Comparison between two modern uncertainty expression and propagation approaches

M Pertile, S Debei
CISAS G. Colombo University of Padova, Italy
marco.pertile@unipd.it

Abstract. Two different uncertainty expression and propagation approaches are presented and compared. In particular, an implementation of the known probabilistic approach and a new Random-Fuzzy Variable (RFV) method based on the theory of Evidence. Both approaches use an explicit time correlation of input quantities to take into account systematic contributions. Numerical results show that both the type of uncertainty contribution (random or systematic) and the owned level of knowledge (Probability density Function PDF or simply a limited interval) must be carefully evaluated in uncertainty analysis. The new RFV approach allows to seamlessly deal with PDFs and limited intervals. This advantage is not present in the probabilistic approach, which yields questionable results in complete ignorance situations.

1. Introduction
In order to perform a measurement, it is necessary to choose a suitable method able to express the uncertainty associated with all the quantities that contribute to form the quantity subject to measurement, both in direct and indirect way by means of a known function (mathematical model). After that, it is necessary to define a method for the propagation of the selected contributions, so as to obtain an evaluation of the uncertainty associated with the measurand.

Known and usually accepted procedures for uncertainty expression and propagation are based on probability theory, e.g. see [1], [2]. Some probabilistic approaches can deal both with random and systematic contributions to uncertainty, provided that some tricks that are briefly described in the present work are applied. However, these probabilistic methods could lead to some questionable results in presence of complete ignorance situations. Expressing a situation of complete ignorance within a range by a uniform Probability Density Function PDF means assign the same probability to all values within the allowed interval, whereas the unknown PDF could have whatever shape. Thus, using an uniform PDF in these cases could mean that not owned information is arbitrarily introduced in the model as explained in [3], [4]. Conversely, complete ignorance situations can be correctly expressed by intervals and the mathematics of the intervals, as described in [4] and [5]. In some recent works [3], [6] - [11], an effective and interesting approach to express and propagate uncertainty has been proposed. This approach expresses uncertainty using Random-Fuzzy Variables RFV and is based on the theory of evidence, joining the most useful features of the probabilistic method and the mathematics of the intervals. A possible limit of this approach is that systematic contributions are always expressed by possibility distributions, assuming for these contributions a complete ignorance.

1 To whom any correspondence should be addressed.
situation within a certain interval. The knowledge only of a limit interval for a systematic contribution is frequent in engineering practice, but there are also systematic contributions whose PDF can be evaluated. An example is the length uncertainty of a ruler for linear measurements: if the same ruler is employed, its deviation from the nominal length yields a systematic contribution to uncertainty, since it is the same for all the acquired length values. Even so, a PDF of the deviation from the nominal ruler length can be evaluated, e.g. from statistical control of the production process.

The present work wishes to improve the known RFV method, in order to take into account also the situation of a systematic uncertainty contribution, that is known by a PDF and not only by a limit interval. The new RFV approach keeps the advantage of being a non-parametric estimation method. This particular characteristic can be very useful, as the application of this method does requires very little modeling or assumptions. The introduced RFV method can successfully be applied also when the indirect measurement involves a functional relation (mathematical model) between the output measurement quantities and their contributing input quantities which comprises numerical models (finite element analyses, dynamic systems simulations, identification algorithms like linear/non-linear regression, optimization algorithms, or other complex relations).

The new RFV method is compared with a probabilistic approach based on the probability theory and the Monte Carlo propagation, which is already known [12], [13]. In this work a particular probabilistic approach is introduced and applied. This approach propagates systematic contributions employing a mathematical propagation model with explicit expression of time and assuming complete time correlation for this type of uncertainty sources. This assumption may greatly modify the propagated uncertainty when the output quantities are affected by past values of one or more input time-varying quantities, as in several dynamical systems. Only if all input quantities instantly yield the output quantities, i.e. each output at time \( t_k \) only depends on inputs at the same time \( t_k \), the time correlation of uncertainty sources does not affect the propagated uncertainty. But, if one output at time \( t_k \) depends on previous inputs at time \( t_1, \ldots, t_{k-1} \), time correlation of the uncertainty contributions to those inputs becomes important. Particularly, when all measured values, from an initial time instant \( t_1 \) to a final instant \( t_N \), of one time-varying input quantity are used in the indirect measurement, the time correlation hypothesis implies that the uncertainty contribution to those input values will be roughly similar for all time instants. Conversely, if there is no time correlation, the uncertainty contribution to the considered input quantity may vary from time \( t_k \) to time \( t_{k+1} \) according to the associated probability density function. Thus, the effect of this behavior depends on the propagation model. Two simple examples: if the propagation model calculates an average of values corresponding to different time instants, which is an additive model, the propagated uncertainty does not decrease in the time correlation case whereas it is reduced when there is no correlation; on the contrary, if the propagation model calculates differences of values corresponding to consecutive time instants, i.e. it is a differential model, the propagated uncertainty is smaller in the time correlation case, since a compensation occurs. Thus, since the time correlation of one or more uncertainty contributions may yield very different results, its presence or absence should explicitly be analyzed during the propagation procedure. In this way, for an indirect measurement, a propagation model whose input quantities may explicitly vary with time is required.

In the present work, the following two methodologies have been examined and introduced:

1. Uncertainty expressed by PDFs and propagated by means of Monte Carlo simulation, [12], [13]; the employed method is known, but in the present paper its application to the propagation of systematic uncertainty contributions using explicit time correlation is described and applied to two examples; the cited literature does not explicitly deals with this subject.

2. Uncertainty expressed by RFVs and propagation in agreement with the Theory of Evidence; the presented method is new and different from the one described in [3], [4], [6] - [11], since in the presented one the uncertainty systematic contributions can be expressed both by PDFs and possibility distributions and are propagated using an explicit time correlation. Furthermore, [3] applies the old RFV only to an additive propagation model, whose behavior is very different from a differential model.
in presence of uncertainty systematic contributions. In the present work, two different examples are illustrated and discussed, one of additive type and one of differential type.

Employed methods are briefly described in section 2, while the results obtained by the two approaches are compared in section 3.

2. Approaches description

2.1. Probabilistic approach

The Guide to the Expression of Uncertainty in Measurement (GUM) [1] is the internationally accepted reference document for uncertainty evaluation. A basic concept expressed in the GUM, and even more strengthened in the supplement [13], is that a distribution of probability is the optimal mean in order to express the available information on possible values of a quantity. Thus, in this approach a PDF is associated to every identified uncertainty source. The aim is to evaluate the uncertainty to be associated with an output quantity (variable) Y, obtainable as indirect measurement of n input quantities (variables) X1, …, Xn. As explained in [2], a good approach in order to obtain a correct confidence interval for the output quantity Y requires the propagation of PDFs of all input quantities X, in order to arrive to the evaluation of the PDF associated with Y. After that, if the PDF of Y is unimodal, an unique smallest confidence interval for a desired coverage probability (e.g. 0.997) can be found out. The PDF of the output Y can be estimated by means of the Monte Carlo method, which is described in [2], [12], [13]. After Niter iterations, the Monte Carlo simulation allows to evaluate a frequency distribution dividing the whole interval in small bins. This frequency can be normalized and employed as an approximation of the PDF of Y, from which the expectation values of Y and one or more intervals associated with desired levels of confidence (probability) can be numerically calculated.

In order to try and take into account the presence of systematic contributions in the uncertainty sources, a slight modification to the Monte Carlo simulation used to propagate the uncertainty can be introduced. The evaluation of the PDF of Y is usually performed using a repeated sampling from each PDF of the input quantities X1, …, Xn. The sampling phase from a PDF employs a random number generator, with rectangular distribution in the interval (0,1), and the value of the corresponding input quantity is calculated using the inverse Cumulative Distribution Function (CDF), see [2], [3] and [13] for details. If the propagation model can employ different values of the generic input quantity Xi, i.e. X(t1), X(t2), …, X(tn) are present in the model instead of a constant Xi for t1,..,tn, then the sampling phase from the PDF can be performed in two ways, depending on the time correlation of Xi.

In the case that value X(tk) is completely independent of values of Xi at all other time instants, a new random number is generated at each time step to calculate the source values from the inverse CDFs in the sampling phase. Conversely, if the value X(tk) is completely correlated with the values of the same quantity at all other time instants, the same random number is used for all measured time instants in the sampling phase of the Monte Carlo simulation. In the case of complete time correlation of an uncertainty source, the random number used to calculate values of the source from the inverse CDF is the same for all time instants, but it is changed randomly at each Monte Carlo iteration. Generally, assigning the same random number to all time instants does not mean to assign the same value over time to the uncertainty source, since the PDF (and thus the inverse CDF) may vary over time, e.g. the expected value may be different over time.

The method outlined above to treat uncertainty sources over time, suggests a simple way to propagate both random and systematic effects using the probabilistic approach. Since a systematic effect should be repeatable and should yield a similar contribution every time a measurement is acquired, a complete time correlation of the corresponding uncertainty source seems more correct than neglecting the time correlation. Moreover, as explained in [14], for systematic contributions the operation of averaging of time series of data should not reduce the uncertainty. This behavior can be obtained by explicitly introducing time correlation for this type of contributions in the probability approach. Thus, for systematic contributions, the same random number is used for all time instants in...
one iteration of the Monte Carlo simulation, and it is changed every iteration. Conversely, a complete
independence over time is assumed for random contributions, since they are supposed to vary every
time a measurement is performed. In this case a different random number is used for each time instant
at each Monte Carlo iteration.

2.2. RFV approach
In some recent works [4], [6] - [11], an approach to express and propagate uncertainty has been
proposed. This approach uses the concept of random-fuzzy variables that are special cases of fuzzy
variables of type 2, and can be defined within the mathematical theory evidence. The latter
encompasses both the probability theory that allows to define random variables, and the possibility
theory which gives theoretical support for fuzzy variables. A detailed description can be found in the
cited papers. In the new approach used here, the uncertainty contributions (both random and
systematic) whose PDF can be evaluated or is known from a priori data, are expressed by their PDFs.
Conversely, if for some uncertainty contributions (both random and systematic) only limit intervals
can be evaluated, a fuzzy variable or possibility distribution is used to express their contribution. Thus,
differently from the known RFV method, in the new one, a possibility distribution is used only to
express uncertainty contributions in presence of complete ignorance situations (i.e. the quantity is
confined within a limited interval with a known level of confidence, but nothing can be said inside that
interval). PDFs and possibility distributions are propagated separately and are merged together only at
the end of the procedure, when the RFV of the output quantity Y is evaluated.

In this new RFV method, the different behavior of random and systematic uncertainty contributions
is taken into account in the same way as in the probabilistic approach: explicitly enforcing complete
time correlation for systematic contributions and complete time independence for random
contributions.

As explained in [3] and [4], a desired number of levels of confidence 1-α is chosen to express the
uncertainty of output Y (or input quantity X_i(tk)). For each level 1-α, the uncertainty of Y (or X_i(tk)) is
expressed by the external interval [ξ_1,α, ξ_4,α] which is the confidence interval of Y (or X_i(tk)) associated
with level 1-α. Furthermore, the expression by means of fuzzy variables of type 2 allows to ascribe,
for each level of confidence, all uncertainty due to complete ignorance situations (only a limit interval
is known and not a PDF) to the internal interval [ξ_2,α, ξ_3,α], while contributions whose PDF can be
evaluated are ascribed to the lateral intervals [ξ_1,α, ξ_2,α] and [ξ_3,α, ξ_4,α]. ξ_1 - ξ_4 are possible values of Y,
arranged in increasing order ξ_1 ≤ ξ_2 ≤ ξ_3 ≤ ξ_4.

Uncertainty propagation from input quantities X_i(tk) to output Y is performed separately for
contributions whose PDF is known and for contributions known only by a limit interval. All
uncertainty contributions to input quantities whose PDF is known are propagated in the same way as
in the probabilistic approach, i.e. with a Monte Carlo simulation. After a PDF for output Y is
evaluated, a probability to possibility transformation is performed; the two sides of the obtained
possibility distribution define the lateral intervals [ξ_1,α, ξ_2,α] and [ξ_3,α, ξ_4,α] and are laterally added to
the internal interval [ξ_2,α, ξ_3,α], as described in [3],[4]. The internal interval [ξ_2,α, ξ_3,α] of Y is obtained
propagating the complete ignorance uncertainty contributions of the input quantities X_i(tk) with the
following procedure: for each uncertainty interval of the inputs X_i(tk) a corresponding uniform PDF of
equal amplitude is defined; these uniform PDFs are propagated with a second Monte Carlo simulation
to evaluate a second PDF of Y, which is different from the first PDF of Y obtained above; the last step
is to define the internal interval [ξ_2,α, ξ_3,α] as the one with level of confidence 1. From a computational
point of view, the method employed to express and propagate uncertainty is the same as that described
in [3] and [4]. The main difference is that now the internal interval is associated only with uncertainty
contributions known by a limit interval (complete ignorance situation), while both random and
systematic contributions are processed employing the same time correlation described in section 2.1.
3. Examples and discussion

The two procedures described above are applied to find out the propagated uncertainty in two very simple examples, in order to explain the major differences between a probability based approach and the new evidence based approach. In both the following examples, two quantities $X_1$ and $X_2$ are assumed to be directly measured in two time instants, i.e. $X_1(t_1), X_1(t_2), X_2(t_1), X_2(t_2)$ are the input quantities. The indirect measurement is performed using two simple propagation models: the first one of additive type $Y=X_1(t_1)+X_1(t_2)+X_2(t_1)+X_2(t_2)$; the second one of differential type $Y=X_1(t_1)-X_1(t_2)+X_2(t_1)-X_2(t_2)$. For each propagation model, $X_1$ is always affected by a random contribution expressed by a uniform PDF (from 3 to 5 for $X_1(t_1)$, from 4 to 6 for $X_1(t_2)$), while four cases are analyzed for the uncertainty of $X_2$:

1) random type with a known uniform PDF;
2) systematic type with a known uniform PDF;
3) random type with a limited interval (complete ignorance case);
4) systematic type with a limited interval;

in all four cases the uniform PDF or the limited interval is from 1 to 3 for $X_2(t_1)$, from 3 to 5 for $X_2(t_2)$.

The obtained results are depicted in the following figures: cases 1 - 4 of the additive model in figures 1-4, case 1 - 4 of the differential model in figures 5-8. In each figure, the magenta line defines the PDF of $Y$ obtained by the probabilistic approach; the red and the blue lines depict respectively the external and internal intervals of the RFV of $Y$ obtained by the evidence approach; the right vertical axis shows the value of $\alpha$ ($1-\alpha$ is the level of confidence); the magenta horizontal intervals are the confidence intervals calculated from the PDF of $Y$ of the probabilistic approach, with the corresponding level of confidence $1-\alpha$ that can be read on the right vertical axis. In this way, for the three selected heights (levels of confidence), the width of the confidence intervals obtained by the probabilistic approach can be directly compared with the corresponding external intervals obtained by the RFV approach having the same level of confidence. The reader should note that the horizontal width of the PDF can not be directly compared with the horizontal width of the RFV, since the latter is a possibility distribution.

![Figure 1. Case 1, model $X_1(t_1)+X_1(t_2)+X_2(t_1)+X_2(t_2)$.

Figure 1. Case 1, model $X_1(t_1)+X_1(t_2)+X_2(t_1)+X_2(t_2)$.

Figure 2. Case 2, model $X_1(t_1)+X_1(t_2)+X_2(t_1)+X_2(t_2)$.

In both figures 1 and 2, there is no difference in the confidence intervals obtained by the two approaches: with the additive propagation model, the two approaches yield the same results when the PDFs of all inputs are known, both in case of random contributions (figure 1) and in case of systematic ones (figure 2).

Comparing figures 1 and 2, the presence of a systematic contribution (figure 2) yields wider confidence intervals than in case 1, when all contributions are of random type. This result is due to the partial compensation that takes place when the sum of random contributions is preformed. This partial compensation in an additive model does not happen for systematic contributions.
In presence of complete ignorance situations (figures 3 and 4), nothing changes for the probabilistic approach with reference to cases 1 and 2: the probabilistic approach does not differentiate between PDF expressed and interval expressed uncertainties, and with an additive model keeps on yielding larger results in presence of systematic contributions. In both cases 3 and 4, the RFV approach gives wider confidence intervals than the probabilistic one, both in presence of random (case 3) and systematic (case 4) contributions. Moreover, the results obtained by the RFV approach are identical in cases 3 and 4. This results can be explained by the following considerations:

- when the four contributions are of random type and are expressed by uniform PDFs (case 1 and 3 in the probabilistic approach, case 1 in the RFV one), the additive propagation model yields partial compensation between the two variables $X_1$ and $X_2$, and between the two time instants $t_1$ and $t_2$; thus, the obtained confidence intervals are the smallest ones.
- When two contributions are of systematic type ($X_2(t_1)$, $X_2(t_2)$) and are expressed by PDFs (case 2 and 4 in the probabilistic approach, case 2 in the RFV one), the additive propagation model does not yield partial compensation between the two contributions $X_2(t_1)$ and $X_2(t_2)$; thus wider confidence intervals are expected.
- When two contributions ($X_2(t_1)$, $X_2(t_2)$) are expressed by intervals (case 3 and 4 in the RFV approach), the additive propagation model does not yield partial compensation between the two contributions $X_2(t_1)$ and $X_2(t_2)$ nor between the two variables $X_1$ and $X_2$; thus the largest confidence intervals are obtained.
Figure 7. Case 3, model \(X_1(t_1) - X_1(t_2) + X_2(t_1) - X_2(t_2)\).

Figure 8. Case 4, model \(X_1(t_1) - X_1(t_2) + X_2(t_1) - X_2(t_2)\).

Similar results are obtained for the differential propagation model. The main difference is that the systematic contributions (case 2 and 4) completely compensate each other, as depicted in figures 6 and 8, yielding the smallest confidence intervals.

4. Conclusion

Two different uncertainty expression and propagation approaches are presented and compared. A particular implementation of the known probabilistic approach is described. This method tries to take into account the repetitive nature of systematic uncertainty sources exploiting time correlation. A new RFV based on the theory of Evidence is presented. Also in this new approach an explicit time correlation of input quantities is used to take in account systematic contributions. This approach allows to analyse and process uncertainty contributions (both of random and systematic type) whose PDFs are known and also situations of complete ignorance (both random and systematic), when only a limited interval is known for each quantity.

These methods are applied to two simple propagation models simulating uncertainty evaluation in indirect measurement. One selected propagation model is of additive type and the other is of differential type. Results obtained greatly depends on the different assumptions made for the input quantity, showing that both the type of contribution (random or systematic) and the owned level of knowledge (PDF or simply a limited interval) must be carefully evaluated in uncertainty analysis. The new RFV approach allows to seamlessly deal with PDFs and limited intervals. This advantage is not present in the probabilistic approach, which yields questionable results in complete ignorance situations.

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML 1995 *Guide to the Expression of Uncertainty in Measurement*, (Geneva, Switzerland: International Organization for Standardization), ISBN 92-67-10188-9;

[2] M Cox, B R L Siebert, 2006, *Metrologia*, vol. 43, S178-S188.

[3] M Pertile, M De Cecco, 2008, *Meas. Sci. Technol.*, 19(055103), 10pp;

[4] S Salicone, *Measurement Uncertainty, An Approach via the Mathematical Theory of Evidence* Springer, 2007, ISBN 978-0-387-30655-1;

[5] G J Klir, B Yuan, 1995, *Fuzzy Sets and Fuzzy Logic. Theory and Applications*, Englewood Cliffs, NJ: Prentice-Hall;

[6] A Ferrero, S Salicone, 2004, *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1370–1377;

[7] A Ferrero, S Salicone, 2006, *IEEE Trans. Instrum. Meas.*, vol. 55, no. 3, pp. 706–712;

[8] A Ferrero, R Gamba, S Salicone, 2004, *IEEE Trans. Instrum. Meas.*, vol.53,no.5,pp.1362–1369;

[9] A Ferrero, S Salicone, 2005, *IEEE Trans. Instrum. Meas.*, vol. 54, no. 4, pp. 1475–1481;

[10] A Ferrero, S Salicone, 2007, *IEEE Trans. Instrum. Meas.*, vol.56, no. 4, pp. 1292 – 1299;

[11] A Ferrero, S Salicone, 2007, *IEEE Trans. Instrum. Meas.*, vol. 56, no. 3, pp. 704 – 716;
[12] M Cox, P Harris, B R L Siebert, 2003, *Measurement Techniques*, Vol.46, no. 9;
[13] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML *Evaluation of measurement data—Supplement 1 to the ‘Guide to the Expression of Uncertainty in Measurement’—Propagation of distributions using a Monte Carlo method*;
[14] M K Urbanskia, J Wasowskib, 2003, *Measurement*, vol. 34, pp. 67–74;
[15] G Mauris, V Lasserre, L Foulloy, 2001, *Measurement*, vol. 29, pp. 165–177;
[16] D Dubois, L Foulloy, G Mauris, H Prade, 2004, *Reliable Computing*, vol. 10, pp. 273–297.