Optimal time step of SCADA data for the power curve of wind turbine

Minh-Thang DO, Julien BERTHAUT-GERENTES
METEODYN, 33 Bd Salvador Allende - 44800 Saint-Herblain – FRANCE
info@meteodyn.com

Abstract: The SCADA system for wind turbine often logs data at 10 minutes time step, mostly due to the technical constraints. However, with the development in measurement and storage technology, the collection of data with higher sampling frequency is feasible. The question about an optimal time step for SCADA data is then become interesting. Our paper aims to answer this question through a statistical analysis of the power curve at different time step. In order to explain and confirm the results of this study, a spectral analysis is performed. Both approaches indicate that the time step of 10 minutes is still the best choice as it gives the best predictability of the output power from wind speed.

1. Introduction

Nowadays, in most SCADA system of wind turbine, signals are first measured at a sampling rate of 1 second; then averaged at every 10 minutes; and finally written to the database. The aggregation at 10 minutes became an industry standard [1], [2]. It provides a good balance between sampling frequency and database size. Additionally, this 10-minutes time step allows comparing and correlating with pre-construction wind data, using in the wind resource assessments [3].

With the development in storage technology, the collection of data with higher sampling frequency is feasible. Several researches are trying to analyze these high frequency data [4], [5]. However, is the time step of 10 minutes still the best choice in terms of describing the relationship between wind speed and output power of a wind turbine? This paper aims to provide a more detailed analysis of this relationship, which is represented through the power curve.

2. Approach

In this study, the data is collected by a SCADA system with a time step of 7 seconds. An algorithm of classification is used to remove anomalies data, such as: measurement failure, stop, curtailment, under and over production. Only normal data are considered.

Figure 1 shows the relationship between the wind speed and the output power of the wind turbine (green dots), each data point representing a 7s time step. This relationship could be quantified by the dispersion of the data points. Less dispersion means the relationship is better and it is easier to estimate the output power from the wind speed. High dispersion means that the estimation of output power from the wind speed is less precise. In the ideal case, each value of wind speed corresponds to one and only one value of output power, this form the power curve of the wind turbine (blue curve). The dispersion of the data points in this case is zero and we obtain a perfect relationship.
In this study, the fluctuation is quantified by the Root Mean Square Error between the measured value of output power and the estimated value obtained from the power curve and the measured wind speed. Note that the power curve is an estimation based on the SCADA data, through an iterative median estimation technique (the manufacture power curve is not used).

In order to explain and confirm the results of this study, a spectral analysis is performed on both signals. This approach is able to distinguish the behavior at different frequencies or time scales. It is well suited for linear phenomena, which is not the case here (non-linearity of the power-curve). By the way, the monotonic shape of the relationship (the power curve is an increasing function) allows using this analysis with caution.

Figure 1. Scada data of wind speed and active power with a time step of 7s

3. Results and analysis
In order to observe the time-dependency of the power curve, both signals (wind speed and output power) are time-averaged under different time steps. Figure 2 shows the relationship between the wind speed and the output power at the following time steps: 7 seconds, 30 seconds, 1 minute, 5 minutes, 10 minutes, 30 minutes, 60 minutes and 120 minutes. It could be observed that the data points are less dispersed at higher time step.

| Wind turbine ref | 7s  | 30s | 1min | 5min | 10min | 30min | 60min | 120min |
|-----------------|-----|-----|------|------|-------|-------|-------|--------|
| F01             | 163 | 121 | 106  | 90   | 72    | 120   | 114   | 126    |
| F02             | 165 | 122 | 118  | 105  | 109   | 117   | 122   | 121    |
| F07             | 195 | 123 | 100  | 67   | 60    | 56    | 59    | 85     |
| F12             | 181 | 131 | 120  | 105  | 99    | 111   | 96    | 113    |
| F13             | 207 | 139 | 125  | 106  | 98    | 110   | 120   | 113    |
| F19             | 190 | 145 | 129  | 114  | 96    | 130   | 134   | 150    |
| F24             | 172 | 130 | 107  | 81   | 88    | 82    | 97    | 126    |
| F25             | 159 | 106 | 90   | 64   | 60    | 57    | 63    | 74     |
| AVERAGE         | 179 | 127 | 112  | 92   | 85    | 98    | 101   | 113    |

Table 1. RMSE of active power [kW]
The RMSE between the measured output power and the estimated value obtained from the power curve is calculated for each time step (first line of Table 1). It can be seen that 7s gives the worst RMSE, while 10min seems to be the best. This process is repeated over 8 machines of a wind farm. Results are presented graphically in Figure 3.

The mean RMSE is very high at small time steps (179kW at time step of 7 seconds), but it decreases rapidly. The best mean RMSE is obtained at the time step of 10 minutes (85kW, about a half of the maximum value). Then, it slowly increases at higher time steps.

Through this analysis, it’s justified that the time step of 10 minutes is the best choice in term of describing the relationship between the wind speed and the output power of wind turbine.

In the next step, we will apply the signal processing techniques with the data to understand this phenomenon [6]. The Power Spectrum Density (PSD) is computed on both signals, with a small window and large overlap to reinforce the smoothing.
Figure 4 presents the ratio between the PSD, in a linear-linear scale. It can be seen that above 0.03Hz (corresponding to phenomena repeating every half a min), this ratio is roughly constant. We suppose that this is only noise: at these frequencies, the variation of both signals is very small, and data are representing only measurement noise.

Figure 5 plots the same quantity in a log-log scale; it is similar to a Bode magnitude plot. As a first estimation, this corresponds to 1st order low pass filter tuned at 0.005 Hz (~3min). “first order” signifies that if the frequency is multiplied by $x$, so the response is divided by the same quantity $x$. In mechanics, this is typical for Mass-Damper system with no Stiffness. 3 minutes is the time-response of the whole machine inertia, including hub, high speed train and electro mechanical parts. This low-pass filter models the fact that the hub rotational speed is not able to follow the rapid fluctuations of wind speed, because of its inertia. In other words, the rapid fluctuations of wind speed (those one during less than some minutes), due to the turbulence, has no direct impact on the output power.

A more detailed exam exhibits an absolute maximum about 0.65 mHz, i.e. 25 min. This signifies that the power conversion is less effective for very slow variations (~2H) than for faster fluctuations (half an hour). This might be linked with the non-linearity of the power curve.

![Graph showing power spectrum density between wind speed and active power in log-log scale.](image)

**Figure 5.** The Power Spectrum Density between wind speed and active power in log-log scale

### 4. Conclusions

This study focuses on the relationship between wind-speed and output power in wind energy, depending on the time scales.

The Bode magnitude plot shows different regions:

- for high frequencies (faster than 30 seconds), noise is dominating the spectra
- for intermediate frequencies, the generator behaves as a 0.005Hz 1st order low-pass filter:
  - response is roughly flat for fluctuations slower than ~3min
  - higher frequencies are filter (the higher the more filtered)
- for very slow fluctuation (several hours), the non-linearity of the Power curve reduced the global power generation apparent efficiency.
As a result, the 10-minute averaging standard is a good choice. At smaller time-step, the generator is not able to follow the wind speed fluctuations. At larger time step, the wind is never constant over the 1-step time span, so the power curve curvature is reduced.

References

[1]. Wang Ke-Sheng, Sharma Vishal S. and Zhang Zhen-You, “SCADA data based condition monitoring of wind turbines”, Advances in Manufacturing, vol. 2, no. 1, p. 61-69, Mar 2014.

[2]. IEC 61400-12-1, "Wind turbines – Part 12-1: power performance measurements of electricity producing wind turbines," IEC – International Electrotechnical Commission, Geneva, Switzerland, 2017.

[3]. A.J. Brand, “A downscaling method for application in wind resource assessment and wind power forecasting”, 13th International Conference on Wind Engineering, Amsterdam, The Netherlands, 11-15 July 2011

[4]. Friis Pedersen T, Wagner R, Demurtas G. “Wind Turbine Performance Measurements by Means of Dynamic Data Analysis”, DTU Wind Energy, 2016. p. 91 (DTU Wind Energy E, Vol. 0082).

[5]. Tanja A. Mücke, Matthias Wächter, Patrick Milan, Joachim Peinke, “Langevin power curve analysis for numerical WEC models with new insights on high frequency power performance”, Wind Energy 18:11, pp. 1953-1971 (2015)

[6]. Hernandez, W.; Maldonado-Correa, J.L.; Mendez, A. Frequency-domain Analysis of Performance of a Wind Turbine. Electron. Lett. 2016, 52, 221–223.