Time Series Models to Simulate and Forecast Wind Speed and Wind Power

BARBARA G. BROWN[1]

Department of Atmospheric Sciences, Oregon State University, Corvallis, OR 97331

RICHARD W. KATZ

Environmental and Societal Impacts Group, National Center for Atmospheric Research,[2] Boulder, CO 80307

ALLAN H. MURPHY

Department of Atmospheric Sciences, Oregon State University, Corvallis, OR 97331

(Manuscript received 30 August 1983, in final form 22 May 1984)

ABSTRACT

A general approach for modeling wind speed and wind power is described. Because wind power is a function of wind speed, the methodology is based on the development of a model of wind speed. Values of wind power are estimated by applying the appropriate transformations to values of wind speed. The wind speed modeling approach takes into account several basic features of wind speed data, including autocorrelation, non-Gaussian distribution, and diurnal nonstationarity. The positive correlation between consecutive wind speed observations is taken into account by fitting an autoregressive process to wind speed data transformed to make their distribution approximately Gaussian and standardized to remove diurnal nonstationarity.

As an example, the modeling approach is applied to a small set of hourly wind speed data from the Pacific Northwest. Use of the methodology for simulating and forecasting wind speed and wind power is discussed and an illustration of each of these types of applications is presented. To take into account the uncertainty of wind speed and wind power forecasts, techniques are presented for expressing the forecasts either in terms of confidence intervals or in terms of probabilities.

1. Introduction

Before wind power can be integrated into a large multisource energy network, it is necessary to obtain good estimates of its potential contribution to that network. Because the data records at potential wind power sites typically are of short length, such determinations usually must be made using computer simulations of long series of wind power values. Once a wind power generator is supplying power to an energy system, a method of forecasting wind power output a few hours in advance is required to ensure efficient utilization of the power. Since wind power is a function of wind speed, simulations (forecasts) of power generally are derived from simulations (forecasts) of speed.

In their simplest form, wind speed simulations may be based on Monte Carlo methods that rely solely on the estimated parameters of the marginal distribution of wind speed (e.g., the mean and variance in the case of the Gaussian distribution), and many studies have been undertaken to fit distributions to wind speed data (e.g., Bardsley, 1980; Hennessey, 1977; Luna and Church, 1974). However, this approach ignores the positive correlations between consecutive observations of wind speed. Failure to take this autocorrelation into account leads to underestimation of the variances of time averages of wind speeds. Furthermore, the long runs of high (and low) wind speeds that are characteristic of such data do not occur frequently enough in simulated data when wind speeds are assumed to be uncorrelated over time.

Several previous studies have attempted to incorporate autocorrelation into wind speed models (e.g., Chou and Corotis, 1981; Goh and Nathan, 1979). However, the approaches taken in these studies have been based on very specific assumptions about the statistical characteristics of wind speed data. Moreover, some of the studies have neglected to consider important properties of such data, including the non-Gaussian shape of wind speed distributions.

The primary purposes of this paper are 1) to describe a general model for wind speed and wind power that takes into account the non-Gaussian

[1] Present affiliation: Environmental and Societal Impacts Group, National Center for Atmospheric Research, Boulder, CO 80307.

[2] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

distribution, the diurnal nonstationarity, and the autocorrelated nature of wind speed (Section 2); and 2) to describe how this model may be used to simulate and forecast hourly wind speed (Sections 4 and 5, respectively). As an example, the methodology is applied to a small set of hourly wind speed data (Section 3), and the model developed for this specific data set is used to illustrate the processes of simulating and forecasting hourly wind power (Sections 4 and 5, respectively).

2. Time series models for wind speed

Measurements of a meteorological variable, such as wind speed, taken over time tend to be positively correlated. To take into account this autocorrelation, we consider a class of parametric time series models called autoregressive (AR) processes (Box and Jenkins, 1976). Such processes have been employed to model many meteorological time series (e.g., see Katz and Skaggs, 1981). The modeling methodology described in this section consists of first fitting AR processes of various orders to hourly wind speed data which have been transformed to make their distribution approximately Gaussian and standardized to remove diurnal nonstationarity. Then a model selection criterion is employed to select the most appropriate AR process. Seasonal nonstationarity is removed by fitting a separate model for each month or season. An example of the exploratory data analysis employed to identify the particular forms of transformation and standardization to apply to a specific data set will be presented in Section 3.

a. Transformation and standardization

Suppose we are given a time series of hourly wind speeds at anemometer level. We let $U_A(t)$, $t = 1, 2, \ldots, 24T$, denote the wind speed for the tth hour of the record, with subscript A denoting the anemometer height above local ground level and T denoting the total number of days in the record. First, a power transformation is applied to adjust for the non-Gaussian distribution of hourly wind speeds, giving transformed data

$$U_A'(t) = [U_A(t)]^m, \quad t = 1, 2, \ldots, 24T. \tag{1}$$

Here the power m may be positive or negative with, by convention, the case of $m = 0$ corresponding to the logarithm. Several techniques are available for choosing the power m of the transformation (1).

Then the diurnal nonstationarity in hourly wind speeds is removed by subtracting the hourly expected values

$$\mu(t) = E[U_A'(t)] \tag{2}$$

and dividing by the hourly standard deviations

$$\sigma(t) = \{\text{var}[U_A'(t)]\}^{1/2}, \tag{3}$$

obtaining standardized transformed wind speeds

$$U_A^*(t) = [U_A'(t) - \mu(t)]/\sigma(t), \tag{4}$$

$t = 1, 2, \ldots, 24T$. Here it is assumed that μ and σ are periodic functions with

$$\mu_h \equiv \mu(h), \quad \sigma_h \equiv \sigma(h), \quad h = 1, 2, \ldots, 24, \tag{5}$$

$$\left.\begin{aligned} \mu(25) = \mu_1, \quad \mu(26) = \mu_2, \text{ etc.} \\ \sigma(25) = \sigma_1, \quad \sigma(26) = \sigma_2, \text{ etc.} \end{aligned}\right\}. \tag{6}$$

Several methods are available for estimating the hourly expected values μ_h and hourly standard deviations σ_h.

b. AR processes

We assume that the standardized transformed hourly wind speeds constitute a pth-order autoregressive [AR(p)] process. This model has the form

$$\left.\begin{aligned} U_A^*(t) = \sum_{k=1}^{p} \phi_k U_A^*(t-k) + \epsilon(t) \\ t = p+1, p+2, \ldots, 24T \end{aligned}\right\}. \tag{7}$$

Here $\phi_1, \phi_2, \ldots, \phi_p$ are unknown AR coefficients that need to be estimated from the data at hand. The error term process $\epsilon(t)$ in (7) is assumed to constitute a "white noise" process; that is, the $\epsilon(t)$ are uncorrelated and each has a Gaussian distribution with zero expected value and constant, but unknown, variance σ_ϵ^2. Constraints must be placed on the possible values of the AR coefficients in (7) for the $U_A^*(t)$ process to be stationary (Box and Jenkins, 1976, p. 53). For example, when dealing with an AR(1) process, the single AR parameter ϕ_1 must satisfy $|\phi_1| < 1$.

Note that an AR(p) process allows the standardized transformed wind speed for the current hour to contain information from more than just the previous p hours. For example, in an AR(2) process, (7) reduces to

$$U_A^*(t) = \phi_1 U_A^*(t-1) + \phi_2 U_A^*(t-2) + \epsilon(t). \tag{8}$$

But

$$U_A^*(t-1) = \phi_1 U_A^*(t-2) + \phi_2 U_A^*(t-3) + \epsilon(t-1), \tag{9}$$

so that $U_A^*(t)$ can be written as

$$U_A^*(t) = (\phi_1^2 + \phi_2) U_A^*(t-2) + \phi_1\phi_2 U_A^*(t-3) + \phi_1\epsilon(t-1) + \epsilon(t), \tag{10}$$

thus introducing information from the $(t-3)$th hour. This method could be applied repeatedly to demonstrate the dependence of $U_A^*(t)$ on $U_A^*(t')$ for any previous hour $t' < t$.

c. Model fitting

One method of obtaining estimates $\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p$ of the AR coefficients is the so-called Yule–Walker recursion (Box and Jenkins, 1976, p. 82). This recursive method of calculation is employed because closed-form expressions for the coefficient estimates do not exist. The Yule–Walker recursion requires only that the sample autocorrelation coefficients

$$r(k) = c(k)/c(0), \tag{11}$$

where

$$c(k) = \sum_{t=k+1}^{24T} U_A^*(t-k)U_A^*(t), \quad k = 0, 1, \ldots, p, \tag{12}$$

be calculated from the data. Other methods of estimating the AR coefficients include multiple regression analysis and Burg's algorithm (see Ulrych and Bishop, 1975, for a description of Burg's algorithm).

In the Yule–Walker recursion, the white noise variance σ_ϵ^2 is estimated by

$$\hat{\sigma}_\epsilon^2(p) = \frac{c(0)}{24T - s} \prod_{k=1}^{p} \{1 - [\hat{\phi}_k(k)]^2\}. \tag{13}$$

Here $\hat{\phi}_k(k)$ denotes the sample kth-order partial autocorrelation coefficient and is calculated from the sample autocorrelation coefficients (11). The parameter s in (13) denotes the total number of parameters required to be estimated in fitting the model. It includes the p coefficients as well as the parameters involved in estimating the diurnal expected value and standard deviation functions μ and σ. In (13), s is subtracted from the total number of observations $24T$ to make $\hat{\sigma}_\epsilon^2(p)$ an unbiased estimator of the white noise variance. The specific example presented in Section 3 will make the definition of s more clear.

A procedure for selecting the appropriate order p of the AR process also is needed. One such procedure introduced by Schwarz (1978) is called the Bayesian information criterion (BIC). It requires that all possible orders p, $p = 0, 1, \ldots, P$, be fitted to the data, where P is an upper bound on the order and $p = 0$ corresponds to an uncorrelated process. The value of p that minimizes the quantity

$$\text{BIC}(p) = 24T \ln\hat{\sigma}_\epsilon^2(p) + s \ln 24T, \quad p = 0, 1, \ldots, P \tag{14}$$

is selected. The first term on the right-hand side of (14) can be thought of as a measure of how well an AR(p) process fits the data, whereas the second term is a penalty function for the s parameters that need to be estimated. This procedure has been used to select time series models for other types of meteorological variables (Katz and Skaggs, 1981).

Alternative model selection procedures, such as Akaike's information criterion (AIC) (Akaike, 1974), could be employed instead of the BIC. The BIC technique is more parsimonious (that is, selects models with fewer parameters) than the AIC technique. However, in some cases the AIC technique may result in the selection of a model that predicts more accurately. In a closely related problem concerning choice of the order of a Markov chain, Katz (1981) compared the performance of the AIC and BIC procedures both on a theoretical basis and with regard to fitting precipitation occurrence time series.

3. Application of the model to a specific set of wind speed data

In this section, the modeling methodology described in Section 2 is illustrated by applying it to a particular set of wind speed data from the Pacific Northwest. The data consist of one month (December 1981) of hourly wind speeds recorded at Goodnoe Hills, a wind power site located in south-central Washington along the Columbia River Gorge. For convenience, these data will be referred to in the text as the "GH8112" data. The 744 hourly observations were electronically digitized from anemometer strip charts and represent hourly averages rather than instantaneous measurements.

It naturally would be very desirable to have more than a single month of data for an actual application of the modeling methodology. In fact, it would be appropriate to use several years of wind speed data for a given month to develop a model for that month. If more than one year of data for a given month were available, the statistical calculations would be slightly more complicated, requiring a revision of the formula (12) for calculating the sample autocovariance. In practice, an independent data set also should be employed to verify the forecasting ability of the model. However, since the goal of this section is simply to provide an example of the application of the modeling methodology, the data have been restricted to a single month and no attempt will be made to verify the performance of the model selected.

As mentioned in Section 2, it is necessary to transform and standardize hourly wind speed data prior to fitting AR processes in order to make their distribution approximately Gaussian and to remove diurnal nonstationarity. The forms of transformation and standardization required for this particular data set are determined in Sections 3a and 3b using some techniques of exploratory data analysis. The appropriate order of the AR process is selected and the AR parameters are estimated in Section 3c. The latter section also contains a summary of the model identified, including parameter estimates.

a. Transformation to obtain approximate Gaussian distribution

The frequency distribution of the GH8112 hourly wind speeds is shown in Fig. 1. It is evident that the

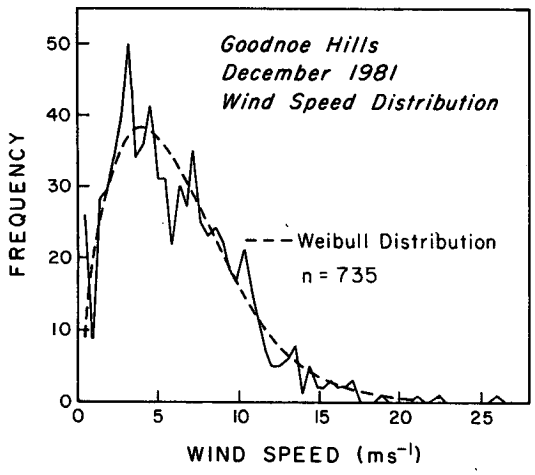

FIG. 1. Frequency distribution of GH8112 wind speeds.

distribution is positively skewed. Similar skewed distributions have been found to characterize wind speeds measured at other sites and over other time periods (Brown *et al.*, 1981).

Several non-Gaussian distributions have been suggested as appropriate models for wind speed. These models include the gamma distribution (e.g., Putnam, 1948; Sherlock, 1951); the lognormal distribution (Luna and Church, 1974); the inverse Gaussian distribution (Bardsley, 1980); the Weibull distribution (Hennessey, 1977; Justus *et al.*, 1976; Stewart and Essenwanger, 1978; Takle and Brown, 1978); and, most recently, the squared normal distribution (Carlin and Haslett, 1982). The results of studies of wind speed data from Goodnoe Hills indicate that the Weibull distribution adequately models such data (Brown *et al.*, 1981). In particular, the Weibull distribution shown in Fig. 1 fits the actual wind speed frequencies quite well.

Unfortunately, no time series models have been rigorously developed for random variables possessing a Weibull distribution. We will take an alternative approach to approximating the fitting of a Weibull distribution by, instead, applying a power transformation. Dubey (1967) showed that, for shape parameters close to 3.6, the Weibull distribution is similar in shape to a Gaussian distribution. Using this result provides one method of selecting an appropriate power transformation to make the distribution of hourly wind speeds approximately Gaussian. This approach is reasonable because a Weibull random variable raised to a power m also is a Weibull random variable. In particular, if wind speed U has a Weibull probability density function

$$f_U(u) = \frac{g}{b}\left(\frac{u}{b}\right)^{g-1} \exp[-(u/b)^g], \quad u > 0, \qquad (15)$$

where g and b are the shape and scale parameters, respectively, then U^m has a Weibull distribution with shape parameter g/m and scale parameter b^m. In order to find an appropriate power m for the transformation (1), it is only necessary to solve the equation $g/m = 3.6$, or

$$m = g/3.6. \qquad (16)$$

The maximum likelihood estimate of g for the Weibull distribution shown in Fig. 1 is 1.63. Substituting this value into (16), one choice for the value of m is 0.45, or nearly $\frac{1}{2}$.

Another approach to selecting a power transformation to obtain an approximately Gaussian distribution has been suggested by Hinkley (1977). Hinkley's method requires the transformation of the data using several values of m (e.g., 1, $\frac{1}{2}$, $\frac{1}{4}$, 0). Then a statistic

$$d_m = \frac{\text{mean} - \text{median}}{\text{scale}}, \qquad (17)$$

which is a measure of the symmetry of a distribution, is calculated for each set of transformed data. Scale parameters that may be used in (17) include the standard deviation and interquartile range (IQ). Since the mean equals the median for a symmetric distribution, the optimal value of d_m is zero. By comparing and interpolating between the values of d_m, the value of m for which d_m is closest to zero can be determined approximately. Figure 2 shows several values of d_m calculated (using IQ as the measure of scale) for the GH8112 hourly wind speed data. From this figure, it appears that the best choice of m (the m for which $d_m = 0$) is slightly greater than $\frac{1}{2}$.

If consideration is given only to common transformations [such as the square root ($m = \frac{1}{2}$), fourth root ($m = \frac{1}{4}$), or logarithm ($m = 0$)], then these results indicate that the square root is the most appropriate choice of transformation for making the distribution of these data approximately Gaussian. The use of this transformation is equivalent to the assumption that the squared normal distribution is an adequate model for the distribution of wind speed, as has been suggested by Carlin and Haslett (1982). The square root also has been found to be an appropriate trans-

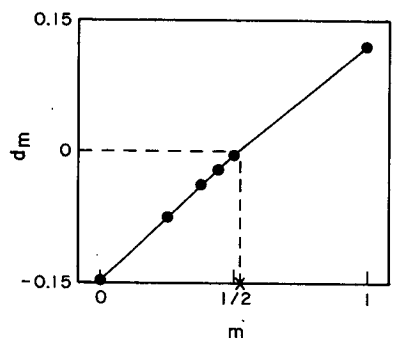

FIG. 2. Values of d_m calculated from GH8112 wind speeds (m = 0, ¼, ⅜, 7/16, ½, 1).

formation for wind speeds measured over other time periods at Goodnoe Hills and other locations in the Pacific Northwest (Brown *et al.*, 1982). For other wind speed data, Essenwanger (1959) showed that the square root or cube root transformation gives an approximately normal distribution.

Figure 3 shows Weibull and squared normal probability plots for the GH8112 wind speed data. A straight line corresponding closely to the 45° line in this figure should be obtained if the choice of theoretical distribution is correct. The Weibull and squared normal curves in Fig. 3 are quite similar to one another and both approximate the 45° line reasonably well. Thus, the square root ($m = \frac{1}{2}$) is selected to transform these data.

b. *Standardization to remove diurnal nonstationarity*

As mentioned in Section 2a, the hourly expected values μ_h and hourly standard deviations σ_h ($h = 1, 2, \ldots, 24$) may be estimated using several different approaches. Here, for simplicity, we will consider only the case where μ_h is estimated by the sample mean of all square-root-transformed wind speeds measured in the hth hour of the day and σ_h is estimated by the sample standard deviation of all square-root-transformed wind speeds measured in the hth hour. A more parsimonious approach (i.e., requiring the estimation of fewer parameters) would consist of using Fourier series to model the functions μ and σ (e.g., Panofsky and Brier, 1958).

The diurnal means and standard deviations of the square-root-transformed GH8112 hourly wind speeds are shown in Fig. 4. These values have been smoothed using a three-point algorithm called hanning [equivalent to applying a second-order moving average twice (Tukey, 1977)] to make any diurnal variations

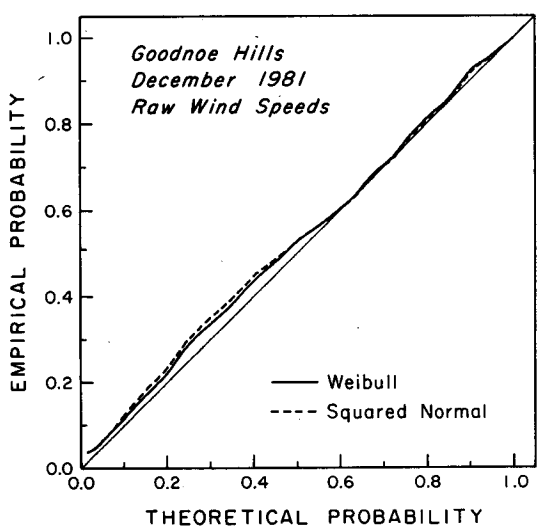

FIG. 3. Probability (*P–P*) plot of GH8112 wind speed distributions, with 45° line included for comparative purposes (see text).

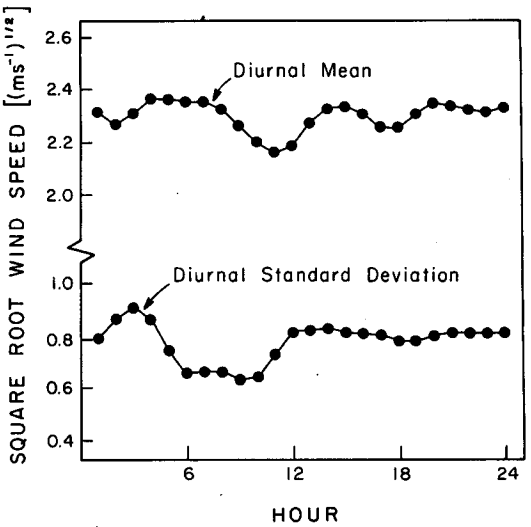

FIG. 4. Diurnal means and standard deviations of square-root-transformed GH8112 wind speeds.

that might be present more apparent. The patterns of the means and standard deviations in Fig. 4 suggest that some diurnal variability is present in the wind speeds. However, the ranges of the values are quite small [0.26 $(\text{m s}^{-1})^{1/2}$ for the diurnal means], suggesting that diurnal nonstationarity probably is not a problem with this very small data set. Examination of the autocorrelation function for the nonstandardized square root wind speed data confirmed that the transformed data do not contain important diurnal variations.

Previous evaluations of longer series of hourly wind speed data from Goodnoe Hills have indicated the need for standardization (Brown *et al.*, 1982). This standardization has been accomplished simply by subtracting the hourly square root means [i.e., (4) becomes $U_A^*(t) = U_A'(t) - \mu(t)$]. For the purposes of this illustration, standardization by subtracting the actual (not smoothed) diurnal means was selected in order to be consistent with previous analyses and to present an example of a model that allows for diurnal variation.

c. *Time series model identification*

Calculated values of $\hat{\sigma}_\epsilon^2(p)$ and BIC(p) for the GH8112 standardized transformed wind speed data for $p = 0, 1, \ldots, 10$ are listed in Table 1 and the sample autocorrelation function is presented in Fig. 5. The order p for which BIC is minimized is two. The estimated white noise variance $\hat{\sigma}_\epsilon^2(p)$ also is minimized for $p = 2$, whereas $\hat{\sigma}_\epsilon^2(1)$ is only slightly larger.

The desirability of using an AR(2) model rather than an AR(1) model can be evaluated in two ways. First, note that along with the sample autocorrelation function, Fig. 5 contains the theoretical autocorrela-

TABLE 1. AR process order selection for GH8112 standardized transformed wind speeds using BIC with $P = 10$.

p	$\hat{\sigma}_\epsilon^2(p)$ (m s^{-1})	BIC(p)
0	0.660	−150.45
1	0.126	−1375.88
2	0.119	−1411.79*
3	0.119	−1405.18
4	0.119	−1398.56
5	0.119	−1391.95
6	0.119	−1385.34
7	0.119	−1378.73
8	0.119	−1372.12
9	0.119	−1365.50
10	0.119	−1358.89

* Minimum.

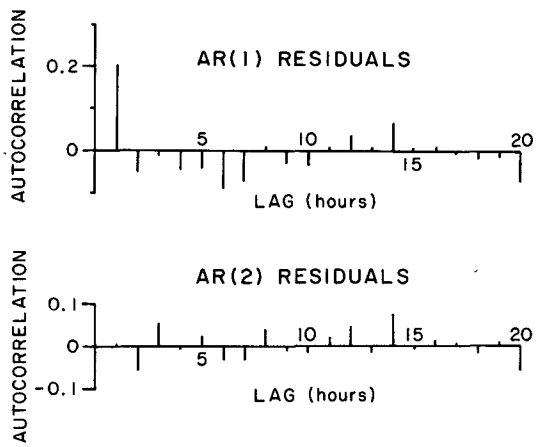

FIG. 6. Autocorrelations of AR(1) and AR(2) residuals for GH8112 data.

tion functions for AR(1) and AR(2) processes. It is evident that the sample autocorrelation function is more closely approximated by the AR(2) process than by the AR(1) process. Second, the sample autocorrelation functions for the residuals of the two fitted models can be examined. If the model selected has adequately taken into account the autocorrelated nature of the wind speeds, then the residuals should have the characteristics of a white noise process according to the model defined in (7). In particular, they should be uncorrelated. The sample autocorrelation functions for the AR(1) and AR(2) residuals are shown in Fig. 6. The pattern of the autocorrelations for the AR(1) residuals indicates that some autocorrelation has not been removed by the AR(1) process. In fact, the sample first-order autocorrelation coefficient of 0.20 for the AR(1) residuals is 5.5 times larger than its estimated large sample standard error (Box and Jenkins, 1976, p. 35), indicating that some autocorrelation has not been taken into account by the AR(1) process. In contrast, the autocorrelations of the AR(2) residuals appear to be random, and all are (in absolute value) less than two times the estimated large sample standard error. Hence, an AR(2) process is selected to model these data.

The estimated AR(2) process parameters for the GH8112 data, calculated using the Yule–Walker recursion, are $\hat{\phi}_1 = 1.1044$ and $\hat{\phi}_2 = -0.2273$. The value of s for this model is 26 (24 estimated hourly means plus two estimated AR coefficients). The final model is summarized in Table 2 which also contains estimates of the parameters μ_h ($h = 1, 2, \ldots, 24$), ϕ_1, ϕ_2 and σ_ϵ^2.

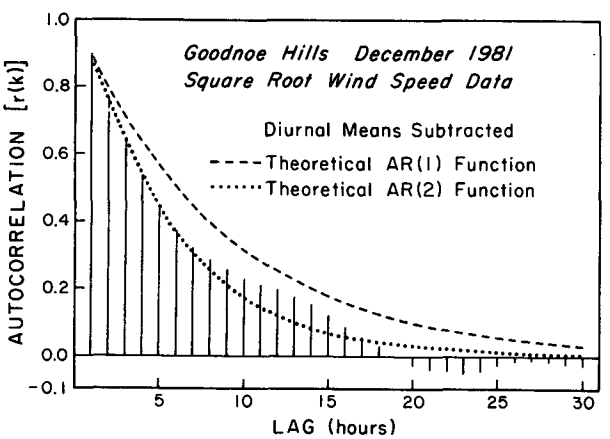

FIG. 5. Autocorrelations of GH8112 square-root-transformed wind speeds standardized by subtracting hourly means.

TABLE 2. Summary of the model and parameter estimates for the GH8112 data.

I. The model:

(a) $U_A(t) \equiv$ wind speed for hour t at anemometer height A,

(b) $U'_A(t) \equiv [U_A(t)]^{1/2}$,

(c) $U^*_A(t) \equiv U'_A(t) - \mu(t)$,

(d) $U^*_A(t) = \phi_1 U^*_A(t-1) + \phi_2 U^*_A(t-2) + \epsilon(t)$,

where the $\epsilon(t)$ are uncorrelated and identically distributed Gaussian random variables with zero expected value and unknown variance σ_ϵ^2.

II. Parameter estimates:

h	$\hat{\mu}_h$ (m s^{-1})$^{1/2}$	h	$\hat{\mu}_h$ (m s^{-1})$^{1/2}$	h	$\hat{\mu}_h$ (m s^{-1})$^{1/2}$
1	2.33	9	2.27	17	2.26
2	2.25	10	2.21	18	2.23
3	2.31	11	2.15	19	2.33
4	2.41	12	2.18	20	2.39
5	2.37	13	2.30	21	2.34
6	2.35	14	2.36	22	2.35
7	2.39	15	2.35	23	2.31
8	2.33	16	2.34	24	2.34

$\hat{\phi}_1 = 1.1044$, $\hat{\phi}_2 = -0.2273$, $\hat{\sigma}_\epsilon^2 = 0.119$ m s^{-1}.

4. Simulation of wind power

The generation of simulated time series of hourly wind power output involves several steps and is discussed in this section. First, the time series model for wind speeds is employed to generate a time series of simulated hourly wind speeds at anemometer level; second, the anemometer-level speeds are converted to wind turbine hub-height speeds; and third, the hub-height wind speeds are converted to a time series of hourly wind power output. To make the presentation more concrete, the algorithm for simulated wind speeds is presented in terms of the specific form of model fitted to the GH8112 data in Section 3. As an example, a simulated sequence of hourly wind power output is generated using the parameter estimates obtained for the GH8112 data.

a. Simulation of wind speed

First, a sequence of independent random variables, denoted by $Z(1), Z(2), \ldots$, having a Gaussian distribution with zero mean and unit variance $[N(0, 1)]$ is generated. We note that most computers have software available so that this sequence may be obtained by repeated use of an existing subroutine or function. The simulated hourly wind speeds at anemometer level are denoted by $\tilde{U}_A(1), \tilde{U}_A(2), \ldots$, where the tildes distinguish them from the actual hourly wind speed observations. These simulated speeds are obtained as follows:

1) The simulated wind speed for the first hour, $\tilde{U}_A(1)$, is determined by

$$\tilde{U}'_A(1) = \hat{\mu}(1) + \hat{\sigma}Z(1), \tag{18}$$

$$\tilde{U}_A(1) = [\tilde{U}'_A(1)]^2, \tag{19}$$

where

$$\hat{\sigma}^2 = \frac{c(0)}{24(T-1)} \tag{20}$$

is the estimated variance of the $U^*_A(t)$-process.

2) The simulated wind speed for the second hour, $\tilde{U}_A(2)$, given $\tilde{U}_A(1)$, is determined by

$$\tilde{U}'_A(2) = \hat{\mu}(2) + r(1)[\tilde{U}'_A(1) - \hat{\mu}(1)] + \hat{\sigma}\{1 - [r(1)]^2\}^{1/2}Z(2), \tag{21}$$

$$\tilde{U}_A(2) = [\tilde{U}'_A(2)]^2, \tag{22}$$

where $r(1)$ is the sample first-order autocorrelation coefficient (11).

3) In general, the simulated wind speed for the tth hour, $\tilde{U}_A(t)$, given the simulated wind speeds for the $(t-1)$th and $(t-2)$th hours, $\tilde{U}_A(t-1)$ and $\tilde{U}_A(t-2)$, is determined by

$$\tilde{U}'_A(t) = \hat{\mu}(t) + \hat{\phi}_1[\tilde{U}'_A(t-1) - \hat{\mu}(t-1)] + \hat{\phi}_2[\tilde{U}'_A(t-2) - \hat{\mu}(t-2)] + \hat{\sigma}_\epsilon Z(t), \tag{23}$$

$$\tilde{U}_A(t) = [\tilde{U}'_A(t)]^2. \tag{24}$$

In Step 3, (23) produces autocorrelation among consecutive wind speeds by means of an AR(2) process. It also introduces diurnal nonstationarity in the hourly expected values. The square transformation (24), the inverse of the square root transformation, converts variables having a Gaussian distribution into variables having a squared normal distribution. Steps 1 and 2 are necessary to initialize the simulation process. For the GH8112 data, $\hat{\sigma}$ used in (18) and (21) equals 0.812 $(\text{m s}^{-1})^{1/2}$ and $r(1)$ used in (21) equals 0.900.

b. Conversion of simulated speed to simulated power

We assume that, applying the method presented in Section 4a, a sequence of simulated hourly wind speeds at anemometer level, $\tilde{U}_A(1), \tilde{U}_A(2), \ldots$, is now available. Next, these simulated speeds are converted to simulated speeds at wind turbine hub height H, $\tilde{U}_H(1), \tilde{U}_H(2), \ldots$, using some type of extrapolation technique such as the power law

$$\tilde{U}_H(t) = [\tilde{U}_A(t)](H/A)^\alpha. \tag{25}$$

Here A equals 13.7 m and H equals 61.0 m at Goodnoe Hills. Studies of low-level wind profiles at this site indicate that the power law (25) provides a reasonable first approximation to the change of wind speed with height under most meteorological conditions.

In general, if the appropriate measurements are available, the parameter α in (25) should be allowed to vary diurnally and seasonally. Such detailed measurements were not available for the particular Goodnoe Hills power site considered in this example. Based on limited data, α was estimated to be about 0.085 for December at Goodnoe Hills. It should be noted that other techniques for extrapolating from anemometer-height wind speed to hub-height wind speed (such as the log law) may be more appropriate than the power law in certain situations. Moreover, if at all possible, wind speed measurements made at hub height should be used in developing the wind speed model so that the uncertainties inherent in the conversion of anemometer-height wind speed to hub-height wind speed can be avoided.

We assume that, by applying (25), a sequence of simulated hourly wind speeds at hub height, $\tilde{U}_H(1), \tilde{U}_H(2), \ldots$, is now available. As the final step in the simulation procedure, these wind speeds at hub height need to be converted to wind power output. This conversion from hub-height wind speed $\tilde{U}_H(t)$ for the tth hour to wind power output $\tilde{P}(t)$ for the tth hour is accomplished using a wind generator performance curve, adjusted for the elevation at the location of the generator. The wind generator performance curve for the MOD-2 wind generator at Goodnoe Hills is presented in Fig. 7. This curve indicates that power output is zero for wind speeds less than 5.8 m s^{-1} and then increases nonlinearly as the wind speed

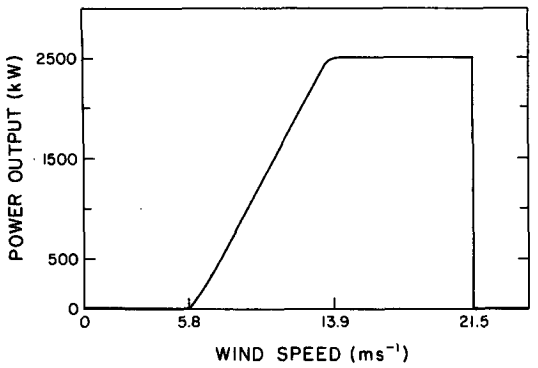

FIG. 7. Wind generator performance curve for MOD-2 wind generator at Goodnoe Hills, Washington.

increases from 5.8 to 13.9 m s^{-1}. Maximum power output is associated with wind speeds between 13.9 and 21.5 m s^{-1}, whereas no power output occurs for wind speeds greater than 21.5 m s^{-1}.

Note that the important short-term fluctuations of wind power output cannot be simulated using hourly wind speed data and this performance curve. In order to incorporate these short-term fluctuations into the simulations, the model could be developed using wind speed data consisting of averages over shorter periods (e.g., 10 minutes) if such data were available. Averages over such shorter time periods could have been electronically digitized from the anemometer strip charts used in this study. However, the averaging period was restricted to one hour in order to limit the amount of time required for digitizing.

c. *Example*

As an example of how the model can be used for simulations, the parameter values estimated for the GH8112 data are used to generate a short sequence of simulated hourly wind power output. First, a sequence of 24 independent $N(0, 1)$ random numbers, $Z(1), Z(2), \ldots, Z(24)$, was generated. These random numbers are listed in the second column of Table 3. Next, the simulated hourly wind speeds at anemometer level were obtained by applying (18), (19) and (21)–(24) using the parameter estimates given in Table 2. These simulated wind speeds at anemometer level, $\tilde{U}_A(1), \tilde{U}_A(2), \ldots, \tilde{U}_A(24)$, are listed in the third column of Table 3. Then the simulated hourly wind speeds at wind turbine hub height were computed by the power law (25). These simulated hourly wind speeds, $\tilde{U}_H(1), \tilde{U}_H(2), \ldots, \tilde{U}_H(24)$, are listed in the fourth column of Table 3. Finally, the simulated hourly wind power values were calculated by means of the wind generator performance curve (Fig. 7). These simulated hourly power values, $\tilde{P}(1), \tilde{P}(2), \ldots, \tilde{P}(24)$, are listed in the last column of Table 3.

It is important to note that a sufficiently long simulated sequence of wind speeds generated using the model described here would, by definition, share the important statistical properties of the original wind speed data. For example, the wind speed values in a long simulation based on the model developed using the GH8112 data would necessarily have a squared normal distribution (with the same hourly means and the same variance as the original data) due to the square transformation in the simulation process. Moreover, the simulated values $\tilde{U}'_A(t) - \hat{\mu}(t)$ would form an AR(2) process with parameters equivalent to those estimated using the GH8112 standardized transformed wind speeds, implying that the first- and second-order autocorrelation coefficients would be the same for both the simulated and real data.

5. Forecasting of wind power

This paper has been concerned primarily with the simulation of wind speed and wind power. However, a time series model developed as described here also could be used to forecast wind speed and wind power one or a few hours in advance. Such forecasts may be quite useful in short-term operational decisions related to the integration of wind power into multiple-source energy systems. In this section, we describe how the form of time series model identified in Section 3 for the small set of Goodnoe Hills data could be used as a forecasting tool (as opposed to a simulation tool), and we present an illustration of this application. A direct assessment of the performance of this particular model will not be considered

TABLE 3. Simulated sequence of hourly wind power output.

t	$Z(t)$	$\tilde{U}_A(t)$ (m s^{-1})	$\tilde{U}_H(t)$ (m s^{-1})	$\tilde{P}(t)$ (kW)
1	0.666	8.22	9.34	1090
2	−0.129	7.26	8.25	720
3	−0.437	6.40	7.26	415
4	0.515	7.48	8.49	800
5	0.825	8.78	9.96	1290
6	0.528	9.70	11.01	1650
7	0.781	11.31	12.84	2245
8	−0.679	9.02	10.24	1390
9	−0.220	7.34	8.33	750
10	−0.512	5.63	6.39	175
11	0.843	6.32	7.17	380
12	2.137	10.80	12.27	2075
13	−0.102	11.59	13.16	2365
14	−1.532	7.83	8.89	950
15	−1.860	3.75	4.26	0
16	−0.787	2.30	2.61	0
17	−0.625	1.51	1.51	0
18	0.428	2.04	2.32	0
19	0.530	3.46	3.93	0
20	−0.153	3.98	4.52	0
21	1.998	7.32	8.31	735
22	1.318	10.86	12.33	2090
23	−0.801	8.97	10.18	1370
24	0.842	10.08	11.44	1780

in this methodological discussion, although, as noted in Section 3, it is desirable in practice to verify a forecasting model using an independent data set.

a. Point forecasts

Like the process of simulating wind power described in Section 4, the process of forecasting wind power one or more hours in advance involves several steps. These steps include determining the forecast value of the standardized transformed wind speed, converting that value to a forecast wind speed, and transforming the forecast wind speed to a forecast wind power value. To start the process, assume that we have a new set of t wind speed observations, $U_A(1)$, $U_A(2)$, ..., $U_A(t)$, with corresponding standardized transformed wind speed values, $U_A^*(1)$, $U_A^*(2)$, ..., $U_A^*(t)$, and that starting at the tth hour we wish to forecast the wind speed and wind power l hours ahead [i.e., for the $(t + l)$th hour].

Given the AR(2) model selected in Section 3c using the GH8112 data and the observed values $U_A^*(1)$, $U_A^*(2)$, ..., $U_A^*(t)$, the l-hours ahead forecast of the standardized transformed wind speed $\hat{U}_A^*(t + l)$ can be expressed as

$$\hat{U}_A^*(t + l) = \hat{\phi}_1 \hat{U}_A^*(t + l - 1) + \hat{\phi}_2 \hat{U}_A^*(t + l - 2), \quad l = 1, 2, \cdots. \tag{26}$$

Here $\hat{U}_A^*(t + l - 1)$ and $\hat{U}_A^*(t + l - 2)$ are the $(l - 1)$-hours and $(l - 2)$-hours ahead forecasts, respectively, with the conventions that $\hat{U}_A^*(t) = U_A^*(t)$ and $\hat{U}_A^*(t - 1) = U_A^*(t - 1)$. To convert $\hat{U}_A^*(t + l)$ in (26) into the wind speed forecast $\hat{U}_A(t + l)$, the relationships

$$\hat{U}_A'(t + l) = \hat{U}_A^*(t + l) + \hat{\mu}(t + l), \tag{27}$$

$$\hat{U}_A(t + l) = [\hat{U}_A'(t + l)]^2 \tag{28}$$

are used (see Part I of Table 2). Note that $\hat{U}_A^*(t + l) \to 0$ as $l \to \infty$ (i.e., the forecast converges to the hourly medians at large lead times). It should be evident that the process of producing wind speed and wind power forecasts is quite similar to the process of producing wind speed and wind power simulations. In fact, comparing (23) and (26), a simulated value of standardized transformed wind speed can be regarded as a one-hour ahead forecast to which a random error has been added.

A forecast wind speed $\hat{U}_A(t + l)$ is transformed into a forecast wind power value $\hat{P}(t + l)$ using the procedures described in Section 4b for transforming simulated wind speeds into simulated power values. This process involves applying the best method available to translate the forecast anemometer-level wind speed into a forecast hub-height wind speed and then using a wind generator performance curve to convert the hub-height wind speed into a forecast wind power value. For the GH8112 data used as an example in this paper, the power law (25) is used to translate the forecast anemometer-level wind speed into hub-height wind speed, and the performance curve shown in Fig. 7 is used to convert the forecast hub-height wind speed into a forecast power value.

For convenience, we consider first the case of a one-hour ahead forecast (i.e., $l = 1$). In this case, the expressions (26), (27) and (28) can be written as

$$\hat{U}_A^*(t + 1) = \hat{\phi}_1 U_A^*(t) + \hat{\phi}_2 U_A^*(t - 1), \tag{29}$$

$$\hat{U}_A'(t + 1) = \hat{U}_A^*(t + 1) + \hat{\mu}(t + 1), \tag{30}$$

$$\hat{U}_A(t + 1) = [\hat{U}_A'(t + 1)]^2, \tag{31}$$

respectively. When the values of the AR(2) parameters for the GH8112 data are introduced into (29), the one-hour ahead forecast becomes

$$\hat{U}_A^*(t + 1) = 1.1044 U_A^*(t) - 0.2273 U_A^*(t - 1). \tag{32}$$

As a specific example of a one-hour ahead forecast, we take $U_A(1) = 8.0$ m s^{-1} and $U_A(2) = 8.9$ m s^{-1} (note that $t = 2$ in this example). Then $U_A^*(1) = 0.51$ and $U_A^*(2) = 0.74$, and from (32), $\hat{U}_A^*(3) = 0.70$. Applying (30) and (31) and noting that $\hat{\mu}(3) = 2.31$ (m s^{-1})$^{1/2}$ (see Table 2), the one-hour ahead wind speed forecast is $\hat{U}_A(3) = 9.1$ m s^{-1}. Finally, using the power law (25) and the wind generator performance curve (Fig. 7), the one-hour ahead power forecast is $\hat{P}(3) = 1402$ kW. As noted earlier in this section, an AR(2) process can be used to forecast l hours in advance by means of (26). To illustrate this use of the model, wind speed and wind power forecasts for $l = 1, 2, \ldots, 6$ hours for this example are presented in Table 4.

b. Prediction intervals and probability forecasts

In addition to the point forecasts of wind speed and wind power such as those presented in Table 4, it is important to have estimates of their uncertainty. Prediction intervals provide such estimates and these intervals can be determined from the variance of the l-hours ahead standardized transformed wind speed forecast $\hat{U}_A^*(t + l)$. This variance, $V(l)$ say, is expressed as $V(1) = \sigma_\epsilon^2$ and

TABLE 4. Wind speed forecasts $\hat{U}_A(2 + l)$ and wind power forecasts $\hat{P}(2 + l)$ for $l = 1, 2, \ldots, 6$ based on the AR(2) model, for the example considered in Section 5a [$U_A(1) = 8.0$ m s^{-1}, $U_A(2) = 8.9$ m s^{-1}].

Lead time l (h)	Wind speed $\hat{U}_A(2 + l)$ (m s^{-1})	Wind power $\hat{P}(2 + l)$ (kW)
1	9.1	1402
2	10.2	1858
3	9.4	1560
4	8.7	1270
5	8.3	1125
6	7.6	825

$$V(l) = [1 + \sum_{j=1}^{l-1} \psi_j^2]\sigma_\epsilon^2, \; l = 2, 3, \ldots, \qquad (33)$$

where $\psi_j = \psi_{j-2}\phi_2 + \psi_{j-1}\phi_1$ with the conventions that $\psi_{-1} = 0$ and $\psi_0 = 1$ (Box and Jenkins, 1976, p. 156). Note that $V(l) \rightarrow \sigma^2$, the variance of the $U_A^*(t)$-process, as $l \rightarrow \infty$. Because the forecast variance $V(l)$ depends on the unknown parameters ϕ_1, ϕ_2 and σ_ϵ^2, it must be estimated. The estimated forecast variance, denoted by $\hat{V}(l)$, is obtained by replacing these unknown parameters by their estimates $\hat{\phi}_1$, $\hat{\phi}_2$ and $\hat{\sigma}_\epsilon^2$. The approximate $100(1 - \alpha)\%$ prediction interval for $\hat{U}_A^*(t + 1)$ is then defined by the interval

$$\{\hat{U}_A^*(t + l) - Z(\alpha/2)[\hat{V}(l)]^{1/2}, \; \hat{U}_A^*(t + l) + Z(\alpha/2)[\hat{V}(l)]^{1/2}\}, \qquad (34)$$

in which $Z(\alpha/2)$ is the deviate that is exceeded by $100(\alpha/2)\%$ of the $N(0, 1)$ distribution.

The approximate 75% prediction intervals for the l-hours ahead forecasts of the standardized transformed wind speeds $\hat{U}_A^*(2 + l)$ for the example considered in Section 5a are shown in Fig. 8. Specifically, this figure contains the point forecasts of $\hat{U}_A^*(2 + l)$ for $l = 1, 2, \ldots, 6$ hours, and the upper and lower limits of the corresponding prediction intervals. Note that these prediction intervals increase in width as lead time l increases, approaching the width of unconditional climatological intervals for longer lead times.

Prediction intervals for $\hat{U}_A^*(t + l)$ can be translated into prediction intervals for wind speed $\hat{U}_A(t + l)$ by applying (27) and (28) to the prediction limits for $\hat{U}_A^*(t + l)$. Then the prediction limits for $\hat{U}_A(t + l)$ can be converted into prediction limits for wind power $\hat{P}(t + l)$ using, for the Goodnoe Hills example, the power law (25) and the wind generator performance curve (Fig. 7). The nonlinear transformations involved in converting $\hat{U}_A^*(t + l)$ into $\hat{U}_A(t + l)$ and $\hat{P}(t + l)$ sometimes can lead to quite unusual prediction intervals for wind power. For example, consider the case in which $l = 1$, $U_A(1) = 17$ m s^{-1} and $U_A(2) = 18$ m s^{-1}. In this case, applying the model developed using the GH8112 data, an approximate 95% prediction interval for $\hat{P}(3)$ includes the range of values from 2440 to 2500 kW and also the value 0 kW. The latter value is included in the prediction interval for wind power because the 95% prediction interval for wind speed includes values greater than 21.5 m s^{-1}, the cut-out speed of the wind power generator (see Fig. 7).

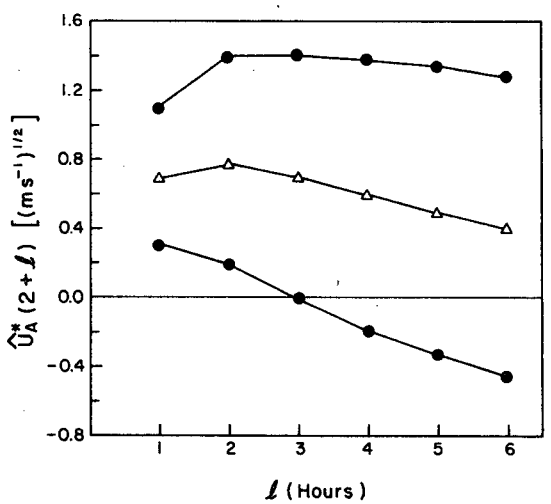

FIG. 8. 75% prediction intervals for standardized transformed wind speed $\hat{U}_A^*(2 + l)$ ($l = 1, 2, \ldots, 6$) for example considered in Section 5a [$U_A(1) = 8.0$ m s^{-1}, $U_A(2) = 8.9$ m s^{-1}]. The triangles represent the point forecasts $\hat{U}_A^*(2 + l)$ and the closed circles represent the upper and lower bounds of the prediction intervals.

TABLE 5. Approximate 75% prediction intervals for wind speed forecasts $\hat{U}_A(2 + l)$ and wind power forecasts $\hat{P}(2 + l)$ for $l = 1, 2, \ldots, 6$ based on AR(2) model for the example considered in Section 5a [$U_A(1) = 8.0$ m s^{-1}, $U_A(2) = 8.9$ m s^{-1}].

Lead time l (h)	75% prediction intervals: Wind speed (m s^{-1})	75% prediction intervals: Wind power (kW)
1	(6.8, 11.6)	(535, 2350)
2	(6.7, 14.3)	(530, 2500)
3	(5.6, 14.3)	(160, 2500)
4	(4.6, 14.0)	(0, 2500)
5	(4.2, 13.9)	(0, 2500)
6	(3.5, 13.1)	(0, 2500)

The approximate 75% prediction intervals for $\hat{U}_A(2 + l)$ and $\hat{P}(2 + l)$ corresponding to the intervals for $\hat{U}_A^*(2 + l)$ depicted in Fig. 8 are presented in Table 5. The widths of these intervals are quite large by the third hour of the forecast. Moreover, the prediction intervals for the power forecasts include the entire range of possible power values by the fourth hour. This rapid degradation in precision for longer lead times suggests that this particular model (for Goodnoe Hills in December) would be useful only for very short-range (i.e., 1 to 3 hour) forecasts of power output. This result probably also would characterize models of this type developed for other sites and/or time periods. Because the AR process relies only on the previous record of wind speeds, it would not be expected to have much skill in forecasting more than a few hours in advance. Other methods based on physical models and which consider information in addition to wind speed observations could possibly be used to forecast wind power for longer lead times.

As an alternative to the prediction intervals for wind power described in the previous paragraphs, probabilities of future wind power values can be estimated from the conditional probability distribution

of the standardized transformed wind speed $U_A^*(t + l)$. Specifically, the assumptions made in defining an AR process (Section 2b) imply that the conditional distribution of a future value $U_A^*(t + l)$, given $U_A^*(1), U_A^*(2), \ldots, U_A^*(t)$, is approximately $N[\hat{U}_A^*(t + l), \hat{V}(l)]$. From this distribution and the relationship between $U_A^*(t + l)$ and $P(t + l)$, it is possible to compute probabilities for specific wind power events (i.e., individual power values or ranges of power values). These probabilities could be very useful in such applications as integrating wind power output into an energy array, where it is of particular importance to know, for example, the likelihood of no wind power output, some wind power output or maximum possible power output in a given hour.

The probability forecasts of three wind power events—no power [$P(2 + l) = 0$ kW], some power [$0 < P(2 + l) < 2500$ kW] and full power [$P(2 + l) = 2500$ kW]—are presented in Table 6 for the example considered in Section 5a. In examining these probabilities, it is of interest to note that for the GH8112 data, the observed relative frequencies (or climatological probabilities) of no power, some power and full power are 0.493, 0.453 and 0.054, respectively. For $l = 1$ (one hour ahead), the probability forecast of some power is very high (0.885), nearly twice as large as the observed relative frequency of some power, and the probability forecast of no power is very low (0.021), only about 4% of the observed relative frequency of no power. As lead time increases, the probability forecast of some power decreases and the probability forecast of no power increases. The forecast probability of full power initially increases from a modest value (0.093 for $l = 1$) to a fairly large value (0.287 for $l = 2$), more than five times as large as the observed relative frequency of full power, but then decreases for longer lead times. These probability forecasts are necessarily consistent with the wind power forecasts $\hat{P}(2 + l)$ in Table 4. Of course, these forecast probabilities of wind power depend on the initial wind speed measurements [$U_A(1)$ and $U_A(2)$] and would differ with different initial conditions.

TABLE 6. Probability forecasts of no power, some power and full power based on AR(2) model for the example considered in Section 5a [$U_A(1) = 8.0$ m s^{-1}, $U_A(2) = 8.9$ m s^{-1}].

	Estimated probabilities		
Lead time l (h)	No power $\Pr\{P(2 + l) = 0$ kW$\}$	Some power $\Pr\{0 < P(2 + l) < 2500$ kW$\}$	Full power $\Pr\{P(2 + l) = 2500$ kW$\}$
1	0.021	0.885	0.093
2	0.056	0.657	0.287
3	0.129	0.628	0.243
4	0.196	0.600	0.204
5	0.236	0.573	0.191
6	0.296	0.551	0.153

6. Discussion and conclusion

This paper has described a methodology for modeling wind speed and wind power that takes into account several characteristics of wind speed; namely, autocorrelation, non-Gaussian distribution and diurnal nonstationarity. Time series models of wind speed developed using this methodology have many potential uses, including the simulation and very short-range forecasting of wind speed and wind power values at a single site, as described in Sections 4 and 5. Such models should be valuable in the process of integrating the power from wind generators into large energy networks.

As an example, the modeling methodology was applied to one month (December 1981) of hourly wind speed data from Goodnoe Hills, Washington. Exploratory analysis of the GH8112 data (Section 3) indicated that application of the square-root transformation results in a distribution that is approximately Gaussian. Diurnal nonstationarity did not seem to be a problem with this small data set. However, for the purpose of illustration, the data were standardized to remove diurnal nonstationarity by subtracting the appropriate diurnal means of the transformed data from each transformed observation. Using an objective model selection procedure, an AR(2) process was found to be the most appropriate time series model for this set of data.

Similar analyses could be performed to develop time series models of wind speed and wind power for other time periods and/or for other locations. It is anticipated that such models would share the general characteristics of the model developed using the GH8112 wind speed data, but that some of the specific details of the models would differ. On the one hand, analyses of many series of hourly wind speed data from the Pacific Northwest have indicated that the square root frequently is an appropriate transformation for making the distribution of the data approximately Gaussian (Brown *et al.*, 1982). On the other hand, it may be found in some cases that an AR(1) process is more appropriate than an AR(2) process for modeling standardized transformed wind speeds. The modeling methodology described in this paper is general in the sense that it allows for the appropriate forms of transformation, standardization and AR process to be determined individually for each location and time period.

As noted in Section 3, it would be desirable in practice to use several years of data to develop the type of model described in this paper. The primary advantage of including more than one year of data in the model development is the increased reliability of the model parameter estimates. In particular, a wind speed model based on many years of January data should be more representative of future January wind speeds than a model based on a single year of January data.

The performance of a model developed for the purpose of forecasting wind power output one or more hours in advance should be evaluated using an independent data set before the model is applied in practice. As suggested in Section 5 for the GH8112 model, this type of model probably is only useful for forecasting wind power output one or two hours in advance. It may be desirable to combine this type of model with other models (e.g., numerical model output statistics) to obtain improved forecasts and/or forecasts for longer lead times.

The modeling methodology described in this paper could easily be used to develop time series models based on wind speeds averaged over periods shorter than one hour. For example, if an adequate amount of 10-minute wind speed data were available, it would be possible to develop a model that would describe the important shorter-term fluctuations of wind power. Such a model would more realistically simulate the time-dependence of the power output of present-day wind power generators.

The time series models described in this paper are appropriate for modeling the power output of individual wind power generators. Since it is likely in the future that most large-scale wind power generators will be located in closely spaced groups (called "wind farms"), it would be desirable to take into account the spatial correlations as well as the temporal correlations of wind speed in future wind power simulations. It should be noted that Carlin and Haslett (1982) have modeled spatial correlations of wind speed, although they ignored the autocorrelations in such data. An anticipated generalization of the model described in this paper would be the incorporation of spatial correlations into the model, perhaps by using multiple time series analysis (e.g., Granger and Newbold, 1977). Such a model would provide simulations and forecasts of wind speed and wind power that are realistic in both time and space.

Acknowledgments. Bruce A. Peterson provided much of the computer and data base management support for this project. His contribution is gratefully acknowledged. The authors also would like to thank the reviewers for their useful comments. This work was conducted at Oregon State University and was supported by the Bonneville Power Administration under Contract DE-AC79-81BP25593, M004.

REFERENCES

Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Autom. Control,* **19,** 716–723.

Bardsley, W. E., 1980: Note on the use of the inverse Gaussian distribution for wind energy applications. *J. Appl. Meteor.,* **19,** 1126–1130.

Box, G. E. P., and G. M. Jenkins, 1976: *Time Series Analysis: Forecasting and Control* (rev. ed.). Holden-Day, 575 pp.

Brown, B. G., R. W. Katz and A. H. Murphy, 1981: An evaluation of statistical distributions of wind power. *Preprints Seventh Conf. on Probability and Statistics in Atmospheric Sciences,* Monterey, Amer. Meteor. Soc., 142–147.

——, ——, —— and B. A. Peterson, 1982: Time series models for simulating hourly wind power. Rep. No. BPA 82-10, DOE/BP-154, Dept. Atmos. Sci., Oregon State University, 51 pp.

Carlin, J., and J. Haslett, 1982: The probability distribution of wind power from a dispersed array of wind turbine generators. *J. Appl. Meteor.,* **21,** 303–313.

Chou, K. C., and R. B. Corotis, 1981: Simulation of hourly wind speed and array wind power. *Sol. Energy,* **26,** 199–212.

Dubey, S. D., 1967: Normal and Weibull distributions. *Naval Res. Logistics Quart.,* **14,** 69–79.

Essenwanger, O. M., 1959: Probleme der Windstatistik. *Meteor. Rundsch.,* **12,** 37–47.

Goh, T. N., and G. K. Nathan, 1979: A statistical methodology for study of wind characteristics from a close array of stations. *Wind Eng.,* **3,** 197–206.

Granger, C. W. J., and P. Newbold, 1977: *Forecasting Economic Time Series.* Academic Press, 333 pp.

Hennessey, J. P., Jr., 1977: Some aspects of wind power statistics. *J. Appl. Meteor.,* **16,** 119–128.

Hinkley, D., 1977: On quick choice of power transformation. *Appl. Statist.,* **26,** 67–69.

Justus, C. G., W. R. Hargraves and A. Yacin, 1976: Nationwide assessment of potential output from wind-powered generators. *J. Appl. Meteor.,* **15,** 673–678.

Katz, R. W., 1981: On some criteria for estimating the order of a Markov chain. *Technometrics,* **23,** 243–249.

——, and R. H. Skaggs, 1981: On the use of autoregressive-moving average processes to model meteorological time series. *Mon. Wea. Rev.,* **109,** 479–484.

Luna, R. E., and H. W. Church, 1974: Estimation of long-term concentrations using a "universal" wind speed distribution. *J. Appl. Meteor.,* **13,** 910–916.

Panofsky, H. A., and G. W. Brier, 1958: *Some Applications of Statistics to Meteorology.* Pennsylvania State University 224 pp.

Putnam, P. C., 1948: *Power from the Wind.* Van Nostrand–Reinhold, 224 pp.

Schwarz, G., 1978: Estimating the dimension of a model. *Ann. Statist.,* **6,** 461–464.

Sherlock, R. H., 1951: Analyzing winds for frequency and duration. *On Atmospheric Pollution, Meteor. Monogr.,* No. 4, Amer. Meteor. Soc., 42–49.

Stewart, D. A., and O. M. Essenwanger, 1978: Frequency distribution of wind speed near the surface. *J. Appl. Meteor.,* **17,** 1633–1642.

Takle, E. S., and J. M. Brown, 1978: Note on the use of Weibull statistics to characterize wind speed data. *J. Appl. Meteor.,* **17,** 556–559.

Tukey, J. W., 1977: *Exploratory Data Analysis.* Addison-Wesley, 688 pp.

Ulrych, T. J., and T. N. Bishop, 1975: Maximum entropy spectral analysis and autoregressive decomposition. *Rev. Geophys. Space Phys.,* **13,** 183–200.