DEVELOPMENT OF LARGE DIRECTIONAL WIND SPEED DATABASES

DongHun Yeo
Engineering Laboratory, National Institute of Standards and Technology, 100 Bureau Dr., Gaithersburg, MD 20899-8611; PH (301) 975-8103; FAX (301) 869-6275; email: donghun.yeo@nist.gov

ABSTRACT

For structures sensitive to wind directionality, methods for the estimation of wind effects require the use of time series of directional wind speeds with length exceeding the length of the MRIs of interest. This study proposes a procedure for generating such time series from relatively short synoptic wind data sets. The focus in this paper is on the estimation of the parameters of a probabilistic model of the wind speeds and on errors in the estimation. The wind speed data being generated can be used in the Database-Assisted Design approach to account for directional wind effects.

Keywords: Directional wind speeds; Mean recurrence interval; Extreme value statistics; Synthetic wind speed data; Monte Carlo simulation; Generalized Pareto distribution.

INTRODUCTION

Inherent in the ASCE 7 Standard (ASCE 2010) wind loading provisions are approximations that may be acceptable for the design of ordinary structures but are deemed unacceptable for the design of special structures, including in particular tall buildings. Among these approximations are those inherent in the use by ASCE 7 Standard of a blanket wind directionality factor. For special structures wind engineering laboratories use more elaborate methods for accounting for directionality effects. In particular, methods that determine wind effects corresponding to specified MRIs by using non-parametric statistical approaches require the use of time series of directional wind speeds that are longer than those MRIs. Hence it is necessary to use simulation techniques to develop long time series of synthetic directional wind speeds from smaller data sets.

The design MRIs of wind effects for wind loads specified in ASCE 7 Standard are 300 years, 700 years, and 1700 years, depending upon the structure’s risk category. If design wind speeds are specified regardless of their direction, as is the case for the ASCE 7 Standard, the MRI of the design wind speeds and the MRI of the design wind effects are the same. This is typically not the case if directional wind speeds are used in design.

The procedure used in this study involves three steps. First, probabilistic models are fitted to the directional wind speeds. Second, the fitted models are used to generate by Monte Carlo simulation synthetic directional wind speed records that have any desired length. Third, the uncertainty in the generated wind speeds is assessed. The procedure proposed in the paper yields simple and transparent probabilistic estimates of directional wind speeds for use in structural design.

MODEL OF DIRECTIONAL WIND SPEEDS

This study proposes a probabilistic model for synoptic winds based on (1) a Poisson process representing the arrival of extreme wind events and (2) a translation vector defining directional wind speeds, consisting of a nonlinear transformation of a direction-dimensional Gaussian vector. In this study we make use of the methodology developed by Grigoriu (2009), which is summarized herein.

Let **T** be a sequential time series of the extreme wind events. The element T_k describes the time of the k^{th} extreme wind event in the time sequence (k = 1, 2, …, n, where n is the number of extreme wind events being considered). If it is assumed that the random times T_k are the jump times of a homogeneous Poisson process $N(t)$ ($t > 0$) of intensity $\lambda > 0$, the average number of the extreme wind events during a time $t > 0$ is equal to the expectation $E[N(t)] = \lambda t$ of $N(t)$, where the intensity λ is the rate of arrival of the extreme wind events. The arrival rate of extreme wind events is simply modeled by a homogeneous Poisson processes with mean rate inferred from observations. For example, if the observed number of wind events at a site during a 30-year period is 66, then the estimated rate of arrival is λ = 2.2 year^{-1}.

Suppose that a matrix of synoptic wind speeds **V** has n rows and d columns, where n is the number of extreme wind events and d is the direction of the synoptic winds. The element $V_i^{(k)}$ describes the wind speed recorded in the i^{th} direction during the k^{th} wind event (i = 1, 2, …, d and k = 1, 2, …, n) at time T_k. If it is assumed that the wind speed vectors $V^{(k)}$ representing the extreme wind events are independent vectors of a d-dimensional random vector **V** with joint distribution F, the proposed model can be characterized by (1) the intensity λ of the Poisson model N and (2) the distribution of **V**. While the estimation of λ of the Poisson process N is simple, the selection of the joint distribution of F of **V** is much more complicated because there are no general models for arbitrary non-Gaussian joint distributions. In this study, we assume that **V** is a translation vector that accounts in an approximate manner for the correlation between directional wind speeds.

Suppose the components V_i of the directional wind speeds in extreme wind events are defined as follows:

$$V_i = F_i^{-1}\left[\Phi\left(G_i\right)\right] \qquad \text{for} \quad i = 1, 2, \ldots, d \tag{1}$$

where F_i denotes the distribution of the i^{th} direction wind speed vector V_i, Φ denotes the i^{th} direction distribution of the standard Gaussian variable with mean 0 and variance 1, G_i is the i^{th} direction correlated standard Gaussian variable with covariance matrix $\rho_{ij} = E[G_i, G_j]$ where i, j = 1, ..., d, and E denotes expectation. Equation 1 establishes a one-to-one correspondence between the wind speed matrix **V** and the Gaussian matrix **G**. For synoptic winds, however, the available data is typically not sufficient for the estimation of the correlation between directional wind speeds. For this reason it is assumed in this study that directional wind speeds are independent of each other. This assumption is conservative from a structural design viewpoint, as is shown in the paper.

MODEL CALIBRATION

To calibrate the probabilistic model for **V**, we estimate the marginal distributions F_i of the directional wind speeds V_i. Since numerous directional wind speed data are considered as zero when they are not above a threshold u, the distribution F_i of a random variable **V** is defined as:

$$F_i(v) = q_i + (1 - q_i)1(v > u)\tilde{F}_i(v) \qquad \text{for} \quad i = 1, 2, \ldots, d \tag{2}$$

where $1(A)$ denotes the indicator of set A and is equal to 1 and 0 when A is true and false, respectively, q_i is the probability $P(V_i \leq u)$ that V_i is less than or equal to the threshold wind speed u in the i^{th} directional wind speed vector, and $\tilde{F}_i(v)$ is a proper distribution expressing the wind speeds of V_i above the threshold.

We assume that the distribution $\tilde{F}_i(v)$ of a random variable **V** is the generalized Pareto distribution (GPD) with parameters (c_i, a_i, u_i), where c_i is the shape (i.e., tail length) parameter, a_i is the scale parameter, and u_i is the location (i.e., threshold) parameter, in the i^{th} direction (the GPD is an appropriate model for the exceedances over a suitable threshold of extreme value variates; see Galambos et al. 1994). The cumulative distribution function of the generalized Pareto distribution is

$$\tilde{F}_i(v) = \begin{cases} 1 - \left[1 + c_i\left(\dfrac{v-u_i}{a_i}\right)\right]^{-1/c_i} & \text{for } c_i \neq 0 \\ 1 - \exp\left(\dfrac{v-u_i}{a_i}\right) & \text{for } c_i = 0 \end{cases} \tag{3}$$

where the domain is $v \geq u_i$ for $c_i \geq 0$ and $u_i \leq v \leq u_i - a_i/c_i$ for $c_i < 0$. The density function is

$$\tilde{f}_i(v_i) = \begin{cases} \dfrac{1}{a_i}\left[1 + c_i\left(\dfrac{v-u_i}{a_i}\right)\right]^{-\left(1+\frac{1}{c_i}\right)} & \text{for } c_i \neq 0 \\ \dfrac{1}{a_i}\exp\left[-\left(\dfrac{v-u_i}{a_i}\right)\right] & \text{for } c_i = 0. \end{cases} \tag{4}$$

In order to estimate the parameters (c_i, a_i, u_i) in the generalized Pareto distribution $\tilde{F}_i(v)$ of above-threshold wind speed data, the study employs the maximum likelihood estimation (MLE, Kotz and Nadarajah 2000) and the de Haan method (de Haan 1994).

To avoid unrealistic or unconservative modeling of non-hurricane wind speeds, we propose that if the estimated value of the GPD tail length parameter is $c_i > -0.01$, the value used in the calculations can be taken as $c_i = -0.01$ (corresponding to within a close approximation to a Gumbel distribution tail), and if the estimated value is $c_i < -0.1$, the value used in the calculations can be taken as $c_i = -0.1$, thereby avoiding distribution tails that may be unconservatively short.

GENERATION OF SYNTHETIC WIND SPEEDS

Once the probability law of the proposed directional synoptic wind speed model has been calibrated to observed wind records at a site, we can apply Monte Carlo simulations to generate directional wind speed data of any length that are consistent with the observed records.

For the generation of directional synoptic wind speeds with annual arrival rate λ over a period of τ years, the following steps are required:

(1) Generate d independent random numbers (r_1, r_2, …, r_d) uniformly distributed in the range of (0, 1). If the random number r_i is less than the probability q_i that wind speeds are not higher than the threshold, the generated wind speed in the i^{th} direction is zero. Otherwise, an above-threshold wind speed v is generated from Eq. (5) by another random number r_i' in (0, 1):

$$v = u_i + \frac{a_i}{c_i}\left[\left(r_i'\right)^{-c_i} - 1\right]. \tag{5}$$

(2) Repeat step (1) $n_\tau = \lambda\tau$ times to generate a time series of directional extreme wind speeds over τ years.

ESTIMATION OF MEAN RECURRENCE INTERVALS

Consider a set of n wind speed data at a site where the mean storm arrival rate is λ year^{-1}. If the rate were λ = 1 storm/year, the estimated probability that the highest speed in the set would be exceeded is $1/(n+1)$, and the corresponding estimated MRI would be $\bar{N} = n+1$ years (Simiu and Miyata 2006). The estimated probability that the q^{th} highest speed in the set is exceeded is $q/(n+1)$, the corresponding estimated MRI in years is $\bar{N} = (n+1)/q$, and the rank of the wind speed with MRI is $q = (n+1)/\bar{N}$.

For a general case of $\lambda \neq 1$, the estimated MRI is $\bar{N} = (n+1)/(\lambda q)$ years. For example, if n = 999 synoptic wind speed data, and λ = 0.5 year^{-1}, the estimated MRI of the event that the highest wind speed in the sample will occur is $\bar{N}$ = 1000/0.5 = 2000 years, the estimated MRI of the second highest speed is 1000 years, and so forth. The rank of the speed with a specified MRI $\bar{N}$ is $q = (n+1)/(\lambda\bar{N})$.

The wind speed with an $\bar{N}-\text{year}$ mean recurrence interval obtained from wind speeds above a threshold u can be estimated by the generalized Pareto distribution as follows:

$$v(\bar{N}) = u_i - \frac{a_i}{c_i}\left[1-\left(\lambda_i\bar{N}\right)^{c_i}\right]. \tag{6}$$

where λ_i is the annual occurrence rate of wind events in the i^{th} direction and can be estimated as $\lambda_i = \lambda \cdot P(V_i > u) = \lambda(1-q_i)$, where λ is the annual occurrence rate of the wind events in all wind directions and q_i is the probability that wind speeds in the i^{th} direction are less than or equal to the threshold u, as previously explained.

ESTIMATION OF WIND EFFECTS WITH SPECIFIED MRIs

For structural design for wind, the estimation of wind effects -- rather than wind speeds -- corresponding to design MRIs, is of concern to structural engineers. This is the case because, as was mentioned earlier, owing to wind directionality effects, the MRIs of the wind effects induced by directional wind speeds typically differ from the MRIs of the corresponding wind speeds regardless of their direction.

The analytical procedure specified in the ASCE 7 Standard uses a building wind directionality factor K_d = 0.85 applied to the wind effect calculated by disregarding wind directionality. This approach is simple, but can either overestimate or underestimate the response. For structures for which directional wind effects are significant alternative approaches have been developed and are currently being used in practice. Database-Assisted Design (DAD), developed by National Institute of Standards and Technology (NIST), is an integrated design tool for structural design of strength and serviceability (Spence 2009; Yeo 2010) which enables the probabilistic estimation of wind effects while accounting for wind directionality as reflected by measured or simulated wind speed data. The estimation procedure entails the following steps:

(1) Develop an $n \times d$ matrix of directional wind speeds from measured data. The number n of rows is equal to the number of storm events or of years of record, and must be sufficiently large to allow the use of non-parametric estimates of wind effects with MRIs of the order of thousands of years. The number d of columns in the matrix is equal to the number of wind directions being considered (e.g., 8, 16, or 36). The matrix of directional wind speeds at a site,

called *climatological wind database*, is developed for long periods exceeding the design MRIs by probabilistic estimates. The procedure of generating synoptic winds from measured data was previously described in the paper.

(2) Transform the $n \times d$ matrix of directional wind speeds into an $n \times d$ matrix of wind effects induced by each directional storms. The wind effects in DAD include demand-to-capacity indexes for structural members, inter-story drifts, and top-floor accelerations. The detailed procedure of DAD is provided in Yeo (2010).

(3) Create a vector of dimension n consisting of the largest wind effect of interest corresponding to each row (i.e., each storm) of the wind effects matrix developed in Step 2. For each storm event only the largest of the directional responses is of interest from a structural design viewpoint and all the other responses are discarded.

(4) Use non-parametric estimates to obtain statistics of the wind effect for which the vector was created in Step 3. This vector is rank-ordered, and the peak responses corresponding to the required mean recurrence intervals are obtained using the non-parametric estimation method (Simiu and Miyata, 2006, p. 33). The peak wind effects of interest can be estimated for the respective specified MRIs.

APPLICATION

We employ the proposed probabilistic model of directional wind speeds to generate synthetic synoptic wind speed series data for large MRIs at Newark, New Jersey. For the calibration of the proposed model for directional synoptic wind speeds, we use observed data from the Automated Surface Observing System (ASOS), a network of about 20000 standardized US weather stations (NCDC 2008). The ASOS data of synoptic winds in Newark have 228 wind events in 36 directions in 10 º increments, threshold wind speed of 35 knots (1 knot ≈ 0.51 m/s), and measuring duration of 19.94 years. Thus, the annual rate of occurrence for the wind events is 228/19.94 year = 11.43 year^{-1}. Figure 1 shows the distribution of directional speeds of synoptic winds at Newark, NJ.

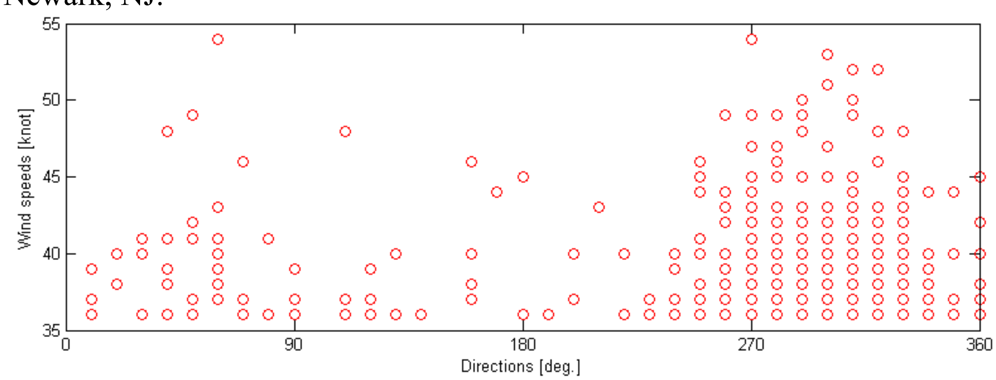

Figure 1. Directional wind speeds of synoptic winds (Newark, NJ)

Because the directional data are not sufficient to calibrate parameters of the generalized Pareto distribution in each direction, this study divides the data by 4 sectors (i.e., 10 º to 90 º, 100 º to 180 º, 190 º to 270 º, and 280 º to 360 º), and estimates the probability of wind speeds less than or equal to a threshold u, $q_i = P(V_i \leq u)$ and the parameters (c_i, a_i, u_i) for each sector by the MLE and the de Haan method. As shown in Table 1, the probability q_i (Eq. 2) significantly depends on sectors: it is 0.93 for sector 2, and 0.22 for sector 4. For the parameters of GPD, 3 sectors have estimated shape parameters c_i lower than -0.10; as indicated earlier, these are

assumed to be −0.10. In contrast, since for section 2 c_i = -0.08, this value of the parameter is used without adjustment. The scale parameters a_i are approximately 4 to 7, and the threshold u_i is 35 knots, regardless of sector.
Synthetic wind speeds for 60000 synoptic wind events in Newark, NJ have been generated in 4 sectors by Monte Carlo simulation using parameters estimated from the MLE and the de Haan method. The adjusted parameters of the sectors are used in the simulation. Estimates of parameters (c_i, a_i) and q_i are shown in Table 1. Where two numbers separated by a slash are shown in Table 1, the first and second number is estimated by the MLE and the de Haan method, respectively. The simulation has reliably generated the probabilities q_i that wind speeds are less than or equal to a threshold u = 35 knots. The MLE method has estimated parameters of the synthetic data that are relatively closer to those of the ASOS data than the de Haan method. Figure 2 shows empirical and fitted cumulative distribution functions (CDFs) $\tilde{F}_i(x)$ of generated above-threshold wind speeds and of adjusted parameters of GPD, respectively, in sector 4 using both estimation methods.

The Monte Carlo simulation has enabled the generation of the time-series of directional synoptic wind speeds of synoptic wind speeds allowing the estimation of wind effects with MRIs of up to 5000 years. Figure 3 shows results based on the MLE and the de Haan method. For any given MRI the wind speeds are generally higher for the MLE method than for the de Haan Method.

Bootstrapping is used to assess the statistical uncertainty in the estimates of the wind speeds v_{MRI} corresponding to MRIs. Classical bootstrapping using empirical distributions to generate replicates of wind speed series is not adequate because v_{MRI} can be out of the range of the wind records and all replicates are constructed from the parameters in the range defined by the record. To overcome this limitation, we use parametric bootstrapping in which replicates are

Table 1. Parameters of the generalized Pareto distributions

	Sector 1 (10 º to 90 º)	Sector 2 (100 º to 180 º)	Sector 3 (190 º to 270 º)	Sector 4 (280 º to 360 º)
q_i	0.90	0.93	0.71	0.22
a_i	5.72 / 5.58	4.84 / 4.06	6.14 / 5.87	6.53 / 6.58
c_i	-0.15 / -0.24	-0.17 / -0.08	-0.24 / -0.31	-0.30 / -0.43
Adjusted c_i	-0.10 / -0.10	-0.10 / -0.08	-0.10 / -0.10	-0.10 / -0.10
q_i (synthetic data)	0.90	0.93	0.71	0.23
a_i (synthetic data)	5.88 / 5.04	4.76 / 3.61	6.14 / 5.11	6.58 / 5.70
c_i (synthetic data)	-0.11 / -0.10	-0.09 / -0.06	-0.10 / -0.08	-0.11 / -0.09

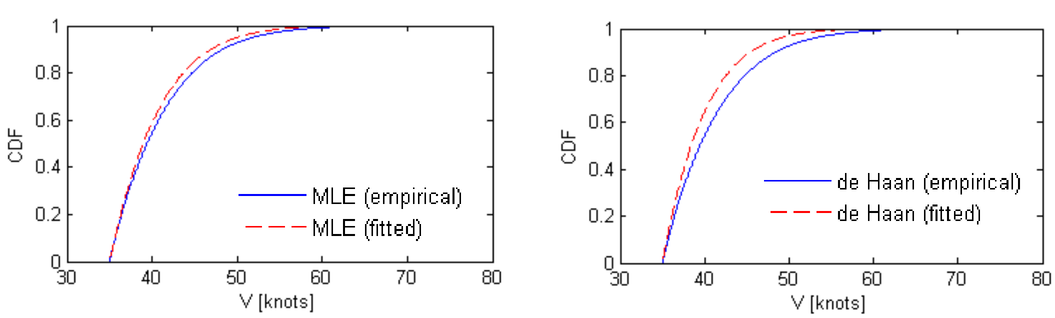

Figure 2. CDF of generated wind speeds (sector 4)

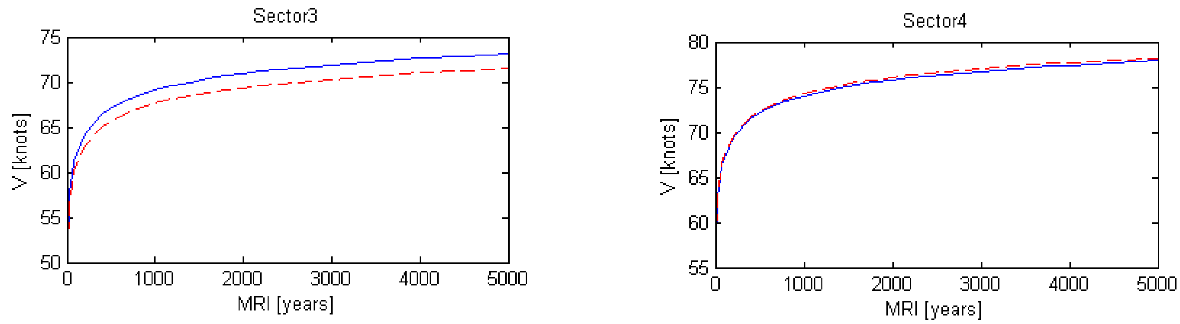

Figure 3. Generated wind speeds with MRIs (MLE; de Haan)

generated form probabilistic models for directional wind speed data calibrated to the record rather than empirical distribution given by the record. The Monte Carlo simulations are repeated 1000 times to generate 1000 replicates of the directional synoptic wind speeds at a site. Figure 4 shows the resulting scatter plot of parameter estimates in sector 4.

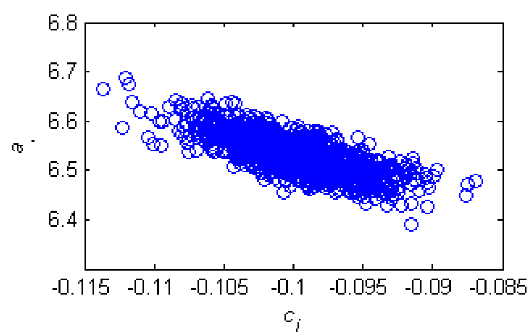

Figure 4. Realizations of estimated parameters (c_i, a_i)

Substituting the generated bootstrap sample values of (c_i, a_i) into Eq. (6) generates the corresponding bootstrap samples of directional wind speeds with any specific MRIs. Sample distributions of the wind speeds v_{MRI} are considered for 20-yr, 100-yr, 2000-yr, and 5000-yr MRIs. For sector 4 their histograms, and statistics of sample wind speeds, are plotted in Figure 5 and summarized in Table 2. The results, reported in detail in the paper, indicate that wind speeds with specified MRIs follow Gaussian distributions (Figure 5), and that the uncertainties in their estimation increase as the MRI increases (Table 2).

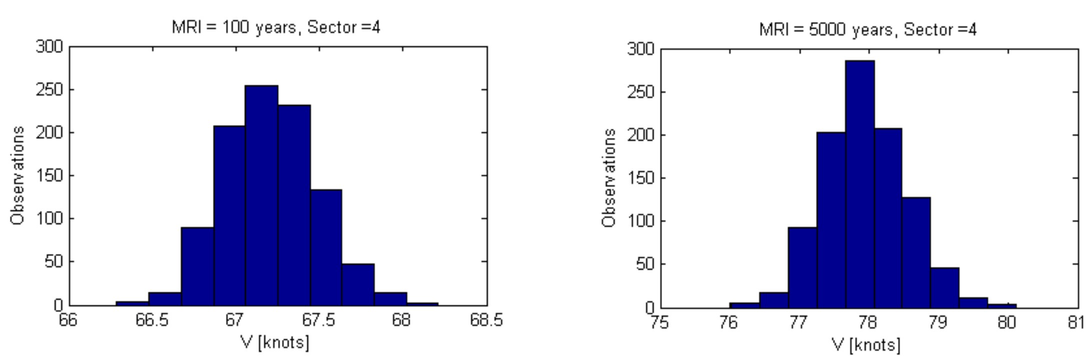

Figure 5. Bootstrap sample distributions of MRIs

Table 2. Estimates of wind speeds with MRIs [knots]

MRI [years]	Mean	Standard error	95% lower limit	95% upper limit
20	61.4	0.005	61.1	61.7
100	67.2	0.009	66.6	67.7
2000	75.8	0.016	74.7	76.7
5000	77.9	0.019	76.7	79.0

Note: Standard error is defined as $\sigma/\sqrt{n}$, where σ is the standard deviation of samples, and n is the number of samples

CONCLUSIONS

We proposed an algorithm to generate directional wind speeds of synoptic winds with large design MRIs by Monte Carlo simulations. The probability model in the study was based on a generalized Pareto distribution with the assumption of directionally independent wind speeds. The parameters of the distribution were estimated from the ASOS data by the maximum likelihood estimation and the de Haan method. Using the Monte Carlo simulation we generated, from the probabilistic model calibrated to the data, synthetic directional wind speeds with large MRIs. Uncertainties in the estimated wind speeds were also estimated. The methodology illustrated in this study is equally applicable to hurricanes and thunderstorm wind speeds.

ACKNOWLEDGEMENTS

This work is based on a methodology developed by Professor Mircea Grigoriu of the Department of Civil and Environmental Engineering, Cornell University. The author would like to thank Dr. Emil Simiu of NIST for useful discussions.

REFERENCES

ASCE (2010). *Minimum design loads for buildings and other structures*, American Society of Civil Engineers, Reston, VA.

de Haan, L. (1994). "Extreme Value Statistics." in *Extreme value theory and applications*, J. Galambos, J. Lechner, and E. Simiu, eds., Kluwer Academic Publishers, 93-122.

Galambos, J., Lechner, J., and Simiu, E. (1994). "Extreme Value Theory and Applications." Kluwer Academic Publishers.

Grigoriu, M. (2009). *Algorithms for generating large sets of synthetic directional wind speed data for hurricane, thunderstorm, and synoptic winds*. NIST Technical Note 1626, National Institute of Standards and Technology, Gaithersburg, MD.

Kotz, S., and Nadarajah, S. (2000). *Extreme value distributions: theory and applications*, Imperial college press, London.

NCDC (2008). *Data documentation for data set 3505 (DSI-3505)*: <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3505.pdf> (accessed 11/30/10).

Simiu, E., and Miyata, T. (2006). *Design of buildings and bridges for wind: a practical guide for ASCE-7 Standard users and designers of special structures*, John Wiley & Sons, Hoboken, NJ.

Spence, S. M. J. (2009). *High-rise database-assisted design 1.1 (HR_DAD_1.1): Concepts, software, and examples*. NIST Building Science Series 181, National Institute of Standards and Technology, Gaithersburg, MD.

Yeo, D. (2010). *Database-Assisted Design of high-rise reinforced concrete structures for wind: Concepts, software, and application*. NIST Technical Note 1665, National Institute of Standards and Technology, Gaithersburg, MD.