

# Generación de series sintéticas de la velocidad del viento mediante especificación de histograma.

Eduardo Ernesto Ordóñez-López\*, Mauricio Alberto Escalante Soberanis, Manuel Israel Flota Bañuelos, Bassam Ali

*Fecha de recepción: 22 de septiembre de 2017 – Fecha de aprobación: 14 de marzo de 2018*

## RESUMEN

En este trabajo se presenta un método para la generación de series sintéticas de datos de velocidad de viento basado en la especificación del histograma. Este método permite generar series sintéticas con base en la transformación de datos obtenidos de manera experimental y la descripción estadística de los datos deseados en términos de su función distribución de probabilidad, la cual puede ser descrita analíticamente o ser estimada mediante el histograma de una serie de datos experimentales. La obtención de series sintéticas por medio de la transformación de datos experimentales de la velocidad del viento es útil en la evaluación del comportamiento dinámico de los sistemas eólicos en redes energéticas porque se preservan algunas características del recurso tales como el patrón promedio diario, mientras se pueden variar otras como la función densidad de probabilidad. Se presentan los resultados de la aplicación del algoritmo propuesto en la transformación de series de tiempo con distribución monomodal, en nuevas series de tiempo con distribución de probabilidad monomodal y en series de tiempo con distribución bimodal.

**Palabras clave:** velocidad del viento, generación de datos sintéticos, especificación de histograma.

## Histogram specification based method for synthetic wind speed time series generation.

### ABSTRACT

A histogram specification based method for synthetic wind speed time series generation is presented. This method allows the generation of synthetic time series based on the transformation of experimentally obtained data and the statistical description of the desired data in terms of its probability distribution function. This distribution can be described analytically or can be estimated from experimental data. The production of synthetic time series for the wind speed by means of the transformation of experimental data is useful in the evaluation of the dynamic behavior of wind energy systems in electric grids since some features of the wind resource such as the average daily pattern are preserved while others such as the probability density function can be varied. We present the results of the

Facultad de Ingeniería, Av. De Industrias no Contaminantes, Periférico Norte, s/n, Universidad Autónoma de Yucatán, Mérida Yucatán. CP. 97203. México.

\*Autor de correspondencia. Correo electrónico: [eduardo.ordonez@correo.uady.mx](mailto:eduardo.ordonez@correo.uady.mx)

**Nota:** Este artículo de investigación es parte de Ingeniería–Revista Académica de la Facultad de Ingeniería, Universidad Autónoma de Yucatán, Vol. 22, No. 1, 2018, ISSN: 2448-8364.

proposed algorithm in the transformation of time series data with unimodal distribution, into new time series with unimodal probability distribution and in time series with bimodal distribution.

**Keywords:** wind speed, synthetic wind data time series generation, histogram specification.

## INTRODUCCIÓN

La conocida variabilidad del recurso eólico, el crecimiento de la industria de generación de electricidad a escala comercial a partir del viento y sus efectos sobre los sistemas de generación, transmisión y distribución de electricidad han motivado el desarrollo de una cantidad importante de métodos para estimar y predecir la producción energética con una gran variedad de enfoques: estadísticos, dinámicos, redes neuronales artificiales, híbridos, etc. (Jung y Broadwater 2014).

La estimación mediante técnicas estadísticas de la producción energética de los sistemas eólicos se basa comúnmente en la combinación de la descripción del sistema (mediante la curva de potencia) y la descripción del recurso eólico (por medio de la función densidad de probabilidad de la velocidad del viento o por su representación mediante un histograma) (Celik 2003). Los métodos estadísticos son adecuados cuando el objetivo es la evaluación del desempeño del sistema en un amplio periodo de tiempo, como por ejemplo para confirmar la viabilidad de un sitio para la ubicación de un sistema eólico (Carta *et al.* 2008).

Para la estimación a corto plazo de la producción energética de los sistemas eólicos y el desarrollo de sistemas de control cada vez más efectivos, no basta describir el sistema mediante su curva de potencia, sino que es importante incorporar modelos dinámicos (Carta *et al.* 2008). En la estimación de potencia el modelo

dinámico es utilizado en conjunto con la representación del recurso eólico mediante series de tiempo de la velocidad del viento y posiblemente también de la dirección y otros parámetros que influyan en la potencia disponible en él.

La obtención experimental de series reales de tiempo para describir el viento para un conjunto amplio de sitios es impráctica debido a la necesidad de incorporar, durante periodos suficientemente largos, torres adecuadamente instrumentadas para la obtención de estos datos.

En este contexto, la generación de series sintéticas de la velocidad del viento, tanto de baja como de alta tasa de muestreo, que se ajusten a una descripción estadística objetivo y que a la vez contengan elementos descriptivos de la región (como por ejemplo el comportamiento estacional, el patrón promedio mensual, el patrón promedio diario, el contenido de turbulencia, etc.) es un elemento indispensable para las simulaciones computacionales orientadas a evaluar preliminarmente el comportamiento de sistemas eólicos en diferentes entornos y sus efectos en la estabilidad de las redes eléctricas (Suomalainen *et al.* 2012).

Diversos enfoques se han aplicado para la obtención de series sintéticas de la velocidad promedio horaria del viento. Nfai *et al.* (1996) emplearon modelos autorregresivos. Carapellucci y Giordano (2013a) utilizan un modelo estocástico que usa como parámetros la velocidad media y

la velocidad máxima en una base anual o mensual, pero también incorporan una componente determinista del recurso eólico, como por ejemplo mediante el perfil diurno de la velocidad del viento (Carapellucci y Giordano 2013b). Los procesos de Markov también son una herramienta empleada para la generación de series sintéticas de la velocidad del viento promedio horaria empleadas en conjunto con descriptores tales como la velocidad media, desviación estándar y distribución de probabilidad (Shamshad *et al.* 2005).

Para una región determinada, la obtención de series sintéticas de la velocidad del viento, para diferentes puntos dentro de la región o a diferentes alturas sobre el nivel del terreno, se puede realizar mediante la transformación de series de datos experimentales de una estación base de monitoreo (por ejemplo una estación meteorológica con intervalo de muestreo de 10 minutos) para ajustar parámetros tales como la velocidad promedio, la cantidad de turbulencia, la función distribución de probabilidad, la frecuencia de muestreo, el contenido espectral, entre otros.

En las regiones costeras es frecuente que la brisa marina tenga un efecto importante en el comportamiento diurno de la velocidad y dirección del viento (Jaramillo y Borja 2004; Soler *et al.* 2009), de esta manera los datos experimentales tomados en una región se pueden transformar para adquirir las características estadísticas de otra región y así poder realizar una estimación preliminar del desempeño de un sistema eólico mediante la simulación computacional de su comportamiento dinámico.

La adaptación de las series de tiempo para que representen las características del

viento a diferentes alturas también es un caso particularmente importante porque en el proceso de optimización económica de los sistemas eólicos es necesario calcular la energía producida para diferentes longitudes de torre. Es conocido que la velocidad promedio del viento puede ser estimada a diferentes alturas mediante modelos logarítmicos. Sin embargo, la aplicación de este modelo a una serie de tiempo, aún cuando representa adecuadamente la velocidad promedio, falla en describir correctamente la potencia que será producida por el sistema eólico porque falla en adecuar la función densidad de probabilidad de la velocidad del viento que se tiene a la nueva altura (Emeis 2013).

La transformación de una serie de tiempo de la velocidad del viento para que su función de densidad de probabilidad se ajuste a una función objetivo permite obtener una nueva serie de tiempo que preservará el comportamiento, estacional, mensual o diurno del viento, pero que a la vez permitirá describir adecuadamente un recurso eólico con diferente densidad de potencia.

La función densidad de probabilidad objetivo puede aproximarse mediante un método experimental (haciendo una toma reducida de datos de velocidad de viento), o se puede estimar mediante diferentes modelos ya ampliamente aceptados como son Weibull, Weibull&Weibull, etc. (Jaramillo y Borja 2004; Ramírez y Carta 2006).

En este trabajo se presenta una nueva metodología basada en la especificación de histograma para transformar series de tiempo correspondientes a periodos reducidos de datos de la velocidad del viento tomadas a altas tasas de muestreo de tal forma que la función de densidad de probabilidad de la señal obtenida se

aproxime a una función densidad de probabilidad objetivo mientras que se preservan características de la serie de tiempo tales como el comportamiento diario.

### Objetivos:

Obtener series de tiempo de la velocidad del viento que contengan las características meteorológicas representativas de una región y se ajusten a una nueva función densidad de probabilidad (función objetivo).

Obtener una estimación de la función densidad de probabilidad de la velocidad del viento, adecuada para el cálculo de la producción energética mediante la curva de potencia del sistema eólico.

### METODOLOGÍA

Especificación de histograma es una metodología ampliamente utilizada para el procesamiento de señales digitales. En procesamiento de imágenes la especificación del histograma y su caso particular de ecualización de histograma, permiten hacer visibles detalles que están enmascarados en regiones oscuras o muy brillantes.

La especificación del histograma es una metodología que permite obtener una variable aleatoria  $y$  (señal transformada) con función de densidad de probabilidad  $f_{dp_y}(y)$  (función objetivo) a partir de una variable aleatoria  $x$  (señal base) con función de densidad de probabilidad  $f_{dp_x}(x)$ . Sean  $DP_x(x)$  y  $DP_y(y)$  las distribuciones de probabilidad de las variables aleatorias  $x$  e  $y$  respectivamente. Si estas variables aleatorias son continuas, entonces sus distribuciones de probabilidad son funciones monótonas estrictamente crecientes con rango  $[0,1]$  y la

especificación de histograma se basa en la siguiente transformación (Coltuc *et al.* 2006):

$$y = DP_y^{-1}(DP_x(x))$$

En la aplicación de especificación de histograma en señales discretas es usual que las variables aleatorias base y objetivo compartan el mismo alfabeto. En las variables aleatorias discretas la densidad de probabilidad es una función discreta, por lo que la especificación de histograma no logra resultados exactos; es decir, la función densidad de probabilidad de la variable aleatoria transformada no es exactamente igual a la función  $f_{dp_y}(y)$

objetivo. Existen diversas propuestas para lograr una especificación de histograma aproximadamente exacta en variables aleatorias discretas con aplicaciones principalmente orientadas al procesamiento de imágenes. En (Cultuc *et al.* 2006) se presenta un método que se basa en considerar un mecanismo para la ordenación estricta de las muestras de la variable aleatoria derivada de la posición del pixel en la imagen.

En este trabajo presentamos una variación de la especificación de histograma en la que la señales base y objetivo no comparten el mismo alfabeto, esto permite que, aún cuando las variables aleatorias son discretas, el problema de transformación de histograma sea una aproximación al caso de las variables aleatorias continuas y por consecuencia la especificación de histograma de la variable aleatoria  $x$  (señal base) produzca una variable aleatoria  $yt$  (señal transformada) cuya función densidad de probabilidad real  $f_{dp_{yt}}(yt)$  sea muy aproximada a la densidad de probabilidad objetivo  $f_{dp_y}(y)$  de una variable aleatoria  $y$  (señal objetivo). Aquí se considera que la función densidad de probabilidad de las señales es continua pero está representada

por un conjunto de muestras. Como se verá más adelante estas muestras no están uniformemente espaciadas en el rango de la función.

**Obtención del alfabeto de la variable aleatoria de la señal base**

En el procesamiento de datos de señales físicas, como las series de tiempo de la velocidad del viento, la inversa de la función distribución de probabilidad  $DP_y(y)$  se debe obtener mediante métodos numéricos debido a que  $fdp_x(y)$ , y posiblemente también  $fdp_y(y)$ , son descritas con base en el histograma de un conjunto de  $M$  muestras de la variable aleatoria.

La medición de la velocidad del viento  $v(t)$  (que es una variable aleatoria continua con función densidad de probabilidad  $fdp_x(v)$ ) mediante un sistema digital (tomando muestras a cada  $t_0$  segundos) puede considerarse (por el proceso de cuantificación escalar inherente al proceso de conversión de analógico a digital) como una variable aleatoria discreta  $v[nt_0]$  con alfabeto  $A_v = \{v_1, v_2, \dots, v_{KT}\}$ . El número de elementos  $KT$  del alfabeto de esta variable aleatoria depende del rango y de la resolución del instrumento de medición de la velocidad del viento, por consecuencia los valores específicos de  $v_i$  y su espaciamiento pueden variar dependiendo del tipo de instrumento o inclusive de la calibración del mismo. Como se trata de una variable aleatoria discreta, la función densidad de probabilidad se representa mediante el histograma  $H_v = \{hv_1, hv_2, \dots, hv_{KT}\}$  en donde  $hv_i$  es el número de veces que se repite el elemento  $v_i$  del alfabeto dentro de un conjunto dado de  $M$  muestras de la variable aleatoria, de tal forma que:

$$M = \sum_{i=1}^{KT} hv_i$$

En un conjunto  $V = \{v[1], v[2], \dots, v[M]\}$  de  $M$  muestras de la velocidad del viento, en donde  $v[i] \in A_v \quad \forall i \in [1, M]$ , es posible que algunos elementos de  $H_v = \{hv_1, hv_2, \dots, hv_{KT}\}$  sean iguales a cero (principalmente en el rango de las velocidades del viento más altas). Tomar como base esta descripción del histograma producirá una distribución de probabilidad acumulativa  $DP_v(v)$  de la velocidad del viento inadecuada para la especificación del histograma porque  $DP_v(v)$  no será monótona estrictamente creciente. Se propone entonces modificar el alfabeto de la variable aleatoria a un subconjunto  $A_x$ , de  $K$  elementos de  $A_v$ , correspondientes a los elementos del alfabeto que cumplen con  $hv_i \neq 0$ :

$$K < KT$$

$$A_x = \{x_1, x_2, \dots, x_K\} \subset \{v_1, v_2, \dots, v_{KT}\}$$

$$x_i < x_j \quad \forall i < j \quad i, j \in [1, K]$$

$$H_x = \{hx_1, hx_2, \dots, hx_K\} \subset \{hv_1, hv_2, \dots, hv_{KT}\}$$

$$hx_i > 0 \quad \forall i \in [1, K]$$

$$M = \sum_{i=1}^K h_{x_i}$$

El conjunto  $A_x$  de  $K$  elementos, en donde  $x_i < x_j \quad \forall i < j \quad i, j \in [1, K]$ , se obtiene tomando muestras no repetidas de la ordenación creciente del conjunto  $V = \{v[1], v[2], \dots, v[M]\}$ .

El histograma  $H_x = \{hx_1, hx_2, \dots, hx_K\}$  se obtiene clasificando a  $V = \{v[1], v[2], \dots, v[M]\}$  en  $K$  intervalos centrados en los elementos del conjunto  $A_x = \{x_1, x_2, \dots, x_K\}$ . De esta manera todos los elementos en  $H_x$  son diferentes de cero y la distribución de

probabilidad acumulativa correspondiente  $DP_x(v)$  será monótona estrictamente creciente.

### Estimación de las funciones de densidad y distribución de probabilidad de la señal base

La estimación de las versiones muestreadas  $fdp_x[x_i]$  y  $DP_x[x_i]$  de  $fdp_x(v)$  y de  $DP_x(v)$  respectivamente, para  $A_x = \{x_1, x_2, \dots, x_K\}$ , pueden derivarse del histograma  $H_x = \{hx_1, hx_2, \dots, hx_K\}$  mediante:

$$fdp_x[x_i] = \frac{hx_i}{M\Delta x_i} \quad \forall i \in [1, K]$$

$$\Delta x_i = \begin{cases} i=1 & x_1 + \frac{x_2 - x_1}{2} \\ 1 < i < K & \frac{x_{i+1} - x_{i-1}}{2} \\ i=K & \frac{x_K - x_{K-1}}{2} \end{cases}$$

$$DP_x[x_i] = \sum_{n=1}^i fdp_x(x_n)\Delta x_n \quad \forall i \in [1, K]$$

$$DP_x[x_j] > DP_x[x_i] \quad \forall j > i \quad i, j \in [1, K]$$

En donde  $\Delta x_i$  representan a las longitudes de los elementos de una partición exhaustiva mutuamente exclusiva del rango de la variable  $v(t)$ ,  $R = [0, v_{\max}]$ , centrados en los elementos del conjunto  $A_x = \{x_1, x_2, \dots, x_K\}$ , de tal forma que:

$$\sum_{i=1}^K \Delta x_i = v_{\max} = \max\{v[1], v[2], \dots, v[M]\}$$

Es importante notar aquí que, debido a que algunos elementos del alfabeto original de la variable aleatoria no aparecen en la serie de datos experimentales, no todos los elementos del conjunto  $A_x = \{x_1, x_2, \dots, x_K\}$  están uniformemente espaciados, por lo

que las muestras de  $fdp_x(v)$  tampoco lo están.

Para utilizar en la práctica esta función de densidad de probabilidad, por ejemplo en el cálculo de la producción energética de un sistema eólico mediante la descripción de su curva de potencia, es importante tomar en cuenta de que las muestras  $fdp_x[x_i]$  representativas de  $fdp_x(v)$  corresponden al conjunto de valores posibles de la variable aleatoria  $A_x = \{x_1, x_2, \dots, x_K\}$ .

### Estimación del alfabeto y la función distribución de probabilidad de la señal objetivo

Caso 1: Descripción analítica de la densidad de probabilidad objetivo.

Si se cuenta con una descripción analítica de  $DP_y = f(y)$  (esta función debe ser monótona creciente estricta), entonces la obtención de los valores  $y_i$  que producen las muestras de  $DP_{y_i}$  que cumplen con  $DP_{x_i} = DP_{y_i}$  se pueden calcular numéricamente mediante el siguiente procedimiento:

a) Estimar el valor  $y_{\max}$  de la variable aleatoria  $y$ , en el que  $f(y)$  sea muy próximo a uno, es decir,  $1 - f(y) < \varepsilon$  con  $\varepsilon \rightarrow 0$  siendo un parámetro definido arbitrariamente, por ejemplo  $\varepsilon = 1 \times 10^{-6}$ . Con este valor se define el rango  $R_y = [0, y_{\max}]$  de la variable aleatoria que contiene casi el 100% de la probabilidad.

b) Obtener un conjunto inicial  $y_{ini}$  de  $L$  valores uniformemente espaciados en el rango  $R_y = [0, y_{\max}]$  y el correspondiente conjunto de valores de distribución de

probabilidad  $DPI_y$  tales que:

$$L \gg K$$

$$y_{ini} = \{0, 1, \dots, L-1\} \times \frac{v_{y\max}}{L-1} = \{y_{ini1}, y_{ini2}, \dots, y_{iniL}\}$$

$$DPI_y = \{f(y_{ini1}), f(y_{ini2}), \dots, f(y_{iniL})\} = \{DPI_{y1}, DPI_{y2}, \dots, DPI_{yL}\}$$

$$(v_i, DPI_{yi}) \quad i \in [0, L]$$

c) Considerando el criterio del vecino más cercano, obtener el subconjunto  $DP_y$ , de  $K$  elementos, de los valores  $DPI_y$  más parecidos a los valores  $DP_x$  y el subconjunto  $A_y$  de  $y_{ini}$  correspondientes a los valores  $DP_y$  tales que:

$$DP_y = \{DP_{y1}, DP_{y2}, \dots, DP_{yK}\} \subset DPI_y = \{DPI_{y1}, DPI_{y2}, \dots, DPI_{yL}\}$$

$$A_y = \{y_1, y_2, \dots, y_K\} \subset y_{ini} = \{y_{ini1}, y_{ini2}, \dots, y_{iniL}\}$$

tales que, para  $n \in [1, K]$

$$DP_{yn} = DPI_{yi} \text{ tq } |DPI_{yi} - DP_{xn}| \leq |DPI_{yj} - DP_{xn}| \forall i, j \in [1, L]$$

$$y_n = y_{ini i}$$

Caso 2: Descripción numérica de la densidad de probabilidad objetivo por medio de una serie de tiempo de la velocidad del viento

Si no se cuenta con una descripción analítica de  $DP_y = f(y)$ , sino que es necesario estimarla a partir de una serie de tiempo de la velocidad del viento adquirida experimentalmente, entonces la obtención de los valores  $y_i$  que producen las muestras de  $DP_{yi}$  que cumplen con

$DP_{xi} = DP_{yi}$  se pueden calcular numéricamente mediante el siguiente procedimiento:

a) Obtener el alfabeto  $A_{yo} = \{yo_1, yo_2, \dots, yo_{KO}\}$  de la señal objetivo (en este caso representada por una serie de tiempo de la velocidad del viento) con un procedimiento similar al descrito para la señal base.

b) Obtener el histograma  $H_{yo} = \{hyo_1, hyo_2, \dots, hyo_{KO}\}$  y con base en él obtener las muestras  $DP_{yo} = \{DP_{yo1}, DP_{yo2}, \dots, DP_{yoKO}\}$  de la función distribución de probabilidad mediante:

$$\Delta yo_i = \begin{cases} i=1 & yo_1 + \frac{yo_2 - yo_1}{2} \\ 1 < i < KO & \frac{yo_{i+1} - yo_{i-1}}{2} \\ i=KO & \frac{yo_K - yo_{KO-1}}{2} \end{cases}$$

$$DP_{yo}[y_i] = \sum_{n=1}^i fdp_{yo}(yo_n) \Delta yo_n \quad \forall i \in [1, KO]$$

$$DP_{yo}[yo_j] > DP_{yo}[yo_i] \quad \forall j > i \quad i, j \in [1, KO]$$

c) Como se indica más adelante, para poder realizar la transformación de la señal se requiere que se cumpla con  $DP_x[x_i] = DP_y[y_i] \forall i \in [1, K]$ , sin embargo, el conjunto de muestras  $DP_{yo} = \{DP_{yo1}, DP_{yo2}, \dots, DP_{yoKO}\}$  en general no cumplirá con estas características, por lo que es necesario remuestrear la distribución de probabilidad objetivo  $DP_{yo}$  y el alfabeto correspondiente  $A_{yo} = \{yo_1, yo_2, \dots, yo_{KO}\}$  obtenidos inicialmente al procesar las muestras de la serie de tiempo de la señal objetivo. Este remuestreo se puede hacer por medio de

interpolación, por ejemplo del tipo lineal como se indica a continuación:

$$DP_{yi} = DP_{xi} \forall i \in [1, K]$$

$$y_i = \frac{(y_{o_{sup}} - y_{o_{inf}})}{(DP_{yosup} - DP_{yoinf})} D_{yi} +$$

$$\frac{y_{o_{inf}} DP_{yosup} - y_{o_{sup}} DP_{yoinf}}{(DP_{yosup} - DP_{yoinf})},$$

con

$$DP_{yoinf} = DP_{yok} tq |DP_{yok} - D_{xi}| \leq |DP_{yol} - D_{xi}|$$

$$\forall k, l \in [1, KO] \text{ para los cuales } DP_{yok}, DP_{yol} < DP_{xi}$$

$$y_{o_{inf}} = y_{ok}$$

$$DP_{yosup} = DP_{yok} tq |DP_{yok} - D_{xi}| \leq |DP_{yol} - D_{xi}|$$

$$\forall k, l \in [1, KO] \text{ para los cuales } DP_{yok}, DP_{yol} \geq DP_{xi}$$

$$y_{o_{sup}} = y_{ok}.$$

Cuando la muestra más pequeña de  $DP_x$  es menor que la muestra más pequeña de  $DP_y$  es necesario hacer la siguiente modificación para evitar indefiniciones en la interpolación en el extremo cercano a cero:

$$DP_{yo} = \{\min(DP_{xi}), DP_{yol}, DP_{yo2}, \dots, DP_{yoko}\}$$

$$A_{yo} = \left\{ y_{o1} \frac{\min(DP_{xi})}{DP_{yol}}, y_{o1}, y_{o2}, \dots, y_{oko} \right\}$$

### Obtención de la señal transformada a partir de la señal base y la distribución de probabilidad objetivo

Sean  $DP_x[x_i]$  y  $DP_y[y_i]$  las versiones representadas por  $K$  muestras de las distribuciones de probabilidad de las variables aleatorias  $x$  (señal base) e  $y$  (señal objetivo) respectivamente. Si estas distribuciones de probabilidad son funciones monótonas estrictamente crecientes con rango  $[0, 1]$  y se cumple con  $DP_x[x_i] = DP_y[y_i] \forall i \in [1, K]$  la especificación de histograma, la cual se basa en la

transformación  $y = DP_y^{-1}(DP_x(x))$  (Coltuc *et al.* 2006), se puede resolver numéricamente considerando que  $DP_x$  y  $DP_y$  estarán representadas por:

$$A_x = \{x_1, x_2, \dots, x_K\} DP_x = \{DP_{x1}, DP_{x2}, \dots, DP_{xK}\},$$

$$x_i \neq x_j \forall i \neq j \text{ tq } i, j \in [1, K]$$

$$A_y = \{y_1, y_2, \dots, y_K\} DP_y = \{DP_{y1}, DP_{y2}, \dots, DP_{yK}\},$$

$$y_i \neq y_j \forall i \neq j \text{ tq } i, j \in [1, K]$$

De esta forma la transformación de  $M$  muestras de la señal base  $x = \{x[1] = v[1], x[2] = v[2], \dots, x[M] = v[M]\}$ , en una señal  $yt = \{yt[1], yt[2], \dots, yt[M]\}$ , de  $M$  muestras también, que cumpla con la función de distribución de probabilidad objetivo  $fdp_y(y)$  se hace mediante:

$$\forall m \in [1, M] \text{ encontrar } n \in [1, K] \text{ tal que}$$

$$|x[m] - x_n| \leq |x[m] - x_l| \forall l \in [1, K]$$

(usar búsqueda binaria)

$$\text{con } x_n, x_l \in A_x = \{x_1, x_2, \dots, x_K\}$$

Alfabeto de la señal base

$$\text{ya que } DP_{xn} = DP_{yn}$$

$\Rightarrow$

$$yt[m] = y_n \in A_y = \{y_1, y_2, \dots, y_K\}$$

Alfabeto de la señal transformada.

### Estimación de las funciones de densidad y distribución de probabilidad de la señal transformada

La estimación de las versiones muestreadas  $fdp_{yt}[y_i]$  y  $DP_{yt}[y_i]$  de  $fdp_{yt}(v)$  y de  $DP_{yt}(v)$  respectivamente, para la señal transformada  $yt = \{yt[1], yt[2], \dots, yt[M]\}$  de  $M$  muestra, con alfabeto  $A_y = \{y_1, y_2, \dots, y_K\}$ , pueden derivarse del histograma  $H_{yt} = \{hyt_1, hyt_2, \dots, hyt_K\}$  (calculado considerando como centroides

de los bins el alfabeto  $A_y$ ) mediante:

$$fdp_{yt}[y_i] = \frac{hyt_i}{M\Delta y_i} \quad \forall i \in [1, K]$$

$$\Delta y_i = \begin{cases} i=1 & y_1 + \frac{y_2 - y_1}{2} \\ 1 < i < K & \frac{y_{i+1} - y_{i-1}}{2} \\ i=K & \frac{y_K - y_{K-1}}{2} \end{cases}$$

$$DP_{yt}[y_i] = \sum_{n=1}^i fdp_{yt}(y_n)\Delta y_n \quad \forall i \in [1, K]$$

$$DP_{yt}[y_j] > DP_{yt}[y_i] \quad \forall j > i \quad i, j \in [1, K]$$

$$fdp_{vw}(vw) = p \left[ \frac{K_1}{A_1} \left( \frac{vw}{A_1} \right)^{K_1-1} e^{-\left(\frac{vw}{A_1}\right)^{K_1}} \right] +$$

$$(1-p) \left[ \frac{K_2}{A_2} \left( \frac{vw}{A_2} \right)^{K_2-1} e^{-\left(\frac{vw}{A_2}\right)^{K_2}} \right] \quad [1]$$

$(0 < p < 1)$   
 $p$  – Factor de poderación  
 $A_1 [m/s]$  – Factor de escala izquierdo  
 $K_1$  – Factor de forma izquierdo  
 $A_2 [m/s]$  – Factor de escala derecho  
 $K_2$  – Factor de forma derecho  
 $vw [m/s]$  – velocidad del viento

## RESULTADOS Y DISCUSIÓN

### Generación de una señal con distribución bimodal a partir de una señal con distribución monomodal

La distribución de probabilidad Weibull&Weibull (W&W) es una distribución bimodal<sup>1</sup> que resulta de la combinación lineal de dos distribuciones monomodales Weibull. Se ha demostrado que esta distribución es más adecuada que la Weibull monomodal en la representación de las características del recurso eólico en sitios en los que existen dos fenómenos distintos que marcan el comportamiento cotidiano de la velocidad del viento, como por ejemplo la combinación de vientos provenientes de un paso montañoso con vientos derivados del fenómeno brisa marina. La función densidad de probabilidad W&W se describe de la siguiente forma (Jaramillo y Borja 2004):

El algoritmo propuesto en el presente trabajo se utilizó para transformar una serie de tiempo de la velocidad del viento (señal base  $x$ ) generada por una estación de monitoreo del recurso eólico en la Facultad de Ingeniería de la Universidad Autónoma de Yucatán (FI-UADY) en una serie de tiempo (señal transformada  $y$ ) con las características del viento descritas por Jaramillo *et al.* (2004) para la temporada de verano en la estación de monitoreo de la Venta en Oaxaca. En ambos sitios se tienen dos fenómenos meteorológicos determinantes del comportamiento cotidiano de la velocidad del viento, de los cuales uno de ellos es la brisa marina. Tanto la señal base  $x$ , como la señal transformada  $y$  (Figura 1) están formada por las muestras de 28 días tomadas a intervalos de 10 segundos (241,920 muestras). La señal base presenta un claro comportamiento monomodal, su ajuste al modelo de Weibull (parámetro de escala  $A$  de 5.7859m/s y parámetro de forma  $K$  de 2.7898) no es conveniente para realizar la transformación puesto que sobrestima la probabilidad de ocurrencia de la velocidad del viento en el rango de 0 a 2.9m/s y en el rango de 5 a 8.6m/s, mientras que

<sup>1</sup> En el sentido estricto es bimodal solo en algunos casos. Sin embargo, generalmente se considera como bimodal.

subestima la probabilidad en el rango de 2.9 a 5m/s (Figura 1 superior derecha). Por lo tanto la transformación de la señal se realizó con la estimación numérica de la distribución de probabilidad para generar una nueva serie de tiempo que se apegara a una distribución de probabilidad objetivo que en este caso es bimodal W&W. Por consecuencia de que la señal base no es continua sino que está representada mediante una señal discreta, la transformación de su función densidad de probabilidad no produce resultados

exactos. En la figura 1 (inferior derecha) se compara la función densidad de probabilidad de la señal transformada (obtenida numéricamente a partir de la serie de tiempo  $y$ ) con la función densidad de probabilidad objetivo obtenida con la ecuación 1. El error cuadrático medio normalizado (ECMN) entre las dos funciones de densidad de probabilidad es de 0.0041 y gráficamente se aprecia la similitud entre ambas funciones.

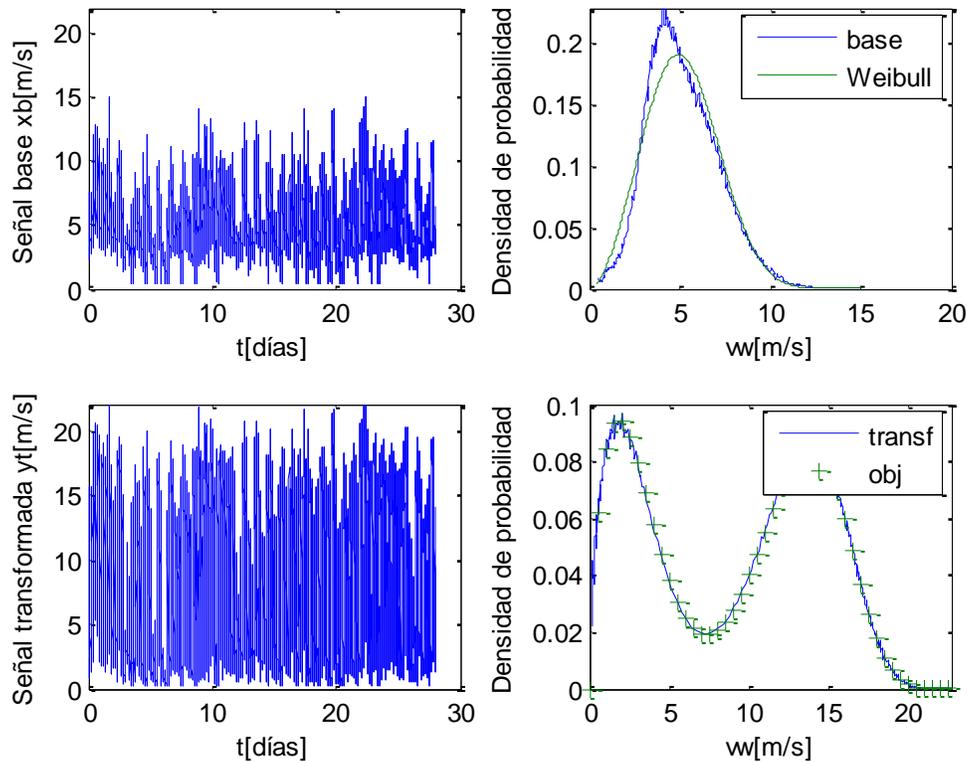


Figura 1. Series de tiempo y funciones densidad de probabilidad de la velocidad del viento en la FI-UADY (señal base) y de su transformación (señal transformada) contemplando las características del viento en el verano en la estación La Venta (señal objetivo).

Para la descripción de la señal objetivo  $y$  con distribución W&W se aplicaron en la ecuación 1, los parámetros que corresponden a la época de verano, descritos por Jaramillo *et al.* (2004), en la estación de monitoreo La Venta y que se especifican a continuación:

$$p = 0.4094, A_1 = 3.285[m/s], K_1 = 1.594, A_2 = 14.308[m/s], K_2 = 5.612.$$

La velocidad promedio del viento para la temporada de verano en la estación de monitoreo La Venta es de 9.017m/s

mientras que la que se obtuvo con la serie de tiempo sintética es de 9.0702m/s. Como es de suponer, la transformación de las series de tiempo con base en sus distribuciones de probabilidad (Figura 2) es una transformación no lineal que tiene efectos sobre el alfabeto de la señal transformada. El alfabeto de la señal de entrada es un subconjunto de las posibles lecturas que el instrumento para la medición de la velocidad del viento puede entregar, y por consecuencia será un conjunto de muestras, en su mayoría, uniformemente espaciadas en un intervalo de velocidad del viento determinado por el valor máximo encontrado en la serie de tiempo de la señal base. Sin embargo, el alfabeto de la señal de salida comprende un conjunto de valores no uniformemente espaciados como se puede ver en la figura 2 inferior izquierda. En este caso la resolución de la señal de salida para el intervalo de velocidad del viento que va de

1m/s a 14m/s tiene una menor resolución que la de la señal base, teniéndose el peor de los casos (>0.4m/s) alrededor de los 7m/s.

La transformación realizada sobre la señal base para obtener la señal  $y = DP_y^{-1}(DP_x(x))$ , al utilizar la relación entre dos funciones monótonamente crecientes, preserva la estructura general del patrón promedio diario de la velocidad del viento (Figura 2 inferior derecha). La señal base corresponde a un sitio en el que la velocidad del viento es baja por la noche, permanece más o menos estable por la mañana (después de haberse incrementado rápidamente al amanecer), se incrementa nuevamente a media tarde por consecuencia de la brisa marina para finalmente empezar a decrecer al caer la noche (Soler *et al.* 2009) y este mismo comportamiento se aprecia en el patrón promedio diario de la señal transformada.

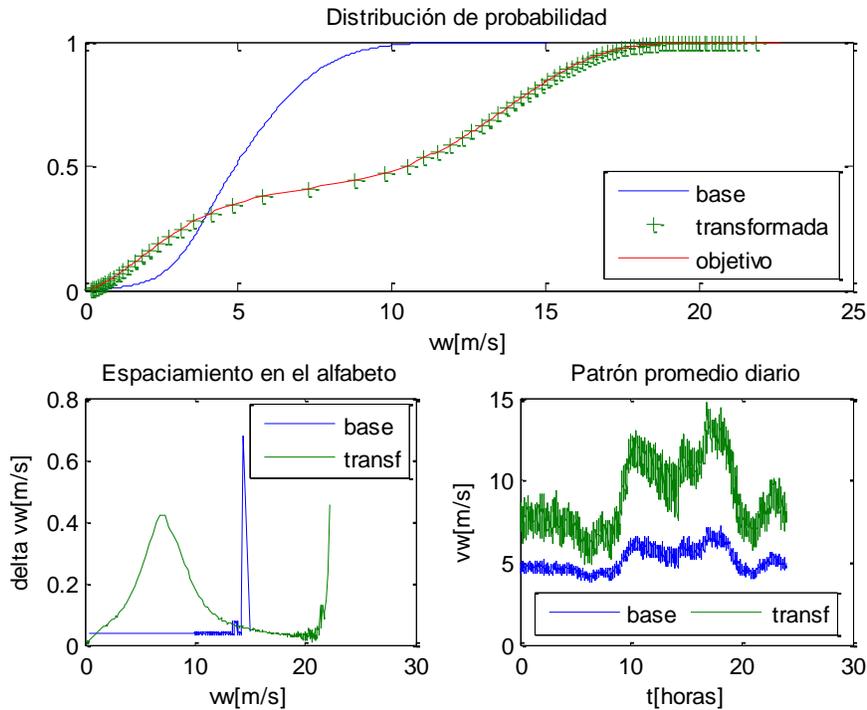


Figura 2. Funciones de distribución de probabilidad para la transformación de la señal base en una señal adecuada a las características de la distribución objetivo y efectos en la cuantificación de la señal ajustada por consecuencia de la transformación.

### Generación de series sintéticas de la velocidad del viento a diferentes alturas sobre el nivel del terreno con una distribución monomodal arbitraria

En el análisis económico del desempeño de un sistema eólico es necesario estimar la producción energética a diferentes alturas sobre el nivel del terreno para optimizar el tamaño de la torre. Este análisis puede estar basado en métodos puramente estadísticos, métodos dinámicos o métodos híbridos. Para los métodos dinámicos es necesario contar con series de tiempo de la velocidad del viento en función de la altura sobre el nivel del terreno. El algoritmo propuesto en el presente trabajo se utilizó para transformar una serie de tiempo de la velocidad del viento (señal base  $x$ ) generada por una estación de monitoreo del recurso eólico en la Facultad de Ingeniería de la Universidad Autónoma de Yucatán (FI-UADY) en dos series de tiempo (señales transformadas  $yt_1$  e  $yt_2$ ) con las

características del viento para las alturas de 15m y 25m sobre el nivel del terreno. La serie de tiempo de la señal base corresponde a mediciones realizadas a 35m sobre el nivel del terreno. Tanto la señal base  $x$ , como las señales transformadas  $yt_1$  e  $yt_2$  (ver figura 3) están formadas por las muestras de 28 días tomadas a intervalos de 10 segundos (241,920 muestras). Las distribuciones de probabilidad de la velocidad del viento para las señales base, objetivo1 y objetivo2, requeridas para la obtención de las señales transformadas a partir de la señal base, han sido estimadas en los tres casos a partir del procesamiento de series de tiempo de la velocidad del viento. Como se puede apreciar en la figura 4, la transformación preserva la estructura general del patrón promedio diario (en este caso calculado sobre un periodo de 28 días) de las señales transformadas para las alturas de 15m y 25m.

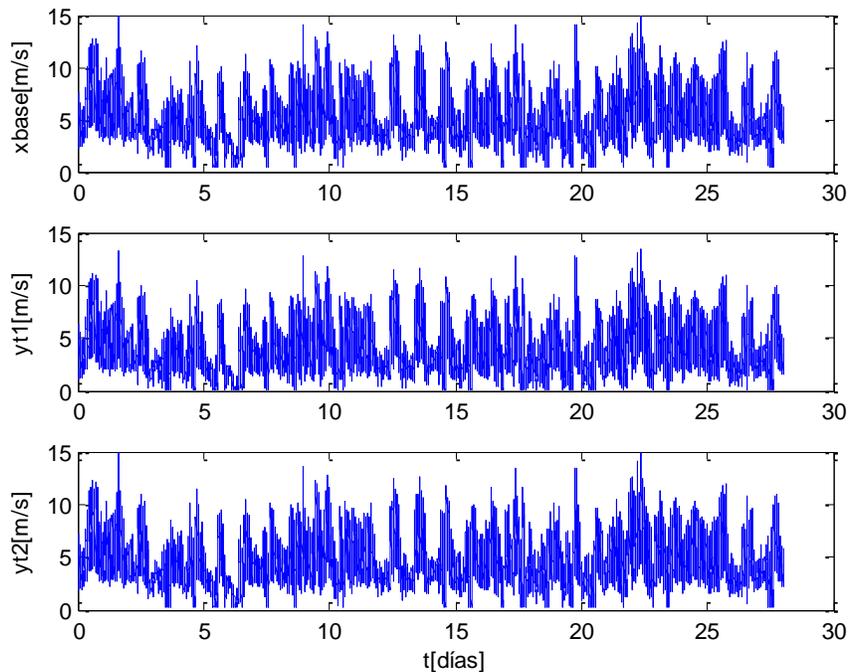


Figura 3. Series de tiempo de la velocidad del viento en la FI-UADY (señal base a 35m) y de su transformación contemplando las características del viento a 15m y 25m de altura sobre el nivel del terreno.

Otro de los parámetros que es reproducido adecuadamente por el algoritmo es el de la velocidad media global de las series de tiempo. La velocidad media de la serie transformada  $yt_1$  es de 3.9728m/s, muy cercano al valor objetivo que es de 3.9767m/s. Para el caso de  $yt_2$  su valor medio es de 4.6992m/s para un valor objetivo de 4.7001m/s. En este caso la transformación ha tenido poco efecto sobre la resolución de las señales en el rango de 1 a 12m/s. En este intervalo la resolución de las señales transformadas es cercano a 0.035m/s como en el caso de la señal base (Figura 4 inferior derecha).

datos de la velocidad del viento al modelo de Weibull (tanto de la señal base como de las señales objetivo) se encontró que el ajuste a modelo presenta el inconveniente de sobreestimar la probabilidad de ocurrencia de valores bajos y altos de velocidad, mientras que subestima la probabilidad de ocurrencia de los valores medios, por consecuencia es evidente que no es conveniente realizar la transformación de la señal base en las señales objetivo utilizando el modelo de Weibull para describir la función de distribución de estas variables aleatorias. Por consecuencia es evidente la ventaja del algoritmo presentado en este trabajo de hacer la transformación con base en la descripción numérica de las distribuciones de probabilidad.

Como se puede ver en la figura 5, las funciones densidad de probabilidad obtenidas numéricamente muestran un comportamiento monomodal. Al ajustar los

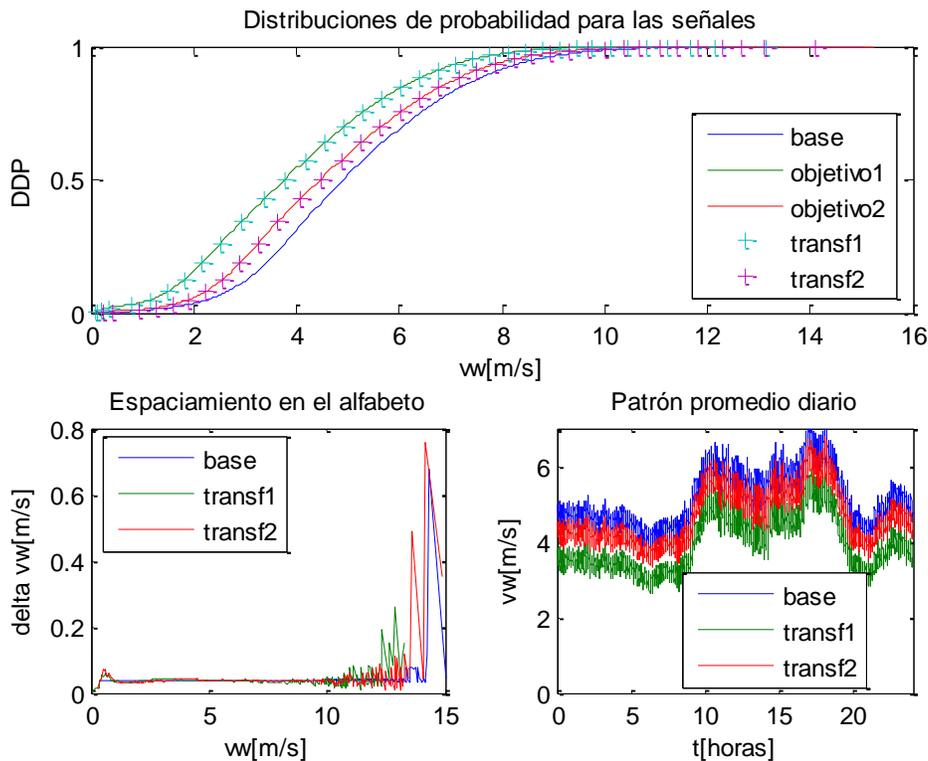


Figura 4. Distribuciones de probabilidad descritas numéricamente empleadas en la transformación, espaciamento de los alfabetos y patrones promedio diario para las señales base, transformada 1 y transformada 2.

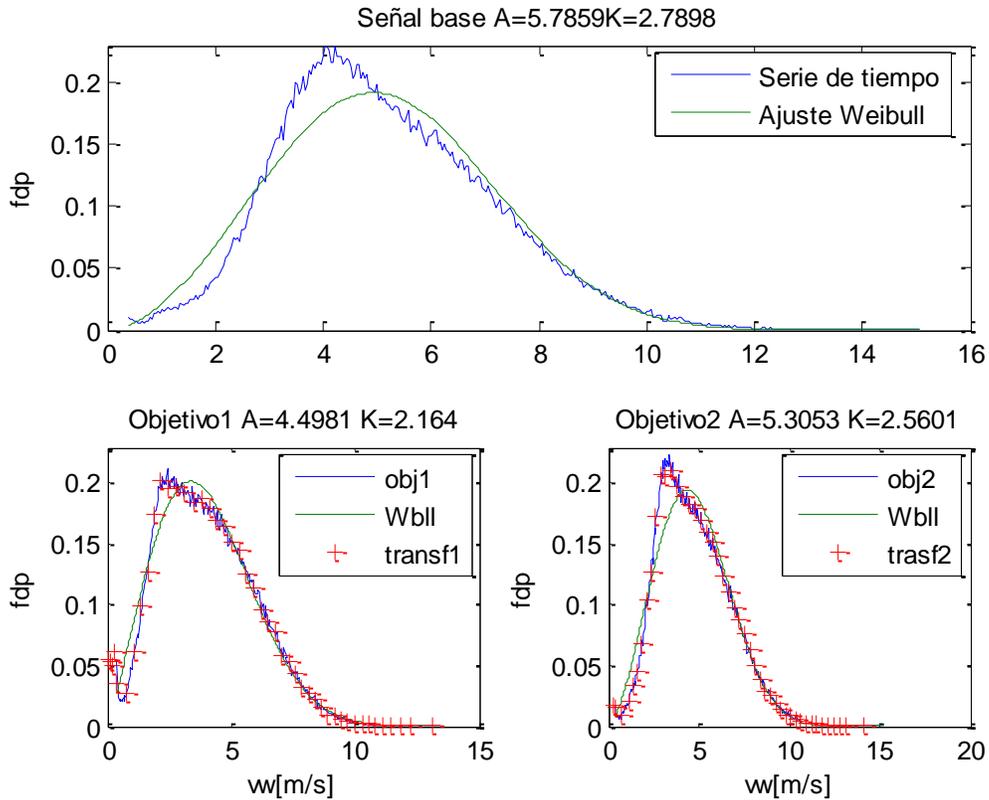


Figura 5. Funciones de densidad de probabilidad obtenidas numéricamente y su ajuste al modelo de Weibull muestran la inconveniencia de hacer la transformación con base en la función que describe a este modelo.

Como los alfabetos de la señal objetivo y el de la señal transformada son diferentes, es imposible hacer una estimación directa del error cuadrático medio normalizado entre la función densidad de probabilidad objetivo y la función densidad de probabilidad de la señal transformada. Sin embargo, se puede apreciar gráficamente en la figura 5 la aproximación entre ambas funciones para las dos señales transformadas.

### CONCLUSIONES

En este documento se presenta un algoritmo para obtener series de tiempo sintéticas a partir de series obtenidas experimentalmente y una de la descripción de la velocidad del viento mediante su función densidad de probabilidad. Este trabajo tiene aplicación para el modelado

del viento en el estudio del comportamiento de sistemas eólicos frente a diferentes condiciones de operación.

Las series de tiempo obtenidas mediante la transformación de una serie de tiempo base por medio del algoritmo propuesto se ajustan a la distribución de probabilidad objetivo.

Se obtuvieron experimentalmente resultados adecuados en la transformación de series de tiempo con distribuciones de probabilidad monomodales en series de tiempo con distribución monomodal y en series de tiempo con distribución bimodal.

La resolución de los datos para la representación de la series de tiempo de la velocidad del viento es afectada por la transformación, por lo tanto es importante profundizar en el estudio de los efectos que tiene la afectación en la resolución de la

serie de tiempo de velocidad del viento de salida en la simulación del comportamiento dinámico de los sistemas eólicos.

Aún cuando las series de tiempo que se obtienen con la metodología descrita solo pueden tener una duración menor o igual a la de la serie de tiempo original, esto no es una limitante si dichas series se utilizan para evaluar el comportamiento dinámico de sistemas eólicos ante las condiciones planteadas por la distribución de

probabilidad objetivo.

Las transformación propuesta preserva la estructura general del patrón promedio diario de la velocidad del viento, lo cual es deseable si se busca evaluar el comportamiento dinámico de un sistema eólico y sus efectos sobre la red de energía eléctrica en un sitio hipotético con comportamiento diurno de la velocidad del viento similar al del sitio en el cual se obtuvo la señal base.

## REFERENCIAS

Carapellucci R. & Giordano L. (2013a). A methodology for the synthetic generation of hourly wind speed time series based on some known aggregate input data. *Applied Energy* 101 pp 541-550.

Carapellucci R. & Giordano L. (2013b). The effect of diurnal profile and seasonal wind regime on sizing grid-connected and off-grid wind power plants. *Applied Energy* 107 pp 364–36.

Carta J. A., Ramírez P. Velázquez S. (2008). Influence of the level of fit of a density probability function to wind-speed data on the WECS mean power output estimation. *Energy Conversion and Management* 49 pp 2647-2655.

Celik A. N. (2003). Energy output estimation for small-scale wind power generators using Weibull-representative wind data. *Journal of Wind Engineering and Industrial Aerodynamics* 91 pp 693-707.

Coltuc D., Bolom P., Chassery J.M. (2006) Exact Histogram Specification. *IEEE Transactions on Image Processing*. Vol. 15, No. 5, pp 1143-1152.

Emeis S. (2013). *Wind Energy Meteorology. Atmospheric Physics for Wind Power Generation*. Springer-Verlag, Berlin Heidelberg.

Jaramillo O. A., Borja M. A. (2004) Wind speed analysis in La Ventosa, México: a bimodal probability distribution case. *Renewable Energy* 29 pp 1613-1630.

Jung J., Broadwater R. P. (2014). Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews* 31 pp 762-777.

Ramírez P, Carta JA. (2006) The use of wind probability distributions derived from the maximum entropy principle in the analysis of wind energy. A case study. *Energy Conversion and Management* 47 pp 2564–77.

Nfaoui H., Buret J. and Sayigh A. A. M. (1996). Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco). *Solar Energy* Vol 56 No 3 pp 301-314.

Shamshad A., Bawadi M. A., Wan Hussin W. M. A., Majid T. A, Sanusi S. A. M. (2005). First and second order Markov chain models for synthetic generation of wind speed time series. *Energy* 30 pp 693-708.

Soler R., Watson S. and Infield D. (2009). Evaluation of the Wind Shear a a Site in the North-West of Yucatan Peninsula, Mexico. *Wind Engineering* 33 No. 1, pp 93-107.

Suomalainen K, Silva C. A., Ferrão P. and Connors S. (2012). Synthetic wind speed scenarios including diurnal effects: Implications for wind power dimensioning. *Energy* 37 pp 41-50.