Expert Systems with Applications 40 (2013) 964–974

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

New developments in uncertainty assessment and uncertainty management

T.J. Ross [a], J.M. Booker [b,*], A.C. Montoya [c]

[a] Department of Civil Engineering, University of New Mexico, USA
[b] Booker Scientific, 2682 Old Harper Road Fredericksburg, TX 78624, USA
[c] Department of Civil Engineering, University of New Mexico, USA

ARTICLE INFO

Keywords:
Uncertainty quantification
Fuzzy sets
Inference
Total uncertainty
Management

ABSTRACT

The paper presents a general method for doing predictions and test planning which can also be used as a tool for managing uncertainty. Uncertainty is generally defined as "that which is not precisely known". This definition permits the identification of different kinds of uncertainty arising from different sources and activities, most of which go unnoticed in analysis. In this paper, the description of uncertainty begins from a historical perspective and concludes with a new perspective based upon making inferences; fuzzy logic can be most helpful in quantifying some inferences. Our assessment of uncertainty begins the identification of the various forms of uncertainty (ambiguity, fuzziness, randomness, non-specificity, ignorance, etc.) and concludes with models and methods for assessing the 'total uncertainty' within an application. The material contained herein is described in the context of physical science and engineering applications; however, nothing presented precludes application to other fields, e.g. economics, social sciences, medicine and business. Uncertainty assessment involves how to identify, classify, characterize, quantify, and combine uncertainties within an application, with the expressed goal of understanding how to manage uncertainties. Uncertainty management presumes that we have a process to quantify uncertainties and to be able to aggregate them in such a way that they can be compared in terms of their individual contributions to the 'total uncertainty'. Managing uncertainties is important, because uncertainties directly affect decision and policy making. An example, using a concept called Quantification of Margins and Uncertainty (QMU), is provided to illustrate our ideas.

1. Introduction

We have spent many decades trying to understand and quantify uncertainty. The more uncertainty in a problem, the less precise or correct we can be in our understanding of that problem. Most science and engineering endeavors do not address the uncertainty in the information, models, and solutions that are conveyed within the problem. We propose that the characterization and quantification of uncertainty within such problems should be done commensurate with what is known or can be determined in the physical world and with an appropriate level of expressed precision. One reason for engaging in such a pursuit is obvious: achieving high levels of precision costs significantly in time, money or both. The more complex a problem or system is, the more imprecise or inexact is the information that we have to characterize that system and hence the greater the uncertainty about it. Thus, uncertainty is related to precision, information, and complexity, making its assessment paramount in the problems we pose for eventual solution. Lotfi Zadeh has a famous quote about the importance of balancing the precision we seek with the uncertainty that exists (Zadeh, 1973), "*we must exploit our tolerance for imprecision*".

This paper suggests that uncertainty of various forms permeates all scientific endeavors, and it exists as an integral feature of all abstractions, models, theory, and solutions. It is our intent to summarize methods to handle some of these forms of uncertainty in our technical problems. Since much of what we have explored in the past 45 years is new and not part of a canonical jargon, we begin first with the definition of some terms and some critical thinking about uncertainty.

2. Some thoughts and definitions

The discussion of the assessment and management of uncertainty begins with some definitions and concepts:

Certainty can operationally be defined as "a state such that evidence to the contrary is below a threshold of disputation." For example, if we say that initial conditions in an experiment are known with certainty, we are understood as saying that initial conditions are indisputably fixed. This definition of certainty coincides with *determinism* or a deterministic solution.

* Corresponding author.
E-mail address: ross@unm.edu (J.M. Booker).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.05.054

Precision refers to abilities in making good predictions, being exact, being correct, maintaining control, operating within specifications, and representing the physical world. These achievements usually coincide with a high degree certainty.

Uncertainty is broadly defined as "what is not known precisely," and manifests itself in numerous ways, most of which are undetected or considered too difficult to assess. A partial list of theories used to assess uncertainty, developed over the past 45 years (beyond probability theory which is over 400 years old!), is provided below in Section 4.

Quantification (as in *uncertainty quantification* or *UQ*) refers to an analysis and assessment process or evaluation based upon models, data, expertise, etc. Quantification does not necessarily require a numerical statement or conversion to a number. Some might distinguish linguistic statements describing uncertainty from quantification by identifying the former as an uncertainty assessment. We understand that numbers, like linguistics, have an interpretation that is provided by a human, often a policy or decision-maker. Thus UQ is equally valid expressed as numbers or words. A non-numeric example of a UQ statement is: *Jet engines operate with small but stable margins.*

Confidence is a commonly used term whose definitions include words like trust, belief, reliance, and certitude; it is the state of feeling sure. "Confidence comes from repetition, from the breath of many mouths," from an anonymous quote. It is interesting to note that even the Greeks were unable to precisely (or mathematically) define what is meant by confidence. In statistics, confidence has a specific meaning in sampling and inference when referring to a *confidence interval* for an unknown parameter (e.g., the mean). The *confidence level*, $1-\alpha$, is defined as the complement of a significance level, α, or the Type I error in statistical hypothesis testing. (Type I error is the chance, e.g., 5%, that a null hypothesis is rejected when it should not have been rejected, i.e., the null hypothesis is true. This is a chosen, and therefore, controlled error in statistical inference.) Outside of the statistical context, there is no modern day definition for the mathematical meaning or quantification of confidence. The use of the term in statistical applications notwithstanding, we propose a technical definition of confidence as having an inverse relationship to uncertainty, which is assessed and/or quantified.

Total uncertainty is the combination or aggregation of all relevant uncertainties within an application or problem. Ideally, an analyst would produce an estimate of the final or system-level overall answer to a question, y, accompanied by a total uncertainty estimate, Δy, or $y \pm \Delta y$. Constructing methods of determining total uncertainty is a relatively new area of research (see Ross, 2003).

3. Brief history of uncertainty assessment & quantification

From an historical point of view the issue of uncertainty has not always been embraced within the scientific community (Klir & Yuan, 1995). In the traditional view of science, uncertainty represents an undesirable state, a state that must be avoided at all costs. This was the state of science until the late nineteenth century when physicists realized that Newtonian mechanics did not address problems at the molecular level. Newer methods, associated with statistical mechanics, were developed which recognized that statistical averages could replace the specific manifestations of microscopic entities, accounting for what was not precisely known (i.e., uncertainty). These statistical quantities, which summarized the activity of large numbers of microscopic entities, could then be connected in a model with appropriate macroscopic variables (Klir & Yuan, 1995). Since then, the role of Newtonian mechanics and its underlying calculus that considered no uncertainty was replaced with statistical mechanics that could be described by a probability theory—a theory which could capture a form of uncertainty arising from random processes. After the development of statistical mechanics, there has been a gradual trend in science during the past century to consider the influence of uncertainty on problems, and to do so in an attempt to make models more robust, in the sense that credible solutions are achievable and at the same time quantify the amount of uncertainty.

Of course, the leading theory in quantifying uncertainty in scientific models from the late nineteenth century until the late twentieth century had been probability theory. Probability has a long history of use, dating back to the 1500s, to the time of Cardano when gamblers recognized the rules of probability in games of chance. By the time of Newton, physicists and mathematicians were formulating different interpretations of probability consist with its axioms and operational theory. The most popular ones remaining today are the relative frequency probability and the subjectivist or personalistic probability. The latter development was based upon Rev. Thomas Bayes' (*circa 1763*) powerful theorem for conditional probabilities. Subjectivist probabilities specified that a human's *degree of belief* or *willingness to bet* was a mathematically coherent interpretation of probability within its theoretical construct.

However, the gradual evolution of the expression of uncertainty using probability theory was challenged, first in 1937 by Max Black (Black, 1937) with his studies in vagueness, then with the introduction of fuzzy sets in 1965 (Zadeh, 1965). Zadeh's work had a profound influence on the thinking about uncertainty because it challenged not only probability theory as the sole representation for uncertainty, but the very foundations upon which probability theory was based: classical binary (two-valued) logic (Klir & Yuan, 1995).

The twentieth century saw the first developments of alternatives to probability theory and to classical Aristotelian logic as paradigms to address more kinds of uncertainty than just the random kind. Jan Lukasiewicz developed a multi-valued, discrete logic (*circa* 1930). In the 1960's Arthur Dempster (Dempster, 1968) developed a *theory of evidence* which, for the first time, included an assessment of ignorance, or the absence of information. In 1965, Lotfi Zadeh introduced his seminal idea in a continuous-valued logic that he called *fuzzy set theory*. In the 1970s Glenn Shafer (Shafer, 1976) extended Dempster's work to produce a complete theory of evidence dealing with information from more than one source, and Zadeh (1973) illustrated a possibility theory resulting from special cases of fuzzy sets. Later in the 1980s other investigators showed a strong relationship between evidence theory, probability theory, and possibility theory with the use of what was called *fuzzy measures* (Klir & Wierman, 1998), and what is now being termed *monotone measures* (Ross, 2010).

4. General theories of uncertainty

Since 1965, there have been numerous developments in mathematical uncertainty theories, such as possibility theory, evidence theory, and the theory of imprecise probabilities, to name a few. These theories can be collectively referred to using the title from the Klir and Wierman (1998) book, *Generalized Information Theories* or *GITs*. A partial list of mathematical uncertainty theories includes:

Probability theory (Dempster, 1968 and Feller, 1968).
Zadeh fuzzy sets and fuzzy logic.
Possibility theory (Dubois & Prade, 1988).
Dempster–Shafer evidence theory;
Imprecise probability theory (Walley, 1991).
Random Intervals (Joslyn & Booker, 2004).

5. Inference uncertainty—A new perspective on uncertainty

A new perspective on uncertainty assessment is emerging from the old concept of making inferences. *Inference* is defined as the difference between what is measured (the observable quantity) and what is desired (the unobserved quantity); said in another way, inference is the difference between what we "want" to know and what we "can" know. There are many inferences that have to be made because we cannot measure directly that which we "want to know". Common examples of making inferences include:

We **infer** the temperature of the Earth in ancient history using tree ring measurements, solar activity (sunspot cycles), ice core samples and geologic features. We impute a source of uncertainty from this inference.

We **infer** the yields of the Hiroshima and Nagasaki bombs by analyzing the radioisotopes in soil samples, estimating fireball size and blast damage, and tracking population illness and lethality. Uncertainty in the yield results from the uncertainty of radiological and blast effects.

We **infer** the safety of a building to an earthquake by measuring small-scale test results of materials and structures under laboratory quake-like conditions. Uncertainty is the result of scaling small-scale tests to the full-size earthquake effects.

We **infer** the diets of the first (native) Americans by measuring carbon isotopic amounts in skeletal remains. Uncertainty is partly from the radiological analysis and the assumptions made about the climate at that time.

We **infer the** likelihood of a terrorist attack that takes down major power grids by relying on knowledge of terrorists' behaviors/plans, and risk/reliability/security studies of our electrical infrastructure. We impute, perhaps incorrectly, logical reasoning to terrorists.

Often these inferences become common practice and their effects are not considered in analyses or in drawing conclusions. Recent research investigates the effects of making inferences and the uncertainties induced by those (Booker & Ross, 2011). Thus, *inference uncertainty* is defined as the uncertainty induced by the act or process of deriving a conclusion about an entity that is unmeasured or unavailable based on what one has been or can be observed and measured or made available.

Langenbrunner, Booker, Ross, and Hemez (2010) identified and proposed methods for estimating the following inference types and their inference uncertainties in engineering and scientific problems:

- Predictive: The process of inferring the future from the past; a forecast.
- Statistical: The process of inferring the whole population from a representative sample.
- Validation: The process of inferring that the model/code matches the data. This is also known as fidelity-to-data or goodness-of-fit.
- Analogical (similarity): The process of inferring degree of similarity between applications or quantities (e.g., variables). System behavior is often inferred from small scale or sub-system tests.
- Proxy (unobservable): The process of inferring an unobservable quantity from an observable one.

Considering many of the different kinds of uncertainties as arising from making inferences provides a much-needed common framework for the standardization of uncertainty assessment. Therefore, the study of inference uncertainty represents a new perspective on uncertainty quantification.

6. The necessary steps in uncertainty assessment

Using recent and historical developments about uncertainty, some basic steps are outlined for its assessment, including quantification. We will discuss the necessary steps in conducting an uncertainty assessment, and then illustrate some of them in a case study. The steps can be enumerated as: (i) identification, (ii) classification, (iii) uncertainty inventory, (iv) quantification methods, (v) combination of uncertainties, and (vi) management of uncertainties.

6.1. Identification

Uncertainty can be thought of as being the inverse of information. Information about a particular engineering or scientific problem may be incomplete, imprecise, fragmentary, unreliable, vague, contradictory, or deficient in some other way. Acquiring more information about a problem tends to produce less uncertainty about its formulation and solution. Problems characterized by little information are said to be ill-posed, complex, or insufficiently known. These problems are imbued with a high degree of uncertainty.

Uncertainty can be manifested in many forms: it can be *fuzzy* (not sharp, unclear, imprecise, approximate), it can be *vague* (not specific, amorphous), it can be ambiguous (too many choices, contradictory), it can be *non-specific* (data are intervals instead of point estimates), it can be of the form of *ignorance* (dissonant, not knowing something), or it can be a form due to natural *variability* (conflicting, random, chaotic, unpredictable). Zadeh (2002) posed a simple example of a person's statements about when they shall return to a current place in time identifying these forms. The statement "*I shall return soon*" is vague, whereas the statement "*I shall return in a few minutes*" is fuzzy. The former is not known to be associated with any unit of time (seconds, hours, days), and the latter is associated with an uncertainty that is at least known to be on the order of minutes. The phrase, "*I shall return within 3 minutes of 5pm*" involves an uncertainty which has a quantifiable imprecision; probability theory could address this form. "*I will return between 5 and 15 minutes*" is a kind of non-specificity; we only know the interval in which we will return. "*I will be back sooner or later*", is a kind of ambivalence, and "*I don't know when I will return*" is a kind of ignorance.

Identification of types and sources of various uncertainties inherent in a problem is the first step. Uncertainty in physical science and engineering applications arises from observation, measurement, recording, poorly understood initial conditions, random effects, uncontrollable effects and unknown effects. However, there are additional sources of uncertainty from incomplete information, lack of knowledge, vagueness, and ambiguity. Sources for these kinds of uncertainties include physical models, mathematical models, statistical models, computational models, currently known theory, decisions, interpretations, extrapolation, interpolation, prediction, indirect observable quantities, inferences being made, and conflicts among data, models, tests and experiments.

6.2. Classification

Various categorization and classification schemes exist to accommodate different kinds of uncertainties. No universal taxonomy or standard definitions are available; therefore it is up to the analyst to construct a scheme that suits the application. Some example classifications are: (i) aleatoric; uncertainties from random or stochastic processes, (ii) epistemic; uncertainties from lack of knowledge, (iii) irreducible; natural variability (which cannot be reduced, but only quantified), (iv) reducible; due to lack of specific information, knowledge (which can be reduced with acquisition of more information), and (v) inference uncertainty.

6.3. Uncertainty inventory

An *uncertainty inventory* should be considered and performed in the earliest stages of uncertainty assessment and quantification. An uncertainty inventory is an organized set of all the information, statements, and questions relating to the different kinds of uncertainties in a given problem/system. The uncertainty inventory includes choices, such as which uncertainty mathematical theory(ies) is (are) appropriate for characterizing each uncertainty, and what data, information, and knowledge are available to characterize each uncertainty. The format of the uncertainty inventory could range in complexity from a list or table, to an interactive, relational knowledge base. More specific details on the process of doing an inventory can be found in Langenbrunner et al. (2009), and specific questions that need to be posed to assist in this inventory were presented recently in Booker and Ross (2011).

6.4. Quantification methods

Construction of the uncertainty inventory involves characterizing the different uncertainties within a problem. Based upon the answers to the questions above, some known UQ metrics and/or mathematical uncertainty theories (GITs) may be deemed appropriate for use. A brief listing includes:

- Quantifying Margin and Uncertainty (QMU) which can be simply defined as margin divided by uncertainty, used for physical system certification and/or qualification Sharp and Wood-Schultz (2003),
- Engineering Index, used for certification and/or qualification to track degradation under uncertainty (Booker et al., 2006),
- Chi-squared goodness-of-fit statistic used for validation and inference estimation (Conover, 1999),
- D_n distance metric, a more general goodness-of-fit metric (Langenbrunner et al. (2007) used for validation and inference.
- Roache's Grid Convergence Index used for estimating grid convergence error in the verification of computer calculations (Roache, 1988),
- Probability density functions, used to characterize distributions of probabilistic uncertainties. Examples include random noise, data dispersion, parametric uncertainty, and uncertainties in Probability Risk Assessments (PRAs),
- Kullback and Leibler (1951) and Jeffreys (1961) statistics used to compare probability density functions,
- Kolmogorov-Smirnov tests non-parametric tests used to compare empirical-based probability distributions (Conover, 1999),
- Possibility distributions used to characterize distributions of possibilistic uncertainties (Ross, 2010).

One of the first modeling tools that focused on assessing uncertainty was Probabilistic Risk Assessment (PRA). PRA (see US NRC, 1975) identifies uncertainties using probability theory, but it lacks formal procedures for quantifying uncertainties that are not probabilistic. In the past decade, risk assessment tools have expanded to include other mathematical theories of uncertainties. For example, possibility theory has used for assessing the risk of terrorism (Darby, 2004). The advantage of using possibility theory over probability theory is the axioms for possibility are more general, less restrictive, than those for probability theory. Possibility is better suited to rare event estimation as evidenced by the common expression "*that which is possible but not probable*." The disadvantage of using an alternative to probability theory is that most experts and decision makers will be unfamiliar with its interpretation and use; this argument, however, will not be sustainable as the alternative theories are now being taught in the universities.

Probability theory's popularity stems from its long history, and many analysts, experts and decision makers have at least heard about it. Some may even understand it (although far fewer than think they do). As Bruno de Finetti (1974) proclaimed in his book on probability, "*Probability does not exist; it is a subjective description of a person's uncertainty,*" and that "*The calculus of probability can say absolutely nothing about reality*."

6.5. Uncertainty combination

It is understood by systems scientists that reduced doubt and reduced uncertainty is equivalent to increased understanding and increased confidence. These notions are opposites, or inversely related; as one increases, the other decreases. This is our concept of confidence, as defined in Section 2. If an analyst is to conduct a test or a numerical simulation for the purpose of predicting similar phenomenon in the full-scale or real world case, they would want to express not only the expected result of the test or simulation, but also the degree of confidence in it. Being able to express a level of confidence in a prediction is analogous to being able to express the level of uncertainty in that prediction.

However, confidence–and its inverse–uncertainty, are relative concepts that must be defined with respect to some standard. An example of a statistical confidence interval would be "*We are 95% confident that the true stress in a mechanical system is within 100 MPa of the predicted stress*." The quantified uncertainty expressed as "*within 100 MPa*" can be reduced by additional repeatable tests under exactly the same test conditions. A different confidence statement might be: "*We are highly confident that the true stress in the system will be 1200 MPa*." In this case the task is to quantify the phrase "*highly confident*", and determine a standard against which this uncertainty is compared. Some standards currently exist, like the 95% relative to a 0–100% scale in the confidence interval. If the confidence for the latter statement is expressed as a possibility, then we must know the range between what is impossible, and what is certain (the complement of impossible).

Previous work, Ross (2003), proposed a new standard by hypothesizing that all uncertainty should scale between two extremes, or boundary conditions, on uncertainty, i.e., between the case of *no uncertainty* and the case of *maximum uncertainty*. If we make a prediction on the response of some mechanical system and the level of uncertainty that we express in that prediction is *close* to the extreme of *no uncertainty* we can say that we are "highly confident" in that prediction. On the other hand, if we are *closer* to the other extreme, the case of *maximum uncertainty*, then we can say that we are not very confident in the prediction. Of more importance, however, is the fact that we can develop a "metric of confidence" that will scale linearly with our quantified level of uncertainty and, in a mathematical sense, measure the degree of *confidence*. We call this measure of uncertainty, *total uncertainty*. We use the term *total uncertainty* because the research has shown a way to combine different types of uncertainty, for example combining probabilities and possibilities; hence the term *total* refers to the aggregation of all forms of uncertainty. It should be noted that Klir & Yuan, 1995, previously proposed a metric for combining possibilistic and probabilistic uncertainty. Their metric is a combination of the measures of: non-specificity, possibilistic strife, and probabilistic conflict.

In assessing *total uncertainty* our initial research attempted to distinguish between two types of uncertainty. The first is the *nat-*

ural variability of things (e.g., minute differences due to manufacturing processes) and cannot be reduced. Another type of uncertainty is that due to a lack of specific information, called *non-specificity,* to distinguish it from natural variability. Non-specificity can be reduced–with the acquisition of more information. Non-specificity can result from ignorance, from scare data, from poor control, from misleading data, and from unknown biases. In our early work, we assessed *total uncertainty* as a combination of these two forms of uncertainty, i.e. *variability* plus *non-specificity*. We quantified *variability* using probability theory, and quantified non-specificity using possibility theory. However, recently we have identified many more forms of uncertainty, most notably inference uncertainty, and we have also identified various characterizations of inference uncertainty; e.g., scaling uncertainty, proxy uncertainty, analogical uncertainty, and others (Langenbrunner et al., 2010)

6.6. Uncertainty management

With no standard for an all-encompassing solution about how to combine all kinds of uncertainties, the key to uncertainty assessment may be to *manage uncertainty*. The first step to managing uncertainties is becoming aware of the uncertainty types as discussed. The uncertainty inventory aids in determining what data, knowledge and theory is available. Choosing appropriate GITs and uncertainty metrics, including methods for combining uncertainties not only account for what is available but also must be done considering how the uncertainties will be understood and interpreted by decision makers. These steps are also part of managing uncertainty.

7. A process for uncertainty management

Uncertainty management involves understanding the relationship of changes in uncertainty (called *robustness*) to other assessment activities, such as goodness-of-fit and predictive capability. There are tradeoffs between these three such that it is not possible to simultaneously optimize them all, as demonstrated by Hemez and Ben-Haim (2004). For example improving predictive capability can adversely affect robustness to uncertainty.

We have identified a process for doing uncertainty management; it is a computational paradigm to combine various forms of uncertainties, such as inference uncertainties, in what we call the *4-box approach* to information integration, shown in Fig. 1 (Langenbrunner et al., 2008). The three gray boxes contain data, knowledge and information that can be combined to solve a problem in the upper right (white) box which has little or no data. The combination of both the data and the uncertainties is done using a weighting scheme. The approach incorporates Saaty's AHP methodology (Saaty, 1980) to solve the age-old problem of how to determine the weights for combining the uncertainties contained inside the 4 boxes and for those between the boxes that relate to making inferences. We will illustrate this paradigm with a simple example using what is called Quantification of Margin and Uncertainty (QMU).

Quantification of Margins and Uncertainties (QMU) is a simple idea used to express confidence in the ability of a manufactured system to perform according to prescribed specifications, or according to design criteria. Fig. 2 shows a schematic of the QMU metric. In its most simple manifestation, it is a confidence ratio (CR), given by,

$$CR = \frac{M}{U} \tag{1}$$

The margin, M, is defined as the distance between a systems nominal operating condition, and the threshold of failure of the system. For example, suppose that a system relies on water to be in its fluid state to be able to operate normally (e.g., as a coolant). Water is a fluid except when it is subjected to a temperature in excess of 212 °F, when is becomes a gas (water vapor) or when it is subjected to a temperature less than 32 °F, when it becomes a solid (ice). Suppose in a "normal operating state" the water temperature is 100 °F. It has a margin from the high threshold of 112 °F and a margin from the low threshold of 68 °F. The uncertainty, U, in Eq. 1 is defined as the uncertainty in assessing the actual operating state of the system. For our example, suppose the uncertainty in the gauge used to measure the temperature of 100 °F is ±10 °F. The uncertainty in a component of a system arises because of (i) an inability to measure a quantity precisely with a specific diagnostic, (ii) an inability to measure the desired quantity of interest using any diagnostic, or (iii) an unknown-ness in what quantity to measure at all. Uncertainty can also arise because we do not know how to assess the interaction of one component with another, to assess the influence of the behavior of one component on another, or to test the assembly of components in any way. Uncertainty also arises because of the processes of aging and other forms of deterioration, such as degradation due to environmental influences like corrosion and oxidation, or influences due to normal operation like cyclic phenomena. Most materials are newly manufactured when put into operation for the first time; and when these materials are new we know their properties, and we know the strengths and rigidities of their connections. However, with time materials decay, connections become loose, and deteriorating environmental and operational influences increase with time.

Usually, a manufactured system is broken down into a number of interacting components, each one of which can have its CR as-

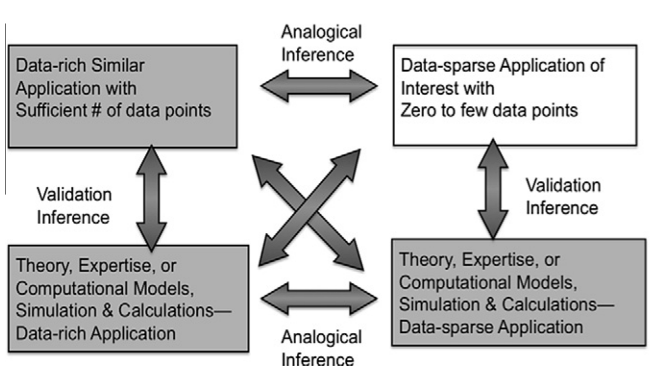


Fig. 1. *4-box approach* of information integration.

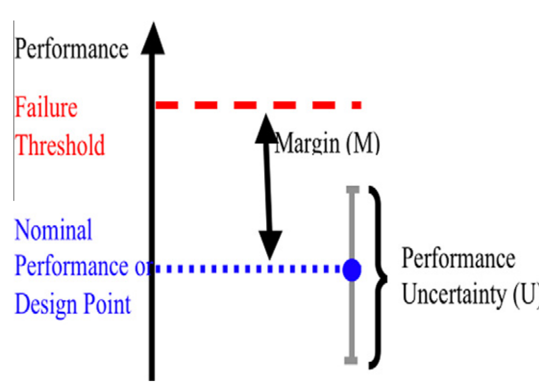


Fig. 2. QMU as defined by margin over uncertainty; this shows a threshold that is a "performance gate".

sessed by numerical or other means. If all the CRs for the components of the (series) system are assessed to be greater than one (CR > 1) then the system is said to be acceptable (or safe) for use in its design environment. If any of the components has a CR < 1, then the system is said to be unacceptable (or unsafe) for operation, and work needs to be done to repair or upgrade that particular component. To continue with the example on water, the CR for the high threshold is M/U = 112/10 = 11.2; and for the low threshold, the CR = M/U = 68/10 = 6.8. Both are greater than one, so the "system" containing water is said to be acceptable at 100 °F. Note that the temperatures of the system at which it would just become unacceptable (too close to being water vapor or ice) would be at 202 °F and 42 °F. Eq. (1) gives the most basic definition of QMU.

Engineering systems are not simple systems like the coolant example, above. Systems like airplanes or automobiles are made up of many components, each of which is critical to the operation of the system. And, each of these components may or may not be easy to test or to assess in terms of performance. Some components may perform quite well when new, but may not operate at all, or in a subnormal way due to the effects of aging. Some or all of the components in a system may behave differently to environmental affects like temperature fluctuations and ambient or natural vibrations, or to the effects of an operational environment. Because many engineered systems are designed to perform operationally for a period of many years it is important that their capacity to perform be monitored periodically. And, when monitored, it is important to be able to diagnose the problem, in order to repair the component or subsystem. The QMU value (CR ratio) can be very helpful in identifying what tests or analyses to perform to upgrade or repair the components in a system.

While the QMU value seems simple at first glance, there are some nuances that produce challenges to determining the CR. For example, suppose the uncertainty we are interested in is not related to the actual performance state of the system, but to the lower bound of the normal operating range. Fig. 2 shows the uncertainty and margin in the context of a performance gate. How is this uncertainty measured, particularly if this region is in a difficult environment to reproduce? Or, suppose our failure threshold is not known with impunity; how do we determine the uncertainty in this quantity if it is difficult to measure or is subject to ambiguity, or some other form of non-numeric uncertainty? In our water coolant example, the actual physical thresholds that can cause failure are temperatures "near boiling" or "near freezing". Water gets more viscous and less fluid-like at temperatures of 33 °F or 34 °F, for example. The additional nuances for the thresholds and the regions of uncertainty make the determination of the CR more difficult, but nonetheless still important in terms of test planning.

The assessment of the CR for a system as it ages is actually a better metric for determining the time to repair, or for making test planning decisions, than is the reliability of a system. Fig. 3 depicts the assessment of a system as it changes over time. In this figure, the IAC is the *initial assessed condition* of the system when new, and the AC(t) is the *assessed condition* of the system at some later time, t. If the PT is the *performance* or operating failure *threshold*, then we see that the system is safe until the dark curve intersects with the PT line. This intersection marks the first time that the reliability of the system would drop below 1.0. If we wait to make a repair decision, or a decision to perform certain tests when the reliability drops below 1.0, then we have missed some valuable time to make adjustments. If we monitor the decreases in CR with time ($t = 1$, $t = 2$, and so forth) we will know that repairs, or testing, is warranted long before the reliability drops below 1.

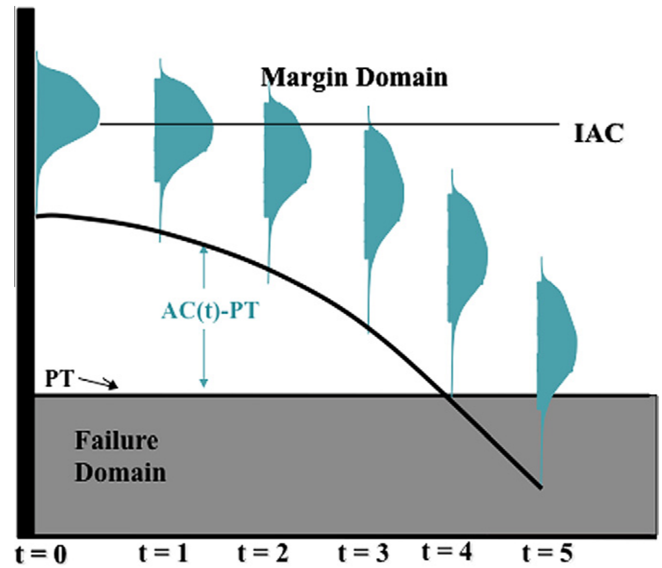


Fig. 3. System monitoring when the performance of the system changes over time.

8. Case study: Magnesium-alloys

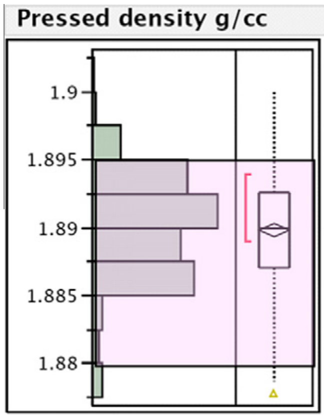
Magnesium (Mg) alloys are seeing increasing use in commercial and industrial applications because they are metals that are: lightweight, low density (2/3 that of aluminum or 1.70 gm/cc to 1.90 gm/cc), good high temperature mechanical properties, and good resistance to corrosion. The density vs. temperature relationship is especially important in understanding formability issues for use in commercial products, so our case study is focused on issues of density and ductility (formability) under various temperatures.

In trying to understand the theoretical relationship between the formability potential of Mg-alloy and the density, we have discovered how little is known about the fundamentals of the Mg-alloy material. For example, a very special property of Mg-alloys is their usefulness in damping. Their damping capacity is 4 times that of aluminum and 2 times that of steel. We need to understand better the theory relating damping to density. The two drawbacks of Mg-alloys are: cost, and low formability at room temperature. Improvements in the understanding of the influence of density on formability can have significant influence on some new technologies, such as batteries and fuel cells. In developing information to assess the average density of the Mg-alloy, 253 tests were conducted for pressed densities. A manufacturing pressing process influences density, which is turn gives an indication of formability. The statistical information for those tests is shown in Fig. 4.

In this graphic we can see the overall average density from all the tests, and statistical quantities like the standard deviation and the various quantiles. Of special note is that the manufacturing specification for the density for all the samples was in the range 1.885 gm/cc to 1.895 gm/cc. The tails of the data distribution show that many of the samples were outside the range of the specification. The light pink box on the graphic indicates our choice of the gates for this problem by highlighting the specification range.

8.1. QMU of Mg-alloy density

The density of Mg-alloy, ρ, has a direct role in determining whether there will be improvements in formability. If the density is too high, then there is low formability, and if the density is too low, then there the Mg-alloy is too ductile for many applications involving vibrations. Fig. 5 indicates the values to be used in our



Quantiles

100.0%	maximum	1.9
99.5%		1.89946
97.5%		1.896
90.0%		1.89466
75.0%	quartile	1.8926
50.0%	median	1.89
25.0%	quartile	1.887
10.0%		1.885
2.5%		1.882
0.5%		1.87804
0.0%	minimum	1.8778

Moments

Mean	1.8898399
Std Dev	0.0037436
Std Err Mean	0.0002354
Upper 95% Mean	1.8903034
Lower 95% Mean	1.8893764
N	253

Fig. 4. Overall average density from 253 tests for formability of Mg-alloy.

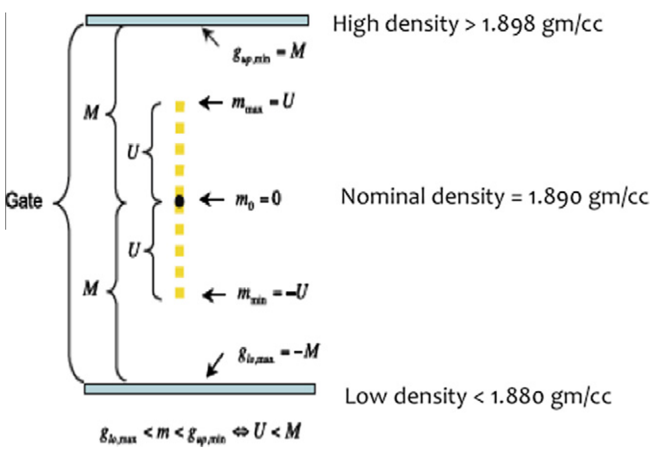


Fig. 5. Performance gate for the density of Mg-alloy.

QMU analysis for the high-density gate and the low-density gate of the Mg-alloy.

We now have some information to calculate a QMU quantity using just the experimental data from the pressed density tests. If we presume that the uncertainty in the density of the Mg-alloy, U, is governed by a 95% probability interval on the histogram data from above, then the CR=M/U for the top and bottom gates, are given by the following:

8.1.1. Top gate

$$\begin{aligned} QMU+ &= M+/U+ = (1.898 - 1.8898)/(1.896 - 1.8898) \\ &= 1.32 > 1 \end{aligned}$$

8.1.2. Bottom gate

$$\begin{aligned} QMU- &= M-/U- = (1.8898 - 1.880)/(1.8898 - 1.882) \\ &= 1.26 > 1 \end{aligned}$$

From these two calculations, since CR > 1, both margins (M) exceed their associated uncertainties (U) and the Mg-alloy material is said to illustrate an acceptable level of formability. We should remind the reader, that the density values used to make this assessment of the CRs is based only on "normal" conditions of formability. If on the other hand, we simply change our definition of U to a probability interval on the histogram data to 99%, then we see that the situation involving the CR dropping below a value of 1.

8.1.3. Top gate

$$\begin{aligned} QMU+ &= M+/U+ = (1.898 - 1.8898)/(1.89946 - 1.8898) \\ &= 0.85 < 1 \end{aligned}$$

8.1.4. Bottom gate

$$\begin{aligned} QMU- &= M-/U- = (1.8898 - 1.880)/(1.8898 - 1.878) \\ &= 0.83 < 1 \end{aligned}$$

Now, the QMU values for density are NOT within the normal formability range for the material. As before, the assessment is valid only for normal ranges of formability.

What this brief analysis reveals is that QMU is sensitive to the definition of how we express U. If we want U to express most, or all, of the tails of a histogram we can get densities from tests that are outside our gates. In the example using a 99% probability interval we get 2.4% of the data being outside the gates. Moreover, some of the data might also be beyond the specifications for density. As seen in Fig. 5, if we define our acceptable density specification in the range 1.885–1.895 gm/cc, then 20% of the density test data falls outside this range. Our specification range (1.885–1.895 gm/cc) is contained within the range of the gates (1.880–1.898 gm/cc), which is typical for most engineered systems. But, we define margins in terms of the maximum or minimum allowed values for a variable.

The QMU analysis reveals that the acceptable formability of the Mg-alloy is sensitive to our definitions of the margin in terms of 95% of the data or 99% of the data. This sensitivity shows that this particular component of a system is NOT robust to uncertainty. However, the analysis did reveal that the theory and issues contributing to this uncertainty is weak. Even though there were over 200 tests performed on the Mg-alloy using a pressed process for density, the data do not reveal the reasons for the weak theory. Because of this we should examine other sources of data (un-pressed process), computational information (3D simulations), and knowledge (expert judgments) to be in a better position to manage the uncertainties.

8.2. Quantifying uncertainties of Mg-alloy density

As mentioned earlier, Saaty (1980) developed a process which computes a similarity relationship between pairs of entities in the 4-box paradigm. The similarities speak to the level of uncertainty in the impact of information that is contained "within" the

4 boxes. For example, as shown in Fig. 6, suppose we have information in the green box on some 1D coupon tests done on samples of steel, and we have information in the gold box on 3D simulations on a system component made from this same steel. Then an interesting question is: if the 3D simulation code simply models the tensile strength of the steel in a 1D loading state, what uncertainty in the 3D simulation is introduced by not using 3D coupon test results? Such an assessment of this impact of using 1D tensile test values in a 3D (e.g., triaxial) material model would address one form of uncertainty that exists between the information content of the green box and the gold box in the 4-box paradigm. What is important is that an "uncertainty-inventory" be conducted by experts and analysts to identify the underlying uncertainties in the QMU process, due to assumptions, lack of information, analytic and experimental judgments, that impact all 4 boxes.

Such an inventory was done for our Information Integration approach (4-box paradigm). This inventory is a collection of known sources of uncertainty in attempting to understand the formability of Mg-alloy as a function of density. The list given in Table 1, enumerates the number of phenomena of interest that we feel generates most of the uncertainty in the total collection of information that we can use to address this question. It is this list that will guide us in determining the similarities using Saaty's AHP process (1980).

Table 1 provides an example of the reasoning used by an expert in developing the similarities among the issues above, which would be useful in quantifying the *analogical inference uncertainty* that exists between the green box (small scale tests) and the red box (3D tests). The similarity values range on a scale of 1 to 9, with 1 meaning *completely similar*, and a 9 meaning *completely dissimilar*.

If we take all the assumptions, and their associated similarity values (obtained from an expert), we get the results shown in Table 1 that represents the combined similarities for the green box and the red box in the 4-box paradigm.

If we sum all these similarity values, we get 66, and if we divide this number by the total of 17 similarity values, we get an average similarity value of 3.89. This is the similarity value that represents the magnitude of the analogical inference using Saaty's scale; a large value means that the information from the green box on Mg-alloy density is not very influential on the information content (regarding formability) in the red box.

If we repeat this process for the other 5 inference arrows for the 4-box paradigm (i.e., for similarities between green-blue, green-gold, blue-red, etc. as seen in Fig. 6), we get the total scores and average similarity for the six inferences (see the values in Fig. 7). Note that even expert judgment (the information in the blue box) can be used in our 4-box paradigm. Sometimes the expert judgment is the most useful information in a problem that is "data-sparse".

Table 1
Combined similarities for the green box and the red box in the 4-box paradigm.

Phenomena of interest	Similarity values
Symmetry effects, 3-D effects	8
Boundary conditions	2
Failure criterion	4
Pre-formable material	1
Physics not modeled	4
Measurement error	7
Manufacturing variability	7
Modeling parameters	2
Failure behavior in formability	5
Pressed/unpressed density	5
Material creep	5
Temperature effect	2
Particle size; crystal morphology	1
Number of diagnostics	3
Pressing anisotropy	1
Analytic chemical Composition	1
Effect of formability	8
Sum	66
Average	3.89

The next step in the uncertainty quantification process is to construct a 4x4 matrix that represents the similarities between the 4-boxes. We use the Saaty (1980) process called the inverse method, where the elements in the upper triangle of the matrix are the values shown in the figure, above, and the lower triangle are the inverses of the values in the upper triangle of the matrix. Since the matrix represents similarity, then the diagonal elements equal unity, necessarily (an entity is completely similar to itself).

We perform a singular-value-decomposition (SVD) on this matrix to find the singular values and the weights associated with these values. The largest singular value is the only one of interest, since this determines which of the 4-boxes contains the most important information in terms of the similarities of the information from among all the boxes. The singular values, in ranked order, are:

Singular values = 14.98, 2.89, 1.34, 0.07

These values are used in conjunction with the singular vectors associated with the largest singular value (14.98) to develop weights to be applied to the information in each of the 4 boxes. We use a formulation for the weights from a recent paper on this subject (Mamat & Daniel, 2007). The final weights for each box are shown in Fig. 7, and the associated mean values, m, and standard deviations, σ, for the uncertainties calculated for each box are now weighted by the weight-values in each box to determine the overall inference uncertainty, as shown in calculations, below.

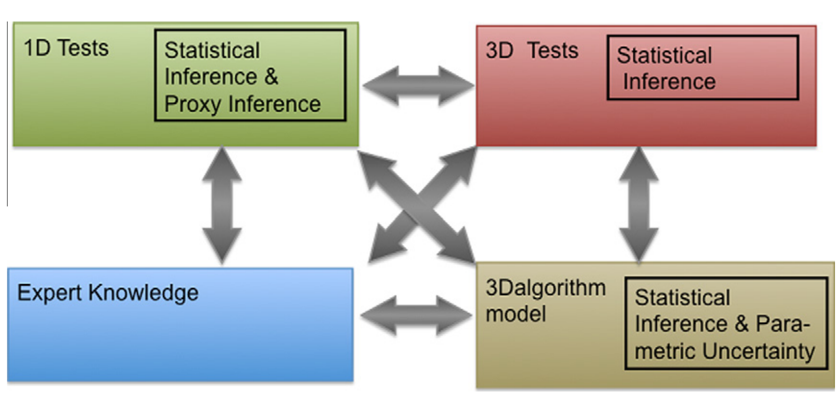


Fig. 6. Example *4-box approach* used for example on steel strength prediction.

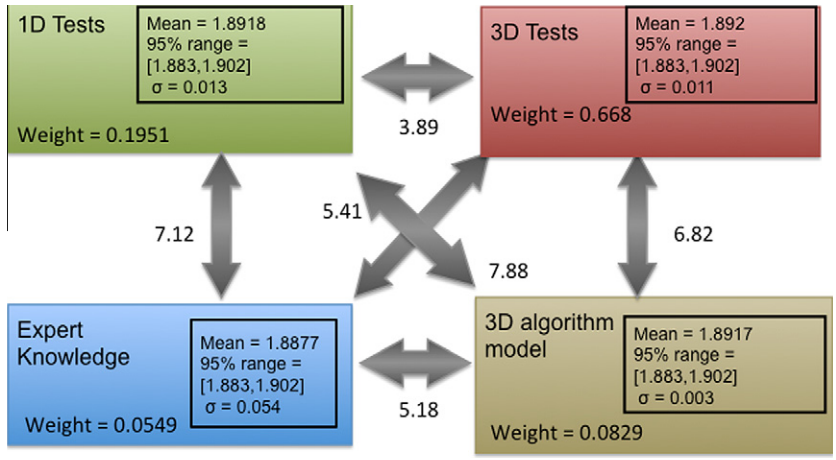


Fig. 7. *4-box* results for Mg-alloy case study.

$$m = \sum_{i=1}^{4} w_i \rho_i = 0.1951 * 1.8918 + 0.688 * 1.892 + 0.0549 * 1.8877 + 0.0829 * 1.8917 = 1.8898\ \text{gm/cm}^3$$

$$\sigma = \sqrt{\left(\sum_{i=1}^{4} w_i \sigma_i^2\right)} = 0.011\ \text{gm/cm}^3$$

Now, all of the foregoing discussion assumed that the densities were measured or calculated under assumed normal operating temperatures and normal ranges of operating densities. How does the formability of the Mg-alloy get altered when the temperatures are room temperature (normal) or very hot, and when the Mg-alloy densities are especially high or low? To answer this, we developed a rule-based linguistic system obtained from experts.

8.3. Evaluating formability

We do understand the relationship between density and temperature of the Mg-alloy and their combined effect on the potential ability to improve formability. These relationships can be cast into the form of linguistic rules. For example,

- IF the Temperature is Hot, and the Mg-Alloy density is Low, THEN the Mg-alloy formability will Probably Fail; or
- IF the Temperature is Medium, and the Mg-alloy density is High, THEN we need to Check the Failure Potential; or
- IF the Temperature is Medium, and the Mg-alloy density is Operating, THEN the formability is OK.

When we refer to the Failure Potential, we are referring to certain combinations of the Temperature and Density where such conditions like planar anisotropy and intensity of texture must be checked first to see if we expect the formability to improve or to fail. These rules describe a functional relationship: a relationship between two inputs (Temperature and Density) and one output (Failure potential of formability), given in a linguistic form. These linguistic rules can be organized into a matrix, generally described in the literature as a *Fuzzy Associative Memory*, or a FAM matrix (Ross, 2010). We use the FAM terminology since we model the linguistic variables such as *High, Low, Hot, Medium, Operating*, etc. as fuzzy membership functions. The rules are collected in the FAM matrix used in this study, and this is shown in Table 2.

The membership functions to describe the temperature input are defined on a cardinal scale in units of Celsius and the membership functions to describe the density input are defined on a cardinal scale of units of gm/cc. The membership functions for the output, Failure Potential, are defined on an ordinal scale of values (μ) on the unit interval $[0,1]$. The membership functions for the inputs and the output are shown in Fig. 8.

Since we have the FAM and the membership functions, we have what is called a "Fuzzy System" which describes the relationship between linguistic descriptions of the inputs and outputs. Often, linguistic descriptions of fundamental variables is all we have from experts, and this fuzzy system can actually be implemented into a simulation using certain logical connectives and implication operators; a rule is an implication (inference), and the inputs for each rule have logical connectors, AND or OR.

So, let's say we have a specific density and a specific temperature for a specific piece of Mg-alloy. Using our fuzzy system we can perform a simulation and determine the failure potential of the formability of Mg-alloy, or determine the failure potential of many different pieces. Fig. 9 illustrates the results of this simulation. As seen in this figure, the failure potential of formability varies considerably with different combinations of the formability temperature and the density. The details of a fuzzy system simulation are described in many references on fuzzy systems (see Ross, 2010).

8.4. Case study summary

The foregoing case study illustrates several ideas. First, when we want to identify and then quantify different uncertainties we

Table 2
Fuzzy Associative Memory (FAM) table for density and formability temperature.

		Temperature		
		Hot	Medium	Normal
Density	Low	Probable failure	*Check*	Ok
	Operating	*check*	Ok	*check*
	High	Ok	*Check*	Probable failure

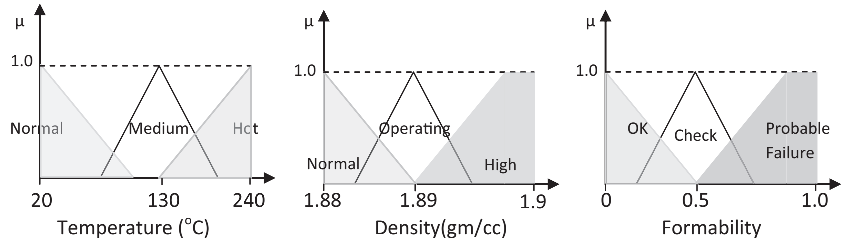


Fig. 8. Membership functions for Mg-alloy formability case study.

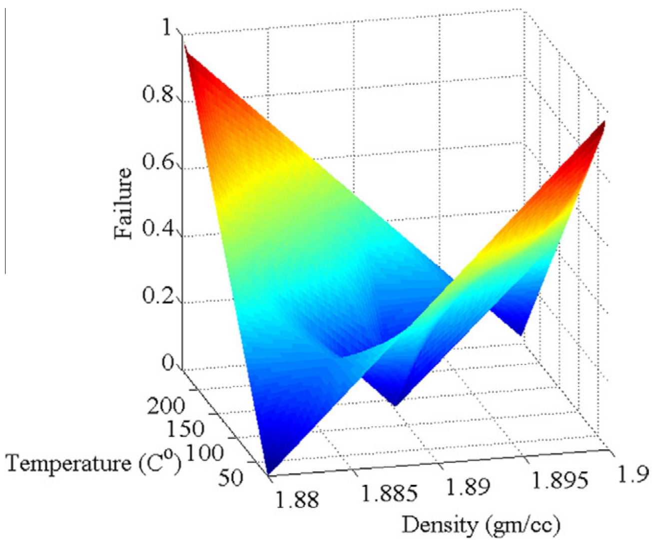


Fig. 9. Fuzzy system simulation of Mg-alloy case study.

may need to use a protocol, called uncertainty inventory, and we may need to use different uncertainty quantification theories, like those in GIT. The case study addressed here only used probability theory to quantify and mean and variance of test and simulation data (see Fig. 7). We can then assess the inference uncertainty among the various pieces of information (the various boxes in Figs. 6 and 7) by using a process developed earlier by Saaty (1980). The inference uncertainty gives us a better way to assess the errors introduced by ignoring some sources of data and the errors introduced by not assessing the interaction of the various pieces of information (the 4 boxes). Fuzzy systems can be used in the process of assessing the inference uncertainties by adding the capability to incorporate linguistic knowledge from experts. The final step in the process involves the management of this uncertainty using other metrics, such as the TU (total uncertainty) metric developed by Ross (2003). The enlargement of this case study to show that final step will be the subject of another paper.

9. Summary

For most of the past 400 years most of what we thought of uncertainty was guided by the principles and axioms of probability theory. Some early-20th century theorists and others talked about non-binary logics, logics which are not constrained by the *excluded middle axioms*, and types of uncertainty which clearly would not be assessed by probability theory—like vagueness. After the seminal work of Lotfi Zadeh in 1965 on fuzzy sets appeared, the uncertainty community has not been the same and has changed significantly. Now, we realize that the largest uncertainty introduced in our predictions or our assessments of reliability or performance is the one introduced by our choice of a model–we have called this the uncertainty in inference. Quantification of such an uncertainty is not easily done with probability theory, and the newer theories can be very useful in terms of assessing the magnitude of the inference uncertainty and in terms of managing this uncertainty to assist in prediction or test planning.

Acknowledgments

The authors are grateful for the collaboration of Dr. James Langenbrunner, Los Alamos National Laboratory (LANL), whose seminal works in the area of uncertainty quantification are referenced in many of the citations in this paper. The authors are also grateful for the discussions of other colleagues at LANL, especially Don Quintana.

References

Black, M. (1937). Vagueness: An exercise in logical analysis. *International Journal of General Systems, 17*, 107–128.

Booker, J., & Ross, T. (2011). An evolution of uncertainty assessment and quantification. *Scientia Iranica D, 18*(3), 669–676.

Booker, J. M., Ross, T. J., Dolin, R. M., Faust, C., Hamada, M., Najera, L., et al. (2006). Engineering index: An engineering certification/qualification metric. *Military Operations Research., 11*(2), 27–44.

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd edition). Somerset, NJ: John Wiley.

Darby, J. L. (2004). Estimating terrorist risk with possibility theory. Los Alamos National Laboratory manuscript LA-12179, Los Alamos, NM.

de Finetti, B. (1974). *Theory of probability*. NY: John Wiley & Sons.

Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society, Series B, 30*, 205–247.

Dubois, D., & Prade, H. (1988). *Possibility theory*. NY: Plenum Press.

Feller, W. (1968). *An introduction to probability and its applications* (3rd ed., Vol. 1). Somerset, NJ: John Wiley.

Hemez, F. M., & Ben-Haim, Y. (2004). The good, the bad and the ugly of predictive science. In *4th Int. Conf. on Sensitivity of Model Output, Santa Fe, NM*.

Jeffreys, H. (1961). *The theory of probability*. Oxford Press, p. 432.

Joslyn, C.A., & Booker, J. M. (2004). Generalized Information Theory for Engineering Modeling and Simulation, Chapter 9, Engineering Design Reliability Handbook, CRC Press LLC, Boca Raton, FL.

Klir, G. J., & Wierman, M. (1998). *Uncertainty-based information: Elements of generalized information theory*. Heidelberg, Germany: Springer Verlag.

Klir, G. J., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic: Theory and applications*. Upper Saddle River, NJ: Prentice Hall.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*, 79–86.

Langenbrunner, J. R., Booker, J. M. Hemez, F. M., & Hench, K. W. (2007). Variance-based metric Dn for baselining, verification and validation. In *9th Annual AIAA Non-Deterministic Approaches Conference, Honolulu, HI*.

Langenbrunner, J. R., Booker, J. M., Hemez, F. M., & Ross, T. J. (2008). Inference uncertainty quantification instead of full-scale testing. In *49th AIAA/ASME/ ASCE/ AHS/ASC Structures, Structural Dynamics, and Materials Conf. and 10th AIAA Non-Deterministic Approaches Conf., Schaumburg, IL*.

Langenbrunner, J., Booker, J., Hemez, F., Ross, T., & Salazar, I. (2009). An uncertainty inventory demonstration: A primary step in uncertainty quantification", Los Alamos National Laboratory report LA-UR 09–1969, Los Alamos, NM; 11th AIAA Non-Deterministic Approaches Conf., Palm Springs, CA.

Langenbrunner, J., Booker, J., Ross, T., & Hemez, F. (2010). Evolving desiderata for validating engineered-physics systems without full scale testing. In *12th AIAA Non-Deterministic Approaches Conf., Orlando, FL*.

Mamat, N. J., & Daniel, J. K. (2007). Statistical analyses on time complexity and rank consistency between singular value decomposition and duality approach in AHP: A case study of faculty member selection. *International Journal of Mathematical and Computer Modeling, 46*, 1099–1106.

Roache, P. J. (1988). Verification of codes and calculations. *AIAA Journal, 36*, 696–702.
Ross, T. J. An enhanced reliability index for assessing margin of safety in structures. In *Conf. on Rel. and Optimization in Structural Sys., Banff, Canada*, A. A. Balkema Publishers, Netherlands, pp 291–298 (2003).
Ross, T. J. (2010). *Fuzzy Logic with Engineering Applications* (3rd ed.). West Sussex, UK: Wiley & Sons.
Saaty, T. L. (1980). *Analytic hierarchy process: Planning, priority setting, and resource allocation*. Columbus, OH: McGraw-Hill.
Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.
Sharp, D. H., & Wood-Schultz, M. M. (2003). QMU and nuclear weapons certification: What's under the hood? *Los Alamos Science, 28*, 47–53.
US Nuclear Regulatory Commission (NRC). (1975). Reactor safety study: An assessment of accident risks in us commercial nuclear power plants, WASH-1400, NUREG 75/014.
Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. London, UK: Chapman Hall.
Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.
Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems Man and Cybernetics, SMC-3*, 28–44.
Zadeh, L. A. (2002). *Forward to fuzzy logic and probability applications: Bridging the gap*. Philadelphia, PA: Society for Industrial and Applied Mathematics.