for predictions up to 365 days. For clarity, the data have been smoothed by computing a 48-hour moving average. See Fig. 2 for more details.

models: persistence and a random sample. Additionally, we examined the dependence of the incurred prediction errors on season and on time of day.

Results of the timescale analysis suggest that the wind system will tend to "forget" previous wind conditions after approximately 5.5 days (see Sec. 4.1). Indeed, results show that, while the autoregressive model consistently outperforms the simpler alternative models, it reaches a plateau in prediction error over approximately that timescale (see Fig. 3). The autoregressive model was also found to yield more accurate predictions in the Summer than in the Winter (see Fig. 5), which suggests that wind conditions in the Summer may be more consistent and "predictable" as a result. Daytime predictions of wind velocity were also found to be slightly more accurate than nighttime predictions (see Fig. 6), although this discrepancy does not appear to be significant.

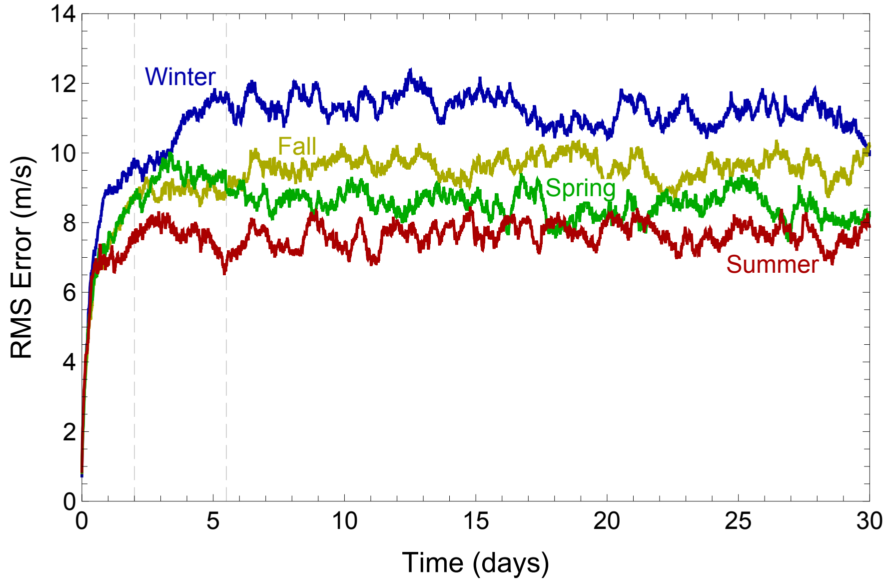

Figure 5: Prediction errors incurred by the autoregressive model in each season. The vertical dashed lines indicate 2 and 5.5 days. Here, "Winter" (blue curve) refers to December through February; "Spring" (green) refers to March through May; "Summer" (red) refers to June through August; and "Fall" (yellow) refers to September through November.

5.1 Future Work

Future investigations should include an implementation of more sophisticated autoregressive models. For example, seasonal [6, Sec. 3.1.6] or Markov-switching [10] autoregressive approaches, in which the autoregression coefficients vary with season, can be expected to achieve improved performance under each condition. Additionally, the performance of the autoregressive model should be compared to numerical weather prediction models, such the Weather Research and Forecasting model,[5] which have been shown to outperform persistence beyond about 6 hours ahead [2, 4].

[5]Available for download here: `https://www.mmm.ucar.edu/weather-research-and-forecasting-model`

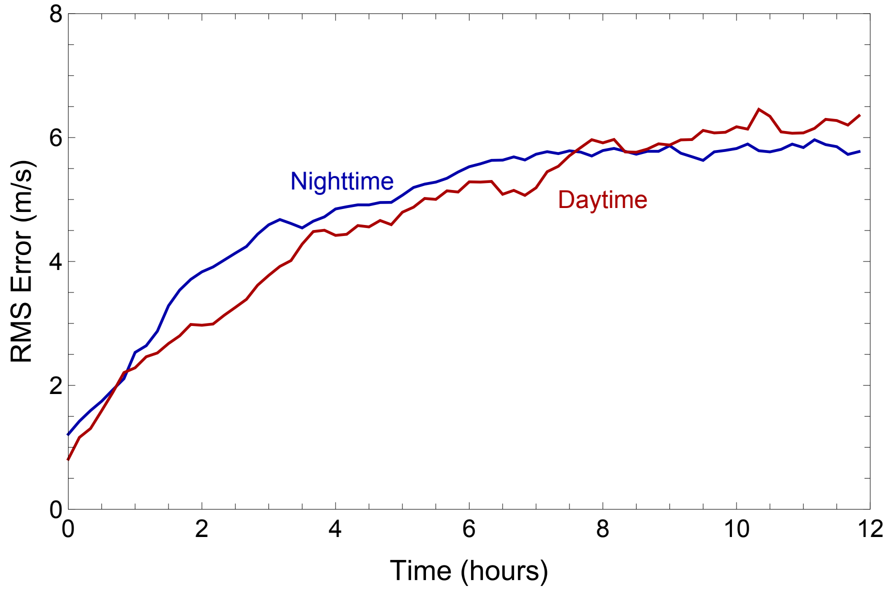

Figure 6: Prediction errors incurred by the autoregressive model under daytime and nighttime conditions. For daytime predictions (red curve), $t = 0$ corresponds to 8AM, while for nighttime predictions (blue), $t = 0$ corresponds to 8PM.

References

[1] Xiaochen Wang, Peng Guo, and Xiaobin Huang. A Review of Wind Power Forecasting Models. *Energy Procedia*, 12:770–778, 2011. ISSN 1876-6102. doi: 10.1016/j.egypro.2011.10.103. URL `http://www.sciencedirect.com/science/article/pii/S1876610211019291`. The Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE) 2011.

[2] Gregor Giebel, Richard Brownsword, George Kariniotakis, Michael Denhard, and Caroline Draxl. *The State-Of-The-Art in Short-Term Prediction of Wind Power: A Literature Overview, 2nd edition*. ANEMOS.plus, 2011. doi: 10.11581/DTU:00000017. Project funded by the European Commission under the 6th Framework Program, Priority 6.1: Sustainable Energy Systems.

[3] Barbara G. Brown, Richard W. Katz, and Allan H. Murphy. Time Series Models to Simulate

and Forecast Wind Speed and Wind Power. *Journal of Climate and Applied Meteorology*, 23 (8):1184–1195, August 1984. doi: 10.1175/1520-0450(1984)023⟨1184:TSMTSA⟩2.0.CO;2. URL `https://doi.org/10.1175/1520-0450(1984)023<1184:TSMTSA>2.0.CO;2`.

[4] Lars Landberg and Simon J. Watson. Short-term prediction of local wind conditions. *Boundary-Layer Meteorology*, 70(1):171–195, July 1994. ISSN 1573-1472. doi: 10.1007/BF00712528. URL `https://doi.org/10.1007/BF00712528`.

[5] Roland B. Stull. *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

[6] Chris Chatfield. *Time-Series Forecasting*. Chapman & Hall/CRC, Boca Raton, Florida, 2000.

[7] A. P. Van Ulden and J. Wieringa. Atmospheric Boundary Layer Research at Cabauw. In J. R. Garratt and P. A. Taylor, editors, *Boundary-Layer Meteorology 25th Anniversary Volume, 1970–1995*, pages 39–69. Springer Netherlands, Dordrecht, 1996. ISBN 978-94-017-0944-6. doi: 10.1007/978-94-017-0944-6_3. URL `https://doi.org/10.1007/978-94-017-0944-6_3`.

[8] Predrag Petrović, Djordje Romanic, and Mladjen Ćurić. Homogeneity analysis of wind data from 213 m high Cabauw tower. *International Journal of Climatology*, 38(S1):e1076–e1090, 2018. doi: 10.1002/joc.5434. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.5434`.

[9] P. L. O'Neill, D. Nicolaides, D. Honnery, and J. Soria. Autocorrelation Functions and the Determination of Integral Length with Reference to Experimental and Numerical Data. In *15th Australasian Fluid Mechanics Conference*, December 2004.

[10] Pierre Ailliot and Valérie Monbet. Markov-switching autoregressive models for wind time series. *Environmental Modelling & Software*, 30:92 – 101, 2012. ISSN 1364-8152. doi: 10.1016/j.envsoft.2011.10.011. URL `http://www.sciencedirect.com/science/article/pii/S1364815211002222`.