A Literature Review of Time-Series Clustering Techniques and Machine Learning Techniques Used for Monitoring of Wind Turbines

Written by
Yohann Jacob Sandvik

Abstract

Here the abstract will be.

Contents

Chapter 1

Introduction

Clustering is a method used to categorize big amounts of data into groups known as *clusters*, when there is little, or no information available about the underlying groups. It is a popular choice for extracting patterns from large datasets, because clustering falls within the category of *unsupervised machine learning*, meaning that it does not require the dataset to be labeled. Clustering is also a common step in data mining algorithms, where the goal is to learn rules relating the different variables in a dataset. As of year 2000(?) clustering started to become more common to use on time-series datasets as they are abundunt, and labeling is often time-costly. Time-series clustering has been applied on financial time series, medical time series and time series from a variety of other industries.

1.1 Motivation

As of 2018 wind power, together with solar power made up 7% of the worlds electricity production,[1] and has been referred to as "the fastest growing source of energy" by the Norwegian company Statkraft[2]. As the effects of climate change steadily are becoming a reality shifting to renewable energy sources is imperative, and wind power will certainly play a bigger part in meeting the worlds energy demand in the future.

To make wind power as a whole more lucrative, a good start would be to reduce the downtime, and improve the performance of turbines. The argument that time-series clustering may be a good approach for this is two-fold.

1. A single wind turbine can have several houndred sensors sampling up to every second, meaning that a wind farm can produce colossal amounts of time-series data. An unsupervised approach is usefull simply because labelling of all this data is cumbersome.
2. When wind farms become big enough it will become to costly to manually inspect every turbine to construct an effective model for condition monitoring, further automation is required [1]. Time-series clustering is then a good alternative for condition monitoring.

1.2 Objective

This literature review has three objectives in the form of questions.

[1] https://www.iea.org/geco/electricity/
[2] https://www.statkraft.com/globalassets/old-contains-the-old-folder-structure/documents/wind-power-aug-2010-eng_tcm9-11473.pdf

Objectives

1. What machine learning methods are currently being used to monitor the condition, and performance of wind turbines?
2. What are the different methods of time-series clustering currently used?
3. What time-series clustering methods (if any) are appropriate to test on time-series data produced by wind turbines?

The literature review is meant to be a preliminary work for a master thesis where select techniques will be evaluated on actual time series data produced by a wind farm in Norway. The project assigment is also a continuation of the master thesis written in the spring of 2019 by Espen Waaga,

1.3 Structure of Review

Chapter 2

Method

2.1 Search terms

To find the relevant literature on the subjects of interest the search engine Oria was used to search the university library of the NTNU. Oria was preferred over other search engines such as Google Scholar because Oria allows one to combine multiple search terms in unison using "AND" or "OR", and because it allows the user to choose whether the search term should be in the title, subject, or other parts of the articles. The review will only consider articles published in peer-reviewed journals. Table 2.1 summarizes the search results. The *Title* and *General content* columns show which terms were used in the different searches; which terms where required to be in the title, and which terms could be in the "general content", meaning any part of the article. Let "×" represent the AND operator between two search terms, and "∧" represent the OR operator. The "*" operator means that the search will include any word beginning with the word before the star, e.g *detect** includes *detection*, *detecting*, *detected*, etc. The N_f and N_i columns show how many results each search yielded and how many articles from each search were included in the review, respectively.

The last three rows in table 2.1 are somewhat disconnected from the rest of the table. The term "forward snowballing" refers to finding other articles by finding articles citing a specific source, and the term "backward snowballing" refers to finding articles by going through the references of a particular source.

Nr.	Title terms	General terms	N_r	N_i
1	time × series × clustering	None	219	121
2	wind × turbine* × (monitor* ∧ detect*) × review	None	32	3
3	wind × turbine* × (monitor* ∧ detect*)	machine × learning	100	48
4	(monitor* ∧ detect*) × clustering	time × series	187	21
			Articles included through backward snowballing	0
			Articles included through forward snowballing	0
			Total number of articles included	193

Table 2.1: Search results

2.2 Screening method

To make sure that the articles used were relevant, the review is limited to articles published 2014 or later. Some older articles are included through backward snowballing for their historical importance. There were three levels of screening, screening of the title, abstract, and full article. Title-screening was primarily for seeding out duplicate articles returned from the search-engine. The screening of the abstract and full-article were to identify the articles that were not relevant for the review and exclude them.

It has been a challenge to include enough literature to get a good overview of the different methods within time series clustering, but also not more literature the author alone could review in the time available. So although the objectives is to get an overview of the different time series clustering methods, and an overview of the current machine learning methods used for monitoring wind turbines, the author does not claim to have made a complete exhaustive summary of all the possible methods.

When screening articles from search one, articles regarding time-series forecasting, articles only using clustering as a preprocessing step, articles working on regression or classification, and articles using image time series were not included. The scope is restricted to work focusing on data mining using time-series clustering techniques.

Search number two was used to find existing literature reviews on methods used for monitoring of wind turbines. Three good literature reviews on the subject where found. One good literature review was on machine learning methods used for condition monitoring of wind turbines was found in search three as well, so when screening the remaining articles from search number three the objective was to get a superficial overview over the different machine-learning techniques applied, find relevant articles not included in the literature reviews, and find work done after the reviews were published.

For the fourth search some of the results overlapped with the results from the first search. Articles not using time-series data where discarded which is the main reason that so many of the results have not been used.

Chapter 3

Time Series Clustering

Time-series clustering (TSC) can be divided into three components. How the time series is represented, which metric is used to measure similarity between time series, and what algorithm is used to cluster a set of time series. This chapter will give a short description of what specific type of TSC that this report will consider, it will go through the aforementioned steps of a TSC system, give an in-depth description of some time-series models and finally show some metrics used to evaluate a TSC system.
In this report we will deal with *discrete time series*. A time series is defined as a set of observations $\{x_t\}$ recorded at a specific time t. A discrete time series is a time series where the set of times when observations are made (T_0) is discrete [2]. A multivariate time series can be viewed as a set of vectors $\{\mathbf{x}_t\}$ where each set of vector elements $\{x_t^i\}$ is an individual time series. This means that the elements of the same vector $[x_t^1, x_t^2, ..., x_t^N]$ are separate observations made at the same time instance t. In a wind turbine, measurements of the temperature in the gear box made every 10 seconds can be considered a univariate time series. While the set of measurements made every 10 seconds of the temperature in the gear box, the power produced by the turbine, and the wind speed ahead of the blades can be considered a multivariate time series.

3.1 Types of Times-Series Clustering

There are three types of TSC, *whole-series TSC*, *subsequence TSC* and *time-point TSC*. Whole-series TSC is when multiple "whole" time series are clustered with respect to each other. Subsequence TSC comprises the clustering of subsequences of the same time series with respect to each other. The defining difference between whole-series and subsequence TSC is that whole-series TSC clusters multiple time series while subsequence TSC clusters clusters different subsequences of the same time series. In time-point clustering one clusters individual data points from multiple time-series with the same timestamp with respect to each other. One can consider it as regular clustering on the datapoints of a time series at a single time-instance.

3.2 Representation Methods

There are numerous ways which a time series can be represented. Aghabozorgi, Shirkhorshidi, and Wah [3] define a time series representation given time series data $\{x_t\} = \{x_1, x_2, ..., x_T\}$ as transforming the time series into another vector $\{x_t\} = \{\hat{x}_1, \hat{x}_2, ..., \hat{x}_L\}$ where $L < T$. In theory $L = T$, but most often the point in transforming a time series is to reduce the amount of information present in the raw time series to more easily reveal patterns of interest. According to Aghabozorgi, Shirkhorshidi, and Wah [3] and Philippe Esling [4] representation methods can broadly be categorized into four categories:

- **Non-data adaptive**: Here the parameters of the transformation are the same for every time series [4]. Spectral transformations such as the Discrete Fourier Transform (DFT) and the Discrete Wavelet Transform (DWT) fall within this category.
- **Data adaptive**: The name here implies that the parameters of the transformation such as the length of the time series can change for every specific time series [4]. These methods focus on approximating a given time series in the best possible manner by reducing the global reconstruction error [3].
- **Model based**: These representations attempt to fit a stochastic model to the data [3]. Examples of models include Autoregressive Moving Average (ARMA) models, and Hidden Markov Models (HMM).
- **Data dictated**: This representation method differs from the aforementioned methods in as much that it automatically chooses the compression ratio based on the characteristics of the raw time series [3].

Data adaptive representations will be better at approximating time series than non-adaptive methods, but since transformation parameters can change for different time series it is harder to compare time series [3]. What is worth noting about model based representation methods is that they make assumptions about the underlying process that is generating the time series data [4]. Hence, it can be a good way of integrating a priori knowledge about the time series into the clustering system. In the sub-section below there is a more in-depth description of the two common time series models ARMA models, and HMM.

3.3 Time Series Models

To select suitable mathematical models for a dataset, we have to allow for the random nature of future observations. This is done by assuming that each observation in a time series x_t is a realization of a particular random variable X_t. The time series can then be modelled as a collection/set of random variables $\{X_t\}$, also known as a *stochastic process* [2].

To define an ARMA model, one needs to have a clear understanding of the terms white noise process, and stationary process. We say that the stochastic process $\{Z_t\}$ is "white noise" with zero mean, and variance σ^2 ($\{Z_t\} \sim WN(0, \sigma^2)$) if and only if $\{Z_t\}$ is zero mean, and every random variable contained in $\{Z_t\}$ is uncorrelated with every other random variable contained in $\{Z_t\}$. A stochastic process $\{X_t\}$ is said to be weakly wide-sense stationary if the mean, and variance are constant for all terms in the process. $\{X_t\}$ is said to be weakly short-term stationary if the mean and variance of terms are constant for distinct time periods within the duration of the process, but are not constant for all terms in the process. For brevity the term "stationary process" will be used when refering to a *weakly wide-sense stationary process*.

3.3.1 Autoregressive Moving Average Models

An ARMA model descirbes a time series in terms of difference equations. It can be considered a combination of two smaller models, an autoregressive (AR) model and a moving average (MA) model. Let $\{X_t\}$ be a stationary process. An MA(q) model will describe every term X_t as a linear combination of q distinct white noise terms as in equation (3.1).

$$X_t = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q} \tag{3.1}$$

Whereas an AR(p) model will describe every term X_t as a linear combination of p previous terms of $\{X_t\}$ as in equation (3.2)

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} \tag{3.2}$$

Putting equations (3.2) and (3.1) together, an ARMA(p, q) model will describe every term X_t as a linear combination of p previous terms, and q white noise terms as in equation (3.3).

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q} \tag{3.3}$$

Given that the polynomials $1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q$ and $1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p$ have no common factors [5].

3.3.2 Hidden Markov Models

Let $\{X_n\}$ be a stochastic process where the random variables contained in $\{X_n\}$ only can take on a finite number of values which we will call states. Let X_n denote the state at time period n. The probability of X_n transitioning from state i to state j at the next time period $n+1$ is denoted p_{ij}. It might seem natural that p_{ij} is conditional on what the state was in the last time period. $\{X_t\}$ is said to be a *Markov chain* if p_{ij} only is conditional on the last past state, as shown in equation (3.4)

$$\begin{aligned} p_{ij} &= P(X_n = i | X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, ..., X_1 = i_1, X_0 = i_0) \\ &= P(X_n = i | X_{n-1} = i_{n-1}) \end{aligned} \tag{3.4}$$

Suppose now that the states that the process is in are hidden from the observer. Instead there exists a finite set of signals $\{S\}$ that are emmitted when the process enters a state. In addition, let the probability of emmitting signal s, at time period n, in state j ($P(S_n = s | X_n = j)$) be independent of previous states, and signals emmitted. A model of this type where the signals $S_1, S_2, ...$ are observed, and the underlying Markov states remain hidden is called a *hidden Markov chain model* [6].

3.4 Similarity Metrics

3.5 Clustering Algorithms

Brief introduction to clustering techniques. Go through how the different algorithms cluster time series based on *similarity*

3.6 Evaluation Indices

Chapter 4

Wind Turbine Monitoring

4.1 Wind turbine Components

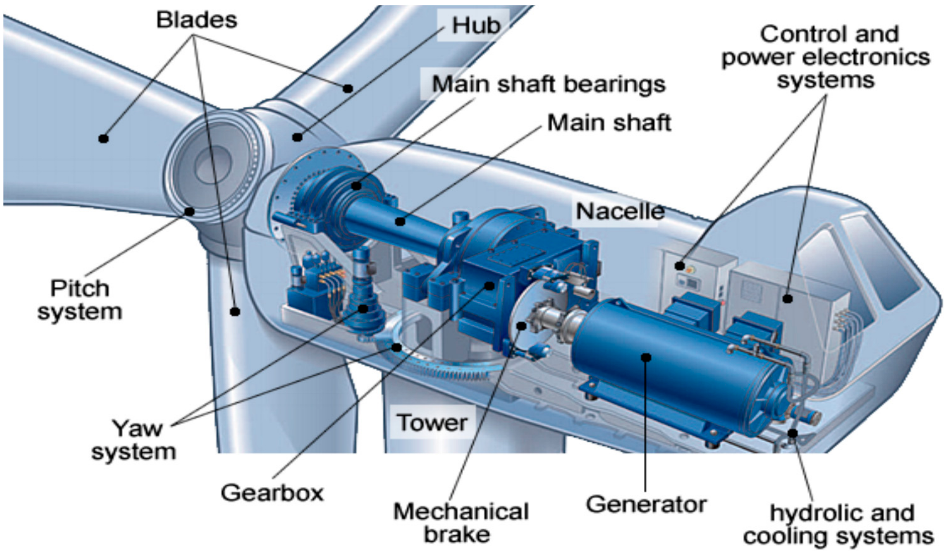

Figure 4.1: Illustration of the different parts of a wind turbine, taken from [7]

Figure 4.1 shows the main parts of a wind turbine which includes the rotor (blades and hub), shafts, gearbox and generator. Simplified a wind turbine works by wind pushing the blades, generating torque that makes the hub rotate. The hub is connected to the gearbox through the main shaft. The gearbox then gears down the torque and gears up the rotational speed to a level that the generator can use to induce current, that goes to a station that transforms the voltage to a level that can be used in the electrical grid.

4.2 Sensors and Data Acquisition

Information about a wind turbine can come from many sources, it can come from external sources such as images from a camera, or from internal sensors measuring operational data. The collective term for systems measuring operational data is supervisory control and data acquisition (SCADA) systems. To choose what algorithm to use, or what model to use, one must first consider what data one has available. From the literature considered, these where the most used forms of data used as input for the model.

- Vibration measurement
- Acustic emission monitoring
- Temperature measurement
- Power signal measurement
- Oil debris monitoring
- Strain monitoring
- Optical fiber monitoring
- Ultrasonic testing
- Image analysis

Analysis of vibration signals is the most common form of condition monitoring used in industry for any form of rotating equipment [8]. By measuring the acustic emission generated by a component of a wind turbine, one can estimate how much damage it has obtained. The temperature of components in a wind turbine is closely correlated with the health of the component, and is therefore used often in condition monitoring applications [9]. The power signal can also say a lot about how well a wind turbine is performing, specifically the wind speed - power curve. When monitoring the debris in the oil of a wind turbine gearbox one is analysing the size, type and number of wear particles present in the lubricant, as they can indicate the degree of damage in the gearbox [10]. Strain monitoring, optical fiber monitoring, ultrasonic testing, and image analysis are all used to detect structural damage in different components of the wind turbine, usually the blades [11, 12, 13, 14, 15, 16], or tower [17]. The most common approach however was to use a combination of multiple sensors-values to make predictions about the condition about the wind turbine.

4.3 Machine learning techniques

A machine learning is a subset of artificial intelligence. Machine learning models that extract rules from data, which can then be applied to classify, or estimate components of another dataset. Machine learning algorithms are formally divided into *supervised learning*, *unsupervised learning* and *semi-supervised learning*. Supervised learning models require labelled datasets to extract information from the dataset, and are usually used to perform classification tasks, or to estimate a variable that is considered dependent on the input variables (regression). Unsupervised learning algorithms do not require labelled datasets. Semi-supervised learning uses a combination of labelled and unlabelled datasets.

4.3.1 Feature extraction

There are two central problems in condition monitoring that can be solved by feature extraction and selection. The first is the sheer volume of information being produced. A wind turbine with only 20 sensors, sampled at 100 Hz will produce 170 MB of information per day. Feature selection is used here to reduce the number of features to only those relevant for condition monitoring. The second problem is that for systems using only one signal such as vibration, there are many components that are superposed to create the signal that is visible, and there might be noise present. Feature extraction is used to separate the interesting components from each other, and cancel the noise. For some machine learning models feature extraction is not

Articles	Extraction method
ARMA models	[18, 19, 11, 14, 20]
Discrete Wavelet Transform (DWT)	[19, 21, 22, 23]
Principal Component analysis (PCA)	[11, 24, 14, 23, 16, 25, 26]
Basic signal statistics	[27, 23, 28]
Neural networks (NN)	[29, 30, 12, 16]

Table 4.1: Feature extraction and selection methods

a neccesary preprocessing step, for others careful though must be given as to how to extract features. Table 4.1 shows the most frequent methods found in the articles.
Time-requency domain analysis [19, 27]

4.3.2 Supervised learning models

The models that this review considers as supervised learning models, are those only given a dataset which is labelled with fault occurance. The models included in this category are mainly (only?) classification models.

4.3.3 Semi-supervised learning models

Chapter 5

Discussion

Chapter 6

Conclusion

Chapter 7

Questions for Pierluigi and Gerthory

Short questions

- this is a short question

Chapter 8

Appendices

8.1 Machine learning methods for wind turbine condition monitoring

Nr	Input	Feature extraction method	Machine learning model
[18]	Vibration signal.	ARMA model and J48 decision tree	Tests a set of (38) meta-, misc-, rule- and tree-based classifiers for FD in blades.
[31]	SCADA data.		Deep auto-encoder made of RBMs to model normal behaviour of SCADA variables. Uses reconstruction error for AD.
[32]	SCADA data.	Spatiotemporal pattern network	Unsupervised use of RBMs for AD.
[19]	SCADA data.	Time-frequency domain analysis, DWT and ARMA model	Uses fusion of MLP, RBF, decision tree, KNN, SVM for FD of wind turbine.
[11]	Ultrasonic testing.	Tests linear and non-linear PCA and ARMA models	Neighbourhood component analysis for feature selection. Tests 20 different supervised classifiers for detecting ice on blades.
[33]	Wind-Power curve.		Uses Gaussian Process regression with extreme fucntion theory to determine whether a particular power curve is an outlier.
[27]	Acustic emission.	FFT.	Uses Distinguishability Measure for feature selection, and logistic regression and SVM for binary blade fault classification.
[34]	Power signal, wind speed and ambient temperature signals.		Hierarchical Extreme Learning Machine
[35]	SCADA data.		Gaussian processes regression to estimate wind-power curve

[36]	SCADA data.	Tests KNN, random forest, and SVR to estimate power curve. Detects anomalies by reconstruction error.	
[37]	Vibration signals.		Uses unsupervised dictionary learning extracting features which are then used to determine fault in drivetrain bearings.
[38]	Oil temperature, wind speed and rotor speed.		Trains ANN to estimate vibration signal, uses reconstruction error for AD.
[24]	SCADA data.	PCA.	Sets up baseline model using multiway PCA, then finds outliers by hypothesis testing whether multivariate distribution is equal to baseline.
[21]	FAST wind turbine simulator.	Image texture analysis tools.	KNN, Linear Discriminant Model, decision trees, bag-tree, linear SVM.
[39]	SCADA data.	K-means for outlier elimination.	Uses Auto-Associative Neural Networks as an auto-encoder, and the Hotel T2 statistic as a dynamic threshold for the reconstruction error.
[40]	SCADA data.	Grey correlation algorithm for eigencevtor extraction.	Use genetic algorithm for feature selection, and SVR for estimating performance curves.
[41]	Rotor speed, gearbox temperature and generator temperature.		Uses three NN for normal behaviour modelling. Reconstruction error sendt to proportional hazard model which sets dynamic threshold.
[29]	SCADA data.		Uses Inductive Transfer Learning and five differen ML classifiers for ice detection on blades.
[30]	Vibration signals.	Variational mode decomposition (VMD).	Uses multi-scale permutation entropy (MPE) used for feature selection, COVAL for domain normalization and an SVM for binary fault classification.
[42]	SCADA data.		Uses bins and SVR to estimate blade angle pitch curve.
[43]	SCADA data.		Tests different architectures of ANNs to for estimating temperature of non-drive end bearing. Uses reconstruction error for anomaly detection.
[12]	Images taken by drones		Recurrent Convolutional Neural Network to classify structural damage in blades.
[13]	Uses images taken from ground level		Convolutional Neural Network, and YOLO-based small object detection approach (YSODA) for damage detection in blades and hub.

[22]	Vibration signals, acustic emission and oil particle analysis	DWT.	Uses decision tree for feature selection, and SVC for assessing fault severity in gearbox
[44]	Uses images taken from ground level		Convolutional Neural Network to detect cracks & damage in blades.
[14]	Ultrasonic testing	PCA and ARMA models.	Neighbourhood Component analysis for feature selection and an ensamble of KNN, linear SVM, decision trees, LDA and subspace discriminant to estimate amount of dirt and mud on blades.
[45]	Uses pitch position, rotor speed and generator speed.		Detects faults with an SVC with parameters optimized by Cuckoo-swarm optimization.
[20]	Vibration signals	ARMA model.	Dominating features selected with J48 decision tree, fault classification done with Bayesian- and lazy classifiers.
[23]	Vibration signals, acustic emission and oil particle analysis	DWT and PCA.	Dominating features selected with decision tree, fault detection done with SVC.
[15]	Images taken from imaging array		Uses a deep neural network for binary classification of blade defects.
[16]	Uses images taken by drone	Uses a CNN trained on an unrelated image dataset to extract general features.	Compress features with PCA, and pass them to a unsupervised one-class SVM.
[46]			This is a literary review of vibration based condition monitoring and fault diagnosis of planetary gearboxes in wind turbines.
[47]	Uses the FAST wind turbine simulator to get SCADA data.	Random forest.	Uses XGBoost to train an ensamble of classifiers for specific faults.
[48]	SCADA data.		Uses several ANN to build a normal behaviour model, then uses reconstruction error together with the age of the age of the turbine to predict anomolous behaviour.
[49]	SCADA data.		Uses Pearson product-moment rank correlation to select features, and applies different ANN structures to predict the active power.
[50]	SCADA data from a simulink model of a wind turbine.		Uses Deep Belief Networks for detecting anomolous behaviour.

[51]	SCADA data.		Uses kolmogorov-smirnov test to compare different turbines at same moment in time combined with the reconstruction error of the gearbox bearing temperature of an ANN to detect anomalies.
[52]	Vibration signals.	Approximates vibration distributions at different rotor speeds with Weibull distribution.	Uses a HMM for statistical fault detection.
[53]	SCADA data.		Compares linear models, ANNs and state-dependent parameter models for fault detection.
[26]	SCADA data.	parallel factor analysis (PARAFAC) as a decomposition method	uses K-means clustering after decomposition for fault detection.
[25]	Uses a wind turbine simulator for SCADA data.		Multiple PCA models are as a statistical reference reflecting the data variability in local zones and used in parallel for online fault detection.
[28]	Vibration signals.	Variational mode decomposition	Uses Fisher score and ReliefF algorithm for feature selection. Feeds selected signals into a multi-class SVC for bearing fault detection.
[54]	SCADA data.		Uses deep neural networks for detection of icing on the blades.
[55]	Wind speed, ambient temperature and air density.		Uses a series of ARMA models (ARIMA, GAS, GASX, NLGASX) to predict the potential for energy production at a geographical location.
[56]	SCADA data.		Combines NN with alarms generated by SCADA system to reduce false alarm rate.

Table 8.1: Summary of machine learning methods for wind turbine condition monitoring

Bibliography

[1] Espen Waaga. *Machine Learning for Automatic Classification of Wind Turbines*. English. Master thesis. 2019.

[2] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods Second Edition*. Springer-Verlag New York, Inc., 1991.

[3] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. "Time-series clustering - A decade review". eng. In: *Information Systems* 53 (2015), p. 16. ISSN: 0306-4379.

[4] Carlos Agon Philippe Esling. "Time-series data mining". In: *ACM Cumputing Surveys* 45 (2012).

[5] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting Third Edition*. Springer International Publishing Switzerland, 2016.

[6] Sheldon M. Ross. *Introduction to Probability Models 11th Edition*. Academic Press Elsevier Inc., 2014.

[7] Ml Hossain, A Abu-Siada, and SM Muyeen. "Methods for Advanced Wind Turbine Condition Monitoring and Early Diagnosis: A Literature Review". English. In: *Energies* 11.5 (2018). ISSN: 1996-1073.

[8] Henrique Dias Machado de Azevedo, Alex Maurício Araújo, and Nadège Bouchonneau. "A review of wind turbine bearing condition monitoring: State of the art and challenges". English. In: *Renewable and Sustainable Energy Reviews* 56 (2016), pp. 368–379. ISSN: 1364-0321.

[9] H. Wang et al. "Early fault detection of wind turbines based on operational condition clustering and optimized deep belief network modeling". In: *Energies* 12.6 (2019). ISSN: 19961073.

[10] P. Qian et al. "A novel condition monitoring method of wind turbines based on long short-term memory neural network". English. In: *Energies* 12.18 (2019). ISSN: 19961073.

[11] Alfredo Arcos Jiménez et al. "Linear and nonlinear features and machine learning for wind turbine blade ice detection and diagnosis". English. In: *Renewable Energy* 132 (2019), pp. 1034–1048. ISSN: 0960-1481.

[12] ASM Shihavuddin et al. "Wind Turbine Surface Damage Detection by Deep Learning Aided Drone Inspection Analysis". English. In: *Energies* 12.4 (2019), p. 676. ISSN: 1996-1073.

[13] Zifeng Qiu et al. "Automatic visual defects inspection of wind turbine blades via YOLO-based small object detection approach". English. In: *Journal of Electronic Imaging* 28.4 (2019), pp. 043023–043023. ISSN: 1017-9909.

[14] Alfredo Arcos Jiménez, Carlos Quiterio Gómez Muñoz, and Fausto Pedro García Márquez. "Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers". English. In: *Reliability Engineering and System Safety* 184 (2019), pp. 2–12. ISSN: 0951-8320.

[15] Ningning Zhang, Chengzhi Lu, and Anmin Wang. "Study on wind turbine blade defect detection system based on imaging array". English. In: *E3S Web of Conferences* 118 (2019). ISSN: 25550403. URL: http://search.proquest.com/docview/2301959230/.

[16] Yinan Wang et al. "Unsupervised anomaly detection with compact deep features for wind turbine blade images taken by a drone". English. In: *IPSJ Transactions on Computer Vision and Applications* 11.1 (2019), pp. 1–7. ISSN: 1882-6695.

[17] Pierre Tchakoua et al. "Wind Turbine Condition Monitoring: State-of-the-Art Review, New Trends, and Future Challenges". English. In: *Energies* 7.4 (2014), pp. 2595–2630. ISSN: 19961073. URL: http://search.proquest.com/docview/1537086932/.

[18] A. Joshuva; V. Sugumaran. "A machine learning approach for condition monitoring of wind turbine blade using autoregressive moving average (ARMA) features through vibration signals: a comparative study". English. In: *Progress in Industrial Ecology, An Int. J.* 12.1/2 (2018). ISSN: 1476-8917. URL: http://www.inderscience.com/link.php?id=95867.

[19] Vahid Pashazadeh, Farzad R Salmasi, and Babak N Araabi. "Data driven sensor and actuator fault detection and isolation in wind turbine using classifier fusion". English. In: *Renewable Energy* 116.PB (2018), pp. 99–106. ISSN: 0960-1481.

[20] A. Joshuva and V. Sugumaran. "Improvement in wind energy production through condition monitoring of wind turbine blades using vibration signatures and ARMA features: a data-driven approach". English. In: *Progress in Industrial Ecology* 13.3 (2019), p. 207. ISSN: 1476-8917.

[21] Magda Ruiz et al. "Wind turbine fault detection and classification by means of image texture analysis". English. In: *Mechanical Systems and Signal Processing* 107.C (2018), pp. 149–167. ISSN: 0888-3270.

[22] Inturi Vamsi, G.R Sabareesh, and P.K Penumakala. "Comparison of condition monitoring techniques in assessing fault severity for a wind turbine gearbox under non-stationary loading". English. In: *Mechanical Systems and Signal Processing* 124 (2019), pp. 1–20. ISSN: 0888-3270.

[23] Vamsi Inturi et al. "Integrated condition monitoring scheme for bearing fault diagnosis of a wind turbine gearbox". English. In: *Journal of Vibration and Control* 25.12 (2019), pp. 1852–1865. ISSN: 1077-5463.

[24] Francesc Pozo, Yolanda Vidal, and Óscar Salgado. "Wind Turbine Condition Monitoring Strategy through Multiway PCA and Multivariate Inference". English. In: *Energies* 11.4 (2018), p. 749. ISSN: 19961073. URL: http://search.proquest.com/docview/2041094406/.

[25] Azzeddine Bakdi, Abdelmalek Kouadri, and Saad Mekhilef. "A data-driven algorithm for online detection of component and system faults in modern wind turbines at different operating zones". English. In: *Renewable and Sustainable Energy Reviews* 103 (2019), pp. 546–555. ISSN: 1364-0321.

[26] Wenna Zhang and Xiandong Ma. "Simultaneous Fault Detection and Sensor Selection for Condition Monitoring of Wind Turbines". English. In: *Energies* 9.4 (2016), p. 280. ISSN: 19961073. URL: http://search.proquest.com/docview/1780819136/.

[27] T. Regan, C. Beale, and M. Inalpolat. "Wind Turbine Blade Damage Detection Using Supervised Machine Learning Algorithms". In: *Journal of Vibration and Acoustics, Transactions of the ASME* 139.6 (2017). ISSN: 10489002.

[28] L. Fu et al. "Condition monitoring for the roller bearings of wind turbines under variable working conditions based on the Fisher score and permutation entropy." In: *Energies* 12.16 (2019). ISSN: 19961073.

[29] Hongguang Yun et al. "An Adaptive Approach for Ice Detection in Wind Turbine With Inductive Transfer Learning". English. In: *IEEE Access* 7.99 (2019), pp. 122205–122213. ISSN: 2169-3536.

[30] H. Ren et al. "A new wind turbine health condition monitoring method based on VMD-MPE and feature-based transfer learning". In: *Measurement: Journal of the International Measurement Confederation* 148 (2019). ISSN: 02632241.

[31] Hongshan Zhao et al. "Anomaly detection and fault analysis of wind turbine components based on deep learning network". English. In: *Renewable Energy* 127 (2018), pp. 825–834. ISSN: 0960-1481.

[32] Wenguang Yang, Chao Liu, and Dongxiang Jiang. "An unsupervised spatiotemporal graphical modeling approach for wind turbine condition monitoring". English. In: *Renewable Energy* 127 (2018), pp. 230–241. ISSN: 0960-1481.

[33] Evangelos Papatheou et al. "Performance monitoring of a wind turbine using extreme function theory". English. In: *Renewable Energy* 113.C (2017), pp. 1490–1502. ISSN: 0960-1481.

[34] Peng Qian et al. "A novel wind turbine condition monitoring method based on cloud computing". English. In: *Renewable Energy* 135 (2019), pp. 390–398. ISSN: 0960-1481.

[35] Ravi Pandit and David Infield. "Gaussian Process Operational Curves for Wind Turbine Condition Monitoring". English. In: *Energies* 11.7 (2018), p. 1631. ISSN: 19961073. URL: http://search.proquest.com/docview/2108516073/.

[36] Elena Gonzalez et al. "Using high-frequency SCADA data for wind turbine performance monitoring: A sensitivity study". English. In: *Renewable Energy* 131 (2019), pp. 841–853. ISSN: 0960-1481.

[37] Sergio Martin-del-Campo, Fredrik Sandin, and Daniel Strömbergsson. "Dictionary learning approach to monitoring of wind turbine drivetrain bearings". In: (2019).

[38] Marcin Straczkiewicz and Tomasz Barszcz. "Application of Artificial Neural Network for Damage Detection in Planetary Gearbox of Wind Turbine". In: *Shock and Vibration* 2016 (2016). ISSN: 1070-9622.

[39] Hsu-Hao Yang, Mei-Ling Huang, and Shih-Wei Yang. "Integrating Auto-Associative Neural Networks with Hotelling T2 Control Charts for Wind Turbine Fault Detection". English. In: *Energies* 8.10 (2015), pp. 12100–12115. ISSN: 19961073. URL: http://search.proquest.com/docview/1732948122/.

[40] Liang Tao et al. "Abnormal Detection of Wind Turbine Based on SCADA Data Mining". In: *Mathematical Problems in Engineering* 2019 (2019). ISSN: 1024-123X.

[41] Peyman Mazidi et al. "A health condition model for wind turbine monitoring through neural networks and proportional hazard models". English. In: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 231.5 (2017), pp. 481–494. ISSN: 1748-006X.

[42] Ravi Pandit and David Infield. "Comparative assessments of binned and support vector regression-based blade pitch curve of a wind turbine for the purpose of condition monitoring". English. In: *International Journal of Energy and Environmental Engineering* 10.2 (2019), pp. 181–188. ISSN: 2008-9163.

[43] MA Rodriguez-Lopez et al. "Methodology for Detecting Malfunctions and Evaluating the Maintenance Effectiveness in Wind Turbine Generator Bearings Using Generic versus Specific Models from SCADA Data". English. In: *Energies* 11.4 (2018). ISSN: 1996-1073.

[44] Abhishek Reddy et al. "Detection of Cracks and damage in wind turbine blades using artificial intelligence-based image analytics". English. In: *Measurement* 147 (2019), p. 106823. ISSN: 0263-2241.

[45] A. Agasthian, Rajendra Pamula, and L. Kumaraswamidhas. "Fault classification and detection in wind turbine using Cuckoo-optimized support vector machine". English. In: *Neural Computing and Applications* 31.5 (2019), pp. 1503–1511. ISSN: 0941-0643.

[46] Tianyang Wang et al. "Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review". English. In: *Mechanical Systems and Signal Processing* 126 (2019), pp. 662–685. ISSN: 0888-3270.

[47] Dahai Zhang et al. "A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost". English. In: *IEEE Access* 6 (2018), pp. 21020–21031. ISSN: 2169-3536.

[48] P Bangalore and M Patriksson. "Analysis of SCADA data for early fault detection, with application to the maintenance management of wind turbines". English. In: *Renewable Energy* 115 (2018), pp. 521–532. ISSN: 0960-1481.

[49] Majid Morshedizadeh et al. "Improved power curve monitoring of wind turbines". In: *Wind Engineering* 41.4 (2017), pp. 260–271. ISSN: 0309-524X.

[50] D Yu et al. "A radically data-driven method for fault detection and diagnosis in wind turbines". English. In: *International Journal of Electrical Power and Energy Systems* 99 (2018), pp. 577–584. ISSN: 0142-0615.

[51] P. Guo, J. Fu, and X. Yang. "Condition monitoring and fault diagnosis of wind turbines gearbox bearing temperature based on kolmogorov-smirnov test and convolutional neural network model". In: *Energies* 11.9 (2018). ISSN: 19961073.

[52] Sangryul Kim and Seo Yun-Ho. "Development of a Fault Monitoring Technique for Wind Turbines Using a Hidden Markov Model". English. In: *Sensors* 18.6 (2018), p. 1790. ISSN: 14248220. URL: `http://search.proquest.com/docview/2108718625/`.

[53] Philip Cross and Xiandong Ma. "Model-based and fuzzy logic approaches to condition monitoring of operational wind turbines". English. In: *International Journal of Automation and Computing* 12.1 (2015), pp. 25–34. ISSN: 1476-8186.

[54] Longting Chen et al. "Learning deep representation of imbalanced SCADA data for fault detection of wind turbines". English. In: *Measurement* 139 (2019), pp. 370–379. ISSN: 0263-2241.

[55] Anil Kushwah and Rajesh Wadhvani. "Performance monitoring of wind turbines using advanced statistical methods". English. In: *Sadhana* 44.7 (2019), pp. 1–11. ISSN: 0256-2499.

[56] Alberto Pliego Marugan, Ana Maria Peco Chacon, and Fausto Pedro Garcia Marquez. "Reliability analysis of detecting false alarms that employ neural networks: A real case study on wind turbines". English. In: *Reliability Engineering and System Safety* 191 (2019), p. 106574. ISSN: 0951-8320.