Data Article

Data set on wind speed, wind direction and wind probability distributions in Puerto Bolivar - Colombia

Guillermo Valencia Ochoa a, *, José Núñez Alvarez b, Marley Vanegas Chamorro a

a Facultad de Ingeniería, Grupo de investigación en gestión eficiente de la energía – kaí. Universidad del Atlántico, Carrera 30 Número 8-49, Puerto Colombia, Área Metropolitana de Barranquilla 080007, Colombia
b Universidad de la Costa CUC, Energy Optimization Research Group GIOPEN, Cl. 58 #66, Barranquilla, Atlántico, Colombia

A R T I C L E   I N F O

Article history:
Received 2 July 2019
Received in revised form 26 October 2019
Accepted 28 October 2019
Available online 4 November 2019

Keywords:
Wind speed
Wind probability distribution
Wind direction

A B S T R A C T

This paper presents wind speed and direction data measured with a weather station located in Puerto Bolivar, department of La Guajira, situated in the extreme north of Colombia, whose geographic coordinates are 12°11′N 71°55′W. A wind speed and direction sensor, a barometric pressure sensor, and a temperature sensor were used to obtain the presented data. These data were taken at the height of 10 m, which is the highest point of the weather station. The data taken by the meteorological station correspond to a period of 20 years (1993–2013), with hourly frequency. For the missing data, a mathematical model to estimate the Julian averages was developed, allowing to calculate the frequency histograms and four types of probability distributions for these data. Also, the representative wind roses were generated, taking into account the averages in each of the 12 months of the year.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Data

The data presented in this paper are meteorological measurements taken at a weather station located in Puerto Bolivar, Colombia. Daily averages of wind speed and direction were obtained. The historical series has some missing data for the measured period. For this reason, in order to complete these data, it was necessary to use an algorithm developed in MATLAB® to calculate the Julian averages for each of the series. This average estimation uses wind speed and wind direction data from the same month, day, and time from other years to calculate the average of these and estimate the missing data.

The data series are presented, complemented by different probability distributions that describe the statistical behavior of the data which parameter are shown in Table 1, together with the superposition of four probability distributions obtained from the monthly wind speed and direction data supplied with this document. These distributions are shown in Fig. 1 and 2. In addition, the monthly wind rose graphs are presented in Fig. 3 and 4 to determine the most likely wind direction in the place studied. The original hourly data with which these roses have been generated are presented in the attached documents. The calculations to generate all these figures and tables were made with the data presented in Appendix A.

2. Experimental design, materials, and methods

2.1. Experiment set up

The weather station was located in Puerto Bolivar, in the department of La Guajira. This department is located in the north extreme of Colombia at coordinates 12°11’N 71°55’W. In this place, was installed a wind speed sensor, wind speed sensor, and pressure sensor, and the parameters of the installed sensors are shown in Table 2.

The wind speed and direction were measured at the height of 10 m, this being the highest point of the weather station. A little further down, the temperature and relative humidity sensors were located. The schematic diagram of the weather station used to measure the data presented in Appendix A, is shown in Fig. 5.
2.2. Method

The data taken by the weather station correspond to 20 years (1993–2013), with hourly frequency. For the missing data, a data mining was performed to calculate the empty spaces and fill them with the averages from the data on the same day and time in different years, which are presented in Appendix A. Thus, from the totality of these data, four types of probability distributions are calculated. Therefore, the wind speed data are presented as a continuous random variable with an associated probability distribution, and the wind roses give the predominant wind direction presented every month. The methodological procedure developed for the data treatment is shown in Fig. 6, which consists of four fundamental stages.

In the first stage, the series of data from the sensors installed at the measurement site was read, followed by the development of software for mathematical data processing, allowing the user to select the variables and desired time interval for analysis. Once the variables have been selected, the statistical treatment of the data is performed to calculate the desired outputs, such as frequency histograms, wind roses, and wind speed probability distributions. Fig. 7 shows the main views of the WindAnalysisUA v1.0 software developed to analyze the wind data.

The four probability distributions used were Normal or Gaussian, Gamma, Weibull and Rayleigh [1], which are presented below along with their mathematical models.

### Table 1
Parameters of the Probability Distributions.

| Month      | Gamma | Gaussian | Rayleigh | Weibull |
|------------|-------|----------|----------|---------|
| January    | Shape 7.75267, Standard dev. 2.01057 | Average 6.42375, Scale 6.63631 | Lower threshold 0.09932, Shape 3.61657 |
| February   | Shape 9.74721, Standard dev. 1.99181 | Average 6.9427, Scale 6.93962 | Lower threshold 0.29785, Shape 3.93514 |
| March      | Shape 10.5084, Standard dev. 1.95163 | Average 6.98236, Scale 7.1543 | Shape 4.0324 |
| April      | Shape 9.1053, Standard dev. 2.01402 | Average 6.84414, Scale 7.04035 | Shape 3.8431 |
| May        | Shape 5.34405, Standard dev. 2.30579 | Average 3.40028, Scale 6.71382 | Shape 3.05405 |
| June       | Shape 7.20571, Standard dev. 2.2335 | Average 7.12602, Scale 7.37403 | Shape 3.6251 |
| July       | Shape 8.55002, Standard dev. 2.14574 | Average 7.35573, Scale 7.57195 | Shape 3.88415 |
| August     | Shape 5.6485, Standard dev. 2.29827 | Average 6.38453, Scale 6.69418 | Shape 3.05505 |
| September  | Shape 3.56666, Standard dev. 5.08969 | Average 5.5498, Scale 5.5498 | Shape 2.31037 |
| October    | Shape 3.5436, Standard dev. 2.32575 | Average 4.53266, Scale 4.9735 | Shape 2.20963 |
| November   | Shape 0.78306, Standard dev. 2.16318 | Average 4.83386, Scale 5.14798 | Shape 2.60483 |
| December   | Shape 6.50515, Standard dev. 1.99666 | Average 5.76972, Scale 6.01059 | Shape 3.11331 |
| Scale      | 1.12755, Standard dev. 1.98683 | Lower threshold 0.09653, Scale 6.41973 |

2.2.1. Gaussian probability distribution

This distribution can be applied in a large number of case studies, and this makes it a distribution of great statistical relevance [2]. Therefore, this distribution adjusts greatly to physical measurements. This distribution is governed by the probability density function for a normally distributed random variable given by Equation (1) [3,4].

\[
f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right]
\]

where \( \sigma \) is the standard deviation, and \(-\infty < x < \infty\), \(-\infty < \mu < \infty\) and \(\sigma > 0\).
Fig. 1. Monthly Probability distributions for Puerto bolivar (January–June).
Fig. 2. Monthly Probability distributions for Puerto bolivar (July–December).
Fig. 3. Puerto bolivar wind rose (January–June).
Fig. 4. Puerto bolivar wind rose (July–December).
Table 2
Sensors technical data.

| Measurement          | Range                   | Precision                                      |
|----------------------|-------------------------|------------------------------------------------|
| Wind speed           | 0–75 m/s                | ±0.1 m/s (0.3–10 m/s); ±1% (10–55 m/s); ±2% (>55 m/s) |
| Wind direction       | 0–360°                  | ±2% FS                                         |
| Barometric pressure  | 600–1100 hPa            | ±1 hPa                                         |
| Temperature          | de 30°C a 60°C          | 30 to +60°C; ±0.3 °C                           |
| Relative humidity    | 0–100 %HR               | ±0.5% RH                                       |

Fig. 5. Weather station schematic diagram.

Fig. 6. Data analysis flowchart.
2.2.2. Gamma probability distribution

This model is commonly used to adjust wind speed distributions. Contrary to the symmetry presented by the Gaussian distribution, the gamma distribution is biased to the right. This function is given by

\[ G(a) = \int_0^\infty x^{a-1}e^{-x}dx \quad \text{for} \quad a > 0 \]

and the probability density function is calculated using Equation (2) [5].

\[
f(x) = \begin{cases} \frac{1}{\beta^a\Gamma(a)}x^{a-1}e^{-x/\beta}, & x > 0 \\ 0 & \end{cases}
\]

where \( \alpha \) is the form parameter, and \( \beta \) the scale parameter. When \( \alpha, \beta > 0 \), the value of \( E(X) = \alpha\beta \), and \( \text{Var}(X) = \alpha\beta^2 \).

2.2.3. Weibull probability distribution

This is another widely applied model for the amplitude distribution of wind speeds over time. This distribution is influenced by the shape parameter (k or \( \alpha \)), which varies between values 1 and 3.6. In addition, it depends on a scale parameter (c, \( \theta \) or \( \beta \)). If a random variable X fits the probability density function expressed by Equation (3), it can be said to have a Weibull distribution [6].

![Fig. 7. Main view of the WindAnalysisUA v1.0 software, (a) Main view, (b) Output selection, (c) Probability distribution analysis, (d) Wind rose analysis.](image)
where $x, \alpha, \theta > 0$.

### 2.2.4. Rayleigh probability distribution

For this probability distribution, there is also a form parameter $\alpha$ and a scale parameter $\theta$, when they take a value of 2 for the form parameter and $\sqrt{2}$ for the scale parameter. The $\sigma$ parameter is obtained from the probability density function of the Rayleigh distribution, whose mathematical expression is given by Equation (4) [7].

$$f(x; \sigma^2) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

where $x > 0$. In the case in which the shape parameter takes a value equal to 3, one arrives at the Gaussian distribution.

**Conflict of Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.104753.

**References**

[1] M.P. Pinto, J.K. Moreno, Y.A. Muñoz, A. Ospino, Technical and economic evaluation of a small-scale wind power system located in Berlin, Colombia, Tecciencia 13 (24) (2018) 63–72.

[2] A.-L.J. Luis, An approximation to the probability normal distribution and its inverse, Ing. Invest. Tecnol. 16 (4) (Oct. 2015) 605–611.

[3] M. Ordaz, A simple approximation to the Gaussian distribution, Struct. Saf. 9 (4) (Jun. 1991) 315–318.

[4] K. Krishnamoorthy, Handbook of the Normal Distribution Distributions with Applications, University of Louisiana Lafayette, 2010.

[5] P.-A. Amaya-Martínez, A.-J. Saavedra-Montes, E.-I. Arango-Zuluaga, A statistical analysis of wind speed distribution models in the Aburrá Valley, Colombia, CT&F - Ciencia, Tecnol. y Futur. 5 (5) (2018) 121–136.

[6] J.A. Carta, P. Ramírez, Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions, Renew. Energy 32 (3) (Mar. 2007) 518–531.

[7] H. Bidaoui, I. El Abbassi, A. El Bouardi, A. Darcherif, Wind speed data analysis using Weibull and Rayleigh distribution functions, case study: five cities northern Morocco, Procedia Manuf. 32 (Jan. 2019) 786–793.