Engineering Management & Systems Engineering
Faculty Publications

Engineering Management & Systems Engineering

2012

Epistemic Uncertainty Analysis: An Approach Using Expert Judgment and Evidential Credibility

Patrick Hester
Old Dominion University

Research Article

Epistemic Uncertainty Analysis: An Approach Using Expert Judgment and Evidential Credibility

Patrick Hester

Engineering Management and Systems Engineering Department, Old Dominion University, Kaufman Hall, Room 241, Norfolk, VA 23529, USA

Correspondence should be addressed to Patrick Hester, pthester@odu.edu

When dealing with complex systems, all decision making occurs under some level of uncertainty. This is due to the physical attributes of the system being analyzed, the environment in which the system operates, and the individuals which operate the system. Techniques for decision making that rely on traditional probability theory have been extensively pursued to incorporate these inherent aleatory uncertainties. However, complex problems also typically include epistemic uncertainties that result from lack of knowledge. These problems are fundamentally different and cannot be addressed in the same fashion. In these instances, decision makers typically use subject matter expert judgment to assist in the analysis of uncertainty. The difficulty with expert analysis, however, is in assessing the accuracy of the expert's input. The credibility of different information can vary widely depending on the expert's familiarity with the subject matter and their intentional (i.e., a preference for one alternative over another) and unintentional biases (heuristics, anchoring, etc.). This paper proposes the metric of evidential credibility to deal with this issue. The proposed approach is ultimately demonstrated on an example problem concerned with the estimation of aircraft maintenance times for the Turkish Air Force.

1. Introduction

Real-world decision making is always performed under uncertainty. This uncertainty is present in the physical attributes of the system being analyzed, the environment in which it operates, and the individuals which operate the system. Decision makers must make decisions which best incorporate these uncertainties. With some problems, such as determining the probability of a terrorist attack on a given target, assigning probabilistic estimations to uncertain parameters is impossible due to the lack of statistical evidence upon which to base probabilistic estimates. Given these complex problems, decision makers often solicit subject matter expert opinion to provide estimates on uncertain parameters within a model. While this is a valid approach, soliciting expert opinions introduces additional uncertainty due to the varying degree of knowledge of the expert about the subject matter (i.e., one individual may truly be the world renowned expert in a field whereas others are merely

seasoned practitioners). Additionally, as human beings, they have the potential for intentional and unintentional biases.

The challenge when performing this type of analysis, in which expert judgment is essential to address uncertainty, is in assigning "weights" to the information provided by different experts in accordance with the level of expertise the expert provides. The credibility of different experts can vary widely depending on the expert's familiarity with the subject matter and their intentional (i.e., a preference for one alternative over another) and unintentional biases (e.g., heuristics, and anchoring). While expert opinion in an area that is of little familiarity to the expert may be not be entirely correct, there is no reason to believe that the information should be ignored completely, as the expert may have a particular insight to bring to the analysis. Further, the principle of complementarity [1] indicates that no one individual has complete knowledge of a complex system; thus, additional perspectives are value-added. Additionally, even though

human beings have inherent biases and prejudices, the information they provide should not be completely discounted. This paper develops an approach to address these problems.

This paper begins with a background discussion about uncertainty analysis, expert judgment elicitation, evidence combination, and expert biases. It then develops an approach which allows the decision maker to determine a level of credibility to use in incorporating each expert's evidence. The proposed approach is demonstrated on an example problem concerned with the estimation of aircraft maintenance times for the Turkish Air Force. Finally, conclusions and recommendations for future work are presented.

2. Background

Uncertainty is typically separated into aleatory uncertainty and epistemic uncertainty (see, e.g., [3, 4]). "Aleatory uncertainty is also referred to as variability, irreducible uncertainty, inherent uncertainty, stochastic uncertainty, and uncertainty due to chance. Epistemic uncertainty is also referred to as reducible uncertainty, subjective uncertainty, and uncertainty due to lack of knowledge" ([5], p. 10-2). Aleatory uncertainty refers to variation which is inherent to a given system, typically as a result of the random nature of model inputs. Aleatory uncertainties are typically modeled as random variables described by probability distributions, where decision makers typically make assumptions about the distribution's descriptive statistics (i.e., its mean and variance). "Epistemic uncertainty as a source of nondeterministic behavior derives from lack of knowledge of the system or the environment" Oberkampf et al. [6]. Oberkampf and Helton [5] elaborate on this definition:

> The key feature stressed in this definition is that the fundamental source of epistemic uncertainty is incomplete information or incomplete knowledge of some characteristic of the system or the environment. As a result, an increase in knowledge or information can lead to a reduction in the predicted uncertainty of the response of the system, all things being equal. Examples of sources of epistemic uncertainty are: little or no experimental data for a fixed (but unknown) physical parameter, a range of possible values of a physical quantity provided by expert opinion, limited understanding of complex physical processes, and the existence of fault sequences or environmental conditions not identified for inclusion in the analysis of a system (p.10-2).

Epistemic uncertainty often becomes an issue when expert opinion is required to solve a problem. In trying to determine the likelihood of a terrorist attack on a given building, a decision maker may solicit many expert opinions due to a lack of sufficient knowledge about the problem. In doing so, the decision maker is introducing additional uncertainty into the analysis, both in the lack of knowledge about the credibility of the experts being solicited and in the experts' own intentional (i.e., a preference for one alternative over another) and unintentional (heuristics, anchoring, etc.)

biases that influence the information they provide. Epistemic uncertainty can be reduced with increased information, but aleatory uncertainty is a function of the problem characteristics itself.

Oberkampf et al. [7] describe various methods for estimating the total uncertainty in a model by identifying all sources of variability and uncertainty. Traditionally, uncertainty has been handled with probability theory, but recent developments maintain that representing all uncertainty information in the same manner is inappropriate and, in order to be analyzed appropriately, several experts believe that aleatory and epistemic uncertainty should be addressed separately (e.g., [8–15]).

Traditional quantification of uncertainty uses probability theory, which represents uncertainties as random variables by utilizing a probability density function which presents the probability information about the variable. Probability theory, however, has problems separating aleatory from epistemic uncertainty [8]. As a result, various techniques including Dempster-Shafer theory [16, 17] have gained increased use in recent years as techniques that can adequately separate differing types of uncertainty. Other theories such as generalized information theory [18] and approximate reasoning [19] have also proven useful in characterizing uncertainty.

Modern approaches to deal with epistemic uncertainty include fuzzy sets [20, 21], Dempster-Shafer theory [16, 17, 22], and possibility theory [23]. Dempster-Shafer theory was chosen for use in this paper due to its strong theoretical basis, large number of recent example problems to draw from, and versatility of Dempster-Shafer theory to represent and combine potentially dissimilar evidence from various sources. A brief discussion of the mathematics of Dempster-Shafer evidence theory is provided in the approach section (Section 3) of this paper.

In addition to dealing with uncertainties present in the problem domain, analysts must also understand what inherent biases are incorporated into an individual's thought process. The following section discusses biases that may influence an expert's judgment.

2.1. Biases. Whenever an expert is utilized as a source of information, their beliefs and experiences bias how they view the problem and what information they choose to provide to help solve the problem. These biases take the form of either intentional or unintentional biases. Intentional biases are a result of the expert's willful decision to bias the results of their assessment. This willful deceit can occur due to preference of one alternative over another. The expert may prefer one alternative over another due to gains that he/she stands to receive as a result of the analysis. An example would be a company that is using expert judgment to assess its building's level of security. If the expert were to have an interest in convincing the company that their security levels were subpar (such as if the expert owned his/her own security company), then the expert may intentionally bias the results. Alternatively, the expert may have a reason not to prefer a particular alternative and may intentionally bias the results

accordingly. Typically, these intentional biases are easier for an outside observer to recognize as strong connections between the expert and his/her intentionally biased choice (such as significant financial connections) should emerge. It is important to note that the vast majority of experts will not exhibit this behavior, but it is important for the analyst to be cognizant of the potential for this bias nonetheless.

It is mistakenly assumed that because an individual is an expert in a particular subject matter, he/she is perfectly capable of providing accurate likelihood estimates for particular events. Even without intentional biases to account for, all human beings have unintentional cognitive biases that affect the information that is elicited from them. These cognitive biases include behaviors such as the availability heuristic, conjunction fallacy, representativeness heuristic, and anchoring.

The availability heuristic [24, 25] refers to the practice of basing probabilistic evidence on an available piece of information in one's own set of experiences. That is to say, humans estimate the likelihood of an event based on a similar event that they can remember. Further, since newer events are fresher in our minds, they influence our reasoning in larger proportion than older events. Since experts have a larger set of experiences to draw from, and thus more available data, it is likely that their propensity for the availability heuristic will decrease as their experience level increases. However, a more naïve expert may be able to provide a better result if he/she has experienced a relevant event recently, whereas an expert in the field with many years of relevant experience (none of which are recent), may not be as likely to provide useful information.

Another bias that humans incorporate when providing uncertainty estimates is the conjunction fallacy. This fallacy occurs when individuals identify specific scenarios as being more likely than general ones. Tversky and Kahneman [26] explored this phenomenon and found that this mistake is commonly committed despite the fact that it is mathematically impossible for the joint probability of two events to be more likely than the probability of either of the individual events. Individuals often make this mistake as the specific scenario seems more realistic to them and it is possible that experts can be prone to this type of fallacy as well. While experts are less prone to this type of behavior, the phenomenon is still something that analysts should be aware of when eliciting expert opinion.

The representativeness heuristic [25] occurs when commonalities between objects are assumed. For example, an expert has estimated the probability of attack against a building before and assumes the current building that is being analyzed is similar to his/her previous work, and therefore, estimates the probabilities to be similar. There may, in fact, be a glaring difference between the two problems that the expert is overlooking.

Another bias is the anchoring and adjustment heuristic, observed by Tversky and Kahneman [24]. Humans anchor their judgments and base subsequent observations on the initial value that was provided to them. In other words, if the expert is provided a baseline value, he/she can be influenced to a degree where subsequent probability values

will be anchored by the provided baseline value. Even experts can be influenced to provide probabilistic values close to values that the analyst desires by anchoring the questions that are asked when eliciting their opinion.

The biases discussed here are only a few of the possible that may affect experts. The important takeaway with respect to biases, both intentional and unintentional, is that decision makers must be cognizant of their effect on results obtained when eliciting expert judgment. Any approach that incorporates expert judgment must take into the account the presence of biases and adjust its approach accordingly.

An approach is developed in the next section which provides a method for dealing with these biases when using expert judgment to address epistemic uncertainty.

3. Solution Approach

3.1. Dempster-Shafer Theory. Dempster-Shafer theory is a mathematical theory of evidence, defined by three important functions: the basic probability assignmentfunction (BPA or m), the belief function (Bel), and the plausibilityfunction (Pl). The seminal work on the subject is [17], which is an expansion of [16]. In evidence theory, uncertainty is separated into belief and plausibility, whereas traditional probability theory uses only the probability of an event to analyze uncertainty. Belief and plausibility provide bounds on probability. In special cases, they converge on a single value, probability. In other cases, such as in the evidence theory representation of uncertainty, belief, and plausibility represents a range of potential values for a given parameter, without any assumptions on the likelihood of the underlying data.

In evidence theory, for a sample space X, degrees of evidence are assigned to subsets (events) of X. A subset (x) with a nonzero degree of evidence is called a focal element. Based on available information, a basic probability assessment (BPA), denoted by $m(x)$, can be defined as

$$m(x) \geq 0 \quad \text{for } x \subset X \tag{1a}$$

$$\sum_{x \subset X} m(x) = 1. \tag{1b}$$

An BPA is provided by experts in lieu of a traditional probability assessment. Imagine a scenario in which experts are being asked to predict weather occurrences in a given city. Two experts (*E1* and *E2*) are providing their opinions on the likelihood of three weather occurrences (W1, W2, and W3). The potential weather phenomena are as follows: W1 is sunny, W2 is cloudy and W3 is rainy.

In this case, the objective is to find the likely weather occurrence. As such, $X = \{W1, W2, W3\}$ and the frame of discernment representing all possible categories of evidence, $2^X = \{W1, W2, W3, W1 \cup W2, W2 \cup W3, W1 \cup W3, W1 \cup W2 \cup W3\}$.

Suppose the following information is collected: Expert one (*E1*) says it will be rainy or sunny with 90% probability, while Expert two (*E2*) says it will be sunny or cloudy with 75% probability.

The BPA for expert one can then be stated as

$$m_{E1}(\text{W1} \cup \text{W3}) = 0.90, \qquad m_{E1}(\text{W1} \cup \text{W2} \cup \text{W3}) = 0.10. \tag{2}$$

The second piece of evidence ($m_{E1}(\text{W1} \cup \text{W2} \cup \text{W3})$) is due to the fact that nothing is known about the remaining evidence, so it must be allocated to what is termed the remaining frame of discernment. That is, since nothing is known (except that weather must occur), a judgment cannot be made about the unknown frame of reference (the remaining 0.10 is not specified by the expert, so it is specified as being either W1 or W2 or W3).

The BPAs for the second expert can be stated as

$$m_{E2}(\text{W1} \cup \text{W2}) = 0.75, \qquad m_{E2}(\text{W1} \cup \text{W2} \cup \text{W3}) = 0.25. \tag{3}$$

In other words, *E1* says the weather phenomena are most probably W1 or W3, while *E2* says the weather phenomena are most probably W1 or W2. The sources of evidence can be combined using the Dempster rule of combination as follows:

$$m_{12}(A) = \frac{\sum\limits_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum\limits_{B \cap C = \varnothing} m_1(B) m_2(C)}. \tag{4}$$

This equation yields the combined evidence of experts 1 and 2 that support *A* (which is composed of the intersection of *B* and *C*). This can be applied recursively to combining the evidence of more than two experts. For example, the results obtained from combining experts one and two can then be combined with expert three in order to determine the combination of the three experts' evidence. The order in which the experts' evidence is combined is irrelevant.

Dempster's rule of combination has been subject to criticism in that it tends to ignore conflict available within the evidence (as pointed out by [27]) and attributes evidence supporting conflict to the null set [22]. Additional combination rules deal with this complication, but they require that the relative importance (in the case of this approach, the evidential credibility values) of each expert is known. Information on additional rules of combination is provided in Agarwal et al. [8] and Sentz and Ferson [28]. For the purposes of this paper, it will be assumed that the evidence provided does not have a large enough level of conflict that using Dempster's rule will adversely affect the results. The approach presented here can be generalized to other rules of combination if desired.

The evidence of the two weather experts can then be combined as in Table 1.

Given this combined BPA, the evidence can now be used to form belief and plausibility bounds on the uncertainty. Belief in any set is the sum of all probabilities of all subsets of that set. It represents any proof that has been provided (it is believed) that a particular event is true.

Belief values for the individual events in the aforementioned problem are

$$\text{Bel}_{E1, E2}(\text{W1}) = m_{E1, E2}(\text{W1}) = 0.675,$$

$$\text{Bel}_{E1, E2}(\text{W2}) = 0, \tag{5}$$

$$\text{Bel}_{E1, E2}(\text{W3}) = 0.$$

On the other hand, plausibility is more general. It represents the degree to which it is plausible that a particular event is true. Another way to look at belief is it is a measure of the degree to which an event will happen, whereas plausibility is a measure of the degree to which an event could happen.

Plausibility values for the individual events in the aforementioned problem are

$$\begin{aligned}
\text{Pl}_{E1, E2}(\text{W1}) &= m_{E1, E2}(\text{W1}) + m_{E1, E2}(\text{W1} \cup \text{W3}) \\
&\quad + m_{E1, E2}(\text{W1} \cup \text{W2}) \\
&\quad + m_{E1, E2}(\text{W1} \cup \text{W2} \cup \text{W3}) \\
&= 0.675 + 0.225 + 0.075 + 0.025 = 1,
\end{aligned}$$

$$\begin{aligned}
\text{Pl}_{E1, E2}(\text{W2}) &= m_{E1, E2}(\text{W1} \cup \text{W2}) \\
&\quad + m_{E1, E2}(\text{W1} \cup \text{W2} \cup \text{W3}) \\
&= 0.075 + 0.025 = 0.10,
\end{aligned}$$

$$\begin{aligned}
\text{Pl}_{E1, E2}(\text{W3}) &= m_{E1, E2}(\text{W1} \cup \text{W3}) \\
&\quad + m_{E1, E2}(\text{W1} \cup \text{W2} \cup \text{W3}) \\
&= 0.225 + 0.025 = 0.25.
\end{aligned} \tag{6}$$

More information on the mathematics and application of evidence theory can be found in Dempster [16] and Shafer [17].

3.2. Evidential Credibility. When expert judgments are elicited, and epistemic uncertainty is introduced due to lack of knowledge about the credibility of the evidence provided by experts, however, a modified version of evidence theory must be developed to deal with this additional layer of uncertainty. In this modified approach, the proposed expert's modified BPA, $m_i^*(x)$, is given as

$$m_i^*(x) = m_i(x) \text{EC}_i(x), \tag{7}$$

where $m_i(x)$ is as defined in (1a) and (1b) and $\text{EC}_i(x)$ is the evidential credibility value, with i and x being the indices corresponding to the particular expert and event, respectively.

Evidential credibility is a measure of the analyst's confidence in the expert's estimated likelihood values; it acts as a weight to adjust the likelihood estimate given by the individual expert. An evidential credibility value of one means that the analyst has complete confidence in the expert's estimate and it should be taken into account fully,

TABLE 1: Combined expert evidence.

	$m_{E2}(\text{W1} \cup \text{W2}) = 0.75$	$m_{E2}(\text{W1} \cup \text{W2} \cup \text{W3}) = 0.25$
$m_{E1}(\text{W1} \cup \text{W3}) = 0.90$	$m_{E1,E2}(\text{W1}) = 0.675$	$m_{E1,E2}(\text{W1} \cup \text{W3}) = 0.225$
$m_{E1}(\text{W1} \cup \text{W2} \cup \text{W3}) = 0.10$	$m_{E1,E2}(\text{W1} \cup \text{W2}) = 0.075$	$m_{E1,E2}(\text{W1} \cup \text{W2} \cup \text{W3}) = 0.025$

whereas an evidential credibility value of less than one demonstrates the analyst's reluctance to place full confidence in the likelihood estimates provided by the expert. The remainder of evidence not attributed to an event by the expert (independent of his/her evidential credibility) is allocated to the remaining frame of discernment as detailed earlier in this section (with this evidence increasing as $\text{EC}_i(x)$ approaches zero). If evidential credibility is calculated as one for an expert, the adjusted BPA in (7) reverts to the original form in (1a) and (1b). If the expert's calculated evidential credibility is zero, then his/her likelihood for the specified event reverts to zero. If an expert is deemed to have an evidential credibility of zero for all possible events, all evidence for the given expert is allocated to complete ignorance ($m^*((\cup x) \forall x \in x) = 1$). This reflects the notion that the analyst has no confidence in the expert's predictions, based on his/her evidential credibility. For the previous weather example, a single expert being polled with no evidential credibility would result in an BPA of $m_{E1}^*(\text{W1} \cup \text{W2} \cup \text{W3}) = 1.0$.

For the proposed approach, evidential credibility is calculated in a manner derived from the Brier score (1950) approach to evaluating experts, which is straightforward and provides a good basis for the development of an evidential credibility measure. Brier's work is predicated on the existence of verifiable data, whereas there is no established "right" answer in many of the applications for which this paper's proposed approach is intended. There are two options to deal with this complication: (1) develop a Brier score based on the information that is present (e.g., historical data of similar systems) or (2) adjust the Brier score to create a new scoring rule which reflects the lack of available data. The author utilizes the second approach in this paper, as the first approach is valid only in simpler systems where the variance between new and old systems is trivial, thus making the necessity of the method developed in this paper unnecessary. The adjusted Brier score, then, is used to calculate the evidential credibility as follows:

$$\text{EC}_i(x) = \left[1 - \left| f_i(x) - \overline{m(x)} \right|^{1/2} \right], \tag{8}$$

where $f_i(x)$ is the ith expert's estimate of the average of all expert forecasts for event x and $\overline{m(x)}$ is the average of actual expert predictions of event x.

Equation (8) reflects an error-function-based approach to scoring experts' evidence. Evidential credibility is not intended to reflect the expert's individual credibility, but rather it provides a discount factor of the individual's knowledge about what the collective judgment of the experts will be. This point is illustrated in (8), where the difference between the expert's prediction of the average estimation and the actual average of the expert's estimations (the error) is

calculated. The closer the two values are together, the lower the prediction error of the expert is, and, thus, the more knowledgeable the expert is proven to be. A true expert would be able to provide an accurate prediction of an event ($m_i(x)$) and an estimate of what other experts would say ($f_i(x)$). If he/she is not accurate in this regard, the resulting $\text{EC}_i(x)$ will be reduced. This makes an intuitive sense as the true expert should understand what knowledge he/she has that would alter his/her prediction relative to other experts. If the expert is not able to do so, the analyst's confidence in his/her predictions should be reduced (as demonstrated by an evidential credibility value of less than one) as they are likely not as knowledgeable as originally predicted. If the expert is not privy to additional information which biases his/her predictions, he/she is likely to assume $f_i(x) = m_i(x)$.

Utilizing the above definition for evidential credibility, individuals are incentivized to report their true predictions for the pool of experts, as shading their predictions would result in their evidential credibility being reduced and their initial prediction, $m_i(x)$, being adjusted significantly by a lower $\text{EC}_i(x)$ value. Thus, less evidential credibility means the evidence provided by the expert (in terms of their likelihood values) will be lessened as there is less confidence in their estimates. It should be noted that if an expert is not comfortable with providing a likelihood estimate for a particular event, he/she may abstain from providing one and the average prediction that is calculated will exclude any individuals who choose to omit a response. As each response provided by the experts has its own evidential credibility value, there is no incentive for a particular expert to answer more questions than he/she is comfortable with to artificially inflate their $\text{EC}_i(x)$. If an event has only one expert providing an estimate for its likelihood, there is no expert to contrast with and therefore their evidential credibility is taken to be one (since they are incentivized to tell the truth, there is no reason to believe that their answer is anything less than completely truthful, as they have no knowledge of how many other experts will answer this particular question).

Additionally, the evidential credibility metric is useful in combating biases facing experts. If an expert is prone to the availability heuristic, conjunction fallacy, representativeness heuristic, or anchoring, and these biases influence their average predictions, their resultant evidence will be reduced through a reduced $\text{EC}_i(x)$ value. Thus, the experts are incentivized to examine their predictions before reporting them and report their true values, free of intentional and unintentional biases.

The approach proposed in this paper is useful for several reasons.

(1) It has the benefit of assigning a separate evidential credibility value for each expert and each event being

estimated so that the opinions of experts with varying levels of experience can be combined without having to ignore experts with less experience.

(2) The evidential credibility measure is flexible and can be utilized regardless of how much information each expert chooses to provide for a particular problem.

(3) There is no need for this approach to have the "right" answer a priori, as experts are ranked relative to one another's estimates and not relative to a correct base-line. This is especially useful for complex problems as the correct answer is often not known until after the analysis occurs, if at all.

(4) This approach incentivizes experts to tell the truth, both in their own estimates and in their predictions of others' estimates, limiting the effect of intentional and unintentional biases on the resulting expert judgments.

Utilizing the proposed approach, two example problems are solved in the next section.

4. Example Discussion

The example demonstration of the proposed method for calculating evidential credibility is drawn from Kudak and Hester [2] and based on an example regarding maintenance time estimation in the Turkish Air Force. Kudak and Hester [2] explain how evidence theory was used to estimate maintenance times for damaged aircraft repairs. Three experts were surveyed and asked to provide their estimates regarding the three major statistically observed failure sources of ignition, fuel, and electrical systems [29].

Hester and Kudak [2] describe the problem:

> *For wartime operations, Maintenance Time (MT) of each specific failure can be divided into three separate time intervals (x_1, x_2 and x_3) as shown in [Figure 1]. Expected Actual Time, $E_a(t)$, is the exact MT as assigned in [29] for normal operation times. The first time interval, x_1, represents times less than the Best Time (defined by the expert as the shortest maintenance time expected to complete the failure, with a default time given as 5–10% less than $E_a(t)$, depending on the system). The second interval, x_2, represents the period between Best Time and Worst Time (defined by the expert as the longest maintenance time expected to complete the failure, with a default time given as 5–10% more than $E_a(t)$, depending on the system), and it includes the Expected Actual Time $E_a(t)$. The third time interval, x_3, represents times greater than the Worst Time. (pp. 56-57).*

This example will be used to explore the evidential credibility metric proposed in this paper.
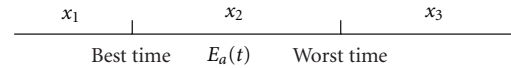

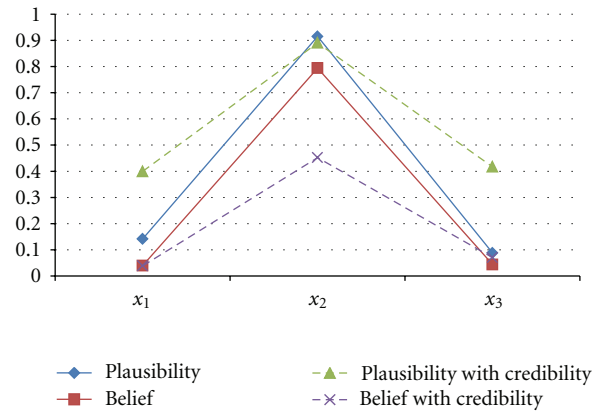Figure 1: Maintenance time intervals (adapted from Kudak and Hester [2]).

Figure 2: Belief and plausibility measures with and without evidential credibility.

5. Results and Analysis

The three experts discussed in Kudak and Hester [2] were asked to provide their BPAs for the three time intervals shown in Figure 1 and their combinations. Their BPAs are provided in Table 2.

Using (4), these estimates can be combined to provide belief and plausibility estimates as shown in Tables 3 and 4, respectively.

Now, let us suppose that each expert is asked to provide an estimate of the average of all expert forecasts, $f_i(x)$. In the absence of greater knowledge about his or her fellow experts, each individual assumes their estimates are equal to the average, that is, $f_i(x) = m_i(x)$, as discussed in Section 3.2. Using (8), we can then calculate the $EC_i(x)$ for each event and expert as shown in Table 5.

Using (7), modified BPAs can be generated for the experts, as shown in Table 6.

Using (4), these modified BPA assignments can be combined to provide belief and plausibility estimates (including evidential credibility) as shown in Tables 7 and 8.

Further, belief and plausibility values for each of the three time ranges (x_1, x_2, and x_3) can be compared graphically as shown in Figure 2.

It is clear from investigating Figure 2 that evidence supporting x_2 has decreased when evidential credibility is taken into account. While baseline belief and plausibility estimates indicate a narrow band at x_2, incorporation of evidential credibility widens this band and allocates associated evidence to both x_1 and x_3. Further investigation of Tables 5 and 2 to determine the cause of this change reveals that Expert 3 is likely the reason behind the adjusted estimates. His/her low evidential credibility (0.42, as shown in Table 5) with respect to estimating x_2 reduces the credibility of his/her

TABLE 2: BPA Assignments (adapted from Kudak and Hester, [2]).

	$m_i\{x_1\}$	$m_i\{x_2\}$	$m_i\{x_3\}$	$m_i\{x_1,x_2\}$	$m_i\{x_1,x_3\}$	$m_i\{x_2,x_3\}$	$m_i\{x_1,x_2,x_3\}$
Expert 1	0	0	0.2	0.6	0	0	0.2
Expert 2	0	0.3	0	0	0	0.3	0.4
Expert 3	0.1	0.65	0	0	0	0	0.25

TABLE 3: Belief values.

$\mathrm{Bel}_{123}\{x_1\}$	$\mathrm{Bel}_{123}\{x_2\}$	$\mathrm{Bel}_{123}\{x_3\}$	$\mathrm{Bel}_{123}\{x_1,x_2\}$	$\mathrm{Bel}_{123}\{x_1,x_3\}$	$\mathrm{Bel}_{123}\{x_2,x_3\}$
0.04	0.794	0.044	0.911	0.085	0.857

TABLE 4: Plausibility values.

$\mathrm{Pl}_{123}\{x_1\}$	$\mathrm{Pl}_{123}\{x_2\}$	$\mathrm{Pl}_{123}\{x_3\}$	$\mathrm{Pl}_{123}\{x_1,x_2\}$	$\mathrm{Pl}_{123}\{x_1,x_3\}$	$\mathrm{Pl}_{123}\{x_2,x_3\}$
0.142	0.915	0.088	0.955	0.205	0.959

TABLE 5: Evidential credibility values.

	$\mathrm{EC}_i\{x_1\}$	$\mathrm{EC}_i\{x_2\}$	$\mathrm{EC}_i\{x_3\}$	$\mathrm{EC}_i\{x_1,x_2\}$	$\mathrm{EC}_i\{x_1,x_3\}$	$\mathrm{EC}_i\{x_2,x_3\}$
Expert 1	0.82	0.44	0.63	0.37	1.00	0.68
Expert 2	0.82	0.87	0.74	0.55	1.00	0.55
Expert 3	0.74	0.42	0.74	0.55	1.00	0.68

TABLE 6: Modified BPA assignments incorporating evidential credibility.

	$m_i^*\{x_1\}$	$m_i^*\{x_2\}$	$m_i^*\{x_3\}$	$m_i^*\{x_1,x_2\}$	$m_i^*\{x_1,x_3\}$	$m_i^*\{x_2,x_3\}$	$m_i^*\{x_1,x_2,x_3\}$
Expert 1	0.00	0.00	0.13	0.22	0.00	0.00	0.65
Expert 2	0.00	0.26	0.00	0.00	0.00	0.17	0.57
Expert 3	0.07	0.27	0.00	0.00	0.00	0.00	0.65

TABLE 7: Modified belief values incorporating evidential credibility.

$\mathrm{Bel}_{123}^*\{x_1\}$	$\mathrm{Bel}_{123}^*\{x_2\}$	$\mathrm{Bel}_{123}^*\{x_3\}$	$\mathrm{Bel}_{123}^*\{x_1,x_2\}$	$\mathrm{Bel}_{123}^*\{x_1,x_3\}$	$\mathrm{Bel}_{123}^*\{x_2,x_3\}$
0.04	0.45	0.07	0.58	0.11	0.60

TABLE 8: Modified plausibility values incorporating evidential credibility.

$\mathrm{Pl}_{123}^*\{x_1\}$	$\mathrm{Pl}_{123}^*\{x_2\}$	$\mathrm{Pl}_{123}^*\{x_3\}$	$\mathrm{Pl}_{123}^*\{x_1,x_2\}$	$\mathrm{Pl}_{123}^*\{x_1,x_3\}$	$\mathrm{Pl}_{123}^*\{x_2,x_3\}$
0.40	0.89	0.42	0.93	0.55	0.96

strong evidence supporting this event (an BPA of 0.65, as shown in Table 2). Thus, it would be worthwhile for the analyst to seek more information from expert 3 to determine why he/she felt so strongly about x_2 and yet did not provide an accurate estimate with respect to his/her fellow experts ($f_i(x_2)$). Similar analyses can be undertaken to support Kudak and Hester's [2] discussion regarding fuel and electrical systems, or in any other problem where evidence theory is an appropriate candidate for uncertainty quantification.

6. Conclusions

This paper proposed an approach for including a measure of evidential credibility into analysis when eliciting expert opinion to estimate epistemic uncertainty in a problem. It is the hope of the author that this approach can be extended to other evidence theory combination rules in the future in order to further explore its usefulness. Other scoring rules (specifically Prelec [30] and Matheson and Winkler [31]) should also be explored for use in this approach. Further,

it is also thought that this approach, while demonstrated on a single case study in this paper, must be further explored to ensure its validity and utility.

References

[1] N. Bohr, "Quantum postulate and the recent development of atomic theory," *Nature*, vol. 121, no. 3050, pp. 580–591, 1928.

[2] H. Kudak and P. Hester, "Application of Dempster-Shafer theory in aircraft maintenance time assessment: a case study," *Engineering Management Journal*, vol. 23, no. 2, pp. 55–62, 2011.

[3] J. C. Helton and D. E. Burmaster, "Treatment of aleatory and epistemic uncertainty in performance assessments for complex systems," *Reliability Engineering and System Safety*, vol. 54, no. 2-3, pp. 91–94, 1996.

[4] W. L. Oberkampf, "Uncertainty quantification using evidence theory," in *Proceedings from the Advanced Simulation & Computing Workshop*, Albuquerque, NM, USA, 2005.

[5] W. L. Oberkampf and J. C. Helton, "Evidence theory for engineering applications," in *Engineering Design Reliability Handbook*, E. Nikolaidis, D. M. Ghiocel, and S. Singhal, Eds., CRC Press, Boca Raton, Fla, USA, 2005.

[6] W. L. Oberkampf, S. M. DeLand, B. M. Rutherford, K. V. Diegert, and Alvin, "Error and uncertainty in modeling and simulation," *Reliability Engineering and System Safety*, vol. 75, no. 3, pp. 333–357, 2002.

[7] W. L. Oberkampf, K. V. Diegert, K. F. Alvin, and B. M. Rutherford, "Variability, uncertainty, and error in computational simulation," in *ASME Proceedings of the 7th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, vol. 2, pp. 259–272, 1998.

[8] H. Agarwal, J. E. Renaud, E. L. Preston, and D. Padmanabhan, "Uncertainty quantification using evidence theory in multidisciplinary design optimization," *Reliability Engineering and System Safety*, vol. 85, no. 1–3, pp. 281–294, 2004.

[9] J. Darby, *Evaluation of Risk from Acts of Terrorism: The Adversary/Defender Model Using Belief and Fuzzy Sets (SAND2006-5777)*, Sandia National Laboratories, Albuquerque, NM, USA, 2006.

[10] J. C. Helton, "Treatment of uncertainty in performance assessments for complex systems," *Risk Analysis*, vol. 14, no. 4, pp. 483–511, 1994.

[11] J. C. Helton, "Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty," *Journal of Statistical Computation and Simulation*, vol. 57, no. 1–4, pp. 3–76, 1997.

[12] F. O. Hoffman and J. S. Hammonds, "Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability," *Risk Analysis*, vol. 14, no. 5, pp. 707–712, 1994.

[13] M. E. Paté-Cornell, "Uncertainties in risk analysis: six levels of treatment," *Reliability Engineering and System Safety*, vol. 54, no. 2-3, pp. 95–111, 1996.

[14] S. N. Rai, D. Krewski, and S. Bartlett, "A general framework for the analysis of uncertainty and variability in risk assessment," *Human and Ecological Risk Assessment*, vol. 2, no. 4, pp. 972–989, 1996.

[15] W. D. Rowe, "Understanding uncertainty," *Risk Analysis*, vol. 14, no. 5, pp. 743–750, 1994.

[16] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Statistics*, vol. 38, no. 2, pp. 325–339, 1967.

[17] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA, 1976.

[18] G. J. Klir and M. J. Wierman, *Uncertainty Based Information: Elements of Generalized Information Theory*, Physica, Heidelberg, Germany, 1998.

[19] D. Padmanabhan and S. M. Batill, "Reliability based optimization using approximations with applications to multidisciplinary system design," in *Proceedings of the 40th AIAA Sciences Meeting & Exhibit*, AIAA-2002-0449, Reno, Nev, USA, 2002.

[20] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[21] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, USA, 1995.

[22] R. R. Yager, "Arithmetic and other operations on Dempster-Shafer structures," *International Journal of Man-Machine Studies*, vol. 25, no. 4, pp. 357–366, 1986.

[23] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, NY, USA, 1988.

[24] A. Tversky and D. Kahneman, "Availability: a heuristic for judging frequency and probability," *Cognitive Psychology*, vol. 5, no. 2, pp. 207–232, 1973.

[25] A. Tversky and D. Kahneman, "Judgment under uncertainty: heuristics and biases," *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974.

[26] A. Tversky and D. Kahneman, "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment," *Psychological Review*, vol. 90, no. 4, pp. 293–315, 1983.

[27] L. A. Zadeh, "Review of Shafer's mathematical theory of evidence," *Artifical Intelligence Magazine*, vol. 5, pp. 81–83, 1984.

[28] K. Sentz and S. Ferson, *Combination of Evidence in Dempster-Shafer Theory (SAND2002-0835)*, Sandia National Laboratories, Albuquerque, NM, USA, 2002.

[29] T. O. F-4E-1-9, "F-4E aircraft jet engine maintenance technical order manual," U.S. Air Force, 1983.

[30] D. Prelec, "A Bayesian truth serum for subjective data," *Science*, vol. 306, no. 5695, pp. 462–466, 2004.

[31] J. E. Matheson and R. L. Winkler, "Scoring rules for continuous probability distributions," *Management Science*, vol. 22, no. 10, pp. 1087–1096, 1976.