Reconstruction of gusty wind speed time series from autonomous data logger records

Javier Amezcua[1], Raúl Muñoz[2] and Oliver Probst*[3]

[1]*Department of Atmospheric and Oceanic Science University of Maryland, College Park 20742-2425, USA*
[2,3]*Physics Department, Instituto Tecnológico y de Estudios Superiores de Monterrey,*
Eugenio Garza Sada 2501 Sur, Monterrey, N.L., CP 64849, Mexico

Abstract. The collection of wind speed time series by means of digital data loggers occurs in many domains, including civil engineering, environmental sciences and wind turbine technology. Since averaging intervals are often significantly larger than typical system time scales, the information lost has to be recovered in order to reconstruct the true dynamics of the system. In the present work we present a simple algorithm capable of generating a real-time wind speed time series from data logger records containing the average, maximum, and minimum values of the wind speed in a fixed interval, as well as the standard deviation. The signal is generated from a generalized random Fourier series. The spectrum can be matched to any desired theoretical or measured frequency distribution. Extreme values are specified through a postprocessing step based on the concept of constrained simulation. Applications of the algorithm to 10-min wind speed records logged at a test site at 60 m height above the ground show that the recorded 10-min values can be reproduced by the simulated time series to a high degree of accuracy.

1. Introduction

Wind speed time series have been studied in recent literature for a series of reasons, including forecasting (Brown *et al*. 1984, Sfetsos 2000, Cadenas and Rivera 2007), analysis of extreme wind events (Durañona *et al*. 2007), and simulation of hourly wind speeds (Sahin and Sen 2001, Nafaoui *et al*. 2003, Aksoy *et al*. 2004). If real-time simulations are required (Shinozuka 1971, Monbet *et al*. 2007, Kareem 2008), e.g., for the analysis of built-up stuctures or wind turbines under stochastic loads, often an appropriate modeling of certain extreme value characteristics such as gust heights and shapes is necessary (Seong and Peterka 1998, Bierbooms and Cheng 2002, Bierbooms 2005).

While real-time information on wind speed, pressure and other related parameters can be obtained through dedicated measurement campaigns, many continuous remote data collection systems such as weather and environmental monitoring stations, wind resource assessment towers, and wind turbine Supervisory Command and Data Acquisition (SCADA) systems rely on

---

* Corresponding Author, Dr., E-mail: oprobst@itesm.mx

autonomous data logging units that generally store information in 1- to 10-min time intervals, retaining representative information of the interval such as the average, maximum, and minimum values, as well as the standard deviation. If the dynamics of a system monitored by a data logger such as a small wind turbine (Morales and Probst 2006, Elizondo *et al.* 2009) is to be studied, the information lost upon data reduction has be reconstructed. One approach is to use stochastical simulation to generate real-time wind speed series compatible with the wind speed information recorded by the data logger, feed the simulated time series into an appropriate simulator, obtain output information from the simulator for a given averaging interval, and then compare the measured and the averaged simulated data to extract free parameters in the model. The short-scale time series can be reconstructed if some plausible assumptions are made: (1) The wind speed is expected to follow a Gaussian process. (2) The autocorrelation function for a given data logger time interval should depend only on the time difference of two data points, i.e., the process is taken to be stationary. (3) The frequency spectrum should be compatible with the one exhibited by turbulent wind and therefore it should be possible to describe it by a Kaimal- or van Karman-type power density spectrum.

Although the conditions mentioned above allow to properly account for the average and standard deviation of the wind speed signal as well as the autocorrelation structure, maximum and minimum values cannot be adequately specified with this information alone. As pointed out by Bierbooms (2005) a desired gust amplitude could in principle be extracted from a sufficiently long simulated time series, but especially for extreme gust values the required time frame to be covered by the simulation may be too long to be practical. A suitable alternative is contrained simulation (Bierbooms and Cheng 2002, Bierbooms 2005), which allows to specify a desired gust amplitude with a realistic gust shape at a specific time location. Constrained simulation is based on random Fourier series first introduced by Shinozuka (1971) and a set of restrictions for the Fourier coefficients. Compliance with the constraints such as a desired gust amplitude (Bierbooms and Cheng 2002) or a wind speed jump with a specified rise time (Bierbooms 2005) is obtained by adding the signal produced by the constrained simulation onto a time series obtained from an unconstrained simulation. In the case of a desired gust amplitude, the knowledge of the autocorrelation function of the unconstrained time series is enough to fully specify the gusty signal (Bierbooms and Cheng 2002). Another approach, put forward by Seong and Peterka (1998), also uses an add-on process to generate a signal with specified spiky features. Their starting point is also a random Fourier series, but specification of the desired (anti-) gust features is achieved through an appropriate manipulation of the phases of the Fourier components, rather than putting constraints on the amplitudes. An exponential first-order autoregressive process is used to first generate a time series with the desired peak structure and frequency of occurrence; the signal is then Fourier-transformed to obtained the phases of the components of the random Fourier series, as opposed to uniformly distributed random phases as in Shinozuka (1971) and Bierbooms and Cheng (2002). While the method proposed by Seong and Peterka allows simulating non-Gaussian processes (as opposed to the one by Bierbooms and Cheng which relies on Gaussianity), its implementation is less straightforward and the extreme values can be adjusted to a lesser degree of accuracy than in Bierboom and Cheng's algorithm.

2. Methodology

The algorithm presented here is an engineering approach based on a three-step process: (1) A

time series compatible with the recorded average and standard deviation, as well as a specified power spectrum, is generated for a given averaging interval using stochastic simulation, using an approach similar to the one put forward by Shinozuka (1971). (2) Extreme values (compatible with the recorded maximum and minimum value) are synthesized by constrained simulation; for computational efficiency an analytical expression for the average autocorrelation function calculated for a Kaimal-type power spectrum is used to generate the gust signal. Gusts and anti-gusts are adjusted in a sequential manner. Two different strategies termed symmetric and asymmetric gust control are discussed. (3) Adjacent intervals are concatenated by a smooth transitioning process allowing for a continuous wind speed signal with a continuous derivative.

The initial wind speed signal is generated from a Fourier sum of the following form (Shinozuka 1971).

$$v(t) = \bar{v} + \sum_{i=1}^{N} A_i \cos(2\pi\, n_i t + \varphi_i)$$ (1)

where $\bar{v}$ is the mean wind speed in the interval and is maintained constant for a given data logger interval. The phase values φ_i are chosen from a uniform distribution on the interval $[0, 2\pi]$. The random amplitude values A_i are generated from a normal distribution centered in zero with a variable standard deviation (section 2.1). Non-Gaussian distributions can be incorporated by applying suitable power-law transformations (Aksoy *et al.* 2004, Brown *et al.* 2002). The frequency values n_i are generated with the help of a truncated Kaimal distribution (section 2.2). Kaimal-distributed n_i values are generated using the Integral Probability Theorem (Saucier 2000).

2.1 Regulation of the standard deviation

The standard deviation of the signal, σ_v, can be conveniently regulated by specifying the standard deviation of the amplitude σ_A. We first observe that from Eq. (1) we have

$$\sigma_v^2 = Var[v(t)] = \sum_i^N E[A_i^2] E[\cos^2(2\pi\, n_i t + \varphi_i)] - (E[A_i \cos(2\pi n_i t + \varphi_i)])$$ (2)

where the second term vanishes. The first term can be re-written as

$$\sum_i^N E[A_i^2] E[\cos^2(2\pi\, n_i t + \varphi_i)] = \frac{N\sigma_A^2}{2} + \sigma_A^2 \sum_i^N E\left[\frac{\cos(4\pi\, n_i t + 2\varphi_i)}{2}\right]$$ (3)

The second term vanishes, because the mean of a periodic function of a uniformly distributed variable is 0. Hence

$$\sigma_v^2 = Var(v(t)) = \frac{N\sigma_A^2}{2}$$ (4)

Therefore, the standard deviation of the wind speed signal is expected to be proportional to the standard deviation of the amplitude, with the proportionality constant depending only on the number of terms in the Fourier sum.

2.2 Specification of the autocorrelation function and power density spectrum

As stated earlier, apart from having the correct mean and standard deviation, the simulated signal is required to have a power density spectrum and (equivalently) an autocorrelation function compatible with turbulent wind. In order to achieve this goal we first observe that the covariance of the wind speed signal can be expressed as

$$\mathrm{Cov}[v(t), v(t + \tau)] = E\left[\sum_{i,j}^{N} A_i A_j \cos(2\pi\, n_i t + \varphi_i)\cos(2\pi\, n_j(t + \tau) + \varphi_j)\right] \tag{5}$$

The double sum can be split up into a diagonal part C (where $i = j$) and a non-diagonal part D (where $i \neq j$).We then obtain

$$C = \sum_{i}^{N} E[A_i^2]E[\cos(2\pi\, n_i t + \varphi_i)\cos(2\pi\, n_i(t + \tau) + \varphi_i)] \tag{6}$$

$$D = \sum_{i=1}^{N}\sum_{j=1}^{N} E[A_j]E[A_j]E[\cos(2\pi\, n_i t + \varphi_i)\cos(2\pi\, n_j(t + \tau) + \varphi_j)] \tag{7}$$

where $D = 0$ due to the fact that A_i and A_j are uncorrelated and have zero mean. With the use of an angle-addition identity, Eq. (6) can be rewritten as

$$C = \sum_{i}^{N} \sigma_A^2 \frac{1}{2}E[\cos(-2\pi\, n_i\tau) + \cos(2\pi\, n_i(2t + \tau) + 2\varphi_i)] \tag{8}$$

Expanding the second cosine term and observing that $E[\cos(2\varphi_i)] = E[\sin(2\varphi_i)] = 0$ we obtain

$$\mathrm{Cov}[v(t), v(t + \tau)] = \sum_{i}^{N} \frac{\sigma_A^2}{2}E[\cos(2\pi\, n_i\tau)] \tag{9}$$

Using the definition of the expectation value of a continuous variable we have

$$\sum_{i}^{N} \frac{\sigma_A^2}{2}E[\cos(2\pi\, n_i\tau)] = \frac{N\sigma_A^2}{2}\int_{0}^{\infty} f(n)\cos(2\pi\, n\tau)dn \tag{10}$$

where *f(n)* is the frequency probability distribution function. If we now recall the definition of the autocorrelation function

$$\kappa(\tau) = \frac{\mathrm{Cov}[v(t), v(t + \tau)]}{\mathrm{Var}[v(t)]} \tag{11}$$

we obtain

$$\kappa(\tau) = \frac{N\sigma_A^2}{2}\int_{0}^{\infty} f(n)\cos(2\pi\, n\tau)dn \times \left(\frac{N\sigma_A^2}{2}\right)^{-1} = \Im_{\cos}[f(n)] \tag{12}$$

i.e., the autocorrelation function is equal to the Fourier (cosine) transform of the frequency distribution. On the other hand, the autocorrelation function and the spectral power density $S_v(n)$ form a Fourier pair (see e.g., Childers (1997))

$$S_v(n)/\sigma_v^2 = 4 \int_0^\infty \kappa(\tau)\cos(2\pi n\tau)\mathrm{d}\tau \tag{13}$$

$$\kappa(\tau) = \int_0^\infty S_v(n)/\sigma_v^2 \cos(2\pi n\tau)\mathrm{d}n \tag{14}$$

which is known as the Wiener-Khintchin relationships (Childers 1997). Therefore, the frequency distribution and the power density function of the signal, normalized to the standard deviation of the wind speed, are equal

$$f(n) = S_v(n)/\sigma_v^2 \tag{15}$$

    Common choices for the power density function are the Kaimal and the van Karman functions (Burton *et al*. 2001), where Kaimal-type expressions are used in most wind engineering codes including wind turbine standards (International Electrotechnical Commission 1998). The Kaimal expression for the spectral power density is given by (Kaimal *et al*. 1972, Kaimal and Finnigan 1994)

$$\frac{S_v(n)n}{\sigma_v^2} = \frac{4nL_1/\bar{v}}{1 + 6nL_1/\bar{v}^{5/3}} \tag{16}$$

where L_1 is a characteristic length of the order of 100 to 250 m. This expression was derived from field data in Kansas and applies for flat terrain and near-neutral surface boundary layer. For mildly complex terrain a perturbation theory approach can be used to incorporate the effect of orographic features, such as a hill (Mann *et al*. 2000, sections 2.4 and 2.6). Similarly, roughness changes can be modeled by using the concept of the internal boundary layer hill (Mann *et al*. 2000, section 2.5). An analytical expression for the autocorrelation function corresponding to Eq. (16) can be derived with the result shown in Eq. (17)

$$\kappa(\tau) = {}_1F_2\left[1, \left\{\frac{1}{6},\frac{2}{3}\right\}, -\frac{\pi^2\tau^2}{36\alpha^2}\right] - \frac{2\pi^{5/3}}{9\Gamma[5/3]}\left(\frac{\tau}{3\alpha}\right)^{2/3}\left(\sqrt{3}\cos\left(\frac{\pi\tau}{3\alpha}\right) - 3\sin\left(\frac{\pi\tau}{3\alpha}\right)\right) \tag{17}$$

where $\alpha = L_1/\bar{v}$, ${}_pF_q$ represents the generalized hypergeometric function and Γ is the gamma function. For numerical purposes the Kaimal distribution is not ideal, however, since it is a heavy-tailed distribution, giving rise to numerical instability problems. A more numerically appropriate version is the truncated Kaimal distribution defined by

$$f_t(n) = \frac{1}{\gamma}\left[\frac{4L/\bar{v}}{(1 + 6L/\bar{v}n)^{5/3}}\right] \quad 0 \le n \le n_{\max} \tag{18}$$

where

$$\gamma \,=\, 1 - \frac{1}{(1 + 6L/\bar{v}n_{\max})^{2/3}} \tag{19}$$

and $n_{\max}$ is a cut-off frequency. For frequencies greater than $n_{\max}$, f_t is set to zero. For typical values of α and $n_{\max}$, the normalization factor γ is generally close to 1. For example, for $\alpha = 12\ s$ and $n_{\max} = 5\ Hz$, we have $\gamma = 0.9355$.

2.3 Regulation of the extreme values of the signal

As stated earlier, the time series generated from Eq. (1), though having defined expressions for mean, standard deviation and frequency distribution, do not have specified values of the maximum and minimum values within the simulation interval. One approach of overcoming this restriction is to run a postprocessing algorithm on each of the sample functions generated from (1), adjusting the initially generated extreme values to the maximum and minimum observed in the experimental (data logger) sequences. As mentioned before, the method used is an adaptation of the concept of constrained simulation put forward by Bierbooms and Cheng (2002). In their paper the authors propose to modify the simulated wind speed time series in the vicinity of a gust occurring in t_0 according to the following formula

$$v_c(t) \,=\, v(t) + (v_{\text{gust}} - v(t_0))\kappa(t - t_0) + \frac{\dot{v}(t_0)\dot{\kappa}(t - t_0)}{-\ddot{\kappa}(0)} \tag{20}$$

where v_{gust} is the desired gust amplitude at t_0, in our case the actually measured maximum wind speed value in the interval, and $k(\tau)$ is the normalized autocorrelation function (ACF) of the simulated signal as before. The autocorrelation function of the originally simulated signal can in principle be calculated numerically. For greater computational efficiency, however, we use the analytical expression for the ensemble average ACF (Eq. (17)).

In our algorithm, we first identify the time locations and the values of the maximum and minimum obtained from the unconstrained simulation of a given sample function and a given logger interval. The simulated values are compared to the recorded ones and are potentially subject to adjustments. Two gust control strategies have been implemented; in both cases first the maximum and minimum wind speed values are determined for a given interval (with a duration of typically 1 or 10 mins) from the data logger time series. The maximum and minimum values of the simulated time series are calculated and compared to the measured values. In case the measured maximum is greater than the simulated maximum and, correspondingly, the measured minimum is smaller than the simulated minimum, then an additional gust signal according to Eq. (20) is generated and summed to the signal.

In case the observed maximum was below the simulated maximum or conversely, the observed minimum was above the simulated minimum, two alternative strategies have been explored: In the first algorithm, termed the asymmetric gust control process, no adjustment was made and the simulated maxima or minima are used as provided by the original simulation. In the second algorithm, which we will call the symmetric gust control strategy an anti-gust is generated to adjust the simulated maximum downward; a corresponding measure is taken in the case of the minima. An illustration of the effect of the gust-controlling post-process is given in Fig. 1 where a piece of a sample function simulated for a 10-min time interval has been plotted in the vicinity

of the highest wind speed value in the interval. The input data for the simulation were a mean of 10 m/s, a standard deviation of 1.5 m/s, a minimum of 2.5 m/s and a maximum of 17.5 m/s. It can be seen that the "natural" process (given by Eq. (1)) falls slightly short of the requirement of the maximum wind speed with a maximum value of only about 16 m/s, as opposed to the required value of 17.5 m/s. It can be seen from Fig. 1 how the gust-controlling post-process adjusts the maximum, while preserving the fluctuations of the signal and converging towards the original signal at a distance of about 15 s from the center location of the gust. The minimum is adjusted in a similar way by treating it as a negative or anti-gust.

2.4 Concatenation of the intervals

As stated earlier, time series are simulated for one data logger averaging interval at a time to account for the recorded values of the mean, maximum, minimum, and standard deviation. Upon concatenating adjacent intervals, however, continuity of each simulated sample function and its derivative has to be assured explicitly, requiring a second type of postprocessing. To this end a smooth transitioning process has been implemented based on the following sequence. First, the wind speed difference between the final values of the interval i and the first value of the interval $i+1$ is computed: $\Delta v = v_1^{i+1} - v_{1023}^i$. A smooth transition between the two intervals is accomplished by modifying the last ten wind speed values of interval i and the first ten values of the interval $i+1$ according to

$$v_{1023-k}^{i,\mathrm{mod}} = v_{1023-k}^i + \frac{\Delta v}{20}(10 - k) \qquad \text{for } k = 9 \text{ to } 0 \tag{21}$$

$$v_{1+k}^{i+1,\mathrm{mod}} = v_{1+k}^{i+1} + \frac{\Delta v}{20}(10 - k) \qquad \text{for } k = 0 \text{ to } 9 \tag{22}$$
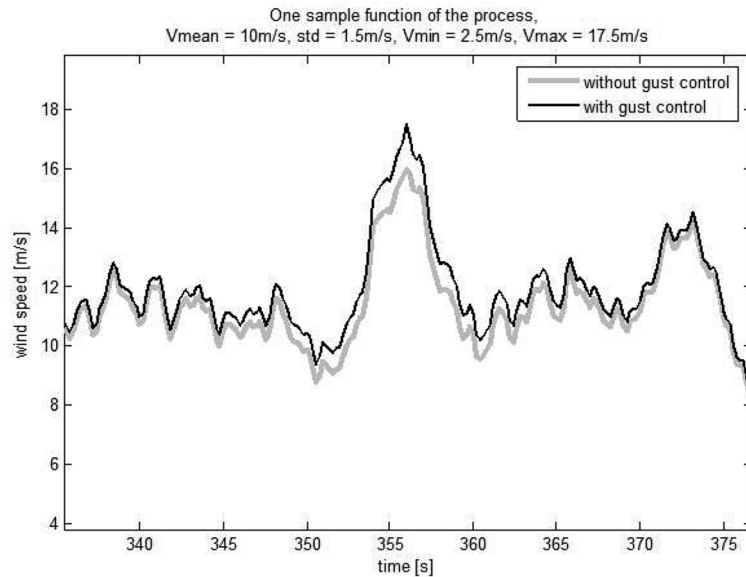

Fig. 1 Illustration of the gust control algorithm applied to one sample function of the [anti-symmetric] process described by Eq. (1)

Since only 20 wind speed values out of 1024 in each data logger averaging interval are subject to modifications the effect of this postprocessing step on the specified values of mean and standard deviation in each interval are generally small, although dependent on difference of wind speeds of adjacent intervals.

3. Results and discussion

3.1 Wind speed distribution and standard deviation

We first verified that the wind speed distribution obtained from Eq. (1) for a given interval is indeed Gaussian and that the standard deviation, σ_v, can be adjusted by specifying the standard deviation, σ_A, of the underlying amplitude distribution. It can be seen from Fig. 2, corresponding to one 10-min sample function with a mean of 9.6 m/s and a standard deviation of 1.4 m/s and 100 sample functions with the same parameters that the wind speed distribution for a sufficiently large number of sample functions indeed approaches normality, consistent with the Central Limit Theorem (Priestly 1981).

As to the specification of the standard deviation of the wind speed within a given interval, we observe from Fig. 3 that indeed a linear relationship between σ_v and σ_A exists as predicted by Eq. (4). In Fig. 3 the average standard deviation of the wind speed calculated from an ensemble of 400 sample functions has been plotted vs. σ_A ; N=100 cosine terms were used to generate the signal. It can be seen from the figure that the theoretical relationship $\sigma_v = (N/2)^{1/2}\sigma_A = 7.07\sigma_A$ is a good description of the data. It should be noted, however, that the spread of σ_v among different sample functions is higher for higher values of σ_A , so that less control exists over this parameter for high turbulence levels. Whereas for low levels of σ_A the wind speed standard deviation σ_v can be specified to an accuracy of a few percent, a spread of the order of 10-20% may occur at higher values.

3.2 Spectral power density

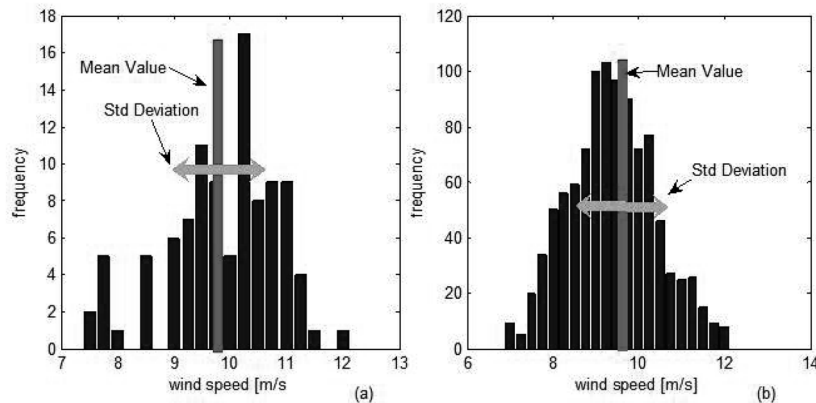In order to verify that the simulated time series generated from Eq. (1) using the truncated

Fig. 2 (a) Wind speed distribution for one 10-minute sample function with a mean speed of 9.6 m/s, and a standard deviation of 1.4 m/s and (b) Wind speed distribution for 100 sample functions
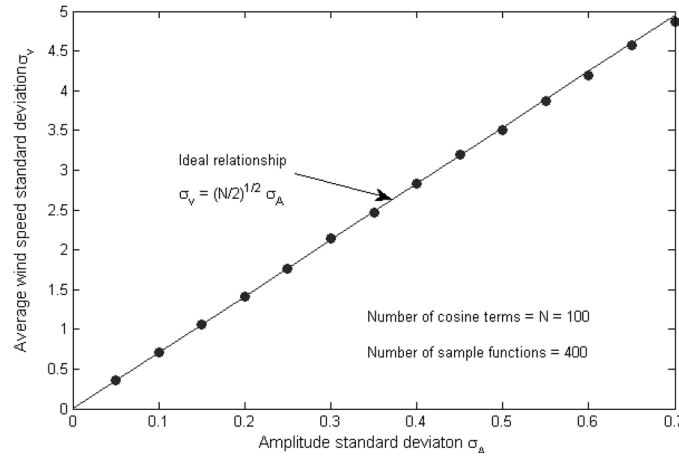
Fig. 3 Standard deviation of the simulated signal as function of the standard deviation of the amplitude values of the random Fourier series

Kaimal distribution function (18) has the spectral characteristics expected from turbulent wind the normalized spectral power density $S_v(n)n/\sigma_v^2$ was calculated numerically and compared to the theoretical curves obtained from Eq. (17). Fig. 4 shows an ensemble average of 100 sample functions of the simulated normalized spectral power density compared to the theoretical prediction obtained from the Kaimal expression. No postprocessing (adjustment of extreme values and concatenation of intervals) was performed in this case. Both for the truncated distribution function f_t and the normalized spectral power density a characteristic length of $L = L_1 = 180$ m was used. The cut-off frequency was taken as $n_{max} = 5$ Hz. It is conspicuous that the spectrum of the simulated time series adjusts itself well to the theoretical expression, even though at high frequencies the simulated power density was found to be slightly lower than the Kaimal prediction. The agreement can be improved if the characteristic length L of the distribution
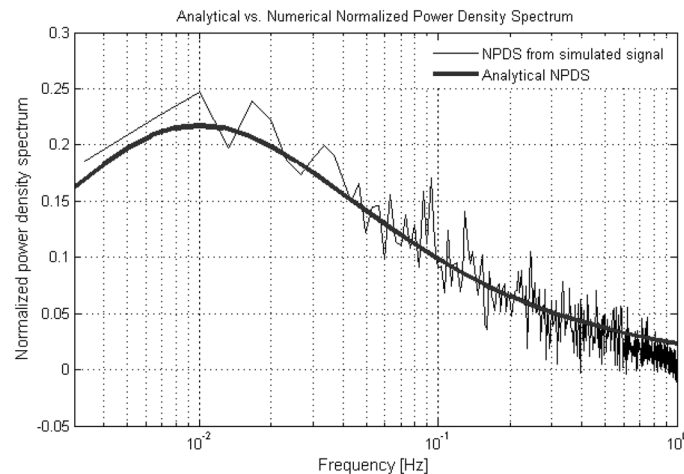

Fig. 4 Ensemble average (100 sample functions) of the normalized spectral power density (NPDS) of the simulated time series with no postprocessing compared to the Kaimal NPDS. In both cases a characteristic length of $L = L_1 = 180$ m was used

function is taken as a free parameter and a least-mean square fit to the theoretical curve is performed. A best fit to the data can be described by the linear relationship $L1 = 1.02L$, where the slight deviation from the equality of the two characteristic lengths can be attributed to the truncation of the frequency distribution.

3.3 Autocorrelation function

As stated earlier, the choice of the frequency distribution $f(n)$ implies a corresponding spectral power density which was found to be consistent with the theoretical expectation. It is interesting to verify that the time-domain equivalent, the normalized autocorrelation function (ACF) is also consistent with the theoretical expression (Eq. (17)). To do this, the ACF was calculated using Matlab for 300 sample functions with time lags between 0 and 512 time units (taken as 600s/ 1024) and subsequently the ensemble average was taken. The result is shown in Fig. 5, where the ensemble-average ACF is compared with the analytical ACF $\kappa(\tau)$ (Eq. (17)) which was calculated from the Kaimal spectrum using the Wiener-Khintchin relationship. It is conspicuous that a good agreement exists, consistent with the good agreement of the theoretical and the simulated spectra.

3.4 Validation of the Kaimal spectrum assumption

Since the time series created by the algorithm described in this work are based on the hypothesis of a Kaimal spectrum (Eq. (16)) it is interesting to compare this assumption against real-time wind speed data. We have recently conducted experiments with an instrumented small-scale wind turbine in the Ventosa region in Southern Mexico where wind speed and turbine performance data were sampled with a frequency of 1 Hz (Elizondo *et al.* 2010). The site is characteristics by wide-open land with sparse vegetation cover. Wind speed measurements were taken at a height above ground level of 20 m. In order to assess the wind speed turbulence spectrum we concentrated on a representative day with no extreme meteorological events, chosen to be July 24, 2010. The available information was grouped into 58 samples of 10 minutes each
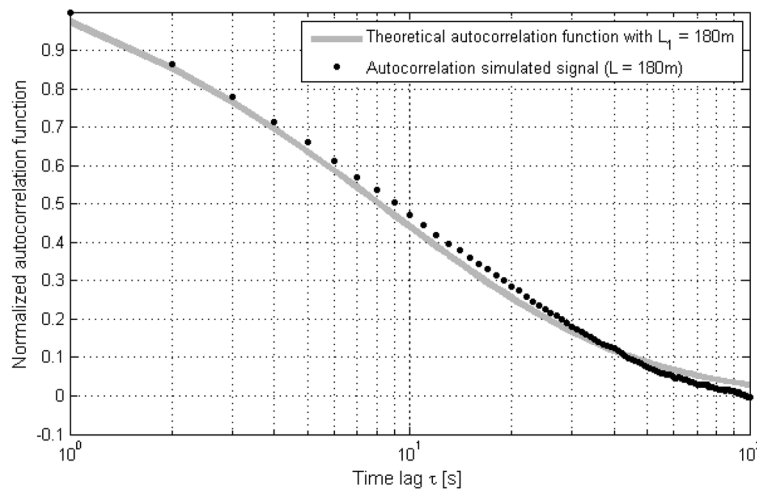

Fig. 5 Normalized autocorrelation function (NACF) calculated from 300 sample functions of the simulated signal compared to the analytical NACF

and the autocorrelation function was calculated for each of these sample functions. The individual autocorrelation functions are shown in Fig. 6(a). In order to obtain a fit of the theoretical expression to the data the empirical ensemble-average autocorrelation function was subjected to a Hamming-Tukey tapering with a window of 300 lags. The corresponding curve is shown as the black line in Fig. 6(a). The best fit of the theoretical expression to the data (heavy gray line in Fig. 6(a)) provide a characteristic length of L_1=124 m, which is a plausible value. For instance, the Danish standard DS 472, 1992 (see Burton *et al*. (2001) for a discussion) recommends a longitudinal length scale of $5z$ (z being the observation height) for heights below 30 m, which translates into L_1=100 m for our observation height of 20 m. From the empirical ensemble-average autocorrelation function the normalized spectrum (Eq. (16)) was calculated, with the result shown in Fig. 6(b). It can be seen that a good agreement with the Kaimal spectrum for L=124 m for a wide range of frequencies from about 2×10^{-2} Hz to 3×10^{-1} Hz was obtained. It should be stated, however, that this good agreement was expected for well-behaved wind climates
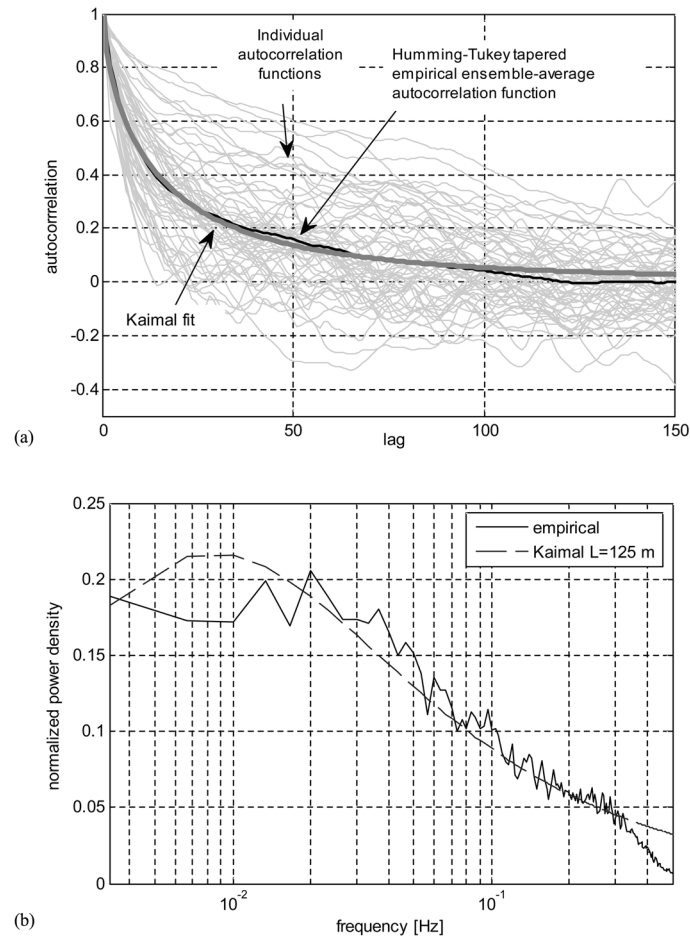

Fig. 6 Turbulence characteristics of wind speed data measured at La Ventosa (Southern Mexico) with a sample frequency of 1Hz. (a) Normalized autocorrelation function for 58 sample functions vs. lag in seconds and Kaimal fit and (b) Normalized power density spectrum and Kaimal fit.

only and is not likely to apply in the case of extreme wind events such as downbursts or tropical storms. A more systematical analysis of the results obtained at the La Ventosa site will be published in a follow-up paper.

3.5 Assessment of the impact of the asymmetric gust-controlling process: simulated data

In order to assess the effect of the gust-controlling post-process on the statistical properties of the signal specified by the original process we have conducted five simulation experiments, the parameters of which are given in Table 1. Only the asymmetric process is considered in this section. An average wind speed of 10 m/s was used for all simulations, the variance was taken as σ^2=2.25 m^2/s^2, N=100 terms in the random Fourier sum (Eq. (1)) were used, and M=500 sample functions were generated for each process. Ensemble averages of the mean and the standard deviation were calculated for all processes and then compared to the original specifications.

We first looked at the number of sample functions that were modified by the gust-controlling post-process. It can be seen from Fig. 7 that in the case of the simulations 1 and 2 (maximum and minimum at one or two standard deviations from mean) practically no modification of the process occurs. This changes drastically if the measured gusts (maxima) and anti-gusts (minima) are three or more standard deviations away from the mean. As we can expect from the Gaussian distribution of the amplitudes in Eq. (1), it will be rare to encounter to encounter such outliers and effectively, as demonstrated by Fig. 7, about half of the 500 sample functions have to be modified to account for either an adjustment in the maximum or the minimum. In about 20% of the cases both maxima and minima are affected. As expected, for extreme values even farther away from the mean (processes 4 and 5), practically all processes have to be modified. In real-world processes, we would of course rarely expect extreme values with these characteristics.

The impact of the gust control algorithm on the mean value and the variance of the process is illustrated in Table 2. As expected for symmetric extreme values, almost no modification of the mean values occurs. The variance, on the other hand, does experience some changes in the cases of the processes 4 and 5; under normal circumstances (extreme values no further away from the mean than three standard deviations), however, the variance does not change by more than 0.3%.

Our final assessment of the effect of the (asymmetric) gust control algorithm in the case of simulated data referred to spectral characteristics, as measured by the normalized autocorrelation function. As demonstrated by Fig. 8, no modification of the autocorrelation function occurs, as long as the extreme values are located at a maximum distance of three standard deviations from the mean. In other words, the spectral characteristics of the signal are not affected by the gust-control algorithm. As before, some modifications of the ACF occur in the case of the processes 4 and 5.

Table 1 Parameters of simulation experiments conducted to assess the effect of the gust-controlling post-process

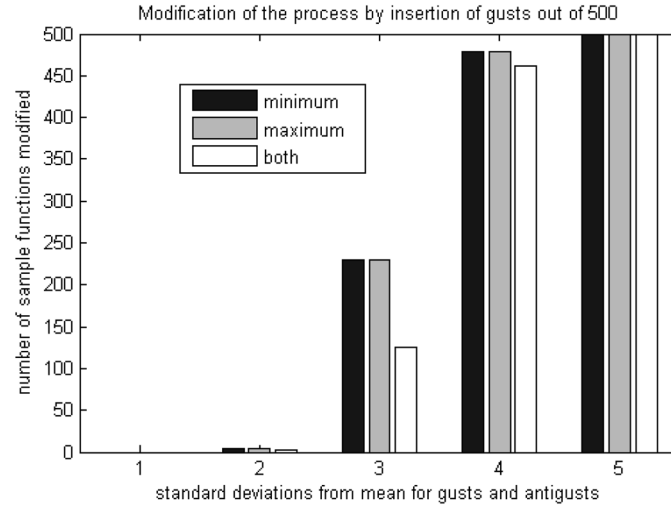| Process	$|v_{min, max} - \bar{v}|$ [m/s]	$|v_{min, max} - \bar{v}|/\sigma$	v_{min} [m/s]	v_{max} [m/s]
1	1.5	1	8.5	11.5
2	3	2	7	13
3	4.5	3	5.5	14.5
4	6	4	4	16
5	7.5	5	2.5	17.5

Fig. 7 Number of sample functions (out of 500) modified by the gust-controlling post-process

Table 2 Modification of the process mean and variance due to the addition of the gust-controlling process

Process	$\bar{v}_{with\ gusts} - \bar{v}_{no\ gusts}$ [m/s]	$\dfrac{\bar{v}_{with\ gusts} - \bar{v}_{no\ gusts}}{\bar{v}_{no\ gusts}}$	$\sigma^2_{with\ gusts} - \sigma^2_{no\ gusts}$ [m²/s²]	$\dfrac{\sigma^2_{with\ gusts} - \sigma^2_{no\ gusts}}{\sigma^2_{no\ gusts}}$
1	0	0	0	0
2	-2.4×10^{-5}	-2.4×10^{-4} %	-0.0001	-0.004%
3	2.6×10^{-4}	2.6×10^{-3} %	0.0072	0.32%
4	-9.5×10^{-4}	-9.5×10^{-3} %	0.1552	6.89%
5	-2.7×10^{-4}	-2.7×10^{-2} %	0.5128	22.79%

3.6 Assessment of the impact of the symmetric and asymmetric gust-generating processes: measured data

After studying the effect of the gust-control process on synthetic time series we focused on the original objective of this work, the reconstruction of time series from logger data streams. As a study case we used data from a wind resource assessment project in Northern Mexico where wind speed data are recorded in 10-min time intervals using a NRG Symphonie data logger and anemometers at three measurement heights. For the purpose of the present study the records at 60 m have been used. The data logger records data at a sample frequency of 1Hz and then stores the average value, the standard deviation, the maximum, and the minimum for each channel and fixed 10-min intervals in an internal SD-type memory. After storage the original records are discarded. Fig. 9 shows the results for the 60 m channel for a 6000-interval period (1000 hours) during the period of June/July of 2008. This period corresponds to the normal wind climate; no extreme wind events such as tropical storms are contained in the time period. We therefore believe that the assumption of a stationary process is reasonable for the complete period.

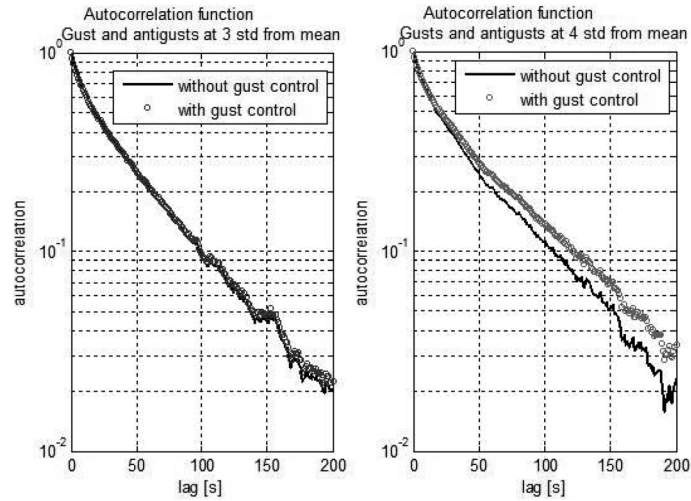In the first step of our analysis we have assumed that the intervals in the observation periods

Fig. 8 Modification of the normalized autocorrelation function due to the gust-controlling post-process for the case of gusts at three and four standard deviations from the mean

can be considered as independent realizations of the same process with varying values of mean, standard deviation, maximum, and minimum. We then created simulated time series for each time interval, both with and without gust control, and analyzed the distributions of the four variables. Both the asymmetric and the symmetric gust control process have been implemented. We will first discuss the results obtained with the asymmetric gust control algorithm.

As shown in Figs. 10 and 11, the distributions of the measured and the simulated data (both with and without asymmetric gust control) are in good agreement for all four variables. It is especially noteworthy that the gust-control process does not significantly alter the overall distribution of the extreme values (Fig. 10). If the discrepancies between the measured and the predicted values of the four variables are analyzed, however, it becomes apparent how the gust-control process significantly improves the agreement between measured and simulated extreme values (Fig. 12), while leaving
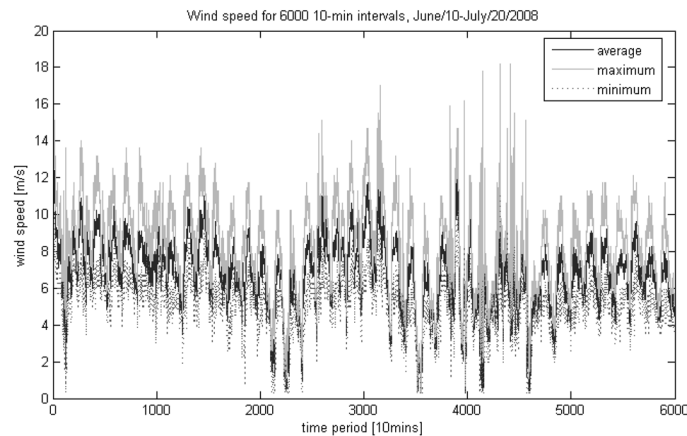

Fig. 9 Data logger records for average, maximum, and minimum wind speed values for 6000 10-min intervals recorded at a potential wind farm development site at 60 m above the ground

the averages and standard deviations unchanged (Fig. 11). Specifically, we note that the gust and anti-gust distributions now peak at a relative error of zero between measured and simulated values and decay more rapidly towards zero than the error distributions calculated for the original process. It is also conspicuous that the gust and anti-gust distributions are skewed due to the inherent asymmetry of the gust control process.

In a second step, the simulation intervals were concatenated to provide a smooth transition between intervals. Even though the extreme values are unlikely to be affected by this procedure, some modification of the interval averages and variances could occur. In Fig. 13 we have assessed this effect by plotting the distributions of alterations of both mean and variance due to the concatenation process. It can be seen that in both cases the error distribution is symmetrical with a large peak at zero error. The standard deviation of the error distribution of the mean is of the order of 0.01 m/s or about 0.1%, so the effect of the concatenation can be largely neglected.
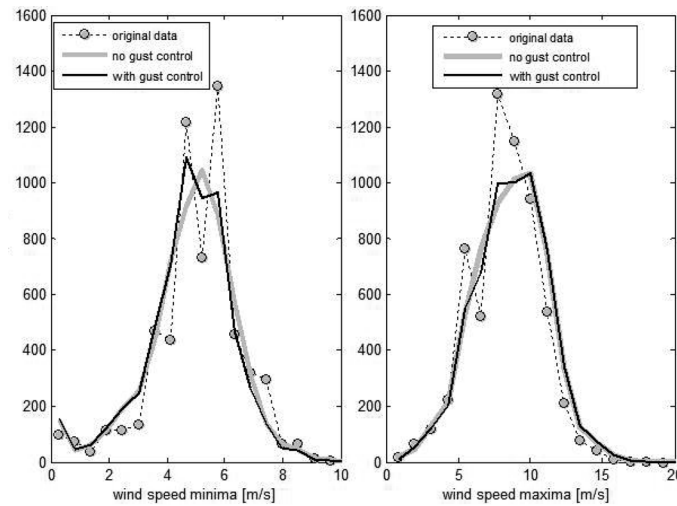

Fig. 10 Distribution of maximum and minimum calculated for 10-min intervals for the original data logger records and the simulated time series, both with and without asymmetric gust control
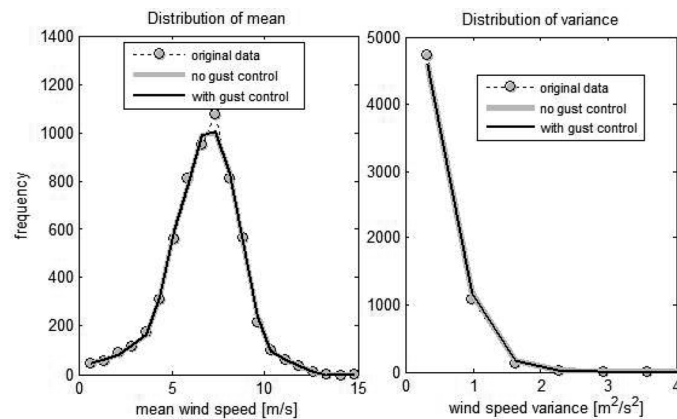

Fig. 11 Distribution of mean and variances calculated for 10-min intervals for the original data logger records and the simulated time series, both with and without asymmetric gust control
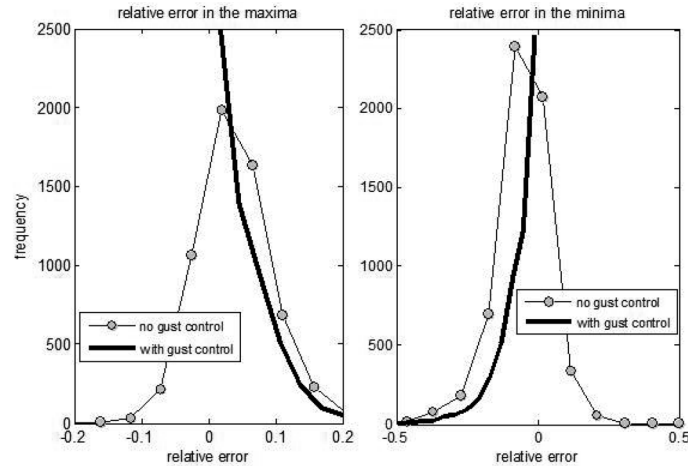
Fig. 12 Distributions of the relative error of the simulated time series (both with and without gust control) compared to the measured data logger records with respect to the variables minimum and maximum. Asymmetric gust control case

Similarly, the modification of the variance is negligible.

The results of the simulations obtained directly from Eq. (1), with asymmetric gust control, and with gust control plus concatenation have been summarized in Table 3 together with the corresponding results for the original logger data. It is conspicuous that the ensemble averages of the mean as well as the ensemble standard deviation are practically identical for the measured data stream and the simulated time series in all three cases. Also, the standard deviation is very well reproduced by the original process and very little modification occurs due to the gust post-processing and concatenation. It can also be observed, however, that all simulated processes overpredict the gusts and underpredict the anti-gusts, respectively. The main reason lies with the original process itself which produces gusts with an average of about 3% higher that observed experimentally. One plausible explanation lies with the low-pass characteristics of the ane⁻mometer and the logging system which tend to reduce the range of the extreme values. In case the length constant of the anemometer and the transfer function of the logging system are well
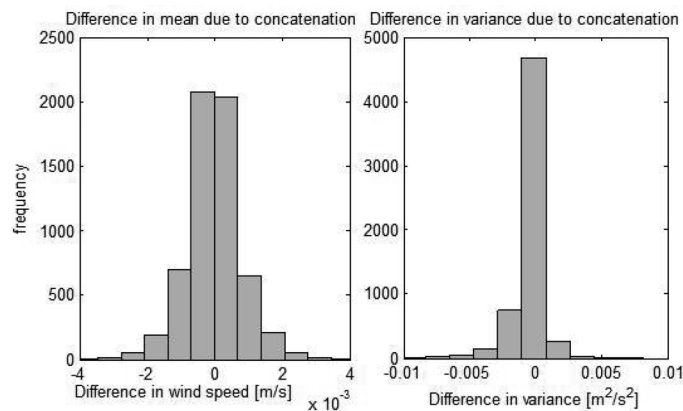

Fig. 13 Effect of the concatenation post-process on the mean and variance of the simulated time series

Table 3 Simulation results based on a real data stream (ensemble averages and standard deviations for 6000 sample functions) for the asymmetric gust control process

	Mean [m/s]		Max [m/s]		Min [m/s]		Std [m/s]
	Ensemble average	Ensemble standard deviation	Ensemble average	Ensemble standard deviation	Ensemble average	Ensemble standard deviation	Ensemble average
Original data	6.7314	1.9353	8.3158	2.4567	5.1062	1.5154	0.6843
Synthetic series	6.7345	1.9395	8.6421	2.5700	4.8386	1.5872	0.6699
Synthetic series with gusts	6.7343	1.9387	8.6994	2.5787	4.7722	1.5515	0.6714
Synthetic series with gusts and concatenation	6.7343	1.9387	8.7006	2.5779	4.7737	1.5520	0.6710

known, then the simulated time series can be filtered accordingly before comparing the simulated and the recorded extreme values. A detailed assessment of these effects is not subject of this article; for a discussion the reader is referred to Verkaik (2000).

An option for correcting the situation described above is the use of the symmetric gust control process. The results are illustrated in Fig. 14 where the relative error distributions for the maxima and minima for the symmetric gust control algorithm have been plotted. It is conspicuous that now the maxima are perfectly reproduced by the algorithm, while the minima (though greatly improved) still have an error distribution with a finite width. The reason for this behavior lies with the necessity of processing the maxima and minima in a sequential way, correcting the minima first and the maxima afterwards. In this case we have chosen to reproduce the maxima in an optimum way, while allowing for a small discrepancy in the minima.

A summary of the results for the symmetric gust control case is shown in Table 4. As before, the ensemble average of the mean and its standard deviation is very close to the measured data and the maxima are reproduced with high accuracy. As mentioned before, the minima have a
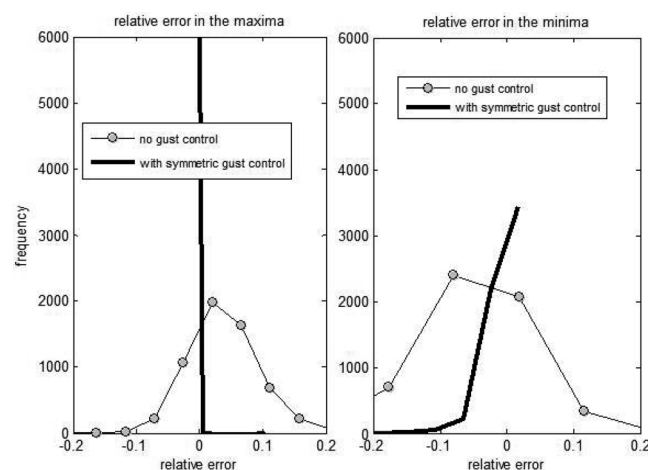

Fig. 14 Distributions of the relative error of the simulated time series (both with and without gust control) compared to the measured data logger records with respect to the variables minimum and maximum. Symmetric gust control case

small discrepancy which amounts to about 0.9%. Generally, applications of gusty wind speed signals focus on the gusts and their distributions, so that this small error should be tolerable for most applications. It is noteworthy, however, that the standard deviation is reduced by the symmetric gust control algorithm by about 10% (as opposed to the asymmetric gust control process). Applications that rely on the precise value of the standard deviation should therefore consider using the asymmetric gust control process.

4. Applications

As stated earlier, the main application of the algorithm presented in this work lies with the study of dynamic systems where the wind is the main driver and all variables (both input and output) are available for fixed logging intervals, typically 10 minutes or 1 minute. An example is the field evaluation of small wind turbines (Morales and Probst 2006, Elizondo *et al.* 2008), where wind speed and direction are recorded together with the AC and/or DC output power of the turbine. If a decoder is available, the turbine yaw angle can be recorded, too. Now, the time scales for the turbine dynamics, such as the change in rotational frequency in response to variations in wind speed or the rotation around the vertical axis in order to reduce the cross section of the swept area perpendicular to the wind (known as *furling*) are of the order of seconds and therefore well below the smallest recording interval of the data logger (typically 1 minute). One possible approach is now to generate an ensemble of sample functions for the wind speed compatible with the recorded values and then feed each sample time series into a dynamic computer model of the system and obtain a corresponding time series for the output variable under study (say, the yaw angle or the power delivered to the battery bank). After calculating average values, standard deviations, and extreme values of the output variable these values can be compared to the values actually recorded and those simulation events that are compatible with both the recorded input and output values can be identified.

As an illustration of this approach, examples of the simulated *furling* dynamics of a small wind turbine are shown in Figs. 15 and 16, respectively, where the input variable is the wind speed and the output variable the turbine yaw angle. Details of the system and the simulator were published elsewhere (Audierne *et al.* 2009). In the event shown in Fig. 15 the wind speed can be seen to

Table 4 Simulation results based on a real data stream (ensemble averages and standard deviations for 6000 sample functions) for the symmetric gust control process

	Mean [m/s]		Max [m/s]		Min [m/s]		Std [m/s]
	Ensemble average	Ensemble standard deviation	Ensemble average	Ensemble standard deviation	Ensemble average	Ensemble standard deviation	Ensemble average
Original data	6.7314	1.9353	8.3158	2.4567	5.1062	1.5154	0.6208
Synthetic series	6.7345	1.9395	8.6421	2.57	4.8386	1.5872	0.6049
Synthetic series with gusts	6.7222	1.9266	8.3159	2.4566	5.06	1.526	0.5503
Synthetic series with gusts and concatenation	6.7222	1.9266	8.3179	2.4548	5.06	1.5271	0.55

fluctuate between about 7 and 13 m/s with almost no response from the yaw angle which remains at its parking position of about -3°, except for small oscillations around the equilibrium value. In the event in Fig. 16, on the other hand, the yaw angle responds to gusts in the range of about 14 to 16 m/s by quickly transitioning to its maximum value of about 68°. Interestingly, the system shows some hysteresis as evidenced by the fact that the turbine remains in the furled position for about 10 to 20 seconds after the gust has died off, indicating the complex interaction of aerodynamic forces of rotor and tail vane, as well as the asymmetric mechanical conditions for entering and exiting the furling regime.

5. Conclusions

An algorithm has been presented capable of reconstructing wind speed time series with a typical resolution of seconds from data logger records providing only the values of the average
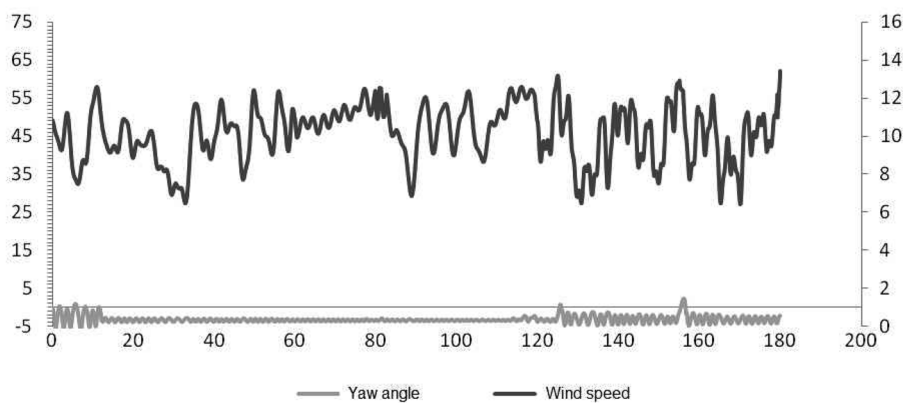

Fig. 15 Simulated wind speed time series (upper line, right scale) for a 3-minute time frame covered by data logger records and the calculated system response of a small furling wind turbine, exemplified by the turbine's yaw angle (lower line, left y-scale)
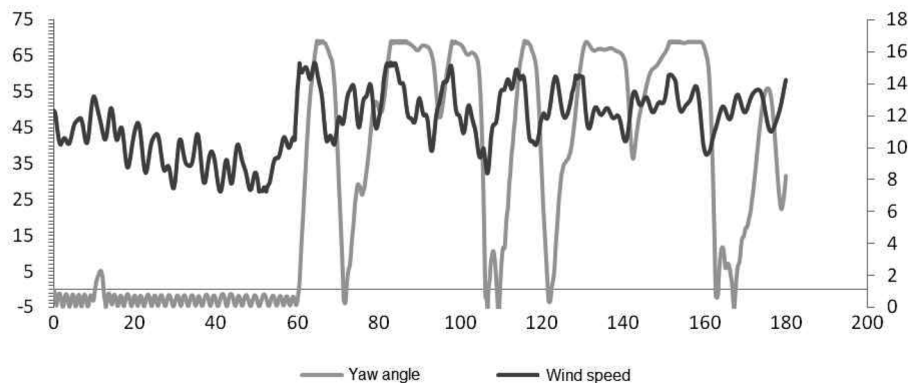

Fig. 16 Simulated wind speed time series (upper line, right scale) for a 3-minute time frame covered by data logger records and the calculated system response of a small furling wind turbine, exemplified by the turbine's yaw angle (lower line, left y-scale)

wind speed, its standard deviation, and its extreme values in a given interval (typically 10 mins or 1 min). The algorithm is based on random Fourier series and uses a post-processing step to tune the values of the extreme values of the simulated time series to the logged values. In an additional post-processing step a smooth transition between adjacent intervals is performed to avoid discontinuities. The wind speed time series comply with the spectral and autocorrelation properties expected for turbulent wind, where the power density spectrum can be specified through the appropriate choice of the frequency distribution function of the random Fourier series. Truncated Kaimal distribution functions with one length scale have been used. The ensemble autocorrelation function was calculated numerically and shown to be consistent with an analytical formula derived from the properties of the Kaimal distribution.

The fine tuning of the extreme values is based on the concept of constrained simulation introduced by Bierbooms and Cheng (2002) and has been applied to the original time series using the analytical expression for the ensemble autocorrelation function. Two strategies for the adjustment of the extreme values by constrained simulation have been used; in the first one, termed the asymmetric gust control process, gusts are only corrected if the simulated maximum is below the recorded maximum value in the interval; a similar strategy is used for the minima. In the other strategy, the symmetric gust control algorithm, corrections have been applied in both directions. The distribution of the mean values is well reproduced in both cases. Extreme value and variance distributions are well reproduced, too, but some differences can be noted between the two approaches: The symmetric algorithm reproduces the extreme values to practically any desired degree of accuracy, but reduces the standard deviation by some 10%, while the asymmetric algorithm leaves the standard deviation at its original value but produces slightly higher gusts than recorded. A possible explanation of this discrepancy lies with the low-pass characteristics of the anemometers and the recording system. While both algorithms may be suitable for most applications, the symmetric strategy may be used if the exact values of the gusts are more important, while the asymmetric algorithm may be used if the exact value of the standard deviation is of foremost importance. The applications of the algorithm have been illustrated by the case of a small wind turbine with *furling* overspeed control.

Acknowledgements

References

Aksoy, H., Toprak, Z.F., Aytek, A. and Ünal, N.E. (2004), "Stochastic generation of hourly mean wind speed data", *Renew. Energ.*, **29**(14), 2111-2131.
Audierne, E., Bergami, L., Elizondo, J. and Probst, O. (2010), "Analysis of the furling behavior of small wind turbines", *Appl. Energ.*, **87**(7), 2278–2292.
Brown, B.G., Katz, R.W. and Murphy, A.H. (1984), "Time series models to simulate and forecast wind speed and wind power", *J. Appl. Meteorol. Clim.*, **23**, 1184-1195.

Bierbooms, W. and Cheng, Po-Wen. (2002), "Stochastic gust model for design calculations of wind turbines", *J. Wind Eng. Ind. Aerod.*, **90**(11), 1237–1251.

Bierbooms, W. (2005), "Investigation of spatial gusts with extreme rise time on the extreme loads of pitchregulated wind turbines", *Wind Energy*, **8**(1), 17–34.

Burton, T., Sharpe, D., Jenkins, N. and Bossanyi, E. (2001), *Wind Energy Handbook.*, John Wiley & Sons, United Kingdom.

Cadenas, E. and Rivera, W. (2007), "Wind speed forecasting in the South Coast of Oaxaca, México", *Renew. Energ.*, **32**(12), 2116-2128.

Childers, D. (1997), Probability and random processes using Matlab with applications to continuous and discrete time series. Irwin/McGraw-Hill, ISBN 0-256-13361-1.

Durañona, V., Sterling, M. and Baker, Ch. (2007), "An analysis of extreme non-synoptic winds", *J. Wind Eng. Ind. Aerod.*, **95**(9-11), 1007-1027.

Elizondo, J., Martínez, J. and Probst, O. (2009), "Experimental study of a small wind turbine for low and medium wind regimes", *Int. J. Energ. Res.*, **33**(3), 309–326.

Elizondo, J., Delgado. A., Martínez, J. and Probst, O. (2010), "Sistema de plegado para aerogenerador de 1.5 kW", *Proceedings of the XXXIV Solar Energy Week*, *National Solar Energy Association (Mexico), Guanajuato*, Mexico, October.

International Electrotechnical Commission. IEC61400-12 (1998): *Wind turbine generator systems - Part 12: Wind turbine power performance testing*, 1st Edition, Golden: National Wind Technology Center.

Kaimal, J.C., Wyngaard, J.C., Izumi Y. and Cote, O.R. (1972), "Spectral characteristics of surface-layer turbulence", *Q. J. Roy. Meteor. Soc.*, **98**, 563-589.

Kaimal, J.C. and Finnigan, J.F. (1994), *Atmospheric boundary layer flows: their structure and measurement*, Oxford University Press US.

Kareem, A. (2008), "Numerical simulation of wind effects: A probabilistic perspective", *J. Wind Eng. Ind. Aerod.*, **96**(10-11), 1472–1497.

Mann, J., Ott, S., Hoffmann Jørgensen, B. and Frank, H.P. (2000), "WasP engineering 2000", Risø–R–1356 (EN)

Monbet, V., Ailliot, P. and Prevosto, M. (2007), "Survey of stochastic models for wind and sea state time series", *Probab. Eng. Mech.*, **22**(2), 113–126.

Morales, A. and Probst, O. (2006), "The field performance of a small wind system", *Proceedings of the WindPower 2006, American Wind Energy Association*, Pittsburg, June.

Nfaoui, H., Essiarab, H. and Sayigh, A.A.M. (2004), "A stochastic Markov chain model for simulating wind speed time series at Tangiers, Morocco", *Renew. Energ.*, **29**(8), 1407-1418.

Priestley, M.B. (1981), *Spectral analysis and time series, Probability and Mathematical Statistics*, Academic Press. London.

Sahina, A.D. and Senb, Z. (2001), "First-order Markov chain approach to wind speed modelling", *J. Wind Eng. Ind. Aerod.*, **89**(3-4), 263–269.

Saucier, R. (2000), "Computer generation of statistical distributions", Army Research Laboratory.

Seong, S.H. and Peterka, J.A. (1998), "Digital generation of surface-pressure fluctuations with spiky features", *J. Wind Eng. Ind. Aerod.*, **73**(2), 181-192.

Shinozuka, M. (1971), "Simulation of multivariate and multidimensional random processes", *J. Acoust. Soc. Am.*, **49**(18), 357-368.

Sfetsos, A. (2000), "A comparison of various forecasting techniques applied to mean hourly wind speed time series", *Renew. Energ.*, **21**(1), 23-35.

Verkaik, J.W. (2000), "Evaluation of two gustiness models for exposure correction calculations", *J. Appl. Meteorol. Clim.*, **39**(9), 1613-1626.

CC