shedding some light on this thing called Risk Assessment

Part I – Risk Analysis Basics

By Gregg A. Scott, P.E.

The Issue

Although dam safety risk assessment has been around for decades, recently it has become a hot topic and has found its way into the dam safety conferences and literature. However, the papers and explanations often assume the audience has some familiarity with probabilistic methodology and terminology. For those who have had limited exposure to these concepts, including most engineers and scientists who had a single college level course in probability and statistics, it is difficult to follow the discussion. This article is intended to provide some basic background information to help dam safety professionals understand what this thing called risk assessment is all about, drawing upon the more practical methods in use at the Bureau of Reclamation.

Definitions

Many terms are bandied about during discussions on risk. As a basis for common understanding, the following definitions are offered for the discussion that follows (Bureau of Reclamation, 2010).

RISK - The probability of adverse consequences. It is normally calculated as the product of the probability of the load, the probability of failure (given the load), and the consequences (given that failure occurs). Thus, by definition, risk includes both likelihood and consequences.

RISK ANALYSIS - A quantitative calculation or qualitative evaluation of risk.

RISK ASSESSMENT - The process of deciding whether existing risks are tolerable and present risk control measures are adequate and if not, whether alternative risk control measures are justified and should be implemented.

Risk Analysis – The Basic Building Blocks

The basic building blocks used for risk analyses at the Bureau of Reclamation are outlined in a best practices document (Bureau of Reclamation, 2010). They include the following, described further below:

- Potential failure mode analysis
- Event trees
- Load frequency analyses
- Probabilistic analyses and models
- Subjective probability and expert elicitation
- Consequence evaluation.

Potential Failure Mode Analysis

The first, and perhaps the most critical step in a risk analysis involves identifying and fully describing potential failure modes based on an evaluation of a dam's vulnerabilities. If this first step is not diligent and thorough, it doesn't matter what is done for the rest of the risk analysis; the results will have significantly less value, and may even lead to incorrect or unsubstantiated conclusions. Attempts to define risk or risk-related parameters without identifying potential failure modes may not produce reliable results (ASDSO, 2003). In addition, generic consequence evaluations performed independently of potential failure mode scenarios can be misleading.

It is important to include, but also think beyond, traditional "standards-based" analyses when identifying potential failure modes. Dams are engineered systems, and significant thought must be put into the details surrounding the interactions between the various features of a particular structure. This includes partitioning the system into various structural and geologic features. For example,

ISSN 1944-9836 - Association of State Dam Safety Officials

potential failure modes for concrete structures, embankments, and mechanical features must be evaluated separately. In the case of long embankments, changes in embankment geometry or foundation conditions may necessitate additional subdivision. Some of the greatest risks for uncontrolled reservoir release may be due to operational issues or potential failure modes that do not lend themselves to standard engineering calculations. Knowledge of dam failure case histories can also be useful in identifying and evaluating potential failure modes.

An adequate job of identifying potential failure modes can only be performed after all relevant background information on a dam is diligently collected and thoroughly reviewed, including information related to geology, design, analysis, construction, flood and seismic loading, operations, dam safety evaluations, and performance monitoring. Photographs, particularly those taken during construction or unusual events, are often key to identifying vulnerabilities. It is essential that the records be reviewed by more than one person, as something might have been overlooked in previous reviews, and one person may pick up on critical information related to their area of expertise that another person might miss.

Identifying potential failure modes is best done in a team setting, with a small but diverse group of qualified people. Input from operating personnel is essential to the process. Team members develop the potential failure modes, based on their understanding of the vulnerabilities of the dam and project from the data review and current field conditions. The potential failure modes must be described fully, from initiation to breach (or uncontrolled reservoir release). The description must include three parts:

- **THE INITIATOR.** This is the loading or physical condition that leads to initiation of a failure mode. For example, this could include increases in reservoir level due to flooding (perhaps exacerbated by a debris-plugged spillway), strong earthquake ground shaking, malfunction of a gate or equipment, an increase in internal water pressures due to drain plugging, or a decrease in strength.
- **FAILURE PROGRESSION.** This includes the step-by-step mechanisms, or states of nature, that lead to the breach or uncontrolled release of the reservoir. The location where the failure is most likely to occur should also be highlighted. For example, this might include the path through which materials will be transported in a piping situation, the location of overtopping in a flood, or anticipated failure surfaces in a sliding situation. If one section of the dam is more susceptible to a potential failure mode, this should be noted.
- **THE BREACH.** The method and expected magnitude of the breach or uncontrolled release of the reservoir is also part of the description. This could include the breach mechanism, and how rapid and large the expected breach would be, all of which aid in assessing the available time to implement emergency actions such as warnings and evacuations to mitigate the potential consequences of failure.

The reasons for completely describing the potential failure modes are: (1) to ensure that the team and reviewers have a common understanding, (2) to ensure that someone reading or using the report well into the future will have a clear understanding of what the team was thinking, and (3) to enable development of an event tree or other means of estimating risks, if warranted. It is often helpful to sketch the potential failure modes to ensure there is a common understanding of each.

Once a potential failure mode is identified and described, it is screened to determine whether it is a candidate for risk analysis. }The primary intent is to identify those potential failure modes that are clearly so remote as to be non-credible. These non-credible failure modes are not carried forward into the quantitative risk analysis. However, it is important that they be documented, along with the reasons for their dismissal.

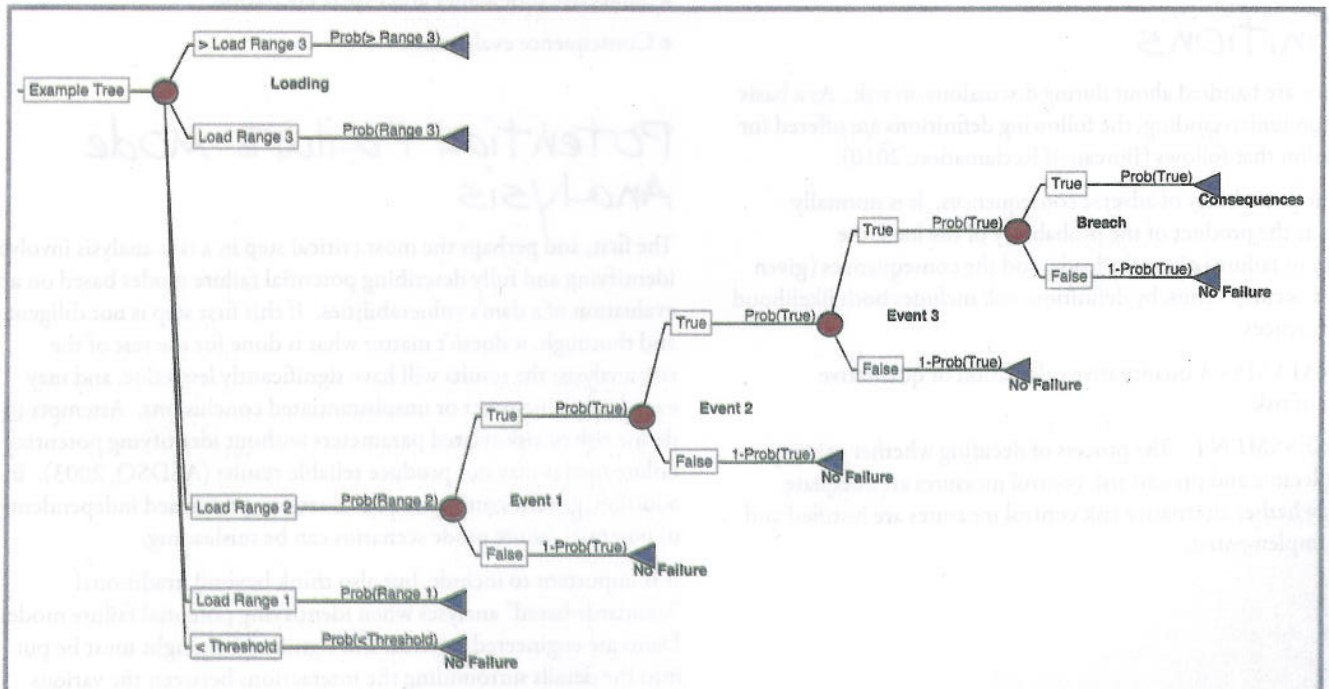


Figure 1 – Example Event Tree (note that only one complete branch is shown)

Event Trees

Event trees allow risk analysts to visualize the progression of events that lead to failure. Since this is how potential failure modes are typically described, most geotechnical, hydraulic, and structural failure modes are evaluated using event trees. The Bureau of Reclamation uses the event tree as the primary tool for estimating risks.

Once a complete description of a potential failure mode is developed, it can usually be broken down into a series of steps or events that lead to failure, which can be modeled using an event tree, as illustrated in Figure 1. Events progress from left to right in the event tree. After the event tree is developed, the likelihood of each node (event or state of nature) on the event tree can be evaluated, using results from numerical modeling, statistical evaluations when sufficient frequency-based information exists, and/or subjective degree-of-belief estimates, depending on the available information, level of study, and question to be answered. These will subsequently be described.

Risk is defined as the product of the load probability, the failure probability given the load, and the consequences given that failure occurs. The loading is typically included near the beginning of the event tree and is typically characterized by annual exceedance probability hazard curves. When using event trees, it is necessary to partition the loading into ranges, and evaluate the remainder of the tree for each load range. The estimates following the loading are referred to as "conditional" probabilities based on the condition that the specified loading has occurred. In addition, the probability of each subsequent node or event is also a conditional probability; that is, the previous nodes are assumed to be true when evaluating subsequent nodes. The probabilities of all branches assigned to a node must sum to 1.0. The probability of failure is calculated by multiplying the branch probabilities (for those branches that lead to dam failure) from left to right through the tree, and summing the products of all branches that terminate in dam failure.

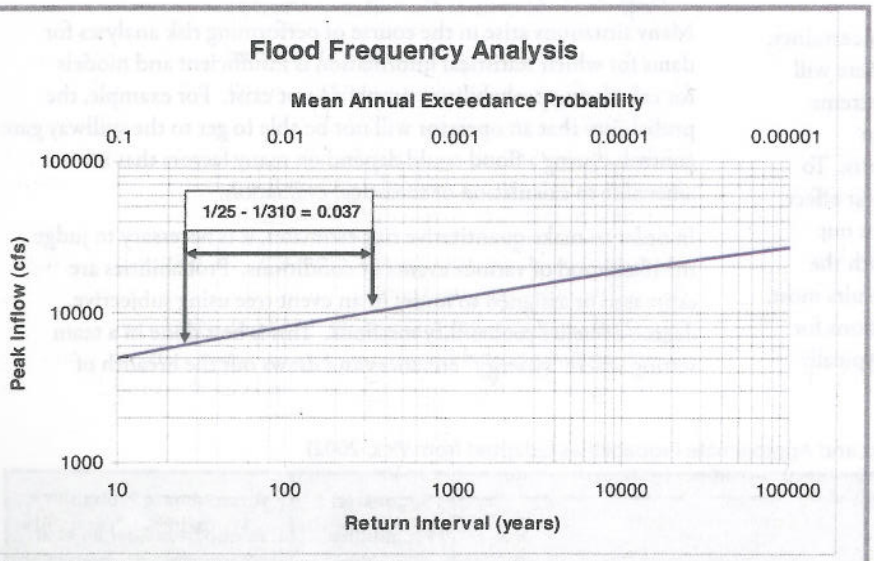


Figure 2 – Example Hazard Curve (often uncertainty bounds are also provided)

Load Frequency Analysis

Seismic and hydrologic hazard curves provided by seismologists and hydrologists are used to calculate load range probabilities for earthquakes and floods. Some potential failure modes are more likely at higher reservoir levels, and reservoir exceedance curves under normal operating conditions may also factor into load probabilities. A load range probability is taken as the exceedance probability of the upper end of the load range subtracted from the exceedance probability of the lower end of the load range, as shown in Figure 2 for the load range from 6,000 to 10,000 ft^3/s. Note that the probabilities for all of the individual load ranges must sum to 1.0, or some range of loading is not being considered, or is being double counted. For the upper load range, the exceedance probability at infinity, or the value of zero, is subtracted from the lower bound of the range.

Selecting the appropriate number of load ranges (load range partitioning), and the appropriate limits for the load ranges is not a trivial matter. It is important to identify the threshold loading, below which failure probability is negligible. This becomes the bottom end of the lowest load range for which risks are estimated and the upper end of the lowest load range which is assumed to be a "no failure" range. While simple in concept, the selected value can have a significant effect on the annual failure probability, and care is needed to ensure the value selected is consistent with the definition of failure. It is important to select enough load ranges that the transitions in conditional failure probability are smoothly captured. On the other hand, too many load ranges can make interpreting the results more difficult, as they make it more difficult to understand where the risks are coming from and why. Load range boundaries should be selected, to the extent possible, at loads where the structural response is expected to change. It is a good idea to examine the mean conditional failure probability at the lower and upper ends of an individual load range. If they are largely different (for example, 0.1 at the low end and 0.9 at the upper end) then the load range is probably too large.

Probabilistic Analysis

With the advent of computer statistical analysis add-in tools, if you can program a deterministic analysis spreadsheet, you can use it to perform probabilistic analyses. Several companies sell such software, and some commercially available stability analysis software programs have options for performing probabilistic analysis.

With these tools, the standard deterministic equations are programmed into a spreadsheet, but instead of defining the input parameters as single values, they are defined as probability distributions. Typical input distributions might consist of normal (defined by a mean and a standard deviation), triangular (defined by a low, high and most likely value), or uniform (defined by a high and low value, with any value in between equally likely). Several other types of distributions are also available in the commercial software. Instead of calculating a single value for the output, a distribution of

ISSN 1944-9836 - Association of State Dam Safety Officials

output is generated from numerous iterations using a Monte-Carlo simulation, whereby each input distribution is sampled in a manner consistent with its shape or probability density function. This output distribution is used to determine the probability of unsatisfactory performance.

Consider, for example, that the equations for calculating the sliding factor of safety (for either a concrete or embankment structure) have been entered into a spreadsheet. After entering the distributions in the spreadsheet cells for the input parameters, the factor of safety cell is selected as the output. Then the simulation is run with the click of a button. For each iteration, the input distributions are sampled and a factor of safety is calculated. This results in a listing of the calculated factors of safety for the simulation. It is a simple matter to sort the listing of output factors of safety in ascending or descending order using the sort command of the spreadsheet program. The probability of FS<1.0 is the number of iterations whose calculated factor of safety is less than 1.0, divided by the total number of iterations. For example, if 228 iterations out of 10,000 produced a factor of safety less than 1.0, the probability of FS<1.0 is 228/10,000 or 0.0228. If no factors of safety are calculated to be less than 1.0, a parameter called the "reliability index" can be used to estimate the probability of FS<1.0. The reliability index is simply the number of "standard deviation" units between the mean value of the distribution and the value representing unsatisfactory performance, and is directly related to the area under the output distribution at values less than that representing FS=1.0 (see Scott, 2007 for additional discussion). There are other methods for performing probabilistic analyses besides the Monte-Carlo approach described here. However, the Monte-Carlo analyses have proven to provide reasonable results for a wide range of problems, and are currently the method of choice at the Bureau of Reclamation.

Although probabilistic analyses attempt to account for uncertainty, when dealing with dam safety engineering it is unlikely there will be sufficient data to define the input distributions with extreme confidence. Therefore, it is usually appropriate to perform sensitivity studies using variations to the input distributions. To help understand which input distributions have the greatest effect on the results, most commercially available programs print out a list of ranking coefficients. Those input distributions with the highest absolute value of ranking coefficients affect the results most. Additional analyses using variations in the input distributions for the parameters with the highest ranking coefficients are typically warranted. This could take the form of increasing the upper and lower limits, or the standard deviation (dispersion) to account for more uncertainty. Each set of input distributions results in a factor of safety output distribution, and its associated probability of unsatisfactory performance. By using different distributions to represent possible ranges in interpreting the available information, a range of probabilities for unsatisfactory performance is obtained. This range in probabilities can then be used to develop a distribution for an event tree node.

Uncertainty also exists as to how well the models used in the calculations actually reflect the real situation. For example, if the results are based on two-dimensional calculations, but there are significant three-dimensional effects that tend to help with stability, the failure probability can be reduced based on supplemental calculations and the expectations related to 3-D improvements to stability. Similarly, if there are uncertainties in the model that may be un-conservative, such as adding shear strengths for different materials along the sliding plane that may not be mobilized at the same shear strain, the failure probability can be increased based on supplemental analyses, assuming the load is carried only by the most rigid segments.

In some cases, other types of probabilistic models are available to assist with assigning probabilities to event trees. For example, probabilistic liquefaction models can be used to estimate the likelihood of liquefaction. If more than one such model is used, they should be weighted according to the degree to which each is judged to represent the case at hand.

Subjective Probability and Expert Elicitation

Many situations arise in the course of performing risk analyses for dams for which statistical information is insufficient and models for calculating probabilities simply do not exist. For example, the probability that an operator will not be able to get to the spillway gate controls during a flood could depend on many factors that are not amenable to calculation or statistical evaluation.

In order to make quantitative risk estimates, it is necessary to judge the likelihood of various events or conditions. Probabilities are estimated or assigned to nodes in an event tree using subjective, degree-of-belief probability methods. This is best done in a team setting where "synergy" enhances and draws out the breadth of

Table 1. Verbal Descriptors and Approximate Probabilities (adapted from Vick, 2002)

Verbal Descriptor	Suggested Probability	Approximate Probability Range from Reagan et al
Virtually Impossible, due to known physical conditions or processes that can be described and specified with almost complete confidence	0.01	0-0.05
Very Unlikely, although the possibility cannot be ruled out	0.1	0.02-0.15
Equally Likely, with no reason to believe that one outcome is more or less likely than the other (when given two outcomes)	0.5	0.45-0.55
Very Likely, but not completely certain	0.9	0.75-0.9
Virtually Certain, due to known physical processes and conditions that can be described and specified with almost complete confidence	0.99	0.9-0.995

experience brought to the table by a group of individuals qualified to make the estimates. A subjective probability estimate may include the range of values judged to be believable based upon the available evidence and experience of the estimators.

Subjective probability estimates can be made to represent the likelihood of events for a potential failure mode that has been decomposed in an event tree. A verbal mapping scheme reported by Vick (2002) is shown in Table 1, based on experiments reported by Reagan et al (1989), which are traceable to a known set of results. These experiments show that, within reasonable limits, people are fairly well calibrated and consistent relative to known probabilities when describing events in common terms.

A key finding of the experiments was that people's ability to judge likelihood does not extend very far out on either end of the probability scale–that is, to more than a couple orders of magnitude as one approaches either zero (0.01) or one (0.99)–even when words like "almost impossible" or "almost certain" are used. This is likely due to the fact that human experience rarely allows us to conceptualize likelihoods at extreme probabilities; thus we do not have words that adequately describe them. In addition, it is difficult for most people to conceptualize how often events happen at remote probabilities, absent actual frequency data. Therefore, it is important to "decompose" each potential failure mode into enough events to allow meaningful estimation of likelihoods.

Once a potential failure mode has been decomposed, probability estimates for an event (node) can be made using the verbal mapping scheme. This is done by listing the adverse factors (factors that make the node more likely) and favorable factors (factors that make the node less likely) associated with each node. These factors are the evidence that is used to reach a range in likelihood estimates and build the case for the numbers that are generated. The process of collecting and listing the "adverse and favorable factors" by a trained and experienced facilitator typically generates significant team discussion, which provides a relative sense of the importance of each factor discussed. Once the facilitator senses discussion has ended, the group is queried for a probability estimate or range of estimates using the verbal descriptors. After the estimates or range of estimates are made, the key factors and justification for those estimates must be captured and documented. The facilitator helps guide the process, attempting to deal with biases and adverse group interactions along the way (see Scott et al, 2009 for additional discussion). In some circumstances it may be beneficial to poll the team anonymously for decision-making rather than eliciting verbal estimates. This can reduce unintended bias, especially if the discussions are dominated by a few outspoken members.

Fragility Curves

Fragility curves are sometimes used to represent the probability of an adverse outcome as a function of some other parameter, such as shown in Figure 3. They can also be used to represent a node on an event tree. Coincidentally, event trees can also be used to develop fragility curves. Fragility curves can be developed using numerical analyses or subjective probability. An advantage of using fragility curves in combination with hazard curves is that very small load range increments can be used in calculating risk (essentially, the curves are multiplied together in a manner similar to an event tree), and thus the selection of the load range intervals becomes relatively unimportant. A disadvantage of using fragility curves is that it becomes more difficult to sort out what is controlling the risk (for example, whether it is from remote floods or earthquakes, or more frequent events). This is particularly true if several curves are multiplied together representing various nodes on an event tree. Thus, the Bureau of Reclamation neither requires nor prohibits the use of fragility curves; their use is decided case-by-case.

Table 2. Empirical Fatality Rates (after Graham, 1999)

Flood Severity	Warning Time	Understanding	Fatality Rate	
			Suggested	Range
High	None	N/A	0.75	0.30 to 1.00
	15 to 60 min	Vague	Use above values - apply to number of people who remain in flood plain after evacuation.	
		Precise		
	> 60 min	Vague		
		Precise		
Medium	None	N/A	0.15	0.03 to 0.35
	15 to 60 min	Vague	0.04	0.01 to 0.08
		Precise	0.03	0.005 to 0.04
	> 60 min	Vague	0.02	0.005 to 0.06
		Precise	0.01	0.002 to 0.02
	None	N/A	0.01	0.0 to 0.02
Low	15 to 60 min	Vague	0.007	0.0 to 0.015
		Precise	0.002	0.0 to 0.004
	> 60 min	Vague	0.0003	0.0 to 0.0006
		Precise	0.0002	0.0 to 0.0003

ISSN 1944-9836 - Association of State Dam Safety Officials

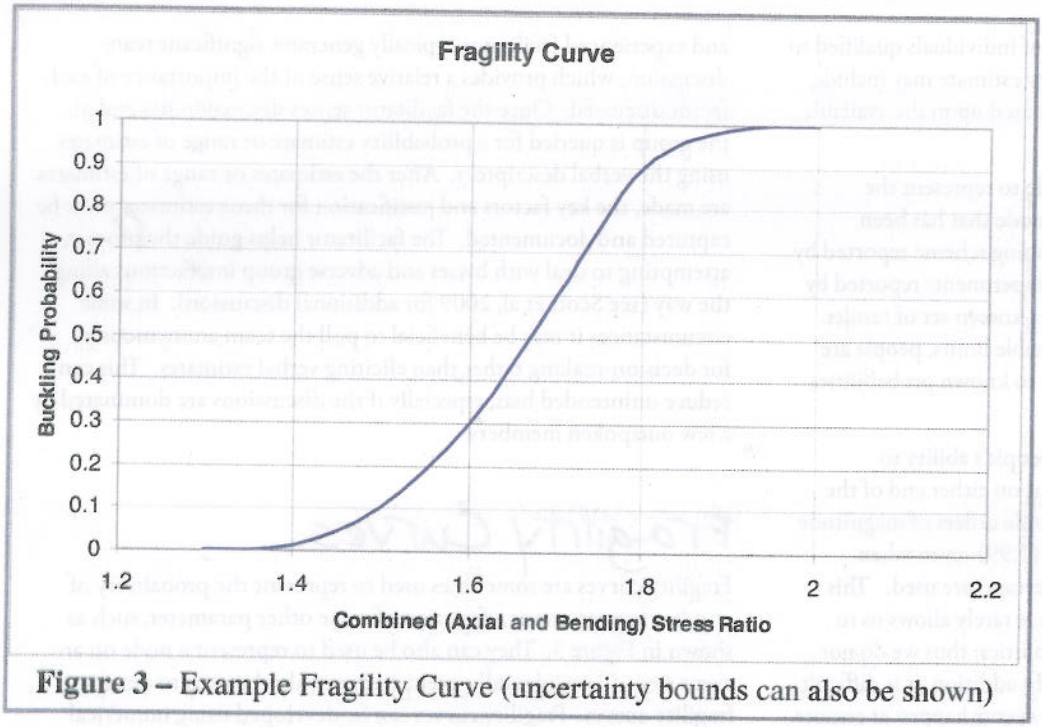


Figure 3 – Example Fragility Curve (uncertainty bounds can also be shown)

Consequences

Consequences associated with dam failure can take many forms, including economic, environmental, sociological, and life safety. Although the Bureau of Reclamation has performed risk analyses that consider all of these, dam safety risk analyses tend to focus on the potential for loss of life. In this regard, the methodology developed by Graham (1999) continues to be the primary means of estimating potential consequences at the Bureau of Reclamation. Although there are some cases when quantifying the number of potential fatalities may not be desirable (these will be covered in later issues), the sad fact of the matter is that when dams fail, people often die, and this must be acknowledged. The methodology developed by Graham is based on case studies of historical dam failures. The data suggests that for a given population at risk and potential failure mode, the level of life loss is dependent primarily on three factors:

- FLOOD SEVERITY. For high severity flooding, the area is wiped clean and the ultimate flood depths appear suddenly (for example as a result of rapid failure of a concrete dam in a narrow canyon). For medium severity flooding, houses are washed off their foundations, but the ultimate flood depths ramp up with time and mangled structures and trees remain for refuge. Houses and buildings are not washed off their foundations for low severity flooding.
- WARNING TIME. The more time people have to evacuate, the more effective the evacuation will be. The time it takes to detect failure and initiate warning is subtracted from the flood wave travel time (determined from inundation mapping) to arrive at the warning time.
- FLOOD SEVERITY UNDERSTANDING. If people understand precisely what is about to descend upon them, they are more likely to evacuate in a timely manner and stay away; they may not do so if their understanding is vague.

The flood plain downstream of a dam is typically divided into reaches representing different population centers. A flood severity, warning time, and flood severity understanding is assigned to each population center based on the characteristics derived from the inundation mapping. A spreadsheet can be set up to use the results from Table 2 in estimating consequences. The likelihood of variations in population at risk (PAR), flood severities, warning times, and understandings can be included in the spreadsheet to account for additional uncertainty. For example, the potential for loss of life at night may be significantly different than during day time hours (due to differences in warning time), or during summer when recreation use is at a peak versus times of the year when recreation use is minimal.

It should be noted that the case histories from which this method was developed included only small to moderately-sized population centers with limited warning time. The effects of evacuation are included in the fatality rates. Additional consideration is required for large population centers with significant warning time. These cases require an evaluation of evacuation routes and the potential for gridlock, and appropriate fatality rates applied to the population that is unable to or does not evacuate.

The database used to develop the method includes no seismic failures. Following an earthquake large enough to fail a dam, the inundation area would likely be seriously damaged with major disruption to communication lines, evacuation routes, and availability of emergency personnel. Under these circumstances, the warning time and ability to evacuate would be reduced, which would increase the potential loss of life estimate. However, the probability of the reservoir being full or at a high level during an earthquake should also be considered.

It should also be noted that models for numerically simulating consequences are being developed at Utah State University and BC Hydro. These models are more complex and take time to learn and use. Regardless of the method used for estimating fatalities, the process involves assessing complex human behavior; therefore estimates must encompass a wide range.

R. Craig Findlay, Ph.D., P.E., G.E.
70 Old Field Road, Yarmouth, ME 04096

FINDLAY ENGINEERING INC.

Phone: 207-846-1465
Fax: 207-846-3434
cfindlay@findlayengineering.com

Dam Safety, Geotechnical and Water Resources Engineering
website: www.findlayengineering.com

REFERENCES

Association of State Dam Safety Officials. *Risk Categorization for Dams: Report of the Steering Committee for ASDSO*. April 2003.

Graham, W.J. *A Procedure for Estimating Loss of Life Caused by Dam Failure*. Report DSO-99-06. U.S. Bureau of Reclamation. Denver, Colorado. September 1999.

Scott, G.A. *Probabilistic Stability Analysis - You Can Do It*. Dam Safety 2007 - Proceedings of the Annual Conference of the Association of State Dam Safety Officials Held in Austin, Texas. September 2007.

Scott, G.A., N.J. Snorteland, and K.M. Dise. *Risk Analysis and Building the Dam Safety Case*. Dam Safety 2009 - Proceedings of the Annual Conference of the Association of State Dam Safety Officials Held in Hollywood, Florida. September 2009.

Reagan, R., F. Mosteller, and C. Youtz. "Quantitative Meanings of Verbal Probability Expressions." *Journal of Applied Psychology* 74, no. 3 (1989): 433-442.

U.S. Bureau of Reclamation. *Best Practices in Dam Safety Risk Analysis*. Version 1.3. Denver. February 2010.

Vick, S.G. *Degrees of Belief, Subjective Probability and Engineering Judgment*. Reston, VA: American Society of Civil Engineers, 2002.

Gregg A. Scott, P.E.

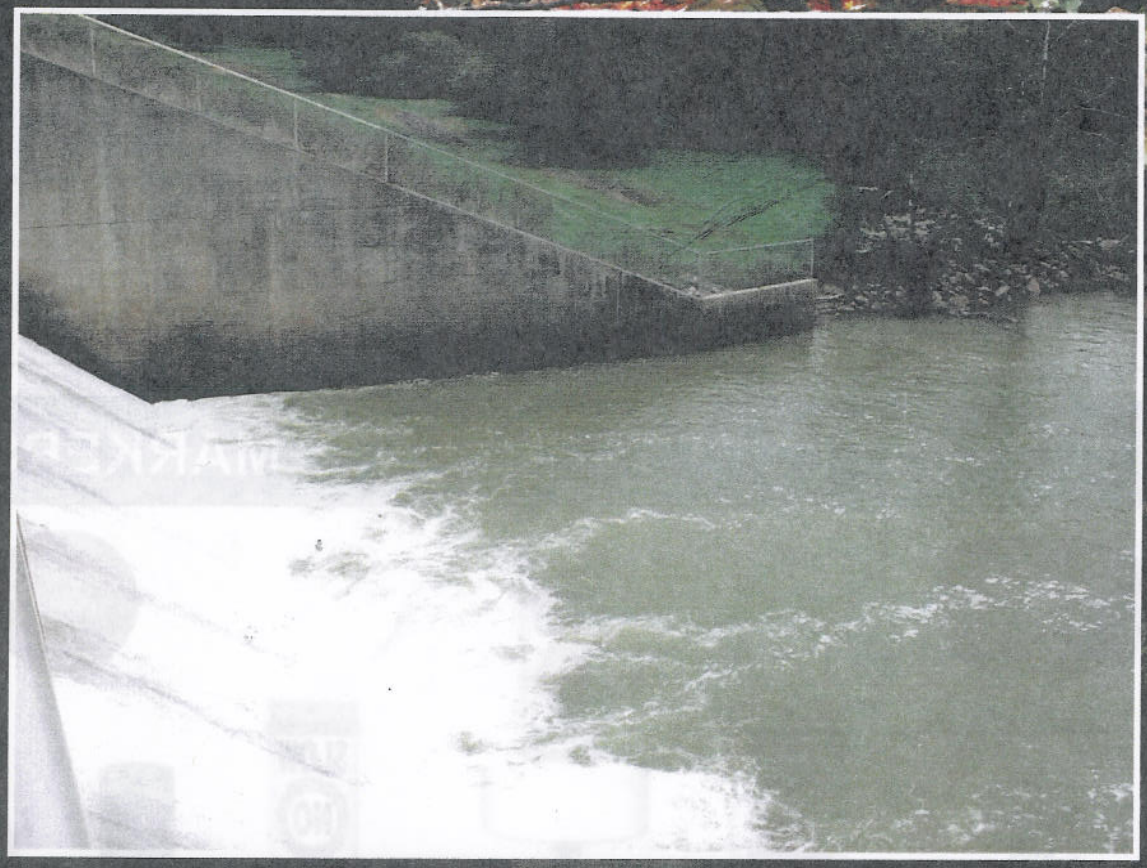
Civil Engineer, Senior Technical Specialist
Bureau of Reclamation
Denver, Colorado
gscott@usbr.gov

Mr. Scott received B.S. and M.S. degrees in Civil Engineering from the University of Colorado, Boulder. He has been at the Bureau of Reclamation since 1976, and is a registered professional engineer. He has been responsible for explorations, testing, analyses, designs, specifications, and construction support for several major dam projects. He has served as Senior Engineer on over 30 comprehensive facility reviews for concrete and embankment dams. While helping to implement Reclamation's risk analysis processes as a member of Reclamation's Risk Cadre, he has facilitated risk analyses for over 30 dams and water conveyance structures. He is also a member of Reclamation's Dam Safety Advisory Team, which reviews all dam safety studies and recommendations, and advises the Dam Safety Office on the appropriate course of action. He is a Fellow of the American Society of Civil Engineers, and a Member of the Association of State Dam Safety Officials and the U.S. Society on Dams.

ISSN 1944-9836 - Association of State Dam Safety Officials

ASSOCIATION OF STATE
DAM SAFETY OFFICIALS

THE JOURNAL OF Dam Safety



ISSN 1944-9836

VOLUME 8 | ISSUE 1 | 2010