SANDIA REPORT

SAND84—1227 • Unlimited Release • UC—60
Printed January 1985

A Markov Method for Simulating Non-Gaussian Wind Speed Time Series

Gerald M. McNerney, Paul S. Veers

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-76DP00789

SF2900Q(8-81)

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

NTIS price codes
Printed copy: A02
Microfiche copy: A01

A MARKOV METHOD FOR SIMULATING NON-GAUSSIAN WIND SPEED TIME SERIES*

Gerald M. McNerney
Wind America, Inc.

Paul S. Veers
Sandia National Laboratories
Albuquerque, NM 87185

ABSTRACT

This paper details a method which can be used to construct a wind simulator capable of generating wind time series with any distribution of hourly averages, exponentially decaying autocorrelation function, and a Gaussian realization of the turbulence. The method is based on a Markov random walk for hourly averages, and an inverse inverse hourly transform of the power spectrum to produce short-term turbulence. The Markov process is discussed in the first section and the turbulence generator is covered in the second section. A description of the applications for which the model was developed follows.

*This work supported by the US Department of Energy at Sandia National Laboratories under Contract No. DE-AC04-76DP00789.

INTRODUCTION

Wind speed time series may be periodic, random non-Gaussian or random Gaussian depending on the time scales involved. Monthly averages appear to have a strong periodic component over twelve months due to the seasonal changes in the global wind phenomenon. Hourly averages are a random process with a probability density which resembles a Weibull distribution. Instantaneous readings taken at one-second intervals strongly resemble a Gaussian random process. The purpose of this wind speed simulation is to provide input to a wind turbine generator control system simulation. The control system for a wind turbine has the task of deciding when to turn the turbine on or off to balance the energy capture with fatigue life consumption. This is currently accomplished by continually sampling the wind, usually at roughly one-second intervals, and using a predetermined algorithm to decide if the turbine should be on or off. The efficiency and protection from fatigue damage afforded by any particular algorithm is dependent on how it performs over a long period of time. Since high frequency wind speed data is not available for long times and a wide variety of wind sites, a wind speed simulation is necessary. The periodic nature of monthly averages is not modeled, but both the non-Gaussian hourly averages and the Gaussian high frequency wind speed variations are modeled.

The method of generating a sample of hourly averages is a Markov process random walk. A transition matrix relates each hourly average with the previous hourly average based upon 1) the desired probability density and 2) autocorrelation of the wind speed process. These two properties of the random process are uncoupled by defining the transition matrix as a normalized product of a diagonal probability matrix, to control the probability density, and a symmetric decay matrix, to control the autocorrelation. Thus, a series of hourly average wind speeds can be generated with any desired probability density function and an exponentially decaying autocorrelation with specified rate of exponential decay.

The high frequency wind speed time series, which represents the atmospheric turbulence, is superimposed on the hourly averages. The turbulence is represented as a Gaussian random process defined by its power spectral density (psd). This frequency domain representation of the turbulence is converted to the time domain using a fast Fourier transform which is computationally efficient. The functional form of the psd reflects changes in mean wind speed and permits adjustments to the turbulence intensity. A tapering scheme is employed to produce a smooth transition between adjacent hours.

GENERATING HOURLY AVERAGES

The basic method for generating a realization of hourly averages is a discrete time random walk over a finite state space of wind speeds. The process is Markovian in that the step to each succeeding state depends only on the preceding state and not on earlier states. Such a process may be represented by a matrix

$$[T] = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1m} \\ t_{21} & t_{22} & \cdots & t_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ t_{m1} & t_{m2} & \cdots & t_{mm} \end{bmatrix} . \qquad (1)$$

called the transition matrix. In this representation, t_{ij} is the one-step probability of transition from state i to state j. The transition matrix must have the following two properties:

i) all the entries are non-negative,

ii) the sum of the entries in each row is one.

Any matrix with these properties is called a stochastic matrix. It is a useful fact that the product of any two stochastic matrices is a stochastic matrix.

The property of transition matrices that will be used is the so-called renewal theorem which implies that if [T] is stationary, irreducible, and aperiodic, then [T] has the ergodic property, which is

$$\lim_{M \to \infty} [T]^M = \begin{bmatrix} r_1 & r_2 & \cdots & r_m \\ r_1 & r_2 & \cdots & r_m \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ r_1 & r_2 & \cdots & r_m \end{bmatrix} . \qquad (2)$$

where the limit is of the Mth exponent of the transition matrix [T], and the elements r_1, r_2, ... , r_m constitute a probability vector {r} called the limiting probability density function or limiting pdf (1). A transition matrix is stationary if it does not change in time, the irreducible property implies there are no closed or absorbing subsets of states, and a state is periodic if a random walk can return to that state only on integral multiples of the period.

The transition matrix can be used to represent a random walk since for a starting probability vector $\{v_0\}$, the nth step probabilities are given by the vector $\{v_n\} = \{v_0\} [T]^n$. To use [T] in realizing a random walk, first form the cumulative distribution matrix [C], where $c_{ij} = \sum_{k=1}^{j} t_{ik}$. Then when in the state i, choose a random number R from a uniform distribution and the next state will be k, where

$$c_{ik-1} < R \leq c_{ik} .$$

For our purposes, [T] needs to be ergodic with the limiting pdf equal to the pdf of hourly average wind speeds, and the resulting time series must have an exponentially decaying autocorrelation function with a specified rate of decay or, equivalently, a specified base.

To do this, define the transition matrix as the normalized product of two matrices, one to determine the probabilities of arriving in each state and the other to induce the desired autocorrelation property. A transition matrix can be found which will produce any combination of autocorrelation and pdf by varying these two matrices.

To begin the construction, the discrete weighting function $g_i = 2^{-|i|}$ is chosen and a decay matrix [G] is defined by

$$[G]_{ij} = g_{i-j} \quad (3)$$

The transition matrix is then defined as the matrix product

$$[T] = [N][G][P], \quad (4)$$

where [P] is the initial pdf matrix consisting only of diagonal entries $p_1, p_2, \ldots, p_m$, and [N] is a diagonal normalization matrix with diagonals

$$1/n_i = 1/(\sum_k p_k \, g_{k-i}) \quad (5)$$

The elements of [T] are then

$$t_{ij} = g_{i-j} \, p_j/n_i \quad (6)$$

[T] has no zero entries so it is aperiodic and irreducible. Since [T] has no time dependence, it is also stationary and therefore ergodic and has a limiting pdf $\{r\} = (r_1, r_2, \ldots, r_m)^T$. In practice, the initial pdf {p} and the limiting pdf {r} are different. In order that the limiting pdf be the desired result, the initial pdf must be compensated. This is done by first finding a functional form of the limiting pdf in terms of the initial pdf and then numerically inverting the relation.

To derive the relation between {r} and {p}, an alternate form of $[T]^M$ will be given which is then compared with equation (2). First, observe that any power of [T] must have the form

$$T^M = [N][A]^{M-1}[P] \quad (7)$$

where $[A]^{M-1}$ is a symmetric matrix of the form

$$[A]^{M-1} = [G]([P][N][G])^{M-1} = ([G][P][N])^{M-1}[G] \quad (8)$$

Since [T] is ergodic,

$$\lim_{M \to \infty} [T]^M = [N][A^\infty][P] , \quad (9)$$

where the limit has the form of equation (2). Let the elements of

$[A^\infty]$ be $A_{ij} = A_{ji}$,

then

$$[N][A^\infty][P] = \begin{matrix} A_{11}p_1/n_1 & A_{12}p_2/n_1 & \cdots & A_{1m}p_m/n_1 \\ A_{12}p_1/n_2 & A_{22}p_2/n_2 & \cdots & A_{2m}p_m/n_2 \\ \vdots & \vdots & & \vdots \\ A_{1m}p_1/n_m & A_{2m}p_2/n_m & \cdots & A_{mm}p_m/n_m . \end{matrix} \quad (10)$$

Comparison with the limiting matrix in equation (2) reveals that

$$\begin{matrix} A_{11}/n_1 = A_{12}/n_2 = \cdots = A_{1m}/n_m = r_1/p_1 \\ A_{12}/n_1 = A_{22}/n_2 = \cdots = A_{2m}/n_m = r_2/p_2 \\ \vdots \\ A_{1m}/n_1 = A_{2m}/n_2 = \cdots = A_{mm}/n_m = r_m/p_m . \end{matrix} \quad (11)$$

Thus, all the A_{ij}'s can be solved in terms of one of them, say A_{11}, or

$$A_{ij} = n_i n_j A_{11}/n_1^2 \quad (12)$$

It follows that

$$r_j = A_{11} p_j n_j / n_1^2 \quad (13)$$

which has the proper form, i.e., no i dependence. Since the limiting matrix is stochastic and {r} is a probability vector, it follows that

$$\sum_k r_k = (A_{11}/n_1^2) \, \Sigma p_k n_k = 1 , \quad (14)$$

and thus

$$A_{11} = n_1^2/\Sigma p_k n_k . \quad (15)$$

With this form for A_{11}, the elements of {r} can be solved for

$$r_j = p_j n_j/(\Sigma p_k n_k) \quad (16)$$

The product $p_j n_j$ can be expressed in vector form, which will be useful in inverting the equation for {r} , thus {r} tentatively will be expressed as

$$\{r\} = [P][G]\{p\} . \quad (17)$$

remembering that this product must be normalized to represent a probability.

Equation (17) may be used in an iterative process to find the initial pdf {p} in terms of the limiting pdf {r} which is known. First choose $\{p\}^\circ = \{r\}$, then calculate

$$\{r\}^1 = [P]^\circ[G]\{p\}^\circ . \quad (18)$$

Then normalize $\{r\}^1$ and calculate

$$\{p\}^1 = \{p\}^\circ + F(\{r\} - \{r\}^1) , \quad (19)$$

where $o < F < 1$ is a relaxation factor which stabilizes the convergence. For successive steps use

$$\{r\}^n = [P]^{n-1}[G]\{p\}^{n-1} \qquad (20)$$

$$\{p\}^n = \{p\}^{n-1} + F(\{r\} - \{r\}^n) \, . \qquad (21)$$

where $\{r\}^n$ must be normalized every step after equation (20). This process converges in 5 to 10 steps. The transition matrix can be calculated since the initial pdf is now known. However, the decay function has been arbitrarily chosen and the resulting autocorrelation may not have the desired rate of decay. To remedy this, a random walk must be conducted and the autocorrelation calculated. If the rate of decay is too fast, then change the decay function from $g_i = 2^{-|i|}$ to $g_i = B^{-|i|}$, where $B < 2$. If the decay is too slow, use use $B > 2$, and repeat the process. All physically reasonable hourly autocorrelations can be reached by starting initially with $B = 2$ and $\Delta B = 0.4$, then bisecting in five steps. The autocorrelation of the random walk will never be an exact exponential but if the first two hours are matched to the exponential, the error on the 12th hour generally will be 10 to 20 percent with the simulated usually greater than exponential decay. This is beneficial in the sense that real wind autocorrelation deviates from the exponential in the same manner. The Markov process cannot be made to give a diurnal cycle. The procedure described above allows the operator to choose any combination of autocorrelation decay and limiting pdf.

As an example, a Markov process was constructed for a Rayleigh pdf with mean wind speed of 8 m/s and autocorrelation with exponential decay on a base of 0.87. The results of the walk are given in Table I. The first column is the wind speed in meters per second, the second column is the desired Rayleigh pdf, the third column is the pdf that was realized by the walk after 8760 steps, and the fourth colunm is the initial pdf that was calculated during the program. The first 20 rows and columns of the associated transition matrix with two significant figures for clarity of presentation are given in Table II. The χ^2 goodness-of-fit test was performed on the resulting distributions. It was found that the resulting pdf just passes at the 0.05 level of confidence, which is normally used with goodness-of-fit tests. Figure 1 is a plot of the first 1000 hours of the random walk on hourly averages.

TABLE 1. DESIRED AND REALIZED PDF'S

Wind Speed (m/s)	Rayleigh pdf	Realized pdf	Initial pdf
1	.0242	.0267	.0381
2	.0466	.0531	.0933
3	.0657	.0667	.0622
4	.0805	.0793	.0678
5	.0901	.0889	.0710
6	.0945	.0945	.0722
7	.0940	.0950	.0716
8	.0894	.0847	.0695
9	.0817	.0807	.0661
10	.0719	.0704	.0617
11	.0612	.0610	.0566
12	.0503	.0555	.0511
13	.0402	.0413	.0454
14	.0311	.0287	.0397
15	.0233	.0204	.0342
16	.0170	.0179	.0290
17	.0121	.0112	.0243
18	.0083	.0089	.0200
19	.0056	.0049	.0163
20	.0036	.0039	.0134
21	.0023	.0018	.0104
22	.0014	.0016	.0081
23	.0009	.0010	.0063
24	.0005	.0007	.0045
25	.0003	.0006	.0032
26	.0002	.0005	.0020
27	.0001	.0001	.0012

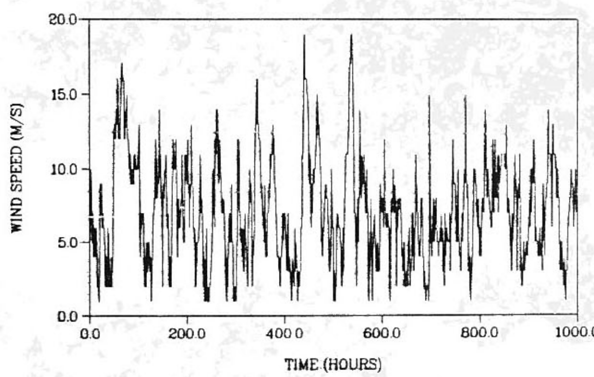

FIG. 1. 1000 HOURS OF A RANDOM WALK OF HOURLY AVERAGES

TABLE II. SAMPLE TRANSITION MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	.33	.26	.17	.10	.06	.03	.02	.01	.01											
2	.13	.33	.22	.13	.08	.05	.03	.01	.01											
3	.06	.16	.32	.20	.12	.07	.04	.02	.01	.01										
4	.03	.08	.16	.31	.18	.11	.06	.03	.02	.01	.01									
5	.02	.04	.09	.17	.31	.18	.10	.05	.02	.02	.01									
6	.01	.02	.05	.09	.17	.30	.17	.09	.05	.03	.01	.01								
7	.01	.01	.03	.05	.09	.17	.30	.16	.09	.05	.02	.01	.01							
8		.01	.02	.03	.05	.10	.17	.29	.16	.08	.04	.02	.01	.01						
9			.01	.02	.03	.06	.10	.17	.29	.15	.08	.04	.02	.01	.01					
10				.01	.02	.03	.06	.10	.18	.29	.15	.08	.04	.02	.01					
11				.01	.01	.02	.04	.06	.11	.18	.29	.15	.07	.04	.02	.01				
12					.01	.01	.02	.04	.07	.11	.18	.28	.14	.07	.03	.02	.01			
13						.01	.01	.02	.04	.07	.11	.18	.28	.14	.07	.03	.01	.01		
14						.01	.01	.02	.03	.04	.07	.11	.18	.28	.13	.06	.03	.01	.01	
15							.01	.01	.02	.03	.05	.07	.12	.18	.27	.13	.06	.04	.01	.01
16								.01	.01	.02	.03	.05	.08	.12	.18	.27	.13	.06	.03	.01
17									.01	.01	.02	.03	.05	.08	.12	.18	.27	.12	.06	.03
18									.01	.01	.01	.02	.03	.05	.08	.12	.18	.26	.12	.05
19										.01	.01	.01	.02	.04	.06	.08	.12	.18	.26	.12
20											.01	.01	.02	.03	.04	.06	.08	.12	.18	.25

TURBULENT WIND SIMULATION

After each hourly average wind speed is generated, second-by-second time series of the Turbulent speed during that hour can be simulated. This high frequency time series during any hour is a Gaussian random process. The second order statistical properties of the process can be controlled using the turbulence power spectral density (psd). In a NASA summary of atmospheric environments by Frost, Long and Turner(2), the following form for the psd of atmospheric turbulence is suggested:

$$S(\omega) = \frac{12.3\bar{V}h[\ln(10/z_o+1)\ln(h/z_o+1)]^{-1}}{1+192[h\omega\ln(10/z_o+1)/V\ln(h/z_o+1)]^{5/3}} \quad (22)$$

where

ω = frequency

h = height above ground

$\bar{V}$ = mean wind speed at h = 10 m

z_o = surface roughness coefficient

An example turbulence psd at a reference height of 30 ft (10 m) with a surface roughness coefficient of 0.1 is shown in Figure 2. *For the wind velocity a Gaussian time* series described by this psd can be obtained by the following method (3). Represent the power in a narrow band, $\Delta\omega$, of the psd by sine and cosine components, at the central frequency and sum these inputs over all the *frequencies of interest. Such a velocity* time history is given by

$$V(t) = \bar{V} + \sum_{j=1}^{n} (A_j \sin\omega_j t + B_j \cos\omega_j t) \quad (23)$$

where A_j and B_j, the real and imaginary parts of the spectrum at frequency ω_j, are randomly chosen from a normally distributed set of mean zero and variance $\sqrt{\frac{1}{2} S_j}$

S_j = magnitude of the psd at frequency ω_j

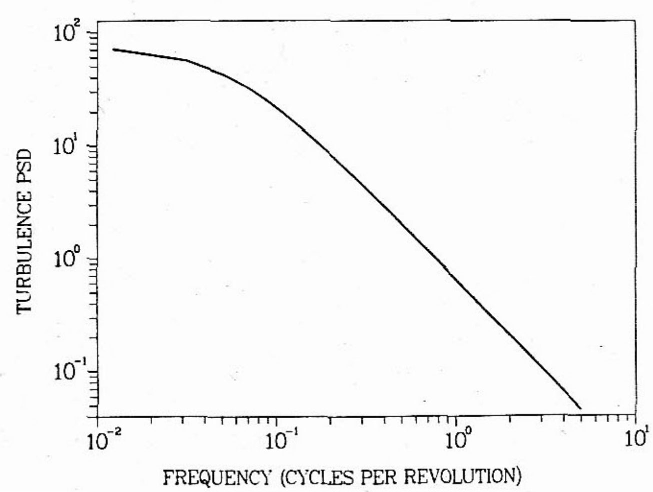

FIG 2. FREQUENCY CONTENT OF ATMOSPHERIC TURBULENCES

This is accomplished by making A_j and B_j sum of 12 or more uniformly distributed random variables and then scaling it to get the proper variance value. The central limit theorem guarantees convergence of V(t) to a Gaussian form as the number of sinusoidal components at different frequencies with random phase becomes large. This is equivalent to band limited, filtered white noise. If the frequency spacing is regular and begins at zero, the computation of the time series can be accomplished with a discrete inverse Fourier transform, F^{-1}, of the complex series,

$$V(t) = \bar{V} + F^{-1} \sum_{j=1}^{n} (A_j + iB_j) \qquad (24)$$

By using the fast Fourier transform, the above equation can be evaluated very efficiently. Adjacent hourly simulations are connected by creating an additional eight minutes of high frequency data (68 minute total) and overlapping into both the previous and subsequent hour. The ends of each hour's data are tapered with a sinusoidal weighting function so that the sum of the two overlapping sections maintains a uniform variance.

APPLICATIONS

The wind model described in this paper was constructed for use in a wind turbine control simulation. The problem of deciding when to turn on and off wind turbines to optimize energy capture while at the same time minimizing turbine fatigue life consumption is still unsolved, especially on the high wind side. Earlier work on this problem on the low wind side came to conclusive results for specific wind sites using real high frequency (0.5 hz) wind time series data taken from two sites: Bushland, Texas, and Albuquerque, New Mexico (4). In general, this kind of wind data is not available, and a wind simulator model was needed to fill the gap. For each wind site characteristic, the turbine control simulation will calculate the percentage of time a turbine is on and connected to the utility grid as well as the number of starts the turbine experiences for each algorithm and choices of parameters. With this information, the percent energy available that is captured and the fatigue life consumption can be calculated. The best algorithms to use in various combinations of turbine-wind site characteristics can then be determined.

Using the wind simulator will allow any number of wind regime types to be tested for essentially unlimited duration, a feature that is not possible using a real wind data base.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy at Sandia National Laboratories via a subcontract to Wind America, Inc.

REFERENCES

1. Isaacson, D. L., and Madsen, R. W., Markov Chains: Theory and Applications, John Wiley & Sons, New York, 1976, pp. 62-82.

2. Frost, W., Long, D. H., and Turner, R. E., "Engineering Handbook on the Atmospheric Environment Guideline for Use in Wind Turbine Generator Development," NASA Technical Paper 1359, December 1979.

3. Shinozuka, M., "Simulation of Multivariate and Multidimensional Random Processes," Journal of the Acoustical Society of America, Vol. 49, No. 1 (Part 2), p. 357, 1971.

4. McNerney, G. M., "Vertical Axis Wind Turbine Control Strategy, SAND81-1156, Sandia National Laboratories, Albuquerque, NM, August 1981.