energies

Article

A CUSUM-Based Approach for Condition Monitoring and Fault Diagnosis of Wind Turbines

Phong B. Dao

Department of Robotics and Mechatronics, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland; phongdao@agh.edu.pl

Abstract: This paper presents a cumulative sum (CUSUM)-based approach for condition monitoring and fault diagnosis of wind turbines (WTs) using SCADA data. The main ideas are to first form a multiple linear regression model using data collected in normal operation state, then monitor the stability of regression coefficients of the model on new observations, and detect a structural change in the form of coefficient instability using CUSUM tests. The method is applied for on-line condition monitoring of a WT using temperature-related SCADA data. A sequence of CUSUM test statistics is used as a damage-sensitive feature in a control chart scheme. If the sequence crosses either upper or lower critical line after some recursive regression iterations, then it indicates the occurrence of a fault in the WT. The method is validated using two case studies with known faults. The results show that the method can effectively monitor the WT and reliably detect abnormal problems.

Keywords: wind turbine; condition monitoring; fault detection; CUSUM test; structural change; multiple linear regression; SCADA data

Citation: Dao, P.B. A CUSUM-Based Approach for Condition Monitoring and Fault Diagnosis of Wind Turbines. *Energies* **2021**, *14*, 3236. https://doi.org/10.3390/en14113236

Academic Editors: Bruce Stephen and Mohamed Benbouzid

Received: 16 February 2021
Accepted: 31 May 2021
Published: 1 June 2021

1. Introduction

In econometrics and statistics, a structural break is an unexpected change over time in the parameters of regression models, which can lead to seriously biased estimates and forecasts and the unreliability of models [1–3]. A structural break could be caused by a shift in mean, variance, or a persistent change in the data property [4]. Generally, macroeconomic and financial data are subject to occasional structural breaks, which can be caused by various economic and political events [5,6]. It has been found in [7] that structural breaks have lowered the global welfare gains from world trade integration by almost 40 percent over the past four decades. Therefore the detection of structural breaks (or structural changes) has drawn ceaseless attention from the theoretical, applied economic and financial fields for many decades, as illustrated in the literature [1–8].

Testing for structural breaks dates back to the seminal paper of Chow [9]. In this paper, Chow developed an F-test for regime shift in parameters and resolved how to detect a single structural change by assuming that such break dates are known. The idea of using residuals calculated recursively to test model misspecification and parameter instability dates to the landmark cumulative sum (CUSUM) test, which was first introduced into the statistics and econometrics literatures in 1975 by Brown et al. [10], and later extended in [11] to dynamic models. The CUSUM test is based on the analysis of the scaled recursive residuals and has a significant advantage over the Chow test [9] of not requiring prior knowledge of the point at which the hypothesized structural break takes place. In essence, the motivation behind CUSUM tests is to provide a diagnostic tool for the detection of unknown structural breaks [12].

An intensive literature search has been conducted by the author and it reveals that CUSUM-based control charts have been used for fault detection and diagnosis in engineering applications. The work in [13,14] examined the feasibility of using CUSUM control charts and artificial neural networks (ANNs) together for fault detection and diagnosis.

The proposed strategy was successfully deployed in [13] for incipient diagnosis of fault conditions of a pumping machinery based on experimental data corresponding to historical pump faults. In [14], it was tested on a model of the heat transport system of a nuclear reactor. The method was able to eliminate all false alarms at steady state and correctly diagnose five out of the six faults. A fault detection, identification and diagnosis methodology for chemical plants, based on a combination of CUSUM and principal component analysis (PCA) tool, was developed in [15,16]. The CUSUM-based chart was used to enhance fault detection under conditions of small fault/signal to noise ratio while the use of PCA facilitated the filtering of noise in the presence of highly correlated data. The method was validated through a particular set of the Tennessee Eastman (TE) process faults that could not be properly detected or diagnosed with other methodologies previously reported. The proposed technique was successful in detecting, identifying and diagnosing both individual and simultaneous occurrences of these faults. The study [17] presented a new algorithm to identify and diagnose stochastic faults in the TE process. The algorithm combined ensemble empirical mode decomposition (EEMD) with PCA and CUSUM control charts to diagnose a group of faults that could not be properly diagnosed with previously reported techniques. Measured variables were first decomposed into different scales using the EEMD-based PCA, from which fault signatures could be extracted for fault detection and diagnosis. CUSUM-based statistics were further used to improve fault detection. The algorithm successfully identified three particular faults in the TE process with small time delay. In [18], the authors proposed an average accumulative (AA)-based time-varying PCA model for early detection of slowly varying faults using numerical simulation data. Through combining the advantage of the CUSUM-based method and the AA-based method, a CUSUM-AA-based method was developed to detect faults at earlier times. A condition monitoring method, based on modified CUSUM and exponentially weighted moving average (EWMA) control charts, was proposed for detecting system or equipment failure trend [19]. The method was validated using an electro pump equipment. A vibration measurement was used to monitor the equipment performance. The proposed method was shown to be effective for vibration trend detection, allowing early interventions planning before catastrophic failures could occur. The work in [20] developed an incipient fault detection and classification method for three-level neutral point clamped (NPC) inverters used in electrical drives. Phase current time series measurements for different operating conditions were used. For the fault detection, the authors used the first four statistical moments as fault features and then applied the CUSUM algorithm as the feature analysis technique to improve the performances. The PCA was then used to perform the fault classification. Recently, the authors in [21] proposed an adaptive fault detection scheme, which merges random forest (RF) with an adaptive CUSUM-based chart. RF was used to obtain the residuals and then the adaptive CUSUM control chart of time-varying shift was applied to detect the changes of residuals. This scheme was found to be superior to other competing methods in capturing faults and reducing false alarms.

It is necessary to discuss here two important issues. First, all previous studies [13–21] used a popular type of CUSUM statistical control chart, first introduced by Page (1954) [22], which is sensitive to persistent changes in mean values. This type of control chart cumulates deviations of the sample averages from the desired value [14]. Whenever the cumulations reach either a high or low limit, an out-of-control signal is triggered. Second, all methods reported in [13–21] used this type of CUSUM control chart as a supporting or auxiliary tool to improve their condition monitoring and/or fault diagnosis. Specifically, in these previous studies the CUSUM control chart technique was combined with other techniques, such as ANNs [13,14], PCA [15,16,20], EEMD-based PCA [17], AA-based PCA [18], EWMA [19], and RF [21], for fault detection and diagnosis.

The major difference between the work presented in this paper and the previous studies [13–21] is that, instead of using the CUSUM statistical control chart discussed above, the CUSUM test (algorithm) first introduced by Brown et al. [10] has been studied, adapted and used as a single tool for the purpose of condition monitoring and fault diagnosis.

In a broad view, this study has aimed at developing a new approach, based on the well-established structural change/break tests from the fields of econometrics and statistics, for condition monitoring and fault diagnosis of engineering systems and/or structures. The main ideas are to first form a multiple linear regression model using data collected while the system or structure of interest is operating in the normal state without fault, then start monitoring the stability of the regression coefficients of the model on new arriving observations, and detect a structural change of the model (in the form of coefficient instability) using structural break tests. It is assumed in this study that any structural change will most likely reflect the occurrence of a fault or an abnormal problem in the engineering system and/or structure of interest.

In the present work, the proposed approach has been investigated for monitoring the operational condition of wind turbines (WTs) using supervisory control and data acquisition (SCADA) data. Since the CUSUM test [10] is one of the most common testing methods used for structural changes (or structural breaks) in economic and financial time series data, the test has been adapted for this purpose. A CUSUM-based computation procedure involving four steps is developed in this paper. The method is relied on multiple linear regression models, which are formed using WT SCADA data. Monitoring the stability of regression coefficients in a multiple linear regression model is used to assess the operating condition of a WT. In essence, coefficient instability means structural change in the regression model, which can be interpreted as the occurrence of a fault in the WT. Temperature-related SCADA data—acquired from a WT drivetrain with a nominal power of 2 MW within 30 days under varying environmental and operational conditions—were used to validate the method. To the best of author's knowledge, condition monitoring and fault detection of WTs based on the CUSUM test, originally introduced by Brown et al. [10], for structural change/break detection in SCADA data has not been previously investigated in the literature.

The remaining parts of the paper are organised as follows: Section 2 introduces the theoretical background of the CUSUM test. An example of CUSUM tests for structural break detection is then given using economic time series data. Section 3 presents a new condition monitoring approach for engineering systems and/or structures based on the CUSUM test. A case study using temperature-related SCADA data for condition monitoring of wind turbines is described in Section 4. Implementation of the proposed method for the case study is presented in Section 5. The results of condition monitoring and abnormality detection are then presented and discussed. Finally, the paper is concluded in Section 6.

2. Testing for Structural Change/Break in Economic Time Series

2.1. Introduction

The problem of detecting structural changes in regression relationships has been an important topic in statistical and econometric research for many years. General speaking, structural change (or break) tests can be categorized into two groups [3,4]. The first group is the classical approach, which employs *retrospective tests* using a historical data set of a given length. These tests are based on F statistics. The Chow test [9] and the supF test [1] belong to this class. The second group is the *fluctuation-type test* in a monitoring scheme. Within this test framework, a regression relationship is known to be stable for a given history period; then one will test whether incoming data are consistent with the previously established relationship [3,4]. Fluctuation tests do not assume a particular pattern of structural change. The best-known example from the fluctuation-type test framework is the cumulative sum (CUSUM) test for parameter stability first introduced in [10], and later extended in [11] to dynamic models. In the following, the CUSUM test is briefly described. It is noted that the introduction avoids using complicated mathematics behind the test. For more theoretical details, potential readers are referred to the original work [10,11].

2.2. CUSUM Test for Structural Change

CUSUM test is based on cumulative sums of residuals resulting from recursive regressions. The test is used to assess the stability of the regression coefficients in a multiple linear regression model of the form:

$$y_t = \beta_0 + \beta_{1t}x_{1t} + \beta_{2t}x_{2t} + \cdots + \beta_{pt}x_{pt} + \varepsilon_t,\ t = 1, \ldots, T, \quad (1)$$

which can be written as:

$$y_t = x_t\beta_t + \varepsilon_t,\ t = 1, \ldots, T, \quad (2)$$

where y_t is the response (or dependent) variable, $x_t = \left(1, x_{1t}, \ldots, x_{pt}\right)$ are p predictor (or independent) variables, β_0 is the intercept term (often labeled as constant), $\beta_t = \left(\beta_0, \beta_{1t}, \ldots, \beta_{pt}\right)^T$ is an $(p+1)$-dimensional vector of regression coefficients, and ε_t are independently distributed normal random errors with mean zero and variance σ^2. It should be noted here that the intercept term β_0 is the expected mean value of the response variable when all predictor variables are equal to zero, the coefficients $\beta_{1t}, \ldots, \beta_{pt}$ are known as slope coefficients for predictor variables, and ε_t is also known as the residual. In Equation (2), x_t and y_t are specified as an $(T \times p + 1)$ numeric matrix and an $(T \times 1)$ numeric vector, respectively, where T is the number of observations (or sample size) and p is the number of predictor (or independent) variables.

CUSUM tests are commonly used in econometrics and statistics to assess whether there are structural changes (or structural breaks) in a regression equation of interest. Inference is based on a sequence of sums of recursive residuals computed iteratively from sequential subsamples of the data. The calculation is relied on standardized one-step-ahead forecast errors [10]. The CUSUM test computes recursive residuals beginning with the first $(k+1)$ observations, where k is the number of regression coefficients. Then it adds one at a time until it reaches the number of observations. In the following, a calculation procedure for the CUSUM test statistics is presented and an approximation of the standard error band for the CUSUM test statistics is derived.

From Equation (2), the *recursive residuals* can be calculated as [12]:

$$w_r = \frac{y_r - x_r\hat{\beta}_{r-1}}{\sqrt{1 + x_r'\left(X_{r-1}'X_{r-1}\right)^{-1}x_r}}, \qquad r = k+1, \ldots, T, \quad (3)$$

where k is the number of regression coefficients, $X_r = (x_1, x_2, \ldots, x_r)$ and $\hat{\beta}_r = \left(X_r'X_r\right)^{-1}X_r'y_r$ is the vector of ordinary least-squares (OLS) estimates for the regression parameters based on data $t = 1, \ldots, r$. These recursive residuals are used to construct the CUSUM test statistics as [12]:

$$W_r = \sum_{t=k+1}^{r} \frac{w_t}{\hat{\sigma}}, \qquad r = k+1, \ldots, T, \quad (4)$$

where:

$$\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{T}\left(y_t - x_t\hat{\beta}_r\right)^2}{T-k}} \quad (5)$$

In the CUSUM test, the null hypothesis (H_0) is that the regression coefficients β_t in Equation (2) are equal (or stable) in all sequential subsamples. In other words, if the null hypothesis is true then it implies that there is no structural break in the observed time series, i.e., the model of interest is stable. On the contrary, the alternative hypothesis (H_1) is that the regression coefficients change during the period of the sample.

The statistical hypothesis testing results in one of two different decisions (i.e., $h = 1$ or $h = 0$). Based on that, we can assess the stability of the data-based regression model over time as follows:

The test decision $h = 1$ indicates rejection of H_0 in favor of the alternative hypothesis H_1, i.e., it rejects the coefficient stability of the model at a certain level of significance.

The test decision $h = 0$ indicates failure to reject H_0, i.e., it fails to reject the coefficient stability of the model at a certain level of significance.

In practice, the sequence of CUSUM test statistics is often used because it brings out departures from parameter constancy in a graphical way. Specifically, under the null hypothesis of coefficient constancy, if the sequence of the CUSUM test statistics crosses into a critical region (also known as a standard error band) then it suggests structural change in the model over time.

Calculation of the Critical Region

The critical region or the standard error band for the CUSUM statistics is specified by nonlinear functions of r which take the form $\pm\gamma\sqrt{r-k}$, where γ is the parameter which determines the size of the test [10]. However, these nonlinear functions are usually approximated by the straight lines passing through the points $\left\{k, \pm a\sqrt{T-k}\right\}$, $\left\{T, \pm 3a\sqrt{T-k}\right\}$, as discussed in [12], where k is the number of coefficients in the regression equation and a is chosen so that these lines are tangential to the nonlinear functions $\pm\gamma\sqrt{r-k}$ at the mid-point $r = (T-k)/2$. This means that the approximate critical values will exceed the true critical values for values of r close to k or to T [12]. For a certain level of significance selected, a specific value is found for a. For example, $a = 0.948$ gives a 5% significance level while $a = 1.143$ gives a 1% significance level. As a result, the critical region or the standard error band for the CUSUM statistics can approximately be specified by these two critical straight lines.

In the following, the use of CUSUM tests for structural break detection is illustrated using economic time series data.

2.3. Example of CUSUM Tests for Structural Break Detection Using Economic Data

This example uses the data of Nelson and Plosser [23], which contains annual measurements of fourteen different macroeconomic indexes of the U.S. from 1915 to 1970. The data set used in this example consists of the real gross national product ($GNPR$), industrial production index (IPI), total employment index (E), and real wages (WR), which are plotted in Figure 1. Suppose that we want to develop an explanatory model for real gross national product as determined by the other three indexes, and then assess its stability over time by means of CUSUM tests.

Consider the multiple linear regression model of the following form:

$$GNPR_t = \beta_0 + \beta_{1t} IPI_t + \beta_{2t} E_t + \beta_{3t} WR_t + \varepsilon_t \tag{6}$$

where β_0 is the intercept term, β_{1t}, β_{2t}, and β_{3t} are the slope coefficients, and ε_t is a Gaussian random variable with mean zero and standard deviation σ^2. The null hypothesis is defined as that the regression coefficients $\beta_t = (\beta_0, \beta_{1t}, \beta_{2t}, \beta_{3t})^T$ in the model given by Equation (6) are identical (or stable) across all sequential subsamples. If at least a value of the sequence of the CUSUM test statistics goes into the critical region (limited by the upper and lower critical lines), it suggests structural change in the model over time (i.e., the null hypothesis is rejected). If all values of the test statistics stay out of the critical region, it indicates failure to reject the null hypothesis, i.e., there is no structural break in the observed data, or in other words, the model is stable.

The CUSUM test is performed to assess the stability of the model at 5% level of significance. The results, plotted in Figure 2, show that the sequence of the CUSUM test statistics starts crossing the upper critical line at the 46th recursive regression (exactly, from the 46th iteration step till the end). As a result, we reject the null hypothesis of coefficient stability of the model, or in other words, the CUSUM test suggests that the model in this example is not stable.

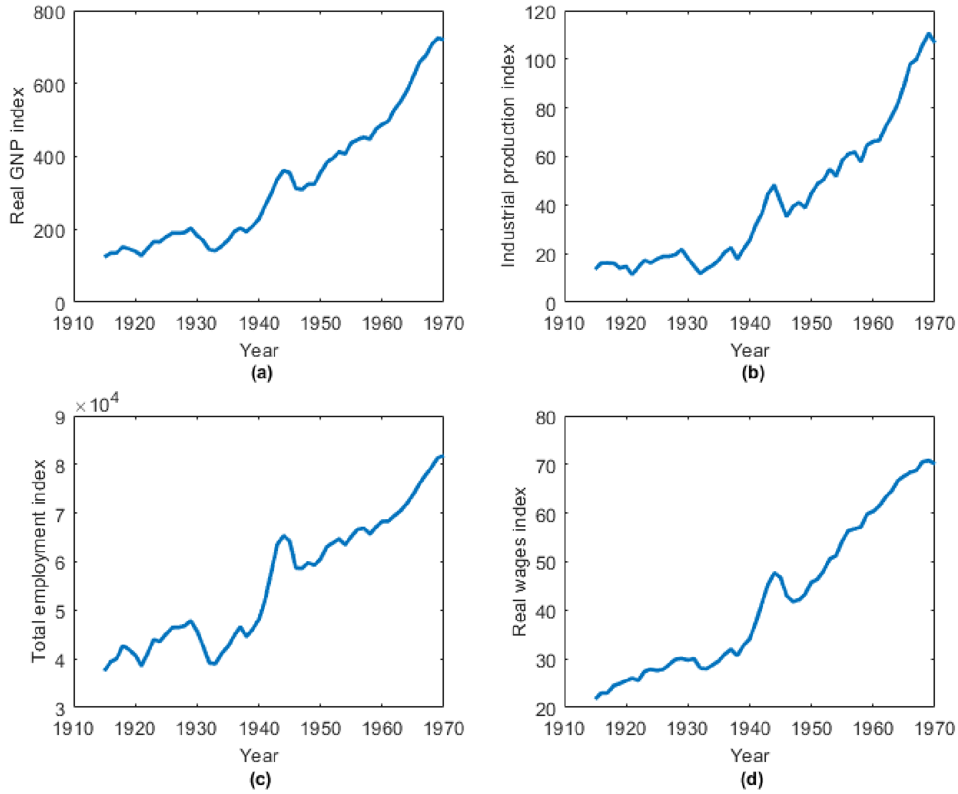


Figure 1. Macroeconomic indexes of the U.S. from 1915 to 1970: (**a**) real gross national product ($GNPR$); (**b**) industrial production index (IPI); (**c**) total employment index (E); (**d**) real wages (WR).

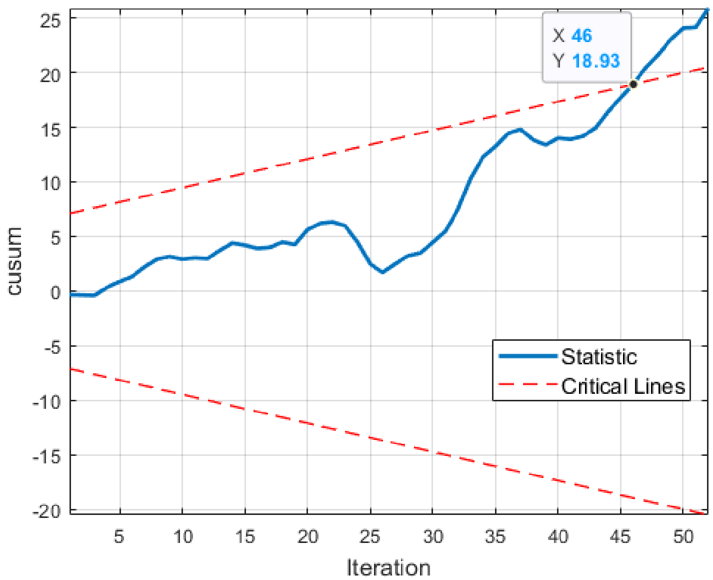


Figure 2. Structural change detection in the model for real gross national product using the CUSUM test.

3. Condition Monitoring and Fault Diagnosis of Engineering Systems and/or Structures Based on CUSUM Tests

A CUSUM-based computation procedure for condition monitoring and fault detection of engineering systems and/or structures is shown in Figure 3. The entire procedure consists of four steps.

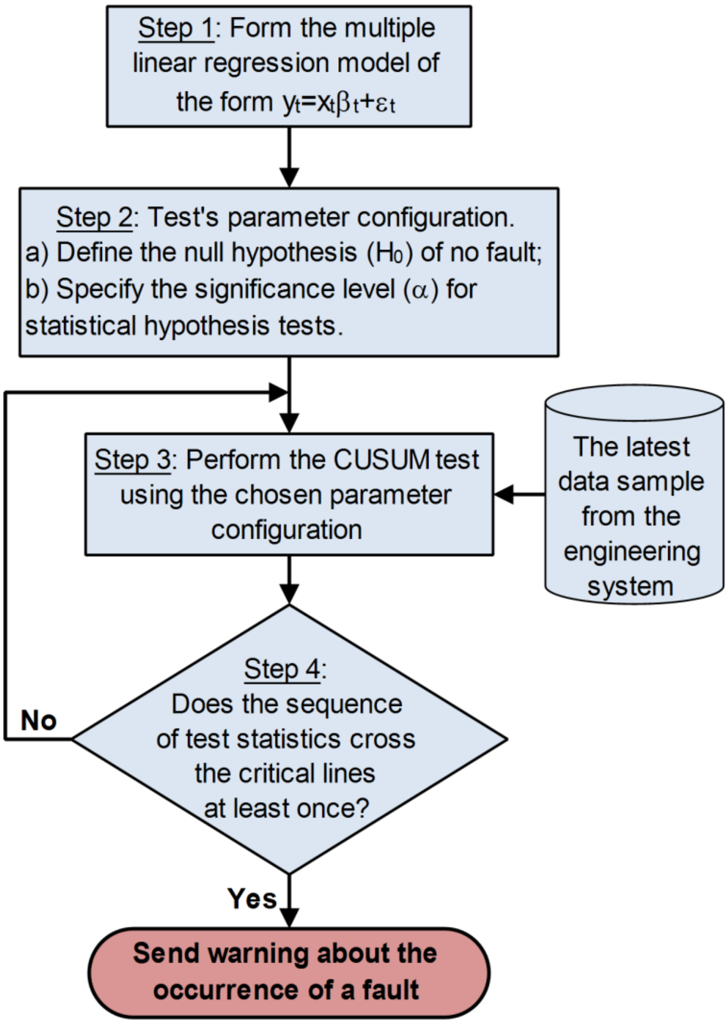


Figure 3. CUSUM-based computation procedure for condition monitoring and fault detection of engineering systems and/or structures.

Step 1: Form a multiple linear regression model of the form $y_t = x_t\beta_t + \varepsilon_t$ for the engineering system and/or structure of interest. It is important to note here that the model is required to be formed with an initial data set acquired in the normal operation state without fault. This initial data set consists of $(k + 1)$ observations or samples, where k is the number of regression coefficients.

Step 2: Tests of the parameter configuration, including:

- Define the null hypothesis (H_0)—i.e., what is known as true for the regression model formed in Step 1. In other words, H_0 is what we want to test. Essentially, H_0 represents the fact that there are not structural changes in the model. This implies in the point of view of condition monitoring that the engineering system and/or structure of interest has no fault or abnormal problems. On the contrary, the alternative hypothesis (H_1)—i.e., what we need to accept if the null hypothesis is not true—means that an abnormal problem would appear in the system.
- Specify the significance level (or confidence level) α for the statistical hypothesis testing, which is performed in Step 3. It is well known that, for statistical hypothesis tests, the significance level α is the probability of rejecting the null hypothesis.

Step 3: Perform the CUSUM test using the chosen parameter configuration.

Before performing the test, the latest (or most recent) data samples, acquired from the engineering system and/or structure of interest, are added into the data set. The CUSUM test is then executed to assess the regression model. All available data, including from the first sample of the current monitoring process to the latest one, are used in the test. Sums of recursive residuals are iteratively computed across all sequential subsamples of the data set. Afterward, the sequence of the CUSUM test statistics is computed, as described in Section 2.2.

It is important to clarify here that this study does not simply apply the CUSUM test. The main idea of the proposed method is to perform the CUSUM test in the loop (as shown in Figure 3) at every acquired data samples (or observations) to assess the stability of the regression coefficients (β_t) in the multiple linear regression model using significance tests or statistical hypothesis tests. This is the key point to make the CUSUM test usable for on-line condition monitoring applications. It is assumed that if there is statistically significant evidence to reject the null hypothesis of coefficient stability at the chosen level of significance, then it implies that a fault or an abnormal problem might occur in the engineering system and/or structure of interest.

Step 4: Does the sequence of the CUSUM test statistics cross into the critical region (limited by two critical lines) at least once?

To ease the interpretation of statistical hypothesis testing results and provide a more convenient way for the condition monitoring process of engineering systems or structures in a graphical way, this study has used the CUSUM test statistics as an effective damage-sensitive feature (or indicator) in a control chart scheme, where the sequence of the test statistics is plotted against the critical region. The fault detection process can be interpreted as follows:

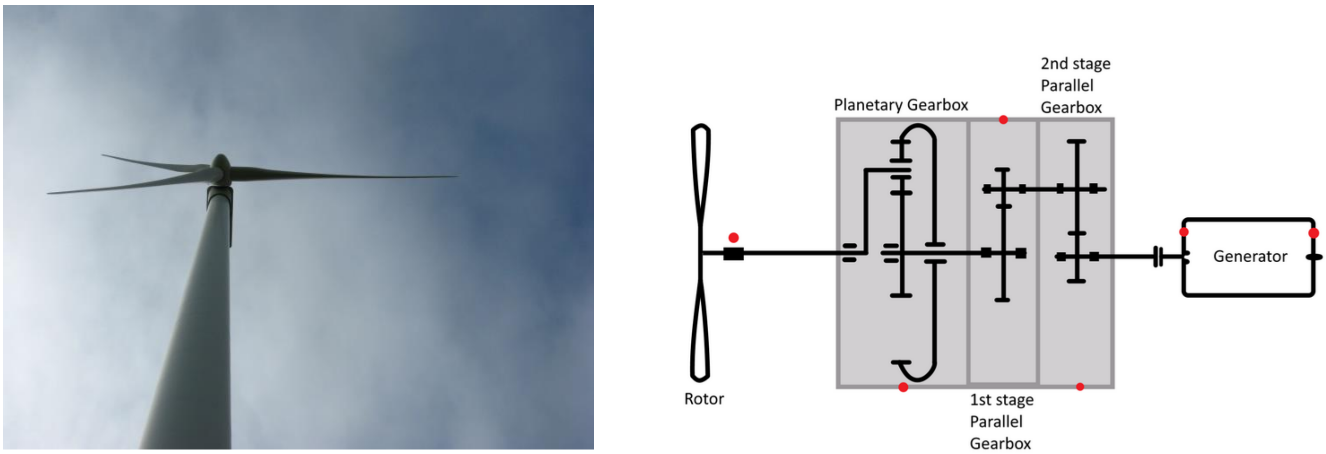
- If the sequence of the CUSUM test statistics crosses into the critical region after some recursive regression iterations, then it indicates the occurrence of a fault or an abnormal problem in the engineering system or structure at the current data sample. Consequently, send the warning information about the possible occurrence of a fault. After the fault/abnormality is identified, a waiting time period must be spent until a certain data sample when the system or structure completely returns to the normal operating state. Then the calculation procedure can start again from Step 1.
- If the sequence of the CUSUM test statistics does not cross into the critical region after all recursive regression iterations, then it is an indication that there is no fault or abnormality in the engineering system or structure from the beginning of the monitoring process till the current data sample. In this case, the calculation procedure returns to Step 3 and continues.

Before closing this section, a few remarks are given hereafter. First, the proposed CUSUM-based method can be considered as a semi-supervised approach because only a data set under the normal operating condition (i.e., not involving data from fault or abnormal states) is used to initially form the model. Second, an advantage of employing the control chart approach in this study is that the proposed method can be automated for on-line condition monitoring and fault detection of engineering systems and structures. Third, it is also important to discuss that the fit of the multiple linear regression model to the observed data is assured by the CUSUM test algorithm originally developed in [10]. More specifically, the recursive regression coefficients in the multiple linear regression model are instantly updated with respect to changing in the acquired data samples. Therefore when using the CUSUM-based method proposed in this paper, the fit of the model to the given data is checked and guaranteed in Step 3. In the following section, a case study using temperature-related wind turbine SCADA data is described.

4. Case Study Using Temperature-Related Wind Turbine SCADA Data

According to a recent report of Global Wind Energy Council (GWEC) [24], total cumulative installations of wind power capacity reached to 591 GW in 2018 and new installations are expected of more than 55 GW each year until 2023. However, unexpected failures of wind turbine components are the main reasons that cause costly repair and long-term machine breakdown, leading to high operation and maintenance cost [25,26]. Therefore condition monitoring (CM) and fault diagnosis of wind turbines (WTs) has become an essential research topic over the past twenty years aiming to increase lifetime expectancy of WTs while reducing operation and maintenance cost [25–27]. Many efforts have been made to develop reliable, efficient and cost-effective CM techniques for WTs, as reviewed in the literature [28–33]. The SCADA-based approach has recently been recognized as an effective solution because it offers great advantages for developing CM systems for WTs, as discussed in [34–42]. Specially, SCADA systems are installed in the majority of WTs for system control and data acquisition therefore the data needed for analysis are readily available and no more hardware investment is required when developing a SCADA-based CM system. This solution is thus cost-effective and easily deployed when compared with other CM techniques. The SCADA system in each WT is equipped with a wide area network (WAN) for data transmission from the WT to the data server (central computer). In a typical WAN, one or multiple gateways convey the traffic between the sensor nodes and the central computer where SCADA data are stored. A gateway in a WAN uses an internet protocol-based backbone communication interface. Depending on the implementation, the backbone can operate over a wired (e.g., Ethernet) or wireless (e.g., 4G or LTE) broadband network. Recently, the survey in [31] summarises the current state of machine learning methods that have been used for condition monitoring in WTs. The work has found that most models use SCADA data, with almost two-thirds of methods using classification and the rest relying on regression. Artificial neural networks, support vector machines and decision trees are most commonly used [31].

The SCADA data used in this paper originate from a series of experimental measurements for a wind turbine drivetrain with a nominal power of 2 MW (see Figure 4). The wind turbine belongs to a wind farm in Poland. It should be noted that the data set was not collected under regular operating phase (i.e., not during electricity production stage) of the wind turbine. The data acquisition was performed in a full-control scheme to collect a 'benchmark' SCADA data set with known faults, which can be used for testing advanced signal/data processing algorithms and/or machine learning techniques with applications to wind turbine condition monitoring and fault diagnosis. The experimental data were acquired at 10-min intervals during thirty days in November 2012. A number of process (or operational) parameters were monitored and recorded under varying operating conditions. The collected data were also influenced by environmental conditions (e.g., wind speed, ambient temperature variations between day and night, and air humidity). The data acquisition process acquired 4320 data samples for each process parameter. It is necessary to note that the data set was created from high quality industrial sensors and did not require outlier cleaning nor de-noising. Therefore we did not apply any preprocessing, as it was not required. Since SCADA systems are very reliable, there were no missing values in the data set. More detailed description of the wind turbine SCADA data can be found in [38].



(**a**) (**b**)

Figure 4. The wind turbine used in this study: (**a**) general view; (**b**) kinematic scheme with locations of vibration sensors (red dots).

Furthermore, it has been broadly discussed in the literature that temperature data and temperature trend analysis of wind turbine components can reliably provide an early indication of generator, bearing, and gearbox faults [34,43–49]. Feng et al. [43] discussed that the gearbox temperature rises when the gearbox efficiency decreases. Moreover, an unexpected increase in component temperature may indicate overload, poor lubrication, or possibly ineffective passive or active cooling [45,46]. Particularly, generator temperature is believed to have direct relation with the electrical loads and ambient conditions [46]. Also, the gearbox main bearing and lubrication oil temperature may offer the possibility of detecting gearbox overheating [34]. Consequently, the analysis of temperature-related parameters of WT's main components has been largely used in the existing condition monitoring techniques, as reported in [34,43–49].

Following this practice, the temperature data of the gearbox and generator bearing have been used in the current work for condition monitoring and fault detection of the wind turbine. The three temperature parameters, plotted in Figure 5b–d, were measured at the generator bearing (one in the front and another in the back of the generator) and at the gearbox bearing. It is assumed in this study that the temperature of the gearbox and generator bearing depends strongly on the generator speed. Therefore this work has investigated the relations between these temperature parameters and the generator speed. The nonlinear relationships between the generator speed and the generator bearing temperature (front), the generator bearing temperature (back), and the gearbox bearing temperature are shown in Figure 6a–c, respectively.

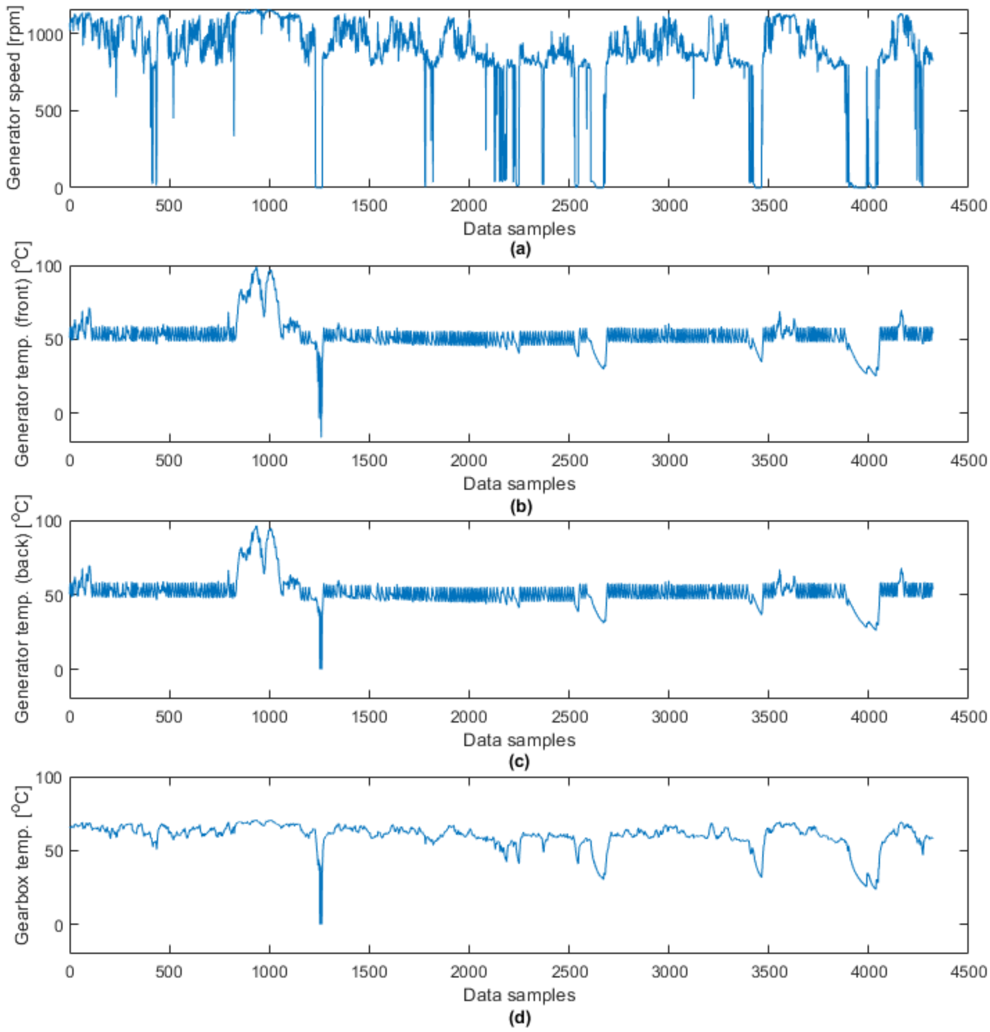


Figure 5. SCADA data set used in the case study: (**a**) generator speed; (**b**) generator bearing temperature (front); (**c**) generator bearing temperature (back); (**d**) gearbox bearing temperature.

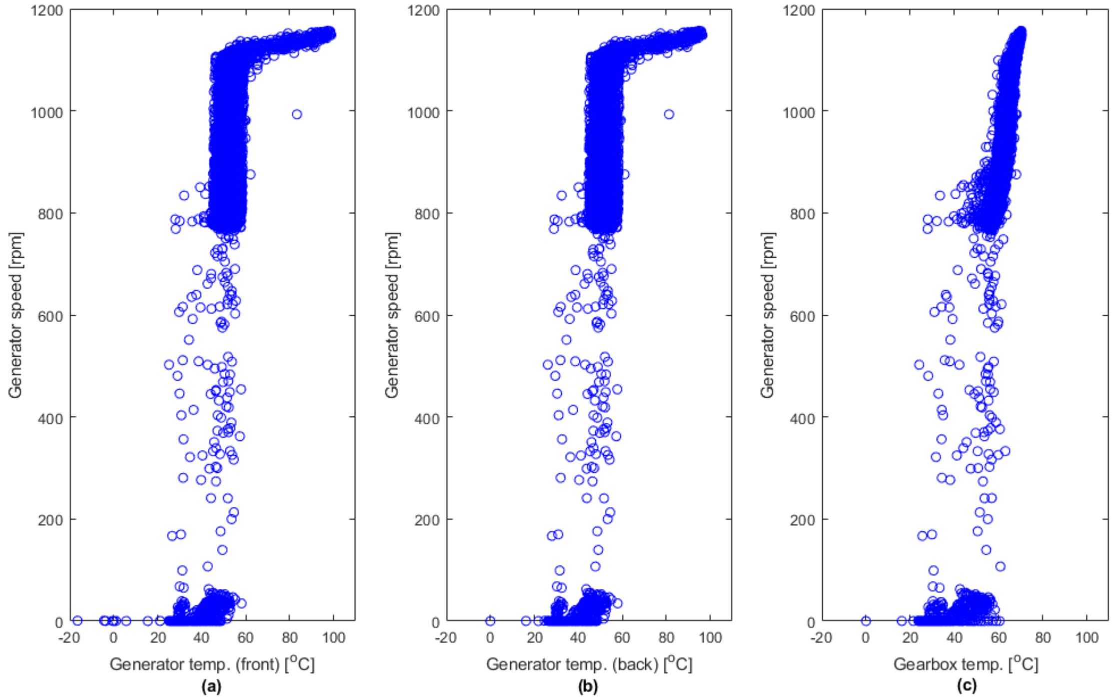


Figure 6. Nonlinear relationships between the generator speed and three temperature parameters: (**a**) generator temperature (front); (**b**) generator temperature (back); and (**c**) gearbox temperature.

The CUSUM-based method (presented in Section 3) has been used for condition monitoring and fault detection of the wind turbine. Additionally, it is expected that the method can reliably detect two known fault events of the wind turbine, which are indicated in the data in Figure 7 as the abnormal operating state (F1) and the gearbox fault (F2).

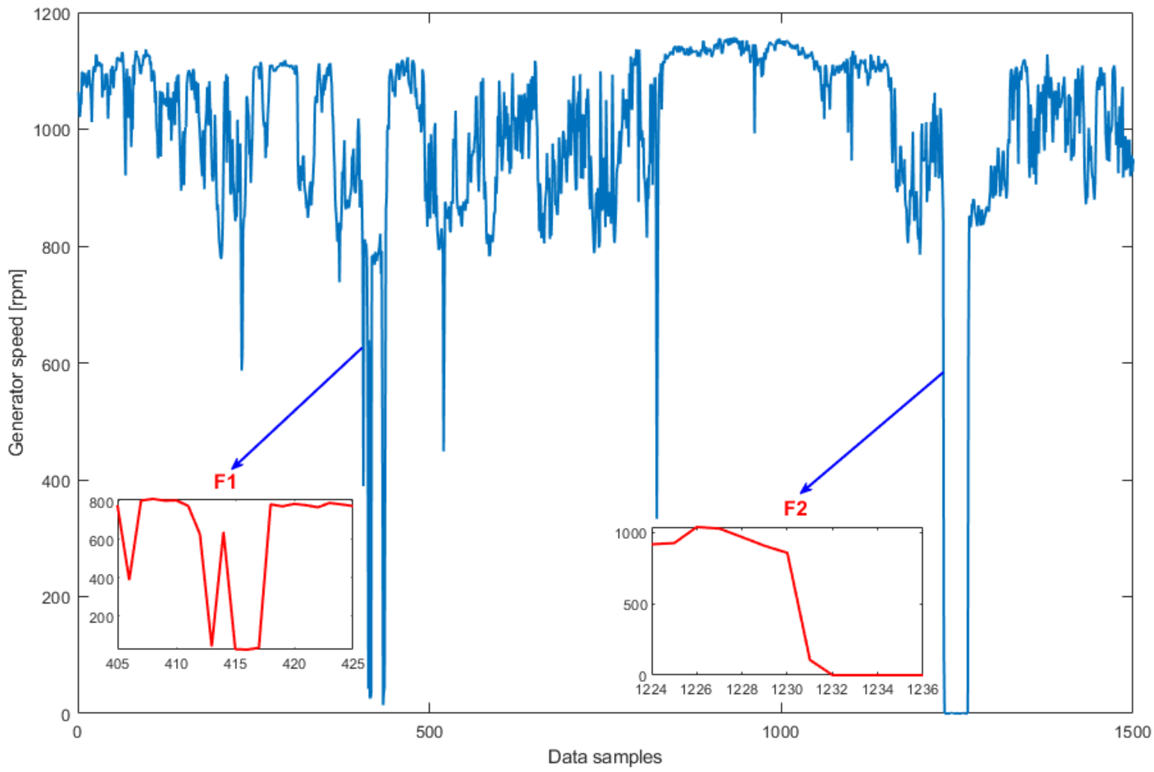


Figure 7. Generator speed data displaying the abnormal operating state (F1) and the gearbox fault (F2) occurred during monitoring process.

These two fault events were made during experimental process and then identified from the event logs for the wind turbine. They are described hereafter:

(1) The abnormal operating state F1 occurred during a time interval (80 min) between the data samples 410 and 418 (see Figure 7). This abnormality happened while the wind speed dropped down from 4.4 meters per second (mps), then stayed around 3.5 mps, and finally increased up to 4.75 mps. In response to the wind speed variations, the generator speed dropped down from 800 revolutions per minute (rpm) to almost stationary state (standstill), then suddenly increased up to more than 600 rpm and afterward rapidly decreased again to 0 rpm, and finally boosted up to the speed nearly 800 rpm. It can be observed that the generator speed varied abnormally with respect to the conditions of the wind speed. It is expected that this abnormality would be early identified to guarantee a proper operating condition of the WT and avoid more serious problems.

(2) The gearbox fault F2 occurred at the data sample 1232 (see Figure 7). It happened when the generator speed was abruptly dropped down to the zero value, while the wind speed was still stable around 5.6 mps (i.e., normal range of wind speed for WT operation). It was reported that this fault might be caused by a bearing failure or damage in the gearbox. So, it is important to detect this fault accurately at the early stage of its occurrence.

In Section 5, a multiple linear regression model—using the temperature data of the gearbox and generator bearing as the predictors (or independent variables) and generator speed data as the response (or dependent variable)—has been formed to validate the proposed method using these two case studies with known faults.

5. Results and Discussion

5.1. Implementation of The Proposed Method for The Case Study

The temperature-related SCADA data of the wind turbine (described in Section 4) were used to validate the condition monitoring and fault detection approach based on the CUSUM test (presented in Section 3). In the following, the four-step calculation procedure (shown in Figure 3) is deployed for the case study. It should be noted that the CUSUM-based computation procedure for condition monitoring and fault detection was implemented using the MATLAB Econometrics Toolbox™ [50]. In particular, the calculations of the CUSUM statistics as well as the critical region have been done using the function 'cusumtest' of the toolbox.

Step 1: A multiple linear regression model is formed using gearbox and generator temperature data as the predictors (i.e., independent or input variables) and generator speed data as the response (i.e., dependent or outcome variable). More specifically, it is assumed that the generator speed (S_t) is a linear function of the front-part generator bearing temperature (T_{1t}), the back-part generator bearing temperature (T_{2t}), and the gearbox bearing temperature (T_{3t}). In other words:

$$S_t = \beta_0 + \beta_{1t}T_{1t} + \beta_{2t}T_{2t} + \beta_{3t}T_{3t} + \varepsilon_t \tag{7}$$

where $\beta_t = (\beta_0, \beta_{1t}, \beta_{2t}, \beta_{3t})^T$ are the regression coefficients, and ε_t is a Gaussian random variable with mean zero and standard deviation σ^2.

Following the explanation in Section 3, since the number of regression coefficients is equal to four in this case, the model is initially formed with a data set consisting of five data samples acquired in the normal operation state without fault for each variable.

Step 2: Specify the parameters:

- The null hypothesis (H_0) is defined as that the regression coefficients β_t in the multiple linear regression model given by Equation (7) are identical (or stable) across all sequential subsamples. In a simple description, if the null hypothesis is true then it implies that the wind turbine is in the normal operation condition (no fault). Otherwise, the null hypothesis is rejected in favor of the alternative hypothesis (H_1), which indicates that a fault occurs in the wind turbine.

- The 1% level of significance is chosen for statistical hypothesis tests.

Step 3: Perform the CUSUM test using the parameters specified in Step 2.

Before performing the CUSUM test to assess the regression model in Equation (7), the latest (or most recent) data samples of the generator speed, generator and gearbox temperature are added into the data set. CUSUM test is then performed using all available data, including from the first sample of the current monitoring process to the latest one. Sums of recursive residuals are iteratively computed across all sequential subsamples of the data set and then the sequence of the CUSUM test statistics is computed.

Step 4: Does the sequence of the CUSUM test statistics cross into the critical region (limited by two critical lines) at least once?

Using the proposed method, the outcome of the fault detection process is simply relied on the answer found for this question.

- If at least a value of the sequence of the CUSUM test statistics goes into the critical region, it suggests structural change in the model over time (i.e., the null hypothesis is rejected at the chosen level of significance). This indicates the occurrence of a fault or an abnormal problem in the wind turbine at the current data sample.
- If all values of the test statistics stay out of the critical region (or the test statistics do not cross into the critical region), it indicates failure to reject the null hypothesis, i.e., there is no structural break in the observed data, or in other words, the model is stable. If this is the case then there is no fault or abnormality in the wind turbine from the beginning of the monitoring process till the current data sample, or in other words, the wind turbine is still in the normal operation condition.

Some remarks are given here. Since the CUSUM test is based on multiple linear regression models, it turns out that linear regression is used for the wind turbine case study in which the SCADA data has nonlinear relations. However, as depicted in Figure 6, the nonlinear relationships between the generator speed and three selected temperature parameters exhibit simple monotonic behaviour. It is thus assumed in this study that these monotonic nonlinear relationships can be significant when modelling using the multiple linear regression model. It is expected that if the proposed CUSUM-based method can be adapted for using with multiple nonlinear regression models, the fault detection results would be improved. In the following, the applicability of the presented method is illustrated through the monitoring and diagnosis results of the abnormal operating state (F1) and the gearbox fault (F2).

5.2. Condition Monitoring and Fault Detection Results

It should be noted that the nth recursive regression iteration is referred to as the calculation of the CUSUM test at the moment when the nth data sample arrives. This implies that the concepts of iterations and data samples are identical and exchangeable.

The monitoring and diagnosis results of the abnormal operating state F1 are shown in Figure 8, where the sequence of CUSUM test statistics is plotted together with two critical lines (specified by the dashed lines). In order to illustrate the detection process of F1 more clearly, the results in Figure 8 are enlarged in Figure 9 for the recursive regression iterations in the range (380, ... , 415). It indicates that F1 could be detected between the iterations 410 and 411 when the sequence of CUSUM test statistics crosses the lower critical line. In other words, the abnormal operating state was detected by the calculation procedure at the data sample 411. As described in Section 4, the abnormal operating state F1 occurred during a 80-min time interval between the data samples 410 and 418. Due to the fact that this abnormal operating state could be detected at the data sample 411, it implies that the proposed method detected the abnormal operating state F1 at the very early stage of its occurrence.

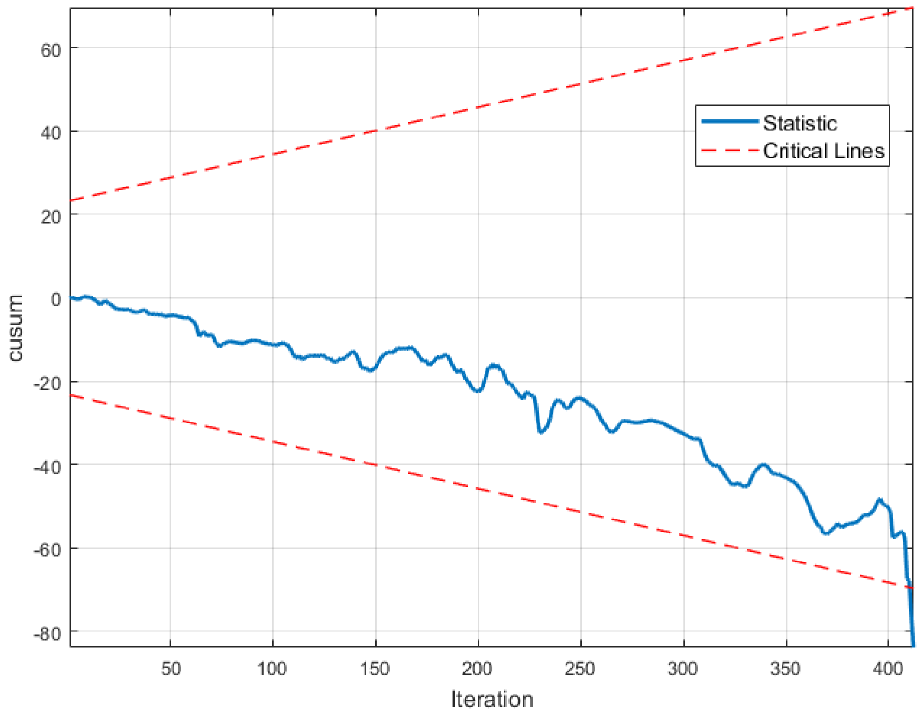


Figure 8. Monitoring and diagnosis results of the abnormal operating state F1 using CUSUM test statistics.

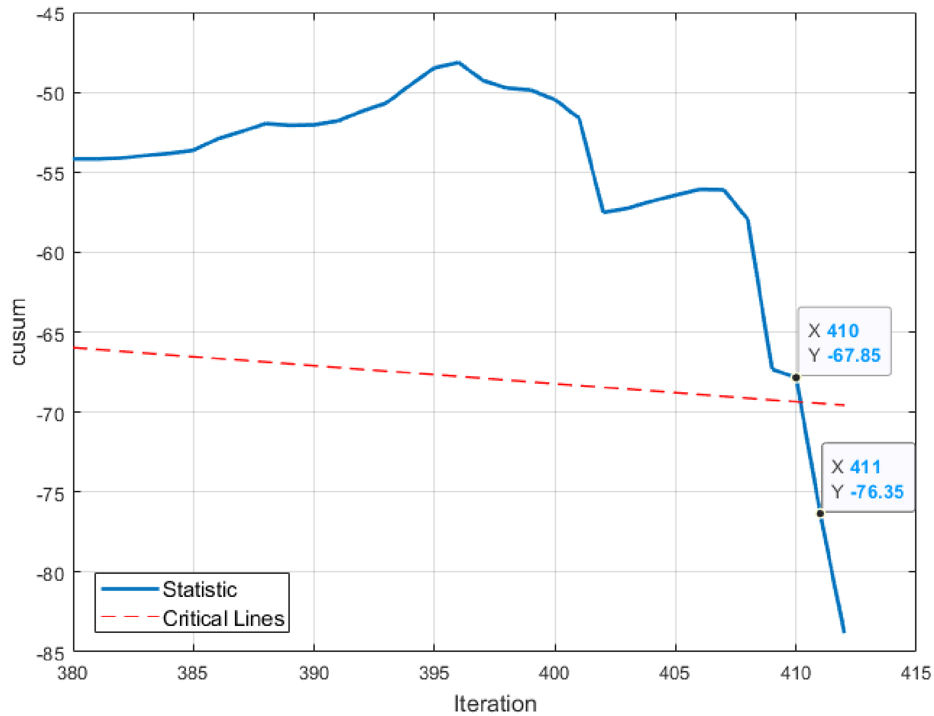


Figure 9. Zoomed data from Figure 8 displaying the detection process of F1.

It should be noted here that after the occurrence of the abnormal operating state F1 between the data samples 410 and 418, we had a waiting time period until the data sample 440 before continuing with the calculation procedure for the gearbox fault F2. It is due to the fact that we needed to wait until the moment when the wind turbine returns fully to the normal operating state. This is confirmed by the fact that the multiple linear regression model given by Equation (7) is again stable from the data sample 440.

Figure 10 presents the monitoring and diagnosis results obtained for the gearbox fault F2. The fault detection process is simply based on the sequence of CUSUM test statistics plotted against two critical lines. Figure 11 enlarges the results for the recursive regression iterations in the range (760, . . . , 795). It shows that F2 could be detected between the iterations 789 and 790 when the sequence of CUSUM test statistics crosses the lower critical line. In other word, the gearbox fault was detected by the calculation procedure at the data sample 790. Moreover, since this detection is computed using data set from the sample 440, as discussed above; and for each iteration, a new data sample is added to the computation. Therefore the exact moment of detection is at the data sample 1230 (i.e., 440 plus 790). As described in Section 4, the gearbox fault F2 really came to effect at the data sample 1232 after the generator speed was dropped down to the zero value. Since this gearbox fault could be detected at the data sample 1230, it can be stated that the proposed method detected in advance (20 min earlier) the occurrence of the gearbox fault. This would provide operators sufficient time to shut down the WT in order to prevent the system from being damaged or destroyed.

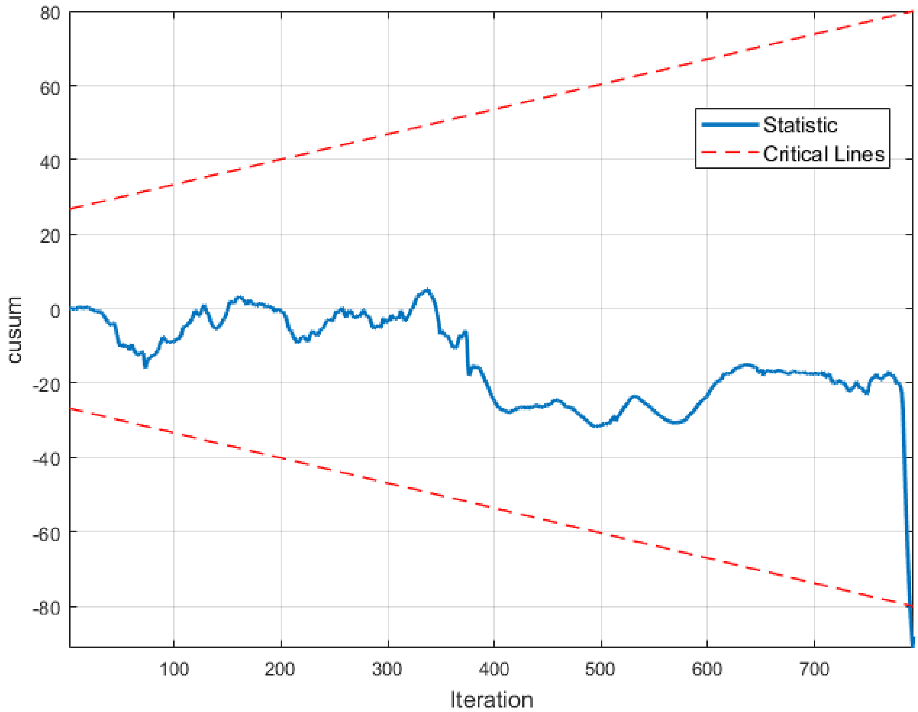


Figure 10. Monitoring and diagnosis results of the gearbox fault F2 using CUSUM test statistics.

5.3. Discussion

It should be mentioned that the same wind turbine SCADA data set was used in the author's previous paper [48], which is based on the cointegration analysis approach. More specifically, three temperature parameters, selected as the same as in this study, were also used in [48]. Apart from the same data set used, cointegration is also a regression-based technique, therefore it would be interesting to compare the method proposed in this paper with the cointegration-based previous study in [48].

In [48], as three temperature variables were used to form the cointegration model, two cointegration residuals were obtained as the results of the cointegration process. Here, it is noted that basically the number of residuals obtained depends on the number of cointegrated variables and the number of common trends presenting in the cointegration system. More detailed explanation for this issue can be found in [38].

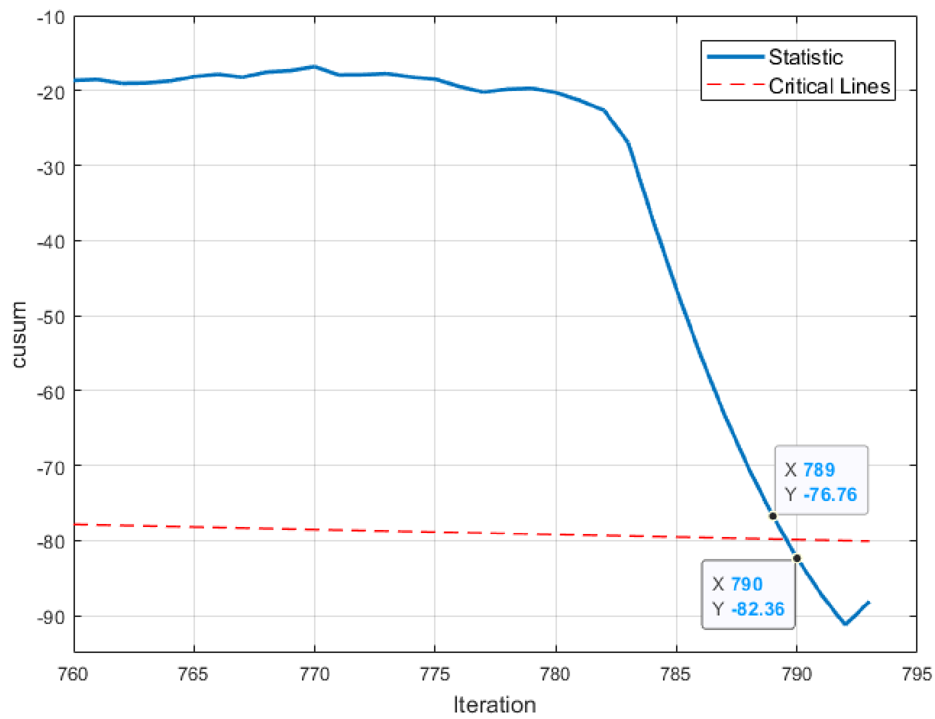


Figure 11. Zoomed data from Figure 10 displaying the detection process of F2.

As shown in Figure 7 in [48], the gearbox fault F2 was detected by the cointegration residual in the middle of the data samples 1230 and 1231. However, it is necessary to mention that this is the results using the first residual. The second residual, not shown in [48], identified the gearbox fault F2 between the data samples 1239 and 1240. In comparison with the results in Figure 11 of this paper, where the fault F2 has been detected at the data sample 1230, one can conclude that the gearbox fault F2 can be detected using the method proposed in this paper as early and accurately as being detected by the cointegration-based approach under the condition that the first residual is used.

The cointegration-based approach has been considered as an efficient tool for structural health monitoring (SHM) and condition monitoring (CM), especially for the removal of the influence of environmental and operational variations on damage-sensitive features. However, when the cointegration residuals are used for condition monitoring and fault/damage detection, the obtained results are diverse depending on which cointegration residual is used. The first cointegration residual has been found to be the best one in terms of early and accurate fault detection ability, as discussed in [38,51].

The CUSUM-based method proposed in this paper offers users a more straightforward and reliable solution for wind turbine condition monitoring, because they need to monitor only one sequence of the CUSUM test statistics (i.e., the damage-sensitive feature) within a control chart scheme.

6. Further Discussion and Conclusions

Given the well-established CUSUM test from the fields of econometrics and statistics for monitoring and detecting unexpected change over time in the parameters of data-driven regression models, the present work has developed a new condition monitoring and fault diagnosis approach for engineering systems and/or structures in general and for wind turbines in particular. A CUSUM-based computation procedure consisting of four steps has been proposed for this purpose. It starts with forming a multiple linear regression model using data collected while the system or structure of interest is operating in the normal condition without fault, then monitoring the stability of the regression coefficients of the model on every acquired observations, and detecting a structural change in the form of coefficient instability using the CUSUM test. The method falls in the category of semi-supervised algorithms because only a data set under the normal operating condition is used to initially form the model. The availability of the proposed method is based on the assumption that any coefficient instability indicates a structural change in the regression model, which can be interpreted as the occurrence of a fault in the system or structure. The proposed CUSUM-based method has been demonstrated through a wind turbine case study using temperature-related SCADA data. A multiple linear regression model is formed using gearbox and generator temperature data as the independent variables and generator speed data as the dependent variable. Two known problems of the wind turbine (an abnormal operating state F1 and a gearbox fault F2) were used to illustrate the fault detection ability of the proposed method. The results show that the method can effectively monitor the wind turbine and reliably detect both F1 and F2 at the early stage of its occurrence.

It is important to mention that the detection of structural change in the model (i.e., the change-point detection problem) using WT SCADA data is not a simple task. One of the main reasons is because wind turbines are subject to harsh environmental conditions. The effects of noise and/or environmental condition variations can cause or increase the problems of false positives and false negatives in fault detection results. For example, both a gearbox fault and an abrupt ambient temperature change can cause the test statistics to cross into the critical region. If this is the case then the CUSUM-based method proposed in this paper should be combined with other techniques for the removal of noise and/or environmental condition variations. Some potential methods are principal component analysis, signal subtraction method, and cointegration analysis.

The work presented in this paper has successfully applied the CUSUM-based approach for wind turbine condition monitoring and fault detection. However, this is still a feasibility study and the results presented in this paper are preliminary in the context of practical wind turbine condition monitoring applications. Therefore further research works are required to test the approach with other WT SCADA database (real-world test cases). Specially, the proposed methodology should be investigated for a large number of wind turbines with different types of fault/abnormal components. Moreover, the qualitative and quantitative comparisons of the proposed method with other existing approaches, such as artificial neural networks, support vector machines and decision trees, will be investigated in the future. It is expected in practice that the early fault detection should be at least some days in advance for preventing wind turbine damages. Therefore future study on adapting the CUSUM-based computation procedure to make it possible for early fault prognosis in wind turbines has been planned. A promising direction is to consider approaches to forecasting time series that are subject to multiple structural breaks, developed in the field of econometrics and statistics for macroeconometric model prediction. Furthermore, if the CUSUM test algorithm can be integrated with multiple nonlinear regression models, the fault detection results would be improved.

As mentioned at the beginning of the paper that this study in a broad view has aimed at developing a new approach, based on the CUSUM test, for condition monitoring and fault diagnosis of different engineering systems and/or structures. Since the proposed methodology is a general approach—which is simply based on the analysis of measurement data in terms of time series responses acquired from investigated processes or structures by sensors—the author believes that this method can be applied to a variety of engineering systems and/or structures. For example, vibration data of rotating machines or vibration responses (natural frequencies) of bridge structures would be suitable to be analysed by the presented approach. In summary, the promising results obtained in this paper suggest that the developed method should be further explored for SHM and CM problems.

Finally, as discussed in [52] the monitoring problem of a system or structure of interest in the common view can be considered as the problem of detecting one or several abrupt changes in the parameters of a static or dynamic stochastic process. If this is the case, then it would be useful for people working in the fields of SHM and CM to consider using change detection algorithms proposed in [52]. The book provides a unified framework for the design and performance analysis of (parametric) statistical approaches for on-line abrupt change detection problems together with the sufficient mathematical background necessary for this purpose.

Funding: This research received no external funding. The APC was funded by AGH University of Science and Technology.

Data Availability Statement: No new data were created in this study. Data sharing is not applicable to this article.

Acknowledgments: The author would like to thank Tomasz Barszcz from the AGH University of Science and Technology for the experimental wind turbine data. The author acknowledges the constructive and useful comments and suggestions of the Anonymous Reviewers and Academic Editor that have been used to improve the overall quality of this paper.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Andrews, D.W.K. Tests for parameter instability and structural change with unknown change point. *Econometrica* **1993**, *61*, 821–856. [CrossRef]
2. Chu, C.S.J.; Stinchcombe, M.; White, H. Monitoring structural change. *Econometrica* **1996**, *64*, 1045–1065. [CrossRef]
3. Zeileis, A.; Leisch, F.; Kleiber, C.; Hornik, K. Monitoring structural change in dynamic econometric models. *J. Appl. Econom.* **2005**, *20*, 99–121. [CrossRef]
4. Wang, C.S.H.; Xie, Y.M. Structural change and monitoring tests. In *Handbook of Financial Econometrics and Statistics*; Lee, C.F., Lee, J.C., Eds.; Springer: New York, NY, USA, 2015; pp. 873–902.
5. Kleiber, C. Structural change in (economic) time series. In *Complexity and Synergetics*; Müller, S., Plath, P., Radons, G., Fuchs, A., Eds.; Springer: Cham, Switzerland, 2018; pp. 275–286.
6. Antoch, J.; Hanousek, J.; Horváth, L.; Hušková, M.; Wang, S. Structural breaks in panel data: Large number of panels and short length time series. *Econom. Rev.* **2019**, *38*, 828–855. [CrossRef]
7. Lewis, L.T.; Monarch, R.; Sposi, M.; Zhang, J. *Structural Change and Global Trade*; Globalization Institute Working Paper No. 333; Federal Reserve Bank of Dallas: Dallas, TX, USA, 2020.
8. Comin, D.A.; Lashkari, D.; Mestieri, M. *Structural Change with Long-Run Income and Price Effects*; NBER Working Paper No. 21595; National Bureau of Economic Research: Cambridge, MA, USA, 2020.
9. Chow, G.C. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **1960**, *28*, 591–605. [CrossRef]
10. Brown, R.L.; Durbin, J.; Evans, J.M. Techniques for testing the constancy of regression relationships over time. *J. R. Stat. Soc. Ser. B* **1975**, *37*, 149–192. [CrossRef]
11. Krämer, W.; Ploberger, W.; Alt, R. Testing for structural change in dynamic models. *Econometrica* **1988**, *56*, 1355–1369. [CrossRef]
12. Turner, P. Power properties of the CUSUM and CUSUMSQ tests for parameter instability. *Appl. Econ. Lett.* **2010**, *17*, 1049–1053. [CrossRef]
13. Ilott, P.W.; Griffiths, A.J. Development of a Pumping System Decision Support Tool Based on Artificial Intelligence. In Proceedings of the 8th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 1996), Toulouse, France, 16–19 November 1996; pp. 260–267.
14. Leger, R.P.; Garland, W.J.; Poehlman, W.F.S. Fault detection and diagnosis using statistical control charts and artificial neural networks. *Artif. Intell. Eng.* **1998**, *12*, 35–47. [CrossRef]
15. Bin Shams, M.A.; Budman, H.M.; Duever, T.A. Fault detection, identification and diagnosis using CUSUM based PCA. *Chem. Eng. Sci.* **2011**, *66*, 4488–4498. [CrossRef]
16. Bin Shams, M.A.; Budman, H.M.; Duever, T.A. Fault Detection Using CUSUM Based Techniques with Application to The Tennessee Eastman Process. In Proceedings of the 9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010), Leuven, Belgium, 5–7 July 2010; pp. 109–114.
17. Du, Y.; Du, D. Fault detection and diagnosis using empirical mode decomposition based principal component analysis. *Comput. Chem. Eng.* **2018**, *115*, 1–21. [CrossRef]
18. Zhou, F.; Park, J.H.; Wen, C.; Hu, P. Average accumulative based time variant model for early diagnosis and prognosis of slowly varying faults. *Sensors* **2018**, *18*, 1804. [CrossRef] [PubMed]
19. Lampreia, S.P.G.F.d.S.; Requeijo, J.F.G.; Dias, J.A.M.; Vairinhos, V.M.; Barbosa, P.I.S. Condition monitoring based on modified CUSUM and EWMA control charts. *J. Qual. Maint. Eng.* **2018**, *24*, 119–132. [CrossRef]
20. Baghli, M.; Delpha, C.; Diallo, D.; Hallouche, A.; Mba, D.; Wang, T. Three-level NPC inverter incipient fault detection and classification using output current statistical analysis. *Energies* **2019**, *12*, 1372. [CrossRef]
21. Xu, Q.; Lu, S.; Zhai, Z.; Jiang, C. Adaptive fault detection in wind turbine via RF and CUSUM. *IET Renew. Power Gener.* **2020**, *14*, 1789–1796. [CrossRef]
22. Page, E.S. Continuous inspection scheme. *Biometrika* **1954**, *41*, 100–115. [CrossRef]

23. Nelson, C.R.; Plosser, C.I. Trends versus random walks in macroeconomic time series: Some evidence and implications. *J. Monet. Econ.* **1982**, *10*, 139–162. [CrossRef]
24. Global Wind Energy Council. Global Wind Report: Annual Market Update 2018. 2019. Available online: https://gwec.net/global-wind-report-2018/ (accessed on 31 January 2021).
25. Hyers, R.W.; McGowan, J.G.; Sullivan, K.L.; Manwell, J.F.; Syrett, B.C. Condition monitoring and prognosis of utility scale wind turbines. *Energy Mater.* **2006**, *1*, 187–203. [CrossRef]
26. Kusiak, A.; Li, W. The prediction and diagnosis of wind turbine faults. *Renew. Energy* **2011**, *36*, 16–23. [CrossRef]
27. Qian, P.; Ma, X.; Zhang, D. Estimating health condition of the wind turbine drivetrain system. *Energies* **2017**, *10*, 1583. [CrossRef]
28. 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. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1–39. [CrossRef]
29. Salameh, J.P.; Cauet, S.; Etien, E.; Sakout, A.; Rambault, L. Gearbox condition monitoring in wind turbines: A review. *Mech. Syst. Signal Process.* **2018**, *111*, 251–264. [CrossRef]
30. Wang, T.; Han, Q.; Chu, F.; Feng, Z. Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review. *Mech. Syst. Signal Process.* **2019**, *126*, 662–685. [CrossRef]
31. Stetco, A.; Dinmohammadi, F.; Zhao, X.; Robu, V.; Flynn, D.; Barnes, M.; Keane, J.; Nenadic, G. Machine learning methods for wind turbine condition monitoring: A review. *Renew. Energy* **2019**, *133*, 620–635. [CrossRef]
32. Hossain, M.L.; Abu-Siada, A.; Muyeen, S.M. Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review. *Energies* **2018**, *11*, 1309. [CrossRef]
33. Zhang, P.; Lu, D. A survey of condition monitoring and fault diagnosis toward integrated O&M for wind turbines. *Energies* **2019**, *12*, 2801.
34. Zaher, A.; McArthur, S.D.J.; Infield, D.G.; Patel, Y. Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy* **2009**, *12*, 574–593. [CrossRef]
35. Kusiak, A.; Zhang, Z. Analysis of wind turbine vibrations based on SCADA data. *ASME J. Sol. Energy Eng.* **2010**, *132*, 031008. [CrossRef]
36. Wilkinson, M.; Darnell, B.; Delft, T.V.; Harman, K. Comparison of methods for wind turbine condition monitoring with SCADA data. *IET Renew. Power Gener.* **2014**, *8*, 390–397. [CrossRef]
37. Sun, P.; Li, J.; Wang, C.; Lei, X. A generalized model for wind turbine anomaly identification based on SCADA data. *Appl. Energy* **2016**, *168*, 550–567. [CrossRef]
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. *Renew. Energy* **2018**, *116*, 107–122. [CrossRef]
39. Tautz-Weinert, J.; Watson, S.J. Using SCADA data for wind turbine condition monitoring—A review. *IET Renew. Power Gener.* **2017**, *11*, 382–394. [CrossRef]
40. Zolna, K.; Dao, P.B.; Staszewski, W.J.; Barszcz, T. Nonlinear cointegration approach for condition monitoring of wind turbines. *Math. Probl. Eng.* **2015**, *2015*, 1–11. [CrossRef]
41. Vidal, Y.; Pozo, F.; Tutivén, C. Wind turbine multi-fault detection and classification based on SCADA data. *Energies* **2018**, *11*, 3018. [CrossRef]
42. Astolfi, D.; Castellani, F. Wind turbine power curve upgrades: Part II. *Energies* **2019**, *12*, 1503. [CrossRef]
43. Feng, Y.; Qiu, Y.; Crabtree, C.J.; Long, H.; Tavner, P.J. Monitoring wind turbine gearboxes. *Wind Energy* **2013**, *16*, 728–740. [CrossRef]
44. Schlechtingen, M.; Santos, I.F. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mech. Syst. Signal Process.* **2011**, *25*, 1849–1875. [CrossRef]
45. Guo, P.; Infield, D.; Yang, X. Wind turbine generator condition-monitoring using temperature trend analysis. *IEEE Trans. Sustain. Energy* **2012**, *3*, 124–133. [CrossRef]
46. Abdusamad, K.B.; Gao, D.W.; Muljadi, E. A Condition Monitoring System for Wind Turbine Generator Temperature by Applying Multiple Linear Regression Model. In Proceedings of the 2013 North American Power Symposium (NAPS), Manhattan, KS, USA, 22–24 September 2013.
47. Astolfi, D.; Castellani, F.; Terzi, L. Fault prevention and diagnosis through SCADA temperature data analysis of an onshore wind farm. *Diagnostyka* **2014**, *15*, 71–78.
48. Dao, P.B. Condition monitoring of wind turbines based on cointegration analysis of gearbox and generator temperature data. *Diagnostyka* **2018**, *19*, 63–71. [CrossRef]
49. Wang, X.; Zhao, Q.; Yang, X.; Zeng, B. Condition monitoring of wind turbines based on analysis of temperature-related parameters in supervisory control and data acquisition data. *Meas. Control.* **2020**, *53*, 164–180. [CrossRef]
50. *Econometrics Toolbox*^{TM}; Release 2019b; The MathWorks Inc.: Natick, MA, USA, 2019.
51. Dao, P.B.; Staszewski, W.J. Cointegration approach for temperature effect compensation in Lamb wave based damage detection. *Smart Mater. Struct.* **2013**, *22*, 095002. [CrossRef]
52. Basseville, M.; Nikiforov, I. *Detection of Abrupt Changes: Theory and Application*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1993.