Journal Paper

“Reliability analysis of detecting false alarms that employ neural networks: A real case study on wind turbines”

- Reliability Engineering & System Safety-

November 2019

Alberto Pliego Marugán
Ingenium Research Group, Castilla-La Mancha University, Spain
Alberto.Pliego@uclm.es

Ana María Peco Chacón
Ingenium Research Group, Castilla-La Mancha University, Spain
AnaMaria.Peco@uclm.es

Fausto Pedro García Márquez
Ingenium Research Group, Castilla-La Mancha University, Spain
FaustoPedro.Garcia@uclm.es

Cite as: Marugán, A. P., Chacón, A. M. P., & Márquez, F. P. G. (2019). Reliability analysis of detecting false alarms that employ neural networks: A real case study on wind turbines. *Reliability Engineering & System Safety*, *191*, 106574.

DOI: https://doi.org/10.1016/j.ress.2019.106574

Reliability analysis of detecting false alarms that employ neural networks: a real case study on wind turbines

Alberto Pliego Marugán[1], Ana María Peco Chacón, Fausto Pedro García Márquez[1]

[1]Ingenium Research Group, Universidad Castilla-La Mancha, 13071 Ciudad Real, Spain.

(alberto.pliego; Ana.Peco; FaustoPedro.Garcia)@uclm.es

Abstract

Operations and maintenance tasks are critical to the reliability of a wind turbine. The state-of-the-art demonstrates the effectiveness of reliability centred maintenance, but there are no research studies that consider false alarms to reliability of the wind turbines. This paper presents a novel approach based on artificial neural networks to reliability centred maintenance. The methodology is employed for false alarm detection and prioritization, training the artificial neural networks over the time to increase the system reliability. The approach is applied to a real dataset from a supervisory control and data acquisition system together with a vibration monitoring system of a wind turbine. The results accuracy is done by confusion matrices, studding real alarms with the estimations provided by the approach, and the results are validated with real false alarms and compared by the results given by a fuzzy logic model. The method provides accuracy results (over 90%). A novelty is to use a two real dataset from a wind turbine to create a redundant response to detect false alarms by artificial neural networks.

Key words: Reliability centred maintenance, condition monitoring, artificial neural network, wind energy conversion systems, false alarms

1. Introduction

Wind energy has become one the most important renewable energy sources due to wind energy conversion systems (WECS), and increases in the complexity and extent of producing electric power [1]. It has led to increases in the efficiency of the system, and therefore in competitiveness within the electrical market. Investment in renewable energy depends on technological conditions, politics and economic decisions [2]. It has generated a growing trend in global, annual installed wind capacity, according to Figure 1 [3,4]. The cumulative global capacity rose from 3.5 GW in 1994 to more than 480 GW in 2016, i.e. 20 years ago, the global capacity was 0.7% of the total current capacity [5]. Wind energy is expected to increase at an annual average rate of 6% until 2035 [3].

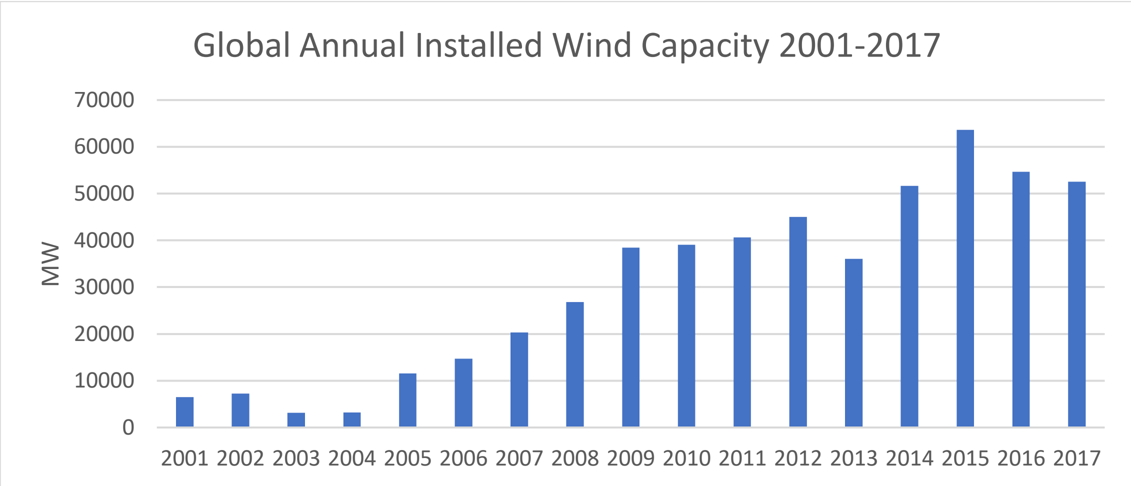

Fig 1: Global Annual Capacity. Source: Global Wind Energy Council [3].

It is estimated that the electricity produced by one wind turbine (WT) is similar to the electricity produced daily by 1.000 Kg of fuel. Experts have estimated that wind energy avoided the emission of 140 million tonnes of CO_2 in 2011 in the EU, the equivalent of one third of all car emissions in the EU [6].

The expected growth of wind energy is a consequence of the new technologies for WECS. Offshore wind farms will play an essential role in the future. Offshore locations allow more energy production than onshore ones because there is more wind and the WTs can be larger, and they also have less environmental impact [7]. Figure 2 shows the expected onshore and offshore capacity until 2030. Onshore capacity indicates a decline in growth, whereas offshore capacity continues to grow.

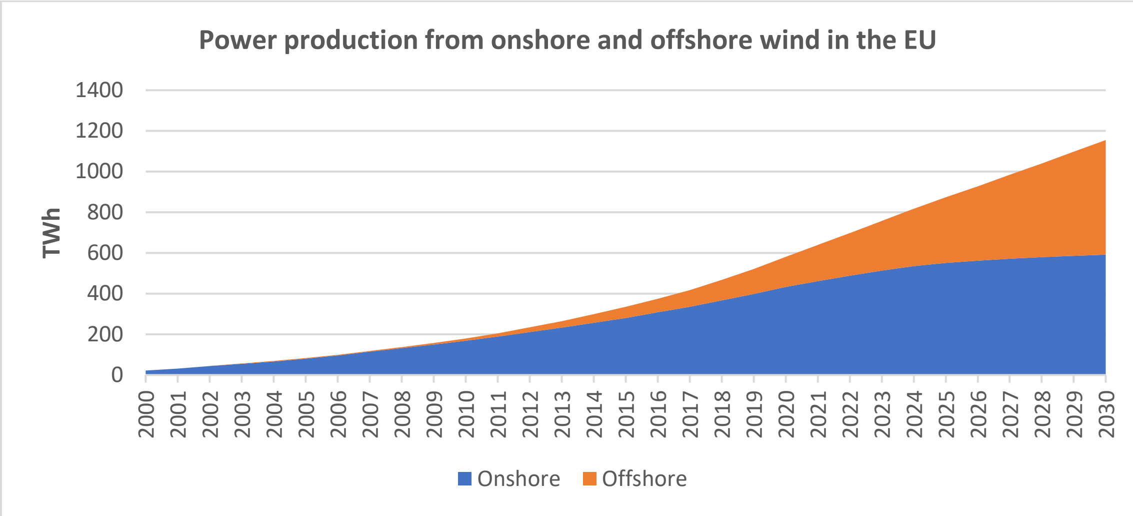

Fig 2: Offshore and onshore capacity estimation [6].

Offshore wind energy involves some drawbacks that must be considered to ensure effective energy production, but it is more difficult to predict and evaluate the wind as a resource. The main difference between offshore and onshore is investment as well as operation and maintenance (O&M) costs. Offshore installation costs are double those of onshore; offshore installation costs being about 1.44 million €/MW [8]. Offshore O&M costs are 23% of total system costs, 12% of which are for onshore wind farms [9].

Wind farms will require a specific infrastructure regarding the location and the transport required for maintenance tasks, e.g. vessels, helicopters, fixed installations, mobile jack-up installations, etc. [9]. The costs are high; therefore it is necessary to optimize preventive and corrective maintenance strategies in order to reduce their use [10].

Wind farms require optimized maintenance processes to avoid losses of production and to increase the system reliability [11]. Many researchers have demonstrated the efficiency of the reliability centred maintenance systems for WTs [12-15]. New techniques and methods for data analysis are employed in reliability centred maintenance [10,16,17]. However, a few studies are focused on the consequences and effects of an incorrect data analysis. False alarms, also known as predictors, can inflict significant economic loss for both offshore and onshore wind farm operators.

This paper presents a novel approach for improving the reliability of wind turbines. In this field, multiple data driven techniques and methods have been developed. These methods are based on risk analysis [18], correlation of faults and weather conditions [19], the analysis of the statistical uncertainty of component reliability [20] or dirt and mud detection [21]. Other studies aim at developing maintenance strategies through opportunistic condition based maintenance [22], or considering production and deterioration uncertainties [23]. This paper proposes a data driven analysis for identifying false alarms.

This function can be incorporated into any maintenance strategy [24], leading to significant improvements in system reliability.

This paper is divided into six sections. The first section is an introduction where the contextual information is provided in order to explain the main objective of the paper. Section 2 presents a state-of-the-art of the techniques and methods employed to improve the reliability of wind turbines. Section 3 provides information about the principal characteristics of the data acquisition systems employed in this paper. Section 4 discusses the approach. The approach is applied in a case study in Section 5. The results of the case study and a validation of the methodology are provided in Section 6. Finally, some conclusions are drawn and presented in Section 7.

2. State-of-the-Art

According to industrial standard ISA-18.2 (2009) [25], "an alarm system is the collection of hardware and software that detects an alarm state, this communicates the indication of that state to operators, and it records changes in the alarm state". Alarm is also defined as an "operational signal or message designed to notify personnel when a selected anomaly, or a logical combination of anomalies, requiring corrective actions is encountered" [15]. Regarding to ISO 13379 [15], descriptors could be identified in some cases as alarms because they are used to express symptoms and anomalies. Descriptor is defined as a "data item derived from raw or processed parameters or an external observation". In some research papers the term "condition" can be found instead of descriptors. This paper employs the term alarm rather than predictor because predictor can also be identified with other terms related to measurement, and alarm also contains details different from symptoms and anomalies.

Condition monitoring (CM) and supervisory control and data acquisition (SCADA) systems have led to improvements in the reliability and the productivity of wind farms [26,27]. CM is defined as "the detection and collection of information and data that indicate the state of a machine" (see reference [28]). However, these systems can generate false alarms due to a wrong data analysis [14], i.e. many alarms can be triggered and turn out to be false. A solution is to prioritize alarms and try to detect any false alarms. Quiu et al.[29] showed the importance of improving the reliability of alarm detection systems, because the data is overwhelming and there are a large number of false alarms. Their research demonstrated that 11.5% of the alarms were reset without any intervention or following any rules, and 12% can be reset by certain "rules".

Hameed et al. [30] showed an implementation of a CM and Diagnosis (CM&D) system, allowing the detection of faults, but also generating false alarms. A general data interpretation and diagnostic techniques of CM&D can be found in reference [15]. They concluded that it is necessary to fix an optimal threshold criterion to detect false alarms. A detailed state-of-the-art of WT reliability based on CM&D is presented in [31,32]. Yang et al. [14] explained the main characteristics and limitations of the current CM&Ds. They concluded that the variation of fault-related parameters, e.g. temperatures or vibration, may not always correspond to a WT fault because they can depend on more variables. They proposed a detailed analysis of the data to find the cause of the variation and, therefore, to increase the reliability of the alarms. Allan May and David [33] showed that the identification of false alarms affects the cost-benefit of the CM&D in offshore WTs. They employed Markov chains and a time-series model to determine the implications of CM&D on detection rates, or false alarms. This work shows that the reliability of the CM&D affects the annual failure rate and the availability of the WTs. They concluded that the CM&D can reduce operational lifetime costs by 20% of preventive maintenance using some known detection rates.

Crabtree et al. analysed the main commercial CM employed in SCADA systems in WT [34]. They distinguish between onshore and offshore cases. Chen et al. [35] showed the main methods employed in SCADA to analyse the data employing statistical methods to data offline and online, and considering a large number of signals.

This paper analyses the signals from the SCADA and CM&D together to increase the accuracy of the system reliability. There are similar studies, but with different approaches, fewer variables analysed or only ones that focus on a component of the WT. Feng et al. [36] study the reliability of a gearbox in WTs, demonstrating that faults can be predicted by SCADA and CM&D together just days before they happen. However, only a signal is used, generating false alarms. They suggested analysing the SCADA and CM dataset to increase accuracy. The systems process the data independently and with different formats, limiting the accuracy of alarm prediction. Chen et al. [37] consider the SCADA dataset to analyse the oil temperature in bearings together with CM to detect gearbox faults weeks in advance. Dao et al. [38] reached a similar conclusion using non-linear data trends to continuously monitor the WT. This paper presents an approach that considers a large number of variables that require a robust approach based on artificial neural networks (ANN).

There are new approaches based on ANN that are improving fault detection accuracy. Su and Chong [39] proposed a ANN fault detection system for modelling an induction motor through the vibration spectra. They observed that vibration harmonics variations can be used for detecting faults, but tracking variations in the fundamental vibration might result in false alarms. A similar approach is employed in this paper, but considers a large number of signals from the SCADA and CM.

The real case study employed in this paper was considered in reference [40]. The approach presented generates alarms based on the collected data. False alarm detection is based on a new approach for the identification of alarms by Fuzzy Logic. Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth, truth values between "completely true" and "completely false”. The creation of fuzzy systems implies the definition of a set of fuzzy rules. In this case, the variables considered correspond to the distance of the value measured by the SCADA to the simple moving average. The greater the distance from the average, the more likely an abnormal measure it is, and, therefore, the greater the likelihood of an alarm being generated. Three different outcomes of the fuzzy system have been considered. Firstly, the values are in range of normal behaviour and no actions are required. Secondly, orange alarms are triggered where the probability of an alarm exceeds a defined threshold, but it is not a critical point. Finally, a red alarm is triggered when the probability is unacceptable and the system needs urgent action. The results of this methodology can become a statistical support for the generation of alarms. The methodology can also be used as complementary information for evaluating the priority of each alarm. A similar type of process is used in neural networks (NN), expert systems and other artificial intelligence applications.

Most research considers the benefits of detecting alarms, but the literature review shows that there are not many studies that consider the disadvantages of false alarms [41]. In this paper, a novel approach is presented to increase the system reliability studying the SCADA together with a CM based on vibration signals. The approach employs ANN to detect false alarms and alarm prioritization. It considers the responses from a SCADA and a CM&D to provide redundant results to increase the accuracy of the results. The data, from a real case study, are studied over time.

3. SCADA and CM to RCM in WTs

WTs are monitored by sensors, where fault detection is generally done by a simple threshold to complex signal processing methods [42]. This paper considers CM based on vibration signals together with a SCADA. The vibration signal is done according to ISO standard 2041 [43].

RCM is a methodology that identifies the functions of a system in a given operational context, the way these functions may fail and then establishes a set of applicable and effective preventive maintenance tasks, based on considerations of system safety and economy. [15]. Signal processing is employed on the CM/SCADA data to set a correct maintenance [44]. Several research studies have demonstrated the benefits of RCM [45-48].

The main WT components that are monitored are blades [49], bearings [50,51], gearboxes [52], electrical or electronic components [53,54] and the tower [49]. Several measurement techniques can be employed for CM [55]: Vibration [56]; acoustic emission [21,57]; ultrasonic testing [58,59]; rotor speed [60]; oil analysis [12],etc. RCM is employed to reduce the energy loss, the downtimes and to optimize the O&M tasks [50]. Vibration analysis is a technique for monitoring rotatory equipment [61,62]. Gearbox, bearings and rotor are the most susceptible components to be analysed with vibrations. The CM is based on vibration signals employing 8 accelerometers, see Figure 3 [5]. A one-second signal from each point is selected and stored every three hours to discretize the continuous vibration signals. The feature parameters are extracted from the one-second signals.

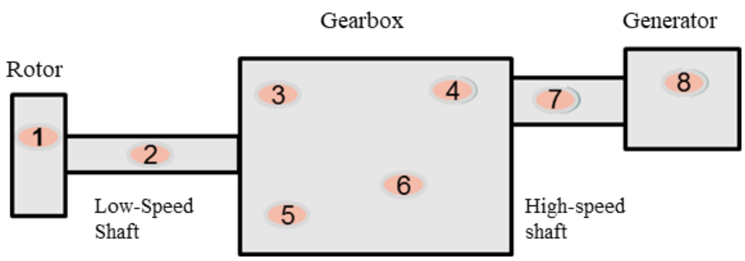

Fig. 3. Location of the different accelerometers of the CM [5]

The SCADA is also employed for fault detection [63,64], but it can be employed for forecasting [65,66] and production assessment [67,68]. It collects data from several sensors, usually every 10 min, and is managed by a central computer. The sensors are employed for finding an optimal control solution and reducing O&M costs [69-71]. There are some advantages, see e.g. [12]; however some studies show that it can present some disadvantages related to reliability and operational conditions [69]. This paper study 34 variables from the SCADA together with the dataset from CM, a novelty regarding the state-of-the-art.

4. Approach

The main objective of this paper is to develop a methodology to detect false alarms in a WT. The proposed strategy is to correlate the condition of the WT and the outcomes of two independent data acquisition systems (SCADA and CM). The approach employs an NN-based structure to establish relationships between data from a CM, that measures vibrations, and a SCADA system. The approach is composed of the multi-stage decision scheme shown in Figure 4, and employs the nomenclature shown in ISO 13374 [72].

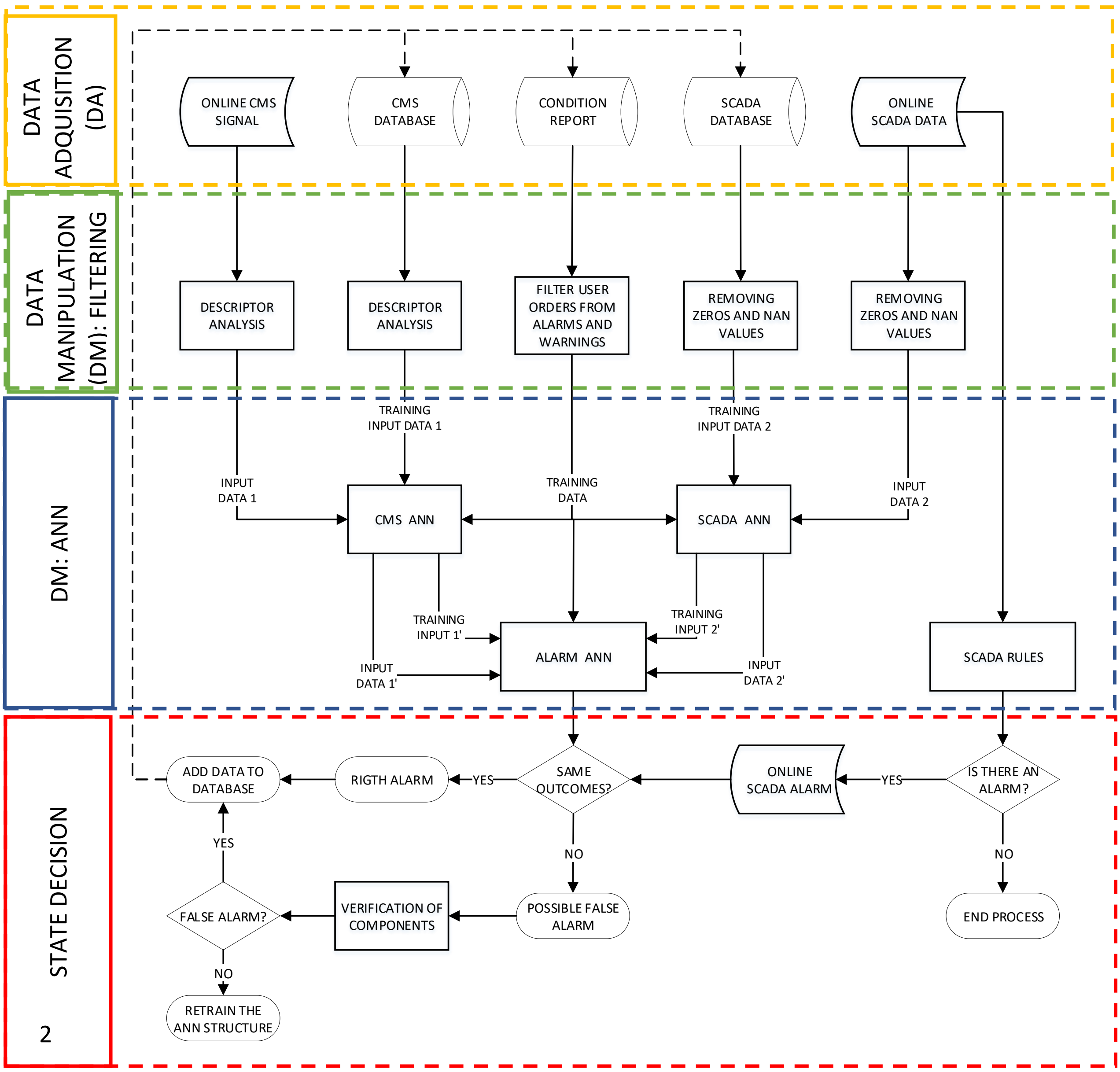

Fig.4. Flowchart of the proposed approach (terminology adapted from [72])

The design of the ANN based structure is based on:

- The analysed data that comes from two independent data sources. The ANNs enable the processing of data of different natures.

- Three different ANNs allow three different solutions to be obtained. The redundancy of the solutions is essential for reinforcing the veracity of a specific outcome and avoiding errors in the detection of false alarms. The output of each ANN can be considered as an independent solution of the structure.

- ANNs allows an intelligent adjustment of the weights of ANNs in order to provide a response adapted to the current conditions of the WT. This adjustment is carried out through a feedback process where the historical database is uploaded.

- The cascade connection of the Alarm-ANN in this structure allows analysis of the SCADA system and the CM from a holistic point of view, considering all the data collected from the WT as a unique data source. This means the effect of possible errors in the SCADA output are attenuated due to the CM outcome and vice versa.

4.1 Data acquisition (DA)

The data is collected in a specific format according to the data acquisition system. It is transformed to a scaled digital representation [72]. Two data sources are studied: historical and online data. The data is formatted to a common structure [73]. In this paper, three different databases have been considered regarding the data source [72]:

- *SCADA database*: It contains the SCADA dataset collected over a long period. This database contains the values and the units of the parameters, and the times when they were measured.
- *Condition report:* This database contains the WT condition over time, e.g. alarms, warnings, orders, etc. Alarms cause the WT to stop working. Warnings indicate that there are some parameters out of range, and the orders request start/stop. This paper is focused on alarms and warnings. Orders are not considered because they cannot assume a false positive. The condition report indicates whether the WT presents an alarm or not, warning and/or any order at any time.
- *CM database*: It contains the historical data from the CM. This information is not completely stored since the CM provides continuous signals and, therefore, there is not enough capacity for it to be stored in the real case study. Feature parameters are extracted from the signals in order to reduce the size of the dataset without losing the most important information.
- The *online data* is the information about the current WT condition. This data is also provided by the SCADA and the CM. The online data will be processed only in the event of an alarm, see Figure 4.

4.2 Data Manipulation: filtering and preparation

The data from the condition report can generate false alarms and separate them from the orders. The filter acts so that the database only has the desired alarms and warnings.

The SCADA dataset is filtered to remove incorrect values. This data is provided by a large number of sensors, where some can be out of service. The SCADA must trigger an alarm when the parameters exceed a determined threshold, but no alarm should be activated when a sensor is not providing any data, i.e. not a number (NaN). The wrong data is removed from the SCADA database before the NNs are trained. The NaN components are converted into 0, and the zero vectors are removed from the database. Figure 5 shows the SCADA before/after being filtered.

	SCADA Data before filtering				
	1	2	3	...	N
Date 1	a_1^1	a_1^2	a_1^3	$a_1^{\cdots}$	a_{1j}^N
Date 2	a_2^1	a_2^2	a_2^3	$a_2^{\cdots}$	a_{2j}^N
Date 3	a_3^1	a_3^2	a_3^3	$a_3^{\cdots}$	a_{3j}^N
Date 4	0	0	0	0	0
Date 5	a_5^1	a_5^2	a_5^3	$a_5^{\cdots}$	a_5^N
Date 6	a_6^1	NaN	a_6^3	$a_6^{\cdots}$	a_6^N
Date 7	a_7^1	a_7^2	a_7^3	$a_7^{\cdots}$	a_7^N

Delete vector

Set component to 0

Fig. 5. Erroneous values from SCADA system

A period is defined for the CM data before the feature parameters are calculated, because they can only be achieved for a certain memory [5]. The parameters employed in this paper regarding ISO 3373 [74] are: mean, root mean square, peak, skewness, standard deviation, shape indicator, Kurtosis, crest factor, impulse and clearance factor.

4.3 Data manipulation: ANN

Once the data have been filtered and set, they will be used by an NN-based structure. This structure is composed of three different networks.

A *CM-ANN* will link the feature parameters extracted from the CM with the condition report. The inputs and outputs of the ANN are provided in Table I. The CM-ANN will recognise descriptors (patterns) in the CM data, and they will be associated to the specific condition of the WT set by the condition report at that time. In other words, the CM-ANN employs the condition report to supervise the training.

Table I. CM-ANN input/output structure

	Input CM Data			Output
	Feature 1	Feature 2	Feature '*j*'	WT Condition
Date 1	e_{11}^k	e_{12}^k	e_{1j}^k	C_1
Date 2	e_{21}^k	e_{22}^k	e_{2j}^k	C_2
Date 3	e_{31}^k	e_{32}^k	e_{3j}^k	C_3
Date 4	e_{41}^k	e_{42}^k	e_{4j}^k	C_4

e_{ij}^{k} represents the value of the feature *'j'* extracted from the signal of the CMS sensor *'k'* at the data acquisition time *'i'*. The features are defined in Table IV. C_i is the condition in *i*.

The CM data are independent of the condition report because the report is generated by the SCADA. Therefore, alarms not related with vibrations will not detected within the CM data. The objective is to employ these data to create redundancies in the diagnostic and, consequently, to support the decision of the SCADA system.

A *SCADA-ANN* will establish the relationship between the condition report and the SCADA data. As aforementioned, the condition report is generated from the SCADA data. Each alarm is activated when a specific parameter, or a set of parameters, exceeds a determined threshold. This ANN considers the specific parameters together with the SCADA data. The SCADA-ANN employs the condition report to supervise the training. The input/output structure of this ANN is indicated in Table II:

Table II. SCADA-ANN input/output structure

	Input SCADA Data			**Output**
	Parameter 1	**Parameter 2**	**Parameter *'i'***	**WT Condition**
Date 1	P_{11}	P_{12}	P_{1j}	C_1
Date 2	P_{21}	P_{22}	P_{2j}	C_2
Date 3	P_{31}	P_{32}	P_{3j}	C_3
Date 4	P_{41}	P_{42}	P_{4j}	C_4

P_{ij} is the value of the parameter *j* collected by the SCADA system at the sampling time *i*. The different parameters are described for a specific case study in Table III.

Health assessment is employed to determine if the system is degraded [72]. In case of *i* presents an abnormal value, the ANN will evaluate these parameters to determine whether an alarm should be activated or not. The SCADA activates an alarm when the wind speed exceeds a threshold. The ANN considers all the variables, i.e. the wind speed, the turbulences, the vibrations, etc. An incorrect work of the anemometer could be detected without activating an alarm.

Finally, the Alarm-ANN joins the outputs "CM-ANN" and "SCADA-ANN" and processes these data in order to obtain a deeper data analysis. At this level, the data inputted into the Alarm-ANN does not come from the WT, but it is processed information provided by the previous ANNs. The Alarm-ANN will recognise patterns and will correlate these patterns with the WT condition. The outcomes of the Alarm-ANN represent the unification of the outcomes of the previous ANNs.

4.4 State Decision

The final objective is to support the detection of false alarms and the prioritization of alarms. The Alarm-ANN output is compared with the online alarm activated by the SCADA. If both outputs are the same, then all the data must be stored. If the outputs are not the same, action is needed according to the SCADA indications. In the event of a false alarm, the ANN based structure works in normal conditions and the data will be added to the database. If the method provides an incorrect result, the structure must be retrained until the correct alarm is generated using the corresponding current data.

The response of the ANN-based structure generates useful information to the operators that decide which alarms could be false and the alarms that should be prioritized. Each ANN has its own precision and sensitivity for alarm detection. The objective is to compare the alarm activated by the SCADA system with the outcomes of the ANNs to validate the results.

The oldest data must be deleted from the databases when the maximum storage capacity has been reached. Consequently, the databases will contain uploaded data, and the NN-based structure will maintain its size.

5. Case study

A real SCADA and CM installed in a large WT is analysed. The data were collected for 2 years and belong to the European project OPTIMUS [75]. The SCADA provides parameters every ten minutes shown in Table III [76].

Table III. SCADA parameters

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

The data shows the maximum, the minimum and the average values of each parameter in periods of ten minutes. Therefore, each SCADA data vector, **Sv**, is defined as:

$$\mathbf{Sv} = \left[t, \min(p_1), \overline{p_1}, \max(p_1), \ldots, \min(p_i), \overline{p_j}, \max(p_j) \ldots \min(p_{34}), \overline{p_{34}}, \max(p_{34})\right]$$

where p_j is the parameter j at the sampling time t.

The SCADA also provides the Health Assessment Report (HAR). HAR is used in this paper to supervise the training stage of the NN. HAR is uploaded if the condition of the WT has any variation. Each Condition Vector (**Cv**) is defined as:

$$\mathbf{Cv}=[\text{date, code, act}] \quad , \quad \text{act}= \begin{cases} 1 & for\ activation \\ 0 & for\ deactivation \end{cases}$$

where *code* is referred to the identifier of each condition, also called descriptor [15], and *act* is a binary variable that indicates if the condition is activated or deactivated.

The occurrence frequency of the different alarms is shown in Figure 6. The alarm codes cannot be explained in detail for confidentiality reasons. More than 90% are concentrated in the first 10 alarms. Therefore, to simplify this case study for the reader, the alarms considered in this paper are the 10 most repeated alarms in the historical data.

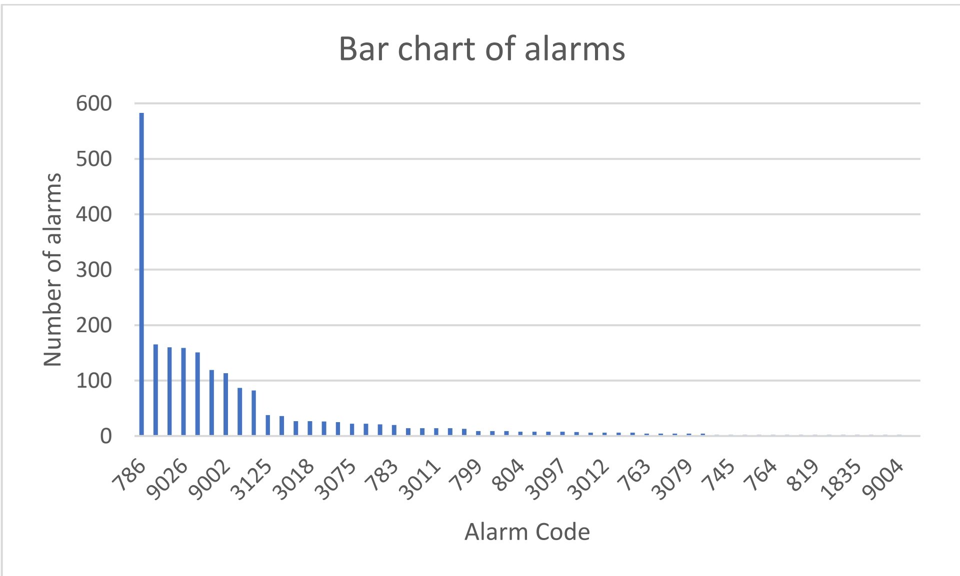

Fig. 6. Bar chart of the alarms

Figure 7 shows an example of a one-second signal provided by the accelerometer 3 (Figure 3) that has been selected from the continuous vibration signal. The sample frequency is 1024 sample/s.

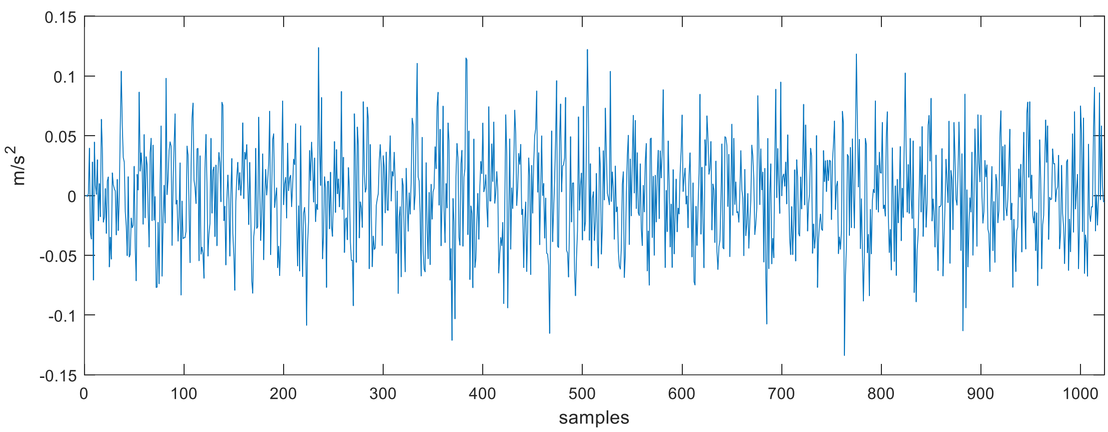

Fig.7. Example of one-second vibration signal

The time domain signal is discarded and replaced by the corresponding parameters once the feature parameters are extracted according to Table IV. Only 1.1% of data is considered but the main information about the signal is not lost.

Table IV. Example of feature parameter extraction

Date	**Point**			
05/03/2014	3			
Mean	RMS	Peak	SD	Skewness
8.4778e-05	0.0413	0.1238	0.0413	0.0623
Kurtosis	Crest Factor	Shape Ind.	Clearance	Impulse Factor
2.8400	3.2417	1.2402	111.7168	3.7186

The information shown in Table V is collected in the monitoring vector (**Mv**), defined as:

$$\mathbf{Mv} = [\text{date}, \text{point}, \text{mean}, \text{Rms}, \text{SD}, \text{peak}, \text{skewness}, \text{kurtosis}, \text{crest factor}, \text{clearance}, \text{shape}\ , \text{impulse}\ ],$$

where all the components, except the date, are repeated 8 times due to the different points of interest, i.e. **Mv** has 89 different elements.

The NNs are trained using the historical data from the SCADA system, the CM and the HAR. The number of neurons in the output layers is 10 because only the 10 most frequent alarms are considered. The Alarm-ANN input layer has 20 neurons that correspond to the sum of the outputs of the previous NNs.

The size of the NNs, i.e. the number of neurons in the hidden layer, is set according to the geometric pyramid rule [77]. The stochastic gradient descent with momentum (SGDM) optimizer is employed for training the ANNs. The hyper-parameters (learning rate, batch size, momentum, regularization) have been selected according to the recommendations given in reference [78]. The maximum number of epochs has been set to 22 for all the ANNs. The main features of the ANNs are shown in Table V.

Table V. Example of feature parameter extraction

NN	INPUT	OUTPUT	Neurons Input layer	Neurons Hidden layer	Neurons Output layer	Initial learning rate	Batch size	Momentum	L2 Regularization
SCADA-ANN	**Sv**	Alarm Code	103	20	10	0.01	128	0.9	0.0001.
CM-ANN	**Cv**	Alarm Code	89	32	10	0.006	98	0.95	0.0001.
Alarm-ANN	Sum of alarms Codes	Alarm predicted	20	16	10	0.01	150	1	0.0001.

6. Results and validation

6.1 Accuracy Analysis by Confusion Matrices

The decision-making process, given in the state decision [72], will provide feedback that allows the ANNs to improve the accuracy of their outputs as health assessments. The training process results of the ANNs are shown in Table VII by three confusion matrices. Each confusion matrix is an ordered representation of the classification provided by a NN, where the predicted values are compared with the real values. Table VI shows a scheme of a confusion matrix.

Table VI. Scheme of confusion matrix

		Target Class	
		Positive	**Negative**
Output Class	**Positive**	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

The accuracy analysis of the NNs is made by the confusion matrix considering [79] [80]:

- Accuracy (ACC): It corresponds to a measure of the degree of coincidence between predictions and the reality. The accuracy of an

ANN is calculated as the quotient between the number of correctly classified samples and the total number of examples. It can be obtained by:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

- Sensitivity or true positive rate (*TPR*): It is the fraction of positive examples predicted correctly by the NN. In this paper, this parameter corresponds to the ability of the ANN to detect a specific alarm. It is calculated by:

$$TPR = \frac{TP}{TP + FN}$$

- Specificity or true negative rate (*TNR*): It is the fraction of negative examples predicted correctly by the NN. In this paper, this parameter represents the capacity of the ANN to refuse the existence of a specific alarm.

$$TNR = \frac{TP}{TN + FP}$$

- Precision or positive predictive value (*PPV*): It is defined as the quotient between the true positives and the number of positives predicted by a NN. In this paper, this measure shows the degree of success of an ANN when a specific alarm is predicted. The main objective of this paper is false alarms detection; therefore, this parameter is employed for aiding the decision-making process. The precision is obtained by:

$$PPV = \frac{TP}{TP + FP}$$

- Negative predicted value (*NPV*): This measurement is the capacity of the ANN to discard the existence of a specific alarm, i.e. the degree of success when the occurrence of an alarm is refused.

$$NPV = \frac{TN}{TN + FN}$$

Table VII shows the results of the different NNs (left) and the accuracy analysis by the confusion matrices (right).

Table VII. Results and validations of the ANNs

Confusion matrix of SCADA-Condition NN									
257	5	0	1	0	0	2	1	1	0
0	52	1	2	1	0	5	0	3	0
0	0	54	0	0	2	0	0	0	0
0	2	3	36	34	0	0	0	0	0
34	3	30	67	69	3	3	4	0	50
1	2	5	1	0	55	0	0	0	0
2	7	0	1	0	0	25	2	16	0
1	1	0	0	0	0	2	30	0	0
0	7	0	0	0	0	17	0	17	0
0	0	0	1	1	0	0	0	0	8

Output Class

Target Class

Alarm	ACC	TPR	TNR	PPV	NPV
786	0.929	0.854	0.964	0.916	0.934
9001	0.950	0.709	0.973	0.709	0.973
3078	0.953	0.581	0.994	0.915	0.955
9026	0.879	0.128	0.979	0.452	0.894
3072	0.754	0.867	0.740	0.298	0.977
3062	0.980	0.833	0.990	0.847	0.988
9002	0.932	0.222	0.976	0.364	0.953
9010	0.985	0.757	0.994	0.848	0.99
9025	0.948	0.378	0.972	0.359	0.974
3125	0.940	0.138	0.993	0.571	0.945

Confusion matrix of CM-Condition NN									
184	20	38	4	0	21	11	12	4	6
19	17	2	0	0	3	11	7	7	0
4	0	18	0	0	6	0	1	0	0
47	5	28	59	59	3	3	5	0	48
4	6	5	46	46	2	0	0	0	6
0	1	3	0	0	8	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	1	1	0	2

Output Class

Target Class

Alarm	ACC	TPR	TNR	PPV	NPV
786	0.758	0.713	0.779	0.613	0.847
9001	0.895	0.34	0.933	0.258	0.954
3078	0.889	0.191	0.984	0.621	0.899
9026	0.684	0.541	0.707	0.23	0.905
3072	0.837	0.438	0.898	0.4	0.912
3062	0.95	0.186	0.995	0.667	0.955
9002	0.967	0	1.000	0	0.967
9010	0.967	0	1.000	0	0.967
9025	0.986	0	1.000	0	0.986
3125	0.92	0.032	0.996	0.4	0.923

Confusion matrix of Alarms-Condition NN									
277	0	0	0	34	1	0	0	3	0
8	57	0	4	3	2	0	0	11	0
1	1	56	2	31	3	0	0	0	0
1	2	1	35	70	0	0	0	0	0
0	0	0	33	72	0	0	0	0	0
1	0	4	0	3	53	0	0	0	0
6	10	0	0	3	0	3	4	32	0
4	2	1	0	3	0	2	34	0	0
5	6	0	0	0	0	0	0	30	0
0	1	1	0	27	0	0	2	0	0

Output Class

Target Class

Alarm	ACC	TPR	TNR	PPV	NPV
786	0.932	0.914	0.941	0.879	0.959
9001	0.947	0.722	0.968	0.671	0.974
3078	0.952	0.889	0.957	0.596	0.992
9026	0.88	0.473	0.915	0.321	0.953
3072	0.781	0.293	0.953	0.686	0.793
3062	0.985	0.898	0.991	0.869	0.993
9002	0.94	0.6	0.941	0.052	0.998
9010	0.981	0.85	0.987	0.739	0.993
9025	0.94	0.395	0.987	0.732	0.949
3125	0.967	0	0.967	0	1.000

The results show that all the alarms cannot be predicted by the CM&D. Only alarms related to vibrations can be predicted by using the CM-ANN. The main advantage of this ANN is that two different and independent datasets are employed. On one hand, the vibration data generated by the CM and, on the other hand, the condition of the WT that is provided by the SCADA. The ANN is able to correlate data collected by two independent systems.

ANNs can be used to support the decision process when a specific alarm is activated. Table VII shows the probabilities of success detecting each type of false alarm. These probabilities have been calculated from the confusion matrices. The method is accurate in identifying most of them, although some alarms cannot be detected by the three ANNs. The most important parameter in this paper is the precision of the ANN because it indicates the success of the ANN when an alarm is identified. Figure 8 shows, in terms of probability, the precisions, the negative predicted values, the sensitivities and the true negative rates of each ANN for the different alarms shown in x -axe.

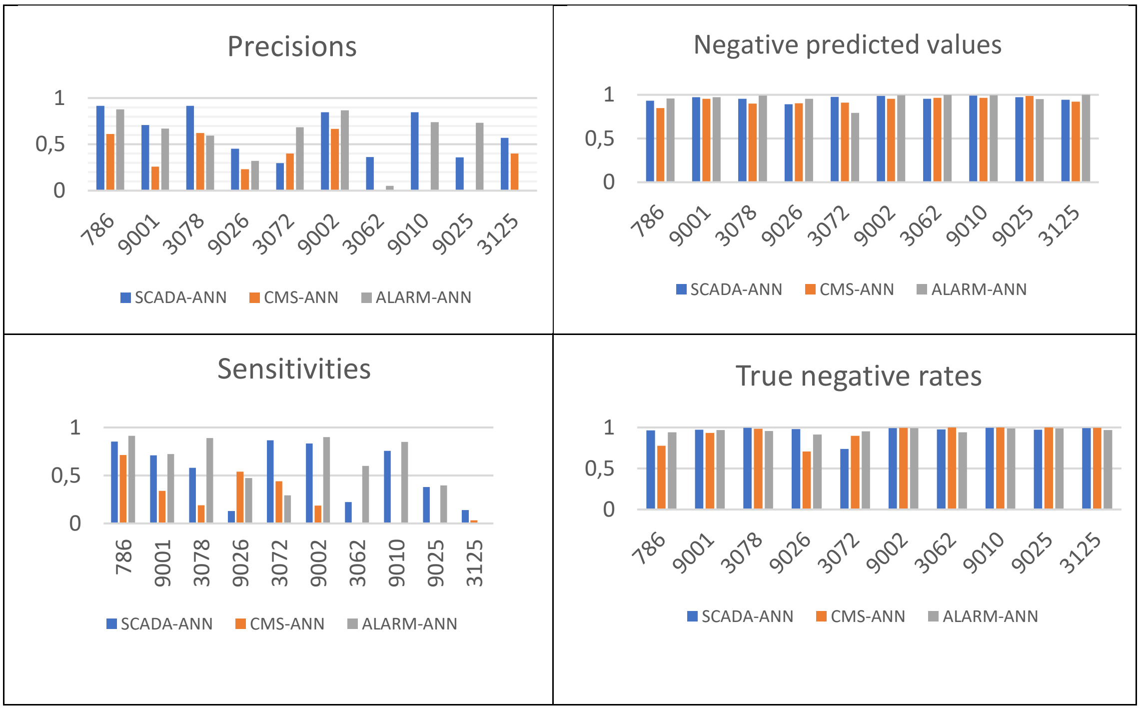

Fig. 8. Results of NN validation

The most suitable ANN will be selected for supporting the decision regarding a specific alarm. The CM-ANN always presents the lowest precision; however, it is employed to obtain the Alarm-ANN, which is in some cases more precise than the SCADA-ANN. The decision should be made considering the best ANN according to the alarm, given by Table VIII:

Table VIII. Best NNs regarding the alarm

Alarm	TPR	TNR	PPV	NPV
786	ALARM-Cond.	SCADA-Cond.	SCADA-Cond.	ALARM-Cond.
9001	ALARM-Cond.	SCADA-Cond.	SCADA-Cond.	ALARM-Cond.
3078	ALARM-Cond.	SCADA-Cond.	SCADA-Cond.	ALARM-Cond.
9026	CM-Cond.	SCADA-Cond.	SCADA-Cond.	ALARM-Cond.
3072	SCADA-Cond.	ALARM-Cond.	ALARM-Cond.	SCADA-Cond.
3062	ALARM-Cond.	CM-Cond.	ALARM-Cond.	ALARM-Cond.
9002	ALARM-Cond.	CM-Cond.	SCADA-Cond.	ALARM-Cond.
9010	ALARM-Cond.	CM-Cond.	SCADA-Cond.	SCADA-Cond.
9025	ALARM-Cond.	CM-Cond.	ALARM-Cond.	CM-Cond.
3125	SCADA-Cond.	CM-Cond.	SCADA-Cond.	ALARM-Cond.

For example, for alarm 9010, the decision will be aided by the sensitivity of the Alarm-NN (0.85), the true negative rate of the CM-NN (1.00), the precision and the negative predicted value of the SCADA-NN (0.84 and 0.99, respectively). A discordance between the activated alarm and the response of the ANN based structure is a statistical indicator of a false alarm. The convergence of the outputs provided by the different ANN means a higher probability of success, and, therefore, a more accurate response.

The results must be employed to feedback these ANNs once the alarm is checked. The data is incorporated into the database in the event of success of the NN-based structure. Otherwise, the structure should be retrained until the correct output for this specific alarm is obtained. This process will adapt the NNs to the real behaviour of the WT, allowing the decision maker to have extra, valuable information.

The outcomes of the method can also be employed to prioritize alarms. If the same alarm is activated from different WTs at the same time, the response of the NN-based structure may be helpful in deciding which WT should be attended to first.

6.2 Validation with Fuzzy Logic.

A Fuzzy logic based methodology is proposed to validate the approach proposed in this paper. It has been employing previously with the SCADA dataset in reference [40]. Figure 9 shows the basic configuration of the fuzzy logic system [81].

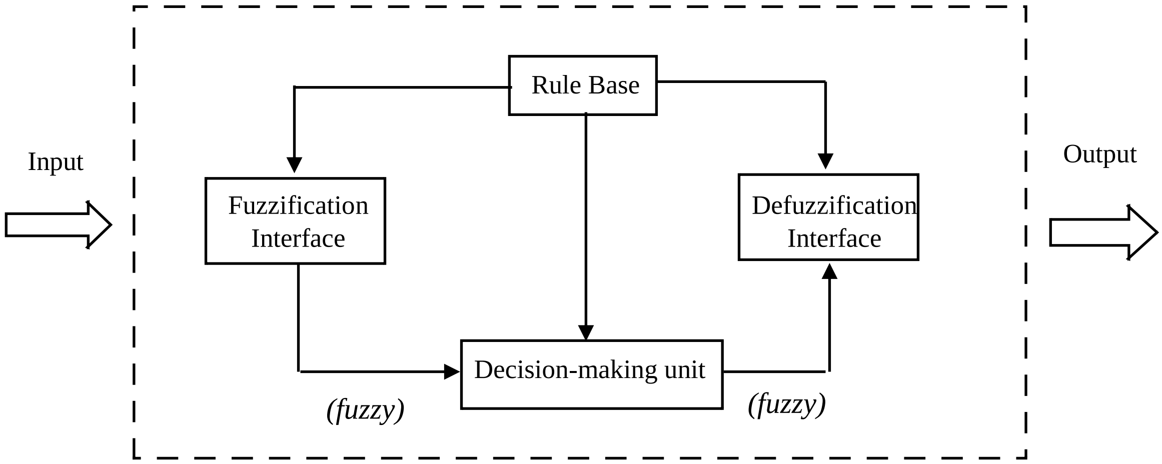

Fig.9: Basic configuration of a fuzzy logic system.

Fuzzy logic is a technique that is associated with the theory of fuzzy sets and the theory of possibilities [82]. The fuzzy system is composed by the fuzzification, fuzzy interference and the defuzzification.

The objective of fuzzification is to convert any numerical value of each input data into a fuzzy subset, i.e. a linguistic value between 0 and 1 [83]. Any fuzzy subset of each input variable requires a membership function whose shape is well defined (sigmoid, hyperbolic tangent, exponential,…) [84]. This paper considers three possible subsets (Good, Acceptable, Unacceptable) for each input variable.

Fuzzy inference is the process of formulating the mapping of input data based on membership functions, fuzzy logic operators (and / or), and If-Then rules at an output using the Fuzzy logic [21]. The Fuzzy rules (IF antecedent THEN consequent) in expert system are usually [85]:

$$IF\ Var\ (1)\ is\ A11\ and/or\ Var\ (2)\ is\ A21 \ldots THEN\ y\ is\ B1$$
$$else$$
$$IF\ Var\ (1)\ is\ A12\ and/or\ Var\ (2)\ is\ A22 \ldots THEN\ y\ is\ B2$$
$$else$$
$$.$$
$$.$$
$$IF\ Var\ (1)\ is\ A1n\ and/or\ Var\ (2)\ is\ A2n \ldots THEN\ y\ is\ Bn$$

where *Var (1)*, *Var (2)*,…, *Var(n)* are the fuzzy input (antecedent) variables, y is a single output (consequent) variable, and *A11* ... *A1n* are the fuzzy sets [10]. The total rules used in the inference system are all possible combinations of the input variables. They depend on the number of linguistic variables that characterize the membership functions of the input data. If, for example, there are n input variables with 3 fuzzy linguistic variables, then the rules are 3^n. There are two main types of fuzzy inference methods [86]: Mamdani-Type and Sugeno-Type. It has been employed in this paper the Sugeno-type, where the Sugeno has an output membership function linear or constant.

Defuzzification is the process of producing a quantifiable result given fuzzy sets and the corresponding membership degrees. There are many types of defuzzification methods, where in this paper has been choose the centroid technique [18][87].

The flowchart of the methodology proposed is shown by Figure 10.

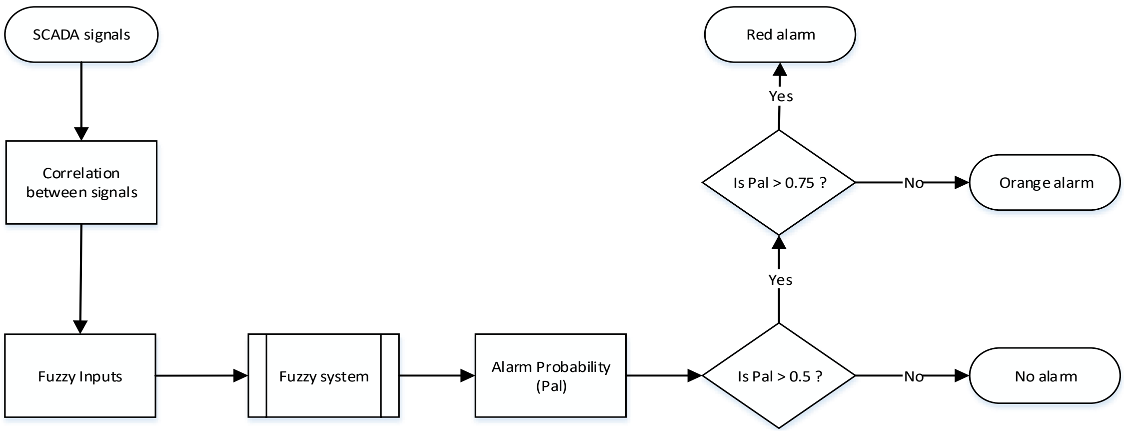

Fig. 10: Alarm identification flowchart.

The large volume of inputs generates a huge number of fuzzy rules and, therefore, fuzzy system is complex. Several techniques can ensure the reduction of this volume such as statistical methods. The Pearson correlation [88] is employed in this paper using a linear correlation between two variables. The correlation coefficient, r, between two discrete variables, x and y, is given by:

$$r = \frac{N(\sum_{n=1}^{N} xy) - (\sum_{n=1}^{N} x)(\sum_{n=1}^{N} y)}{\sqrt{(N\sum_{n=1}^{N} x^2 - (\sum_{n=1}^{N} x)^2)(N\sum_{n=1}^{N} y^2 - (\sum_{n=1}^{N} y)^2)}}$$

The type of correlation can be determined according to the following criteria [76,89]:

- Weak correlation $0.3 \leq |r| < 0.5$
- Moderate correlation : $0.5 \leq |r| < 0.7$
- Strong correlation : $|r| \geq 0.7$
- Perfect correlation $|r| = 1$

The reduction of the fuzzy system inputs is done considering the variables with perfect and strong correlation. The variables that have a strong correlation can be represented by only one variable considering their common behavior.

The fuzzy inference system is based on different rules to generate the occurrence probabilities of the alarms in the output. The output of the fuzzy logic will correspond to three different scenarios [40]:

- *No alarms*: The parameters have values under control and the condition of the WT is correct. The output of the fuzzy logic is less than 0.5.
- *Orange alarms*: Probable faults that do not cause problems for maintenance planning and can be attended by programmed with daily or weekly preventive maintenance tasks. This alarm will be considered when the output of fuzzy system is from 0.5 to 0.75.
- *Red alarms*: Critical states (maximum values for more than one physical variable). It requires diagnosis and urgent intervention to return the status parameters to acceptable levels. This alarm will be considered when the output of fuzzy system is > 0.75.

These parameters can be divided as: the variables related to the condition of the kinematic chain; the parameters related to the condition of safety systems, divided into two subgroups: the pitch control system and the braking safety system.

A total of 76810 inputs have been analyzed through the created fuzzy system. The outcomes were 42.46 % of green alarms, 54.60% of orange alarms and 2.94 % of red alarms

The method is employed to detect false alarms by analyzing the alarms provide by the SCADA system with the alarm probability provide by the fuzzy logic.

A period of 3 months, where the false alarms are known, has been considered to compare the results. The results of the Fuzzy Logic have been compared with the results of the approach presented in this paper in Table IX.

Table IX. Accuracy results given by the NN approach and the Fuzzy Logic (FL)

Alarm	ANN-Approach	FL
786	0.932	0.915
9001	0.950	0.820
3078	0.953	0.724
9026	0.880	0.813
3072	0.837	0.620
3062	0.985	0.905
9002	0.967	0.795
9010	0.967	0.825
9025	0.948	0.813
3125	0.967	0.785

Table IX shows an accuracy for the approach similar to the accuracy found by the confusion matrix. It is better in every alarm than the fuzzy logic accuracy.

7. Conclusions

The detection of false alarms in a wind turbine increases the system reliability. A novel approach based on artificial neural networks has been developed in this paper to detect false alarms and prioritize the alarms. The artificial neural network is composed of three different multilayer perceptrons that analyse the dataset from both a supervisory control and data acquisition system and a condition monitoring system. The dataset is analysed by pattern recognition when it has been filtered. The pattern recognition considers the historical database employed to train the artificial neural networks. The approach can analyse different alarms, the ten most repeated alarms being discussed in this paper. It has been applied to a real dataset from the OPTIMUS European project. Precision and sensitivity are more than 80%, and in some cases more than 90%. The approach accuracy has been studied by confusion matrices, comparing the estimated response of the neural network based structure with real alarms. Finally, fuzzy logic method has been also employed in order to validate the results given by the approach.

Acknowledgements

The work reported herewith has been supported by the Spanish Ministerio de Economía y Competitividad, under Research Grants DPI2015-67264-P

8. References

1. Marugán, A.P.; Márquez, F.P.G.; Perez, J.M.P.; Ruiz-Hernández, D. A survey of artificial neural network in wind energy systems. *Applied Energy* **2018**, *228*, 1822-1836.

2. Pérez, J.M.P.; Márquez, F.P.G.; Hernández, D.R. Economic viability analysis for icing blades detection in wind turbines. *Journal of Cleaner Production* **2016**, *135*, 1150-1160.
3. Council, G.W.E. Global wind statistics 2016, 10 february 2017. Brussels: Global Wind Energy Council (http://www. gwec. net/wp-content/uploads/vip/GWEC-PRstats-2015_LR. pdf: 2016.
4. Council, G.W.E. Global wind 2015 report: Annual market update. *Global Wind Energy Council (GWEC), Brussels, Belgium* **2016**.
5. Marugán, A.P.; Márquez, F.P.G.; Papaelias, M. In *Multivariable analysis for advanced analytics of wind turbine management*, Proceedings of the Tenth International Conference on Management Science and Engineering Management, 2017; Springer: pp 319-328.
6. (EWEA), E.W.e.A. Wind energy's frequently asked questions [online]. *http://www.ewea.org/wind-energy-basics/faq/ (accessed 20.07.2017).* .
7. Esteban, M.D.; Diez, J.J.; López, J.S.; Negro, V. Why offshore wind energy? *Renewable Energy* **2011**, *36*, 444-450.
8. Lloyd, G. Guideline for the certification of condition monitoring systems for wind turbines, 2007. *Hamburg, Germany*.
9. Tavner, P. Offshore wind turbine reliability. **2012**.
10. Pliego Marugán, A.; García Márquez, F.P.; Pinar Pérez, J.M. Optimal maintenance management of offshore wind farms. *Energies* **2016**, *9*, 46.
11. Jiménez, A.A.; Márquez, F.P.G.; Moraleda, V.B.; Muñoz, C.Q.G. Linear and nonlinear features and machine learning for wind turbine blade ice detection and diagnosis. *Renewable Energy* **2019**, *132*, 1034-1048.
12. Márquez, F.P.G.; Tobias, A.M.; Pérez, J.M.P.; Papaelias, M. Condition monitoring of wind turbines: Techniques and methods. *Renewable Energy* **2012**, *46*, 169-178.
13. Papatheou, E.; Dervilis, N.; Maguire, E.; Worden, K. In *Wind turbine structural health monitoring: A short investigation based on scada data*, EWSHM-7th European Workshop on Structural Health Monitoring, 2014.
14. Yang, W.; Tavner, P.J.; Crabtree, C.J.; Feng, Y.; Qiu, Y. Wind turbine condition monitoring: Technical and commercial challenges. *Wind Energy* **2014**, *17*, 673-693.
15. 13379-1:2012, I. Condition monitoring and diagnostics of machines -- data interpretation and diagnostics techniques -- part 1: General guidelines. *ISO* **2012**.
16. Mazidi, P.; Du, M.; Bertling Tjernberg, L.; Sanz Bobi, M.A. A health condition model for wind turbine monitoring through neural networks and proportional hazard models. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* **2017**, 1748006X17707902.
17. García, F.P.; Pedregal, D.J.; Roberts, C. Time series methods applied to failure prediction and detection. *Reliability Engineering & System Safety* **2010**, *95*, 698-703.
18. Brouwer, S.R.; Al-Jibouri, S.H.; Cárdenas, I.C.; Halman, J.I. Towards analysing risks to public safety from wind turbines. *Reliability Engineering & System Safety* **2018**, *180*, 77-87.
19. Reder, M.; Yürüşen, N.Y.; Melero, J.J. Data-driven learning framework for associating weather conditions and wind turbine failures. *Reliability Engineering & System Safety* **2018**, *169*, 554-569.
20. Scheu, M.N.; Kolios, A.; Fischer, T.; Brennan, F. Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability. *Reliability Engineering & System Safety* **2017**, *168*, 28-39.
21. Jiménez, A.A.; Muñoz, C.Q.G.; Márquez, F.P.G. Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers. *Reliability Engineering & System Safety* **2018**.

22. Shafiee, M.; Finkelstein, M.; Bérenguer, C. An opportunistic condition-based maintenance policy for offshore wind turbine blades subjected to degradation and environmental shocks. *Reliability Engineering & System Safety* **2015**, *142*, 463-471.
23. Ghamlouch, H.; Fouladirad, M.; Grall, A. The use of real option in condition-based maintenance scheduling for wind turbines with production and deterioration uncertainties. *Reliability Engineering & System Safety* **2017**.
24. Shafiee, M.; Sørensen, J.D. Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies. *Reliability Engineering & System Safety* **2017**.
25. ANSI, I. Isa-18.2-2009 management of alarm systems for the process industries. *Durham, NC, USA: International Society of Automation* **2009**.
26. Gomez Munoz, C.; De la Hermosa Gonzalez-Carrato, R.; Trapero Arenas, J.; Garcia Marquez, F. A novel approach to fault detection and diagnosis on wind turbines. *Global Nest Journal* **2014**, *16*, 1029-1037.
27. Papaelias, M.; Cheng, L.; Kogia, M.; Mohimi, A.; Kappatos, V.; Selcuk, C.; Constantinou, L.; Muñoz, C.Q.G.; Marquez, F.P.G.; Gan, T.-H. Inspection and structural health monitoring techniques for concentrated solar power plants. *Renewable Energy* **2016**, *85*, 1178-1191.
28. 13372:2012, I. Condition monitoring and diagnostics of machines - vocabulary. *ISO* **2012**.
29. Qiu, Y.; Feng, Y.; Tavner, P.; Richardson, P.; Erdos, G.; Chen, B. Wind turbine scada alarm analysis for improving reliability. *Wind Energy* **2012**, *15*, 951-966.
30. Hameed, Z.; Hong, Y.S.; Cho, Y.M.; Ahn, S.H.; Song, C.K. Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable Energy Reviews* **2009**, *13*, 1-39.
31. Tchakoua, P.; Wamkeue, R.; Ouhrouche, M.; Slaoui-Hasnaoui, F.; Tameghe, T.; Ekemb, G. Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges. *Energies* **2014**, *7*, 2595.
32. Pérez, J.M.P.; Márquez, F.P.G.; Tobias, A.; Papaelias, M. Wind turbine reliability analysis. *Renewable and Sustainable Energy Reviews* **2013**, *23*, 463-472.
33. May, A.; McMillan, D. Condition based maintenance for offshore wind turbines: The effects of false alarms from condition monitoring systems. *ESREL 2013* **2013**.
34. Crabtree, C.J.; Zappalá, D.; Tavner, P.J. Survey of commercially available condition monitoring systems for wind turbines. *Technical Report. Durham University School of Engineering and Computing Sciences and the SUPERGEN Wind Energy Technologies Consortium* **2014**.
35. Chen, B.; Zappalá, D.; Crabtree, C.J.; Tavner, P.J. Survey of commercially available scada data analysis tools for wind turbine health monitoring. *Technical Report. Durham University School of Engineering and Computing Sciences* **2014**.
36. Feng, Y.; Qiu, Y.; Crabtree, C.J.; Long, H.; Tavner, P.J. Monitoring wind turbine gearboxes. *Wind Energy* **2013**, *16*, 728-740.
37. Feng, Y.; Qiu, Y.; Crabtree, C.J.; Long, H.; Tavner, P.J. In *Use of scada and cms signals for failure detection and diagnosis of a wind turbine gearbox*, European Wind Energy Conference and Exhibition 2011, EWEC 2011, 2011; Sheffield: pp 17-19.
38. Dao, P.B.; Staszewski, W.J.; Barszcz, T.; Uhl, T. Condition monitoring and fault detection in wind turbines based on cointegration analysis of scada data. *Renewable Energy* **2018**, *116*, 107-122.
39. Su, H.; Chong, K.T. Induction machine condition monitoring using neural network modeling. *IEEE Transactions on Industrial Electronics* **2007**, *54*, 241-249.
40. Benmessaoud, T.; Marugán, A.P.; Mohammedi, K.; Márquez, F.P.G. In *Fuzzy logic applied to scada systems*, International conference on management science and engineering management, 2017; Springer: pp 749-757.

41. Chacón, A.M.P.; Márquez, F.P.G. False alarms management by data science. In *Data science and digital business*, Springer: 2019; pp 301-316.
42. Kim, K.; Parthasarathy, G.; Uluyol, O.; Foslien, W.; Sheng, S.; Fleming, P. In *Use of scada data for failure detection in wind turbines*, Proceedings of 2011 energy sustainability conference and fuel cell conference, 2011; pp 2071-2079.
43. 2041:2018, I. Mechanical vibration, shock and condition monitoring — vocabulary. *ISO* **2018**.
44. Márquez, F.P.G.; Schmid, F. A digital filter-based approach to the remote condition monitoring of railway turnouts. *Reliability Engineering & System Safety* **2007**, *92*, 830-840.
45. McMillan, D.; Ault, G.W. Quantification of condition monitoring benefit for offshore wind turbines. *Wind Engineering* **2007**, *31*, 267-285.
46. Márquez, F.P.G.; Pérez, J.M.P.; Marugán, A.P.; Papaelias, M. Identification of critical components of wind turbines using fta over the time. *Renewable Energy* **2016**, *87*, 869-883.
47. Marugán, A.P.; Márquez, F.G. System management for remote condition monitoring in railway systems. *6th IET Conference on Railway Condition Monitoring (RCM 2014)* **2014**.
48. de la Hermosa González, R.R.; Márquez, F.P.G.; Dimlaye, V. Maintenance management of wind turbines structures via mfcs and wavelet transforms. *Renewable and Sustainable Energy Reviews* **2015**, *48*, 472-482.
49. Tsai, C.-S.; Hsieh, C.-T.; Huang, S.-J. Enhancement of damage-detection of wind turbine blades via cwt-based approaches. *IEEE Transactions on energy conversion* **2006**, *21*, 776-781.
50. de Azevedo, H.D.M.; Araújo, A.M.; Bouchonneau, N. A review of wind turbine bearing condition monitoring: State of the art and challenges. *Renewable and Sustainable Energy Reviews* **2016**, *56*, 368-379.
51. Hamadache, M.; Lee, D. In *Wind turbine main bearing fault detection via shaft speed signal analysis under constant load*, 2016 16th International Conference on Control, Automation and Systems (ICCAS), 2016; IEEE: pp 1579-1584.
52. Guo, P.; Bai, N. Wind turbine gearbox condition monitoring with aakr and moving window statistic methods. *Energies* **2011**, *4*, 2077-2093.
53. Chen, Z.; Guerrero, J.M.; Blaabjerg, F. A review of the state of the art of power electronics for wind turbines. *IEEE Transactions on power electronics* **2009**, *24*, 1859-1875.
54. García Márquez, F.P.; Quiterio, C.; Muñoz, G.; Lev, B.; Arcos Jimenez, A. New pipe notch detection and location method for short distances employing ultrasonic guided waves. **2017**.
55. 17359:2018, I. Condition monitoring and diagnostics of machines — general guidelines. *ISO* **2018**.
56. de da González-Carrato, R.R.; Márquez, G.; Pedro, F.; Papaelias, M. Vibration-based tools for the optimisation of large-scale industrial wind turbine devices. *International Journal of Condition Monitoring* **2016**, *6*, 33-37.
57. Arcos Jiménez, A.; Gómez Muñoz, C.Q.; García Márquez, F.P. Machine learning for wind turbine blades maintenance management. *Energies* **2017**, *11*, 13.
58. Muñoz, C.Q.G.; Jiménez, A.A.; Márquez, F.P.G. Wavelet transforms and pattern recognition on ultrasonic guides waves for frozen surface state diagnosis. *Renewable Energy* **2018**, *116*, 42-54.
59. Muñoz, C.Q.G.; Marquez, F.P.G.; Lev, B.; Arcos, A. New pipe notch detection and location method for short distances employing ultrasonic guided waves. *Acta Acustica United with Acustica* **2017**, *103*, 772-781.

60. Hamadache, M.; Lee, D.; Veluvolu, K.C. Rotor speed-based bearing fault diagnosis (rsb-bfd) under variable speed and constant load. *IEEE Transactions on Industrial Electronics* **2015**, *62*, 6486-6495.
61. 10816:2014, I. Vibration monitoring non rotating machines. **2014**.
62. Hameed, Z.; Hong, Y.; Cho, Y.; Ahn, S.; Song, C. Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable energy reviews* **2009**, *13*, 1-39.
63. Mylaraswamy, D.; Olson, L.; Nwadiogbu, E. In *Engine performance trending*, AIAC12-HUMS Conference, 2007.
64. Kusiak, A.; Li, W. The prediction and diagnosis of wind turbine faults. *Renewable Energy* **2011**, *36*, 16-23.
65. Song, Y.; Dhinakaran, B.; Bao, X. Variable speed control of wind turbines using nonlinear and adaptive algorithms. *Journal of Wind Engineering and Industrial Aerodynamics* **2000**, *85*, 293-308.
66. Song, Z.; Jiang, Y.; Zhang, Z. Short-term wind speed forecasting with markov-switching model. *Applied Energy* **2014**, *130*, 103-112.
67. Zeng, Z.; Tao, N.; Feng, L.; Li, Y.; Ma, Y.; Zhang, C. Breakpoint detection of heating wire in wind blade moulds using infrared thermography. *Infrared Physics & Technology* **2014**, *64*, 73-78.
68. Zhang, H.; Yu, Y.-J.; Liu, Z.-Y. Study on the maximum entropy principle applied to the annual wind speed probability distribution: A case study for observations of intertidal zone anemometer towers of rudong in east china sea. *Applied Energy* **2014**, *114*, 931-938.
69. Yang, W.; Tavner, P.J.; Crabtree, C.J.; Wilkinson, M. Cost-effective condition monitoring for wind turbines. *Industrial Electronics, IEEE Transactions on* **2010**, *57*, 263-271.
70. Wymore, M.L.; Van Dam, J.E.; Ceylan, H.; Qiao, D. A survey of health monitoring systems for wind turbines. *Renewable and Sustainable Energy Reviews* **2015**, *52*, 976-990.
71. Pliego Marugán, A.; García Márquez, F.P.; Lev, B. Optimal decision-making via binary decision diagrams for investments under a risky environment. *International Journal of Production Research* **2017**, 1-16.
72. 13374:2015, I. Condition monitoring and diagnostics of machine systems -- data processing, communication and presentation. *ISO* **2015**.
73. 19501, I.I. Information technology -- open distributed processing -- unified modeling language (uml) *ISO* **2005**.
74. 13373:2002, I. Condition monitoring and diagnostics of machines -- vibration condition monitoring. *ISO* **2002**.
75. Project, E.R. Demonstration of methods and tools for the optimisation of operational reliability of large-scale industrial wind turbines, optimus project. . **2014**.
76. Marugán, A.P.; Márquez, F.P.G. In *Scada and artificial neural networks for maintenance management*, International Conference on Management Science and Engineering Management, 2017; Springer: pp 912-919.
77. Niculescu, S.P. Artificial neural networks and genetic algorithms in qsar. *Journal of Molecular Structure: THEOCHEM* **2003**, *622*, 71-83.
78. Bengio, Y. Practical recommendations for gradient-based training of deep architectures. In *Neural networks: Tricks of the trade*, Springer: 2012; pp 437-478.
79. Sammut, C.; Webb, G.I. *Encyclopedia of machine learning*. Springer Science & Business Media: 2011.
80. Fawcett, T. An introduction to roc analysis. *Pattern recognition letters* **2006**, *27*, 861-874.

81. Abreu, G.L.; Ribeiro, J.F. On-line control of a flexible beam using adaptive fuzzy controller and piezoelectric actuators. *Sba: Controle & Automação Sociedade Brasileira de Automatica* **2003**, *14*, 377-383.
82. Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems* **1999**, *100*, 9-34.
83. Jamshidi, M. Tools for intelligent control: Fuzzy controllers, neural networks and genetic algorithms. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **2003**, *361*, 1781-1808.
84. Soria-Olivas, E.; Martín-Guerrero, J.D.; Camps-Valls, G.; Serrano-López, A.J.; Calpe-Maravilla, J.; Gómez-Chova, L. A low-complexity fuzzy activation function for artificial neural networks. *IEEE Transactions on Neural Networks* **2003**, *14*, 1576-1579.
85. Khanna, V.; Cheema, R.K. Fire detection mechanism using fuzzy logic. *International Journal of Computer Applications* **2013**, *65*.
86. Gomez, C.; Marquez, F.P.G.; Arcos, A.; Cheng, L.; Kogia, M.; Mohimi, A.; Papaelias, M. A heuristic method for detecting and locating faults employing electromagnetic acoustic transducers. *Eksploatacja i Niezawodnosc - Maintenance and Reliability* **2017**, *19*, 493-500.
87. Zhang, P. *Advanced industrial control technology*. William Andrew: 2010.
88. MARUGÁN, A.P. Approches for qualitative and quantitative analysis of complex systems: Algorithms and case studies. . Castilla La Mancha, Ciudad Real, July 2016.
89. Schlechtingen, M.; Santos, I.F.; Achiche, S. Wind turbine condition monitoring based on scada data using normal behavior models. Part 1: System description. *Applied Soft Computing* **2013**, *13*, 259-270.