axioms

MDPI

Article

A Combination of Fuzzy Techniques and Chow Test to Detect Structural Breaks in Time Series

Vilém Novák * and Thi Thanh Phuong Truong

Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic
* Correspondence: vilem.novak@osu.cz

Abstract: In a series of papers, we suggested a non-statistical method for the detection of structural breaks in a time series. It is based on the applications of special fuzzy modeling methods, namely Fuzzy transform (F-transform) and selected methods of Fuzzy Natural Logic (FNL). In this paper, we combine our method with the principles of the classical Chow test, which is a well-known statistical method for testing the presence of a structural break. The idea is to construct testing statistics similar to that of the Chow test which is formed from components of the first-degree F-transform. These components contain an estimation of the average values of the tangents (slopes) of the time series over an imprecisely specified time interval. In this paper, we illustrate our method and its statistical test on a real-time series and compare it with three classical statistical methods.

Keywords: time series; structural breaks; Chow test; fuzzy transform; F-transform; evaluative linguistic expressions; fuzzy natural logic

check for updates

Citation: Novák, V.; Truong, T.T.P. A Combination of Fuzzy Techniques and Chow Test to Detect Structural Breaks in Time Series. *Axioms* **2023**, *12*, 103. https://doi.org/10.3390/axioms12020103

Academic Editor: Radko Mesiar

Received: 21 December 2022
Revised: 11 January 2023
Accepted: 12 January 2023
Published: 19 January 2023

1. Introduction

Structural breaks in time series are sudden unexpected changes in their course triggered by a structural change in the system (for example, opening a new factory), by an attack from outside, or by some other outer cause. Unlike single outliers, structural breaks are characterized by a longer duration. There are various statistical methods for detection and testing whether a structural break occurs in the given time point, see [1–5], and also non-statistical methods such as genetic algorithms [6].

However, a structural break hardly occurs just at one point. It is a phenomenon lasting usually over several time points. Hence, we argue that techniques based on methods of fuzzy modeling are better suited for this purpose. In [7] and elsewhere, we presented an effective non-statistical method for finding structural breaks. It is based on the application of the so-called *fuzzy transform* (F-transform) in combination with methods of *fuzzy natural logic* (FNL). They enable us to detect intervals of specific monotonic behavior and to discover existing structural breaks in a time series. The main idea consists of the fact that the F-transform provides an estimation of the average value of the slope in an imprecisely specified time interval (cf. [8]). The former is then evaluated using a specific *evaluative linguistic expression* whose semantics is modeled inside FNL (for the details, see [9]). The method is very effective and we argue that it can find the real structural breaks. However, it is non-statistical and so it is a challenge to prove that the detected structural breaks can also be statistically verified. This is the topic of this paper. We modify the classical Chow test [1]. First, we form a modified null hypothesis stating that there is no structural break in the given area and then prove that it is rejected if a structural break is indeed present and detected using our fuzzy method.

The paper is structured as follows. Section 2 contains preliminaries in which we recall the principles of the Chow test, fuzzy transform, and the theory of evaluative linguistic expressions (as a part of FNL). In Section 3, we introduce the algorithm for finding structural breaks using our method, suggest modification of Chow test, and prove that if a structural break is detected then the null hypothesis stating that there is no structural break is rejected. Section 4 contains experimental verification of our results and also a comparison with three statistical techniques, namely the classical Chow, Bai-Perron, and Pettitt's tests.

2. Preliminaries

2.1. Chow Test

This is one of the first statistical tests (see [1,4]) using which it is possible to decide whether we are facing a structural break or not. Its idea consists of splitting the data having n observations into two parts: n_1 observations before the break in time t_0 and n_2 observations after it. Then we construct two linear regressions models

$$\begin{aligned} y_1(t) &= \alpha_1 + \beta_1 X_1(t) + \varepsilon_1, \qquad t = 1, \dots, n_1 \\ y_2(t) &= \alpha_2 + \beta_2 X_2(t) + \varepsilon_2, \qquad t = n_1 + 1, \dots, n_1 + n_2 \end{aligned}$$

which are compared with the general regression model

$$y(t) = \alpha + \beta X(t) + \varepsilon, \qquad t = 1, \dots, n$$

$(n = n_1 + n_2)$.

The ordinary least squares method is applied to the models above and then we form squares of residuals

- RSS_1 which is the residual of squares before the break:

$$\mathrm{RSS}_1 = \sum_{t=1}^{n_1} (y_1(t) - (\hat{\alpha}_1 + \hat{\beta}_1 X_1(t)))^2$$

- RSS_2 is the residual of squares after the break:

$$\mathrm{RSS}_2 = \sum_{t=n_1+1}^{n_2} (y_2(t) - (\hat{\alpha_2} + \hat{\beta_2} X_1(t)))^2$$

- RSS_3 is the residual of squares of the general regression model:

$$\mathrm{RSS}_3 = \sum_{t=1}^{n} (y(t) - (\hat{\alpha} + \hat{\beta} X(t)))^2.$$

The test of hypothesis related to the test of structural break is conducted by testing the null hypothesis

$$H_0 : \alpha_1 = \alpha_2, \beta_1 = \beta_2.$$

The structural change is caused due to different intercept terms as well as different regression coefficients. Under the assumption that the probability distribution of y_1, y_2 is normal, we can test the null hypothesis by the statistics

$$F_{test} = \frac{(\mathrm{RSS}_3 - (\mathrm{RSS}_1 + \mathrm{RSS}_2))/p}{(\mathrm{RSS}_1 + \mathrm{RSS}_2)/(n_1 + n_2 - 2p)} \tag{1}$$

which has the F-distribution $F(p, n_1 + n_2 - 2p)$. The null hypothesis H_0 is rejected if

$$F_{test} > F_{crit}$$

where F_{crit} is a critical value of the F-distribution for a corresponding significance level α (as usual, we consider $\alpha \in \{0.1, 0.05, 0.01\}$).

2.2. Fuzzy Transform (F-Transform)

The fuzzy transform is a technique for the approximation of continuous functions. In our case, it can be effectively applied to analysis and forecasting of time series. Let a bounded real continuous function $f : [a, b] \to [c, d]$ be given, where $a, b, c, d \in \mathbb{R}$.

Definition 1. *Let $c_0 < \cdots < c_n$ be fixed nodes in the interval $[a, b]$ where $c_0 = a$, $c_n = b$ with $n \geq 2$ and $a, b \in \mathbb{R}$. The set $\mathscr{A} = \{A_0, \ldots, A_n\}$ of fuzzy sets on $[a, b]$ is called a* fuzzy partition *of $[a, b]$ if the following conditions are fulfilled:*

- $A_k : [a, b] \to [0, 1]$, $A_k(c_k) = 1$;
- $A_k(x) = 0$ *if* $x \notin (c_{k-1}, c_{k+1})$ *(for $c_{-1} = a$ and $c_{n+1} = b$);*
- A_k *is continuous;*
- A_k *strictly increases on $[c_{k-1}, c_k]$ for $k = 1, \ldots, n$ and A_k strictly decreases on $[c_k, c_{k+1}]$ for $k = 0, \ldots, n-1$;*
- $\sum_{k=0}^{n} A_k(x) = 1$ *for all $x \in [a, b]$;*
- *If $c_k = a + hk$, where $h = (b - a)/n$, then fuzzy partition $\mathscr{A}$ is called* uniform *and the following holds for the fuzzy sets forming it: $A_k(c_k - x) = A_k(c_k + x)$, $x \in [0, h]$, $A_k(x) = A_{k-1}(x - h)$, $A_{k+1}(x) = A_k(x - h)$ where $k = 1, \ldots, n-1$ and $x \in [c_k, c_{k+1}]$.*

The fuzzy sets $A_k \in \mathscr{A}$ are often called *basic functions*. Their shape can be arbitrary (but fulfilling Definition 1). Most often they are simple triangles. The F-transform has two phases: direct and inverse.

Definition 2. *Let the set $\mathscr{A} = \{A_0, \ldots, A_n\}$ be a uniform fuzzy partition with triangular basic functions, h be the distance between nodes and $f : [a, b] \to [c, d]$ be a continuous function on $[a, b]$. The $(n+1)$-tuple $\mathbf{F}[f] = (F_0[f], \ldots, F_n[f])$ is a* direct fuzzy transform *of f where the elements $F_0[f], \ldots, F_n[f]$ are called components.*

- *The zero-degree fuzzy transform has components of the form*

$$F_k^0[f] = \frac{\int_a^b f(x) A_k(x) dx}{h}, \quad k = 0, \ldots, n. \tag{2}$$

- *The first-degree fuzzy transform has components of the form*

$$F_k^1[f](x) = (\beta_k^0[f] + \beta_k^1[f] \cdot (x - c_k)) \tag{3}$$

where

$$\beta_k^0[f] = \frac{\int_{c_{k-1}}^{c_{k+1}} f(x) A_k(x) dx}{h}, \tag{4}$$

$$\beta_k^1[f] = \frac{6 \int_{c_{k-1}}^{c_{k+1}} f(x)(x - c_k) A_k(x) dx}{h^3}. \tag{5}$$

The coefficient $\beta_k^1[f]$ provides estimation of an average value of the tangent (slope) of f over the area characterized by the fuzzy set A_k.

Remark 1. *Note that the coefficients $\beta_k^0[f]$ are identical with the components (2) of the zero-degree F-transform.*

Definition 3. *Let $\mathbf{F}^m[f] = (F_0^m[f], \ldots, F_n^m[f])$ be a direct F-transform of f due to Definition 2, where $m \in \{0,1\}$. (In fact, we can define F-transform of arbitrary degree. For our purposes, however, zero and first degrees are sufficient.). The* inverse *F-transform of f is a function*

$$\hat{f}_h^m(x) = \sum_{k=0}^{n} F_k^m[f] \cdot A_k(x), \quad x \in [a,b]. \tag{6}$$

It can be proved that $\hat{f}_h^m$ approximates the original function f with arbitrary precision (depending on h). We can set the parameters so that the approximating function $\hat{f}_h^m$ has desired properties. The computational complexity of the F-transform is linear. More details can be found in [8–10].

2.3. Fuzzy Natural Logic

Our applications to time series require selected methods of *Fuzzy Natural Logic* (FNL). This is a class of mathematical models characterizing some parts of human common sense thinking that is based on the use of natural language. It includes, besides others, the theory of *evaluative linguistic expressions* and their semantics and fuzzy/linguistic IF-THEN rules. In a form suitable for time series processing is FNL described in [9]. In this section, we will recall some of the main points needed below.

Evaluative Linguistic Expressions (We will often omit the adjective "linguistic".) are special expressions of natural language in the form

$$\langle \textit{linguistic hedge} \rangle \langle \textit{TE-adjective} \rangle \tag{7}$$

where $\langle$*linguistic hedge*$\rangle$ is a special adverb standing before $\langle$*TE-adjective*$\rangle$, that makes the adjective more or less specific. We will consider linguistic hedges *extremely (Ex), significantly (Si), very (Ve), rather (Ra), more or less (ML), roughly (Ro), very roughly (VR)*. The TE-adjectives are canonical adjectives *zero (Ze), small (Sm), medium (Me), big (Bi)*. Note that they can be replaced by many other adjectives, for example *shallow, medium deep, deep*, etc.

Remark 2. *Evaluative linguistic expressions (7) are called* simple. *We can introduce also complex ones that are formed using connective (and, or). Their syntax and semantics, however, are more complicated since they are not just boolean expressions. We need not consider them in this paper.*

The model of the semantics of evaluative expressions requires the concept of *context*. In our case, this is the interval $w = [v_L, v_S] \cup [v_S, v_R]$ where $v_L, v_S, v_R \in R$. The numbers have the following meaning: v_L is the left bound, v_S is a typical middle value, and v_R is the right bound.

Let $\mathscr{B}$ be an evaluative linguistic expression. The mathematical model of its meaning is a function $W \to \mathscr{F}(R)$ where W is a set of all contexts. Such a function is called *intension* of $\mathscr{B}$. If a context $w \in W$ is given then *extension* of $\mathscr{B}$ w.r.t. w is a fuzzy set from $\mathscr{F}(R)$. The details of this model are described in [9].

Let a value $x \in \mathbb{R}$ be given. We may now ask, what is a proper evaluative linguistic expression $\mathscr{B}$ using which we can characterize linguistically the value x? Of course, this depends on the context $w \in W$. For example, 100 USD may be a big money in a poor country but very small in a rich one. Therefore, we consider a special function of *local perception*

$$\mathscr{B} = LPerc(x, w) \tag{8}$$

which assigns an evaluative expression $\mathscr{B}$ to the value x w.r.t. the context $w \in W$. (*LPerc* is implemented in the special software LFL Forecaster described, e.g., in [9,11]).

A special class of evaluative expressions are those characterizing the trend of time series:

$$\text{Trend is } \langle \textit{direction} \rangle \tag{9}$$

where

- $\langle direction \rangle$:= stagnating | $\langle special\ hedge \rangle \langle sign \rangle$,
- $\langle sign \rangle$:= increasing | decreasing,
- $\langle special\ hedge \rangle$:= $\emptyset$ | negligibly | slightly | somewhat | clearly | roughly | sharply | quite largely | fairly large| hugely| significantly.

We must also consider the context w_{tg} for tangent that is here extended to have two parts: positive w^+_{tg} for the increase in time series and negative w^-_{tg} for its decrease.

3. Processing of Time Series Using Methods of Fuzzy Modeling

3.1. Processing of Time Series Using F-transform

A time series X is a mapping [12–15]

$$X : \mathbb{T} \times \Omega \to \mathbb{R},$$

where $\mathbb{T}$ is a set of numbers interpreted as time moments, Ω is a nonempty set of elementary random events and $(\Omega, \mathscr{C}, P)$ is a probability space, where $\mathscr{C}$ is a σ-algebra over Ω and P is a probability measure. In general, $\mathbb{T} \subset \mathbb{R}$ can be an arbitrary set. For our purposes, we will consider $\mathbb{T}$ to be a finite set of natural numbers $T = \{0, \ldots, n\} \subset \mathbb{N}$.

We assume that the time series can be decomposed into 4 components, namely

$$X(\omega, t) = Tr(t) + C(t) + S(t) + R(\omega, t), \qquad t \in T, \omega \in \Omega \tag{10}$$

where $Tr(t)$ and $C(t)$ are trend and cyclic components of the time series. These two components are usually combined into one component called *trend-cycle* $TC(t) = Tr(t) + C(t)$.

The $S(t)$ is the seasonal component and $R(\omega, t)$ is a random noise. The trend, cycle and seasonal components are ordinary functions not having stochastic character. The noise $R(\omega, t)$ is assumed to be a sequence of independent random variables with the mean $\mu = 0$ and variance $\sigma^2 < +\infty$.

It has been proved (see [9,16,17]) that using the F-transform, we can estimate trend Tr or trend-cycle TC with high fidelity which means that the seasonal component S is almost "wiped out" (i.e., its inverse F-transform is close to zero) and the noise R is significantly reduced [18].

3.2. Detection of Structural Breaks in a Time Series

Detection of a structural break in a time series means to determine a time interval in which the course of the time series abnormally changes in comparison with its previous/subsequent development. Our detection method presented in [7] is based on finding short intervals with a steep slope of trend (big tangent) preceded or followed by an interval with a small slope.

Let X be a time series, $\bar{\mathbb{T}} \subseteq T$ be a time interval. Let $\mathscr{A}$ be a basic function due to Definition 1 with the support $\bar{\mathbb{T}}$ over which $\beta^1[X|\bar{\mathbb{T}}]$ is the slope of the trend of X computed using (5). Let w^-_{tg}, w^+_{tg} be the corresponding negative and positive parts of the context, respectively. Then, the evaluative expression $\pm Ev[\mathrm{X}|\bar{\mathbb{T}}]$ obtained using the function of local perception (8)

$$\pm Ev[X|\bar{\mathbb{T}}] = LPerc(\pm\beta^1[X|\bar{\mathbb{T}}], w^{\pm}_{tg}). \tag{11}$$

evaluates the trend of the time series X in the interval $\bar{\mathbb{T}}$. Using evaluation (11), we can decompose the time domain T into a set of intervals

$$\mathscr{T} = \{\bar{T}_i \mid i = 1, \ldots, s\}, \qquad \bigcup \mathscr{T} = \mathbb{T}$$

in which the slope is evaluated by specific evaluative expressions. The intervals need not necessarily be disjointed.

Definition 4. *Let* $\pm Ev[X|\bar{\mathbb{T}}_{i-1}], \pm Ev[X|\bar{\mathbb{T}}_i], \pm Ev[X|\bar{\mathbb{T}}_{i+1}]$ *be evaluative expressions computed using (11). An interval* $\bar{\mathbb{T}}_i \in \mathscr{T}$ *is an area of a* structural break *in the course of* X *if* $\pm Ev[X|\bar{\mathbb{T}}_i] \in \{\pm Ve\,Bi, \pm Si\,Bi, \pm Ex\,Bi\}$ *and* $\pm Ev[X|\bar{\mathbb{T}}_k] \in \{Ze, \pm Ve\,Sm, \pm Si\,Sm, \pm Ex\,Sm\}$ *where* $k = i-1$ *or* $k = i+1$.

This definition is the basis for the following algorithm for finding structural breaks in a time series X.

Algorithm for Finding Structural Breaks

- Set the distance $h > 0$ and determine a uniform fuzzy partition $\mathscr{A}$ over the time domain $\mathbb{T}$ due to Definition 1.
- Set the context w_{tg}^-, w_{tg}^+ for evaluation of the trend in the areas determined by the basic functions.
- Compute the direct first-degree fuzzy transform $\mathbf{F}^1[X] = (F_1^1[X], \dots, F_{n-1}^1[X])$ over the fuzzy partition $\mathscr{A}$.
- Localize all pairs of components $(F_k^1[X], F_{k+1}^1[X])$ (cf. (3)) with the following properties:
  - $LPerc(\beta_k^1[X], w_{tg}^\pm) \in \{Ze, Ve\,Sm, Si\,Sm, Ex\,Sm\}$, i.e., the coefficient β_k^1 is close to zero.
  - $LPerc(\beta_{k+1}^1[X], w_{tg}^\pm) \in \{Ve\,Bi, Si\,Bi, Ex\,Bi\}$, i.e., the coefficient β_{k+1}^1 is unexpectedly big.

  Alternatively, k and $k+1$ can be interchanged, i.e., β_{k+1}^1 is close to zero and β_k^1 is unexpectedly big.
- The interval $\mathbb{T}_k$ (or $\mathbb{T}_{k+1}$) which is a support of the basic function $A_k \in \mathscr{A}$ (or $A_{k+1} \in \mathscr{A}$) is the area of a structural break due to Definition 4.

  Without loss of generality, we will in the sequel assume that β_k^1 is close to zero and β_{k+1}^1 is unexpectedly big.

3.3. Combination of Fuzzy Techniques and Chow Test

Suppose that a time series X contains a structural break that occurs in the intersection of two adjacent areas represented by the triangular basic functions (fuzzy sets) A_k, $A_{k+1} \in \mathscr{A}$ in the sense of the algorithm presented in the previous section. Moreover, we also consider a third triangular basic function denoted by $A_{k,k+1}$ over the nodes c_{k-1}, c_{k+2} where the node

$$c_{k,k+1} = \frac{c_k + c_{k+1}}{2}$$

is added. Note that the fuzzy set $A_{k,k+1}$ is defined over nodes with the distance $\frac{3}{2}h$. The situation is depicted in Figure 1 (the fuzzy set $A_{k,k+1}$ is depicted upside-down for better visibility).

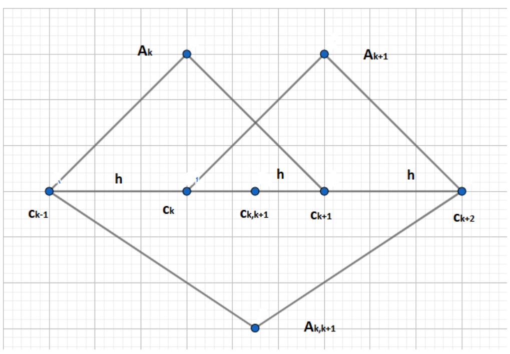


Figure 1. Fuzzy sets covering the structural break occurring in the interval $[c_k, c_{k+2}]$ (the support of the fuzzy set A_{k+1}).

Remark 3. *The triangular fuzzy set A_k w.r.t. the node c_k is defined by the equation*

$$A_k(t) = 1 - \frac{|t - c_k|}{h}, \qquad t \in [c_{k-1}, c_{k+1}]$$

and analogously also the fuzzy sets A_{k+1} and $A_{k,k+1}$. Note also that

$$t - c_{k,k+1} = \frac{t - c_k}{2} + \frac{t - c_{k+1}}{2}.$$

Lemma 1. *Let the triangular fuzzy sets depicted in Figure 1 be given. Then*

$$A_{k,k+1}(t) \leq A_k(t) + A_{k+1}(t)$$

for all $t \in T$.

Proof. It is enough to consider only the interval $[c_{k-1}, c_{k+2}]$. By simple computation, we verify that $A_{k,k+1}(t) \leq A_k(t)$ for $t \in [c_{k-1}, c_k]$ and $A_{k,k+1}(t) \leq A_{k+1}(t)$ for $t \in [c_{k+1}, c_{k+2}]$. For $t \in [c_k, c_{k+1}]$ the inequality follows from the property $A_k(t) + A_{k+1}(t) = 1$. □

Let us now consider the following F^1-transform components:

$$F^1_k[X](t) = \beta^0_k[X] + \beta^1_k[X] \cdot (t - c_k), \tag{12}$$

$$F^1_{k+1}[X](t) = \beta^0_{k+1}[X] + \beta^1_{k+1}[X] \cdot (t - c_{k+1}), \tag{13}$$

$$F^1_{k,k+1}[X](t) = \beta^0_{k,k+1}[X] + \beta^1_{k,k+1}[X] \cdot (t - c_{k,k+1}) \tag{14}$$

where the coefficients β^0, β^1 are determined by the corresponding basic functions due to Figure 1.

Lemma 2. *Let the triangular fuzzy sets depicted in Figure 1 be given. Then*

$$\beta^0_{k,k+1} \leq \beta^0_k + \beta^0_{k+1}.$$

Proof. Using (4) and Lemma 1 we obtain

$$\beta^0_{k,k+1} = \frac{1}{\frac{3}{2}h} \int_{c_{k-1}}^{c_{k+2}} X(t) A_{k,k+1}(t) dt \leq \frac{1}{h} \int_{c_{k-1}}^{c_{k+2}} X(t)(A_k(t) + A_{k+1}(t)) dt =$$
$$\frac{1}{h} \int_{c_{k-1}}^{c_{k+1}} X(t) A_k(t) dt + \frac{1}{h} \int_{c_k}^{c_{k+2}} X(t) A_{k+1}(t) dt = \beta^0_k + \beta^0_{k+1}.$$

□

Lemma 3. *Let the triangular fuzzy sets depicted in Figure 1 be given. Then*

$$\beta^1_{k,k+1} \leq \beta^1_k + \beta^1_{k+1}.$$

Proof. Using (5) and Lemma 1 we obtain

$$\begin{aligned}
\beta^1_{k,k+1} = \frac{6}{(\frac{3}{2}h)^3}\int_{c_{k-1}}^{c_{k+2}} X(t)(t-c_{k,k+1})A_{k,k+1}(t)dt \leq \\
\frac{6}{h^3}\int_{c_{k-1}}^{c_{k+2}} X(t)(t-c_{k,k+1})(A_k(t)+A_{k+1}(t))dt = \\
\frac{1}{2}\left[\frac{6}{h^3}\int_{c_{k-1}}^{c_{k+1}} X(t)(t-c_k)A_k(t)dt + \frac{6}{h^3}\int_{c_{k-1}}^{c_{k+1}} X(t)(t-c_{k+1})A_k(t)dt\right] + \\
\frac{1}{2}\left[\frac{6}{h^3}\int_{c_k}^{c_{k+2}} X(t)(t-c_k)A_{k+1}(t)dt + \frac{6}{h^3}\int_{c_k}^{c_{k+2}} X(t)(t-c_{k+1})A_{k+1}(t)dt\right] = \\
\left[\frac{1}{2}\beta^1_k + \frac{1}{2}\beta^1_k - \frac{h}{2}\frac{6}{h^3}\int_{c_{k-1}}^{c_{k+1}} A_k(t)\right] + \left[\frac{1}{2}\beta^1_{k+1} + \frac{1}{2}\beta^1_{k+1} + \frac{h}{2}\frac{6}{h^3}\int_{c_k}^{c_{k+2}} A_k(t)\right] = \\
\beta^1_k - \frac{6}{h} + \beta^1_{k+1} + \frac{6}{h} = \beta^1_k + \beta^1_{k+1}.
\end{aligned}$$

□

Lemma 4. *Let (12)–(14) be components of the* F^1 *transform according to the fuzzy partition depicted in Figure 1. Let* $\beta^1_k = 0$ *and* $\beta^1_{k+1} = K \neq 0$*. Then* $\beta^1_{k,k+1} \neq 0$*.*

Proof. It is sufficient to consider the interval $[c_{k+1}, c_{k+2}]$, in which the fuzzy set $A_k(t) = 0$ and, therefore, it cannot affect the size of β^1_{k+1} and $\beta^1_{k,k+1}$ (cf. formula (5)).

Put $H = \min\{X(t) \neq 0 \mid t \in [c_{k+1}, c_{k+2}]\}$. By the assumption, if there is a structural break over $[c_k, c_{k+2}]$ then $H \neq 0$ and it should be large since the corresponding $X(t)$ are not covered by the fuzzy set A_k over which $\beta^1_k = 0$ by the assumption. Then,

$$\begin{aligned}
L &= \int_{c_{k+1}}^{c_{k+2}} H(t-c_{k+1})A_{k+1}(t)dt \leq \int_{c_{k+1}}^{c_{k+2}} X(t)(t-c_{k+1})A_{k+1}(t)dt, \\
L' &= \int_{c_{k+1}}^{c_{k+2}} H(t-c_{k,k+1})A_{k,k+1}(t)dt \leq \int_{c_{k+1}}^{c_{k+2}} X(t)(t-c_{k,k+1})A_{k,k+1}(t)dt.
\end{aligned}$$

After computation, we obtain $L = \frac{Hh^2}{6}$ and $L' = \frac{5Hh^2}{18}$ which means that $0 \neq L < L'$. Consequently, $\beta^1_{k,k+1} \neq 0$. □

Let us now define the following sums of squares of differences between values of the time series and the F^1-transform components corresponding to the basic functions from Figure 1:

$$\mathrm{RSS}_k = \sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F^1_k[X](t))^2 A_k(t), \tag{15}$$

$$\mathrm{RSS}_{k+1} = \sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F^1_{k+1}[X](t))^2 A_{k+1}(t), \tag{16}$$

$$\mathrm{RSS}_{k,k+1} = \sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F^1_{k,k+1}[X](t))^2. \tag{17}$$

Then, analogously to the classical Chow test (1) we construct the statistics

$$F_{test} = \frac{(\mathrm{RSS}_{k,k+1} - (\mathrm{RSS}_k + \mathrm{RSS}_{k+1}))/p}{(\mathrm{RSS}_k + \mathrm{RSS}_{k+1})/(2n-2p)} \tag{18}$$

where n is the number of time points between $[c_{k-1}, c_{k+2}]$ and p is the number of parameters (note that $p = 2$ in our case). The statistics (18) has the F-distribution $F(p, 2n-2p)$. The structural break is tested by the null hypothesis:

$$H_0 : \beta^0_k = \beta^0_{k+1}, \beta^1_k = \beta^1_{k+1}. \tag{19}$$

The null hypothesis H_0 is rejected if $F_{test} > F_{crit}$ where F_{crit} is a critical value.

The statistics (17) can be rewritten as

$$F_{test} = \frac{2(n-p)}{p}\left(\frac{\mathrm{RSS}_{k,k+1}}{\mathrm{RSS}_k + \mathrm{RSS}_{k+1}} - 1\right). \tag{20}$$

We will refer to (18) or (20) as *fuzzy Chow test*.

Theorem 1. *Let $X(t)$ be a time series over the time domain $\mathbb{T}$ and $\mathscr{A}$ be a uniform triangular fuzzy partition of $\mathbb{T}$ with the distance h between nodes. Then*

$$\left(\frac{\mathrm{RSS}_{k,k+1}}{\mathrm{RSS}_k + \mathrm{RSS}_{k+1}}\right) > 1. \tag{21}$$

Proof. By ([8], Corollary 2), $X(t) = F_k^1(t) + O(h^2)$ for $t \in [c_{k-1}, c_{k+1}]$. Then,

$$(X(t) - F_k^1(t))^2 \le Mh^4$$

for some constant M and arbitrary h. Similar inequality holds also for F_{k+1}^1 and $t \in [c_k, c_{k+2}]$. Hence,

$$\mathrm{RSS}_k + \mathrm{RSS}_{k+1} = \textstyle\sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F_k^1(t))^2 A_k(t) + \sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F_{k+1}^1(t))^2 A_{k+1}(t) < M'h^4 \tag{22}$$

for some constant M' and $t \in [c_{k-1}, c_{k+2}]$. Considering a wider fuzzy partition with $h' = \frac{3}{2}h$, in a similar way we obtain that

$$\mathrm{RSS}_{k,k+1} = \sum_{t=c_{k-1}}^{c_{k+2}} (X(t) - F_{k,k+1}^1[X](t))^2 < M'\left(\frac{3}{2}h\right)^4 \tag{23}$$

(in both cases we can consider the same constant M'). We argue that

$$\mathrm{RSS}_k + \mathrm{RSS}_{k+1} < \mathrm{RSS}_{k,k+1}\,.$$

Indeed, let the opposite inequality hold. Then we choose h so that $\mathrm{RSS}_{k,k+1} < M'(\frac{3}{2}h)^4 < \mathrm{RSS}_k + \mathrm{RSS}_{k+1}$ and, by the assumption and (22), we obtain

$$\mathrm{RSS}_{k,k+1} < M'(\frac{3}{2}h)^4 < \mathrm{RSS}_k + \mathrm{RSS}_{k+1} < M'h^4$$

which is a contradiction. □

Theorem 2. *Let $X(t)$ be a time series over the time domain $\mathbb{T}$ and $\mathscr{A}$ be a uniform triangular fuzzy partition of $\mathbb{T}$ with the distance h between the nodes. Let F_k^1, F_{k+1}^1 be two components identifying structural break in the areas characterized by A_k, A_{k+1} according to the algorithm presented in Section 3.2 (cf. Figure 1). Let*

$$|X(t) - F_k^0(t)| \le \varepsilon, \quad t \in [c_{k-1}, c_{k+1}] \quad \text{and} \quad |X(t) - F_{k+1}^1(t)| \le \varepsilon, \quad t \in [c_k, c_{k+2}]. \tag{24}$$

If

$$\sum_{t \in [c_{k-1}, c_{k+2}]} (X(t) - F_{k,k+1}^1(t))^2 > m_h \varepsilon^2$$

where m_h is the number of time points between two nodes, then $F_{test} >> 1$.

Proof. Let us consider formula (20) and denote the formulas after the sum symbol in (15)–(17) by $\mathrm{RSS}_k(t)$, $\mathrm{RSS}_{k+1}(t)$, $\mathrm{RSS}_{k,k+1}(t)$, respectively.

(a) Let $t \in [c_{k-1}, c_k]$. Then

$$\text{RSS}_{k+1}(t) = 0 \quad \text{and} \sum_{t \in [c_{k-1}, c_k]} \text{RSS}_k(t) \le m_h \varepsilon^2,$$

i.e., $\sum_{t \in [c_{k-1}, c_k]} \text{RSS}_k(t) + \sum_{t \in [c_{k-1}, c_k]} \text{RSS}_{k+1}(t) \le m_h \varepsilon^2$. Similarly we obtain

$$\sum_{t \in [c_{k+1}, c_{k+2}]} \text{RSS}_k(t) + \sum_{t \in [c_{k+1}, c_{k+2}]} \text{RSS}_{k+1}(t) \le m_h \varepsilon^2.$$

(b) Let $t \in [c_k, c_{k+1}]$. Then

$$\sum_{t=c_k}^{c_{k+1}} (X(t) - F_k^1[X](t))^2 A_k(t) + \sum_{t=c_k}^{c_{k+1}} (X(t) - F_{k+1}^1[X](t))^2 A_k(t) \le \sum_{t=c_k}^{c_{k+1}} (\varepsilon^2 A_k(t) + \varepsilon^2 A_{k+1}(t)) = m_h \varepsilon^2.$$

Consequently,

$$\sum_{t=c_{k-1}}^{c_{k+2}} \text{RSS}_k(t) + \sum_{t=c_{k-1}}^{c_{k+2}} \text{RSS}_{k+1}(t) \le m_h \varepsilon^2. \tag{25}$$

Using the assumption, we thus obtain

$$1 < \frac{\text{RSS}_{k,k+1}}{m_h \varepsilon^2} \le \frac{\text{RSS}_{k,k+1}}{\text{RSS}_k + \text{RSS}_{k+1}}.$$

Realizing that for $p = 2$, the multiplicative constant in (20) can be fairly high, the value of F_{test} is also high. □

So far, we did not see whether our method for detection of the structural breaks has an impact on the statistics (20). In the following theorem we will show that when detecting a structural break using the algorithm from Section 3.2, the nominator of (20) significantly increases. For this purpose, we consider two time series X, X' which differ only in the interval $t \in [c_{k,k+1}, c_{k+2}]$ where the structural break occurs, i.e., $X'(t) = X(t)$ for $t \in [c_{k-1}, c_{k,k+1}]$. We will denote the formulas (15)–(17) for the time series X' by $\text{RSS}'_k, \text{RSS}'_{k+1}, \text{RSS}'_{k,k+1}$, respectively.

In the proof of the following theorem, we meet the formula

$$W = \sum_{[c_{k-1}, c_{k+2}]} (t - c_{k,k+1})^2 = \frac{1}{2} m_h (\frac{3}{2} m_h + 1)(3m_h + 1).$$

Theorem 3. *Let m_h be the number of time points between two nodes. Furthermore, we will assume:*

(a) $\beta_k^1[X] = \beta_{k,k+1}^1[X] = \beta_{k+1}^1[X] = 0$*,*
(b) *the inequalities (24) hold for some* $\varepsilon > 0$*,*
(c) $|\beta_{k+1}^1[X']| > 0$*,*
(d) *there are minimal* $\varepsilon', \varepsilon'' > 0$ *such that the following holds:*

$$\max\{|X(t) - X(c_{k,k+1})| \mid t \in [c_{k-1}, c_{k+2}]\} \le \frac{\varepsilon'}{2},$$
$$\max\{|X'(t) - X'(c_{k,k+1})| \mid t \in [c_{k-1}, c_{k+2}]\} \le \frac{\varepsilon''}{2}.$$

If the slope $\beta_{k,k+1}^1[X']$ *fulfills the inequality*

$$|\beta_{k,k+1}^1[X']| \ge \frac{\sum_{[c_{k-1}, c_{k+2}]} |t - c_{k,k+1}| \cdot |X(t) - \beta_{k,k+1}^0[X]|}{W} \tag{26}$$

then

$$\mathrm{RSS}_{k,k+1} \leq \mathrm{RSS}'_{k,k+1}\,. \tag{27}$$

Proof. Note that $X'(c_{k,k+1}) = X(c_{k,k+1})$ and so, it follows from the assumptions (c), (d) that $\varepsilon' < \varepsilon''$ because (c) says that there is a structural break in $[c_k, c_{k+2}]$. Therefore, there must be $t_0 \in [c_k, c_{k+2}]$ such that $|X'(t_0) - X(c_{k,k+1})| > \varepsilon'$.

In the same way as is in the proof of ([10], Theorem 5), we can show that

$$|X(t) - \beta^0_{k,k+1}[X]| \leq \varepsilon', \quad \text{and} \quad |X'(t) - \beta^0_{k,k+1}[X']| \leq \varepsilon''$$

for $t \in [c_{k-1}, c_{k+2}]$. Then we have

$$\sum_{t\in[c_{k-1},c_{k+2}]} (X(t) - \beta^0_{k,k+1}[X])^2 \leq 3m_h\varepsilon'^2 < \sum_{t\in[c_{k-1},c_{k+2}]} (X(t) - \beta^0_{k,k+1}[X'])^2 -$$
$$2|\beta^1_{k,k+1}[X']| \sum_{t\in[c_{k-1},c_{k+2}]} |t - c_{k,k+1}| \cdot |X(t) - \beta^0_{k,k+1}[X']| + (\beta^1_{k,k+1}[X'])^2 \sum_{t\in[c_{k-1},c_{k+2}]} (t - c_{k,k+1})^2.$$

The third inequality is assured if

$$2|\beta^1_{k,k+1}[X']| \cdot \sum_{t\in[c_{k-1},c_{k+2}]} |t - c_{k,k+1}| \cdot |X(t) - \beta^0_{k,k+1}[X']| \leq$$
$$(\beta^1_{k,k+1}[X'])^2 \sum_{t\in[c_{k-1},c_{k+2}]} (t - c_{k,k+1})^2,$$

which implies that

$$|\beta^1_{k,k+1}[X']| \geq \frac{\sum_{[c_{k-1},c_{k+2}]} |t - c_{k,k+1}| \cdot |X(t) - \beta^0_{k,k+1}[X]|}{W}.$$

This assures inequality (27). □

In this theorem, assumption (b) says that $F^0_k(t)$ and $F^1_{k+1}(t)$ well approximate X and X', respectively. Note that due to assumption (a), we may consider only F^0_k instead of $F^1_k(t)$. Note that Lemma 4 justifies assumption (c); assumption (d) is justified by the results of [8,10].

Corollary 1. *There is a context $w_{tg} = \langle 0, v_S, v_R\rangle$ such that $LPerc(\beta^1_k[X'], w_{tg}) \in \{zero\}$ and $LPerc(\beta^1_{k+1}[X'], w_{tg}) \in \{big\}$ and the null hypothesis using the fuzzy Chow-test for the time series X' is rejected.*

Proof. We take the context w_{tg} such that v_S is equal to the right-hand side of (26) and v_R big enough to assure that $\beta^1_{k+1}[X']$ is in (11) evaluated as *big*. Following the proof of Theorem 2, we can assure that (21) holds true and, consequently, that F_{test} in (20) is sufficiently large. □

4. Experiments

4.1. Detection of Structural Breaks

We will demonstrate our results on real data taken from a Micro subset of time series from M4-Competition published on the Internet. For computations of the F-transform components, we used the experimental software FT-studio. (It was developed in the Institute for Research of Applications of Fuzzy Modeling of the University of Ostrava, Czech Republic. Its author is Radek Valášek. Let us remark that in the R-repository is available R-package lfl providing algorithms used in this paper, see [19].).

In Figure 2, a real time series is depicted together with two fuzzy partitions with equidistant triangular basic functions. The structural breaks are detected using our method.

Namely, the time series contains structural breaks occurring in areas characterized by couples of fuzzy sets (A_{48}, A_{52}), (A_{204}, A_{208}), (A_{244}, A_{248}).

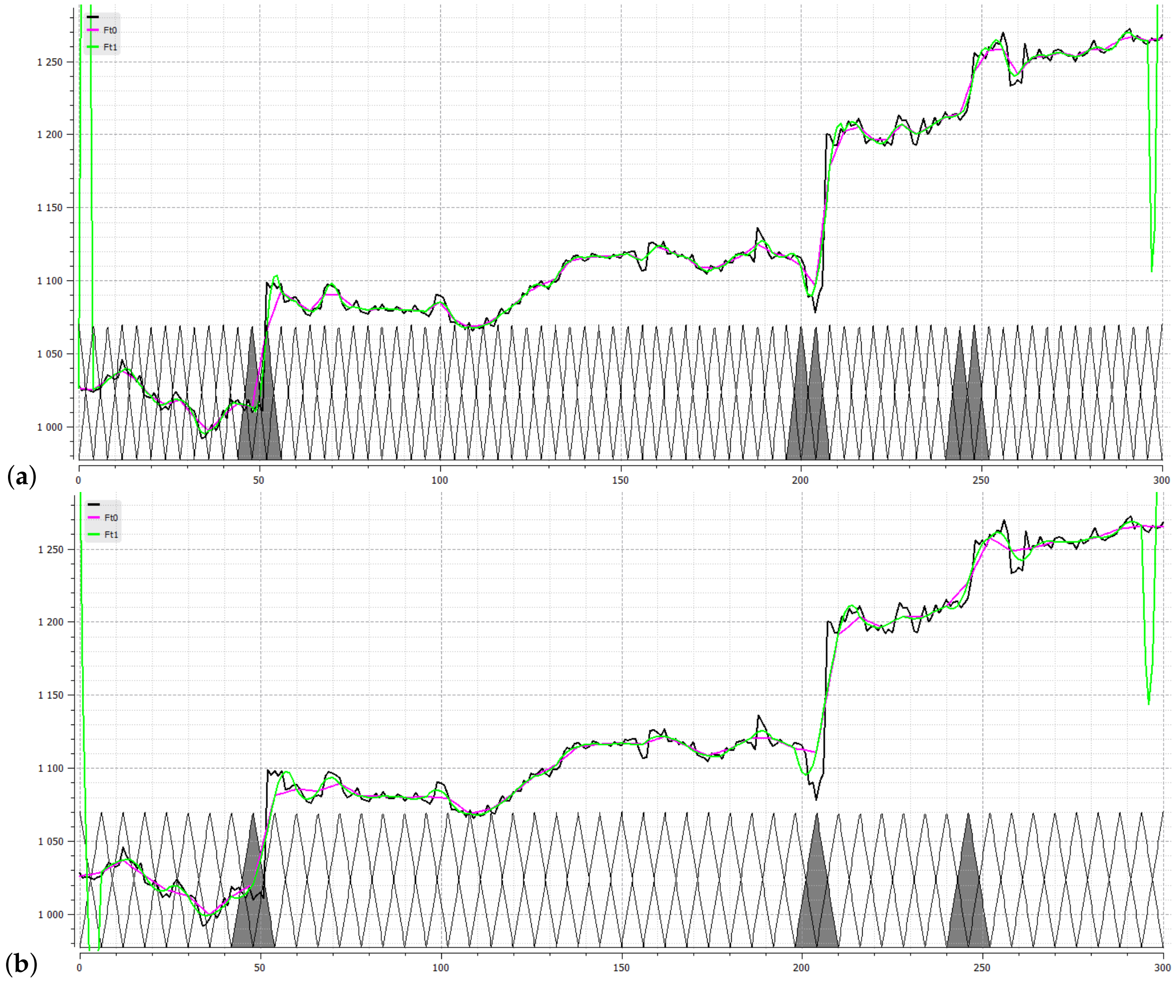


Figure 2. The time series is depicted together with fuzzy partitions with the basic functions of the width: (**a**) $2h = 8$ and (**b**) $2h = 12$. In both images are marked basic functions over which the structural breaks are detected.

We apply Fuzzy Chow test to the found structural breaks using fuzzy technique. The first one is obtained from two F^1-transform components

$$F^1_{48}[X] = \beta^0_{48}[X] + \beta^1_{48}[X] \cdot (t - c_{48}) = 1013.19 - 0.38 \cdot (t - 48),$$
$$F^1_{52}[X] = \beta^0_{52}[X] + \beta^1_{52}[X] \cdot (t - c_{52}) = 1065.41 + 20.83 \cdot (t - 52).$$

To verify the found structural break also statistically, we will test the null hypothesis

$$H_0 : \beta^0_{48}[X(t)] = \beta^0_{52}[X(t)] \quad and \quad \beta^1_{48}[X(t)] = \beta^1_{52}[X(t)]$$

The squares of residuals RSS_{48}, RSS_{52}, $\mathrm{RSS}_{48,52}$ are computed as follows:

$$\mathrm{RSS}_{48} = \sum_{t=44}^{56} (X_t - F^1_{48}[X](t))^2 \cdot A_{48}(t) = 38.9604,$$
$$\mathrm{RSS}_{52} = \sum_{t=44}^{56} (X_t - F^1_{52}[X](t))^2 \cdot A_{52}(t) = 2441.03,$$
$$\mathrm{RSS}_{48,52} = \sum_{t=44}^{56} (X_t - F^1_{48,52}[X](t))^2 = 9406.196.$$

The Fuzzy Chow test is based on the statistics

$$F_{test} = \frac{(\mathrm{RSS}_{48,52} - \mathrm{RSS}_{48} - \mathrm{RSS}_{52})/p}{(\mathrm{RSS}_{48} + \mathrm{RSS}_{52})/(2n-2p)} = 30.72$$

where $n = 13$ and $p = 2$.

We obtain $F_{test} = 30.72 > F_{crit} = F(p, 2n-2p) = 4.3828$ for $\alpha = 0.025$ and so, we reject the null hypothesis. This means that we have sufficient evidence to say that a structural break detected using our method indeed occurs in the interval $t \in [44, 56]$.

We apply the same algorithm with two subsequent by fuzzy sets (A_{204}, A_{208}) and (A_{244}, A_{248}), the structural breaks confirmed by testing the null hypothesis:

$$H_0 : \beta^0_{204}[X(t)] = \beta^0_{208}[X(t)] \quad \textit{and} \quad \beta^1_{204}[X(t)] = \beta^1_{208}[X(t)],$$
$$H_0 : \beta^0_{244}[X(t)] = \beta^0_{248}[X(t)] \quad \textit{and} \quad \beta^1_{244}[X(t)] = \beta^1_{248}[X(t)].$$

The results are summarized in Table 1.

Table 1. Fuzzy Chow Test results.

t	Fuzzy Sets $[A_k, A_{k+1}]$	F_{test}	F_{crit} ($\alpha = 0.025$)	Decision
[44,56]	$[A_{48}, A_{52}]$	30.72	4.3828	Reject H_0
[200,212]	$[A_{204}, A_{208}]$	26.52	4.3828	Reject H_0
[240,252]	$[A_{244}, A_{248}]$	39.77	4.3828	Reject H_0

All of the results show that we have sufficient evidence to demonstrate that structural breaks occur in the data in the time intervals stated above.

4.2. When Fuzzy Chow Test Does Not Reject the Null Hypothesis?

Our method for detection of structural breaks is non-statistical. Therefore, we should also check whether the fuzzy Chow test does not falsely accept the null hypothesis about the existence of a structural break in place where our method detects none.

For example, lets us check time intervals $[80, 92]$, $[112, 124]$, $[152, 164]$, $[228, 240]$ characterized by fuzzy sets (A_{84}, A_{88}), (A_{116}, A_{120}), (A_{156}, A_{160}), (A_{232}, A_{236}) where no structural break is detected. We apply Fuzzy Chow test and so, we compute two F-transform components:

$$F^1_{84}[X] = \beta^0_{84}[X] + \beta^1_{84}[X] \cdot (t - c_{84}),$$
$$F^1_{88}[X] = \beta^0_{88}[X] + \beta^1_{88}[X] \cdot (x - c_{88})$$

and form the null hypothesis:

$$H_0 : \beta^0_{84}[X] = \beta^0_{88}[X] \quad \textit{and} \quad \beta^1_{84}[X] = \beta^1_{88}[X]$$

The statistics of the Fuzzy Chow test is

$$F_{test} = \frac{(\mathrm{RSS}_{84,88} - \mathrm{RSS}_{84} - \mathrm{RSS}_{88})/p}{(\mathrm{RSS}_{84} + \mathrm{RSS}_{88})/(2n-2p)} = 3.13$$

where $n = 13$ and $p = 2$.

We may verify that $F_{test} = 3.13 < F_{crit} = F(p, 2n-2p) = 4.3828$ for $\alpha = 0.025$. Hence, we cannot reject the null hypothesis. This means that we have no sufficient evidence to say that a structural break occurs in the interval $[80, 92]$.

Let us apply same algorithm with fuzzy sets (A_{116}, A_{120}), (A_{156}, A_{160}) and (A_{236}, A_{240}) and form the following null hypotheses:

$$\begin{aligned} H_0 :& \beta^0_{116}[X] = \beta^0_{120}[X] \quad \textit{and} \quad \beta^1_{116}[X] = \beta^1_{120}[X], \\ H_0 :& \beta^0_{156}[X] = \beta^0_{160}[X] \quad \textit{and} \quad \beta^1_{156}[X] = \beta^1_{160}[X], \\ H_0 :& \beta^0_{236}[X] = \beta^0_{240}[X] \quad \textit{and} \quad \beta^1_{236}[X] = \beta^1_{240}[X]. \end{aligned}$$

The results are summarized in Table 2.

Table 2. Fuzzy Chow Test results and Decisions.

Time Interval	**Fuzzy Sets** (A_k, A_{k+1})	F_{test}	F_{crit} ($\alpha = 0.025$)	**Decision**
[112,124]	(A_{116}, A_{120})	2.81	4.3828	H_0 is not rejected
[152,164]	(A_{156}, A_{160})	1.46	4.3828	H_0 is not rejected
[232,244]	(A_{236}, A_{240})	1.86	4.3828	H_0 is not rejected

We conclude that the fuzzy Chow test indeed statistically verifies structural breaks only on places in which our method described in Section 3.2 detects them.

4.3. Comparison with Classical Statistical Tests

In this section, we will compare our method with three classical statistical methods for the detection of structural breaks. Namely, we will consider the classical Chow test described in Section 2.1, Pettitt's, and Bai-Perron tests.

4.3.1. Pettitt'S Test (1979)

The Pettitt's test [20,21] is a way to find out when there was a big change in the mean of a set of numbers over time when the exact time of the change is unknown. The data collect n observations, and if it contains a change point at time t, then the change point will split the data into two parts, each of which has a distinct distribution $F_1(x)$ and $F_2(x)$. The test statistic K is defined by $K = \max|U_t|$ where

$$U_t = \sum_{i=1}^{t} \sum_{j=t+1}^{n} \mathrm{sign}(x_i - x_j) \qquad \text{and} \qquad \mathrm{sign}(x_i - x_j) = \begin{cases} 1 & \text{if } (x_i - x_j) > 0, \\ 0 & \text{if } (x_i - x_j) = 0, \\ -1 & \text{if } (x_i - x_j) < 0. \end{cases}$$

The confidence level for n samples is $p = \exp\left(\frac{-K}{n^2+n^3}\right)$. The null hypothesis is rejected if the value of p does not exceed the given confidence level.

4.3.2. Bai-Perron Test (1998)

Another well known test for identification of several structural breaks in a time series is Bai-Perron one [22–24].

With the null hypothesis: There is no structural break in the time series, the Bai-Perron proccess suggests a multiple linear regression with k changes. The change points are $(T_1, \ldots, T_k)$ and are explicitly handled as unknown:

$$\begin{aligned} y_t &= x_t\beta + z_t\sigma_1 + \varepsilon_t, && t = 1, \ldots, T_1, \\ y_t &= x_t\beta + z_t\sigma_2 + \varepsilon_t, && t = T_1 + 1, \ldots, T_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_t &= x_t\beta + z_t\sigma_{k+1} + \varepsilon_t, && t = T_{k+1}, \ldots, T \end{aligned}$$

where y_t is the dependent variable, β and σ are coefficients corresponding to covariates x_t and z_t, and u_t is the residue term. The purpose is to estimate the unknown regression coefficients and the break dates $(\beta, \sigma_1, \ldots, \sigma_{k+1}, T_1, \ldots, T_k)$ when T observations on (y_t, x_t, z_t) are given.

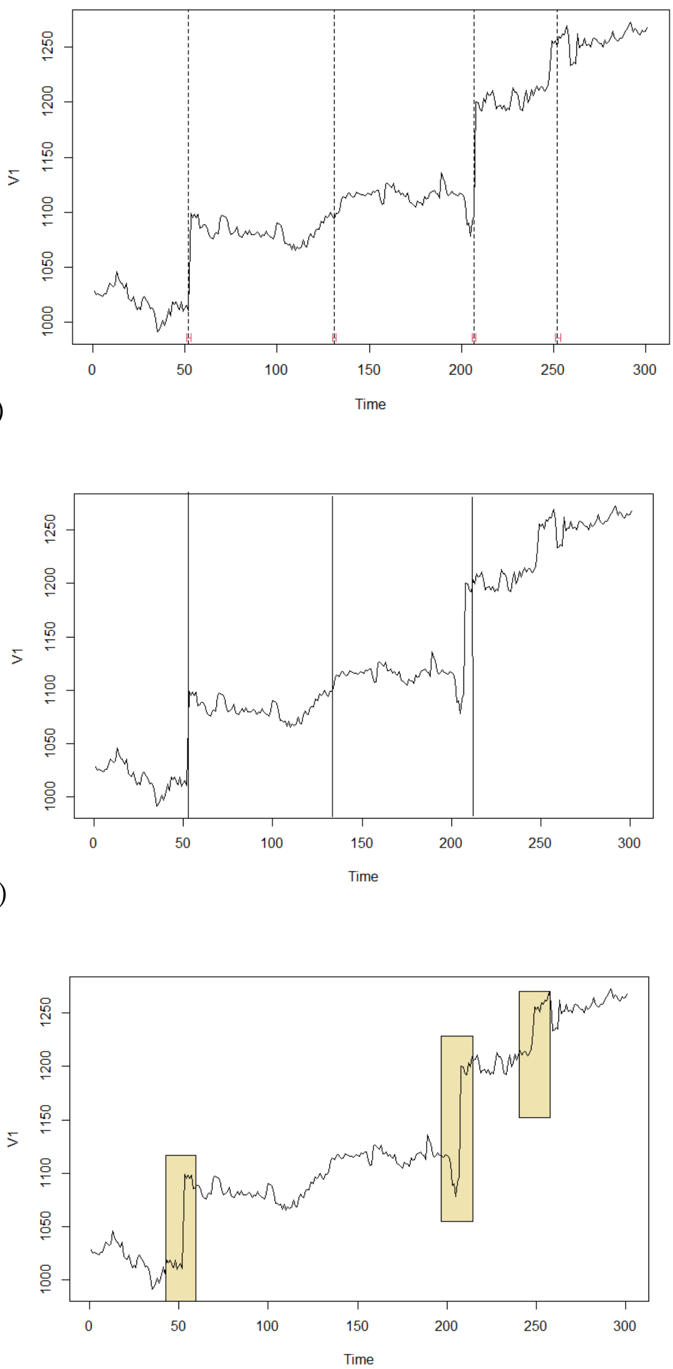
4.3.3. Comparison of Statistical Methods with Fuzzy Chow Test

The results of the comparison can be seen in Figure 3. The computations have been performed using R. The classical Chow test detected change points in $t = 51, 130, 206, 251$ and Bai-Perron test in $t = 52, 131, 207, 252$. Thus, the differences in the graph in Figure 3 are invisible. The Pettitt test found only three points in $t = 51, 130, 206$.

Our method detected three structural breaks in the intervals $[44, 56]$, $[200, 212]$, $[240, 252]$ which, as shown in Section 4.1, are also statistically significant by the fuzzy Chow test. Note that these intervals cover the points found by Chow and Bai-Perron tests with the exception of $t = 130$ ($t = 131$) where, however, there is no structural break. We argue that, as is apparent from Figure 3, our method correctly detected real structural breaks occurring in the time series. The following conclusions imply from the comparison with our method:

- Statistical methods detect the breaks as if occurring in one time point. However, it is clear from the practice that this rarely happens; the break usually shows in a wider area. Hence, the result of a statistical test should be taken as a point *around which* the break occurs. But this is equal to what our fuzzy method provides explicitly, and thus we argue that more realistically.
- Statistical methods are less robust and thus, more sensitive to the data. Hence, they are prone to the detection of non-existing breaks. In our example, this happened to all three statistical methods in point $t = 130$ ($t = 131$) where apparently there is no structural break. Let us emphasize that our fuzzy method ignored this point/area due to its robustness.
- Statistical methods require optimization and therefore, they are computationally expensive. Our method is very simple and computationally effective since the fuzzy transform has linear complexity.

It is important to emphasize that our goal is not to beat statistical methods. Vice-versa, we want to employ maximally the synergy of both approaches.



(**a**)

(**b**)

(**c**)

Figure 3. Results of statistical tests as well as of our method to finding structural breaks in the real time series: (**a**) Chow and Bai-Perron tests (in the graph, there is no visible difference between them), (**b**) Pettitt's test, and (**c**) our method verified by the fuzzy Chow test.

5. Conclusions

In this paper, we focus on methods for the detection of structural breaks in time series. In [7], we suggested a method based on the fuzzy transform (F-transform) and selected methods of Fuzzy Natural Logic. In this paper, we suggested a combination of our method with a statistical text that stems from the well-known Chow test [1]. Statistical significance is demonstrated on the data. We have shown that our technique is effective in spotting abnormalities and gives statistically significant results. Further research will be focused on the combination of our fuzzy method with other kinds of statistical testing.

Author Contributions: Methodology, V.N.; Software, computations T.T.P.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

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

References

1. Chow, G.C. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* **1960**, *28*, 591–605. [CrossRef]
2. De Wachter, S.; Tzavalis, D. Detection of structural breaks in linear dynamic panel data models. *Comput. Stat. Data Anal.* **2012**, *56*, 3020–3034. [CrossRef]
3. Fischer, P.; Hilbert, A. Fast detection of structural breaks. In Proceedings of the 21th International Conference on Computational Statistics, Geneva, Switzerland, 19–22 August 2014; pp. 9–16.
4. Nielsen, B.; Whitby, A. A Joint Chow Test for Structural Instability. *Econometrics* **2015**, *3*, 156–186. [CrossRef]
5. Preuss, P.; Puchstein, R.; Detter, H. Detection of multiple structural breaks in multivariate time series. *J. Am. Stat. Assoc.* **2015**, *110*, 654–668. [CrossRef]
6. Doerr, B.; Fischer, P.; Hilbert, A.; Witt, C. Detecting structural breaks in time series via genetic algorithms. *Soft Comput.* **2017**, *21*, 4707–4720. [CrossRef]
7. Truong, P.; Novák, V. An Improved Forecasting and Detection of Structural Breaks in Time Series using Fuzzy Techniques. In *Theory and Applications of Time Series Analysis and Forecasting*; Valenzuela, O., Rojas, F., Herrera, L., Pomares, H., Rojas, I., Eds.; Springer: Berlin/Heidelberg, Germany, 2022.
8. Perfilieva, I.; Daňková, M.; Bede, B. Towards a Higher Degree F-transform. *Fuzzy Sets Syst.* **2011**, *180*, 3–19. [CrossRef]
9. Novák, V.; Perfilieva, I.; Dvořák, A. *Insight into Fuzzy Modeling*; Wiley & Sons: Hoboken, NJ, USA, 2016.
10. Perfilieva, I. Fuzzy Transforms: Theory and applications. *Fuzzy Sets Syst.* **2006**, *157*, 993–1023. [CrossRef]
11. Novák, V.; Mirshahi, S.; Pavliska, V. LFL Forecaster: Analysis, Forecasting and Mining Information from Time Series. In Proceedings of the 2019 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2019, New Orleans, LA, USA, 23–26 June 2019; pp. 1–6. [CrossRef]
12. Anděl, J. *Statistical Analysis of Time Series*; SNTL: Praha, Czech Republic, 1976. (In Czech)
13. Bovas, A.; Ledolter, J. *Statistical Methods for Forecasting*; Wiley: New York, NY, USA, 2003.
14. Kedem, B.; Fokianos, K. *Regression Models for Time Series Analysis*; Wiley: New York, NY, USA, 2002.
15. Hamilton, J. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.
16. Nguyen, L.; Novák, V.; Holčapek, M. Gold Price: Trend-cycle Analysis Using Fuzzy Techniques. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part III*; Lesot, M.-J., Vieira, S., Reformat, M.Z., Carvalho, J.P., Wilbik, A., Bouchon-Meunier, B., Yager, R.R., Eds.; Springer Nature: Cham, Switzerland 2020; pp. 254–266.
17. Novák, V. Fuzzy vs. Probabilistic Techniques in Time Series Analysis. In *Econometrics for Financial Applications*; Anh, L., Dong, L., Kreinovich, V., Thach, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 213–234.
18. Nguyen, L.; Holčapek, M. Suppression of High Frequencies in Time Series Using Fuzzy Transform of Higher Degree. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems: 16th International Conference, IPMU 2016*; Carvalho, J., Lesot, M.-J., Kaymak, U., Vieira, S., Bouchon-Meunier, B., Yager, R.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; Volume 2, pp. 705–716.
19. Burda, M.; Štěpnička, M. lfl: An R Package for Linguistic Fuzzy Logic. *Fuzzy Sets Syst.* **2022**, *431*, 1–38. [CrossRef]
20. Pettitt, A. A non-parametric approach to the change-point problem. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1979**, *28*, 126–135. [CrossRef]
21. Conte, L.; Bayer, D.; Bayer, F. Bootstrap Pettitt test for detecting change points in hydroclimatological data: Case study of Itaipu Hydroelectric Plant, Brazil. *Hydrol. Sci. J.* **2019**, *64*, 1312–1326. [CrossRef]
22. Bai, J.; Perron, P. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* **1998**, *66*, 47–78. [CrossRef]
23. Bai, J.; Perron, P. Computation and Analysis of Multiple Structural Change Models. *J. Appl. Econom.* **2003**, *18*, 1–22. [CrossRef]
24. Hall, A.R.; Han, S.; Boldea, O. Inference regarding multiple structural changes in linear models with endogenous regressors. *J. Econom.* **1998**, *170*, 281–302. [CrossRef] [PubMed]