2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Time series grouping on the basis of F^1-transform

Anton Romanov
Department of Information Systems,
Ulyanovsk State Technical University,
Ulyanovsk, Russia
Email: romanov73@gmail.com

Irina Perfilieva
Institute for Research and Applications
of Fuzzy Modeling University of Ostrava,
Ostrava, Czech Republic
Email: irina.perfilieva@osu.cz

Nadezhda Yarushkina
Department of Information Systems,
Ulyanovsk State Technical University,
Ulyanovsk, Russia
Email: jng@ulstu.ru

***Abstract*—The contribution is focused on a new method of grouping time series according to their local tendency indicator that is expressed by a linear coefficient of the F^1 -transform. The useful consequence of grouping is an effective procedure of forecasting such that only one time series from a group is forecasted.**

***Keywords*—F-transform; fuzzy time series; fuzzy tendency; time series trend; express analysis**

I. INTRODUCTION

In a time series analysis and forecast, the use of the F- transform (in its direct and inverse form) was reported in several publications, e.g., [1], [2], [3], [4], [5]. In [1], [4], this technique was used to extract a low-frequency trend component, whereas in [2] it was used for the modeling of an autoregression function. In [3], the inverse F-transform was used as a technical indicator in a stock market instead of the commonly used simple and exponential moving averages. In [1], [4], the inverse F -transform was used in combination with perception based logical deduction, where the latter provides forecasts of the future F-transform component(s).

In [5], we made two principal changes compared to the methods described in [1], [4]. First, we decomposed a time series based on the direct F-transform (as opposed to the inverse F-transform, which was used in [1], [4]). We relax some constraints (e.g., the Ruspini condition) to make the F - transform more flexible and adjustable to various time series (in general, functions). Second, we proposed forecasting a future F -transform component based on the theory of fuzzy time series tendencies ([6]).

In this contribution, we keep the direction of [5], but decompose a time series based on the direct F^1 -transform. The reason is that linear coefficients of F^1 -transform components characterizes local tendencies which we are going to forecast. Let us explain in more details.

If we consider economic applications of time series, then we often encounter a situation where a certain process is characterized by a set of various time series. In many cases, they are connected, i.e. a dynamic behavior of one influences the others. It appears in similar behavior of their local tendencies, where the latter characterizes a stable direction of changes over a certain period, e.g., increase, decrease, etc. Even though time series are connected, their absolute values may be different. Therefore, it is reasonable to select one time series from a group, forecast its local tendency and use it for others.

The evident advantages of this type of analysis are as follows:

1) Reducing the forecasting time,
2) Creating groups of similar time series and choosing a distinctive element of each group.

The paper considers examples of two sets of time series. The first set contains 14 time series of statistics collected between 1970 and 2000 in Russian Federation in areas of economics, production, social, health and culture indicators. The second set consists of time series selected from the NN3 competition (http:// www.neural-forecasting-competition.com/ downloads/NN3/datasets/NN3_REDUCED.xls).

We perform the analysis on the basis of F^1 -transform components for both sets and find groups of similar time series. We choose a representative of each group, forecast its local tendency and show that this forecast is valid for all other elements of the same group.

II. FIRST DEGREE F^1-TRANSFORM

In this section, we recall the basic fact about F-transforms

A. Fuzzy partition with Ruspini condition

The *fuzzy partition with the Ruspini condition* (1) (simply, *Ruspini partition*) was introduced in [7]. This condition implies normality of the respective fuzzy partition, i.e., the partition-of-unity. It then leads to a simplified version of the inverse F-transform. In later publications [8], the Ruspini condition was weakened to obtain an additional degree of freedom and a better approximation by the inverse F-transform.

Definition 1. *Let $x_1 < ... < x_n$ be fixed nodes within $[a, b]$ such that $x_1 = a$, $x_n = b$ and $n \geq 2$. We say that the fuzzy sets $A_1, ..., A_n$, identified with their membership functions defined on $[a, b]$, establish a Ruspini partition of $[a, b]$ if they fulfill the following conditions for $k = 1, ..., n$:*

1) $A_k : [a, b] \to [0, 1], A_k(x_k) = 1$;
2) $A_k(x) = 0$ *if* $x \in (x_{k-1}, x_{k+1})$, *where for uniformity of notation, we set* $x_0 = a$ *and* $x_{n+1} = b$;
3) $A_k(x)$ *is continuous*;
4) $A_k(x)$, *for* $k = 2, ..., n$, *strictly increases on* $[x_{k-1}, x_k]$ *and* $A_k(x)$, *for* $k = 1, ..., n-1$, *strictly decreases on* $[x_k, x_{k+1}]$;
5) *for all* $x \in [a, b]$,

$$\sum_{k=1}^{n} A_k(x) = 1 \quad (1)$$

If the nodes $x_1, ..., x_n$ are $h-$equidistant, i.e. for all $k = 1, ..., n+1, x_k = x_{k-1} + h$ where $h = (b-a)/(n-1)$, we say that the fuzzy partition $A_1, ..., A_n$ is $h - uniform$.

The condition (1) is known as the Ruspini condition. The membership functions $A_1, ..., A_n$ are called *basic functions*. The shape of the basic functions is not predetermined and therefore, it can be chosen according to additional requirements (e.g., smoothness).

B. F^1-transform of function of one variable

The F-transform of a higher degree $m \geq 1$ was introduced in [8]. In this section, we give a short description of the F^1-transform of functions of one variable. The F^1 -transform is a generalization of the F-transform where the constant components are replaced by linear components (polynomials of the first degree).

Throughout this section, we assume that $A_1, ..., A_n, n > 2$ is a fuzzy partition of $[a, b]$ with nodes $x_k, k = 0, ..., n+1$. Let k be a fixed integer from $\{1, ..., n\}$ and let $L_2(A_k)$ be a normed space of square-integrable functions $f : [x_{k-1}, x_{k+1}] \to \mathbb{R}, k = 1, ..., n$.

By $L_2(A_1, ..., A_n)$ we denote a set of functions $f : [a, b] \to \mathbb{R}$ such that for all $k = 1, ..., n, f|_{[x_{k-1}, x_{k+1}]} \in L_2(A_k)$, where $f|_{[x_{k-1}, x_{k+1}]}$ is the restriction of f on $[x_{k-1}, x_{k+1}]$. For any function f from $L_2(A_1, ..., A_n)$ we define the F^1 -transform of f with respect to $A_1, ..., A_n$ as a vector

$$F^1[f] = [F_1^1, ..., F_n^1]$$

where the components $F_k^1, k = 1, ..., n$ are linear functions

$$F_k^1(x) = c_{k,0} + c_{k,1}(x - x_k), \quad (2)$$

with the coefficients $c_{k,0}, c_{k,1}$ given by

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

$$c_{k,1} = \frac{\int_{x_{k-1}}^{x_{k+1}} f(x)(x - x_k)A_k(x)}{(\int_{x_{k-1}}^{x_{k+1}} (x - x_k)^2 A_k(x)dx)}$$

Here we emphasize the following property of the F^1 -transform, more can be found in [8].

If f is four-times continuously differentiable on $[a, b]$, then for each $k = 1, ..., n$

$$c_{k,0} = f(x_k) + O(h^2),$$

$$c_{k,1} = f'(x_k) + O(h^2).$$

In the discrete case, a function f is assumed to be known only at points $p_i \in [a, b]$, where $i = 1, ..., N$. For the h-uniform fuzzy partition and the triangular-shaped basic functions, the coefficients $c_{k,0}^1, c_{k,1}^1, k = 1, ..., n$, are computed as follows:

$$c_{k,0} = \frac{1}{h} \sum_{i=1}^{N} f(p_i)A_k(p_i) \quad (3)$$

$$c_{k,1} = \frac{12}{h^3} \sum_{i=1}^{N} f(p_i)(p_i - x_k)A_k(p_i) \quad (4)$$

III. Time series analysis

Usually, a time series is a set of observations of an econometric quantity at various moments in time. In statistics, a time series is connected with the notion of a stochastic process, and from this concept, it is formalized by a sequence $\{x_t, t \in [1, T]\}$ of random variables. Specifically, x_t can be considered as one possible observation of the equally named random variable, such that in this case, a time series is a sequence of real numbers. In practice, real numbers x_t are known to have errors, and therefore, a time series cannot be modeled using a real function of the discrete variable t.

Depending on the statistical properties of a time series, its models are divided into two groups: stationary and non-stationary. Although stationarity is a desirable property (a model of a time series is rather simple), real-life time series are usually non-stationary. Because in our experiments, we focus on time series with economic indicators, we assume that an analyzed time series is non-stationary and can be decomposed as follows:

$$x_t = f(t) + y_t. \quad (5)$$

In (5), $f(t)$ is a deterministic part, which is usually called a *trend*, and y_t is a random part, which is additionally assumed to be stationary with a zero mean value and a constant variance. If a trend is represented by a linear function, i.e. $f(t) = c_1 t + c + 0$, then c_1 is a numeric characterization of a *tendency*. If $c_1 > 0 (c1 < 0)$, then the tendency is “increase” (“decrease”), or if $c_1 = 0$, then the tendency is “stagnation”.

A. F^1 -transform for trend extraction

We assume that a time series $\{x_t, t \in [1, T]\}$ can be decomposed in accordance with (5). Therefore, our task is to extract its trend. For this purpose, we propose the technique F^1 -transform that was described above. By applying the direct and inverse F^1 -transform, we represent a trend of a time series using the combination of $2 \leq n \leq T$ “basic functions” $A_1, ..., A_n$, which results in

$$f(t) = \sum_{k=1}^{n} F_k^1(t)A_k(t). \quad (6)$$

A justification that (6) can be chosen as a trend follows from the fact (local stationarity) that for all $k = 1, ..., n$, the F^0 - transform component of the function $x_t - F_k^1(t)$ is equal to zero. Based on the representation (2), we take a sequence of linear coefficients $\{c_{1,1}, ..., c_{n,1}\}$ of the corresponding F^1 -transform components and call it “sequence of local tendencies”.

In Figure 1, we show the time series #104 NN3 and its F^1 -transform component for the last interval of the partition.

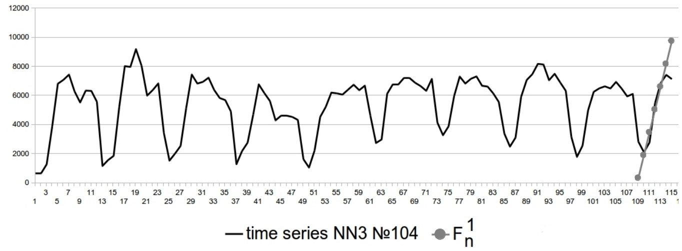

Fig. 1. F_n^1 component for time series example

IV. Creating groups of similar time series

Groups of similar time series will be created by analyzing local tendencies represented by coefficients $c_{k,1}$. For each pair x'_t , x''_t of time series from a set we calculate the Pearson correlation coefficient (linear correlation coefficient) of the corresponding sequences of local tendencies:

$$r_{x'_t,x''_t} = \frac{\sum_{k=1}^{n}(c'_{k,1} - \overline{c'_{k,1}})(c''_{k,1} - \overline{c''_{k,1}})}{\sqrt{\sum_{k=1}^{n}(c'_{k,1} - \overline{c'_{k,1}})^2 \sum_{k=1}^{n}(c''_{k,1} - \overline{c''_{k,1}})^2}} \tag{7}$$

We say that time series x'_t , x''_t are similar, if $r_{x'_t,x''_t} > \theta$ where θ is a chosen threshold.

An important requirement is equal lengths of all time series in a set. Moreover, absolute values of F^1 -transform components were normalized to the interval $[0, 1]$. Grouping of time series requires a choice of a threshold.

V. Experiments with the proposed method of grouping

Let us choose two sets of time series. The first set contains 14 time series with 37 observations each, collected between 1970 and 2000 in Russian Federation in areas of economics, production, social, health and culture indicators (RF statistics). The second set consists of 8 time series with 126 observations each, selected from the NN3 competition (http:// www.neural-forecasting-competition.com/ downloads/NN3/datasets/NN3_REDUCED.xls). For both sets, partitions were chosen in such a way that each basic function covers 7 points. The threshold θ for the correlation coefficient is set to 0.9.

In Table I, we show how many similar time series from the first set correspond to the chosen one.

Because the relation "to have the correlation coefficient greater than a threshold" is not transitive, we establish grouping by choosing a representative time series with the biggest number of similar ones, deleting its group from the set and repeating until the set is empty. In the first set, the biggest first group has representative time series #1 and consists of similar time series whose order numbers are: 1,2,3,7,8,9,10,11,12,13,14. Time series from the first group are shown below in Figure 2.

In Figure 3, we show time series that are not included into the first group. Their order numbers are: 4,5,6. All of them are similar with the representative time series #5.

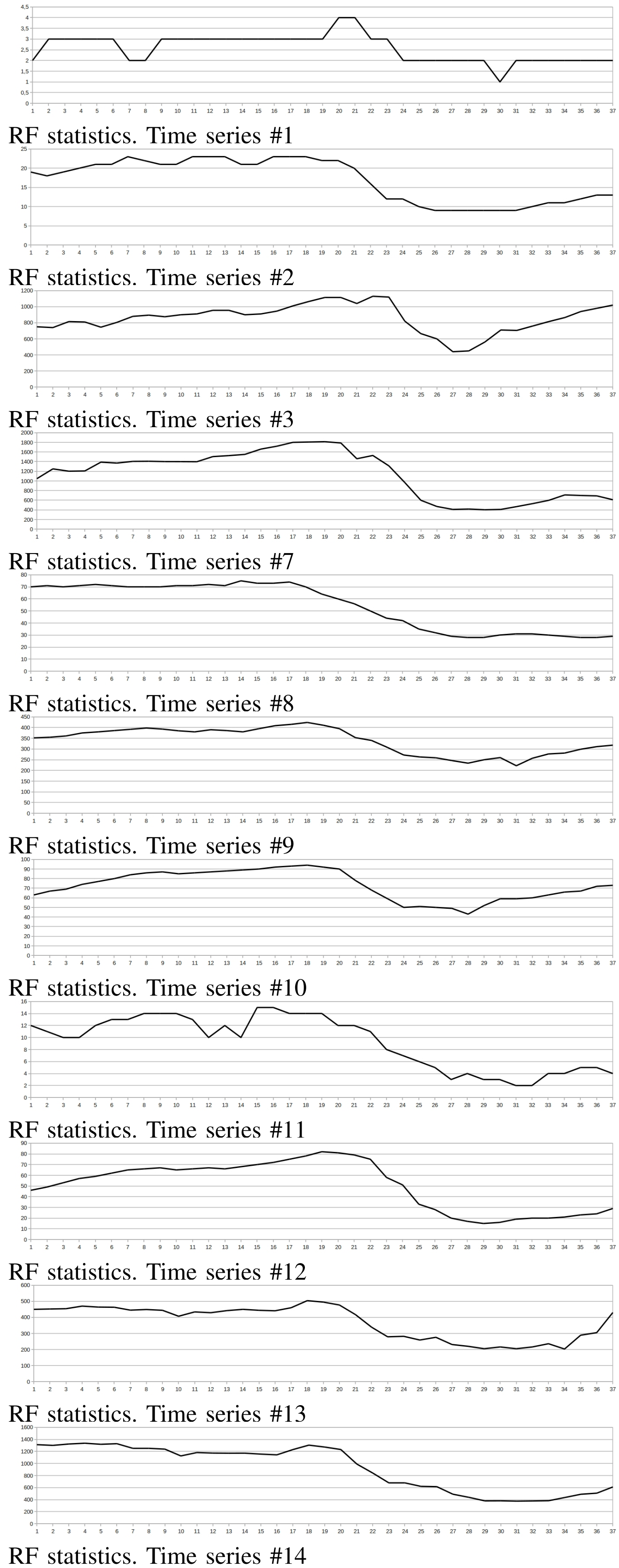

Fig. 2. The group of 11 similar time series from the first set (RF statistics) with the representative time series #1

In the second set (NN3 competition), there is only one group of similar time series that can be selected by the threshold 0.9. This group consists of time series with order numbers 1,4,5,6. In Figure 4, we show time series that are included into this group.

TABLE I. GROUPING OF SIMILAR TIME SERIES WITH CORRELATION COEFFICIENT GREATHER THAN 0.9

Time series	Number of similar time series
1	11
2	11
3	8
4	3
5	5
6	2
7	11
8	9
9	10
10	9
11	10
12	9
13	11
14	9

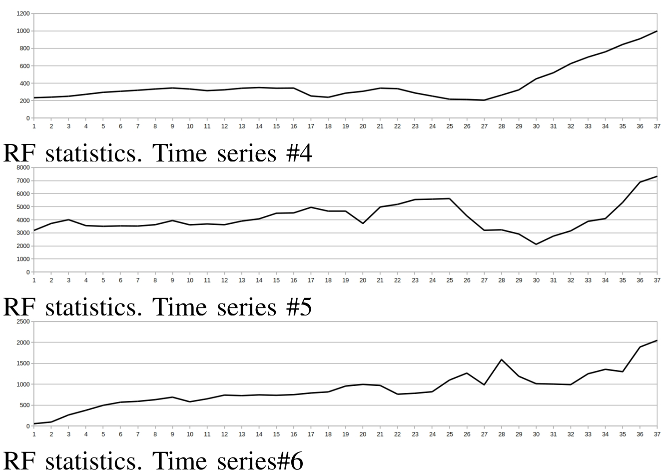

Fig. 3. The second group of 3 similar time series from the first set (RF statistics) with the representative time series #5

In Figure 5, we show time series from the second set (NN3 competition) that are not included into the group of similar ones. Their order numbers are: 2,3,7,8. All of them are not pairwise similar.

VI. FORECASTING OF LOCAL TENDENCIES

In this section, we show that prediction of a local tendency of a representative time series from a group of similar ones can be used for all time series from the group. By this we mean prediction of a time series of linear coefficients $c_{k,1}$, $k = 1, ..., n$, of the F^1 -transform components. Below, we give an illustrative example where we choose three similar time series with order numbers 1,2,3 from the first group of the set RF statistics with the representative time series #1. We made prediction of the coefficient $c_{n,1}$ from the time series $c_{k,1}, k = 1, ..., n-1$. This predicted value has been used in computation of the last F^1 -transform component F^1_n for time series #2 and #3. The results are illustrated in Figures 6, 7. It is obvious that both predicted components F^1_n of time series #2 and #3 can be used instead of actual ones.

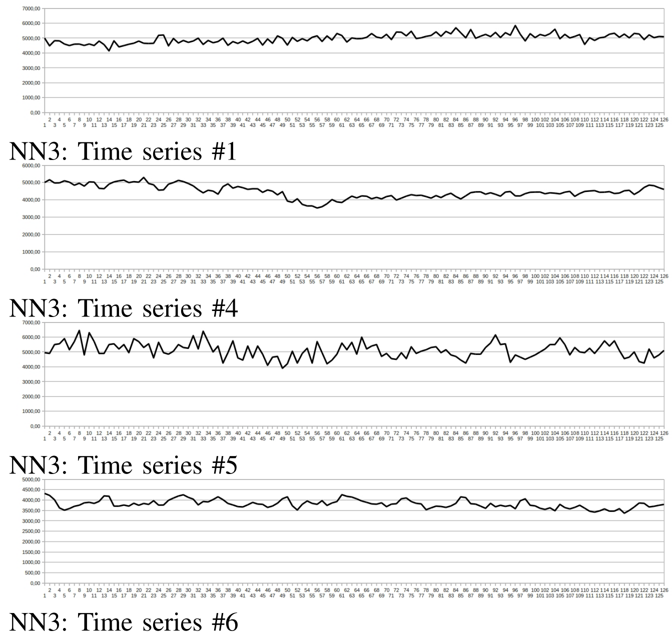

Fig. 4. The group of 4 similar time series from the second set (NN3 competition) with the representative time series #1.

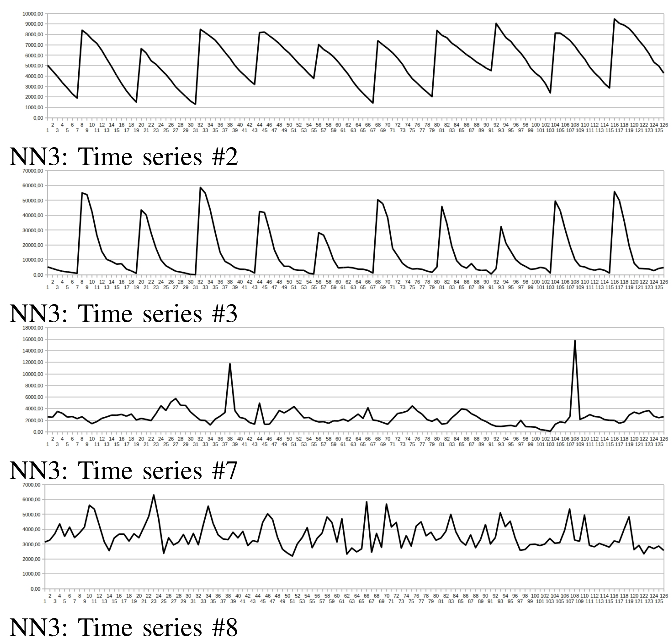

Fig. 5. 4 non - pairwise similar time series from the second set (NN3 competition).

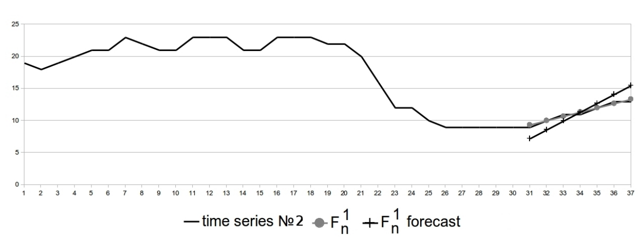

Fig. 6. Time series #2 (RF statistics). the last F^1 - transform component F^1_n was computed with the predicted from time series #1 coefficient $c_{n,1}$.

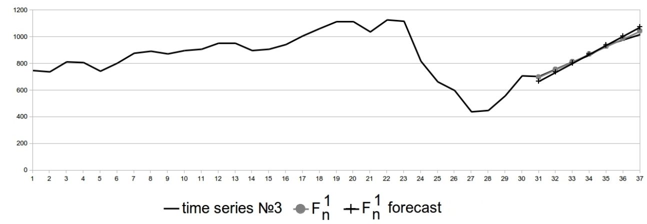

Fig. 7. Time series #3 (RF statistics). the last F^1 - transform component F_n^1 was computed with the predicted from time series #1 coefficient $c_{n,1}$.

VII. Conclusion

In the proposed contribution, we were focused on the new method of grouping time series according to their local tendency indicator that is expressed by a linear coefficient of the F^1 -transform. Groups of similar time series are created on the basis of the Pearson correlation coefficient of the corresponding sequences of local tendencies represented by coefficients $c_{k,1}$. The useful consequence of grouping is an effective procedure of forecasting such that only one time series from a group is forecasted. Two sets of real time series were chosen for experiments.

VIII. ACKNOWLEDGEMENT

The authors acknowledge that this paper was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and under a partial support of the projects MŠMT-7026/2012-36 (KONTAKT II, LH 12229 Research and development of methods and means of intelligent analysis of time series for the strategic planing problems) and RFFI-10-01-00183.

References

[1] Novák, V., M. Štěpnićka, A. Dvořák, I. Perfilieva, V. Pavliska, and L. Vavřičková, "Analysis of seasonal time series using fuzzy approach," Int. Journ. of General Systems, vol. 39, pp. 305-328, 2010.

[2] F. Di Martino, V. Loia, and S. Sessa, "Fuzzy transforms method in prediction data analysis," Fuzzy Sets and Systems, vol. 180, pp. 146-163, 2011.

[3] L. Troiano and P. Kriplani, "Supporting trading strategies by inverse fuzzy transform," Fuzzy Sets and Systems, vol. 180, pp. 121145, 2011.

[4] M. Štěpnićka, A. Dvořák, V. Pavliska, and L. Vavřičková, "A linguistic approach to time series modeling with the help of F-transform," Fuzzy Sets and Systems, vol. 180, pp. 164-184, 2011.

[5] I. Perfilieva, N. Yarushkina, T. Afanasieva, and A. Romanov, "Time series analysis by soft computing methods," Int. Journ. of General Systems, vol. 42, pp. 687705, 2013.

[6] Yarushkina, N. G., Principles of the theory of fuzzy and hybrid systems, Finances and statistics, Moscow, 2004.

[7] I. Perfilieva, "Fuzzy transforms: Theory and applications," Fuzzy Sets and Systems, vol. 157, pp. 9931023, 2006.

[8] I. Perfilieva, M. Danková, and B. Bede. "Towards a Higher Degree F-Transform." Fuzzy Sets and Systems, vol. 180, p. 319, 2011.