Detection of Structural Breaks in Time Series Using Fuzzy Techniques

Vilém Novák
Institute for Research and Applications of Fuzzy Modeling, NSC IT4Innovations, University of Ostrava, Ostrava, Czech Republic

Abstract

In this paper we suggest to use special “fuzzy” techniques for detection of structural breaks in time series, namely the fuzzy (F)-transform and one method of fuzzy natural logic (FNL). The idea is based on application of the $F^1$-transform which makes it possible to estimate effectively slope of time series (ignoring its possible volatility) and its evaluation by a suitable evaluative linguistic expression. The method is computationally very effective.

Keywords: Fuzzy transform, F-transform, Evaluative linguistic expressions, Fuzzy natural logic

1. Introduction

This paper addresses interesting problem related to the area of mining information from time series (cf. [1]), namely detection of structural breaks in the latter. These are unexpected shifts in the course of the time series that can be caused, e.g., by economic development, global shifts in capital and labor, changes in resource availability due to war or natural disaster, discovery or depletion of natural resources, or a change in political system, etc. From the general point of view, a structural change is possible because of the dynamic nature of the economic system and it occurs when an industry or market changes the way how it functions or operates. Structural breaks are quite frequent and concern not only time series analysis but also general signal processing, for example searching faults in a wire, irregularities in heart functioning, etc.

There are a lot of methods for detection of structural breaks in time series. Most of them are statistical, for example [2–4]. There are also non-statistical approaches, e.g., [5] where the authors suggest to use genetic algorithms.

In this paper we suggest to use special “fuzzy” techniques for detection of structural breaks in time series, namely the fuzzy (F)-transform and one method of fuzzy natural logic (FNL). Application of these techniques to the analysis and forecasting of time series was described in several papers and also in the book [6]. The first application of the F-transform to detection of changes in a signal was described in [7] where the problem was to recognize network attacks. Let us recall that the F-transform makes it possible to analyze time series, extract its trend and trend-cycle, and to detect specific areas, for example intervals of monotonous behavior. The FNL stems from these results and provides methods for forecasting of the future behavior of time series, automatic generation of linguistic description of specific parts of it.
and for obtaining other kinds of information.

One of the important outcomes of the F-transform is its ability to estimate derivatives of the time series (taken as a function) over a larger area that can be specified imprecisely, and in which the volatility is ignored. Then, using methods of FNL, we can generate comments in natural language to various characteristic features of the time series (cf. [8, 9]).

In this paper, we suggest methods how the above mentioned techniques can be used for detection of the structural breaks in time series. We will apply estimation of the first derivative and the theory of evaluative linguistic expressions. The main outcome of this method is its relative simplicity, transparency and computation effectiveness because the complexity of the F-transform is linear.

The structure of the paper is the following. In Section 2 we will briefly overview the main principles of the fuzzy transform with the focus on those parts that are used in our paper. In Section 3 we will mention FNL and especially the theory of evaluative linguistic expressions that is applied in the sequel. In Section 4 we describe how the F-transform is applied in the analysis of time series. Finally, Section 5 is devoted to description of the method for finding structural breaks in time series. The paper is finished by demonstration of the method on four concrete examples of structural breaks and its comparison with statistical methods.

By a fuzzy set $A \subseteq U$ we understand a function $A : U \rightarrow [0, 1]$ where $U$ is an ordinary set called the universe. The $[0, 1]$ is the support of a suitable algebra. In this paper we will consider the standard Łukasiewicz MV-algebra (see [10, 11]).

2. The Principle of F-Transform

The fuzzy F-transform is a universal approximation technique introduced in [12, 13]. Its fundamental idea is to map a bounded continuous function $f : [a, b] \rightarrow \mathbb{R}$ to a finite vector $F[f]$ of components and then to transform it back. The former is called the direct F-transform and the latter the inverse one. The result of the inverse F-transform is a function $\hat{f}$ that approximates the original function $f$. The parameters of the F-transform can be set in such a way that the approximating function $\hat{f}$ has desired properties.

The power of the F-transform stems from the proof of its approximation power, its ability to filter out high frequencies and reduce noise [14], and the possibility to estimate average values of derivatives over a given imprecisely specified area [15]. This property is especially useful in mining information from time series.

2.1 Fuzzy Partition

The first step of the F-transform procedure is to form a fuzzy partition of the domain $[a, b]$. It consists of a finite set of fuzzy sets

$$A = \{A_0, \ldots, A_n\}, \quad n \geq 2,$$

defined over nodes

$$a = c_0, \ldots, c_n = b.$$  

The properties of the fuzzy sets from $A$ are specified by five axioms, namely: normality, locality (bounded support), continuity, unimodality, and orthogonality. The support of $k$-the fuzzy set $A_k$, $k = 1, \ldots, n - 1$, is formed by the interval $[c_{k-1}, c_k] \cup [c_k, c_{k+1}]$. A fuzzy partition $A$ is called $h$-uniform if the nodes $c_0, \ldots, c_n$ are $h$-equidistant, i.e., for all $k = 0, \ldots, n - 1$, $c_{k+1} = c_k + h$, where $h = (b - a)/n$ and the fuzzy sets $A_1, \ldots, A_{n-1}$ are shifted copies of a generating function $A : [-1, 1] \rightarrow [0, 1]$ such that for all $k = 1, \ldots, n - 1$

$$A_k(x) = A\left(\frac{x - x_k}{h}\right), \quad x \in [c_{k-1}, c_{k+1}]$$

(for $k = 0$ and $k = n$ we consider only half of the function $A$, i.e. restricted to the interval $[0, 1]$ and $[-1, 0]$, respectively). Membership functions $A_0, A_1, \ldots, A_n$ of the fuzzy sets forming the fuzzy partition $A$ are usually called basic functions.

The orthogonality is formally defined as

$$\sum_{i=0}^{n} A_i(x) = 1, \quad x \in [a, b]$$

(Equation (3) is sometimes called Ruspini condition).

2.2 Zero Degree F-Transform

Once the fuzzy partition $A_0, \ldots, A_n \in A$ is determined, we define a direct F-transform of a continuous function $f$ as a vector $F[f] = (F_0[f], \ldots, F_n[f])$, where each $k$-th component $F_k[f]$ is equal to

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

Clearly, the $F_k[f]$ component is a weighted average of the functional values $f(x)$ where weights are the membership degrees.
\( A_k(x) \). The inverse F-transform of \( f \) with respect to \( F[f] \) is a continuous function \( \hat{f} : [a, b] \rightarrow \mathbb{R} \) such that
\[
\hat{f}(x) = \sum_{k=0}^{n} F_k[f] \cdot A_k(x), \quad x \in [a, b].
\]

**Theorem 1.** The inverse F-transform \( \hat{f} \) has the following properties:

- The sequence of inverse F-transforms \( \{ \hat{f}_n \} \) determined by a sequence of uniform fuzzy partitions based on uniformly distributed nodes with \( h = (b - a)/n \) uniformly converges to \( f \) for \( n \rightarrow \infty \).
- The F-transform is linear, i.e., if \( f = \alpha u + \beta v \) then \( \hat{f} = \alpha \hat{u} + \beta \hat{v} \).

All the details and full proofs can be found in [12][13].

### 2.3 Higher Degree F-Transform

The F-transform introduced above is \( F^0 \)-transform (i.e., zero-degree F-transform). Its components are real numbers. If we replace them by polynomials of arbitrary degree \( m \geq 0 \), we arrive at the higher degree \( F^m \) transform. This generalization has been in detail described in [13]. Let us remark that the \( F^1 \) transform enables to estimate also derivatives of the given function \( f \) as weighted average values over a specified area.

The direct \( F^1 \)-transform of \( f \) with respect to \( A_1, \ldots, A_{n-1} \) is a vector \( F^1[f] = (F^1_1[f], \ldots, F^1_{n-1}[f]) \) where the components \( F^1_k[f], k = 1, \ldots, n - 1 \) are linear functions
\[
F^1_k[f](x) = \beta^0_k + \beta^1_k(x - c_k)
\]
(5)

with the coefficients \( \beta^0_k, \beta^1_k \) given by
\[
\beta^0_k = \frac{\int_{c_{k-1}}^{c_k} f(x) A_k(x) dx}{\int_{c_{k-1}}^{c_k} A_k(x) dx} , \quad \beta^1_k = \frac{\int_{c_{k-1}}^{c_k} (x - c_k) A_k(x) dx}{\int_{c_{k-1}}^{c_k} A_k(x) dx} .
\]
(6)\( (7)

Note that \( \beta^0_k = F_k[f] \), i.e. the coefficients \( \beta^0_k \) are just the components of the \( F^0 \) transform given in (4). The \( F^1 \)-transform has also the properties stated in Theorem 1 (see [13]).

We may also use the \( F^2 \)-transform. Its components are the functions
\[
F^2_k[f](x) = \beta^0_k + \beta^1_k(x - c_k) + \left( \beta^2_k(x - c_k)^2 - \frac{h^2}{6} \right)
\]
(provided that the basic functions are triangles) and the coefficient \( \beta^2_k \) is estimation of the second derivative.

**Theorem 2.** If \( f \) is four-times continuously differentiable on \([a, b]\) then for each \( k = 1, \ldots, n - 1 \),
\[
\beta^0_k = f(c_k) + O(h^2), \quad \beta^1_k = f'(c_k) + O(h^2), \quad \beta^2_k = \frac{f''(c_k)}{2} + O(h^2).
\]
(8)\( (9)\)\( (10)

Thus, each F-transform component provides a weighted average of values of the function \( f \) in the area around the node \( c_k \) (8), and also a weighted average of slopes (9) of \( f \) and that of its second derivatives (10) in the same area. Let us remark that only the nodes \( c_1, \ldots, c_{n-1} \) should be considered when dealing with the F-transform.

### 3. Fuzzy Natural Logic

#### 3.1 Definition of FNL

The fuzzy natural logic is a class of formal logical theories developed on unique grounds whose goal is to develop a mathematical model of special human reasoning schemes that employ natural language. So far, the following theories are included:

- A formal theory of evaluative linguistic expressions [16].
- A formal theory of fuzzy/linguistic IF-THEN rules and approximate reasoning [17][20].

- A formal theory of intermediate and generalized fuzzy quantifiers [21][25].

In this paper, we will apply only the theory of evaluative linguistic expressions.

**Remark 1.** Of course, the fuzzy IF-THEN rules or fuzzy and generalized quantifiers are studied by many authors, for example [26][27] or [28][31]. Though a lot of their results are source of inspiration for the development of FNL, they lack a sound formal theory and so, we do not consider them as part of FNL.

#### 3.2 Evaluative Linguistic Expressions

Evaluative linguistic expressions are special expressions of natural language that have the general form
\[
\langle \text{linguistic hedge} \rangle \langle \text{TE-adjective} \rangle
\]
(11)
where \( \langle \text{TE-adjective} \rangle \) is one of the canonical adjectives “small, medium, big”, or “zero” and possibly also a symmetric fuzzy number. Instead of the canonical adjectives, we can use other kinds of adjectives having similar nature (they belong to specific classes such as gradable or evaluative adjectives), for example “shallow, medium deep, deep”, “weak, medium strong, strong”, etc. The \( \langle \text{linguistic hedge} \rangle \) is a special expression that makes the meaning of the \( \langle \text{TE-adjective} \rangle \) more specific. Quite often it is represented by an intensifying adverb such as “very, roughly, approximately, significantly”, etc. Linguistic hedges can have narrowing (“extremely, significantly, very, typically”) and widening effect (“more or less, roughly, quite roughly, very roughly”) on the meaning of the \( \langle \text{TE-adjective} \rangle \).

If \( \langle \text{linguistic hedge} \rangle \) is not present (expressions such as “weak, large”, etc.) then we take it as presence of empty linguistic hedge. Thus, all the simple evaluative expressions have the same form \((11)\). Since they characterize values on an ordered scale, we may consider also scales divided into two parts that are usually interpreted as positive and negative. Hence, the evaluative expressions may have also a sign, namely “positive” or “negative”. Simple evaluative expressions of the form \((11)\) can also be combined using logical connectives (usually “and” and “or”) to obtain compound ones but we do not consider them in this paper. The set of all (considered) evaluative expressions is denoted by \( \text{EvExpr} \).

We distinguish abstract evaluative expressions from more specific evaluative predications. The latter are expressions of natural language of the form ‘\( X \) is \( A \)’ where \( A \) is an evaluative expression and \( X \) is a variable which stands for objects, for example “degrees of temperature, height, length, speed”, etc. Examples are “temperature is high”, “speed is extremely low”, “medium value and etc. The meaning of the \( X \) is \( A \)”.

When using evaluative expressions, we must first specify the linguistic context. In our theory it is an interval \([v_L, v_S] \cup [v_S, v_R]\) determined by a triple of numbers \( w = \langle v_L, v_S, v_R \rangle \) where \( v_L \) is the leftmost typically small value, \( v_S \) is typically medium value and \( v_R \) is the rightmost typically big value. For example, when speaking about temperature of water, we may set \( v_L = 15^\circ\text{C} \) (water temperature in the crane), \( v_S = 50^\circ\text{C} \) and \( v_R = 100^\circ\text{C} \). In the sequel, we will consider a set of all linguistic contexts

\[
W = \{ w = \langle v_L, v_S, v_R \rangle \mid v_L, v_S, v_R \in \mathbb{R}, v_L < v_S < v_R \}. \quad (12)
\]

The element \( x \) belongs to a context \( w \in W \) if \( x \in [v_L, v_R] \) and we may write \( x \in w \).

The meaning of an evaluative linguistic expression or predication is represented by its intension

\[
\text{Int}(X \text{ is } A) : W \rightarrow \mathcal{F}(\mathbb{R}) \quad (13)
\]

where \( \mathcal{F}(\mathbb{R}) \) is a set of all fuzzy sets on \( \mathbb{R} \). This is a function that to each context \( w \in W \) assigns the extension \( \text{Ext}_w(X \text{ is } A) \) where the latter is a specific fuzzy set on \( \mathbb{R} \). Example of extensions of several evaluative linguistic expressions is in Figure 1.

Let us emphasize that their shapes have been established on the basis of logical analysis of the meaning of the corresponding evaluative expressions (for the details, see [16]).

Moreover, there is a natural ordering \( \ll \) of these expressions obtained as lexicographic ordering based on the natural ordering \( Ze \) (zero) \( \ll Sm \) (small) \( \ll Me \) (medium) \( \ll Bi \) (big) and the following ordering of hedges:

\[
Ex \text{ (extremely)} \ll Si \text{ (significantly)} \ll Ve \text{ (very)}
\]

\[\ll \text{ (empty hedge)} \ll ML \text{ more or less} \ll QR \text{ (quite roughly)} \ll VR \text{ (very roughly)} \ll VV \text{ (very very roughly)}. \quad (14)
\]

The linear lexicographic ordering \( \ll \) of simple evaluative expressions \((11)\) can be interpreted as “to be more precise” (or, “to have a sharper meaning”).

Figure 1. Shapes of extensions of few evaluative expressions in the context \( (0, 0.5, 1) \). The hedges are \{Extremely, Significantly, Very, empty hedge\} for “small” and “big” and \{More-or-Less, Roughly, Quite Roughly, Very Roughly\} for “small”, and “big”. Extension of “medium” is not modified but it can be modified by widening hedges.
The function due to the following definition is needed in the sequel.

**Definition 1.** The function of local perception is a partial function $L_{\text{Perc}} : \mathbb{R} \times W \rightarrow \text{EvExpr}$ defined as follows:

$$L_{\text{Perc}}(u, w) = \begin{cases} X & \text{if } u \in w, \\ \text{undefined} & \text{otherwise.} \end{cases} \quad (15)$$

If (15) is defined, then “$X$ is $A$” is the sharpest evaluative predication in the sense of the lexicographic ordering $\ll$, so that $u \in w \subset \mathbb{R}$ is the most typical element for the extension $\text{Ext}_w(X) = A$.

4. Time Series and F-Transform

4.1 Decomposition of Time Series

A time series is a stochastic process (see [32,33]) $X : Q \times \Omega \rightarrow \mathbb{R}$ where $\Omega$ is a set of elementary random events and $Q = \{0, \ldots, p\} \subset \mathbb{N}$ is a finite set whose elements are interpreted as time moments. Statistical models assume that each $X(t), t \in Q$ is a random variable having a specific distribution function. This assumption is the basis of many kinds of probabilistic methods for analysis and forecasting of time series.

Fuzzy techniques stem from the following decomposition model:

$$X(t) = Tr(t) + C(t) + S(t) + R(t), \quad t \in \mathbb{T}, \quad (16)$$

where $Tr$ is the trend, $C$ is a cyclic component, $S$ is a seasonal component that is a mixture of periodic functions. Note that $Tr, C, S$ are ordinary functions not being stochastic. Only $R$ is a random noise, i.e., a sequence of independent random variables $R(t)$ such that for each $t \in \mathbb{T}$, the $R(t)$ has zero mean and finite variance. Let us remark that trend and cyclic component are quite often joined into one component $TC(t) = Tr(t) + C(t)$ called trend-cycle.

Let us now assume (without loss of generality) that the time series (16) contains periodic subcomponents with frequencies $\lambda_1 < \cdots < \lambda_r$. These frequencies correspond to periodicities

$$T_1 > \cdots > T_r, \quad (17)$$

respectively (via the equality $T = 2\pi/\lambda$).

The following theorem demonstrates the power of the F-transform for time series analysis.

**Theorem 3.** Let $X(t)$ be realization of the stochastic process (16). Let us assume that all subcomponents with frequencies $\lambda$ lower than $\lambda_q$ are contained in the trend-cycle $TC$. If we construct a fuzzy partition over the set of equidistant nodes $\mathbb{N}$ with the distance $h = dT_q$ where $d \in \mathbb{N}$ and $T_q$ is a periodicity corresponding to $\lambda_q$ then the corresponding inverse F-transform $\hat{X}$ of $X(t)$ gives the following estimation of the trend-cycle:

$$|\hat{X}(t) - TC(t)| \leq 2\omega(h, TC) + D \quad (18)$$

for $t \in [c_1, c_{n-1}]$, where $D$ for $d \geq 2$ is a certain small number and $\omega(h, TC)$ is a modulus of continuity of $TC$ w.r.t. $h$.

The precise form of $D$ and the detailed proof of this theorem can be found in [14,34]. The theorem holds both for $F^0$ as well as for $F^1$-transform. It follows from it that the F-transform makes it possible to filter out frequencies higher than a given threshold and also reduces the noise $R$.

The periodicity $T_q$ can be found using the well known periodogram — see [32,33]. Then, by setting a proper fuzzy partition due to Theorem 3, we first compute the F-transform of $X(t)$ (either zero or first degree)

$$F[X] = (F_1[X], \ldots, F_n[X]).$$

Then the estimation of the trend-cycle $TC$ is obtained using the inverse F-transform: $TC(t) \approx \hat{X}(t)$.

Theorem 3 tells us how the distance between nodes of the fuzzy partition should be set to obtain good estimation of the trend-cycle. Namely, we put it equal to $h = dT_q$ where $d$ is in practice equal to 1 or 2. The fuzzy partition is constructed over an interval $[0, p]$ of real numbers, i.e., it is not constructed over the discrete set of natural numbers. Consequently, we obtain $\hat{X} \approx TC$ which means that we can estimate the trend cycle $TC$ with high fidelity. Of course, the estimation depends on the course of $TC$ and it is the better the smaller is the modulus of continuity $\omega(h, TC)$ (which is usually the case).

4.2 Characterization of Trend Using Expressions of Natural Language

In this section, we will briefly describe the principles of linguistic evaluation of the (direction of) trend of time series in a local time slot (for example, quarter of year, production period, etc.). Such evaluation is later used in the detection of structural breaks.

The trend is estimated using the coefficients $\hat{B}_k$, determined in (7). Their values over the areas characterized by the respective
fuzzy sets $A_k$ (from the fuzzy partition $\bar{\Pi}$) are evaluated by a specific evaluative expressions. The procedure, that is based on the theories of F-transform and FNL, is outlined below and in detail described in [6,9]. Note that a similar problem has been discussed also by Kacprzyk et al. [35]. Their approach is based on a construction of a properly inclined line and selection of several angles that might represent specific inclinations of the trend line. The whole procedure, however, is purely heuristic and is not well justified mathematically.

First, we must specify a context in which the (direction of) the trend is evaluated. This is clear because, for example, the increase of temperature in Sweden by $3^\circ \text{C}$ in, say, 2 hours can be taken as a “sharp increase”, while the same in Africa is a “very small increase”. Therefore, must specify the three distinguished values $v_L,v_S,v_R$ of the tangent where $v_R$ is the extreme increase (decrease), while the smallest one is typically (but not always) $v_L = 0$. The typical medium value $v_S$ is determined analogously as $v_R$. The result is the linguistic context $w_{tg} = [v_L,v_S,v_R]$ for the trend. Moreover, because we distinguish between increase and decrease of time series, we must also distinguish positive context $w^+_{tg} = [v^+_L,v^+_S,v^+_R]$ from the negative one $w^-_{tg} = [v^-_R,v^-_S,v^-_L]$. We usually put $v^-_L = v^+_S = 0$.

The trend can be evaluated by the following sentence:

\[
\text{Trend is } \langle \text{direction} \rangle \quad (19)
\]

where

\[
\langle \text{direction} \rangle := \text{stagnating} \langle \text{special hedge} \rangle \langle \text{sign} \rangle, \quad (20)
\]

\[
\langle \text{sign} \rangle := \text{increasing} | \text{decreasing} \quad (21)
\]

and

\[
\langle \text{special hedge} \rangle := \emptyset | \text{negligibly} | \text{slightly} | \text{somewhat} | \text{clearly} | \text{roughly} | \text{sharply} | \text{significantly}.
\]

The empty hedge $\emptyset$ is used if no more specific characterization of the direction of the trend is required. Note that (19) is not the only possibility. Other possibilities are described in [6].

The evaluative predication (19) is translated into the following standard form:

\[
\text{Trend of } X | \bar{\Pi} \text{ is } \pm Ev[X | \bar{\Pi}] \quad (22)
\]

where $Ev[X | \bar{\Pi}]$ is an evaluative expression in which the TE-adjective is canonical (i.e., either of “small, medium, big” or

Table 1. The translation table between special and canonical evaluative expressions

| Special hedge, direction | $Ev[X | \bar{\Pi}]$ from (22) |
|--------------------------|-------------------------------|
| stagnating               | extremely small or zero       |
| negligibly               | significantly small           |
| slightly                 | very small                    |
| somewhat                | rather small                  |
| clearly                  | very roughly small or medium  |
| roughly                 | very roughly big              |
| fairly large             | roughly big                   |
| quite large              | rather big                    |
| large                    | big                           |
| sharply                  | very big                      |
| significantly            | significantly big             |
| huge                     | extremely big                 |

Figure 2. Artificial time series with four types of structural breaks. “zero”). This leads to the following definition.

Definition 2. Let $X$ be a time series and $\bar{\Pi} \subset \Pi$ be a time interval. Let $\beta^1[X | \bar{\Pi}]$ be coefficient (7) that provides estimation of the slope of trend of $X$ in the area determined by a fuzzy set $\bar{\Pi}$ with the support $\bar{\Pi}$. Finally, let $w^+_{tg},w^-_{tg}$ be the corresponding
context. Then, the evaluative expression \( \pm Ev[X|\bar{T}] \) in (22) is obtained using the function of local perception as follows:

\[
\pm Ev[X|\bar{T}] := LPer(x^1, w_{tg} \cup w_{tg}^+).
\] (23)

Note that (23) characterizes trend over the time period \( \bar{T} \). Translation of the evaluative expressions (23) into the expressions due to (19), (20) and (21) is in Table 1.

From (23) we can construct extension in the context of either \( w_{tg}^+ \) or \( w_{tg}^- \), depending on the sign of the evaluative expression (23).

\[
Ext_{w_{tg}}(\text{ Trend of } X|\bar{T} \text{ is } \langle \text{direction} \rangle) \subseteq \mathbb{R}.
\] (24)

Recall that (24) is a fuzzy set of values of the slopes of \( X|\bar{T} \) estimated by (7). More details can be found in [6].

5. Detection of Structural Breaks in Time Series

5.1 The Principle of Detection of Structural Breaks

The fundamental idea for detection of structural breaks consists in the application of the \( F^1 \)-transform over properly defined fuzzy partition. Note that the supports of the basic functions cannot be too wide because then the structural breaks might disappear among other values of the time series \( X(t) \) (cf. Theorem 3).

**Definition 3.** Let the realization \( \{X(t) \mid t \in \bar{T}\} \) of the time series \( X \) be given. Let \( w_{tg} \) be a context for the trend of \( X \). A structural break is the sudden change of the values \( X(t) \) in an interval \( \bar{T} \) such that \( \pm Ev[X|\bar{T}] \) in (22) is “\( \pm \) very big” or bigger w.r.t. ordering \( \ll \) defined in (14).

As mentioned, the interval \( \bar{T} \) is supposed to be “short”. According to the experiments, its proper value should be \( |\bar{T}| \in \{4, \ldots, 10\} \) (by \( |\cdot| \) we denote length of the corresponding interval).

The method for detection of structural breaks using fuzzy modeling methods is the following:

1. Specify the distance \( h \) between nodes of the fuzzy partition \( A \) (cf. Subsection 2.1). This means that width of the basic functions is equal to \( 2h \). Recall that the fuzzy partition is constructed over the interval \( \bar{T} \).

2. Set the context \( w_{tg} \) for evaluation of the trend in areas determined by the basic functions due to item 1. This can be determined as follows: \( v_L = 0, v_R = (\pm \sigma)/(2h) \) where \( \sigma \) is the standard deviation of the whole time series. Note that it covers both normal course of the time series as well as parts of the structural breaks. Alternatively, we can set \( v_R \) on the basis of the knowledge of the character of time series or, replace \( \sigma \) by:

\[
r = |\max_{t \in \bar{T}} X(t) - \min_{t \in \bar{T}} X(t)|.
\]

3. Compute the \( F^1 \)-transform \( F^1[f] = (F^1_1[X], \ldots, F^1_{n-1}[X]) \) over the fuzzy partition determined in item 1.

4. Locate all components \( F_k[X] \), where the coefficient \( \beta^1_k \) in (7) is characterized as fairly large increase (decrease), or larger w.r.t. ordering \( \ll \) determined in (14), and the context \( w_{tg} \) from step 2. In other words, this means that

\[
LPer_{\beta^1_k}(X, w_{tg} \cup w_{tg}^+) \in \{Ro Bi, Ra Bi, ML Bi, Ve Bi, Si Bi, Ex Bi\}
\]

To make the location more convincing, it is possible to shift the corresponding basic function(s) \( A_k \) slightly to the left or right. Let us emphasize, however, that the shift can be only little and, as the whole procedure is very robust, we do not need to find the maximal value of \( \beta^1_k[X] \).

5. The located components point to the areas of structural breaks in the time series. The areas are determined by the corresponding fuzzy sets (basic functions) \( A_k \in A \) from the fuzzy partition (1).

5.2 Demonstration

The method presented in the previous subsection is demonstrated on examples. We prepared 4 artificial time series with typical structural breaks — see Figure 2. The time series were obtained by combination of few real time series taken from the INDUSTRY subset of time series on a monthly basis from the M3-Competition published on the Internet.

5.2.1 Detection using fuzzy modeling methods

Each time series in Figure 3 is depicted together with the fuzzy partition (Figure 3 was prepared using the experimental software FT-studio which makes it possible to apply fuzzy transform to a function given either by a precise formula, or by the data. The program was developed in the Institute for Research of Applications of Fuzzy Modeling of the University.
The standard deviations are $\pm 57$.

Figure 3. Artificial time series from Figure 2 with fuzzy partition and marked basic functions detecting structural breaks. Note that the fuzzy partition is constructed over the interval $[0, q]$. The evaluative expression over the interval determined by the marked basic functions is always huge increase (decrease). The green line is approximation of the time series using the inverse F-transform.

One can see that the places of structural breaks are well detected. Moreover, we also immediately know whether it is jump up or down (in correspondence with the sign of $\beta^1$):

(a) There are 2 structural breaks determined by 3 basic functions. The values of the corresponding coefficients are $\beta^1_{10} = 813$, $\beta^1_{19} = -592$ and $\beta^1_{20} = -465$ (from left-to-right) with peaks at the time moments $t = 41, 77, 81$, respectively. It is clear from the signs of $\beta^1$ that the structural break is shift up in the area around $t = 41$ and shift down at $t \in \{77, 81\}$.

Linguistic evaluation of the course of time series between the structural breaks (the interval $[41, 77]$) is “stagnating” (Evaluation of the course of time series was obtained using the software LFL Forecaster developed in the Institute for Research of Applications of Fuzzy Modeling of the University of Ostrava, Czech Republic. Its author is Viktor Pavliska).

(b) There are 4 structural breaks determined by 5 basic functions. The values of the corresponding coefficients are $\beta^1_4 = 1430$, $\beta^1_5 = 2299$, $\beta^1_{11} = -3898$, $\beta^1_{18} = 2408$ and $\beta^1_{27} = -1842$ with peaks at the time moments $t = 16, 20, 44, 52, 108$, respectively.

Linguistic evaluations of the course of time series between the structural breaks are in both intervals $[20, 44]$ and $[52, 108]$ “stagnating”.

(c) There are 2 narrow structural breaks, each determined by 2 basic functions, one for shift up and the second one for shift down. The values of the corresponding coefficients are $\beta^1_{10} = 616$, $\beta^1_{11} = -695$, $\beta^1_{27} = 762$ and $\beta^1_{28} = -628$ with peaks at the time moments $t = 40, 44, 108, 112$, respectively.

(d) There are 2 structural breaks determined by 2 basic functions. The values of the corresponding coefficients are $\beta^1_{12} = 2595$, $\beta^1_{21} = -2412$, with peaks at the time moments $t = 49, 85$, respectively. As the right break is wider, the neighboring basic function provides the coefficient $\beta^1_{22} = -690$, $t = 89$. Moreover, because the time series is very volatile for $t \in \{49, 89\}$, the volatility is reflected also by the coefficients $\beta^1$ whose values range around $\pm 500$. This provides additional information about the course of the time series.

Figure 3. Artificial time series from Figure 2 with fuzzy partition and marked basic functions detecting structural breaks. Note that the fuzzy partition is constructed over the interval $[0, q]$. The evaluative expression over the interval determined by the marked basic functions is always huge increase (decrease). The green line is approximation of the time series using the inverse F-transform.

of Ostrava, Czech Republic. Its author is Radek Valášek). All the fuzzy partitions are equidistant with $h = 4$, formed by triangular basic functions with the support equal to $2h$. The linguistic contexts of each time series were obtained by setting $\pm v_R = \sigma/(2h)$, $v_L = 0$ and $v_S = 0.4v_R$. The standard deviations are $\sigma_{(a)} = 978, \sigma_{(b)} = 1711, \sigma_{(c)} = 6307, \sigma_{(d)} = 4921$, respectively. The results are depicted in Figure 3.
Figure 4. The result of Chow test applied to the artificial time series from Figure 2. The upper part of each subfigure is the given time series, in the lower part are the corresponding $p$-values.
Linguistic evaluation of the course of time series between the structural breaks (the interval $[49, 89]$) is also “stag-nating” which means that the volatility does not lead to change of the trend. On the other hand, inside this area places with specific course can be detected, e.g., “clear decrease” in $[52, 64]$ and “clear increase” in $[64, 75]$.

5.2.2 Detection using Chow test

For comparison, we also applied the classical Chow test on the above four artificial time series. It is based on statistical test of significance of non-zero change of regression coefficients

$$y_i = \beta_i x_i + u_i, \quad i = 1, \ldots, n.$$ 

The null hypothesis is $H_0 : \beta_0 = \beta_i$. If it is rejected then there is a structural break detected. Note that Chow test requires longer time series.

The results of the Chow test applied to the above 4 artificial time series is in Figure 4. One can see that the detection using this method is not convincing. As a matter of fact, all $p$-values are significant which means that the structural breaks are detected in each time moment. Of course, it does not mean that other statistical methods cannot be successful. At any case, however, we may conclude that our method has the power to detect structural breaks in time series and sometimes is able to give even more convincing results than the standard statistical techniques.

6. Conclusions

In this paper we suggested the method how structural breaks in time series can be detected. The method is based on application of the first degree fuzzy transform that provides estimation of the slope of time series in an imprecisely determined area. The structural breaks are easily detected if we use fuzzy partition whose basic functions are sufficiently narrow. The method is robust w.r.t. the starting position of the fuzzy partition. It is also important to emphasize that the detection is extremely fast because time complexity of the fuzzy transform is linear.

Finally let us remark that this method is related to detection of intervals of monotonous course of time series described in [9]. Using this algorithm, we can obtain information about length and character of the behavior of the time series between structural breaks.
[9] V. Novák, “Linguistic characterization of time series,” *Fuzzy Sets and Systems*, vol. 285, pp. 52-72, 2016. [https://doi.org/10.1016/j.fss.2015.07.017](https://doi.org/10.1016/j.fss.2015.07.017)

[10] V. Novák, I. Perfilieva, and J. Mockor, *Mathematical Principles of Fuzzy Logic*. New York, NY: Springer, 1999.

[11] P. Hajek, *Metamathematics of Fuzzy Logic*. Dordrecht: Kluwer Academic Publisher, 1998.

[12] I. Perfilieva, “Fuzzy transforms: theory and applications,” *Fuzzy Sets and Systems*, vol. 157, no. 8, pp. 993-1023, 2006. [https://doi.org/10.1016/j.fss.2005.11.012](https://doi.org/10.1016/j.fss.2005.11.012)

[13] I. Perfilieva, M. Dankova, and B. Bede, “Towards a higher degree F-transform,” *Fuzzy Sets and Systems*, vol. 180, no. 1, pp. 3-19, 2011. [https://doi.org/10.1016/j.fss.2010.11.002](https://doi.org/10.1016/j.fss.2010.11.002)

[14] V. Novák, I. Perfilieva, M. Holcapek, and V. Kreinovich, “Filtering out high frequencies in time series using F-transform,” *Information Sciences*, vol. 274, pp. 192-209, 2014. [https://doi.org/10.1016/j.ins.2014.02.133](https://doi.org/10.1016/j.ins.2014.02.133)

[15] I. Perfilieva and V. Kreinovich, “Fuzzy transforms of higher order approximate derivatives: a theorem,” *Fuzzy Sets and Systems*, vol. 180, no. 1, pp. 55-68, 2011. [https://doi.org/10.1016/j.fss.2011.05.005](https://doi.org/10.1016/j.fss.2011.05.005)

[16] V. Novák, “A comprehensive theory of trichotomous evaluative linguistic expressions,” *Fuzzy Sets and Systems*, vol. 159, no. 22, pp. 2939-2969, 2008. [https://doi.org/10.1016/j.fss.2008.02.023](https://doi.org/10.1016/j.fss.2008.02.023)

[17] A. Dvorak, M. Stepnicka, and L. Stepnickova, “On redundancies in systems of fuzzy/linguistic IF-THEN rules under perception-based logical deduction inference,” *Fuzzy Sets and Systems*, vol. 277, pp. 22-43, 2015. [https://doi.org/10.1016/j.fss.2014.10.002](https://doi.org/10.1016/j.fss.2014.10.002)

[18] V. Novák, “Perception-based logical deduction,” in *Computational Intelligence, Theory and Applications*. Heidelberg: Springer, 2005, pp. 237-250. [https://doi.org/10.1007/3-540-31182-3_21](https://doi.org/10.1007/3-540-31182-3_21)

[19] V. Novák and S. Lehmke, “Logical structure of fuzzy IF-THEN rules,” *Fuzzy Sets and Systems*, vol. 157, no. 15, pp. 2003-2029, 2006. [https://doi.org/10.1016/j.fss.2006.02.011](https://doi.org/10.1016/j.fss.2006.02.011)

[20] V. Novák and I. Perfilieva, “On the semantics of perception-based fuzzy logic deduction,” *International Journal of Intelligent Systems*, vol. 19, no. 11, pp. 1007-1031, 2004. [https://doi.org/10.1002/int.20034](https://doi.org/10.1002/int.20034)

[21] M. Holcapek, “Monadic L-fuzzy quantifiers of the type < 1," Fuzzy Sets and Systems*, vol. 159, no. 14, pp. 1811-1835, 2008. [https://doi.org/10.1016/j.fss.2008.03.028](https://doi.org/10.1016/j.fss.2008.03.028)

[22] A. Dvorak and M. Holcapek, “L-fuzzy quantifiers of the type <1> determined by measures,” *Fuzzy Sets and Systems*, vol. 160, no. 23, pp. 3425-3452, 2009. [https://doi.org/10.1016/j.fss.2009.05.010](https://doi.org/10.1016/j.fss.2009.05.010)

[23] V. Novák, “A formal theory of intermediate quantifiers,” *Fuzzy Sets and Systems*, vol. 159, no. 10, pp. 1229-1246, 2008. [https://doi.org/10.1016/j.fss.2007.12.008](https://doi.org/10.1016/j.fss.2007.12.008)

[24] P. Murinova and V. Novák, “A formal theory of generalized intermediate syllogisms,” *Fuzzy Sets and Systems*, vol. 186, no. 1, pp. 47-80, 2012. [https://doi.org/10.1016/j.fss.2011.07.004](https://doi.org/10.1016/j.fss.2011.07.004)

[25] P. Murinova and V. Novák, “The structure of generalized intermediate syllogisms,” *Fuzzy Sets and Systems*, vol. 247, pp. 18-37, 2014. [https://doi.org/10.1016/j.fss.2014.01.009](https://doi.org/10.1016/j.fss.2014.01.009)

[26] D. Dubois and H. Prade, “What are fuzzy rules and how to use them,” *Fuzzy Sets and Systems*, vol. 84, no. 2, pp. 169-185, 1996. [https://doi.org/10.1016/0165-0114(96)00066-8](https://doi.org/10.1016/0165-0114(96)00066-8)

[27] D. Ruan and E. E. Kerre, *Fuzzy If-Then Rules in Computational Intelligence: Theory and Applications*. New York, NY: Springer, 2000.

[28] M. Delgado, M. Ruiz, D. Sanchez, and M. Vila, “Fuzzy quantification: a state of the art,” *Fuzzy Sets and Systems*, vol. 242, pp. 1-30, 2014. [https://doi.org/10.1016/j.fss.2013.10.012](https://doi.org/10.1016/j.fss.2013.10.012)

[29] I. Glockner, *Fuzzy Quantifiers: A Computational Theory*. Heidelberg: Springer, 2006.

[30] S. Peters and D. Westerstahl, *Quantifiers in Language and Logic*. Oxford: Oxford University Press, 2006.

[31] L. A. Zadeh, “A computational approach to fuzzy quantifiers in natural languages,” *Computers & Mathematics with Applications*, vol. 9, no. 1, pp. 149-184, 1983. [https://doi.org/10.1016/0898-1221(83)90013-5](https://doi.org/10.1016/0898-1221(83)90013-5)
[32] J. Andel, *Statistical Analysis of Time Series*. Praha: SNTL, 1976.

[33] J. D. Hamilton, *Time Series Analysis*. Princeton, NJ: Princeton University Press, 1994.

[34] L. Nguyen and V. Novak, “Filtering out high frequencies in time series using F-transform with respect to raised cosine generalized uniform fuzzy partition,” in *Proceedings of International Conference on Fuzzy Systems*, Istanbul, Turkey, 2015, pp. 1-8, https://doi.org/10.1109/FUZZ-IEEE.2015.7337864

[35] J. Kacprzyk, A. Wilbik, and S. Zadrozny, “Linguistic summarization of time series using a fuzzy quantifier driven aggregation,” *Fuzzy Sets and Systems*, vol. 159, no. 12, pp. 1485-1499, 2008, https://doi.org/10.1016/j.fss.2008.01.025

Vilem Novak received M.Sc. in System engineering at Mining Technical University, Ostrava in 1975, Post-graduate course of theoretical cybernetics at Charles University in Prague, Czechoslovakia, 1982. Also, he received Ph.D. in mathematical logic at Charles University, Prague, 1988, DSc. (Doctor of Sciences) in mathematics at Polish Academy of Sciences, Warsaw in 1995, associate professor habilitation, at Palacky University, Olomouc, in 1995, and became full professor in Masaryk University, Brno in 2001. He is professor of the University of Ostrava, Czech Republic and director of the Institute for Research and Applications of Fuzzy Modeling. His research areas are mathematical fuzzy logic, mathematical modeling of linguistic semantics and fuzzy natural logic, fuzzy modeling, modeling with words and their applications. He has been working as a general chair of international conferences IFSA 1997 and EUSFLAT 2007. He works as an editor and guest editor of some academic journals such as *Fuzzy sets and systems*, *Soft computing*, and several other ones. He is the author or co-author of more than 270 scientific research articles, 6 scientific monographs and has known for more than 2,000 citations of his books and papers.