



Article

# Why Fuzzy Partition in F-Transform?

Vladik Kreinovich <sup>1,†,‡</sup> , Olga Kosheleva <sup>1,‡</sup>  and Songsak Sriboonchitta <sup>2,‡\*</sup>

<sup>1</sup> University of Texas at El Paso; vladik@utep.edu, olgak@utep.edu

<sup>2</sup> Chiang Mai University, Thailand; songsakecon@gmail.com

\* Correspondence: vladik@utep.edu; Tel.: +1-915-747-6951 (V.K.)

† Current address: Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA

‡ These authors contributed equally to this work.

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

10 **Keywords:** F-transform; fuzzy partition; measurements; measurement uncertainty

## 11 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

$$x_i = \int A_i(t) \cdot x(t) dt, \quad i = 0, \dots, n, \quad (1)$$

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

13 In many applications, a very specific form of these functions are used: namely,  $A_i(t) = A(t - t_i)$   
 14 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  
 15 to 0 outside the interval  $[-h, h]$ . However, more general families of functions  $A_i(t)$  are also sometimes  
 16 efficiently used.

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

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

20 **The general idea behind F-transform is very reasonable.** From the general measurement viewpoint,  
 21 F-transform makes perfect sense – it corresponds to the results of measuring the signal. Indeed, in  
 22 practice, a measuring instrument cannot measure the exact value  $x(t)$  of the signal at a given moment  
 23  $t$ . No matter how fast the processes within the measuring instrument, it always has some inertia. As a  
 24 result, the value  $m_i$  measured at each measurement depends not only on the value  $x(t)$  of the signal at  
 25 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

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

26 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

$$m_i = \int A_i(t) \cdot x(t) dt. \quad (3)$$

27 This is exactly the form used in F-transform. Thus, the F-transform is indeed a very natural procedure:  
28 it replaces the original signal  $x(t)$  with the simulated results of measuring this signal – and the results  
29 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

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

30 for all moments  $t$  from the corresponding time interval.

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

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

## 36 2. Main Idea

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

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

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

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

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

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

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

### 61 3. Case of Probabilistic Uncertainty

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

65 We will consider the usual way measurement uncertainties are treated in this approach (see, e.g.,  
 66 [20]): namely, we will assume that each measurement error  $\Delta m_i$  is normally distributed with 0 mean  
 67 and known standard deviation  $\sigma$ , and measurement errors  $\Delta m_i$  and  $\Delta m_j$  corresponding to different  
 68 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  $\sigma$ . 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

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

and thus,

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

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

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

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

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

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

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

**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  $\tilde{x}_i(t) \approx x(t)$  with the accuracy  $\sigma_i$  described by the formula (9).

$$\begin{aligned} x(t) &\approx \tilde{x}_0(t), \\ x(t) &\approx \tilde{x}_1(t), \\ &\dots \\ x(t) &\approx \tilde{x}_n(t). \end{aligned} \quad (10)$$

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

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

Since the measurement errors  $\Delta m_i$  of different measurements are independent, the resulting estimation errors  $\Delta x_i(t) = \tilde{x}_i(t) - x(t)$  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:

$$\rho(\tilde{x}_0(t), \dots, \tilde{x}_n(t)) = \frac{1}{(\sqrt{\pi})^{n+1} \cdot \prod_{i=0}^n \sigma_i(t)} \cdot \exp\left(-\sum_{i=0}^n \frac{(\tilde{x}_i(t) - x(t))^2}{2(\sigma_i(t))^2}\right). \quad (12)$$

70 As a combined estimate  $\tilde{x}(t)$  for  $x(t)$ , it is reasonable to select the value for which the corresponding  
71 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

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

72 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

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

and thus, that

$$\tilde{x}(t) = \frac{\sum_{i=0}^n \tilde{x}_i(t) \cdot (\sigma_i(t))^{-2}}{\sum_{i=0}^n (\sigma_i(t))^{-2}}. \quad (15)$$

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

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

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

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

i.e., equivalently, that

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

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

$$\tilde{\sigma}(t) = \frac{\sigma(t)}{\sqrt{N}}, \quad (19)$$

73 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

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

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

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

74

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

#### 83 4. How Uncertainties Can Be Combined in Different Approaches

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

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

91 In the probabilistic approach, we can use the variance of  $\Delta x$  – which is described in different units  
92 than  $x$  – but we can also take the square root of this variance and consider standard deviation  $\sigma$ , which  
93 is already in the same units.

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

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

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

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

104 **Associativity.** If we have three estimates, then:

- 105     • we can first combine the first and the second ones, and then combine the result with the third  
 106         one,  
 107     • or we can first combine the second and the third ones, and then combine the result with the first  
 108         one.

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

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

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

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

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

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

125         If we replace the original measuring unit with a new one which is  $\lambda$  times smaller, then all the  
 126 numerical values are multiplied by  $\lambda$ . For example, if we replace meters by centimeters, then all the  
 127 numerical values are multiplied by 100. The corresponding transformation  $x \rightarrow \lambda \cdot x$  is known as  
 128 *scaling*.

129         Suppose that in the original units, we had accuracies  $a$  and  $b$  and the combined accuracy was  $a * b$ .  
 130 Then, in the new units – since accuracies are described in the same units as the quantity itself – the  
 131 original accuracies become  $\lambda \cdot a$  and  $\lambda \cdot b$ , and the combined accuracy is thus  $(\lambda \cdot a) * (\lambda \cdot b)$ . This is  
 132 the combined accuracy in the new units. It should be the same as when we transform the old-units  
 133 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  
 134 under scaling is known as *scale-invariance*.

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

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

- 138     • for all  $a$  and  $b$ , we have  $a * b = b * a$  (commutativity);  
 139     • for all  $a$ ,  $b$ , and  $c$ , we have  $(a * b) * c = a * (b * c)$  (associativity);  
 140     • for all  $a$  and  $b$ , we have  $a * b \leq a$  (first monotonicity requirement);  
 141     • for all  $a$ ,  $b$ ,  $a'$ , and  $b'$ , if  $a \leq a'$  and  $b \leq b'$ , then  $a * b \leq a' * b'$  (second monotonicity requirement);  
 142     • if  $a > 0$  and  $b > 0$ , then  $a * b > 0$  (non-degeneracy); and  
 143     • for all  $a$ ,  $b$ , and  $\lambda > 0$ , we have  $(\lambda \cdot a) * (\lambda \cdot b) = \lambda \cdot (a * b)$  (scale-invariance).

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

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

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

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

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

$$a^{-\beta} + b^{-\beta} = c^{-\beta}. \quad (22)$$

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

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

$$(\tilde{\Delta})^{-\beta} = N \cdot \Delta^{-1/\beta}. \quad (23)$$

Thus, we have

$$\tilde{\Delta} = \frac{\Delta}{N^{1/\beta}}. \quad (24)$$

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

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

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

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

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

$$\Delta(t) = \min_i \Delta_i(t) = \frac{\Delta}{\Delta t} \cdot \frac{1}{\max_i A_i(t)}. \quad (26)$$

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

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

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

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

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

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

157

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

161 **Which value  $\beta$  should we use in the case of fuzzy uncertainty.** In the fuzzy case (see, e.g., [2,6,9–  
162 11,15,23]), the usual way of propagating uncertainty – Zadeh extension principle – is equivalent to  
163 applying interval computations for each  $\alpha$ -cut. Thus, for analyzing fuzzy data, it makes sense to use  
164 the value of  $\beta$  corresponding to interval uncertainty – which, as we have mentioned, is  $\beta = 1$ . For



165  $\beta = 1$ , the formula (29) becomes the partition property. Thus, *the partition property indeed makes perfect*  
 166 *sense when analyzing fuzzy data.*

167 **Author Contributions:** The authors contributed equally to this work.

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

172 **Acknowledgments:** The authors are thankful to Irina Perfilieva for encouragement and valuable discussions.

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

## 174 References

- 175 1. Autcharyapanitkul, A.; Kosheleva, O.; Kreinovich, V.; Sriboonchitta, S. Quantum econometrics: how to  
 176 explain its quantitative successes and how the resulting formulas are related to scale invariance, entropy, and  
 177 fuzziness, In: *Proceedings of the International Symposium on Integrated Uncertainty in Knowledge Modelling and*  
 178 *Decision Making IUKM'2018*, V.-N. Huynh, M. Inuiguchi, D.-H. Tran, and T. Denoeux (eds.), Hanoi, Vietnam,  
 179 March 13–15, 2018.
- 180 2. Belohlavek, R.; Dauben, J.W.; Klir, G.J. *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University  
 181 Press: New York, 2017.
- 182 3. Feynman, R.; Leighton, R.; Sands, M. *The Feynman Lectures on Physics*, Addison Wesley: Boston, Massachusetts,  
 183 2005.
- 184 4. Gholamy, A.; Parra, J.; Kreinovich, V.; Fuentes, O.; Anthony, E. How to best apply deep neural networks in  
 185 geosciences: towards optimal 'averaging' in dropout training, In: *Smart Unconventional Modelling, Simulation*  
 186 *and Optimization for Geosciences and Petroleum Engineering*, J. Watada, S. C. Tan, P. Vasant, E. Padmanabhan,  
 187 and L. C. Jain (eds.), Springer Verlag, 2019, pp. 15–26.
- 188 5. Jaulin, L.; Kiefer, M.; Didrit, O.; Walter, E. *Applied Interval Analysis, with Examples in Parameter and State*  
 189 *Estimation, Robust Control, and Robotics*, Springer: London, 2001.
- 190 6. Klir, G.; Yuan, B. *Fuzzy Sets and Fuzzy Logic*, Prentice Hall: Upper Saddle River, New Jersey, 1995.
- 191 7. Kreinovich, V.; Lakeyev, A.; Rohn, J.; Kahl, P. *Computational Complexity and Feasibility of Data Processing and*  
 192 *Interval Computations*, Kluwer: Dordrecht, 1998.
- 193 8. Mayer, G. *Interval Analysis and Automatic Result Verification*, de Gruyter: Berlin, 2017.
- 194 9. Mendel, J.M. *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer: Cham, Switzerland,  
 195 2017.
- 196 10. Nguyen, H.T.; Kreinovich, V. Nested intervals and sets: concepts, relations to fuzzy sets, and applications, In:  
 197 *Applications of Interval Computations*, Kearfott, R.B.; Kreinovich, V. (eds.), Kluwer: Dordrecht, 1996, pp. 245–290.
- 198 11. Nguyen, H.T.; Walker, C.; Walker, E.A. *A First Course in Fuzzy Logic*, Chapman and Hall/CRC: Boca Raton,  
 199 Florida, 2019.
- 200 12. Moore, R.E.; Kearfott, R.B.; Cloud, M.J. *Introduction to Interval Analysis*, SIAM: Philadelphia, 2009.
- 201 13. Novak, V.; Perfilieva, I.; Holcapek, M.; Kreinovich, V. Filtering out high frequencies in time series using  
 202 F-transform, *Information Sciences* **2014**, 274, 192–209.
- 203 14. Novak, V.; Perfilieva, I.; Kreinovich, V. F-transform in the analysis of periodic signals, In: *Proceedings of the 15th*  
 204 *Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Inuiguchi, M., Kusunoki,  
 205 Y., Seki, M. (eds.), Osaka, Japan, September 24–27, 2012.
- 206 15. Novák, V.; Perfilieva, I.; Močkoř, J. *Mathematical Principles of Fuzzy Logic*, Kluwer: Boston, Dordrecht, 1999.
- 207 16. Perfilieva, I. Fuzzy transforms: Theory and applications, *Fuzzy Sets and Systems* **2006**, 157, 993–1023.
- 208 17. Perfilieva, I. F-transform, In: *Springer Handbook of Computational Intelligence*, Springer Verlag, 2015, pp. 113–130.
- 209 18. Perfilieva, I., Danková, M., Bede, B. Towards a higher degree F-transform", *Fuzzy Sets and Systems* **2011**, 180(1),  
 210 3–19.
- 211 19. Perfilieva, I.; Kreinovich, V.; Novak, V. F-transform in view of trend extraction, In: *Proceedings of the 15th*  
 212 *Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Inuiguchi, M., Kusunoki,  
 213 Y., Seki, M. (eds.), Osaka, Japan, September 24–27, 2012.
- 214 20. Rabinovich, S.G. *Measurement Errors and Uncertainties: Theory and Practice*, Springer: New York, 2005.



- 215 21. Sheskin, D.J. *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC: Boca  
216 Raton, Florida, 2011.
- 217 22. Walster, G.W.; Kreinovich, V. For unknown-but-bounded errors, interval estimates are often better than  
218 averaging, *ACM SIGNUM Newsletter* **1996**, 31(2), 6–19.
- 219 23. Zadeh, L.A. Fuzzy sets, *Information and Control* **1965**, 8, 338–353.