

## Article

# An Outlier Cleaning Based Adaptive Recognition Method for Degradation Stage of Bearings

Jingsong Xie <sup>1</sup>, Yujie Xie <sup>1</sup>, Tiantian Wang <sup>2</sup>  and Yougang Xiao <sup>1,\*</sup><sup>1</sup> School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China<sup>2</sup> School of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China

\* Correspondence: csuxyg@163.com

**Abstract:** Accurate identification of the degradation stage is key to the prediction of the remaining useful life (RUL) of bearings. The  $3\sigma$  method is commonly used to identify the degradation point. However, the recognition accuracy is seriously disturbed by the random outliers in the normal stage. Therefore, this paper proposes an adaptive recognition method for the degradation stage based on outlier cleaning. Firstly, an improved multi-scale kernel regression outlier detection method is adopted to roughly search the abnormal signal segments. Then, a method for the accurate locating of the start and end points of abnormal impulses is established. After that, indexes are constructed for screening abnormal segments and an iterative strategy is proposed to achieve an accurate and efficient removal of abnormal impulses. After outlier cleaning, the  $3\sigma$  approach is used to set the degradation warning threshold adaptively to realize the degradation stage recognition of the bearings. The PHM 2012 rotating machinery dataset is used to verify the effectiveness of the proposed method. Experimental results show that the proposed method can accurately locate and remove the outliers adaptively. After the cleaning of the outliers, the identification of the degradation stage is no longer disturbed by the selection of the reference signal of the normal stage and the robustness and the accuracy of the degradation stage identification have been improved significantly.

**Keywords:** degradation stage recognition; outlier cleaning; accurate impulse locating; RUL prediction



**Citation:** Xie, J.; Xie, Y.; Wang, T.; Xiao, Y. An Outlier Cleaning Based Adaptive Recognition Method for Degradation Stage of Bearings. *Sensors* **2022**, *22*, 6480. <https://doi.org/10.3390/s22176480>

Academic Editor: Hossam A. Gabbar

Received: 12 July 2022

Accepted: 24 August 2022

Published: 28 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Prognostics and health management (PHM) is an important technology in modern industries that can enhance the reliability of mechanical equipment, reduce unnecessary maintenance costs and avoid catastrophic accidents [1–3]. The performance of bearings (one of the key components of rotating machinery) directly affects the safe and reliable operation of mechanical equipment. In order to maintain running time and to prevent abnormal failure, it is necessary to monitor the performance degradation of bearings in the operation state.

Generally, the degradation process of bearings can be divided into two stages, namely the normal stage and the irreversible degradation stage [4]. Ideally, in the normal operation stage, the bearing is in a healthy state and the monitoring data fluctuate randomly in a small range with no significant difference. In the irreversible degradation stage, the bearing will present early failure and then begin to degrade, and the degradation degree becomes more and more serious as the running time increases. In order to estimate the degradation accurately and to improve the prediction accuracy of the RUL, different methods were proposed. In some research [5,6], scholars artificially and subjectively determined the degradation stage of bearings. Li et al. [7] determined that when the health index of bearings exceeded the  $3\sigma$  interval, this was the dividing point between the normal and the degradation stage. Kong et al. [8] proposed an adaptive time selection method for the first prediction based on the  $3\sigma$  approach. Singh et al. [9] applied Chebyshev's inequality to the identification of the locations of the bearing health state change points.

In the above  $3\sigma$  approach, a piece of healthy signal data is selected as the standard to build the  $3\sigma$  threshold for the degradation warning. However, due to inappropriate selections of the health stage, it is easy to make an identification error of the degradation stage. Therefore, Pan et al. [10] determined the degradation threshold according to the continuous gradient value of health indicators with running time. Gao et al. [11] proposed a change point degradation model with random jump and considered a random change point that was influenced by the accumulative effect of a shock process. In their work, human involvement and the application of professional knowledge are required. Different working conditions will bring various performance degradation trends of bearings. Therefore, it is not universal to choose fixed parameters under specific working conditions. In addition, some scholars introduced the neural network into the recognition of the bearing's degradation stage. Li et al. [12] generated an adjournment network to learn the data distribution of bearing health stages to determine the change point of the degradation state. Baptista et al. [13] used a recursive neural network to find the location of the degradation point. Wu et al. [14] classified a system health stage based on the trained model and real-time sensory data. However, in actual work, vibration signals in the normal operation stage of bearings will be covered by noise and random abnormal shocks [15] and the constructed degradation health indicators show serious shocks and poor stability and monotonicity, thus affecting the accurate recognition of the degradation stage.

In order to reduce the influence of random abnormal impulses in the normal operation stage, scholars used empirical mode decomposition (EDM), wavelet threshold denoising, Fourier transform (FFT) and other signal processing methods to reduce the noise of bearing vibration signals [16–18]. Although such signal processing methods may work well in some cases, they will weaken the original signal to some extent or remove some key signals altogether, which may lead to low prediction accuracy. In addition, Li et al. [19] used the exponential weighted moving average (EWMA) method to smooth the health indicators. If the oscillation and deviation of the current point were large, the EWMA method could not obtain a good smoothing effect. Chen et al. [20] proposed a fault impulse method named the Intelligent Impulse Finder to accurately identify and locate the fault impulses; however, there are cases of missed or false detections. M. Gupta et al. [21] used the outlier detection method to look for local abnormal areas, but since the detected abnormal areas were formed by some continuous outlier points, which sometimes did not exist, the outlier detection method could only eliminate the oscillation phenomenon of health indicators locally according to the outlier area.

To solve the above problems, this paper proposes a degradation stage recognition method based on outlier cleaning. In the global perspective, an improved outlier detection method based on the local outlier factor is used to detect abnormal signal segments adaptively. For the local abnormal signal segment, according to the kernel density estimation and the analysis of the gradient value of the abnormal impulse envelope gradient value to achieve an accurate impulse location. Then, a screening index of abnormal signal segments and an iterative removal strategy are constructed to removing abnormal impulses, thereby reducing the time and obtaining a better effect for data cleaning. Finally, a health index representing the degradation characteristics of the mechanical equipment is constructed for the vibration signals after abnormal processing. The degradation threshold set by the  $3\sigma$  method is used to identify the degradation stage of the bearings.

The main contributions are as follows:

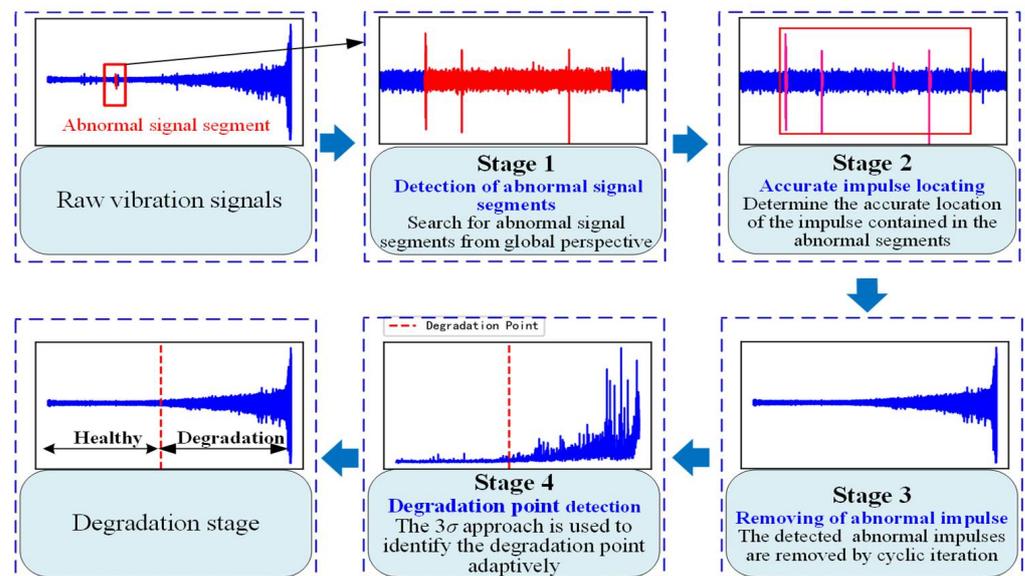
- (1) An outlier detection method is constructed by combining the global search of abnormal signal segments and the accurate locating of local abnormal impulses. Adopting the constructed method, the normal signals misjudged as outliers in abnormal signal segments can be greatly reduced and the normal data can be fully retained.
- (2) A screening criteria and an iteration removal strategy are proposed, achieving an adaptive detection of segments that contain abnormal impulses and a quick removal of abnormal segments.

- (3) After outlier cleaning, the  $3\sigma$  approach is used to set the degradation warning threshold adaptively to realize the degradation stage recognition of bearings. The outlier cleaning avoids the interference of the selection of reference signals and improves the accuracy and stability of the degradation stage recognition.

This paper is organized as follows: Section 2 describes the algorithm principle and the method flow of the proposed method. In Section 3, the proposed method is verified by the degradation data of rolling bearings. Some conclusions are made in Section 4.

## 2. The Proposed Method

The proposed method is shown in Figure 1, including four key steps: (1) global detection of abnormal signal segments; (2) accurate locating of local abnormal impulses; (3) outlier removal method; (4) degradation point adaptive detection.



**Figure 1.** The flow chart of the proposed method.

### 2.1. Detection Method for Abnormal Signal Segments

The kernel regression based abnormal detection method is adopted to search for abnormal signal segments globally that are contained in the degraded full life data. Firstly, the time-domain statistical features are extracted from the raw vibration signals. Then, the local outlier factor of these features based on kernel regression are calculated to assess the abnormality degree of each data segment. Finally, the abnormal data segments can be detected according to the local outlier factor and the anomaly determination threshold.

#### 2.1.1. Data Preprocessing

The raw vibration signal data are divided into segments by the sliding window, and the time-domain features are extracted from each segment. In order to obtain great detection performance, appropriate features should be selected to constitute a feature set for calculating the local outlier factor. The features adopted include mean, maximum, minimum, peak-to-peak, wave peak factor, variance, root mean square, mean amplitude, root amplitude, waveform factor, impulse factor, margin factor, kurtosis, peak and skewness [22].

### 2.1.2. Local Outlier Factor

The local outlier factor detects outliers by comparing the density of each sample in the dataset with the density of sample points in the field. The local outlier factor of a point  $p$  in the dataset is calculated as follows:

$$LOF_k = \frac{\sum_{o \in N_k(p)} \frac{lrd_k(o)}{lrd_k(p)}}{|N_k(p)|} = \frac{\sum_{o \in N_k(p)} lrd_k(o)}{|N_k(p)|} / lrd_k(p) \quad (1)$$

where  $k$  is the  $k$ th neighborhood of  $p$ , and  $o$  is a point in the  $k$ th neighborhood of  $p$ .  $lrd_k(\cdot)$  is the locally accessible density; it is obtained according to:

$$lrd_k(p) = 1 / \left( \frac{\sum_{o \in N_k(p)} reach - dist_k(p, o)}{|N_k(p)|} \right) \quad (2)$$

where  $N_k(p)$  is the  $k$ th distance neighborhood of  $p$ :  $N_k(p) = \{q \in D \mid d(p, q) \leq kd(p)\}$ ;  $reach-dist_k(p, o)$  is the reachability distance of the point  $p$  to  $o$ :  $reach-dist_k(p, o) = \max\{kd(o), d(p, o)\}$ ;  $kd(p)$  is the  $k$ th distance between  $p$  and a point  $o \in D$ , which satisfies the following two criteria: (1) there are at least  $k$  objects  $q$ , such that  $d(p, q) \leq d(p, o)$ ; (2) there are at most  $k - 1$  objects  $q$ , such that  $d(p, q) < d(p, o)$ .  $d(p, q)$  is the distance between point  $p$  and  $q$  and defined as:

$$d(p, o) = \sqrt{\sum_{i=1}^{\dim} (f_i(p) - f_i(o))^2} \quad (3)$$

where  $f_i(p)$  and  $f_i(o)$  are the time-domain features of points  $p$  and  $o$ , respectively, and  $\dim$  is the dimension of features. Some basic concepts in this expression can be referred to in the paper [23].

### 2.1.3. Local Outlier Factor Based on Multi-Scale Kernel Regression

The detection of outliers by the local outlier factor method is based on local density estimation, which indicates that a good detection performance requires an accurate density estimation. The outlier method roughly estimates the local density by calculating the reciprocal of the average reachable distance between a given object and its neighborhood points; hence, the method may fail to detect true outliers. Moreover, the local outlier factor is calculated in the  $k$ -distance neighborhood, so the local outlier factor is very sensitive to parameter  $k$ . If the outliers are randomly distributed in several clusters with different densities, the local outlier factor method may obtain poor outlier detection results because the parameter  $k$  cannot meet the detection requirements of all outliers. In order to overcome the above shortcomings, Nadaraya–Watson kernel regression can be used to calculate regression estimators in the neighborhood of  $k$  distance to construct a more robust local outlier factor [23] in order to construct the final index for anomaly evaluation. Based on the Nadaraya–Watson kernel regression, the local outlier factor expression of  $p$  is calculated as follows:

$$LOF_s(p) = \frac{\sum_{o \in N_k(p)} \frac{1}{(kd(o))^\gamma} K\left(\frac{d(p, o)}{kd(o)}\right) LOF_k(o)}{\sum_{o \in N_k(p)} \frac{1}{(kd(o))^\gamma} K\left(\frac{d(p, o)}{kd(o)}\right)} \quad (4)$$

where  $s$  is the step of iteration;  $\gamma$  is the sensitivity, which defaults to 2;  $d(p, o)$  is the distance between point  $p$  and  $o$ ;  $kd(\cdot)$  is the  $k$ th distance;  $LOF_k(\cdot)$  is a local outlier factor;  $K(\cdot)$  is the kernel function, and the kernel function of parameter  $x$  is shown as follows:

$$K(x) = \begin{cases} \beta & , \text{ if } \|x\| \leq 1 \\ \beta \cdot \exp\left(-(\|x\| - 1)^2 / 2\right) & , \text{ otherwise} \end{cases} \quad (5)$$

where  $\beta$  is a constant and is set to 1.

#### 2.1.4. Outliers Determination

The abnormal index is constructed based on the multi-scale kernel regression function representing the deviation degree from the normal sample. During the normal operation stage of bearings, most of the signal data are in a normal state and belong to the same cluster. The index value of the abnormal data is much higher than that of the normal data. In order to detect outliers more accurately, when the outlier index is greater than the set anomaly threshold  $T$ , it is judged as an outlier.

$$T = \mu + 3\sigma \quad (6)$$

where  $\mu$  is the average value and  $\sigma$  is the standard deviation.

#### 2.2. Accurate Locating Method for Abnormal Impulses

The abnormal signal segments obtained from the global perspective contain a large number of normal signal data and a small number of abnormal impulses. In order to retain more effective raw vibration signal data, the impulse contained in the abnormal segments will be located accurately by the determining of the start and end points.

##### 2.2.1. Start Point Identification Methods for Impulses

In the abnormal signal segment, the distribution of the normal signal and the abnormal impulse is unknown. In order to estimate the distribution of the population sample data in the abnormal signal segment, the kernel density estimation is used to estimate the probability distribution of the normal signal and the abnormal impulse. In the degradation process of the bearings, the local extreme value of the outlier is stronger than that of the signal in the normal stage, so the local extreme value can well reflect the difference between the normal signal and the abnormal impulse.

First, the local extreme values ( $x_1, x_2, \dots, x_n$ ) for each sliding window of the abnormal signal segments are extracted. Then, the kernel density estimate of the extreme value sequence is calculated. The formula of the kernel density estimate is as follows:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (7)$$

where  $h$  is the bandwidth and  $K(\cdot)$  is the kernel function and

$$K(X) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}X^2\right) \quad (8)$$

$$h = c \cdot n^{-1/5} \quad (9)$$

where  $c = 1.05 \cdot \sigma$  and  $\sigma$  is the standard deviation [24].

Because the abnormal signal segment contains a normal signal and an abnormal impulse, the probability distribution calculated by the kernel density estimation shows a bimodal distribution [25]. Therefore, the minimum at the trough between these two peaks is regarded as the threshold for separating the normal signal and the abnormal impulse. When the local extremum exceeds the threshold for the first time, the corresponding position is judged as the start point of the abnormal impulse.

##### 2.2.2. End Point Identification Methods for Impulses

The amplitude of an impulse rises sharply after the start point and then is attenuated in an exponential form approximately. Therefore, the envelope of the abnormal impulse can be fitted by the exponential function. In order to determine the end point of the abnormal impulse, Hilbert transform is adopted to obtain the envelope of the impulse and then the exponential function is used to fit the envelope.

The end point of the abnormal impulse is considered as the point where the gradient of the envelope is less than the critical value. Therefore, the end points of the abnormal impulse are determined by analyzing the gradient of the fitting function.

With the step size of 1, the gradient value of the signal segments with  $n$  points of the fitting function is calculated successively:

$$g_i = \frac{y_i(n) - y_i(1)}{n - 1} \quad (10)$$

where  $y_i$  is the  $i$ -th signal segment of the fitting function and  $g_i$  is the gradient of the  $i$ -th signal segment. When the gradient  $g_i$  reaches the set threshold, the corresponding position is considered as the end point of the abnormal impulse.

### 2.3. Removing Method for Outliers

After the detection of the abnormal signal segments and the accurate locating of impulses, how to remove the outliers efficiently and accurately is the key problem to be solved.

The local outlier factor detection method based on kernel regression can only detect the abnormal signal segments that have a higher anomaly degree compared with normal data relatively, and it cannot identify all abnormal segments in one detection process. In addition, the detection effect of the outliers depends on the parameters of the sliding window and the  $k$  value of the  $k$ -distance neighborhood. If the parameter steps are too small in the search process, the calculation efficiency will be reduced and the time of outlier cleaning will be greatly increased.

Therefore, the iterative removal strategy is proposed to overcome the above difficulties. As shown in Figure 2, the window length  $w$  and the  $k$  value are constantly updated in each iteration to obtain the optimum parameter. In this strategy, the detection accuracy and the detection efficiency of the abnormal signal segments are improved.

The abnormal segments identified by the kernel-regression-based method are not only the abnormal segments of impulse types that have significant interference with the identification of degradation stage but also other types of abnormal signal segments with almost no interference in the identification of the degradation stage. Therefore, a screening criteria is constructed to realize the selection of abnormal signals, so that only the abnormal signal segments containing significant impulses are selected for further locating, thus avoiding the time-consuming process of locating other abnormal signal segments with insignificant impulses.

The screening criteria for impulse-type abnormal signal segments are as follows, that is, the maximum vibration signals in an abnormal segment should be greater than that of the vibration signals in the optimal sliding window length at the two adjacent ends:

$$\begin{aligned} \max(\{x_1, x_2, \dots, x_n\}w_i) &\geq \max(\{x_1, x_2, \dots, x_n\}w_{i-1}) \\ \max(\{x_1, x_2, \dots, x_n\}w_i) &\geq \max(\{x_1, x_2, \dots, x_n\}w_{i+1}) \end{aligned} \quad (11)$$

where  $\{x_1, x_2, \dots, x_n\}$  is the vibration signal data of the bearings;  $n$  is the sliding window length;  $\{\cdot\}w_i$  is the  $i$ -th sliding window;  $\max(\cdot)$  is the maximum.

Finally, using the abnormal signal segment detection method and the screening criteria for impulse-type abnormal signal segments, combined with the accurate locating method for impulses and the iterative removal strategy, the impulse-type outliers with significant interference in the identification of the degradation stage can be cleared efficiently and accurately.

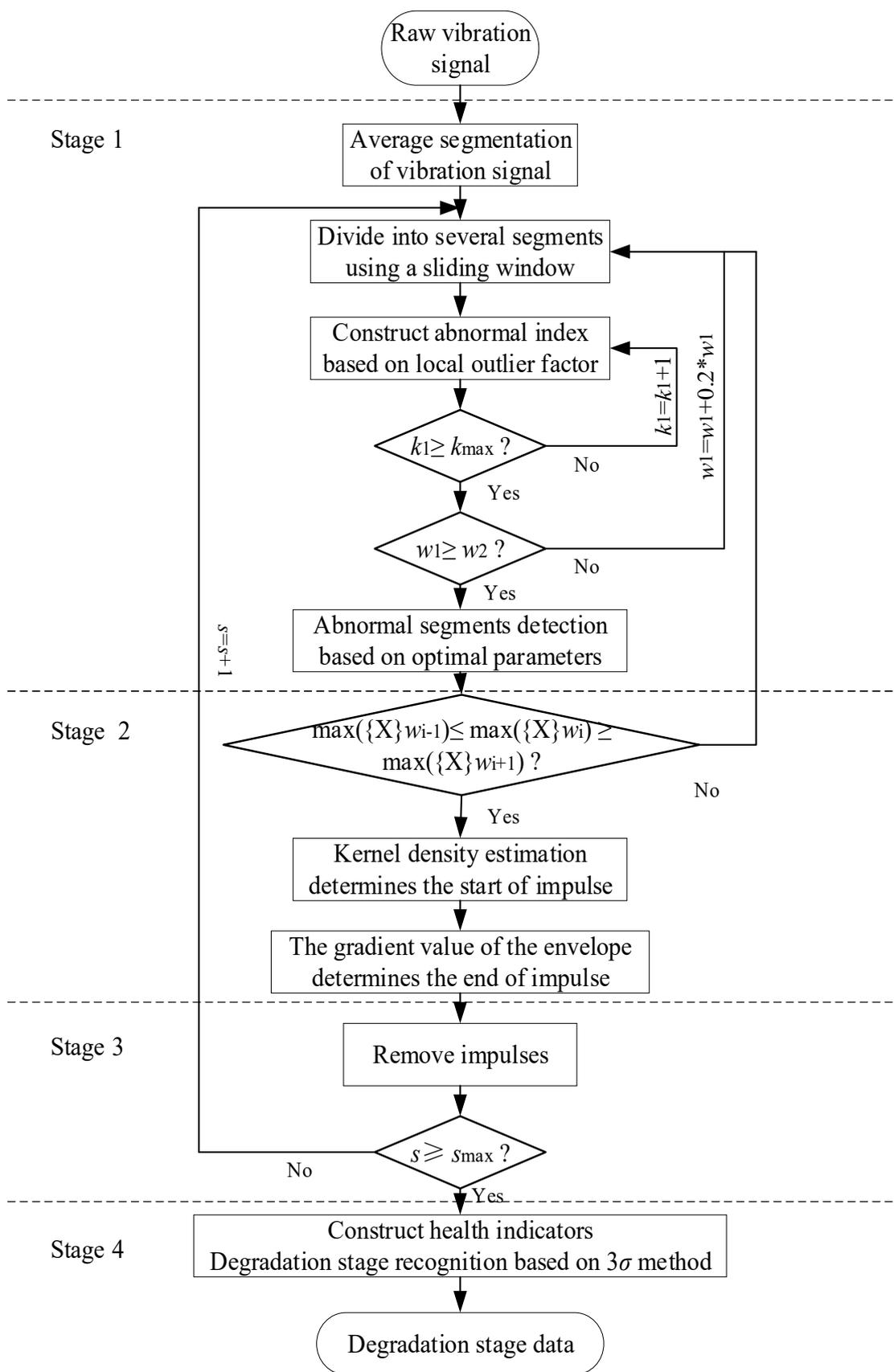


Figure 2. The algorithm flow chart of the proposed method.

#### 2.4. Degradation Point Detection

The  $3\sigma$  approach has been widely used in the degradation stage recognition of bearings [26] and it can set the degradation warning threshold adaptively according to the degradation characteristics of the bearing itself. The method is based on the theory of Chebyshev inequality [27]:

$$P(|X - \mu| \geq K\sigma) \leq \frac{1}{K^2} \quad (12)$$

where  $P$  is the occurrence probability of event;  $X$  is a random variable;  $K$  is any positive real number;  $\mu$  is the mean;  $\sigma$  is the standard deviation.

According to Formula (12), the probability of sample points in the random variable  $X$  falling outside  $K\sigma$  is not greater than  $1/K^2$ .

Therefore, it can be inferred from Formula (12) that when  $K$  is set to 3, the probability of any sample falling outside the interval of  $3\sigma$  (i.e.,  $[\mu - 3\sigma, \mu + 3\sigma]$ ) is less than or equal to  $1/9$ . At the same time, in order to further limit the interference caused by random anomalies, an improved  $3\sigma$  strategy is adopted [8], that is,  $t_i$  is the degradation point only when the following conditions is satisfied:

$$|X(t_i) - \mu| \geq 3\sigma, i = l, l + 1, \dots, l + v - 1 \quad (13)$$

where  $v$  is usually taken as 5,  $\mu$  is the mean and  $\sigma$  is the standard deviation of the reference signals in the normal stage.

In the detection of the degradation point of bearings, the health indicators should be constructed first. Based on the improved  $3\sigma$  approach, when the health indicators exceed the  $3\sigma$  interval for five times in a row, the corresponding point can be detected as the degradation point to divide the normal and the degradation stage.

The algorithm flow of degradation stage recognition based on outlier cleaning is shown in Figure 2. In Stage 1, the parameters that affect the detection results of the abnormal signal segments include  $k$  in the local outlier factor and the sliding window length  $w$ .  $k_1$  denotes the initial iteration value of parameter  $k$ , and  $k_{max}$  denotes the maximum of parameter  $k$ . The initial window length  $w_1$  is determined as 500 and the step length is 20% of the window length.  $w_2$  denotes the maximum of the sliding window and it depends on the length of the whole data  $l$ :  $w_2 = l/k_1$ . These parameters are constantly updated to search for the best abnormal signal segments from a global perspective. In Stage 2, if the signal segments meet the screening conditions of the impulse-type abnormal signal segments, the start and end points of the abnormal impulse can be further determined.

### 3. Experiments and Analysis

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

In this section, a PHM 2012 dataset is used to verify the proposed degradation stage recognition method. The detection effects of outliers and the degradation stage are compared with that of the current methods.

#### 3.1. PHM 2012 DataSet

The PHM 2012 challenge dataset used in this study was collected on a PRONOSTIA platform that accelerated the degradation of bearings, as shown in Figure 3. The motor, bearing, gearbox and two pulley assemblies were used to change the speed of the rolling bearing. The acceleration sensor was fixed on the outer ring of the bearing and collected vibration signals in the horizontal and vertical directions, respectively. Similarly, only horizontal vibration data are used in this study. This dataset contains 17 bearing failure data from run to failure under different working conditions [28]. The sampling frequency is 25.6 kHz and 2560 data points are collected every 10 s. For safety, the test was stopped when the amplitude of the vibration data exceeded 20 g.

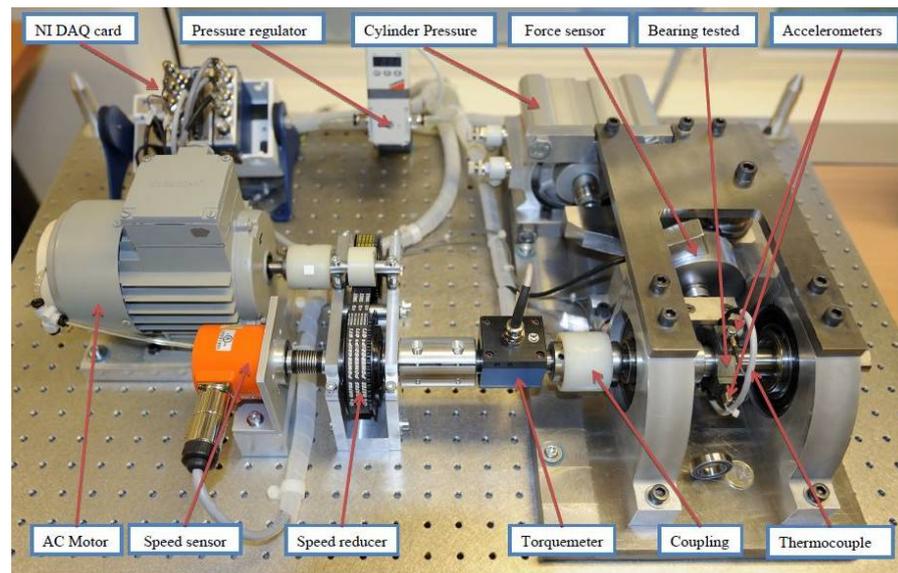


Figure 3. PRONOSTIA test platform.

### 3.2. XJTU-SY Data Set

The XJTU-SY dataset is provided by Xi'an Jiaotong University and Changxing Sumyong Technology Company [29] and the test platform is shown in Figure 4. The test platform consists of a series of components that produce the run-to-failure vibration data under different operating conditions. The radial force generated by the hydraulic loader is applied to the bearing housing of the bearing under test to simulate different working conditions, and the shaft speed is controlled by the motor speed controller. The details of the run-to-failure data for the tested bearings can be found in [29]. The model of the test bearings is LDK UER204, and 32,768 data points in 1.28 s are collected every 1 min with a sampling rate of 25.6 kHz. For this dataset, only horizontal vibration data are used in this study. The accelerated degradation bearing tests are stopped when the vibration amplitude exceeds 20 g for safety, and that moment is considered as the failure time for the bearings.

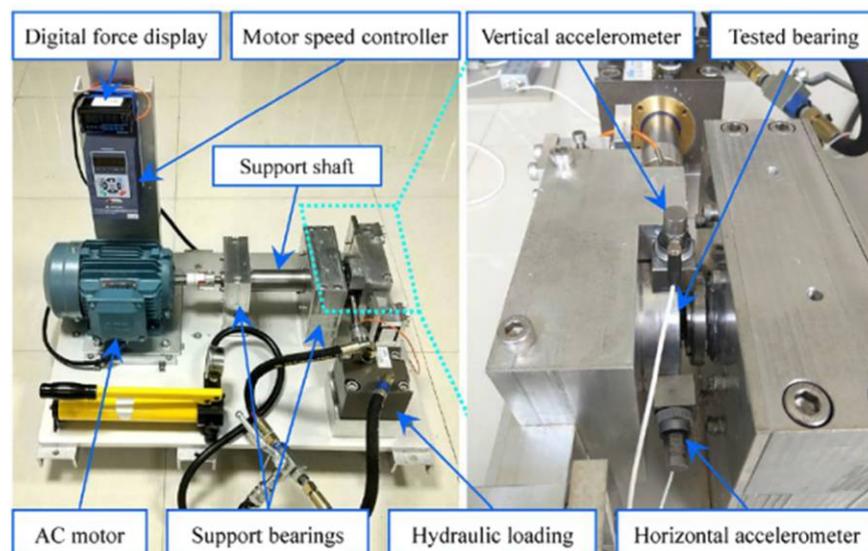
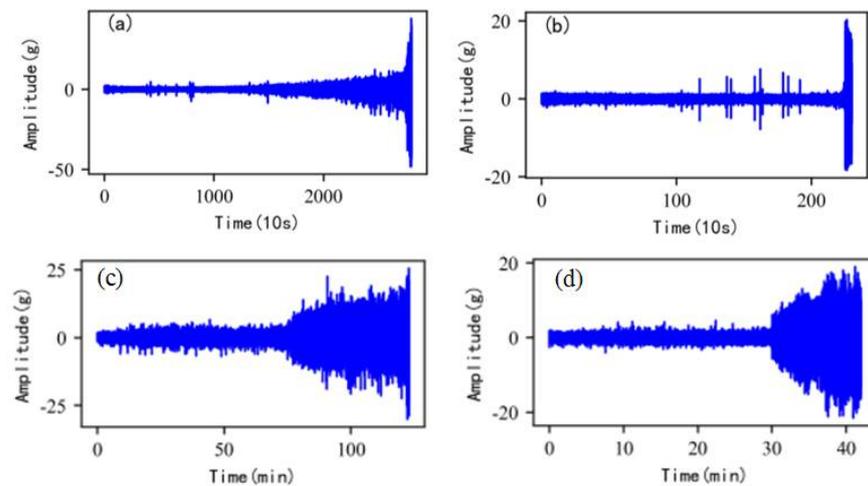


Figure 4. XJTU-SY test platform.

As shown in Figure 5, the full lifecycle accelerated degradation data of bearing 1 and bearing 2 (PHM 2012 dataset) and bearing 3 and bearing 4 (XJTU-SY dataset) are used to identify the proposed method.



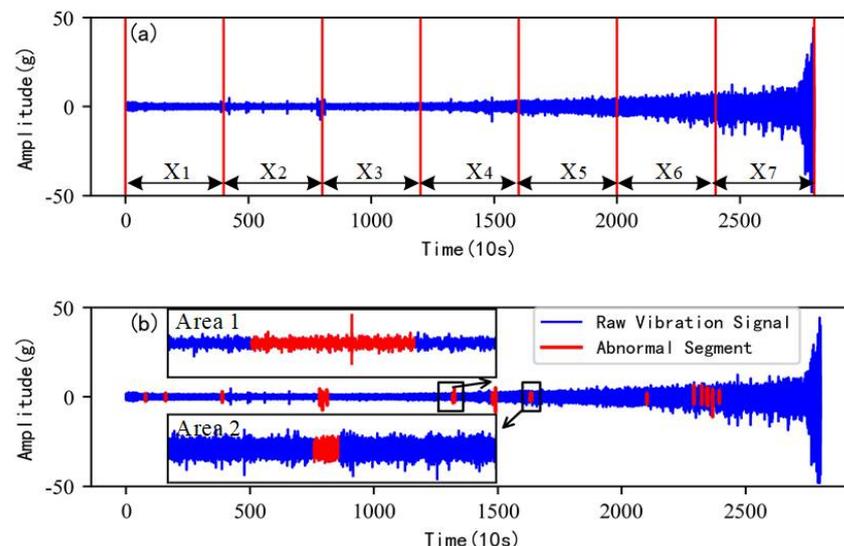
**Figure 5.** Full lifecycle vibration accelerated degradation data of bearings. (a) Bearing 1. (b) Bearing 2. (c) Bearing 3. (d) Bearing 4.

### 3.3. Results and Analysis

#### 3.3.1. Outliers Detection

In order to accurately detect the abnormal segments in each stage, the whole-life vibration signals of the bearings are divided into a number of segments with the same time interval, and the outliers are detected, respectively.

Figure 6a shows the segmented results of the raw vibration signals of bearing 1. At this time, the full lifecycle vibration signals of bearing 1 are divided into segments on average. According to the global detection method of abnormal segments introduced in Section 2, the initial length of the sliding window  $w_1$  is set to 500 and the initial  $k_1$  of the local outlier factor is set to 5. The sliding window length and the  $k$  are constantly updated in iteration to find the optimum parameters.



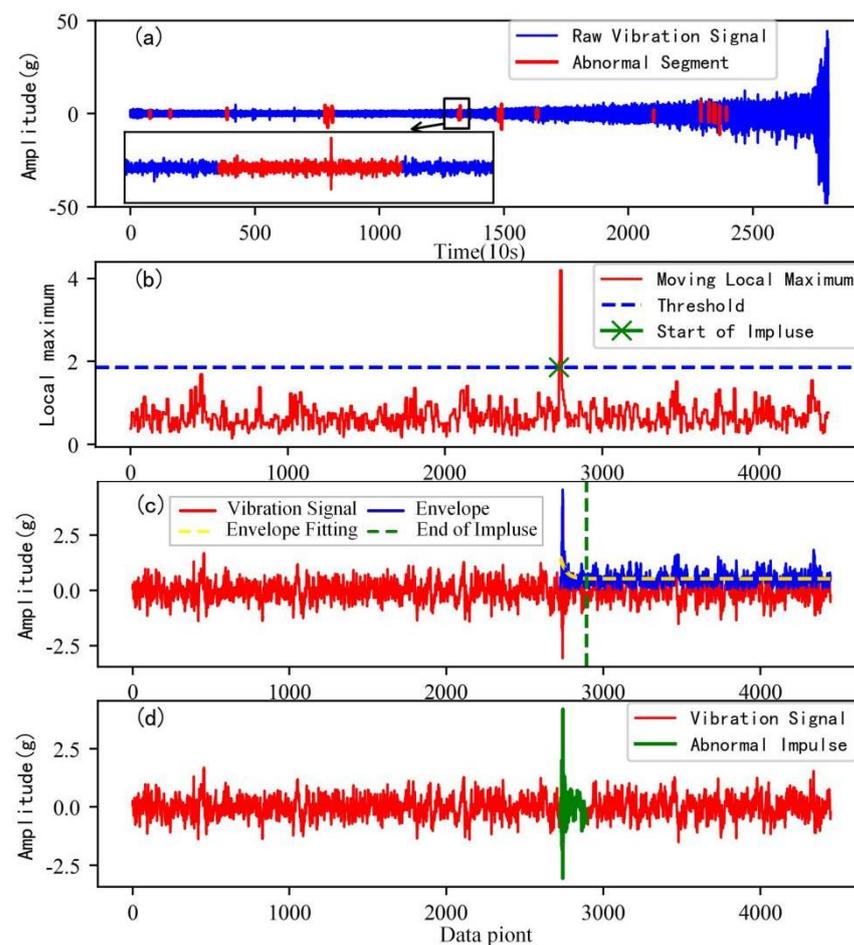
**Figure 6.** Abnormal signal segments detection of bearing 1. (a) Segmented result of the raw vibration signal. (b) An iteration of abnormal signal segment detection.

After a round of iteration, the global detection results of the abnormal segments are shown in Figure 6b and the signals marked in red are considered as abnormal segments. As can be seen in the local zoom area, the abnormal segments are not only the impulse-type signals (Area 1) that have significant interference with the identification of the degradation

stage but also other types of abnormal signal segments (Area 2) with almost no interference with the identification of the degradation stage.

Based on the proposed removing method in Section 2.3, the abnormal segment screening criteria in Formula (11) is adopted to select the impulse-type abnormal segments for the further locating and removing of abnormal impulses.

Figure 7 shows the process of the accurate locating of an abnormal impulse. As can be seen in the local zoom impulse-type abnormal segments in Figure 7a, the detected abnormal segments contain a lot of normal signal data and a very short impulse signal. It is necessary to determine the accurate start and end points of the impulses in order to retain the original normal data to the greatest extent. The local zoom impulse-type abnormal segments in Figure 7a are taken as an example to detect the accurate locating of abnormal impulses.



**Figure 7.** Accurate locating of abnormal impulse of bearing 1. (a) Detection of abnormal signal segments; (b) the start point of the abnormal impulse; (c) the end point of the abnormal impulse; (d) accurate locating of abnormal impulse.

As shown in Figure 7b, the threshold to judge the start point of abnormal impulses is obtained by the kernel density estimation method introduced in Section 2.2. The start point of an abnormal impulse is detected when the sliding local extremum exceeds the threshold for the first time.

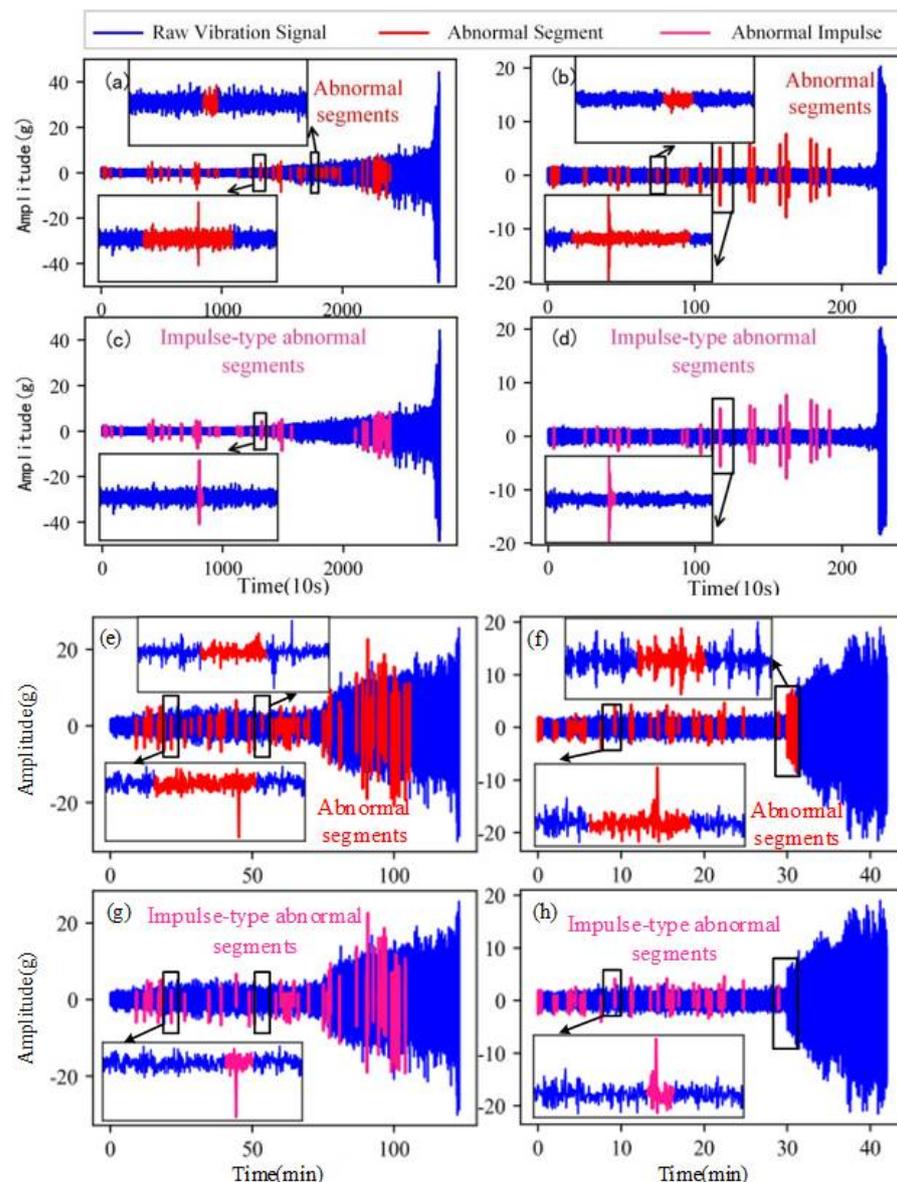
Figure 7c shows the detection result of the end point of the abnormal impulse. According to the proposed method in Section 2.2, the gradient threshold is set as 0.3. In this paper, when the gradient  $g$  of the envelope fitting function reaches the set threshold, the corresponding position is detected as the end point of the abnormal impulse.

Figure 7d shows the accurate impulse detection results of the abnormal segments. It indicates that the proposed method can accurately identify the location of abnormal

impulses and can avoid a large number of normal data being mistakenly removed as abnormal data.

### 3.3.2. Comparison of Outlier Detection Effects

Figure 8 is the comparison of the outlier detection results. Figure 8a,b,e,f are the outlier detection results based on the local outlier factor [30] of bearing 1, bearing 2, bearing 3 and bearing 4, respectively. A total of 72, 25, 63 and 37 abnormal signal segments are detected in bearing 1, bearing 2, bearing 3 and bearing 4, respectively, containing impulse-type abnormal segments and non-impulse-type abnormal segments. Each impulse-type abnormal segment (marked with red) only contains a lot of normal signal data and a very short impulse signal.



**Figure 8.** Comparison of outlier detection effects. (a) Detection result by proposed method of bearing 1. (b) Detection result by proposed method of bearing 2. (c) detection result based on local outlier factor of bearing 1. (d) Detection result based on local outlier factor of bearing 2. (e) Detection result by proposed method of bearing 3. (f) Detection result by proposed method of bearing 4. (g) Detection result based on local outlier factor of bearing 3. (h) Detection result based on local outlier factor of bearing 4.

Figure 8c,d,g,h are the outlier detection results based on the proposed method. After the selection of the impulse-type abnormal segments by the proposed removing method in Section 2.3, the number of the abnormal signal segments is reduced to 54, 23, 47 and 29, in bearing 1, bearing 2, bearing 3 and bearing 4, respectively. Furthermore, adopting the proposed accurate locating method in Section 2.2, the impulse positions are detected accurately and a large number of normal data avoid being removed.

The important principle of data cleaning is not to significantly change the original data. Compared with the outlier detection method based on the local outlier factor (LOF), the proposed method greatly reduces the removing of normal data. The comparison results of data points after outlier removing are shown in Table 1. It can be seen that the proposed method saves 315,662, 34,052, 9437 and 3834 data points on bearing 1, bearing 2, bearing 3 and bearing 4, respectively.

**Table 1.** Comparative statistics of vibration data points by outlier processing.

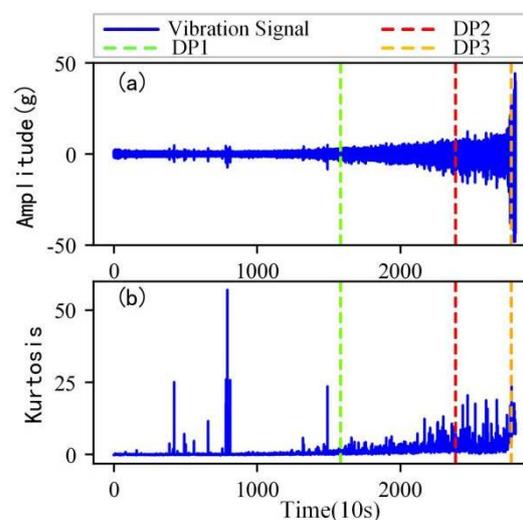
	Data Point			
	Raw Vibration Signal	Method Based on LOF	Proposed Method	Savings
Bearing 1	7,175,680	6,803,081	7,118,743	315,662
Bearing 2	588,800	552,524	586,576	34,052
Bearing 3	314,880	299,469	308,906	9437
Bearing 4	107,520	100,125	103,959	3834

### 3.3.3. Comparative of Degradation Stage Recognition

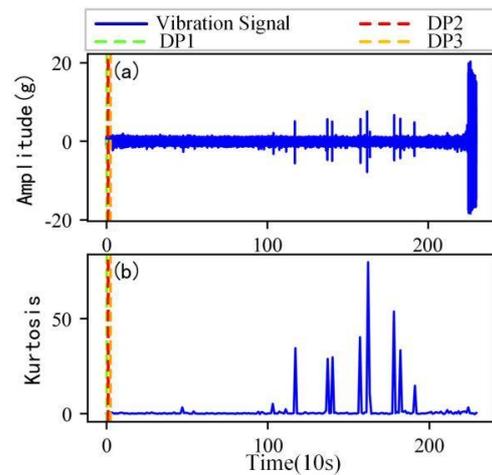
Kurtosis is extremely sensitive to the early failure of the bearing; it is generally considered that mechanical components begin to degrade once the early failure occurs [7]. So, kurtosis is selected as an exemplar indicator to characterize the degradation state and the improved  $3\sigma$  method is used to identify the degradation point, as shown in Figures 9–16. The kurtosis expression is calculated as follows:

$$K_v = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n x_i^2\right)^2} \quad (14)$$

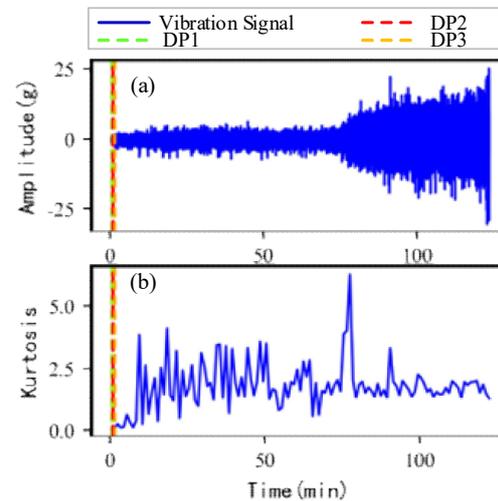
where  $x_i$  is the vibration signal data of bearings,  $i = 0, 1, 2, \dots, n$ ;  $\bar{x}$  is the mean.



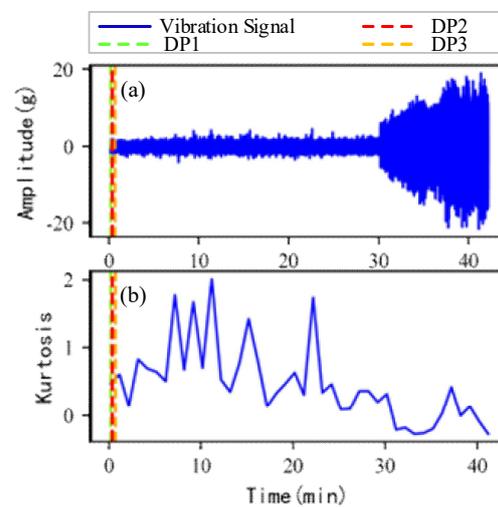
**Figure 9.** Degradation point identification of bearing 1 before outlier cleaning. (a) Vibration signal. (b) Kurtosis.



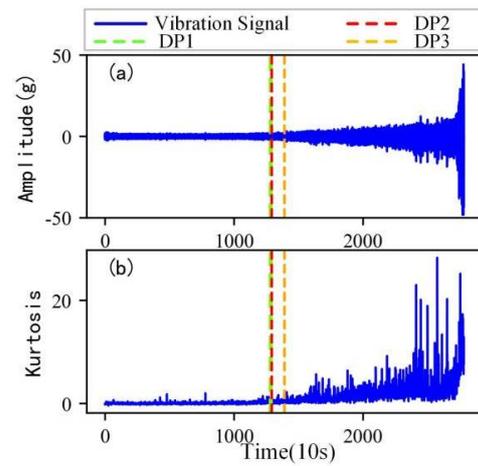
**Figure 10.** Degradation point identification of bearing 2 before outlier cleaning. (a) Vibration signal. (b) Kurtosis.



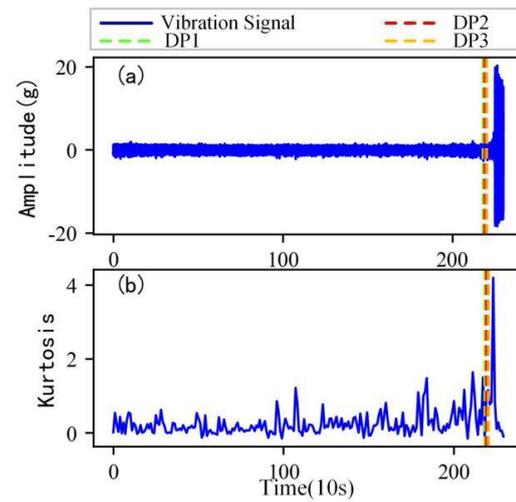
**Figure 11.** Degradation point identification of bearing 3 before outlier cleaning. (a) Vibration signal. (b) Kurtosis.



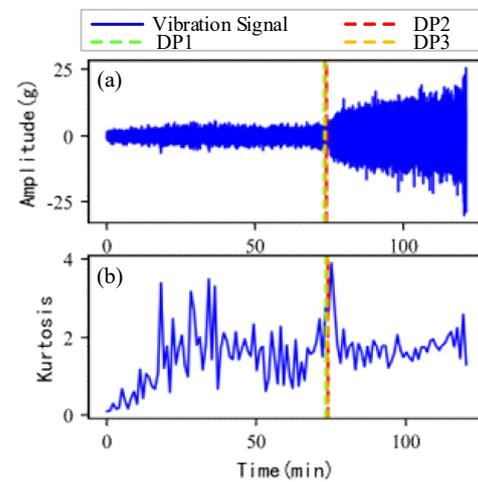
**Figure 12.** Degradation point identification of bearing 4 before outlier cleaning. (a) Vibration signal. (b) Kurtosis.



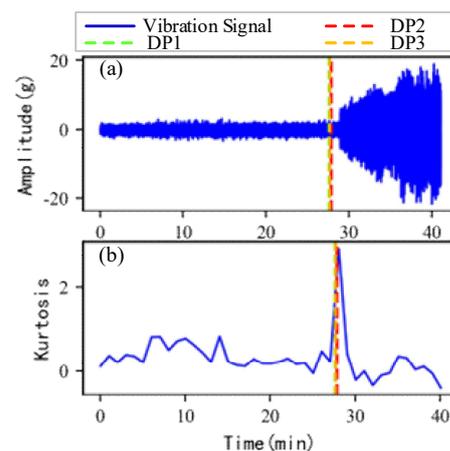
**Figure 13.** Degradation point identification of bearing 1 after outlier cleaning. (a) Vibration signal. (b) Kurtosis.



**Figure 14.** Degradation point identification of bearing 2 after outlier cleaning. (a) Vibration signal. (b) Kurtosis.



**Figure 15.** Degradation point identification of bearing 3 after outlier cleaning. (a) Vibration signal. (b) Kurtosis.



**Figure 16.** Degradation point identification of bearing 4 after outlier cleaning. (a) Vibration signal. (b) Kurtosis.

Figures 9–12 are the degradation point identification results of bearing 1, bearing 2, bearing 3 and bearing 4 based on raw vibration signals. Figures 13–16 are the degradation point recognition results of bearing 1, bearing 2, bearing 3 and bearing 4 after outlier cleaning.

After outlier cleaning, the stability of the health indicators and the robustness of the degraded point identification have been significantly improved.

#### (1) Stability comparison of health indicator

As can be seen in Figures 9b and 10b, in the normal stage without bearing failure, the degradation health indicator kurtosis is generally kept at a low level and the amplitude fluctuates in a small range. When the bearings begin to degrade, the amplitude of kurtosis will increase significantly. This shows that kurtosis is an effective health indicator of bearing degradation. Figures 11b and 12b show the health indicators of bearing 3 and bearing 4, from which it can be observed that due to the great influence of outliers in the normal stage, the variation range of health indicators becomes smaller when the bearing is really degraded. However, the random impulse-type outliers will cause great interference to the amplitude of kurtosis. Therefore, the impulse-type outliers will reduce the stability of the health index. After outlier cleaning, as shown in Figures 13b, 14b, 15b and 16b, the stability of the kurtosis indicator has been significantly improved. At this time, the amplitude of the health index will not be disturbed by the outliers, which provides important help for health monitoring and degradation stage identification.

#### (2) Robustness comparison of degraded point identification

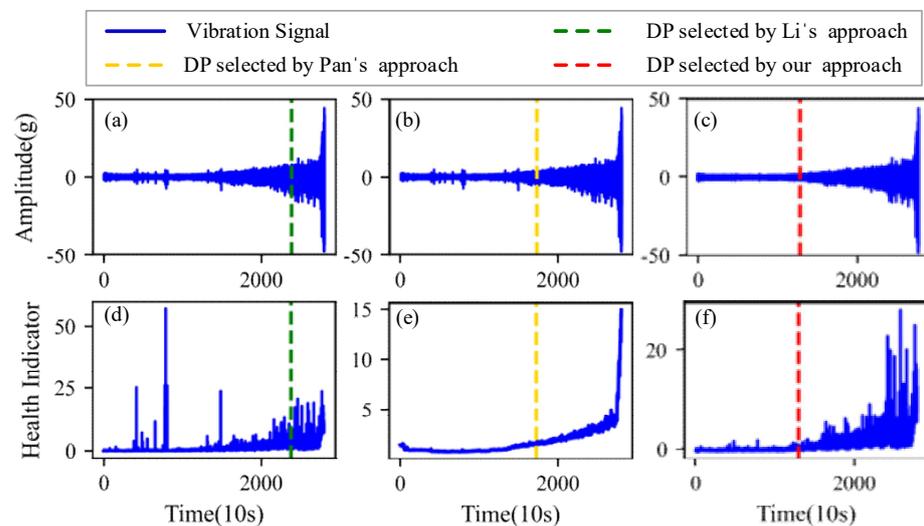
According to Formula (12), in the  $3\sigma$  method for degradation stage identification, the mean  $\mu$  and standard deviation  $\sigma$  of the reference signal are the key parameters to determining the degradation point. The method with good performance should be robust to the selection of the reference signal, that is, for different reference signals, as long as they are contained in the normal stage, the degradation point should be identified at the similar positions.

The robustness of the degradation point identification before and after outlier cleaning with three reference signals is compared in Figures 9–16. These three reference signals are the earliest 15%, 20%, 30% data points of the full lifecycle degradation data. The corresponding three degradation points are DP1, DP2 and DP3.

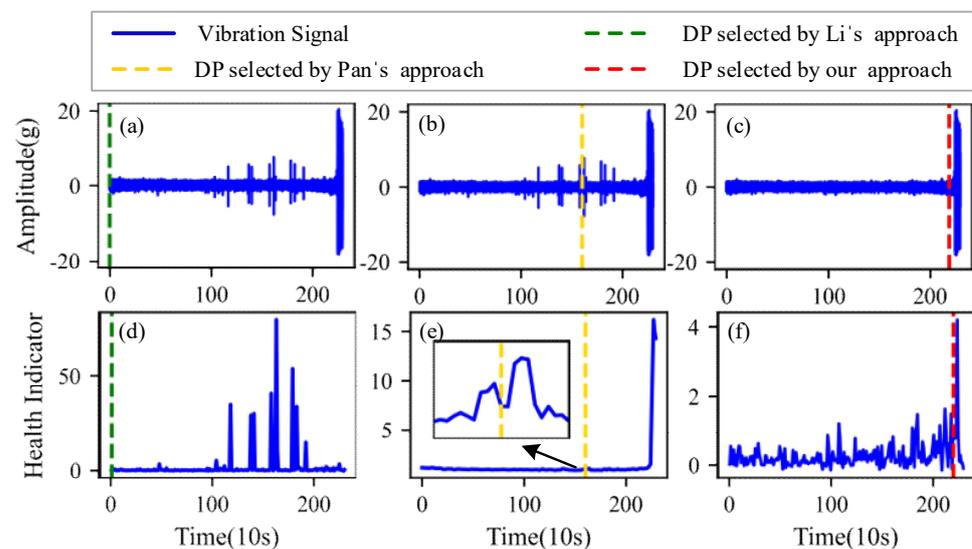
As shown in Figures 9–12, before outlier cleaning, when different reference signals are selected, the degradation points are identified in places with large time differences. That means the selection of the reference signal will cause significant interference to the identification of the degradation points, and the robustness of the  $3\sigma$  method without outlier cleaning is very poor.

Conversely, after outliers are cleaned, as shown in Figures 13–16, the identified three degradation points (DP1, DP2, DP3) corresponding to the earliest 15%, 20% and 30% data points have small time errors. That means the robustness of the  $3\sigma$  method after outlier cleaning increases significantly. Moreover, after the identified degradation point, the vibration amplitude increases significantly, which verifies the accuracy of the degradation point identification.

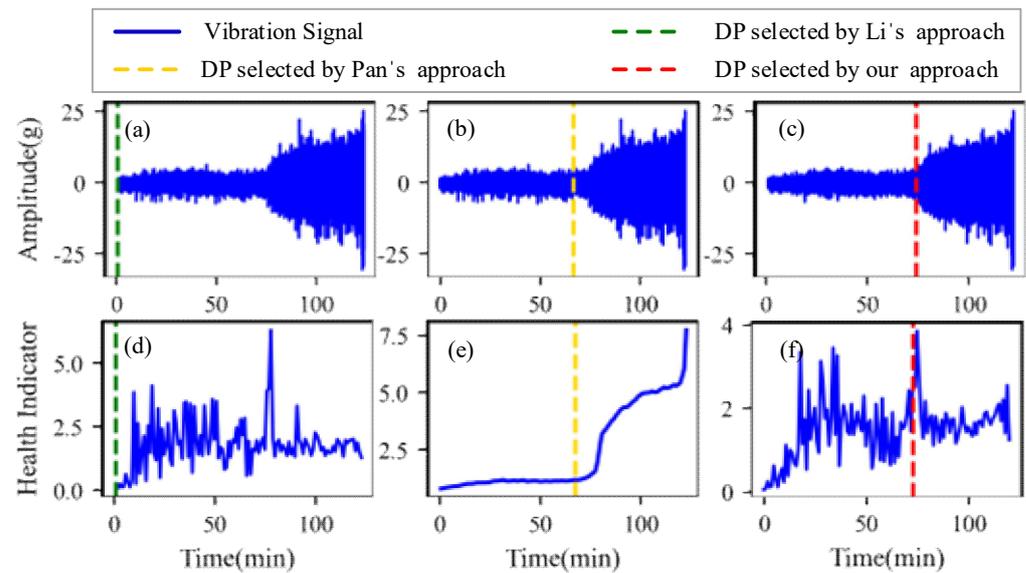
Furthermore, to identify the effectiveness of the proposed method, the degradation point identification results of the proposed method are compared with those of the methods proposed by Li et al. [7] and Pan et al. [10]. The health indicators in Li's approach and our approach are kurtosis characteristics, while the health indicator in Pan's approach is relative root mean square. These reference signals for the three approaches are the earliest 20% data points of the full lifecycle degradation data. The degradation point identification results of the four bearings are shown in Figures 17–20, respectively, and the degradation point identification values are given in Table 2.



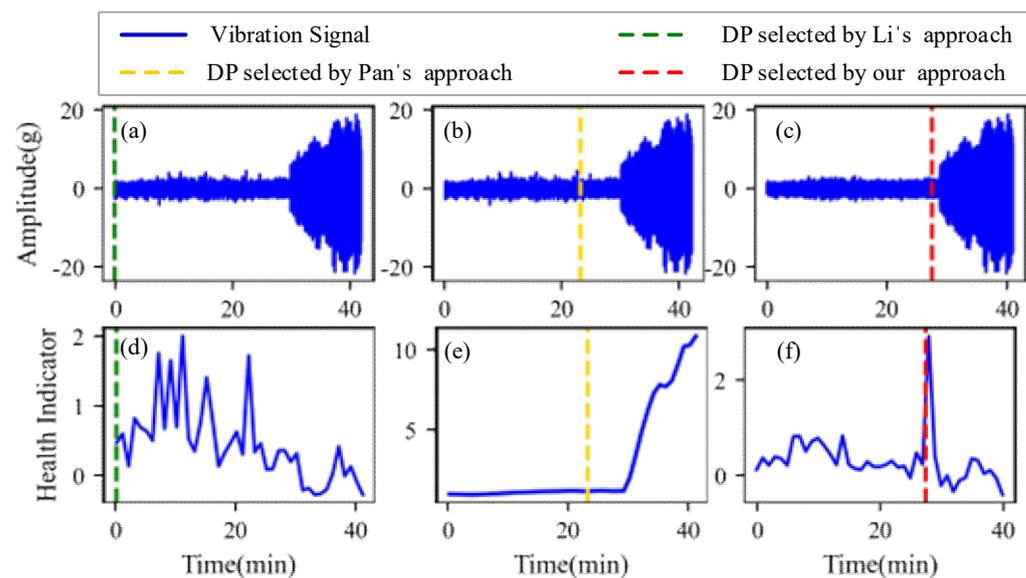
**Figure 17.** Degradation point identification results of bearing 1. (a) Li's approach. (b) Pan's approach. (c) Our approach. (d) Li's approach. (e) Pan's approach. (f) Our approach.



**Figure 18.** Degradation point identification results of bearing 2. (a) Li's approach. (b) Pan's approach. (c) Our approach. (d) Li's approach. (e) Pan's approach. (f) Our approach.



**Figure 19.** Degradation point identification results of bearing 3. (a) Li's approach. (b) Pan's approach. (c) Our approach. (d) Li's approach. (e) Pan's approach. (f) Our approach.



**Figure 20.** Degradation point identification results of bearing 4. (a) Li's approach. (b) Pan's approach. (c) Our approach. (d) Li's approach. (e) Pan's approach. (f) Our approach.

**Table 2.** Comparison results of degradation point identification by different methods of two bearings.

Case	Li's Approach	Pan's Approach	Our Approach
Bearing 1	23,810 s	17,230 s	12,910 s
Bearing 2	0 s	1600 s	2190 s
Bearing 3	0 min	67 min	73 min
Bearing 4	0 min	23 min	27 min

Figure 17 plots the corresponding degradation point identification results of bearing 1 by the three approaches. Figure 17a,d are the raw vibration signals and the health indicators of Li's approach, respectively. Figure 17b,e are the raw vibration signals and the health indicators of Pan's approach, respectively. Figure 17c,f are the vibration signals and the health indicators after outlier cleaning of our approach, respectively. As can be seen from

Figure 17, during the normal stage, there are some abnormal states caused by random noises instead of faults, which interfere with the identification of the degradation points. The degradation points identified by Li's approach and Pan's approach have a little time delay. Thanks to the anti-interference ability of our approach, the time when degradation occurs is appropriately selected as the degradation point.

Figures 18–20 plot the corresponding degradation point identification results of bearing 2, bearing 3 and bearing 4 by the three approaches; the meanings of the subgraphs of Figures 18–20 are shown in Figure 17. For bearing 2, bearing 3 and bearing 4, they behave in abrupt degradation processes. Therefore, it is easier to exactly give the degradation stage of bearing 2, bearing 3 and bearing 4, compared with bearing 1. However, due to the intense interference of outliers during the normal stage, the degradation characteristics of the bearings are obviously weaker than those of the outliers, and it is seen from Figures 18–20 that Li's approach fails to identify the degradation point. In the experiments using Pan, weak changes in the gradient of the health indicators caused misjudgments of the degradation points (the location of the degradation point was determined in advance). Similarly, Pan's approach also fails to distinguish the normal and the abnormal states. Our approach still performs better than Li's and Pan's approaches in identifying the degradation point of bearing 2, bearing 3 and bearing 4.

#### 4. Conclusions

In this paper, a degradation stage recognition method of bearings is proposed based on outlier cleaning. The outlier detection method, combining global abnormal segment detection and accurate locating of abnormal impulses, is constructed, realizing the accurate and quick removing of impulse-types outliers that have significant interference with the identification of the degradation points. The main conclusions of this paper are as follows:

The method for locating the start point and the end point of the impulse are constructed to realize the precise identification of the abnormal impulse. While removing abnormal impulses, the normal data are fully retained.

The screening criteria and the iterative removal strategy for abnormal segments are proposed to realize the fast recognition of segments that contain abnormal impulses.

The outlier cleaning avoids the interference of the selection of reference signals when using the  $3\sigma$  method to identify the degradation stage. The stability of the health indicators and the robustness of the degraded point identification have been significantly improved.

**Author Contributions:** Conceptualization, J.X. and Y.X. (Yujie Xie); methodology, Y.X. (Yujie Xie); validation, J.X. and Y.X. (Yujie Xie); formal analysis, Y.X. (Yujie Xie); investigation, T.W.; resources, T.W. and Y.X. (Yougang Xiao); writing—original draft preparation, J.X. and Y.X. (Yujie Xie); visualization, J.X. and Y.X. (Yujie Xie); supervision, T.W. and Y.X. (Yougang Xiao); project administration, Y.X. (Yougang Xiao); funding acquisition, J.X., T.W. and Y.X. (Yougang Xiao). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number 2020YFA0710904-02 and the Natural Science Foundation of Hunan Province China, grant number 2021JJ30847 and 2021JJ40765.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Wu, T.; Ma, X.; Yang, L.; Zhao, Y. Proactive maintenance scheduling in consideration of imperfect repairs and production wait time. *J. Manuf. Syst.* **2019**, *53*, 183–194. [[CrossRef](#)]
2. Zhou, F.; Yang, S.; Fujita, H.; Chen, D.; Wen, C. Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. *Knowl.-Based Syst.* **2020**, *187*, 104837. [[CrossRef](#)]

3. Wang, B.; Lei, Y.; Yan, T.; Li, N.; Guo, L. Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery. *Neurocomputing* **2020**, *379*, 117–129. [\[CrossRef\]](#)
4. Jin, X.; Zhao, M.; Chow, T.W.; Pecht, M. Motor Bearing Fault Diagnosis Using Trace Ratio Linear Discriminant Analysis. *IEEE Trans. Ind. Electron.* **2013**, *61*, 2441–2451. [\[CrossRef\]](#)
5. Che, Y.; Chen, Y.; Cui, Y. Rolling element bearing remaining useful life estimation based on KPCA and improved long-short-term memory network. *J. Electron. Meas. Instrum.* **2021**, *35*, 109–114.
6. Miao, M.; Yu, J. A Deep Domain Adaptative Network for Remaining Useful Life Prediction of Machines Under Different Working Conditions and Fault Modes. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–14. [\[CrossRef\]](#)
7. Li, N.; Lei, Y.; Lin, J.; Ding, S.X. An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Ind. Electron.* **2015**, *62*, 7762–7773. [\[CrossRef\]](#)
8. Kong, X.; Yang, J. Remaining useful life prediction of rolling bearings based on RMS-MAVE and dynamic exponential regression model. *IEEE Access* **2019**, *7*, 169705–169714. [\[CrossRef\]](#)
9. Singh, J.; Darpe, A.K.; Singh, S.P. Bearing remaining useful life estimation using an adaptive data-driven model based on health state change point identification and K-means clustering. *Meas. Sci. Technol.* **2020**, *31*, 085601. [\[CrossRef\]](#)
10. Pan, Z.; Meng, Z.; Chen, Z.; Gao, W.; Shi, Y. A two-stage method based on extreme learning machine for predicting the remaining useful life of rolling-element bearings. *Mech. Syst. Signal Process.* **2020**, *144*, 106899.
11. Gao, H.; Cui, L.; Dong, Q. Reliability modeling for a two-phase degradation system with a change point based on a Wiener process. *Reliab. Eng. Syst. Saf.* **2020**, *193*, 106601. [\[CrossRef\]](#)
12. Li, X.; Zhang, W.; Ma, H.; Luo, Z.; Li, X. Data alignments in machinery remaining useful life prediction using deep adversarial neural networks. *Knowl.-Based Syst.* **2020**, *197*, 105843. [\[CrossRef\]](#)
13. Baptista, M.L.; Henriques, E.M.P.; Goebel, K. More effective prognostics with elbow point detection and deep learning. *Mech. Syst. Signal Process.* **2020**, *146*, 106987. [\[CrossRef\]](#)
14. Wu, J.-Y.; Wu, M.; Chen, Z.; Li, X.; Yan, R. A joint classification-regression method for multi-stage remaining useful life prediction. *J. Manuf. Syst.* **2021**, *58*, 109–119. [\[CrossRef\]](#)
15. Kong, X.; Yang, J.; Li, L. Remaining useful life prediction for degrading systems with random shocks considering measurement uncertainty. *J. Manuf. Syst.* **2021**, *61*, 782–798. [\[CrossRef\]](#)
16. Zhang, S.; Liu, H.; Hu, M.; Jiang, A.; Zhang, L.; Xu, F.; Hao, G. An Adaptive CEEMDAN Thresholding Denoising Method Optimized by Nonlocal Means Algorithm. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 6891–6903.
17. Chen, W.; Li, J.; Wang, Q.; Han, K. Fault feature extraction and diagnosis of rolling bearings based on wavelet thresholding denoising with CEEMDAN energy entropy and PSO-LSSVM. *Measurement* **2020**, *172*, 108901. [\[CrossRef\]](#)
18. Yang, D.; Karimi, H.R.; Sun, K. Residual wide-kernel deep convolutional auto-encoder for intelligent rotating machinery fault diagnosis with limited samples. *Neural Netw.* **2021**, *141*, 133–144. [\[CrossRef\]](#)
19. Li, X.; Jiang, H.; Xiong, X.; Shao, H. Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network. *Mech. Mach. Theory* **2019**, *133*, 229–249. [\[CrossRef\]](#)
20. Chen, J.; Chang, Y.; Qu, C.; Zhang, M.; Li, F.; Pan, J. Intelligent Impulse Finder: A boosting multi-kernel learning network using raw data for mechanical fault identification in big data era. *ISA Trans.* **2020**, *107*, 402–414. [\[CrossRef\]](#)
21. Gupta, M.; Gao, J.; Aggarwal, C.C.; Han, J. Outlier Detection for Temporal Data: A Survey. *IEEE Trans. Knowl. Data Eng.* **2014**, *9*, 2250–2267. [\[CrossRef\]](#)
22. Zhang, Y.; Wang, A. Remaining Useful Life Prediction of Rolling Bearings Using Electrostatic Monitoring Based on Two-Stage Information Fusion Stochastic Filtering. *Math. Probl. Eng.* **2020**, *2020*, 1–12. [\[CrossRef\]](#)
23. Gao, J.; Hu, W.; Li, W.; Zhang, Z.; Wu, O. Local Outlier Detection Based on Kernel Regression. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; Volume 148, pp. 585–588.
24. Wang, X.; Li, H.; Fan, X.; Liu, S. Probabilistic interval prediction of wind power based on VMD-GRU and nonparametric kernel density estimation. *J. Beijing Inf. Sci. Technol. Univ.* **2021**, *36*, 59–65.
25. Gaşior, K.; Urbańska, H.; Grzesiek, A.; Zimroz, R.; Wylomańska, A. Identification, decomposition and segmentation of impulsive vibration signals with deterministic components—A sieving screen case study. *Sensors* **2020**, *20*, 5648. [\[CrossRef\]](#)
26. Shakya, P.; Kulkarni, M.S.; Darpe, A.K. A novel methodology for online detection of bearing health status for naturally progressing defect. *J. Sound Vib.* **2014**, *333*, 5614–5629. [\[CrossRef\]](#)
27. Hayter. *Probability and Statistics for Engineers and Scientists*; Cengage Learning: Boston, MA, USA, 2012.
28. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, Beijing, China, 23–25 May 2012; pp. 1–8.
29. Wang, B.; Lei, Y.; Li, N.; Li, N. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [\[CrossRef\]](#)
30. Hu, W.; Gao, J.; Li, B.; Wu, O.; Du, J.; Maybank, S.J. Anomaly Detection Using Local Kernel Density Estimation and Context-Based Regression. *IEEE Trans. Knowl. Data Eng.* **2020**, *32*, 218–233. [\[CrossRef\]](#)