# Point Downscaling of Surface Wind Speed for Forecast Applications

BRIAN H. TANG

*Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, New York*

NICK P. BASSILL

*New York State Mesonet, Albany, New York*

(Manuscript received 23 May 2017, in final form 26 December 2017)

ABSTRACT

A statistical downscaling algorithm is introduced to forecast surface wind speed at a location. The downscaling algorithm consists of resolved and unresolved components to yield a time series of synthetic wind speeds at high time resolution. The resolved component is a bias-corrected numerical weather prediction model forecast of the 10-m wind speed at the location. The unresolved component is a simulated time series of the high-frequency component of the wind speed that is trained to match the variance and power spectral density of wind observations at the location. Because of the stochastic nature of the unresolved wind speed, the downscaling algorithm may be repeated to yield an ensemble of synthetic wind speeds. The ensemble may be used to generate probabilistic predictions of the sustained wind speed or wind gusts. Verification of the synthetic winds produced by the downscaling algorithm indicates that it can accurately predict various features of the observed wind, such as the probability distribution function of wind speeds, the power spectral density, daily maximum wind gust, and daily maximum sustained wind speed. Thus, the downscaling algorithm may be broadly applicable to any application that requires a computationally efficient, accurate way of generating probabilistic forecasts of wind speed at various time averages or forecast horizons.

## 1. Introduction

Forecasting surface wind speed has broad importance for aviation, wind energy, engineering, public safety, and other applications (Young and Kristensen 1992; Peterka and Shahid 1998; Ashley and Black 2008; Emeis 2014). Accurate sustained and gust wind speed forecasts at the surface at a variety of spatial and temporal resolutions are necessary for planning and warning decisions in these applications on both short and long time scales (Okumus and Dinler 2016). One challenge in predicting wind speed is that available numerical weather prediction (NWP) models and general circulation models (GCMs) do not have the necessary temporal resolution needed to inform or make decisions. For example, current operational NWP models have model time steps ranging from about 20 s to several minutes, depending on the resolution of the model, and usually only have output available at a much larger time step (e.g., 1 h). This challenge has led to the development of downscaling techniques to produce wind information at higher spatial and/or temporal resolution.

One possible method is dynamical downscaling. Dynamical downscaling involves running a high-resolution model over a limited domain using the initial and boundary conditions from a coarser NWP model or GCM (e.g., Horvath et al. 2012; Cao and Fovell 2016; Daines et al. 2016). The spatial and temporal resolution can be controlled to get tailored output for a particular application. However, dynamical downscaling can be computationally expensive, especially if there is a need to explicitly resolve the turbulent eddies within the boundary and surface layers (e.g., Talbot et al. 2012; Mirocha et al. 2014), resolve winds around complex terrain features (e.g., Horvath et al. 2012; Cao and Fovell 2016), or generate many years of wind data for climate change studies (e.g., Daines et al. 2016).

Another possible method is statistical downscaling. Statistical downscaling involves deriving transfer functions that relate NWP model or GCM fields to a more realistic representation of the local- to regional-scale surface wind speed or wind speed distribution.

*Corresponding author*: Brian H. Tang, btang@albany.edu

DOI: 10.1175/JAMC-D-17-0144.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

A number of methods have been developed, including varieties of regression (de Rooy and Kok 2004; Pryor et al. 2005; Cheng et al. 2012; Curry et al. 2012; Huang et al. 2015; Winstral et al. 2017), generalized linear models (Yan et al. 2002; Kirchmeier et al. 2014), cumulative distribution function transformations (Michelangeli et al. 2009), and Kalman filters (Cassola and Burlando 2012). Oftentimes, statistical downscaling methods are applied to the model output from dynamical downscaling in order to correct model biases and gain further detail (e.g., Daines et al. 2016).

When simulating or forecasting the wind speed, especially when trying to compare with observations, one has to be mindful of the averaging periods used to define the wind speed. Downscaling studies have encompassed a range of time averages, such as hourly averages (Curry et al. 2012; Winstral et al. 2017), multiple-hourly averages (Daines et al. 2016), daily averages (Kirchmeier et al. 2014; Huang et al. 2015), and daily maxima (Yan et al. 2002; Hewston and Dorling 2011). Sustained wind speeds, conventionally consisting of an average of instantaneous wind speeds over the duration of minutes, and wind gusts, conventionally consisting of an average of instantaneous wind speeds over the duration of seconds, are also needed at higher frequencies for some applications (e.g., Spark and Connor 2004). Given that NWP models and GCMs lack the ability to directly simulate wind gusts, a number of statistical and physical methods have been developed to parameterize wind gusts (Wieringa 1973; Brasseur 2001; Sheridan 2011; Cheng et al. 2012; Suomi et al. 2013). It would be advantageous to create a downscaling method that could encompass all the above time averages.

The goal is to present a statistical downscaling technique that is capable of realistically simulating the wind at a required temporal resolution to extract any flavor of wind speed, from wind gusts to daily averages (or potentially longer), applied at a point. While other methods have been employed to generate high-resolution time series of the wind at a point, such as simulating the wind as an autoregressive process (Brown et al. 1984; Poggi et al. 2003) or using single-column models (Traiteur et al. 2012), these methods are typically only accurate over short periods of time (less than 1 h). Our goals are to provide a method that can be used over longer periods of time, is computationally cheap and accurate, and can be used to make probabilistic forecasts of wind speed.

The remainder of the paper is organized as follows: Section 2 gives an example in order to motivate the problem. Section 3 introduces the downscaling algorithm. Section 4 describes the data used to train the algorithm. Section 5 discusses the training of parameters that make up the algorithm. Section 6 gives a few examples of the simulated wind speed produced by the algorithm. Section 7 reviews verification metrics for an independent sample of days. Section 8 ends with conclusions.

## 2. Motivating example

Let us presume we have a time series of wind speed, sampled every 5 s, at an observation site, as shown by the black line in Fig. 1a. The details of these data will be elaborated upon in section 4, but the details are not critical for the motivation. Let us also presume we have a model forecast of the wind speed that perfectly matches the 2-min sustained wind speed at each hour (pink dots in Fig. 1a), which we call a perfect-prog forecast. However, we do not have forecasts between each hour, which is usually the case when using actual operational NWP model forecasts. As a first guess to get information between hours, we cubicly interpolate the perfect-prog forecasts to a 5-s time step (pink, dotted line in Fig. 1a).

This interpolated wind speed, which we call the ''resolved'' wind speed, clearly is not representative of the variability of the observed wind speed. Figure 1b shows the error, that is, the observed wind minus the interpolated perfect-prog wind. The error, which we call the ''unresolved'' wind speed, is what we seek to simulate through the downscaling algorithm presented in the next section.

## 3. Downscaling algorithm

The total wind speed $v$ is the sum of the resolved $\overline{v}$ and unresolved components $v'$:

$$v = \overline{v} + v'. \tag{1}$$

The unresolved component (e.g., Fig. 1b) contains the subhourly variability in the wind speed due to a combination of the unresolved boundary layer turbulence and forecast (interpolation) errors. The statistical properties of the unresolved component, particularly its standard deviation and power spectral density, may be used to generate a synthetic time series that mimics the unresolved component, such that

$$v' = \sigma\left\{\left(\sum_{n=1}^{N/2}\Phi\right)^{-1/2}\left[\sum_{n=1}^{N/2}a_n\cos\left(\frac{2\pi nt}{T}\right) + \sum_{n=1}^{N/2}b_n\sin\left(\frac{2\pi nt}{T}\right)\right]\right\}, \tag{2}$$

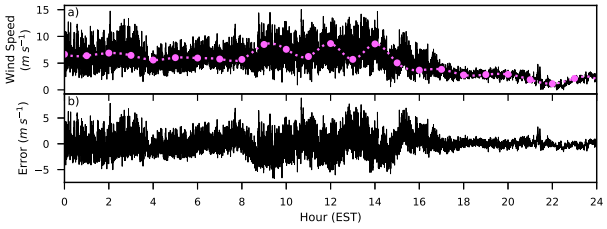


FIG. 1. (a) Observed wind speed (black; $m\,s^{-1}$), perfect-prog 2-min sustained wind speed at each hour (pink dots), and interpolated perfect-prog 2-min sustained wind speed (pink dotted line) at Fayetteville on 23 Oct 2016. (b) Error between observed wind speed and interpolated perfect-prog wind speed ($m\,s^{-1}$).

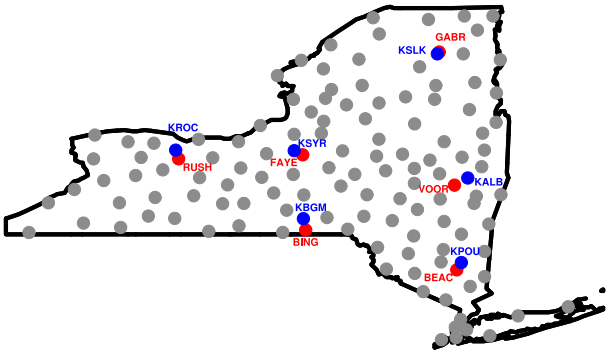


FIG. 2. NYSM standard site locations (gray dots), NYSM sites used herein (red dots), and corresponding nearby ASOS sites (blue dots).

where $\sigma$ is the standard deviation of the wind speed, $N$ is the total number of points in the time series, $\Phi$ is the power spectral density of the normalized wind speed ($v'\sigma^{-1}$), $a_n$ and $b_n$ are random numbers drawn from a Gaussian distribution with a mean of zero and a standard deviation equal to $\Phi$, $T = N\Delta t$ is the total time, and $\Delta t$ is the time resolution of $v$. Note that $\Phi$ is a function of frequency, $s = nT^{-1}$. The term inside the curly braces is constrained to have unit standard deviation and a power spectral density that is given by $\Phi$. A similar method was used by Emanuel et al. (2006) to generate synthetic time series of the flow at different levels of the troposphere to advect hypothetical tropical cyclones in a risk assessment study.

Because of the stochastic nature of the last term in Eq. (2), multiple evaluations of Eq. (2) yield different time series simulating the unresolved component. This method is advantageous because it allows probabilistic approaches, since one does not know the unresolved component deterministically.

When Eq. (2) is added to the resolved component, explained in section 5c, the combined result yields a synthetic wind speed forecast with the time step of $\Delta t$ over a total time of $T$. The synthetic wind speed forecast has subhourly variance and a power spectral density that mimics observations of the unresolved component of the wind speed. To apply Eqs. (1) and (2), one needs to estimate $\Phi$, $\sigma$, and $\overline{v}$.

## 4. Data

To estimate these parameters, we use one year (October 2016–October 2017) of New York State Mesonet (NYSM) observations from "standard" sites that observe common meteorological variables (Fig. 2). The wind data are produced by an RM Young 05103 Wind Monitor, which measures wind speed and direction at 10 m above ground level. This device is a propeller-style anemometer with an expected accuracy of $\pm 0.3\,m\,s^{-1}$ for wind speed and $\pm 3°$ for wind direction. The data sampling rate is 3 s, and the data are quality controlled. More information about the NYSM can be found in Brotzge et al. (2017).

We selected a subset of NYSM sites based on two criteria. The first criterion is that the site has data going back to 2016. A full year of data allows for a sufficiently large sample size to begin to assess how the downscaling algorithm performs in different seasons. The second criterion is that the site be in close proximity to a long-term climate reporting Automated Surface Observing System (ASOS) station. For an NYSM site to be paired to an ASOS station, they have to be less than 20 km apart and have an elevation difference less than 100 m. The proximity to long-term climate reporting ASOS stations allows us to determine precipitation type, which is not observed by the NYSM. Days that had freezing rain or drizzle are filtered out to eliminate any possibility of ice accretion leading to erroneous wind readings. Additionally, the proximity to ASOS stations means that archived forecasts of wind speed are available to estimate $\overline{v}$ at these sites. Based on these two criteria, six standard sites are selected: Beacon (BEAC), Binghamton (BING), Fayetteville (FAYE), Gabriels (GABR), Rush (RUSH), and Voorheesville (VOOR). Note that the selection of these six sites is not meant to be representative of all environments, such as topography, surface roughness, or sheltering in the NYSM network.

Surface and point sounding forecasts (Hart et al. 1998; Hart and Forbes 1999) from the North American Mesoscale Forecast System (NAM) model are obtained from archived Buffalo Toolkit for Lake Effect Snow (BUFKIT; Niziol and Mahoney 1997) input files at the six corresponding ASOS stations: Poughkeepsie (KPOU), Binghamton (KBGM), Syracuse (KSYR), Saranac Lake (KSLK), Rochester (KROC), and Albany (KALB) (Fig. 2). For each 0000 UTC NAM BUFKIT input file

TABLE 1. Wind speed variables, names, and definitions.

| Variable | Name | Definition |
|---|---|---|
| $v$ | Wind speed | 5-s sample wind speed |
| $v_o$ | Observed wind speed | 5-s sample wind speed from NYSM sites |
| $v_s$ | Sustained wind speed | 2-min avg of $v$ |
| $v_{s,o}$ | Observed sustained wind speed | 2-min avg of $v_o$ |
| $v_{s,\mathrm{pp}}$ | Perfect-prog sustained wind speed | Interpolated hourly $v_{s,o}$ |
| $v_{s,f}$ | Forecast sustained wind speed | Interpolated hourly NAM 10-m wind speed |
| $v_{s,f}^{\mathrm{bc}}$ | Bias-corrected forecast sustained wind speed | Bias-corrected interpolated hourly NAM 10-m wind speed |
| $v_g$ | Gust wind speed | Max $v$ within a 2-min interval |
| $v_{g,o}$ | Observed gust wind speed | Max $v_o$ within a 2-min interval |

from October 2016 to October 2017, forecasts of NAM 10-m wind are extracted for the first 24-h period beginning at midnight eastern standard time (EST; 0500 UTC). Wind forecasts are available at hourly resolution. As in the motivating example, the forecast data are cubicly interpolated between hours, which we define as $v_{s,f}$, to generate NAM forecasts of the resolved wind speed. A cubic interpolation is chosen because it represents increases and decreases in wind speed between hours better than a lower-degree interpolation, but the methodology is not sensitive to the choice of interpolation.

In preparing the NYSM wind data for the estimation of parameters in the downscaling algorithm, we conform to past ASOS standards (Nadolski 1998). The NYSM wind data are downsampled from a 3-s interval to a 5-s interval by applying a Fourier transform to the 3-s time series, phase shifting the components of the transform, and then applying an inverse Fourier transform to yield the downsampled 5-s time series (Prichard and Theiler 1994). The raw 5-s wind data are defined as $v_o$. The sustained wind speed, $v_{s,o}$, is a moving average of $v_o$ over a 2-min period. The gust wind speed $v_{g,o}$ is the maximum value of $v_o$ within the same 2-min period.

Additionally, we define perfect-prog forecasts of the sustained wind speed $v_{s,\mathrm{pp}}$. As in the motivating example, the perfect-prog wind speed matches the observed wind speed at the top of every hour, and the perfect-prog wind speeds are then cubicly interpolated between hours. The purpose of having a perfect-prog forecast is to evaluate the downscaling algorithm without the presence of NAM forecast errors. Table 1 summarizes all the wind speed variables used herein.

To train and validate the downscaling algorithm, we divide the wind data at each site into six portions. Each portion is noncontiguous (i.e., every sixth element is selected) and nonoverlapping. The first five portions are used to conduct a fivefold cross validation, in which four-fifths of the data is used to train the downscaling algorithm, and one-fifth is used for validation. This procedure is repeated 5 times for all possible combinations of training and validation folds. The validation is used to arrive at a selected downscaling algorithm for each NYSM site. Finally, the sixth portion is used to test and verify the selected downscaling algorithm at each site.

## 5. Parameters

### a. *Power spectral density* $\Phi$

The power spectral density (Chatfield 2003) is computed from 1-h segments of the observed, normalized wind speed at each NYSM site. The choice of 1-h segments represents the spacing between consecutive NAM or perfect-prog wind forecasts, and is hence the time interval over which we wish to simulate the unresolved wind speed. The normalized wind speed is defined as

$$v^* = v'/\sigma_{v'}, \tag{3}$$

where $v^*$ is the normalized wind speed, $v' = v_o - v_{s,\mathrm{pp}}$, and $\sigma_{v'}$ is the standard deviation of $v'$. It is necessary to perform the normalization before calculating the power spectral densities; otherwise the power spectral density for each segment would be offset in magnitude because of conservation of total energy in both time and frequency space (Parseval's theorem).

The power spectral density is computed using Welch's method (Welch 1967). Each 1-h segment is broken into five pieces with 50% overlap using a Hanning window. The power spectral density is computed for each piece and then averaged.

The mean power spectral density of the 1-h segments of $v^*$ (colored, dashed lines in Fig. 3) serves as an empirical estimate of $\Phi$ at each site. It is important to emphasize that these estimates are not the power spectral densities of the observed wind itself but rather the error between an *imperfect* forecast and the observed wind speed, and therefore they have different characteristics. Namely, the power spectral density of the normalized wind represents a combination of unresolved subhourly

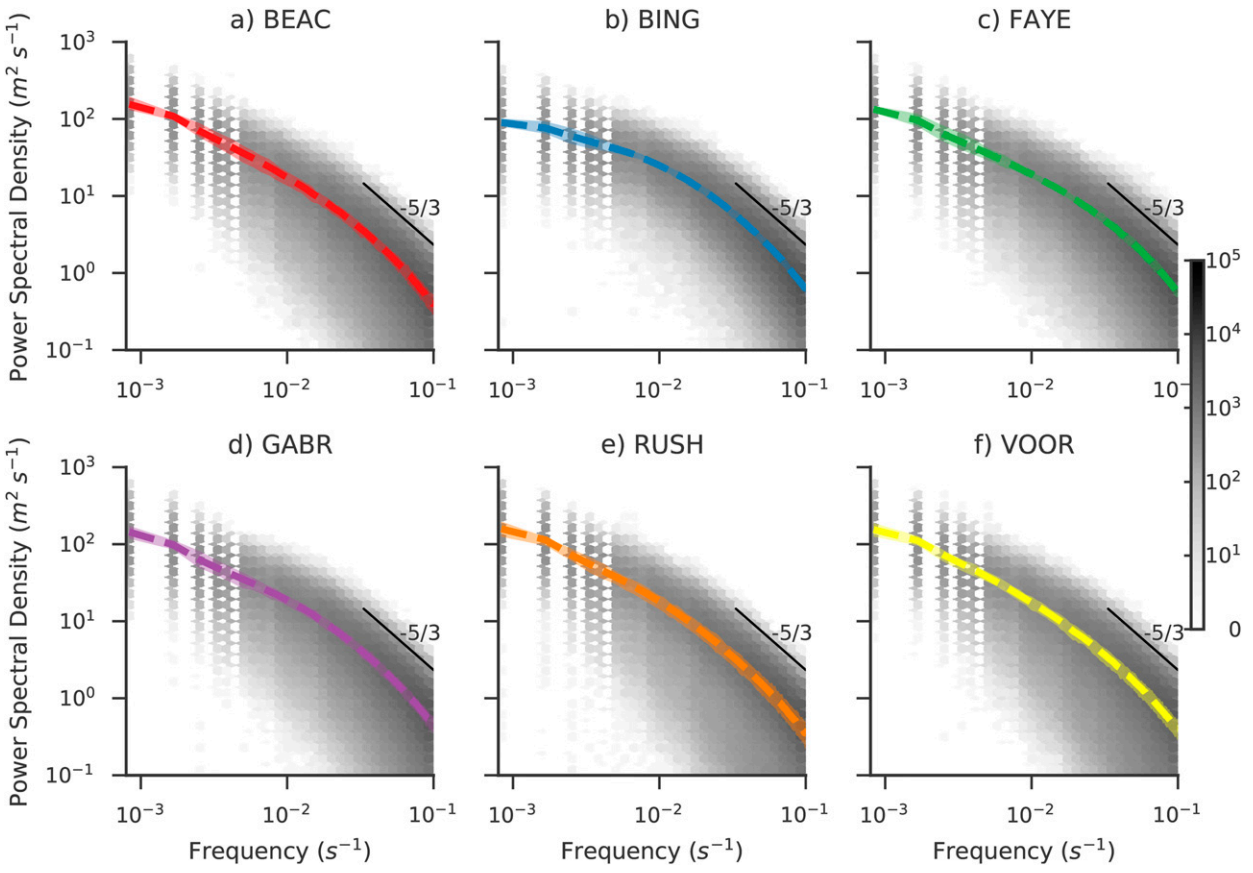


FIG. 3. Mean power spectral densities of 1-h segments of the normalized wind at (a) BEAC, (b) BING, (c) FAYE, (d) GABR, (e) RUSH, and (f) VOOR (dashed lines; $m^2\ s^{-1}$). Shading of the same color about each dashed line gives the range of mean power spectral densities for all hours and all cross-validation folds. Gray shading shows the bin counts (log scale) of power spectral density points for all 1-h segments. The thin black line is a reference −5/3 power law.

variability, such as turbulence in the eddy-containing range and in the inertial subrange where the power spectral density parallels the −5/3 power law (Stull 1988), and forecast errors caused by the interpolation that do not have any physical relationship to boundary layer physics. These forecast errors are unavoidable and should be incorporated into the algorithm design.

The range in the empirical estimates of Φ across all hours and across the five cross-validation folds is small. In fact, the range is barely discernible in Fig. 3. Therefore, there is confidence that the estimates of Φ do not depend on the hour or sample of days. However, there are differences in Φ between sites, which are possibly due to unique mesoscale and microscale heterogeneities in the local environment at each site. Note that there is spread about the mean when considering all the power spectral densities (gray shading in Fig. 3). It is not possible to capture all the spread with a single estimate of Φ at each site, so it represents a typical power spectral density of the unresolved component of the wind speed.

### b. *Standard deviation σ*

Recall that the term inside the curly braces in Eq. (2) is constrained to have unit standard deviation. The standard deviation of the wind speed is needed to denormalize this term. Following Harper et al. (2010) and Suomi et al. (2013), the standard deviation of the wind speed can be estimated as

$$\sigma = \frac{v_g - v_s}{\alpha}. \tag{4}$$

The empirically estimated parameter $\alpha$ is called the normalized gust. Using the observed wind data from the NYSM observations, $\alpha$ can be estimated by regressing $(v_{g,o} - v_{s,o})$ on $\sigma_{v_o}$ and calculating the slope of the least squares linear fit. Figure 4 shows that the mean $\alpha$ for all the cross-validation folds ranges between 1.98 and 2.04 at the six sites. In other words, the gust wind speed tends to be about two standard deviations away from the sustained wind speed. There is a very small range

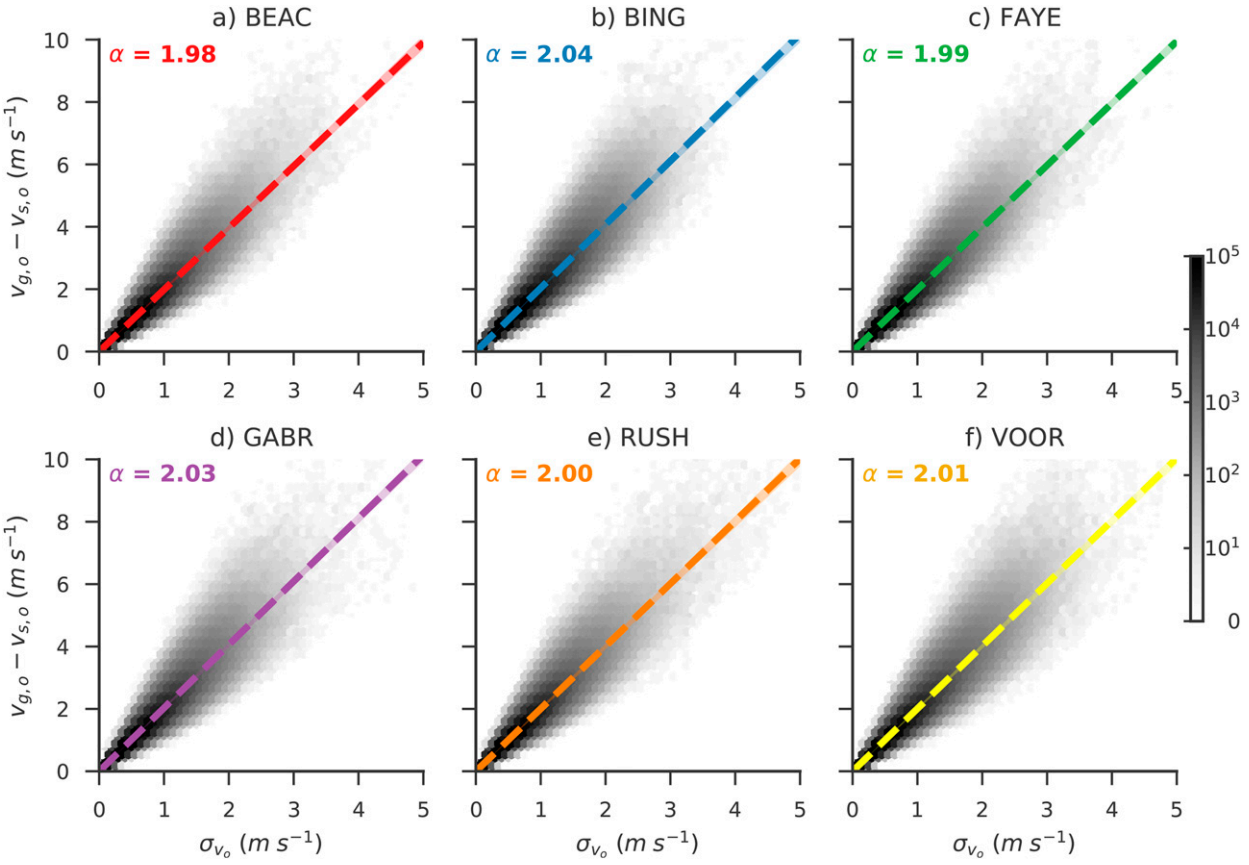


FIG. 4. Linear regression (dashed line) of $(v_{g,o} - v_{s,o})$ on $\sigma_{v_o}$ for (a) BEAC, (b) BING, (c) FAYE, (d) GABR, (e) RUSH, and (f) VOOR. Shading of the same color about each dashed line gives the range of linear regressions for all cross-validation folds. Gray shading shows bin counts (log scale) of all the data in $[\sigma_{v_o}, (v_{g,o} - v_{s,o})]$ parameter space. The mean $\alpha$ (slope of the linear regression line) is given in the upper left.

(<0.01) in $\alpha$ across the folds, and in the 95% confidence interval of $\alpha$ in each fold, lending high confidence in the estimation of $\alpha$. Similar results are obtained if using the perfect-prog winds in place of the observed winds on the rhs of Eq. (4).

For forecasting applications, we must parameterize $v_g$. The simplest parameterization is to use a gust factor $\beta$, where

$$v_g = \beta v_s . \tag{5}$$

A similar approach may then be taken by regressing $v_{g,o}$ on $v_{s,o}$, as shown in Fig. 5. Figure 5 shows that the mean $\beta$ for all the cross-validation folds ranges between 1.40 and 1.57 at the six sites. In other words, the gust wind speed tends to be about 40%–57% higher than the sustained wind speed. There is also a very small range (<0.01) in $\beta$ across the folds, and in the 95% confidence interval of $\beta$ in each fold, lending high confidence in the estimation of $\beta$.

Both $\alpha$ and $\beta$ have diurnal variability that is important to consider (Gallagher 2016). To explore this diurnal variability, the regressions used to calculate $\alpha$ and $\beta$ are repeated, but controlling for hour, as shown in Fig. 6. The highest values of $\alpha$ occur in the evening and overnight hours, and the lowest values occur in the morning hours. At BEAC, BING, FAYE, and GABR, the lowest values of $\beta$ occur overnight, and the highest values occur during the day. There is little diurnal variability in $\beta$ at RUSH and VOOR. The combined diurnal variability of $\alpha$ and $\beta$ can result in about a 10% difference in $\sigma$ between night and day, which is nontrivial. Therefore, we incorporate this diurnal variability of $\alpha$ and $\beta$ into the downscaling algorithm by making both parameters a function of hour.

There are a few of caveats that should be kept in mind. Expressing $\sigma$ with Eqs. (4) and (5) is a major simplification that circumvents the physical relationship between surface wind variability (including wind gusts) and boundary layer structure. A more physically based expression relating $\sigma$ to boundary layer parameters, like the vertical stability, would be more useful in situations that deviate from typical diurnal variability. In addition to diurnal variability, there is likely seasonal variability

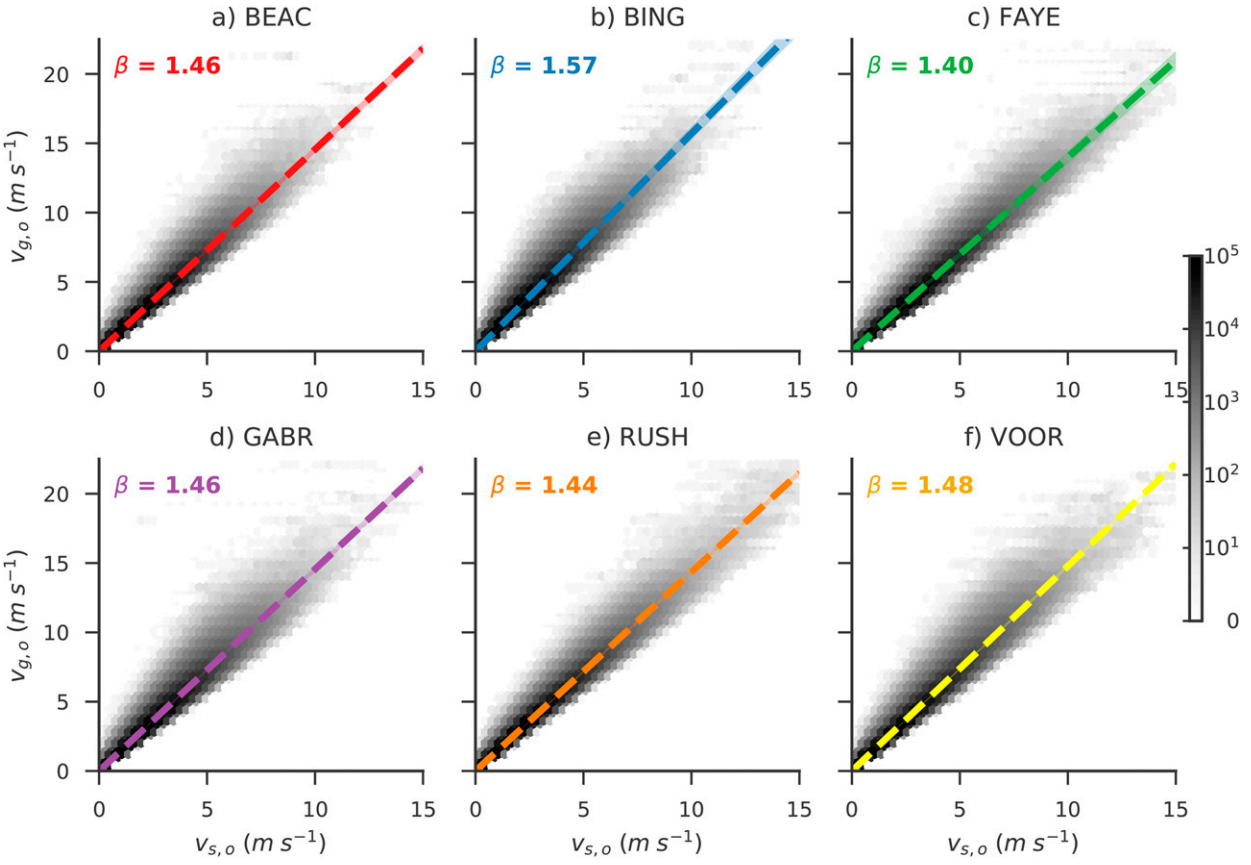


FIG. 5. Linear regression (dashed line) of $v_{g,o}$ on $v_{s,o}$ for (a) BEAC, (b) BING, (c) FAYE, (d) GABR, (e) RUSH, and (f) VOOR. Shading of the same color about each dashed line gives the range of linear regressions for all cross-validation folds. Gray shading shows bin counts (log scale) of all the data in ($v_{s,o}$, $v_{g,o}$) parameter space. The mean $\beta$ (slope of the linear regression line) is given in the upper left.

in both $\alpha$ and $\beta$ (Gallagher 2016). We do not have multiple years of NYSM wind data to rigorously explore the seasonal variability further, since the NYSM was recently installed. Additionally, both $\alpha$ and $\beta$ depend on the time averages used to calculate both the sustained and gust wind speeds (Suomi et al. 2013; Gallagher 2016), so the values here are particular to the past ASOS standards that we have adopted.

### c. *Resolved wind speed* $\overline{v}$

The last piece needed for the downscaling algorithm is the choice of $\overline{v}$. One may simply use $v_{s,f}$ directly as $\overline{v}$, but

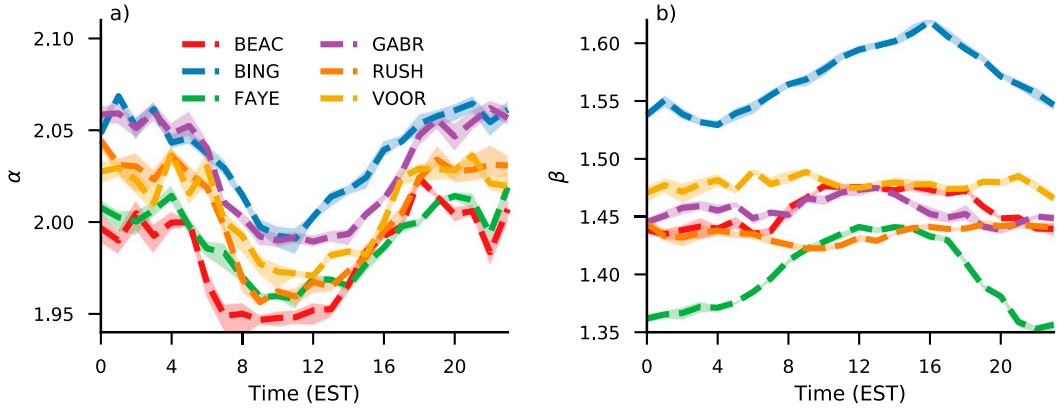


FIG. 6. (a) $\alpha$ and (b) $\beta$ as a function of hour for each of the six sites. Shading of the same color about each dashed line gives the range of $\alpha$ and $\beta$ for all cross-validation folds.

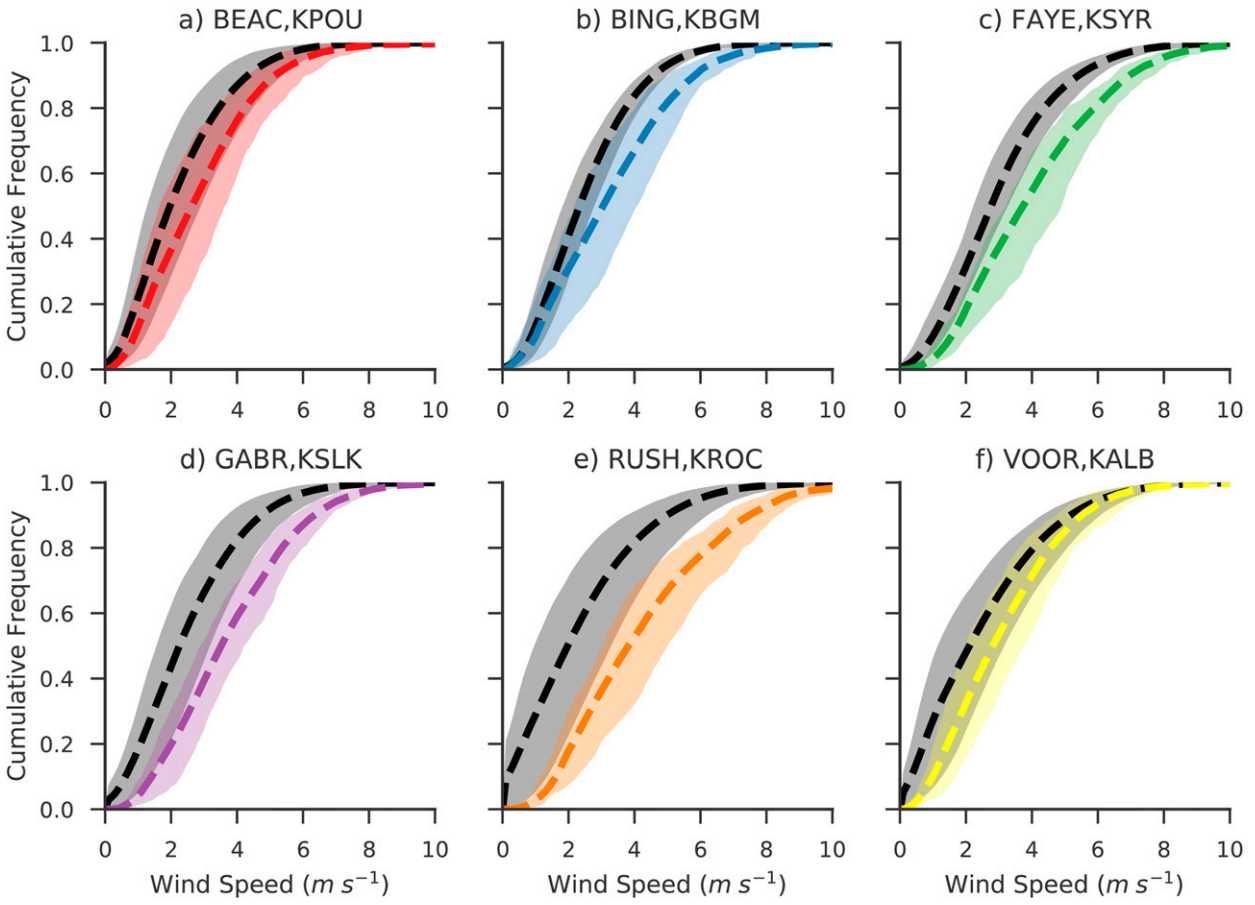


FIG. 7. CDFs of the observed sustained wind speed $v_{s,o}$ (dashed black lines) at (a) BEAC, (b) BING, (c) FAYE, (d) GABR, (e) RUSH, and (f) VOOR. Also shown are CDFs of the NAM forecast 10-m wind speed, $v_{s,f}$, (dashed colored lines) at (a) KPOU, (b) KBGM, (c) KSYR, (d) KSLK, (e) KROC, and (f) KALB. Shading of the same color about each dashed line gives the range of CDFs for all hours and all cross-validation folds.

one must be cautious that $v_{s,f}$ is not biased. Biases in $v_{s,f}$ would introduce the same biases in the synthetically generated wind time series and cause systematic errors in $\sigma$, resulting in potentially poor forecasts.

We employ an empirical quantile mapping bias correction (Michelangeli et al. 2009; Gudmundsson et al. 2012) on $v_{s,f}$:

$$v_{s,f}^{\mathrm{bc}} = F_o^{-1}[F_f(v_{s,f})], \tag{6}$$

where $F_f$ is the cumulative distribution function (CDF) operator on the sample of forecasts, $F_o$ is the CDF operator on the sample of observations, and $v_{s,f}^{\mathrm{bc}}$ is the bias-corrected forecast sustained wind speed. The CDFs are estimated at each individual NYSM site from the samples of $v_{s,o}$ and $v_{s,f}$ within the cross-validation folds (Fig. 7).

Since the bias may have diurnal dependence, the samples of $v_{s,o}$ and $v_{s,f}$ are partitioned into hourly segments, and the CDFs are recomputed. Figure 7 shows that there is substantial spread in the hourly CDFs, as given by the shading about the dashed lines, in both the forecast and observed sustained wind speeds because of diurnal variability. Very little of the spread is due to variability in CDFs between cross-validation folds (not shown).

At all six sites, the $v_{s,f}$ CDF is to the right of the $v_{s,o}$ CDF, meaning that the NAM 10-m sustained wind speeds tend to be too high relative to observations. We speculate that the positive bias in the NAM wind speed may be because all six sites are located in valleys (Winstral et al. 2017). Application of Eq. (6) tends to reduce the NAM 10-m wind speeds at these sites. The bias-corrected $v_{s,f}^{\mathrm{bc}}$ is then substituted where $\overline{v}$ appears in Eq. (1) and where $v_s$ appears in Eqs. (4) and (5).

## 6. Examples of simulated wind

We now have all the pieces needed to apply the downscaling algorithm to simulate the wind speed. We arbitrarily choose one day at three of the sites. The first

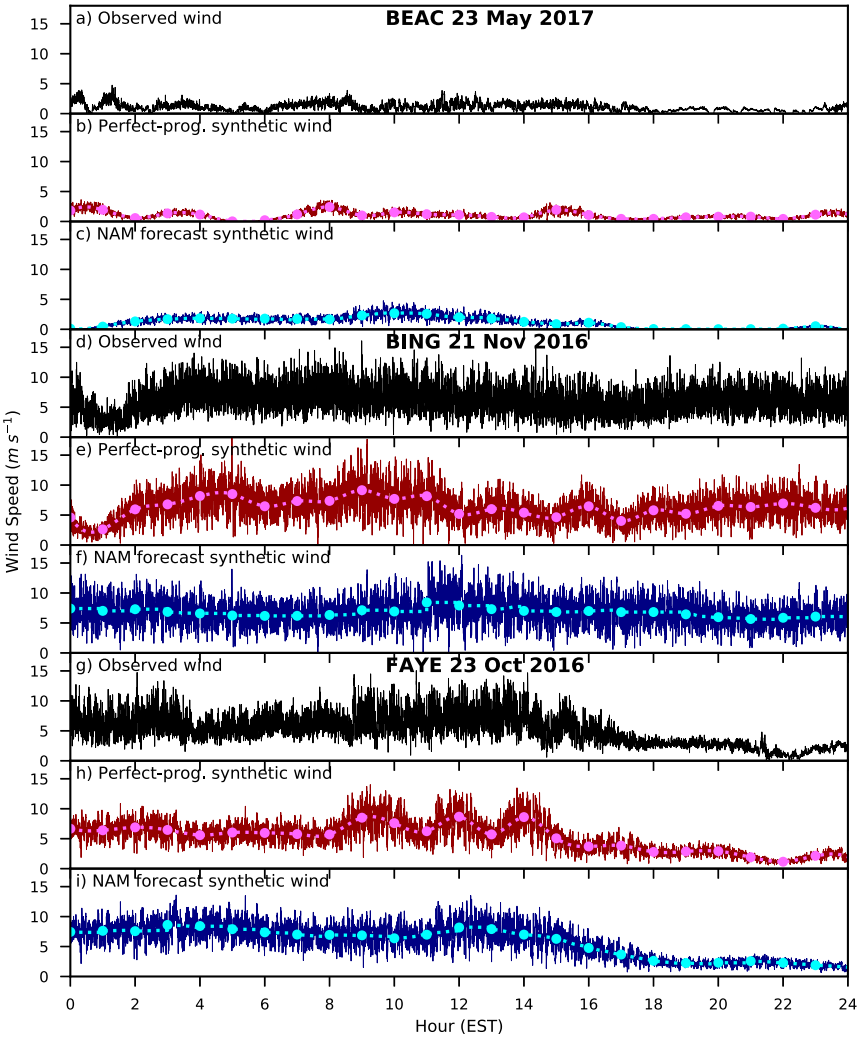


FIG. 8. (a) Observed wind speed ($m\,s^{-1}$) at BEAC on 23 May 2017. (b) Interpolated perfect-prog sustained wind speed (pink) and perfect-prog synthetic wind speed (dark red). (c) Interpolated, bias-corrected NAM forecast sustained wind speed (cyan) and NAM forecast synthetic wind speed (dark blue), using the NAM initialized at 0000 UTC 23 May 2017 at KPOU. (d)–(f) As in (a)–(c), but for BING on 21 Nov 2016 and the NAM initialized at 0000 UTC 21 Nov 2016 at KBGM in (f). (g)–(i) As in (a)–(c), but for FAYE on 23 Oct 2016 and the NAM initialized at 0000 UTC 23 Oct 2016 at KSYR in (i).

example is 23 May 2017 at BEAC, which is a relatively calm day (Figs. 8a–c). The second example is 21 November 2016 at BING, which is a relatively windy day (Figs. 8d–f). The third example is 23 October 2016 at FAYE, which starts out relatively windy, and then the wind lessens later in the day (Figs. 8g–i).

The parameters $\Phi$, $\sigma$ ($\alpha$ and $\beta$), and $\overline{v}$ are obtained from the cross-validation fold that has the particular site and day in the validation portion. Also recall that $\Phi$, $\alpha$, and $\beta$ are functions of hour and differ for each site. For one iteration, we use the perfect-prog winds $v_{s,\mathrm{pp}}$ in Eqs. (4) and (5), and as $\overline{v}$ in Eq. (1), to yield the "perfect-prog synthetic wind." For the second iteration, we use the bias-corrected NAM 10-m winds $v_{s,f}^{\mathrm{bc}}$ in a similar manner to yield the "NAM forecast synthetic wind." After obtaining the necessary parameters, we simulate the wind speed over a full day by applying Eqs. (1) and (2) with a 1-h interval ($T = 3600\,\mathrm{s}$) and a 5-s time step ($\Delta t = 5\,\mathrm{s}$), and then concatenate the intervals.

Both synthetic winds reproduce the qualitative features of the observed winds. The perfect-prog synthetic winds (Figs. 8b,e,h) and the NAM forecast synthetic winds (Figs. 8c,f,i) appear to approximately match the range of observed winds (Figs. 8a,d,g). Additionally, the

standard deviation tends to be higher as the wind speed increases, which matches observations. However, the NAM forecast synthetic winds have less interhourly variability.

To conduct a more rigorous evaluation of these examples, we compare their distributions and power spectral densities. The distributions of synthetic winds are similar to the distribution of observed winds on these three example days (Figs. 9a,c,e). As expected, the perfect-prog synthetic wind distribution is closer to the observed distribution, because there are forecast errors in the resolved component of the NAM forecast synthetic wind. The power spectral densities of the synthetic winds mimic the shape of the observed power spectral densities (Figs. 9b,d,f). We will show verification metrics for a larger sample of days in the next section.

The synthetic winds shown in Fig. 8 are single realizations of a plausible evolution, but the exact wind speed at a given time cannot be forecast deterministically. One principal advantage of the downscaling algorithm is that it can be repeated to yield an ensemble of plausible evolutions, yielding probabilistic forecasts of wind speed.

For instance, say we wish to forecast the daily maximum 1-, 2-, and 10-min average sustained wind speeds at FAYE on 23 October 2016. The algorithm is repeated 30 times to yield a 30-member ensemble of NAM forecast synthetic winds. Figure 10 shows the forecast probabilities of daily maximum 1-, 2-, and 10-min average sustained wind speeds obtained from the ensemble. The mean of the forecast distribution may serve as a "best guess" forecast, and the spread about the mean gives a sense of the uncertainty.

In summary, this algorithm can be applied to any application that requires probabilistic wind guidance at high time resolution or over specific averaging periods. For example, in an educational application the daily maximum 2-min average sustained wind speed is a forecast variable in the WxChallenge forecast competition (Illston et al. 2013). This algorithm could be used to provide probabilistic guidance of the daily maximum 2-min average sustained wind speed at WxChallenge forecast cities.

## 7. Verification

The cross validation shows that there is little variation in the parameters for each fold, as indicated by little difference in the estimates of $\Phi, \alpha, \beta$, and the CDFs used to bias correct $v_{s,f}$ when considering differences across folds only (controlling for hour) in Figs. 3–7. Therefore, these parameters are stable for different samples of days used to train the algorithm. We average the parameters across all folds to arrive at the selected downscaling algorithm, since the individual downscaling algorithms from each fold are nearly identical.

Recall that one-sixth of the data (56–58 days at each site) was withheld from the training and validation in order to verify the selected downscaling algorithm. For each of these days at each site, we use the downscaling algorithm to generate perfect-prog synthetic winds and NAM forecast synthetic winds, following the examples in the previous section. Both synthetic winds are verified in multiple ways against the observed winds. The first verification metric is the Bhattacharyya distance (Bhattacharyya 1943), which measures differences between two probability distribution functions. The two probability distribution functions are that of the observed wind speeds $p(v)$ and that of the synthetic wind speeds $q(v)$. The Bhattacharyya distance $B_d$ is defined as

$$B_d = -\log\left[\sum_v \sqrt{p(v)q(v)}\right], \tag{7}$$

where the sum is taken over all $v$ in 0.1 m s$^{-1}$ bin widths. Thus, if the two distributions are identical, $B_d = 0$. If the two distributions have no overlap, then $B_d \to \infty$. The Bhattacharyya distance is calculated for each day in the verification dataset and then averaged over all days.

We also generate a 30-member ensemble of perfect-prog synthetic winds and a 30-member ensemble of NAM forecast synthetic winds for each day and site. The ensemble mean of the daily maximum wind gust, and the daily maximum 1-, 2-, and 10-min average sustained wind speeds are verified against their corresponding observed values by calculating the mean absolute error (MAE) across all days.

The results of the verification are shown in Table 2. The Bhattacharyya distance is nearly zero, which indicates that there is almost complete overlap between the normalized distributions of the observed winds and synthetic winds. The mean absolute errors for the daily maximum wind at various averaging periods are also reasonably low. For the perfect-prog synthetic wind, the mean absolute errors are between 0.48 and 1.66 m s$^{-1}$. The error is lower for longer averaging periods. For the NAM forecast synthetic wind, the Bhattacharyya distance and mean absolute errors are all larger than the perfect-prog counterparts, but this is expected in the presence of NAM forecast error. However, the mean absolute errors are still small (0.95–2.66 m s$^{-1}$), which indicates that the downscaling algorithm is useful for predicting wind speeds over a variety of averaging periods characterizing different operational standards.

Figure 11 shows the predicted versus observed daily maximum wind speeds for each of the verification days

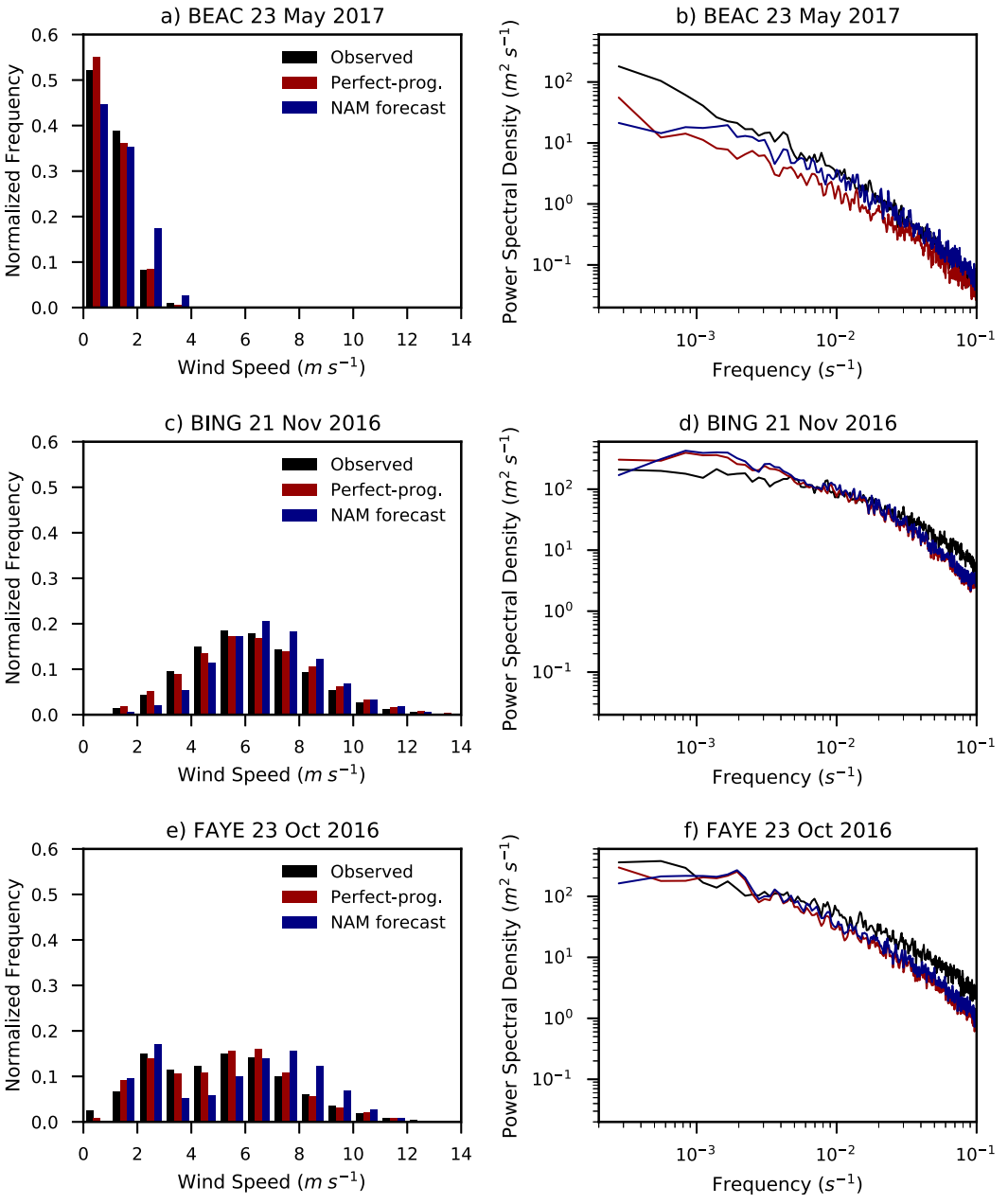


FIG. 9. (left) Normalized frequency histogram and (right) power spectral density ($m^2\ s^{-1}$) of the observed wind (black), perfect-prog synthetic wind (dark red), and NAM forecast synthetic wind (dark blue) at (a),(b) BEAC on 23 May 2017, (c),(d) BING on 21 Nov 2016, and (e),(f) FAYE on 23 Oct 2016.

at each site. The predictions using the perfect-prog synthetic wind (Figs. 11a,c) are tightly clustered about the one-to-one line, consistent with the lower mean absolute errors in Table 2. There does not appear to be any particular bias at specific sites or different wind speeds, at least in the range observed, in the absence of NWP forecast error. When using the NAM forecast synthetic winds (Figs. 11b,d), there is more spread about the one-to-one line, consistent with the higher mean absolute errors in Table 2 with the addition of NWP forecast error. Additionally, there is tendency for the NAM forecast synthetic winds to slightly underpredict the daily maximum wind speeds, especially the wind gust speed. Some of the outlier points with large error at BEAC and FAYE are associated with transient high winds associated with convection (not shown), which the NAM model does not represent well.

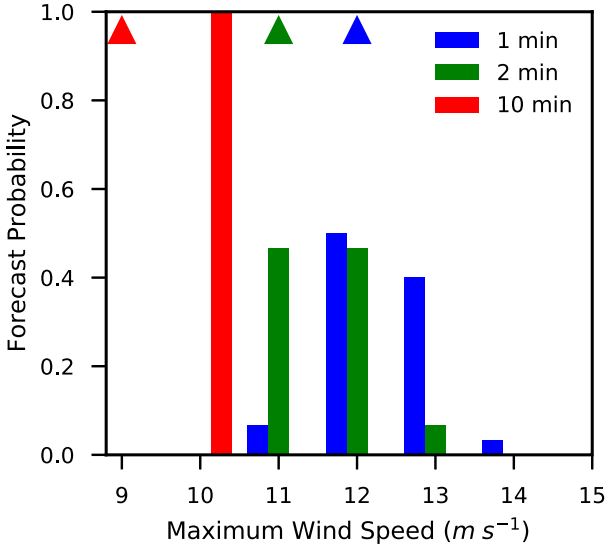


FIG. 10. Ensemble forecast probabilities of the daily max 1-min (blue), 2-min (green), and 10-min (red) sustained wind speed at FAYE on 23 Oct 2016. Triangles show the verifying sustained wind speeds, rounded to the nearest meter per second.

Since we have 1 year of wind data, we can also begin to assess how the downscaling algorithm performs as a function of month and hour in order to determine the seasonal and diurnal performance of the downscaling algorithm. Figure 12 shows there are generally higher errors in the hourly maximum wind speed in meteorological winter and spring (December–May) than in meteorological summer and fall (June–November). The errors tend to be lower in the overnight hours and higher in the daytime hours. The higher errors tend to be associated with when the wind speeds are typically also higher, that is, during the daytime and when there is a larger frequency of strong synoptic-scale storm systems and convection. Nonetheless, over all months and hours, the mean absolute errors of the hourly maximum wind speed are less than $3\,\mathrm{m\,s^{-1}}$.

## 8. Conclusions

We have presented a method to downscale NWP wind forecasts to yield synthetic, high-time-resolution wind speed forecasts at a point. These synthetic wind speeds are designed to match past ASOS specifications, namely 10-m instrument height, 5-s wind speed sampling interval, 2-min sustained wind speed averaging interval, and the gust wind speed defined as the maximum 5-s wind speed within the 2-min interval. From NYSM observations with the same specifications, the empirical power spectral density, normalized gust, and gust factor are estimated and used to simulate the unresolved component of the wind. The NAM 10-m forecast wind speeds are bias corrected with a quantile mapping algorithm and used to determine the resolved component of the wind. The combination of the resolved and unresolved components results in a synthetic wind speed forecast.

The unresolved component of the wind is stochastic in the downscaling algorithm, because the unresolved component cannot be predicted deterministically. As a result, the downscaling algorithm can be repeated to yield an ensemble of synthetic wind speed forecasts. Ensemble wind speeds at various averaging periods may be extracted; yielding probabilistic forecasts of gusts and $n$-min averages, where $n$ is a specified time length, as a function of time. Thus, the algorithm may be used to inform or make decisions for any number of applications that may benefit from probabilistic wind speed forecasts tailored to the needs of the application.

TABLE 2. Verification statistics for the downscaling algorithm. The first column gives the site. The second column indicates whether the synthetic winds are generated using the perfect-prog ($v_{s,\mathrm{pp}}$) or bias-corrected NAM forecast sustained winds ($v_{s,f}^{\mathrm{bc}}$). The remaining columns give verification metrics: Bhattacharyya distance $B_d$; the MAE of the daily max wind gust; and the daily max 1-, 2-, and 10-min avg sustained wind speeds ($\mathrm{m\,s^{-1}}$). See the text for details on the verification metrics.

| Site | $v_s$ input | $B_d$ | MAE gust | MAE 1-min | MAE 2-min | MAE 10-min |
|---|---|---|---|---|---|---|
| BEAC | $v_{s,\mathrm{pp}}$ | 0.016 | 1.32 | 0.86 | 0.74 | 0.57 |
| BEAC | $v_{s,f}^{\mathrm{bc}}$ | 0.081 | 2.66 | 1.71 | 1.45 | 1.07 |
| BING | $v_{s,\mathrm{pp}}$ | 0.012 | 1.19 | 0.78 | 0.69 | 0.48 |
| BING | $v_{s,f}^{\mathrm{bc}}$ | 0.058 | 2.05 | 1.34 | 1.14 | 0.95 |
| FAYE | $v_{s,\mathrm{pp}}$ | 0.014 | 1.66 | 1.02 | 0.87 | 0.59 |
| FAYE | $v_{s,f}^{\mathrm{bc}}$ | 0.084 | 2.52 | 1.92 | 1.73 | 1.31 |
| GABR | $v_{s,\mathrm{pp}}$ | 0.018 | 1.47 | 0.92 | 0.82 | 0.60 |
| GABR | $v_{s,f}^{\mathrm{bc}}$ | 0.089 | 2.57 | 1.83 | 1.61 | 1.38 |
| RUSH | $v_{s,\mathrm{pp}}$ | 0.014 | 1.11 | 0.71 | 0.64 | 0.52 |
| RUSH | $v_{s,f}^{\mathrm{bc}}$ | 0.086 | 2.23 | 1.36 | 1.22 | 0.99 |
| VOOR | $v_{s,\mathrm{pp}}$ | 0.018 | 1.18 | 0.75 | 0.69 | 0.56 |
| VOOR | $v_{s,f}^{\mathrm{bc}}$ | 0.094 | 2.44 | 1.68 | 1.49 | 1.24 |

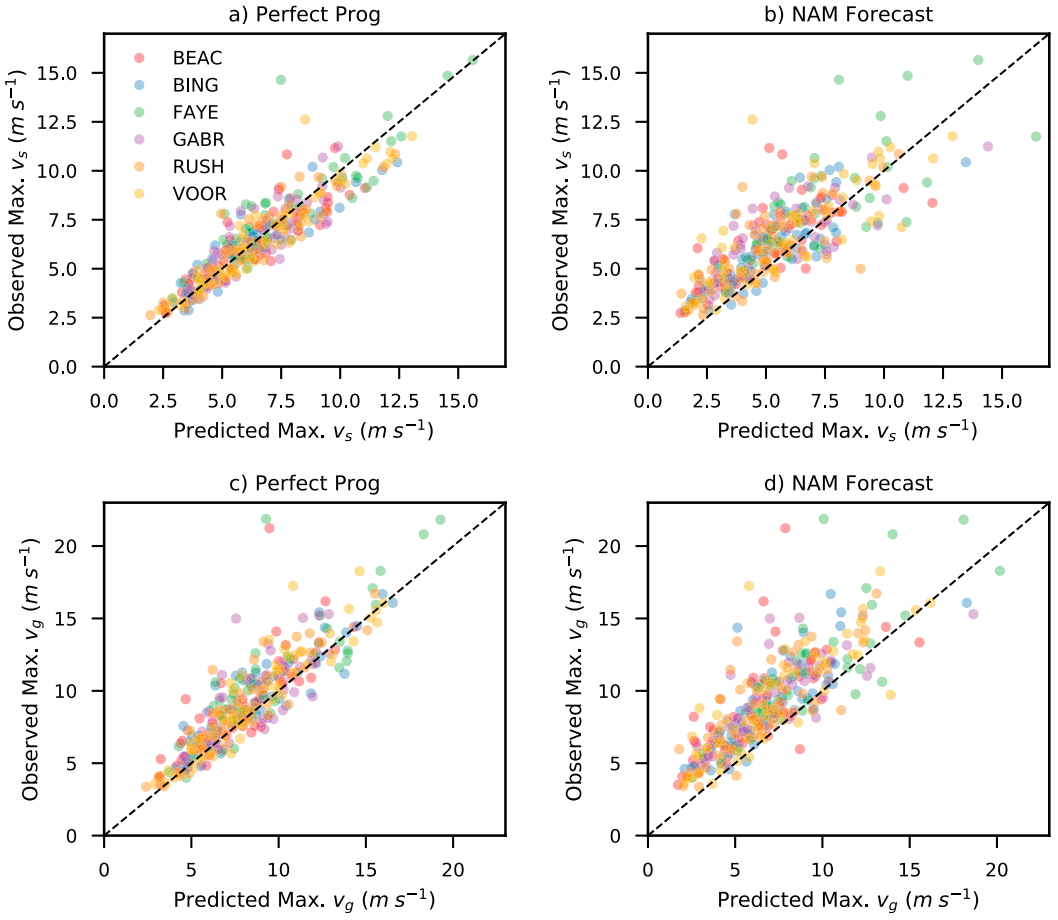


FIG. 11. Predicted vs observed (a), (b) daily max 2-min avg sustained wind speed ($m\,s^{-1}$) and (c), (d) daily max wind gust speed ($m\,s^{-1}$) at the six NYSM sites. Predicted wind speeds are generated using the perfect-prog approach in (a) and (c) and the bias-corrected NAM forecast approach in (b) and (d).

Verification metrics show that the algorithm is accurate. The algorithm closely reproduces the observed distributions of wind speed, provided the NAM wind forecast is skillful itself, and closely parallels the observed power spectral densities. When comparing the forecast daily maximum wind gust, and the daily maximum 1-, 2-, and 10-min average sustained wind speeds with the observed values, mean absolute errors are 0.95–2.66 $m\,s^{-1}$. Mean absolute errors of the hourly maximum wind gust and 2-min average sustained wind speeds are generally $<3\,m\,s^{-1}$ over all months and hours. It would be interesting to compare this algorithm with other dynamical and statistical downscaling algorithms in terms of its accuracy in forecasting measures of the daily wind and wind speed distributions (Horvath et al. 2012; Kirchmeier et al. 2014; Huang et al. 2015).

One advantage of this algorithm is that it is modular. The time period of simulation $T$ and the time step $\Delta t$ may be changed provided one has estimates of $\Phi$ for the range of frequencies encompassed by the choice of $T$ and $\Delta t$. In lieu of using a simple gust factor, one may use a more sophisticated algorithm based on the turbulent kinetic energy and buoyancy profiles within the boundary layer to determine wind gusts (Brasseur 2001). Instead of the quantile mapping bias correction, one can employ any number of bias correction methods, including model output statistics (Glahn and Lowry 1972). One can generate more realistic ensemble uncertainty in the resolved wind speed by using bias-corrected, 10-m winds from a NWP ensemble forecast, like the Global Ensemble Forecast System, or combinations of NWP model forecasts. Although we have focused on NWP and forecasts on short time scales, one could also train this algorithm and repeat the methods herein with a GCM to study winds on longer time scales.

There are several caveats to keep in mind. First, since the algorithm requires observed data to train the algorithm, it requires reliable and high-resolution observed wind data to be available at the desired forecast point. Thus, the algorithm cannot be used in a data-void area.

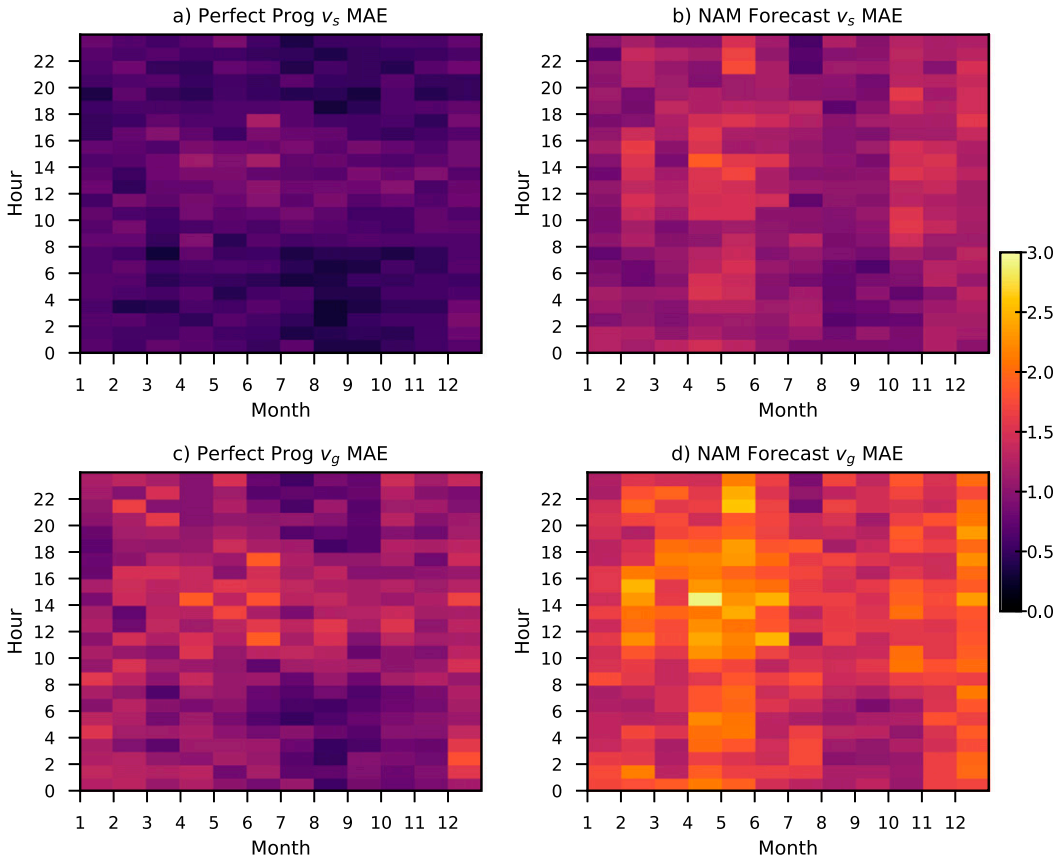


FIG. 12. MAE ($m\,s^{-1}$) of the predicted (a), (b) hourly max 2-min avg sustained wind speed and (c), (d) hourly max wind gust speed, averaged over the six NYSM sites, as a function of month and hour (EST). Predicted wind speeds are generated using the perfect-prog approach in (a) and (c) and the bias-corrected NAM forecast approach in (b) and (d).

One possible way around this problem is to use high-resolution dynamical downscaling to determine the covariance between points that have observations and points that do not have observations. One can then extend the methods presented here from a single point to a local area if the covariances are well constrained and meaningful (Emanuel et al. 2006). Second, different wind measurement specifications would yield different parameters in the downscaling algorithm, so the algorithm is not directly transferable to forecast winds between networks with different specifications or at nonconforming sites within networks. Third, we do not attempt to account for heterogeneity in the normalized gust and gust factor due to season (Gallagher 2016), terrain (Howard and Clark 2007; Winstral et al. 2017), and surface roughness (de Rooy and Kok 2004; Paulsen and Schroeder 2005; Huang et al. 2015), which may be important to account for when estimating these parameters. Fourth, the algorithm would likely have trouble simulating extreme events, when the wind speed forecast is beyond the tail of the CDF of the observed wind speeds, and convective wind gusts, when there are sudden increases in wind that are not representative of typical variability. Regardless of these caveats, we have presented a computationally efficient, accurate method for forecasting wind speed at high time resolution, and in a probabilistic manner, that may be useful for a variety of applications that require wind forecasts at various time averages and forecast horizons.

*Acknowledgments.* This research is made possible by the NYSM. Nathan Bain supplied the NYSM data, which may be requested at http://www.nysmesonet.org/data/requestdata. Original funding for the NYSM was provided by Federal Emergency Management Agency Grant FEMA-4085-DR-NY, with the continued support of the NYS Division of Homeland Security and Emergency Services; the state of New York; the Research Foundation for the State University of New York (SUNY); the University at Albany, SUNY; the Atmospheric Sciences Research Center at SUNY Albany; and the Department of Atmospheric and Environmental Sciences at SUNY Albany. Iowa State University and the BUFKIT Warehouse supplied the

archived BUFKIT data. Three anonymous reviewers and Katherine Klink provided valuable feedback that improved the manuscript.

## REFERENCES


Ashley, W. S., and A. W. Black, 2008: Fatalities associated with nonconvective high-wind events in the United States. *J. Appl. Meteor. Climatol.*, **47**, 717–725, https://doi.org/10.1175/2007JAMC1689.1.

Bhattacharyya, A., 1943: On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, **35**, 239–258.

Brasseur, O., 2001: Development and application of a physical approach to estimating wind gusts. *Mon. Wea. Rev.*, **129**, 5–25, https://doi.org/10.1175/1520-0493(2001)129<0005:DAAOAP>2.0.CO;2.

Brotzge, J. A., C. D. Thorncroft, and E. L. Joseph, 2017: The New York State Mesonet: A general overview. *Special Symp. on Meteorological Observations and Instrumentation*, Seattle, WA, Amer. Meteor. Soc., 1A.2, https://ams.confex.com/ams/97Annual/webprogram/Paper313185.html.

Brown, B. G., R. W. Katz, and A. H. Murphy, 1984: Time series models to simulate and forecast wind speed and wind power. *J. Climate Appl. Meteor.*, **23**, 1184–1195, https://doi.org/10.1175/1520-0450(1984)023<1184:TSMTSA>2.0.CO;2.

Cao, Y., and R. G. Fovell, 2016: Downslope windstorms of San Diego County. Part I: A case study. *Mon. Wea. Rev.*, **144**, 529–552, https://doi.org/10.1175/MWR-D-15-0147.1.

Cassola, F., and M. Burlando, 2012: Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. *Appl. Energy*, **99**, 154–166, https://doi.org/10.1016/j.apenergy.2012.03.054.

Chatfield, C., 2003: *The Analysis of Time Series: An Introduction.* 6th ed. Chapman and Hall/CRC, 352 pp.

Cheng, C. S., G. Li, Q. Li, H. Auld, and C. Fu, 2012: Possible impacts of climate change on wind gusts under downscaled future climate conditions over Ontario, Canada. *J. Climate*, **25**, 3390–3408, https://doi.org/10.1175/JCLI-D-11-00198.1.

Curry, C. L., D. d. Kamp, and A. H. Monahan, 2012: Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. I. Predicting wind speed. *Climate Dyn.*, **38**, 1281–1299, https://doi.org/10.1007/s00382-011-1173-3.

Daines, J. T., A. H. Monahan, and C. L. Curry, 2016: Model-based projections and uncertainties of near-surface wind climate in western Canada. *J. Appl. Meteor. Climatol.*, **55**, 2229–2245, https://doi.org/10.1175/JAMC-D-16-0091.1.

de Rooy, W. C., and K. Kok, 2004: A combined physical–statistical approach for the downscaling of model wind speed. *Wea. Forecasting*, **19**, 485–495, https://doi.org/10.1175/1520-0434(2004)019<0485:ACPAFT>2.0.CO;2.

Emanuel, K., S. Ravela, E. Vivant, and C. Risi, 2006: A statistical deterministic approach to hurricane risk assessment. *Bull. Amer. Meteor. Soc.*, **87**, 299–314, https://doi.org/10.1175/BAMS-87-3-299.

Emeis, S., 2014: Current issues in wind energy meteorology. *Meteor. Appl.*, **21**, 803–819, https://doi.org/10.1002/met.1472.

Gallagher, A. A., 2016: The network average gust factor, its measurement and environmental controls, and role in gust forecasting. M.S. thesis, Dept. of Atmospheric and Environmental Sciences, University at Albany, State University of New York, 139 pp., http://search.proquest.com/pqdtlocal1005758/docview/1862184964/33C786C77CE44A30PQ/1.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, 2012: Downscaling RCM precipitation to the station scale using statistical transformations—A comparison of methods. *Hydrol. Earth Syst. Sci.*, **16**, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012.

Harper, B. A., J. D. Kepert, and J. D. Ginger, 2010: Guidelines for converting between various wind averaging periods in tropical cyclone conditions. WMO Tech. Rep. WMO/TD-1555, 64 pp., https://www.wmo.int/pages/prog/www/tcp/documents/WMO_TD_1555_en.pdf.

Hart, R. E., and G. S. Forbes, 1999: The use of hourly model-generated soundings to forecast mesoscale phenomena. Part II: Initial assessment in forecasting nonconvective strong wind gusts. *Wea. Forecasting*, **14**, 461–469, https://doi.org/10.1175/1520-0434(1999)014<0461:TUOHMG>2.0.CO;2.

——, ——, and R. H. Grumm, 1998: The use of hourly model-generated soundings to forecast mesoscale phenomena. Part I: Initial assessment in forecasting warm-season phenomena. *Wea. Forecasting*, **13**, 1165–1185, https://doi.org/10.1175/1520-0434(1998)013<1165:FTTUOH>2.0.CO;2.

Hewston, R., and S. R. Dorling, 2011: An analysis of observed daily maximum wind gusts in the UK. *J. Wind Eng. Ind. Aerodyn.*, **99**, 845–856, https://doi.org/10.1016/j.jweia.2011.06.004.

Horvath, K., D. Koracin, R. Vellore, J. Jiang, and R. Belu, 2012: Sub-kilometer dynamical downscaling of near-surface winds in complex terrain using WRF and MM5 mesoscale models. *J. Geophys. Res.*, **117**, D11111, https://doi.org/10.1029/2012JD017432.

Howard, T., and P. Clark, 2007: Correction and downscaling of NWP wind speed forecasts. *Meteor. Appl.*, **14**, 105–116, https://doi.org/10.1002/met.12.

Huang, H.-Y., S. B. Capps, S.-C. Huang, and A. Hall, 2015: Downscaling near-surface wind over complex terrain using a physically-based statistical modeling approach. *Climate Dyn.*, **44**, 529–542, https://doi.org/10.1007/s00382-014-2137-1.

Illston, B. G., J. B. Basara, C. Weiss, and M. Voss, 2013: The WxChallenge: Forecasting competition, educational tool, and agent of cultural change. *Bull. Amer. Meteor. Soc.*, **94**, 1501–1506, https://doi.org/10.1175/BAMS-D-11-00112.1.

Kirchmeier, M. C., D. J. Lorenz, and D. J. Vimont, 2014: Statistical downscaling of daily wind speed variations. *J. Appl. Meteor. Climatol.*, **53**, 660–675, https://doi.org/10.1175/JAMC-D-13-0230.1.

Michelangeli, P.-A., M. Vrac, and H. Loukos, 2009: Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophys. Res. Lett.*, **36**, L11708, https://doi.org/10.1029/2009GL038401.

Mirocha, J., B. Kosovi, and G. Kirkil, 2014: Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **142**, 806–831, https://doi.org/10.1175/MWR-D-13-00064.1.

Nadolski, V. L., 1998: Automated Surface Observing System user's guide. NOAA Tech. Rep., 74 pp., http://www.nws.noaa.gov/asos/pdfs/aum-toc.pdf.

Niziol, T. A., and E. A. Mahoney, 1997: The use of high resolution hourly soundings for the prediction of lake effect snow. Preprints, *13th Int. Conf. on Interactive Information and*

*Processing Systems for Meteorology, Oceanography, and Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 92–95.

Okumus, I., and A. Dinler, 2016: Current status of wind energy forecasting and a hybrid method for hourly predictions. *Energy Convers. Manage.*, **123**, 362–371, https://doi.org/10.1016/j.enconman.2016.06.053.

Paulsen, B. M., and J. L. Schroeder, 2005: An examination of tropical and extratropical gust factors and the associated wind speed histograms. *J. Appl. Meteor.*, **44**, 270–280, https://doi.org/10.1175/JAM2199.1.

Peterka, J. A., and S. Shahid, 1998: Design gust wind speeds in the United States. *J. Struct. Eng.*, **124**, 207–214, https://doi.org/10.1061/(ASCE)0733-9445(1998)124:2(207).

Poggi, P., M. Muselli, G. Notton, C. Cristofari, and A. Louche, 2003: Forecasting and simulating wind speed in Corsica by using an autoregressive model. *Energy Convers. Manage.*, **44**, 3177–3196, https://doi.org/10.1016/S0196-8904(03)00108-0.

Prichard, D., and J. Theiler, 1994: Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.*, **73**, 951–954, https://doi.org/10.1103/PhysRevLett.73.951.

Pryor, S. C., J. T. Schoof, and R. J. Barthelmie, 2005: Empirical downscaling of wind speed probability distributions. *J. Geophys. Res.*, **110**, D19109, https://doi.org/10.1029/2005JD005899.

Sheridan, P., 2011: Review of techniques and research for gust forecasting and parameterisation forecasting. UKMO Forecasting Research Tech. Rep. 570, 19 pp.

Spark, E., and G. J. Connor, 2004: Wind forecasting for the sailing events at the Sydney 2000 Olympic and Paralympic Games. *Wea. Forecasting*, **19**, 181–199, https://doi.org/10.1175/1520-0434(2004)019<0181:WFFTSE>2.0.CO;2.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic, 666 pp.

Suomi, I., T. Vihma, S.-E. Gryning, and C. Fortelius, 2013: Wind-gust parametrizations at heights relevant for wind energy: A study based on mast observations. *Quart. J. Roy. Meteor. Soc.*, **139**, 1298–1310, https://doi.org/10.1002/qj.2039.

Talbot, C., E. Bou-Zeid, and J. Smith, 2012: Nested mesoscale large-eddy simulations with WRF: Performance in real test cases. *J. Hydrometeor.*, **13**, 1421–1441, https://doi.org/10.1175/JHM-D-11-048.1.

Traiteur, J. J., D. J. Callicutt, M. Smith, and S. B. Roy, 2012: A short-term ensemble wind speed forecasting system for wind power applications. *J. Appl. Meteor. Climatol.*, **51**, 1763–1774, https://doi.org/10.1175/JAMC-D-11-0122.1.

Welch, P., 1967: The use of the fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, **15**, 70–73, https://doi.org/10.1109/TAU.1967.1161901.

Wieringa, J., 1973: Gust factors over open water and built-up country. *Bound.-Layer Meteor.*, **3**, 424–441, https://doi.org/10.1007/BF01034986.

Winstral, A., T. Jonas, and N. Helbig, 2017: Statistical downscaling of gridded wind speed data using local topography. *J. Hydrometeor.*, **18**, 335–348, https://doi.org/10.1175/JHM-D-16-0054.1.

Yan, Z., S. Bate, R. E. Chandler, V. Isham, and H. Wheater, 2002: An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. *J. Climate*, **15**, 2073–2088, https://doi.org/10.1175/1520-0442(2002)015<2073:AAODMW>2.0.CO;2.

Young, G. S., and L. Kristensen, 1992: Surface-layer gusts for aircraft operation. *Bound.-Layer Meteor.*, **59**, 231–242, https://doi.org/10.1007/BF00119814.