

F-transform and Fuzzy Natural logic in Time Series Analysis

Vilém Novák, Viktor Pavliska, Irina Perfilieva, Martin Štěpnička

Centre of Excellence IT4Innovations, division of the University of Ostrava
Institute for Research and Applications of Fuzzy Modeling, 30. dubna 22, 701 03 Ostrava 1, Czech Republic

Abstract

This paper continues the development of the innovative method for time series analysis and forecasting using special soft-computing techniques: fuzzy (F-) transform and Fuzzy Natural Logic. We will demonstrate that the F-transform is a proper technique for extraction of the trend-cycle of time series. Furthermore, we will elaborate in more detail automatic generation of linguistic evaluation of its behavior in arbitrary time slots. Thanks to the first-degree F-transform (F^1 -transform), this works even if the graph of the time series visually does not suggest a clear tendency.

Keywords: Fuzzy transform; F-transform; evaluative linguistic expressions; fuzzy natural logic

1. Introduction

In [1, 2, 3], advanced application of special soft computing methods, namely the *fuzzy transform* (*F-transform*) and *fuzzy natural logic*^{*} to analysis and forecasting of time series was proposed. The method was implemented in the software LFL Forecaster. Its results are convincing since, as demonstrated in [2], precision of the forecast using this software is fully comparable with precision of the top professional systems such as ForecastPro®. In comparison with them, however, our system provides added value — automatically generated linguistic comments explaining how forecast of the trend-cycle of the time series has been obtained. The comments are generated using methods of fuzzy natural logic.

In [5], we proposed the idea to generate automatically also linguistic evaluation of the trend of the time series that is, its general tendency in a given time slot. A similar problem has been solved in [6]. The authors suggest to approximate the general trend by a properly inclined line that is obtained by a heuristic and quite complicated algorithm.

In this paper, we follow the idea from [3] where we proposed to use the F^1 -transform because it provides estimation of the average value of a first derivative of a given function over a specified area.

^{*}Fuzzy natural logic continues the program initiated by the concept of *fuzzy logic in broader sense* (FLb-logic) introduced in [4].

This is very useful because the trend of the time series is often not clear even when viewing the graph. The situation can be even more difficult if we are to characterize many (tens to thousands) time series. The evaluation can be useful, for example, when managers are to make decision about future directing of their company, for global economical analysis, and in many other occasions. Below, we elaborate this idea in more detail from linguistic point of view.

The notion of F-transform was in detail described in [7]. It has a lot of interesting properties (approximation ability, optimality and others) and great potential in various kinds of applications, for example, in image processing, mining dependencies from numerical data, signal processing, special numerical methods and other ones (see [8, 9] and elsewhere).

The *fuzzy natural logic* (FNL) is an extension of mathematical fuzzy logic in narrow sense. Its goal is to develop a *formal theory of human reasoning* whose essential feature is the use of natural language. Therefore, FNL includes also mathematical model of semantics of certain special natural language expressions as well as theory of generalized quantifiers. At present, it consists of formal theories of evaluative linguistic expressions, intermediate and generalized quantifiers and their syllogisms and formal theory of the meaning of fuzzy/linguistic IF-THEN rules and reasoning based on the latter (cf., e.g., [10]). Recall that the concept of *natural logic* appears already since seventies of the previous century (cf. [11]).

By a fuzzy set, we understand a function $A : U \rightarrow [0, 1]$ where U is a universe and $[0, 1]$ is a support set of some standard algebra of truth values. The set of all fuzzy sets over U is denoted by $\mathcal{F}(U)$. If A is a fuzzy set in U then we will sometimes write $A \subseteq U$.

2. Theory of F-transform

The fundamental idea of this theory is to transform a continuous function $f : [a, b] \rightarrow \mathbb{R}$ to a finite vector of numbers (*direct F-transform*) and then to transform it back (*inverse F-transform*). The result is a function \hat{f} with interesting properties that approximates the original function f .

2.1. The principle of F-transform

The first step is to form a *fuzzy partition* of the domain $[a, b]$. It consists of a finite set of fuzzy sets $\mathcal{A} = \{A_0, \dots, A_n\}$, $n \geq 2$, defined over nodes $a = c_0, \dots, c_n = b$. Properties of the fuzzy sets from \mathcal{A} are specified by five axioms, namely: *normality*, *locality*, *continuity*, *unimodality*, and *orthogonality*. A fuzzy partition \mathcal{A} is called *h-uniform* if the nodes c_0, \dots, c_n are *h-equidistant*, i.e., for all $k = 0, \dots, n-1$, $c_{k+1} = c_k + h$, where $h = (b-a)/n$ and the fuzzy sets A_1, \dots, A_{n-1} are shifted copies of a *generating function* $A : [-1, 1] \rightarrow [0, 1]$ such that for all $k = 1, \dots, n-1$

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

(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). We also put $s_0 = \int_{-1}^1 A(x) dx^\dagger$.

The membership functions A_0, \dots, A_n in a fuzzy partition \mathcal{A} are called *basic functions*. Once the basic functions $A_0, \dots, A_n \in \mathcal{A}$ are selected, we define a *direct F-transform* of a continuous function f as a vector $\mathbf{F}[f] = (F_0[f], \dots, F_n[f])$, where each k -th component $F_k[f]$ is equal to

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

The meaning of $F_k[f]$ component is *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 $\mathbf{F}[f]$ is a continuous function[‡] $\hat{f} : [a, b] \rightarrow \mathbb{R}$ such that

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

It is proved that the function \hat{f} differs from f (unless f is a constant function) but, under certain conditions, the sequence $\{\hat{f}_n\}$ uniformly converges to f for $n \rightarrow \infty$. All the details and full proofs can be found in [7, 12].

2.2. F^1 -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 F^m transform. This generalization has been in detail described in [12].

Let us remark that extension of the F-transform to higher degree is not autotelic. First of all, we

[†]If the partition fulfils the orthogonality condition (sum of all the membership functions at each point $x \in [a, b]$ is equal to 1), then $s_0 = 1$.

[‡]By abuse of language, we call by direct F-transform both the procedure as well as its result \hat{f} .

can achieve better approximation properties. Other nice property is the possibility to estimate also derivatives of the given function f as average values over wider area. In this paper, we need only F^1 -transform whose brief description follows.

Definition 1

Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function and $\mathcal{A} = \{A_0, \dots, A_n\}$, $n \geq 2$ be a fuzzy partition of $[a, b]$. The vector of linear functions

$$\mathbf{F}^1[f] = (\beta_1^0 + \beta_1^1(x - c_1), \dots, \beta_{n-1}^0 + \beta_{n-1}^1(x - c_{n-1})) \quad (1)$$

is called the F^1 -transform of f with respect to the fuzzy partition \mathcal{A} , where

$$\beta_k^0 = \frac{\int_{c_{k-1}}^{c_{k+1}} f(x) A_k(x) dx}{hs_0}, \quad (2)$$

$$\beta_k^1 = \frac{\int_{c_{k-1}}^{c_{k+1}} f(x)(x - c_k) A_k(x) dx}{\int_{c_{k-1}}^{c_{k+1}} (x - c_k)^2 A_k(x) dx} \quad (3)$$

for every $k = 1, \dots, n-1$.

The fuzzy partition in the above definition needs not be uniform. If it is uniform, then the following simplification holds true.

Theorem 1

Let an *h-uniform* partition of $[a, b]$ be given by the triangular-shaped basic functions $A_1, \dots, A_{n-1} \in \mathcal{A}$ with the generating function $A_0 = 1 - |x|$. Then the coefficients β_k^0 and β_k^1 in the representation (1) of $\mathbf{F}^1[f]$ are given by

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

$$\beta_k^1 = \frac{12 \int_{c_{k-1}}^{c_{k+1}} f(x)(x - c_k) A_k(x) dx}{h^3}, \quad (5)$$

for every $k = 1, \dots, n-1$.

The following theorem plays an important role in our application to time series trend evaluation.

Theorem 2

Let $\mathcal{A} = \{A_1, \dots, A_{n-1}\}$ be an *h-uniform* partition of $[a, b]$, let functions f and $A_k \in \mathcal{A}$, $k = 1, \dots, n-1$, be four times continuously differentiable on $[a, b]$. Finally, let $\mathbf{F}^1[f]$ be the F^1 -transform (1) of f . Then

$$\beta_k^1 = f'(c_k) + O(h), \quad k = 1, \dots, n-1. \quad (6)$$

According to Theorem 2, the coefficient β_k^1 is a convenient average estimation of the first derivative of f in the interval $[c_{k-1}, c_{k+1}]$. We will use this result in the evaluation of the trend of time series.

2.3. Inverse F^1 -transform

The inverse F^1 -transform of a function f is defined as a linear combination of basic functions with “coefficients” given by the F^1 -transform components.

Definition 2

Let $f : [a, b] \rightarrow \mathbb{R}$ be a given function and $\mathbf{F}^1[f] = (F_1^1[f], \dots, F_n^1[f])$ be the F^1 -transform of f with respect to $\mathcal{A} = \{A_0, \dots, A_n\}$. Then, the function $\hat{f}^1 : [a, b] \rightarrow \mathbb{R}$ defined by

$$\hat{f}^1(x) = \sum_{k=1}^{n-1} F_k^1[f](x) A_k(x) \quad (7)$$

is called inverse F^1 -transform of f with respect to $\mathbf{F}^1[f]$.

Similarly as the F^0 transform, the sequence $\{\hat{f}^1\}$ of inverse F^1 -transforms of f uniformly converges to f (see [12] for the details).

3. Analysis of time series using F-transform

3.1. Decomposition of time series

Let us consider a stochastic process (see [13, 14])

$$X : [a, b] \times \Omega \rightarrow \mathbb{R} \quad (8)$$

where $[a, b] \subset \mathbb{R}$ is an interval of reals and $\langle \Omega, \mathcal{A}, P \rangle$ is a probabilistic space. For simplicity, we will usually suppose that $a = 0$. By a *time series* we understand a stochastic process where $[a, b]$ is replaced by a finite set $Q = \{0, \dots, p\} \subset \mathbb{N}$. It follows from (8), each $X(t, \omega)$ for $t \in [a, b]$ and $\omega \in \Omega$ is a random variable. If we fix $\omega \in \Omega$ then we obtain one realization of (8) and in this case, we will write $X(t)$ only. Our basic assumption is that $X(t, \omega)$ can be decomposed into three constituent components (cf. [14, 15]), namely

$$X(t, \omega) = TC(t) + S(t) + R(t, \omega), \quad t \in [a, b], \omega \in \Omega, \quad (9)$$

where $TC(t)$ is a *trend-cycle* and $S(t)$ is a seasonal component of the time series $X(t)$. The $TC(t)$ is assumed to be an ordinary real function. The seasonal component $S(t)$ is considered to be a mixture of complex periodic functions

$$S(t) = \sum_{j=1}^r P_j e^{i(\lambda_j t + \varphi_j)} \quad (10)$$

for some finite r where λ_j are frequencies, φ_j phase shifts and P_j are amplitudes. The $R(t, \omega)$ is a random noise, i.e. each $R(t)$ for $t \in [a, b]$ is a random variable with the zero mean value finite variance. We assume that μ is equal to 0 and has a finite variance. Moreover, in this paper we consider the noise to be represented by the simplest possible type of a stationary stochastic process (see, e.g., [16, Example 1]), namely, as the process of the type

$$R(t) = \xi e^{i\lambda t + \varphi} \quad (11)$$

where ξ is a random variable with zero mean value and λ a real number. It is known (see, e.g., [16]) that, under reasonable conditions, every stationary random process with zero mean can be represented as a linear combination of processes of type (11).

Thus, (9) becomes

$$X(t) = TC(t) + \sum_{j=1}^r P_j e^{i(\lambda_j t + \varphi_j)} + \xi e^{i\lambda t + \varphi}, \quad t \in [a, b]. \quad (12)$$

3.2. Extraction of the trend-cycle and trend

We will suppose the following:

- (i) The stochastic process X can be decomposed as in (12) where the seasonal constituent S consists of periodic functions having periodicities T_j , $j \in \{1, \dots, r\}$ (see (10)). The longest of the latter is denoted by \bar{T} .
- (ii) A number $\bar{d} \in \mathbb{N}$ is chosen and the distance h in (28) between the nodes is set to

$$h = \bar{d} \bar{T} \quad (13)$$

so that $n \geq 2$. The corresponding triangular fuzzy partition A_0, \dots, A_n is fixed.

- (iii) The trend-cycle TC is a function with no clear periodicity or its periodicity is much longer than \bar{h} . Moreover, the modulus of continuity $\omega(h, TC)$ is small.

Put $\bar{\xi} = \sup\{\xi(t) \mid t \in [a, b]\}$ and $\underline{\xi} = \inf\{\xi(t) \mid t \in [a, b]\}$, and

$$\tilde{\xi} = \begin{cases} \bar{\xi} & \text{if } \bar{\xi} \geq 0, \\ \underline{\xi} & \text{if } \bar{\xi} < 0. \end{cases} \quad (14)$$

where $\xi(t)$ is a realization of the random variable ξ in (11) at the given time moment t . Finally, put

$$D^0 = \sum_{j \in I} \left| \frac{P_j \sin^2(d'_j \pi)}{d_j^2 \pi^2} \right| + \frac{|\tilde{\xi}| \sin^2(d \pi)}{d^2 \pi^2}, \quad (15)$$

$$D^1 = \sum_{j \in I} \sum_{j \in I} \eta_j(t) + |\tilde{\xi}| \cdot \eta(t) \quad (16)$$

where $d_j = \frac{h}{T_j}$ and $I \subset \{1, \dots, r\}$ is the set of all subscripts, for which $d'_j = d_j - [d_j] \in (0, 1)$, $\tilde{\xi}$ is determined in (14) and $\eta(t)$ is equal to

$$\eta(t) = \frac{\sin^2(d \pi)}{d^4 \pi^4} \cdot \max\{(d^4 \pi^4 \sin^2 d \pi + 9(\lambda t - 2dk\pi)^2 (\sin d \pi - d \pi \cos d \pi)^2)^{\frac{1}{2}} \mid k = 1, \dots, n-1\}. \quad (17)$$

Let us remark that for $d \geq 2$, D^0 in (15) as well as D^1 in (16) are *very small*.

Theorem 3

Let $X(t)$ be realization of the stochastic process in (9) considered over the interval $[0, b]$. If we construct a fuzzy partition over the set of equidistant nodes (28) with the distance (13) then the corresponding inverse F -transform \hat{X}^m , $m \in \{0, 1\}$, of $X(t)$ gives the following estimator of the trend-cycle:

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

for $t \in [c_1, c_{n-1}]$, where D^m is the error (15) or (16), respectively depending on the degree of the applied F -transform.

Proofs of the theorems above and more details can be found in [17].

The trend-cycle is generally a function with very small changes (low frequency) and so, its modulus of continuity $\omega(h, TC)$ is also small. Since, as noted, D^0 and D^1 are also very small, we may conclude from (18) that $\hat{X}^m \approx TC$, i.e., the F -transform makes it possible to extract trend-cycle of time series with high precision.

To verify experimentally the above methodology, we artificially formed a time series $X(t)$ defined on the set of integers $\{0, \dots, 100\}$ as follows:

$$X(t) = TC(t) + 5\sin(0.63t + 1.5) + 5\sin(1.26t + 0.35) + 15\sin(2.7t + 1.12) + 7\cos(0.41t + 0.79) + R(t). \quad (19)$$

The function $TC(t)$ in (19) is the trend-cycle given by artificial data without clear periodicity (it is depicted by dotted line in Figure 1; its modulus of continuity is $\omega(30, TC) = 3.22$). The other four sine members form the seasonal constituent $S(t)$. Their periodicities are $T_1 = 10$, $T_2 = 5$, $T_3 = 2.3$, $T_4 = 15.4$, respectively. Therefore, we set $\bar{T} = T_4$ and $d = 1$, i.e. the distance (13) is $h = 15$. Consequently, the width of basic functions is $2h = 30$ (the time axis is discrete and so, fractions are neglected). Since all d_1, \dots, d_4 are close to natural numbers, the error D^0 in (15) is practically 0. The $R(t)$ is a random noise with (almost) zero average. The result of application of the F^0 -transform is depicted in Figure 1. One can see from it that the both the whole seasonal constituent as well as the noise were almost completely removed. Maximal difference $|TC(t) - \hat{X}^0(t)| = 3.32$ and so, we may conclude that the trend-cycle is estimated with the error corresponding to (18).

4. Linguistic evaluation of the behavior of time series

Important feature of the trend-cycle is its *trend* (tendency). We may distinguish it on the whole time series as well as in local time slots (for example, quarter of year, production period, etc.). The course of the time series, however, can be largely

volatile and so, it may be quite difficult to recognize its trend. Thus, it might be useful to have an objective tool using which the trend can be clearly recognized even on the volatile time series. A suitable tool is the F^1 -transform because it enables us to estimate the tangent. On the basis of Theorem 2, we will formulate the following definition.

Definition 3

Let $F^1[X]$ be a direct F^1 -transform of the time series X in (9) w.r.t. a fuzzy partition \mathcal{A} . Then trend $T(A_k)$ of X in the area characterized by the fuzzy set $A_k \in \mathcal{A}$ is

$$T(A_k) = \beta_k^1 \quad (20)$$

where β_k^1 is the coefficient (3) (or, alternatively, (5)).

Hence, $T(A_k)$ is a weighted average tangent of the function $X(t)$ over the area determined by the fuzzy set (basic function) $A_k \in \mathcal{A}$.

4.1. Evaluative linguistic expressions

One of methods of the fuzzy natural logic applied in [2] is the *perception-based logical deduction* (PbLD). It is applied to forecasting of the trend cycle. The basic theory behind is the formal theory of special natural language expressions called *evaluative linguistic expressions* which was in detail described in [18]. Recall that the latter are expressions such as *small*, *very big*, *rather medium*, *extremely strong*, etc. On the basis of our theory a mathematical model of the meaning of evaluative expressions was developed.

Essential concept in this theory is that of (linguistic) *context*. For evaluative expressions, it is determined by a triple of real numbers $\langle v_L, v_S, v_R \rangle$ where $v_L < v_S < v_R$ ($\in \mathbb{R}$). These numbers represent the *smallest*, *typically medium*, and the *largest* thinkable values, respectively. The context is thus a set

$$w = \{x \mid v_L \leq x \leq v_R\} \quad (21)$$

together with three distinguished points $DP(w) = \langle v_L, v_S, v_R \rangle$. By W we denote the set of all contexts (21) and by $EvExpr$ the set of all considered evaluative expressions. Each evaluative expression $Ev \in EvExpr$ is assigned the meaning which is a function

$$\text{Int}(Ev) : W \longrightarrow \mathcal{F}(\mathbb{R}).$$

We will call this function *intension* of the evaluative expression Ev . It assigns to each context $w \in W$ a fuzzy set $\text{Ext}_w(Ev) \subseteq w$ called *extension* of the expression Ev in the context $w \in W$.

We will distinguish abstract evaluative expressions, i.e. expressions such as *small*, *weak*, *very strong*, etc., that alone do not talk about any specific objects and *evaluative linguistic predications* such as “temperature is high, expenses are extremely low, the building is quite ugly”, etc. In general, the latter have the surface form

$$\langle \text{noun} \rangle \text{ is } \langle \text{simple evaluative expression} \rangle \quad (22)$$

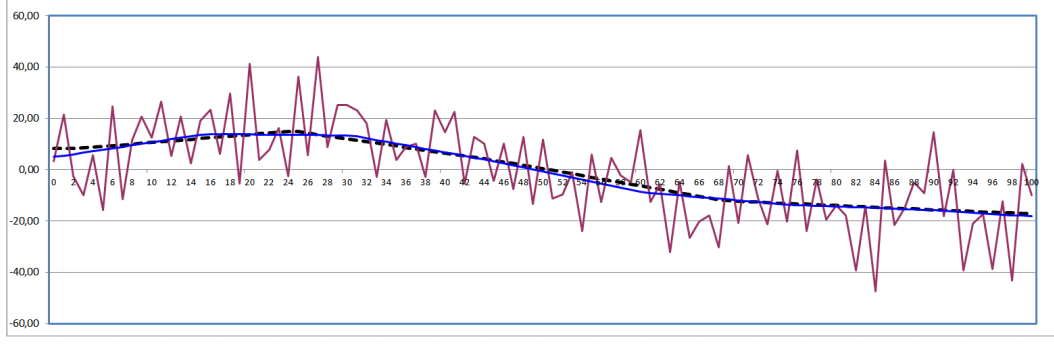


Figure 1: F-transform of the artificial time series $X(t)$, $t \in \{0, \dots, 100\}$. The dotted line depicts the original trend-cycle $TC(t)$ given by the data. Using (19), it is completed by values of the four sine members and the noise $R(t)$ to form the artificial time series depicted also in the figure. After application of the F-transform to $X(t)$, we obtain approximation $\hat{X}(t)$ of the trend-cycle that is depicted by the solid line. Note that $\hat{X}(t)$ is almost identical with the original $TC(t)$.

where

$$\langle \text{simple evaluative expression} \rangle := \langle \text{hedge} \rangle \langle \text{TE-adjective} \rangle,$$

$\langle \text{hedge} \rangle$ is a linguistic hedge (for example *very*, *rather*, *extremely*, *more or less*, *roughly*, etc.) and $\langle \text{TE-adjective} \rangle$ is a trichotomous evaluative adjective (for example *small*, *medium*, *big*, *large*, *weak*, *good*, etc.). The “is” takes here the role of a copula assigning property to objects and is not treated as a genuine verb.

4.2. Linguistic evaluation of the time series trend

The theory of evaluative expressions provides a special function of *local perception*

$$LPerc : w \times W \longrightarrow EvExpr. \quad (23)$$

This function assigns to each value $x \in w$ in each context $w \in W$ an evaluative expression of the form (22). The function is constructed in such a way that given a linguistic context $w \in W$ and a value $x \in w$, the result of (23) is the *most plausible* evaluative expression that characterizes x in the given context w . This function makes it possible to *learn* linguistic description (a set of fuzzy/linguistic IF-THEN rules) characterizing behavior of the trend cycle so that its course can be predicted.

Using (23), we can also generate linguistic evaluation of the trend of time series (cf. Definition 3). First we must specify the context. We start with specification of what does it mean “extreme increase (decrease)”. In practice, it can be determined as the largest acceptable difference of time series values with respect to a given (basic) time interval. Hence, mathematically we speak about the tangent. The usual basic time interval is 12 months, 31 days, etc. depending on the kind of the time series. Thus, the context is determined by the three distinguished

values v_L, v_S, v_R of the tangent. The largest tangent v_R is determined in the way mentioned above while the smallest one is usually $v_L = 0$. The typical medium value v_S is determined analogously as v_R . The result is the context $w_{tg} = \langle v_L, v_S, v_R \rangle$.

Now, we can linguistically characterize the trend $T(A_k)$ in (20) with respect to the context w_{tg} , i.e. we will automatically generate evaluative linguistic expressions using the function of local perception (23):

$$LPerc(T(A_k), w_{tg}). \quad (24)$$

This is justified by the fact that $T(A_k)$ is an average tangent over an area covered by the basic function $A_k \in \mathcal{A}$.

Predications using which we linguistically evaluate time series trend have specific form. The basic characteristic is *sign* of the trend. This is characterized by a special word, namely “+” is expressed by the word *increasing* (or *increase*) and “−” by *decreasing* (or *decrease*). This can further be completed by special expressions characterizing its *gradient*. Moreover, the obtained expressions are apparently subject to ordering that is similar to the natural ordering of the “standard” evaluative expressions. We conclude that the general syntactic form of expressions characterizing trend is either (a) or (b) specified below:

$$(a) \quad \text{Trend is } \langle \text{gradient} \rangle \quad (25)$$

where

$$\langle \text{gradient} \rangle := \text{stagnating} | \langle \text{hedge} \rangle \langle \text{sign} \rangle, \quad (26)$$

$$\langle \text{sign} \rangle := \text{increasing} | \text{decreasing} \quad (27)$$

and

$$\langle \text{hedge} \rangle := \text{negligibly} | \text{slightly} | \text{somewhat} | \text{clearly} | \text{roughly} | \text{sharply} | \text{significantly}.$$

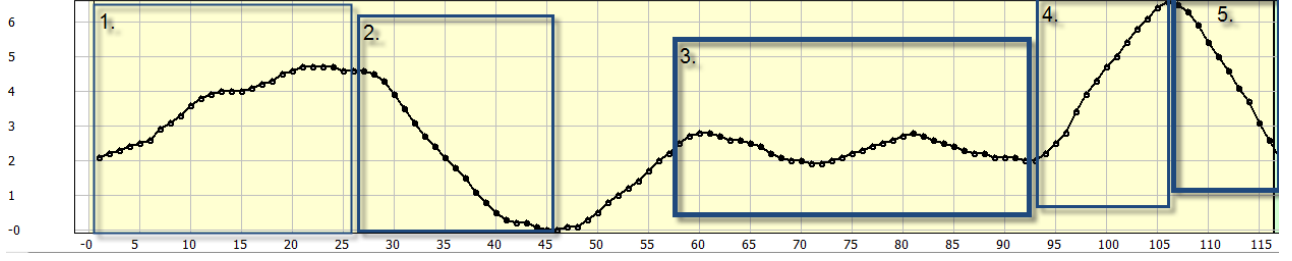


Figure 2: Automatically generated linguistic evaluations of monthly trend of inflation measure over 10 years. The context for tangent is $\langle v_L = 0, v_S = 2/12, v_R = 5/12 \rangle$. Evaluation of trend in the marked areas is the following: Slot 1.: *clear increase*, Slot 2.: *fairly large decrease*, Slot 3. *stagnating*, Slot 4. *significant increase* and Slot 5. *huge decrease*.

- (b) In some cases, however, only the feature *increase (decrease) of trend* is evaluated:

$$\langle \text{sign of trend} \rangle \text{ is } \langle \text{special hedge} \rangle \quad (28)$$

where $\langle \text{sign of trend} \rangle := \text{increase|decrease}$ and

$$\begin{aligned} \langle \text{special hedge} \rangle := & \text{negligible|slight|} \\ & \text{small|clear|rough|large|fairly large|} \\ & \text{quite large|significant|huge.} \end{aligned}$$

Note that, in fact, the *increase (decrease) (of trend)* is evaluated both in (25) as well as in (28). The difference in their use depends on the syntactic specificities but not on their semantics. Some cases, for example “trend is slightly increasing” and “increase of trend is slight” are even synonymous.

This suggests the idea that the above special evaluative predications (25) and (28) are semantically tantamount[†] to the standard form

$$\langle \text{sign of trend} \rangle \text{ is } \mathcal{B} \quad (29)$$

where \mathcal{B} is an evaluative expression generated by the function (23). Therefore, we can first generate the predications (29) linguistically characterizing the given tangent and then replace them by (25) or (28) according to the following tables:

Case (a)

Tantamount linguistic expressions	
$\langle \text{gradient} \rangle$	\mathcal{B}
stagnating	Ze, \pm extremely small
negligibly $\langle \text{sign} \rangle$	significantly small
slightly $\langle \text{sign} \rangle$	very small
somewhat $\langle \text{sign} \rangle$	rather small
clearly $\langle \text{sign} \rangle$	medium, quite roughly small, very roughly small
roughly $\langle \text{sign} \rangle$	quite roughly big, very roughly big
sharply $\langle \text{sign} \rangle$	very big
significantly $\langle \text{sign} \rangle$	significantly big

[†]It is not clear whether the expressions (25) and (28) are indeed synonymous in the strict sense. We need further linguistic research to answer this question.

Case (b)

Tantamount linguistic expressions	
$\langle \text{special hedge} \rangle$	\mathcal{B}
negligible	significantly small
slight	very small
small	small
clear	medium, quite roughly small, very roughly small
rough	quite roughly big, very roughly big
fairly large	roughly big, more or less big
quite large	rather big
large	big
sharp	very big
significant	significantly big
huge	extremely big

5. Examples

As mentioned, practical realization of the idea presented above is based on Definition 3. Depending on the fuzzy partition, the generated evaluation may concern either the whole time series or an arbitrary part of it.

The way how evaluation is generated is demonstrated on the case of smooth curve of monthly inflation measure in Figure 2. The context for tangent is $w_{tg} = \langle v_L = 0, v_S = 2/12, v_R = 5/12 \rangle$. This means that typically extremely small tangent is 0, typically medium is 2/12 (i.e., increase 2% per 12 months) and typically extremely big is 5/12 (i.e., increase 5% per 12 months). Note that Slot 2 is wider than Slot 5 and so, the program generated the comment *fairly large decrease* for Slot 2 and *huge decrease* for Slot 5. The evaluations are generated on the basis of objectively computed average tangent and we argue that they comply well with the course of the time series.

More complicated is demonstration of comments generated to one time series taken from *M3-Competition* provided by the *International Institute of Forecasters* (the real content of the time series is not known). The time series is depicted in Fig. 3.

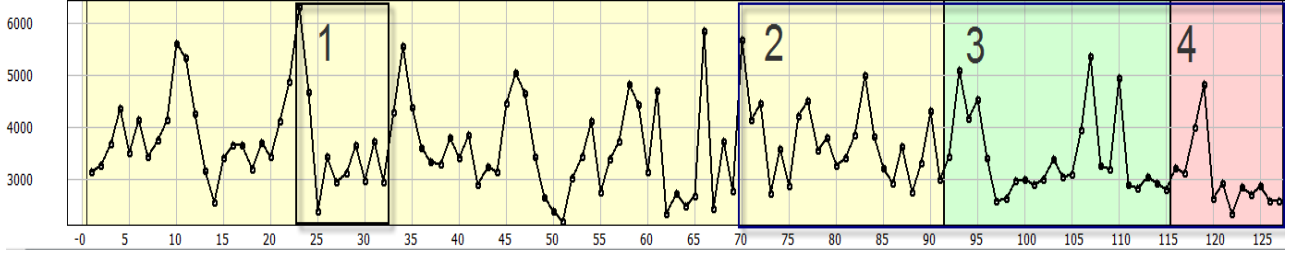


Figure 3: Demonstration of evaluation of trend of various parts of a complicated time series. Trend of the whole series is *stagnating*. Slot 1 (time 23-32): *clear decrease*, Slot 2 (time 70-127): *negligible decrease*, Slot 3 (time 92-115): *small increase*, Slot 4 (time 116-127): *fairly large decrease*.

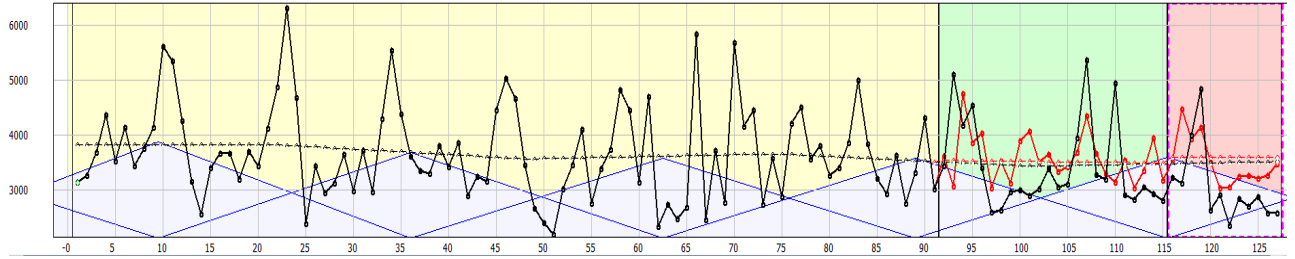


Figure 4: Analysis and forecast of the time series from Fig. 3. Slot 3 (time 92-115 — validation part) and Slot 4 (time 116-127 — testing part) contain computed and predicted trend cycle, and also real and predicted values the time series itself. The generated comments in Slot 4 are: *rough decrease* for the predicted data and *fairly large decrease* for the real data.

One can see that its trend is by no means clear. Slot 3 (time 92-115) of this time series is validation part, on which the quality of the forecast is tested and the best one is chosen. Slot 4 (time 116-127) is testing part that is not used for computation of the forecast but only for comparison of the forecast with the real data. In Fig. 4 the same time series is analyzed and its forecast is computed.

The linguistic context for the trend evaluation was set to $w_{tg} = \langle v_L = 0, v_S = 1200/12, v_R = 3000/12 \rangle$ since the time series demonstrates clear periodicity of $T = 12$ (this was obtained using periodogram — cf., e.g., [13]). The generated evaluation of trend of the predicted values in the testing part is *rough decrease* while evaluation of trend of the real data is *fairly large decrease*. Both evaluations are in good agreement which is another support for the quality of our forecasting method. Thus, instead of presenting concrete predicted numbers, the manager might be satisfied with the information that “*rough decrease is expected*”.

6. Conclusion

In this paper, we presented special soft computing methods that can be used for analysis and forecasting of time series. The used methods are F-transform (both zero and first degree) and one of the theories of the fuzzy natural logic, namely the formal theory of evaluative linguistic expressions.

F-transform method is used especially for extraction of the trend-cycle of the time series while methods of the fuzzy natural logic are applied to prediction of the latter. Furthermore, by combination of the F^1 -transform and fuzzy natural logic we can generate linguistic comments to the trend of the time series in arbitrary time slots. We believe that such comments can be useful, for example, in situations when it is difficult to see the global trend because the time series is too much varying.

One of the directions for further research is application of other theories of fuzzy natural logic, for example formal theory of *intermediate quantifiers* developed in [19, 20]. The latter are linguistic expressions such as *most*, *many*, *almost all*, *few*, *some*, *a large part of*, etc. Using this theory, we can model the meaning of sentences, such as

“most (many, few) analyzed time series stagnated recently but their future trend is slightly increasing”,

“huge decrease of trend of almost all time series in the recent quarter of the year”,

and similar. Moreover, we can also apply syllogistic reasoning with such expressions, for example

*Few analyzed time series are not
from IT industry*

*Many analyzed time series are clearly raising
Few clearly raising time series are not
from IT industry*

It is important to realize that the latter is example of valid generalized Aristotle's syllogism which means that it is true in all situations (models).

Another possibility is to mine interesting information from the given set of time series, summarize their properties and summarize also their possible future development. Namely, we start with analysis and forecasting of all the time series. Then we generate comments to interesting time slots, summarize them using the intermediate quantifiers and derive further properties on the basis of valid syllogisms. We will discuss this topic in some of future papers.

Acknowledgment

This paper was supported by the program MŠMT-KONTAKT II, project LH 12229 "Research and development of methods and means of intelligent analysis of time series for the strategic planning problems". Additional support was given also by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

References

- [1] V. Novák, M. Štěpnička, I. Perfilieva, and V. Pavliska. Analysis of periodical time series using soft computing methods. In Da Ruan, J. Montero, J. Lu, L. Martínez, P. D'hondt, and E.E. Kerre, editors, *Computational Intelligence in Decision and Control*, pages 55–60. World Scientific, New Jersey, 2008.
- [2] V. Novák, M. Štěpnička, A. Dvořák, I. Perfilieva, V. Pavliska, and L. Vavříčková. Analysis of seasonal time series using fuzzy approach. *Int. Journal of General Systems*, 39:305–328, 2010.
- [3] V. Novák, I. Perfilieva, and V. Pavliska. The use of higher-order F-transform in time series analysis. In *World Congress IFSA 2011 and AFSS 2011*, pages 2211–2216, Surabaya, Indonesia, 2011.
- [4] V. Novák. Towards formalized integrated theory of fuzzy logic. In Z. Bien and K.C. Min, editors, *Fuzzy Logic and Its Applications to Engineering, Information Sciences, and Intelligent Systems*, pages 353–363. Kluwer, Dordrecht, 1995.
- [5] V. Novák, I. Perfilieva, and N. G. Jarushkina. A general methodology for managerial decision making using intelligent techniques. In E. Rakus-Anderson, R.R. Yager, N. Ichalkaranje, and L.C. Jain, editors, *Recent Advances in Fuzzy Decision-Making*, pages 103–120. Springer, Heidelberg, 2009.
- [6] J. Kacprzyk, A. Wilbik, and Zadrozny. Linguistic summarization of time series using a fuzzy quantifier driven aggregation. *Fuzzy Sets and Systems*, 159:1485–1499, 2008.
- [7] I. Perfilieva. Fuzzy transforms: theory and applications. *Fuzzy Sets and Systems*, 157:993–1023, 2006.
- [8] I. Perfilieva, V. Novák, and A. Dvořák. Fuzzy transform in the analysis of data. *Int. Journal of Approximate Reasoning*, 48:36–46, 2008.
- [9] I. Perfilieva. Fuzzy transforms: A challenge to conventional transforms. In P. W. Hawkes, editor, *Advances in Images and Electron Physics*, 147, pages 137–196. Elsevier Academic Press, San Diego, 2007.
- [10] V. Novák. Reasoning about mathematical fuzzy logic and its future. *Fuzzy Sets and Systems*, 192:25–44, 2012.
- [11] G. Lakoff. Linguistics and natural logic. *Synthese*, 22:151–271, 1970.
- [12] I. Perfilieva, M. Daňková, and B. Bede. Towards a higher degree F-transform. *Fuzzy Sets and Systems*, 180:3–19, 2011.
- [13] J. Anděl. *Statistical Analysis of Time Series*. SNTL, Praha, 1976 (in Czech).
- [14] J.D. Hamilton. *Time Series Analysis*. Princeton, Princeton University Press, 1994.
- [15] A. Bovas and J. Ledolter. *Statistical Methods for Forecasting*. Wiley, New York, 2003.
- [16] A.M. Yaglom. *An introduction to the theory of stationary random functions. Revised English ed. Translated and edited by Richard A. Silverman*. Englewood Cliffs, NJ: Prentice-Hall, Inc. XIII, 1962.
- [17] V. Novák, I. Perfilieva, M. Holčapek, and V. Kreinovich. Filtering out high frequencies in time series using f-transform. *Information Sciences*, (submitted).
- [18] V. Novák. A comprehensive theory of trichotomous evaluative linguistic expressions. *Fuzzy Sets and Systems*, 159(22):2939–2969, 2008.
- [19] V. Novák. A formal theory of intermediate quantifiers. *Fuzzy Sets and Systems*, 159(10):1229–1246, 2008.
- [20] P. Murinová and V. Novák. A formal theory of generalized intermediate syllogisms. *Fuzzy Sets and Systems*, 186:47–80, 2012.