

Research Article

Applying Grey Relational Analysis to Detect Change Points in Time Series

Yi-Chung Hu ¹, Shu-hen Chiang ², and Yu-Jing Chiu ¹

¹Department of Business Administration, Chung Yuan Christian University, Taoyuan City, Taiwan

²Department of Finance, Chung Yuan Christian University, Taoyuan City, Taiwan

Correspondence should be addressed to Yi-Chung Hu; ychu@cycu.edu.tw

Received 11 July 2022; Revised 19 August 2022; Accepted 26 August 2022; Published 30 September 2022

Academic Editor: Xiangfeng Yang

Copyright © 2022 Yi-Chung Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The goal of detecting change points is to recognize abrupt changes in time series data. This is suitable, for instance, to find events that characterize the financial market or to inspect data streams of stock returns. Regression models categorized as supervised methods have played a significant role in change-point detection. However, since change points might not be available beforehand to train the model, and because the series data might be statistically atypical, the applicability of regression models is limited. To avoid statistical assumptions, this study uses the grey theory, a kind of artificial intelligence tools, to measure the relationships between sequences by grey relational analysis (GRA). This paper contributes to propose an unsupervised method to detect possible change points in time series by GRA. Change-point analysis of the proposed method was performed on S&P100 stock returns. Experimental results from evaluating the recognition accuracy rate show that the proposed method performs well compared to other methods considered for change-point detection.

1. Introduction

A change point represents a transition between different sequences or states of a time series [1]. From the perspective of statistics, the probability density functions of two consecutive sequences resulting from a change point are different [2, 3]. Detecting transitions in time series data has become increasingly important. Many change-point detection (CPD) methods have been proposed for a range of real-world problems to detect and react to interesting events, such as climate change detection [4, 5], image analysis [6], hydrology [2, 7, 8], medical issues [9, 10], and tourism forecasting [11, 12]. Also, statistical methods, such as the likelihood ratio test, the standard normal homogeneity procedure [13], and the regression, have taken a significant role in CPD.

Learning methods applied to CPD problems can be either supervised or unsupervised. Regression models such as the logistic regression [14] and support vector machines [15, 16] can be treated as a supervised approach, with which sufficient training data with labels need to be provided for

reasonable performance [1, 17, 18]. However, it is possible that change points are totally unknown or that there are only few available prior to training. Furthermore, the collected time series data might well not conform to statistical properties such as homogeneity and a normality of errors. As such, regression is restricted when applied to CPD. In contrast to supervised methods, unsupervised learning uses unlabeled data to find desired patterns [19]. Thus, to expand the applicability of regression, we try to use a regression-like method to measure the relationship between response and explanatory variables in order to develop an unsupervised method for CPD.

Given a time series, each variable in the series can be extended as a sliding window with a sequence of time series variables [14, 20, 21]. It is found that the grey theory, categorized into artificial intelligence tools [22], can effectively measure the degrees of relationships among sequences by grey relational analysis (GRA) [23–27]. To estimate the relationship between a reference sequence and a set of comparative sequences, GRA treats the reference sequence as the desired goal or the response variable [28]. Indeed,

GRA has been widely applied to diverse real-world problems (e.g., [29–40]). In practice, GRA assigns a so-called grey relational grade (GRG) to each comparative sequence, such that the greater the GRG, the closer the relationship to the reference sequence. For CPD, given a reference sequence, when the respective GRGs of time series variables are obtained from a set of comparative sequences, we can investigate how these ratings can be used to determine all possible change points. Insofar as it does not require prior training, this paper contributes to propose an unsupervised CPD method for detecting multiple change points in time series with GRA.

Despite the usefulness of GRA, few studies have addressed its adoption for CPD. For instance, Wong et al. [2] used GRA to analyze change points in hydrological time series. The main difference between the proposed grey CPD method and Wong et al.'s GRA-based method (WGM) is that the proposed method can detect multiple change points from the collected data, but only one change point could be detected by WGM. Since multiple change points often exist in time series data, this makes WGM restricted. To evaluate the performance of the proposed grey CPD method in terms of whether change points can be detected, several metrics can be adopted, including the sensitivity, specificity, and G -mean [1].

The rest of the paper is organized as follows. Section 2 introduces regression-like GRA. Section 3 presents the proposed grey CPD method. In Section 4, we examine the detection performance of the proposed method using real data from daily log returns of stocks listed on the Standard and Poor's 100 (S&P100) index. Section 5 presents a discussion and the conclusions of this study.

2. Regression-Like GRA

Let (x_1, x_2, \dots, x_n) denote a time series sequence of length n , where x_i ($1 \leq i \leq n$) represents a variable at time stamp i . In a time series, sliding window $X_i = (x_i, x_{i+1}, \dots, x_{i+s-1})$ with length s ($1 \leq s \leq n$) can be used instead of x_i . Furthermore, $i+s-1$ cannot be greater than n . The reference sequence $X_i = (x_i(1), x_i(2), \dots, x_i(s)) = (x_i, x_{i+1}, \dots, x_{i+s-1})$ and $m+1$ comparative sequences, $X_{i+1} = (x_{i+1}(1), x_{i+1}(2), \dots, x_{i+1}(s)) = (x_{i+1}, x_{i+2}, \dots, x_{i+s})$, $X_{i+2} = (x_{i+2}(1), x_{i+2}(2), \dots, x_{i+2}(s)) = (x_{i+2}, x_{i+3}, \dots, x_{i+s+1})$, \dots , $X_{i+m} = (x_{i+m}(1), x_{i+m}(2), \dots, x_{i+m}(s)) = (x_{i+m}, x_{i+m+1}, \dots, x_n)$, are prepared for CPD, where $m = n - s - i + 1$. For instance, given $n = 6$, $i = 3$, and $s = 3$, it follows that $m = 1$ and $i + s - 1 \leq n$ hold. Thus, $X_3 = \{x_3, x_4, x_5\}$ is the reference sequence, and both X_3 and $X_4 = \{x_4, x_5, x_6\}$ become the comparative sequences.

From the perspective of multiple attribute decision making, GRA can be used to evaluate a decision problem with $m+1$ alternative and s attributes. Therefore, s should not be smaller than two. The relationship between X_i on $x_i(k)$ and X_j ($i \leq j \leq i+m+1$) on $x_j(k)$ ($1 \leq k \leq s$) can be derived by the grey relational coefficient (GRC), denoted by $\xi_k(X_j, X_i)$, as follows [24]:

$$\xi_k(X_i, X_j) = \frac{\Delta_{\min} + \rho \Delta_{\max}}{\Delta_{ijk} + \rho \Delta_{\max}}, \quad (1)$$

where ρ ($0 \leq \rho \leq 1$) is the discriminative coefficient and

$$\begin{aligned} \Delta_{\min} &= \min_j \min_l |x_i(l) - x_j(l)|, \quad i \leq j \leq i+m+1, i \leq l \leq s, \\ \Delta_{\max} &= \max_j \max_l |x_i(l) - x_j(l)|, \quad i \leq j \leq i+m+1, i \leq l \leq s, \\ \Delta_{ijk} &= |x_i(k) - x_j(k)|. \end{aligned} \quad (2)$$

It is noted that ρ is often specified as 0.5 [23–25], but this is apparently not an optimal setting.

The overall relationship, $r(X_i, X_j)$, between X_i and X_j can be obtained by aggregating $\xi_k(X_i, X_j)$ as follows:

$$r(X_i, X_j) = \frac{1}{s} \sum_{k=1}^s \xi_k(X_i, X_j). \quad (3)$$

This means that each variable in a sequence is of equal weight. X_i and X_j are similar to dependent and independent variables in traditional regression, and $r(X_i, X_j)$ is analogous to the regression coefficient of X_j to X_i .

3. The Proposed Grey CPD Method

$r(X_i, X_j)$ is the foundation of the proposed grey CPD method. Among the set of variables $\{x_i, x_{i+1}, \dots, x_{i+m}\}$, CPD can be conducted by inspecting the absolute ratio of variety with respect to x_j , $\delta(X_j)$, formulated as:

$$\delta(X_j) = \left| \frac{r(X_i, X_j) - r(X_i, X_{j+1})}{r(X_i, X_j)} \right|, \quad (4)$$

$\delta(X_j)$ also measures the degree of variety of $r(X_i, X_j)$. If x_q is a candidate change point, then $\delta(X_q)$ has the maximum value among $\delta(X_k)$ such that

$$X_q = \arg \max \delta(X_k), \quad (5)$$

and vice versa. x_j can be judged as a change point when $\delta(X_q) \geq \theta$, where θ denotes a nonnegative cut value. The greater the value of θ , the fewer possible change points that can be discovered.

The next possible change point can be detected among $\{x_{q+1}, x_{q+2}, \dots, x_{q+m}\}$, where $m = n - s - q$, since the previous candidate changes point (x_q). In practice, $X_{q+1} = (x_{q+1}(1), x_{q+1}(2), \dots, x_{q+1}(s)) = (x_{q+1}, x_{q+2}, \dots, x_{q+s})$ can serve as the reference sequence, and X_{q+1b} $X_{q+2} = (x_{q+2}(1), x_{q+2}(2), \dots, x_{q+2}(s)) = (x_{q+2}, x_{q+3}, \dots, x_{q+s+1})$, $X_{q+3} = (x_{q+3}(1), x_{q+3}(2), \dots, x_{q+3}(s)) = (x_{q+3}, x_{q+4}, \dots, x_{q+s+2})$, \dots , $X_{q+m+1} = (x_{q+m+1}(1), x_{q+m+1}(2), \dots, x_{q+m+1}(s)) = (x_{q+m+1}, x_{q+m+2}, \dots, x_n)$ serve as comparative sequences. Then, i is set to $q+1$. Given the sequence length s between s_1 and s_2 ($1 < s_1 < s_2$), this CPD process is iteratively performed until $i+s-1 > n$ for each possible value of ρ ranging from 0 to 1. Figure 1 demonstrates the flowchart of the proposed grey CPD method.

4. Empirical Results

4.1. CPD Methods Considered. Since the main applications for GRA are alternative evaluation and clustering, two

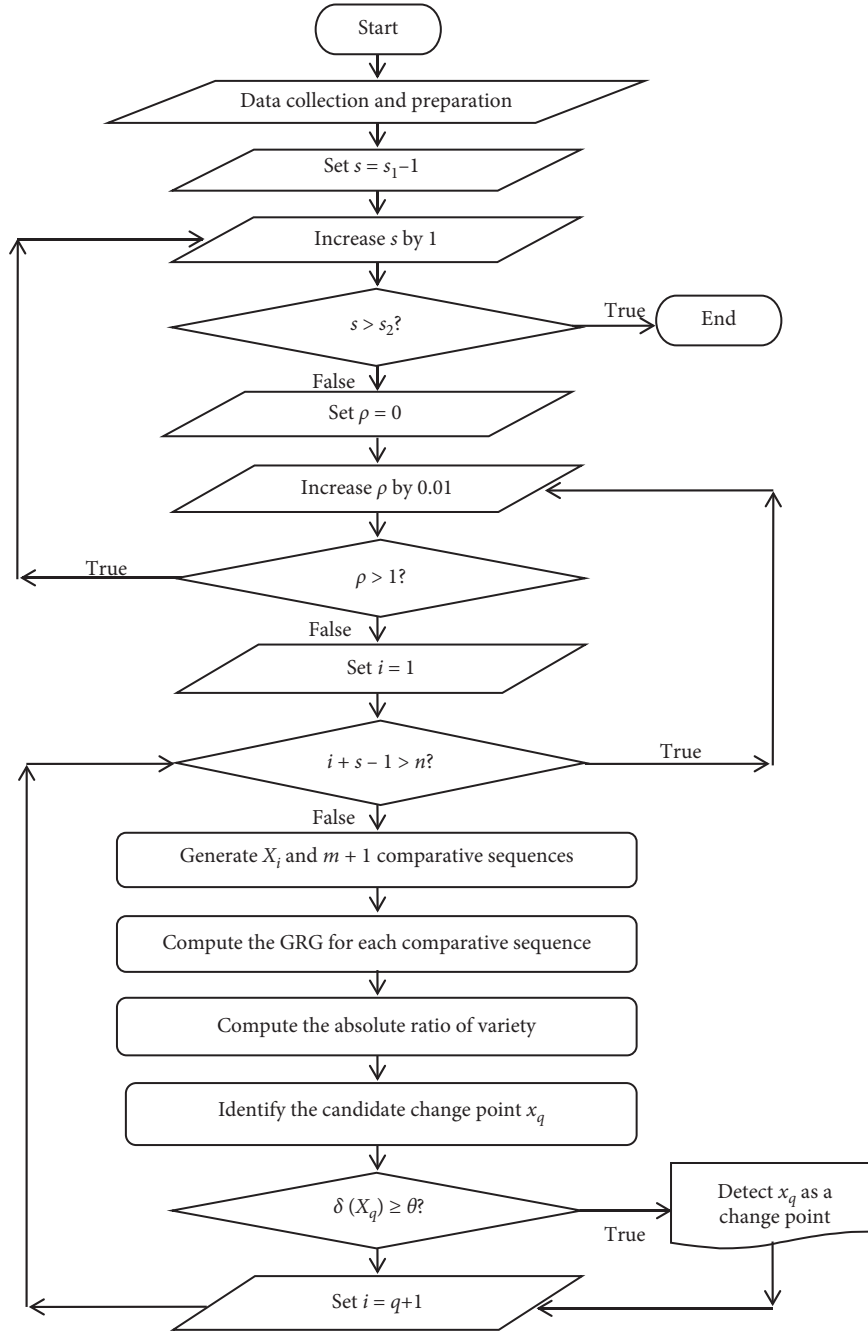


FIGURE 1: Flowchart of the proposed grey CPD method.

unsupervised clustering methods were considered in the empirical analysis with distinct features for CPD methods, namely, Wong et al.'s GRA-based method (WGM) [2], the clustering-based change detector (CBCD) [41], and the piecewise linear function. These methods are briefly described as follows.

4.1.1. WGM. WGM was originally designed to detect change points in hydrological time series. The method considers the reference sequence $X_1 = (x_1, x_2, \dots, x_s)$, and $n - 2s + 1$ comparative sequences, $X_{s+1} = (x_{s+1}, x_{s+2}, \dots, x_{2s})$, $X_{s+2} = (x_{s+2}, x_{s+3}, \dots, x_{2s+1})$, \dots , and $X_{n-s+1} = (x_{n-s+1},$

$x_{n-s+2}, \dots, x_n)$. After computing $r(X_1, X_{s+1})$, $r(X_1, X_{s+2})$, \dots , and $r(X_1, X_{n-s+1})$, the relational degree of X_1 to all comparative sequences is defined as

$$r(s) = \frac{1}{n - 2s + 1} \sum_{k=1}^{n-2s+1} r(X_1, X_k). \quad (6)$$

Subsequently, X_1 is replaced with $(x_1, x_2, \dots, x_{s+1})$, whereas $X_{s+2} = (x_{s+2}, x_{s+3}, \dots, x_{2s+2})$, $X_{s+3} = (x_{s+3}, x_{s+4}, \dots, x_{2s+3})$, \dots , and $X_{n-s} = (x_{n-s}, x_{n-s+2}, \dots, x_n)$ become comparative sequences, and $r(s+1)$ can be thus obtained from $n - 2s - 1$ comparative sequences on average. This process is performed until $r(n/2)$ is obtained. CPD can then

be conducted by inspecting the relative variety ratio of the relational degree

$$\eta(k) = \left| \frac{r(k) - r(k+1)}{r(k)} \right| \times 100, \quad (7)$$

where $k = s, s+1, \dots, n/2-1$. WGM detects x_j as a change point when $\eta(j)$ satisfies

$$j = \arg \max_k \eta(k). \quad (8)$$

It is obvious that only a single change point can be detected in a time series sequence with WGM, regardless of the length of the sequence. Moreover, generating comparative sequences for a time series variable with the proposed method differ from WGM. However, since multiple change points often exist in time series data, WGM was not considered in the empirical study.

4.1.2. CBCD. The CBCD performs CPD by K -means clustering. Initially, a reference window (x_1, x_2, \dots, x_s) is given for which the K clusters are created. The centroid c_p and the radius r_p of the cluster p ($1 \leq p \leq K$) can be computed as follows:

$$c_p = \frac{1}{n_p} \sum_{k=1}^{n_p} x_{p,k}, \quad (9)$$

$$r_p = \sqrt{\frac{1}{n_p} \sum_{k=1}^{n_p} (x_{p,k} - c_p)^2},$$

where n_p denotes the size of the cluster p , and $x_{p,k}$ represents the sample k in the cluster p . Then, a current window $(x_2, x_3, \dots, x_{s+1})$ is generated that replaces x_1 with a new incoming x_{s+1} , and the distance between x_{s+1} and c_p is computed as

$$d(x_{s+1}, c_p) = (x_{s+1} - c_p)^2. \quad (10)$$

If $d(x_{s+1}, c_p) > r_p$, then x_{s+1} is not a member of the cluster p .

As a result, x_{s+1} can be considered a change point when it cannot be categorized into any cluster. At this time, $(x_2, x_3, \dots, x_{s+1})$ becomes the reference window, and K new clusters can be created from the new reference window. Subsequently, a current window catches x_{s+2} to generate $(x_3, x_4, \dots, x_{s+2})$, and x_{s+2} is inspected with regard to whether it is a change point. This process terminates after checking x_n .

4.1.3. Piecewise Linear Function. The piecewise linear function has been used for CPD by finding the joints of the pieces by making the approximation function continuous [42]. Li and Yu [43] proposed piecewise regression analysis that requires users to prespecify the number of change points although this might well be unknown beforehand. Keogh et al. [44] presented several useful CPD methods that do not entail prespecifying the number of change points, among which the bottom-up method seems to perform best. Since the bottom-up method determines a

piece to approximate each state, the property of continuity in the approximation function does not exist. We slightly revise the bottom-up method to conform to the perspective of continuity.

At first, a piece denoted $\text{seg}(x_{i-1}, x_i)$ is generated for (x_{i-1}, x_i) by connecting x_{i-1} and x_i ($2 \leq i \leq n$), and the cost of merging each pair of adjacent pieces is then calculated. For instance, if the cost of merging $\text{seg}(x_a, x_{a+1})$ and $\text{seg}(x_{a+2}, x_{a+3})$ is the lowest, and if it is less than a prespecified merging threshold, then a new piece $\text{seg}(x_a, x_{a+3})$ can be generated by deleting $\text{seg}(x_a, x_{a+1})$ and $\text{seg}(x_{a+2}, x_{a+3})$. The cost of generating $\text{seg}(x_a, x_{a+3})$ is computed by summing all the r between $\text{seg}(x_a, x_{a+3})$ and $x_a, x_{a+1}, x_{a+2}, x_{a+3}$. The cost of merging $\text{seg}(x_{a-2}, x_{a-1})$ and $\text{seg}(x_a, x_{a+3})$ and that of merging $\text{seg}(x_a, x_{a+3})$ and $\text{seg}(x_{a+4}, x_{a+5})$ can be computed. Then, the method iteratively merges the pair with the lowest cost until the cost of merging any pair of adjacent pieces is greater than the threshold.

A new method, called the sliding window and bottom-up (SWAB) method, was developed for online detection [44]. SWAB efficiently produces results that are identical to those of the bottom-up method. We used the SWAB to implement the piecewise linear function, but we omit an introduction to this method for simplicity.

4.2. CPD Performance Evaluation. In order to compare alternative CPD methods, appropriate measures of performance are needed. Since the ratio of change points to total data is small, CPD is typically involved in a learning problem with an unbalanced class distribution. When treating CPD as one kind of pattern classification problem, the G -mean ends with a commonly used indicator of CPD performance [16].

A confusion matrix used to evaluate the performance Table 1 of a CPD method is represented as follows.

G -mean then utilizes both sensitivity and specificity measures to assess the performance, where the sensitivity refers to the ratio of correctly recognized change points, and the specificity refers to the ratio of correctly recognized nonchange points.

$$G\text{-mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}}, \quad (11)$$

where sensitivity and specificity are formulated as

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (12)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}.$$

4.3. Application to S&P100 Stock Returns

4.3.1. Data Collection and Preparation. Empirical studies were conducted using a real dataset to compare the CPD ability of the proposed method to CBCD and SWAB. The task of CPD was performed on log-returns of the daily closing values of stocks consisting of the S&P100 index.

TABLE 1

	True change points	True non-change points
Classified as change points	True positive (TP)	False positive (FP)
Classified as nonchange points	False negative (FN)	True negative (TN)

Change points found in stock returns can better indicate events that characterize the financial market. In terms of 88 series obtained from different institutions, Barigozzi et al. [45] utilized time series factor models to derive multiple primary change points in daily log returns between 4 January 2000 and 10 August 2016. The change points they discovered can be taken as true change points. Furthermore, among these 88 series, two representative series mentioned in [45]—namely, Goldman Sachs (GS) and Bank of America (BAC)—were taken into account. All data are available from Yahoo Finance.

The aim of the empirical study was to examine the G -mean of the CBCD, SWAB, and the proposed grey CPD method by carefully tuning parameter specifications. Because time periods that include too few or no change points are unhelpful for finding parameter specifications, data from 2004–2006 and 2011–2016 were excluded from the GS and BAC series. Therefore, 8 yearly time series remained for GS and BAC with 31 change points (indicated by dates for each series).

Besides, since change points are correlated with events, it is reasonable to analyze the change points and their corresponding events by month, rather than the specific days reported in [45]. For instance, the burst of the dot-com bubble occurred between March 2000 and October 2002. Thus, 31 change points originally indicated by a specific date were reduced to 21 change points indicated by month. For instance, both 4 May 2000 and 10 May 2000 were detected as change points, but we replaced these two days with a single change point: May 2000.

4.3.2. CPD Results. The programs for implementing the proposed grey CPD method were coded in Delphi 7.0 on a personal computer with an Intel Core i3-8100 CPU, Microsoft Windows 10, 8 GB RAM, and a clock rate of 3.60 GHz. Two parameters significantly influenced the performance of the proposed grey CPD method with the eight time segments: the sequence length s and the cut value θ . For the segment k ($1 \leq k \leq 8$), given s between s_1 and s_2 , θ ranging from zero to 6 was carefully tuned so as to maximize the G -mean $_k$, as follows:

$$G - \text{mean}_k = \sqrt{\text{Sensitivity}_k \times \text{Specificity}_k}, \quad (13)$$

where Sensitivity_k and Specificity_k denote the sensitivity and specificity for the segment k , respectively. For instance, G -mean $_2$, Sensitivity_2 , and Specificity_2 are associated with the time segment between January 2001 and December 2001. The maximum G -mean $_2$ value can be found by tuning θ . Finally, the optimal results from all eight segments were used to summarize the overall CPD results. The results of the proposed method on the GS series are summarized in Table 2. G -mean = 0.713 was the best performance obtained by the proposed method.

TABLE 2: Performance evaluation of the proposed method for the GS series.

(s_1, s_2)	Sensitivity	Specificity	G -mean
(3, 5)	0.762	0.667	0.713
(3, 8)	0.810	0.573	0.681
(3, 12)	0.810	0.547	0.665

Table 2 shows that the performance of the proposed method can be improved by choosing appropriate s_1 and s_2 . Therefore, the proposed method was further applied to the BAC series by finding appropriate parameter specifications within the range of s ($3 \leq s \leq 12$). As a result, when $s_1 = 5$ and $s_2 = 12$, the sensitivity and specificity were 0.714 and 0.72, respectively, obtaining the best G -mean (0.717) of the proposed method on the BAC series.

(1) *Comparison with the CBCD.* To improve CPD results, the rule of determining if x_{s+1} is not a member of the cluster p should be revised. In our design, if $d(x_{s+1}, c_p)$ is greater than a user-specified radius threshold rather than r_p , then x_{s+1} is not a member of the cluster p , and vice versa. By contrast, the CBCD is sensitive to the number of clusters K , the sequence length s , and the radius threshold. With radius thresholds of 0.005 and 0.01, the performance values of the CBCD on the GS series are depicted in Figures 2–7, in which the dashed lines denote the performance values of the proposed method. The performance values on the BAC series are depicted in Figures 8–13. Because the best G -mean on the GS series was 0.598 with $K = 4$, $s = 15$, and a radius threshold of 0.01, and 0.556 on the BAC series with $K = 3$, $s = 5$, and a radius threshold of 0.01, the proposed method outperformed the CBCD in terms of the G -mean.

Moreover, a greater radius threshold (0.01) can lead to the discovery of fewer change points with higher specificity. By contrast, a lower radius threshold (0.005) leads to better sensitivity at the expense of specificity. This is the reason why the sensitivity and specificity of the proposed method were considerably superior to those of the CBCD in the cases of higher or lower radius thresholds, respectively.

(2) *Comparison with the SWAB.* The merging threshold significantly influenced the performance of SWAB. The performance of SWAB at different merging thresholds on the GS and the BAC series is depicted in Figures 14 and 15, respectively. The greater the merging threshold, the fewer the number of change points that can be discovered, along with lower sensitivity and higher specificity. The best G -mean on the GS series was 0.690 with a merging threshold of 1.8 and 0.661 on the BAC series with a merging threshold of 1.9. Therefore, in terms of the best G -mean, the proposed method outperformed the SWAB with both series.

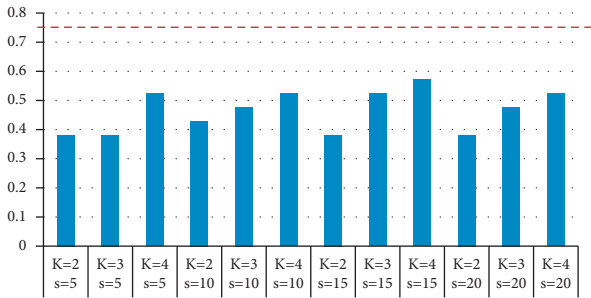


FIGURE 2: Sensitivity of CBCD on GS with a threshold 0.01.

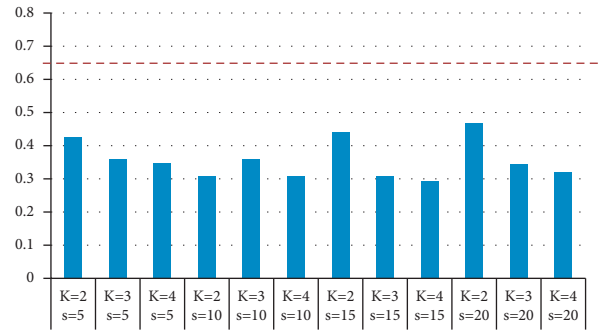


FIGURE 6: Specificity of CBCD on GS with a threshold 0.005.

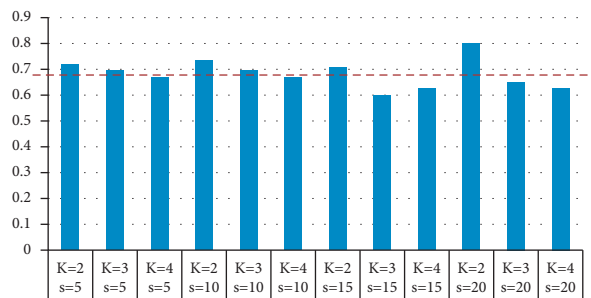


FIGURE 3: Specificity of CBCD on GS with a threshold 0.01.

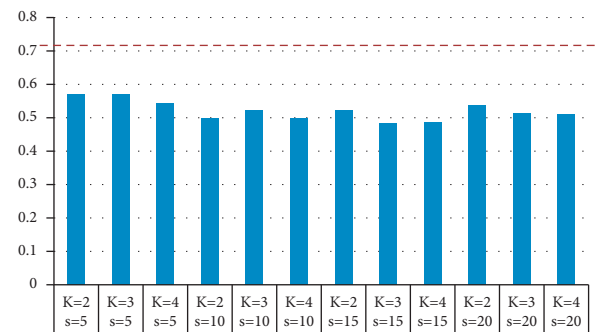


FIGURE 7: G-mean of CBCD on GS with a threshold 0.005.

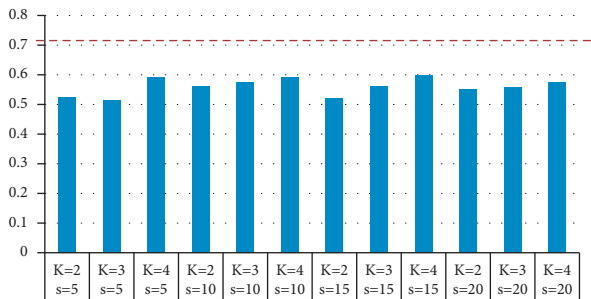


FIGURE 4: G-mean of CBCD on GS with a threshold 0.01.

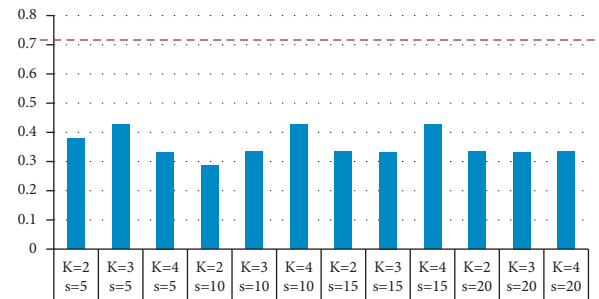


FIGURE 8: Sensitivity of CBCD on BAC with a threshold 0.01.

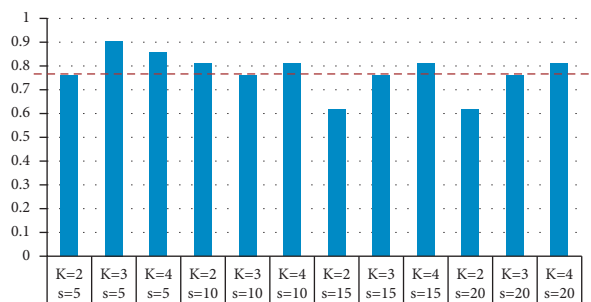


FIGURE 5: Sensitivity of CBCD on GS with a threshold 0.005.

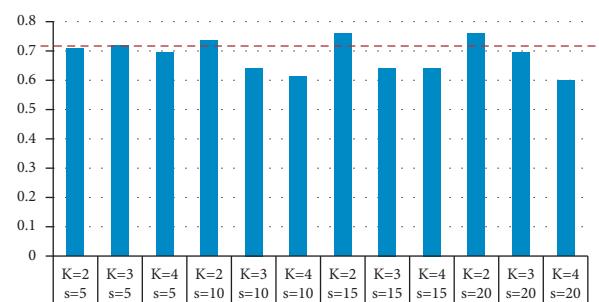


FIGURE 9: Specificity of CBCD on BAC with a threshold 0.01.

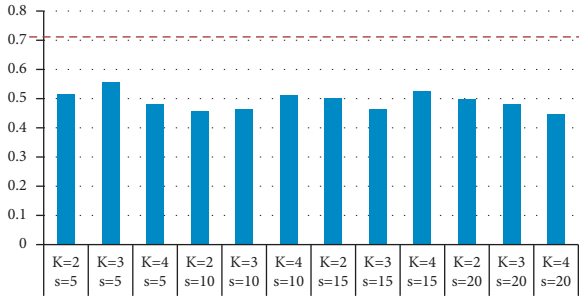


FIGURE 10: G-mean of CBCD on BAC with a threshold 0.01.

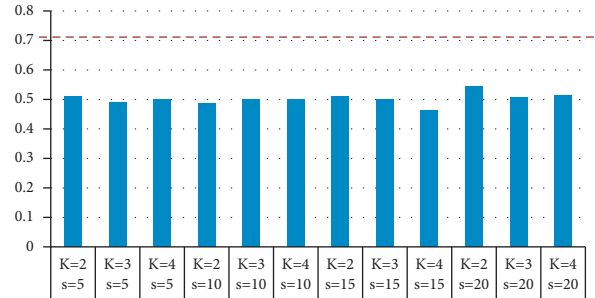


FIGURE 13: G-mean of CBCD on BAC with a threshold 0.005.

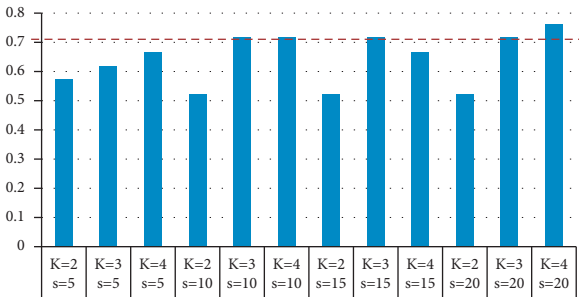


FIGURE 11: Sensitivity of CBCD on BAC with a threshold 0.005.

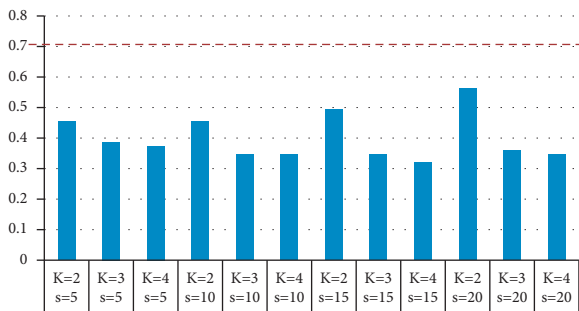


FIGURE 12: Specificity of CBCD on BAC with a threshold 0.005.

5. Discussion and Conclusion

In the financial market, CPD can be applied to discover abnormal volatility in a stock returns series. Using the detected change points, we can find and account for events that characterize this volatility. This can help relevant authorities examine or set up managerial mechanisms to cope with these extraordinary situations. To correctly identify change points, it is crucial to develop an accurate and high-performance CPD method. This paper addresses the case where variables in the collected series cannot be labeled because the change points are unknown. In such a case, traditional regression methods cannot perform CPD. Thus, we proposed an unsupervised regression-like method using GRA. GRA is a kind of multiple attribute decision-making methods, which are capable of effectively evaluating the overall performance of alternatives [46].

There are several advantages to using the proposed grey CPD method. First, there are no constraints on the time series data, such as the stationarity, and the data do not need to be independent and identically distributed. Second, it is not necessary to prespecify the number of change points. Finally, the proposed method is sufficiently simple to implement as a program without any statistical assumptions. Empirical results on two representative series of stock returns were encouraging in terms of the performance obtained using the proposed grey CPD method. This demonstrates that using GRA to measure the relationship between response and explanatory time series variables can boost the performance of the proposed method. It should be noted that both the CBCD and SWAB could obtain better sensitivity than the proposed method, but only at the expense of specificity for instance.

- (1) For the GS series, the best sensitivity of the CBCD was 0.905 with $K = 3, s = 5$, and a radius threshold of 0.005. That of the SWAB was 1 with a merging threshold of 0.5. The specificity of the former and the latter was 0.36 and 0.053, respectively.
- (2) For the BAC series, the best sensitivity of the CBCD was 0.762 with $K = 4, s = 20$, and a radius threshold of 0.005. That of the SWAB was 0.857 with a merging threshold of 0.5. The specificity of the former and the latter was 0.347 and 0.147, respectively.

This is the reason why the G-mean, which combines sensitivity and specificity, is commonly used to measure the performance of CPD methods from the perspective of classification. As such, it is not possible to conclude whether any classification method was “best” insofar as there is no such thing as a best classifier [47].

This study motivated us to explore further studies. First, the proposed grey CPD method can be further applied to other real-world problems. For instance, it can be used to discover abnormal trades in the stock market. Detecting anomalies in the volatility of stock prices can provide regulators with useful information about investments and prevent crime. We will explore this application in future research. Furthermore, the GRG was implemented using a weighted-average method, where noninteraction was assumed among the attributes involved. Nevertheless, the assumption of additivity may not be realistic with many applications [48]. Thus, our future work will explore the

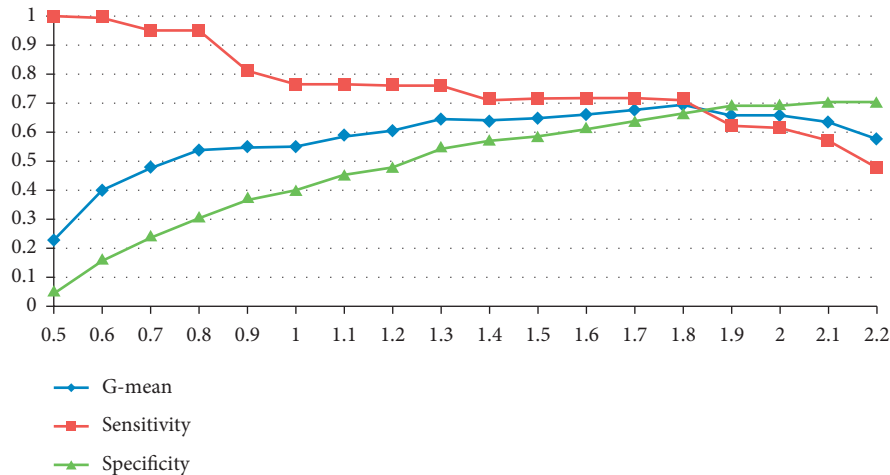


FIGURE 14: Performance values of SWAB on GS with different merging thresholds.

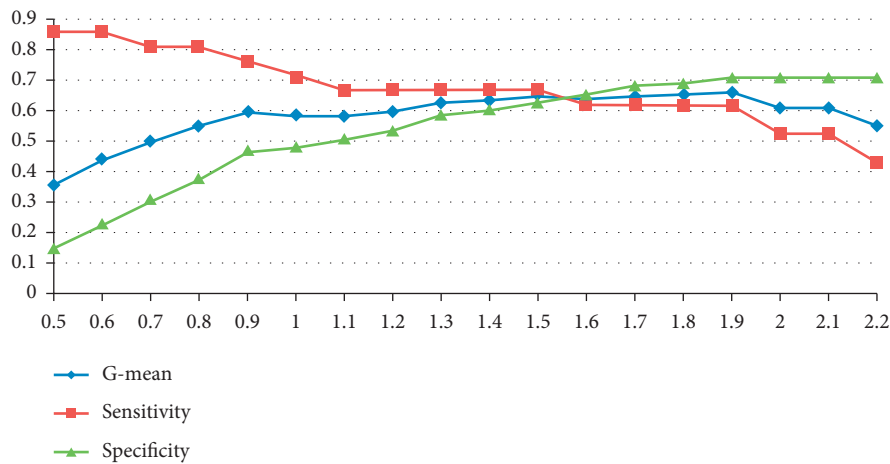


FIGURE 15: Performance values of SWAB on BAC with different merging thresholds.

development of a nonadditive grey CPD method using nonadditive GRG with fuzzy integral [39, 49] and check the resultant impact on performance. It should be noted that the fuzzy integral has proven to be effective in dealing with performance evaluation and preferential dependence among attributes [50–52].

Data Availability

The authors declare that they have no conflict of interest. This article does not contain any studies with human participants performed by the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was financially supported by the Ministry of Science and Technology, Taiwan, under grant MOST 110-2410-H-033-013-MY2.

References

- [1] S. Aminikhanghahi and D. J. Cook, “A survey of methods for time series change point detection,” *Knowledge and Information Systems*, vol. 51, pp. 339–367, 2017.
- [2] H. Wong, B. Q. Hu, W. C. Ip, and J. Xia, “Change-point analysis of hydrological time series using grey relational method,” *Journal of Hydrology*, vol. 324, no. (1-4), pp. 323–338, 2000.
- [3] H. Chen and N. Zhang, “Graph-based change-point detection,” *Annals of Statistics*, vol. 43, no. 1, pp. 139–176, 2015.
- [4] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, “A review and comparison of changepoint detection techniques for climate data,” *Journal of Applied Meteorology and Climatology*, vol. 46, no. 6, pp. 900–915, 2007.
- [5] H. Alexandersson, “A homogeneity test applied to precipitation data,” *Journal of Climatology*, vol. 6, pp. 661–675, 1986.
- [6] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, “Image change detection algorithms: a systematic survey,” *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 294–307, 2005.
- [7] L. Perreault, J. Bernier, B. Bobée, and E. Parent, “Bayesian change-point analysis in hydrometeorological time series.

- Part 1 The normal model revisited,” *Journal of Hydrology*, vol. 235, no. 3-4, pp. 221–241, 2000.
- [8] L. Perreault, J. Bernier, B. Bobée, and E. Parent, “Bayesian change-point analysis in hydrometeorological time series. Part 2 Comparison of change-point models and forecasting,” *Journal of Hydrology*, vol. 235, no. 3-4, pp. 242–263, 2000.
- [9] M. Staudacher, S. Telser, A. Amann, H. Hinterhuber, and M. Ritsch-Martel, “A new method for change-point detection developed for on-line analysis of the heart beat variability during sleep,” *Physica A: Statistical Mechanics and its Applications*, vol. 349, no. 3–4, pp. 582–596, 2005.
- [10] P. Yang, G. Dumont, and J. M. Ansermino, “Adaptive change detection in heart rate trend monitoring in anesthetized children,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2211–2219, 2006.
- [11] C. K. Chan, S. F. Witt, Y. C. E. Lee, and H. Song, “Tourism forecast combination using the CUSUM technique,” *Tourism Management*, vol. 31, no. 6, pp. 891–897, 2010.
- [12] Y. Sun, J. Zhang, X. Li, and S. Wang, “Forecasting tourism demand with a new time-varying forecast averaging approach,” *Journal of Travel Research*, Article ID 004728752110612, 2021.
- [13] D. M. Hawkins, “Testing a sequence of observations for a shift in location,” *Journal of the American Statistical Association*, vol. 72, no. 357, pp. 180–186, 1977.
- [14] K. D. Feuz, D. J. Cook, C. Rosasco, K. Robertson, and M. Schmitter-Edgecombe, “Automated detection of activity transitions for prompting,” *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 5, pp. 575–585, 2015.
- [15] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, “Using mobile phones to determine transportation modes,” *ACM Transactions on Sensor Networks*, vol. 6, no. 2, pp. 1–27, 2010.
- [16] Y. Zheng, L. Liu, L. Wang, and X. Xie, “Learning transportation mode from raw gps data for geographic applications on the web,” in *Proceedings of the 17th international conference on World Wide Web*, pp. 247–256, Beijing, 2008.
- [17] V. S. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” *Energy Policy*, vol. 35, no. 3, pp. 1701–1708, 2007.
- [18] D. C. Li, C. J. Chang, C. C. Chen, and W. C. Chen, “Forecasting short-term electricity consumption using the adaptive grey-based approach—An Asian case,” *Omega*, vol. 40, no. 6, pp. 767–773, 2012.
- [19] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, Berlin, 2011.
- [20] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, “Change-point detection in time-series data by relative density-ratio estimation,” *Neural Networks*, vol. 43, pp. 72–83, 2013.
- [21] E. Keogh and J. Lin, “Clustering of time-series subsequences is meaningless: implications for previous and future research,” *Knowledge and Information Systems*, vol. 8, no. 2, pp. 154–177, 2005.
- [22] H. Song, R. T. R. Qiu, and J. Park, “A review of research on tourism demand forecasting: launching the Annals of Tourism Research Curated Collection on tourism demand forecasting,” *Annals of Tourism Research*, vol. 75, pp. 338–362, 2019.
- [23] D. Ju-Long, “Control problems of grey systems,” *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.
- [24] S. Liu and Y. Lin, *Grey Information: Theory and Practical Applications*, Springer-Verlag, London, 2006.
- [25] S. Liu, Y. Yang, and J. Forrest, *Grey Data Analysis: Methods, Models and Applications*, Springer, Berlin, 2017.
- [26] Y. C. Hu, “A multivariate grey prediction model with grey relational analysis for bankruptcy prediction problems,” *Soft Computing*, vol. 24, no. 6, pp. 4259–4268, 2020.
- [27] Y. C. Hu and G. Wu, “The impact of Google Trends index and encompassing tests on forecast combinations in tourism,” *Tourism Review*, vol. 77, no. 5, pp. 1276–1298, 2022.
- [28] Y. C. Hu, R. S. Chen, Y. T. Hsu, and G. H. Tzeng, “Grey self-organizing feature maps,” *Neurocomputing*, vol. 48, no. 1–4, pp. 863–877, 2002.
- [29] Y. Huang, L. Shen, and H. Liu, “Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China,” *Journal of Cleaner Production*, vol. 209, pp. 415–423, 2019.
- [30] S. J. Huang, N. H. Chiu, and L. W. Chen, “Integration of the grey relational analysis with genetic algorithm for software effort estimation,” *European Journal of Operational Research*, vol. 188, no. 3, pp. 898–909, 2008.
- [31] P. Wang, P. Meng, J. Y. Zhai, and Z. Q. Zhu, “A hybrid method using experiment design and grey relational analysis for multiple criteria decision making problems,” *Knowledge-Based Systems*, vol. 53, pp. 100–107, 2013.
- [32] H. Jiang and W. He, “Grey relational grade in local support vector regression for financial time series prediction,” *Expert Systems with Applications*, vol. 39, no. 3, pp. 2256–2262, 2012.
- [33] Y. C. Hu, “Pattern classification using grey tolerance rough sets,” *Kybernetes*, vol. 45, no. 2, pp. 266–281, 2016.
- [34] Y. C. Hu, Y. J. Chiu, Y. L. Liao, and Q. Li, “A fuzzy similarity measure for collaborative filtering using nonadditive grey relational analysis,” *Journal of Grey System*, vol. 27, no. 2, pp. 93–103, 2015.
- [35] P. Jiang, Y. C. Hu, G. F. Yen, and S. J. Tsao, “Green supplier selection for sustainable development of the automotive industry using grey decision making,” *Sustainable Development*, vol. 26, pp. 890–903, 2018.
- [36] Y. C. Hu, “Constructing grey prediction models using grey relational analysis and neural networks for magnesium material demand forecasting,” *Applied Soft Computing*, vol. 93, Article ID 106398, 2020.
- [37] Y. C. Hu, “Grey prediction with residual modification using functional-link net and its application to energy demand forecasting,” *Kybernetes*, vol. 46, no. 2, pp. 349–363, 2017.
- [38] Y. C. Hu, “Grey relational analysis and radial basis function network for determining costs in learning sequences,” *Applied Mathematics and Computation*, vol. 184, no. 2, pp. 291–299, 2007.
- [39] Y. C. Hu, Y. J. Chiu, and J. F. Tsai, “Establishing grey criteria similarity measures for multi-criteria recommender systems,” *Journal of Grey System*, vol. 30, no. 1, pp. 192–205, 2018.
- [40] P. Jiang, W. B. Wang, Y. C. Hu, Y. J. Chiu, and S. J. Tsao, “Pattern classification using tolerance rough sets based on non-additive grey relational analysis and DEMATEL,” *Grey Systems: Theory and Application*, vol. 11, no. 1, pp. 166–182, 2020.
- [41] D. H. Tran, “Automated change detection and reactive clustering in multivariate streaming data,” 2013, <https://arxiv.org/abs/1311.0505>.
- [42] J. R. Yu, G. H. Tzeng, and H. L. Li, “Interval piecewise regression model with automatic change-point detection by quadratic programming,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 13, no. 03, pp. 347–361, 2005.
- [43] H. L. Li and J. R. Yu, “A piecewise regression analysis with automatic change-point detection,” *Intelligent Data Analysis*, vol. 3, no. 1, pp. 75–85, 1999.

- [44] E. Keogh, S. Chu, D. Hart, and M. Pazzani, "An online algorithm for segmenting time series," in *Proceedings of the 2001 IEEE International Conference on Data Mining*, pp. 289–296, San Jose, CA, 2001.
- [45] M. Barigozzi, H. Cho, and P. Fryzlewicz, "Simultaneous multiple change-point and factor analysis for high-dimensional time series," *Journal of Econometrics*, vol. 206, no. 1, pp. 187–225, 2018.
- [46] Y. C. Hu, J. S. Hu, R. S. Chen, and G. H. Tzeng, "Assessing weights of product attributes from fuzzy knowledge in a dynamic environment," *European Journal of Operational Research*, vol. 154, no. 1, pp. 125–143, 2004.
- [47] L. I. Kuncheva, *Fuzzy Classifier Design*, Physica-Verlag, Heidelberg, 2000.
- [48] Z. Wang, K. S. Leung, and G. J. Klir, "Applying fuzzy measures and nonlinear integrals in data mining," *Fuzzy Sets and Systems*, vol. 156, no. 3, pp. 371–380, 2005.
- [49] Y. C. Hu, "Nonadditive similarity-based single-layer perceptron for multi-criteria collaborative filtering," *Neurocomputing*, vol. 129, pp. 306–314, 2014.
- [50] G. H. Tzeng and K. Y. Shen, *New Concepts and Trends of Hybrid Multiple Criteria Decision Making*, CRC Press, New York, NY, USA, 2017.
- [51] Y. C. Hu and H. C. Chen, "Choquet integral-based hierarchical networks for evaluating customer service perceptions on fast food stores," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7880–7887, 2010.
- [52] Y. C. Hu and W. B. Wang, "Nonlinear interval regression analysis with neural networks and grey prediction for energy demand forecasting," *Soft Computing*, vol. 26, no. 14, pp. 6529–6545, 2022.