COMPARISON OF THE WEIBULL MODEL WITH MEASURED WIND SPEED DISTRIBUTIONS FOR STOCHASTIC WIND GENERATION

S. J. van Donk, L. E. Wagner, E. L. Skidmore, J. Tatarko

ABSTRACT. *Wind is the principal driver of the Wind Erosion Prediction System (WEPS), which is a process−based computer model for the simulation of wind−blown sediment loss from a field. WEPS generates wind using a stochastic wind generator. The objectives of this study were to improve the stochastic generation of wind speed and direction and to update the wind statistics used by the generator with statistics derived from more recent, quality−controlled data for the 48 contiguous states of the U.S. Erosive wind power density (WPD) was chosen to evaluate how well wind is generated, since it is proportional to sediment transport by wind. It is important that WPD calculated from stochastically generated data (WPD_g) closely reproduces WPD calculated from the underlying measured data (WPD_m). The commonly used two−parameter Weibull model did not fit wind speed distributions well enough for application in wind erosion models. WPD_g deviated more than 20% from WPD_m for 168 out of the 332 stations having $WPD_m > 5$ W m^{-2}. Fitting the model to the high wind speeds only, with the expectation of a better curve fit, resulted in some generated wind speeds exceeding 100 m s^{-1}, which is unacceptable. A more direct method uses the wind speed distributions themselves instead of the Weibull model that describes them. Wind speeds are then generated directly from the distributions using linear interpolation between data points. With this more robust direct approach, there was only one station (down from 168 stations) where WPD_g deviated more than 20% from WPD_m. The direct method of wind speed generation reproduces wind speeds more accurately than the Weibull model, which is important for wind erosion prediction and may be important for other applications as well.*

Keywords. *Erosive wind power density, Weibull, Wind erosion, Wind speed distribution.*

The Wind Erosion Prediction System (WEPS) is a process−based computer model for the simulation of wind−blown sediment loss from a field. It has been designated to replace the more empirical Wind Erosion Equation (WEQ) for use by the USDA−Natural Resources Conservation Service (NRCS) in the U.S. Wind is the principal driver of WEPS. However, it is not practical to use measured historical wind data with WEPS, since many wind records have missing data. Researchers may also want to simulate wind erosion for a longer period than the length of the measured data record, e.g., for 30 years, which is a typical WEPS simulation run. In addition, the measured data require much more computer disk space than wind summary statistics combined with a stochastic wind generator. Therefore, a stochastic wind generator is often more appropriate for use with WEPS than using the measured data directly.

Distributions of weather variables are needed by stochastic weather generators in order to generate data. Wind speed distributions have been described by the two−parameter Weibull model (Takle and Brown, 1978; Corotis et al., 1978; Skidmore and Tatarko, 1990), the two−parameter gamma model (Nicks and Lane, 1989), and the one−parameter Rayleigh model, which is a special case of the Weibull (Hennessey, 1977; Corotis et al., 1978). The Weibull is the most widely used model.

Erosive wind power density (WPD, W m^{-2}) is proportional to sediment transport by wind (Bagnold, 1941; Chepil, 1945; Skidmore, 1998). There exist different formulations of WPD. The definition that is used in WEPS is (Hagen et al., 1999):

$$\begin{aligned} \mathrm{WPD} &= 0.5\rho(u-u_t)u^2 \quad & u > u_t \\ \mathrm{WPD} &= 0 & u \le u_t \end{aligned} \tag{1}$$

where ρ is air density (kg m^{-3}), u is wind speed (m s^{-1}) at 10 m height, and u_t is the threshold wind speed (m s^{-1}) above which sediment starts to move. A threshold of 8 m s^{-1} is often thought of as a minimum threshold, but for less erodible fields, u_t may be 10 or 12 m s^{-1} or even higher.

Wind erosion models need accurate estimates of all wind speeds above the threshold. Estimates of wind speeds barely above the threshold may be the most important, since the same relative error in wind speed produces a larger relative error in WPD in this range than at much greater wind speeds. For instance, using u_t = 10 m s^{-1}, a 10% error in a 13 m s^{-1} wind speed produces a 73% error in WPD, whereas a 10% error in a 30 m s^{-1} wind speed produces "only" a 39% error in WPD.

The range of wind speeds that is of interest to wind erosion researchers is unique compared to that of other users of wind

Article was submitted for review in December 2003; approved for publication by the Soil & Water Division of ASAE in January 2005.

The authors are **Simon J. van Donk, ASAE Member Engineer,** Agricultural Engineer, **Larry E. Wagner, ASAE Member Engineer,** Agricultural Engineer, **Edward L. Skidmore,** Soil Scientist, and **John Tatarko,** Soil Scientist; USDA−ARS, Manhattan, Kansas. **Corresponding author:** Simon J. van Donk, USDA−ARS, 1515 College Avenue, Manhattan, KS 66502; phone: 785−537−5538; fax: 785−537−5507; e−mail: sdonk@weru.ksu.edu.

data. The wind power industry is also interested in high wind speeds, but only up to a certain value, e.g., 8 m s^{-1}, at which a wind turbine generates its maximum power (Justus et al., 1976; Hennesey, 1977). Thus, for the wind power industry, it is not critical that a stochastic wind generator exactly reproduces wind speeds beyond the design maximum power of the wind turbine. The construction industry is interested only in the maximum wind speed that a structure may experience during its lifetime (Mayne, 1979; Cook, 1982). For the prediction of evaporation and transpiration, one is interested in the entire range of wind speeds, but again, exact simulation of high wind speeds is not critical. Therefore, a stochastic wind generator that is satisfactory for these other fields may not be satisfactory for application in wind erosion prediction.

A stochastic wind generator has been developed for use with WEPS (Skidmore and Tatarko, 1990). It generated wind speeds from Weibull parameters, but they were suspect in certain locations. Another reason for revisiting this subject was the availability of an updated, quality−controlled, hourly wind database obtained from the National Climatic Data Center (NCDC). The objectives of this study were to improve the stochastic generation of wind speed and direction and to update the wind statistics used by the generator with statistics derived from more recent data from NCDC.

METHODS

Erosive wind power density (eq. 1) was chosen to evaluate how well wind is simulated (generated) because it is proportional to sediment transport by wind. In this article, u_t = 10 m s^{-1} was used, because it may represent an average threshold wind speed for agricultural fields. We used a constant value of 1.2 kg m^{-3} for ρ, although it varies with elevation and air temperature. However, for our purposes, using a constant is justified, since the principal use of WPD in this study is to compare WPD_g with WPD_m for any given station.

It is important that WPD calculated from stochastically generated data (WPD_g) closely reproduces WPD calculated from the underlying measured data (WPD_m). In this article, WPD_g is based on 30 years of generated data. This period is long enough, since WPD_g based on 200 years showed only very small differences. WPD_m is based on the length of record of the measured data, ranging from 5 to 65 years.

DATA SET

A quality−controlled hourly wind data set (TD−6421, version 1.1), including 1304 stations in the 48 contiguous states of the U.S., was obtained from NCDC. It includes data up to the year 2000 (1976 for the previous WEPS data set), providing longer data records and more recent data. The longest record is 65 years. Stations with less than 5 years of data were excluded, leaving 971 stations for use with a stochastic generator. This provided a denser network than the 673 stations previously used in WEPS.

The data set contains both Automated Surface Observing System (ASOS; Lockhart, 2000; McKee et al., 2000) data and data collected before ASOS. The ASOS data are 2 min averages and the before−ASOS data are 1 min averages. ASOS coverage has only begun recently. More than 800 stations have no ASOS data at all, and none of the stations has more than 8 years of ASOS data. Analysis of 28 stations with the longest ASOS records showed that on average for the ASOS data the mean wind speed was 3.6 m s^{-1}, WPD_m was 2.9 W m^{-2}, and the percentage of wind speeds exceeding 10 m s^{-1} was 1.8%. For the before−ASOS data, these figures were 4.0 m s^{-1}, 5.3 W m^{-2}, and 2.8%, respectively. Despite these differences, we decided to use all data in order to have the benefit of the full data record, rather than reducing the record length by excluding either the ASOS data or the before−ASOS data. The longer the record, the better the real distribution is captured. Records are already short since we do not use just one record for a location, but 192 records: one record for each month−direction combination (12 months × 16 directions).

Wind speeds are generated by month and by wind direction, since some months and directions have greater wind speeds than others. Wind speed by month is important because a field may be protected against wind erosion in one month, but not in another. For instance, most winter wheat fields in the U.S. Great Plains will be better protected with biomass in May than in February. Wind speed by direction is important for determining distances to non−erodible field boundaries. The longer this distance, the more a wind erosion avalanche effect can be expected. Wind direction relative to the direction of tillage operations and row crops is also important for wind erosion. Ridges and rows offer more protection against perpendicular winds than against parallel winds. In addition, the proper placement of wind barriers depends on wind direction.

STOCHASTIC WIND GENERATOR

There are two steps in the stochastic generation of wind data from measured data. First, statistics need to be created from the measured data, describing the distributions of wind direction and speed. Second, the wind data are generated from the statistics.

Creation of Statistics to be Used for Wind Generation

Wind direction frequencies were calculated for each of 16 directions plus calm for each month (table 1). Wind speeds that were not calm were sorted into 25 classes (table 2) for each month−direction combination. These classes are the same as the ones used by NCDC for the Wind Energy Resource Information System. Two different methods were used to summarize wind speed statistics: (1) fitting the two−parameter Weibull model to the measured wind speed distribution (fig. 1), and (2) using the measured wind speed distribution directly with linear interpolation between measurements (fig. 2). This second method will be referred to as the "direct" method.

The cumulative two−parameter Weibull distribution function, $F(u)$, (fig. 1) is defined by:

$$F(u) = 1 - \exp\left[-\left(\frac{u}{c}\right)^k\right] \tag{2}$$

where k is the Weibull shape parameter, and c is the Weibull scale parameter (m s^{-1}). The corresponding probability density function, $f(u)$, is defined by:

$$f(u) = \frac{dF(u)}{du} = \left(\frac{k}{c}\right)\left(\frac{u}{c}\right)^{k-1} \exp\left[-\left(\frac{u}{c}\right)^k\right] \tag{3}$$

Table 1. Wind direction frequencies (%) for Sidney, Nebraska, for 12 months and 16 cardinal wind directions plus calm. If a wind speed is less than or equal to 0.5 m s^{-1}, it is assigned to calm; otherwise, it is assigned to one of the 16 cardinal directions. Each month (column) adds up to 100%.

Direction	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
North	6.0	7.5	10.1	10.8	9.9	8.4	7.3	7.2	8.9	8.3	7.5	6.6
NNE	2.0	2.7	3.6	4.2	4.3	4.0	4.2	4.0	4.3	3.3	2.8	2.5
NE	1.6	2.0	2.4	3.3	3.8	3.5	3.9	3.9	3.3	2.4	1.6	1.5
ENE	1.5	1.8	2.4	2.9	3.2	3.7	3.8	3.3	2.9	1.8	1.3	1.2
East	1.9	2.3	3.3	4.0	4.4	6.0	5.7	4.4	3.8	3.0	1.6	1.5
ESE	1.5	2.0	3.7	3.9	4.6	5.4	5.3	4.7	3.4	2.7	1.4	1.0
SE	2.4	3.0	4.7	5.4	6.6	7.4	8.8	7.7	5.4	4.1	2.2	1.5
SSE	3.9	4.7	6.3	7.2	10.4	10.1	11.5	11.8	9.4	6.3	3.3	2.7
South	5.7	6.4	7.9	7.7	11.9	13.4	13.9	14.7	13.1	8.4	6.1	5.4
SSW	4.4	4.7	4.2	4.3	4.7	5.3	6.1	6.6	6.3	5.0	5.7	5.2
SW	6.6	5.5	4.8	3.9	3.7	4.4	4.4	4.8	5.1	5.4	6.3	7.1
WSW	6.6	5.2	4.1	3.3	2.9	3.5	3.4	3.4	3.8	4.8	6.3	7.2
West	17.5	15.1	10.6	8.0	5.7	6.0	5.6	6.5	8.0	12.9	16.7	17.3
WNW	17.4	14.7	11.1	10.3	6.9	5.9	5.0	6.0	8.2	12.0	14.9	17.8
NW	11.2	11.4	9.6	9.7	7.2	5.8	4.6	5.0	7.0	10.0	11.5	12.2
NNW	8.5	9.4	9.5	9.6	7.6	5.6	4.6	4.2	5.8	8.1	8.9	8.1
Calm	1.4	1.7	1.5	1.4	2.1	1.6	2.0	1.6	1.6	1.4	1.9	1.2

Table 2. The 25 classes used for sorting non−calm wind speeds. The upper wind speed limit is inclusive, e.g., a wind speed of 2.5 m s^{-1} goes into class 2. For class 25, the central wind speed was chosen as 43 m s^{-1}, as if the upper limit were 45.5 m s^{-1}. In reality, the upper limit of this last class is infinity.

	Wind Speeds (m s^{-1})		
Class	Lower	Upper	Central
1	0.5	1.5	1.0
2	1.5	2.5	2.0
3	2.5	3.5	3.0
.	.	.	.
.	.	.	.
.	.	.	.
17	16.5	17.5	17.0
18	17.5	18.5	18.0
19	18.5	19.5	19.0
20	19.5	20.5	20.0
21	20.5	25.5	23.0
22	25.5	30.5	28.0
23	30.5	35.5	33.0
24	35.5	40.5	38.0
25	40.5	inf.	43.0

The Weibull parameters are determined for each month−direction combination (tables 3 and 4). This can be accomplished using several methods, including the linearized least squares (LS) method (Justus et al., 1976; Skidmore and Tatarko, 1990) and the modified maximum likelihood (MML) method (Seguro and Lambert, 2000). With MML, the Weibull shape parameter (k) is estimated as:

$$k = \left[\frac{\sum_{i=1}^{n} u_i^k \ln(u_i) P(u_i)}{\sum_{i=1}^{n} u_i^k P(u_i)} - \sum_{i=1}^{n} \ln(u_i) P(u_i) \right]^{-1} \tag{4}$$

where u_i is the wind speed central to class i (table 2), $P(u_i)$ is the frequency with which the wind speed falls within class i, and n is the number of classes (n = 25). Equation 4 is solved iteratively, using the bracketing and bisection method (Press et al., 1992). The Weibull scale parameter (c) is estimated as:

$$c = \left[\sum_{i=1}^{n} u_i^k P(u_i) \right]^{1/k} \tag{5}$$

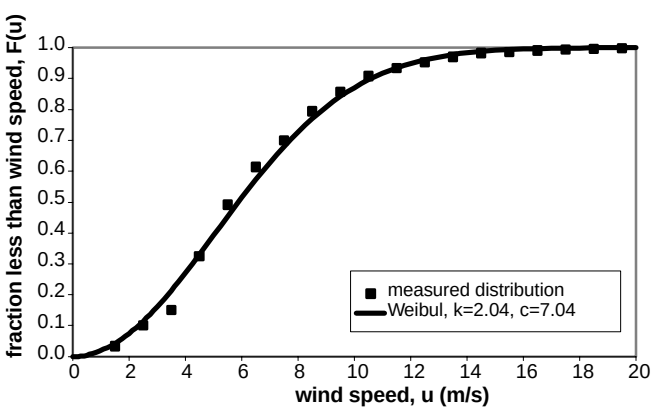

Figure 1. Cumulative distribution of measured wind speed data and fitted Weibull curve for Sidney, Nebraska, for December with winds coming from the northwest. The modified maximum likelihood (MML) method was used for the curve fit.

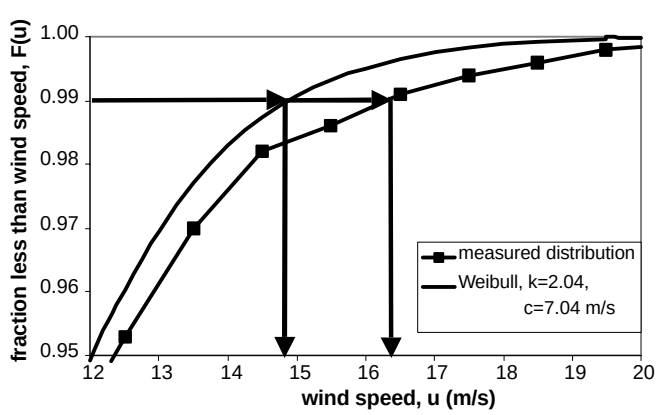

Figure 2. Same as figure 1, but zoomed in at higher wind speeds. In this example, when drawing a random number of 0.99, the Weibull curve would generate a wind speed of approximately 14.8 m s^{-1} and the direct method would generate a wind speed of approximately 16.3 m s^{-1}.

Table 3. Values of the Weibull shape parameter (k) for Sidney, Nebraska, for 12 months and 16 cardinal wind directions. The modified maximum likelihood (MML) method was used for the curve fits. The highlighted values in tables 3 and 4 define the Weibull curve in figures 1 through 3.

Direction	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
North	1.93	1.96	2.24	2.21	2.09	2.05	2.02	1.98	2.07	2.04	2.05	2.18
NNE	1.95	1.74	2.03	2.18	2.13	2.12	2.18	2.17	2.07	1.90	2.02	1.95
NE	1.79	1.91	2.04	2.08	2.06	2.11	2.04	2.04	2.20	1.98	1.69	2.04
ENE	2.16	1.97	1.76	2.07	1.98	1.98	2.15	2.04	1.81	2.11	1.88	1.80
East	1.90	1.88	2.12	2.04	1.87	2.04	1.95	2.03	1.94	2.09	1.86	2.00
ESE	1.63	1.98	2.17	2.12	1.98	2.18	2.32	2.23	2.24	2.22	2.10	1.77
SE	2.12	2.29	2.16	2.20	2.04	2.31	2.51	2.39	2.37	1.98	2.18	1.52
SSE	2.36	2.31	2.17	2.20	2.20	2.27	2.25	2.48	2.37	2.12	2.37	2.01
South	2.10	2.15	2.08	2.10	2.04	2.17	2.25	2.21	2.36	1.86	2.04	1.93
SSW	2.16	2.20	2.15	1.94	2.21	2.04	2.14	2.22	2.01	2.05	2.05	2.29
SW	2.24	2.09	2.38	2.12	1.94	2.01	1.99	2.01	2.18	2.35	2.43	2.28
WSW	2.31	2.23	2.07	1.96	1.70	1.77	1.63	1.90	1.84	2.10	2.04	2.34
West	2.26	2.14	1.79	1.93	1.64	1.65	1.79	1.76	1.87	2.07	2.02	2.22
WNW	2.25	2.00	1.90	1.86	1.85	1.67	1.72	1.95	2.08	1.94	2.10	2.21
NW	2.14	1.86	1.89	1.82	1.86	1.81	1.85	1.86	1.90	1.79	2.01	**2.04**
NNW	2.08	2.10	2.14	2.21	2.00	1.95	2.00	1.87	1.99	1.99	2.07	1.83

Table 4. Values of the Weibull scale parameter (c, m s^{-1}), for Sidney, Nebraska, for 12 months and 16 cardinal wind directions. The modified maximum likelihood (MML) method was used for the curve fits. The highlighted values in tables 3 and 4 define the Weibull curve in figures 1 through 3.

Direction	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
North	6.14	5.93	6.78	6.76	6.01	5.37	5.06	4.95	5.50	6.12	6.15	5.59
NNE	5.12	4.75	5.49	5.71	5.26	4.87	4.95	4.66	5.04	5.11	4.90	5.14
NE	4.55	4.52	4.87	5.15	5.43	4.84	5.06	4.61	4.80	4.49	4.76	4.68
ENE	3.82	4.10	4.41	4.75	4.58	4.70	4.82	4.19	4.35	3.97	3.58	3.68
East	4.36	4.09	4.69	4.98	4.71	4.81	4.72	4.47	4.32	3.99	3.73	3.62
ESE	4.72	4.48	5.44	5.77	5.69	5.27	5.09	4.83	4.72	4.82	3.99	4.13
SE	4.97	5.08	5.84	6.34	6.35	5.58	5.55	5.21	5.22	5.34	4.43	4.76
SSE	5.31	5.76	6.49	6.58	6.98	6.16	5.86	5.66	5.77	5.86	4.64	5.29
South	4.87	5.47	5.81	5.74	6.11	5.88	5.42	5.37	5.62	5.43	4.84	5.15
SSW	5.04	4.98	4.96	5.22	5.00	4.82	4.82	4.82	5.02	4.86	4.68	4.95
SW	4.81	4.79	4.79	4.73	4.60	4.44	4.46	4.36	4.53	4.41	4.68	4.94
WSW	4.57	4.43	4.42	4.66	4.06	4.31	3.93	3.88	4.22	4.05	4.34	4.45
West	5.59	5.26	5.20	5.27	4.69	4.54	3.99	4.27	4.63	4.82	5.14	5.46
WNW	6.55	6.38	6.34	6.56	5.37	5.22	4.79	4.76	5.10	5.53	6.17	6.48
NW	6.95	7.30	7.03	7.37	6.32	5.43	4.72	4.50	5.38	6.40	6.93	**7.04**
NNW	7.65	7.52	7.83	7.80	6.71	5.60	5.03	4.79	5.87	6.78	7.47	7.49

Equation 5 can be solved explicitly. A disadvantage of MML is that the curve fit has to be to the entire wind speed distribution. One is not free to select only a part of the distribution, since the weights are not accessible (N. J. Cook, personal communication). In our case, a partial fit may be desirable, since we are interested in the high wind speeds only.

With LS, the cumulative distribution function (eq. 2) is first transformed (fig. 3) according to:

$$y = a + bx \tag{6}$$

where $y = \ln[-\ln(1 - F(u))]$, $x = \ln(u)$, $a = -k \ln(c)$, and $b = k$. Intercept a and slope b, and thus k and c, are determined from linear regression. LS can be carried out with different weighting schemes, e.g., uniform (no) weighting (LS_{uni}) or weighting with the number of observations in the corresponding class (LS_{obs}). With LS, partial fitting is possible, which is a special case of weighted fitting: one or more points receive a weight of 0.

With the direct method, the measured distribution itself is used without fitting to the Weibull or any other model. Linear interpolation is used between the measured distribution points (fig. 2). In this case, for the generation of wind speeds, the entire distribution needs to be available for each month−direction combination (table 5), rather than just the Weibull parameters (tables 3 and 4).

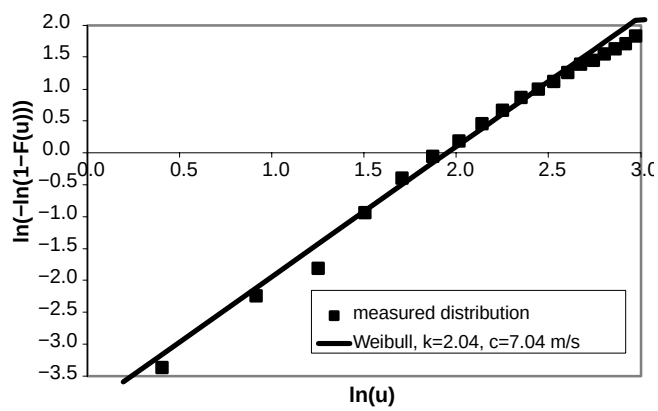

Figure 3. Same as figure 1, but data and Weibull curve were transformed using equation 6, making the fitted Weibull curve a straight line.

Table 5. Cumulative wind speed distributions (fraction less than wind speed) for Sidney, Nebraska, for 12 months and 16 cardinal wind directions. The highlighted distribution is shown in figures 1 through 3.

Month and Wind Direction	Wind Speed (m s^{-1})											
	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	…
Jan. North	0.074	0.188	0.256	0.404	0.539	0.672	0.771	0.868	0.924	0.951	0.965	…
Jan. NNE	0.075	0.259	0.375	0.544	0.675	0.784	0.866	0.944	0.975	0.991	0.994	…
Jan. NE	0.117	0.340	0.449	0.619	0.769	0.883	0.915	0.943	0.964	0.988	0.988	…
Jan. ENE	0.141	0.389	0.509	0.735	0.885	0.979	0.996	0.996				
Jan. East	0.161	0.339	0.426	0.613	0.781	0.906	0.942	0.977	0.990	0.997		
Jan. SE	0.139	0.325	0.398	0.597	0.736	0.844	0.900	0.961	0.970	0.991	0.996	…
Jan. SE	0.090	0.227	0.339	0.561	0.713	0.840	0.899	0.956	0.987	0.995	0.997	
Jan. SSE	0.064	0.195	0.268	0.478	0.651	0.802	0.890	0.957	0.987	0.998		
Jan. South	0.080	0.232	0.353	0.583	0.761	0.859	0.906	0.954	0.976	0.993	0.997	…
Jan. SSW	0.057	0.197	0.306	0.543	0.735	0.860	0.915	0.969	0.980	0.994	0.996	…
Jan. SW	0.079	0.238	0.333	0.576	0.761	0.870	0.923	0.973	0.990	0.999		
Jan. WSW	0.058	0.240	0.359	0.657	0.822	0.906	0.951	0.978	0.988	0.997		
Jan. West	0.035	0.135	0.214	0.437	0.673	0.825	0.896	0.942	0.961	0.978	0.982	…
Jan. WNW	0.023	0.082	0.133	0.304	0.533	0.702	0.798	0.883	0.914	0.945	0.960	…
Jan. NW	0.028	0.092	0.152	0.326	0.503	0.617	0.696	0.801	0.861	0.921	0.942	…
Jan. NNW	0.033	0.097	0.146	0.269	0.392	0.523	0.623	0.729	0.807	0.865	0.906	…
Feb. North	0.068	0.189	0.262	0.415	0.586	0.710	0.788	0.887	0.931	0.967	0.980	…
Feb. NNE	0.134	0.312	0.413	0.602	0.718	0.829	0.894	0.952	0.975	0.990	0.995	…
Feb. NE	0.119	0.317	0.409	0.617	0.785	0.865	0.927	0.980	0.997	0.997	0.997	…
.	.	.	.	.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	.	.	.	.	
Dec. WNW	0.022	0.092	0.156	0.331	0.550	0.708	0.793	0.864	0.909	0.941	0.959	…
Dec. NW	**0.034**	**0.101**	**0.151**	**0.324**	**0.491**	**0.613**	**0.699**	**0.794**	**0.857**	**0.908**	**0.934**	…
Dec. NNW	0.037	0.122	0.185	0.314	0.438	0.534	0.630	0.743	0.815	0.876	0.914	…

Wind Generation

First, one of the 16 cardinal wind directions or calm is selected for the current month from table 1 using a random number generator. The selected direction is applied for an entire day. Next, 24 hourly wind speeds are generated for this day. If calm was selected in the previous step, then 24 wind speeds of 0 m s^{-1} are generated. Otherwise, if one of 16 directions was selected, then 24 wind speeds are generated from either the Weibull parameters or the cumulative distribution. If using the Weibull model, the parameters *k* (table 3) and *c* (table 4) are selected for the current month and direction, and a wind speed is generated from the Weibull curve, using a random number generator (fig. 2). If using the direct method, the distribution for the current month and direction (table 5) is selected, and a wind speed is generated from the linearly interpolated distribution, using a random number generator (fig. 2).

The average duststorm lasts 6.6 h in the U.S. Great Plains (Hagen and Woodruff, 1973), but there is no auto−correlation for the generated hourly wind speeds in WEPS. Instead, as a first approximation, the hourly wind speeds are rearranged to create more realistic windstorms that last longer than 1 h. Preliminary tests have shown that simulated wind erosion is not very sensitive to how the winds speeds are rearranged. In addition, there is no cross−correlation with other weather elements in WEPS, although it does exist. Wind speed may be correlated with precipitation (Visser et al., 2003) and with change in maximum air temperature from one day to the next (G. L. Johnson, 2003, personal communication). Auto− and cross−correlation may be incorporated into a future version of WEPS. Computer programs to convert measured wind data into the summary statistics described above and to generate winds from these statistics are available upon request.

Results

Figure 1 illustrates a Weibull curve fit to the cumulative distribution of measured wind speed data for Sidney, Nebraska, for December with winds coming from the northwest. The MML curve fit seems reasonable, but it resulted in a substantial underprediction of WPD (WPD_m = 40.0 W m^{-2} and WPD_g = 26.9 W m^{-2}). In addition, a chi−squared goodness of fit test showed that it is extremely unlikely ($p < 10^{-9}$) that the measured distribution was "drawn" from the Weibull curve. Upon closer inspection of the high wind speeds (fig. 2), differences between the measured distribution and the Weibull curve are seen more clearly. They were typical for many month−direction combinations, resulting in fewer high wind speeds in the generated data than in the measured data. For Sidney, Nebraska (all months and directions combined), this caused a considerable underprediction of WPD (WPD_m = 13.3 W m^{-2} and WPD_g = 8.6 W m^{-2}). Out of the 332 stations having WPD_m > 5 W m^{-2}, 164 stations underpredicted WPD by more than 20%, and maximum wind speeds were frequently underpredicted as well (table 6). Average wind speeds were predicted much better. Under− or overprediction was less than 5% at all stations.

Results were worse using LS, trying several weighting schemes. It was expected that other models than the Weibull would give similar challenges, given the fact that relatively small deviations between measured and fitted distributions cause such large deviations in WPD. Thus, it was decided not

Table 6. Number of stations (out of the 332 stations having WPD_m > 5.0 W m^{-2}) that over− or underpredict erosive wind power density or maximum wind speed by more than 5%, 10%, 15% 20%, 25%, or 30%. Comparison of the Weibull modified maximum likelihood method (Weibull) and the method of generating values directly from the measured distribution (Direct).

Prediction Error	Number of Stations							
	Erosive Wind Power Density				Maximum Wind Speed			
	Underprediction		Overprediction		Underprediction		Overprediction	
	Weibull	Direct	Weibull	Direct	Weibull	Direct	Weibull	Direct
>5%	270	18	23	50	252	61	47	129
>10%	238	5	11	10	229	38	34	48
>15%	200	1	7	0	197	24	19	9
>20%	164	1	4	0	157	19	13	1
>25%	122	0	2	0	126	9	8	0
>30%	86	0	1	0	82	6	7	0

to try other models, but rather to attempt fitting (still using the Weibull) to the high wind speeds only, ignoring low and medium wind speeds, expecting a closer fit at the high wind speeds, where it is the most critical for wind erosion prediction.

When fitting to high wind speeds only, WPD_g came closer to WPD_m for many station−month−direction combinations. However, unrealistic wind speeds were generated for others. For example, for Omak, Washington, a partial curve fit using LS_{obs} (figs. 4 and 5) resulted in k = 0.39 and c = 0.34 m s^{-1}, which generated wind speeds as high as 107 m s^{-1}. A partial curve fit using LS_{uni} generated even higher wind speeds. When including some lower wind speeds in the curve fit, generated wind speeds were not as high. However, the same problem appeared as when fitting to the entire wind speed range: the high wind speeds did not fit well enough for many stations. With more than 185,000 curve fits (971 stations, 12 months, 16 wind directions), it was not practical to consider each one individually. One algorithm had to be applied to all curve fits, but it became clear that it would be difficult to devise an algorithm that could prevent cases like that of Omak.

The direct method is more robust, and WPD_g closely reproduced WPD_m. With this method, there was only one station (out of the 332 stations having WPD_m > 5 W m^{-2}) where WPD_g deviated more than 20% from WPD_m, and maximum wind speeds were also predicted much better than with the Weibull model (table 6). As with the Weibull model, under− or overprediction of average wind speed was less than 5% at all stations.

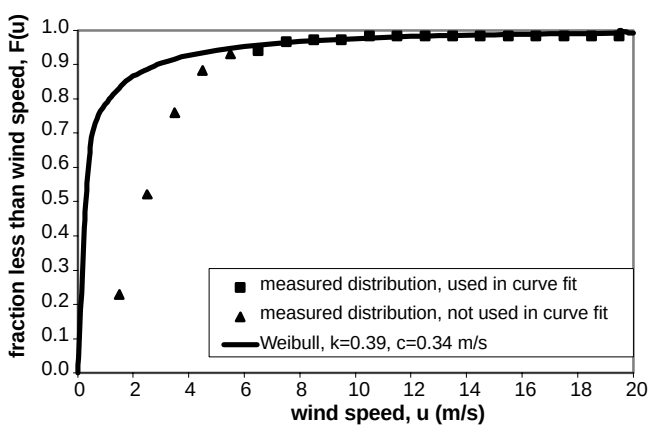

Figure 4. Cumulative distribution of measured wind speed data and fitted Weibull curve for Omak, Washington, for March with winds coming from the WNW. The method of least squares, weighted by the number of observations in the corresponding class (LS_{obs}), including only the data points depicted as squares, was used for the curve fit.

Discussion

The direct method is more convenient for spatial interpolation between stations, since the "parameters" in table 5 have a physical meaning as opposed to the Weibull k and c parameters. Take the simple example of a location situated exactly halfway between two wind stations. With the direct method, it makes sense to assign averages of the parameters at the two wind stations to this location. For instance, if 80% of the wind speeds at one station are less than 9.5 m s^{-1}, compared to 90% at the other station, then it seems reasonable to assign 85% to a location halfway between these stations. Averaging the Weibull parameters does not provide an acceptable solution, and the interpolation procedure becomes more complex. For both methods, more sophisticated algorithms using additional information, such as geographical and terrain data, can improve spatial interpolation, especially in mountainous regions. The U.S. Department of Energy (2004) has prepared wind power maps of high spatial resolution from such data.

Temporal interpolation could be accomplished in the same manner as spatial interpolation. For instance, if 80% of the wind speeds in January are less than 9.5 m s^{-1}, compared to 90% in February, then 81% could be assigned to 18 January, 82% to 21 January, 85% to 31 January, etc. As with spatial interpolation, temporal interpolation would be more complex using the Weibull model. Wind speed distributions may not change much from one month to the next; consequently, temporal interpolation may not be very important.

An additional advantage of the direct method is that the implementation can be simplified. Although we maintained

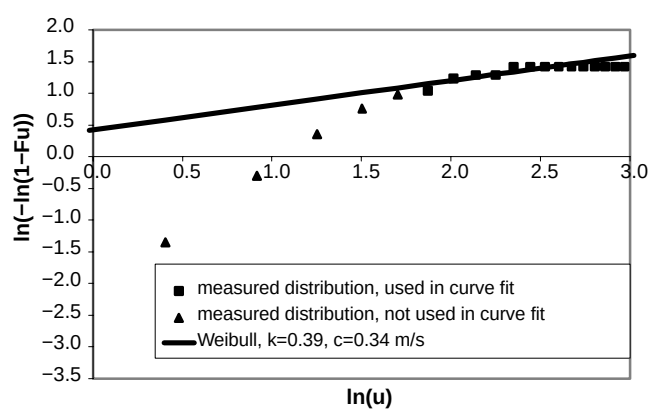

Figure 5. Same as figure 4, but data and Weibull curve were transformed using equation 6, making the fitted Weibull curve a straight line.

the separation of calm and non−calm winds when implementing the direct method, this is not necessary. This is only required when fitting the wind speed distribution to a model such as the Weibull, since the model does not fit the calm region of a distribution that includes calm.

A slight disadvantage of the direct method is that the wind statistics file necessary for WEPS (tables 1 and 5) is about four times larger than the file containing the Weibull coefficients (tables 1, 3, and 4). However, the extra disk space required is worth the improved accuracy.

If the number of wind speed classes were increased beyond the 25 classes used for this work, then the results would be expected to further improve with the direct method, since there would be more and shorter linear pieces between more numerous data points (fig. 2). Of course, these linear pieces are only approximations of the real wind speed behavior between data points. With more and shorter linear pieces, the errors in these approximations would be reduced. More wind speed classes would not necessarily improve results when using the Weibull model. This may be understood by considering figure 3: the same data points would still be there, with additional points in between. The data in figure 3 would still show the same non−linearity at high wind speeds, and the Weibull fit would not change much. The deviation at the high wind speeds (figs. 2 and 3) would still be there.

A measured wind speed distribution is not the "true" population distribution, but only a sample. It would be the true distribution if the number of measurements were infinite. However, the measured distribution is the best approximation available, and the more measurements used, the better it becomes. Using a mathematical model such as the Weibull would be appropriate if there were a theory that showed that wind speeds are "drawn" from a Weibull distribution. Take the example of one perfect die. It is easy to see that the observations are drawn from a uniform distribution and that the sample distribution approaches this uniform distribution as the number of throws of the die approaches infinity. In this case, it is better to use the uniform model rather than throwing the die a finite number of times and then using the sample distribution. In the case of natural winds, it is not known from what theoretical distribution the wind speeds are drawn. The processes that create winds are far too complex for this.

Sometimes there are atypical data in a measured data record. For instance, a measured data set with a length of five years may include a very high wind speed, which occurs only once every 100 years. The likelihood of this 100−year maximum wind speed will be overpredicted using the direct method. The overprediction is probably less using the Weibull model because of a smoothing effect of the curve fit. However, it is impossible to know that this is, in fact, the 100−year maximum wind speed. Perhaps this is not an atypical data point, but simply a data point that does not fit the Weibull curve. In this case, the Weibull model does not represent the true distribution, which leads to errors such as the substantial systematic underprediction seen in table 6. Models with more parameters than the two used by the Weibull model will most likely reduce this systematic error. The direct method, which can be considered a model with 25 parameters, is an extreme implementation of using a model with more parameters.

Conclusions

The Weibull model did not fit wind speed distributions well enough for application in WEPS. Fitting the Weibull model to only the high wind speeds resulted in some generated wind speeds exceeding 100 m s^{-1}, which is unacceptable. The decision was made to store the cumulative wind speed distributions themselves, instead of storing the Weibull parameters describing them, and then generate wind speeds directly from the distributions. This direct method is robust and very closely reproduced the erosive wind power density of the measured data. With this method, there was only one station (out of the 332 stations having $WPD_m > 5$ W m^{-2}) where WPD_g deviated more than 20% from WPD_m, compared to 168 stations using the Weibull model. Thus, the direct method of wind speed generation reproduces wind speeds more accurately than the Weibull model, which is important for wind erosion prediction and may be important for other applications as well.

References

Bagnold, R. A. 1941. *The Physics of Blown Sand and Desert Dunes*. London, U.K.: Chapman and Hall.

Chepil, W. S. 1945. Dynamics of wind erosion: III. The transport capacity of the wind. *Soil Science* 60(6): 475−480.

Cook, N. J. 1982. Towards better estimation of extreme winds. *J. Wind Eng. and Ind. Aerodynamics* 9(3): 295−323.

Corotis, R. B., A. B. Sigl, and J. Klein. 1978. Probability models of wind velocity magnitude and persistence. *Solar Energy* 20(6): 483−493.

Hagen, L. J., L. E. Wagner, and E. L. Skidmore. 1999. Analytical solutions and sensitivity analyses for sediment transport in WEPS. *Trans. ASAE* 42(6): 1715−1721.

Hagen, L. J., and N. P. Woodruff. 1973. Air pollution from duststorms in the Great Plains. *Atmospheric Environ.* 7(3): 323−332.

Hennessey, J. P. 1977. Some aspects of wind power statistics. *J. Applied Meteorology* 16(2): 119−128.

Justus, C. G., W. R. Hargraves, and A. Yalcin. 1976. Nationwide assessment of potential output from wind−powered generators. *J. Applied Meteorology* 15(7): 673−678.

Lockhart, T. J. 2000. A summary of wind climate continuity with ASOS. In *Proc. 12th AMS Conference on Applied Climatology*, 131−133. Boston, Mass.: American Meteorological Society.

Mayne, J. R. 1979. The estimation of extreme winds. *J. Wind Eng. and Ind. Aerodynamics* 5(1−2): 109−137.

McKee, T. B., N. J. Doesken, C. A. Davey, and R. A. Pielke. 2000. Climate data continuity with ASOS. Climatology Report No. 00−3. Fort Collins, Colo.: Colorado State University, Department of Atmospheric Science, Colorado Climate Center.

Nicks, A. D., and L. J. Lane. 1989. Weather generator. In *USDA Water Erosion Prediction Project: Hillslope Profile Model Documentation*, 2.1−2.19. NSERL Report No. 2. L. J. Lane, and M. A. Nearing, eds. West Lafayette, Ind.: USDA−ARS National Soil Erosion Research Laboratory.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. Cambridge, U.K.: Cambridge University Press.

Seguro, J. V., and T. W. Lambert. 2000. Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis. *J. Wind Eng. and Ind. Aerodynamics* 85(1): 75−84.

Skidmore, E. L. 1998. Wind erosion processes. In *Wind Erosion in Africa and West Asia: Problems and Control Strategies*, 137−142. M. V. K. Sivakumar, M. A. Zöbisch, S. Koala, and T. Maukonen, eds. Aleppo, Syria: ICARDA.

Skidmore, E. L., and J. Tatarko. 1990. Stochastic wind simulation for erosion modeling. *Trans. ASAE* 33(6): 1893−1899.

Takle, E. S., and J. M. Brown. 1978. Note on the use of Weibull statistics to characterize wind speed data. *J. Applied Meteorology* 17(4): 556−559.

U.S. Department of Energy. 2004. U.S. State Maps of Wind Resources. Available at: www.eere.energy.gov/windpoweringamerica/wind_resources.html.

Visser, S. M., J. K. Leenders, and M. Leeuwis. 2003. Farmers' perceptions of erosion by wind and water in northern Burkina Faso. *Land Degradation and Development* 14(1): 123−132.