Why Fuzzy Partition in F-Transform?

Vladik Kreinovich[1], Olga Kosheleva[2], and Songsak Sriboonchitta[2]
[1]University of Texas at El Paso
El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu
[2]Faculty of Economics, Chiang Mai University, Thailand
songsakecon@gmail.com

Abstract

In many application problems, F-transform algorithms are very efficient. In F-transform techniques, we replace the original signal or image with a finite number of weighted averages. The use of weighted average can be naturally explained, e.g., by the fact that this is what we get anyway when we measure the signal. However, most successful applications of F-transform have an additional not-so-easy-to-explain feature: the partition requirement, that the sum of all the related weighting functions is a constant. In this paper, we show that this seemingly difficult-to-explain requirement can also be naturally explained in signal-measuring terms: namely, this requirement can be derived from the natural desire to have all the signal values at different moments of time estimated with the same accuracy.

1 Formulation of the Problem

F-transform: a brief reminder. In many practical applications, it turns out to be beneficial to replace the original continuous signal $x(t)$ defined on some time interval with a finite number of "averaged" values

$$x_i = \int A_i(t) \cdot x(t)\, dt, \quad i = 0, \ldots, n, \tag{1}$$

where $A_i(t) \geq 0$ are appropriate functions; see, e.g., [13, 14, 16, 17, 18, 19].

In many applications, a very specific form of these functions are used: namely, $A_i(t) = A(t - t_i)$ for some function $A(t)$ and for $t_i = t_0 + i \cdot h$, where t_0 and $h > 0$ are numbers for which $A(t)$ is equal to 0 outside the interval $[-h, h]$. However, more general families of functions $A_i(t)$ are also sometimes efficiently used.

The transition from the original function $x(t)$ to the tuple of values $\langle x_0, x_1, \ldots, x_n \rangle$ is known as the *F-transform*; [13, 14, 16, 17, 18, 19].

A similar 2-D transformation is very useful in many image processing problems.

The general idea behind F-transform is very reasonable. From the general measurement viewpoint, F-transform makes perfect sense – it corresponds to the results of measuring the signal. Indeed, in practice, a measuring instrument cannot measure the exact value $x(t)$ of the signal at a given moment t. No matter how fast the processes within the measuring instrument, it always has some inertia. As a result, the value m_i measured at each measurement depends not only on the value $x(t)$ of the signal at the given moment of time, it also depends on the values at nearby moments of time; see, e.g., [20].

The signal is usually weak, so the values $x(t)$ are small. Thus, we can expand the dependence of m_i on $x(t)$ in Taylor series and safely ignore terms which are quadratic or of higher order with respect to $x(t)$ and conclude that the value m_i is a linear function of different values $x(t)$; this is a usual technique in applications; see, e.g., [3]. The general form of a linear dependence is

$$m_i = m_i^{(0)} + \int A_i(t) \cdot x(t)\, dt \tag{2}$$

for some coefficients $A_i(t)$.

A measuring instrument is usually calibrated in such a way that in the absence of the signal, when $x(t) = 0$, the measurement result is 0. After such a calibration, we get $m_i^{(0)} = 0$ and thus, the expression (2) gets a simplified form

$$m_i = \int A_i(t) \cdot x(t)\, dt. \tag{3}$$

This is exactly the form used in F-transform. Thus, the F-transform is indeed a very natural procedure: it replaces the original signal $x(t)$ with the simulated results of measuring this signal – and the results of measuring the signal is exactly what we have in real life.

But why partition? So far, everything has been good and natural, but there is one aspect of successful applications of F-transform that cannot be explained so easily: namely, in most such applications, the corresponding functions $A_i(t)$ form a *partition*, in the sense that

$$\sum_{i=1}^{n} A_i(t) = 1 \tag{4}$$

for all moments t from the corresponding time interval.

What we do in this paper. In this paper, we show that the partition requirement (4) can also be naturally explained in the measurement interpretation of F-transform.

To be more precise, we show that what naturally appears is a 1-parametric family of similar requirements of which the partition requirement is a particular case, and then we explain that in the fuzzy cases, it is indeed reasonable to use the partition requirement.

2 Main Idea

What if we can exactly measure instantaneous values? In the idealized case, when inertia of measuring instruments is so small that it can be safely ignored, we can measure the exact values $x(t_1)$, $x(t_2)$, ..., of the signal $x(t)$ at different moments of time.

In this case, we get perfect information about the values of the signal at these moments of time t_1, t_2, ..., but practically no information about the values of the signal $x(t)$ at any other moment of time. In other words, we reconstruct the values $x(t_1)$, $x(t_2)$, ..., with perfect accuracy (0 measurement error), while the values $x(t)$ corresponding to all other moments of time t are reconstructed with no accuracy at all (the only bound on measurement error is infinity).

Even if we take into account that measurements are never 100% accurate, and we only measure the values $x(t_i)$ with some accuracy, we will still get the difference between our knowledge of values $x(t)$ corresponding to different moments of time:

- we know the values $x(t_i)$ with finite accuracy, but
- for all other moments of time t, we know nothing (i.e., the only bound of measurement error is infinity).

This difference does not fit well with the fact that we want to get a good representation of the *whole* signal $x(t)$, i.e., a good representation of its values at all moments of time. Thus, we arrive at the following idea.

Main idea. To adequately represent the original signal $x(t)$, it is desirable to select the measurement procedures in such a way that, based on these measurements, we reconstruct each value $x(t)$ with the same accuracy.

Comment. At this moment, we have presented this idea informally. In the following sections, we will show how to formalize this idea – and we also show that this idea leads to the partition requirement (or, to be more precise, to a general formula that includes the partition requirement as a particular case – with the argument of why namely the partition requirement should be selected in the fuzzy case).

3 Case of Probabilistic Uncertainty

Description of the case. Let us start with the most studied case of probabilistic uncertainty, when we have probabilistic information about the measurement error $\Delta m_i \stackrel{\text{def}}{=} \widetilde{m}_i - m_i$ of each measurement, where $\widetilde{m}_i$ denotes the result of measuring the quantity m_i.

We will consider the usual way measurement uncertainties are treated in this approach (see, e.g., [20]): namely, we will assume that each measurement error Δm_i is normally distributed with 0 mean and known standard deviation σ, and measurement errors Δm_i and Δm_j corresponding to different measurements $i \neq j$ are independent.

How accurately can we estimate $x(t)$ based on each measurement. Based on each measurement, we know each value $m_i = \int A_i(t) \cdot x(t)\, dt$ with accuracy σ. The integral is, in effect, a large sum, so we have

$$m_i = \sum_t A_i(t) \cdot x(t) \cdot \Delta t.$$

Thus, for each moment t, we have

$$A_i(t) \cdot x(t) \cdot \Delta t = m_i - \sum_{s \neq t} A_i(s) \cdot x(s) \cdot \Delta s, \tag{5}$$

and thus,

$$x(t) = \frac{1}{A_i(t) \cdot \Delta t} \cdot m_i - \frac{1}{A_i(t) \cdot \Delta t} \cdot m_i \cdot \sum_{s \neq t} A_i(s) \cdot x(s) \cdot \Delta s. \tag{6}$$

The measurement result $\widetilde{m}_i$ is an estimate for the quantity m_i, with mean 0 and standard deviation σ. Thus, if we know all the values $x(s)$ corresponding to $s \neq t$, then, based on the result $\widetilde{m}_i$ of the i-th measurement, we can estimate the remaining value $x(t)$ as

$$x(t) \approx \widetilde{x}_i(t) \stackrel{\text{def}}{=} \frac{1}{A_i(t) \cdot \Delta t} \cdot \widetilde{m}_i - \frac{1}{A_i(t) \cdot \Delta t} \cdot m_i \cdot \sum_{s \neq t} A_i(s) \cdot x(s) \cdot \Delta s. \tag{7}$$

By comparing the formulas (6) and (7), we can conclude that the approximation error $\Delta x_i(t) \stackrel{\text{def}}{=} \widetilde{x}_i(t) - x(t)$ of this estimate is equal to

$$\Delta x_i(t) = \frac{1}{A_i(t) \cdot \Delta t} \cdot \Delta m_i. \tag{8}$$

Since the measurement error Δm_i is normally distributed, with 0 mean and standard deviation σ, the approximation error $\Delta x_i(t)$ is also normally distributed, with 0 mean and standard deviation

$$\sigma_i(t) = \frac{\sigma}{A_i(t) \cdot \Delta t}. \tag{9}$$

How accurately can we estimate $x(t)$ based on all the measurements. For each moment t, based on each measurement i, we get an estimate $\widetilde{x}_i(t) \approx x(t)$ with the accuracy σ_i described by the formula (9).

$$\begin{aligned} x(t) &\approx \widetilde{x}_0(t), \\ x(t) &\approx \widetilde{x}_1(t), \\ &\ldots \\ x(t) &\approx \widetilde{x}_n(t). \end{aligned} \tag{10}$$

For each estimate, since the distribution of the measurement error is normal, the corresponding probability density function has the form

$$\rho_i(\widetilde{x}_i(t)) = \frac{1}{\sqrt{\pi}\cdot\sigma_i(t)}\cdot\exp\left(-\frac{(\widetilde{x}_i(t)-x(t))^2}{2(\sigma_i(t))^2}\right). \tag{11}$$

Since the measurement errors Δm_i of different measurements are independent, the resulting estimation errors $\Delta x_i(t) = \widetilde{x}_i(t) - x(t)$ are also independent. Thus, the joint probability density corresponding to all the measurements is equal to the product of all the values (11) corresponding to individual measurements:

$$\rho(\widetilde{x}_0(t),\ldots,\widetilde{x}_n(t)) = \frac{1}{(\sqrt{\pi})^{n+1}\cdot\prod\limits_{i=0}^{n}\sigma_i(t)}\cdot\exp\left(-\sum_{i=0}^{n}\frac{(\widetilde{x}_i(t)-x(t))^2}{2(\sigma_i(t))^2}\right). \tag{12}$$

As a combined estimate $\widetilde{x}(t)$ for $x(t)$, it is reasonable to select the value for which the corresponding probability (12) is the largest possible. This is known as the *Maximum Likelihood Method*; see, e.g., [21].

To find such a maximum, it is convenient to take the negative logarithm of the expression (12) and use the fact that $-\ln(z)$ is a decreasing function – so the original expression is the largest if and only if its negative logarithm is the smallest. Thus, we arrive at the need to minimize the sum

$$\sum_{i=0}^{n}\frac{(\widetilde{x}_i(t)-x(t))^2}{2\sigma_i^2}; \tag{13}$$

this minimization is known as the *Least Squares* approach.

Differentiating the expression (13) with respect to the unknown $x(t)$ and equating the derivative to 0, we conclude that

$$\sum_{i=0}^{n}\widetilde{x}_i(t)\cdot(\sigma_i(t))^{-2} = \widetilde{x}(t)\cdot\sum_{i=0}^{n}(\sigma_i(t))^{-2}, \tag{14}$$

and thus, that

$$\widetilde{x}(t) = \frac{\sum\limits_{i=0}^{n}\widetilde{x}_i(t)\cdot(\sigma_i(t))^{-2}}{\sum\limits_{i=0}^{n}(\sigma_i(t))^{-2}}. \tag{15}$$

The accuracy $\widetilde{\sigma}(t)$ of this estimate can be determined if we describe the expression (12) in the form

$$\frac{1}{\sqrt{\pi}\cdot\widetilde{\sigma}(t)}\cdot\exp\left(-\frac{(x(t)-\widetilde{x}(t))^2}{2(\widetilde{\sigma}(t))^2}\right). \tag{16}$$

By comparing the coefficients at $(x(t))^2$ under the exponent in the formulas (12) and (16), we conclude that

$$\frac{1}{2(\widetilde{\sigma}(t))^2} = \sum_{i=0}^{n}\frac{1}{2(\sigma_i(t))^2}, \tag{17}$$

i.e., equivalently, that

$$(\widetilde{\sigma}(t))^{-2} = \sum_{i=0}^{n} (\sigma_i(t))^{-2}. \tag{18}$$

In particular, if all the estimation errors were equal, i.e., if we had $\sigma_i(t) = \sigma(t)$ for all i, then, from (18), we would conclude that

$$\widetilde{\sigma}(t) = \frac{\sigma(t)}{\sqrt{N}}, \tag{19}$$

where $N \stackrel{\text{def}}{=} n + 1$ is the overall number of combined measurements.

Substituting expression (9) for $\sigma_i(t)$ into the formula (18), we conclude that

$$(\widetilde{\sigma}(t))^{-2} = \frac{(\Delta t)^2}{\sigma^2} \cdot \sum_{i=0}^{n} (A_i(t))^2. \tag{20}$$

Thus, the requirement that we get the same accuracy for all moments of time t, i.e., that $\widetilde{\sigma}(t) = \text{const}$ means that we need to have

$$\sum_{i=0}^{n} (A_i(t))^2 = \text{const}. \tag{21}$$

Discussion. The formula (21) is somewhat similar to the partition requirement but it is different: in the partition requirement, we demand that the sum of the functions $A_i(t)$ be constant, but here we have the sum of the squares. On the other hand, the formula (21) is based on the probabilistic uncertainty, for which the measurement error decreases with repeated measurements as $1/\sqrt{N}$, while, e.g., for interval uncertainty (see, e.g., [5, 7, 8, 12, 20]), the measurement error resulting from N repeated measurements decreases as $1/N$; see, e.g. [22]. So maybe by considering different types of uncertainty, we can get the partition formula? To answer this question, let us consider a general case of how uncertainties can be combined in different approaches.

4 How Uncertainties Can Be Combined in Different Approaches

Towards a general formulation of the problem. In the general case, be it probabilistic or interval or any other approach, we can always describe the corresponding uncertainty in the same unit as the measured quantity.

In the interval approach, a natural measure of uncertainty is the largest possible value Δ of the absolute value $|\Delta x|$ of the approximation error $\Delta x = \widetilde{x} - x$, where x is the actual value of the corresponding quantity and $\widetilde{x}$ is the measurement result. This value is clearly measured in the same units as the quantity x itself.

In the probabilistic approach, we can use the variance of Δx – which is described in different units than x – but we can also take the square root of this variance and consider standard deviation σ, which is already in the same units.

In the general case, let us denote the corresponding measure of accuracy by Δ. The situation when we have no information about the desired quantity corresponds to $\Delta = \infty$. The idealized situation when we know the exact value of this quantity corresponds to $\Delta = 0$.

If Δ and Δ' are corresponding measures of accuracy for two different measurements, then what is the accuracy of the resulting combined estimate? Let us denote this combined accuracy by $\Delta * \Delta'$.

In these terms, to describe the combination, we need to describe the corresponding function $a * b$ of two variables. What are the natural properties of this function?

Commutativity. The result of combining two estimates should not depend on which of the two estimates is listed first, so we should have $a * b = b * a$. In other words, the corresponding combination operation must be commutative.

Associativity. If we have three estimates, then:

- we can first combine the first and the second ones, and then combine the result with the third one,
- or we can first combine the second and the third ones, and then combine the result with the first one.

The result should not depend on the order, so we should have $(a*b)*c = a*(b*c)$. In other words, the corresponding operation should be associative.

Monotonicity. An additional information can only improve the accuracy. Thus, the accuracy of the combined estimate cannot be worse than the accuracy of each of the estimates used in this combination. So, we get $a * b \le a$.

Similarly, if we increase the accuracy of each measurement, the accuracy of the resulting measurement will increase too: if $a \le a'$ and $b \le b'$, then we should have $a * b \le a' * b'$.

Non-degenerate case. If we start with measurements of finite accuracy, we should never get the exact value, i.e., if $a > 0$ and $b > 0$, we should get $a * b > 0$.

Scale-invariance. In real life, we deal with the actual quantities, but in computations, we need to describe these quantities by their numerical values. To get a numerical value, we need to select a measuring unit: e.g., to describe distance in numerical terms, we need to select a unit of distance.

This selection is usually arbitrary. For example, for distance, we could consider meters, we could consider centimeters, and we could consider inches or feet. It is reasonable to require that the combination operation remains the same if we keep the same quantities but change the measuring unit. Let us describe this requirement in precise terms.

If we replace the original measuring unit with a new one which is λ times smaller, then all the numerical values are multiplied by λ. For example, if we

replace meters by centimeters, then all the numerical values are multiplied by 100. The corresponding transformation $x \to \lambda \cdot x$ is known as *scaling*.

Suppose that in the original units, we had accuracies a and b and the combined accuracy was $a * b$. Then, in the new units – since accuracies are described in the same units as the quantity itself – the original accuracies become $\lambda \cdot a$ and $\lambda \cdot b$, and the combined accuracy is thus $(\lambda \cdot a) * (\lambda \cdot b)$. This is the combined accuracy in the new units. It should be the same as when we transform the old-units accuracy $c = a * b$ into the new units, getting $\lambda \cdot (a * b)$: $(\lambda \cdot a) * (\lambda \cdot b) = \lambda \cdot (a * b)$. This invariance under scaling is known as *scale-invariance.*

Discussion. Now, we are ready to formulate the main result.

Definition 1. *By a* combination operation*, we mean a function $a * b$ that transforms two non-negative numbers a and b into a new non-negative number and for which the following properties hold:*

- *for all a and b, we have $a * b = b * a$ (*commutativity*);*
- *for all a, b, and c, we have $(a * b) * c = a * (b * c)$ (*associativity*);*
- *for all a and b, we have $a * b \le a$ (*first monotonicity requirement*);*
- *for all a, b, a', and b', if $a \le a'$ and $b \le b'$, then $a * b \le a' * b'$ (*second monotonicity requirement*);*
- *if $a > 0$ and $b > 0$, then $a * b > 0$ (*non-degeneracy*); and*
- *for all a, b, and $\lambda > 0$, we have $(\lambda \cdot a) * (\lambda \cdot b) = \lambda \cdot (a * b)$ (*scale-invariance*).*

Proposition 1. *Every combination operation has either the form $a * b = \min(a, b)$ or the form $a * b = (a^{-\beta} + b^{-\beta})^{-1/\beta}$ for some $\beta > 0$.*

Proof of this result is, in effect, described in [1] (see also [4]).

Comment. The proof shows that if we do not impose the non-degeneracy condition, the only other alternative is $a * b = 0$. Thus, the non-degeneracy condition can be weakened: instead of requiring that $a * b > 0$ for *all* pairs of positive numbers a and b, it is sufficient to require that $a * b > 0$ for at least *one* such pair.

Discussion. The form $\min(a, b)$ is the limit case of the second form when $\beta \to \infty$.

In the generic case $\beta < \infty$, $a * b = c$ is equivalent to

$$a^{-\beta} + b^{-\beta} = c^{-\beta}. \tag{22}$$

Thus, the probabilistic case corresponds to $\beta = 2$.

In the situation when we have N measurement results with the same accuracy $\Delta_1 = \ldots = \Delta_N = \Delta$, the combined accuracy $\widetilde{\Delta}$ can be determined from the formula

$$\left(\widetilde{\Delta}\right)^{-\beta} = N \cdot \Delta^{-1/\beta}. \tag{23}$$

Thus, we have

$$\widetilde{\Delta} = \frac{\Delta}{N^{1/\beta}}. \tag{24}$$

In the probabilistic case, we indeed have this formula with $\beta = 2$. The above interval-case formula corresponds to the case $\beta = 1$; thus, this is the value of the parameter β corresponding to interval uncertainty.

5 Which Functions $A_i(t)$ Should We Choose: Case of General Uncertainty

Analysis of the problem. If we measure m_i with accuracy Δ, then, due to the formula (8) (and similar to the case of probabilistic uncertainty), the estimate $\widetilde{x}_i(t)$ is known with accuracy

$$\Delta_i(t) = \frac{\Delta}{A_i(t) \cdot \Delta t}. \tag{25}$$

For the case of min combination formula, the combined accuracy is equal to

$$\Delta(t) = \min_i \Delta_i(t) = \frac{\Delta}{\Delta t} \cdot \frac{1}{\max\limits_i A_i(t)}. \tag{26}$$

Thus, the requirement that we estimate all the values $x(t)$ with the same accuracy means that

$$\max_i A_i(t) = \text{const}. \tag{27}$$

For the generic case $\beta < \infty$, from the formula (22), we conclude that

$$(\Delta(t))^{-\beta} = \frac{(\Delta t)^\beta}{\Delta^\beta} \cdot \sum_{i=1}^{n} (A_i(t))^\beta. \tag{28}$$

Thus, the requirement that we get the same accuracy for all moments of time t means that we need to have

$$\sum_{i=1}^{n} (A_i(t))^\beta = \text{const}. \tag{29}$$

General conclusion. The requirement that we get the same accuracy for reconstructing the value of the signal at each moment of time of t leads to condition (27) or (29). In particular, for $\beta = 1$, we get the partition property.

Which value β should we use in the case of fuzzy uncertainty. In the fuzzy case (see, e.g., [2, 6, 9, 10, 11, 15, 23]), the usual way of propagating uncertainty – Zadeh extension principle – is equivalent to applying interval computations for each α-cut. Thus, for analyzing fuzzy data, it makes sense to use the value of β corresponding to interval uncertainty – which, as we have mentioned, is $\beta = 1$. For $\beta = 1$, the formula (29) becomes the partition property. Thus, *the partition property indeed makes perfect sense when analyzing fuzzy data.*

Acknowledgments

This work was supported by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand, and by the US National Science Foundation via grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] K. Autchariyapanitkul, O. Kosheleva, V. Kreinovich, and S. Sriboonchitta, "Quantum econometrics: how to explain its quantitative successes and how the resulting formulas are related to scale invariance, entropy, and fuzziness", In: V.-N. Huynh, M. Inuiguchi, D.-H. Tran, and T. Denoeux (eds.), *Proceedings of the International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making IUKM'2018*, Hanoi, Vietnam, March 13–15, 2018.

[2] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[3] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[4] A. Gholamy, J. Parra, V. Kreinovich, O. Fuentes, and E. Anthony, "How to best apply deep neural networks in geosciences: towards optimal 'averaging' in dropout training", In: J. Watada, S. C. Tan, P. Vasant, E. Padmanabhan, and L. C. Jain (eds.), *Smart Unconventional Modelling, Simulation and Optimization for Geosciences and Petroleum Engineering*, Springer Verlag, 2019, pp. 15–26.

[5] L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2001.

[6] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[7] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.

[8] G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.

[9] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[10] H. T. Nguyen and V. Kreinovich, "Nested intervals and sets: concepts, relations to fuzzy sets, and applications", In: R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[11] H. T. Nguyen, C. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[12] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.

[13] V. Novak, I. Perfilieva, M. Holcapek, and V. Kreinovich, "Filtering out high frequencies in time series using F-transform", *Information Sciences*, 2014, Vol. 274, pp. 192–209.

[14] V. Novak, I. Perfilieva, and V. Kreinovich, "F-transform in the analysis of periodic signals", In: M. Inuiguchi, Y. Kusunoki, and M. Seki (eds.), *Proceedings of the 15th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Osaka, Japan, September 24–27, 2012.

[15] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[16] I. Perfilieva, "Fuzzy transforms: Theory and applications", *Fuzzy Sets and Systems*, 2006, Vol. 157, pp. 993–1023.

[17] I. Perfilieva, "F-transform", In: *Springer Handbook of Computational Intelligence*, Springer Verlag, 2015, pp. 113–130.

[18] I. Perfilieva, M. Danková, and B. Bede, "Towards a higher degree F-transform", *Fuzzy Sets and Systems*, 2011, Vol. 180, No. 1, pp. 3–19.

[19] I. Perfilieva, V. Kreinovich, and V. Novak, "F-transform in view of trend extraction", In: M. Inuiguchi, Y. Kusunoki, and M. Seki (eds.), *Proceedings of the 15th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Osaka, Japan, September 24–27, 2012.

[20] S. G. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer, New York, 2005.

[21] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

[22] G. W. Walster and V. Kreinovich, "For unknown–but–bounded errors, interval estimates are often better than averaging", *ACM SIGNUM Newsletter*, 1996, Vol. 31, No. 2, pp. 6–19.

[23] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.