Combining Interval, Probabilistic, and Fuzzy Uncertainty: Foundations, Algorithms, Challenges – An Overview

Vladik Kreinovich[1], Daniel J. Berleant[2], Scott Ferson[3], and Weldon A. Lodwick[4]

[1] University of Texas at El Paso vladik@utep.edu
[2] Iowa State University berleant@iastate.edu
[3] Applied Biomathematics scott@ramas.com
[4] University of Colorado at Denver wlodwick@math.cudenver.edu

Summary. Since the 1960s, many algorithms have been designed to deal with interval uncertainty. In the last decade, there has been a lot of progress in extending these algorithms to the case when we have a combination of interval, probabilistic, and fuzzy uncertainty. We provide an overview of related algorithms, results, and remaining open problems.

1 Main Problem

Why indirect measurements? In many real-life situations, we are interested in the value of a physical quantity y that is difficult or impossible to measure directly. Examples of such quantities are the distance to a star and the amount of oil in a given well. Since we cannot measure y directly, a natural idea is to measure y *indirectly*. Specifically, we find some easier-to-measure quantities $x_1, \ldots, x_n$ which are related to y by a known relation $y = f(x_1, \ldots, x_n)$; this relation may be a simple functional transformation, or complex algorithm (e.g., for the amount of oil, numerical solution to an inverse problem). Then, to estimate y, we first measure the values of the quantities $x_1, \ldots, x_n$, and then we use the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of these measurements to to compute an estimate $\widetilde{y}$ for y as $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

For example, to find the resistance R, we measure current I and voltage V, and then use the known relation $R = V/I$ to estimate resistance as $\widetilde{R} = \widetilde{V}/\widetilde{I}$.

Computing an estimate for y based on the results of direct measurements is called *data processing*; data processing is the main reason why computers were invented in the first place, and data processing is still one of the main uses of computers as number crunching devices.

Comment. In this paper, for simplicity, we consider the case when the relation between x_i and y is known exactly; in some practical situations, we only known an approximate relation between x_i and y.

Why interval computations? From computing to probabilities to intervals. Measurement are never 100% accurate, so in reality, the actual value x_i of i-th measured quantity can differ from the measurement result $\widetilde{x}_i$. Because of these *measurement errors* $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$, the result $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ of data processing is, in general, different from the actual value $y = f(x_1, \ldots, x_n)$ of the desired quantity y.

It is desirable to describe the error $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$ of the result of data processing. To do that, we must have some information about the errors of direct measurements.

What do we know about the errors Δx_i of direct measurements? First, the manufacturer of the measuring instrument must supply us with an upper bound Δ_i on the measurement error. If no such upper bound is supplied, this means that no accuracy is guaranteed, and the corresponding "measuring instrument" is practically useless. In this case, once we performed a measurement and got a measurement result $\widetilde{x}_i$, we know that the actual (unknown) value x_i of the measured quantity belongs to the interval $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$, where $\underline{x}_i = \widetilde{x}_i - \Delta_i$ and $\overline{x}_i = \widetilde{x}_i + \Delta_i$.

In many practical situations, we not only know the interval $[-\Delta_i, \Delta_i]$ of possible values of the measurement error; we also know the probability of different values Δx_i within this interval. This knowledge underlies the traditional engineering approach to estimating the error of indirect measurement, in which we assume that we know the probability distributions for measurement errors Δx_i.

In practice, we can determine the desired probabilities of different values of Δx_i by comparing the results of measuring with this instrument with the results of measuring the same quantity by a standard (much more accurate) measuring instrument. Since the standard measuring instrument is much more accurate than the one use, the difference between these two measurement results is practically equal to the measurement error; thus, the empirical distribution of this difference is close to the desired probability distribution for measurement error. There are two cases, however, when this determination is not done:

- First is the case of cutting-edge measurements, e.g., measurements in fundamental science. When a Hubble telescope detects the light from a distant galaxy, there is no "standard" (much more accurate) telescope floating nearby that we can use to calibrate the Hubble: the Hubble telescope is the best we have.
- The second case is the case of measurements on the shop floor. In this case, in principle, every sensor can be thoroughly calibrated, but sensor calibration is so costly – usually costing ten times more than the sensor itself – that manufacturers rarely do it.

In both cases, we have no information about the probabilities of Δx_i; the only information we have is the upper bound on the measurement error.

In this case, after we performed a measurement and got a measurement result $\widetilde{x}_i$, the only information that we have about the actual value x_i of the measured quantity is that it belongs to the interval $\mathbf{x}_i = [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. In such situations, the only information that we have about the (unknown) actual value of $y = f(x_1, \ldots, x_n)$ is that y belongs to the range $\mathbf{y} = [\underline{y}, \overline{y}]$ of the function f over the box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$:

$$\mathbf{y} = [\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \,|\, x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\}.$$

The process of computing this interval range based on the input intervals $\mathbf{x}_i$ is called *interval computations*; see, e.g., [6, 7].

Interval computations techniques: brief reminder. Historically the first method for computing the enclosure for the range is the method which is sometimes called "straightforward" interval computations. This method is based on the fact that inside the computer, every algorithm consists of elementary operations (arithmetic operations, min, max, etc.). For each elementary operation $f(a, b)$, if we know the intervals $\mathbf{a}$ and $\mathbf{b}$ for a and b, we can compute the exact range $f(\mathbf{a}, \mathbf{b})$. The corresponding formulas form the so-called *interval arithmetic*. For example,

$$[\underline{a}, \overline{a}] + [\underline{b}, \overline{b}] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]; \quad [\underline{a}, \overline{a}] - [\underline{b}, \overline{b}] = [\underline{a} - \overline{b}, \overline{a} - \underline{b}];$$

$$[\underline{a}, \overline{a}] \cdot [\underline{b}, \overline{b}] = [\min(\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}), \max(\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b})].$$

In straightforward interval computations, we repeat the computations forming the program f step-by-step, replacing each operation with real numbers by the corresponding operation of interval arithmetic. It is known that, as a result, we get an enclosure $\mathbf{Y} \supseteq \mathbf{y}$ for the desired range.

In some cases, this enclosure is exact. In more complex cases (see examples below), the enclosure has excess width.

There exist more sophisticated techniques for producing a narrower enclosure, e.g., a centered form method. However, for each of these techniques, there are cases when we get an excess width. Reason: as shown in [10], the problem of computing the exact range is known to be NP-hard even for polynomial functions $f(x_1, \ldots, x_n)$ (actually, even for quadratic functions f).

Practical problem. In some practical situations, in addition to the lower and upper bounds on each random variable x_i, we have some additional information about x_i. So, we arrive at the following problem:

- we have a data processing algorithm $f(x_1, \ldots, x_n)$, and
- we have some information about the uncertainty with which we know x_i (e.g., measurement errors).

We want to know the resulting uncertainty in the result $y = f(x_1, \ldots, x_n)$ of data processing.

In interval computations, we assume that the uncertainty in x_i can be described by the interval of possible values. In real life, in addition to the intervals, we often have some information about the probabilities of different values within this interval. What can we then do?

2 What is the Best Way to Describe Probabilistic Uncertainty?

How is the partial information about probabilities used in decision making? One of the main objectives of data processing is to make decisions. A standard way of making a decision is to select the action a for which the expected utility (gain) is the largest possible. This is where probabilities are used: in computing, for every possible action a, the corresponding expected utility. To be more precise, we usually know, for each action a and for each actual value of the (unknown) quantity x, the corresponding value of the utility $u_a(x)$. We must use the probability distribution for x to compute the expected value $E[u_a(x)]$ of this utility.

In view of this application, the most useful characteristics of a probability distribution would be the ones which would enable us to compute the expected value $E[u_a(x)]$ of different functions $u_a(x)$.

Which representations are the most useful for this intended usage? General idea. Which characteristics of a probability distribution are the most useful for computing mathematical expectations of different functions $u_a(x)$? The answer to this question depends on the type of the function, i.e., on how the utility value u depends on the value x of the analyzed parameter.

Smooth utility functions naturally lead to moments. One natural case is when the utility function $u_a(x)$ is smooth. We have already mentioned, in Section I, that we usually know a (reasonably narrow) interval of possible values of x. So, to compute the expected value of $u_a(x)$, all we need to know is how the function $u_a(x)$ behaves on this narrow interval. Because the function is smooth, we can expand it into Taylor series. Because the interval is narrow, we can safely consider only linear and quadratic terms in this expansion and ignore higher-order terms: $u_a(x) \approx c_0 + c_1 \cdot (x - x_0) + c_2 \cdot (x - x_0)^2$, where x_0 is a point inside the interval. Thus, we can approximate the expectation of this function by the expectation of the corresponding quadratic expression: $E[u_a(x)] \approx E[c_0 + c_1 \cdot (x - x_0) + c_2 \cdot (x - x_0)^2]$, i.e., by the following expression: $E[u_a(x)] \approx c_0 + c_1 \cdot E[x - x_0] + c_2 \cdot E[(x - x_0)^2]$. So, to compute the expectations of such utility functions, it is sufficient to know the first and second moments of the probability distribution.

In particular, if we use, as the point x_0, the average $E[x]$, the second moment turns into the variance of the original probability distribution. So, instead of the first and the second moments, we can use the mean E and the variance V.

In decision making, non-smooth utility functions are common. In decision making, not all dependencies are smooth. There is often a threshold x_0 after which, say, a concentration of a certain chemical becomes dangerous.

This threshold sometimes comes from the detailed chemical and/or physical analysis. In this case, when we increase the value of this parameter, we see the drastic increase in effect and hence, the drastic change in utility value. Sometimes, this threshold simply comes from regulations. In this case, when we increase the value of this parameter past the threshold, there is no drastic increase in effects, but there is a drastic decrease of utility due to the necessity to pay fines, change technology, etc. In both cases, we have a utility function which experiences an abrupt decrease at a certain threshold value x_0.

Non-smooth utility functions naturally lead to CDFs. We want to be able to compute the expected value $E[u_a(x)]$ of a function $u_a(x)$ which changes smoothly until a certain value x_0, then drops it value and continues smoothly for $x > x_0$. We usually know the (reasonably narrow) interval which contains all possible values of x. Because the interval is narrow and the dependence before and after the threshold is smooth, the resulting change in $u_a(x)$ before x_0 and after x_0 is much smaller than the change at x_0. Thus, with a reasonable accuracy, we can ignore the small changes before and after x_0, and assume that the function $u_a(x)$ is equal to a constant u^+ for $x < x_0$, and to some other constant $u^- < u^+$ for $x > x_0$.

The simplest case is when $u^+ = 1$ and $u^- = 0$. In this case, the desired expected value $E[u_a^{(0)}(x)]$ coincides with the probability that $x < x_0$, i.e., with the corresponding value $F(x_0)$ of the cumulative distribution function (CDF). A generic function $u_a(x)$ of this type, with arbitrary values u^- and u^+, can be easily reduced to this simplest case, because, as one can easily check, $u_a(x) = u^- + (u^+ - u^-) \cdot u^{(0)}(x)$ and hence, $E[u_a(x)] = u^- + (u^+ - u^-) \cdot F(x_0)$.

Thus, to be able to easily compute the expected values of all possible non-smooth utility functions, it is sufficient to know the values of the CDF $F(x_0)$ for all possible x_0.

3 How to Represent Partial Information about Probabilities

General idea. In many cases, we have a complete information about the probability distributions that describe the uncertainty of each of n inputs.

However, a practically interesting case is how to deal with situations when we only have partial information about the probability distributions. How can we represent this partial information?

Case of cdf. If we use cdf $F(x)$ to represent a distribution, then full information corresponds to the case when we know the exact value of $F(x)$ for every x. Partial information means:

- either that we only know approximate values of $F(x)$ for all x, i.e., that for every x, we only know the interval that contains $F(x)$; in this case, we get a *p-box*;
- or that we only know the values of $F(x)$ for some x, i.e, that we only know the values $F(x_1)$, ..., $F(x_n)$ for finitely many values $x = x_1, \ldots, x_n$; in this case, we have a *histogram*.

It is also possible that we know only approximate values of $F(x)$ for some x; in this case, we have an *interval-valued histogram*.

Case of moments. If we use moments to represent a distribution, then partial information means that we either know the exact values of finitely many moments, or that we know intervals of possible values of several moments.

Resulting problems. This discussion leads to a natural classification of possible problems:

- If we have complete information about the distributions of x_i, then, to get validated estimates on uncertainty of y, we have to use Monte-Carlo-type techniques; see, e.g., [12, 13].
- If we have p-boxes, we can use methods from [4].
- If we have histograms, we can use methods from [1, 2].
- If we have moments, then we can use methods from [9].

There are also additional issues, including:

- how we get these bounds for x_i?
- specific practical applications, like the appearance of histogram-type distributions in problems related to privacy in statistical databases.

4 Case Study

Practical problem. In some practical situations, in addition to the lower and upper bounds on each random variable x_i, we know the bounds $\mathbf{E}_i = [\underline{E}_i, \overline{E}_i]$ on its mean E_i. Indeed, in measurement practice (see, e.g., [11]), the overall measurement error Δx is usually represented as a sum of two components:

- a *systematic* error component $\Delta_s x$ which is defined as the expected value $E[\Delta x]$, and
- a *random* error component $\Delta_r x$ which is defined as the difference between the overall measurement error and the systematic error component: $\Delta_r x \stackrel{\text{def}}{=} \Delta x - \Delta_s x$.

In addition to the bound Δ on the overall measurement error, the manufacturers of the measuring instrument often provide an upper bound Δ_s on the systematic error component: $|\Delta_s x| \le \Delta_s$.

This additional information is provided because, with this additional information, we not only get a bound on the accuracy of a single measurement, but we also get an idea of what accuracy we can attain if we use repeated measurements to increase the measurement accuracy. Indeed, the very idea that repeated measurements can improve the measurement accuracy is natural: we measure the same quantity by using the same measurement instrument several (N) times, and then take, e.g., an arithmetic average $\bar{x} = \dfrac{\widetilde{x}^{(1)} + \ldots + \widetilde{x}^{(N)}}{N}$ of the corresponding measurement results $\widetilde{x}^{(1)} = x + \Delta x^{(1)}, \ldots, \widetilde{x}^{(N)} = x + \Delta x^{(N)}$.

- If systematic error is the only error component, then all the measurements lead to exactly the same value $\widetilde{x}^{(1)} = \ldots = \widetilde{x}^{(N)}$, and averaging does not change the value – hence does not improve the accuracy.
- On the other hand, if we know that the systematic error component is 0, i.e., $E[\Delta x] = 0$ and $E[\widetilde{x}] = x$, then, as $N \to \infty$, the arithmetic average tends to the actual value x. In this case, by repeating the measurements sufficiently many times, we can determine the actual value of x with an arbitrary given accuracy.

In general, by repeating measurements sufficiently many times, we can arbitrarily decrease the random error component and thus attain accuracy as close to Δ_s as we want.

When this additional information is given, then, after we performed a measurement and got a measurement result $\widetilde{x}$, then not only we get the information that the actual value x of the measured quantity belongs to the interval $\mathbf{x} = [\widetilde{x} - \Delta, \widetilde{x} + \Delta]$, but we can also conclude that the expected value of $x = \widetilde{x} - \Delta x$ (which is equal to $E[x] = \widetilde{x} - E[\Delta x] = \widetilde{x} - \Delta_s x$) belongs to the interval $\mathbf{E} = [\widetilde{x} - \Delta_s, \widetilde{x} + \Delta_s]$.

If we have this information for every x_i, then, in addition to the interval $\mathbf{y}$ of possible value of y, we would also like to know the interval of possible values of $E[y]$. This additional interval will hopefully provide us with the information on how repeated measurements can improve the accuracy of this indirect measurement. Thus, we arrive at the following problem:

Precise formulation of the problem. Given an algorithm computing a function $f(x_1, \ldots, x_n)$ from R^n to R, and values $\underline{x}_1$, $\overline{x}_1$, …, $\underline{x}_n$, $\overline{x}_n$, $\underline{E}_1$, $\overline{E}_1$, …, $\underline{E}_n$, $\overline{E}_n$, we want to find

$$\underline{E} \stackrel{\text{def}}{=} \min\{E[f(x_1, \ldots, x_n)] \mid \text{ all distributions of } (x_1, \ldots, x_n) \text{ for which}$$

$$x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_n \in [\underline{x}_n, \overline{x}_n], E[x_1] \in [\underline{E}_1, \overline{E}_1], \ldots E[x_n] \in [\underline{E}_n, \overline{E}_n]\};$$

and $\overline{E}$ which is the maximum of $E[f(x_1, \ldots, x_n)]$ for all such distributions.

In addition to considering all possible distributions, we can also consider the case when all the variables x_i are independent.

How we solve this problem. The main idea behind straightforward interval computations can be applied here as well. Namely, first, we find out how to solve this problem for the case when $n = 2$ and $f(x_1, x_2)$ is one of the standard arithmetic operations. Then, once we have an arbitrary algorithm $f(x_1, \ldots, x_n)$, we parse it and replace each elementary operation on real numbers with the corresponding operation on quadruples $(\underline{x}, \underline{E}, \overline{E}, \overline{x})$.

To implement this idea, we must therefore know how to, solve the above problem for elementary operations.

For *addition*, the answer is simple. Since $E[x_1 + x_2] = E[x_1] + E[x_2]$, if $y = x_1 + x_2$, there is only one possible value for $E = E[y]$: the value $E = E_1 + E_2$. This value does not depend on whether we have correlation or nor, and whether we have any information about the correlation. Thus, $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2$.

Similarly, the answer is simple for *subtraction:* if $y = x_1 - x_2$, there is only one possible value for $E = E[y]$: the value $E = E_1 - E_2$. Thus, $\mathbf{E} = \mathbf{E}_1 - \mathbf{E}_2$.

For *multiplication*, if the variables x_1 and x_2 are independent, then $E[x_1 \cdot x_2] = E[x_1] \cdot E[x_2]$. Hence, if $y = x_1 \cdot x_2$ and x_1 and x_2 are independent, there is only one possible value for $E = E[y]$: the value $E = E_1 \cdot E_2$; hence $\mathbf{E} = \mathbf{E}_1 \cdot \mathbf{E}_2$.

The first non-trivial case is the case of multiplication in the presence of possible correlation. When we know the exact values of E_1 and E_2, the solution to the above problem is as follows:

Theorem 1. *For multiplication $y = x_1 \cdot x_2$, when we have no information about the correlation,*

$$\underline{E} = \max(p_1 + p_2 - 1, 0) \cdot \overline{x}_1 \cdot \overline{x}_2 + \min(p_1, 1 - p_2) \cdot \overline{x}_1 \cdot \underline{x}_2 + \min(1 - p_1, p_2) \cdot \underline{x}_1 \cdot \overline{x}_2 + \max(1 - p_1 - p_2, 0) \cdot \underline{x}_1 \cdot \underline{x}_2;$$

$$\overline{E} = \min(p_1, p_2) \cdot \overline{x}_1 \cdot \overline{x}_2 + \max(p_1 - p_2, 0) \cdot \overline{x}_1 \cdot \underline{x}_2 + \max(p_2 - p_1, 0) \cdot \underline{x}_1 \cdot \overline{x}_2 + \min(1 - p_1, 1 - p_2) \cdot \underline{x}_1 \cdot \underline{x}_2,$$

where $p_i \stackrel{\text{def}}{=} (E_i - \underline{x}_i)/(\overline{x}_i - \underline{x}_i)$.

Theorem 2. *For multiplication under no information about dependence, to find $\underline{E}$, it is sufficient to consider the following combinations of p_1 and p_2:*

- *$p_1 = \underline{p}_1$ and $p_2 = \underline{p}_2$; $p_1 = \underline{p}_1$ and $p_2 = \overline{p}_2$; $p_1 = \overline{p}_1$ and $p_2 = \underline{p}_2$; $p_1 = \overline{p}_1$ and $p_2 = \overline{p}_2$;*
- *$p_1 = \max(\underline{p}_1, 1 - \overline{p}_2)$ and $p_2 = 1 - p_1$ (if $1 \in \mathbf{p}_1 + \mathbf{p}_2$); and*
- *$p_1 = \min(\overline{p}_1, 1 - \underline{p}_2)$ and $p_2 = 1 - p_1$ (if $1 \in \mathbf{p}_1 + \mathbf{p}_2$).*

The smallest value of $\underline{E}$ for all these cases is the desired lower bound $\underline{E}$.

Theorem 3. *For multiplication under no information about dependence, to find $\overline{E}$, it is sufficient to consider the following combinations of p_1 and p_2:*

- $p_1 = \underline{p}_1$ *and* $p_2 = \underline{p}_2$*;* $p_1 = \underline{p}_1$ *and* $p_2 = \overline{p}_2$*;* $p_1 = \overline{p}_1$ *and* $p_2 = \underline{p}_2$*;* $p_1 = \overline{p}_1$ *and* $p_2 = \overline{p}_2$*;*
- $p_1 = p_2 = \max(\underline{p}_1, \underline{p}_2)$ (*if* $\mathbf{p}_1 \cap \mathbf{p}_2 \neq \emptyset$)*; and*
- $p_1 = p_2 = \min(\overline{p}_1, \overline{p}_2)$ (*if* $\mathbf{p}_1 \cap \mathbf{p}_2 \neq \emptyset$)*.*

The largest value of $\overline{E}$ *for all these cases is the desired upper bound* $\overline{E}$*.*

For the *inverse* $y = 1/x_1$, the finite range is possible only when $0 \notin \mathbf{x}_1$. Without losing generality, we can consider the case when $0 < \underline{x}_1$. In this case, we get the following bound:

Theorem 4. *For the inverse* $y = 1/x_1$*, the range of possible values of* E *is* $\mathbf{E} = [1/E_1, p_1/\overline{x}_1 + (1 - p_1)/\underline{x}_1]$.

(Here p_1 denotes the same value as in Theorem 1).

Similar formulas can be produced for max and min, and also for the cases when there is a strong correlation between x_i: namely, when x_1 is (non-strictly) increasing or decreasing in x_2; see, e.g., [9].

Additional results. The above techniques assume that we already know the moments etc., but how can we compute them based on the measurement results? For example, when we have only interval ranges $[\underline{x}_i, \overline{x}_i]$ of sample values $x_1, \ldots, x_n$, what is the interval $[\underline{V}, \overline{V}]$ of possible values for the variance V of these values?

It turns out that most such problems are computationally difficult (to be more precise, NP-hard), and we provide feasible algorithms that compute these bounds under reasonable easily verifiable conditions [5, 11].

5 Fuzzy Uncertainty: In Brief

In the fuzzy case, for each value of measurement error Δx_i, we describe the degree $\mu_i(\Delta x_i)$ to which this value is possible.

For each degree of certainty α, we can determine the set of values of Δx_i that are possible with at least this degree of certainty – the α-*cut* $\{x \mid \mu(x) \geq \alpha\}$ of the original fuzzy set. Vice versa, if we know α-cuts for every α, then, for each object x, we can determine the degree of possibility that x belongs to the original fuzzy set [3, 8, 13, 14, 15]. A fuzzy set can be thus viewed as a nested family of its α-cuts.

If instead of a (crisp) interval $\mathbf{x}_i$ of possible values of the measured quantity, we have a fuzzy set $\mu_i(x)$ of possible values, then we can view this information as a family of nested intervals $\mathbf{x}_i(\alpha)$ – α-cuts of the given fuzzy sets.

Our objective is then to compute the fuzzy number corresponding to this the desired value $y = f(x_1, \ldots, x_n)$. In this case, for each level α, to compute the α-cut of this fuzzy number, we can apply the interval algorithm to the α-cuts $\mathbf{x}_i(\alpha)$ of the corresponding fuzzy sets. The resulting nested intervals form the desired fuzzy set for y.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328, by NIH grant 3T34GM008048-20S1, and by the Army Research Lab grant DATM-05-02-C-0046.

This work was partly supported by a research grant from Sandia National Laboratories as part of the Department of Energy Accelerated Strategic Computing Initiative (ASCI).

References

1. Berleant D and Zhang J (2004), Using Pearson correlation to improve envelopes around the distributions of functions. Reliable Computing 10(2):139–161.
2. Berleant D and Zhang J (2004), Representation and Problem Solving with the Distribution Envelope Determination (DEnv) Method, Reliability Engineering and System Safety 85(1–3).
3. Bojadziev G and Bojadziev M (1995), *Fuzzy Sets, Fuzzy Logic, Applications*, World Scientific, Singapore.
4. Ferson S (2002), *RAMAS Risk Calc 4.0*, CRC Press, Boca Raton, Florida.
5. Ferson S, Ginzburg L, Kreinovich V, Longpré L, and Aviles M (2005), Exact Bounds on Finite Populations of Interval Data. Reliable Computing, 11(3):207–233.
6. Jaulin L, Kieffer M, Didrit O, and Walter E (2001), *Applied interval analysis*, Springer Verlag, London.
7. Kearfott RB and Kreinovich V, eds. (1996), *Applications of Interval Computations*, Kluwer, Dordrecht.
8. Klir G, Yuan, B (1995), *Fuzzy sets and fuzzy logic*, Prentice Hall, New Jersey.
9. Kreinovich V (2004), Probabilities, Intervals, What Next? Optimization Problems Related to Extension of Interval Computations to Situations with Partial Information about Probabilities. Journal of Global Optimization 29(3):265–280.
10. Kreinovich V, Lakeyev A, Rohn J, Kahl P (1997), *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht.
11. Kreinovich V et al. (to appear), Towards combining probabilistic and interval uncertainty in engineering calculations: algorithms for computing statistics under interval uncertainty, and their computational complexity. Reliable Computing
12. Lodwick WA and Jamison KD (2003), Estimating and Validating the Cumulative Distribution of a Function of Random Variables: Toward the Development of Distribution Arithmetic. Reliable Computing 9(2):127–141.
13. Moore RE and Lodwick WA (2003), Interval Analysis and Fuzzy Set Theory. Fuzzy Sets and Systems 135(1):5–9.
14. Nguyen HT and Kreinovich V (1996), Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications, In [7], pp. 245–290
15. Nguyen HT and Walker EA (1999), *First course in fuzzy logic*, CRC Press, Boca Raton, Florida.