Towards a More Realistic Representation of Uncertainty: an Approach Motivated by Info-Gap Decision Theory

Daniel Berleant
Department of Information Science
University of Arkansas at Little Rock
Little Rock, Arkansas 72204
email: jdberleant@ualr.edu

Karen Villaverde
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003
email: kvillave@cs.nmsu.edu

Olga M. Koseheleva
Department of Teacher Education
University of Texas at El Paso
El Paso, TX 79968
email: olgak@utep.edu

***Abstract*—In the traditional statistical approach, we assume that we know the exact cumulative distribution function (CDF) $F(x)$. In practice, we often only know the envelopes $[\underline{F}(x), \overline{F}(x)]$ bounding this CDF, i.e., we know the interval-valued "p-box" which contains $F(x)$. P-boxes have been successfully applied to many practical applications. In the p-box approach, we assume that the actual CDF can be any CDF $F(x) \in [\underline{F}(x), \overline{F}(x)]$. In many practical situations, however, we know that the actual distribution is smooth. In such situations, we may wish our model to further restrict the set of CDFs by requiring them to share smoothness (and similar) properties with the bounding envelopes $\underline{F}(x)$ and $\overline{F}(x)$. In previous work we used ideas from Info-Gap Decision Theory to propose heuristic methods for selecting such distributions. In this paper, we provide justifications for this heuristic approach.**

I. Traditional approach: a brief reminder

In the traditional statistical techniques typically used in science and engineering applications, we assume that we know the exact probability distributions of measurement errors, of different population quantities, etc. (e.g. [11], [12]).

For each quantity ξ, this distribution can be described, e.g., by its cumulative distribution function (CDF)

$$F(x) \stackrel{\text{def}}{=} \text{Prob}(\xi \le x).$$

Computationally, the CDF is often represented by its *quantiles*, i.e., by the values $x(\alpha)$ for which $F(x(\alpha)) = \alpha$ for some pre-selected values α: e.g., $\alpha = 0, 0.1, 0.2, \ldots, 1.0$.

In mathematical terms, the representation means, crudely speaking, that instead of discussing the original CDF function $F(x)$ directly, we discuss the inverse function $x(\alpha)$. (This is exactly true when the CDF is strictly monotonic.)

II. P-box approach

A. Main idea

In practice, we rarely know the exact values of the probabilities of different events. In particular, for real-life quantities, we rarely know the exact values $F(x)$ of the probability $\text{Prob}(\xi \le x)$. Instead, we have approximate knowledge of $F(x)$.

In some situations, we know the bounding envelopes $[\underline{F}(x), \overline{F}(x)]$ for the unknown CDF $F(x)$. In other situations, we have expert estimates for values $F(x)$ – which can be naturally described as fuzzy numbers.

Comment. It is well known that a fuzzy number can be represented as a nested family of its α-cuts (intervals), and that processing these fuzzy numbers can be reduced to processing the corresponding α-cuts. Moreover, this is usually how fuzzy numbers are processed (see e.g. [3], [5], [7], [8], [10]).

In view of this fact, in the following text we will concentrate on interval uncertainty.

B. Mathematical representation

For each dimension ξ, uncertainty about probability

$$F(x) = \text{Prob}(\xi \le x)$$

can be described by an interval

$$\mathbf{F}(x) = [\underline{F}(x), \overline{F}(x)]$$

that is guaranteed to contain the unknown actual value of $F(x)$.

The function $\mathbf{F}$ that maps each real number x into the interval $\mathbf{F}(x)$ is called a *probability box*, or, for short, *p-box* [4].

C. Computer representation

As we have mentioned, in the traditional statistical approach a probability distribution is usually represented in the computer by its quantiles. In the p-box case, the fact that we do not know the exact values of $F(x)$ means that we do not know the exact values of the quantiles either.

Instead of the actual value of a quantile $x(\alpha)$, we only know the bounds on the quantile. Namely, from the fact that

$$\underline{F}(x) \le F(x) \le \overline{F}(x),$$

we can conclude that

$$\underline{x}(\alpha) \le x(\alpha) \le \overline{x}(\alpha),$$

where $\underline{x}(\alpha)$ are the quantiles corresponding to $\overline{F}(x)$, and $\overline{x}(\alpha)$ are the quantiles corresponding to $\underline{F}(x)$.

As a result, in the computer, a p-box is represented by its *interval-valued quantiles*, i.e., by the intervals $[\underline{x}(\alpha), \overline{x}(\alpha)]$ which are guaranteed to contain the actual (unknown) values $x(\alpha)$. These interval-valued quantiles are given for some pre-selected values α: e.g., $\alpha = 0, 0.1, 0.2, \ldots, 1.0$ (see e.g. [4]).

III. Need for a more realistic representation of uncertainty

A. The meaning of p-boxes: reminder

In short, a p-box expresses the information that for every x, the actual (unknown) value $F(x)$ of the CDF is contained in the interval $[\underline{F}(x), \overline{F}(x)]$.

B. Limitations of a p-box interpretation

In many practical situations, the extreme-case bounds $\underline{F}(x)$ and $\overline{F}(x)$ correspond to smooth distributions such as Gaussian, uniform, etc. In such situations, it is often reasonable to expect that the actual distribution $F(x)$ is also smooth. However, in the p-box approach, the only limitation on $F(x)$ is that

$$F(x) \in [\underline{F}(x), \overline{F}(x)].$$

This limitation permits, in addition to smooth functions, very non-smooth – and thus (for many problems) unrealistic – CDF functions $F(x)$.

To take such situations into account, it is desirable to be able to limit ourselves to smooth bounds $\underline{F}(x)$ and $\overline{F}(x)$ and to distributions $F(x)$ which share the same smoothness characteristics as the bounds.

IV. An approach motivated by Info-Gap Decision Theory

A. Main idea

To solve this problem, in [2] we described a new approach which is motivated by Info-Gap Decision Theory (see e.g. [1]). Specifically, once we have two possible distributions $F_1(x)$ and $F_2(x)$ which can be described using a set of quantiles $x_1(\alpha)$ and $x_2(\alpha)$ for various values of α, we then assume that for every value $\beta \in [0, 1]$, the distribution corresponding to the quantiles

$$x(\alpha) = \beta \cdot x_1(\alpha) + (1 - \beta) \cdot x_2(\alpha)$$

is also possible. Once we fix $F_1(x)$ and $F_2(x)$, we get a 1-parameter class which is much sparser than the p-box of all distributions between $F_1(x)$ and $F_2(x)$.

Let us show that this idea indeed allows us to avoid non-smooth combinations of smooth distributions.

B. An approach motivated by Info-Gap Decision Theory avoids non-smooth distributions

In many practical situations, the uncertainty is in the values of the parameters: we know the shape of the distribution, but we do not know the exact values of the corresponding parameters. For example, we may know that the distribution is Gaussian (or uniform), but not the exact values of the corresponding parameters.

How can we describe this situation in precise terms? For example, let $F_0(x)$ be the CDF of the "standard" normal distribution, with 0 mean and standard deviation 1. Then, the CDF of a general normal distribution, with mean a and standard deviation σ, can be described as

$$F(x) = F_0\left(\frac{x - a}{\sigma}\right).$$

A similar expression describes a general uniform distribution, etc.

For such distributions $F(x)$, as one can easily check, the corresponding quantiles $x(\alpha)$ are linearly related to the quantiles $x_0(\alpha)$ of the just-mentioned standard normal distribution $F_0(x)$:

$$x(\alpha) = a + \sigma \cdot x_0(\alpha).$$

If we have two distributions with the same property, i.e., if we have

$$x_1(\alpha) = a_1 + \sigma_1 \cdot x_0(\alpha)$$

and

$$x_2(\alpha) = a_2 + \sigma_2 \cdot x_0(\alpha),$$

then their convex combination

$$x(\alpha) = \beta \cdot x_1(\alpha) + (1 - \beta) \cdot x_2(\alpha)$$

also has the same form:

$$x(\alpha) = a + \sigma \cdot x_0(\alpha),$$

with $a = \beta \cdot a_1 + (1 - \beta) \cdot a_2$ and $\sigma = \beta \cdot \sigma_1 + (1 - \beta) \cdot \sigma_2$.

Thus, when applied to two distribution of the same shape, the above procedure leads to the distribution of this same shape: a combination of Gaussian distributions is Gaussian, a combination of uniform distributions is uniform, etc.

C. Remaining open problem.

The above procedure seems to work well, but is too *ad hoc*, requiring more justification.

D. Objective of this paper

In this paper, we provide a justification for this procedure.

Specifically, we want to describe an operation $I(x_1, x_2)$ that for every α, given two values $x_1(\alpha)$ and $x_2(\alpha)$, returns a suitable intermediate value

$$x(\alpha) = I(x_1(\alpha), x_2(\alpha)).$$

We will call such an operation an *intermediate value operation*.

Comment. The main ideas behind our justification are based on the natural notions of symmetry. Similar ideas have been used, e.g., in [6], [9].

V. Relevant types of invariance

The three types of invariance described in this section provide background for the definitions given in the following section.

A. Scale invariance

The values x often come from measurements. In this case, if we change the unit of measurement (e.g. from centimeters to meters), numerical values will be multiplied by a constant $\lambda > 0$. It is natural to require that the result of the intermediate value operation not depend on the choice of unit.

How does replacing a unit change the intermediate value operation function $I(x_1, x_2)$? If we replace a unit by a one that is λ times smaller, then the quantity that was initially described by the value x_1 will be described by a new value $x_1' = \lambda \cdot x_1$, and the quantity that was initially described by the value x_2 will be described by a new value $x_2' = \lambda \cdot x_2$. When we combine these values by using the intermediate value operation I, we get the resulting value

$$x' = I(x_1', x_2') = I(\lambda \cdot x_1, \lambda \cdot x_2).$$

This is the expression of the combined quantile in the new units. In the old units, its expression is

$$x = \lambda^{-1} \cdot x' = \lambda^{-1} \cdot I(\lambda \cdot x_1, \lambda \cdot x_2).$$

We will denote the resulting "re-scaled" intermediate value operation

$$x_1, x_2 \to \lambda^{-1} \cdot I(\lambda \cdot x_1, \lambda \cdot x_2)$$

by $S_\lambda(I)$.

In these terms, the intermediate value operation I is scale invariant if and only if $S_\lambda(I) = I$ for all λ.

B. Reverse invariance

In addition to changing the units, there can also be changes in sign. For example, when measuring a spatial coordinate, we can change the direction and that will change the sign, or when measuring an electric charge, we usually follow the convention that an electron's charge is negative, but we can also consider electron charges as positive numbers. This possibility is equivalent to a re-scaling with $\lambda = -1$. Therefore we wish to consider not only positive values λ, but in fact arbitrary non-zero values λ.

C. Shift invariance

When measuring quantities like time or location, we can also change the starting point. In this case, a constant will be added to all numerical values: $x \to x + a$.

Then the quantity that was initially described by the value x_1 will be described by a new value $x_1' = x_1 + a$, and the quantity that was initially described by the value x_2 will be described by a new value $x_2' = x_2 + a$. When we combine these values by using the intermediate value operation I, we get the resulting value

$$x' = I(x_1', x_2') = I(x_1 + a, x_2 + a).$$

This is the expression of the combined quantile in the new units. In the old units, its expression is

$$x = x' - a = I(x_1 + a, x_2 + a) - a.$$

We will denote the resulting "shifted" intermediate value operation

$$x_1, x_2 \to I(x_1 + a, x_2 + a) - a$$

by $T_a(I)$.

It is natural to require that the intermediate value operation is invariant w.r.t. these symmetries as well, i.e., that $T_a(I) = I$ for all possible real values a.

VI. Invariance: definitions and the main result

Definition 1.

- *By an* intermediate value operation*, we mean a function $I : R^2 \to R$ from the set of pairs of real numbers into real numbers for which the value $I(x_1, x_2)$ is always located in between x_1 and x_2:*

  $$\min(x_1, x_2) \le I(x_1, x_2) \le \max(x_1, x_2).$$

  The set of all possible intermediate value operations will be denoted by A.
- *For every intermediate value operation I, and for every $\lambda \neq 0$, by a* re-scaled *intermediate value operation $S_\lambda(I)$, we mean an intermediate value operation*

  $$x_1, x_2 \to \lambda^{-1} \cdot I(\lambda \cdot x_1, \lambda \cdot x_2).$$

- *For every intermediate value operation I, and for every a, by a* shifted *intermediate value operation $T_a(I)$, we mean an operation*

  $$x_1, x_2 \to I(x_1 + a, x_2 + a) - a.$$

- *We say that the intermediate value operation I is* scale-invariant *if for all λ, we have $S_\lambda(I) = I$.*
- *We say that the intermediate value operation I is* shift-invariant *if for all a, we have $T_a(I) = I$.*

Proposition 1. *For an intermediate value operation I, the following two conditions are equivalent to each other:*

- *I is scale-invariant and shift-invariant;*
- *I is described by the expression*

  $$I(x_1, x_2) = \beta \cdot x_1 + (1 - \beta) \cdot x_2$$

  for some $\beta \in [0, 1]$.

Comment. As a consequence of Proposition 1, the naturalness of scale- and shift-invariance implies the naturalness of the equivalent intermediate value operation.

VII. Proof of Proposition 1

It is easy to check that for every $\beta \in [0, 1]$, the formula $I(x_1, x_2) = \beta \cdot x_1 + (1 - \beta) \cdot x_2$ indeed describes a scale- and shift-invariant intermediate value operation.

Let us therefore move on to showing that every scale- and shift-invariant intermediate value operation $I(x_1, x_2)$ has the above form. Indeed, let I be such an operation, and let us define $\beta \stackrel{\text{def}}{=} I(1, 0)$.

For arbitrary $x_1 \neq x_2$, we can apply shift-invariance with $a = -x_2$, and conclude that

$$I(x_1, x_2) = I(x_1 - x_2, 0) + x_2.$$

Now, scale-invariance with $\lambda = 1/(x_1 - x_2)$ implies that

$$I(x_1 - x_2, 0) = (x_1 - x_2) \cdot I(1, 0).$$

By definition of β, we conclude that

$$I(x_1 - x_2, 0) = (x_1 - x_2) \cdot \beta$$

and, because I was given as shift-invariant, that

$$I(x_1, x_2) = (x_1 - x_2) \cdot \beta + x_2.$$

One can easily see that this expression is exactly equal to $\beta \cdot x_1 + (1 - \beta) \cdot x_2$. So, we have proven that

$$I(x_1, x_2) = \beta \cdot x_1 + (1 - \beta) \cdot x_2$$

for all $x_1 \neq x_2$.

For $x_1 = x_2$, this equality follows from the fact that I is an intermediate value operation and thus, $T(x_1, x_1) = x_1$, just like the convex combination $\beta \cdot x_1 + (1 - \beta) \cdot x_2$ is equal to x_1. The proposition is proven.

VIII. Towards an alternative justification based on optimality

A. Main idea

Instead of requiring that the intermediate value operation be invariant, it is reasonable to look for *optimal* operations, i.e., operations which are the best in the sense of some optimality criterion.

B. What is an "optimality criterion"?

When we say that some *optimality criterion* is given, we mean that, given two different intermediate value operations, we can decide whether the first or the second one is better, or if these operations are equivalent w.r.t. the given criterion. In mathematical terms, this means that we have a *pre-ordering relation* $\preceq$ on the set of all possible intermediate value operations.

C. The need to enumerate optimal intermediate value operations

One way to approach the problem of choosing the "best" intermediate value operation function is to select *one* optimality criterion, and to find an intermediate value operation which is the best with respect to this criterion. The main drawback of this approach is that there can be different optimality criteria, and they can lead to different optimal solutions. It is, therefore, desirable not only to describe an intermediate value operation that is optimal relative to some criterion, but to describe *all* intermediate value operations that are optimal relative to any member of a set of natural criteria.

Comment. The word "natural" is used informally. We merely want to say that from the purely mathematical viewpoint, there can be weird ("unnatural") optimality criteria. We will only consider criteria that satisfy some requirements that we would, from a common sense viewpoint, consider reasonable and natural.

D. Examples of optimality criteria

Pre-ordering is the general formulation of optimization problems in general, not just of the problem of choosing an intermediate value operation. In general optimization theory, in which we are comparing arbitrary *alternatives* $a, b, \ldots$, from a given set A, the most frequent case of a pre-ordering is when a *numerical criterion* is used, i.e., when a function $J : A \to R$ is given for which $a \preceq b$ if and only if $J(a) \leq J(b)$.

Various natural numerical criteria can be proposed for choosing the intermediate value operations. For example, we could consider cases in which we are given the class of all distributions classified as possible for the given problem, and we have "weights" assigned to different distributions from this class, so that these weights add up to 1. In this case, for some pairs of distributions from this class – characterized by their quantiles $x_1(\alpha)$ and $x_2(\alpha)$ – the distribution corresponding to the quantiles $x(\alpha) = I(x_1(\alpha), x_2(\alpha))$ also belongs to the given class. For some other pairs of distributions $x_1(\alpha)$ and $x_2(\alpha)$, the distribution corresponding to the $x(\alpha)$ does not belong to the given class. We can then take, as $J(I)$, the "ratio" (total weight) of such pairs of distributions for which $x(\alpha)$ also belongs to the given class.

Many other criteria can be proposed. What should be done if there are several different alternatives that perform equally well? In this case, it makes sense to choose the alternative for which the computations are the fastest. This natural idea leads to an optimality criterion that is not describable by a single numerical optimality criterion $J(a)$: in this case, we need *two* functions: $J_1(a)$ describes the "ratio", $J_2(a)$ describes the computation time, and $a \preceq b$ if and only if either $J_1(a) < J_1(b)$, or $J_1(a) = J_1(b)$ and $J_2(a) \geq J_2(b)$.

We could further specify the described optimality criterion so that it disambiguates cases where both $J_1(a) = J_1(b)$ and $J_2(a) = J_2(b)$ with another function J_3, etc. However, as we have already mentioned, the goal of this paper is not to find a single intermediate value operation that is optimal relative to some criterion, but to describe *all* intermediate value operations that are optimal relative to any of a set of natural optimality criteria. In view of this goal, in the following, we will not specify the criterion, but rather describe a general class of natural optimality criteria.

So, let us formulate what "natural" means.

E. Which optimality criteria are natural?

We have already mentioned that the values x often come from measurements, and that for such values, changing the unit of measurement (e.g. from meters to centimeters) multiplies the measured values by a constant λ. It is natural to require the relative quality of two intermediate value operations not depend on the choice of units. In other words, we require that if I is better than I', then the "re-scaled" I (i.e., $S_\lambda(I)$) should be better than the "re-scaled" I' (i.e., $S_\lambda(I')$).

It is also natural to require the optimality criterion to be invariant w.r.t. shift transformations. In other words, if I is better than I', then $T_a(I)$ should be better than $T_a(I')$.

There is one more reasonable requirement for a criterion, based on the following idea. If the criterion does not select a single optimal intermediate value operation, i.e., if it considers more than one intermediate value operations equally good, then we can always use some other criterion to help select among them, thus designing a two-step criterion. If this new criterion still does not select a unique intermediate value operation, we can continue this process as many steps as necessary to get only one optimal intermediate value operation. Such a multi-step criterion can always be *final* in this sense.

IX. An optimization approach: definitions and the main result

Definition 2. *By an* optimality criterion*, we mean a* pre-ordering *(i.e., a transitive, reflexive relation) $\preceq$ on the set A.*

- *An optimality criterion $\preceq$ is called* scale-invariant *if for all I, I', and $\lambda \neq 0$, $I \preceq I'$ implies $S_\lambda(I) \preceq S_\lambda(I')$.*
- *An optimality criterion $\preceq$ is called* shift-invariant *if for all I, I', and a, $I \preceq I'$ implies $T_a(I) \preceq T_a(I')$.*
- *An optimality criterion $\preceq$ is called* final *if there exists one and only one intermediate value operation I that is preferable to all the others, i.e., for which $I' \preceq I$ for all $I' \neq I$.*

Proposition 2.

- *If an intermediate value operation I is optimal w.r.t. some scale-invariant, shift-invariant, and final optimality criterion, then for some $\beta \in [0,1]$, the operation I is described by a formula*

$$I(x_1, x_2) = \beta \cdot x_1 + (1-\beta) \cdot x_2.$$

- *For every $\beta \in [0,1]$, there exists a scale-invariant, shift-invariant, and final optimality criterion for which the only optimal intermediate value operation is the operation*

$$I(x_1, x_2) = \beta \cdot x_1 + (1-\beta) \cdot x_2.$$

Comment. In other words, if the optimality criterion satisfies the above-described natural properties, then the optimal intermediate value operation coincides with one of β-operations.

A. Proof of Proposition 2

1. To prove the first part of Proposition 2, we will show that the optimal intermediate value operation $I_{\rm opt}$ is scale-invariant and shift-invariant, i.e., that $S_\lambda(I_{\rm opt}) = T_a(I_{\rm opt}) = I_{\rm opt}$ for all $\lambda \neq 0$ and a. Then, the result will follow from Proposition 1.

Indeed, let X be either a scale or a shift transformation. Let us first determine the invertibility of these transformations. Indeed:

- if $X = S_\lambda$, then $X^{-1} = S_{1/\lambda}$;
- if $X = T_a$, then $X^{-1} = T_{-a}$.

Now, from the optimality of $I_{\rm opt}$, we conclude that for every $I' \in A$, $X^{-1}(I') \preceq I_{\rm opt}$. From the invariance of the optimality criterion provided as a given in the proposition statement, we next conclude that $I' \preceq X(I_{\rm opt})$. This is true for all $I' \in A$ and, therefore, the intermediate value operation $X(I_{\rm opt})$ is optimal. But since the criterion is final (as given in the proposition statement), there is only one optimal intermediate value operation; hence, $X(I_{\rm opt}) = I_{\rm opt}$. So, the optimal intermediate value operation is indeed invariant and hence, due to Proposition 1, it coincides with one of the β-expressions. The first part is proven.

2. Let us now prove the second part of Proposition 2. Let $\beta \in [0,1]$ be fixed, and let I_β be the corresponding intermediate value operation. We will then define the optimality criterion as follows: $I \preceq I'$ if and only if $I' = I_\beta$.

Since the intermediate value operation I_β is scale-invariant and shift-invariant, the just-defined optimality criterion is also scale-invariant and shift-invariant. It is also by definition final. The intermediate value operation I_β is clearly optimal w.r.t. this scale-invariant, shift-invariant, and final optimality criterion. The proposition is proven.

Acknowledgments

This work was supported in part by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006-08 and by the Max Planck Institut für Mathematik.

The authors are thankful to Yakov Ben-Haim, Scott Ferson, Vladik Kreinovich, and to all the participants of the NSF Workshop on Reliable Engineering Computing at Georgia Tech (February 2008) for valuable discussions, and to the anonymous referees for valuable suggestions.

References

[1] Y. Ben-Haim, *Info-Gap Decision Theory: decisions under severe uncertainty*, Academic Press, New York, 2006.

[2] D. Berleant, L. Andrieu, J.-P. Argaud, F. Barjon, M.-P. Cheong, M. Dancre, G. Sheble, and C.-C. Teoh, "Portfolio management under epistemic uncertainty using stochastic dominance and Information-Gap Theory", *International Journal of Approximate Reasoning*, to appear.

[3] D. Dubois and H. Prade, Operations on fuzzy numbers, *International Journal of Systems Science*, 1978, Vol. 9, pp. 613–626.

[4] S. Ferson. *RAMAS Risk Calc 4.0*, CRC Press, Boca Raton, Florida, 2002.

[5] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, Prentice Hall, New Jersey, 1995.

[6] V. Kreinovich, G. Mayer, and S. Starks, "On a Theoretical Justification of The Choice of Epsilon-Inflation in PASCAL-XSC", *Reliable Computing*, 1997, Vol. 3, No. 4, pp. 437–452.

[7] R. E. Moore and W. Lodwick, "Interval Analysis and Fuzzy Set Theory", *Fuzzy Sets and Systems*, 2003, Vol. 135, No. 1, pp. 5–9.

[8] H. T. Nguyen and V. Kreinovich, Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications, In: R. B. Kearfott and V. Kreinovich, eds., *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[9] H. T. Nguyen, V. Kreinovich, and Q. Zuo, "Interval-valued degrees of belief: applications of interval computations to expert systems and intelligent control", *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS)*, 1997, Vol. 5, No. 3, pp. 317–358.

[10] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

[11] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[12] H. M. Wadsworth, (ed.), *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York, 1990.