Statistical Science
2006, Vol. 21, No. 4, 460–462
DOI: 10.1214/088342306000000556
Main article DOI: 10.1214/088342306000000510

Rejoinder: Expert Elicitation for Reliable System Design

Tim Bedford, John Quigley and Lesley Walls

First of all, we would like to thank the discussants for the care and thoughtfulness that they have taken in preparing their comments.

Koehler presents a helpful discussion, putting forward a number of different ideas that generalize the approach taken. A taxonomy for technical system elicitation would provide useful guidance for practitioners and serve to codify applicable assumptions during the different systems engineering phases. Although more research is needed here, one could see the emergence of international standards that rely on such a taxonomy.

We acknowledge that the elicitation problem varies greatly depending on the technical system as pointed out by Koehler and we have sought to generalize our experience in studying complex systems, including aerospace, rail and naval for both commercial and defense markets. This explains our bias toward the "closed loop" case. We agree with the two extra areas of expert elicitation identified for "waterfall" cases: lack of expertise continuity and the problem of "forward casting" requirements for an existing system. Both of these relate to discontinuous changes in system operation. Such changes have occurred most obviously in military systems and other projects with long lead times. However, in the commercial world, such discontinuities can be forced by regulatory or market changes, or by outsourcing decisions. These may make historic data collection taxonomies less relevant to the reliability questions posed to support new operational decisions and, therefore, provide new areas of application for expert judgement techniques.

The final point raised by Koehler about the difficulties imposed by system complexity is well made and the notion of multiple concurrent reliability models is intriguing. This does partially link into the notion of expert weighting. However, it also requires a good understanding of the notion of model "expertise" as distinct from expert "expertise." One might argue that if sufficient understanding exists to be able to quantify model expertise, then one should be able to directly build a meta model that incorporates the best of each model. In practice, the need to be cost-efficient will usually mitigate against such a strategy, and model combination is an interesting alternative.

Wang rightly observes that we have not tried to give a survey of expert judgement methodologies. The main reason for this is that several surveys have been undertaken, including a recent one with a wide coverage (Jenkinson, 2005). It has not been our purpose to survey these methods again. Instead we aim to discuss the context in which such models may be used in the engineering design process and to show that the expert problem in this context frequently is more demanding than a "straightforward" probability elicitation.

Having said this, Wang is right to identify empirical Bayes (EB) as an interesting method with potential application in the area under discussion. There is, however, more than one way to utilize this approach. The approach discussed by Wang explicitly uses expert information as data, hence forcing the analyst to choose priors and likelihoods for the expert data given the parameters. This is a fundamental problem because it forces the analyst into the role of meta expert. In this case, the specification of $p(x|\Theta)$ is going to be problematic whether or not we use EB. In our own work with EB (Quigley, Bedford and Walls, 2006, 2007) we have integrated expert judgement into the approach through the selection of pools that comprise different types of events whose data are merged in the EB process. The use of EB allows us to increase the quantity of data available to make estimates of reliability parameters through expert judgements about which events should have similar order of magnitude behavior.

Wang's proposal for using evidential reasoning in reliability combines a number of different questionable features. For the purposes of this rejoinder, we propose distinguishing three different issues contained in the discussion:

- Nonprobabilistic representations of uncertainty.
- Imprecise uncertainties.
- Multicriteria decision models.

Nonprobabilistic representations of uncertainty: We are yet to be convinced that these play a useful role.