# International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

RESEARCHERID
THOMSON REUTERS

ISSN: 2277-9655
[Bokovi * *et al.*, 8(8): August, 2019] Impact Factor: 5.164
IC™ Value: 3.00 CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# WAVELET METHOD FOR ARTIFICIAL GENERATION OF WIND SPEED: CASE OF LOME SITE

**Yao Bokovi [*1], Eyouléki Tcheyi Gnadi Palanga[2] & Adekunlé Akim Salami[3]**
Laboratory of Research in Sciences of the Engineer, National Superior School of Engineers, Department of Electricity, University of Lome, Togo

**DOI**: 10.5281/zenodo.3360426

## ABSTRACT

The use of wind speed data is of undeniable interest in the operation of wind farms. As such, the techniques of artificial generation of wind data are used in case the need to have a large database is required. Several methods are used for this purpose. The most classic are those of normal and Weibull distributions. In this study, a new approach based on wavelet transformation is introduced. It is compared to classical methods. The results show that the wavelet method is an adequate alternative to conventional methods.

**KEYWORDS:** Normal Distribution; Weibull distribution; wavelet.

## 1. INTRODUCTION

Rapid growth in energy requirements, the depletion of hydrocarbon stocks and environmental degradation have led men to turn to clean, renewable energy sources, including wind energy [1,2]. Nevertheless, the exploitation of wind energy, to be sustainable, must take into account certain requirements. These requirements include both qualitative and quantitative aspects of wind speed data [3]. The size of a database is relative to the time period during which the data was saved. More often than not, a larger base is needed, as in the case of the installation of a wind farm or a civil engineering structure. The wind speed, to be exploitable, must be within a required range [4]. An artificial generation of data, which to a large extent reflect the reality, is a necessary step to achieve satisfactory results. To do this, modeling techniques have been developed. The present work introduces a new approach to artificial generation based on the wavelet transform. Statistical methods such as normal distribution and Weibull will also be presented. The results of these different methods will be the subject of a comparative study.

## 2. MATERIALS AND METHODS

In this study, statistical methods were used to account for the characteristics of wind speed. The comparison of the reproduction performances of the long series of hourly speeds proves decisive for the validation of the method. In this regard we have hourly wind speed data collected by the US University Woayoming, a department of NASA, concerning the Lomé site. The choice of normal and Weibull distributions was made in order to generate uncorrelated and random data respecting the distribution. The statistical methods mentioned above are briefly presented. Particular emphasis will be placed on the wavelet method in this paper.

### a. Normal distribution

The normal law, or Gaussian law, is the most widespread and useful statistical law because it represents many random phenomena [5,6].

A normal distribution corresponds to the probability distribution of a continuous random variable whose curve is perfectly symmetrical, unimodal and bell-shaped. In addition, many other statistical laws can be approximated by the normal law, especially in the case of large samples. It also has the advantage of relying only on two parameters: the mean and the standard deviation. Its mathematical expression is given by relation (1).

$$n(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(x-\mu)^2 / 2\sigma^2\right] \quad (1)$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

**b. Distribution of weibull**

The Weibull distribution [7,8] presents itself as a special case of generalized gamma distribution. There are two forms of Weibull distribution:
In this study, we will use the two-parameter Weibull distribution.
The expression of Weibull distribution law (probability density) with two parameters is given by relation (2).

$$f(V)=\left(\frac{K}{C}\right)\left(\frac{V}{C}\right)^{K-1}.\exp\left[-\left(\frac{V}{C}\right)^{K}\right] \qquad (2)$$

V is the wind speed in m / s;
C, the scale factor (dimension of a speed);
K, the form factor (dimensionless) characterizing the dissymmetry of the distribution.
C and K are the parameters of the distribution.
The distribution law F (V) is given by relation (3).

$$F(V)=1-\exp\left[-\left(\frac{V}{C}\right)^{K}\right] \qquad (3)$$

F (V) is the probability that the wind speed is less than V.
The parameters of the Weibull distribution are determined by a numerical method based on the least squares method.

The artificial generation by the Weibull distribution is similar to that of the normal distribution. Only the curve of the Weibull distribution is obtained as a result of the numerical determination of the Weibull parameters (form factor K and scale factor C) characterizing the empirical series.

**c. The wavelet method**

The wavelet transformation was introduced in the early 1980s by the French geophysicist Morlet [9]. Interested in the study of seismic signals involved in oil exploration, he realizes that the "classical" spectral decompositions (Fourier transform and Fourier transform with sliding window [10]) are unsuited to the analysis of signals combining several characteristic scales. To remedy this, he proposes a transformation that allows a representation of the signal simultaneously in time (or in space) and in scales. The signal is no longer decomposed into frequency components, but into a linear combination of elementary functions located at different points of space and having different sizes. These elementary functions are all constructed from a single parent function, by dilations and translations of it according to relation (4).

$$\psi_{b,a}(x)=\psi\left(\frac{x-b}{a}\right). \qquad (4)$$

In 1984, with Grossman, Morlet gives a rigorous framework to the concepts of this new space-scale decomposition [11,12]. In particular, it demonstrates that for the initial signal to be effectively decomposed as a linear combination of elementary functions, the mother function must have some oscillations and thus resemble a wavelet (small wave). This is how the wavelet theory is born.

**i. Definition and properties**

Due to the richness of the concepts it brings into play and the efficiency of its algorithmic implementation, wavelet analysis has been successfully exploited in fields as varied as functional analysis, theoretical physics, analysis and processing of signal and images, fractal theory, etc.

The wavelet decomposition consists in breaking down the signal s (which is the wind speed in our case) as a linear combination of wavelets [13,14]. This decomposition is a function of the two variables a and b, which "evaluates" the relevance of using the wavelet in the description of s.
It is naturally defined by the relation (5).

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

$$T_\psi[s](b,a) = \frac{1}{a} < \psi_{b,a} | s > = \frac{1}{a}\int \overline{\psi}\left(\frac{x-b}{a}\right) s(x)dx \quad (5)$$

where is the scalar product of two functions and the conjugate of. The parameter b is a position parameter (or temporal) and varies in IR while a is a scale parameter (or frequency) and is strictly positive. Thus, the more the wavelet "resembles" the velocity signal s (x) locally (i.e., over a distance proportional to a) around the point x = b, the more the absolute value of the wavelet transform will be great.

By linearly recombining each wavelet weighted by the wavelet coefficient associated with it, we obtain the reconstruction formula expressed by the relation (6).

$$s(x) = \frac{1}{C_\psi}\iint T_\psi[s](b,a).\psi_{b,a}(x).\frac{dbda}{a^2} \quad (6)$$

where is a constant number (which depends only on the chosen wavelet).
Note that this reconstruction formula is correct only in the case where an admissibility condition verifies that it has "enough" oscillations. We will content ourselves here with supposing that the mother function is of average zero (relation 7).

$$\int \psi(x)dx = 0 \quad (7)$$

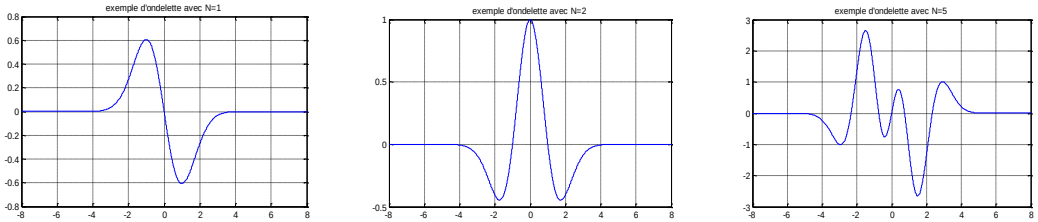
Such a function is called an analyzer wavelet.
Condition (8) leaves a lot of freedom on the choice of the wavelet analyzer. This will be chosen according to the problem dealt with; thus, as we shall see later, the class of wavelet analyzers that constitute the successive derivatives of the Gaussian function is generally used (Relation 8).

$$\psi^{(N)}(x) = \frac{d^N}{dx^N}\left(e^{-\frac{x^2}{2}}\right) \quad (8)$$

These functions are and have the advantage of being well localized both in the direct space and in the Fourier space. We have shown, in Figure 3, the functions for N = 1, 2 and 5. It should be noted that the wavelet is often used in the literature; it is called "Mexican hat" because of its characteristic shape (Fig.3). As we can see, the more N increases, the more is oscillating. This can be explained by the fact that the relation (9) is satisfied.

$$\int x^q \psi^{(N)}(x)dx = 0, \quad 0 \le q < N. \quad (9)$$

The function thus has its N first null moments. As we will see later, this property proves to be crucial in the analysis of the singularities of a signal. Figure 3 shows the wavelet $\psi^{(1)}$ in (a), the wavelet $\psi^{(2)}$ in (b) and the wavelet $\psi^{(5)}$ in (c).




*(a) wavelet $\psi^{(1)}$* *(b) wavelet $\psi^{(2)}$* *(c) wavelet $\psi^{(5)}$*

***Figure 3: Three examples of wavelet analyzers of the family $\psi^{(N)}$ of Gaussian derivatives (rel.9)***

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Moreover, there are other forms of wavelets such as the wavelet of Morlet and the wavelet of Haar. The latter will be the subject of our study on the pages to come.

**ii. Algorithm**

The model proposed in this paper measures, by multiresolution analysis, the weights of the wavelets associated with the speed signal at each level of resolution. The next step is the simulation of the data. To achieve this step you have to choose between two alternatives:

- simulate the data from the measured weights;
- simulate the data from the generation of details.

To appreciate the real performance of the model, we compute the average error between the speed signal and that generated by the model. Here is the algorithm itself.

**a. Calculation of weights**

The multiresolution analysis as described above, reveals coefficients of scale (or approximation) and coefficients of wavelet (or detail). The scaling coefficients are obtained by convolution of the signal s (x) with a scaling function noted in relation (10) :

$$c_{j,k} = 2^{j/2} \int s(x)\phi(2^j x - k)dx \qquad (10)$$

Where j is the resolution index and k is a position parameter.
In the case of this project, it is used the Haar wavelet [4], shown in Figure 5, for its simplicity and the orthogonality of its base. Thus, the associated scale function is given by relation (11).

$$\phi(x) = \begin{cases} 1 & \text{pour } 0 \le x \le 1 \\ 0 & \text{ailleurs} \end{cases} \qquad (11)$$

It is represented in figure 4.

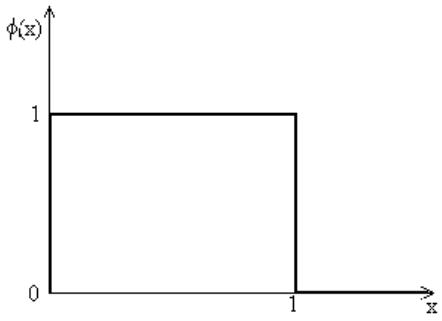



***Figure 4: Scale function for the Haar wavelet***

The translated and expanded version of this function is commonly referred to as (12).

$$\phi_{j,k}(x) = 2^{j/2}\,\phi(2^j x - k) \qquad (12)$$

Like the scale coefficients, the wavelet coefficients are derived from convolution products of the speed signal s (x) with a function, this time called the wavelet analyzer or wavelet function. This is indeed the Haar wavelet defined by the relation (13) and represented in figure 5.

$$\psi(x) = \begin{cases} +1 & \text{pour } 0 \le x \le 1/2 \\ -1 & \text{pour } 1/2 \le x \le 1 \\ 0 & \text{ailleurs} \end{cases} \qquad (13)$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Its representation is recorded in Figure 5.

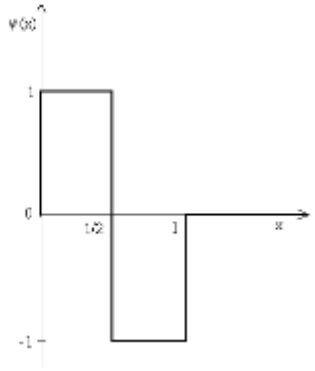



***Figure 5: Wavelet of Haar.***

The coefficients of detail (or wavelet) noted $d_{j,k}$ are obtained by the relation (14).

$$d_{j,k} = 2^{j/2} \int s(x)\psi(2^j x - k)dx \qquad (14)$$

**b. Simulation**

From the scale and wavelet coefficients previously calculated, respectively find the approximation and detail signals associated with each level of analysis. Thus, the approximation is obtained by the relation (15).

$$A_j = \sum_{k \in Z} c_{j,k} * 2^{j/2} \phi(2^j x - k) \qquad (15)$$

The details are found by the relation (16).

$$D_j = \sum_{k \in Z} d_{j,k} * 2^{j/2} \psi(2^j x - k) \qquad (16)$$

From one approximation to another, we lose information; this loss corresponds to the detail of the higher level and results in the relation (17).

$$A_j - A_{j+1} = D_{j+1} \qquad (17)$$

Recurrently we have the relationship (18)

$$\begin{cases} A_0 - A_1 = D_1 \\ A_1 - A_2 = D_2 \\ A_2 - A_3 = D_3 \\ A_3 - A_4 = D_4 \\ \quad . \quad . \quad . \quad . \\ \quad . \quad . \quad . \quad . \\ \quad . \quad . \quad . \quad . \\ A_{n-1} - A_n = D_n \end{cases} \qquad (18)$$

where, n is a natural integer such that: $2^n \leq N < 2^{n+1}$, N being the length of the series. The member-to-member sum leads to the relation (19).

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

$$A_0 - A_n = \sum_{j=1}^{n} D_j \qquad (19)$$

Which is the equivalent of the relation (20).

$$A_0 = A_n + \sum_{j=1}^{n} D_j \qquad (20)$$

being logically the approximation of level 0, it is identical to the signal itself.

$$A_0 = s(x) \qquad (21)$$

In turn, the signal is obtained from the last approximation and the details following the relation (22).

$$s(x) = A_n + \sum_{j=1}^{n} D_j \qquad (22)$$

This simulation method is very interesting because it leads to a modeling error, probably, very small. This accounts for the reliability of the model. Nevertheless, there is a question of parsimony. Indeed, the sum of the lengths of all the coefficients entering in this simulation method is equal to the length of the signal. This seems to greatly reduce the interest of this approach. To answer this question of parsimony, one benefits from an important characteristic of the wavelets and particularly of the wavelet of Haar.

As noted above, the approximations, with the Haar wavelet, are nothing more than averages of the speed signal. Which imposes an almost zero average on the details. They are therefore assimilated to turbulence signals. The sum of the details usually has a distribution close to a known distribution. We can thus characterize by the scale coefficients at the maximum level (often two coefficients) and the distribution parameters of the sum of the details. In the case of this study, the sum of the details is characterized by Gaussian law. It goes without saying that we have two distribution parameters. The number of parameters of the model is thus reduced to four (4), hence parsimony.

The wavelet transform is a very valuable tool that can not be circumvented today in signal processing. Besides its representation as faithful as possible, its sharp analysis offers a sort of real microscope for the analysis of signals. The time-frequency representation and the power in detection of singularity are his weapons. All these benefits will allow us to realize the model of the wind speed based on the algorithm developed here. The application of the model to the Lomé site will essentially be the subject of the following section.

## 3. RESULTS AND DISCUSSION

Statistical methods and wavelet methods will be used and compared.

**Presentation of the data**

The empirical data on which this study is based are provided by the WOAYOMING website. These are hourly data. These data include, among others, the date and time of recording, temperature, wind direction, pressure and wind speed.

The data collected covers a period of eighteen months and cover the period from July 2017 to December 2018. The parameter that interests this project is the speed. Figure 6 shows the frequency polygon of the wind speed at the Lomé site.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

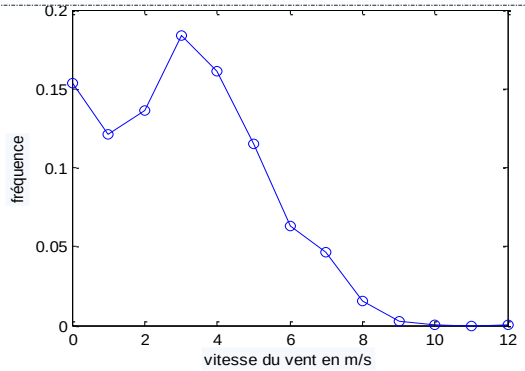



***Figure 6: Frequency polygon***

The frequency polygon is the curve that at each possible velocity value associates the corresponding frequency in the velocity series. On this polygon it appears that the most frequent speed is 3 m / s. The speed of 11m / s is nonetheless non-existent in the data, since its frequency is strictly zero. The maximum speed is 12 m / s. Table 1 shows the statistical parameters of the observed data.

***Table 1: Statistical parameters of the wind speed at the Lomé site.***

| Période | Number of data | Average (m/s) | Mode (m/s) | Median (m/s) | Maximum speed (m/s) | Standard deviation (m/s) | Coefficient of asymmetry (Skewness) | Flattening coefficient (Kurtosis) |
|---|---|---|---|---|---|---|---|---|
| From July 1st, 2004 to December 31st, 2005 | 13176 | 3.0221 | 3.0000 | 3.0000 | 12.0000 | 2.1063 | 0.3055 | 2.4185 |

The speed signal is temporally as shown in figure 7.

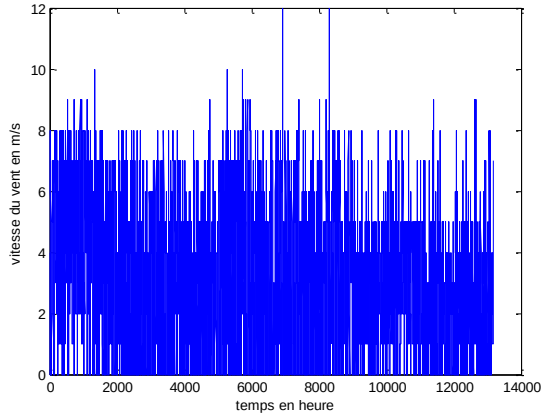



***Figure 7: The wind speed signal at the Lomé site***

**Statistical methods**

***Adjustment by the normal distribution***

Normal distribution plays a very important role in statistics, and many tests can only be applied if the data is reasonably normal. It is therefore important to be able to compare the observed empirical distribution with a theoretical normal distribution. We will thus superimpose the histogram obtained from the real speed data and a curve representing a normal distribution of same average and standard deviation. If the normal curve passes exactly through the vertices of the histogram, the distribution is perfectly normal. This is of course only exceptionally the case, and all empirical distributions depart more or less from normality. This technique is

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

called adjustment of a distribution by the normal distribution. Applied to wind speed, the adjustment by the normal distribution is shown in Figure 8.

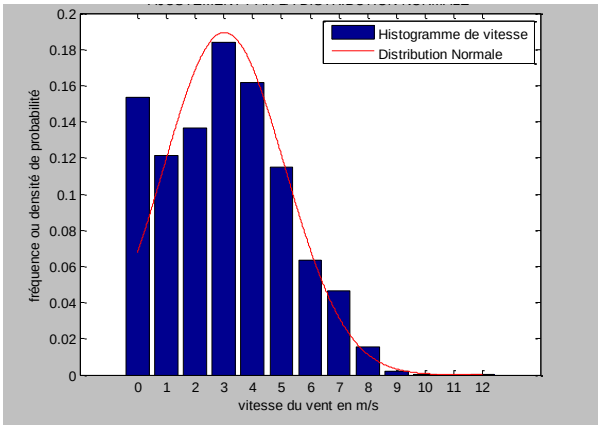



***Figure 8: Adjustment of the empirical distribution by the normal distribution***

It appears clearly, with regard to the result of the adjustment recorded in figure 8, that the probability density function of the normal law of the data series nearly matches the histogram. On the other hand, the probability of occurrence of zero velocity (0m / s) does not respect the normal law.

However, the state of the approximation will be much more revealed in the comparative study of the two methods. Now let's look at the adjustment by the Weibull distribution.

***Adjustment by the Weibull distribution***

The model derived from the Weibull distribution receives as input two parameters characterizing the real series of velocity: the scale factor C and the form factor K. These are parameters intrinsic to the site. For the Lomé site. These parameters were determined from the available speed data.

***Result***

The normal distribution and Weibull distribution parameters are calculated and stored in Table 2.

***Table 2: Normal and Weibull Distribution Parameters***

| | **Settings** | |
|---|---|---|
| **Weibull distribution** | C= 3.6195 | K = 1.7558 |
| **Normal Distribution** | μ= 3.0221 | σ = 2.1063 |

Figure 9 illustrates the probability density function of the Weibull distribution compared to the histogram of the observed data series.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

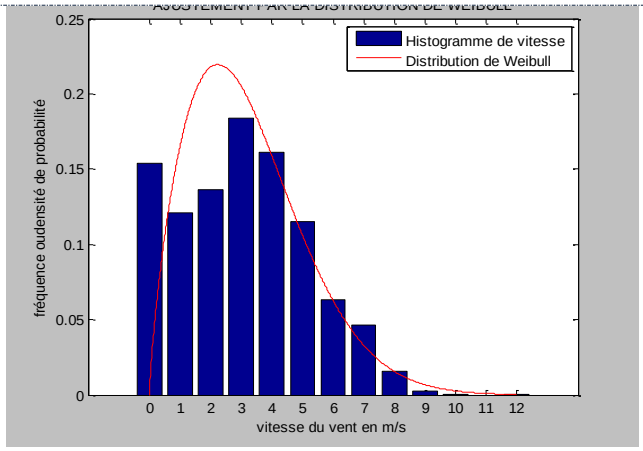



***Figure 9: Weibull distribution function fitted to the histogram of observed data***

It is found that the Weibull distribution also adjusts the empirical distribution of the data set. However it would be difficult even clumsy to say, visibly, anything about the best approximation. The comparative study of the two models will draw a good conclusion.

***Comparative study of the probability density functions of the normal distribution and Weibull***

In Figure 10, the distribution functions of Normal and Weibull are juxtaposed in order to stand out, which is better adjusted to the histogram corresponding to the series of data observed.

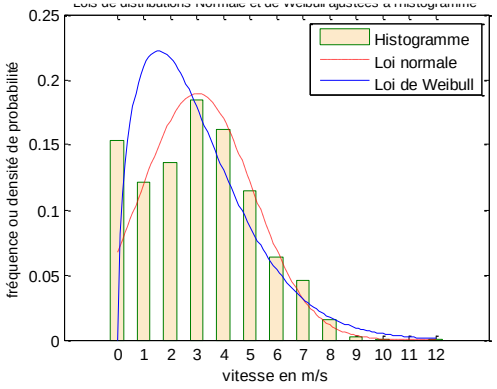



***Figure 10: Normal and Weibull Distribution Functions Adjusted to the Data Series***

From the juxtaposition of the two distributions and the histogram of the actual distribution, it is clear that it is the normal distribution that adjusts for the best.

After the results of the statistical models, we approach the wavelet method.

**Wavelet method**

It will first be presented a brief analysis of the speed signal by the Haar wavelet. Then, we will discuss the detection of singularities and finally the validation results will be presented.

***Signal analysis by the Haar wavelet***

There is a series of eighteen (18) month data from July 2017 to December 2018, giving a total of 13,176 data, which are hourly speeds. The maximum level of wavelet analysis is important. At this level, the number of coefficients is as small as possible. It is equal to two (2). From one level to another directly higher, the ratio of

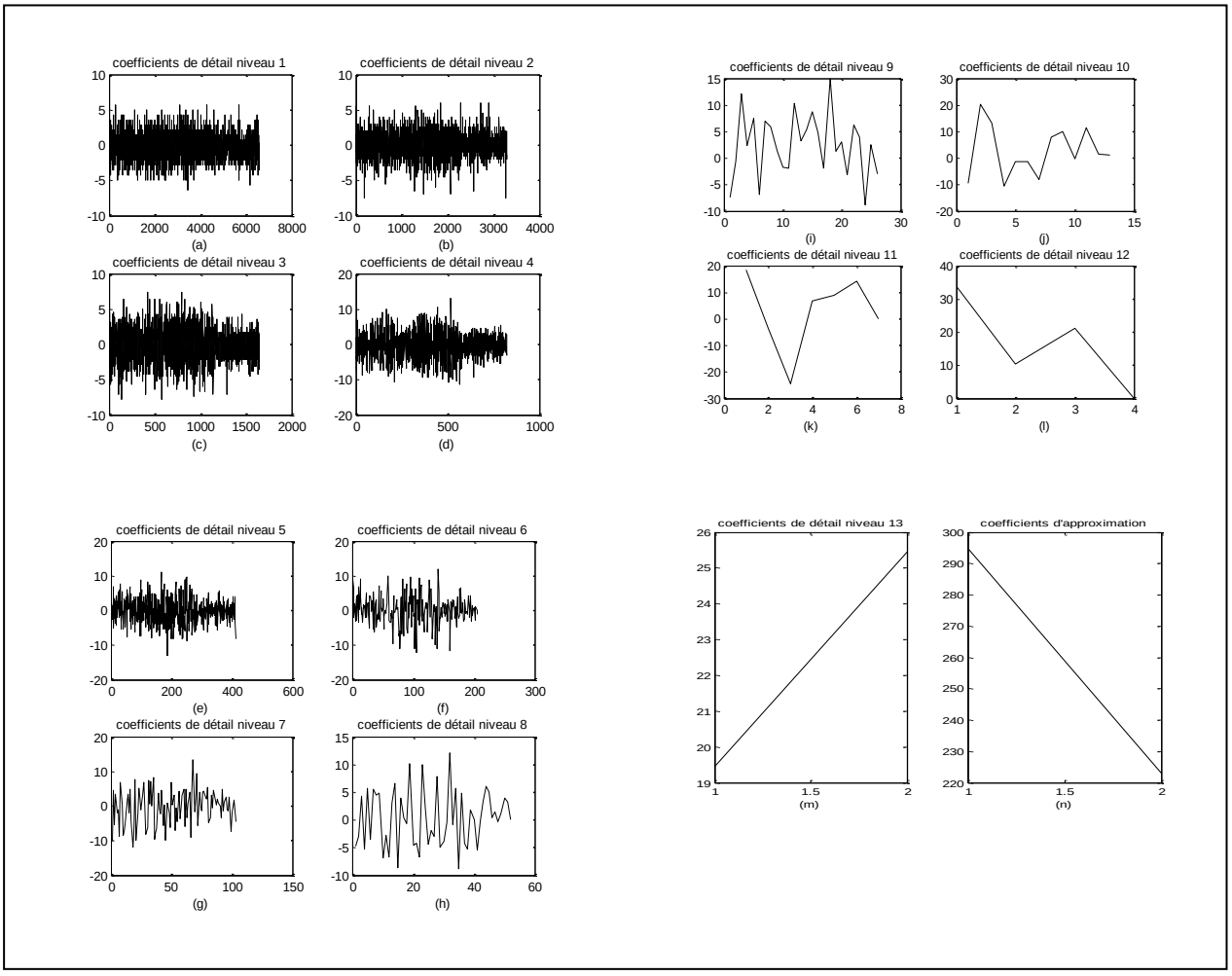
ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

the numbers of coefficients is two. Which leads to the relation (23) giving the number of possible decompositions.

$$n = E[\log_2(N)] \quad (23)$$

where N = 13176 in our case (the number of data); and E symbolizes the whole party function. Which gives n = 13.

The results of the analysis are shown in Figure 11 which represents the wavelet coefficients at the different levels and the scaling coefficients at the thirteen level.



ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

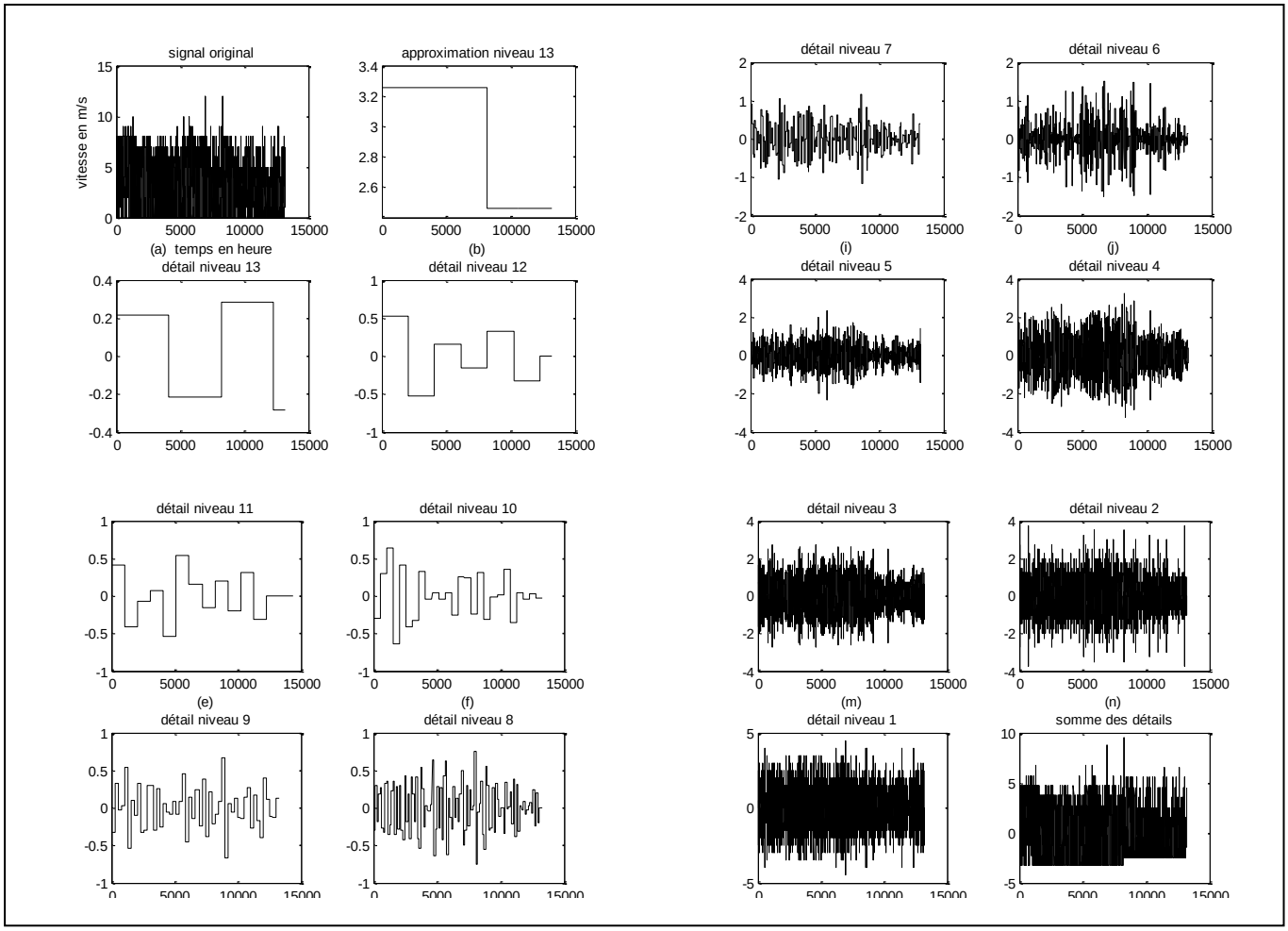



***Table 3: Validation of the wavelet model by the statistical parameters***

| Settings | Real series | Simulated series | Absolute relative error (%) |
|---|---|---|---|
| Average | 3.0221 | 3.1011 | 2.6 |
| Médian | 3.0000 | 3.0385 | 1.2 |
| Mode | 3.0000 | 3.0000 | 0 |
| Maximum | 12.0000 | 11.0205 | 8.1 |
| Minimum | 0.0000 | 0.0000 | - |
| Standard deviation | 2.1063 | 1.9832 | 5.8 |
| Skewness | 0.3055 | 0.3167 | 3.6 |
| Kurtosis | 2.4185 | 2.6326 | 8.8 |

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

| | | | |
|---|---|---|---|
| Scale factor (C) | 3.6195 | 3.4636 | 4.3 |
| Form factor (K) | 1.7558 | 1.6250 | 7.4 |

This table shows that the model provides a signal much like the reference. All absolute relatives errors are less than 10%. Figure 14 compares the histogram of the signal simulated, by the mathematical model of wavelets, with that of the real signal.

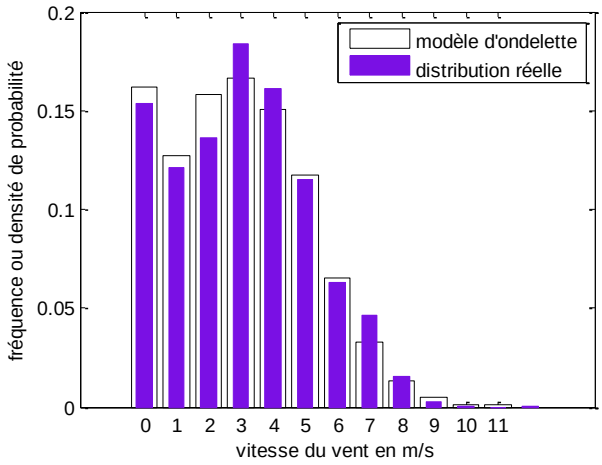



***Figure 14: Validation of the wavelet model***

## 4. CONCLUSION

In this paper, an algorithm based on wavelet decomposition has been proposed to generate the series of wind speeds on a wind power site. This algorithm is applied to the dataset collected at the Lomé site. The results obtained show us that the proposed algorithm presents a weak modeling error compared to the models of artificial generation from the Gauss and Weibull law. Thus, the model based on the proposed wavelet decomposition can be used to generate hourly wind speed time series at the Lomé wind farm with an absolute relative error less than 10%.

## REFERENCES

[1] Boyle, G. (2004). Renewable energy. Renewable Energy, by Edited by Godfrey Boyle, pp. 456. Oxford University Press, May 2004. ISBN-10: 0199261784. ISBN-13: 9780199261789, 456.
[2] European Wind Energy Association. (2009). The economics of wind energy. EWEA.
[3] Ramírez, P., & Carta, J. A. (2005). Influence of the data sampling interval in the estimation of the parameters of the Weibull wind speed probability density distribution: a case study. Energy Conversion and Management, 46(15-16), 2419-2438.
[4] Aksoy, H., Toprak, Z. F., Aytek, A., &Ünal, N. E. (2004). Stochastic generation of hourly mean wind speed data. Renewable energy, 29(14), 2111-2131.
[5] Wang, Li, et al. "Active contours driven by local Gaussian distribution fitting energy." Signal Processing 89.12 (2009): 2435-2447.
[6] Davia, F. G., Herrera, S. E., &Calvo, E. J. (2019). Gaussian Distribution of Electron Transfer Distance in Redox Terminated Self-Assembled Thiol Monolayers. The Journal of PhysicalChemistry C
[7] Pishgar-Komleh, S. H., Keyhani, A., &Sefeedpari, P. (2015). Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh country of Iran). renewable and sustainable energy reviews, 42, 313-322.
[8] Dokur, E., Ceyhan, S., &Kurban, M. (2019). Comparative Analysis of Wind Speed Models Using Different Weibull Distributions. Electrica, 19(1), 22-28.
[9] Morlet, J., Arens, G., Fourgeau, E., &Glard, D. (1982). Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media. Geophysics, 47(2), 203-221.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[10] Fogel, I., &Sagi, D. (1989). Gabor filters as texture discriminator. Biological cybernetics, 61(2), 103-113.
[11] Grossmann, A., &Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM journal on mathematical analysis, 15(4), 723-736.
[12] Goupillaud, P., Grossmann, A., &Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. Geoexploration, 23(1), 85-102.
[13] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis& Machine Intelligence, (7), 674-693.
[14] [Rao, R. (2002). Wavelet transforms. Encyclopedia of Imaging Science and Technology.
[15] [Chen, S. S., Donoho, D. L., & Saunders, M. A. (2001). Atomic decomposition by basis pursuit. SIAM review, 43(1), 129-159