SPECIFYING DESIGN CONSERVATISM: WORST CASE VERSUS PROBABILISTIC ANALYSIS

Ralph F. Miles, Jr.

Jet Propulsion Laboratory
Pasadena, California, USA

ABSTRACT

Design conservatism is the difference between specified and required performance, and is introduced when uncertainty is present. The classical approach of worst-case analysis for specifying design conservatism is presented, along with the modern approach of probabilistic analysis. The appropriate degree of design conservatism is a tradeoff between the required resources and the probability and consequences of a failure. A probabilistic analysis properly models this tradeoff, while a worst-case analysis reveals nothing about the probability of failure, and can significantly overstate the consequences of failure. Two aerospace examples will be presented that illustrate problems that can arise with a worst-case analysis.

Key Words: Conservatism, design, analysis, worst-case, probability.

1. INTRODUCTION

Design conservatism is the difference between specified and required performance, and is introduced when uncertainty is present. It can appear in several forms--as a margin added to the performance requirements, or as an overstatement in the associated risk of failure, either in the consequences or the probability of failure. In any form, it forces more resources into some aspect of the design.

In practice, design conservatism is specified and implemented by a wide range of people, from top managers to technicians. In protecting against failures, the designer must ask:

- What can go wrong?
- How likely is it?
- What are the consequences?
- What can be done to reduce the likelihood?
- What can be done to mitigate the consequences?

It is out of this questioning process that candidates for design conservatism arise. Where the consequences of a power outage are severe, facilities supply their own backup power. Uncertainty exists as to how long the power would be out, and how much power would be needed during this period. The design conservatism would be the overcapacity added to the backup power source.

An out-of-control nuclear power plant can go super-critical with a subsequent catastrophic emission of radioactive material. While this has not happened in the U.S., nevertheless Chernobyl makes this a credible scenario. The consequences are sickness, loss of life, and the environmental damage and cleanup costs. Conservatism could be implemented by overstating the likelihood of an accident and the severity of the consequences, and designing to meet environmental requirements consistent with the overstated risk.

2. WORST-CASE ANALYSIS

Worst-case analysis for a design is a technique used to resolve uncertainties, and is one way to implement design conservatism. A design margin is added to bring the minimum performance level of the design up to the maximum performance level that could be required. An additional design margin may be added to account for unknown effects not included in the worst-case estimates. In principle, worst-case analyses negate any requirement to consider uncertainty.

A worst-case analysis for specifying a design may be appropriate if the following conditions can be met:

(1) The worst-case scenario for the analysis is truly a worst case.
(2) The adverse consequence associated with the worst-case scenario is acceptable.
(3) The design consistent with the worst-case analysis is feasible.

For example, if it were not possible for a bird weighing more than 12 pounds to impact an aircraft windshield at more than 600 miles per hour, then a design that could be guaranteed to survive this impact would be consistent with a worst-case analysis. If it were merely improbable that the bird would weigh more, or improbable that the plane would be traveling at a higher speed, then the design would not truly correspond to a worst-case analysis.

With respect to Condition (1), stating that it is not credible that the bird would weigh more, or that it is not credible that the plane would be traveling at a higher speed, is not sufficient. The touted virtue of a worst-case analysis is that you don't have to specify the uncertainty. A statement that a scenario is not credible only says that the scenario has a very small probability of occurrence, not that the scenario is impossible. Thus the worst-case analysis implicitly becomes a probabilistic analysis. Similarly, with respect to Condition (2), the risk associated with the worst-case analysis must be acceptable because the adverse consequence associated with the worst-case analysis is acceptable, not because the probability of occurrence is acceptably small.

What's wrong with designs consistent with worst-case analyses? Nothing, if you can satisfy Conditions (1) and (2), retain feasibility, and spare the additional resources. But this rarely happens in practice. What often is claimed to be a design consistent with a worst-case analysis is more often a design by fiat--the design will perform under stated conditions which may or may not be exceeded--or is a design consistent with an analysis for all credible conditions masking as a worst-case analysis. For example, worst-case analyses for electronic equipment assume that the environment and the power source will be within specification.

Worst-case analyses are most appropriate at the lowest level of systems, such as specifications for electronic parts and circuits, where a detailed analysis of the environment for each individual part would be impractical and unnecessary (Ref. 1). At higher levels in a system, designs consistent with worst-case analyses impact limited resources.

All designs exist in a resource-constrained environment. Thus, exceeding the requirements of one aspect of the design inevitably comes at the expense of something else. The increase in the strength of a structural element comes at the expense of weight--thus reducing payload. A triply-redundant design costs more than simple redundancy.

3. MEASURING PROBABILITY

It is a maxim of good management that in order to manage something you must be able to measure it (Ref. 2). To manage uncertainty and subsequently apply conservatism to a probabilistic design, you must be able to measure the risk--probability and consequences--associated with the uncertainty. While there are undeniable problems with measuring consequences, the greatest controversy arises in the measurement of probability.

Can probabilities be measured--or do they even exist--for events where the probability of occurrence is so low that few if any events have ever been observed? The answer to this question comes from theories of probability. There is little disagreement on how to use probabilities once they have been assigned. Nearly all analyses claiming to be probabilistic use the three Kolmogorov axioms (Ref. 3): (1) Probabilities are numbers equal to or greater than 0, (2) the probability of the universal event (sure thing--something happens) is 1, and (3) probabilities of mutually exclusive events add.

Simple arguments justify using Kolmogorov axioms for the long-run ratio of successes to trials, and probability is defined this way in many probability texts (Ref. 4). This ratio lies between 0 and 1, and these ratios add for mutually exclusive events.

Where few or no events have been observed, probabilities must be based on degrees of belief. One might believe based on some evidence that the probability of rain today is 2/3. Consistency with the Kolmogorov axioms, for example, requires that the probability of no-rain be 1/3. The most commonly used justification for the use of the Kolmogorov axioms for degrees of belief comes from game theory (Refs. 5-7). It can be demonstrated that if you violate the Kolmogorov axioms with respect to your degree of belief about an uncertain event, it is possible for a broker to place a series of bets against you such that you will always lose, independent of the outcome of the event. Contrariwise, if you are consistent with the Kolmogorov axioms, such a series of bets is not possible.

4. PROBABILISTIC ANALYSIS

Apostolakis, in his *Science* article, "The Concept of Probability in Safety Assessments of Technological Systems," lists four steps to doing a probabilistic analysis for a system design (Ref. 8):

(1) Structure the problem.
(2) Quantify uncertainties.
(3) Quantify preferences.
(4) Choose among alternatives.

A probabilistic analysis proceeds by constructing a mathematical model representing the alternatives, the uncertain parameters, the consequences, and their interrelationships, and a model representing preferences for consequence attributes. Uncertainties are represented by random variables with associated probability distributions assessed by technical experts.

Preferences for attributes of the consequences are quantified by means of a value function, and it is in the value function that preferences for conservatism in the design should be expressed, e.g., the probability shall be less than one in a million that the design fails, or that one attribute of the consequences exceeds a stated amount. Finally, the alternative is chosen that maximizes the value.

Probability assessment techniques are discussed in the literature (Ref. 9), and have been applied in the nuclear power industry (Ref. 10 and 11). For a book exhaustive in techniques for modeling uncertainty, see Martz and Waller's *Bayesian Reliability Analysis* (Ref. 12). For introductions to probability assessment, see Merkhofer (Ref. 13) and Apostolakis (Ref. 14).

The product of a probabilistic analysis should include a statement of the central tendencies and spreads of the uncertain quantities. The central tendencies may be expressed as means or modes, and the spreads as standard deviations or fractile ranges. Even more desirable would be complete probability distributions, which can be derived from analysis or from Monte Carlo simulations.

The analysis should include all uncertainties, not just those easily measurable. It should incorporate, in order of increasing difficulty:

(1) Uncertainties in the model parameters.
(2) Uncertainties in the model selection.
(3) Uncertainties in the decision process.

Unfortunately, too often design analyses are called "probabilistic" when only the uncertainties in the parameters of the model are addressed. Rarely are uncertainties in the model selection treated probabilistically, and almost never uncertainties in the decision process.

5. APOLLO RELIABILITY ASSESSMENT

President Kennedy's pronounced goal of getting a man to the moon and returning him safely before the end of the decade of the sixty's initiated the Apollo Project. At times the goal seemed unattainable. Problems and uncertainties abounded--solar-particle radiation and meteoroid hazards, an unknown lunar surface, and issues in designing and testing the Saturn V launch vehicle. The Apollo fire on the launch pad highlighted the difficulty of this project.

A lesser-known problem involved assessing the reliability of the Apollo Mission. George Mueller, the NASA Associate Administrator for Manned Space Flight during this period, says that an attempt was made to estimate the probability of mission success using bottom-up, piece-part failure rates and single-point-failure analyses (Ref. 15). When the original estimates for all the subsystems were combined, the probability of a successful lunar landing was about 5% (Ref. 16). Mueller ordered that a top-down estimate be made by looking at the success probability of each major event. The revised number came out to be 90% for the mission and 99% for getting the crew back, a success probability acceptably high.

The subsequent success rate for crews launched and safely returned has shown that the extreme pessimism of the bottom-up analysis was unwarranted. Whether this pessimism was the result of conservative estimates or true beliefs can't be known. Given the limited experience with space vehicles at that time, it seems reasonable that the engineers produced overly-conservative numbers to cover the uncertainties. As the senior manager, Mueller recognized the implications to the project if it really were true that the missions were likely to fail. Here conservatism as applied by the engineers didn't just yield a suboptimal design, it rendered the project infeasible.

This incident has had a lasting effect to this day on how NASA views numerical estimates of reliability. Quoting from a recent NASA Management Instruction (Ref. 17):

"It is crucial that when quantitative risk assessment methods are used, care must be exercised to ensure that all uncertainties are properly included in the analysis, and clearly displayed to avoid misinterpretation and misuse of the results."

6. THE GALILEO EXPERIENCE

The Galileo spacecraft was launched on October 18, 1989 on a 6-year flight to Jupiter, where it will release a probe into its atmosphere, and then go into orbit around the planet. As Jupiter is too far from the Sun for solar panels to be effective, plutonium-fueled radioisotope thermoelectric generators provide spacecraft power.

The National Environmental Policy Act (NEPA) and Presidential Directive/National Security Council PD/NSC-25 require that all launches with nuclear payloads undergo an extensive environmental impact analysis and review process. NASA, as the implementing agency, produced an Environmental Impact Statement (EIS) and the Department of Energy (DOE) issued a Safety Analysis Report (SAR). The Interagency Nuclear Safety Review Panel (INSRP), an independent review group, produced a Safety Evaluation Report (SER).

The conclusions of these reports differed in some respects. These differences were used in support of a motion for preliminary injunction against the launch of Galileo by litigating parties (Ref. 18):

"The SER's risk assessments are far more pessimistic than those in the Final EIS. In several instances, risk and radiological effects are 10 to 20 times greater than those in the Final EIS. . . . When, however, the estimates of risk in the SER proved to be higher, by magnitudes, than in the . . . SAR . . . NASA simply rejected the conclusions of the SER . . . these actions violated NEPA . . ."

The Court subsequently denied the motion (Ref. 19):

"While the SER is slightly more pessimistic in its risk analysis, there is no evidence that the difference is statistically significant since both documents indicate that the risk is small. . . . The motion for a temporary restraining order is denied."

Were the differences between the SER, the EIS, and the SAR due to basic differences in engineering judgment, or were they due to conservative assumptions? Subsequent to the release of the SAR, the DOE issued a supplement to the SAR stating (Ref. 20):

"The source terms (plutonium releases) have been analyzed anew but with conservatism removed where it was judged to be excessive."

Similarly, for INSRP's SER (Ref. 21):

(Accident response and source terms incorporate) "reasonable conservatism . . . where understanding or physical descriptions were seriously lacking. . . . the source terms calculated tend to lean towards an upper bound of expectation." . . . (The reentry analysis) "is considered conservative."

While DOE's SAR and INSRP's SER differ, it is not possible to know the source of the differences. The conservative assumptions of both reports masked any differences that may have been due to engineering judgment.

In this and the preceding example, unwarranted conservatism was present. Further-

more, it was introduced in the analysis, which is the wrong place. We all want conservatism in designs--in the buildings we occupy, the planes we fly, and the medicine we take. What we shouldn't want is conservatism in the analysis. In the words of Zeckhauser and Viscusi, it amounts "to lying to ourselves about what we expect" (Ref. 22).

7. CONCLUSION

Imagine that you are the manager for a system undergoing a design review. One of your technical experts says he has done a worst-case analysis of a design that protects against all contingencies. Ask yourself these questions:

- Is the analysis truly worst-case?
- If it is a worst-case analysis, can I afford the design consistent with the analysis?
- If it is not a worst-case analysis, then how did the technical expert specify the most-unfavorable case for the analysis?
- What conservatism would I pick for the design?
- What would be the design if my conservatism were specified?

Design conservatism, properly applied, is necessary in that it protects against failure. This conservatism should appear as a specification for the design, and not a condition on the analysis. The degree of conservatism should relate to the probability and severity of the adverse consequences of a failure and should be traded off against the required resources. A probabilistic analysis properly models this tradeoff, while a worst-case analysis reveals nothing about the probability of failure, and can significantly overstate the consequences of failure.

8. REFERENCES

1. Design and Evaluation Inc. 1991. Worst Case Circuit Analysis Training Course. Laurel Springs, New Jersey. Third edition.

2. Daniel S. Goldin. 5 December 1991. Total Quality Management--A Culture Change. *TQM Lessons Learned.* Goddard Space Flight Center.

3. Kolmogorov, A. N. 1956. *Foundations of the Theory of Probability.* New York: Chelsea Publishing Co. Translation of original 1933 German edition.

4. Papoulis, Athanasios. 1984. The Meaning of Probability. *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill Book Company, I, Second Edition.

5. Earman, John. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory.* Cambridge, Massachusetts: The MIT Press.

6. Kemeny, J. G. 1955. Fair betting and inductive probabilities. In *Journal of Symbolic Logic*, 20:263-273.

7. Shimony, A. 1955. Coherence and the axiom of confirmation. In *Journal of Symbolic Logic*, 20: 1-28.

8. George Apostolakis. 7 December 1990. The Concept of Probability in Safety Assessments of Technological Systems. In *Science*, 250: 1359-1364.

9. M. Granger Morgan and Max Henrion. 1990. Human Judgment About and With Uncertainty. In *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* New York: Cambridge University Press, 6.

10. J. W. Hickman, et. al. January 1983. *PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessment.* NUREG/CR-2300, I and II. U.S. Nuclear Regulatory Commission.

11. U.S. Nuclear Regulatory Commission. September 1984. *Probabilistic Risk Assessment (PRA) Reference Document. Final Report.* NUREG-1050-F. Washington, DC.

12. Harry F. Martz and Ray A. Waller. 1982. *Bayesian Reliability Analysis.* New York: John Wiley.

13. Miley W. Merkhofer. 1987. Criticisms and Limitations of Decision-Aiding Approaches. In *Decision Science and Social Risk Management.* Dordrecht, Holland: D. Reidel Publishing Company, 3.

14. George Apostolakis. 1988. Expert Judgment in Probabilistic Safety Assessment. In *Accelerated Life Testing and Experts' Opinions in Reliability.* Edited by C. A. Clarotti and D. V. Lindley. Elsevier Science Publishing Company Inc.: 116-131.

15. George Mueller. 1973. The Apollo Mission. In *Systems Concepts: Lectures on Contemporary Approaches to Systems.* Edited by Ralph F. Miles, Jr. New York: John Wiley.

16. Ben Buchbinder. March 19, 1992. "It is common knowledge at NASA Headquarters that the number was about 5%, but the study may never have been documented." Washington, D.C. NASA Office of Safety and Mission Quality. Phone conversation. A subsequent literature search failed to identify such a study. The number of 50% in Ref. 15 probably was a misquote.

17. NASA QS/Safety Division. 3 February 1988. *Risk Management Policy for Manned Flight Programs.* NASA Management Instruction NMI 8070. Washington, DC. National Aeronautics and Space Administration.

18. Florida Coalition for Peace and Justice, Foundation on Economic Trends, and Christic Institute vs. George Herbert Walker Bush, et. al. 4 October 1989. Memorandum of Points and Authorities in Support of Motion for Preliminary Injunction. United States District Court for the District of Columbia, Civil Action No. 89-2682-OG, 13.

19. Judge Oliver Gasch. 10 October 1989. Memorandum Concerning Civil Action No. 89-2682-OG. United States District Court for the District of Columbia: 11, 12, and 17.

20. General Electric Astro-Space Division. August 1989. *Supplement to the Final Safety Analysis Report for the Galileo Mission.* Philadelphia, Pennsylvania, Document No. 89SDS4221: 43.

21. Interagency Nuclear Safety Review Panel. September 1989. *Safety Evaluation Report for Galileo,* Document INSRP 89-01, I: VI-19.

22. Richard J. Zeckhauser and W. Kip Viscusi. 4 May 1990. Risk Within Reason. *Science,* 248: 559-564.