# SCADA and Artificial Neural Networks for Maintenance Management

Alberto Pliego Marugán(✉) and Fausto Pedro García Márquez

Ingenium Research Group, Universidad Castilla-La Mancha, 13071 Ciudad Real, Spain
alberto.pliego@uclm.es

**Abstract.** Nowadays, the reliability of the wind turbines is essential to ensure the efficiency and the benefits of the wind energy. The SCADA system installed in a wind turbine generates lot of data that need to be processed. The information obtained from these data can be used for improving the operation and management, obtaining more reliable systems. The SCADA systems operate through different control rules that are predefined. However, a static control of the wind turbine can generate a miscorrelation between the control and the real conditions of the wind turbine. For example, two wind turbines can be separated several kilometers in the same wind farm, therefore, the operation conditions must be different and the control strategy should not be unique. This research work presents a method based on neural networks for a dynamic generation of the control strategy. The method suggests that the thresholds used for generating alarms can vary and, therefore, the control of the wind turbine will be adapted to each specific wind turbine.

**Keywords:** Wind turbine · Reliability · SCADA systems · Advanced control analytics · Artificial neural networks

## 1 Introduction

The wind energy is currently the most important renewable energy, the capacity installed currently is more than 420 GW and it is estimated to be more than 1000GW in 2030 [3]. The maintenance and operation costs of conventional wind turbines are 12% of the total costs, but the wind energy is evolving towards the offshore location. It causes an important increase of these costs, being for offshore wind farms around 23% of the total costs [12].

SCADA systems are widely introduced in wind turbines (WTs) due to their effectiveness has been proved in other industries for detection and diagnostics of failures [9,11,16]. They are presented as an inexpensive and optimal solution [20] to control feedback for the health monitoring while reducing the operation and management costs [19]. Nevertheless, they also present some minor disadvantages due to the operational or reliability conditions [14,21]. These systems consider a large amount of measurements such as temperatures or wind and energy conversion parameters [18]. Data have raised considerable interest in different areas, e.g. wind power forecasting [17], production assessment [22] or fault detection [4,6,8,10].


J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_75

In the case of WTs, the introduction of SCADA systems verifies the efficiency when their components are deteriorated. This degradation can indicate problems of different nature such as misalignments in the drive-train, friction caused by bearing or gear faults. The basic elements of the performance monitoring consist on a first collection of raw values by the sensors. After the application of the appropriate filters, anomalies are detected. Finally, a diagnosis will be provided. The anomaly detection includes a series of techniques that range from simple threshold checks to statistical analyses [5,7,15].

It has been demonstrated that the WT suffers a gradual loss of production. Figure 1 shows that the power has a decrease year by year (this study has been developed by Sheffield university in the OPTIMUS project [2]). The SCADA data has been considered over 5 years and the wind speed-power curve has been estimated.

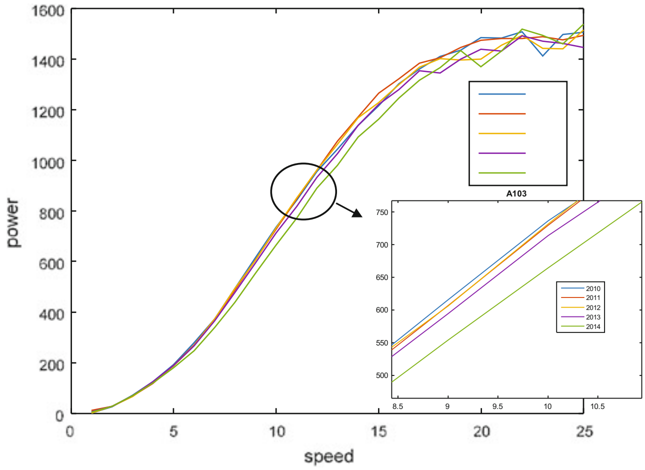


**Fig. 1.** Loss of power of a WT over the time

This paper present a signal processing considers the minor changes in the behavior of the WT. Some alarms will be affected by the decrease of the power shown in Fig. 1. The following alarms will be activated considering the power as a cause.

- Activation of the ice safe mode: One of the statements considered for this alarm is that the power is low for the measured wind.
- High aerodynamic deterioration: It is considered for low power with high wind.

If the power presented a minor reduction, then some false alarms would be activated because the SCADA was prepared for a higher power. However, if the control laws are dynamic, this problem will be solved.

## 2 Proposed Method

The method proposed aims to improve the control of the system by following a supervised iterative process. The objective is to determine if the SCADA control is coherent

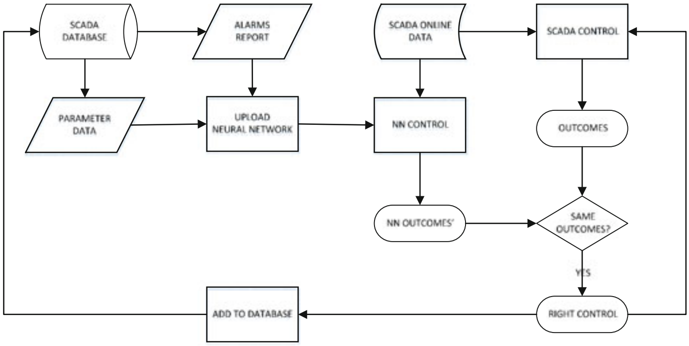


**Fig. 2.** Flowchart of the method proposed

with the historical data and, therefore, if the system has changed. The main capability of the method is to find differences in the behavior of the system over the time. Figure 2 shows a flowchart of the method proposed.

The iterative process starts with a database of the historical SCADA data. In this paper, two different types of data are considered: the value of the measured parameters and an alarm report. This data is employed to build an artificial neural network (NN) with a supervised training. The values of the parameters are employed as the inputs and the alarms will represent the outputs of the NN.

NN are used in problems that cannot be formulated as an exact method or an analytical solution. NN learns by itself and provides a good solution for the problem simulating the biological neurons in a reasonable time. An artificial NN consists of neurons that are simple processing units and weighted connections between those neurons [1].

It can be used for improving the control of the WT when the NN that defines the logic of the system has been generated. The SCADA system will provide online data of the condition of the WT. These data will be processed following two different evaluations: the SCADA system processing and the NN processing. Both outcomes will be compared. If the outcomes are equal, then the data will be added to the database, there for the NN processing will became an adaptive process.

## 3 Case Study

The European Optimus project [2] has provided the SCADA data used in this work. It measures a lot of parameters, but after a filtering process, only 34 parameters (see Table 1) have been considered for the analysis presented hereby. The values of this parameters will be considered as inputs for the NN.

Basically, the NN receives a dataset and make a training process to recognize several patterns. The training process fits the different weights to provide the output. If the output is known, then the training is defined as supervised, otherwise, it is called unsupervised training. The condition of the WT corresponds to the desired outputs of the NN. The data used to design the NN is divided in following groups:

**Table 1.** Inputs and outputs for neural network

| No | Signal | No | Signal |
|---|---|---|---|
| 1 | General accumulator blade 1 pressure | 18 | Environmental temperature |
| 2 | General accumulator blade 2 pressure | 19 | Drive end side generator bearing temperature |
| 3 | General accumulator blade 3 pressure | 20 | Non-drive end side generator bearing temperature |
| 4 | Phi cosine | 21 | Generator winding temperature |
| 5 | Turbulence level | 22 | Nacelle temperature |
| 6 | Oscillation level | 23 | Lower gearbox radiator |
| 7 | Vibration level | 24 | Upper gearbox radiator |
| 8 | Pitch 1 angle | 25 | Gearbox bearing temperature |
| 9 | Pitch 2 angle | 26 | Transformer 1 temperature |
| 10 | Pitch 3 angle | 27 | Transformer 2 temperature |
| 11 | Active power | 28 | Transformer 3 temperature |
| 12 | General accumulator pressure | 29 | Grid voltage |
| 13 | Brake pressure | 30 | Total reactive power |
| 14 | Hydraulic group pressure | 31 | Generator speed |
| 15 | SP pitch angle | 32 | Rotor speed |
| 16 | Hydraulic group oil temperature | 33 | Wind speed |
| 17 | Gearbox oil temperature | 34 | Yaw |

- Training set: Around 75% of the total amount of data.
- Validation set: Around 15% of the total amount of data.
- Testing set: Around 15% of the total amount of data.

The NN is expected to be able to learn from data and predict the output when a generic input is considered.

The output of the NN will be specified by a Report of Alarms generated by the SCADA [2]. The data available consists of a serial of alarms registered from 01/10/2012 to 15/05/2015, where there are different types of alarms identified by a specific code. The codes are not indicated for reasons of confidentiality. In this work, only 5 different scenarios have been considered to simplify the example. These scenarios correspond to 4 different alarms and the absence of alarms.

Therefore, the NN will be created following the structure shown in Table 2.

The design of the NN has been carried out using the geometrical pyramid rule [13]. Figure 3 shows the structure of the NN built. This rule express an approximation of the number of neurons h at the hidden layer. It is defined by:

$$h = \sqrt{m \times n},$$

where $m$ is the number of elements of each input and $n$ is the number of possible outputs, being $m = 34$, $n = 5$, and therefore, $h = 13.03$.

Figure 4 shows the statistics of the NN generated by using the SCADA data. This is a global confusion matrix and corresponds to the sum of the training, validation and testing sets. The rows of the matrix show the outcomes of the NN for the dataset used.

**Table 2.** Inputs and outputs for NN

| | Input data | | | Output data | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | $\cdots$ | 34 | Alarm 1 | Alarm 2 | Alarm 3 | Alarm 4 | None Alarm |
| Date 1 | $a_1^1$ | $a_1^{\cdots}$ | $a_{1j}^{34}$ | 1 | 0 | 0 | 0 | 0 |
| Date 2 | $a_2^1$ | $a_2^{\cdots}$ | $a_{2j}^{34}$ | 0 | 1 | 0 | 0 | 0 |
| Date 3 | $a_3^1$ | $a_3^{\cdots}$ | $a_{3j}^{34}$ | 0 | 0 | 1 | 0 | 0 |
| Date 4 | $a_4^1$ | $a_4^{\cdots}$ | $a_{4j}^{34}$ | 0 | 0 | 0 | 1 | 0 |
| Date 5 | $a_5^1$ | $a_5^{\cdots}$ | $a_5^{34}$ | 0 | 0 | 0 | 0 | 0 |

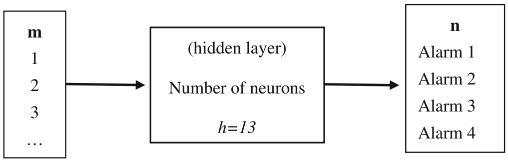


**Fig. 3.** Structure of the neural network

| Output \ Target | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 28<br>28.6% | 0<br>0.0% | 0<br>0.0% | 0<br>0.0% | 1<br>1.0% | 96.6%<br>3.4% |
| 2 | 0<br>0.0% | 9<br>9.2% | 1<br>1.0% | 0<br>0.0% | 1<br>1.0% | 81.8%<br>18.2% |
| 3 | 0<br>0.0% | 0<br>0.0% | 9<br>9.2% | 0<br>0.0% | 1<br>1.0% | 90%<br>10% |
| 4 | 0<br>0.0% | 0<br>0.0% | 0<br>0.0% | 9<br>9.2% | 0<br>0.0% | 100%<br>0.0% |
| 5 | 0<br>0.0% | 2<br>2.0% | 1<br>1.0% | 0<br>0.0% | 36<br>36.7% | 92.3%<br>7.7% |
| 6 | 100%<br>0.0% | 81.8%<br>18.2% | 81.8%<br>18.2% | 100%<br>0.0% | 92.3%<br>7.7% | 92.9<br>7.1% |

**Fig. 4.** Results of the neural network: confusion matrix

The columns show the real output, i.e. the output established by the alarms. The diagonal of this matrix (grey cells) contains the desired solutions. The sixth row and the sixth column show a summary of the rights (green percentage) and wrongs (red percentage).

The outcomes of the NN agree more than 90% with the alarms that the SCADA system generates. The method proposed establishes that the control of the SCADA needs to be uploaded when the alarms cannot be predicted by the NN with enough accuracy. If the SCADA control and this NN have the same response, then the data processed will be added to the database. Therefore, a "healthy" dataset will be created and the control will be adapted to the real conditions over the time.

## 4 Alarm Prediction

The method proposed here by also allows for predicting a possible alarm. In the previous Sect. 3, a NN was generated by inputting the data that the SCADA system employs for generating the alarms. In this section, the input will be the data obtained before the alarm is activated, i.e. the objective is to predict an alarm before the SCADA system generates it. The NN is performed using the same techniques that in previous section. In this case the NN will be designed to distinguish whether alarm will be activated or not. Figure 5 shows the scheme used for the prediction. The inputs employed for performing the NN are the dataset collected before the alarms are activated.

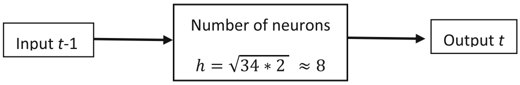


**Fig. 5.** Structure of the second neural network

Following the structure of Fig. 5, a NN has been performed obtaining the results shown in Fig. 6. This confusion matrix shows that the NN can predict a 20.1% of the alarms. Despite it is a low percentage when the NN suggests an alarm, it has a success of 62.1%.

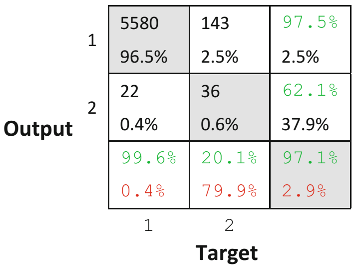


**Fig. 6.** Results of the second neural network: confusion matrix

This method can be used for determining the predictability of some alarms. This can be a useful tool to identify possible alarms before the WT can be damaged.

## 5 Conclusions

In this work, a new methodology is presented for extracting information from SCADA dataset. The methodology is based on the generation of neural network from the quantitative (value of parameters) and qualitative dataset (alarms) of the SCADA system. The methodology has two different purposes. In first place, the adaptation of the SCADA control rules to the variable condition of the wind turbines. A neural network has been generated to determine which data should be added to the database. The creation of a "healthy" database allows for adapting the SCADA control rules to the real condition of the wind turbine over the time. Secondly, an additional neural network has been created for making predictions of the activation of alarms. This can be used to identify an abnormal state of the wind turbine earlier than the SCADA.

**Acknowledgement.** The work reported herewith has been financially by the Spanish Ministerio de Economía y Competitividad, under Research Grant Ref.: DPI2015-67264-P.

## References

1. Arbib MA (2003) The handbook of brain theory and neural networks. MIT Press
2. Camacho E, Requena V et al (2014) Demonstration of methods and tools for the optimisation of operational reliability of large-scale industrial wind turbines. In: International conference on renewable energies offshore renew
3. Council GWE (2015) Global wind report annual market update 2014. Technical report, http://www.gwec.net/wp-content/uploads/2015/03/GWEC_Global_Wind_2014_Report_LR.pdf
4. Gómez Muñoz CQ, Márquez FPG (2016) A new fault location approach for acoustic emission techniques in wind turbines. Energies 9(1):40
5. González-Carrato RRDLH, Dimlaye V, Ruiz-Hernández D (2014) Pattern recognition by wavelet transforms using macro fibre composites transducers. Mech Syst Signal Process 48(1):339–350
6. González-Carrato RRDLH, Márquez FPG, Dimlaye V (2015) Maintenance management of wind turbines structures via mfcs and wavelet transforms. Renew Sustain Energy Rev 48:472–482
7. Kim K (2011) Use of scada data for failure detection in wind turbines. In: ASME 2011 5th International conference on energy sustainability, pp 2071–2079
8. Kusiak A, Li W (2011) The prediction and diagnosis of wind turbine faults. Renew Energy 36(1):16–23
9. Márquez FPG, García-Pardo IP (2010) Principal component analysis applied to filtered signals for maintenance management. Qual Reliab Eng Int 26(6):523–527
10. Márquez FPG, Muñoz JMC (2012) A pattern recognition and data analysis method for maintenance management. Int J Syst Sci 43(6):1014–1028
11. Márquez FPG, Pérez JMP, et al. (2016) Identification of critical components of wind turbines using fta over the time. Renew Energy 87
12. Marugán AP, Márquez FPG, Pérez JMP (2016) Optimal maintenance management of offshore wind farms. Energies 9(1):46
13. Masters T (1993) Practical neural network recipes in C++. Academic Press
14. Muñoz CQG, Márquez FPG, Tomás JMS (2016) Ice detection using thermal infrared radiometry on wind turbine blades. Measurement 93:157–163
15. Munoz CQG, Arenas JRT, Marquez FPG (2014) A novel approach to fault detection and diagnosis on wind turbines. Global Int J 16(6):1029–1037
16. Mylaraswamy D, Olson L, Nwadiogbu E (2007) Engine performance trending. In: AIAC12-HUMS conference
17. Song Z, Jiang Y, Zhang Z (2014) Short-term wind speed forecasting with markov-switching model. Appl Energy 130(5):103–112
18. Sun P, Li J et al (2016) A generalized model for wind turbine anomaly identification based on scada data. Appl Energy 168:550–567

19. Wymore ML, Dam JEV et al (2015) A survey of health monitoring systems for wind turbines. Renew Sustain Energy Rev 52:976–990
20. Yang W, Tavner PJ et al (2010) Cost-effective condition monitoring for wind turbines. IEEE Trans Industr Electron 57(1):263–271
21. Yang W, Court R, Jiang J (2013) Wind turbine condition monitoring by the approach of scada data analysis. Renew Energy 53(9):365–376
22. Zang X, Sun H, Trivedi KS (1998) A bdd-based algorithm for reliability graph analysis. Department of Electrical Engineering