Control and Cybernetics

vol. **22** (1993) No. 4

Risk of extreme events and the fallacy of the expected value

by

Yacov Y. Haimes

Lawrence R. Quarles Professor of Systems Engineering
and Civil Engineering, Director, Center for
Risk Management of Engineering Systems
University of Virginia
Charlottesville, Virginia 22903
USA

Most, if actually not all, managerial decisions are characterized by (a) elements of risk, uncertainty, and inadequate information and (b) multiple, noncommensurate, competing, and often conflicting objectives. To manage risk, professionals must assess it. This is usually done by a process of its identification, quantification, and evaluation. Thus, the assessment and management of risk is essentially a synthesis and amalgamation of the empirical and the normative, the quantitative and the qualitative, and the objective and the subjective. This paper will address the process of risk assessment and management, focusing on the trade–offs that must be made among all costs, benefits, and risks.

In the process of risk assessment, however, extreme and catastrophic events are often underestimated and commensurated with other less consequential events. Managers and decisionmakers are often most concerned with the risk associated with a specific case under consideration, and not with the likelihood of the average outcomes that may result from various risk situations. In this sense, the expected value of risk, which has until recently dominated most risk analysis in the field, is not only inadequate, but can lead to fallacious results and interpretations. A modification of this approach through the use of conditional expectation will be shown to better capture the risk of extreme and catastrophic events. This paper will focus on the importance of addressing extreme and catastrophic events explicitly and within the overall risk-based decisionmaking process, where trade–offs among costs and risks can be generated and evaluated.

Introduction

The interest of the professional community in risk–based decisionmaking and risk–based approaches to decisionmaking has gained a remarkable growth during the last decade or two. This is not surprising given the fact that most, if not all, engineering systems are designed, manufactured, marketed, distributed, and maintained under conditions of risk and uncertainty. Technological risk, risk to human health and safety, and environmental risk – both man–made and natural – have also been of much concern to the public during the last two decades. Dam failures, nuclear power plant failures, ocean dumping of waste, groundwater contamination, hazards associated with nuclear and toxic waste disposal, and such natural calamities as earthquakes, landslides, tsunamis, windstorms, floods, volcano eruptions, and wildfires have claimed our attention recently along with the less visible but potentially adverse effects of acid rain and possible global warming. Whether the hazards are man–made, natural, or a combination of both, and whether the primary risks are to humans or the environment, fundamental and difficult choices must be made about which adaptive strategies will be the most effective and efficient. And, these choices cannot be made simply on the basis of the average level of adverse effects resulting from these natural and man–made hazards. Budgetary constraints and fears of adverse economic consequences add to the difficulty of making the right choices.

With the increase of public interest in risk–based decisionmaking and the involvement of a growing number of professionals in the field, this relatively new professional niche of risk analysts has gained maturity as well as numbers during the last decade. The professionals involved in risk–based decisionmaking are experiencing the same evolutionary process that systems analysts and systems engineers went through a decade or two ago and may be are still going through. That is, risk analysts are realizing and appreciating both the efficacy and also the limitations of mathematical tools and systematic analysis. In fact, there are many who simply see risk analysis as a specialized extension of the body of knowledge and evaluation perspectives that have come to be associated with systems analysis. Professionals from diverse disciplines are responding much more forcefully and knowledgeably to risks of all kinds as well and, in many instances, are leading what has ultimately come to be a political debate. This professional community is more willing to accept the premise that a truly effective risk analysis study must, in most cases, be crossdisciplinary, relying on social and behavioral scientists, engineers, regulators, and lawyers. And, this professional community has become more critical of the tools that it has developed because it recognizes their ultimate importance and usefulness in the resolution of critical societal problems. Clearly, for these risk methodologies and tools to be useful and effective, they must be representative, that is, they must capture not only the average risks but also the extreme and catastrophic ones.

We are also able to acknowledge the fact that the ultimate utility of decision analysis, including risk–based decisionmaking, is not necessarily to articulate the best policy option, but rather to avoid the extreme, the worst, and the most disastrous policies – those actions in which the cure is worse than the disease.

Thus, risk assessment and management must be an integral part of the decisionmaking process, rather than a gratuitous add–on technical analysis. Viable risk management must be done within a multiobjective framework, where trade–offs can be explicitly evaluated to reflect social risk preferences, professional or expert opinion, and of course, the economic consequences of alternative courses of action.

Risk is a term that is by nature compound in meaning. It connotes the likelihood of adverse events associated with bad consequences. Indeed, risk is commonly defined as a measure of the probability and severity of adverse effects, Lowrance W. (1976). While this definition of risk is widely adopted by many disciplines, its translation into quantitative terms has been a major source of misunderstanding and misguided use, and has often led to erroneous results and conclusions. The most common quantification of risk – the use of the mathematical construct known as the expected value – is probably the dominant reason for this chaotic situation in the quantification of risk. This is so in spite of use of the expected value in conjunction with some other notion, such as "risk versus return". Whether the probabilities associated with the universe of events are viewed by the analyst as discrete or continuous, the expected value of risk is an operation that essentially multiplies each event by its probability of occurrence and sums (or integrates) all these products over the entire universe of events. This operation literally commensurates adverse events of high consequences and low probabilities of exceedance with events of low consequences and high probabilities of exceedance. This paper addresses the misuse, misinterpretation, and fallacy of the expected value when it is used as the sole criterion for risk in decisionmaking. Many experts who are becoming more and more convinced of the grave limitations of the traditional and commonly used expected–value concept are augmenting this concept with a supplementary measure to the expected value of risk – the conditional expectation – by which decisions about extreme and catastrophic events are not averaged with more commonly occurring high–frequency/low consequence events.

It should be noted that expected value is often used as a criterion for utility or benefit in decisionmaking. Using it alone is universally recognized as a risk–neutral position in which it is implied that NO criterion for risk is used.

1. The risk of extreme events

In a society that is slowly adjusting to the risks of everyday life, most analysis and decision theorists are beginning to recognize a simple yet fundamental philosophical truth. In the face of such unforeseen calamities as bridges falling, dams bursting, and airplanes crashing, we are more willing to acknowledge the importance of studying "extreme" events. Modern decision analysts are no longer asking questions about expected risk, but are instead asking questions about expected maximum risk. These analysts are focusing their efforts on forming a more robust (in both a theoretical and a practical sense) treatment of extreme events. Furthermore, managers and decisionmakers are most concerned with the risk associated with a specific case under consideration, and not with the likelihood of the average adverse outcomes that may result from various risk situations. In this sense, the expected value of risk, which until recently has dominated most risk analysis in the field, is not only inadequate, but can lead to fallacious results and interpretations. Indeed, people in general are not risk–neutral. They are often more concerned with low–probability catastrophic events than with more frequently occurring but less severe accidents. In some cases, a slight increase in the cost of modifying a structure might have a very small effect on the unconditional expected risk (the commonly used business–as–usual measure of risk), but would make a significant difference to the conditional expected catastrophic risk. Consequently, the expected catastrophic risk can be of a significant value in many multiobjective risk problems.

Two difficult questions – how safe is safe enough, and what is an acceptable risk? – underlie the normative, value–judgment perspectives in risk–based decisionmaking. No mathematical, empirical knowledge base today can adequately model the perception of risks in the minds of decisionmakers. In the study of multiple–criteria decisionmaking (MCDM), we clearly distinguish between the quantitative element in the decisionmaking process, where efficient (Pareto optimal) solutions and their corresponding trade–off values are generated, and the normative value–judgment element, where the decisionmakers make use of these efficient solutions and trade–off values to determine their preferred (compromise) solution, Chankong V. and Haimes Y.Y. (1983). In many ways, risk–based decisionmaking can and should be viewed as a type of stochastic multiple–criteria decisionmaking in which some of the objective functions represent risk functions. This analogy can be most helpful in making use of the extensive knowledge already generated by MCDM (witness the wealth of publications and conferences on the subject).

It is worth noting that there two modalities to the considerations of risk–based decisionmaking in a multiobjective framework. One is viewing risk (e.g., the risk of dam failure) as an objective function to be tradedoff with the cost functions and the benefit function(s). The second modality concerns the treatment of damages of different magnitudes and of different probabilities of occurrence as noncommensurate objectives, which thus must be augmented by a finite, but small, number of risk functions (e.g., a conditional expected–value function, as will be formally introduced in subsequent discussion). Probably the most demonstrable aspect of the importance of considering risk–based decisionmaking within a stochastic MCDM framework is the handling of extreme events.

To dramatize the importance of understanding the risk of extreme events and the centrality of adequate quantification and evolution of the risk associated with extreme events, the following statements are adopted from Runyon R.P. (1977): Imagine What Life Would Be Like If:

- Our highways were constructed to accommodate the average traffic load of vehicles of average weight.
- Mass transit systems were designed to move only the average number of passengers (i.e., total passengers per day divided by 24 hours) during each

hour of the day.

- Bridges, homes, and industrial and commercial buildings were constructed to withstand the average wind or the average earthquake.
- Telephone lines and switchboards were sufficient in number to accommodate only the average number of phone calls per hour.
- Your friendly local electrical utility calculated the year–round average electrical demand and constructed facilities to provide only this average demand.
- Emergency services provided only the average number of personnel and facilities during all hours of the day and all seasons of the year.
- Our space program provided emergency procedures for only the average type of failure.

Chaos is the word for it. Utter chaos.

Lowrance (1976) makes an important observation on the importance of and imperative distinction between the quantification of risk, which is an empirical process, and the determination of safety, which is a normative process. In both of these processes, which are seemingly dichotomous, the influence and imprints of the analyst cannot and should not be overlooked. The essential role of the analyst, sometimes hidden but often explicit, is not unique to risk assessment and management: rather, it is indigenous to the process of modelling and decisionmaking.

The major problem for the decisionmaker remains one of information overload: for every policy (action or measure) adopted by the decisionmaker there will be a vast array of potential damages as well as benefits and costs with their associated probabilities. It is at this stage that most analysts are caught in the pitfalls of the unqualified expected–value analysis. In their quest to protect the decisionmaker from information overload, analysts precommensurate catastrophic damages that have a low probability of happening with minor damages that have a high probability of occurrence. From the perspective of public policy, it is obvious that a catastrophic dam failure, which might cause flooding of, say, 10^6 acres of land with associated damage to human life and the environment, but which has a very low probability of, say, 10^{-6}, of happening, cannot be viewed by decisonmakers in the same vein as minor flooding of, say, 10^2 acres of land that has a high probability of 10^{-2} of happening. Yet this is exactly what the expected–value function would ultimately generate. Most importantly, the analyst's precommensuration of these low–probability of exceedance/high damage events with high probability of exceedance/low–damage events into one expectation function (indeed some kind of a utility function) markedly distorts the relative importance of these events and consequences as they are viewed, assessed, and evaluated by the decisionmakers. This is similar to the dilemma that used to face theorists and practitioners in the field of MCDM, Haimes Y.Y., Tarvainen K., Shima T., and Thadathil J. (1990b).

2. The fallacy of the expected value

One of the most dominating steps in the risk assessment process is the quantification of risk, yet the validity of the approach most commonly used to quantify risk – its expected value – has received neither the broad professional scrutiny it deserves nor the hoped–for wider mathematical challenge that it mandates. The conditional expected value of the risk of extreme events (among other conditional expected values of risks) generated by the partitioned multiobjective risk method (PMRM), Asbeck E. and Haimes Y.Y. (1984), is one of the few exceptions.

Let $P_x(x)$ denote the probability density function of the random variable X, where X is, for example, the concentration of the contaminant trichloroethylene (TCE) in a groundwater system, measured in parts per billion (ppb). The expected value of the containment concentration (the risk of the groundwater being contaminated by TCE at an average concentration of TCE), is $E(x)$ *ppb*. If the probability density function is discretized to n regions over the entire universe of contaminant concentrations, then $E(x)$ equals the sum of the product of p_i and x_i, where p_i is the probability that the i^{th} segment of the probability regime has TCE concentration of x_i. Integration (instead of summation) can be used for the continuous case. Note, however, that the expected–value operation commensurates contaminations (events) of low concentration and high frequency with contaminations of high concentration and low frequency. For example, events $x_1 = 2$ ppb and $x_2 = 20,000$ ppb that have the same contribution to the overall expected value: $(0.1)\ (2) + (0.00001)\ (20,000) = 0.2 + 0.2$. The relatively low likelihood of a disastrous contamination of the groundwater system with 20,000 ppb of TCE, however, cannot be equivalent, in the mind of the decisionmaker in charge of the integrity of the groundwater system, to the contamination at a low concentration of 0.2 ppb, even with a very high likelihood of such contamination. Due to the nature of mathematical smoothing, the averaging function of the contaminant concentration in this example does not lend itself to prudent management decisions. This is because the expected value of risk does not accentuate the catastrophic events and their consequences, thus misrepresenting what would have been a perceived unacceptable risk.

It is worth nothing that the number of "good" decisions managers make during their tenure is not only a basis for rewards, promotion, and advancement; rather, they are likely to be penalized for any disastrous decisions, no matter how few, made during their career. The notion of "not on my watch" clearly demonstrates the point. In this and other senses, the expected value of risk fails to represent a measure that truly communicates the manager's or the decisionmaker's intentions and perceptions. The conditional expected value of the risk of extreme events generated by the PMRM, when used in conjunction with the (unconditional) expected value, can markedly contribute to the total risk management approach. In the above case, the manager must make trade–offs not only between the cost of the prevention of contamination by TCE vs. the expected value of such contamination, but also the cost of the prevention of contamination vs. the conditional expected value of an extreme level of contamination by TCE. Such a dual multiobjective analysis provides the manager with more complete, more factual, and less–aggregated information about all viable policy options and their associated trade–offs, Haimes Y.Y. (1991).

This act of commensurating the expected value operation is analogous in some sense to the commensuration of all benefits and costs into one monetary unit. Indeed, few today would consider benefit–cost analysis, where all benefits, costs, and risks are commensurated into monetary units, as an adequate and acceptable measure for decisionmaking when it is used as the sole criterion for excellence. Multiple–objective analysis has been demonstrated as a superior approach to benefit–cost analysis, Haimes Y.Y. and Hall W.A. (1974). In many respects, the expected value of risk is similar in its theoretical–mathematical construct to the commensuration of all costs, benefits, and risks into monetary units.

Consider the dam safety problem mentioned earlier, with precipitation being the single random variable evaluated in the damage risk function. Let us discretize the universe of events for this random variable into J segments. Then the damage expectation function, $U(\mathbf{x})$, can be written as, Haimes Y.Y. (1988),

$$U(\mathbf{x}) = \sum_{j=1}^{J} p_j f_j(\mathbf{x})$$

where $f_j(\mathbf{x})$ is the damage associated with the j^{th} segment given as a function of the decision variables $\mathbf{x}, p_j$ is the probability associated with the j^{th} segment, and

$$p_j \geq 0, \ \sum_{j=1}^{J} p_j = 1$$

One might argue that if it were practical, the decisionmakers would rather consider the risk–based decisionmaking problem in the following multiobjective optimization format:

$$\min_{\mathbf{x} \in \mathbf{X}} \{f_1(\mathbf{x}), \ldots, f_j(\mathbf{x})\}$$

where $f_j(\mathbf{x})$ represents a specific range of the damage function that corresponds to a specific range of probabilities of exceedance. Additional objectives representing costs and benefits should be added as appropriate.

Each damage function, $f_j(\mathbf{x})$, associated with the j^{th} segment of the probability axis can be viewed as a noncommensurate objective function. And, in their totality, these damage functions constitute a set of noncommensurate objective functions. Clearly, at one extreme one may consider an infinite number of such objective functions, and at the other extreme one may consider a single objective function – namely, the expected–value function. A compromise between the two extremes must be made for tractability and also for the benefit of the decisionmakers.

Table 1. **DESIGN DATA AND RESULTS**

Option Number	Cost	Mean (μ) expected value	Standard Deviation σ
1	\$100,000	5	1
2	80,000	5	2
3	60,000	5	3
4	40,000	5	4

In the field of MCDM, the profession has accepted the fact that trade–offs exist between the consideration of a very large number of objectives and the ability of the decisionmakers to comprehend these objectives. Consequently, the relevant objective functions, which are commonly numerous, are either organized into a hierarchy of objectives and subobjectives, Peterson D.F., et al. (1974), or some augmentation is followed in order to limit the number of objectives to between five and seven.

To further demonstrate the fallacy of the expected–value approach, consider the following design problem. Four design options are being considered. Associated with each option are cost, the mean of a failure rate (i.e., the expected value of failures for a normally distributed probability density function of a failure rate), and the standard deviation (see Table 1). Figure 1 depicts the normally distributed probability density functions of failure rates for each of the four designs. Clearly on the basis of the expected value alone, the least–cost design (Option 4) seems to be preferred, at a cost of \$40,000. Consulting the variances, which provide an indication of extreme failures, reveals, however, that this choice might after all not be the best and calls for a more in–depth trade–off analysis.

3. The partitioned multiobjective risk method (PMRM)

The partitioned multiobjective risk method (PMRM) is a risk analysis method developed for solving multiobjective problems of a probabilistic nature, Asbeck E. and Haimes Y.Y. (1984). Instead of using the traditional expected value of risk, the PMRM generates a number of conditional expected–value functions, termed risk functions, which represent the risk given that the damage falls within specific ranges of the probability of exceedance. Before the PMRM was developed, problems with at least one random variable were solved by computing and minimizing the unconditional expectation of the random variable representing damage. In contrast, the PMRM isolates a number of damage ranges (by specifying so–called partitioning probabilities) and generates conditional expectations of damage, given that the damage falls within a particular range. In this

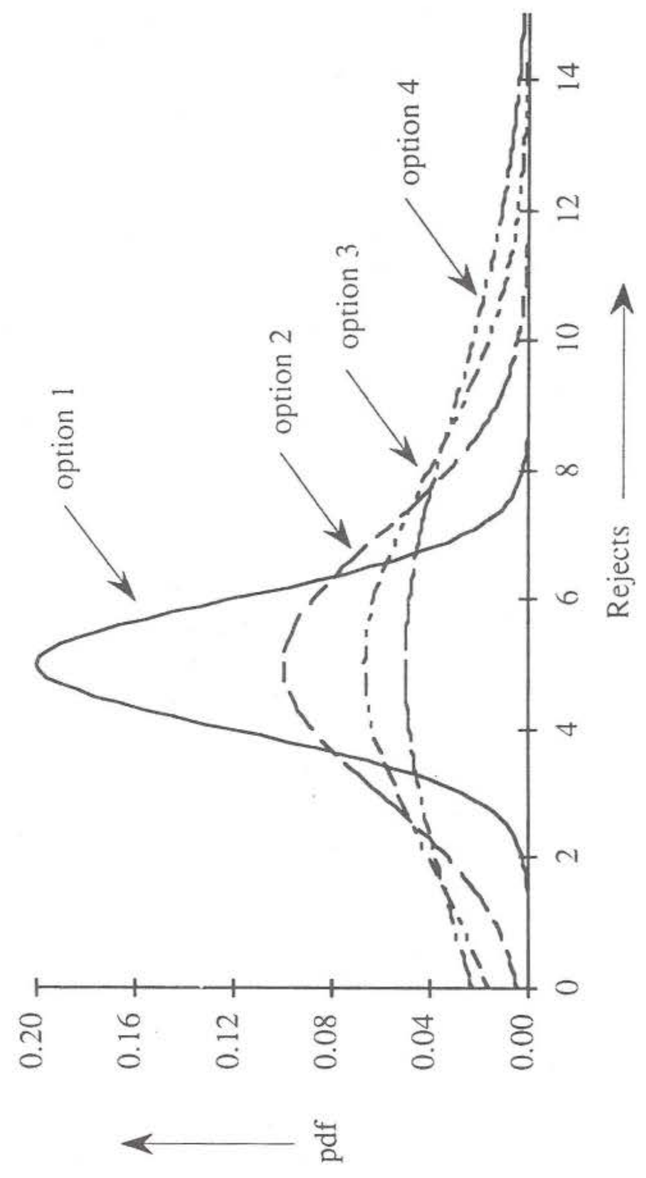

Figure 1 pdf of distributions

manner, the PMRM generates a number of risk functions, one for each range, which are then augmented with the original optimization problem as new objective functions.

The conditional expectations of a problem are found by partitioning the problem's probability axis and mapping these partitions onto the damage axis. Consequently, the damage axis is partitioned into corresponding ranges. A conditional expectation is defined as the expected value of a random variable given that this value lies within some prespecified probability range. Clearly, the values of conditional expectations are dependent on where the probability axis is partitioned. The choice of where to partition is made subjectively by the analyst in response to the extreme characteristics of the decisionmaking problem. If, for example, the analyst is concerned about the once–in–a–million–years catastrophe, the partitioning should be such that the expected catastrophic risk is emphasized. Although no general rule exists to guide the partitioning, Asbeck and Haimes (1984) suggest that if three damage ranges are considered for a normal distribution, then the $+1\sigma$ and $+4\sigma$ partitioning values provide an effective rule of thumb. These values correspond to partitioning the probability axis at 0.84 and 0.99968: that is, the low–damage range would contain 84% of the damage events, the intermediate range would contain just under 16%, and the catastrophic range would contain about 0.032% (probability of 0.00032). In the literature, catastrophic events are generally said to be events with probability of exceedance of 10^{-5} (see, for instance, the NRC Report on dam safety, NRC (1985)). This probability corresponds to events exceeding $+4\sigma$.

A continuous random variable X of damages has a cumulative distribution function (CDF) $P(x)$ and a probability density function (PDF) $p(x)$, which are defined by the relationships

$$P(x) = \text{prob }[X \leq x] \tag{1}$$

and

$$p(x) = \frac{dP(x)}{dx} \tag{2}$$

The CDF represents the *nonexceedance probability* of x. The *exceedance probability* of x is defined as the probability that X is observed to be greater than x and is equal to one minus the CDF evaluated at x.

The expected value, average, or mean value of the random variable X is defined as

$$E[X] = \int_0^\infty xp(x)dx \tag{3}$$

In the PMRM, the concept of the expected value of damage is extended to generate multiple *conditional expected–value functions*, each associated with a particular range of exceedance probabilities or their corresponding range of damage severities. The resulting conditional expected value provides a family of risk measures associated with a particular policy.

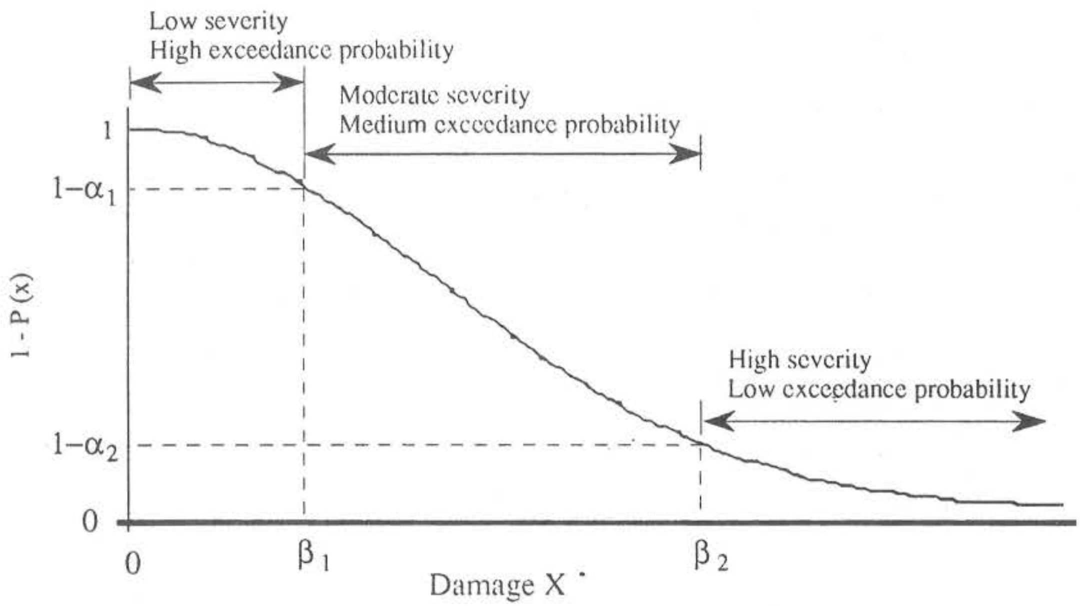

Figure 2 Mapping of the probability partitioning onto the damage axis

Let $1-\alpha_1$ and $1-\alpha_2$, where $0<\alpha_1<\alpha_2<1$, denote exceedance probabilities that partition the domain of X into three ranges, as follows. On a plot of exceedance probability, there is a unique damage β_1 on the damage axis that corresponds to the exceedance probability $1-\alpha_1$ on the probability axis. Similarly, there is a unique damage β_2 that corresponds to the exceedance probability $1-\alpha_2$. Damages less than β_1 are considered to be of low severity, and damages greater than β_2 are of high severity. Similarly, damages of a magnitude between β_1 and β_2 are considered to be of moderate severity. The partitioning of risk into three severity ranges is illustrated in Fig. 2. If the partitioning probability α_1 is specified, for example, to be 0.05, then β_1 is the 5th percentile. Similarly, if α_2 is 0.95, i.e., $1-\alpha_2$ is to be equal to 0.05, then β_2 is the 95th percentile.

For each of the three ranges, the conditional expected damage (given that the damage is within that particular range) provides a measure of the risk associated with the range. These measures are obtained through the definition of the *conditional expected value*. Consequently, the new measures of risk are: $f_2(\bullet)$, of high exceedance probability and low severity; $f_3(\bullet)$, of medium exceedance probability and moderate severity; and $f_4(\bullet)$, of low exceedance probability and high severity. The function $f_2(\bullet)$ is the expected value of X, given that x is less than or equal to β_1:

$$\begin{aligned} f_2(\bullet) &= E[X \| x \leq \beta_1] \\ &= \frac{\int_0^{\beta_1} xp(x)dx}{\int_0^{\beta_1} p(x)dx} \end{aligned} \tag{4}$$

Similarly, for the other two risk functions, $f_3(\bullet)$ and $f_4(\bullet)$:

$$f_3(\bullet) = E[X \| \beta_1 \leq x \leq \beta_2]$$

$$f_3(\bullet) = \frac{\int_{\beta_1}^{\beta_2} xp(x)dx}{\int_{\beta_1}^{\beta_2} p(x)dx} \tag{5}$$

and

$$f_4(\bullet) = E[X \| \beta_2 \leq x]$$

$$f_4(\bullet) = \frac{\int_{\beta_2}^{\infty} xp(x)dx}{\int_{\beta_2}^{\infty} p(x)dx} \tag{6}$$

Thus, for a particular policy option, there are three measures of risk, $f_2(\bullet)$, $f_3(\bullet)$, and $f_4(\bullet)$, in addition to the traditional expected value denoted by $f_5(\bullet)$. The function $f_1(\bullet)$ is reserved for the cost associated with the management of risk. Note that

$$\begin{aligned} f_5(\bullet) &= \frac{\int_0^{\infty} xp(x)dx}{\int_0^{\infty} p(x)dx} \\ &= \int_0^{\infty} xp(x)dx \end{aligned} \tag{7}$$

since the probability of the sample space of X is necessarily equal to one. In the PMRM, all or some subset of these five measures are balanced in a multiobjective formulation. The details are made more explicit in the next two sections.

4. General formulation of the PMRM

Assume that the damage severity associated with the particular policy $s_j, j \in 1, \ldots, q$ can be represented by a continuous random variable X, where $p_X(x; s_j)$ and $P_X(x; s_j)$ denote the PDF and the CDF of damage, respectively. Two partitioning probabilities, $\alpha_i, i = 1, 2$, are preset for the analysis and determine three ranges of damage severity for each policy s_j. A unique damage, β_{ij}, corresponding to the exceedance probability $(1 - \alpha_i)$, can be found due to the monotonicity of $P_X(x; s_j)$. The policies s_j, the partitions α_i, and the bounds β_{ij} of damage ranges are related by the expression

$$P_X(\beta_{ij}; s_j) = \alpha_i \qquad i = 1, 2 \tag{8}$$

This partitioning scheme is illustrated in Fig. 3 for two hypothetical policies s_1 and s_2. The ranges of damage severity include high exceedance probability and low damage, $\{X : x \in [\beta_{0j}, \beta_{1j}]\}$, the set of possible realizations of X for which it is true that $x \in [\beta_{0j}, \beta_{1j}]$; medium exceedance probability and medium damage, $\{X : x \in [\beta_{1j}, \beta_{2j}]\}$; and low exceedance probability and high damage (extreme event), $\{X : x \in [\beta_{2j}, \beta_{3j}]\}$, where β_{0j} and β_{3j} are the lower and upper bounds of damage X.

The conditional expected–value risk functions $f_i, i = 2, 3, 4$, are given by

$$\begin{gathered} f_i(s_j) = E\{X \| p_X(x; s_j), x \in [\beta_{i-2,j}, \beta_{i-1,j}]\} \\ i = 2, 3, 4;\ j = 1, \ldots, q \end{gathered} \tag{9}$$

and, equivalently,

$$f_i(s_j) = \frac{\int_{\beta_{i-2,j}}^{\beta_{i-1,j}} x p_X(x; s_j) dx}{\int_{\beta_{i-2,j}}^{\beta_{i-1,j}} p_X(x; s_j) dx} \qquad i = 2, 3, 4; \quad j = 1, \ldots, q \tag{10}$$

If the unconditional expected value of the damage from policy s_j is defined to be $f_5(s_j)$ and the denominator of Eq. (10) is defined to be $\Theta_i, i = 2, 3, 4$, the following relationship holds:

$$f_5(s_j) = \Theta_2 f_2(s_j) + \Theta_3 f_3(s_j) + \Theta_4 f_4(s_j) \tag{11}$$

with the $\Theta_1 \geq 0$ and $\Theta_2 + \Theta_3 + \Theta_4 = 1$. The Θ_i are the probabilities that X is realized in each of the three damage ranges and are independent of the policies s_j.

The preceding discussion has described the partitioning of three damage ranges by fixed exceedance probabilities α_i, $i = 1, 2$. Alternatively, the PMRM provides for the partitioning of damage ranges by preset thresholds of damage.

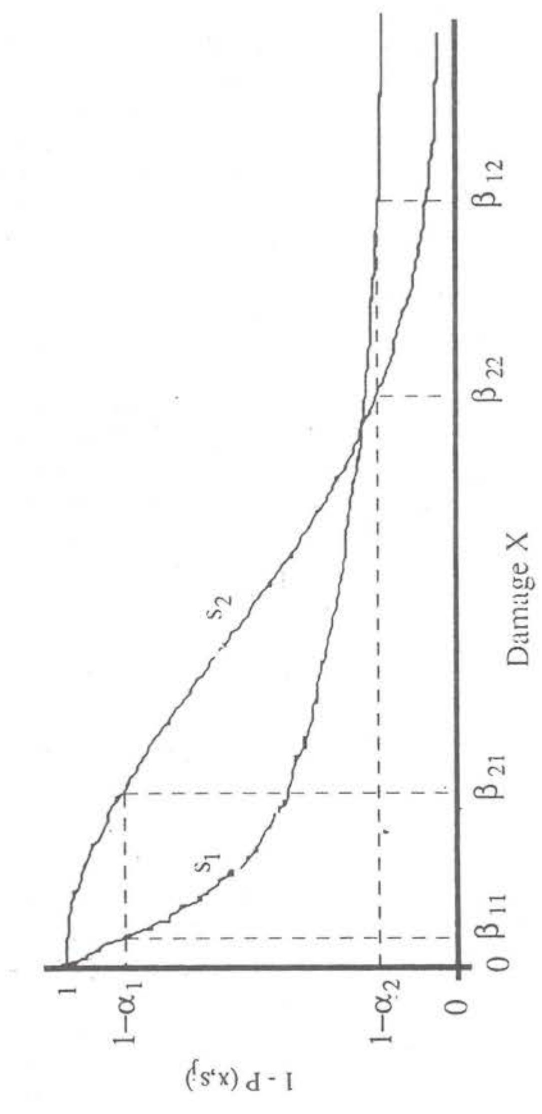

Figure 3 Mapping of the probability partitioning onto the damage axis for two policies

For example, the meaning of $f_4(s_j)$ in partitioning by a fixed damage becomes the expected damage resulting policy j, given that the damage exceeds a fixed magnitude. For further details on the partitioning of damage ranges, see Asbeck E. and Haimes Y.Y. (1984), Karlsson P.-O. and Haimes Y.Y. (1988a), and Karlsson P.-O. and Haimes Y.Y. (1988b).

The conditional expected–value functions in the PMRM are multiple, noncommensurate measures of risk, each associated with a particular range of damage severity. In contrast, the traditional expected value commensurates risks from all ranges of damage severity and represents only the central tendency of the damage.

Combining any one of the generated conditional expected risk functions or the unconditional expected risk functions with the cost objective function f_1 creates a set of multiobjective optimization problems:

$$\min[f_1, f_i]', \quad i = 2,3,4,5$$

This formulation offers more information about the probabilistic behavior of the problem than the single formulation $\min[f_1, f_5]'$. The trade–offs between the cost function f_1 and any risk function $f_i, i \in \{2,3,4,5\}$, allow decisionmakers to consider the marginal cost of a small reduction in the risk objective, given a particular risk assurance for each of the partitioned risk regions and given the unconditional risk functions f_5. The relationship of the trade–offs between the cost function and the various risk functions is given by

$$1/\lambda_{15} = \Theta_2/\lambda_{12} + \Theta_3/\lambda_{13} + \Theta_4/\lambda_{14} \tag{12}$$

where

$$\lambda_{1i} = \partial f_1/\partial f_i \tag{13}$$

with Θ_2, Θ_3 and Θ_4 as defined earlier. A knowledge of this relationship among the marginal costs provides the decisionmakers with insights that are useful for determining an acceptable level of risk. Any multiobjective optimization method (e.g., the surrogate worth trade–off (SWT) method, Haimes Y.Y. and Hall W.A. (1974)), can be applied at this stage.

It has often been observed that the expected catastrophic risk is very sensitive to the partitioning policy. This sensitivity may be quantified using the statistics of extremes approach suggested by Karlsson P.-O. and Haimes Y.Y. (1988a), Karlsson P.-O. and Haimes Y.Y. (1988b). In many applications, if given a data base representing a random process (e.g., hydrological data related to flooding), it is very difficult to find a specific distribution that represents this data base. In some cases one can exclude some probability distribution functions (pdfs) or guess that some are more representative than others. Quite often, one is given a very limited data base that does not contain information about the extreme events. In flood control, for example, records have only been kept for the last 50–100 years, and it is virtually impossible to draw any definite conclusions about floods with return periods exceeding 100 years. In particular, nothing can be said with certainty about the probable maximum flood (PMF), which corresponds to a flood with a return period between 10^4 and 10^6 years. Events of a more extreme character are very important because they determine the expected catastrophic risk. The conditional expectations in the PMRM are dependent on the probability partitions and on the choice of the pdf representing the probabilistic behavior of the data, Karlsson P.-O. and Haimes Y.Y. (1988a), Karlsson P.-O. and Haimes Y.Y. (1988b).

To illustrate the usefulness of the additional information provided by the PMRM, consider Fig.4, where the cost of prevention of groundwater contamination f_1 is plotted against (a) the conditional expected value of contaminant concentration at the low probability of exceedance/high–concentration range f_4 and (b) the unconditional expected value of contaminant concentration f_5. Note that with policy A, an investment of \$2 $\times 10^6$ in the prevention of groundwater contamination results in an expected value of contaminant concentration of 30 parts per billion (ppb); however, under the more conservative view (as represented by f_4), the conditional expected value of contaminant concentration (given that the state of nature will be in a low probability of exceedance/high–concentration region) is twice as high (60 ppb). Policy B, \$$10^6$ of expenditure, reveals similar results: 60 ppb for the unconditional expectation f_5, but 110 ppb for the conditional expectation f_4. Also note that the slopes of the noninferior frontiers with policies A and B are not the same. The slope of f_5 between policies A and B is smaller than that of f_4, indicating that a further investment beyond \$$10^6$ would contribute more to a reduction of the extreme–event risk f_4 than it would to the unconditional expectation f_5. The trade–offs λ_{1i} provide a most valuable piece of information. More specifically, the decisionmaker is provided with an additional insight into the risk trade–off problem through f_4 (similarly through f_2 and f_3). The expenditure of \$$10^6$ may not necessarily result in a contaminant concentration of 60 ppb; it may instead have a nonnegligible probability resulting in a concentration of 100 ppb. (If, for example, the partitioning were made on the probability axis and in addition a normal probability distribution were assumed, then this likelihood can be quantified in terms of a specific number of standard deviations.) Furthermore, with an additional expenditure of \$$10^6$ (policy A), even the extreme event of likely concentration is 60 ppb – within the range of acceptable standards. It is worth remembering that the additional conditional risk functions provided by the PMRM do not invalidate the traditional expected–value analysis per se – they improve on it by providing additional insight into the nature of risk assessment and management.

Let us revisit the design problem with its four alternatives. Table 2 summarizes the values of the conditional expected value of extreme failure, f_4. Figure 5 depicts the cost of each design vs. the unconditional expected value, f_5, and the cost vs. the conditional expected value, f_4. Clearly, the conditional expected value f_4 provides more valued additional information on the associated risk than the unconditional expected value f_5, where the impact of the variance of each

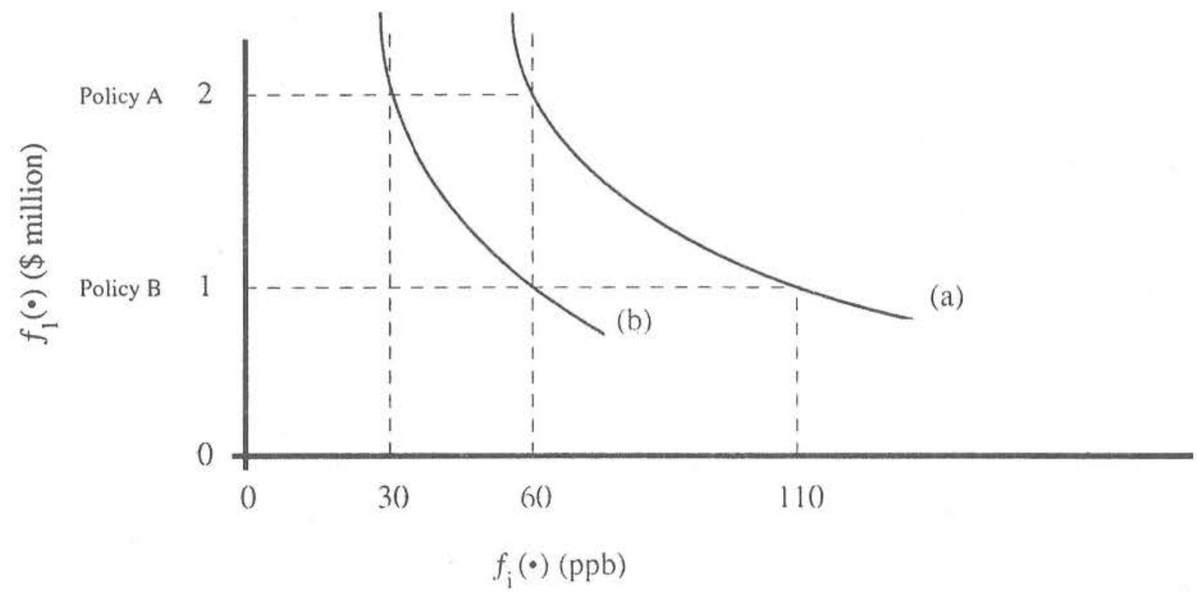

Figure 4 (a) Cost function vs. conditional expected value of contaminant concentration $f_4(\cdot)$; and (b) Cost function vs. expected value of contaminant concentration $f_5(\cdot)$

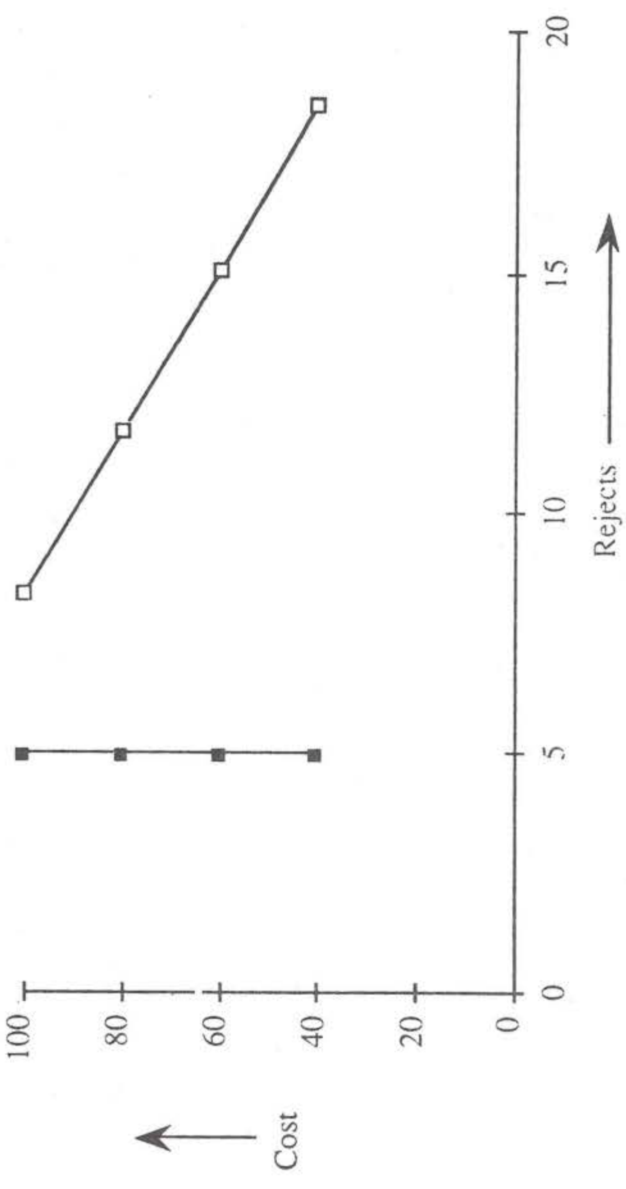

Figure 5 Pareto optimal frontier

Table 2. **DESIGN DATA AND RESULTS**

Option Number	Cost	Mean μ Expected Value	Standard Deviation σ	$f_4(\cdot)$ Conditional Expected Value
1	\$100,000	5	1	8.37
2	80,000	5	2	11.73
3	60,000	5	3	15.10
4	40,000	5	4	18.47

Table 3. **COST OF IMPROVING THE DAM'S SAFETY AND THE CORRESPONDING CONDITIONAL AND UNCONDITIONAL EXPECTED DAMAGES**

Scenarios	$f_1(x)$ \$ 10^6	$f_4(x)$ \$ 10^6	$f_5(x)$ \$ 10^6
1	0	1260	161.7
2	20	835	161.6
3	26	746	161.6
4	36	719	161.5

alternative design is captured by f_4.

To further demonstrate the value of the additional information provided by the conditional expected value $f_4(\mathbf{x})$, consider the following results obtained by Petrakian R., Haimes Y.Y., Stakhiv E.Z. and Moser D.A. (1989) on the Shoohawk dam study. Two decision variables are considered: (a) raising the dam's height and (b) increasing the dam's spillway capacity. Although Petrakian et al. considered several policy options (scenarios), only a few are discussed here. Table 3 presents the values of $f_1(\mathbf{x})$ (the cost associated with increasing the dam's height and the spillway capacity), and $f_4(\mathbf{x})$ and $f_5(\mathbf{x})$ (the conditional and unconditional expected value of damages, respectively). These values are listed for each of the selected scenarios. Note that the range of the unconditional expected value of the damage, $f_5(\mathbf{x})$, is between \$161–162 million for the various scenarios. On the other hand, the range of the low-frequency high-damage conditional expected value, $f_4(\mathbf{x})$, varies between \$719 million and \$1,260 million – a marked difference. Thus, while an investment in the safety of the dam at a cost, $f_1(\mathbf{x})$, ranging from \$0 to \$36 million, does not appreciably reduce the unconditional expected value of damages, such an investment markedly reduces the conditional expected value of extreme damage from about \$1,260 million to

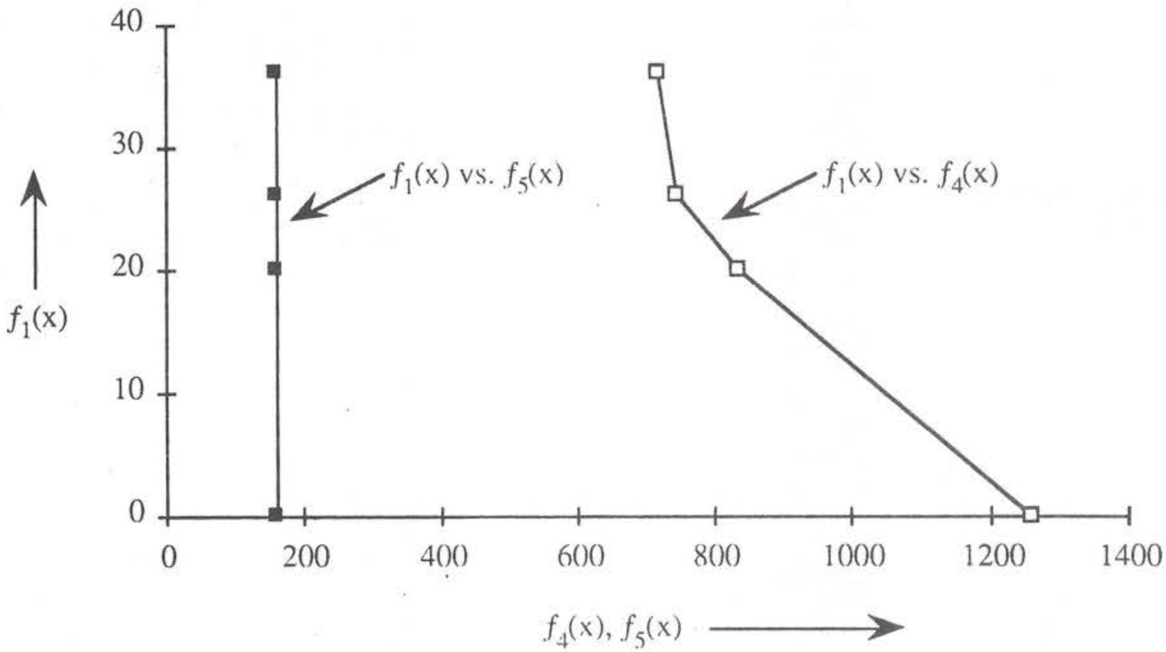

Figure 6 Pareto optimal frontiers of $f_1(x)$ vs. $f_4(x)$ and $f_1(x)$ vs. $f_5(x)$

about \$720 million. This significant insight into the probable effect of different policy options on the safety of the Shoohawk dam would have been completely lost without the consideration of the conditional expected value derived by the PMRM. Figure 6 depicts the plotting of $f_1(\mathbf{x})$ versus $f_4(\mathbf{x})$ and $f_5(\mathbf{x})$. Note that the unusually high values of $f_1(x)$, in the order of \$160 million, are attributed to the assumptions concerning antecedent flood conditions (in compliance with the guidelines and recommendations established by the U.S. Army Corps of Engineers).

5. Statistics of extremes

This section summarizes some research results that relate the theory of the statistics of extremes to the conditional expected value of extreme events, $f_4(\bullet)$.

The subject of the statistics of extremes is concerned with studying the largest (or smallest) value realized from a sample of t independent, identically distributed (i.i.d.) random variables, Ang A.H.S. and Tang W.H. (1984). The study of $f_4(\bullet)$ has been advanced by the correspondence between the function $f_4(\bullet)$, the statistics of extremes, and hydrologic time series phenomena, Karlsson P.-O. and Haimes Y.Y. (1988a), Karlsson P.-O. and Haimes Y.Y. (1988b).

Just as statistical moments parametrize random variables, the statistics of extremes defines two parameters that characterize the largest value of a sample of t i.i.d. variables, which is itself a random variable. The characteristic largest value, $u_t(s_j)$ is of the magnitude that is exceeded on the average once in t realizations of X, and it is implicitly defined by

$$t[1 - P_x(u_t(s_j))] = 1 \tag{14}$$

A second parameter, $\delta_t(s_j)$, which measures the sensitivity of the characteristic largest value, $u_t(s_j)$, to the sample size t, is defined by

$$\frac{du_t(s_j)}{d[\ln(t)]} = \frac{1}{\delta_t(s_j)} \tag{15}$$

From Equation (14) it can be seen that if X is an annual maximum flood, then t is the return period in years of the event of magnitude $u_t(s_j)$. The return period corresponding to the partition probability α_2 of extreme events is given by

$$t = \frac{1}{(1 - \alpha_2)} \tag{16}$$

From Equations (8), (14), and (16), it can be seen that the damage partition β_{2j} for extreme events is exactly the characteristic largest value, $u_t(s_j)$. Equation (10), with $i = 4$ is rewritten via the statistics of extremes as

$$f_4(s_j) = t \int_t^{\infty} \frac{1}{\tau^2} u_\tau(s_j) d\tau \tag{17}$$

It can be derived from Equation (17) that $f_4(s_j)$ can be approximated, with an appropriate number of terms, through the expression

$$f_4(s_j) = u_t(s_j) + \sum\nolimits_{k=1}^{\infty} \frac{d^k u_t(s_j)}{d[\ln t]^k} + c(s_j)t \tag{18}$$

where the term $c(s_j)$ is typically equal to zero. Equation (18), apart from being theoretically interesting, is particularly useful for engineering models in which monetary or other damage is a transformation of a hydrologic random variable.

The statistics–of–extremes approach additionally gives important results concerning the sensitivity of $f_4(\bullet)$ to the range (partitioning point) of extreme events and the relationships between $f_4(\bullet)$ and the asymptomatic nature of the damage distribution, Karlsson P.-O. and Haimes Y.Y. (1988a), Karlsson P.-O. and Haimes Y.Y. (1988b).

6. Risk management

To be effective and meaningful, risk management must be an integral part of the overall management of a system. This is particularly important in the management of the risk of extreme events, where the failure of the system can have catastrophic consequences. The term management may vary in meaning according to the discipline involved and/or the context; this is true also of risk. Risk management is commonly distinguished from risk assessment, even though some may use the term risk management to encompass the entire process of risk assessment and management. In risk assessment the analyst often attempts to answer the following three questions, Kaplan S. and Garrick B.J. (1981): What can go wrong? What is the likelihood that it would go wrong? And what are the consequences? Answers to these questions help risk analysts identify, measure, quantify, and evaluate risks and their consequences and impacts. Risk management builds on the risk assessment process by seeking answers to a second set of three questions, Haimes Y.Y. (1991): What can be done? What options are available and what are their associated trade–offs in terms of all costs, benefits, and risks? And what are the impacts of current management decisions on future options? Only when these questions are addressed in the broader context of management, where all options and their associated trade–offs are considered within the hierarchical organizational structure, can a total risk management be realized. Indeed, evaluating the total trade–offs among all important and relevant system objectives in terms of costs, benefits, and risks cannot be done seriously and meaningfully in isolation from the broader resource–allocation perspectives of the overall organization. Good management must thus incorporate and address risk management within a holistic and all–encompassing framework that incorporates and addresses all relevant resource–allocation and other related management issues. A total risk management approach that harmonizes risk management with the overall system management must also address the following four sources of failure: (a) hardware failure, (b) software failure, (c) organizational failure, and (d) human failure.

This set of sources of failure is intended to be internally comprehensive, i.e., comprehensive within the system's own internal environment. External sources of failure are not addressed here because they are commonly system–dependent. These four internal elements are not necessarily independent of each other, however. The distinction between software and hardware is not always straightforward, and separating human and organizational failure is often not an easy task. Nevertheless, these four categories of sources of failure provide a meaningful foundation upon which to build a total risk management framework.

In first–rate book on quality control, Kaizen, Imai, Masaaki (1986) says, "The three building blocks of business are hardware, software, and humanware". He further states that total quality control "means that quality control efforts must involve people, organization, hardware, and software." In her extensive research on the off–shore oil industry, Pate–Cornell M.E. (1990) found that organizational failures accounted for more than 90% of the accidents.

Total risk management can be defined as a systemic, statistically based, and holistic process that builds on formal risk assessment and management (answering the previously introduced two sets of triplet questions for risk assessment and risk management) and that addresses the set of four sources of failures within a hierarchical–multiobjective framework. The term hierarchical–multiobjective framework can be explained in the context of total risk management. Most if not all organizations are hierarchical in their structure and consequently in the decisionmaking process that they follow. Furthermore, multiple, conflicting, competing, and noncommensurate objectives drive the decisionmaking process at each level of the organizational hierarchy. Thus, within the organization, there are commonly several sets of objectives, subobjectives, and sub–subobjectives corresponding to the levels of the hierarchial structure and to its various units or subsystems, Haimes Y.Y., Tarvainen K., Shima T. and Thadathil J. (1990b). At the heart of good management decisions is the "optimal" allocation of the organization's resources among its various hierarchical levels and its various subsystems. The optimal allocation is meant in the Pareto optimal sense, where trade–offs among all costs, benefits, and risks are evaluated in terms of hierarchical objectives (and sub–objectives) and in terms of their temporal impacts on future options. Methodological approaches for such a hierarchical framework are discussed in Haimes Y.Y., Tarvainen K., Shima T. and Thadathil J. (1990b).

Conclusions

New metrics to represent and measure the risk of extreme events are needed to supplement and complement the expected value measure of risk, which represents the central tendency of events. There is much work to be done in this area, including the extension of the PMRM. Research efforts directed at making use of results from the area of statistics of extremes in representing risk of extreme events, however, have been proven to be very promising and should be continued.

Acknowledgements

The author would like to thank Duan Li, Jim Lambert, and Vijay Tulsiani for their comments and suggestions.

The research summarized in this paper was supported in part by the National Science Foundation under Grant BCS–8912630, entitled "Integrating the Statistics of Extremes With Conditional Expectation." Editorial assistance was provided by Virginia Benade.

References

Ang A.H.S. and Tang W.H. (1984) Probability Concepts in Engineering Planning and Design, Volume II: Decision, Risk, and Reliability, John Wiley and Sons, New York, NY.

Asbeck E. and Haimes Y.Y. (1984) The partitioned multiobjective risk method, *Large Scale Systems,* **6,** (1), 13–38.

Chankong V. and Haimes Y.Y. (1983) Multiobjective Decision Making: Theory and Methodology, North–Holland, New York.

Haimes Y.Y. (1988) Alternatives to the precommensuration of costs, benefits, risks, and time, in The Role of Social and Behavioral Sciences in Water Resources Planning and Management, D.D.Bauman and Y.Y.Haimes, Eds., ASCE, New York.

Haimes Y.Y. (1991) Total Risk Management, Editorial, *Risk Analysis,* **11,** (2).

Haimes Y.Y. and Hall W.A. (1974) Multiobjectives in Water Resources Systems Analysis: The Surrogate Worth Trade–off Method, *Water Resources Research,* **10,** (4).

Haimes Y.Y., Li D., Karlsson, and Mitsiopoulos (1990a) Extreme events: risk management, in System and Control Encyclopedia, supplementary, vol. 1, M.G. Singh, Ed., Pergamon Press, Oxford.

Haimes Y.Y., Tarvainen K., Shima T., and Thadathil J. (1990b) Hierarchical Multiobjective Analysis of Large Scale Systems, Hemisphere Publishing Corp., New York.

Haimes Y.Y., Lambert J.H., and Li D. (1992) Risk of extreme events in a multiobjective framework, *Water Resources Bulletin,* **28,** *1.*

Imai, Masaaki, and Kaizen (1986) *The Key to Japan's Competitive Success, McGraw–Hill Publishing Company,* New York.

Kaplan S. and Garrick B.J., (1981) On the quantitative definition of risk, *Risk Analysis,* **1,** (1).

Karlsson P.-O. and Haimes Y.Y. (1988a) Probability distributions and their partitioning, *Water Resources Research,* **24,** (1).

KARLSSON P.-O. AND HAIMES Y.Y. (1988B) Risk–based analysis of extreme events, *Water Resources Research*, **24**, (1).

LOWRANCE W. (1976) Of Acceptable Risk, William Kaufmann, Inc., Los Altos, California.

NRC (1985) National Research Council (NRC), Committee on Safety Criteria for Dams, Safety of Dams – Flood and Earthquake Criteria, National Academy Press, Washington, DC.

PATE–CORNELL M.E. (1990) Organizational aspects of engineering system safety: The case of offshore platforms, *Science*, **250**, 1210–1217.

PETERSON D.F., ET AL. (1974) Water Resources Planning, Social Goals, and Indicators: Methodological Development and Empirical Tests, Utah Water Research Laboratory, Utah State University, Logan, Utah, PRWG 131–1.

PETRAKIAN R., HAIMES Y.Y., STAKHIV E.Z., AND MOSER D.A. (1989) Risk analysis of dam failure and extreme floods, in Risk Analysis and Management of Natural and Man–Made Hazards, Y.Y. Haimes and E.Z. Stakhiv, Eds., ASCE, New York.

RUNYON R.P. (1977) Winning the Statistics, Addison–Wesley, Reading, Massachusetts.