Table of Contents

List of Figures

List of Tables

Introduction

This appendix describes the computations used by the HEC-FDA package to obtain the following analysis variables: 1) exceedance probability curves; 2) project reliability; 3) expected annual damage, 4) flood damage reduction benefits, and 5) probable flood stages conditional on the occurrence of a particular exceedance probability event. These variables are computed from various relationships that represent watershed runoff and economic factors important to estimating flood damage (e.g., discharge-exceedance probability, stage-discharge and stage-damage curves) . The contributing relationships are characterized by both a best estimate and the uncertainty in this estimate.

Contents

Monte Carlo Simulation

Overview

Monte Carlo simulation (Davis and Rabinowitz 1967) is used in HEC-FDA to derive the expected annual damage corresponding to a particular plan/analysis year for a damage reach. The expected annual damage (EAD) is the mean damage obtained by integrating the damage exceedance probability curve for the damage reach. The damage-exceedance probability function is obtained from the discharge-exceedance probability, stage-discharge, and damage-stage functions derived at a damage reach index location. The inclusion of uncertainty for these variables requires a numerical integration approach be applied. Without uncertainty, the damage-exceedance probability curve can be obtained directly without resorting to numerical simulation approaches.

Monte Carlo simulation is the numerical integration approach. It relies on an exceedance probability analysis of samples of the contributing random variables obtained from the generation of random numbers. Although inelegant, the technique is computationally efficient in comparison with other techniques as the number of contributing variables exceeds about five.

Numerical Integration with Monte Carlo Simulation

Expected annual damage is the probability weighted average of all possible peak annual damages. It is also termed the mean or expected annual damage. As a simple example of computing a probability weighted average, consider the rolling of a die. The probability of obtaining any outcome of any roll of a die is 1/6, since the probability of obtaining any face of the die is considered equally likely (at least if the die is fair). The probability weighted average is then computed as:

$$\sum_{i=1}^{i=6} d_i p_i = \frac{1}{6}(1+2+3+4+5+6) = 3.5 \qquad \textbf{(1)}$$

where d_i is the possible outcome of rolling a die, and p_i is the probability of the outcome. The probability weighted average or expected outcome of 3.5 obtained in equation (1) could be obtained by performing a die rolling experiment. The experiment would just involve many trials of rolling the die and averaging the outcome. As the number of trials becomes large the average obtained will equal 3.5.

Performing trials with the die is an application of a Monte Carlo simulation to obtain an average. In rolling the die, random integers are obtained in the inclusive interval 1 to 6, and a statistical analysis of the outcome is performed to obtain an average. Consequently, Monte Carlo simulation or application of equation (1) are equivalent procedures for obtaining the mean or expected value.

Other statistical characteristics of rolling a die could be obtained, such as by performing a class category analysis on the outcomes to determined the probability of obtaining any outcome. If this were done, the probability of obtaining any die face in a single trial would be found to be 1/6.

This same type of sampling experiment can be performed to obtain EAD. Computation of EAD is somewhat more difficult in that damage is a continuous random variable, unlike the outcome of rolling a die, which has discrete outcomes. Consequently, damage probability is either stated for an interval, or more typically as, the probability of exceeding a particular value. These probabilities are defined by the damage exceedance probability function or equivalently, the cumulative distribution function as defined by:

$$P[D>d] = F(D) = \int_{d}^{\infty} f(D)\,dD \qquad \textbf{(2)}$$

where D is the annual damage, F(D) is a function defining the damage exceedance probability curve, f(D) is the probability density function (units of probability per increment of damage), and P[D>d] is read as "the probability that D exceeds d."

The probability density function can be used to calculate the EAD or equivalently the probability weighted average damage by performing the following numerical integration:

$$EAD = \int_{0}^{\infty} D f(D)\,dD \sim \sum_{i=1}^{i=N} D_i \Delta p \qquad \textbf{(3)}$$

where the integral in equation (3) is approximated by a sum as in equation (1), Δp is the probability of damage being in an interval, D_i is the midpoint damage of this interval, and N is the number of intervals (see Figure F.1). The approximation turns the integration of a continuous random variable into that of a discrete variable much as in the computation of the average outcome for rolling a die shown in equation (1). The difference between the equations is that equation (1) is exact and the probability is for a discrete outcome; whereas, equation (3) is approximate and Δp is an interval probability.

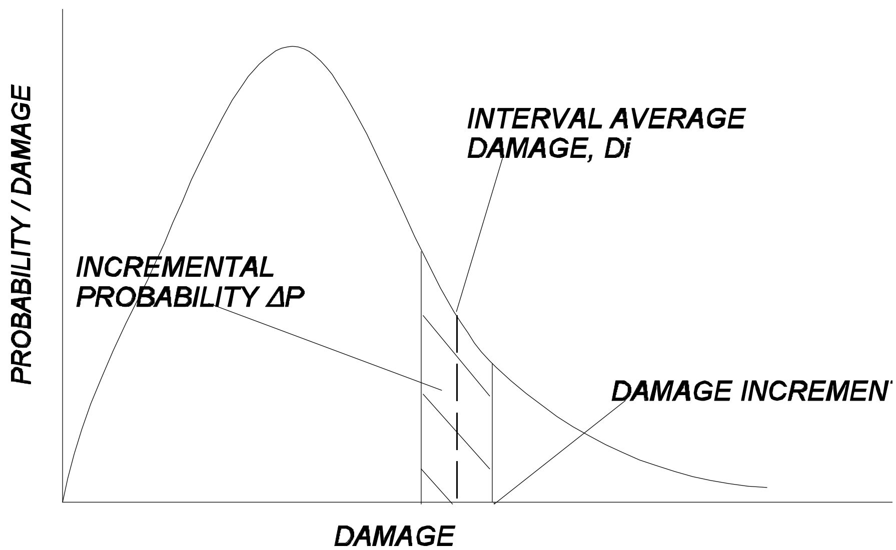

Figure F.1 Numerical Integration of Probability Density Function to Obtain EAD

The numerical integration is necessary because the damage-exceedance probability function is not defined by a continuous analytic function making an analytic integration impossible. Given that an exact analytic value cannot be obtained, how good is the approximation provided in equation (3)? The approximation can be made as accurate as possible by decreasing the interval Δp, or equivalently, increasing the number of intervals shown in Figure F.1.

Recognizing that equal probability increments implies that Δp=1/N, where N is the number of increments in Figure F.1, Equation (3) can be rewritten as:

$$EAD = \sum_{i=1}^{i=N} D_i \Delta p = \sum_{i=1}^{i=N} \frac{D_i}{N} \qquad \textbf{(4)}$$

which is the same as taking the average of a sample of N occurrences of annual damage. Monte Carlo simulation produces as a large a sample as desired to obtain a sufficiently accurate numerical integration to obtain EAD or other statistical characteristics of a probability distribution. The key aspect of the

Motel Carlo simulation is to obtain a random sample from a particular distribution. The algorithms used to generate samples from a probability distribution are discussed later.

Computing Expected Annual Damage, Exceedance Probability, and Event Probabilities

The inclusion of uncertainty in estimates of the variable contributing to damage makes it possible to obtain both a best estimate of expected annual damage and a distribution of possible values about this best estimate. Additionally, an expected set of exceedance probability functions and event conditional stages can be computed as a consequence of providing these estimates of uncertainty.

The relationship between estimation uncertainty and the distribution of EAD can be understood by considering a sensitivity analysis application to computing EAD with a flow-exceedance probability curve, rating curve and stage-damage relationship as shown in Figure F.2. The figure shows that high-bound, low-bound and best estimates of each relationship are combined to obtain a corresponding range in estimates of EAD. This range in estimates could be thought of as defining a rough distribution of possible EAD estimates. The difficulty with this sensitivity analysis approach is that the relative likelihood of the range in estimates is not known.

Monte Carlo simulation is used to improve on the sensitivity analysis by integrating all possible random occurrences of the contributing relationships as shown in Figure F.3. This differs from the basic Monte Carlo application described in the previous section by obtaining a random sample of relationships or random functions instead of obtaining a random sample of individual values. The algorithm used to obtain random samples of each relationship is described later.

The Monte Carlo algorithm used to obtain the distribution and best estimate of EAD, expected exceedance probability curves and event related conditional stage exceedance probability proceeds as follows:

1. <u>Obtain a random sample of the contributing relationships</u>

Each relationship is sampled to obtain a single realization of the discharge-exceedance probability, the stage-discharge (rating) and the stage-damage functions.

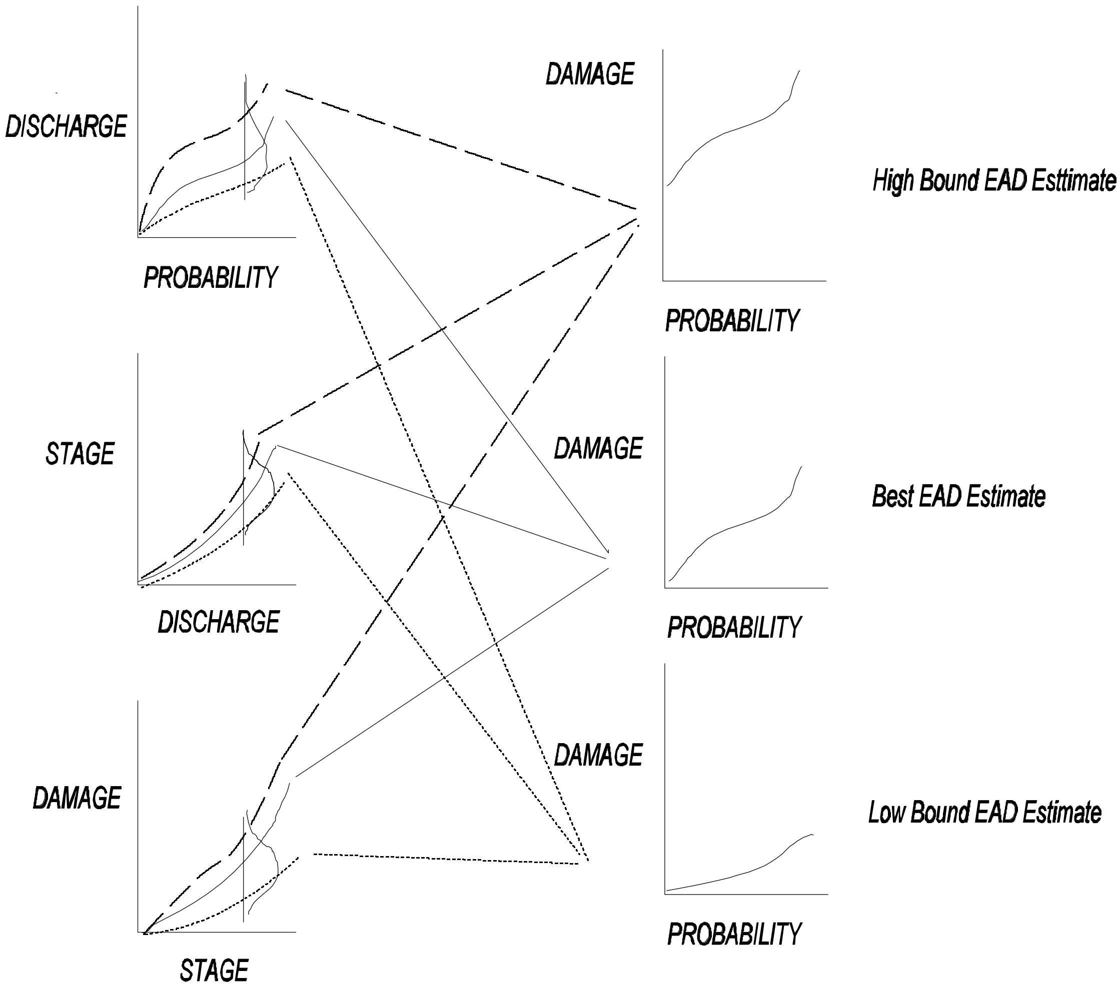

Figure F.2 EAD Computation Sensitivity Analysis

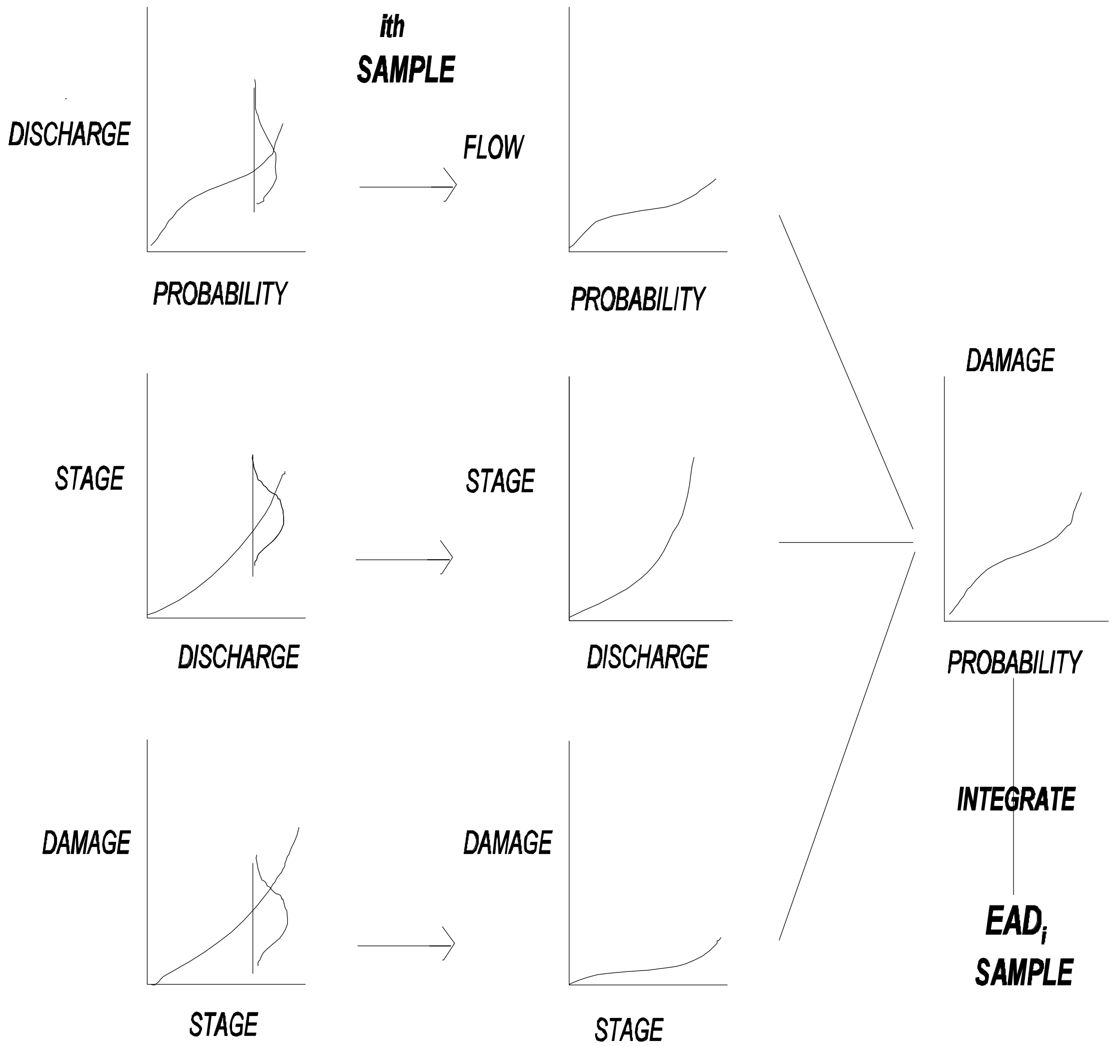

Figure F.3 Monte Carlo Simulation Algorithm for Estimating EAD

2. Compute exceedance probability curves

Compute the stage-exceedance probability function by using the rating curve to transform the sample discharge-exceedance probability function into a stage-exceedance probability curve; and, compute the damage exceedance probability function by using the sample stage-damage function to transform the stage-exceedance probability curve into a damage-exceedance probability function.

3. Save intermediary results for computing expected exceedance probability curves

Intermediary results are saved for the computation of expected exceedance probability functions by adding discharges, stages and damages for specified probabilities to values summed for previous simulation.

4. Save intermediary results for computing event conditional stage probabilities

Event conditional stages are saved for later estimation of conditional stage exceedance probabilities. The stages are conditional on specified exceedance probabilities (e.g., conditional on the 0.1, 0.02, 0.01 stage being exceeded). The stage for each of the events of interest is saved in a stage class interval. For example, consider that a stage of 21.56 corresponds to the 0.01 exceedance probability for the sample stage exceedance probability curve obtained in step 2. This value is saved in a predetermined class interval that may have minimum and maximum limits of respectively, 21.0 and 22.0.

5. Save intermediary results for computation of EAD

The EAD for the sample contributing relationships is computed by integrating the damage exceedance probability curve. This value is both added to a sum of EAD values from previous iterations and saved in a damage class interval.

6. Repeat sampling steps 1 through 5

Additional samples of exceedance probability curves and EAD are obtained by repeating steps 1 through 5. Sampling ceases when an accuracy criterion is met.

7. Compute expected exceedance probability curves

Divide the summed values obtained in step 3 for discharge, stage and damage for each exceedance probability by the number of samples.

8. Compute conditional event stage distributions

The process in step 4 of placing stages in class intervals results in a exceedance probability histogram of stages for each exceedance probability event of interest. Table F.1 provides an example of some possible results for the 0.01 exceedance probability event. As shown in the table, the exceedance probability histogram is converted into an event conditional exceedance probability function.

9. Compute best estimate of EAD and Distribution of EAD

The best estimate of EAD is computed as the average of the samples summed in step 5. The class interval exceedance probabilities for EAD are converted to a exceedance probability distribution using the same procedure for event conditional stages (see Table F.1).

In performing this simulation, only the stage vs total damage relationship is used to obtain the damage exceedance probabilities function and corresponding EAD. Damage-exceedance probability functions and EAD for damage categories are proportioned in the same ratio as the traditional (no uncertainty) category damage is to the tradition total damage values.

Table F.1
Calculating Event Conditional Stage Exceedance Probability from Monte Carlo Simulation Frequencies

Lower Limit Stage	Upper Limit Stage	Frequency	Cumulative Frequency	Cumulative Probability	Exceedance Probability
<21.0	21.0	200	200	0.01	0.99
21.0	22.0	5000	5200	0.26	0.74
22.0	23.0	10000	15200	0.75	0.25
23.0	24.0	5000	20200	0.99	0.01
24.0	25.0	100	20300	1.0	0.0
25.0	25.0>	0	20300	1.0	0.0

Monte Carlo Simulation Options for Calculating EAD

The Monte Carlo simulation can be expanded to include other contributing relationships in the calculation of EAD. Table F.2 describes the options for including other relationships. Notice that some relationships involve uncertainty calculations and others (levee effects and interior stage versus exterior stage relationships) are specified without uncertainty. The inclusion of additional relationships does not require any new aspect of performing the simulation except to require the creation of additional random samples of another relationship. For example, Figure F.4 displays the additional step of using the flow transform to convert a reservoir inflow-exceedance probability curve to a regulated exceedance probability curve.

Table F.2
Contributing Relationships Used in EAD Calculation

Contributing Relationship	Uncertainty Distribution
Flow/stage frequency curve	yes
Flow transform	yes
Rating curve	yes
Wave overtopping of flood wall or levee	yes
Levee impact on damage	no
[1]Exterior versus interior stage	no
Stage versus damage	yes

[1]Used to directly convert exterior river stage , interior levee failure stage, or with wave overtopping

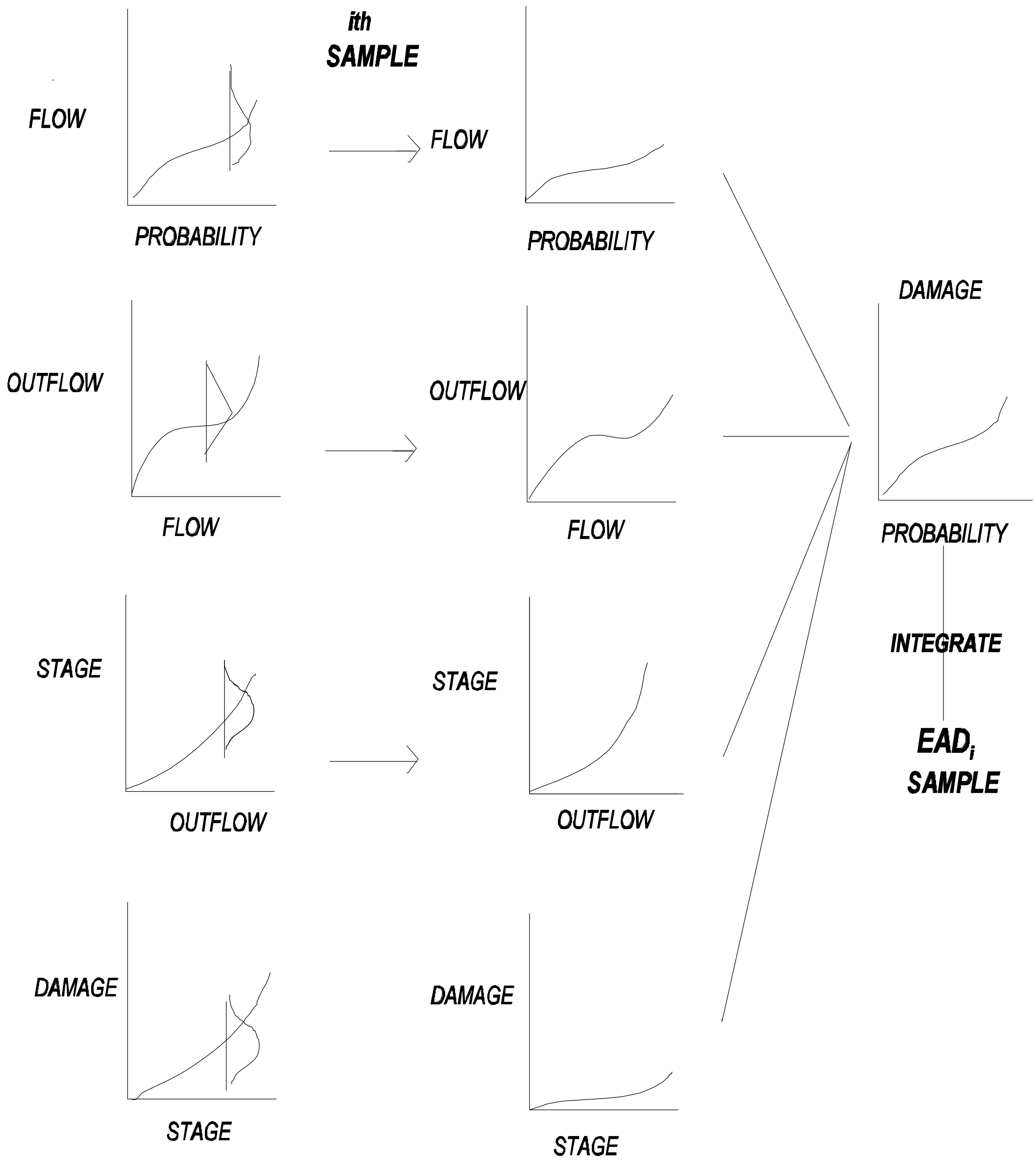

Figure F.4 Adding Computation of Regulated Outflow to Monte Carlo Algorithm for Computing EAD

Sampling Algorithm for Numeric Intergration

Overview

Application of Monte Carlo simulation requires a method for producing random samples and criteria for determining the number of samples needed to obtain a numerical integration with pre-specified accuracy. The algorithms (previously described) produce random samples of the contributing relationships that are combined to obtain samples of EAD, exceedance probability functions and event conditional stage probabilities. This sampling depends on the algorithm for generating random numbers. The generation of random numbers and the random sampling of contributing relationships is the means by which Monte Carlo simulation performs a numerical integration. As previously discussed, the numerical integration accuracy increases with the number of simulations. The criteria used to determine the number of simulations for a desired level of accuracy is described in the next section. The related problem of obtaining a numerically accurate integration of the damage-exceedance probability function is also discussed later.

Sampling from the Log-Pearson III Distribution

Random samples of a log-Pearson III (LPIII) exceedance probability curve are obtained from random samples of the mean and standard deviation of the logarithm of the flow, computing a log-normal relationship and adjusting for the skew of the distribution. This scheme produces the same sampling variability inherent in the calculation of confidence limits and expected probability as described in Bulletin 17B (IACWD, 1982), the federal guidelines for performing flood-flow exceedance probability analysis.

The random sampling is based on a Bayesian statistical approach for assessing uncertainty (see Stedinger, 1983). A goal of Bayesian estimation is to develop the distribution of possible population parameters (the posterior distribution) by combining statistics of the observed sample (e.g., observed stream flows), and other information on the probable range of population parameters (the prior distribution). In this instance, the prior distribution is based on the assumption that an equally likely set of parent populations could have produced the estimated sample mean, standard deviation and resulting log-normal distribution. The resulting posterior distribution of the population mean and standard deviation is given by:

$$P[\mu > m] = F(\mu) = \Phi(\overline{X}, \frac{S}{\sqrt{N}}) \quad \textbf{(5)}$$

$$P[\sigma^2 > s] = F(\sigma^2) = \frac{(N-1)S^2}{\chi^2_{(N-1)}} \qquad \textbf{(6)}$$

where $\overline{X}$ and S are respectively the sample mean and standard deviation of the logarithm of flow values obtained from a record length of N years, μ is the population mean, Φ( ) is the normal distribution defined by the parameters shown, σ is the population standard deviation, and $\chi^2_{(N-1)}$ is the chi-square distribution with N-1 degrees of freedom. Random estimates of the log-normal distribution are obtained by generating random estimates of normal and chi-square numbers, applying equations (5) and (6) to obtain μ and σ and computing the distribution (see Figure F.5).

This scheme for computing uncertainty does not account for the effect of shape or skew that is a characteristic of the LPIII distribution. This omission of the sampling uncertainty in skew is in keeping with the approach taken in the Bulletin 17B guidelines where sampling error is only estimated for a log-normally distributed variate. Consequently, the sampling scheme used for the LPIII distribution follows the Bulletin 17B method of computing uncertainty for a log-normally distributed variate and applying this uncertainty to an LPIII distribution with the same mean and standard deviation as the log-normal distribution. Given this estimation of uncertainty, the sampling of the LPIII distribution (see Figure F.6) proceeds as follows:

1. Compute log-normal and LPIII distributions from sample statistics

The log-normal and LPIII distributions are calculate using the following frequency factor equations:

$$\log_{10} Q^s = \overline{X} + Z_P S \qquad \textbf{(7)}$$

$$\log_{10} Q_G^s = \overline{X} + K_{G,P} S \qquad \textbf{(8)}$$

where Q^s and Q_G^s are respectively the flows for the log-normal and LPIII distribution, Z_p is the standard normal deviate and $K_{G,P}$ is the LPIII deviate for a sample skew G, and exceedance probability P.

2. Randomly select a sample normal distribution

Utilize equations (5) and (6) to obtain a sample of the population mean and standard deviation. Compute the log-normal distribution from the population values as:

$$\log_{10} Q^{r} = \mu + Z_P \sigma \qquad \textbf{(9)}$$

3. Calculate the random probabilities resulting from the randomly selected normal distribution

Compute the random probability associated with the randomly selected normal distribution for a discharge with exceedance probability computed from equation (7) as:

$$P_r = \Phi^{-1}\left(\frac{\log_{10} Q_P^{s} - \mu}{\sigma}\right) \qquad \textbf{(10)}$$

where $Q_p^s = Q^r$ is the flow value computed by equation (7) for exceedance probability P and Φ^{-1} is the inverse normal distribution (i.e., given a flow value, the inverse provides the exceedance probability).

4. Utilize the random probabilities to obtain a random sample of the LPIII frequency curve

Assign the random probability P_r to a flow value $Q_G^r = Q_G^s$, where Q_G^s was obtained from equation (8). Compute as many pairs of P_r, Q_G^r values as needed to adequately define the sample LPIII exceedance probability curve.

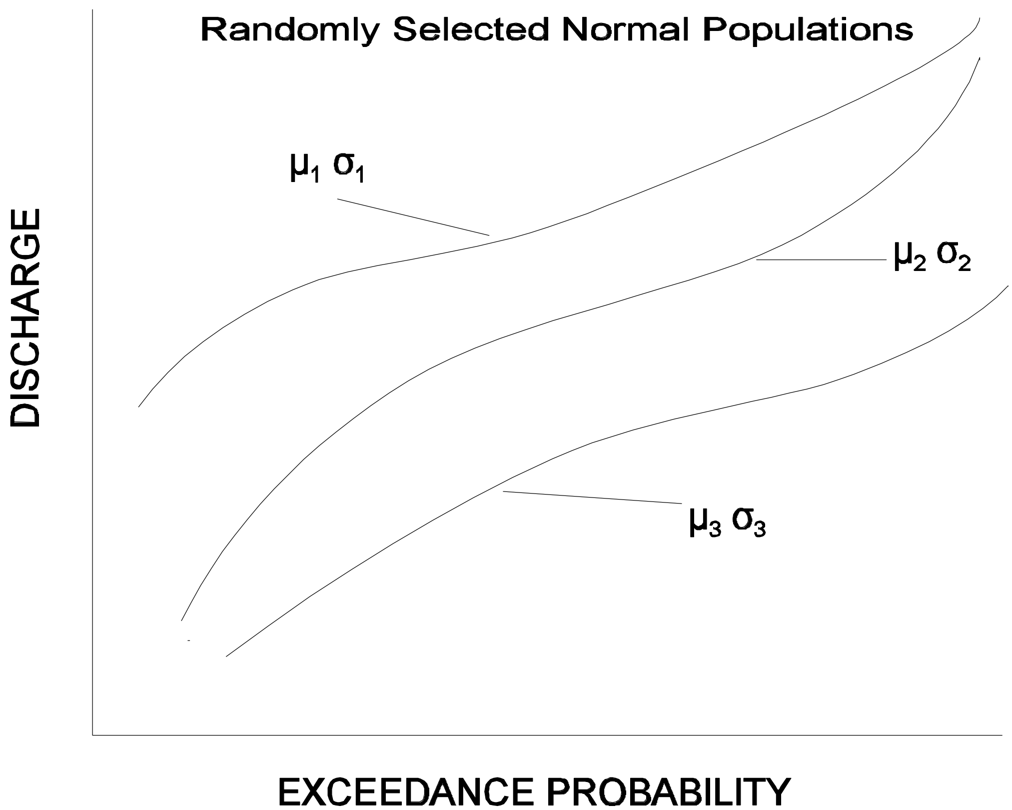

Figure F.5 Random Samples of Normal Populations from Population Parameters μ, σ

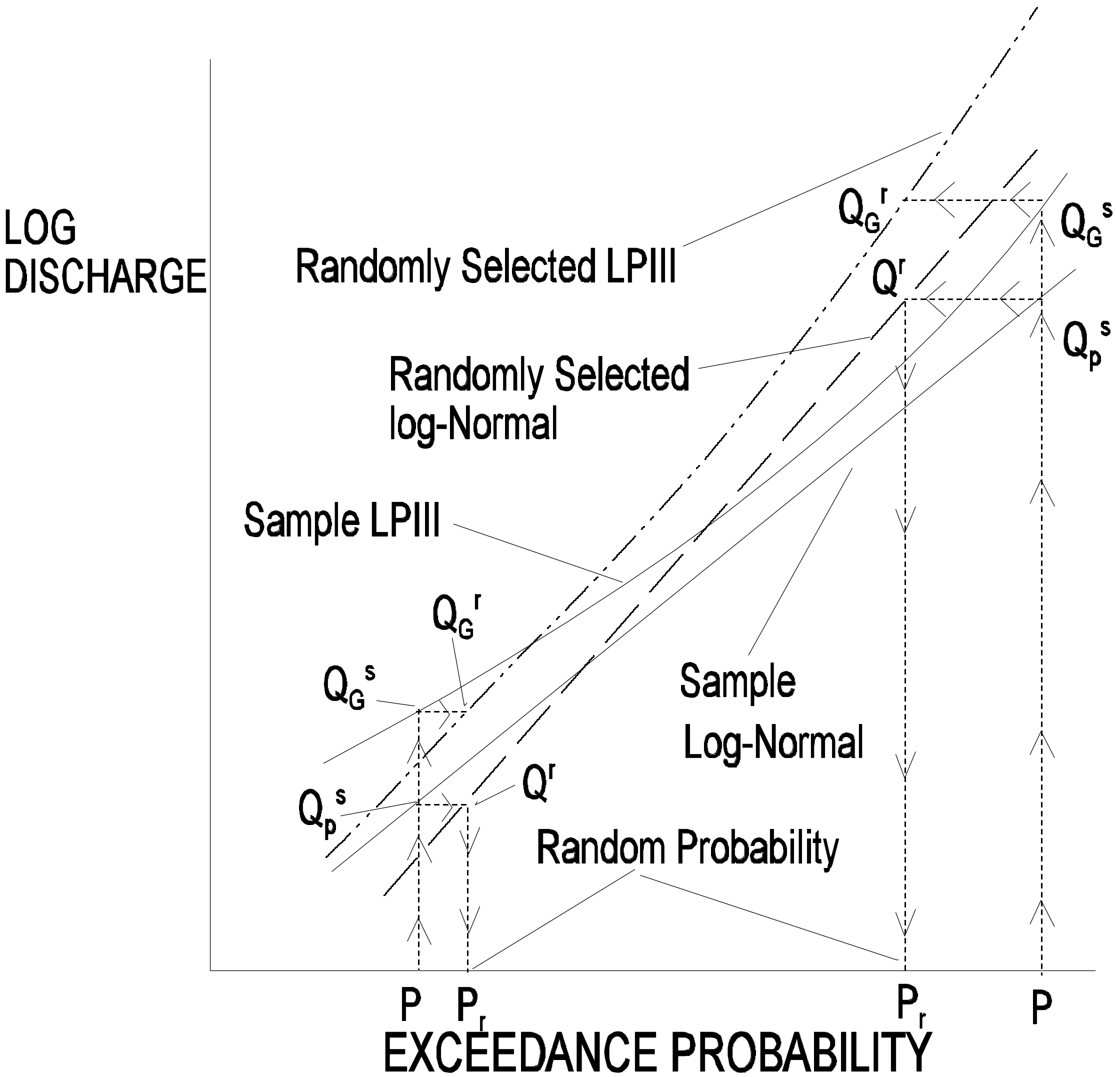

Figure F.6 Random Selection of LPIII Distribution from Random Log-Normal Distribution

Random Sampling of Graphical or Non-Analytic Relationships

The sampling of non-analytic or graphical relationships is necessarily ad hoc because a statistical sampling theory is not available. The algorithm used in this instance applies to any of the other contributing relationships used in the computation of EAD: 1) non-analytic stage or graphical exceedance probability curves; 2) discharge transforms; 3) rating curves; 4) wind waves and 5) stage damage relationships.

Random sampling of any of the graphical relationships is done by calculating the values for a particular confidence limit (see Figure F.7). The algorithm is simply employed by: 1) generating a uniform random number between 0 and 1; and 2) calculating the confidence limit values for the particular relationship of interest. For example, if 0.95 is the value resulting from the randomly selected value, then the 95% chance confidence level confidence limit is calculated as the randomly selected relationship for the algorithm described previously. Note, that the confidence limit for a contributing relationship is randomly selected independently of other confidence limits randomly selected for other contributing relationship used in the Monte Carlo simulation.

Classical statistical theory cannot be used to justify sampling possible population values from confidence limits as is done with this algorithm. Instead, justification for this algorithm must be sought from the sampling of the log-Normal distribution described in the previous section, which relies on a Bayesian approach. As was pointed out, the Bayesian approach results in the same uncertainty distribution for population values as is obtained with a classical statistical approach to obtain the uncertainty distribution used in the 17B guidelines. In the case of the approach for graphical exceedance probability curves, the sampling from confidence limits obtained from an uncertainty distribution might be justified in analogy with this Bayesian approach.

The difficulty with this algorithm is that the sampling based on confidence limit values is very restrictive on the possible shapes of the graphical relationship. This restriction on shape results in some overestimation in the variance of the derived distribution of EAD. However, generalizing the shapes used in the sampling algorithm depends on some parametric representation of the graphical relationships. The representation is not available, leaving the current algorithm as the best available at this time.

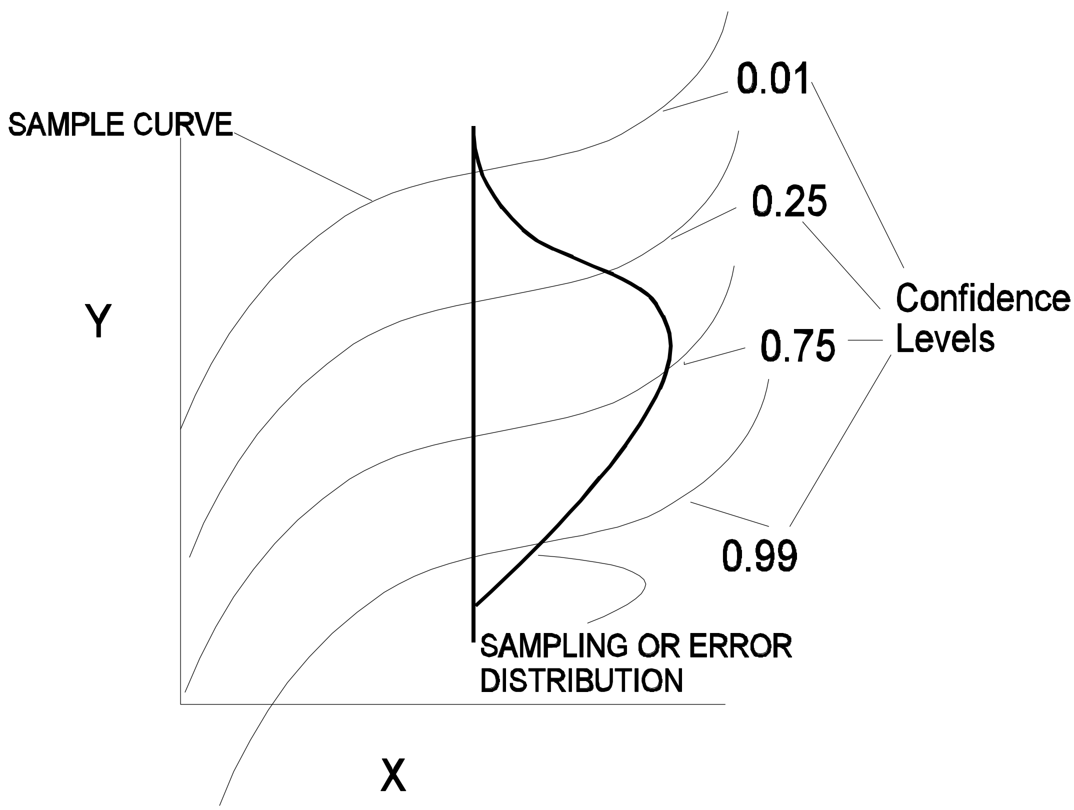

Figure F.7 Sampling of Non-Analytic or Graphical Relationships

Random Sampling of Uncertainty Relationships Using a Random Number Generator

The sampling of uncertainty distributions depends on the generation of uniform random numbers in the range 0.0 to 1.0 by the linear congruential method (Davis and Rabinowitz, 1967) and the transformation of the uniform numbers to the distribution desired. The linear congruential method takes the form:

$$X_{n+1} = \frac{(aX_n + b) \bmod m}{m} \qquad \textbf{(11)}$$

where X_n is the previous number selected, X_{n+1} is the current number to be generated, a and b are constants, *m* is a constant known as the modulus, and "mod" is the modulus or remainder function. The sequence is started for n=1 by a seed value that is set to a default value within the software. The selection of the constants and seed value are critical for an effective generations scheme. This generation scheme, as well as any other using a computer algorithm, is considered to produce pseudo-random numbers because the sequence repeats with period depending on the selection of the constants in equation (11). The constants are selected as shown in Table F.3 to obtain a long period of random numbers that is approximately equal to the size of the modulus, *m*. The resulting sequence of numbers has characteristics that are effective for performing numerical integration with Monte Carlo simulation.

The uniform random numbers can be used to randomly sample the graphical relationship directly. As described in the previous section, a number selected at random between 0.0 and 1.0 can be used to select the confidence level for selecting a graphical curve.

The application to the LPIII distribution requires that deviates from both a normal distribution and a chi-square distribution be obtained from a transformation of the numbers randomly sampled from a uniform distribution. The normal deviates can be obtained from the following transform due to Box and Muller (1958) (also see, Press et al., 1989):

$$n_i = u_i[-\frac{2\ln(s)}{s}] \qquad \textbf{(12)}$$

$$n_{i+1} = u_{i+1}[-\frac{2\ln(s)}{s}] \quad \textbf{(13)}$$

where u_i and u_{i+1} are numbers randomly selected from a uniform distribution defined between -1.0 and 1.0, n_i and n_{i+1} are numbers that will be normally distributed, and *s* is computed as:

$$s=(u_i^2+u_{i+1}^2)^{½} \qquad s \geq 1.0 \quad \textbf{(14)}$$

The application of this transform is accomplished by converting the uniform numbers generated over the range 0.0 to 1.0 in equation (11) by letting u_i = $2(X_i)$ - 1.0. When the resulting uniformly distributed numbers result in $s < 1.0$, the current pairing is discarded and a new pair is generated. On the average, about 1.27 uniform random variates are needed to generate a single normally distributed variate.

Chi-square deviates are obtained by applying the inverse theorem (see Mood et al., 1969, theorem 12, Chapter 5). This theorem is applied by interpolating a chi-square variate from a table of the chi-square cumulative distribution function given a random probability equal to a number generated from the uniform distribution using equation (11). The algorithm used to compute the chi-square distribution was obtained from Press et al. 1989, pg 160. The algorithm utilizes the following relationship between the chi-square and incomplete gamma function:

$$P[\chi^2_{N-1} < y] = G(a,x) = \int_0^x e^{-t} t^{a-1} dt \qquad 0 \leq x \leq \infty \quad \textbf{(15)}$$

where N is the period of record used to compute the sample standard deviation of the LPIII distribution, a = (N-1)/2, x = (y/2), and G( ) is the incomplete gamma function.

Table F.3
Constants for Linear Congruential Method[1]

seed	1331124727
a	65539
b	0
m	2147483647

[1]Constants appropriate for 32-bit machine. Used in Equation 11.

Numerical Error Tolerance for Simulations

The numerical integration accuracy of the Monte Carlo simulation improves with the number of simulations. The accuracy criteria developed for the simulation relies on the central limit theorem for the mean and the asymptotic normality of uncertainty distributions about exceedance probability curves. The central limit theorem (see Mood et al., 1969) states that the sample mean of any random variable is asymptotically normally distributed about the population value. In the case of this application of Monte Carlo simulation, the sample EAD results from a finite number of simulations, and the population value is the value that would be obtained from an infinite number of simulations (i.e., the no numerical error solution).

The following confidence limit results from asymptotic normality of the sample EAD:

$$P[-z_{1-} \leq \frac{M_{EAD} - \mu_{EAD}}{\frac{S}{\sqrt{n}}} \leq z_{1-}] \sim 1.0 - \tag{16}$$

where M_{EAD} is the average EAD obtained from n simulations, μ_{EAD} is the numerical error EAD, S is the standard deviation of the damage exceedance probability curve estimated after n simulations, and $z_{1-\alpha}$ is the standard normal deviate for confidence level . This confidence limit can be rearranged to produce an error bound of the numerical integration error:

$$\frac{z_{1-} S}{M_{EAD}\sqrt{n}} = \frac{M_{EAD} - \mu_{EAD}}{M_{EAD}} \leq \tag{17}$$

where is a tolerance for the confidence level . The error bound is set in the software such that =0.95, =0.01 and n$\leq$200,000. If the limiting number of simulations is reached the computation of EAD terminates with a warning.

A similar error bound is computed for exceedance probability function. In this case, the computed quantile (e.g., flow, stage or damage) is the mean value derived for the exceedance probability of interest. The error bound focuses on the exceedance probability where the corresponding quantile has the largest estimation standard error. This estimation standard error is set to S in equation (17) and computed as part of the simulation. The confidence limit and tolerance are set equal to that used for the error bound of EAD. The simulations will terminate only when the error tolerance for both estimating exceedance probability function and EAD is met or when the maximum number of simulations is reached.

The error bounds constrain the numerical integration error of the simulation but does not reduce the uncertainty in estimates of EAD or exceedance probability curves. The uncertainty in estimate is a function of the error in models and estimates of parameters as indicated by the uncertainty distributions provided. The uncertainty shown by the sensitivity analysis depicted in Figure F.2 is not altered by the number of simulations performed. Rather, the number of simulations reduces the numerical error involved in combining the relationships via the algorithm depicted in Figure F.3.

Integrating the Damage-Exceedance Probability Function to Obtain EAD

The final computation in an individual Monte Carlo simulation is to integrate the damage-exceedance probability function to obtain a sample value of EAD_i as shown in Figure F.3. The damage-exceedance probability function is not analytic being derived from rating curves, stage-damage relationships, etc., that are not analytic. Consequently the following trapezoidal integration scheme is used to obtain an estimate of EAD_i:

$$EAD_i = \int_0^\infty D f_i(D)\, dD \sim \sum_{j=1}^{j=h} \overline{D_j}\,\overline{f_{i,j}}\, \Delta D_j \sim \sum_{j=2}^{j=h-1} \overline{D_j}\,(p_j - p_{j+1}) + D_1 p_1 + D_h p_h \qquad \textbf{(18)}$$

where $f_i(D)$ is the probability density function (PDF) obtained from the ith simulation, for annual damage, D; h is the number of incremental intervals of size ΔD used to approximate the differential dD; $\overline{D}_j$ and $\overline{f}_{i,j}$ are the average values of D and $f_i(D)$ over this interval, and the difference of exceedance probabilities over this interval $(p_j - p_{j+1}) = \overline{f}_{i,j}\, \Delta D$; and, $D_1 p_1$ and $D_h p_h$ are end point approximations to the end intervals of integration, zero and infinity. The assumption is made in the software that $D_1 = 0$, resulting in $D_1 p_1 = 0$.

The trapezoidal rule approximation accuracy improves with increasing number of intervals, h. The number of intervals is determined by computing EAD for damage exceedance probability curves determined by a sensitivity analysis such as shown in Figure F.2 prior to performing the Monte Carlo simulation. The sensitivity analysis is performed by obtaining damage exceedance probability curves by combining confidence limit estimates of the contributing relationships at the same confidence level. The confidence limits investigated are obtained for confidence levels, 0.5, 0.75, 0.25, 0.9, 0.1, 0.99, 0.01, 0.999, 0.001.

The number of intervals, h, is obtained by performing a recursive integration for each confidence limit investigated in the sensitivity analysis. The recursive procedure involves: 1) selecting an interval size; 2) computing EAD; 3) dividing the interval size in half, where appropriate, and re-computing EAD; 4) computing the relative difference between EAD values obtained in steps (2) and (3); and 5) determining if the relative difference in step (4) is less than 1%; if this tolerance is met; then the interval used in step (2) is selected; otherwise steps, 2-4 are repeated with the interval size used in step (3) used in step (2). The division of interval sizes in step (3) is only performed when the interval size reduction will make a significant difference to the computation of EAD. This limits the number of intervals used which is important to the computational efficiency of Monte Carlo simulation. The more intervals used, the more computational time required to perform a simulation. Intervals are divided until the error tolerance is met or the maximum number of 200 are obtained. Experience has shown that 200 intervals provides sufficient accuracy given the data typically available.

Uncertainty Distributions

General

The estimation of uncertainty distributions for the contributing relationships will involve a certain amount of judgment, except for the case of a flow or stage exceedance probability curve where the uncertainty is determined from the length of record. The judgment used in estimating uncertainty for other contributing variables should correspond to the same factors contributing to uncertainty in the exceedance probability curves. The uncertainty in the exceedance probability functions is due to the estimation uncertainty in the parameters, which are the mean and standard deviation for the LPIII (the skew being ignored).

This focus on parameter uncertainty effectively examines the uncertainty in the mean relationship given a set of scattered observations. In other words, the focus is on the uncertainty in fitting a exceedance probability function to an observed set of plotting positions and does not reflect the scatter of the plotting positions about the best estimates.

To understand the difference between uncertainty in fitted relationships and the uncertainty due to scatter, consider a split sample exceedance probability analysis of a gage having 100 years of record. Estimate both pairs of frequency curves and determine the top ranked event from separate 50-year records. In general, the difference between the 1% chance flow estimated by the frequency curves will be considerably less than the difference between the top ranked events. The smaller variation in the fitted relationships, as compared to the plotting positions, represents the difference between uncertainty for best fit relationships and that for scatter about these relationships. If uncertainty in the contributing relationships such as rating and stage-damage curves is based on scatter, then the specified uncertainty will be too great. This in turn will probably increase the magnitude of the EAD best estimate and certainly increase the variance of the EAD distribution.

Therefore, the principle focus of estimating uncertainty should be on the potential variation in the best estimate of the contributing relationship. Consequently, if a sensitivity analysis is performed to determine the uncertainty in a contributing relationship, such as in varying Manning n to determine errors in rating curves, then the parameters varied should be reasonably likely to occur together. Combining extreme parameter values probably reflects scatter rather than the reasonable variation in a fitted relationship.

The error distribution about exceedance probability curves is determined by the effective record length and the type of exceedance probability curve specified. In the case of the LPIII distribution, the uncertainty is computed as described previously. Also, refer to ETL 1110-2-537 (Corps of Engineers, 1995) for the method used to calculate the uncertainty distribution for non-analytic (graphical exceedance probability curves). Normal, log-normal and triangular error distributions are available for specifying uncertainty about other contributing relationships, as is described in the next two sections.

Triangular Error Distribution

The triangular distribution is the simplest available for use with contributing relationships that are not exceedance probability functions (see Figure F.8). This triangular distribution is specified for either: 1) each paired value describing the contributing relationship (e.g., discharge-stage function); or 2) for a specified value in the paired relationship (e.g., for 1000 cfs corresponding to a stage of 10.0 feet). In the case of the specified value, the bounds on the error distribution are linearly interpolated to zero for values less than this specified value and remains unchanged for values greater than this value.

The parameters of the distribution are the mode and the range. The mode is the most frequently occurring value, or the peak of the probability density function for the triangular distribution. The range is simply defined by the minimum and maximum possible values for the dependent variable in the paired relationship.

Inspection of Figure F.8 shows that the triangular distribution need not be symmetric. The effect of the asymmetry is to cause the mean or expected value associated with the triangular distribution to be different than that for the mode. Consequently, Monte Carlo simulation will produce on the average a contribution relationship that is different than might be assumed to occur when specifying the mode as a no uncertainty estimate of the relationship.

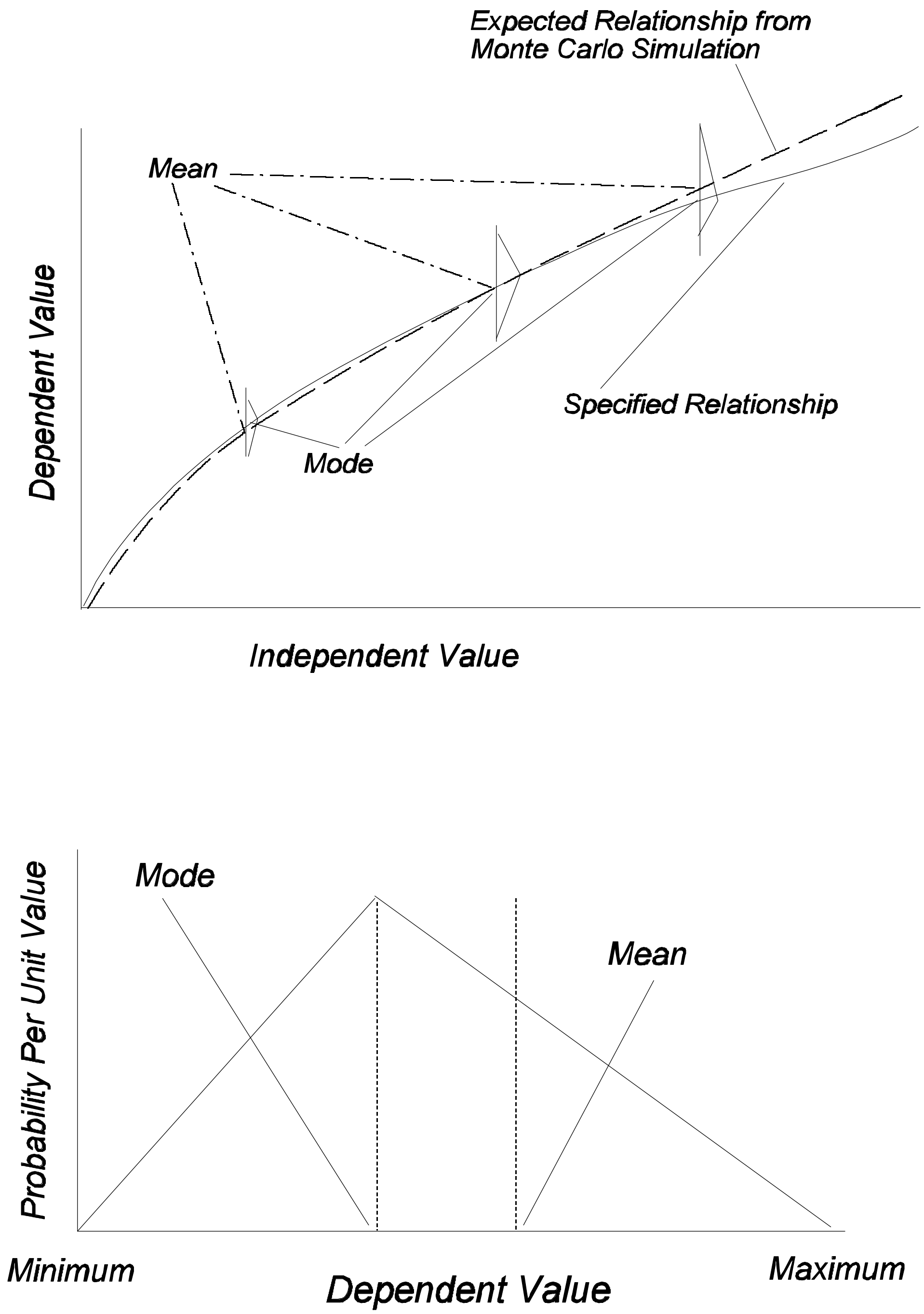

Figure F.8 Triangular Distribution Application

Normal and Log-normal Distributions

The specification of the normal and log-normal distributions is analogous to that of the triangular distribution. These distributions are specified to represent uncertainty for contributing relationships that are not exceedance probability curves and can be specified for each paired value or for a single specified value.

The normal distribution is specified by a mean and standard deviation of the errors (see Figure F.9). The log-normal distribution also is specified by a mean and standard deviation of the logarithms (base 10) of interest. Consequently, estimation of the errors needs to be performed in log space for this distribution. For example, the paired values of discharge and stage should be plotted on $\log_{10}$-$\log_{10}$ scale; and the best fit relationship and the errors should be determined from this scale. The relationship is then specified by the untransformed best fit values (i.e. by taking anti-logs of the best fit) together with the standard errors of the logarithms.

The normal distribution is symmetric with respect to the mean. Consequently, the mean or expected relationship obtained from the Monte Carlo simulation will be the same as the specified relationship. This differs from the average result obtained with an asymmetric triangular uncertainty distribution as explained in the previous section and shown in Figure F.8. The estimation of the log-normal distribution is most conveniently performed in log-space, thus reducing the problem in estimating a normally distributed log variate. However, the log-normal uncertainty distribution is asymmetric when plotted on a linear scale, and, like an asymmetric triangular distribution, will result in an average relationship that differs from the specified relationship when performing a Monte Carlo simulation.

Application to Stage Versus Damage Relationships

The Monte Carlo simulation algorithm reduces the computational effort required by only computing total damage. However, stage versus damage is specified for each damage category with a corresponding uncertainty in the estimates. The total damage is easily obtained by aggregating the specified (no uncertainty) estimates in the case of triangular and normally distributed uncertainty distributions. Logarithms of the specified estimates are added in the case of log-normally distributed uncertainty distributions.

The uncertainty distributions are not so easily aggregated. The assumption is made that the uncertainty estimates are uncorrelated. Consequently, the standard errors of the normal distribution and the log standard errors for the log-normal distribution can be added by summing these standard errors squared and taking the square root (variances added). The triangular distribution is handled in the same manner in that the maximum and minimum ranges are added to obtain the range of an equivalent triangular distribution.

Although normal and log-normal distributions can be added to obtain the same distributions, the same is not true of triangular distributions. If enough triangular distributions were involved in obtaining the total, then the resulting distribution would be normal according to the central limit theorem.

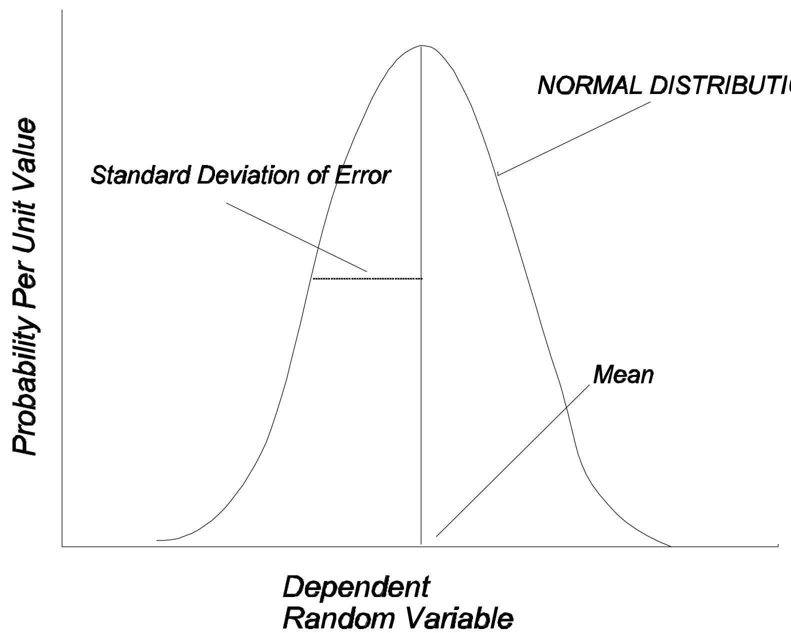

Figure F.9 Normal Distribution of Errors

However, since the number of categories involved is not large, the resulting distribution is not likely to be normal. Consequently, the assumption is made that the distribution of uncertainty for the total damage in the stage versus damage relationship is triangular if the category damage assumes a triangular uncertainty distribution.

Levee Analysis

Computation of damage exceedance probability functions with levees is straight forward when the levee only fails due to overtopping, but requires some additional computations when geotechnical failure can occur. The computation of the damage exceedance probability curve for levee failure due to overtopping only is easily done by setting the zero damage point to a stage corresponding to the top of levee. The integration of the damage exceedance probability curved using equation (18) to obtain EAD is then applied as without a levee.

The computation of the damage exceedance probability curve when geotechnical failure is possible needs to consider the probability of failure below the top of levee. The damage exceedance probability curve is calculated in this situation as follows (see Figure F.10):

$$P[d_j \leq D \prec d_{j+1}] = (p_j - p_{j+1})p^f_{j+½} \qquad p_j \leq p_m \tag{19}$$

where P[$d_{j-} \leq$ D $\prec$ d_{j+1}] is read as “the probability that the annual damage, *D*, will be in the interval $d_{j-1\leq}$ to d_j,”; p_m is the exceedance probability corresponding to the stage that cannot cause damage due to geotechnical or overtopping failure; p_j and p_{j+1} are the exceedance probabilities for stages that cause damage corresponding to $d_{j\leq}$ and d_{j+1} in the absence of the levee; and $p^f_{j+½}$ is the failure probability of the levee for the stage with exceedance probability midway between p_j and p_{j+1}. Equation (18) then can be applied to this damage exceedance probability curve to obtain EAD by letting:

$$\overline{D_j} = \frac{d_j + d_{j+1}}{2} \tag{20}$$

and substituting:

$$(p_j - p_{j+1})p^f_{j+½} \rightarrow (p_j - p_{j+1}) \tag{21}$$

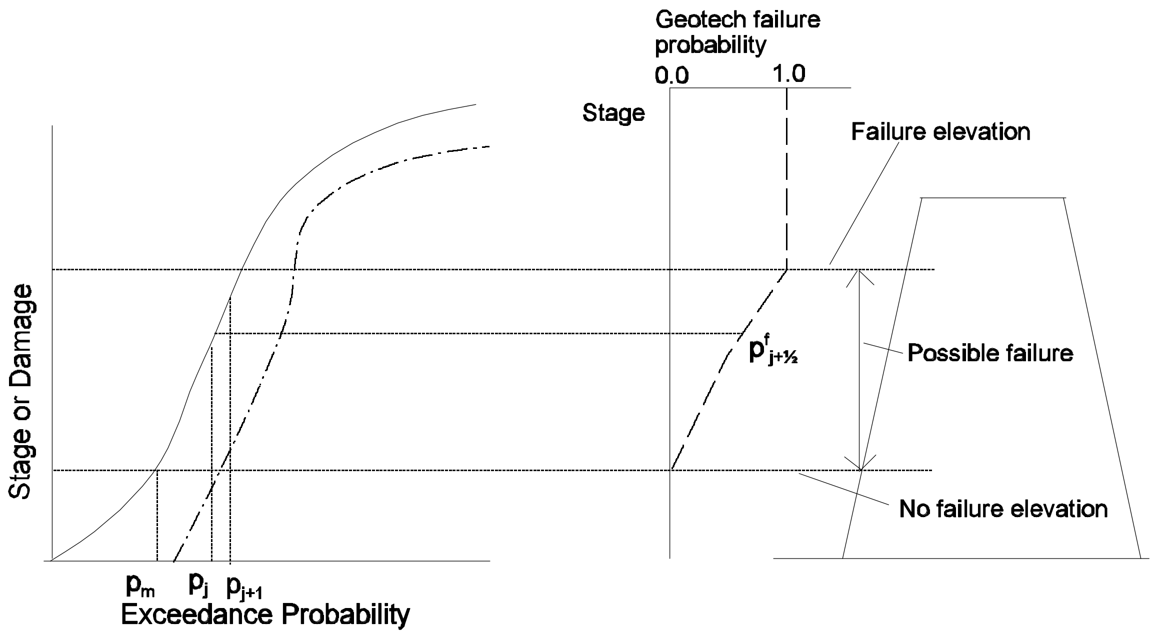

Figure F.10 Damage Considering Levee Geotechnical Failure

Project Reliability and Flood Risk Computations

Reliability is computed as the exceedance probability for a target stage or the likelihood of levee failure. Flood risk is defined as the probability of one or more exceedances of the target stage or levee failures in a specified number of years.

The target stage is determined by interpolation from the stage versus damage relationship using a specified fraction of a damage for a specified exceedance probability . This damage is determined from a damage-exceedance probability function obtained by combining traditional estimates of the contributing relationships (i.e., contributing relationships without uncertainty) for the without-project condition.

The exceedance probability for this stage or the levee failure probability is specified as both a “median” and “expected” value. The median value is obtained from the stage-exceedance probability curve obtained by the traditional (no uncertainty) method. The expected value is obtained by averaging the target stage or levee failure probability over all the Monte Carlo simulations.

The risk of flooding one or more times in N_R years is computed as:

$$R = 1 - (1 - p)^{N_R} \tag{22}$$

where p is either the probability of exceeding the target stage or levee failure. An expected value of R is reported as the average over all Monte Carlo simulations.

Computation of Equivalent Annual Damage

Equivalent annual damage is computed by discounting future EAD values given the appropriate interest rate and time for discounting. The computation is described in detail elsewhere (see Hydrologic Engineering Center, 1984). This computation is applied to not only the best estimate of EAD but to the distribution of possible EAD values obtained as part of the Monte Carlo simulation. This results in a distribution of equivalent annual damage.

Inundation Reduction Benefit Computations

Inundation reduction benefits are computed as the difference between with-and without-project equivalent annual damage. This differencing is performed between the distribution of equivalent annual damage values obtained for both with-and without-project condition resulting in a distribution of equivalent annual damage.

The differencing of uncertainty distributions in this manner recognizes that irrespective of the plan, the future exceedance probability of events causing floods will be the same for all plans. Consequently, differencing these distributions results in the same answer as would be obtained by obtaining the distribution of net benefits by performing Monte Carlo simulation of damage differences.

References:

Box, G.E.P., and Muller, M.E., (1958). A note on the generation of random normal deviates, Ann. Mathematical Statistics, V29, p610-611.

Corps of Engineers, 1995. Uncertainty estimates for non-analytic frequency curves, ETL 1110-2-537,Department of the Army, Washington, D.C., 15p.

Hydrologic Engineering Center, 1984. Expected annual flood damage computation - EAD, Users Manual, U.S. Army Corps of Engineers, Davis, CA. 128p.

IACWD, 1982. Guidelines for determining flood flow frequency analysis, Bulletin 17B, Interagency Advisory Committee on Water Data, U.S. Department of the Interior, Geological Survey, Office of Water Data Coordination, Reston, VA.

Davis, Philip, J., and Rabinowitz, P., 1967. Numerical Integration, Ginn Blaisdell, Waltham, Massachusetts, 230p.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., 1989: Numerical Recipes (FORTRAN), Cambridge University Press, New York, 677p.

Stedinger, J.R., 1983. Design events with specified flood risk, Water Resources Research, V19(2), April, p511-522.