Constructing Belief Functions Using the Principle of Minimum Uncertainty

1st Yanyan He
Dept. of Math, and of CSE
University of North Texas
Denton, USA
yanyan.he@unt.edu

2nd M. Yousuff Hussaini
Dept. of Math
Florida State University
Tallahassee, USA
yousuff@fsu.edu

***Abstract*—This paper presents the MinUnc method to construct m-/belief functions in the framework of the Dempster-Shafer theory to represent the uncertainty in a given body of evidence. Using the principle of minimum uncertainty and the concepts of entropy and non-specificity, the MinUnc method specifies a partition of a finite interval on the real line and assigns belief masses to the uniform subintervals. The proposed MinUnc method is illustrated using a simple example and applied to uncertainty representation of air flight arrival delay data set.**

***Index Terms*—Dempster-Shafer theory, Basic belief assignments, Principle of minimum uncertainty, Entropy, Non-Specificity, Predictive belief function**

I. Introduction

Uncertainty quantification has been intensively applied to various real-world applications to study the uncertainty in the underlying physical system, and consequently provide more accurate and reliable behavior predictions [1], [2]. The mathematical theories for uncertainty quantification include probability theory [3], interval analysis [4], Dempster-Shafer (DS) theory [5] and etc. If the uncertainty is due to the inherent randomness in the system, traditional probability theory approach is most suitable and random variables with probability density functions (base on sufficient amount of information) can be used to represent the uncertainty [6]. If the system is not stochastic, it is quite possible to describe the uncertainty in systems using uncertain variables with upper and lower bounds based on the limited available information, and the concepts of interval analysis can be applied. If the information associated with the uncertain variable is insufficient for the construction of a complete probabilistic framework, but richer than an interval can characterize, the m-/belief function in the DS theory can be utilized for uncertainty quantification.

The m-functions are normally constructed from experts' opinions or statistical data [7]. For example, Yaghlane et al. have solved optimization problem to construct the least informative belief functions from elicited expert opinions expressed in terms of qualitative preference relations [8]. Aguirre et al. have applied statistical inference methods to construct belief functions of reliability parameters of components from statistical data about reliability [9]. However, the research on the development of a general and rigorous algorithm to construct belief function for a quantity of interest Y based on a set of finite data points (i.e., the collection of all possible values of Y) is limited. Denœux has proposed to construct predictive belief function for a random variable based on its finite number of samples [25] [26]. However, the focal elements of the belief function need to be specified in advance. In this work, we propose a general algorithm to construct a belief function for an uncertain variable Y based on a finite set of data. The proposed algorithm does not assume randomness of Y, and will provide both focal elements and their belief masses. It is especially applicable to uncertainty quantification in simulation output given observed data.

We recall here the basic notions of DS theory. Let the universal set or the frame of discernment be a collection of possible values $\mathbf{X} = \{x_1, x_2, ..., x_n\}$ for the quantity of interest X. Let any subset $A \subseteq \mathbf{X}$ represent the proposition "X belongs to A." In DS theory, there are two important measures: belief (Bel) and plausibility (Pl). Different interpretations of these measures are available in the literature [10]–[16]. In this work, we adopt the interpretation due to Shafer [5]: the belief function $Bel(A)$ measures the strength of evidence supporting the proposition A while the plausibility function $Pl(A)$ quantifies the maximum possible support from the evidence to the proposition A. Belief and plausibility function can be derived from each other using $Pl(A) = 1 - Bel(\overline{A})$ (where $\overline{A}$ is the complement of A). In addition, belief and plausibility functions can also be derived from basic belief assignment (BBA) or called m-function, which satisfies $m(\emptyset) = 0$ and $\sum_{A \subseteq \mathbf{X}} m(A) = 1$. Subset A is called a focal element if $m(A) \neq 0$. In this work, we assume the number of focal elements is finite.

Quantifying the uncertainty in a body of evidence (e.g., a finite data set in the current work) in the framework of DS theory amounts to constructing m-/belief functions assigning belief masses to any proposition A. To this end, we introduce important concepts related to the current work in section II and propose the MinUnc method and its algorithm in details in section III. In section IV, the proposed MinUnc method is illustrated using a simple example and the obtained belief function is compared to Denœux's predictive belief function. The MinUnc method is applied for uncertainty representation of air flight arrival time delay data set in section V.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

II. Background

In this section, we introduce a few important concepts in the framework of DS theory.

A. Concave m-Function

When the quantity of interest takes real-number values, the m-function is constructed on the real line. In this work, we assume the m-function has finite number of intervals as the focal elements. For the purpose of decision making, it is natural and convenient to further assume that the m-function representing one body of evidence has one and only one maximum. We call this type of m-function a concave m-function and define it as follows.

Definition 1: Let $\mathcal{P} = \{A_1, A_2, ..., A_l\}$ be a uniform partition of a finite interval $\mathbf{Y}$, i.e., $\mathbf{Y} = \cup_i A_i$, $A_i \cap A_j = \emptyset$ for $1 \leq i \neq j \leq l$, and the subintervals A_i $(1 < i < l)$ have the same length Δ. Then an m-function defined over the universal set $\mathbf{Y}$ is a concave m-function if:

1) $Pl(\{y\}) = m(A_i)+m(\mathbf{Y})$ for all $1 \leq i \leq l$ and $y \in A_i$.
2) $\exists\, y' \in \mathbf{Y}$ such that $Pl(\{y\}) \geq Pl(\{y + \delta\})$ for $y \geq y'$ and $Pl(\{y\}) \geq Pl(\{y - \delta\})$ for $y \leq y'$, where $\delta > 0$.

The requirement (a) states that the consecutive disjoint intervals A_is $(1 \leq i \leq l)$ and the universal set $\mathbf{Y}$ are possible focal elements; and (b) states that the plausibility function $Pl(y)$ is a concave (or concave down) function. Figure 1(a) shows one example of a concave m-function with intervals A_is and real numbers $m(A_i)$s as focal elements and the corresponding belief masses, respectively. The focal elements have the same length and A_2 is assigned with the most degree of belief mass. For the purpose of simplification, we visualize the m-function as in Fig. 1(b) where each stair (interval) represents a focal element and its height represents the corresponding belief mass.

A concave m-function has the following characteristics: i) the subinterval length Δ, which indicates how specific is the support of evidence for a given proposition, the smaller the length is, the more specific the evidence is; ii) the total range of the m-function, i.e., the core of the m-function (the union of all the focal elements of the m-function), the smaller the total range is, the more focused is the contribution of the uncertainty source; iii) the maximum mass m_{max}, which indicates the degree of belief that favors a certain subinterval over all others, a larger $m_{max}(A)$ shows more confidence in the proposition that the true value of Y is in A; and iv) the focal element A_{max} with the maximum mass, which is the interval the evidence (or data) favors most.

B. Cumulative Belief Function and Cumulative Plausibility Function

The cumulative belief function (CBF) and cumulative plausibility function (CPF) has been defined by Oberkampf et al. as follows [17]:

$$CBF(y) = Bel(Y \leq y), \qquad CPF(y) = Pl(Y \leq y).$$

Note: Independently, Yager [18] proposed identical concepts called "belief cumulative distribution functions."

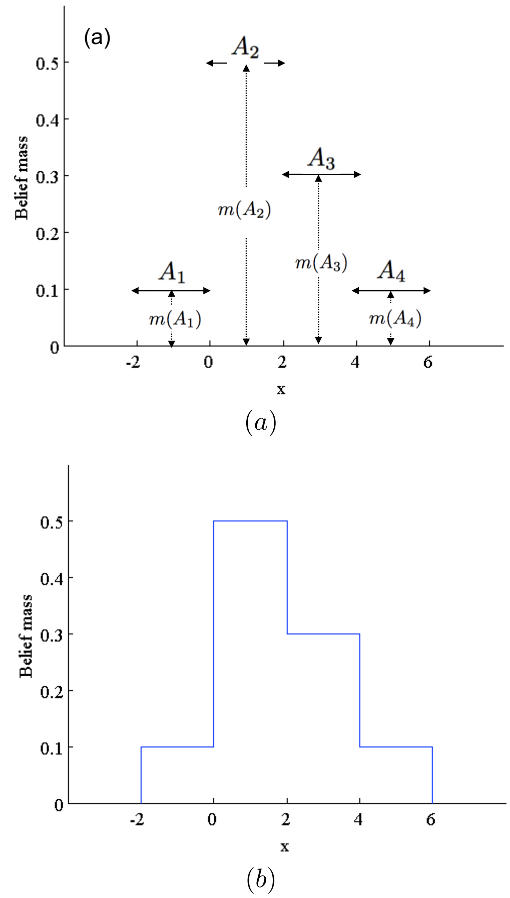

Fig. 1. Concave m-function: (a) focal elements and belief masses; and (b) simplified visualization.

C. Entropy and Specificity

Different formulations have been proposed to measure the uncertainty in a data set [7], [19], [20]. They are mainly classified into three categories: measures of imprecision (or non-specificity), measures of conflict (or dissonance, or entropy-like measures), and measures of total uncertainty. In the current work, we adopt the formulas from Klir and Wierman's book for non-specificity ([20] p55, Eq. (3.51), [21]) and entropy-like measure ([20] p81, Eq. (3.187), [19]).

Definition 2: If Bel is a belief function over the power set $2^{\mathbf{Y}}$ with the corresponding basic belief assignment, m, and plausibility function, Pl, then the entropy-like measure (denoted as Em) and non-specificity (denoted as Nm) of Bel are defined as :

$$Em = -\sum_{A \subset \mathbf{Y}} m(A) log_2(Pl(A)), \tag{1}$$

$$Nm = \sum_{A \subset \mathbf{Y}, A \neq \emptyset} m(A) log_2 L_A, \tag{2}$$

where L_A is the number of elements in the set A, i.e., $L_A = |A|$. If A is a finite interval, L_A is the interval length.

The entropy-like measures (we use entropy directly hereafter for simplicity) indicate the dissonance in the evidence. The

Bayesian structure, which is of a highly dissonant type, has a high entropy value. Specificity relates to the degree to which the evidence is pointing to a specific element. That the plausibility structure is closer to the belief structure indicates that the evidence is more specific. The Bayesian structure is most specific (with minimal non-specificity).

III. Constructing Belief Functions Using the Principle of Minimum Uncertainty

In this section, we propose a novel method (called MinUnc method) specifically designed for constructing m-functions on the real line from a finite data set. The constructed m-function, which serves as a mathematical representation of a data set, can be further used for different application purposes, such as uncertainty propagation through a simulation model, and prediction of occurrence of future events. Before describing the MinUnc method, we discuss the principles of uncertainty, which must be appropriately satisfied while partial knowledge is encoded in an m-function.

A. Principles of Uncertainty

The principle of maximum uncertainty ensures "using all information available but making sure that no additional information is unwittingly added" [22]. Essentially, it is the same idea as the principle of maximum entropy in information science [22] or the least commitment principle. The principle of maximum uncertainty suggests that in order to set up a probability distribution that honestly represents a state of incomplete knowledge, one has to maximize the uncertainty subject to all the information one has. It is applicable when PDFs are constructed to represent the partial knowledge. The principle of maximum uncertainty guarantees recognition of the ignorance and the PDF that satisfies this principle is maximally noncommittal with regard to the assumed information, which is not contained in the partial knowledge. On the other hand, the principle of minimum uncertainty is used when "some of the initial information is inevitably reduced in the solutions to various degrees" [22]. It can be expressed as "using as much information we have as possible, i.e., making the reduction of information in solutions as small as possible." When a belief function is constructed on the real line to represent the partial knowledge (a finite data set), no additional information is added, but some information may inevitably be lost. The principle of minimum uncertainty requires that this loss is minimized; thus we should obtain the belief function with the minimum increase of uncertainty (i.e., with minimum uncertainty subject to the available information).

Lower entropy and lower non-specificity indicate less uncertainty with which the evidence supports a unique outcome [19]. Following the principle of minimum uncertainty, we propose the MinUnc method, which minimizes both non-specificity (equivalent to maximizing the specificity) and entropy, to construct an m-function for a given finite data set. The objective function is

$$J(m) = \beta Em(m) + Nm(m),\ \beta > 0, \tag{3}$$

where Em and Nm are the defined entropy and non-specificity, respectively, and the positive parameter β controls the trade off between entropy and specificity.

B. Constraints on the m-Function

For a finite set with N elements $S = \{y_1, y_2, ..., y_N\}$, a single interval $\mathbf{Y} = [\min\{S\}, \max\{S\}] = [\min_i\{y_i\}, \max_i\{y_i\}]$ on the real line can be always specified to include all the possible values of Y currently available. One can always construct an m-function with the interval $\mathbf{Y}$ as the only focal element, where $m(\mathbf{Y}) = 1$. This leads to a vacuous belief function (see Fig. 2) $Bel(A) = 0$ for all $A \subset \mathbf{Y}$, $A \neq \mathbf{Y}$ and $Bel(\mathbf{Y}) = 1$, which is the least informative.

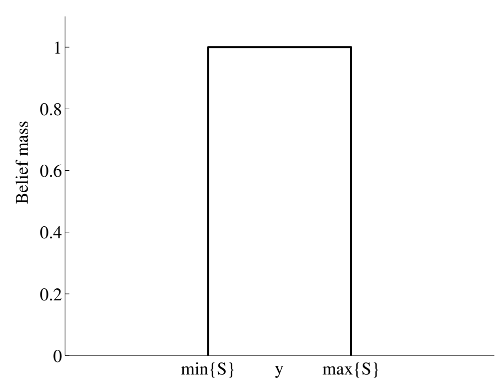

Fig. 2. Vacuous belief function

On the other hand, the evidence may favor a specific subset of $\mathbf{Y}$. Therefore it is reasonable to construct a belief function carried by a partition $\mathcal{P}$ of the interval $\mathbf{Y}$, i.e., belief masses are assigned to the disjoint subintervals $A_1, A_2, ..., A_l$ where $\cup_{i=1}^{l} A_i = \mathbf{Y}$. Here, for practical purposes, $\mathbf{Y}$ is divided into uniform subintervals with interval length Δ and with $Y = 0$ being a boundary of one of the subintervals. The left-most and right-most subintervals may have lengths different from Δ. Each y_j will support one of the subintervals $A_1, A_2, ..., A_l$.

The m-function is defined as $m(A_i) = n_i/N$ $(1 \leq i \leq l)$, where n_i is the number of data points falling inside the subinterval A_i and N is the cardinality of the data set S (i.e., $|S|$). If N is small, or if Δ is small, the distribution of data points may be scattered: unsupported subintervals may alternate with supported ones producing more than one local maximum. A scattered distribution yields little, if any, useful information. In the current work, we construct concave m-functions, i.e., with one subinterval with the maximum evidence support and with the evidence to support subintervals on both sides of this subinterval monotonically decreasing.

Due to experimental or measurement errors, there may exist outliers in a data set that can cause serious problems in statistical analyses. Therefore we identify and exclude the outliers from the considered data set before constructing m-function. Various ways are available in the literature to detect the outliers [23], [24]. We adopt the commonly used Tukey's rule in the current work. Let Q_1 and Q_3 be the lower and upper quartiles of the data set, John Tukey proposed to consider any data point outside the range

$$[Q_1 - 1.5(Q_3 - Q_1), Q_3 + 1.5(Q_3 - Q_1)] \tag{4}$$

as an outlier. It is worth mentioning that the partial belief supporting the outliers will be reassigned to the universal set, therefore the consideration of part of the data set as outliers does not deleteriously influence the results but just adds to the uncertainty.

C. The MinUnc Method

Since the basic belief assignment, m, is a function of Δ, the objective function Eq. (3) becomes

$$J(m(\Delta)) = \beta Em(m(\Delta)) + Nm(m(\Delta)), \tag{5}$$

and the MinUnc method minimizes this objective function with respect to Δ

$$\Delta_{opt} = \underset{0<\Delta=k\Delta_{ini}\leq\Delta_{max}}{\arg\min} \{\beta Em(\Delta) + Nm(\Delta)\}, \tag{6}$$

where $\Delta_{ini} \in \{1, 0.1, 0.01, 0.001, 0.0001, ...\}$ is the minimum desirable length of the subintervals (which can be considered as the tolerance for the accuracy of the prediction results) to be specified in advance. Smaller Δ_{ini} indicates less tolerance for the accuracy, and it requires more computational time to solve the optimization problem. Since Δ_{ini} is considered as the desired precision and smaller scale is not necessary, we constraint the search space of the optimization to be multiples of Δ_{ini} (i.e., $\Delta = k\Delta_{ini}$) for the simplicity of the optimization. The parameter $\beta > 0$ is introduced to control the trade-off between the smaller entropy and larger specificity (smaller non-specificity). The smaller the value of β is, the constructed m-function intend to be more dissonant but more specific as well. Without any further requirement for the emphasis on entropy or specificity from specific application problems, we consider both are equally important and chose $\beta = \min\{0.1, N\Delta_{ini}\}$ so that the magnitudes of entropy and non-specificity are at similar level. The parameter Δ_{max} is chosen so that the concave m-function with the smallest subinterval length is obtained as the optimal solution.

The procedure of the MinUnc method is defined as follows:

1) Specify the initial values of the parameters Δ_{ini}.
2) Remove the outliers from the data set.
   Specifically, calculate the quantiles Q_1 and Q_3 of the data set S, let $\hat{S}$ be the set containing the data points in the interval $[Q_1 - 1.5(Q_3 - Q_1), Q_3 + 1.5(Q_3 - Q_1)]$, and let $\mathbf{Y} = [\min\{\hat{S}\}, \max\{\hat{S}\}]$ and set $j = 1$.
3) Solve the optimization for subinterval length.
   Specifically, set $\Delta_{max} = j \times \Delta_{ini}$ and solve the optimization problem (6) for Δ_{opt}. (Note: The values of the objective function (5) are calculated for all $\Delta_{ini} \leq \Delta \leq \Delta_{max}$ and the chosen Δ_{opt} corresponds to the minimum value.)
4) Obtain the focal elements and their belief masses.
   With the optimal subinterval length Δ_{opt} and $Y = 0$ being a boundary of one of the subintervals, one can obtain a unique partition $\mathcal{P}$ of $\mathbf{Y}$, i.e., the consecutive disjoint subintervals A_i $(1 < i < l_j)$ with length Δ_{opt} (with $A_1 = [\min\{\hat{S}\}, \min\{A_2\}]$ and $A_{l_j} = [\max\{A_{l_j-1}\}, \max\{\hat{S}\}]$). The belief masses of the m-function are calculated as
   $$m(A_i) = n_i/N, \qquad m(\mathbf{Y}) = (N - |\hat{S}|)/N. \tag{7}$$
5) Stop if a concave m-function is obtained; otherwise $j = j + 1$, go to Step 3.

The flowchart of the procedure to construct m-function using MinUnc method is provided in Fig. 3.

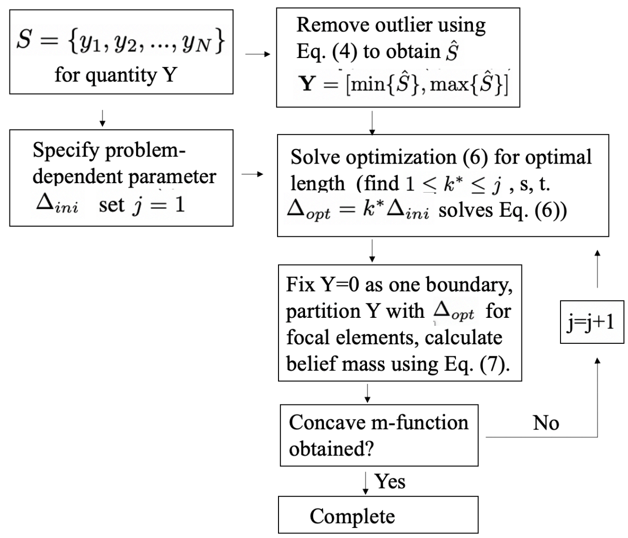

Fig. 3. Flowchart for the MinUnc method.

IV. Illustration of the MinUnc Method on a Simple Example

We consider an example data set $S = \{y_j\}_{j=1}^N$ (with size $N = 1000$) constituting samples from a Gaussian distribution $\mathcal{N}(0, 1)$ as the available information about the uncertain variable Y. We construct an m-function on the data set using the MinUnc method, and analyze its characteristics as a function of the parameter Δ_{ini}.

A. Construction of the m-Functions

The m-function with $\Delta_{ini} = 0.00001$ is constructed for the data set with $N = 1000$ (Fig. 4(a)). It shows that the quantity of interest Y falls inside the interval $A_{\max} = [-0.34837, 0]$ with the maximum degree of belief $Bel(A_{\max}) = m(A_{\max}) = 0.123$. To consider the possibility of Y less than a fixed value, for example $Y < 0.5$, the CBF and CPF are constructed (Fig. 4(b)). It shows that the possibility of the proposition $Y \in (\infty, 0.5)$ being true is bounded by $CBF(0.5) = 0.63$ and $CPF(0.5) = 0.752$.

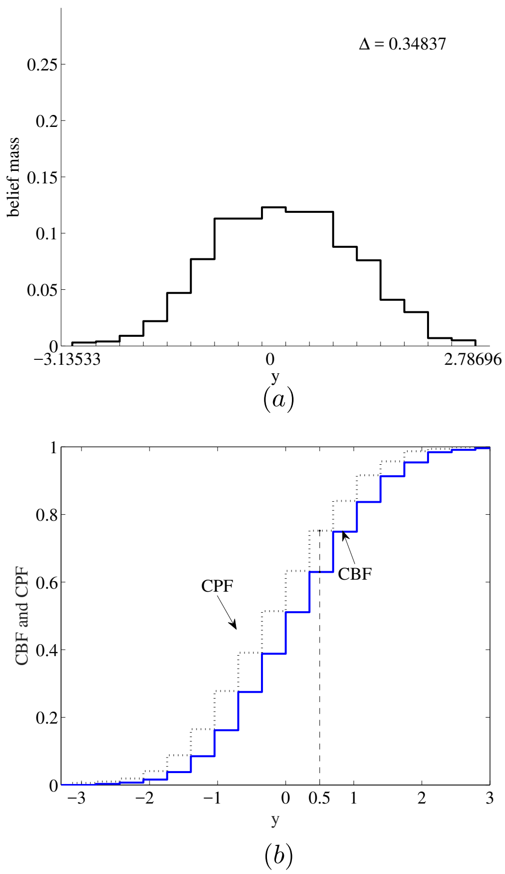

Fig. 4. Data set S with $N = 1000$, $\Delta_{ini} = 0.00001$: (a) BBA, (b) CBF and CPF.

B. The Characteristics of the m-Functions with respect to Initial Subinterval Length

The characteristics, entropy and non-specificity of the m-functions constructed from the example data set S ($|S| = 1000$) are shown in Table I in detail. From the results, one can observe that as Δ_{ini} decreases (i.e., $-log_{10}\Delta_{ini}$ increases), the optimal subinterval size Δ and the non-specificity are decreasing. The obtained m-function is more accurate and precise, which matches our intuition since smaller Δ_{ini} represents finer scale and higher precision.

TABLE I
THE CHARACTERISTICS OF THE m-FUNCTIONS WITH DIFFERENT Δ_{ini}

Δ_{ini}	Δ	A_{max}	m_{max}	Em	Nm
0.1	0.4	$[0, 0.4]$	0.144	3.3217	-1.3154
0.01	0.36	$[-0.36, 0]$	0.126	3.4648	-1.4697
0.001	0.349	$[-0.349, 0]$	0.123	3.5046	-1.5121
0.0001	0.3484	$[-0.3484, 0]$	0.123	3.5056	-1.5144
0.00001	0.34837	$[-0.34837, 0]$	0.123	3.5056	-1.5145

C. Comparison of the MinUnc Method and Denœux's Method

In this section, the belief function obtained from the MinUnc method is compared to the predictive belief function proposed by Denœux [25] [26], which is suitable for a special case where the uncertain quantity of interest Y is a random variable with unknown probability P_Y, and a finite collection of samples of Y (i.e., $\mathbf{Y}_N = \{y_j\}_{j=1}^N$) is available as the data set.

The Predictive Belief Function. Suppose $\mathcal{F}$ is the collection of the focal elements: $A_k (1 \le k \le l)$ and the unions of A_k's, where A_k's are ordered consecutive intervals, and "ordered" means that the elements in A_{k1} are no larger than the elements in A_{k2} if $k1 < k2$. Let n_k be the number of samples falling inside A_k and

$$P_k^- = \frac{a + 2n_k - \sqrt{D_k}}{2(N+a)}, \qquad P_k^+ = \frac{a + 2n_k + \sqrt{D_k}}{2(N+a)},$$

where a is the quantile of order $1-\alpha$ of the chi-square distribution with one degree of freedom (for the degree of freedom $l > 2$, we take a as the quantile of order $1-\alpha/l$ of the chi-square distribution with one degree of freedom suggested by Goodman [25]) and

$$D_k = a(a + \frac{4n_k(N-n_k)}{N}). \tag{8}$$

Let $A_{k,j}$ denote the union $A_k \cup A_{k+1} \cup ... \cup A_j$, then we have the predictive belief function $Bel(A_{k,j}) = P^-(A_{k,j})$ and

$$\begin{aligned}
m(A_{k,j}) &= P_k^-, \text{ if } k = j,\\
m(A_{k,j}) &= P^-(A_{k,j}) - P^-(A_{k+1,j}) - P^-(A_{k,j-1}),\\
&\quad \text{if } j = k+1,\\
m(A_{k,j}) &= P^-(A_{k,j}) - P^-(A_{k+1,j}) - P^-(A_{k,j-1})\\
&\quad +P^-(A_{k+1,j-1}), \text{ if } j > k+1.
\end{aligned}$$

The predictive belief function satisfies:

1) $\forall A \in \mathcal{F}$, $Bel[\mathbf{Y}_N](A) \xrightarrow{P} P_Y(A)$, i.e., for any $\epsilon > 0$, $P(|Bel[\mathbf{Y}_N](A) - P_Y(A)| < \epsilon) \to 1$ as $N \to \infty$.
Remark: The belief function constructed with the MinUnc method also satisfies the property: $Bel(A) \to P_Y(A)$ as $n \to \infty$ since $Bel(A_k) = n_k/N$ approaches $P_Y(A_k)$ as N goes to infinity.
2) $P(Bel[\mathbf{Y}_N](A) \le P_Y(A), \forall A \in \mathcal{F}) = 1 - \alpha$.
Remark: Denœux mentioned "since we have less information than in the asymptotic case (i.e., $N \to \infty$), it seems natural to impose that $Bel[Y_n](A)$ be less committed than $P_Y(A)$ as a consequence of the Least Commitment Principle (LCP)," i.e., $Bel[Y_N](A) \le P_Y(A)$. (Note: According to Smets [27], the LCP formalizes the idea "one should never give more support than justified to any subset of the universal set.") In the situation that the extra information in the asymptotic case is conflicting with the data set S, this requirement 2 will enlarge the belief mass assigned to unions of focal elements and consequently increase the uncertainty. The proposed MinUnc method does not require the existence of $P_Y(A)$ and consequently the condition 2 .

Results of the Example. The MinUnc method specifies the partition of the universal set, i.e., the disjoint subintervals as the focal elements. Denœux's predictive belief function (with $\alpha = 0.025$) is derived with the same subintervals and the unions of these subintervals as the focal elements. Since Denœux's method assigns partial belief to both disjoint subintervals and the unions of the subintervals, it is difficult to compare belief functions directly (Fig. 5(a) shows only the belief masses of the disjoint intervals). The cumulative belief and plausibility functions from both the MinMax and Denœux's methods are constructed and compared.

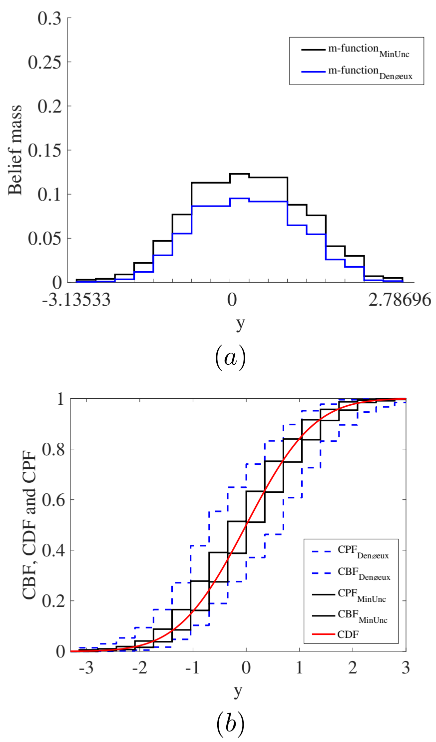

Fig. 5. The comparison of the MinMax method and Denœux's method for the data set S with $N = 1000$: (a) the BBAs, (b) CDF, CBFs and CPFs.

Figure 5(b) shows CBF/CPF from the MinMax method, CBF/CPF from the predictive belief function and the true cumulative distribution function (CDF). One can observe that the predictive belief function introduces more uncertainty (less specific due to the condition 2).

V. Mathematical Representation of Air Flight Arrival Delay Data

In this section, we implement the proposed approach to construct m-function for uncertainty representation in air flight arrival delay data set. Specifically, we focus on the flight NK1704 from an American ultra-low-cost carrier Spirit Airlines. The flight NK1704 is scheduled to departure from CLT (Charlotte, NC) at 11:30AM and arrive at BWI (Baltimore, MD) at 12:50PM. The data set contains the departure delay time and arrival delay time from 9/1/2019 to 9/30/2019, which is obtained from the public database of the United States Department of Transportation website. Since the departure and arrival delay time in the data set is recorded as integers, and also there is no need to consider the time delay in a more precise scale (such as 0.1, 0.01 minutes) in daily life, therefore, it is natural to set $\Delta_{ini} = 1$ for this specific

TABLE II
Departure and Arrival Time Delay (in Minutes) for NK1704 from 9/1/2019 to 9/30/2019

Date	9/1	9/2	9/3	9/4	9/5	9/6	9/7	9/8	9/9	9/10
Dep	19	20	4	0	0	0	0	0	0	0
Arr	32	25	30	0	12	0	11	0	19	9
Date	11	12	13	14	15	16	17	18	19	20
Dep	0	0	7	17	0	0	0	0	0	1
Arr	0	6	12	38	0	8	0	1	22	13
Date	21	22	23	24	25	26	27	28	29	30
Dep	0	0	0	0	0	0	0	0	0	12
Arr	9	0	1	15	7	15	4	12	5	51

problem. Using the proposed approach with $\beta = 0.1$, m-function is constructed to represent the uncertainty in the arrival time delay data set. From Fig. 6, one can conclude that the arrival time delay for NK1704 in September of 2019 is less than 7 minutes with the highest degree of belief 0.4 (i.e., $Bel([0, 7]) = m([0, 7]) = 0.4$). We also analyze the departure time delay data set to check the causality between departure time delay and arrival time delay. Excluding the outliers with Turkey's rule, the departure time delay for NK1704 in September of 2019 is zero with degree of belief 1. Therefore, the conclusion that the arrival delay for NK1704 in September of 2019 is not caused by the departure delay can be drawn and other factors (after departure) needs to be explored.

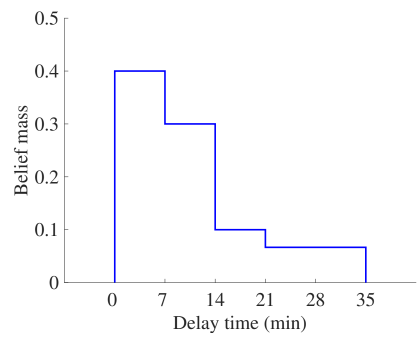

Fig. 6. m-function of arrival time delay for NK1704 in September of 2019.

VI. Summary and Conclusion

A novel algorithm called MinUnc method is proposed to construct the m-/belief function for uncertainty representation/modeling in a finite data set. Before the construction, the outliers is first detected and excluded from the data set using Tukey's rule. Then with the principle of minimum uncertainty and the concepts of entropy and non-specificity, the MinUnc method specifies a partition of a finite interval on the real line and assigns belief masses to the uniform subintervals.

The proposed MinUnc method is suitable for uncertainty representation in any systems/problems involving an uncertain quantity Y associated with a finite data set. The obtained belief function can help to make conclusions about the uncertain quantity of our interest Y. Moreover, it can help to numerically study the uncertainty in function outputs propagated from the uncertainty in input Y, which is of critical importance to simulations.

Although MinUnc method has significant practical advantage, there are limitations as well. For example, if the data favors two or more non-adjacent subintervals mostly with similar degrees of belief, the constructed belief function intends to have large subinterval size and consequently becomes less informative. In addition, enforcing zero to be boundary of one focal element (subinterval) may not match the reality and consequently may also enlarge the subinterval size.

In future, we would like to relax the constraint of zero being boundary of one focal element and let the optimization process to choose the optimal boundaries (instead of the optimal length in our current algorithm) of the focal elements. With more degrees of freedom in the optimization process, the constructed belief function will represent the data better.

References

[1] C. J. Roy and W. L. Oberkampf, "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 25, pp. 2131 – 2144, 2011.

[2] Y. He, M. Razi, C. Forestiere, L. D. Negro, and R. M. Kirby, "Uncertainty quantification guided robust design for nanoparticles' morphology," *Comput. Methods Appl. Mech. Engrg.*, vol. 336, pp. 578–593, 2018.

[3] A. N. Kolmogorov, *Foundations of the Theory of Probability*. Martino Fine Books, 2013.

[4] G. Alefeld and G. Mayer, "Interval analysis: theory and applications," *Journal of Computational and Applied Mathematics*, vol. 121, no. 1, pp. 421 – 464, 2000.

[5] G. Shafer, *A Mathematical Theory of Evidence*. N.J.,US: Princeton University Press, 1976.

[6] Y. He, M. Mirzargar, and R. M. Kirby, "Mixed aleatory and epistemic uncertainty quantification using fuzzy set theory," *International Journal of Approximate Reasoning*, vol. 66, pp. 1–15, 2015.

[7] T. Denœux, "Methods for building belief functions," in *Fourth School on Belief Functions and their Applications*, Xi'an, China, July 2017.

[8] A. B. Yaghlane, T. Denœux, and K. Mellouli, "Elicitation of expert opinions for constructing belief functions," *Uncertainty and Intelligent information Systems*, pp. 75–89, 2008.

[9] F. Aguirre, M. Sallak, and W. Schön, "Construction of belief functions from statistical data about reliability under epistemic uncertainty," *IEEE Transactions on Reliability*, vol. 62, no. 3, pp. 555–568, 2013.

[10] J. Halpern and R. Fagin, "Two views of belief: Belief as generalized probability and belief as evidence," *Artificial Intelligence*, vol. 54, pp. 275–318, 1992.

[11] J. Kohlas and P. Monney, *A Mathematical Theory of Hints: An Approach to Dempster-Shafer Theory of Evidence*. Berlin: Springer-Verlag, 1995.

[12] D. Dubois and H. Prade, "Random sets and fuzzy interval analysis," *Fuzzy Sets and Systems*, vol. 42, pp. 87–101, 1991.

[13] T. Denaeux, "Extending stochastic ordering to belief functions on the real line," *Information Sciences*, vol. 179, pp. 1362–1376, 2009.

[14] F. Tonon, A. Bernardini, and A. Mammino, "Reliability analysis of rock mass response by means of random set theory," *Reliability Engineering and System Safety*, vol. 42, pp. 263–282, 1991.

[15] I. R. Goodman and H. T. Nguyen, *Uncertainty Models for Knowledge-Based Systems; A Unified Approach to the Measurement of Uncertainty*. New York, NY, USA: Elsevier Science Inc., 1985.

[16] H. T. Nguyen, "On random sets and belief functions," *Journal of Mathematical Analysis and Applications*, vol. 65, pp. 531–542, 1978.

[17] W. L. Oberkampf, J. C. Helton, and K. Sentz, "Mathematical representation of uncertainty," *Non-Deterministic Approaches Forum*, April 2001.

[18] R. R. Yager, "Cumulative distribution functions from Dempster-Shafer belief structures," *IEEE Tran. on Systems, Man and Cybernetics*, vol. 34, no. 5, pp. 2080–2087, 2004.

[19] ——, "Entropy and specificity in a mathematical theory of evidence," *J. of General Systems*, vol. 9, no. 4, pp. 249–260, 1983.

[20] G. Klir and M. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory, 2nd edition*. Berlin: Springer-Verlag, 1999.

[21] D. Dubois and H. Prade, "A note on measures of specificity for fuzzy sets," *J. of General Systems*, vol. 10, no. 4, pp. 279–283, 1985.

[22] G. Klir, "Principles of uncertainty: What are they? why do we need them?" *Fuzzy Sets and Systems*, vol. 74, pp. 15–31, 1995.

[23] J. Tukey, *Exploratory Data Analysis*. Addison-Wesley, 1977.

[24] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.

[25] T. Denœux, "Constructing belief function from sample data using multinomial confidence regions," *International Journal of Approximate Reasoning*, vol. 42, pp. 228–252, 2006.

[26] A. Aregui and T. Denœux, "Constructing predictive belief functions from continuous sample data using confidence bands," in *Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications*, ser. ISIPTA 2007, Prague, Czech Republic, 2007, pp. 11–20.

[27] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of Approximate Reasoning*, vol. 9, pp. 1–35, 1993.