



Article Fault Detection for CNC Machine Tools Using Auto-Associative Kernel Regression Based on Empirical Mode Decomposition

Seunghwan Jung ¹, Minseok Kim ¹, Baekcheon Kim ¹, Jinyong Kim ¹, Eunkyeong Kim ¹, Jonggeun Kim ², Hyeonuk Lee ² and Sungshin Kim ³,*¹

- ¹ Department of Electrical and Electronics Engineering, Pusan National University, Busan 46241, Republic of Korea
- ² Artificial Intelligence Research Center, Korea Electrotechnology Research Institute, Changwon 51543, Republic of Korea
- ³ Department of Electrical Engineering, Pusan National University, Busan 46241, Republic of Korea
- * Correspondence: sskim@pusan.ac.kr; Tel.: +82-51-510-2374

Abstract: In manufacturing processes using computerized numerical control (CNC) machines, machine tools are operated repeatedly for a long period for machining hard and difficult-to-machine materials, such as stainless steel. These operating conditions frequently result in tool breakage. The failure of machine tools significantly degrades the product quality and efficiency of the target process. To solve these problems, various studies have been conducted for detecting faults in machine tools. However, the most related studies used only the univariate signal obtained from CNC machines. The fault-detection methods using univariate signals have a limitation in that multivariate models cannot be applied. This can restrict in performance improvement of the fault detection. To address this problem, we employed empirical mode decomposition to construct a multivariate dataset from the univariate signal. Subsequently, auto-associative kernel regression was used to detect faults in the machine tool. To verify the proposed method, we obtained a univariate current signal measured from the machining center in an actual industrial plant. The experimental results demonstrate that the proposed method successfully detects faults in the actual machine tools.

Keywords: machine tool; fault detection; empirical mode decomposition; auto-associative kernel regression

1. Introduction

Computerized numerical control (CNC) refers to a method of automating the control of machine tools by inputting programmed milling information into a microcomputer without a manual operator. The programmed milling specification is stored in the memory of the computer to process large amounts of work efficiently. Furthermore, by flexibly controlling various milling conditions (rotating speed, cutting force, etc.), high-quality material products can be produced at low costs. However, when the milling process is running, machine tools frequently suffer from faults, because these are operated in extreme environments to cut materials [1].

A fault is defined as an unpermitted deviation of at least one characteristic property of a variable from an acceptable behavior [2,3]. Although CNC machines have been well developed, they lack a function that diagnoses the condition of the tools and replaces faulty tools with new ones. The failure of machine tools requires maintenance time (also known as downtime) to replace the machine tool by stopping the machine. This may reduce the quality of the products and efficiency of the target process and increases the maintenance cost. In addition, if the damaged tool is used continuously to process the material, it may cause severe physical damage owing to the failure of the CNC machine. In fact, 79.6% of the maintenance time for CNC machines in modern industry is observed to be caused by damage to machine tools [4]. To solve these problems in the industrial field, when the life



Citation: Jung, S.; Kim, M.; Kim, B.; Kim, J.; Kim, E.; Kim, J.; Lee, H.; Kim, S. Fault Detection for CNC Machine Tools Using Auto-Associative Kernel Regression Based on Empirical Mode Decomposition. *Processes* **2022**, *10*, 2529. https://doi.org/10.3390/ pr10122529

Academic Editor: Jie Zhang

Received: 31 October 2022 Accepted: 24 November 2022 Published: 28 November 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/).

of the tool reaches approximately 60–70%, the operator replaces the tool with a new one. However, this traditional maintenance method cannot respond to unexpected faults, and maintenance costs may increase because of frequent tool replacement. The fault detection techniques can provide the precise process operating conditions to operators, and can help them take properly remedial actions.

Brief Review of Fault-Detection Approaches

The characteristic of the tool condition is sensitively reflected in the measured data because the machine tool rotates at a high speed with a strong force during the cutting process. To consider these properties, many studies have been conducted to detect the faults of machine tool using signal processing techniques [5,6], ensemble methods [7,8], and convolutional neural networks (CNNs) [9–12]. In signal processing techniques, the characteristics of the measured signals were analyzed using time-frequency analysis. The ensemble method combined a signal processing method and machine learning. In the case of using CNN, the samples measured from the target system were converted into a large number of images. Subsequently, CNN-based tool-condition monitoring was performed by training the configured image dataset. This approach has been used to detect faults in various industrial process systems such as motors [13], bearings [14], and drills [15].

Although the related methods mentioned above can successfully detect faults in the machine tool, they have several limitations. First, the characteristics of the data can change frequently during an actual milling process because CNC machines operate in various machining modes. Next, the deep learning algorithms require high-quality and massive image data for effective learning to ensure a competent performance. However, obtaining large amounts of normal and abnormal image data is practically difficult. This can lead to a data imbalance problem. Finally, converting the measured signals, obtained from a precisely operated machine into an image may result in the loss of information (feature) carried by the signal, or the CNN model may not be able to train the feature properly. Hence, fault-detection technology for machine tools is required to solve these problems. To address these challenges, we monitored the status of machine tools using empirical mode decomposition (EMD) and auto-associative kernel regression (AAKR). EMD can extract the intrinsic mode functions (IMFs) inherent in the original signal and is a suitable method for nonlinear and non-stationary signals, such as industrial process data. The current signal measured from the machine operating in the actual industrial field was converted into a multivariate dataset through EMD, because AAKR is an effective algorithm for multivariate datasets for detecting faults.

Generally, fault-detection methods for industrial processes can be divided into modelbased and data-driven approaches. The model-based approach is a method of detecting faults using a mathematical model built based on physical information about the target process, which primarily uses approaches to parity equations, parameter estimation, state observers, and signal models [2]. For example, dynamic property and parameter uncertainty in vehicle dynamic system were considered using mathematical model [16,17]. Shi and Zang diagnosed steering actuator faults in automated vehicles using model-based support vector machine (SVM) [18]. Mathematical models represent the dependencies of various signals that can be measured by a target system. If the physical information is insufficient or inaccurate, the performance of the designed model can be reduced. Furthermore, designing a complex or large-scale process using mathematical models is difficult. Therefore, data-driven methods have been extensively studied recently. The data-driven approach is an efficient alternative method, in which the important process information can be extracted from measured a sizable process data. These methods can be applied to various industrial processes because they do not require any prior physical knowledge of the target process. The multivariate statistical methods are widely used for process monitoring in various industrial applications [19–21].

Statistical process monitoring (SPM), which detects fault using multivariate statistics and machine-learning models, is the most popular data-driven fault-detection method [22].

Multivariate statistics for detecting abnormal scenarios have been studied intensively in the research on multivariate quality control [23]. Principal component analysis (PCA), independent component analysis (ICA), and AAKR have commonly been used for multivariate statistics. PCA and ICA can efficiently manage multivariate process data through a dimensionality reduction. However, these algorithms have a limitation in terms of the distribution of the latent variables. For example, PCA assumes that the hidden variables follow a Gaussian distribution. These statistical assumptions can degrade model performance because of the difficulty in satisfying actual industrial process data. In contrast, AAKR can effectively detect failures in various industrial processes due to no restriction on the abovementioned assumptions.

AAKR is a nonparametric multivariate regression method (also known as lazy, memorybased, and instance-based learning) that compares the similarity between the training data stored in memory and query vector. Subsequently, high weights are assigned to the training vectors with high similarity to compute the estimated vector. AAKR is updated online using a local model; therefore, extensive studies were conducted in [24–29] online monitoring of multivariate processes because of the advantage that it can be applied to time-varying processes operating in different process modes.

Huang proposed EMD in 1998 for decomposing the unique features contained in a univariate signal into amplitude modulation (AM)/frequency modulation (FM) components [30]. This method is not based on any assumption regarding linearity or stationary signals. Thus, many studies have been conducted to analyze nonlinear and non-stationary signals, such as biomedical signal [31,32] and industrial process data [33,34]. EMD is compared with the short-time Fourier transform (STFT) and wavelet transform (WT), a classical time-frequency analysis method. STFT provides a constant resolution for the entire signal through a fixed window. However, the resolution has limitations (also known as the uncertainty principle) owing to the trade-off between frequency and time. Therefore, this method is suitable for quasi-stationary signals. WT can extract the time–frequency characteristics of a signal through expansion and transformation. When using WT, multiscale signal analysis is possible, and it is more suitable for non-stationary signals than STFT. Nevertheless, WT has a critical disadvantage in that its analysis results depend on the selection of the wavelet basis function. In other words, only the signal characteristics correlated with the shape of the selected wavelet function generate high-value coefficients, and other characteristics may be masked or completely ignored. In contrast to WT, EMD is a self-adaptive signal processing method [35]. It can decompose the original signal into IMFs that reflect the vibrational mode features of the original signal without requiring a predefined basis function for signal analysis. Therefore, EMD is more suitable for analyzing industrial process data, which are nonlinear and non-stationary signals, than STFT and WT.

2. Preliminary

Most of the milling processes using a CNC machine have a smaller scale than chemical processes and power plants, and obtaining multivariate process data by installing a large number of sensors is difficult. EMD can extract some IMFs that can obtain multivariate data from univariate signals in the same time domain. This can address the inapplicability of multivariate fault detection models in the environment where only one signal is measured from the target system. In other words, multivariate statistical methods can be applied to univariate signals using EMD. In practice, PCA [36,37] and ICA [38,39] have been employed with EMD for process monitoring. In this study, EMD was used to extract IMFs from univariate signals measured using an actual CNC machine. Subsequently, we conducted an AAKR-based fault detection by constructing a multivariate dataset.

To demonstrate the efficiency of the proposed method, we used the current signal obtained from the long-term operation of an actual machine tool. Several tools of the same type were repeatedly utilized for straight and spiral cutting processes to obtain normal and fault data. The straight-cutting process provided tool breakage datasets. Artificial fault data were generated by applying disturbances (bias and drift) because the tools used in the spiral cutting process did not break. The proposed method was applied to the following two types of process datasets mentioned above: straight cutting process (actual tool breakage data) and spiral cutting process (artificial fault data through bias and drift). The experimental results demonstrated that the proposed method can effectively detect faults in a machine tool for two types of cutting processes. The contributions of this study are summarized as follows:

- 1. Existing signal processing and deep learning methods used for detecting faults in machine tools require the characteristics of the machining process and massive data, respectively. Furthermore, when deep learning is used, the features of the original signal can be lost during the process of converting a signal into an image. In contrast, the proposed method can efficiently perform tool-state diagnosis by avoiding these problems.
- 2. By using EMD, the proposed method has the advantage of applying a multivariate fault-detection model whose performance has been verified through prior related studies, even in a limited environment where only univariate signals can be acquired.
- 3. AAKR was employed to detect the fault in a machine tool for the first time. In addition, it has never been used in combination with EMD to detect faults in various industrial processes.
- 4. To obtain the actual machine tool data, we repeatedly conducted some experiments through straight parallel and spiral circular cutting, and then the proposed method was validated using massive data obtained from an actual operating CNC machine.

The remainder of this paper is organized as follows. Section 2 describes the proposed method for fault detection of a machine tool. The target systems and datasets used in this study are described in Section 3. The experimental results and discussion are presented in Section 4. Finally, Section 5 presents our conclusions.

3. Fault Detection of Machine Tools Using EMD and AAKR

Figure 1 shows the overall procedure for the fault detection of machine tools using EMD and AAKR. First, IMFs are extracted from the univariate current signal measured in a CNC machine using EMD. In this study, the number of IMFs was determined by repeatedly conducting experiments (trial and error) with the best AAKR performance. Subsequently, AAKR is trained using a dataset consisting of two extracted IMFs and original current signal. The first step in AAKR is to calculate the similarity between training and query vectors. Among the various distance functions for computing similarity, the most used Euclidean distance function was selected in this study. Next, a Gaussian weight function is used to assign weights to the training data vector. In this step, the appropriate value of the bandwidth parameter *h* in the Gaussian weighting function is determined using k-fold cross-validation. Subsequently, we calculate the estimated and residual vectors, and then, to obtain the detection indices, SPE, the multivariate statistic, is computed as the square of the residual vector. The confidence limit (also called the threshold and control limit) for declaring the fault was determined using kernel density estimation (KDE) in this study. More details of the AAKR algorithm for fault detection are presented in Figure 2 and Section 3.2.



Figure 1. Framework of fault-detection method for machine tool based on EMD and AAKR.



Figure 2. Schematic block diagram of fault-detection method based on AAKR.

3.1. Empirical Mode Decomposition

The EMD technique was proposed by Huang [30] in 1998 and can decompose a signal into an IMF. As a nonlinear, multiresolution, self-adaptive decomposition technique, EMD has been an effective signal processing tool for signals emanating from nonlinear and nonstationary systems [39]. The IMFs represent the natural oscillatory mode embedded in the signal and function as the basis functions, which are determined by the signal itself, rather than the pre-determined kernels [35]. IMFs must satisfy the following two conditions [30]: (i) in the entire dataset, the number of extrema and number of zero crossings must either be equal or differ at most by one, and (ii) at any point, the mean value of the envelope defined by the local maxima and envelope defined by the local minima is zero. EMD is based on the assumption that any signal contains different simple IMF components. With this definition, a signal x(t) can be decomposed as follows [30]:

- Step 1. Identify all local extrema of *x*(*t*).
- Step 2. Extract the *i*th IMF candidate c_i .
 - (a) Interpolate all local maxima and minima using cubic spline line. The connected lines are called the upper envelope $e_{max}(t)$ and lower envelope $e_{min}(t)$.

- Design the mean of the upper and lower envelope values as $m_1 = (e_{\max}(t) + e_{\min}(t))/2$.
- (c) Calculate the difference between the signals x(t), and m_1 is the first component, h_1 can be obtained as follows:

$$h_1 = x(t) - m_1. (1)$$

- Step 3. Verify that a₁ is correct for the IMF conditions. Ideally, if h₁ is an IMF, h₁ is the first component of x(t).
- Step 4. If h_1 is not an IMF, h_1 is considered the input signal, and repeat Steps 1–3, then

$$h_{11} = h_1 - m_{11}. \tag{2}$$

After repeated this shifting procedure *k* times, until h_{1k} is an IMF, that is

1

$$h_{1k} = h_{1(k-1)} - m_{1k}, \tag{3}$$

then h_{1k} is designated as

(b)

$$h_{1k} = c_1, \tag{4}$$

the first IMF component from the original signal.

• Step 5. Separate *c*₁ from the remainder of the data by

$$c_1 = x(t) - c_1, (5)$$

where r_1 is considered new data; the above processes are resumed, and the second IMF component c_2 can be obtained. We can repeat the above steps to achieve *n*-IMFs of signal x(t), as follows:

$$r_2 = r_1 - c_2, r_3 = r_2 - c_3.$$
 (6)

The shifting process can be stopped by any of the following predefined stopping criteria: (i) the component, c_n , or residue, r_n , becomes so small that it is less than the predetermined value, and (ii) the residue r_n becomes a monotonic function when no more IMF can be extracted. By summing Equations (5) and (6), we obtain

$$x(t) = \sum_{i=1}^{n} c_i + r_n.$$
 (7)

Finally, EMD is completed, and we can decompose the original signal into *n* empirical modes. A residue, r_n , can be either a mean trend or a constant of signal x(t). A more detailed explanation of the EMD is available in [30,35].

3.2. Fault-Detection Method Based on AAKR

Figure 2 shows the AAKR procedure used in this study. The fault-detection procedure using AAKR can be roughly divided into offline and online processes for training and testing the algorithm, respectively. In the offline process, we applied z-score normalization to the multivariate training dataset constructed by EMD to set the average and variance of each variable to 0 and 1, respectively. The bandwidth parameter of the weighting function was selected using *k*-fold cross-validation, where we used the Gaussian weighting function as the kernel function. The residual vector for the training data was then calculated using the leave-one-out method. The confidence limit (also known as the threshold value) for identifying faults was determined in advance. In this study, the threshold value for the fault declaration was obtained using the KDE. In the online process, similarity and weight vectors were computed for the query vector (testing data), and the estimated and residual vectors were computed based on them. Fault detection was then conducted by calculating multivariate statistics and comparing them with a pre-defined threshold in an offline process. A detailed description of fault detection is as following subsections.

3.2.1. AAKR Algorithm

AAKR is a multivariate regression method that stores training data in memory and assigns high weights to vectors having high similarity with the test vector to compute the score [40]. When multivariate data are collected from the target system, data matrix **X** is constructed as follows:

$$\mathbf{X} = [\mathbf{x}_{1}, \dots, \mathbf{x}_{j}]^{T} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{j1} & x_{j2} & \cdots & x_{jp} \end{bmatrix} \in \Re^{j \times p},$$
(8)

where *j* and *p* are the number of samples and process variables, respectively. In AAKR, if the training data matrix **X** is large, the algorithm requires longer training time. However, after the completion of the training phase, the appropriate weights were assigned to training vectors and AAKR can distinguish between normal and abnormal for query vectors. Generally, each variable of the multivariate data measured in the industrial process has a different mean and standard deviation. Therefore, if this dataset is trained on the model without data normalization, it can result in biased training results. For example, any variable with a large range of values is not important but can be considered more important than the other variables. To avoid this problem, data normalization is essential for setting equal mean and variance of all variables. In this study, we employed the z-score normalization technique to set the mean and variance of each variable to 0 and 1, respectively. The z-score normalization is defined as follows:

$$\mathbf{X}'_{l,m} = \frac{\mathbf{X}_{l,m} - \mu_m}{\sigma_m}, \text{ for } l = 1, \dots, j, m = 1, \dots, p,$$
 (9)

where μ_m and σ_m are the average and variance corresponding to the *m*th variable of the training dataset, respectively. In the AAKR algorithm, the similarity between the training and query vectors is calculated using the distance function to compute the estimated vector for the query vector. Several distance functions (such as the Euclidean distance and the Mahalanobis distance) are available for measuring the distance. To select the proper distance function, we compared the performance of the AAKR by applying Euclidean distance and Mahalanobis distance. Mahalanobis distance function required longer than Euclidean distance function. In addition, application of Euclidean distance improved performance compared with Mahalanobis distance. In this study, therefore, we employed the Euclidean distance defined in (10).

$$d_l(\mathbf{x}_l, \mathbf{x}(t)) = \sqrt{(\mathbf{x}_l - \mathbf{x}(t))^T (\mathbf{x}_l - \mathbf{x}(t))}, \text{ for } l = 1, \dots, j,$$
(10)

where $\mathbf{x}(t)$ is the query vector for time t, and \mathbf{x}_l is the training vector. If the training and query vector are highly similar, the value of the distance function reduces. The magnitude of the weight for each data vector is determined according to the degree of similarity measured using the distance function. In this study, we employed a Gaussian weighting function to assign weights as follows:

$$K_h(d_l) = \frac{1}{\sqrt{2\pi h}} \exp\left[-\frac{(d_l)^2}{2h^2}\right], \ l = 1, \dots, j,$$
 (11)

As the value of similarity increases, K_h assigns less weight to training data \mathbf{x}_i . In (11), h is a bandwidth parameter of the weighting function. In AAKR, multivariate regression performance depends on the bandwidth parameter value [41]. If the value of parameter h is too small, then we obtain crude estimation results because a small amount of data is used to estimate the query vector. However, if the value of parameter h is too large, we obtain smooth estimation results. The parameter value is typically set by following

a trial-and-error procedure for validation data [42]. For example, the objective function was used to select h with minimal error by varying the values h [43,44]. In this study, the optimal bandwidth parameter was determined using k-fold cross-validation, which is a method of repeatedly performing training and validation to prevent bias in the training data. This method is widely used to tune the hyperparameters of the models with inherent uncertainties. A more detailed explanation of parameter determination is available in [40]. In this study, the mean squared prediction error (MSPE) was calculated to determine the optimal parameter value as follows:

$$MSPE(q,s) = \frac{1}{\mathbf{X}_s} \sum_{l=1}^{\mathbf{X}_s} (\mathbf{x}_l - \hat{\mathbf{x}})^T (\mathbf{x}_l - \hat{\mathbf{x}}), \qquad (12)$$

where *q* is the number of repetitions up to $1 + (h_{max} - h_{min})/\Delta h$, and X_s is a subset of *k* training data randomly divided by *k* without duplication. h_{max} , h_{min} , and Δh are the maximum, minimum, and incremental values of *h*, respectively. Thus, we determine the value of bandwidth parameter *h* with the minimum MSPE as the optimal parameter value. When the proper parameter value is determined, the estimated vector $\hat{\mathbf{x}}(t)$, and the residual vector, $\mathbf{e}(t)$, are calculated as follows:

$$\hat{\mathbf{x}}(t) = \frac{\sum_{l=1}^{j} K_{h}(d_{l}) \mathbf{x}_{l}}{\sum_{l=1}^{j} K_{h}(d_{l})},$$
(13)

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t), \tag{14}$$

where $K_h(d_j)$ is the weight computed by the weighting function, and $\mathbf{x}(t)$ is the query vector at time *t*. The estimated vector, $\hat{\mathbf{x}}(t)$, is obtained using the weighted average. The residual vector, $\mathbf{e}(t)$, can be obtained by calculating the difference between the query vector, $\mathbf{x}(t)$, and estimated vector, $\hat{\mathbf{x}}(t)$. In this study, as presented in Algorithm 1, the residual vector for training data was calculated using the leave-one-out method [45]. More details of the leave-one-out method are presented in [40,45].

Algorithm 1: Leave-one-out method for calculating the residual vector

Input: Training data $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_j]$ $h \leftarrow$ bandwidth parameter determined by *k*-fold cross validation **for** *l* from 1 to *j* $\mathbf{X}' \leftarrow \mathbf{X}/{\{\mathbf{x}_l\}}$ Calculate the distance function values $d_l (\mathbf{x}_l, \mathbf{x}_i), l = \{1, \dots, j\} \setminus \{l\}$ between data vectors in \mathbf{X}' and \mathbf{x}_i Generate weights $K_{l_l}(d_l), l = \{1, \dots, j\} \setminus \{l\}$ of each data vector Obtain *l*th estimated vector $\hat{\mathbf{x}}_l$ Calculate residual vectors $\mathbf{e}_l = \mathbf{x}_l - \hat{\mathbf{x}}_l = [e_l^1, \dots, e_l^p]^T$ **end return** $\{\mathbf{e}^1, \dots, \mathbf{e}^j\}$

3.2.2. Detection Indices and Confidence Limit

Hotelling's T^2 statistic and *SPE* are typically used as multivariate statistics for fault detection. In this study, *SPE* was employ as a multivariate statistic and is calculated using the residual vector $\mathbf{e}(t)$ as follows:

$$SPE(t) = \mathbf{e}(t)^T \mathbf{e}(t).$$
 (15)

If the target system is normal, the magnitude of SPE(t) is small. In contrast, the values of SPE(t) begin to increase rapidly when a fault occurs in the system. Thus, SPE(t) represents the condition of the target system for multivariate monitoring. In fault-detection steps using multivariate statistics, the threshold should be defined in advance. In this study,

we employed KDE because it is widely used to define the threshold value for declaring the fault [29,46,47]. In KDE, a univariate kernel estimator with kernel function *K* is defined as

$$\hat{f}_{h}(x) = \frac{1}{jh} \sum_{l=1}^{J} K\left(\frac{x - x_{l}}{h}\right),$$
 (16)

$$\hat{F}_{h}(x) = \frac{1}{j} \sum_{l=1}^{j} W\left(\frac{x - x_{l}}{h}\right),$$
(17)

where $K(\cdot)$ is the kernel function; *j* is the number of samples; *h* is the smoothing parameter; and $W(t) = \int_{-\infty}^{t} K(u) du$. The kernel estimator, $\hat{f}_h(\cdot)$, is the sum of the bumps located in each sample. The shape of the bumps is determined using the kernel function $K(\cdot)$ [29]. Various kernel functions are available, such as Epanechnikov, uniform, and Gaussian. Among these kernel functions, we employed the most commonly used Gaussian kernel function. In practice, the confidence limit determined in KDE is influenced by the value of the smoothing parameter, *h*. To estimate cumulative kernel function, we used the 'ksdensity' MATLAB function built into the Statistic and Machine Learning Toolbox. More details of KDE and its smoothing parameter are available in [48,49].

3.2.3. Performance Indices

In this study, the false alarm rate (FDR) and miss detection rate (MDR) (also known as type I and type II errors, respectively) were used as performance indices to quantify the fault detection performance of the proposed method, as summarized in Table 1. They are based on statistical hypothesis tests and are widely used to quantify the performance of fault detection. In Table 1, H_0 is the null hypothesis, which means that the target system is normal, and H_1 is an alternative hypothesis that means that the target system is abnormal. That is, FAR indicates the rates of false alarms, in which the model has detected a fault, but the fault has not occurred in the actual target system. MDR refers to the rates of miss detected the fault. If the MDR value is high, the model does not detect faults occurring in the real target system. Therefore, MDR is considered to be much more important than FAR in this research field.

		Decision				
		Reject H ₀ (Accept H ₁)	Accept H ₀ (Reject H ₁)			
Truth	H_0 is true (H_1 is false)	FAR (Type I error)	Correct decision			
	$(H_1 \text{ is true})$	Correct decision	(Type II error)			

Table 1. The performance indices based on statistical hypothesis test.

4. Data Acquisition

In this section, we introduce the target system and milling process used to measure the current data. Current data are collected from the Internet of Things (IoT) sensor in an actual CNC machining center. The sampling frequency of the current data was 6 Hz. The models of the CNC machining center and cutting tool were Fanuc Oi, Mynx 5400, and YG-1 SUS-CUT, respectively. We used a CNC machining center operated in an actual manufacturing plant and stainless steel to obtain milling process data. Figure 3 shows the CNC machining center used to obtain the milling process data.



Figure 3. CNC machining center used in this study for data acquisition.

To obtain two types of milling process data, we conducted spiral circular and straight parallel cuttings, as illustrated in Figure 4. These cutting processes are primarily used to machine rigid materials during machining. In the spiral circular cutting process, we did not obtain tool breakage data despite repeatedly performing spiral circular cutting processes. Therefore, in this study, we generated artificial fault data by applying two types of disturbances (bias and drift) to normal data measured through a spiral cutting process. For the straight parallel cutting process, machine tools were broken during long-term operation. Thus, we first validated the proposed method by using two types of artificial fault data, and then conducted a real-world application by applying the proposed method to actual tool breakage data. The details of each machining method and the acquired dataset are described below.



Figure 4. Blueprints of process shape for spiral circular cutting and straight parallel cutting: (**a**) for spiral circular cutting process; (**b**) for straight parallel cutting process.

4.1. Case 1. Spiral Circular Cutting Process

Spiral cutting was performed clockwise until the length of the center of the processed stainless steels materials was 16 mm. We used three tools with a diameter of 10 phi, and 73 stainless materials were processed during spiral circular cutting. Notwithstanding the long-term operation of an actual machine tool, we could not obtain actual fault data. Hence, to obtain the fault data in this study, three types of artificial faults for each bias and drift disturbance (i.e., a total of six cases) were applied to the normal data measured by spiral circular cutting, respectively. After sampling approximately 23,500 normal data points among the datasets obtained through the spiral cutting process, 16,500 samples were used

for learning, and the remaining 7000 samples were used for testing. The normal dataset was sampled by considering the period immediately after replacing the tool with the new tool as the normal condition. To generate the artificial fault data, the bias and drift faults were applied to the test samples from the 4001st sample to the end.

- Bias fault: a step change of the current variable by N was introduced from the 4001st sample to the end (N = 0.02, 0.025, and 0.03, respectively).
- Drift fault: the original current variable was linearly increased from the 4001st sample by adding M(k 4000) to the current variable of each sample in this range, where *k* is the sample number (M = 0.0005, 0.00075, and 0.001, respectively).

The bias fault data have the form of a step function. These fault data describe a scenario in which fluctuations in the measured data are significant because of a fatal fault in the actual industrial process. In contrast, the drift fault data are in the form of a ramp function. This describes a phenomenon in which the size of a fault rapidly increases owing to the small faults in the facility. The generated artificial fault data in this manner were used as the test data to verify the proposed method.

4.2. Case 2. Straight Parallel Cutting Process

Figure 5 shows the status of straight cutting of stainless-steel materials before and after milling. As shown in Figure 5a, after the straight parallel cutting of processed stainless steel material 86 times, straight cutting was performed under the same conditions on the opposite side of the processed material. In this way, we used five tools with a diameter of 10 phi, and 68 stainless materials were machined during the straight cutting process.



Figure 5. Condition of straight parallel cutting and images of materials before and after the straight cutting process: (**a**) condition of straight cutting process; (**b**) before the straight milling process; (**c**) after the milling process.

Four of the five tools used in the straight cutting process broke during the process. Among the tool breakage cases, we used three fault datasets of tool breakage to verify the proposed method. Similar to the spiral circular process, the training dataset was sampled by considering the period immediately after replacing the used tool with a new tool as a normal state.

A summary of the datasets measured using the five tools is presented in Table 2. During the use of the first tool, several values were missing owing to communication errors between the machining center and IoT sensor. In addition, the first tool used in the straight cutting process did not break despite processing 18 materials being processed. The second tool cut a total of 16 materials, and broke when cutting the 34th material. The third tool suddenly broke when processing the first material. Because cause of breakage cannot be clearly explained, and it differed from the general failure of the running tool, this dataset was not used as test data. In these experiments, 14 and 15 materials were processed using the fourth and fifth tools, respectively, and each tool broke when the last material was processed. In this manner, we constituted the dataset used in this study for detecting faults in machine tools.

	First Tool	Second Tool	Third Tool	Fourth Tool	Fifth Tool
Tool failure type Material index	- #01~#18	Breakage #19~#34	Breakage #35~#39	Breakage #40~#53	Breakage #54~#68
Material index at the time of failure	-	#34	#35	#53	#68
Important issue	Communication issue	-	Unexpected early failure	-	-

Table 2. Summary of the datasets measured using five tools in straight cutting process.

5. Experimental Results and Discussion

In this section, we explain the results of the fault detection of the proposed method for the actual tool data introduced in Section 3. To perform comparative studies, we employed the local outlier factor (LOF), PCA, and ICA as comparison methods. LOF is a distancebased method frequently used to detect faults in multivariate processes, and can calculate local outlier factors for multivariate data by k-nearest neighbors (k-NN) using a predefined distance function. In this study, we set the value of *k* to 10 and employed the Euclidean distance to compute the distance between the nearest neighbors. PCA and ICA are the most well-known multivariate statistical techniques used primarily in large-scale processes that require the analysis of many process variables. They are suitable models for multivariate process data, owing to dimensionality reduction for the selection of useful process variables. In this study, Hotelling's T^2 and SPE statistics were used as detection indices for PCAbased process monitoring. The T^2 statistic is calculated based on the principal components (PCs) selected for dimensionality reduction. In the case of ICA, we employed the SPE and I_d^2 statistic that have commonly been used. The I_d^2 statistic is computed to consider the dominant part of independent components (ICs). For more details regarding the I_d^2 statistic, readers are invited to read the paper [46,47]. In ICA, we used the FastICA algorithm, which is a very simple and highly efficient fixed-point algorithm proposed in [50]. As mentioned above, the comparison methods are effective for a multivariate dataset. To consider the property, we combined EMD with comparative methods (i.e., EMD-LOF, EMD-PCA, and EMD-ICA) to verify the proposed method. The experimental results are discussed below.

5.1. Artificial Fault Cases (Spiral Circular Cutting Process)

In this subsection, we introduce the experimental results, in which artificial fault datasets generated by two types of disturbances were used to verify the performance of the proposed method. Artificial fault data were generated by applying drift and bias to normal data that were not used for training. Figure 6 shows the behavior of the faulty data generated by three types of artificial faults for each bias and drift disturbance. In these figures, the blue line describes the normal behaviors of the faulty variable, and the red line represents the abnormal behaviors of the faulty variable owing to disturbances. As shown in Figure 6a,c,e, the magnitude of current value has risen steeply at the 4001th sample by 0.02, 0.025, and 0.03, respectively. In Figure 6b,d,f, the current value gradually increases from time t = 4001 drift faults. From now on, only the results of the best cases for each disturbance will be explained to save space. The best cases for each bias and drift fault are N = 0.03 and M = 0.001, respectively.



Figure 6. Faulty variables generated by disturbances: (a) bias (N = 0.02); (b) bias (N = 0.025); (c) bias (N = 0.03); (d) drift fault (M = 0.0005); (e) drift fault (M = 0.00075); and (f) drift fault (M = 0.001).

Figure 7 depicts the IMF signals extracted from the normal and artificial fault data (bias and drift) using EMD. Each first-row signal in Figure 7 is the input signal of EMD, and the other row signals are the outputs of EMD. Some signals included in the red box in Figure 7 are the training (Figure 7a) and test datasets (Figure 7b,c) of AAKR. An appropriate number of IMFs was selected while increasing the number of IMFs from one to five in increments of one. When IMF₁ and IMF₂ were added to the dataset, AAKR exhibited the best fault-detection performance. That is, the final constituted dataset X = [bias faulty variable, IMF₁, IMF₂]. Thus, we constructed a multivariate dataset by extracting the IMFs from the univariate current signal.



Figure 7. (Spiral cutting data) Results of EMD for constructing the training and test datasets: (**a**) for normal training data, (**b**) for bias fault data, (**c**) for drift fault data.

Figure 8 shows the results of the bandwidth parameter selection using *k*-fold cross-validation. In this study, we set the value of *k* to 10, and the values of h_{max} , h_{min} , and Δh were set at 0.01, 1.3, and 0.01, respectively. As shown in the figure, the MSPE values decreased when h = 0.01, and then increased rapidly for the smallest value of *h*. From the *k*-fold cross-validation, the appropriate bandwidth parameter value was determined to be h = 0.06.



Figure 8. (Spiral cutting data) Result of bandwidth parameter selection using k-fold cross validation.

Figure 9 shows the empirical and estimated cumulative distribution functions (CDFs) for the detection indices. The blue and dashed green lines indicate the empirical and estimated CDFs, respectively. The inset of Figure 8 shows that the estimated CDF is adequately estimated for the empirical CDF. Therefore, we determined the confidence limits with a significance level of $\alpha = 0.01$ from the estimated CDFs.



Figure 9. (Spiral cutting data) Empirical and estimated CDFs of SPE for training samples.

Figures 10 and 11 show monitoring charts that are the results of fault detection using the comparison and proposed method for bias and drift data, respectively. The black line denotes the detection indices, and the dotted purple line indicates the confidence limits determined using KDE. In Figures 10f and 11, each monitoring chart at time t = 1 to 5000 is enlarged because the detection indices for the normal region are very low compared with that for the abnormal area. First, as shown in Figure 10, the LOF values exceeded the confidence limit before the bias fault were occurred. Furthermore, many LOF values after failure were often lower than the threshold. In the case of PCA and ICA, T^2 (PCA) and SPE (ICA) statistics did not entirely detect faults in the machine tool. SPE (PCA) and I_d (ICA) statistics can be distinguished between normal and abnormal areas, but each threshold is set to be high. This can lead to a high value of MDR. In contrast, the proposed method (EMD-AAKR) completely detected the faults in all areas except when a small degree of false alarm was declared at approximately time t = 2000. Next, in Figure 11, although the faults occurred from time 4001, the LOF values continuously exceeded the threshold from roughly t = 4500. In the case of PCA and ICA, T^2 (PCA) and SPE (ICA) statistics consecutively exceeded the threshold at approximately t = 4500. This increased the MDR value. SPE (PCA) and I_d (ICA) were expected to exhibit a lower MDR than other statistics (T^2 and SPE statistics). As shown in Figure 11f, detection indices for proposed method quickly exceeded the threshold at t = 4000 compared to other comparative methods, but some false alarms were declared in normal area. The false alarms were negligible considering the improvement of MDR for the proposed method.



Figure 10. (Bias fault, N = 0.03) Monitoring chart for the comparison and proposed method: (a) EMD-LOF (LOF), (b) EMD-PCA (T^2), (c) EMD-PCA (SPE), (d) EMD-ICA (I_d), (e) EMD-ICA (SPE), (f) EMD-AAKR (SPE).



Figure 11. (Drift fault, M = 0.001) Monitoring chart for the comparison and proposed methods: (a) EMD-LOF (LOF), (b) EMD-PCA (T^2), (c) EMD-PCA (SPE), (d) EMD-ICA (I_d), (e) EMD-ICA (SPE), (f) EMD-AAKR (SPE).

The performances of the proposed and comparison methods for the two types of artificial fault data are summarized in Table 3. The lowest values in each performance index are expressed in bold. As shown in Table 3, if the magnitude of each bias and drift is larger, then the performance of the proposed method and comparison methods are better; in case of bias 3 (N = 0.03), SPE (PCA) and I_d (ICA) have significantly improved performance compared to the other bias cases. LOF has a higher FAR than the other methods for both the bias and drift fault datasets. PCA and ICA have significantly lower FAR values for all datasets; however, the MDRs for T^2 (PCA) and SPE (ICA) were relatively high. In particular, the MDRs for all bias fault dataset were higher than 90%. This implies that the detection indices did not detect faults in the target system. In the case of drift fault dataset, SPE (PCA) and I_d (ICA) achieved lower FARs and MDRs than other comparative methods,

but the MDRs for bias fault data were still high. The proposed method outperformed the comparison methods; the FAR for proposed method were slightly higher than that of the other comparative methods except LOF, and the MDRs were the lowest than other methods. In particular, the MDR of the proposed method for the bias 3 (N = 0.03) was calculated as 0%. As the results of process monitoring, Table 3 shows that the proposed method successfully detects the artificial faults.

Disturbances	Indices	EMD-LOF	EMD-PCA		EMD	EMD-AAKR	
		LOF	T^2	SPE	I _d	SPE	SPE
Bias 1	FAR	2.7	0.025	0	0	0.05	0.25
(N = 0.02);	MDR	83.23	98.67	77.43	89.93	98.6	20.23
Bias 2	FAR	2.73	0.025	0	0	0.05	0.28
(N = 0.025);	MDR	77.93	97.7	48.43	62.01	98.23	0.5
Bias 3	FAR	2.7	0.03	0	0	0.05	0.3
(N = 0.03);	MDR	71.33	96.67	17	31.47	97.63	0
Drift 1	FAR	2.45	0.03	0	0	0.05	0.7
(M = 0.0005);	MDR	27.4	48.97	15.93	17.53	76.06	10.8
Drift 2	FAR	2.45	0.03	0	0	0.05	0.7
(M = 0.00075);	MDR	16.87	32.47	9.7	10.97	52.37	6.7
Drift 3	FAR	2.6	0.03	0	0	0.05	0.25
(M = 0.001);	MDR	9.35	23.8	7.8	8.07	39.32	3.21

Table 3. (Spiral cutting data.) Performance indices of comparison and proposed methods.

5.2. Actual Tool Fault Cases (Straight Parallel Cutting)

In this subsection, we introduce the fault-detection results of the proposed method for the tool breakage dataset. As mentioned in Section 4.2., straight parallel cutting was performed with five machine tools, and four of them broke. Among the broken tools, the dataset for fourth tool was excluded from the test dataset because of a sudden failure, which was different from other tool failures and difficult to explain. Thus, we used three types of tool breakage data measured using the second, fourth, and fifth tools to demonstrate the performance of the proposed method.

Figure 12 shows the original and IMFs signals extracted from the normal and tool breakage data (second, fourth, and fifth tool) using EMD. In Figure 12, each first-row signal is the original signal, and the others are the IMF extracted using EMD from the original current signal. By comparing with each first-row signal, we can confirm that the fault signal gradually increased owing to tool breakage compared with the normal signal. In Figure 12, the signals contained in the red box are composed of a set of training and test data of the AAKR for fault detection., In this case, the number of IMFs was determined using the method introduced in Section 5.1. Finally, we constructed the dataset **X** = [actual faulty variable, IMF1, IMF2] to verify the proposed method.

Figure 13 shows the results of the bandwidth parameter selection for straight parallel cutting data using *k*-fold cross-validation. We selected the value of *k* as 10 for *k*-fold cross validation. h_{max} , h_{min} , and Δh were set to 0.01, 1.3, and 0.01, respectively. As shown in this figure, when the value of *h* increased from the predefined initial value h_{min} , the corresponding MSPE decreased exponentially. Subsequently, MSPE attained minimal at h = 0.081. Therefore, we determined the value of the bandwidth parameter *h* to be 0.081.



Figure 12. (Straight cutting data) Results of EMD for constructing the training and test dataset: (a) for normal training data, (b) for second tool fault data, (c) for forth tool fault data, (d) for fifth tool fault data.



Figure 13. (Straight cutting data) Result of bandwidth parameter selection using k-fold cross validation.

Figure 14 depicts the empirical and estimated CDFs for detection indices. The blue and dashed green lines indicate the empirical and estimated CDFs, respectively. In Figure 8, the enlarged figure shows that the estimated CDF was reasonably estimated for the empirical CDF. Therefore, we determined the confidence limits with significance level of $\alpha = 0.01$ from the estimated CDFs.



Figure 14. (Straight cutting data) Empirical and estimated CDFs of SPE for training samples.

Next, we provide the results of fault detection for three tool-breakage datasets gathered from an actual CNC machine. As mentioned in Section 4.2., the datasets using second, fourth, and fifth tools were used to verify the proposed method. Figures 15-17 show the monitoring charts applying the comparison and proposed methods to the three types of tool breakage datasets, as mentioned above. The black line denotes the detection indices, and the dotted purple line represents the thresholds determined using KDE. In Figures 16f and 17c,d,f, the monitoring charts at t = 1 to 5000 are enlarged because the detection indices for the normal region were very small compared with those for the abnormal area. First, for the results of second tool data, the values of SPE (PCA), I_d (ICA), and proposed method statistics began to increase at t = 400. From this, we can determine the occurrence of the first fault. However, the others did not identify the time points; the majority of LOF, T^2 (PCA), and SPE (ICA) statistics were below each threshold. Compared with the proposed method, since the confidence limits of SPE (PCA) and I_d (ICA) statistics were set too low, FAR was expected to be high. In the case of the fourth tool dataset, we identified the start time (t = 280) of consecutive faults through the monitoring chart of SPE (PCA), I_d (ICA), and SPE (AAKR) statistics. In addition, the proposed method was expected to have smaller false alarms. In contrast, LOF, T^2 (PCA), and SPE (ICA) statistics still did not detect the faults. As shown in Figure 17, the comparison methods outperformed each result for the second and fourth tool. Nevertheless, because LOF, T^2 (PCA), and SPE (ICA) could not detect the start time of continuous faults, the MDRs of these methods were expected to be high. On the other hand, SPE (PCA), I_d (ICA), and proposed method identified the outset (t = 420) of the faults. In particular, the proposed method achieves a satisfactory performance for every tool dataset.

The performance indices of the proposed and comparison methods for the actual tool breakage datasets are summarized in Table 4. The lowest values for each performance index are written in bold. For the results of every tool dataset, the FARs and MDRs of the LOF, T^2 (PCA), and SPE (ICA) were calculated as very low and high, respectively; in the case of second and fourth tool, the MDR of these methods were approximately 100%. As can be seen from Figures 15 and 16, these methods could not detect the faults. SPE (PCA) and I_d (ICA) achieved better performance than the other comparison methods; for the fifth tool dataset, these methods outperformed the proposed method, but FARs were the highest in the case of second and fourth tools. The proposed method for every tool dataset achieved satisfactory performance in comparison to SPE (PCA) and I_d (ICA); although FAR for the second tool was high, except this case, both FAR and MDR were relatively low.



Figure 15. (Second tool.) Monitoring chart for the comparison and proposed methods: (a) EMD-LOF (LOF), (b) EMD-PCA (T^2), (c) EMD-PCA (SPE), (d) EMD-ICA (I_d), (e) EMD-ICA (SPE), (f) EMD-AAKR (SPE).



Figure 16. (Fourth tool.) Monitoring chart for the comparison and proposed methods: (a) EMD-LOF (LOF), (b) EMD-PCA (T^2), (c) EMD-PCA (SPE), (d) EMD-ICA (I_d), (e) EMD-ICA (SPE), (f) EMD-AAKR (SPE).



Figure 17. (Fifth tool.) The monitoring chart for the comparison and proposed methods: (**a**) EMD-LOF (LOF), (**b**) EMD-PCA (T^2), (**c**) EMD-PCA (SPE), (**d**) EMD-ICA (I_d), (**e**) EMD-ICA (SPE), (**f**) EMD-AAKR (SPE).

Tab	le 4.	(Straigl	ht cutting o	data)	Pe	rformance inc	lices of	compar	ison and	l pro	posed	meth	ods.
-----	-------	----------	--------------	-------	----	---------------	----------	--------	----------	-------	-------	------	------

Cases	Indices	EMD-LOF	EMD	D-PCA	EME	EMD-AAKR	
		LOF	T^2	SPE	I _d	SPE	SPE
Second tool	FAR	0.73	0	37.96	35.74	0	16.5
	MDR	95.61	99.83	0	0	99.4	0.09
Fourth tool	FAR	0.71	0	10	5.86	0	2.5
	MDR	84.49	98.01	0.16	0.27	97.14	1.35
Fifth tool	FAR	0.77	0.77	0.26	0.51	0.1	1.02
	MDR	36.79	59.9	0.25	0.47	77.68	2.04

5.3. Discussion

In this subsection, we summarize the advantages of our study. The main strength of our research is that we can select various multivariate models by using EMD in scenarios where only univariate process data must be used. EMD has been rarely used for detecting faults in machine tools, and this study is the first to combine EMD with AAKR and detect faults in industrial processes. Moreover, we conducted experiments to obtain actual tool fault data: spiral circular and straight parallel cutting processes. These cutting processes are primarily used to machine rigid materials during machining. During the machining process, we acquired real tool breakage datasets and validated the performance of proposed method using various cases of actual tool breakage data. As explained in Sections 5.1 and 5.2, the proposed method achieved better performance than the comparison methods.

6. Conclusions

Machine tools play an important role in manufacturing processes using CNC machining center. They operate repeatedly for a long period for machining hard and difficult-tomachine materials, such as stainless steel. These operating conditions can easily result in tool breakage. The failure of operating tools requires a downtime to install a new tool by stopping the machine. Moreover, machine tool failure significantly reduces the product quality and efficiency of the target process. Therefore, fault-detection techniques for machine tools are required to reduce the downtime and improve the stability of the process. In this study, EMD and AAKR algorithms were applied to various scenarios of actual tool breakage datasets obtained from a CNC machining center in an actual industrial plant. To obtain the fault data, we conducted several experiments: spiral circular and straight parallel cutting process. EMD was employed to construct a multivariate dataset from the univariate current signal measured using the CNC machining center. After constituting the multivariate datasets, AAKR was used to monitor the machine tool and detect the fault. The experimental results demonstrated that the proposed method successfully detects faults in the machine tools.

In future research, we will consider the following topics: The first is the consideration of multiple cutting modes. Many cutting modes are used in actual manufacturing processes using a CNC machining center. However, in this study, only two cutting modes, spiral circular and straight parallel cutting, were considered to monitor the machine tool. Thus, we will conduct further research to consider multiple cutting processes. The second topic is the consideration of dynamic properties. CNC machines are operated by a time-varying process. In addition, most industrial processes are time-varying. Therefore, dynamic properties should be considered when monitoring industrial processes.

Author Contributions: Conceptualization, S.J. and M.K.; methodology, S.J.; software, B.K.; validation, S.J. and M.K.; formal analysis, J.K. (Jinyong Kim), and E.K.; investigation, M.K.; resources, J.K. (Jonggeun Kim) and H.L.; data curation, H.L. and J.K. (Jonggeun Kim); writing—original draft preparation, S.J. and M.K.; writing—review and editing, J.K. (Jinyong Kim), B.K., E.K. and S.K.; visualization, S.J.; supervision, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Korea Electrotechnology Research Institute (KERI) Primary Research Program through the National Research Council of Science & Technology (NST) funded by the Ministry of Science and ICT (MSIT) (No. 22A01019).

Data Availability Statement: Not applicable.

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

References

- 1. Luo, B.; Wang, H.; Liu, H.; Li, B.; Peng, F. Early fault detection of machine tools based on deep learning and dynamic identification. *IEEE Trans. Ind. Electron.* **2018**, *66*, 509–518. [CrossRef]
- 2. Isermann, R. Model-based fault-detection and diagnosis—Status and applications. Annu. Rev. Control 2005, 29, 71–85. [CrossRef]

- 3. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000.
- Teti, R.; Jemielniak, K.; O'Donnell, G.; Dornfeld, D. Advanced monitoring of machining operations. CIRP Ann. 2010, 59, 717–739. [CrossRef]
- 5. Zhou, C.; Yang, B.; Guo, K.; Liu, J.; Sun, J.; Song, G.; Zhu, S.; Sun, C.; Jiang, Z. Vibration singularity analysis for milling tool condition monitoring. *Int. J. Mech. Sci.* 2020, *166*, 105254. [CrossRef]
- 6. Zhou, C.; Guo, K.; Sun, J.; Yang, B.; Liu, J.; Song, G.; Sun, C.; Jiang, Z. Tool condition monitoring in milling using a force singularity analysis approach. *Int. J. Adv. Manuf. Technol.* 2020, 107, 1785–1792. [CrossRef]
- 7. Madhusudana, C.K.; Gangadhar, N.; Kumar, H.; Narendranath, S. Use of discrete wavelet features and support vector machine for fault diagnosis of face milling tool. *Struct. Durab. Health Monit.* **2018**, *12*, 111.
- Mohanraj, T.; Yerchuru, J.; Krishnan, H.; Aravind, R.N.; Yameni, R. Development of tool condition monitoring system in end milling process using wavelet features and Hoelder's exponent with machine learning algorithms. *Measurement* 2021, 173, 108671. [CrossRef]
- Aghazadeh, F.; Tahan, A.; Thomas, M. Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process. *Int. J. Adv. Manuf. Technol.* 2018, 98, 3217–3227. [CrossRef]
- 10. Cao, X.; Chen, B.; Yao, B.; Zhuang, S. An intelligent milling tool wear monitoring methodology based on convolutional neural network with derived wavelet frames coefficient. *Appl. Sci.* **2019**, *9*, 3912. [CrossRef]
- Valtierra-Rodriguez, M.; Rivera-Guillen, J.R.; Basurto-Hurtado, J.A.; De-Santiago-Perez, J.J.; Granados-Lieberman, D.; Amezquita-Sanchez, J.P. Convolutional neural network and motor current signature analysis during the transient state for detection of broken rotor bars in induction motors. *Sensors* 2020, 20, 3721. [CrossRef] [PubMed]
- Zafar, T.; Kamal, K.; Kumar, R.; Sheikh, Z.; Mathavan, S.; Ali, U. Tool Health Monitoring Using Airborne Acoustic Emission and a PSO-Optimized Neural Network. In Proceedings of the 2015 IEEE 2nd International Conference on Cybernetics (CYBCONF), Gdynia, Poland, 24–26 June 2015; pp. 271–276.
- 13. Hsueh, Y.-M.; Ittangihal, V.R.; Wu, W.-B.; Chang, H.-C.; Kuo, C.-C. Fault diagnosis system for induction motors by CNN using empirical wavelet transform. *Symmetry* **2019**, *11*, 1212. [CrossRef]
- 14. Shao, H.; Xia, M.; Han, G.; Zhang, Y.; Wan, J. Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and thermal images. *IEEE Trans. Ind. Inform.* 2020, 17, 3488–3496. [CrossRef]
- 15. Tran, T.; Lundgren, J. Drill fault diagnosis based on the scalogram and mel spectrogram of sound signals using artificial intelligence. *IEEE Access* **2020**, *8*, 203655–203666. [CrossRef]
- 16. Liu, Y.; Chen, R.; Na, X.; Luo, Y.; Zhang, H. Robust fault estimation of vehicular yaw rate sensor using a type-2 fuzzy approach. *IEEE Trans. Ind. Electron.* **2020**, *68*, 10029–10039. [CrossRef]
- 17. Zhang, G.; Yu, Z.; Wang, J. Correction of contaminated yaw rate signal and estimation of sensor bias for an electric vehicle under normal driving conditions. *Mech. Syst. Signal Process.* **2017**, *87*, 64–80. [CrossRef]
- 18. Shi, Q.; Zhang, H. Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets. *IEEE Trans. Ind. Electron.* **2020**, *68*, 6248–6256. [CrossRef]
- 19. Russell, E.L.; Chiang, L.H.; Braatz, R.D. Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000.
- 20. Dai, X.; Gao, Z. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2226–2238. [CrossRef]
- 21. Zhang, Y.; Dudzic, M.S. Online monitoring of steel casting processes using multivariate statistical technologies: From continuous to transitional operations. J. Process Control 2006, 16, 819–829. [CrossRef]
- Qin, S.J. Data-driven fault detection and diagnosis for complex industrial processes. *IFAC Proc. Vol.* 2009, 42, 1115–1125. [CrossRef]
- 23. Jackson, J.E. A User's Guide to Principal Components; John Wiley & Sons: Hoboken, NJ, USA, 2005.
- 24. Di Maio, F.; Baraldi, P.; Zio, E.; Seraoui, R. Fault detection in nuclear power plants components by a combination of statistical methods. *IEEE Trans. Reliab.* 2013, *62*, 833–845. [CrossRef]
- 25. Hines, J.W.; Garvey, D.R. Development and application of fault detectability performance metrics for instrument calibration verification and anomaly detection. *J. Pattern Recognit. Res.* **2006**, *1*, 2–15. [CrossRef] [PubMed]
- Yang, D.; Usynin, A.; Hines, J.W. Anomaly-Based Intrusion Detection for SCADA Systems. In Proceedings of the 5th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies (NPIC & HMIT 05), Albuquerque, NM, USA, 12–16 November 2006; pp. 12–16.
- 27. Baraldi, P.; Di Maio, F.; Pappaglione, L.; Zio, E.; Seraoui, R. Condition monitoring of electrical power plant components during operational transients. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2012**, *226*, 568–583. [CrossRef]
- Baraldi, P.; Canesi, R.; Zio, E.; Seraoui, R.; Chevalier, R. Signal Grouping for Condition Monitoring of Nuclear Power Plant Components. In Proceedings of the Reliability and Risk Management ESREL 2011, Troyes, France, 18–22 September 2011; pp. 1–13.
- 29. Yu, J.; Yoo, J.; Jang, J.; Park, J.H.; Kim, S. A novel hybrid of auto-associative kernel regression and dynamic independent component analysis for fault detection in nonlinear multimode processes. *J. Process Control* **2018**, *68*, 129–144. [CrossRef]

- Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1998**, 454, 903–995. [CrossRef]
- 31. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Signal Process. Control* **2014**, *14*, 19–29. [CrossRef]
- 32. Park, C. Data-Driven Multivariate and Multiscale Methods for Brain Computer Interface. Ph.D. Thesis, Imperial College London, London, UK, 2012.
- 33. Yu, D.; Cheng, J.; Yang, Y. Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. *Mech. Syst. Signal Process.* **2005**, *19*, 259–270. [CrossRef]
- 34. Li, R.; He, D. Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantification. *IEEE Trans. Instrum. Meas.* **2012**, *61*, 990–1001. [CrossRef]
- 35. Lei, Y.; Lin, J.; He, Z.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126. [CrossRef]
- Jiang, L.; Gong, G.; Zhang, Y.; Liu, Z.; Cui, J. Fault Detection of Rolling Bearing Based on Emd-Dpca. In Proceedings of the 26th Chinese Control and Decision Conference (2014 CCDC), Changsha, China, 31 May–2 June 2014; pp. 3207–3211.
- Shi, H.; Guo, J.; Yuan, Z.; Liu, Z.; Hou, M.; Sun, J. Incipient fault detection of rolling element bearings based on deep EMD-PCA algorithm. *Shock Vib.* 2020, 2020, 8871433. [CrossRef]
- Miao, Q.; Wang, D.; Pecht, M. Rolling Element Bearing Fault Feature Extraction Using EMD-Based Independent Component Analysis. In Proceedings of the 2011 IEEE Conference on Prognostics and Health Management, Denver, CO, USA, 20–23 June 2011; pp. 1–6.
- Žvokelj, M.; Zupan, S.; Prebil, I. EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis. J. Sound Vib. 2016, 370, 394–423. [CrossRef]
- Yu, J.; Jang, J.; Yoo, J.; Park, J.H.; Kim, S. Bagged auto-associative kernel regression-based fault detection and identification approach for steam boilers in thermal power plants. *J. Electr. Eng. Technol.* 2017, 12, 1406–1416.
- 41. Hines, J.W.; Seibert, R. *Technical Review of On-Line Monitoring Techniques for Performance Assessment*; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 2006; Volume 1.
- Al-Dahidi, S.; Baraldi, P.; Di Maio, F.; Zio, E. Quantification of Signal Reconstruction Uncertainty in Fault Detection Systems. In Proceedings of the PHM Society European Conference, Nantes, France, 8–14 July 2014; Volume 2.
- Brandsæter, A.; Manno, G.; Vanem, E.; Glad, I.K. An Application of Sensor-Based Anomaly Detection in the Maritime Industry. In Proceedings of the 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), Ottawa, ON, Canada, 20–22 June 2016; pp. 1–8.
- 44. Baraldi, P.; Canesi, R.; Zio, E.; Seraoui, R.; Chevalier, R. Genetic algorithm-based wrapper approach for grouping condition monitoring signals of nuclear power plant components. *Integr. Comput.-Aided Eng.* **2011**, *18*, 221–234. [CrossRef]
- Wong, T.-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognit.* 2015, 48, 2839–2846. [CrossRef]
- 46. Lee, J.-M.; Yoo, C.; Lee, I.-B. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chem. Eng. Sci.* 2004, *59*, 2995–3006. [CrossRef]
- 47. Lee, J.-M.; Yoo, C.; Lee, I.-B. Statistical process monitoring with independent component analysis. *J. Process Control* **2004**, *14*, 467–485. [CrossRef]
- 48. Silverman, B.W. Density Estimation for Statistics and Data Analysis; CRC Press: Boca Raton, FL, USA, 1986; Volume 26.
- 49. Wand, M.P.; Jones, M.C. Kernel Smoothing; CRC Press: Boca Raton, FL, USA, 1994.
- 50. Hyvarinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **1999**, *10*, 626–634. [CrossRef]