

Measurement of Continuous Quantities and their Statistical Evaluation

Reinhard Viertl

Vienna University of Technology

Abstract

Measurement results of continuous quantities are always more or less imprecise. This imprecision is different from errors. The most suitable mathematical model to describe imprecision is by special fuzzy subsets of the set of real numbers \mathbb{R} , called characterizing functions. The statistical analysis of fuzzy measurement data is subject of this paper.

Keywords: characterizing function, data uncertainty, fuzzy models, fuzzy vectors, non-precise numbers.

1. Introduction

Usually the analysis of measurement uncertainty is done by stochastic models and statistical methods. This is also realized in norms like DIN EN ISO 20988, VDI 4219(E), and GUM EVN (1999).

More recent methods for the analysis of measurement results of continuous quantities are using a combination of fuzzy models and stochastic models. The fuzziness of individual measurement results is described by so-called *fuzzy numbers*, and the variability by stochastic models. Based on that the analysis of repeated measurements is possible by suitably generalized statistical methods.

2. Measurement uncertainty

Measurement results of continuous quantities are connected with unavoidable uncertainties. The best one can do is to try keeping the uncertainty as small as possible, and to make reasonable (reliable) assessments (estimates) for the considered quantities.

The most important measurement uncertainties are errors and the fuzziness of individual measurement results. Errors are divided into systematic errors and random errors.

Usually measurement results of one-dimensional quantities are described by real numbers x_i . Moreover it is assumed that the measurement result x_i is the sum of the so-called *true value* x , a systematic error y , and a random error ϵ_i , i.e.

$$x_i = x + y + \epsilon_i \quad \text{with} \quad x_i \in \mathbb{R}.$$

For multiple measurements ϵ_i is assumed to be the realisation of a one-dimensional stochastic quantity with expectation 0.

The distribution of the stochastic quantity $\tilde{\epsilon}$, which is modelling the random error, can be estimated from repeated measurements of the observed quantity.

The analysis of the systematic error is more complicated and needs a deep scientific analysis of the measurement procedure.

The above assumption that measurement results are real numbers is not realistic because a real number is determined by all its infinitely many decimals which never can be known (except in trivial cases). Therefore it turns out that measurement results are better described by so-called *fuzzy numbers*.

3. Mathematical description of fuzziness

What is the result of an individual measurement of an one-dimensional continuous quantity like length, time, concentration etc.? For digital measurement equipments the measurement result x of an individual measurement is a decimal number with finitely many digits. Concerning the missing infinitely many decimals almost nothing is known. Therefore all possible digits from 1 to 9 are possible for the remaining decimals. In this case the measurement result is an interval $[\underline{x}; \bar{x}]$, where \underline{x} is the real number which is generated if all unknown decimals of x are assumed to be 0. The real number \bar{x} is generated when all unknown decimals of x are set to be 9.

Therefore the measurement result is an *interval*, i.e. a subset of the set of real numbers \mathbb{R} .

The analysis of multiple measurements of the same quantity is than a problem of interval analysis.

The situation is more general in case of using an analog measurement equipment like measurement rods, pointers, and imaging systems like oscilloscopes. In this case individual measurement results are pictures, light points, functions or colour intensity pictures. In order to describe mathematically this kind of measurement results so-called *fuzzy numbers* x^* are used. These are special fuzzy subsets of the set \mathbb{R} .

Coming back to the measurement result $[\underline{x}; \bar{x}]$ in case of a digital measurement equipment, this interval can be characterized logically equivalent by its *indicator function* $I_{[\underline{x}; \bar{x}]}(\cdot)$. The indicator function is defined by its values

$$I_{[\underline{x}; \bar{x}]}(x) := \begin{cases} 1 & \text{for } \underline{x} \leq x \leq \bar{x} \\ 0 & \text{otherwise} \end{cases} \quad \forall x \in \mathbb{R}.$$

Indicator functions can be identified with subsets of general sets M .

In reality the boundaries of subsets can be difficult. Therefore in 1951 K. Menger [Menger \(2003\)](#) published a generalization of indicator functions describing generalized sets called *ensembles flous*. These generalized indicator functions were later called *membership functions* $\zeta(\cdot)$, which can assume all values from the unit interval $[0; 1]$, i.e. $\zeta : M \rightarrow [0; 1]$. These generalized sets were called *fuzzy sets* in a paper by L. Zadeh in 1965. Taking care of a continuum of truth values mathematical methods connected with fuzzy sets are called *fuzzy logic*.

If A^* is a fuzzy subset of an arbitrary set M , having membership function $\zeta(\cdot)$, for elements $x \in M$ the value $\zeta(x)$ is called *degree of membership* of x in A^* .

For the description of measurement results of one-dimensional quantities by analog measurement equipments special fuzzy subsets of \mathbb{R} , so-called *fuzzy numbers* are most suitable. The corresponding membership function $\zeta(\cdot)$ must meet the following conditions:

- (1) $\text{supp}[\zeta(\cdot)] = \{x \in \mathbb{R} : \zeta(x) > 0\}$ is a bounded subset of \mathbb{R}

- (2) $\forall \delta \in (0; 1]$ the so-called δ -cut $C_\delta [\zeta(\cdot)] := \{x \in \mathbb{R} : \zeta(x) \geq \delta\}$ must be non-empty and a finite union of compact intervals, i.e.

$$C_\delta [\zeta(\cdot)] = \bigcup_{j=1}^{k_\delta} [a_{\delta,j}; b_{\delta,j}].$$

Membership functions obeying (1) und (2) are called *characterizing functions*.

Characterizing functions of fuzzy numbers will be denoted by $\xi(\cdot)$ in the following.

In figure 1 some characterizing functions are depicted.

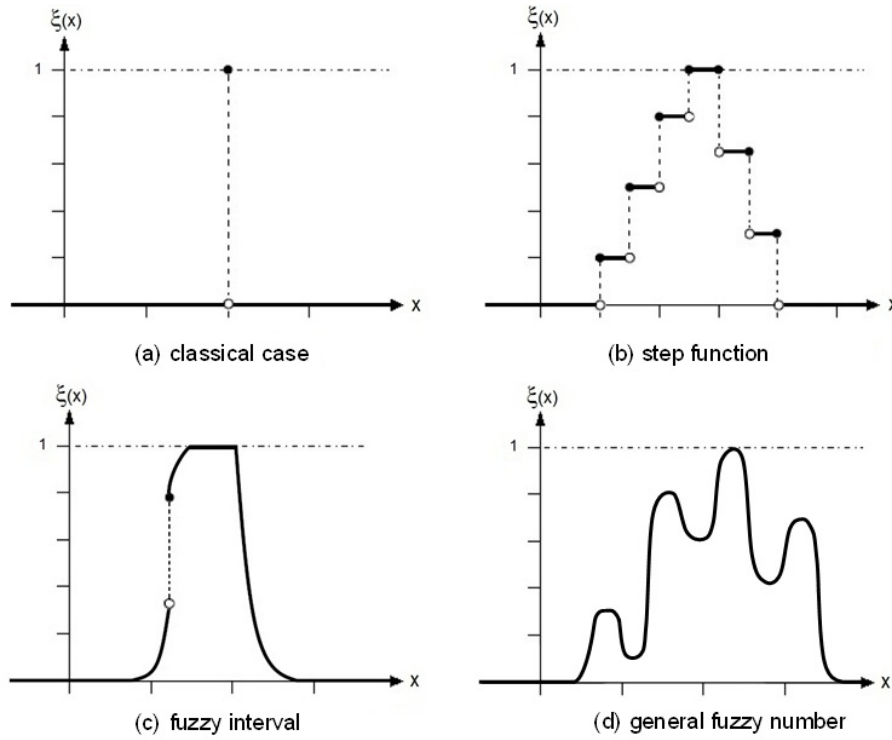


Figure 1: Examples of characterizing functions

4. Analysis of fuzzy measurement results

Based on the description of measurement results in section 3 a measurement result is a fuzzy number x_i^* , whose characterizing function $\xi_i(\cdot)$ in case of a colour intensity picture is a function of the colour intensity $h(x)$, $x \in \mathbb{R}$ and its derivative $h'(\cdot)$. The values of the characterizing function are obtained in the following way:

$$\xi_i(x) = \frac{|h'(x)|}{\sup \{|h'(x)| : x \in \mathbb{R}\}} \quad \forall \quad x \in \mathbb{R}$$

Considering errors we have $x_i^* = x + y + \epsilon_i$, which means, that at least one of the quantities x , y , and ϵ_i is fuzzy. The systematic error is assumed to be a real number, therefore one of the quantities x and ϵ_i must be fuzzy. Based on the standard statistical analysis ϵ_i is the realization of a one-dimensional stochastic quantity. Therefore x must be fuzzy, which means the so-called *true value* is a fuzzy number x^* . It should be noted that the systematic error y can practically only be determined as fuzzy number.

Assuming a measurement procedure has no systematic error, i.e.

$$x_i^* = x^* + \epsilon_i, \quad i = 1(1)n,$$

the determination, i.e. estimation of x^* based on fuzzy data x_1^*, \dots, x_n^* is possible.

In generalization of the classical mean value the mean value (which has to be defined) of the fuzzy observations has to be calculated:

$$\bar{x}_n^* = \frac{1}{n}(x_1^* \oplus \dots \oplus x_n^*).$$

The generalized addition operation \oplus of fuzzy numbers is based on the so-called *extension principle* from the theory of fuzzy sets. \bar{x}_n^* is a fuzzy number whose characterizing function $\psi(\cdot)$ can be calculated from the characterizing functions $\xi_1(\cdot), \dots, \xi_n(\cdot)$ of the fuzzy numbers x_1^*, \dots, x_n^* . The fuzzy number \bar{x}_n^* is the best information which can be obtained concerning the true value x^* , based on the fuzzy sample x_1^*, \dots, x_n^* .

In case of continuous quantities the unavoidable fuzziness of measurements x_i^* appears in the estimation of x^* .

5. Calculation of the characterizing function of the fuzzy sample mean

In order to determine the characterizing function $\psi(\cdot)$ of \bar{x}_n^* it is necessary to combine the fuzzy numbers x_1^*, \dots, x_n^* , having characterizing functions $\xi_1(\cdot), \dots, \xi_n(\cdot)$, into a so-called *fuzzy vector* \underline{x}^* of the n -dimensional Euclidean space \mathbb{R}^n , in order to apply the so-called *extension principle* of the theory of fuzzy sets. This combination is done by the so-called *minimum-t-norm* (compare [Viertl and Hareter \(2006\)](#)) in the following way.

A fuzzy vector \underline{x}^* is defined by its so-called *vector-characterizing function* $\zeta(\cdot, \dots, \cdot)$, which has to obey the following:

- (1) $\zeta : \mathbb{R}^n \longrightarrow [0; 1]$
- (2) $\text{supp}[\zeta(\cdot, \dots, \cdot)]$ is a bounded subset of \mathbb{R}^n
- (3) $\forall \delta \in (0; 1]$ the δ -cut $C_\delta[\zeta(\cdot, \dots, \cdot)] := \{\underline{x} \in \mathbb{R}^n : \zeta(\underline{x}) \geq \delta\}$ has to be a non-empty, closed and bounded subset of \mathbb{R}^n , which is a finite union of simply connected sets.

In order to extend functions $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ to the situation of fuzzy arguments \underline{x}^* , the so-called *extension principle* is used.

Let M and N be arbitrary sets and $g : M \rightarrow N$ an arbitrary function. The function $g(\cdot)$ can be extended to fuzzy argument values x^* of M , where $\zeta(\cdot)$ is the membership function of x^* . The value $g(x^*)$ is a fuzzy subset of N , whose membership function $\psi(\cdot)$ is defined in the following way:

$$\psi(y) := \begin{cases} \sup \{\zeta(x) : x \in M, g(x) = y\} & \text{if } g^{-1}(\{y\}) \neq \emptyset \\ 0 & \text{if } g^{-1}(\{y\}) = \emptyset \end{cases} \quad \forall y \in N$$

In order to apply the extension principle to the function $g(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i$ for fuzzy numbers x_1^*, \dots, x_n^* these fuzzy numbers have to be combined into a n -dimensional fuzzy vector. This is done by application of the minimum- t -norm, which determines the vector-characterizing function $\zeta(\cdot, \dots, \cdot)$ in the following way:

$$\zeta(x_1, \dots, x_n) := \min \{\xi_1(x_1), \dots, \xi_n(x_n)\} \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n$$

The reason for using the minimum- t -norm is explained in the following proposition.

Proposition 1: Let n fuzzy numbers x_1^*, \dots, x_n^* be combined by the minimum- t -norm, then the resulting function $\zeta : \mathbb{R}^n \rightarrow [0; 1]$ is a vector-characterizing function.

Proof: Condition (1) is trivially fulfilled.

Condition (2) follows from the fact that all supports of $\xi_i(\cdot)$ are bounded and the following holds: $\text{supp}[\zeta(\cdot, \dots, \cdot)] = \bigcap_{i=1}^n \text{supp}[\xi_i(\cdot)]$.

Condition (3) follows from the fact that the δ -cuts of $\zeta(\cdot, \dots, \cdot)$ are the Cartesian products of the δ -cuts of $\xi_i(\cdot)$.

Based on this combined fuzzy vector the characterizing function of the fuzzy arithmetic mean of n fuzzy numbers can be determined.

Proposition 2: The arithmetic mean of n fuzzy numbers with corresponding characterizing functions $\xi_1(\cdot), \dots, \xi_n(\cdot)$ is a fuzzy number, whose characterizing function $\psi(\cdot)$ is given by its values $\psi(x) \forall x \in \mathbb{R}$ in the following way:

$$\psi(x) = \sup \left\{ \min \{ \xi_1(x_1), \dots, \xi_n(x_n) \} : \text{for } \frac{1}{n} \sum_{i=1}^n x_i = x \right\}$$

The δ -cuts $C_\delta[\psi(\cdot)]$ for all $\delta \in (0; 1]$ are finite unions of compact intervals given by

$$C_\delta[\psi(\cdot)] = \bigcup_{\sum_{i=1}^n x_i = x} \left\{ \frac{x_1 + \dots + x_n}{n} : x_i \in C_\delta[\xi_i(\cdot)] \right\}.$$

Proof: The δ -cuts of the combined fuzzy vector \underline{x}^* are finite unions of simply connected and compact subsets of \mathbb{R}^n . By the continuity of the function $g(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i$ also the δ -cuts of $g(\underline{x}^*)$ from the extension principle are finite unions of compact simply connected subsets of \mathbb{R} , and therefore finite unions of compact intervals. By that reason the fuzzy mean value is a fuzzy number. For the δ -cuts of the fuzzy mean value the following holds:

$$C_\delta[\psi(\cdot)] = g\left(C_\delta[\xi_1(\cdot)], \dots, C_\delta[\xi_n(\cdot)]\right) = \left\{ \frac{1}{n} \sum_{i=1}^n x_i : x_i \in C_\delta[\xi_i(\cdot)] \forall i = 1(1)n \right\}$$

The result of measurements of continuous quantities are fuzzy numbers and the measurement uncertainty of the calculated value is presented by the characterizing function of the fuzzy mean value.

Example: Let 5 measurement results be given as fuzzy numbers with characterizing functions $\xi_1(\cdot), \dots, \xi_5(\cdot)$ which are depicted in figure 2. Then the characterizing function $\psi(\cdot)$ of the fuzzy arithmetic mean is determined with the help of δ -cuts.

The characterizing function $\psi(\cdot)$ of the arithmetic mean of the fuzzy observations is depicted in figure 3.

Remark: The characterizing function $\psi(\cdot)$ is the most realistic information concerning the measured quantity. The area under the function $\psi(\cdot)$ is characteristic for the measurement uncertainty.

It is possible to analyse the dispersion of the fuzzy measurements. In order to do that the generalized sample spread of the fuzzy measurement results can be calculated. This is the fuzzy quantity s_n^* , whose characterizing function can be calculated by application of the extension principle to the classical sample spread

$$s_n = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \right]^{1/2}.$$

For this the extension principle is applied to the function

$$g(x_1, \dots, x_n) = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \right]^{1/2}.$$

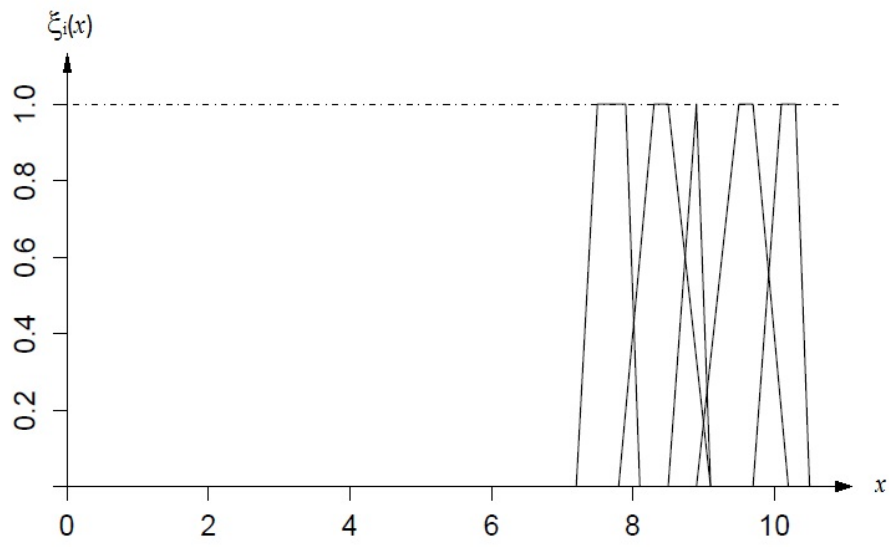


Figure 2: Fuzzy measurement results

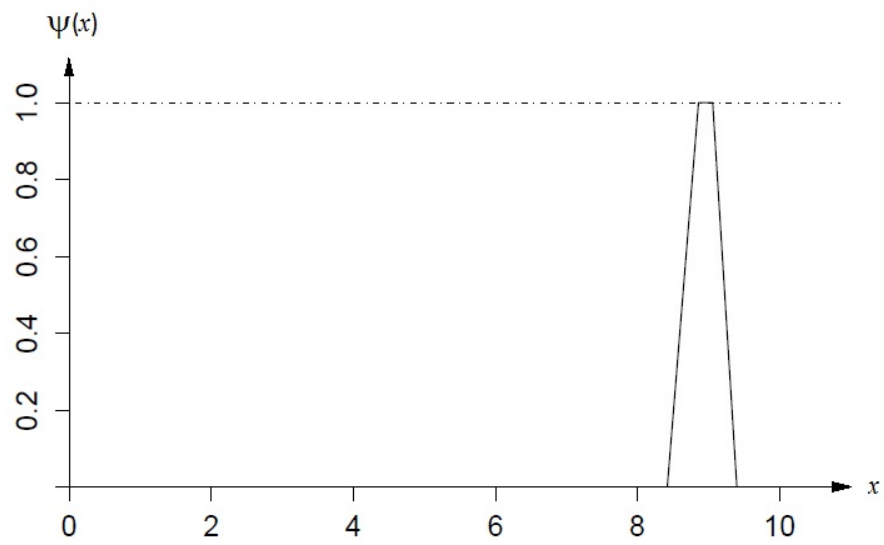


Figure 3: Fuzzy arithmetic mean

The δ -cuts $C_\delta[s_n^*]$ of the generalized (fuzzy) estimate s_n^* of the generalized standard deviation can be obtained by the following equation:

$$C_\delta[s_n^*] = \left[\min_{(x_1, \dots, x_n) \in C_\delta[x_1^*] \times \dots \times C_\delta[x_n^*]} g(x_1, \dots, x_n); \max_{(x_1, \dots, x_n) \in C_\delta[x_1^*] \times \dots \times C_\delta[x_n^*]} g(x_1, \dots, x_n) \right]$$

The characterizing function $\varphi(\cdot)$ of s_n^* can be obtained by the so-called *characterization lemma for characterizing functions* (compare [Viertl \(2011\)](#)):

$$\varphi(x) = \max \left\{ \delta \cdot \mathbb{1}_{C_\delta[s_n^*]}(x) : \delta \in [0; 1] \right\} \quad \forall x \in \mathbb{R}.$$

For the fuzzy sample x_1^*, \dots, x_5^* from figure 2 the characterizing function of the fuzzy estimate s_n^* of the generalized standard deviation is depicted in figure 4.

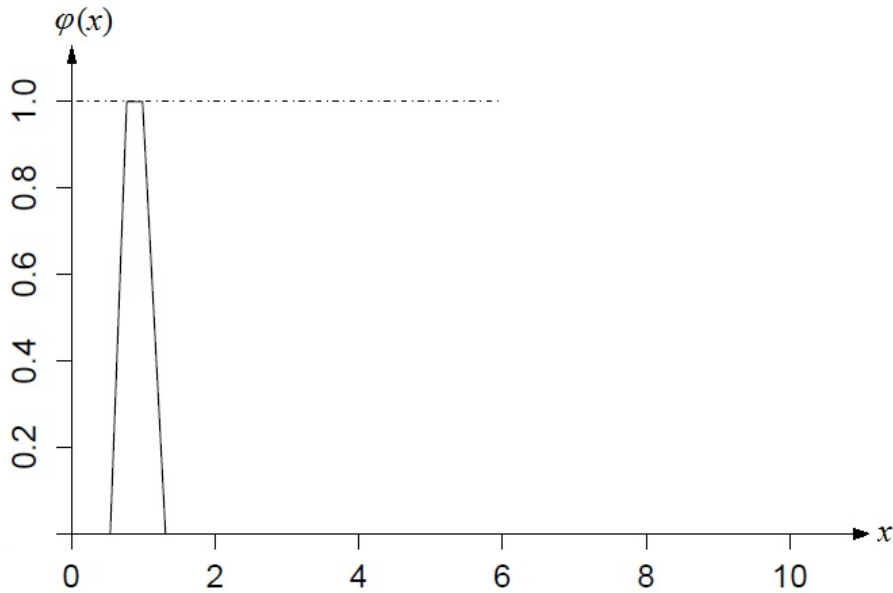


Figure 4: Fuzzy sample standard deviation

Details on the algorithmic realisation of the calculation of s_n^* are given in [Viertl and Hareter \(2006\)](#).

6. Summary and Outlook

In the paper the mathematical description of measurement results of one-dimensional continuous quantities is explained. This is possible by so-called fuzzy numbers. Repeating the measurement yields a finite sequence of observations in form of fuzzy numbers x_1^*, \dots, x_n^* with corresponding characterizing functions $\xi_1(\cdot), \dots, \xi_n(\cdot)$. The generalization of averaging the sample is explained, and a generalized fuzzy estimator for the standard deviation of the quantity is given. For multivariate continuous quantities the measurement results are also more or less imprecise. The results could be modelled by so-called k -dimensional fuzzy vectors which are special fuzzy subsets of the k -dimensional Euclidean space \mathbb{R}^k . The statistical analysis of multivariate fuzzy data is an interesting topic for future research.

References

- (1999). “Guide to the Expression of Uncertainty in Measurement.” EVN 13005.
- Menger M (2003). “Ensembles flous et fonctions aléatoires.” In B Schweizer, A Sklar, K Sigmund, P Gruber, E Hlawka, L Reich, L Schmetterer (eds.), *Selecta Mathematica*, pp. 445–447. Springer Vienna. ISBN 978-3-7091-7294-0.
- Viertl R (2011). *Statistical Methods for Fuzzy Data*. Wiley. ISBN 9780470974568.
- Viertl R, Hareter D (2006). *Beschreibung und Analyse unscharfer Information: Statistische Methoden für unscharfe Daten*. Springer, Wien. ISBN 9783211238776.

Affiliation:

Reinhard Viertl
Department of Statistics and Probability Theory
Vienna University of Technology
A-1040 Vienna, Austria
E-mail: r.viertl@tuwien.ac.at
URL: <http://www.statistik.tuwien.ac.at/public/viertl>