

Article

Unsupervised Anomaly Detection of Intermittent Demand for Spare Parts Based on Dual-Tailed Probability

Kairong Hong ¹, Yingying Ren ¹, Fengyuan Li ¹, Wentao Mao ^{2,*}  and Yangshuo Liu ²¹ China Railway Tunnel Group, Zhengzhou 450001, China² School of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China; 2008183003@stu.htu.edu.cn

* Correspondence: maowt@htu.edu.cn; Tel.: +86-15037301821

Abstract: The quick development of machine learning techniques provides a superior capability for manufacturing enterprises to make effective decisions about inventory management based on spare parts demand (SPD) data. Since SPD sequences in practical maintenance applications usually show an intermittent distribution, it is not easy to represent the demand pattern of such sequences. Meanwhile, there are some aspects like manual report errors, environmental interference, sudden project changes, etc., that bring large and unexpected fluctuations to SPD sequences, i.e., anomalous demands. The inventory decision made based on the SPD sequences with anomalous demands is not trusted by enterprise engineers. For such SPD data, there are two great concerns, i.e., false alarms in which sparse demands are recognized to be anomalous and missing alarms in which the anomalous demands are categorized as normal due to their adjacent demands having extreme values. To address these concerns, a new unsupervised anomaly-detection method for intermittent time series is proposed based on a dual-tailed probability. First, the multi-way delay embedding transform (MDT) was applied on the raw SPD sequences to obtain higher-order tensors. Through Tucker tensor decomposition, the disturbance of extreme demands can be effectively reduced. For the reconstructed SPD sequences, then, the tail probability at each time point, as well as the empirical cumulative distribution function were calculated based on the probability of the demand occurrence. Second, to lessen the disturbance of sparse demand, the non-zero demand sequence was distilled from the raw SPD sequence, with the tail probability at each time point being calculated. Finally, the obtained dual-tailed probabilities were fused to determine the anomalous degree of each demand. The proposed method was validated on the two actual SPD datasets, which were collected from a large engineering manufacturing enterprise and a large vehicle manufacturing enterprise in China, respectively. The results demonstrated that the proposed method can effectively lower the false alarm rate and missing alarm rate with no supervised information provided. The detection results were trustworthy enough and, more importantly, computationally inexpensive, showing significant applicability to large-scale after-sales parts management.



Citation: Hong, K.; Ren, Y.; Li, F.; Mao, W.; Liu, Y. Unsupervised Anomaly Detection of Intermittent Demand for Spare Parts Based on Dual-Tailed Probability. *Electronics* **2024**, *13*, 195. <https://doi.org/10.3390/electronics13010195>

Academic Editors: Hao Xue and Du Huynh

Received: 26 October 2023

Revised: 8 December 2023

Accepted: 14 December 2023

Published: 2 January 2024

Keywords: anomaly detection; intermittent time series; unsupervised learning; spare parts demand; tensor decomposition



Copyright: © 2024 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

In recent years, the after-market service of large manufacturing enterprises has become a key focus for business transformation and value increase [1]. Leveraging operational data acquired throughout the product lifecycle to optimize and upgrade core processes in the after-market service can effectively reduce enterprise's time and labor costs. Intelligent maintenance can then promote the transformation of manufacturing's after-market services in the direction of intelligence and automation. With the rapid development of machine learning techniques, enterprises are starting to utilize historical spare parts demand (SPD) data to forecast the quantity of parts needed in the future period, enabling intelligent spare

parts planning [2]. Accurate SPD prediction often relies on the quality of historical demand data. The demand for spare parts can then be accumulated as a type of sensing data for the after-market maintenance of an enterprise. Having a somewhat different form of the well-known sensing data, like vibration signals and monitoring images, the SPD data are capable of reflecting the operational status of manufacturing enterprises. The analysis and utilization of SPD data for manufacturing enterprises have tremendous research scope.

The demand for after-sales parts in manufacturing enterprises typically arises from the replacement of faulty parts or the initiation of new projects. However, there are also “unusual demands” that occur outside of typical business situations hidden within the sequences of demand [3]. There are several factors that can result in anomalous demand for spare parts, such as manual report errors, which lead to the accumulation of historical repair orders, or environmental interference like seasonal high temperatures, pollen, sandstorms, etc. The quantity of such anomalous demands often exceeds normal demands several times over. Especially when multiple factors are combined, this can lead to an anomalous demand with an extreme value (also named extreme demand in this paper), e.g., ten to even dozens of times higher than the usual demand. Anomalous demand will lead to poor model robustness and bring great deviation to spare parts planning. Therefore, anomaly detection in SPD data is of great significance in enhancing the tolerance of predictive models and realizing intelligent spare parts planning [4].

For large manufacturing enterprises, the demand for spare parts occurs randomly, while the interval of demand occurrence is not fixed. Consequently, the SPD data show an intermittent distribution, and SPD sequences over time can be regarded as intermittent time series. The demand for spare parts is influenced by various factors, such as the lifespan of the equipment, the failure rate, seasonal factors, and changes in production output. These factors make the prediction of spare parts demand more complex and difficult. Therefore, to understand and predict these requirements, time series analysis is introduced to help us forecast evolutionary patterns and trends in SPD data. By analyzing data that changes over time, we can better understand the seasonality, trends, and other characteristics of spare parts demand. The uniqueness of this process lies in its ability to analyze the impact of various factors on demand. For example, we can analyze historical data on equipment failure rates to analyze their impact on spare parts demand. In addition, we can also identify and predict seasonal demand patterns. However, although time series analysis has many advantages in predicting spare parts demand, it also has its limitations. First, the demand for historical data is the foundation of this kind of study, but for some emerging enterprises, it may become a challenge with the limited data volume. In this case, it might be necessary to use interpolation techniques or expert opinions to fill in the data gap. Second, time series analysis assumes that future demand patterns will be similar to those of the past, which is too strict for all situations. Especially in industries with rapidly changing technology or market conditions, this assumption may not hold true. As a kind of time series analysis, time series anomaly detection aims to identify a small portion of the data points with outliers, fluctuations, or other exceptional conditions. Since manual labeling is costly, most time series data lack annotated information about anomalies, leading to unsupervised time series anomaly detection. Unsupervised anomaly-detection methods can typically be categorized into three classes: partition-based methods, prediction-based methods, and reconstruction-based methods. Partition-based anomaly-detection methods include shallow anomaly detection models such as one-class support vector machines (OCSVMs) [5], local outlier factors (LOFs) [6], K-nearest neighbors (KNN) [7], isolation forest (IForest) [8], and others. These models measure the outlier degree of data points by means of density, statistics, distance, etc. These models have a fast calculation speed and are suitable for small-sample data. One representative method is deep support vector data description (Deep SVDD) [9]. It maps data representations to a minimal hypersphere inside which the mapping of normal values falls and outside which the mapping of anomalies is located. Prediction-based anomaly-detection methods assign anomaly scores by measuring the distance between the predicted values of a fore-

casting model and the actual values [10]. Classical anomaly-detection methods employ the variants of autoregressive moving average models for time series [11]. On the basis of classical autoregressive methods, Bontemps et al. [12] modeled the correlation between different sequences by employing recurrent neural networks (RNNs) with long short-term memory (LSTM) units, referred to as LSTM-RNN. Hundman et al. introduced LSTMs and nonparametric dynamic thresholding [13], which utilizes LSTM for time series prediction and employs an adaptive error threshold to determine outliers without any threshold set in advance. Specifically, LSTM-NDT treats the error values as a non-parametric distribution and dynamically estimates the probability density function of errors using kernel density estimation. Within this probability density function, this method identified anomaly points by finding specific confidence intervals. Garg et al. [14] used the transposed convolution to replace the convolutional filter in temporal convolutional networks and proposed an autoencoder (AE) network to realize time series reconstruction. Deldari et al. [15] introduced a contrastive prediction encoding approach that utilizes temporal convolutional networks for feature extraction and detects anomalies in time series data. Reconstruction-based anomaly-detection methods are primarily based on AEs [16], such as the variational AE (VAE) proposed by Kingma et al. [17] and the recursive AE (RAE) introduced by Cho et al. [18]. Park et al. combined LSTM with the VAE to build a reconstruction-based anomaly-detection method, known as the LSTM-VAE [19], by analyzing reconstruction errors. Niu et al. [20] proposed a time series anomaly-detection method based on a hybrid model called LSTM-VAE-GAN. This approach utilizes LSTM networks for training and detects anomalies based on reconstruction differences and discriminative results. Dan et al. [21,22] proposed two anomaly-detection methods based on generative adversarial networks (GANs), called GAN-AD and MAD-GAN. These methods utilize an RNN to capture the distribution of the time series data and detect potential anomalies by computing the error between the reconstruction data from the GAN model's discriminator and the real values. Fan et al. [23] constructed a new abnormal fluctuation similarity matrix and introduced it into the support vector machine model for hypersphere training. Schmidl et al. [24] conducted a wide-ranging literature survey by evaluating state-of-the-art anomaly-detection algorithms based on their commonalities and performance metrics, such as effectiveness, efficiency, robustness, etc. Li et al. [25] categorized time series anomalies into three types, namely abnormal time points, time intervals, and time series, and used LSTM and an autoencoder to detect abnormal time points and abnormal time intervals. Kim et al. [26] fused the Transformer model to predict anomalies by inputting global trends and the local matching of time series. Ren et al. [27] borrowed spectral residuals (SRs) and convolutional neural networks (CNNs) from the field of visual saliency detection for time series anomaly detection for the first time. Such deep-learning-based anomaly-detection methods generally require a large amount of training data and will not perform well in small-scale environments. According to our literature survey, there are a few works on few-shot anomaly detection. For instance, Bashar et al. [28] introduced the TAnoGAN model for detecting anomalies on small-sample data. But, it still suffers from uncertainty in the data. In summary, although the above-mentioned methods have achieved promising results on some datasets, when applied the SPD data in manufacturing enterprises, these methods are prone to identifying normal intermittent data as anomalous demand, leading to false alarms, or recognize anomalous demand as normal, resulting in missing alarms.

Based on the analysis mentioned above, the key to improving intermittent time series anomaly detection lies in: (1) how to identify and reasonably correct extreme demand to avoid false alarms; (2) how to handle the intermittent distribution of SPD data to avoid missing alarms. Following this idea, this paper proposes an unsupervised intermittent time series anomaly-detection method based on dual-tailed probabilities. This method designs a dual methodology: one for addressing missing alarms and the other one for false alarms. Specifically, this method reconstructs the original SPD sequence into a non-zero demand sequence and a tensor reconstruction sequence. Then, the tail probability for each sequence is calculated respectively. Here, we comprehensively considered two factors: the quantity

of anomalous demand and its occurrence time. This way, we aimed to uncover potential (covered by extreme demands) demand anomalies in the sequence by introducing tensor decomposition to handle extreme demands in the sequence. The algorithmic procedure was as follows: First, we employed the multi-way delay embedding transform (MDT) to extract temporal information along the time dimension. The MDT is able to convert the one-dimensional SPD sequence into a higher-order tensor. We utilized Tucker tensor decomposition to project this tensor onto a compressed core tensor and reconstructed it as a new time series. This process can effectively eliminate extreme demand from the original SPD sequence, which achieves a reasonable correction of the demand distribution and provides high-quality data support for the subsequent algorithmic steps. Second, we utilized dual-tailed probabilities for anomaly detection in the sequence. We separately calculated the empirical cumulative distribution functions for both the non-zero demand quantity sequence and the reconstruction sequence. Utilizing these functions, we determined the tail probabilities on each time point (i.e., demand occurrence) and obtained two sets of anomaly-detection results from each sequence. By combining them with appropriate weighting, we obtained the final anomaly detection result. This methodology takes into account both the abnormality of demand occurrence and the demand quantity at each time point. This approach can effectively avoid excessively strict detection outcomes caused by the sparse distribution of the original SPD sequence while maintaining high computational efficiency.

The theoretical contribution of this paper lies in proposing a new unsupervised anomaly-detection method for intermittent time series data. In comparison to traditional methods, this method can accurately identify and effectively correct the extreme demands in the original SPD sequence through Tucker tensor decomposition. It successfully addresses the interference of extreme demands in the original sequence, thereby solving the problem of missing alarm. Furthermore, this method considers anomalies from two perspectives: the time of anomalous demand occurrence and the quantity of anomalies. The false alarms can then be mitigated by calculating the tail probabilities within the intermittent time series. To the best of our knowledge, the research on anomaly detection for intermittent time series is still in its infancy.

2. Background

2.1. Demand Patterns

In general, demand patterns can be categorized based on the average demand interval (*ADI*) and the squared coefficient of variation (CV^2) [29]. Due to the intermittent distribution of SPD data, the demand for different spare parts fluctuates greatly. It will be helpful to construct an effective detection model by classifying spare parts demands according to their intermittent characteristics. Through the categorization with the *ADI* and CV^2 , one can better understand and manage spare parts requirements, thereby improving the maintenance efficiency of equipment and the accuracy of anomaly detection, for instance choosing appropriate parameters for the detection model for different types of SPD distributions. For an intermittent time series $X = \{x_1, x_2, \dots, x_n\}$ with n periods, the *ADI* and CV^2 are calculated by:

$$ADI = \frac{n}{d} \quad (1)$$

$$CV^2 = \left(\frac{S_d}{\bar{x}_d} \right)^2 \quad (2)$$

where d and S_d represent the number of periods and the standard deviation of the non-zero demand sequence in X , respectively, and \bar{x}_d is the average of the non-zero demand sequence. According to *ADI* and CV^2 , the demand sequence can be divided into four categories: stable demand, unstable demand, intermittent demand, and blocky demand. The specific classification criteria are as follows:

- (1) Stable demand ($ADI < 1.32, CV^2 < 0.49$): this category of demand is relatively stable with few zero demand periods.
- (2) Unstable demand ($ADI < 1.32, CV^2 \geq 0.49$): the demand is unstable with high variability and occurs frequently.
- (3) Intermittent demand ($ADI \geq 1.32, CV^2 < 0.49$): the demand is irregular and scattered, but relatively stable.
- (4) Blocky demand ($ADI \geq 1.32, CV^2 \geq 0.49$): this category has a random demand pattern, with a large number of time periods having no demand and the demands varying greatly from period to period, accompanied by a significant number of zero demand stages.

2.2. Multi-Way Delay Embedding Transform

Multi-way delay embedding transform (MDT) technology can embed low-rank data into high-dimensional space and can be used to construct Hankel matrices or block Hankel tensors [30]. The tensors obtained from the MDT have low-rank characteristics, which smooth the original data and facilitate training. Assuming that the Hankel matrix of vector $v = (v_1, \dots, v_L)^T \in R^L$ with a delay of τ is shown in Equation (3), this process is called the Hankel transformation of the vector:

$$\mathcal{H}_\tau(v) := \begin{pmatrix} v_1 & v_2 & \cdots & v_{L-\tau+1} \\ v_2 & v_3 & \cdots & v_{L-\tau+2} \\ \vdots & \vdots & \ddots & \vdots \\ v_\tau & v_{\tau+1} & \cdots & v_L \end{pmatrix} \in R^{\tau \times (L-\tau+1)} \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (3)$$

First, the duplication matrix $S \in \{0, 1\}^{\tau \times (L-\tau+1) \times L}$ is constructed with a delay of τ as:

$$S^T = \begin{pmatrix} I_\tau & & & \\ & I_\tau & & \\ & & \ddots & \\ & & & I_\tau \end{pmatrix}_{\tau \times \tau} T \quad (4)$$

Second, the vector v is transformed into a Hankel matrix, denoted as $\mathcal{H}_\tau(v)$. The replication matrix S is essentially a linear transformation. The specific way of expanding to quantization is:

$$vec(\mathcal{H}_\tau(v)) = Sv, \quad Sv \in R^{\tau \times (L-\tau+1)} \quad (5)$$

where $vec()$ is a matrix that extends along the column direction, and the Hankel matrix obtained through delayed embedding can be represented as:

$$\begin{aligned} \mathcal{H}_\tau(v) &= fold_{(L,\tau)}(Sv) := v_H, \\ fold_{(L,\tau)} &: R^{\tau \times (L-\tau+1)} \rightarrow R^{\tau \times (L-\tau+1)} \end{aligned} \quad (6)$$

where $fold_{(L,\tau)}$ is the process of folding a vector into a matrix.

The inverse transformation of the multi-way delay embedding of vectors can convert data from a high-dimensional space to a low-dimensional target space, and its calculation method is as follows:

$$\mathcal{H}_\tau^{-1}(V_H) = S^\dagger vec(V_H) \quad (7)$$

$$S^\dagger := (S^T S)^{-1} S^T \quad (8)$$

where \dagger is the Moore–Penrose inverse matrix.

2.3. Tucker Tensor Decomposition

The process of decomposing high-order tensor data into a set of low-rank matrices or vectors is called tensor decomposition, which is often applied to tasks such as data

compression, dimensionality reduction, and feature extraction. Tucker tensor decomposition decomposes an Nth-order tensor $\chi \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ into the product of the core tensor $\varsigma \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ and N factor matrices $U^{(n)} \in \mathbb{R}^{I_n \times J_n}$, as shown in Equation (9). The factor matrix obtained by Tucker tensor decomposition represents the principal components of tensor modal expansion, while the kernel tensor captures the correlation between these components [31].

$$\chi = \varsigma \times_1 U^{(1)} \times_2 U^{(2)} \dots \times_N U^{(N)} \tag{9}$$

where $\varsigma \times_1 U^{(n)}$ is the n-mode product of the modular (n) expansion of tensor S and matrix $U^{(n)} \in \mathbb{R}^{I_n \times J_n}$:

$$\begin{aligned} [\varsigma \times U^{(n)}]_{j_1 \dots j_{n-1} i_n j_{n+1} \dots j_N} &= \sum_{j_n=1}^{J_n} \varsigma_{j_1 \dots j_{n-1} i_n j_{n+1} \dots j_N} u_{i_n j_n} \\ \varsigma \times U^{(n)} &\in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N} \end{aligned} \tag{10}$$

According to the above equation, any specific point in the tensor can be extended to:

$$x_{i_1 i_2 \dots i_N} = \sum_{j_1, j_2, \dots, j_N} \varsigma_{j_1 \dots j_N} u_{i_1 j_1}^{(1)} u_{i_2 j_2}^{(2)} \dots u_{i_N j_N}^{(N)} \tag{11}$$

For the ease of understanding, Figure 1 shows the use of Tucker tensor decomposition to decompose a third-order tensor, resulting in a smaller kernel tensor and the product of three factor matrices.

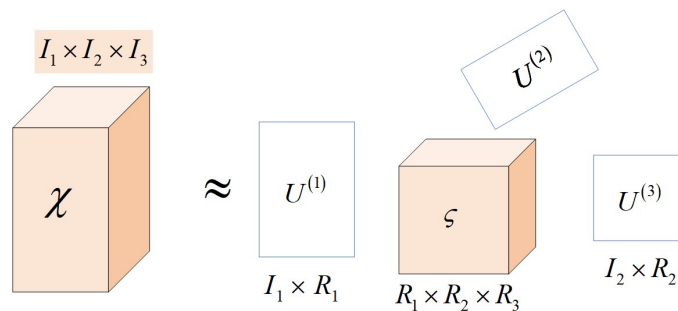


Figure 1. Illustration of Tucker tensor decomposition.

3. The Proposed Method

To address the problem of false alarms and missing alarms caused by the intermittent distribution characteristic, this section proposes the unsupervised anomaly detection of intermittent demand for spare parts based on a dual-tailed probability, which comprehensively judges the abnormality level of demand from two perspectives: demand occurrence time and demand quantity.

First, the original SPD sequences are transformed into high-order tensors using the MDT technique. The obtained tensors are further reconstructed by Tucker tensor decomposition to reduce the interference of extreme demand. The empirical cumulative distribution function is calculated according to the frequency of the demand occurrence in the reconstructed sequence at each time point, and then, the tail probability of each demand can be obtained. Second, to avoid sparse demand interference, the non-zero demand sequence is extracted from the original SPD sequence to calculate the tail probability of each time point. Finally, two probability values are used to judge the anomaly degree of the sample points. The specific workflow of the proposed method is shown in Figure 2.

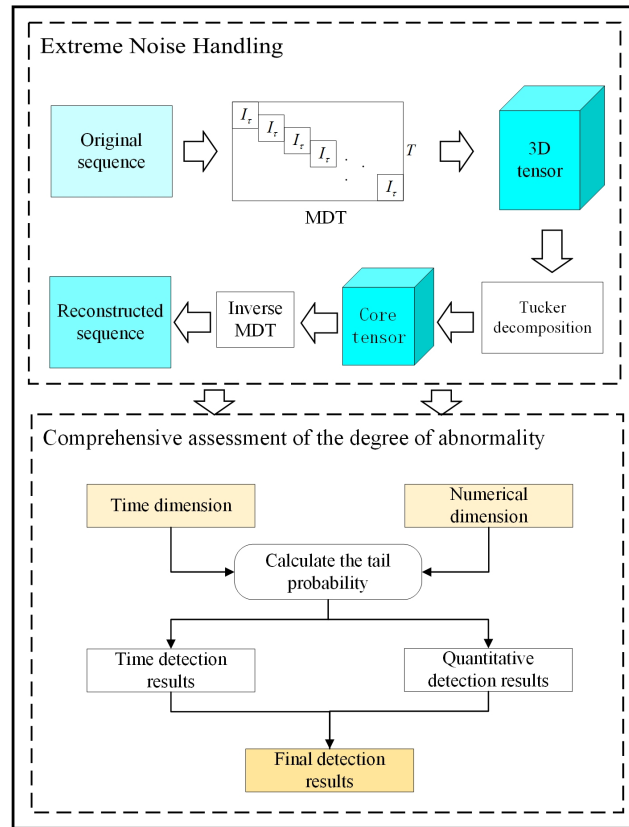


Figure 2. Flowchart of the proposed method.

3.1. Data Pre-Processing

Since the intermittent distribution of the SPD sequence has a huge impact on the results of anomaly detection, we chose to split the intermittent time series in order to obtain a non-zero demand quantity sequence. The obtained non-zero sequence is able to better express the information of demand quantity and expose a more-pronounced periodicity pattern compared to the original SPD sequence. An example of an intermittent time series before and after sequence segmentation is shown in Figure 3. Obviously, the temporal information is exposed more after splitting, which can well support the following anomaly detection.

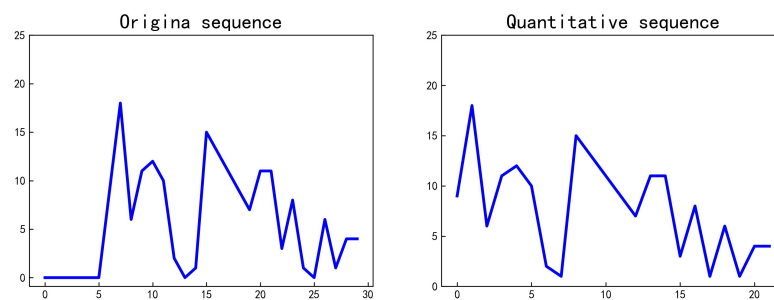


Figure 3. An example of the intermittent time series before and after sequence segmentation.

According to the intermittency assessment metrics listed in Section 2.1, the distribution of intermittency between the original sequence and the non-zero quantity sequence is shown in Table 1. According to the categorization criteria introduced in Section 2.1, the non-zero quantity sequence is more stable, while the original sequence is rather unstable. This case shows that sequence segmentation can change the characteristics of the original SPD sequence such as intermittency and instability and convert the original sequence into a more-stable one.

Table 1. Comparison of serial indicators before and after segmentation.

	Original Sequence	Sequence after Segmentation	Category of Sequence
ADI	0.733	1.000	Unstable sequence
CV ²	0.878	0.373	Stable sequence

3.2. Anomaly Detection Based on Dual-Tailed Probability

In this section, the original SPD sequence needs to be reconstructed by using the MDT technique and Tucker tensor decomposition. The one-dimensional demand sequence is first converted to high-dimensional data using the MDT technique in order to expand the temporal information. The high-dimensional data are further decomposed and reconstructed to extract the core tensor using Tucker tensor decomposition under orthogonal constraints. Finally, the core tensor is converted again to a one-dimensional sequence with the original size by using the inverse MDT technique. This process is shown in Figure 4.

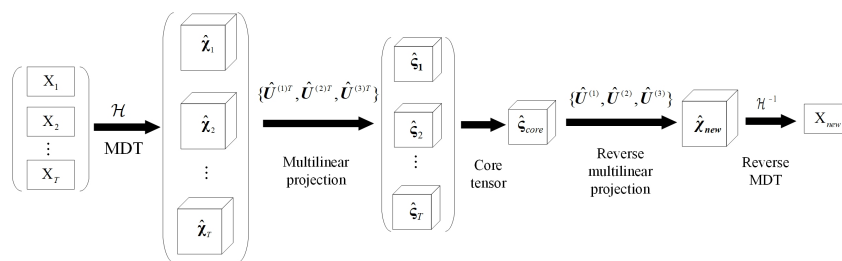


Figure 4. Flowchart of the reconstruction process based on the MDT and Tucker tensor decomposition.

Specifically, the aforementioned process begins by converting multiple time series into a higher-order block Hankel tensor using the MDT. The input data are the SPD sequence set $X \in I \times T$. The three-dimensional block Hankel tensor $\hat{\chi} \in I \times \tau \times (T - \tau + 1)$ is then obtained by performing the MDT in the time dimension. Here, the block Hankel tensor is assumed to be low rank and smooth in the embedding space. Note that the proposed method applies the MDT only along the time direction, which is due to the fact that the strength of the correlation between multiple demand sequences for spare parts is usually much weaker than the temporal correlation within the sequences. Next, the core tensor $\hat{\zeta}_t$ is obtained by Tucker tensor decomposition on $\hat{\chi}$:

$$\hat{\zeta}_t = \hat{\chi}_t \times_1 \hat{U}^{(1)T} \dots \times_M \hat{U}^{(M)T} \tag{12}$$

where the factor matrix \hat{U} maximally preserves the time continuity. $\hat{\zeta}_t$ can represent the essential information from the original Hankel tensor, avoiding the interference of non-essential information and effectively realizing noise reduction. Finally, $\hat{\zeta}_t$ is converted to a time series in the original space, i.e., reconstructed sequence, by means of the inverse MDT.

In order to avoid the interference by sparse demands, the proposed method jointly determines whether the demand for spare parts is abnormal or not from the two perspectives: demand occurrence time and demand quantity. With the reconstructed sequence, the quantity information of demand is now more obvious and regular. Meanwhile, the intermittent distribution characteristic of the quantity sequence is weakened, which can effectively reduce the influence of sparse demand on the detection results. The tail probability of the sequence is then calculated by using the empirical cumulative distribution function for both the quantity sequence and the reconstructed sequence. Finally, the abnormality degree of the SPD data can be determined by the tail probability.

Tail probability refers to the probability of a point being distributed in extreme positions, which is divided into the left-tailed probability and the right-tailed probability. Suppose x_i follows a distribution function F_X and $F_X(x_i) = P(X \leq x_i)$ is the left-tailed probability of x_i , while $1 - F_X(x_i) = P(X \geq x_i)$ is the right-tailed probability of x_i . If the tail probability is very small, this indicates that the probability of observing this value

is very small, that is this value should not occur frequently, and this point is considered as an outlier. Meanwhile, the weights are calculated using both the number of non-zero cycles and the total cycles, and then, the weights are added to the anomaly scores of the two sequences to obtain the final anomaly-detection results. The benefit of introducing the weights lies in the fact that they enable the algorithm to detect different types of demand sequences, which improves the applicability. The process of anomaly detection based on the tail probability is shown in Figure 5.

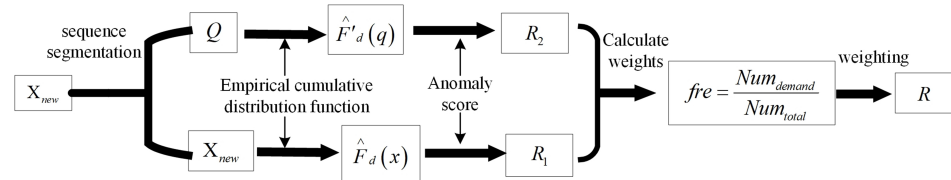


Figure 5. Diagram of anomaly detection for intermittent time series based on tail probability.

The detailed calculation is as follows. First, the reconstructed sequence $X = [x_1, x_2, \dots, x_n]$ is divided into a non-zero demand quantity sequence $Q = [q_1, q_2, \dots, q_m] (m \leq n)$. Second, the empirical left-tailed cumulative distribution function and empirical right-tailed cumulative distribution function are, respectively, calculated for both sequences, which are approximately regarded as tail probabilities. The left-tailed probability can be approximated by calculating the empirical cumulative distribution function [32]. The detailed calculation of the empirical cumulative distribution function for X is shown in Equation (13), where $\hat{F}(x)$ represents the probability of observing sample points. The right-tailed probability can then be calculated by Equation (14).

$$\hat{F}(x) = P((-\infty, x)) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \tag{13}$$

$$\hat{\bar{F}}(x) = \frac{1}{n} \sum_{i=1}^n I(-X_i \leq -x) \tag{14}$$

By taking the negative logarithm of the tail probabilities, the maximum value is regarded as the outlier score $O(X) = [X_1, X_2, \dots, X_n]$, as shown by Equations (15)–(17). Intuitively, the smaller the tail probability, the larger its negative logarithm. So, a point with a small left-tailed probability or a small right-tailed probability has a large probability to be an outlier.

$$p_l = -\log\left(\hat{F}(x)\right) \tag{15}$$

$$p_r = -\log\left(\hat{\bar{F}}(x)\right) \tag{16}$$

$$O(x_i) = \max\{p_l, p_r\} \tag{17}$$

Following the aforementioned process, the detected outliers for the reconstructed sequence $\hat{O}(X) = [\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n]$ and the outliers for the quantity sequence $\hat{O}(Q) = [\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_m]$ can be obtained, respectively. The anomaly detection results R_1 and R_2 are obtained based on the relative values of such outliers. Then, the final results R can be obtained by weighting the two results according to the frequency of the demand occurrence:

$$fre = \frac{Num_{demand}}{Num_{total}} \tag{18}$$

$$R = (1 - fre) \cdot R_2 + fre \cdot R_1 \tag{19}$$

where Num_{demand} is the number of non-zero demand occurrences and Num_{total} is the total number of cycles in the sequence. From Equation (18), a larger value of fre indicates a greater number of non-zero demand occurrences and, certainly, fewer zero values in the original SPD sequence.

Overall, the proposed anomaly-detection method can be summarized as shown in Algorithm 1.

Algorithm 1: Unsupervised anomaly detection of intermittent demand for spare parts based on dual-tailed probability.

Input: An SPD sequence $D = [d_1, d_2, \dots, d_n]$.

Output: Detected anomalous demands $R = [r_1, r_2, \dots, r_n]$.

Step1: Run Tucker tensor decomposition to obtain the reconstructed sequence $X = [x_1, x_2, \dots, x_n]$ by Equation (12).

Step2: Split X to obtain a sequence of non-zero demand quantity $Q = [q_1, q_2, \dots, q_m] (m \leq n)$, and perform the following steps for X and Q :

(1) Calculate the empirical left-tailed cumulative distribution function and the empirical right-tailed cumulative distribution function by Equations (13) and (14), approximated as the tail probability.

(2) Calculate by Equation (17) the outlier score $\hat{O}(X) = [\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n]$ for each data point in X and $\hat{O}(Q) = [\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_m]$ for each data point in Q .

Step3: Obtain the anomaly detection results R_1 and \hat{R}_2 on X and Q based on the relative values of $\hat{O}(X)$ and $\hat{O}(Q)$. Calculate the final result R by Equation (19).

4. Experimental Results

4.1. Dataset Introduction

In this section, the proposed method was validated using two real-life spare parts datasets. One was collected from a large engineering manufacturing enterprise in China, referred to as Dataset 1. For a fair evaluation, we also introduced an open inventory dataset provided by the Zoomlion Heavy Industry Science & Technology Co., Ltd. in China (Changsha, China) (<https://www.industrial-bigdata.com> (accessed on 7 December 2023)), referred to as Dataset 2. Dataset 1 encompasses historical demand data for 1687 categories of spare parts, including 407 sequences in which the anomalous demands were annotated by the enterprise engineers. The demands came from central and site warehouses, spanning from November 2018 to September 2021, totaling 34 months. Dataset 2 comprises the actual demand data for 366 categories of spare parts over a period of 30 months, including 165 sequences in which the anomalous demands were annotated. The implementation of the proposed method and the data used are available at <https://github.com/MMAIGX/gxl/tree/master> (accessed on 7 December 2023).

Figures 6 and 7 illustrate the monthly demand quantity distribution for the two datasets. In Figure 6, ADI and CV^2 are presented to show the demand type. The red horizontal and vertical lines represent the intermittent criteria $ADI = 1.32$ and $CV^2 = 0.49$ [33], respectively. It is noticeable that Dataset 1 predominantly exhibits intermittent and lumpy demand patterns, while Dataset 2 is characterized by lumpy and non-stationary demand patterns.

Figure 7 displays the distribution of demand quantity in both datasets. In Dataset 1, the demand quantities are primarily concentrated in the range of 0–15, with some instances of anomalous demand. In contrast, Dataset 2 exhibits larger demand quantities, concentrated between 1 and 100. Overall, both datasets show intermittent demand patterns with occasional anomalies, including some extreme anomalies.

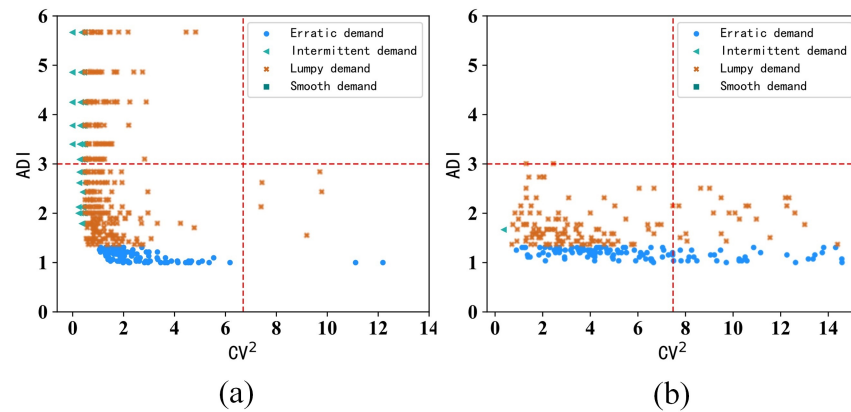


Figure 6. Intermittent distribution characteristic of (a) Dataset 1 and (b) Dataset 2. The metrics utilized for this analysis are the *ADI* (see Equation (1)) and CV^2 (see Equation (2)).

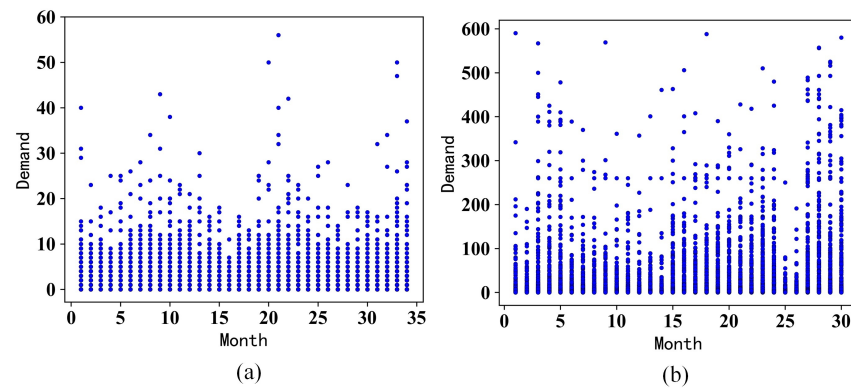


Figure 7. Distribution of demand quantities from (a) Dataset 1 and (b) Dataset 2.

4.2. Evaluation Metric

To evaluate the detection performance of proposed method, we employed the following evaluation metrics: *Precision*, *Recall*, and *F1-score*. These indicators are commonly used in classification problems to comprehensively quantify the performance of classifiers. Accuracy refers to the proportion of correctly classified samples to the total number of samples, which reflects the overall accuracy of the classifier. Recall, also known as the recall rate, refers to the proportion of correctly classified positive samples to true positive samples, which reflects the classifier’s ability to find all true positive examples. The *F1-score* is the harmonic average of the accuracy and recall, which comprehensively considers these two indicators and is a comprehensive indicator for evaluating classifier performance. The definitions of these three metrics are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

$$Recall = \frac{TP}{TP + FN} \tag{21}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{22}$$

where *TP* represent instances correctly detected as anomalies, *FP* represent instances incorrectly detected as anomalies when they are normal, *TN* represent instances correctly recognized as normal, and *FN* represent instances incorrectly recognized as normal when they are anomalies.

4.3. Comparative Experiment

In this section, a total of eight comparative methods were chosen, as listed in Table 2. We believe these eight methods cover the classical strategies of time series anomaly detection. The implementation introduction will be placed in the discussion of the experimental results.

Table 2. Introduction of the anomaly-detection methods for comparison.

Type	Name
Probability Model	COPOD [32]
Probability Model	ECOD [34]
Probability Model	PCA [35]
Partition-Based Method	IForest [8]
Distance-Based Method	KNN [7]
Density-Based Method	LOFs [6]
Classification-Based Method	OCSVM [5]
Deep Learning Method	DeepSVDD [9]

In enterprise maintenance, the SPD data typically are unlabeled due to the high cost of manual annotation. To have ground-truth labels, we selected two categories of spare parts from the two datasets and entrusted several enterprise engineers to annotate anomalous demand based on actual results. The final anomaly labeling (ground-truth) was obtained by merging the results from each expert. To improve understanding, we provide the statistical information of the annotated anomalous demands in Figure 8. It is clear that each sequence in Dataset 1 and Dataset 2 contains a maximum of 4 and 3 anomalous demands, respectively. The quantity of anomalous demands will directly generate a discrete value of the three metrics. For instance, the value of recall for some sequences may be 100% or 75%, rather than a continuous value, which will be shown in the following figures.

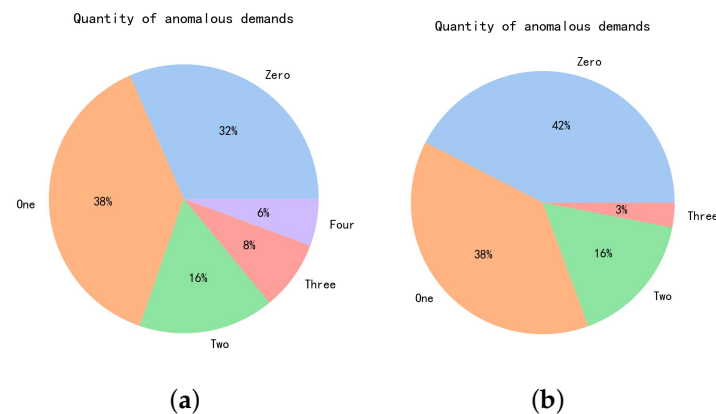


Figure 8. Statistical information of anomalous demands in (a) Dataset 1 and (b) Dataset 2.

For an overall evaluation, Figure 9 shows the detection accuracy of all nine methods on Dataset 1. Figure 10 provides the performance comparison of all nine methods on the two datasets in terms of the three metrics, accuracy, recall, and F1-score. Due to space limitations, we will not provide the accuracy results on Dataset 2 here. From Figure 9, the proposed method obtained much higher detection accuracy than the other eight methods. On some categories, the proposed method can obtain 100% accuracy. The comparative results validated the effectiveness of the dual-tailed probability in the unsupervised anomaly detection. From Figure 11, the proposed method outperformed all the other methods on the three metrics. We observed an interesting phenomenon in which Deep SVDD obtained a very low recall value on both datasets. The reason came from the adaptive feature extraction in Deep SVDD, which easily includes the anomalous demands into the constructed hypersphere. Consequently, the anomalous demands will be recognized

incorrectly. This also indicates that, for intermittent time series, deep learning techniques are not applicable due to the potential model bias.

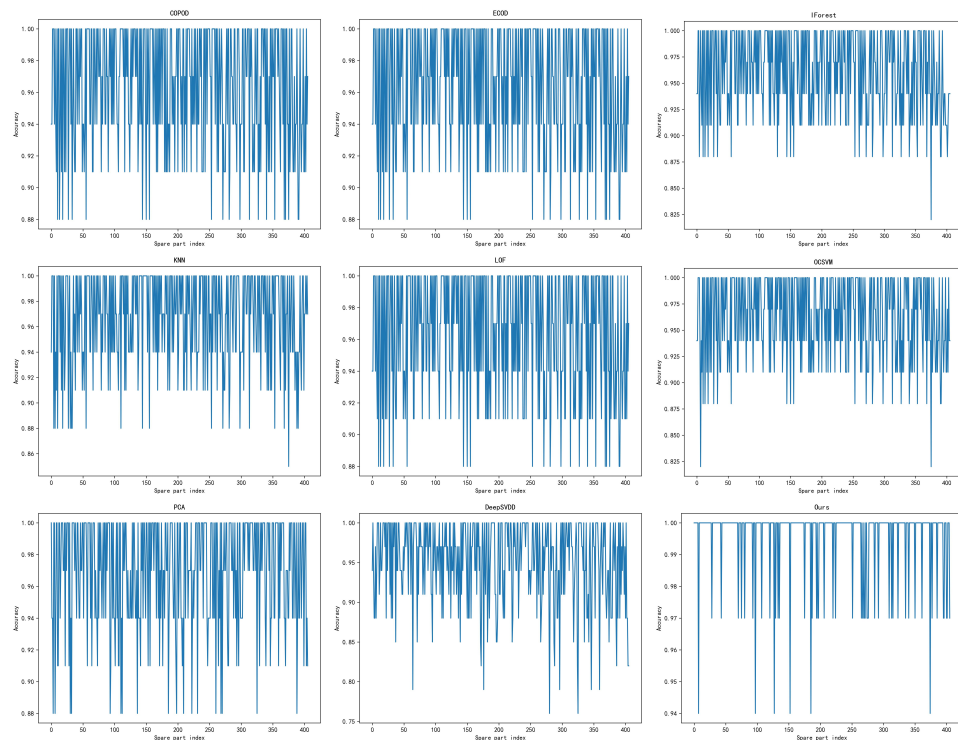
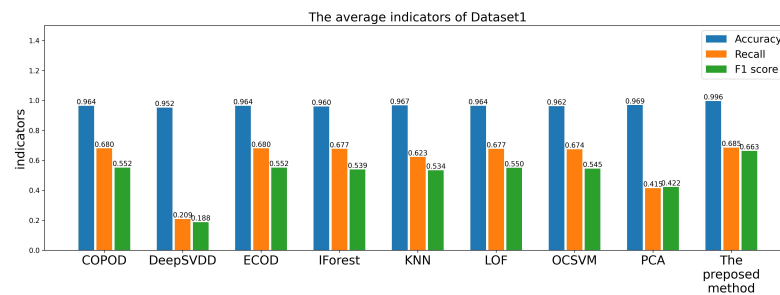
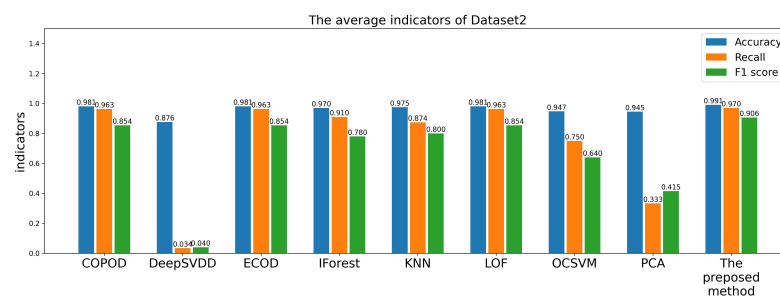


Figure 9. Detection accuracy of all 9 methods on the total of 407 sequences from Dataset 1.



(a)



(b)

Figure 10. Performance comparison of all 9 methods in terms of the the three metrics, accuracy, recall, and F1-score on (a) Dataset 1 and (b) Dataset 2.

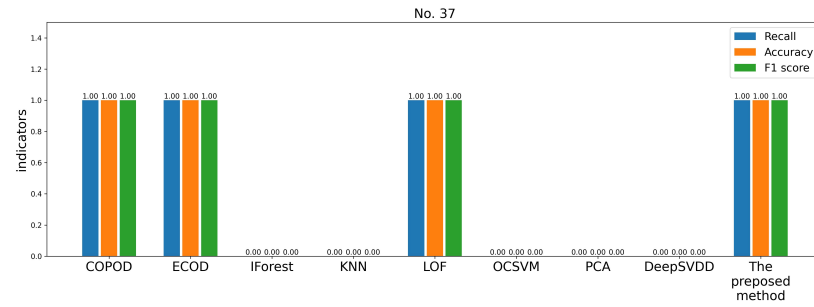
Here, we chose some sequences for a detailed comparison, as shown in Figure 11. To provide a visual evaluation, Figure 12 shows the detection results of all nine methods on these four sequences, accompanied by the expert-annotated anomalies. We also evaluated the operational efficiency of each compared method for a comprehensive evaluation. Table 3

gives the average execution time of 30 repeated trials for each method. Please note that the methods COPOD and ECOD are Croston-like methods, which essentially are statistical techniques. Since these two methods and our method were all built on probability analysis, they had no explicit formulation of time complexity for the comparison. So, we chose the execution time to evaluate the computational complexity. We believe that the average time of 30 repeated trials is able to provide an unbiased estimate for the complexity analysis. From the results listed, the proposed method can comprehensively consider the operation efficiency and detection accuracy of abnormal demand detection for spare parts and obtain the best performance in balancing the computational complexity and algorithmic effectiveness.

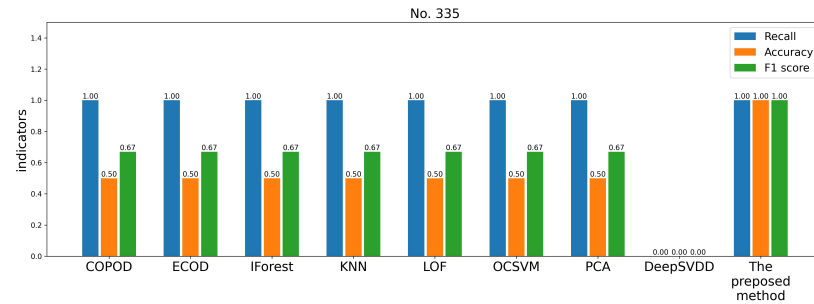
Table 3. Execution time of different methods (unit: seconds).

Dataset	COPOD	ECOD	IForest	KNN	LOFs	OCSVM	PCA	DeepSVDD	Proposed Method
Dataset 1	0.4548	0.4538	85.608	1.7463	1.7154	2.2751	1.7513	1064.5993	1.131
Dataset 2	0.2204	0.1865	50.1218	0.9828	0.8273	0.6951	1.166	449.343	0.4708

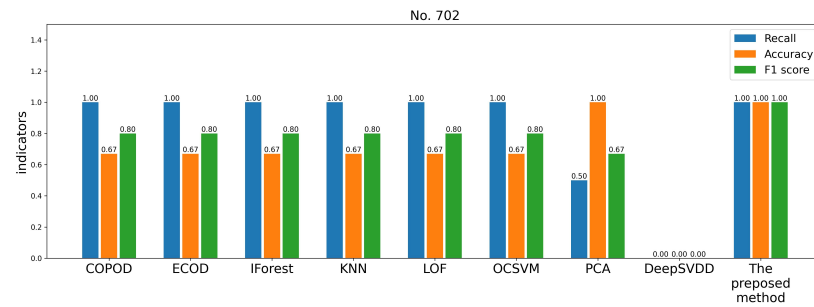
From Figures 11 and 12, COPOD and ECOD, the two unsupervised anomaly-detection methods, performed well on some sequences, effectively identifying anomalies with a fast execution time. Since both methods use empirical cumulative distribution functions, the anomaly detection based on the probability calculation can be proven valid. However, for the sequences including sparse demand, their detection results were unsatisfactory, indicating overly strict detection. In this case, they can easily recognize the normal demands as anomalies. IForest, an ensemble learning algorithm, is suitable for continuous data, but it performed poorly on the sequences with sparse demand. It failed to accurately identify anomalies, and its detection results for the No. 37 part did not align with the actual values. Additionally, its execution time increased rapidly with more base learners. KNN achieved unsupervised anomaly detection by evaluating the distribution distance between sample points. It exhibited high accuracy in detecting anomalies with prominent local demand variations. However, it introduced big errors in detecting sequences that were continuous and demand-dense. LOFs use a density estimation approach to search for the nearest neighbors and marked the sample points located in the sparse regions as anomalies. This method also performed well in detecting anomalies with prominent local variations in the demand sequence. OCSVM's detection results on most SPD sequences were poorer compared to the other methods. This method struggled to train appropriate support vectors from short-length, highly fluctuating sequences for the classification. PCA detected sequence anomalies by contrasting the changes in sample points before and after running eigenvalue decomposition. Its execution time was also higher than the methods with the probability model. It was sensitive to the maximum value in the sequence and lacked sensitivity to the anomalies, resulting in insufficient detection results. Deep SVDD is an anomaly-detection method based on deep neural networks. It had a significantly higher execution time compared to the other methods, and unfortunately, it was not able to effectively detect the anomalies within the sequences. It struggled to achieve high detection precision for intermittent sequences, with a higher number of false alarms and false alarms. Although the proposed method had a longer execution time than COPOD and ECOD, it was still more efficient than the other methods. From Figure 11 and Table 3, it can be concluded that the proposed method is a highly efficient and accurate detection method. Although the COPOD and ECOD methods have lower computational complexity than the proposed method, their detection results were worse than this method. The proposed method provided good detection results on all four categories of spare parts, which verified the effectiveness of the dual-tailed probability used.



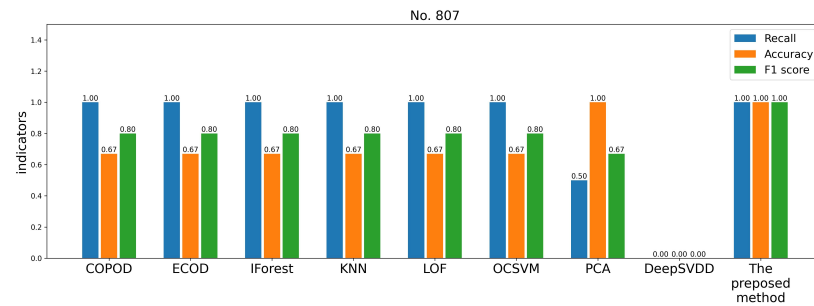
(a)



(b)



(c)



(d)

Figure 11. Detection performance by different methods, where (a–d) are for the spare parts numbered 37, 335, 702, and 807 from Dataset 1, respectively. Since each spare part generally has 2–4 abnormal demands (please refer to Figure 8), the value of the metric recall is limited, so the indicator values look identical.

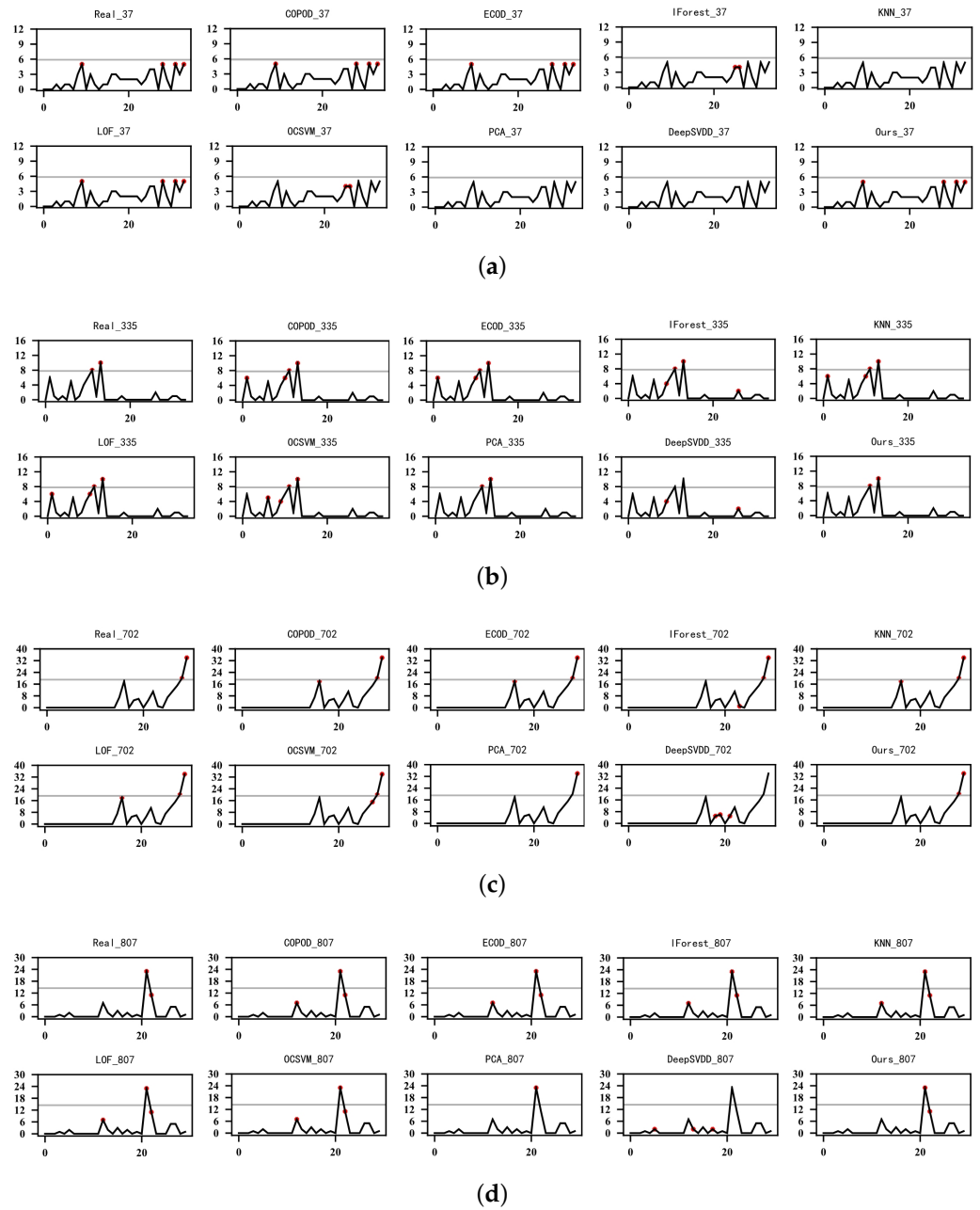


Figure 12. Visual detection results by different methods on Dataset 1, where (a–d), respectively, represent the results on the spare parts numbered 37, 335, 702, and 807. The x-axis and y-axis represent the month index and the number of spare parts required, respectively. Each image displays the algorithm’s name and the part index at the top. The red dot indicates the anomalous demand. The anomalies annotated by business experts are also provided.

4.4. Ablation Experiment

To analyze the influence of tensor decomposition and sequence segmentation on the detection results, a set of ablation experiments was conducted on Dataset 1, as shown in Table 4:

Table 4. Settings of ablation experiments.

Group	Fixed Part	Implementation
Experiment 1	Remove sequence segmentation and tensor decomposition	The original SPD sequences are dealt with only by COPOD
Experiment 2	Remove tensor decomposition	The original SPD sequences and the non-zero quantity sequences are dealt with by COPOD.
Experiment 3	Remove sequence segmentation	The sequences after tensor decomposition are dealt with by COPOD.
Experiment 4	None	The proposed method

Figure 13 displays the demand distribution for the No. 335 and No. 264 parts before and after running the tensor decomposition, with red points indicating changes in the demand values. It can be observed that the combination of the MDT with tensor decomposition effectively reduced the extreme demand in the sequence, providing a solid data support for the subsequent anomaly detection.

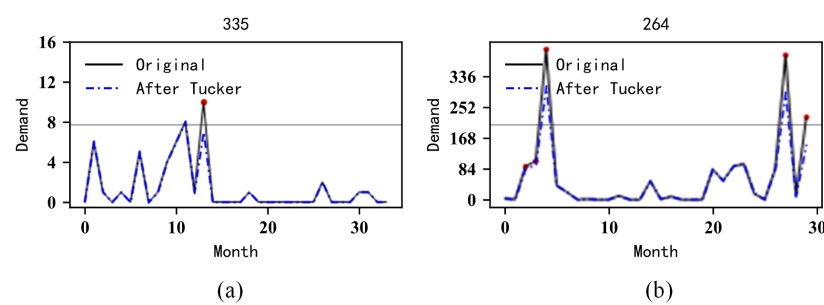


Figure 13. Demand changes before and after running tensor decomposition for the (a) No. 335 part and (b) No. 264 part. The red dot indicates the anomalous demand.

Figure 14 shows the comparative results of the ablation experiments that are shown in Table 4. Due to space limitations, we only provide the results on the two spare parts indexed by No. 264 and No. 335. The results indicated that sequence segmentation can effectively avoid overly strict detection results caused by sparse demand and address the problem of false alarms. For the No. 335 part, Experiment 1 and Experiment 3 detected many anomalous demands from the original SPD sequence. However, these anomalies showed almost no difference from the normal values, resulting in overly strict results that did not align with the actual business requirements. Moreover, the introduction of tensor decomposition was helpful to avoid missing alarms caused by extreme anomalous demand. In the No. 264 sequence, there were two distinct extreme demands with quantities several times higher than the daily demand. In Experiment 1 and Experiment 2, only extreme demands were recognized, while some anomalous demands were unable to be identified, leading to false alarms. In contrast, Experiment 3 and Experiment 4, which included tensor reconstruction, effectively mitigated the impact of extreme demands and accurately identified all abnormal points. The ablation results demonstrated that the proposed method can effectively address the issues of false alarms and missing alarms encountered in intermittent time series anomaly detection and provide a solid data support for spare parts planning and demand forecasting in intelligent maintenance.

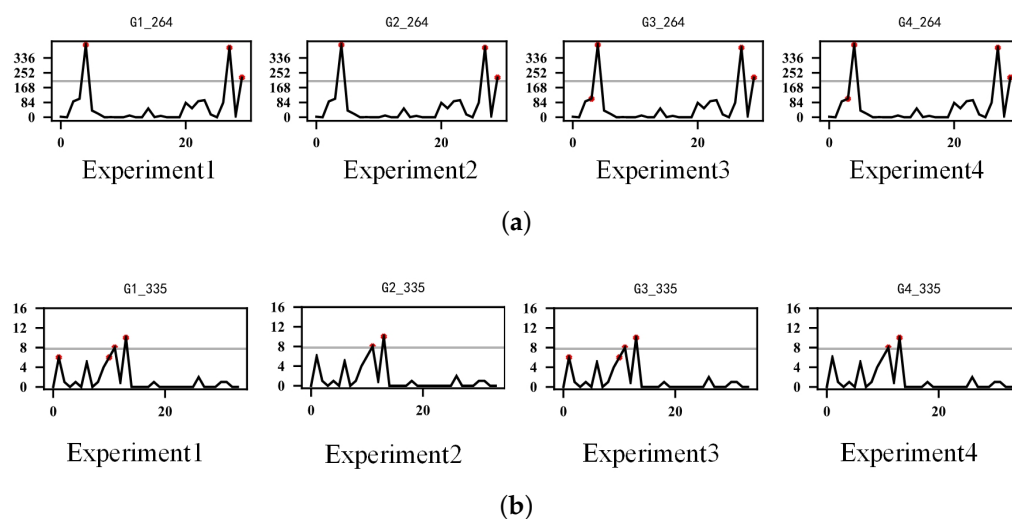


Figure 14. Results of ablation experiment on the (a) No. 264 part and (b) No. 335 part. The red dot indicates the anomalous demand. Due to space limitations, we do not provide the results on the other spare parts, which are similar to the results listed here.

5. Conclusions

In this paper, a new unsupervised anomaly-detection method was proposed to recognize the anomalous demands from intermittent SPD sequences. This method solves the missing alarms caused by extreme demands by employing the MDT technique and Tucker tensor decomposition. This method further solves the problem of false alarms caused by sparse demand through comprehensively evaluating the anomalous degree from the perspectives of demand occurrence time and demand quantity. We believe this method is of superior practical significance in the intelligent maintenance of large manufacturing enterprises. The specific conclusions are as follows:

- (1) The dual-tailed probability is suitable to detect anomalous demand from real-world SPD sequences since it does not require label information. Both false alarms and missing alarms can be effectively recognized in unsupervised mode, which can broaden the application range.
- (2) The proposed method has a very low computational cost by avoiding complex model training like deep learning techniques. A fast anomaly-detection method is of great importance to practical applications.

We would like to point out that the proposed method can not only be an effective solution, but also serve as a feasible framework for intermittent time series anomaly detection. The techniques used, e.g., Tucker tensor decomposition and tail probability calculation, can be replaced by more-efficient methods, providing good flexibility for future improvement.

In our future work, we plan to introduce the reliability concept to the proposed anomaly-detection method, since for practical applications, the maintenance decision should be trustworthy to engineers and managers. Machine learning techniques can also be introduced to conduct detection from the dynamic evolution of intermittent time series. To tackle the SPD sequences of multiple spare parts, a transfer learning algorithm can be used to reduce the concern about data dependency. Interpretability should be paid more attention to in the unsupervised anomaly-detection method.

Author Contributions: Conceptualization, K.H. and F.L.; methodology, W.M.; software, Y.L.; validation, Y.R. and W.M.; formal analysis, W.M.; investigation, W.M.; resources, K.H.; data curation, W.M.; writing—original draft preparation, Y.L.; writing—review and editing, W.M.; visualization, Y.R.; supervision, K.H.; funding acquisition, K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Key R&D Program of China under Grant 2020YFB1712105 and in part by the Open Project of State Key Laboratory of Shield and Tunneling Technology (SKLST-2021-K04).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Kairong Hong, Yingying Ren and Fengyuan Li were employed by the company China Railway Tunnel Group. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Bao, Y.; Wang, W.; Zou, H. SVR-based method forecasting intermittent demand for service parts inventories. In Proceedings of the Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing: 10th International Conference, RSFDGrC 2005, Regina, SK, Canada, 31 August–3 September 2005; Proceedings, Part II 10; Springer: Berlin/Heidelberg, Germany, 2005; pp. 604–613.
2. Van Horenbeek, A.; Buré, J.; Cattrysse, D.; Pintelon, L.; Vansteenwegen, P. Joint maintenance and inventory optimization systems: A review. *Int. J. Prod. Econ.* **2013**, *143*, 499–508. [[CrossRef](#)]
3. Moore, J.R., Jr. Forecasting and scheduling for past-model replacement parts. *Manag. Sci.* **1971**, *18*, B-200. [[CrossRef](#)]
4. Saaksvuori, A.; Immonen, A. *Product Lifecycle Management Systems*; Springer: Berlin/Heidelberg, Germany, 2008.
5. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural Comput.* **2001**, *13*, 1443–1471. [[CrossRef](#)] [[PubMed](#)]
6. Breunig, M.M.; Kriegel, H.P.; Ng, R.T.; Sander, J. LOF: Identifying density-based local outliers. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, Dallas, TX, USA, 16–18 May 2000; pp. 93–104.
7. Ramaswamy, S.; Rastogi, R.; Shim, K. Efficient algorithms for mining outliers from large data sets. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, Dallas, TX, USA, 16–18 May 2000; pp. 427–438.
8. Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation-based anomaly detection. *ACM Trans. Knowl. Discov. Data (TKDD)* **2012**, *6*, 1–39. [[CrossRef](#)]
9. Ruff, L.; Vandermeulen, R.; Goernitz, N.; Deecke, L.; Siddiqui, S.A.; Binder, A.; Müller, E.; Kloft, M. Deep one-class classification. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 4393–4402.
10. Laptev, N.; Amizadeh, S.; Flint, I. Generic and scalable framework for automated time-series anomaly detection. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 1939–1947.
11. Yaacob, A.H.; Tan, I.K.; Chien, S.F.; Tan, H.K. Arima based network anomaly detection. In Proceedings of the 2010 Second International Conference on Communication Software and Networks, Singapore, 26–28 February 2010; pp. 205–209.
12. Bontemps, L.; Cao, V.L.; McDermott, J.; Le-Khac, N.A. Collective anomaly detection based on long short-term memory recurrent neural networks. In Proceedings of the Future Data and Security Engineering: Third International Conference, FDSE 2016, Can Tho City, Vietnam, 23–25 November 2016; Proceedings 3; Springer: Berlin/Heidelberg, Germany, 2016; pp. 141–152.
13. Hundman, K.; Constantinou, V.; Laporte, C.; Colwell, I.; Soderstrom, T. Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 387–395.
14. Garg, A.; Zhang, W.; Samaran, J.; Savitha, R.; Foo, C.S. An evaluation of anomaly detection and diagnosis in multivariate time series. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 2508–2517. [[CrossRef](#)] [[PubMed](#)]
15. Deldari, S.; Smith, D.V.; Xue, H.; Salim, F.D. Time series change point detection with self-supervised contrastive predictive coding. In Proceedings of the Web Conference 2021, Ljubljana, Slovenia, 19–23 April 2021; pp. 3124–3135.
16. Socher, R.; Pennington, J.; Huang, E.H.; Ng, A.Y.; Manning, C.D. Semi-supervised recursive autoencoders for predicting sentiment distributions. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, 27–29 July 2011; pp. 151–161.
17. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
18. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
19. Park, D.; Hoshi, Y.; Kemp, C.C. A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder. *IEEE Robot. Autom. Lett.* **2018**, *3*, 1544–1551. [[CrossRef](#)]
20. Niu, Z.; Yu, K.; Wu, X. LSTM-based VAE-GAN for time-series anomaly detection. *Sensors* **2020**, *20*, 3738. [[CrossRef](#)] [[PubMed](#)]
21. Li, D.; Chen, D.; Goh, J.; Ng, S.K. Anomaly detection with generative adversarial networks for multivariate time series. *arXiv* **2018**, arXiv:1809.04758.
22. Li, D.; Chen, D.; Jin, B.; Shi, L.; Goh, J.; Ng, S.K. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In *International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 703–716.
23. Fan, L.; Zhang, J.; Mao, W.; Cao, F. Unsupervised Anomaly Detection for Intermittent Sequences Based on Multi-Granularity Abnormal Pattern Mining. *Entropy* **2023**, *25*, 123. [[CrossRef](#)] [[PubMed](#)]

24. Schmidl, S.; Wenig, P.; Papenbrock, T. Anomaly detection in time series: A comprehensive evaluation. *Proc. Vldb Endow.* **2022**, *15*, 1779–1797. [[CrossRef](#)]
25. Li, G.; Jung, J.J. Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges. *Inf. Fusion* **2023**, *91*, 93–102. [[CrossRef](#)]
26. Kim, J.; Kang, H.; Kang, P. Time-series anomaly detection with stacked Transformer representations and 1D convolutional network. *Eng. Appl. Artif. Intell.* **2023**, *120*, 105964. [[CrossRef](#)]
27. Ren, H.; Xu, B.; Wang, Y.; Yi, C.; Huang, C.; Kou, X.; Xing, T.; Yang, M.; Tong, J.; Zhang, Q. Time-series anomaly detection service at microsoft. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 3009–3017.
28. Bashar, M.A.; Nayak, R. TAnoGAN: Time series anomaly detection with generative adversarial networks. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 1–4 December 2020; pp. 1778–1785.
29. Boukhtouta, A.; Jentsch, P. Support vector machine for demand forecasting of canadian armed forces spare parts. In Proceedings of the 2018 6th International Symposium on Computational and Business Intelligence (ISCBI), Basel, Switzerland, 27–29 August 2018; pp. 59–64.
30. Yokota, T.; Erem, B.; Guler, S.; Warfield, S.K.; Hontani, H. Missing slice recovery for tensors using a low-rank model in embedded space. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8251–8259.
31. Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. *SIAM Rev.* **2009**, *51*, 455–500. [[CrossRef](#)]
32. Li, Z.; Zhao, Y.; Botta, N.; Ionescu, C.; Hu, X. COPOD: Copula-based outlier detection. In Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM), Sorrento, Italy, 17–20 November 2020; pp. 1118–1123.
33. Croston, J.D. Forecasting and stock control for intermittent demands. *J. Oper. Res. Soc.* **1972**, *23*, 289–303. [[CrossRef](#)]
34. Li, Z.; Zhao, Y.; Hu, X.; Botta, N.; Ionescu, C.; Chen, G. Ecod: Unsupervised outlier detection using empirical cumulative distribution functions. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 12181–12193. [[CrossRef](#)]
35. Shyu, M.L.; Chen, S.C.; Sarinnapakorn, K.; Chang, L. A novel anomaly detection scheme based on principal component classifier. In Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, Melbourne, FL, USA, 19–22 November 2003; pp. 172–179.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.