

# A Wind Turbine Blade Damage Detection System Based on Data Analysis: An Academic Example on the Use of Histograms

Leonardo Acho<sup>1</sup> and Pablo Buenestado<sup>1</sup>

<sup>1</sup> Department of Mathematics  
E.E.B.E., Polytechnic University of Catalonia  
BarcelonaTech, C. d'Eduard Maristany, 19, 08939 Barcelona (Spain)  
Corresponding author e-mail: [leonardo.acho@upc.edu](mailto:leonardo.acho@upc.edu)

**Abstract.** The main objective of this paper is to introduce an academic example on the use of histograms to develop a wind turbine blade fault (or damage) detection system. This approach merely uses data analysis obtained from a generated wind turbine power. Then, and based on the power curve of a wind turbine, histograms are obtained, first, for the healthy wind turbine operation, and then, for two cases on faulty wind turbine operations. By comparing them, a faulty behaviour of a wind turbine is detectable. Finally, according to our numerical example, the proposed technique seems a realistic case to students on the statistic and probability fields as a useful academic example.

## Key words

Blade Wind Turbine, Data Analysis, Histograms, Fault Detection System.

## 1. Introduction

Lately, the technological significance of electrical power production by employing wind turbines has increased and will keep doing it over the next decades [1]-[4], [8]. Moreover, the total installed capacity of wind turbines has also notably growth, for instance, in China, America, and Europe [4], [5]. Hence, to make wind turbines authentic, one key point is to reduce the maintenance of these turbines. Therefore, the development of new fault detection systems is a today's important challenger on this objective.

On the other hand, the pitch blade system is an important mechanical part of the wind turbine to warranty main stability of the overall closed-loop system, load reduction, vibrational mechanical reduction, and good quality of the producer electrical power, etc. [3], [4]. Finally, data analysis seems to be an efficient technique to study mechanical systems behaviour without requiring strong mathematical modelling. Even more, in some cases, this mathematical system model is not required at all, just the input and output data of the corresponding system [5], [7]. Hence, the main objective of this paper is to present a wind

turbine blade fault (or *damage*) detection system by just using statistical data analysis via histograms.

Additionally, statistical concepts and methods are useful to provide mechanisms to understand a system phenomenon. This is the main philosophy given in statistical and probability academic lectures [9]. Moreover, these academic disciplines teach us how to make intelligent judgments in the presence of uncertainty and variation by just employing data generated from the system under study [10]-[14]. Furthermore, motivating future engineering students is a mayor challenge in introductory statistics lectures. Therefore, new academic examples seem motivating, and, especially, when these examples are linked to a state-of-art technology as the wind turbines are.

The rest of the paper is organized as follows. Section 2 describes the probabilistic wind speed model based on the Weibull distribution. A summary of the power curve of a wind turbine is stated in Section 3. Section 4 gives our main contribution. Finally, the Conclusions are given in Section 5.

## 2. Wind modelling

In the corresponding literature, it is well known that the *Weibull* distribution is the most used probabilistic model to represent the wind speed random behaviour to wind turbines [1],[6],[8]. This distribution is given by:

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}, & x \geq 0, \\ 0, & x < 0 \end{cases} \quad (1)$$

where  $\alpha$  and  $\beta$  are the probabilistic model parameters to be estimated according to the regional environment where the wind turbine is located. For instance, the estimated values  $\alpha = 2.05$  and  $\beta = 9.16$  are obtained from a real wind speed data [5]. Here after, we are going to use these values. Figure 1 shows the corresponding Weibull

distribution. On the other hand, Figures 2 and 3 show a discrete-time sample of the corresponding Weibull type wind speed and its related histogram, respectively.

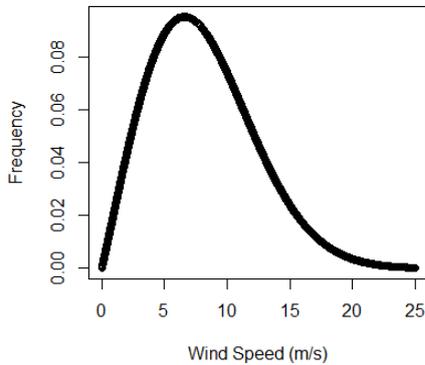


Fig. 1 Weibull distribution of a wind speed random variable.

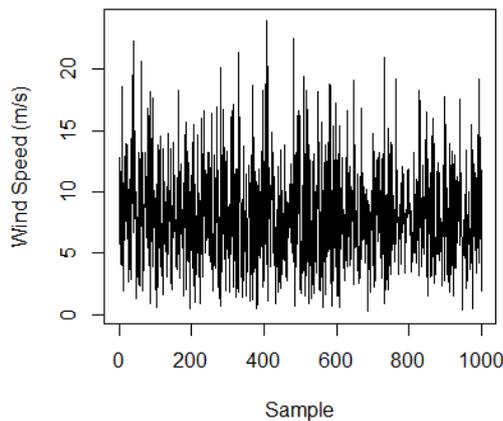


Fig. 2 Sample of the random wind data generated by using the given Weibull probabilistic model.

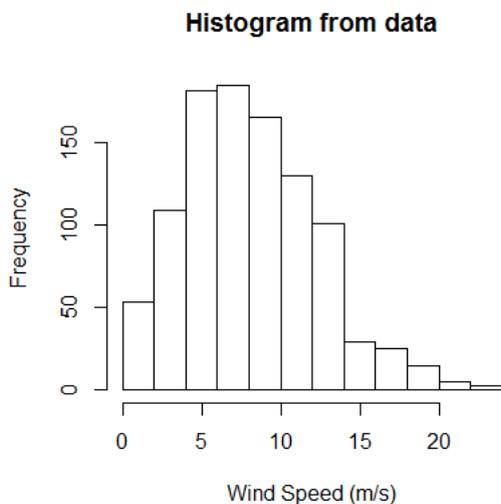


Fig. 3 The histogram obtained from the data shown in Figure 2.

### 3. The Wind turbine power curve

A wind turbine normally works in different operating modes along the wind speed range [3]. These operating modes are illustrated in the power wind speed curve as shown in Fig. 4. This region of operations is usually divided into four zones [3]. Once the uncertain nature of the wind speed is characterized by a random variable, a generated wind turbine power probabilistic model can also be characterized through an appropriate transformation [5], [8]. Finally, during the wind turbine operation, blades encounter complex loading with high number of cycles, such as aerodynamic loads, changing gravity moments, accidental impact, braking force, wind gust, etc., that may damage them [3].

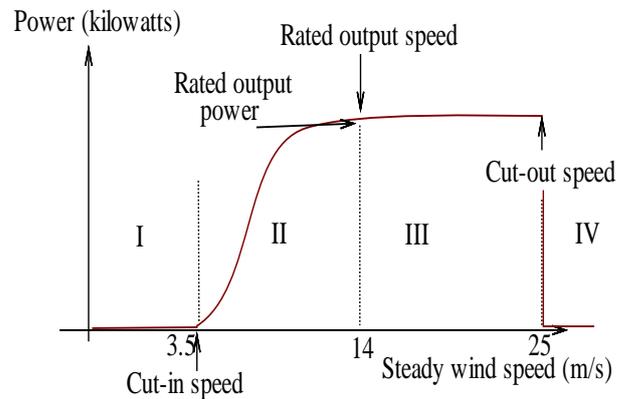


Fig. 4 Typical wind turbine power curve.

From the above power output characteristics of a wind turbine (see Fig. 4), it can be assumed that electrical power starts generating at the cut-in wind speed. Then, the power cubically increases to the rated output power. On the other hand, if the wind speed reaches the Cut-out speed value, for safety reasons, the turbine is shunted-down [3].

In our numerical example, we use the wind turbine parameters given in [5], and described as follows. The output power can be represented as:

$$P_a(u, \sigma) = \begin{cases} 0, & u < u_c, u > u_s \\ cu^3 + N(0, \sigma^2), & u_c \leq u \leq u_r \\ P_r, & u_r \leq u \leq u_s \end{cases} \quad (2)$$

$$c = \frac{1}{2} C_p \rho \pi R^2, \quad (3)$$

where  $u$  represents the wind speed,  $u_c$ ,  $u_r$ , and  $u_s$  are the Cut-in, rated, and Cut-out wind speeds, respectively. In our scenario, we use following data [5]. The power curve coefficient  $c = 0.03906 \text{ kg/m}$ ,  $P_r = 20 \text{ kW}$ ,  $u_c = 2 \text{ m/s}$ ,  $u_r = 8 \text{ m/s}$ ,  $u_s = 18 \text{ m/s}$  and  $\sigma$  is the standard deviation for the added Gaussian noise set to  $0.5 \text{ kW}$ . Moreover,  $\rho$  is the air density,  $R$  is the radius of the rotor determining its swept area, and  $C_p$  is a factor that accounts for the conversion rate of the total wind power to electrical power. Furthermore, see Fig. 6 for a picture representation of this

type of wind turbine and one of its benefit applications in today technology. That is, the use of wind turbines to battery charging of electric vehicles. This is a technological solution, for instance, in remote regions where the access to power lines is missing, such as, for instance, in economical emergent countries. Predominantly, however, 20kW wind turbines units are often used in agricultural and light industrial applications. It has also been used for years in remote village coverings, where diesel power systems supply electricity for a small grid is poor. Finally, because of we are using a ‘small’ wind turbine; the pitch angle is fixed structurally.

#### 4. A Blade fault detection system

Because of the power curve coefficient depends on the blade aerodynamic structure, any structure fault (or damage) in it will affect it. See Fig. 5. We are going to concentrate on the minor faults when the operation of the wind turbine is not affected but the damage may be move to a worse stage if it is not detected opportunely. Moreover, it is a realistic situation to assume that the wind speed is *not exactly* measurable.



Fig. 5 Some blade damage examples. Some of them are minors and others are severs.

Our fault detection design starts with the data acquisition of the generated wind turbine power in its healthy condition. Then, a probabilistic model is obtained by using a histogram. For instance, in our scenario, under the Weibull wind speed behaviour, this histogram is shown in Fig. 7. This stage may be considered as the *training* phase. We call it the reference healthy power  $P_a(U)$ , where  $U$  is the *random variable* of the wind speed. Then, in a new wind turbine operation, we obtain new generated wind power and here represented as  $\hat{P}_a(U)$ . For instance, in healthy condition, the difference histogram of the random process  $\hat{P}_a(U) - P_a(U) = \Delta(U)$  is shown in Fig. 8. From this figure, we can appreciate a symmetric description of the obtained distribution data. On the other hand, Figures 9 and 10 show, from others new run data of the corresponding wind turbine but on the non-healthy cases, the related histograms. From

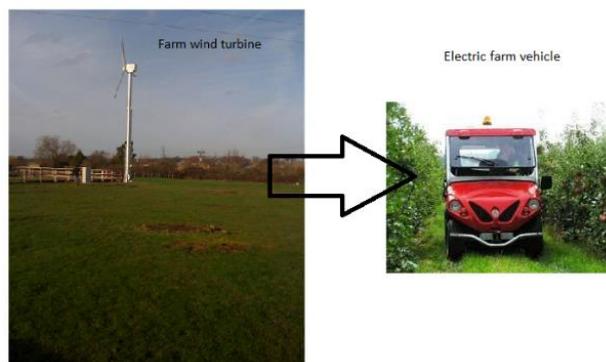


Fig. 6 A 0.5kW wind turbine locate at a farm to possibly charging the battery of an electric farm vehicle.

these figures, we may note a drastic change on the obtained histogram for  $\Delta(U)$ . In resume, we have the following algorithm on wind turbine blade fault detection system based on data analysis:

- 1) Obtain the healthy generated wind turbine power histogram based on data training.
- 2) Obtain a new generated wind turbine power data and create its corresponding histogram.
- 3) Find the difference  $\Delta(U)$  histogram.
- 4) Decide if there exists a fault by studying the histogram of  $\Delta(U)$ .

Additionally, there have been many techniques related to fault (or damage) detection of wind turbine blades [3]. The statistical based on methods usually employ the Principal Component Analysis (*PCA*). However, our approach seems simpler. Additionally, Fig. 11 gives the power curve of our turbine obtained by using our numerical simulated data. Finally, Fig. 12 shows the corresponding diagram box (or *boxplots*) of each case illustrated in Figures 8, 9, and 10, respectively. Clearly, from this figure, the boxplot is sensitive to the healthy and non-healthy wind turbine condition.

Histogram: Healthy Case

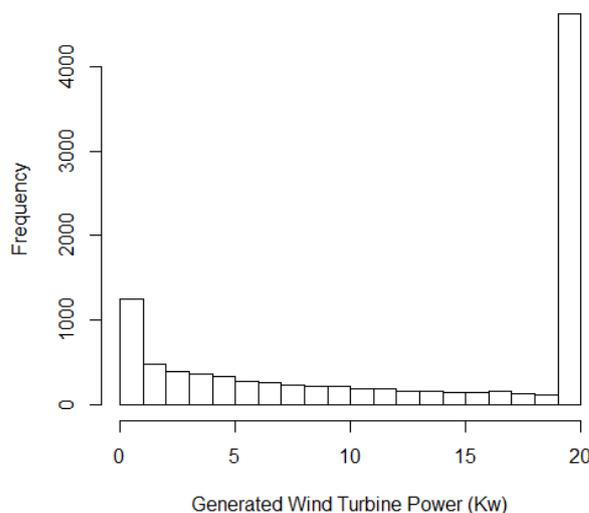


Fig. 7 Histogram of the generated wind power in its healthy condition.

**Difference Histogram when c=0.03906**

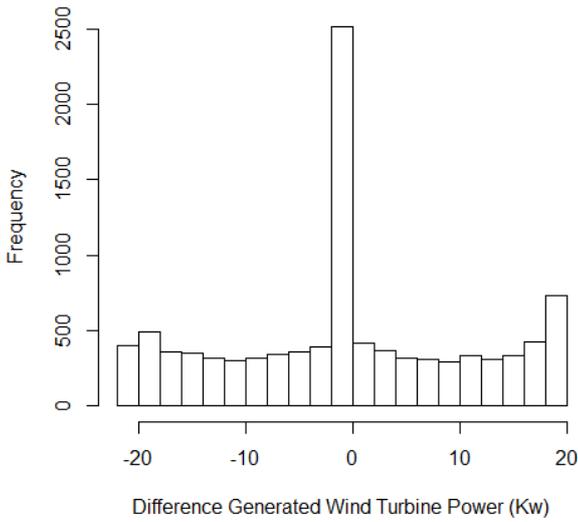


Fig. 8 Difference histogram  $\Delta(U)$  on healthy condition of the wind turbine.

**Difference Histogram when c=0.02**

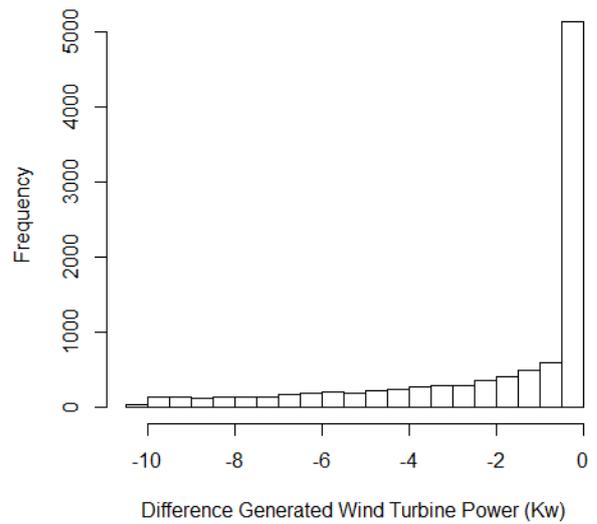


Fig.10 Difference histogram  $\Delta(U)$  on another non-healthy condition of the wind turbine.

**Difference Histogram when c=0.05**

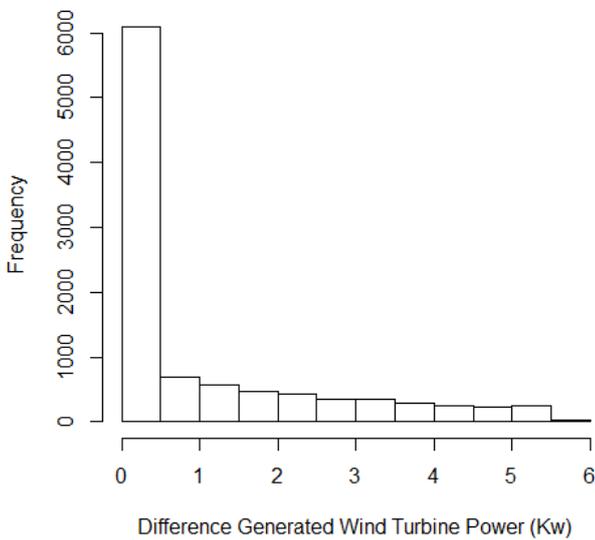


Fig. 9 Difference histogram  $\Delta(U)$  on non-healthy condition of the wind turbine.

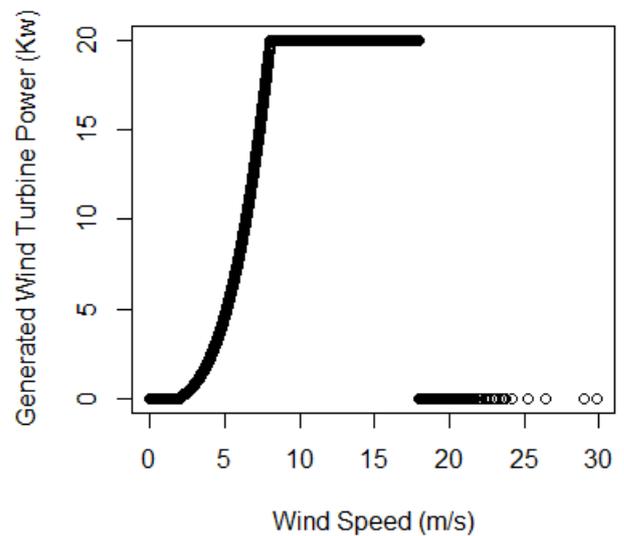


Fig. 11 The wind turbine power curve obtained from the numerical simulated healthy data.

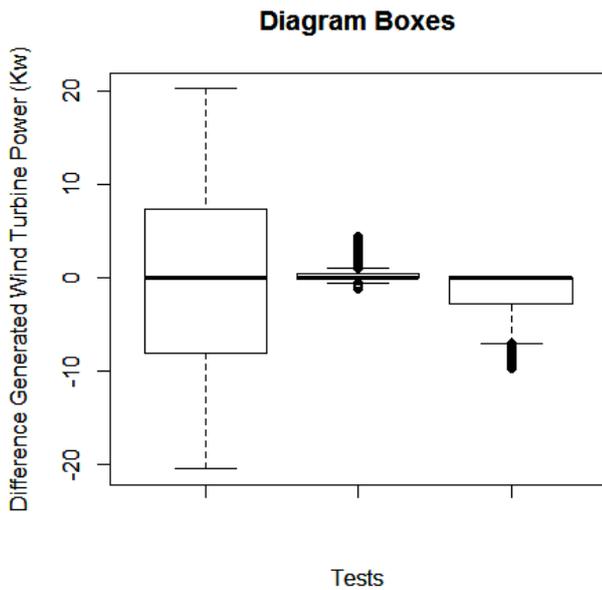


Fig. 12 Diagram boxe for each analysed case (a comparative illustration).

## 5. Conclusion

A wind turbine blade fault detection system based on data analysis has been proposed. This analysis can be realized even when the wind turbine is on real-time operation. Specially, our propose assumes that any minor blade damage affects the power curve coefficient. Obviously, there may exist other kind of faults that can affect, precisely, this coefficient. Moreover, even when we presented simulation data by using the Weibull probabilistic model distribution for the wind speed random variable, our overall scheme should work for any other type of wind speed probabilistic distribution because of the initial training stage phase.

On the other hand, and from the academic point of view, we have built up an interesting example on the use of histograms and probabilistic models applied to a today's technology on renewable energies. In contrast, in real applications, we are interested in detecting the fault within the shortest possible time period. From our data, it is possible to estimate a time fault diagnosis in about 100 seconds if the sampling rate is 0.1 s (From Fig.2). Obviously, this topic can inspire a future work on it.

Finally, the numerical experiments were realized by using *R*, an *open-access* high-level language and an environment for data analysis and graphics [15].

## Acknowledgement

This work was partially supported by the Spanish Ministry of Economy and Competitiveness under Grant DPI2015-64170-R.

## References

- [1] Wang, Jianzhou, Jianming Hu, and Kailiang Ma. "Wind speed probability distribution estimation and wind energy assessment." *Renewable and Sustainable Energy Reviews* 60 (2016): 881-899.
- [2] Arslan, Talha, Sukru Acitas, and Birdal Senoglu. "Generalized Lindley and Power Lindley distributions for modeling the wind speed data." *Energy Conversion and Management* 152 (2017): 300-311.
- [3] Luo, Ningsu; Vidal, Yolanda; Acho, Leonardo (ed.). *Wind turbine control and monitoring*. Berlin: Springer, 2014.
- [4] Vidal, Yolanda, et al. "Fault diagnosis and fault-tolerant control of wind turbines via a discrete time controller with a disturbance compensator." *Energies* 8.5 (2015): 4300-4316.
- [5] Janssens, Olivier, et al. "Data-driven multivariate power curve modeling of offshore wind turbines." *Engineering Applications of Artificial Intelligence* 55 (2016): 331-338.
- [6] Jung, Christopher, and Dirk Schindler. "Global comparison of the goodness-of-fit of wind speed distributions." *Energy Conversion and Management* 133 (2017): 216-234.
- [7] Ouyang, Tinghui, Xiaoming Zha, and Liang Qin. "A combined multivariate model for wind power prediction." *Energy Conversion and Management* 144 (2017): 361-373.
- [8] Marčiukaitis, Mantas, et al. "Non-linear regression model for wind turbine power curve." *Renewable Energy* (2017).
- [9] Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. Cengage learning, 2011.
- [10] Leon-Garcia, Alberto. *Probability, statistics, and random processes for electrical engineering*. USA, 2017.
- [11] Zaher, A. S. A. E., et al. "Online wind turbine fault detection through automated SCADA data analysis." *Wind Energy* 12.6 (2009): 574-593.
- [12] EFRON, Bradley; TIBSHIRANI, Robert. Statistical data analysis in the computer age. *Science*, 1991, p. 390-395.
- [13] KARIP, Emin. Multilevel Analysis of Hierarchically Structured Educational Data. *Kuram ve Uygulamada Egitim Yönetimi Dergisi*, 1996, vol. 2, no 3, p. 409-420.
- [14] ANOVA, Split-plot. *Statistical computing: an introduction to data analysis using S-Plus*. 2002.
- [15] Crawley, Michael J. *The R book*. John Wiley & Sons, 2012.