Use of Partial Cumulative Sum to Detect Trends and Change Periods for Nonlinear Time Series

Berlin Wu and Liyang Chen[*]

Because the structural change of a time series from one pattern to another may not switch at once but rather experience a period of adjustment, conventional change point detection may be inappropriate under some circumstances. Furthermore, changes in time series often occur gradually so that there is a certain amount of fuzziness in the change point. For this, considerable research has focused on the theory of change period detection for improved model performance. However, a change period in some small time interval may appear to be negligible noise in a larger time interval. In this paper, we propose an approach to detect trends and change periods with fuzzy statistics using partial cumulative sums. By controlling the parameters, we can filter the noises and discover suitable change periods. Having discovered the change periods, we can proceed to identify the trends in the time series. We use simulations to test our approach. Our results show that the performance of our approach is satisfactory.

Keywords: fuzzy time series, change periods, partial cumulative sums, trend, noise
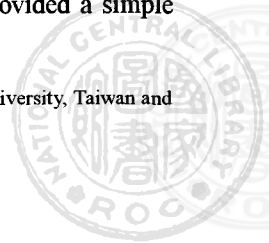**JEL classification:** C32, C42

1   Introduction

An interesting topic in time series analysis is the detection of trends and the measurement of change points. An extensive literature has been proposed to find change points. For example, Chow (1960), Nyblom (1989), Ploberger *et al.* (1989), Bleaney (1990), Lin and Terasvirta (1994), and others have proposed various methods to detect change points. Broemeling and Tsurumi (1987) used a Bayesian procedure to solve inferential problems of structural shift. They provided a simple

[*]Authors are respectively at: Department of Mathematics, National Chengchi University, Taiwan and Department of Mathematics, National Chengchi University, Taiwan.
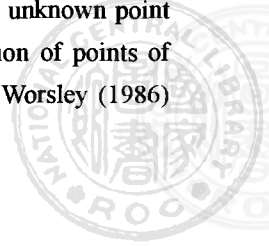
way to analyze data and did not rely heavily on asymptotic distribution theory in making statistical inference. Tsay (1988) proposed a procedure to detect outliers, level shifts, and variance changes in a univariate time series. Balke (1993) pointed out that Tsay's procedures do not always perform satisfactorily when level shifts are present. Barry and Hartigan (1993) also presented a Bayesian analysis for change point problems.

Before one attempts to detect change points, several fundamental questions arise. What does a change point mean? Can we give a clear definition of a change point? How do we determine change points if the economic structure for a time series changes gradually? How do we smooth or get rid of an unstable and uncertain intervention in a time series? How do we address poorly defined economic keywords? These problems involving semantic interpretation and fuzzy statistical analysis have bothered many researchers for a long time. For this reason, Zadeh (1965) proposed fuzzy set theory, a new tool to generalize the classical notion of a set and to accommodate semantic and conceptual fuzziness in statements. Fuzzy theory has the intrinsic property of linguistic variables. This property can help us to reduce the difficulties of uncertain problems. Fuzzy theory is widely used in various areas.

In this paper we use fuzzy logic to deal with change periods and trend problems in time series analysis. These interesting problems have been investigated by many researchers. Custem and Gath (1993) suggested a useful approach based on fuzzy clustering for the detection of outliers and for the robust estimation of underlying parameters. Hathaway and Bezdek (1993) established fuzzy c-regression models as a promising technique for switching regression parameter estimation and clustering. Yoshinari *et al.* (1993) developed a new method to build fuzzy models through clustering methods based on linear varieties. Inclan and Tiao (1994) proposed an iterative procedure to detect variance changes based on a centered versions of the cumulative sums of squares. Wu and Chen (1999) suggested an algorithm for fuzzy time series classification.

Hinkley (1971) proposed the modified page and the cumulative sum methods. Hsu (1979, 1982) investigated the detection of a variance shift at an unknown point in a sequence of independent observations, focusing on the detection of points of change one at a time because of the heavy computational burden. Worsley (1986)

used maximum likelihood methods to test for a change in the mean of a sequence of independent exponential family random variables. Sastri *et al.* (1989) presented a performance comparison for six time series change detection procedures.

However, these detection techniques are based on the assumption that the underlying time series exhibits a significant change point characteristic (Wu and Chen, 1999). Using the concept of fuzzy set theory, Wu and Chen (1999) proposed a procedure for change period detection for nonlinear time series. Nevertheless, in dealing with time series with switching regimes, we should consider not only change point detection but also the properties of change periods. Because many patterns of changing structure in time series occur over a time interval, these phenomena should not be treated as a mere sudden turn at a certain time.

Another problem is that a change period in a time series over a certain time interval may seem like a noise in a larger time interval. In our research, we propose a procedure based on fuzzy logic to detect change periods. Our approach enables us to filter the noise and to locate change periods by controlling parameters. Moreover, we don't require any initial knowledge about the structure in the data to apply this method.

This paper is organized as follows. In Section 2, we introduce the basic concept of fuzzy logic and introduce our approach with examples. In section 3, simulations illustrate our method. Empirical examples of three foreign exchange rates are studied in Section 4. Section 5 provides the conclusion and suggestions.

2   Detection of trends and change periods

2.1   Fuzzy time series

A time series is a set of observed values recorded over time. These observed values could be either continuously observed, called a continuous time series, or observed at discrete time points, called a discrete time series. A time series is usually denoted $\{X_t\}$, where $X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ refer to the observed values at times $t_1, t_2, \cdots, t_n$.

Time series analysis plays a very important role in forecasting and is very successful in many applications. Each observation is taken to be a single, precise value in traditional time series analysis. However, the measurement error of

collecting data, the time lag in observation, or the interaction between variables may turn the single value into a range of possible values. For example, when we talk about the stock index of a day, which value do we specify, the index at beginning of the day, the end of the day, the high point, or the low point?

Conventional time series analysis is based on the concept that the observed data are random with certain measurement errors or noise. However, in empirical studies we often encounter the situation that the data reveal not only randomness but also fuzziness. In this case, the application of fuzzy time series leads to improved inference. We begin with a definition of a fuzzy time series.

Definition 2.1 A fuzzy time series

Let $\{X_t \in \Re; t = 1, 2, \cdots\}$ be a time series and U be the universe of discourse. Let $\{P_i; i = 1, 2, \cdots, m\}$ be an ordered partition of U on which linguistic variables $\{L_i; i = 1, 2, \cdots, m\}$ are defined. For each X_t, the corresponding fuzzy set on U, $F(X_t)$, consists of membership functions $\{\mu_{t1}, \mu_{t2}, \cdots, \mu_{tm}\}$ for $\{L_1, L_2, \cdots, L_m\}$. Then $\{F(X_t)\}$ is a fuzzy time series corresponding to $\{X_t\}$ and is denoted:

$$F(X_t) = \frac{\mu_{t1}(X_t)}{L_1} + \frac{\mu_{t2}(X_t)}{L_2} + \cdots + \frac{\mu_{tm}(X_t)}{L_m} \tag{1}$$
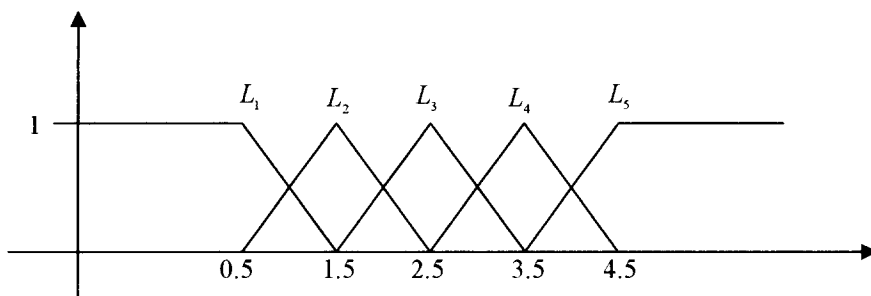
where the addition symbol denotes the connection and $\mu_{ti}(X_t)/L_i$ specifies the corresponding relation of the membership function $\mu_{ti}(X_t)$ of X_t with respect to L_i satisfying $\mu_{ti} : \Re \to [0,1]$ and $\sum_{i=1}^{m} \mu_{ti} = 1$.

For simplicity, we write $F(X_t) = (\mu_{t1}, \mu_{t2}, \cdots, \mu_{tm})$ instead of (1) and take the triangular or trapezoidal membership function in this paper. The set $\{L_i; i = 1, 2, \cdots, m\}$ is regarded as a sequence of linguistic variables, and the fuzzy time series $\{F(X_t); t = 1, 2, \cdots, n\}$ consists of their memberships. That is, any $F(X_t)$ ($t = 1, 2, \cdots, n$) contains the memberships corresponding to each L_i.

Example 2.1

Let $\{X_t(t)\} = \{0.8, 1.7, 2.9, 4.1, 3.5, 3.2, 4.3, 3.6\}$ and $U = \{[0,1], [1,2], [3,4], [4,5]\}$. Define the linguistic variables to be $L_1 =$ very low, $L_2 =$ low, $L_3 =$ middle, $L_4 =$ high, and $L_5 =$ very high. Moreover, we take the average number of the intervals as our typical values. The typical values corresponding to $L_1, L_2 ..., L_5$ are now

defined as 0.5, 1.5, 2.5, 3.5, 4.5. Figure 1 shows the membership functions of these linguistic variables.

Figure 1: The membership functions for Example 2.1

Thus, we have the fuzzy time series $\{F(X_t)\}$ corresponding to $\{X_t\}$ as shown in Table 1.

Table 1: The fuzzy time series for Example 2.1

	very low	low	middle	high	very high
$F(X_1)$	0.7	0.3	0	0	0
$F(X_2)$	0	0.8	0.2	0	0
$F(X_3)$	0	0	0.6	0.4	0
$F(X_4)$	0	0	0	0.4	0.6
$F(X_5)$	0	0	0	1	0
$F(X_6)$	0	0	0.2	0.8	0
$F(X_7)$	0	0	0	0.2	0.8
$F(X_8)$	0	0	0	0.9	0.1

2.2   Detection of change period

Because the structural change of a time series from one pattern to another may not switch at once but rather experience a period of adjustment, it is natural for us to apply the concept of change period instead of change points when analyzing a structural change process. Taking a view different from change points, the concept of a change period provides us with a more reasonable, more comprehensible, and more flexible way to analyze real world problems.

In order to present an approach to locate the change periods in a fuzzy time

series, the following definitions are required.

Definition 2.2 An indicator series

Let $F(X_t) = (\mu_{t1}, \mu_{t2}, \cdots, \mu_{tm})$ be a fuzzy time series and $a_1, a_2, \cdots, a_m$ ($a_i \in \Re, i = 1, 2, \cdots, m$) be the weights with respect to the linguistic variables $L_1, L_2, \cdots, L_m$. Then the series $IF(t) = a_1 \mu_{t1}(X_t) + a_2 \mu_{t2}(X_t) + \cdots + a_m \mu_{tm}(X_t)$ is called the indicator series for the fuzzy time series.

Example 2.2

In Example 2.1, we transformed the time series $\{X_t\}$ into a fuzzy time series $\{F(X_t)\}$. Now letting the fuzzy weights of $\{F(X_t)\}$ be $a_1 = -2$, $a_2 = -1$, $a_3 = 0$, $a_4 = 1$, and $a_5 = 2$, we have:

$$IF(1) = (-2) \times 0.7 + (-1) \times 0.3 + 0 \times 0 + 1 \times 0 + 2 \times 0 = -1.7$$
$$IF(2) = (-2) \times 0 + (-1) \times 0.8 + 0 \times 0.2 + 1 \times 0 + 2 \times 0 = -0.8$$
$$IF(3) = (-2) \times 0 + (-1) \times 0 + 0 \times 0.6 + 1 \times 0.4 + 2 \times 0 = 0.4$$
$$IF(4) = (-2) \times 0 + (-1) \times 0 + 0 \times 0 + 1 \times 0.4 + 2 \times 0.6 = 1.6$$
$$IF(5) = (-2) \times 0 + (-1) \times 0 + 0 \times 0 + 1 \times 1 + 2 \times 0 = 1$$
$$IF(6) = (-2) \times 0 + (-1) \times 0 + 0 \times 0.2 + 1 \times 0.8 + 2 \times 0 = 0.8$$
$$IF(7) = (-2) \times 0 + (-1) \times 0 + 0 \times 0 + 1 \times 0.2 + 2 \times 0.8 = 1.8$$
$$IF(8) = (-2) \times 0 + (-1) \times 0 + 0 \times 0 + 1 \times 0.9 + 2 \times 0.1 = 1.1$$

Definition 2.3 A detection sequence for change periods

Let $\{IF(t); t = 1, 2, \ldots, n\}$ be a fuzzy trend indicator series and define $d_k(t) = \sum_{t'=t}^{t+k} IF(t')$. We call $d_k(t)$ a change period detection sequence of degree 3.

Example 2.3

Let $\{IF(t); t = 1, 2, \ldots, 8\}$ be the same as that in Example 2.2. The change period detection sequence of degree 3 is constructed as follows:

$$d(3) = (-1.7) + (-0.8) + 0.4 = -2.1$$
$$d(4) = (-0.8) + 0.4 + 1.6 = 1.2$$
$$d(5) = 0.4 + 1.6 + 1 = 2.1$$
$$d(6) = 1.6 + 1 + 0.8 = 3.4$$
$$d(7) = 1 + 0.8 + 1.8 = 3.6$$

$$d(8) = 0.8 + 1.8 + 1.1 = 3.7$$

Note that the change period detection sequence of degree n is the sequence of the partial cumulative sums of n consecutive elements in the fuzzy trend indicator series. The choice of the degree of a change period detection sequence and the fuzzy weights of a fuzzy time series determine what kind of change periods can be found.

The first step is to take first differences. If $\{X_t\}$ is the time series, we obtain its first difference time series $\{Y_t\}$ by taking $Y_t = X_t - X_{t-1}$. Intuitively, Y_t is the change in the series at time t relative to time $t-1$. That is to say, $Y_t > 0$ indicates that X_{t-1} rises to X_t at time t by the amount Y_t. Conversely, $Y_t < 0$ indicates that X_{t-1} falls to X_t at time t by the amount Y_t. Obviously, $Y_t = 0$ means that there is no change in the original series between times $t-1$ and t.

After forming the first difference time series $\{Y_t\}$, the second step is to transform $\{Y_t\}$ into the corresponding fuzzy time series $F(Y_t)$ with L_i as its linguistic level, $i = 1, 2, \cdots, m$. If a time series $\{X_t\}$ reflects the linguistic fuzziness, then so does its first difference $\{Y_t\}$.

The third step is to construct the fuzzy trend indicator series $FI(t) = a_1 \mu_{t1}(Y_t) + a_2 \mu_{t2}(Y_t) + \cdots + a_m \mu_{tm}(Y_t)$ of the fuzzy time series $F(Y_t)$. The rule to decide the fuzzy weights of the fuzzy time series is that if a_i is the fuzzy weight of some linguistic variable L_i where the negative Y_t makes the membership function $\mu_{ti}(Y_t)$ greater than zero, then we assign a_i a nonpositive value. Conversely, if a_i is the fuzzy weight of some linguistic variable L_i where the positive Y_t makes the membership function $\mu_{ti}(Y_t)$ greater than zero, then we assign a_i a nonnegative value.

The next step is to compute the change period detection sequence from this fuzzy trend indicator series. We are going to detect trends and change periods in a time series with this detection sequence. The degree of the sequence plays a very important role in controlling what types of change periods can be found.

As noted above, sometimes an obvious change in trend or a change period in a small time interval may appear to be noise caused by randomness and fuzziness in a larger time interval. The observation range heavily influences our recognition of a change period. Furthermore, even under the same observation scope, two researchers may disagree about the existence of a change period. Therefore, an essential question arises: Can we always find the appropriate type of change periods

that are meaningful in a given application?

In this paper, we propose a method to detect empirically relevant types of change periods by controlling the degree of the change period detection sequence and the fuzzy weights. Intuitively, the greater n, the greater change period one is able to find. Next, we formalize the definition of a change period.

Definition 2.4 A change period

Suppose that $\{d(t)\}_{t=n}^{\infty}$ is a change period detection sequence of degree n. For a given $h > 0$, if there is a time interval $T_1 = \{t_l, t_{l+1}, \cdots, t_{l+m}\}$ such that $-h < d(t) < h$ for all $t \in T_1$ and if an immediately preceding time interval $T_2 = \{t_{l-j}, t_{l-(j-1)}, \cdots, t_{l-1}\}$ and an immediately subsequent time interval $T_3 = \{t_{l+m+1}, t_{l+m+2}, \cdots, t_{l+m+k}\}$ are such that the signs of $d(t)$ are all the same for all $t \in T_2$ (T_3) but are opposite to the signs of $d(t)$ for all $t \in T_3$(T_2), then we call $T = \{t_{[l-(n+1)/2]}, t_{[l-(n+1)/2]+1}, t_{[l-(n+1)/2]+2}, \cdots, t_{[l-(n+1)/2]+m}\}$ a change period, where $[x]$ denotes the ceiling function.

Example 2.4

Suppose that $d(t)$ is the change period detection sequence of degree 10 shown in Table 2.

Table 2: The change period detection sequence for Example 2.4

t	$d(t)$	t	$d(t)$	t	$d(t)$	t	$d(t)$	t	$d(t)$
1	6.768	6	8.822	11	2.085	16	−4.230	21	−6.017
2	8.779	7	5.719	12	−0.467	17	−4.272	22	−6.944
3	7.807	8	6.760	13	0.903	18	−9.525	23	−9.651
4	5.699	9	6.998	14	0.557	19	−8.633	24	−10.074
5	9.157	10	3.923	15	−4.167	20	−9.614	25	−7.815

Letting $h = 4$, we can see that at times 10, 11, 12, 13, and 14, we observe $-h < d(t) < h$. We find that for the immediately preceding time interval $T_2 = \{1, 2, \ldots, 9\}$ $d(t)$ is positive for all $t \in T_2$ and that for the immediately subsequent time interval $T_3 = \{10, 11, \ldots, 25\}$ $d(t)$ is negative for all $t \in T_3$. Thus, the change period is $T = \{10 - [(10+1)/2] = 4, 5, 6, 7, 8\}$.

Here is a summary of the steps involved in detecting change periods in a time series $\{X_t\}$:

Step 1.   Calculate the first difference time series $\{Y_t\}$, where $Y_t = X_t - X_{t-1}$.

Step 2.   Transform the difference time series $\{Y_t\}$ into a fuzzy time series $\{F(Y_t)\}$ with L_i as its linguistic level, $i = 1,2,\cdots,m$.

Step 3.   Choose the weights of this fuzzy time series $\{F(Y_t)\}$ and calculate the corresponding fuzzy trend indicator series $IF(t)$.

Step 4.   Construct the change period detection sequence of degree n $\{d(t)\}$ from the fuzzy trend indicator series $IF(t)$.

Step 5.   Observe the change period detection sequence $\{d(t)\}$ and search for a time intervals that satisfy the conditions in Definition 2.4. These time intervals are the change periods.

2.3   Detecting trends

Detecting trends is of great importance in many applications of both practical and theoretical areas. The ability to recognize the beginnings and the ends of trends helps us to make correct decisions and take appropriate actions. In this paper, we propose an approach to detect trends using change period detection sequences. Before introducing this method, we define a trend in a time series.

Definition 2.5 A trend

Suppose that $\{d(t)\}_{t=n}^{\infty}$ is a change period detection sequence of degree n , if we can find an time interval $T = \{t_m, t_{m+1}, \cdots, t_{m+n}\}$ at which $d(t)$ is positive or zero for every $t \in T$ and some $m, n \in N$, then $\{X(t_m), X(t_{m+1}), \cdots, X(t_{m+n})\}$ (henceforth abbreviated $X(T)$) is called an upward trend. Conversely, if we can find an time interval $T = \{t_m, t_{m+1}, \cdots, t_{m+n}\}$ at which $d(t)$ is negative or zero for every $t \in T$ and some $m, n \in N$, then $X(T)$ is called a downward trend.

Property 2.1 A trend must occur between two change periods.

Proof: Suppose that $\{d(t)\}_{t=n}^{\infty}$ is a change period detection sequence of degree n and T_1 is a change period. Then, by Definition 2.4, there is an immediately preceding time interval T_2 at which $d(t)$ takes the same sign for every $t \in T_2$.

Without loss of generality, we assume that $d(t)$ is positive on T_2 and thus T_2 is an upward trend. There are only three possibilities for the sign of the next observation: positive, negative, or zero. If $d(t)$ is also positive or zero at the next observation, then we can enlarge T_2 to include this observation and create a longer upward trend. If $d(t)$ is negative at the next observation, the there be an h and a time interval T_3 such that $-h < d(t) < h$ on T_3. Clearly, T_3 immediately follows T_2 and by Definition 2.4 is a change period. Therefore an upward trend must occur between two change periods. The result for a downward trend is similar.

Property 2.1 is an important result. It tells us that when we detect a change period in a time series, it is the end of a trend.

3 Simulations

Figure 2 illustrates time series generated from each of six models, with lengths ranging from 100 to 400. All error terms are standard normal.

$$X_t = \begin{cases} t + \varepsilon_t & 1 \le t \le 200 \\ 200 - t + \varepsilon_t & 201 \le t \le 400 \end{cases}$$

1

$$X_t = \begin{cases} 200 - t + \varepsilon_t & 1 \le t \le 200 \\ t + \varepsilon_t & 201 \le t \le 400 \end{cases}$$

2

$$X_t = \begin{cases} t^2 + t + \varepsilon_t & 1 \le t \le 50 \\ 2550 - t^2 - t + \varepsilon_t & 51 \le t \le 100 \end{cases}$$

3

$$X_t = \begin{cases} 2550 - t^2 - t + \varepsilon_t & 1 \le t \le 50 \\ t^2 + t + \varepsilon_t & 51 \le t \le 100 \end{cases}$$

4

$$X_t = \begin{cases} 0.2 + 1.5X_{t-1} - 0.5X_{t-2} + \varepsilon_t - 0.2\varepsilon_{t-1} \\ \qquad 1 \le t \le 200 \\ 90 - 0.2 + 1.5X_{t-1} - 0.5X_{t-2} + \varepsilon_t - 0.2\varepsilon_{t-1} \\ \qquad 201 \le t \le 400 \end{cases}$$

5

$$X_t = \begin{cases} 14 - 0.2 + 1.5X_{t-1} - 0.5X_{t-2} + \varepsilon_t - 0.2\varepsilon_{t-1} \\ \qquad 1 \le t \le 200 \\ 0.2 + 1.5X_{t-1} - 0.5X_{t-2} + \varepsilon_t - 0.2\varepsilon_{t-1} \\ \qquad 201 \le t \le 400 \end{cases}$$

6

Figure 2: Simulation data for models 1-6

In transforming the difference time series $\{Y_t\}$ into the fuzzy time series, we choose the linguistic values set $\{L_i; i = 1,2,3,4,5\}$={fall sharply, fall, unchanged, rise, rise sharply}, and the membership function for model 1 to 6 are shown in Figures 3-8.

Figure 3: Membership function for model 1

Figure 4: Membership function for model 2

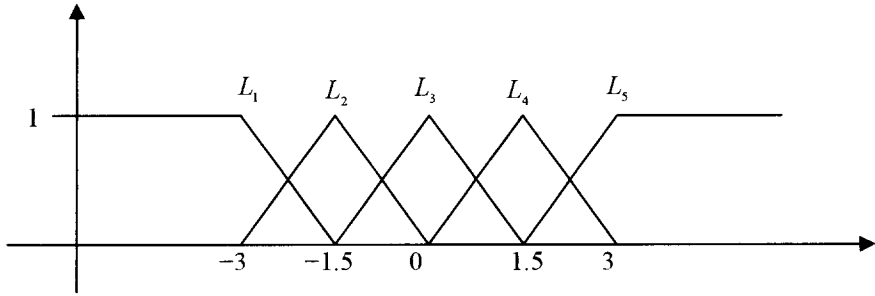**Figure 5: Membership function for model 3**

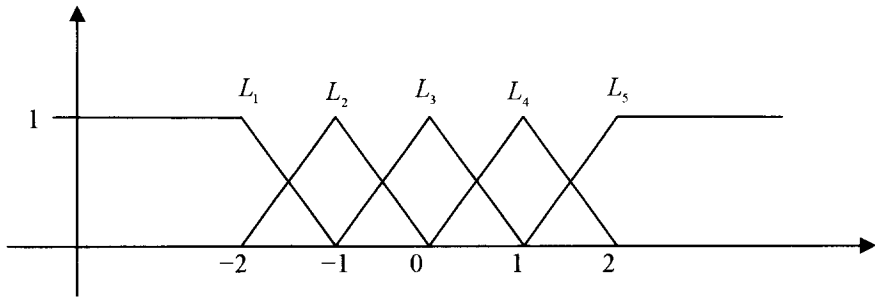**Figure 6: Membership function for model 4**
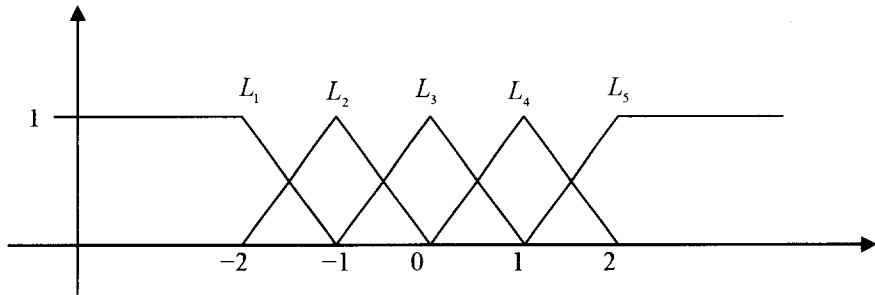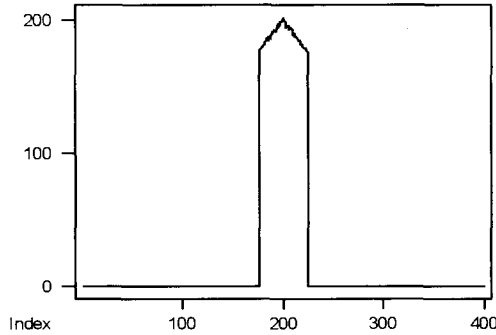

Figure 7: Membership function for model 5

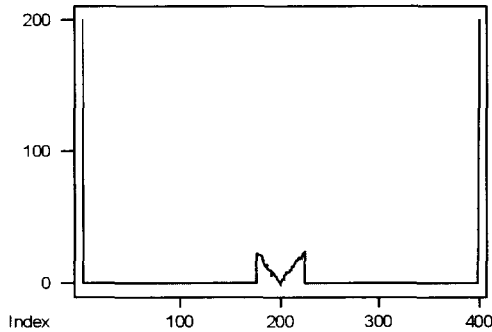**Figure 8: Membership function for model 6**

In all models we set the fuzzy weights $a_1, a_2, \ldots, a_5$ at -2, -1, 0, 1, 2. The degree of the change period detection sequence and the h in Definition 2.4 are set at 10 and 6. It is obvious that there is a clear change at t_{200} and t_{201} which are the end of an upward trend and the beginning of a downward trend. The change periods we found are expected to contain at least these two points. Following our proposed

method, we find the change period $T = \{t_{197}, t_{198}, \cdots, t_{204}\}$ which contains t_{200} and t_{201}. We show the partial time series on T in Figure 9, the twenty observations in the preceding and subsequent intervals are also included for a clearer visual grasp.

Figure 9: Partial time series at times 177 to 224 in model 1

In models 2, a change period is again detected at $T = \{t_{197}, t_{198}, \cdots, t_{204}\}$ when $h = 6$. Figure 10 shows X(177) to X(224) of model 2.

Figure 10: Partial time series at times 177 to 224 in model 2

In models 3 and 4, we set $h = 15$ and find a change period at $T = \{t_{47}, t_{48}, \cdots, t_{51}\}$. We show the results in Figures 11 and 12.
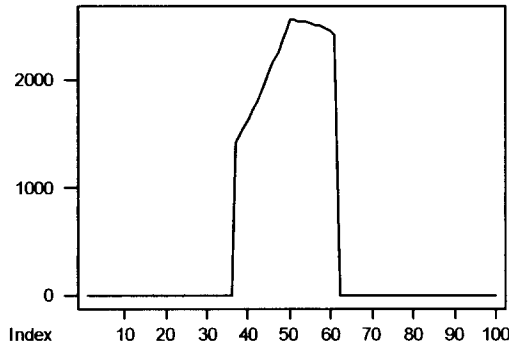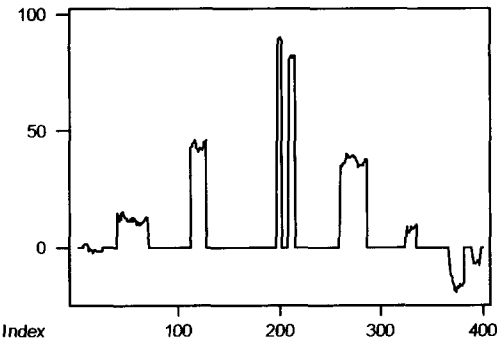
Figure 11: Partial time series at times 37 to 61 in model 3

Figure 12: Partial time series at times 37 to 61 in model 4

We can see from Figure 6 that there are many small change periods in model 5. A researcher can decide whether or not to treat these as noise just by controlling the degree of the change period detection sequence. Letting $h = 6$, we identify the change periods $T_1 = \{t_6, t_7, \cdots, t_{24}\}$, $T_2 = \{t_{40}, t_{41}, \cdots, t_{70}\}$, $T_3 = \{t_{113}, t_{114}, \cdots, t_{127}\}$, $T_4 = \{t_{198}, t_{199}, \cdots, t_{201}\}$, $T_5 = \{t_{209}, t_{210}, \cdots, t_{215}\}$, $T_6 = \{t_{260}, t_{261}, \cdots, t_{286}\}$, $T_7 = \{t_{325}, t_{326}, \cdots, t_{335}\}$, $T_8 = \{t_{375}, t_{376}, t_{377}\}$, and $T_9 = \{t_{391}, t_{392}, t_{393}, t_{394}\}$ in model 5.
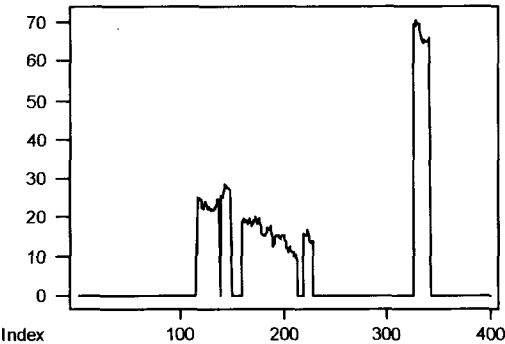
Figure 13: Partial time series at change periods of model 5

Similarly, setting $h = 6$ in model 6, we identify the change periods $T_1 = \{t_{119}, t_{120}, \cdots, t_{137}\}$, $T_2 = \{t_{140}, t_{141}, \cdots, t_{149}\}$, $T_3 = \{t_{160}, t_{161}, \cdots, t_{213}\}$, $T_4 = \{t_{220}, t_{221}, \cdots, t_{228}\}$, and $T_5 = \{t_{326}, t_{327}, \cdots, t_{341}\}$.

Figure 14: Partial time series at change periods of model 6

We change the data structure at t_{200} and t_{201} in models 1, 2, 5, and 6, and at t_{50} and t_{51} in models 3 and 4. In each model, the change periods found by our approach accurately contain those points. As you can see in Figures 13 and 14, we can even filter the noise and still accurately identify change periods.
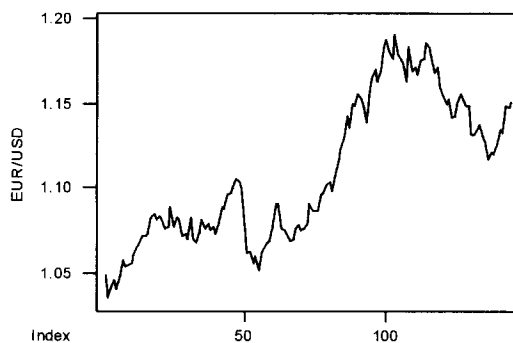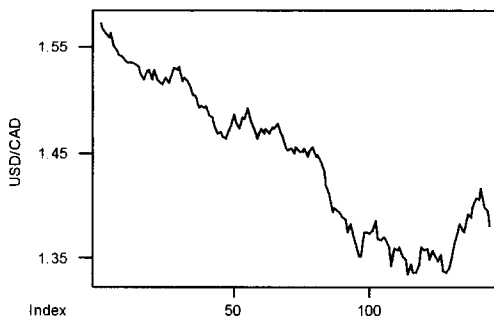
4   Empirical study

To investigate an application of our approach, three sets of exchange rate data were chosen: exchange rates for EUR (Europe) against USD (USA), USD against CAD (Canada), and GBP (UK) against USD. We chose these three series because they exhibit different structures. We arbitrarily chose the starting point January 1, 2003, and ending point July 25, 2003.

Each series contains 144 observations. Our goal is to detect change periods and to test the performance of our approach. Figures 15-17 illustrate the exchange rate series of EUR against USD, USD against CAD, and GBP against USD.
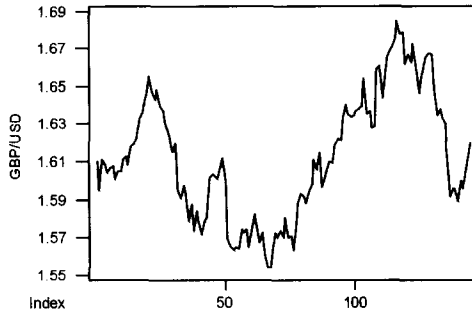

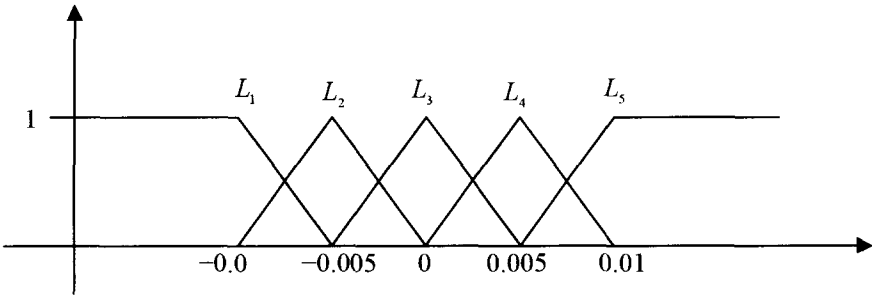**Figure 15: The exchange rate of EUR against USD**

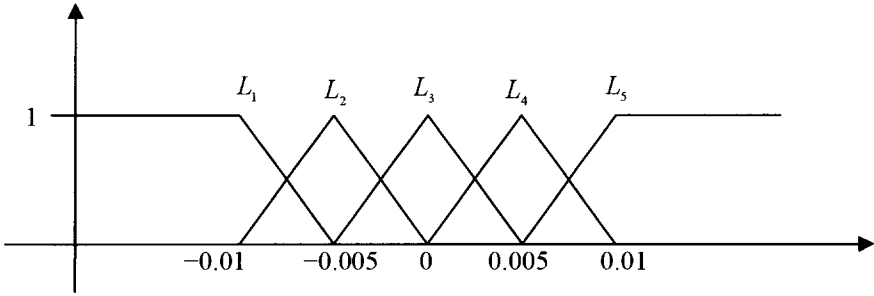**Figure 16: The exchange rate of USD against CAD**

Figure 17: The exchange rate of GBP against USD

As outlined in Section 2, for each series we first construct the first difference series. Then we transform the first difference series into a fuzzy time series with the linguistic values set $\{L_i; i = 1, 2, 3, 4, 5\}$ ={fall sharply, fall, unchanged, rise, rise sharply} and the membership functions shown in Figures 18-20. We set the fuzzy weights $a_1, a_2, \ldots, a_5$ at -3, -1, 0, 1, 3. The degree of the change period detection sequence and the h in Definition 2.4 are set at 10 and 5. Figures 21-23 plot the partial time series used to identify the change periods.
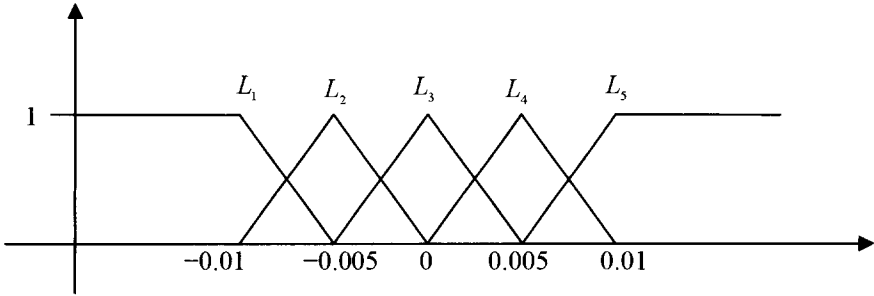

Figure 18: Membership function of the exchange rate of EUR against USD

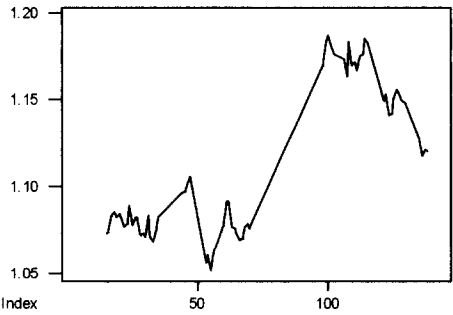**Figure 19: Membership function of the exchange rate of USD against CAD**

Figure 20: Membership function of the exchange rate of GBP against USD

Following our proposed method, we found change periods in the exchange rate series of EUR against USD at $T_1 = \{t_{15}, t_{16}, \cdots, t_{35}\}$, $T_2 = \{t_{44}, t_{45}, \cdots, t_{47}\}$, $T_3 = \{t_{53}, t_{54}, \cdots, t_{57}\}$, $T_4 = \{t_{60}, t_{61}, \cdots, t_{70}\}$, $T_5 = \{t_{98}, t_{99}, \cdots, t_{102}\}$, $T_6 = \{t_{106}, t_{107}, \cdots, t_{115}\}$, $T_7 = \{t_{121}, t_{122}, \cdots, t_{129}\}$, and $T_8 = \{t_{134}, t_{135}, \cdots, t_{138}\}$.
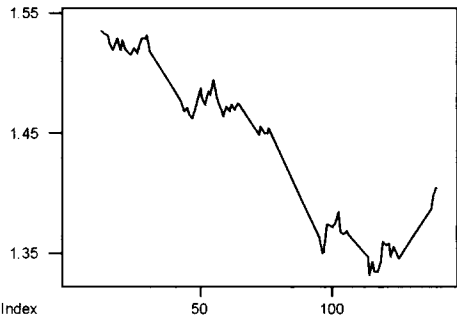
In the exchange rate series of USD against CAD, we found change periods $T_1 = \{t_{13}, t_{14}, \cdots, t_{31}\}$, $T_2 = \{t_{43}, t_{44}, \cdots, t_{48}\}$, $T_3 = \{t_{50}, t_{51}, \cdots, t_{65}\}$, $T_4 = \{t_{72}, t_{73}, \cdots, t_{76}\}$, $T_5 = \{t_{95}, t_{96}, \cdots, t_{106}\}$, $T_6 = \{t_{112}, t_{113}, \cdots, t_{125}\}$, and $T_7 = \{t_{137}, t_{138}, t_{139}\}$.

In the exchange rate series of GBP against USD, the change periods were detected at $T_1 = \{t_5, t_6, \cdots, t_{10}\}$, $T_2 = \{t_{19}, t_{20}, \cdots, t_{23}\}$, $T_3 = \{t_{33}, t_{34}, \cdots, t_{41}\}$, $T_4 = \{t_{43}, t_{44}, \cdots, t_{53}\}$, $T_5 = \{t_{54}, t_{55}, \cdots, t_{59}\}$, $T_6 = \{t_{62}, t_{63}, \cdots, t_{70}\}$, $T_7 = \{t_{98}, t_{99}, \cdots, t_{110}\}$, $T_8 = \{t_{114}, t_{115}, \cdots, t_{125}\}$, and $T_9 = \{t_{136}, t_{137}, t_{138}\}$.
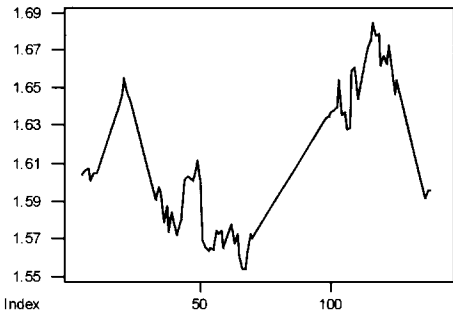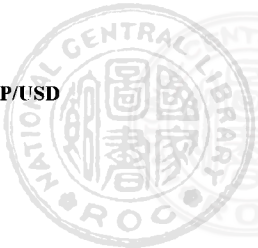
Figure 21: Partial time series at change periods of exchange rate of EUR/USD

Figure 22: Partial time series at change periods of exchange rate of USD/CAD

Figure 23: Partial time series at change periods of exchange rate of GBP/USD

Comparing Figures 15 and 21, Figures 16 and 22, and Figures 17 and 23, we can see that the results are satisfactory. The change periods that we detected in these three foreign exchange rates contains almost all the change periods that we visually recognize. Since the length of the series was selected arbitrarily, the performance of our approach shows that our proposed procedure can be applied in general models. Different scales of change periods can be obtained by adjusting the parameters of the procedure.

5   Conclusion

Economic and financial analysts often need to know when changes occur in a time series. In this research we formalize the concept of change periods in contrast with traditional change points as more realistic structural features of certain time series. We present an approach to detect change periods by partial cumulative sums of fuzzy statistics, allowing us to identify the beginnings and ends of trends.
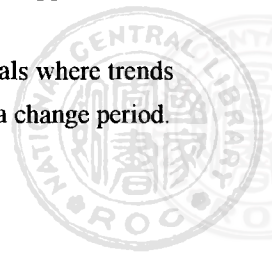
The application of fuzzy theory avoids the potential hazards of over fitting which might occur in traditional analysis with single observations. Through the use of fuzzy statistics, our proposed change period detection approach is able to systematically address fuzziness in the data. As a consequence, its results are more meaningful in practice.

The key contribution of this paper is that we provide a new method to detect change periods. In comparison with conventional methods, our approach offers several advantages:

(1)   Initial knowledge about the structure in the data is not required, so we can take full advantage of the model-free approach.

(2)   We can select standards for change periods by controlling the parameters to detect change periods at the scale desired and filter noise in a time series.

(3)   The fuzzy data can be handled.

Although the simulation and empirical results show that our approach of change period detection is visually satisfactory and can be generally applied, there remain several points to note and problems to be solved:

(1)   Because the change periods we consider are defined as intervals where trends change, the stationary part of a time series may be viewed as a change period.

(2)   Future study should address sensitivity of its results to the parameter choice.

References

Adlassing, K. D., (1986), "Fuzzy Set Theory in Medical Diagnostics," *IEEE Transactions on Systems, Man and Cybernetics*, 16, 260-264.

Balke, N. S., (1993), "Detecting Level Shifts in Time Series," *Journal of Business and Economic Statistics*, 11, 81-92.

Barry, D. and J. A. Hartigan, (1993), "A Bayesian Analysis for Change Point Problems," *Journal of the American Statistics Association*, 88, 309-319.

Bleaney, M., (1990), "Some Comparisons of the Relative Power of Simple Tests for Structural Change in Regression Models," *Journal of Forecasting*, 9, 437-444.

Broemeling, L. D. and H. Tsurumi, (1987), *Econometrics and Structural Change*, New York, NY: Marcel Dekker Inc.

Brown, R. L., J. Durbin and J. M. Evans, (1975) "Techniques for Testing the Constancy of Regression Relationships over Time," *Journal of the Royal Statistical Society B*, 37, 149-163.

Chow, G. C., (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591-605.

Clymer, J. R., P. D. Corey and J. Gardner, (1992), "Discrete Event Fuzzy Airport Control," *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 343-351.

Custem, B. V. and I. Gath, (1993), "Detection of Outliers and Robust Estimation Using Fuzzy Clustering," *Computational Statistics and Data Analysis*, 15, 47-61.

Gupta, M. M., P. N. Nikiforuk, Y. Tsukamoto and R. Martin-Clouaire, (1982), "An Application of Fuzzy Integral to Medical Diagnosis," *Proceedings of IFAC Congress*, New Dehli, 16-21.

Hathaway, R. J. and J. C. Bezdek, (1993), "Switching Regression Models and Fuzzy Clustering," *IEEE Transactions on Fuzzy Systems*, 1, 195-204.

Hinkley, D. V., (1971), "Inference about the Change-Point from Cumulative Sum Tests," *Biometrika*, 58, 509-523.

Hsu, D. A., (1979), "Detecting Shifts of Parameter in Gamma Sequences with Applications to Stock Price and Air Traffic Flow Analysis," *Journal of the*

American Statistical Association, 74, 31-40.

Hsu, D. A., (1982), "A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns," *Journal of the American Statistical Association*, 77, 29-39.

Inclan, C. and G. C. Tiao, (1994), "Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance," *Journal of the American Statistical Association*, 89, 913-923.

Klir, G. F. and T. A. Folger, (1988), *Fuzzy Sets, Uncertainty and Information*, Englewood Cliffs, NJ: Prentice Hall.

Lin, C. F. and T. Terasvirta, (1994), "Testing the Constancy of Regression Parameters against Continuous Structural Change," *Journal of Econometrics*, 62, 211-228.

Nyblom, J., (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223-230.

Page, E. S., (1955), "A Test for Change in a Parameter Occurring at an Unknown Point," *Biometrika*, 42, 523-527.

Ploberger, W., W. Kramer and K. Kontrus, (1989), "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics*, 40, 307-318.

Sastri, T., B. Flores and J. Valdes, (1989), "Detecting Points of Change in Time Series," *Computers Open Research*, 16, 271-293.

Tsay, R. S., (1988), "Outliers, Level Shift, and Variance Changes in Time Series," *Journal of Forecasting*, 7, 1-20.

Worsley, K. J., (1986), "Confidence Regions and Tests for a Change-Point in a Sequence of Exponential Family Random Variables," *Biometrika*, 73, 91-104.

Wu, B. and M. Chen, (1999), "Use of Fuzzy Statistical Methods in Change Periods Detection," *Applied Mathematics and Computation*, 99, 241-254.

Yoshinari, Y., W. Pedrycz and K. Hirota, (1993), "Construction of Fuzzy Models through Clustering Techniques," *Fuzzy Sets and Systems*, 54, 157-165.

Zadeh, L. A., (1965), "Fuzzy Sets," *Information and Control*, 8, 338-353.