Michał URBAŃSKI[1], Kinga WÓJCICKA[1], Paweł WÓJCICKI[2]

Warsaw University of Technology, Faculty of Physics (1), Faculty of Mathematics and Information Sciences (2)

doi:10.15199/48.2023.11.39

Fuzzy approach with t-norm arithmetic to analyze measurement uncertainty - concept and algorithm

Abstract. *The paper shows algorithms for determining uncertainty using methods A and B in the fuzzy set model.The compounding of type A uncertainty with a systematic component of a data is given as an arithmetic sum.*

Streszczenie. *W pracy pokazano algorytmy wyznaczania niepewności metodą A i B w modelu zbiorów rozmytych. Składanie niepewności typu A ze składową systematyczną dane jest sumą arytmetyczną.* ***(Analiza Niepewności Pomiaru w Modelu Zbiorów rozmytych z arytmetyką opartą na t-normach - koncepcja i algorytm)***

Keywords: fuzzy sets, fuzzy variable, uncertainty, theory of measurements, probability to possibility transformation
Słowa kluczowe: zbiory rozmyte, zmienna rozmyta, niepewność, teoria pomiarów, transformacja prawdopodobieństwo - zbiór rozmyty

Introduction

In accordance with the recommendations of the „Guide to the Expression of Uncertainty in Measurement" (in short GUM [2]) estimation methods of uncertainty can be divided into two groups depending on the source of information: A – series of repeated measurements treated as a random sample of a certain distribution, B – other sources of information which form the basis of expert analysis. The implementation of these recommendations is possible in both probabilistic and fuzzy sets models.

The uncertainty estimation method A assumes that the measured quantity is a random variable whose distribution and parameters can be estimated from the measurement data treated as a sample of a random variable representing the measurement process. Statistical methods provide estimators of the probability distribution of a random variable as well as estimators of distribution parameters such as the expected value, median, standard deviation or confidence interval of these estimators. In order to construct a fuzzy set uncertainty model, it is proposed to use a transformation from probability to possibility to construct fuzzy variable distribution and α cuts of fuzzy sets estimators. We have proposed such a transformation in previous papers [5, 6, 8].

Method B consists in determining, on the basis of expert knowledge on measurement system, an a priori distribution of errors. The probability distribution determined in this way describes the decision-making process of choosing this a-priori distribution, but it does not represent the random processes taking place in the measurement system. The metrologist's choice of a probability distribution to describe B-errors is not equivalent to obtaining an empirical distribution from a data series. The choice of the distribution is a view, not an empirical experience, and fuzzy sets have been created to mathematically describe the views. Fuzzy sets were proposed precisely to build a mathematical model describing the aggregation processes of various empirical inaccurate data, and the fuzzy variable distribution function describes the processes of expert analysis rather than an error probability distribution. Therefore, fuzzy set mathematics has tools appropriate for describing B-type errors.

In our opinion, the application of the probability theory based on an additive measure to type B errors is unjustified, because the decision-making process is not a random process, but the choice of a better solution as described by the maximal measure.

Method A is based on the assumption that the measurement data can be treated as a random sample of a stochastic process representing random phenomena in the tested object and measuring instruments.

A combination of both methods is possible in fuzzy sets if we assume a probability-to-possibility transformation that transforms the extended probability uncertainty into the radius of the corresponding $-\alpha$ cut of the fuzzy set representing the measurement results.

In this paper, we assume that the probabilistic model differs from the fuzzy model in the property of a measure describing measurement errors: in the probabilistic model, the probability distribution measure describing observed data dispersions is an additive measure, while in the fuzzy model, the possibility of occurrence of these data is described by a maximal measure.

Measurement

Measurement one can define as a process of assigning values, most often numbers to the objects, what is described as a mapping:

$$\Phi : \Omega \to \mathbb{R}, \tag{1}$$

where Ω is a set of measurable objects of reality, Φ is a mapping which represents the measurement scale, $\mathbb{R}$ is the sets of reals that represent values of measured quantities.

The equation (1) can describe both an ideal measurement (with no errors) and a measurement whose result is inaccurate. This equation can be looked at in several ways:

1. the measurement is an ideal when the measurement result is represented by one number,
2. the equation (1) defines a random variable, then the result of the measurement is represented by the distribution of the random variable,
3. the equation (1) defines the fuzzy variable([8]), the distribution of the fuzzy variable value is defined by the membership function of a fuzzy set - i. e. a fuzzy set,

– The measurement in the interval representation is a special case of a fuzzy variable model when the membership function has a uniform distribution over a interval (rectangular distribution).

In the probabilistic model, the measurement mapping Φ is a random variable with normalized and additive measure, while in the fuzzy set model, Φ is a fuzzy variable with a possibility distribution which is normalized and maximal measure. The ideal measure is if the distribution is one-point, where the distribution can be both the point distribution of probability and the point distribution of possibility.

Measurement at the interval scale can be described as a special case of fuzzy distribution when the distribution is uniform over a compact section (within the interval), i. e. when the membership function (the fuzzy set) is a characteristic function of the interval. Uniform probability distribution does not impose isomorphic arithmetic with interval arithmetic [1]. For example, the distribution of the sum of two independent random variables with the same uniform distribution has a triangular distribution. According to arithmetic, the interval sum of the intervals is an interval, i.e. the sum is described as a uniform distribution. This is only possible if the added random variables are closely correlated.

Measurement error

Within both the probabilistic and fuzzy model, the measurement can be described by the equation:

$$X = x_0 + \delta X \quad (2)$$

This equation means that there is a number x_0 that characterizes an ideal measurement, but due to measurement errors that always occur the result of the measurement is different from the true value, this value of difference we call error δ. The value of x_0 is called the true value.

The error δX is not known but we describe it with the error distribution function. Such a distribution function should have mathematical properties adequate to describe the phenomenon describing the mechanisms of error generation. In the literature, two mathematical models can be found describing measurement data dispersion and measurement inaccuracies: probabilistic and fuzzy sets. These models differ in the property of the measure we use to determine the distribution of errors. The probabilistic measure is additive but the measure of possibility (which is the basis of fuzzy sets) is maximal [4, 8]. These models differ in their interpretation of the mechanisms of error generation. In the probabilistic model, the measurement is understood as a draw, with a certain probability distribution from a set of values of a random variable representing the measured quantity.

In the fuzzy set model, we do not deal with the process of randomized trial from the sample space [8]. Measurement consists in comparison the measured value with the value of the standard. From the point of view of fuzzy set theory, the measure of possibility describes the degree of possibility (membership function) in which the standard represents the object under study. The measure of matching the pattern to the object should represent this matching process. If the object under study is subject to random fluctuations, it is natural to use a probabilistic additive measure to describe this phenomenon. If we do not observe such fluctuations and estimate the uncertainty using an expert method based on knowledge of the measurement system, we will describe the process as a maximal measure. A maximal measure describes the process of choosing the best, and that's usually how experts work. The membership function describes the extent to which the pattern represents the object under study and the extent to which the value of the measured quantity represents the object under study. Estimating the degree of possibility requires the use of general knowledge in a way that is called expert knowledge.

In papers [6, 8] there was proposed transformation from the probability distribution into a fuzzy set that retains the value of the extended uncertainty. This transformation gives an equivalence relation of two different approaches to the construction of a measure of the distribution of measured quantities and measurement errors (fig. 1). The proposed transformation transforms the probability distribution into a fuzzy set so that the extended uncertainty calculated for the probabilistic model equals the uncertainty in the fuzzy approach. Figure 1 shows qualitatively the probabilistic distribution and the fuzzy set. The confidence interval $[a, b]$ at p is defined as the interval that satisfies the condition $P(X \in [a, b]) = p$. The probability to possibility transformation assumes that the condition $[\bar{A}_X]^\alpha = [a, b]$ is fulfilled (see below).

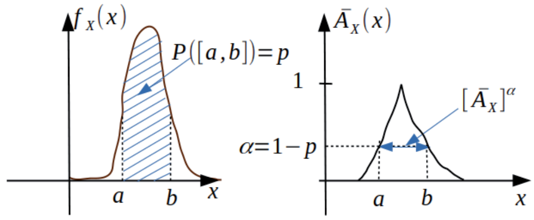

Fig. 1. The probability density distribution $f_X(x)$ of the random variable X and the probability distribution $\bar{A}_X(x)$ of the fuzzy variable. The interval $[a, b]$ is the confidence interval (on the left) and the α-cut of the fuzzy set $\bar{A}_X$ (on the right)

Fuzzy variable

In the original concept, Zadeh did not construct a fuzzy set as a distribution of a fuzzy variable. The fuzzy variable approach was proposed by Nahmias [4], proposing a structure for the definition of fuzzy measures analogous to the axiomatics of probabilistic theory. In this approach the measurement mapping (1) is treated as a fuzzy variable ξ. A fuzzy variable ξ is a real valued function defined on pattern set Γ: $\xi : \Gamma \to [0, 1]$. The pattern set is the set of measurement standards. The fuzzy variable is a measure of the measured quantity: $\xi(\gamma)$ is value of measured quantity for standard $\gamma \in \Gamma$.

Analoguosly to probabilistics, the possibility measure $\Pi : 2^\Gamma \to [0, 1]$ is defined, where Γ: is pattern set. Possibility measure is normalized and maxitive: $\Pi(\emptyset) = 0$, $\Pi(\Gamma) = 0$ and $\Pi(D \cup B) = \max(\Pi(D), \Pi(B))$ for any $D, B \in \Gamma$, where T-norm represents the logical connective „AND" in fuzzy logic. The membership function is the distribution function of a fuzzy variable ξ:

$$\bar{A}_\xi(x) = \Pi(\xi^{-1}(x)) \quad (3)$$

The measurement in the approach of fuzzy variable consists in determining the measure of similarity of the object under study to the set of patterns Γ. The comparator allows to compare the tested object with the elements of pattern set.

By the definition the membership function is a pushforward measure, induced by a function ξ and possibility measure Π.

Fuzzy uncertainty on level α is defined as radius of α cut of a fuzzy set. The level set of a fuzzy set $\bar{A}$ on level α is called α-cut for short and is defined as:

$$\bar{A}^\alpha = \langle x \in \mathbb{R} : \bar{A}(x) \geq \alpha \rangle \quad (4)$$

Probability to possibility transformation

A probability space is a mathematical triple $(\Omega, \mathcal{B}, P)$ where Ω is a sample space, $\mathcal{B}$ is a σ-algebra of subsets of Ω and P is a probability measure.

A possibility space is mathematical 2-tuple (Γ, Π), where Γ is a set of standards and Π is a possibility measure. Note that sigma algebra is not needed here since a possibility measure is maxitive.

In order to construct the transformation of a probabilistic structure into a fuzzy one structure, one needs to note that both can be considered as algebraic structures $(\mathcal{F}, \oplus)$, where $\mathcal{F}$ is a set of random variables or fuzzy variables and $\oplus$

[1] the property of interval arithmetic is obtained if the systematic error is described as a stochastic process with constant values and strong autocorrelation

is addition of random variables or fuzzy variables. The addition operation depends on properties of a random measure or fuzzy measure respectively for random and fuzzy variables. There is a variety of such transformations in literature (see e.g. [1]) but in this paper we assume that this transformation preserves the addition and radius of confidence interval of the median estimator as well [6, 5].

The confidence interval at the probability level can be found by means of the equation:

$$P(I_p) = p \quad (5)$$

Expanded uncertainty at confidence level p is equal to the radius (rad) of a confidence interval. Denote $U_p = \text{rad}(I_p)$. The definition of expanded uncertainty in the Guide is ambiguous: it is assumed that the confidence interval can be applied only to the method A. This is an inconsistency that we want to avoid. Instead we proposed a uniform description in the structure of fuzzy sentences. We assume that a confidence interval corresponds to α-cut (level set on α level) of a fuzzy set representing the measurement result. In some papers it is assumed that one can relate α-cut to the confidence level given by $p = 1 - \alpha$. This may be generalized for other cases as well. Assuming that the relation between confidence level and fuzzy level takes the form $p = 1 - \alpha$ we get the fuzzy set $\bar{A}_X$ which represents a random variable X. The fuzzy set $\bar{A}_X$ is given from the cumulative probabilistic distribution function F_X according to the equation:

$$\bar{A}_X(x) = \begin{cases} 2F_X(x) & for \quad x \le M(X) \\ 2 - 2F_X(x) & for \quad x > M(X) \end{cases} \quad (6)$$

where $M(X)$ is the median of a random variable X and F_X is the cumulative probabilistic distribution of a random variable X.

Fuzzy sets derived from the transformation probability to possibility describe the components of uncertainty of type A, whereas the equations (6) defines a fuzzy set of type $L - R$ (for unimodal distributions).

A distribution of a fuzzy variable representing type B errors is assessed by an expert decision, i. e. a fuzzy set representing type B errors is assessed on the basis of past experience and on general knowledge as well. For example, if we consider a type B error to be systematic, we can represent it by a characteristic function of interval. The reason for this is that systematic errors are modeled in intervals, while the uniform distribution represents an interval itself. A fuzzy set $\bar{A}$ that represents the sum of errors of type A and B has a distribution given by sum of fuzzy sets $\bar{A}_A$ and $\bar{A}_B$, if $\bar{A}_B$ represent interval (rectangular distribution) then and for $\alpha-$cut we have:

$$[\bar{A}]^\alpha = [\bar{A}_A \oplus_T \bar{A}_B]^\alpha = [\bar{A}_A]^\alpha + [\bar{A}_B]^\alpha \quad (7)$$

If a type B error has an rectangular distribution, we get an arithmetic sum of the uncertainty for the combined uncertainty.

Calculation of membership function from measurement data

Suppose that as a result of measurement we obtain a series of N data $\{x_1, \ldots x_n\}$.

Determination of the membership function can be described by the following algorithm (fig. 2):

1. We sort the measurement data, obtaining a series of $x_1 < \ldots < x_n$.
2. For the initial value of x_1 we assume $\bar{A}(x_1) = 0$.

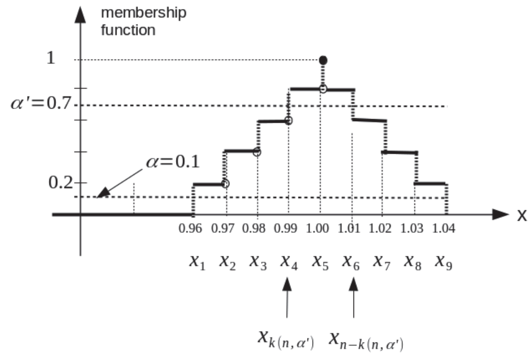

Fig. 2. Empirical estimation of fuzzy set and α-cut in the case $n = 9$

3. We determine $\bar{A}(x_i)$ on the basis of the recursive formula:
   For even n, $i = 2, \ldots, \frac{n}{2}$: $\bar{A}(x_{i+1}) = \bar{A}(x_i) + \frac{2}{n}$
   for $i = \frac{n}{2} + 1, \ldots, n$ we have: $\bar{A}(x_{i+1}) = \bar{A}(x_i) - \frac{2}{n}$
   thus: $\bar{A}(x_{\frac{n}{2}}) = \bar{A}(x_{\frac{n}{2}+1}) = 1$.
   For odd n: $i = 2, \ldots, \frac{n+1}{2}$ $\bar{A}(x_{i+1}) = \bar{A}(x_i) + \frac{2}{n}$
   for $i = \frac{n+1}{2} + 1, \ldots, n$ we have: $A(x_{i+1}) = \bar{A}(x_i) - \frac{2}{n}$
   so: $\bar{A}(x_{\frac{n+1}{2}}) = 1$.
4. For $i = n$ we have: $\bar{A}(x_n) = 0$.
5. Between measuring points we approximate the function with simple horizontal lines and obtain the step function. Other methods of interpolation require additional assumptions.

Since the uncertainty is determined by radius of $\alpha-$cut, we have

$$[\bar{A}]^\alpha = [x_{k(n,\alpha)}, x_{n-k(n,\alpha)}] \quad (8)$$

where $\{x_i\}_{i=1,n}$ is a sequence of sorted data, and $k(n, \alpha)$ is the sequential number of the data forming the left end-points of the α-cut (see 2)

$$k(n, \alpha) = \begin{cases} \left\lceil \frac{n}{2}\alpha \right\rceil & \text{for even } n \\ \left\lceil \frac{(n+1)}{2}\alpha \right\rceil & \text{for odd } n \end{cases} \quad (9)$$

If the measurand is defined as the average value of the data series, then α-cut of the membership function representing the average value for a particular T-norm should be determined. In fuzzy case, averaging results (as in probabilistic approach) in a narrowing of the distribution. We can derive the formula providing the α-cut of the distribution $\bar{A}_{\text{Ave}}$ of the average values:

$$[\bar{A}_{\text{Ave}}]^\alpha = \left[\text{Ave}_n(\bar{A})\right]^\alpha = \left[\bar{A}\right]^{\left(t^{[-1]}\left(\frac{1}{n}t(\alpha)\right)\right)} \quad (10)$$

This formula means that in order to determine the α-cut of the distribution of average values on level α it is necessary to determine the α-cut of non-averaged values on level $\alpha' = t^{[-1]}\left(\frac{1}{n}t(\alpha)\right)$ (where t is additive generator of T-norm).

To apply this formula, you practically need to set a new value for the number of a given (ordered) data:

$$k(n, \alpha) = \frac{1}{2}n\alpha' = \frac{1}{2}t^{[-1]}\left(\frac{1}{n}t(\alpha)\right)n \quad (11)$$

Fuzzy uncertainty will be then given by:

$$U_\alpha(a) = \frac{1}{2}\left(x_{n-k(n,\alpha)} - x_{k(n,\alpha)}\right) \quad (12)$$

The value of $k(n, \alpha)$ depends on the T-norm, which describes the correlation in a series of measurement data. The

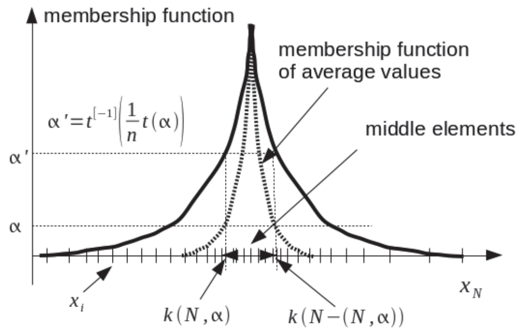

Fig. 3. Middle elements defining the α-cut of membership function of average values

probabilistic model usually assumes no correlation (which is not always true), with this assumption one can derive approximate formulae [5]. The interval $[x_{k(n,\alpha)}, x_{n-k(n,\alpha)}]$ from (8) can be written by means of:

$$(13) \quad K(n) = n - 2\,k(n,\alpha) = n\left(1 - t^{[-1]}\left(\frac{1}{n}t(\alpha)\right)\right)$$

The approximate formulae (reflecting the lack of correlation) are: for a small number of measurements: $K(n) = \dfrac{1}{\alpha + \frac{1}{n}}$, for n large: $K(n) = \sqrt{2n}$.

Example

To illustrate the application of the presented algorithms, consider a series of N =9 voltage measurements with a multimeter with a resolution $0,02V$ on the range 2V. The results of the measurements after sorting are $0.096, 0.097, 0.098, 0.099, 1.00, 1.01, 1.02, 1.03, 1.04$. These data are created artificially to facilitate the use of the algorithm. The average value of this data series is $X_{Av} = 1V$, and the median is $M = 1V$, a membership function is illustrated in 3, α-cut on level $\alpha = 0.1$ is equal to interval $[0.96, 1.04]$. In order to calculate cuts for the averaged values, we determine cuts of a fuzzy set from the measurement data on level $\alpha' = t^{[-1]}\left(\frac{1}{n}t(\alpha)\right)$ (equation (10)). Uncertainty analysis according to GUM [2] assumes, that the measurement data string is not internally correlated, which is not always true. In this example, we will use the T-norm of the Hamacher type with the parameter $p = 0,5$ (research shows that such a t-norm describes an analog-to-digital converter well [6]), then $\alpha' = 0,7$. Section on level $\alpha' = 0,7$ is $[0.99, 1.01]$. This calculated uncertainty on level $\alpha = 0.1$ is the radius of the calculated α-cut and is equal to $U_\alpha = 0,01$V.

The standard deviation calculated from the data series is $s(x) = 0,027$, the standard uncertainty of the average estimator $u_A(x) = 0,009$, the uncertainty extended at the probability level $p = 0,9$ will be approximately $U_{0,9} = 0,015$ (all calculations in Volts).

The approximate value of the interval is obtained by taking $K(n) = \dfrac{1}{\alpha + \frac{1}{n}} = \frac{1}{0,7+1/9} \simeq 1.25$ midpoints. This gives a similar estimate because at least one minimum range has to be taken into account (with such a small number of data, the estimate is inaccurate).

Combined uncertainty

The final step for the theory is a combination of type A and type B uncertainties.

We assume that two components of errors can be distinguished: type A errors and type B errors, but we identify them with a systematic error $\Delta_r X$ and an unsystematic error $\Delta_r X$, which we write as $\Delta X = \Delta_r X + \Delta_s X$

In the language of fuzzy sets we can write it as a composition (addition) of fuzzy sets $\bar{A}_r$ i $\bar{A}_s$ ([7]), where: A_r is a non-systematic component (type A) and A_s is a systematic component (type B) described by the characteristic function of the interval $[-\Delta_m x, +\Delta_m x]$ (equation (7)):

$$\bar{A} = \bar{A}_r \oplus_T \bar{A}_s,$$

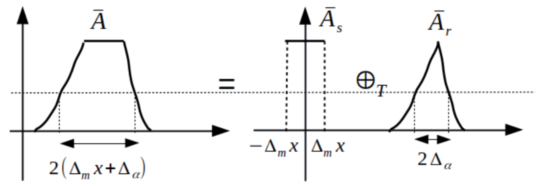

Fig. 4. Fuzzy set describing error and its components

$$U(\bar{A}) = U(\bar{A}_r \oplus_T \bar{A}_s) = U(\bar{A}_r) + U(\bar{A}_s),$$

where: $\Delta_m x$ is the biggest permissible error, Δ_α - the radius of cut at level α [7].

Conclusion

The method presented for calculating the extended uncertainty by the fuzzy set method is based on determining an α-cut of a fuzzy set representing a data series. The combined uncertainty is equal to the arithmetic sum of uncertainty of A and B type. Such a result can also be justified in a probabilistic model if the systematic error is described by a series of identical strongly correlated data. The derivation of the formulas used above can be found in the monograph [5], more complete justification and explanation needs some more papers to be published.

***Authors**: dr hab. inż. Michał Urbański, email: michal.urbanski@pw.edu.pl, mgr inż. Kinga Wójcicka, email: kinga.wojcicka.dokt@pw.edu.pl, Faculty of Physics, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland, dr Paweł Wójcicki, email: Pawel.Wojcicki@pw.edu.pl, Faculty of Mathematics and Informational Sciences, Warsaw University of Technology.*

REFERENCES

[1] Dubois D., Foulloy L., Mauris G., Prade H., Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities, Reliable Computing, 10 273–297, 2004.
[2] Guide to the Expression of Uncertainty in Measurement [web page] https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf, [Access 31.06.2022]
[3] Mauris G., Lasserre V., Foulloy L., A fuzzy approach for the expression of uncertainty in measurement, Measurement 29 pp. 165–177, 2001.
[4] Nahmias S. Fuzzy Variables, Fuzzy Sets and Systems, 1 pp. 97–110, 1978.
[5] Urbański M. K., Modelowanie pomiarów w algebraicznych strukturach rozmytych, Prace Naukowe Politechniki Warszawskiej, Fizyka z. 55, Warszawa, 2011.
[6] Urbański M.K., Bzowski A., Error analysis using fuzzy arithmetic based on t-norm, 12th IMEKO TC1 and TC7 Joint Symposium Man Science & Measurement, l'Universitde Savoie, September 3–5, Annecy, France, 2008.
[7] Urbański M.K., Wąsowski J., Fuzzy approach to the theory of measurement inexactness, Measurement, 34 (2003) 61–74.
[8] Urbański M. K., Wójcicka K., Wójcicki P., Probability to possibility transformation as a way to uniform description of systematic and randoms errors. Przegląd Elektrotechniczny, R. 98 NR 12/2022 pp.159–162.