The Mathematical Theory of Evidence and Measurement Uncertainty

Comparison of Measurement Results Expressed in Terms of Random-Fuzzy Variables

Simona Salicone

In the previous papers [1], [2], it was shown how possibility distributions can be effectively employed to represent and propagate uncertainty in measurements. In particular, it was shown how the Random-Fuzzy variables (RFVs) can be effectively employed to represent a measurement result. The effects of the systematic and random contributions to uncertainty can be well identified in the RFV, and all confidence intervals at all confidence levels are provided, so that complete information about the measurement result is given. Moreover, this distinction also allows one to model the propagation of the systematic and the random contributions in two different ways, according to their different nature and different behavior when they combine.

In most practical applications, the final aim of a measurement procedure is to take a decision on the basis of the comparison of the obtained measurement result with a given threshold. Moreover, the threshold could be either a fixed value or a measurement results itself, thus affected by measurement uncertainty.

The aim of this paper is to show that also this final step can be done in terms of RFVs.

Why Do We Need a Comparison?

In all industrial applications, products must be inspected at the different stages of the production process, in order to assess whether they fulfill specific requirements or not. This operation is called conformity assessment, and is defined as an "*activity to determine whether specified requirements relating to a product, process, system, person or body are fulfilled*" [3].

To perform a conformity assessment of an item, it is first necessary to define:

1. The measurable property of the item that must be controlled;
2. The interval of permissible values of the property, specified by one or two tolerance limits;
3. The decision rule, that is, the "*documented rule that describes how measurement uncertainty will be accounted for with regard to accepting or rejecting an item, given a specified requirement and the result of a measurement*" [3].

These first steps strictly depend on the considered item and on the industrial strategy (business and policy decisions). It is necessary to decide which is the acceptable level of risk and which is the right trade-off between costs and risks: the costs of reducing measurement uncertainty against the risks of too much waste, due to false rejection. Once these important decisions, which are hence not universal but depend on the specific item in a specific factory, are taken, it is possible to perform the conformity assessment.

The conformity assessment of the item is based on the following steps:

a) The measure of the property of interest of the item;
b) The comparison of the obtained measurement result with the tolerance limits;
c) The decision of acceptance/rejection of the item (according with the decision rule).

The present Guide on conformity assessment [3] provides all indications to perform the comparison of the measurement result with the tolerance limits and accept (or reject) the item. Different solutions are also given, by introducing the acceptance limits and the guard bands [3], in order to reduce the risks of accepting a non-conforming item or rejecting a confirming item. It is not the aim of this paper to report here all indications, for which the interested readers are addressed directly to [3], but only to recall that the indications presume that the property of interest of the item

> *…has been measured, with the result of the measurement expressed in a manner compatible with the principles described in the GUM. In particular, it is assumed that corrections have been applied to account for all recognized significant systematic effects* [3].

(NoA: GUM refers to the Standard *Guide to the Expression of Uncertainty in Measurement* [4]). More explicitly, the measurement result must be

> *…expressed in a manner consistent with the principles of the GUM, so that knowledge of the value of the property can*

be reasonably described by (a) a probability density function (PDF), (b) a distribution function, (c) numerical approximations to such functions, or (d) a best estimate, together with a coverage interval and an associated coverage probability [3].

The indications given in [3] are hence based on the probability theory, and the final decision of acceptance or rejection of the item is provided together with (in some particular situations) a probability of having made an incorrect conformance decision.

In this paper, it will be shown, without entering too much the mathematical details, for which the readers are addressed to [5], how a comparison can be performed and a decision can be taken when the measurement result is expressed by possibility distributions. In fact, if the theory of evidence is considered and the measurement results are expressed in terms of RFVs, with the advantages already shown in [1], [2], [6]–[9], then, the conformity assessment must be done in this mathematical context. This requires a review of the previous steps 3) and b), while leaving the others steps unchanged, as well as all theoretical considerations about the conformity assessment.

The Comparison of an RFV with a Tolerance Limit

As stated in [3], a tolerance limit is a value

…that separates intervals of permissible values of the measurand from intervals of non-permissible values.

Moreover:

Intervals of permissible values, called tolerance intervals, are of two kinds: a one-sided tolerance interval with either an upper or a lower tolerance limit; a two-sided tolerance interval with both upper and lower tolerance limits. In either case, an item conforms to the specified requirement if the true value of the measurand lies within the tolerance interval and is nonconforming otherwise.

And:

Seemingly one-sided tolerance intervals often have implied additional limits, for physical or theoretical reasons, that are not explicitly stated.

From the above statements, it follows that the Guide considers limits that are crisp values, not affected by uncertainty. So, let us start by this simplified assumption. Let us consider Fig. 1: the red line represents the given tolerance limit (also called, in the following, threshold) for the property of interest of the considered item, while the blue line is the RFV representing the result of the measurement of that property. The RFV provides all information about the measurement result, giving all confidence intervals and associated levels of confidence [2], [6]–[9]. All these pieces of information can be taken

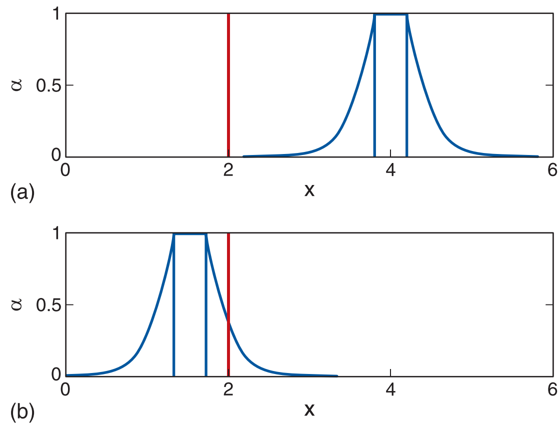

Fig. 1. Comparison of a measurement result (blue line RFV) with a given fixed threshold (red line). In the upper figure, the RFV is completely above the threshold and the credibility coefficient that the RFV is greater than the threshold is 1. In the lower figure, the measurement result is decreased: in this case, the credibility coefficient that the RFV is greater than the threshold is 0.07 while the credibility coefficient that the RFV is lower than the threshold is 0.93.

into account when comparing the measurement result with the threshold. In particular, the comparison is made by considering the external possibility distribution (PD) of the RFV, since the effects of both the random and systematic contributions to uncertainty, combined together, shall be considered in assessing whether the tolerance limit has been exceeded or not.

Let us assume, as an example, that the considered property of the item must not exceed the given tolerance limit. In the case of the upper plot in Fig. 1, the RFV has completely exceeded the tolerance limit, while in the lower figure the RFV is across the tolerance limit. In the first case, since the RFV has completely exceeded the tolerance limit, it is intuitive to state that the considered item must be rejected and that this decision is taken with full certainty, that can be quantified by a credibility factor equal to 1. The credibility factor is a numerical value, in the [0,1] range, quantifying how much one is sure about the taken decision. The credibility factor has a similar meaning of the probability of taking a right/wrong decision in the Guide approach [3]. Since the RFVs are represented by PDs, it is not mathematically correct, of course, to speak about probability and a different term is used.

In the case of the lower plot in Fig. 1, since the RFV is across the tolerance limit, it is not possible to take a decision with full certainty. It is intuitive, according to the figure, that most of the values that can be reasonably attributed to the measurand do not exceed the threshold and this intuition can be quantified by a credibility factor. In particular, it is possible to define two factors: the first one is obtained by evaluating the area exceeding the threshold normalized to the total area subtended by the RFV; the second one is obtained by evaluating the area not exceeding the threshold normalized to the total area subtended by the RFV. It can be readily perceived that the first obtained factor ($C_{A>B}$) reflects the credibility that the considered RFV (A) is greater than the threshold (B), while the second factor ($C_{A<B}$) reflects the credibility that the considered RFV is lower than the threshold. In the considered example, it is: $C_{A>B} = 0.07$ and $C_{A<B} = 0.93$. The obtained values of $C_{A>B}$ and $C_{A<B}$ allow one to state that the considered item should be accepted, and the credibility associated to this decision is equal to 0.93.

In general, however, the decision-making process must obey to the decision rule defined on the basis of business and policy issues, which requires a compromise among costs, risks and benefits [3]. Depending on the considered item, the decision rule can be more or less cogent. These definitions, however, are out of the scope of this paper, being a matter of industrial policy and economy. What is important to underline here is that the above proposed credibility factors can be useful in the decision-making process.

The Comparison of an RFV with Another RFV

In the previous section, it has been shown how an RFV can be compared with a given threshold. Two credibility factors are defined, the credibility that the RFV is below the threshold and the credibility that the RFV is above the threshold. The proposed method is very similar to the one based on probabilities [3] and represents the alternative solution to the guidelines given in [3] when the measurement result is not expressed in terms of a probability distribution, but by an RFV, that is, by possibility distributions.

As also stated in the previous section, the definition given in [3] for the tolerance limit is that of a threshold that is crisp, not affected by uncertainty. It is possible, however, to generalize the definition of tolerance limits. In fact, a tolerance limit can be set on the basis of some theoretical considerations, thus leading to a crisp threshold, but can be also set on the basis of some previous measurements and, in this case, the tolerance limit is affected itself by measurement uncertainty. In this last, more general case, the tolerance limit is not a crisp value but a measurement result itself, and the problem of the comparison of a measurement result with a threshold becomes the problem of comparing two measurement results. This more general case is not considered in [3], but it can be easily dealt with when the measurement results are represented by RFVs.

By extending the considerations given in the previous section, the comparison of two RFVs is made by considering the external PDs of the RFVs themselves. Let us consider Fig. 2. Fig. 2 is the generalization of Fig. 1, where the tolerance limit (red line) is not a crisp value any longer, but an RFV. In this case, it is possible to define three credibility factors: the credibility factor $C_{A>B}$ that RFV A is greater than RFV B; the credibility factor $C_{A<B}$ that RFV A is lower than RFV B; and the credibility factor $C_{A=B}$ that RFV A is compatible with RFV B (in measurement science, because of the presence of the measurement uncertainty, equality does not exist, in a strict mathematical sense, so that we refer to *compatible* measurement results when there is no sufficient evidence to state that two measurement results identify two different measurands [10]).

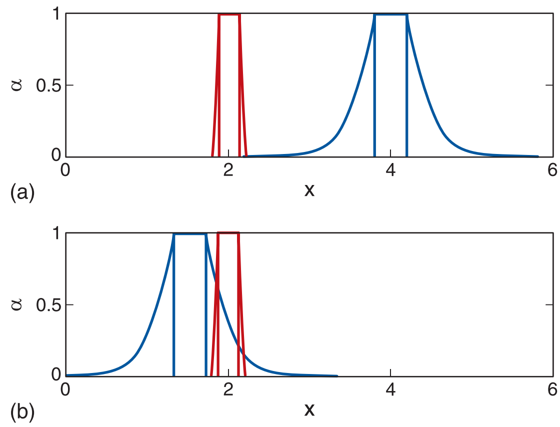

Fig. 2. Comparison of a measurement result (blue line RFV) with a given threshold (red line RFV). In the upper figure, the RFV is completely above the threshold and the credibility coefficient that the RFV is greater than the threshold is 1. In the lower figure, the measurement result is decreased: in this case, the credibility coefficient that the RFV is greater than the threshold is 0.02; the credibility coefficient that the RFV is lower than the threshold is 0.86; the credibility coefficient that the RFV is equal to the threshold is 0.12.

These credibility factors can be evaluated thanks to the use of some well-known and widely employed fuzzy operators: the fuzzy intersection, the fuzzy union, the Hamming distance, the fuzzy-max [7]. Not to enter the mathematical details, Figs. 3–6 show graphically the meanings of these operators. In these figures, the considered RFVs and their relative position is taken in order to have the more general possible situation (as will be shown later in Fig. 7).

Let us consider the two RFVs A (magenta line) and B (black line). Fig. 3 shows the standard fuzzy intersection $Int(A, B)$, numerically represented by the green area. Fig. 4 shows the standard fuzzy union $Un(A, B)$, represented by the green area. Fig. 5 shows the Hamming distance $d(A, B)$, represented by the green area. Fig. 6 shows the fuzzy-max operator $MAX(A, B)$, represented by the dashed green line; the fuzzy-max operator provides a new fuzzy variable.

Once defined the above fuzzy operators, it is possible to define the three credibility factors $C_{A>B}$, $C_{A<B}$, $C_{A=B}$ as:

$$C_{A>B} = \frac{d(B, MAX(A, B))}{Un(A, B)} \tag{1}$$

$$C_{A<B} = \frac{d(A, MAX(A, B))}{Un(A, B)} \tag{2}$$

$$C_{A=B} = \frac{Int(A, B)}{Un(A, B)} \tag{3}$$

from which $C_{A>B} + C_{A<B} + C_{A=B} = 1$ follows. $d(B, MAX(A, B))$ and $d(A, MAX(A, B))$ are the Hamming distances between RFVs B, and A, respectively, and RFV $MAX(A, B)$ (Fig. 6). The graphical interpretation of these distances is given in Fig. 7. In this figure, the two RFVs have been chosen in order to show the more

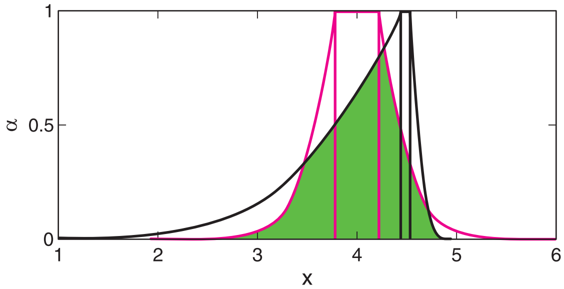

Fig. 3. The standard fuzzy intersection between the two RFVs (magenta and black lines) is represented by the green area.

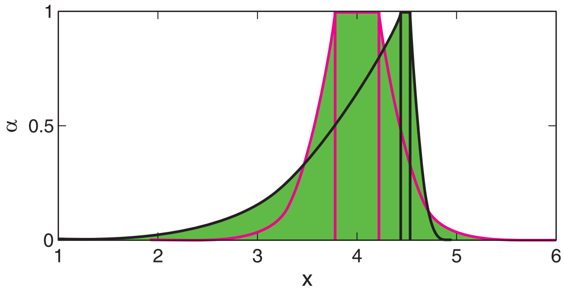

Fig. 4. The standard fuzzy union between the two RFVs (magenta and black lines) is represented by the green area.

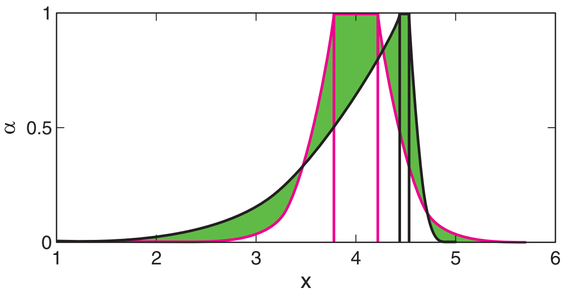

Fig. 5. The Hamming distance between the two RFVs (magenta and black lines) is represented by the green area.

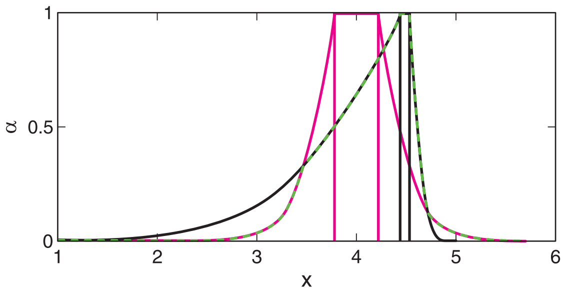

Fig. 6. The fuzzy-max operator between the two RFVs (magenta and black lines) is represented by the dashed green line.

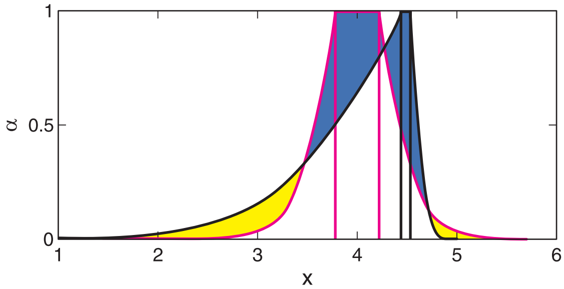

Fig. 7. The more general case of the comparison of two measurement results: the magenta line is RFV *A* and the black line is RFV *B*. The yellow area represents $d(B, MAX(A, B))$, the numerator of $C_{A>B}$, that is, the Hamming distance between *B* and $MAX(A, B)$ (shown in Fig. 6). The blue area represents $d(A, MAX(A, B))$, the numerator of $C_{A<B}$, that is, the Hamming distance between *A* and $MAX(A, B)$. The white area represents $Int(A, B)$, the numerator of $C_{A=B}$, as also shown in Fig. 3.

general situation where two areas determine $d(B, MAX(A, B))$ (yellow areas) and two areas determine $d(A, MAX(A, B))$ (blue areas). As far as the yellow areas are concerned, let us consider the right and the left ones: the numerical value of the area on the right quantifies how much *A* is on the right of *B*, while the numerical value of the area on the left quantifies how much *B* is on the left of *A*. Hence, these two areas together contribute to define factor $C_{A>B}$. Similarly, as far as the blue areas are concerned, let us consider the right and the left ones: the numerical value of the area on the right quantifies how much *B* is on the right of *A*, while the numerical value of the area on the left quantifies how much *A* is on the left of *B*. Hence, these two areas together contribute to define factor $C_{A<B}$.

In the case of the upper plot in Fig. 2, it is: $C_{A>B} = 1$; $C_{A<B} = 0$; $C_{A=B} = 0$, so RFV *A* is considered greater than *B* with full certainty. In the case of the lower plot in Fig. 2, it is: $C_{A>B} = 0.02$; $C_{A<B} = 0.12$; $C_{A=B} = 0.86$, so the available evidence shows that the two measurement results can be considered compatible with a credibility of 0.86. On the basis of the values of these three factors and on the defined decision rule [3], it is hence always possible to take a decision of conformity or non-conformity [7].

Fig. 8 shows an example of comparison between the same two RFVs in Fig. 7, but with different relative position. In the case of the upper plot, it can be also intuitively stated that the two RFVs are compatible. In fact, it follows : $C_{A>B} = 0.10$, $C_{A=B} = 0.58$, $C_{A<B} = 0.32$, that is, the greatest credibility factor is indeed represented by $C_{A=B}$. In the case of the lower plot, it can be also intuitively stated that two RFV *A* is lower than *B*. In fact, it follows: $C_{A>B} = 0$, $C_{A=B} = 0.01$, $C_{A<B} = 0.99$.

It can be proven [7] that the simpler case of a comparison of an RFV with a crisp threshold represents a particular case of this general one, so that equations (1) – (3) can be used in all possible situations. Of course, when the threshold is a crisp value, it is always: $C_{A=B} = 0$ and $C_{A>B} + C_{A<B} = 1$.

The readers can perform comparisons between RFVs by opening this web page (optimized for view in Internet Explorer): http://131.175.120.11:8000/RFVcalculator.html. The

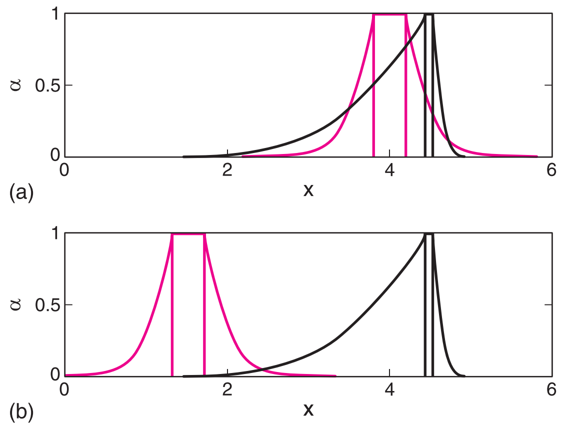

Fig. 8. The more general case of the comparison of two measurement results: the magenta line represents RFV *A* and the black line represents RFV *B*. In the case of the upper plot, it follows: $C_{A>B} = 0.10$, $C_{A=B} = 0.58$, $C_{A<B} = 0.32$, In the case of the lower plot, it follows: $C_{A>B} = 0$, $C_{A=B} = 0.01$, $C_{A<B} = 0.99$.

front panel of a remotely controlled application is shown, which allows the readers to create two RFVs, combine and compare them.

Conclusions

This paper has shown how measurement results can be compared when they are represented by RFVs. This allows one to perform conformity assessments of items and take acceptance or rejection decisions according to the credibility factors provided by the proposed method and to the decision rule defined by the industrial policy.

The advantages of this approach can be synthesized as follows. Since measurement results are represented by RFVs, conformity assessment can be performed also when both random and systematic contributions to uncertainty affect the measurement results themselves, and it is not limited to the case of only random contributions, as required by [3]. Since the RFV provides all confidence intervals associated to the measurand, conformity assessment can be always performed by taking into account all information about the measurand, and not only partial information, as happens, for instance, when only "*a best estimate, together with a coverage interval and an associated coverage probability*" is known [3]. The availability of the three credibility factors $C_{A>B}$, $C_{A<B}$, $C_{A=B}$ can help the decision-making process, once the decision rule has been set.

There is another important advantage in the proposed method. Until now, the comparison of two RFVs has been justified as the main step of conformity assessment. However, the comparisons of measurement results or the comparison of a measurement result with a threshold may not represent the final step of a measurement procedure, when it is used in a decision-making process. It may happen that the measurement result is obtained as the output value of a measurement algorithm and that this algorithm contains *if…then…else…* structures. In such a case, comparisons must be performed throughout the measurement procedure and not only in the final step. As an example, in [11], where the problem of measuring a specific power quality index is considered, it has been shown how the proposed method can provide a significant measurement result expressed in terms of an RFV, also when the measurement algorithm is complex and contains multiple *if…then…else…* structures.

Final Discussion about the Series

This paper concludes the series of three papers published in *I&M Magazine* since August 2014. A quick survey has been given on how possibility distributions can be used to represent, propagate and compare measurement results. Even if the presentation has been purposely done in an intuitive way, without entering too much the mathematical details, I hope I achieved the main goal to show the advantages and potentials of this new approach.

I want to underline, once again, that this approach represents a generalization of the present approach to uncertainty, based on statistics and probability, and is perfectly compliant with the concepts on which the GUM is grounded. It allows one to consider uncertainty contributions of both random and systematic nature, and to propagate them taking into account their different nature and physical behavior. Moreover, despite the quite complex mathematics used to define it, the practical implementation of this method requires to simply compute algebraic operations. Therefore, the entire RFV of the final measurement result is obtained in a relatively simple way, if compared to the time-consuming Monte Carlo simulations that would be required in most cases to obtain the distribution of the values that could be reasonably attributed to the measurand [12].

As one of the authors that contributed to the development of this approach, I'm glad to see that several scientists around the world have already considered and applied it in their measurement applications in different fields [13]–[17].

References

[1] S. Salicone, "The mathematical theory of evidence and measurement uncertainty," *IEEE Instrum. Meas. Mag.*, vol. 17, n. 4, pp. 39-46, Aug. 2014.

[2] S. Salicone, "The mathematical theory of evidence and measurement uncertainty. Expression and combination of measurement results via the random-fuzzy variables," *IEEE Instrum. Meas. Mag.*, vol. 17, no. 5, pp. 36-45, Oct. 2014.

[3] JCGM 106: 2012, *Evaluation of Measurement Data – The role of measurement uncertainty in conformity assessment*, Joint Committee for Guides in Metrology, 2012.

[4] JCGM 100: 2008, *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement, (GUM 1995 with minor corrections)*, Joint Committee for Guides in Metrology, 2008. [Online]. Available: http://www.bipm.org/en/publications/guides/gum.html.

[5] A. Ferrero and S. Salicone, "An original fuzzy method for the comparison of measurement results represented as random-fuzzy variables," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 4, pp. 1292-1299, Apr. 2007.