Probabilistic Methods for Uncertainty Quantification

N. Chugunov
Institute for Systems Analysis – Russian Academy of Sciences, Russia

G. Shepelyov
Institute for Systmes Analysis – Russian Academy of Sciences, Russia

M. Sternin
Institute for Systems Analysis – Russian Academy of Sciences, Russia

INTRODUCTION

The complexity and interdisciplinary nature of modern problems are often coupled with uncertainty inherent to real-life situations. There is a wide class of real-world problems described by well-formulated quantitative models for which a decision maker (DM) has to deal with uncertainty in values of initial parameters for these models. A good example of such a problem is hydrocarbon reservoir assessment in the exploration stage, which requires the involvement and joint consideration of geological, petroleum engineering, and financial models of reservoir exploration. The consequences of some unreasonable decisions can lead to millions of dollars in loss to the companies as it happens in the oil business, where industry sources on investment decision analysis continue to report surprise values (outside the [P10;P90] range) far more than the 20% indicated by this interval (Welsh, Begg, Bratvold, & Lee, 2004).

In many situations, it is an expert whose knowledge and relevant professional experience turns out to be the primary (and sometimes the only) source of information on uncertain values of properties and parameters involved in problem analysis. Therefore, there is a need for specific methods and tools to elicit and represent expert knowledge in an adequate way (Morgan & Henrion, 1992). From a DM perspective, these methods and tools should also be able to interpret and process information obtained from experts.

It should be noted that uncertainty is also inherent in expert judgments. This aspect, being very important for making justified decisions under uncertainty, still needs further investigation and seems to be a promising direction to improve decision support tools, which would allow experts and DMs to develop a deeper understanding of the problem.

The main objective of this article is to provide an analysis of the state-of-the-art methods and models for uncertainty quantification and propagation with the main focus on probabilistic methods. We start with an overview of existing approaches to expert-knowledge representation. Then we motivate our choice of probabilistic formalism in the later section and describe a new method to elicit and represent expert knowledge on quantitative parameters in problems under uncertainty: the method of generalized interval estimations (GIEs; Chugunov, Shepelyov, & Sternin, 2006; Shepelyov & Sternin, 2003). We introduce arithmetic operations on GIE structures as a first step to propagate uncertainty through models under consideration. Then we present two interpretations of GIE structure to provide analysis-oriented and decision-oriented perspectives. The relation of the proposed method to those widely used for uncertainty quantification is discussed.

Finally, we consider a practical example of a West Siberian oil reservoir assessment to demonstrate probabilistic mono-interval, interval probability, and GIE methods and analyze their pros and cons from both the expert and DM perspective.

BACKGROUND

Expert Knowledge

In this article, by expert knowledge we mean estimates, judgments, and patterns that are elicited from experts to describe values of parameters being analyzed and their interrelations.

Two Components of Uncertainty

In the context of expert analysis, there are two main types of uncertainty (Bedford & Cooke, 2001; Ferson, Kreinovich, Ginzburg, Sentz, & Myers, 2002):

1. **Aleatory:** This is objective, stochastic uncertainty (variability) resulting from inherent properties of the analyzed quantity (e.g., environmental stochasticity, inhomogeneity of materials, fluctuations in time, variation in space, or heterogeneity).
2. **Epistemic:** This is subjective uncertainty that comes from scientific ignorance, inobservability, censoring, or other lack of knowledge. This type of uncertainty is characterized by the state of expert knowledge on the analyzed quantity.

Other variants of uncertainty classification can be found in papers by Cullen and Frey (1999), DeLaurentis and Mavris (2000), Morgan and Henrion (1992), and Vose (2000).

Methods for Uncertainty Quantification

Uncertainty quantification has been traditionally performed by probabilistic methods, which proved to be sufficient in situations when uncertainty is purely aleatory. In this approach, every input parameter is treated as a random variable defined on a certain interval by its probability distribution function. Some researchers claim that probability-based methods are not appropriate to handle epistemic uncertainty due to unreasonable restrictions of probability theory axioms (Yager, 1980). This stimulates development of alternative approaches. In 1965, Lofti Zadeh introduced fuzzy set theory. Fuzzy numbers, as a special kind of fuzzy set on a real line, can be seen as a generalization of interval arithmetic (Moore, 1979), which is the simplest way to quantify uncertainty of both types. Interval couples undergo worst-case and best-case analysis in uncertainty propagation. However, the interval approach cannot allow using additional information on estimated quantities, which can be obtained from an expert. As a generalization of the interval approach, expert judgments on possible interval scenarios can be represented by a membership function to form a fuzzy number or by a possibility measure to form a possibility distribution (Zadeh, 1978).

In the probability-based framework, the traditional mono-interval probabilistic approach was extended to interval probability and later to imprecise probability (Walley, 1991). Similar to interval assessment, the main idea behind this approach is that an expert would feel more comfortable if asked to express her or his estimate by intervals rather than a single value. Thus, if she or he estimates a possible range of values and assigns a certain probability value to this range, this could be interpreted as a Dempster-Shafer (DS) structure (Dempster, 1967; Shafer, 1976). Every specified range of values can be interpreted as a focal element of the DS structure. If an expert feels more comfortable estimating ranges of cumulative probabilities for certain values of an estimated quantity, this would result in probability bounds (probability box or p-box) structure, which represents bounds (P_{min} and P_{max}) on cumulative probability distribution function $P(D<D_s)$ for estimated quantity D and thus characterizes uncertainty in assessment of a true distribution (Figure 1).

As was shown by Regan, Ferson, and Berleant (2004), the p-box and Dempster-Shafer methods are equivalent for uncertainty propagation of real-valued random variables. The uncertainty of expert estimation (epistemic component) is characterized by the width of the probability box.

During last 20 years, research in the uncertainty quantification field continued to explore novel formalisms, but also to generalize and synthesize existing formalisms and investigate the relations among them, which finally has led to a new field named generalized information theory (Booker & Joslyn, 2004; Klir, 1991).

On the other hand, generalization of experience from many real-life cases has shown that the efficiency of expert analysis significantly increases if decision support tools use terminology peculiar to the considered application domain and familiar to the expert (Larichev & Petrovskiy, 1987; Petrovskiy, 1996). For many problem domains, it is the probabilistic approach that proves to be the most appropriate for expert-knowledge representation. For oil and gas reserves estimation, this terminology necessarily includes probability theory also because the official Society of Petroleum Engineers (SPE), World Petroleum Congresses (WPC), and American Association of Petroleum Geologists (AAPG) reserves classification is essentially probability-based (SPE, WPC, & AAPG, 2001). An extensive review of decision and risk analysis frameworks for hydrocarbon