SCADA alarms processing for wind turbine component failure detection

E Gonzalez, M Reder and J J Melero

CIRCE-Universidad de Zaragoza, C/ Mariano Esquillor 15, 50018, Zaragoza, Spain

E-mail: egonzalez@fcirce.es

Abstract. Wind turbine failure and downtime can often compromise the profitability of a wind farm due to their high impact on the operation and maintenance (O&M) costs. Early detection of failures can facilitate the changeover from corrective maintenance towards a predictive approach. This paper presents a cost-effective methodology to combine various alarm analysis techniques, using data from the Supervisory Control and Data Acquisition (SCADA) system, in order to detect component failures. The approach categorises the alarms according to a reviewed taxonomy, turning overwhelming data into valuable information to assess component status. Then, different alarms analysis techniques are applied for two purposes: the evaluation of the SCADA alarm system capability to detect failures, and the investigation of the relation between components faults being followed by failure occurrences in others. Various case studies are presented and discussed. The study highlights the relationship between faulty behaviour in different components and between failures and adverse environmental conditions.

1. Introduction

The rapid growth of wind turbines (WT) in terms of size, number of installations and rated capacity has a huge impact on the operation and maintenance (O&M) costs. Indeed, turbine failure and downtime can often compromise the profitability of a wind farm (WF). This is especially valid for offshore WFs, where unexpected failures result in even more severe consequences, in terms of additional downtime and costs. Early detection of these failures can facilitate the changeover from corrective maintenance towards a predictive approach. Avoiding unexpected failures will enable operators to reach their aim to maximise the annual energy production (AEP) with the highest possible reliability of operation and the lowest O&M costs.

Over the last decade, many important contributions have been made to the study of WT component reliability and failure predictability [1, 2]. Nonetheless, nor data collection on WT failures is standardised, neither a standard taxonomy terminology has been defined yet. Condition monitoring (CM) techniques have proved to be very effective in detecting incipient failures of a particular component, and some have even turned into real industrial applications [3, 4]. Research on Supervisory Control and Data Acquisition (SCADA) signals has also been successful in deriving systems for detecting developing faults [5, 6, 7].

The SCADA system provides a comprehensive overview of the historical and present status of a WT. In addition to the operational and environmental parameters, it includes a detailed record of alarms and fault logs, providing valuable information about the sub-assemblies and components. Moreover, since this information is available for any WT, there is no need for additional investments while CM systems often require installations of specific sensors and equipment. Recent studies have shown their success in predicting pitch failures using a combination of operational SCADA data and alarm logs, by applying diverse methods like Artificial Neural Networks [8], Bayesian Networks [9] or the RIPPER algorithm [10]. Data-mining techniques were also applied to this data combination to predict the turbine status in [2]. Additionally, the same combination has been explored to assess WT reliability [11]. However, very little research has been conducted using only SCADA alarms [12].

In this context two shortcomings can be highlighted. The WT operates as a whole, and a particular component fault may result in a failure of a different component. Consequently, there is a need to improve the knowledge of the WT from a holistic approach in order to understand the relation between different components and their failure mechanisms. The strongest drawback posed by the alarm analysis is the overwhelming amount of data. The huge amount of alarms triggered in a short period of time makes them unmanageable and no relevant component-related information can be extracted. This paper combines previous research findings to provide a cost-effective methodology for processing and analysing SCADA alarms in order to detect any component failure. The main objectives of the present research work are:

- Categorise the WT SCADA alarms according to a reviewed taxonomy; the taxonomy developed for the ReliaWind project [13] has been used as a basis, incorporating certain revisions.
- Analyse the component-related alarms by applying two methods, as suggested in [12], for two purposes:
  - Investigation of the relation between different components faults being followed by failure occurrences in others;
  - Demonstration of failure detection capability in various case studies, with known component failures.

2. Data description and processing

2.1. Description of the SCADA alarm system

The SCADA system is primarily used in WTs for simple monitoring and control. It acquires data to track environmental conditions, key performance parameters (power output, rotor speed), and to measure several aspects of the components. Concerning the SCADA alarm system, an alarm is activated when an incident occurs or to keep records of any operational status change. In general, different categories of alarms can be found in the SCADA alarm logs:

- *Grid conditions:* WT performance can slightly deteriorate during grid events. For this reason, grid-related alarms are recorded when some parameters exceed threshold limits, *e.g.* during voltage dips or frequency fluctuations.
- *Environmental conditions:* Wind conditions determine WT control and operation. Alarms are therefore activated when adverse environmental conditions are detected, *e.g.* high wind speed, icing, low or high ambient temperature.
- *WT operational state:* Many SCADA alarm systems include state-related alarms to inform about WT state, *i.e.* emergency, failure, stop, pause, warning or running. This information is essential to understand if an urgent action is required from the operator.
- *Manual stops or restrictions:* Power output curtailments, due to grid restrictions, noise problems or wind sector management strategies, are also frequently informed through alarms activation.
- *Maintenance activities:* Some WT manufacturers include maintenance-related alarms, to keep records of the work carried out.

- *Component malfunction:* Modern WTs include more or less sophisticated condition monitoring systems (CMS). Either way, when any signal (temperature, vibration, etc.) from the CMS or the SCADA exceeds the threshold corresponding to normal operating conditions, an alarm is triggered. In this manner, the alarm system can inform about the component health status avoiding complex signal processing.

2.2. Historical SCADA alarm data

The data available for the research work presented in this paper was collected by SCADA systems at 23 onshore WFs covering a period of three years of operation. The selected fleet aims to be representative of very different situations since the considered WFs are located in geographic regions with diverse climates and present a wide range of sizes. In addition, seven different turbine types were included. A general description of the fleet is presented in table 1.

Table 1. Characteristics of the selected wind turbine fleet

WT make	Technology	Rated power (kW)	Number of WTs	Alarm system
A	Geared Generator	1500	55	1
B	Direct Drive	2000	37	2
C	Direct Drive	2000	19	2
D	Geared Generator	850	77	3
E	Geared Generator	2000	133	4
F	Geared Generator	1800	9	5
G	Geared Generator	2000	76	5

Turbine types A, D, E, F and G correspond to three-blades turbines, with geared-drive technology and doubly fed induction generators (DFIG). On the other hand, turbine types B and C are three-blades turbines, with direct-drive technology and synchronous generators. The number of WTs for each turbine type is also specified in table 1, as well as the rated capacity and the SCADA system used.

As mentioned, alongside comprehensive signal information, SCADA systems record historical alarm and fault log files detailing alarm codes and their corresponding description, start date and duration. These files were collected from the described fleet covering a period of three years of operation. Typical data instances of the SCADA alarm systems are shown in table 2. Alarm IDs cannot be revealed for confidentiality reasons.

Table 2. Illustration of SCADA alarm data instances

WT affected	Alarm ID	Description	Start date	End date
15	XXXX	Lack of wind	2013-04-22 09:21:55	2013-04-22 10:36:10
9	XXXX	Generator heating	2013-04-22 09:23:06	2013-04-22 09:35:08

2.3. Data processing methodology

- *Alarms categorisation*
  In order to establish standard categories for the different turbine components, there is a need to define a specific structure. The ReliaWind project [13] has been identified as the most complete study on WT reliability to date. As part of the work accomplished within

this project, a new WT taxonomy was developed, based on the terminology of *system*, *sub-system*, *assembly* and *sub-assembly*. In order to adapt this taxonomy to the WTs selected for the present work, certain revisions have been incorporated. More information about these modifications as well as the full WT taxonomy can be found in an independent study developed by the authors [14].

In addition to the operational SCADA alarm data, original technical documentation prepared by the corresponding manufacturers was gathered. These alarm allocation lists document all alarm codes and the following related information: detailed description of the problem/failure, component affected and turbine status. Based on these records , each unique alarm code is manually associated to a specific *sub-system* and *assembly*, and even a *sub-assembly* when possible. Since the same WT taxonomy is used for all the different turbine types, it allows to create a homogeneous and standard categorisation.

It has to be mentioned that non component-related alarms are also associated to their corresponding category, *i.e.* grid conditions, environmental conditions, WT operational state, manual restrictions and maintenance activities (see section 2.1).

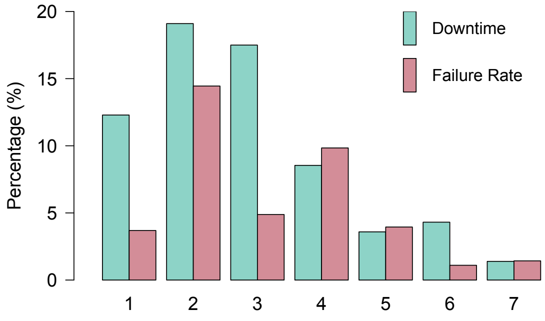
In short, a "dictionary" is created where every alarm code, from any manufacturer and turbine type, is mapped to a *sub-system*, *assembly* and *sub-assembly* or to a non component-related category. In this manner, historical alarm logs can be automatically translated into valuable information assessing WT status.

Moreover, when an alarm is triggered by the SCADA system it can lead to a WT shutdown. In addition to the association of each alarm code to its corresponding *sub-system* and *assembly*, it is categorised as braking or not, in order to distinguish between alarms incurring in WT downtime and warnings.

- *Failure data*

  In a parallel study [14], failure data were collected from more than 4300 WTs over the same operational period of three years. The reviewed taxonomy was also manually applied to the failure logs, allowing the combination of both datasets during the analysis, by locating real failures in time.

  Failure logs were analysed to assess failure occurrence and the consequent downtime, allowing to highlight the most critical cases. As a result, six assemblies were identified as the most severe: the gearbox, the generator, the pitch system, the yaw system, the frequency converter and the transformer. Moreover, expertise was sought from specialists in the WF O&M field, in order to gain insight into the incurred maintenance costs in case of full replacement. Following the consultation, WT blades were included as a critical component. Consequently, the most severe failures would involve the components listed in figure 1, alongside their normalised failure rates and downtime. Further details can be found in [14].



1	Blades
2	Gearbox
3	Generator
4	Pitch System
5	Yaw System
6	Transformer
7	Frequency Converter

Figure 1. Critical components in terms of normalised downtime and failure rate

Hence, investigation of the relation between different component-related alarms before real failures of the listed components is prioritised against other less severe cases.

A scheme of the proposed methodology is shown in figure 2. Historical alarms are automatically translated into component-related alarms, based on the categorisation established previously. Historical failure logs allow then to locate component failures in time. Finally component-related alarms prior to particular failures can be analysed using the techniques selected in [12]: time-sequence and probability-based analysis.

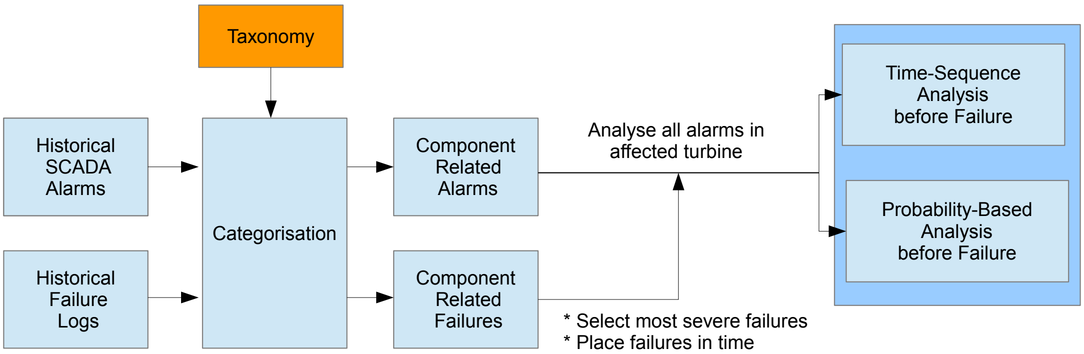

Taxonomy
Historical SCADA Alarms
Historical Failure Logs
Categorisation
Component Related Alarms
Component Related Failures
Analyse all alarms in affected turbine
* Select most severe failures
* Place failures in time
Time-Sequence Analysis before Failure
Probability-Based Analysis before Failure

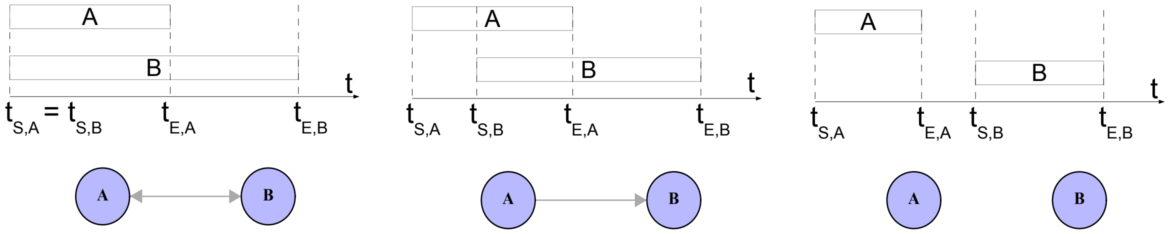
Figure 2. Scheme of the methodology for SCADA alarms and failure data processing

3. SCADA alarm analysis

Historical alarm logs can be understood as a succession of events. As suggested in [8], a specific fault is usually preceded by a certain alarm pattern. The purpose of analysing SCADA alarms is therefore the recognition of alarm patterns before real failures. On-line fault diagnosis could then be accomplished by relating present known alarm patterns to failures that are more likely to occur. Qiu *et al.* [12] applied time-sequence and probability-based methods separately to analyse SCADA alarms. In this paper, the time-sequence approach is used to recognise any alarm pattern. Then, the most frequent patterns are identified before the occurrence of actual failures. Both approaches are briefly described hereafter. All the analyses in the present study have been carried out using the R statistical software platform [15].

3.1. Time-sequence analysis

Given two different alarms A and B, the relationship between them is directly governed by their succession over time, as illustrated in figure 3; t_S and t_E state for start time and end time.



$$A \leftrightarrow B \qquad (1)$$

$$A \rightarrow B \qquad (2)$$

$$A \cap B = 0 \qquad (3)$$

Figure 3. Time-sequence analysis of alarms A and B

Two alarms starting at the same time would imply a two-way relationship of causality (1). On the contrary, if B starts while A is activated, A could be assumed to be the cause of B (2). Finally, if B is triggered after the end of A, no relationship would be established (3). As a result, this analysis technique allows us to identify any alarm pattern of the type $A \rightarrow B$ and $A \leftrightarrow B$. It is important to mention that the first assumption made for situation (2) could be confirmed by the analysis of historical data, as presented in section 4.1.

3.2. Probability-based analysis

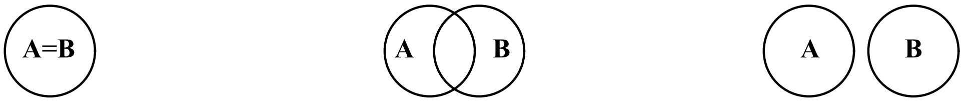
The probabilistic approach is used here to reveal the more frequent alarm patterns before real failures. Less frequent patterns are then considered to be unrelated to failures. Over a period of time T ahead of a specific failure, a total number of N alarms are triggered. The probability of occurrence of alarms A and B is:

$$P(A) = \frac{N(A)}{N} \; and \; P(B) = \frac{N(B)}{N}$$

Then, the probability of occurrence of the pattern $A \rightarrow B$ can be estimated as follows:

$$P(A \rightarrow B) = \frac{N(A \rightarrow B)}{N} = \frac{N(A \rightarrow B)}{N(B)} \frac{N(B)}{N} = P(A \mid B) P(B)$$

The dependency of both alarms before a failure can be illustrated through a Venn diagram.



$A \leftrightarrow B$ (1) $A \rightarrow B$ (2) $A \cap B = 0$ (3)

Figure 4. Venn diagrams for different grade of dependency between alarms A and B

4. Results and Discussion

4.1. Alarm pattern recognition

Time-sequence analysis was applied to historical SCADA alarms on a WT basis. All the WTs included in the selected fleet, described in table 1, were analysed over a period of three years of operation. Nevertheless, some alarms were excluded before performing the analysis. Since alarms related to manual stops or restrictions and to maintenance activities (see section 2.1) result from manual interventions, they were removed. Similarly, alarms informing about the WT operational state were excluded from the analysis.

For each WT, all the alarm patterns of type $A \rightarrow B$ and $A \leftrightarrow B$ were recognised and then grouped by the alarm system used, in order to asses the capability of failure detection. The assumption that A may be the cause of B can be confirmed by the pattern occurrence. Consequently, patterns with a low occurrence were excluded. The final number of different patterns identified per alarm system is shown in table 3.

As one can see, systems 3 and 4 show an important number of alarm patterns detected, whereas a lower number of patterns were identified for systems 1, 2 and 5. From this, an important conclusion can be drawn: the proposed technique is highly dependent on the characteristics of the SCADA system, directly corresponding to the WT manufacturer. Moreover, the number of alarm patterns detected could be correlated to the number of failures experienced by each group of WTs and their age, which was not considered in the present study.

Table 3. Alarm patterns identified per system

Alarm system	Number of WTs	Number of alarm patterns
1	55	36
2	56	12
3	77	990
4	133	2499
5	85	60

At first glance, systems 3 and 4 show a higher capability of failure detection through a correlation with alarm patterns.

At this stage, a shortcoming has to be highlighted. Some alarm patterns may remain unrecognised due to the accuracy of the alarm system. Indeed, a delay might exist between the incident start and its detection, followed by the corresponding alarm activation. This problem is illustrated in figure 5. Alarm A should be recognised as the cause of alarm B. However, the suggested methodology interpret both alarms to be independent due to the delay between the real incident start time, $t_{S*,B}$ and the recorded alarm start time, $t_{S,B}$, when the incident is acknowledged by the system. This could explain the low number of alarm patterns observed for WTs using SCADA alarm systems 1, 2 and 5.

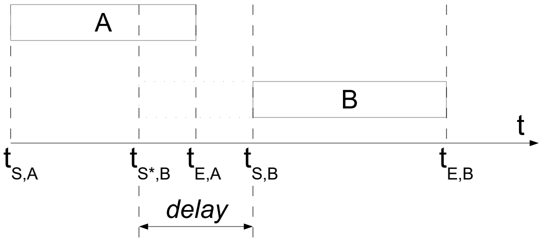


Figure 5. Alarm pattern unrecognised due to a slow response of the system

In future research work, this problem is intended to be overcome by applying more complex pattern recognition algorithms, based on machine learning techniques.

4.2. Case studies

From the previous analysis, SCADA alarm systems 3 and 4 show a higher capability of failure detection through alarm pattern recognition. These systems correspond to WT makes D and E (see table 1). In order to relate the recognised patterns to real failures, several WTs experiencing failures of critical components were selected. For the sake of clarity, two case studies from WT make E are presented here below.

- *Case 1: Gearbox failure*
  The selected WT is located in a WF which comprises 30 WTs. The site, which has been in operation for more than eight years, gives an installed capacity of 60 MW. It is located in a terrain of moderated complexity and no areas of important forest can be founded in the surroundings. Layout spacing generally maintains a distance of five rotor diameters in the predominant wind direction; this is considered sufficient spacing for wake recovery on a WF of this size.
  Historical alarm logs from the affected WT were analysed over a period of one month prior

to the gearbox failure. The period of time before the failure was determined to be sufficient to schedule a necessary intervention. Three main gearbox-related alarms were identified (a13, a14, a15). Alarm patterns leading to these alarms appearance are shown in figure 6 together with their corresponding occurrences. Each alarm is represented as a vertex in the network diagram characterised by its code and the component affected. Alarm codes are replaced by random numbers and the corresponding descriptions are not revealed, in order to ensure confidentiality.

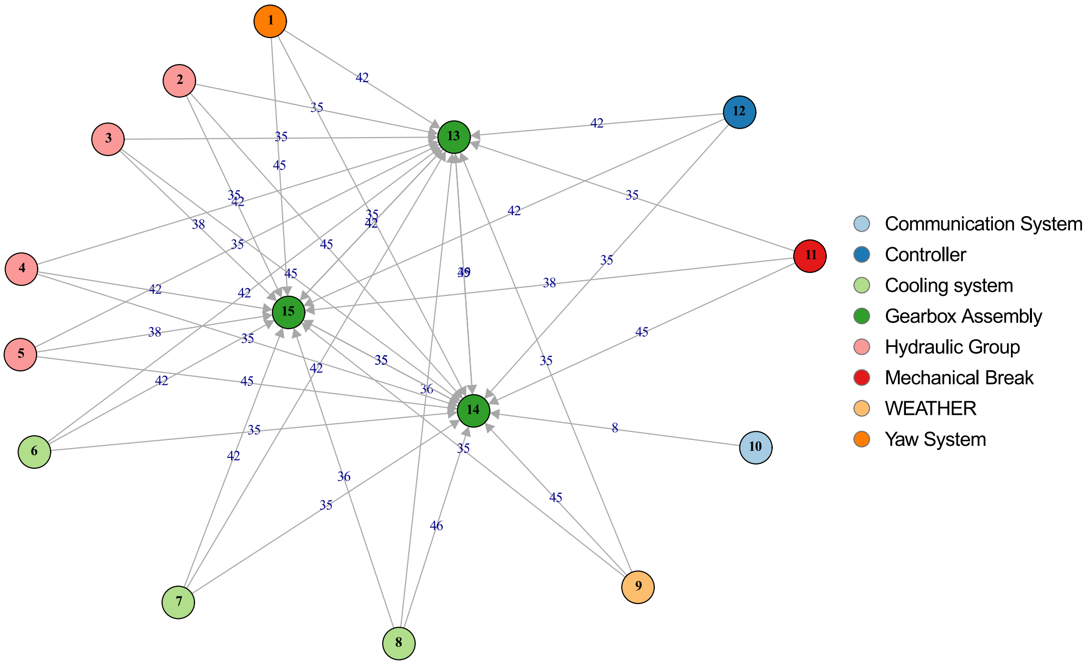


Figure 6. Network diagram of the SCADA alarm patterns before a specific gearbox failure

As can be seen, the gearbox failure did occur following many alarms involving the hydraulic system (a2, a3, a4, a5) and the cooling system (a6, a7, a8). Based on the number of occurrences, a causal relationship between these assemblies and the gearbox could be assumed. Alarms of a certain component resulted in a failure of another component. Moreover, the high number of occurrences of each pattern over a month before the gearbox failure confirms the assumption that a failure can cause a previous alarm pattern. We can conclude that this technique is successful in showing the relation between different components, by providing a holistic overview of the system before a certain failure.

- *Case 2: Yaw system failure*
The affected WT is part of a WF, with similar characteristics to the site described in the previous case study. Analogously, historical alarm logs were analysed over a period of one month prior to a failure of the yaw system. Alarm patterns recognised throughout the month ahead of the failure are shown in 7, together with their corresponding occurrences. Contrary to the previous case, a significantly lower number of patterns were detected. The three yaw-related alarms are linked to three other alarms (a20, a25, a26), being the weather-related the most significant. Indeed, alarm 26 indicates that high wind speed was detected. Nevertheless, this technique does not appear to be as robust as in the previous case, since it does not show any clear relationship between different alarms. The Venn Diagram showing the relation between the yaw alarms and the high wind speed alarm is presented in figure

8, alongside the alarm statistics. Each circle corresponds to a specific alarm; the area of the circle is proportional to the number of occurrences, which is also indicated.

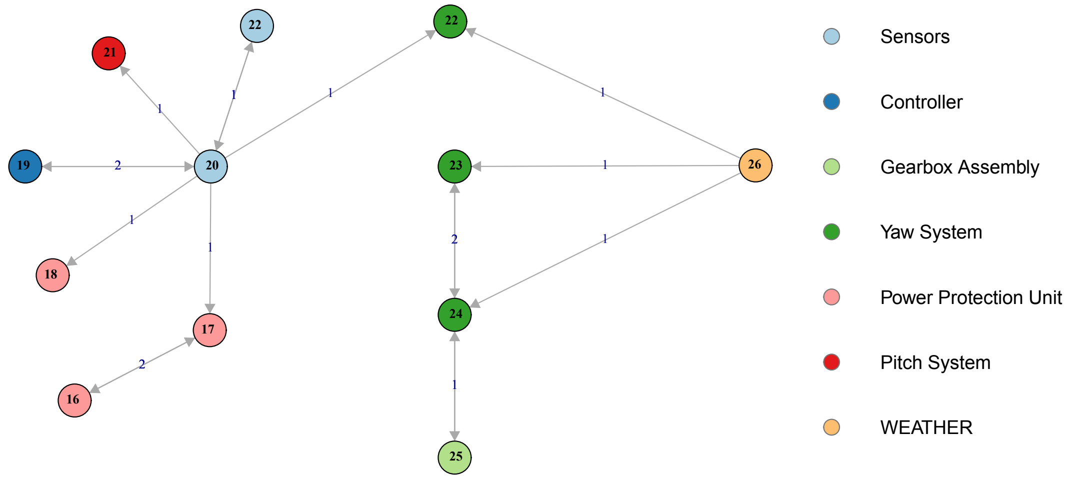


Figure 7. Network diagram of the SCADA alarm patterns before a failure of the yaw system

	Number of occurrences
N_{TOTAL}	71
$N(a23)$	2
$N(a26)$	26
$N(a26 \rightarrow a23)$	2
	Probabilities
$P(a23)$	2.82%
$P(a26)$	36.62%
$P(a26 \rightarrow a23)$	2.82%

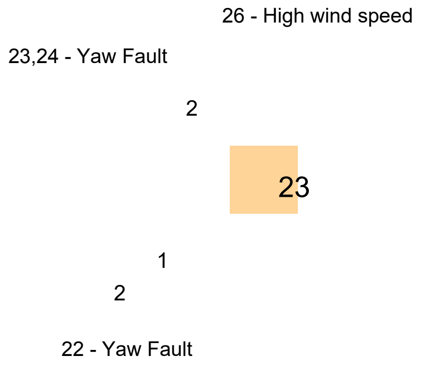


Figure 8. Alarm statistics and Venn diagram before a failure of the yaw system

As can be seen, the high wind speed alarm was activated several times, regardless the time-sequence patterns with the yaw-related alarms. In terms of probability, there is a much greater likelihood of emergence of the alarm 26 than the alarm pattern. Moreover, two of the three alarms indicating a yaw fault (23 and 24) were triggered whenever the high wind speed alarm was activated.

As a result, this probabilistic approach seems to be more convenient for detecting the yaw system failure before its appearance. Indeed, the analysis of the historical alarms reveals the link between alarms related to the affected system and alarms informing about adverse environmental conditions. In this particular case, the alarm pattern detected before the failure would not be of type $A \rightarrow B$; the significant number of times a certain alarm is triggered could be considered as an indicator of a failure likely to occur in the future; the high rate of high wind speed detection could be understood as the cause for the yaw failure.

Both case studies show the capability of the WT SCADA alarm system to provide a holistic overview of the condition of the turbine and so that failures at some components could be predicted way ahead of their actual appearance. Depending on the case, one technique seems to be more successful than the other in providing an alarm pattern before the failure. Nevertheless, since the methodology has only been applied to two case studies, a large number of thoroughly executed case studies would be needed to confirm the conclusion for other components.

Future research work will focus on identifying the most suitable technique for detecting failures of the seven critical components listed in section 2.3. This future study will rely on several historical failures of each component. In addition, specific alarm patterns will be related to particular failure modes issued from the analysis of the available failure data.

5. Conclusion

This paper presents a cost-effective methodology to process and analyse WT SCADA alarms for component failure detection purposes. The suggested categorisation has showed its success in improving alarm handling process, by translating overwhelming data into component-related valuable information. Its cost-effectiveness relies on the automatic use of SCADA data, avoiding additional equipment and complex signal analysis processes. As illustrated through case studies, the present work highlights the relationship between faulty behaviour in different components and between failures and adverse environmental conditions. In other words, certain alarms could be directly related to upcoming component failures. The results show also the capability of the WT SCADA alarm system to provide a holistic overview of the condition of the turbine and so that failures at some components could be predicted way ahead of their actual appearance.

Nevertheless, the suggested analysis technique, based on time-sequence pattern recognition, presents some drawbacks; further research work will aim at overcoming these problems by applying more complex machine learning techniques.

Finally, the authors are also currently exploring other applications of the proposed methodology for SCADA alarms processing. Component-related alarms are being correlated to deviations from normal WT operating conditions. In other words, the contribution to WT underperformance events of specific component failures is being investigated.

6. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642108.

References

[1] Spinato F, Tavner P, van Bussel G and Koutoulakos E 2009 *IET Renewable Power Generation* **3** 387–401
[2] Kusiak A and Li W 2011 *Renewable Energy* **36** 16–23
[3] García Márquez F P, Tobias A M, Pinar Pérez J M and Papaelias M 2012 *Renewable Energy* **46** 169–178
[4] Yang W, Tavner P J, Crabtree C J, Feng Y and Qiu Y 2014 *Wind Energy* **17** 673–693
[5] Yang W, Court R and Jiang J 2013 *Renewable Energy* **53** 365–376
[6] Schlechtingen M, Santos I F and Achiche S 2013 *Applied Soft Computing* **13** 259–270
[7] Zaher A, McArthur S, Infield D and Patel Y 2009 *Wind Energy* **12** 574–593
[8] Chen B, Qiu Y, Feng Y, Tavner P and Song W 2011 *IET Conference on Renewable Power Generation*
[9] Chen B, Tavner P J, Feng Y, Song W W and Qiu Y 2012 *EWEA 2012*
[10] Godwin J L and Matthews P 2013 *International Journal of Prognostics and Health Management* **4**
[11] Kaidis C, Uzunoglu B and Amoiralis F 2015 *IET Renewable Power Generation* **9** 892–899
[12] Qiu Y, Feng Y, Tavner P J, Richardson P, Erdos G and Chen B 2012 *Wind Energy* **15** 951–966
[13] Wilkinson M, Hendriks B, Spinato F, Gomez E, Bulacio H, Tavner P, Feng Y and Long H 2010 *EWEC 2010*
[14] Reder M, Gonzalez E and Melero J J 2016 *Journal of Physics: Conference Series* Manuscript accepted for publication
[15] R Core Team 2014 *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing Vienna, Austria