Journal of Wind Engineering & Industrial Aerodynamics 171 (2017) 202–206

Contents lists available at ScienceDirect

Journal of Wind Engineering & Industrial Aerodynamics

journal homepage: www.elsevier.com/locate/jweia

Short communication

A simulation method for macro-meteorological wind speeds with a Forward Weibull parent distribution of general index

R. Ian Harris

RWDI, Tilers Road, Milton Keynes, MK11 3LH, UK

ARTICLE INFO

Keywords:
Simulated time series
General Weibull parent
Correlated data
Arbitrary auto-correlation
Five million years record length

ABSTRACT

This short communication extends the previous method for generating correlated synthetic mean wind data with a Rayleigh parent to the case where the parent can be a Forward Weibull distribution of arbitrarily chosen index with arbitrarily chosen auto-correlation.

1. Introduction

All countries that are members of the W.M.O. produce wind statistics in the form of mean wind speeds taken over an averaging time of between 10min and 1 h. In the UK an averaging period of an hour is used so, for convenience in all that follows, means will be described as hourly means, with the understanding that in other countries where a different averaging period is used, that average is implied.

There is a considerable body of empirical evidence which suggests that a good model for the probability distribution of hourly mean wind speeds, particularly those arising from extra-tropical depressions, is the Forward Weibull form with an index lying between 1.8 and 2.2. In a recent paper, Harris and Cook(2014) introduced a new distribution, the Offset Elliptic Normal(OEN), which allows the hourly mean wind vectors at any site and arising from combinations of physical causes to be modelled as a combination of these OEN distributions, one distribution for each physical cause. They also showed that the wind speeds from such a combination can always be accurately modelled by a combination of Forward Weibull distributions, all with indexes in the range between 1.8 & 2.2. The middle of this range is a Forward Weibull with index 2, otherwise known as a Rayleigh distribution. Production of design wind speeds for insertion in building codes requires some form of extreme value analysis of measured wind speeds. Verifying the accuracy of any such method using real measured data is difficult, because the true parent distribution of measured data is always unknown and even now it is rare to have available a record more than fifty years long. With records of this length, the sampling errors are likely to swamp the potential defects being investigated. Thus there is considerable merit in first testing extreme value methods on synthetic data of known parent distribution and auto-correlation. In a recent paper Harris (2014) presented a method for simulating data from a Rayleigh parent distribution with an arbitrary auto-correlation. The objective of the present paper is to demonstrate how this method can be extended to generate a time series, still of arbitrary auto-correlation, but with a parent distribution of general Forward Weibull form.

2. Summary of method used to produce simulated wind speeds with a Raleigh parent

The method described in Harris (2014) relies on the result that if $x(t)$ and $y(t)$ are two independent N(0,1) random sequences, each with an auto-correlation $\rho(\tau)$ and are regarded as Cartesian components of a 2-dimensional vector, then the modulus of that vector, $r(t)$, has a CDF given by:

$$P(r) = 1 - \exp\left(-r^2/2\right) \tag{2.1}$$

where:

$$r = \sqrt{(x^2 + y^2)} \tag{2.2}$$

Note that since this relationship is non-linear, the relationship between $\rho(\tau)$, the auto-correlation of the components $x(t)$ and $y(t)$, and $\zeta(\tau)$, the auto-correlation of the modulus $r(t)$, is not simple. In Harris (2014) it is shown that for any value of the timelag, τ:

E-mail address: harris-r1@virginmedia.com.

https://doi.org/10.1016/j.jweia.2017.10.004
Received 6 May 2016; Received in revised form 4 October 2017; Accepted 4 October 2017

$$\zeta(\tau)=\frac{{}_2F_1\left[-\frac{1}{2},-\frac{1}{2};1;\rho^2(\tau)\right]-1}{4/\pi-1} \tag{2.3}$$

where the ${}_2F_1$ is the Gauss Hypergeometric Function. [See Harris (2014) for details].

Thus, given a choice of the target auto-correlation, $\zeta(\tau)$, which can be arbitrary, $\rho(\tau)$, the required auto-correlation for the *x*- and *y*-components can be calculated from (2.3). This calculation has to be done only once/simulation. Thus given two sequences of N(0,1) random numbers independent of one another and also serially independent, each sequence is transformed into *x(t)* or *y(t)* with an auto-correlation $\rho(\tau)$, by applying an auto-regressive and/or moving average transformation [See Box et al., 2008] and then combined according to (2.2) to yield a sequence *r(t)* with auto-correlation $\zeta(\tau)$ and CDF given by (2.1).

3. Generalisation to produce simulated wind speeds with an arbitrary Forward Weibull parent and auto-correlation

When a Forward Weibull parent CDF is used to fit measured data, it is usually in the form:

$$P(V)=1-\exp[-(V/C)^w] \tag{3.1}$$

Here *w* is the Weibull index and *C* the scale parameter. Since the sequence *r(t)* produced by the computer simulation is dimensionless, it is useful to introduce the dimensionless variable $z=V/C$, so that the CDF becomes:

$$P(z)=1-\exp[-(z^w)] \tag{3.2}$$

Comparison of this result with (2.1) gives:

$$z=\left(\frac{r}{\sqrt{2}}\right)^{2/w} \tag{3.3}$$

$$=\left(\frac{x^2+y^2}{2}\right)^{1/w} \tag{3.4}$$

Thus, from any sequence *r(t)* with a Rayleigh CDF generated by the method described in Harris (2014), a sequence *z(t)* with a Weibull CDF with index *w*, can be obtained by applying the transformation (3.3) to each member of the sequence. However, (3.4) shows that the relation between *z(t)* and the components *x(t)* and *y(t)* is also non-linear, and is not the same as (2.2) for *r(t)*. Therefore if *z(t)* is to have the target auto-correlation, $\zeta(\tau)$, then the auto-correlation of the components, $\rho(\tau)$, needs to be changed by amending (2.3).

From (3.2) the probability density of *z(t)* is given by:

$$p(z)=wz^{w-1}\exp(-z^w) \tag{3.5}$$

Gradshteyn & Ryzhik (1994) give:

$$\int_0^\infty x^{\nu-1}\exp(-x^p)dx=\frac{1}{p}\Gamma\left(\frac{\nu}{p}\right) \tag{3.6}$$

and thus if < > denotes the expectation operator:

$$\langle z\rangle=\int_0^\infty zp(z)dz=\Gamma(1+1/w) \tag{3.7}$$

and:

$$\langle z^2\rangle=\int_0^\infty z^2p(z)dz=\Gamma(1+2/w) \tag{3.8}$$

In Harris (2014) it is shown that if *r(t)* is a stationary random sequence with a Rayleigh CDF, and r_1 & r_2 are any two members of that sequence separated by a time lag $\Delta\tau$, then the second order joint pdf of r_1 & r_2 is given by:

$$p(r_1,r_2)=\frac{r_1r_2}{(1-\rho^2)}\exp\left[\frac{-(r_1^2+r_2^2)}{2(1-\rho^2)}\right]I_0\left[\frac{r_1r_2\rho}{(1-\rho^2)}\right] \tag{3.9}$$

I_0 is a modified Bessel Function of the first kind and for brevity ρ is used to denote $\rho(\Delta\tau)$.

It then follows that if z_1 & z_2 are two values of the Weibull sequence *z(t)* corresponding to *r(t)* via (3.3), then:

$$\langle z_1z_2\rangle=\int_0^\infty\int_0^\infty\frac{1}{2^{(2/w)}}\frac{r_1^{(1+2/w)}r_2^{(1+2/w)}}{(1-\rho^2)}\exp\left[\frac{-(r_1^2+r_2^2)}{2(1-\rho^2)}\right]I_0\left[\frac{r_1r_2\rho}{(1-\rho^2)}\right]dr_1dr_2 \tag{3.10}$$

To evaluate the integrals in (3.10), the Bessel function is expanded into its series representation, Gradshteyn & Ryzhik (1994):

$$I_0\left[\frac{r_1r_2\rho}{(1-\rho^2)}\right]=\sum_{k=0}^\infty\frac{1}{\Gamma(1+k)\Gamma(1+k)}\left(\frac{r_1r_2\rho}{2(1-\rho^2)}\right)^{2k} \tag{3.11}$$

It is also useful to introduce the changes of variables:

$$\lambda_1=r_1\Big/\sqrt{2(1-\rho^2}\quad\lambda_2=r_2\Big/\sqrt{2(1-\rho^2)} \tag{3.12}$$

These changes lead to:

$$\langle z_1z_2\rangle=4\left(1-\rho^2\right)^{1+2/w}\sum_{k=0}^\infty\frac{\rho^{2k}}{\Gamma(1+k)\Gamma(1+k)}\int_0^\infty\lambda_1^{(2k+1+2/w)}\exp(-\lambda_1^2)d\lambda_1\int_0^\infty\lambda_2^{(2k+1+2/w)}\exp(-\lambda_2^2)d\lambda_2 \tag{3.13}$$

Again use is made of result (3.6) from Gradshteyn & Ryzhik (1994), with $p=2$ and $\nu=(2k+2+2/w)$, and leads to:

$$\langle z_1z_2\rangle=\left(1-\rho^2\right)^{1+2/w}\sum_{k=0}^\infty\frac{\rho^{2k}\Gamma(k+1+1/w)\Gamma(k+1+1/w)}{\Gamma(1+k)\Gamma(1+k)} \tag{3.14}$$

$$=\left(1-\rho^2\right)^{1+2/w}\Gamma(1+1/w)\Gamma(1+1/w)_2F_1\left(1+1/w,\ 1+1/w;\ 1;\rho^2\right) \tag{3.15}$$

$$=\Gamma(1+1/w)\Gamma(1+1/w)_2F_1\left(-1/w,\ -1/w;\ 1;\rho^2\right) \tag{3.16}$$

Here the ${}_2F_1$ denotes the Gauss Hypergeometric Function which is defined by a power series. (3.15) follows from (3.14) by virtue of this definition, and the simplification to (3.16) follows from a further result for this function. [See (Gradshteyn and Ryzhik, 1994) for the derivation of all three of these results].

By definition, $\zeta(\tau)$, the auto-correlation of *z(t)*, is given by:

$$\zeta(\tau)=\frac{\langle z_1z_2\rangle-\langle z\rangle\langle z\rangle}{\langle z^2\rangle-\langle z\rangle\langle z\rangle} \tag{3.17}$$

so that from (3.7), (3.8) and (3.16):

$$\frac{\Gamma(1+1/w)\Gamma(1+1/w)_2F_1(-1/w,-1/w;1;\rho^2(\tau))-\Gamma(1+1/w)\Gamma(1+1/w)}{\Gamma(1+2/w)-\Gamma(1+1/w)\Gamma(1+1/w)}=\zeta(\tau) \tag{3.18}$$

(3.18) can be further simplified by using identities for the Hypergeometric and Gamma Functions available in Gradshteyn & Ryzhik (1994). If the function *R(w)* is defined as:

$$R(w)=2w\Gamma(2/w)/[\Gamma(1/w)]^2 \tag{3.19}$$

Then (3.18) reduces to:

$$\frac{{}_2F_1[(-1/w,-1/w;1;\rho^2(\tau)]-1}{R(w)-1}=\zeta(\tau) \tag{3.20}$$