NUREG/CR-5032
SAND87-2428
RG, 1S
Printed January 1988

Modeling Time to Recovery and Initiating Event Frequency for Loss of Off-Site Power Incidents at Nuclear Power Plants

Ronald L. Iman, Stephen C. Hora

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-76DP00789

Prepared for
U. S. NUCLEAR REGULATORY COMMISSION

Available from

Superintendent of Documents
U.S. Government Printing Office
Post Office Box 37082
Washington, D.C. 20013-7082

and

National Technical Information Service
Springfield, VA 22161

NUREG/CR-5032
SAND87-2428
RG,1S

MODELING TIME TO RECOVERY AND INITIATING EVENT FREQUENCY FOR LOSS OF OFF-SITE POWER INCIDENTS AT NUCLEAR POWER PLANTS

Ronald L. Iman[1]
Stephen C. Hora[2]

January 1988

Sandia National Laboratories
Albuquerque, NM 87185
operated by
Sandia Corporation
for the
U.S. Department of Energy

Prepared for
Division of Risk Analysis and Operations
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington. DC 20555
Under Memorandum of Understanding DOE 40-550-75
NRC FIN A-1393

---

[1]Reactor Safety Integration Division 6415
Sandia National Laboratories
Albuquerque, NM 87185

[2]Division of Business and Economics
University of Hawaii at Hilo
Hilo, HI 96720

ABSTRACT

Industry data representing the time to recovery of loss of off-site power at nuclear power plants for 63 incidents caused by plant-centered losses, grid losses, or severe weather losses are fit with exponential, lognormal, gamma and Weibull probability models. A Bayesian analysis is used to compare the adequacy of each of these models and to provide uncertainty bounds on each of the fitted models. A composite model that combines the probability models fitted to each of the three sources of data is presented as a method for predicting the time to recovery of loss of off-site power. The composite model is very general and can be made site specific by making adjustments on the models used, such as might occur due to the type of switchyard configuration or type of grid, and by adjusting the weights on the individual models, such as might occur with weather conditions existing at a particular plant. Adjustments in the composite model are shown for different models used for switchyard configuration and for different weights due to weather. Bayesian approaches are also presented for modeling the frequency of initiating events leading to loss of off-site power. One Bayesian model assumes that all plants share a common incidence rate for loss of off-site power, while the other Bayesian approach models the incidence rate for each plant relative to the incidence rates of all other plants. Combining the Bayesian models for the frequency of the initiating events with the composite Bayesian model for recovery provides the necessary vehicle for a complete model that incorporates uncertainty into a probabilistic risk assessment.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The authors would like to acknowledge the assistance of John Flack of the NRC for conversations to clarify many issues with respect to modeling time to recovery of loss of off-site power, F. Eric Haskin of Sandia National Laboratories for encouraging us to write this report and for his valuable comments, to Arthur Payne of Sandia National Laboratories for helping to clarify many engineering issues, Michael J. Shortencarier of Sandia National Laboratories for his assistance with the many figures that appear in this report, and to Ruby Cochrell of Sandia National Laboratories for her editorial and format suggestions that greatly improved the final form of this report. This report was prepared in support of the Phenomenology and Risk Uncertainty Evaluation Program sponsored at Sandia National Laboratories by the Nuclear Regulatory Commission.

1. INTRODUCTION

Industry data representing the time to recovery of loss of off-site power (LOSP) at nuclear power plants are presented in Appendix A of the Nuclear Regulatory Commission Report NUREG-1032. The loss of off-site power as defined in that NRC report is

> *"... the interruption of the preferred power supply to the essential and nonessential switchgear buses necessitating or resulting in the use of emergency AC power supplies."*

NUREG-1032 further states

> *"... total loss of offsite power is a relatively infrequent occurrence at nuclear power plants ... Historically, a loss of offsite power has occurred with a frequency of about once per 10 site years. The typical duration of these events has been on the order of one-half hour."*

The data reported in NUREG-1032 were obtained from the Licensee Event Report System and are classified as being plant-centered (PC), grid (G), or caused by severe weather (W). Again citing from NUREG-1032,

> *"Plant-centered events are those in which the design and operational characteristics of the plant itself play a role in the likelihood or duration of the loss-of-offsite-power event. Area or weather effects include the reliability of the grid and external influences on the grid or at the site (such as severe weather) that have an effect on the likelihood and duration of the loss of offsite power."*

The data reported in NUREG-1032 consist of the time (in hours) to recovery of loss of off-site power for 59 reported incidents through 1935. Since that time, four additional incidents have been reported. All data are shown in Table 1, with the most recent values identified with an asterisk. The data are current through June 1987. Summary statistics for the LOSP are shown in Table 2.

The LOSP data are displayed by each category in Figure 1 as "true" boxplots. The actual box portion of the plot is formed by the lower and upper quartiles of the sample data, with the median indicated by a vertical line within the box. The left-hand endpoint of the boxplot is located at the maximum of

$$X_{.25} - 1.5 \times (\text{interquartile range}) \text{ and } X_{min}.$$

For example, consider the first boxplot in Figure 1, which is for the plant-centered data. Since $X_{min} = .002$ for the plant-centered data, the left-hand endpoint of the boxplot for the plant-centered data is found as

Table 1. Time to Recovery (in hours) for 63 Recorded Incidents of Loss of Off-Site Power at Nuclear Power Plants

Plant Centered (n=43)

0.002	0.067	0.200	0.400	0.767*	1.750
0.003	0.070	0.250	0.430	0.900	2.750
0.003	0.080	0.250	0.480	0.900	7.467*
0.004	0.083	0.250	0.500	0.930	
0.013	0.130	0.270	0.500	1.030	
0.015	0.150	0.280	0.500	1.150	
0.017	0.167	0.330	0.570	1.480	
0.020	0.183	0.334	0.670	1.667*	

Grid (n=13)

0.130	0.300	0.330	0.920	1.500	2.083
0.180	0.330	0.550	1.030	2.000	6.470
0.250					

Severe Weather (n=7)

1.750	4.000	4.317*	5.000	5.500	8.900
2.667					

Table 2. Summary Statistics for the LOSP Data in Table 1

	Plant	Grid	Weather	All Data
Lower quartile ($X_{.25}$):	0.080	0.275	2.667	0.167
Median ($X_{.50}$):	0.280	0.550	4.317	0.430
Upper quartile ($X_{.75}$):	0.767	1.750	5.500	1.480
Sample mean ($\bar{X}$):	0.651	1.236	4.591	1.210

$$\max\{0.080 - 1.5(0.767 - 0.080) = -0.951, .002\} = .002.$$

The right-hand endpoint of the boxplot is located at the minimum of

$$X_{.75} + 1.5 \times (\text{interquartile range}) \text{ and } X_{max}.$$

Since $X_{max} = 7.467$ for the plant-centered data, the right-hand endpoint of the boxplot for the plant-centered data is found as

$$\min\{0.767 + 1.5(0.767 - 0.080) = 1.798, 7.467\} = 1.798.$$

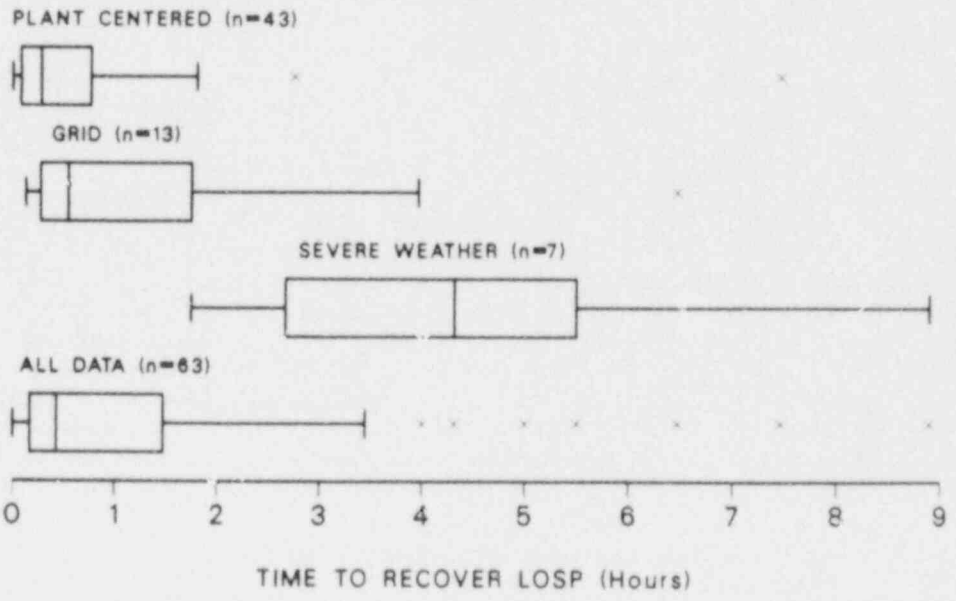

Figure 1. Boxplots of Data for the Time to Recovery LOSP

Any data values beyond the right-hand endpoint (or below the left-hand endpoint) indicate a high degree of skewness, and for symmetric distributions, would be considered as outliers. The basis for the identification of outliers is a boxplot for normal populations, where 1.5 × (interquartile range) subtracted from the lower quartile and added to the upper quartile will cover 99.3% of the distribution. In Figure 1 the symbol "×" is used to identify the outliers. For the plant-centered data, the two largest observations, 2.750 and 7.467 are identified as outliers. More importantly, the general form of each of the boxplots (except for severe weather, which has only seven observations) is consistent with data that are highly skewed.

Another approach to displaying the LOSP data is to use an empirical distribution function (e.d.f.) such as shown in Figures 2 to 4 (the weather data are not shown since there are only seven observations). The steps in the e.d.f. occur at each of the data values in the sample, and the step heights used in graphing the e.d.f.s are equal to 1/n, where n is the number of sample observations being plotted. One immediate advantage of the e.d.f. representation over an approach based on grouping the data in a histogram or boxplot display is that each data value is displayed.

In addition, the e.d.f. removes the subjectivity associated with grouping the data in a histogram, thus its construction is not analyst dependent. Moreover, the distribution function of a candidate probability model can be plotted jointly with the e.d.f., and a goodness-of-fit test designed for use with distribution functions can be used to determine the adequacy of fit. Like the boxplots, the e.d.f.s show the PC data to consist of mostly small values and the overall data to be more spread out.

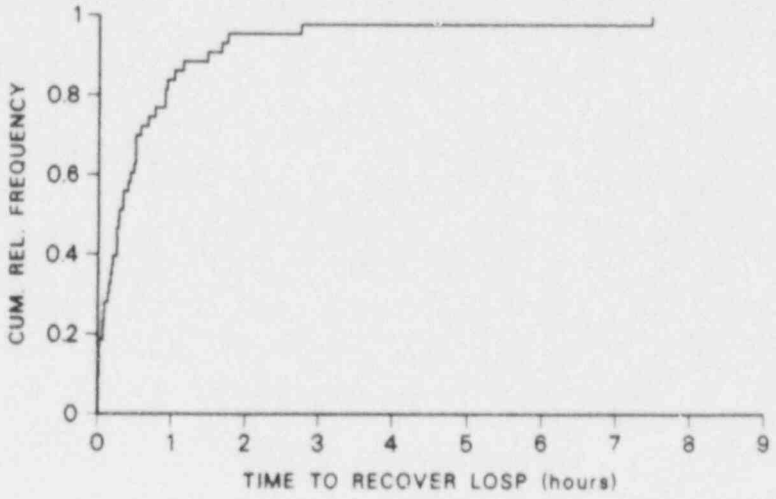

Figure 2. E.D.F. for Time to Recover LOSP (Plant Centered)

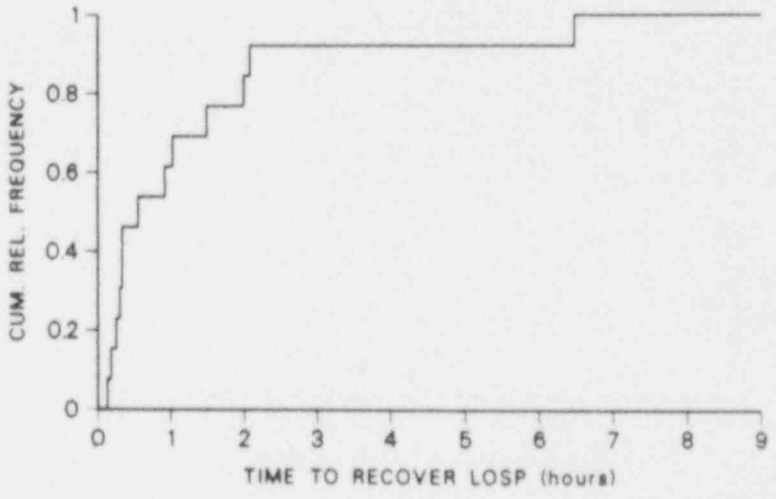

Figure 3. E.D.F. for Time to Recover LOSP (Grid)

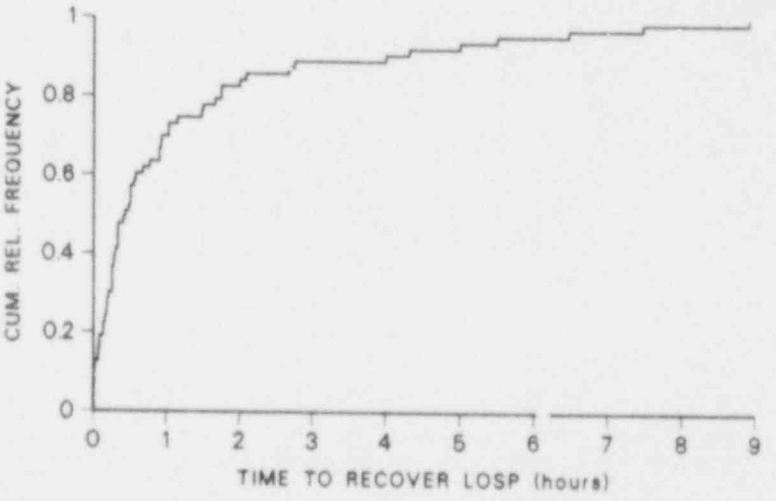

Figure 4. E.D.F. for Time to Recover LOSP (All Data)

Section 2 of this report considers modeling the skewed LOSP data with exponential and lognormal distributions. Sections 3 and 4 provide details of modeling the LOSP data with gamma and Weibull distributions respectively, while Section 5 presents a Bayesian comparison of the various probability models. Section 6 presents a Bayesian approach to modeling the uncertainty in the probability estimates provided by the gamma and Weibull distributions. Section 7 provides a composite model for the LOSP data with plant specific modifications of the composite model presented in Section 8. The frequency of the initiating events leading to LOSP is modeled in Section 9.

2. FITTING EXPONENTIAL AND LOGNORMAL PROBABILITY MODELS

Two probability models frequently used for failure data are the exponential and lognormal distributions. Each of these models will be briefly discussed. The probability density function for the exponential distribution is given as follows:

$$f(t) = \alpha e^{-\alpha t} \qquad t \geq 0,\ \alpha > 0. \tag{1}$$

The maximum likelihood estimate of the parameter α, where $1/\alpha$ is the mean time to recovery, is $\hat{\alpha} = n/\Sigma t_i$ where $t_1,\ldots,t_n$ represent the sample data.

The estimate of the parameter α can used to obtain the estimated distribution function $F(t) = P(T \leq t) = \int_0^t f(x)dx$, which gives the probability that LOSP will be recovered within t hours. The graph of each cumulative distribution function (c.d.f.) can be added to the graph of the respective e.d.f. in Figures 2 to 4 for a visual comparison of how well the exponential model fits the data. However, since the $P(T \geq t)$ is of interest in this application, that is the probability that the time to recover LOSP will be longer than some time t, Figures 5 to 7 have been constructed with the graph of $1 - F(t)$, along with the correspondingly constructed complementary counterpart of the e.d.f. Note that the horizontal scale used in Figure 5 is different than the one used for both Figures 6 and 7 because of the shorter times of plant-centered losses.

The graphs in Figures 5 and 7 both show the fitted exponential model to be consistently above the counterpart of the e.d.f. and then to be consistently below. Such behavior indicates a definite bias in the fit. In Figure 6 the bias is less apparent, but there are only 13 observations. And as will be shown, many different distributions can be fit to small data sets.

The probability density function for the lognormal distribution is given as follows:

$$f(t) = 1/(t\sigma\sqrt{2\pi})\ \exp(-[(\ln t - \mu)/\sigma]^2/2) \quad t\geq 0,\ -\infty<\mu<\infty,\ \sigma>0 \tag{2}$$

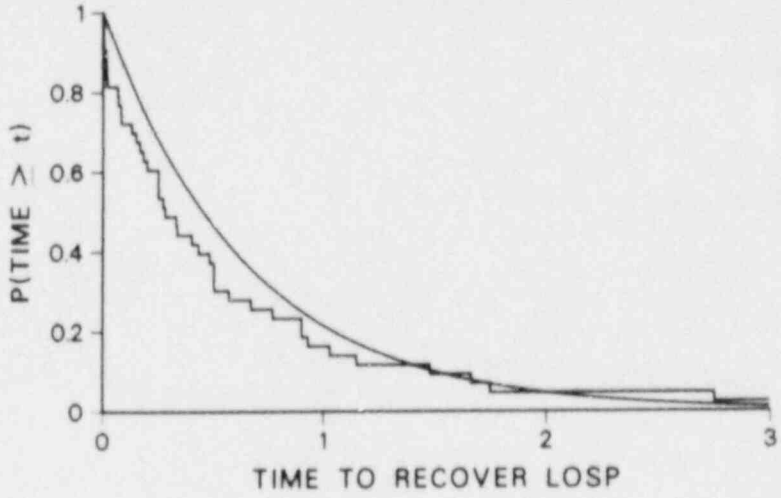

Figure 5. Exponential Probability Model for PC LOSP Data

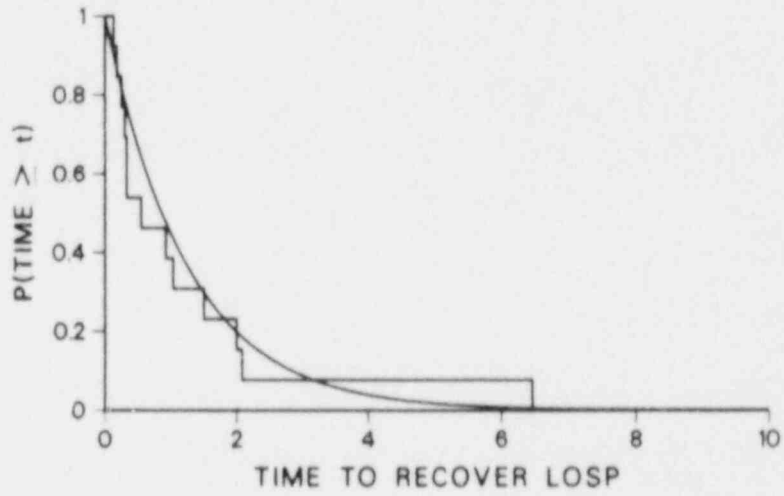

Figure 6. Exponential Probability Model for Grid LOSP Data

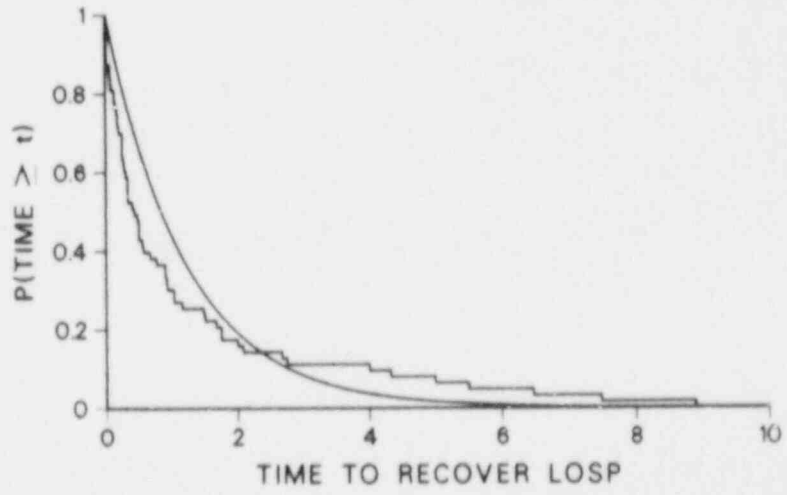

Figure 7. Exponential Probability Model for All LOSP Data

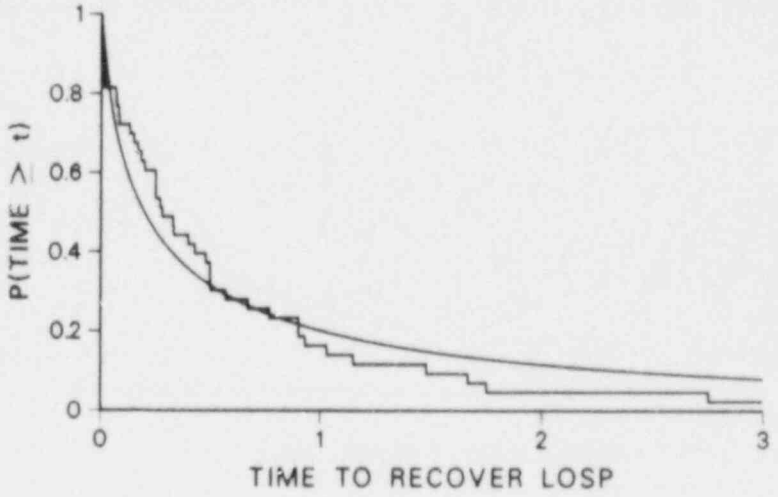

Figure 8. Lognormal Probability Model for PC LOSP Data

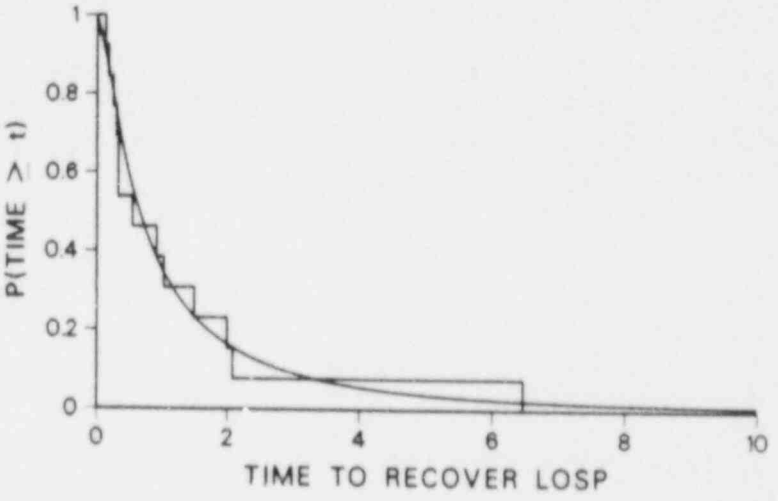

Figure 9. Lognormal Probability Model for Grid LOSP Data

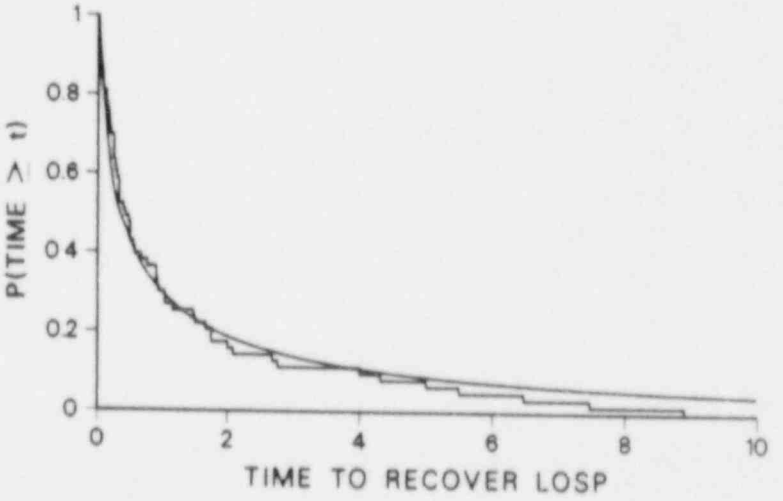

Figure 10. Lognormal Probability Model for All LOSP Data

The maximum likelihood estimates of the location and scale parameters, μ and σ, are $\hat{\mu} = \Sigma \ln t_i/n$ and $\hat{\sigma}^2 = \Sigma (\ln t_i - \hat{\mu})^2/n$.

The fitted lognormal probability models are shown in Figures 8 to 10 in a manner analogous to the exponential models in Figures 5 to 7. Bias in the lognormal model is also apparent in Figures 8 and 10, but the bias is just opposite of that shown in Figures 5 and 7. In Figure 9 it could be argued that the lognormal model provides an adequate fit for the data just as the exponential model did in Figure 6. Again this result is due to the small sample size.

3. FITTING THE GAMMA PROBABILITY MODEL

Another candidate for modeling time to recover LOSP is the gamma distribution, which is skewed to the right, and is indexed by a shape parameter (α) and a scale parameter (β). This distribution has positive probability on the interval $[0,\infty)$ and has the flexibility to represent many different shapes of probability distributions for fitting data. The probability density function for the two-parameter gamma distribution is given as follows:

$$f(t) = \beta^{\alpha}[\Gamma(\alpha)]^{-1}\, t^{\alpha-1}\, e^{-\beta t} \qquad t \geq 0,\ \alpha > 0,\ \beta > 0 \qquad (3)$$

where $\Gamma(\alpha)$ is the gamma function and t is the recovery time. The discussion in the following paragraph concerns the procedure for estimating the parameters α and β.

Lawless (1982, pages 204-206) provides a discussion of the maximum likelihood procedure used for estimating the parameters α and β in the density function given in Equation (3). His procedure can be summarized using the following steps.

1. Find the arithmetic mean ($\bar{t}$) and the geometric mean ($\tilde{t}$) for the sample data.

2. Calculate $s = \log(\bar{t} / \tilde{t})$.

3. Compute $\hat{\alpha} \approx s^{-1}(17.79728 + 11.968477\, s + s^2)^{-1}$

   $\times\ (8.898919 + 9.059950\, s + .9775373\, s^2)$ if $.5772 < s \leq 17$

   or $\hat{\alpha} \approx s^{-1}(.5000876 + .1648852\, s - .0544274\, s^2)$

   if $0 < s \leq .5772$.

4. Compute $\hat{\beta} = \hat{\alpha} / \bar{t}$.

The results of the calculations used to find the estimates of the parameters α and β for each of the plant-centered, grid, and entire data sets are given in Table 3.

Table 3. Maximum Likelihood Estimates for the Parameters α and β in a Gamma Distribution

	Plant Centered	Grid	Entire Data Set
$\bar{t}$:	0.6514	1.2364	1.2099
$\tilde{t}$:	0.1978	0.6543	0.3547
s:	1.1917	0.6364	1.2270
$\hat{\alpha}$:	0.5197	0.9161	0.5057
$\hat{\beta}$:	0.7977	0.7409	0.4180

Graphs of the three fitted probability models appear in Figures 11 to 13. Figure 11 makes it clear that the gamma probability model provides a very good model for the time to recovery of LOSP for the plant-centered data and does not show the bias of the exponential and lognormal probability models. Kolomogrov goodness-of-fit tests (Conover, 1980) support this conclusion. For the grid data in Figure 12 the result of the small sample size is again apparent, as the gamma probability model provides about the same fit as the previous models. Figure 13 shows that the gamma probability model provides a better fit than the exponential model for the overall data. The gamma probability model for the overall data fits the small and large times better than the lognormal probability model, but is not as good in the middle of the distribution.

An important point with respect to probability models used to fit the overall data is that the overall data represents a mixture of distributions; and as such, probability models that are appropriate for the plant-centered or grid portions of the data will usually not be appropriate for the overall data. Thus if the plant-centered data and the grid data are both appropriately modeled with gamma distributions, then the overall data would represent a mixture of gamma distributions, and as such are not properly modeled with a gamma distribution. In addition, the overall data also contains the weather data for which a probability model has not yet been specified.

4. FITTING THE WEIBULL PROBABILITY MODEL

An alternative formulation for modeling time to recovery of LOSP employs the Weibull distribution, which is skewed to the right and is indexed by a shape parameter (β) and a scale parameter (λ). This distribution has positive probability on the interval $[0,\infty)$ and has found widespread use in life testing studies because it offers more flexibility than the traditional exponential model. Like the gamma model, the Weibull model has the flexibility to represent many different shapes of probability distributions for fitting data. The Weibull probability model is

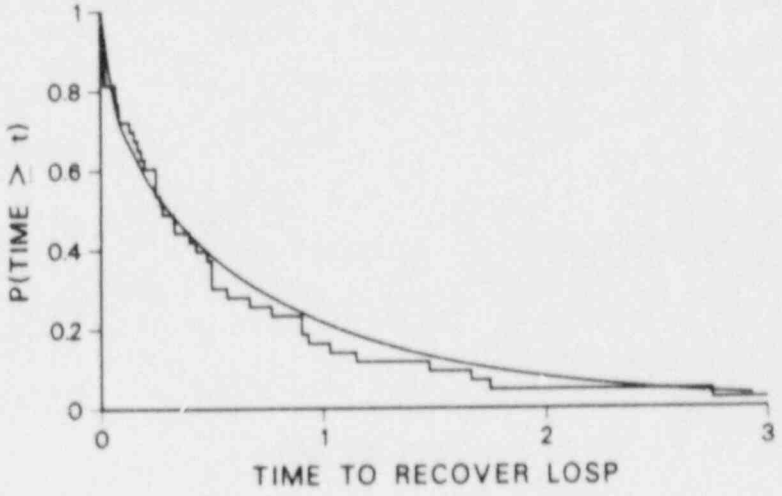

Figure 11. Gamma Probability Model for PC LOSP Data

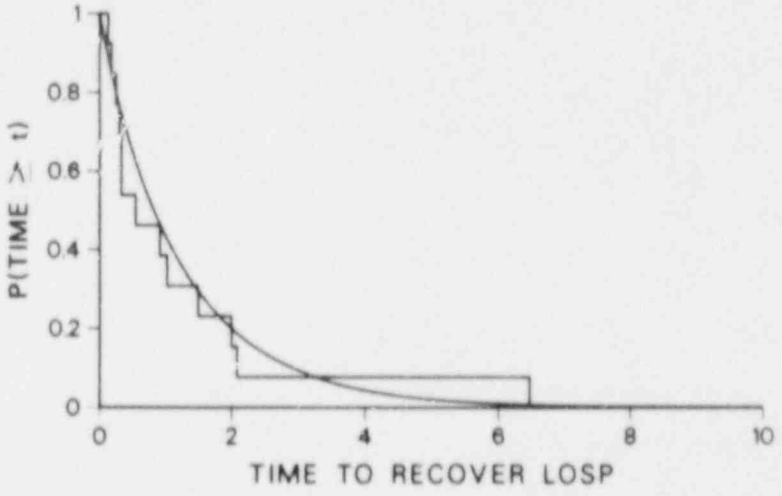

Figure 12. Gamma Probability Model for Grid LOSP Data

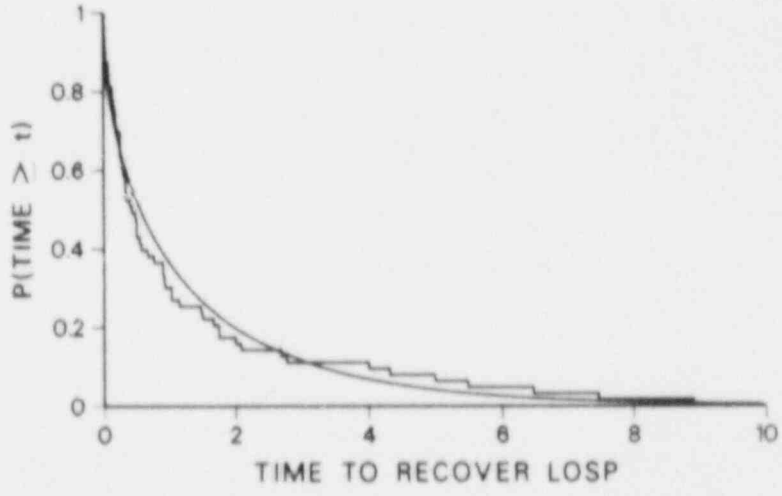

Figure 13. Gamma Probability Model for All LOSP Data

related to the exponential model in the following manner. If T is a Weibull random variable, then T^{β} is an exponential random variable.

The probability density function for the two-parameter Weibull distribution is given as follows:

$$f(t) = \lambda^{\beta}\beta t^{\beta-1} \exp(-\lambda t)^{\beta} \qquad t \geq 0 \tag{4}$$

where $\lambda>0$, $\beta>0$, and t is the recovery time. The discussion in the following paragraph concerns the procedure for estimating the parameters λ and β.

Lawless (1982, pages 142-143) provides a discussion of the maximum likelihood procedure used for estimating the parameters λ and β in the density function given in Equation (4). His procedure requires numerical methods to solve the following equations for $\hat{\lambda}$ and $\hat{\beta}$.

$$\hat{\lambda} = \left[\frac{1}{n} \sum_{i=1}^{n} t_i^{\hat{\beta}} \right]^{-1/\hat{\beta}} \tag{5}$$

$$\sum_{i=1}^{n} t_i^{\hat{\beta}} \log t_i \Big/ \sum_{i=1}^{n} t_i^{\hat{\beta}} - 1/\hat{\beta} - \frac{1}{n} \sum_{i=1}^{n} \log t_i = 0 \tag{6}$$

The results of the calculations used to find the estimates of the parameters λ and β for each of the plant-centered, grid, and entire data sets are given in Table 4.

Table 4. Maximum Likelihood Estimates for the Parameters λ and β in a Weibull Distribution

	Plant Centered	Grid	Entire Data Set
$\hat{\lambda}$:	2.1125	0.8662	1.1576
$\hat{\beta}$:	0.6544	0.8891	0.6396

The estimates of the parameters λ and β are used to obtain the Weibull probability model for LOSP as was done for the gamma probability model. Figures 14 to 16 show the analogous results for the Weibull distribution as Figures 11 to 13 showed for gamma distribution. The interpretation of the graphs is similar to that for the gamma model with the exception that the Weibull appears to provide a better fit to the overall data.

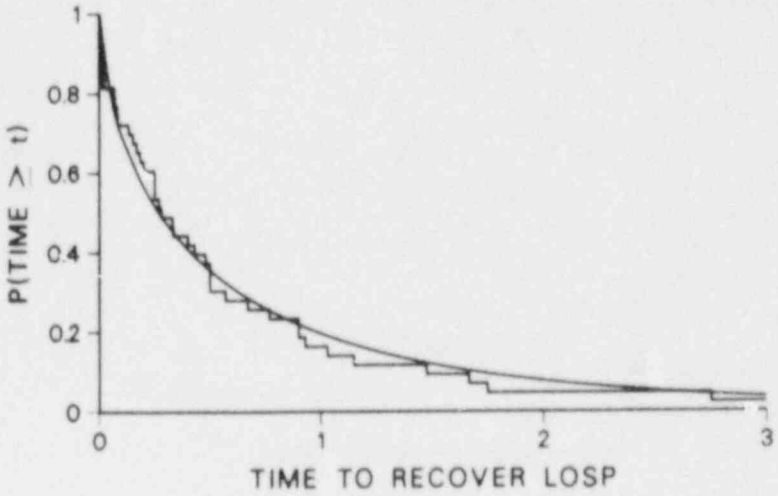

Figure 14. Weibull Probability Model for PC LOSP Data

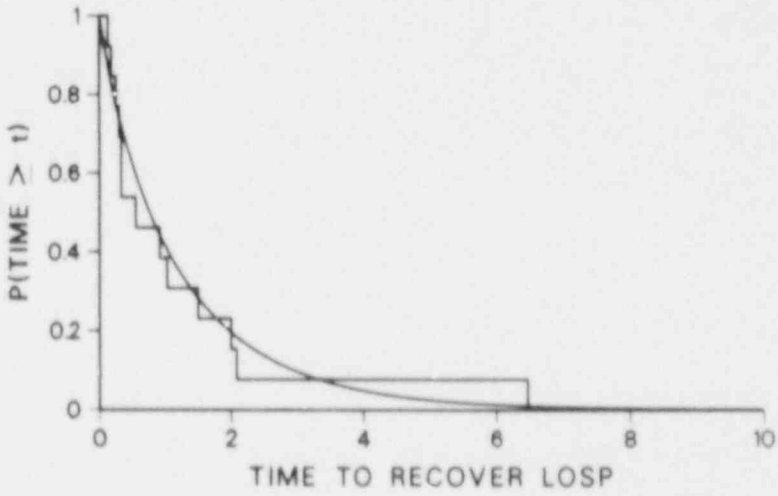

Figure 15. Weibull Probability Model for Grid LOSP Data

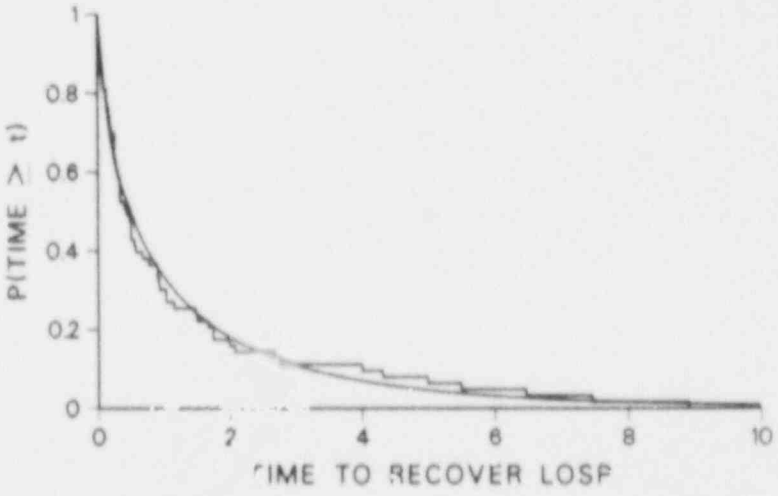

Figure 16. Weibull Probability Model for All LOSP Data

5. A BAYESIAN COMPARISON OF ALTERNATIVE PROBABILITY MODELS

The Bayesian method for comparing alternative probability models produces posterior probabilities that each of the models is the correct model. It is assumed, a priori, that each of the models is equally likely to be the correct model and that one of the models is the correct model. Therefore, when comparing the gamma and Weibull models, each model is assigned a prior probability of .5. The revision of these prior probabilities is conducted using Bayes' Theorem. For example, the probability of the gamma family being correct is obtained from Bayes' Theorem as

$$P(\text{Gamma}|t_1,\ldots,t_n) = \frac{.5P(t_1,\ldots,t_n|\text{Gamma})}{.5P(t_1,\ldots,t_n|\text{Gamma}) + .5P(t_1,\ldots,t_n|\text{Weibull})}. \quad (7)$$

The term $P(t_1,\ldots,t_n|\text{Gamma})$ is calculated from the prior density of the parameters α and β, $p(\alpha,\beta)$, and the likelihood function of the data

$$L(t_1,\ldots,t_n|\alpha,\beta) = \Pi\, f_\gamma(t_i|\alpha,\beta), \quad (8)$$

where $f_\gamma(t_i|\alpha,\beta)$ is the gamma density function of Equation (3). From $p(\alpha,\beta) = 1/(\alpha\beta)$ and $L(t_1,\ldots,t_n|\alpha,\beta)$, $P(t_1,\ldots,t_n|\text{Gamma})$ is calculated, as

$$P(t_1,\ldots,t_n|\text{Gamma}) = \int_0^\infty \int_0^\infty p(\alpha,\beta)L(t_1,\ldots,t_n|\alpha,\beta)d\alpha\, d\beta. \quad (9)$$

The calculation of $P(\text{Weibull}|t_1,\ldots,t_n)$ is accomplished in the same manner with the roles of the Weibull and gamma densities interchanged in the above discussion and $p(\beta,\lambda) = 1/(\beta\lambda)$.

The above analysis extends to encompass other probability models and to the comparison of more than two probability models at one time. In the results given below the posterior probabilities have been calculated for the exponential, lognormal, gamma, and Weibull probability models, using the plant-centered event data, the grid centered event data, and all of the data combined into a single data set. Each probability model in Equation (7) is given a prior probability of .25, although different weights could be used for each model. The results of the comparison of the four probability models are given in Table 5. For the exponential model and lognormal models the priors, $p(\alpha) = 1/\alpha$, and $p(\sigma) = 1/\sigma$ were used, respectively.

Table 5. Posterior Probabilities for the Probability Models Used with Each Category of Loss of Off-Site Power

Probability Model	Posterior Probability		
	Plant Centered	Grid	All Data
Exponential	4E-4	.181	4E-6
Lognormal	.041	.533	.019
Gamma	.466	.175	.430
Weibull	.492	.111	.551

The posterior probabilities of the gamma and Weibull families for the plant-centered data almost sum to 1.0 by themselves, as is also the case for the entire data set, thus either of these models would be favored over either the lognormal or exponential models. Specifically, in the case of plant-centered data, the Weibull model would be essentially the same as the gamma, but favored over both the lognormal model by 12 to 1 (.492 to .041), and heavily favored over the exponential model by 1230 to 1 (.492 to 4E-4). For the grid data that involved only 13 data values, none of the models differed greatly. This corresponds to comments made earlier with respect to the difficulty of choosing a model using a small number of observations. Thus the choice among the four models is not clear, and any one of the family of distributions will serve almost as well as the other. The Weibull and gamma are favored over the other two models for the entire data set with the Weibull being favored over the exponential model by 143,169 to 1.

6. BAYESIAN ANALYSIS OF THE UNCERTAINTY ASSOCIATED WITH $P(T \leq t)$

The calculation of the cumulative probability,

$$F(t) = P(T \leq t) = F(t|\alpha,\beta)$$

depends on knowing α and β. However, if α and β are uncertain, then $P(T \leq t)$ is uncertain. If α and β possess a joint probability distribution, as in Bayesian statistics, then $P(T \leq t)$ is a random variable and as such has a probability distribution that is induced by the joint distribution on α and β. This distribution is of interest in probabilistic risk analysis since it is desirable not only to have an estimate of $P(T \leq t)$, but also to have an expression of the uncertainty in $P(T \leq t)$ for each given value t. Thus a Bayesian analysis is used as the vehicle to quantify the uncertainty.

The Bayesian analysis for LOSP is predicated on the assumption that recovery times follow a two parameter gamma distribution with neither the shape parameter, α, nor the scale parameter, β, known. (The Weibull distribution is considered later in

this section.) The distributions have been verified through the examination of data sets in Sections 3 to 5.

Let T denote the time to recovery from a LOSP incident. The probability that recovery will occur during the first t hours of the incident is given by:

$$F(t|\alpha,\beta) = \int_0^t f(x)\,dx = \int_0^t \beta^{\alpha} x^{\alpha-1} e^{-\beta x}/\Gamma(\alpha)\,dx \qquad (10)$$

where $t \geq 0$, $\alpha > 0$, and $\beta > 0$.

Once point estimates of α and β have been developed following the steps outlined in Section 3, inserting the point estimates of α and β into Equation (10) produces an estimate of the function $F(t|\alpha,\beta)$. This method produces only a point estimate of the distribution of recovery time and does not produce information about the uncertainty in the estimate of $F(t|\alpha,\beta)$. The problem of producing confidence limits or Bayesian uncertainty intervals for the function $F(t|\alpha,\beta)$ is difficult to solve since the probability function contains the complete gamma function $\Gamma(\alpha)$, which introduces severe analytic and numerical difficulties.

The Bayesian Posterior Density

The Bayesian analysis yields explicit, probabilistic uncertainty bounds and is based on the joint posterior probability distribution of α and β calculated using data from LOSP incidents and Bayes' Theorem. The data are denoted by $t_1,\ldots,t_n$. Each datum is the time to recovery for an independent LOSP incident. Thus, measurements for n distinct incidents comprise the data set. The joint probability density function, or <u>likelihood function</u>, for the n independent observations is

$$L(t_1,\ldots,t_n) = \prod_{i=1}^{n} f(t_i|\alpha,\beta)$$

$$= \beta^{n\alpha} p^{\alpha-1} e^{-\beta S}/\Gamma(\alpha)^n \qquad (11)$$

where $f(t|\alpha,\beta)$ is the derivative of $F(t|\alpha,\beta)$ with respect to t, $P = \prod t_i$ and $S = \Sigma t_i$.

Bayes' Theorem combines the information in the likelihood function with prior, or subjective, information concerning α and β. In this analysis, the use of prior information is forgone, and instead, a noninformative prior distribution will be used to avoid injecting subjective information into the analysis. Following Jeffreys (1961), the <u>joint prior density</u> for α and β is taken to be proportional to $1/(\alpha\beta)$. This expression does not, in fact, define a proper probability distribution. Its usefulness arises, instead, from the lack of information about the relative likelihoods of various values of α and β.

Combining the likelihood function and the prior distribution using Bayes' Theorem amounts to multiplying the likelihood function and the prior distribution and then finding a normalizing constant that makes the area under the resulting joint function of α and β equal to unity. The resulting joint posterior density of α and β is given by:

$$h(\alpha,\beta|t_1,\ldots,t_n) = [\beta^{n\alpha}p^{\alpha-1}e^{-\beta S}]/[k\alpha\beta\Gamma(\alpha)^n] \qquad (12)$$

where k is the constant

$$k = \int_0^\infty \int_0^\infty [\beta^{n\alpha}p^{\alpha-1}e^{-\beta S}]/[\alpha\beta\Gamma(\alpha)^n]\, d\alpha\, d\beta \qquad (13)$$

Uncertainty Bounds

Engineering analyses often employ the complementary probability function $G(t|\alpha,\beta) = 1 - F(t|\alpha,\beta) = P(T > t)$, which is the probability that recovery will not be achieved before t hours. Each pair of parameter values, (α, β), defines a different recovery time complementary cumulative probability function. At a given value of t, the pairs (α, β) can be conceptually ordered in such a way that each successive pair produces a larger value of $G(t|\alpha,\beta)$. Let the set of pairs (α, β) that produce values of $G(t|\alpha,\beta) \leq q$ be denoted by $R(t,q)$ for $0 \leq q \leq 1$. Formally,

$$R(t,q) = \{(\alpha,\beta)|G(t|\alpha,\beta) \leq q\}. \qquad (14)$$

$R(t,q)$ then defines a region of values in the α, β parameter space for which the probability of the recovery time exceeding t is not greater than q.

The uncertainty bounds for the recovery time distribution are found from $R(t,q)$ and $h(\alpha,\beta|t_1,\ldots,t_n)$. For example, the 90% upper uncertainty bound is the value for which there is only a 10% chance that the true recovery probability exceeds this value. The value of the 90% uncertainty bound is the value of q that satisfies:

$$\int_{R(t,q)} \int h(\alpha,\beta|t_1,\ldots,t_n)\, d\alpha\, d\beta = .90. \qquad (15)$$

Similarly, bounds other than the 90% bourd are generated using various values for the right side of Equation (15). Note that since t is fixed, the uncertainty bounds for other values of t must be found through the set $R(t,q)$ that is specific to that t.

Monte Carlo Implementation

In practice the procedure described above is virtually impossible to implement. The difficulty is in the need to define $R(t,q)$ for all values of t and q. This must be done so that the various values of q can be searched to find the one value that satisfies Equation (15). An alternative methodology that is capable of providing the same uncertainty bounds is a Monte Carlo implementation of the Bayesian analysis.

The Monte Carlo implementation of the Bayesian analysis requires that values of α and β be sampled from the joint posterior density of α and β given in Equation (12). For each sampled pair (α,β), $G(t|\alpha,\beta)$ is found for the values of t that cover the range of interest. The process is repeated many times until a sufficient number of observations are obtained to estimate the probability distribution function, $G(t|\alpha,\beta)$. This distribution forms the required uncertainty bound for a given value of t. While there is some error due to the finite size of the Monte Carlo sample, the error can be controlled to any desired level by setting an appropriate Monte Carlo sample size.

Factoring the Joint Posterior Density

While the Monte Carlo implementation of the Bayesian analysis provides a means for estimating the probability distribution function of $G(t|a,b)$, there are some technical issues concerning sampling from the joint distribution of α and β that must be resolved. Recall that the joint posterior distribution is given by Equation (12). One may sample from $h(\alpha,\beta|t_1,\ldots,t_n)$ by factoring the joint density $h(\alpha,\beta|t_1,\ldots,t_n)$ into the product of the conditional density function of β given α and the unconditional or marginal density of α. Thus,

$$h(\alpha,\beta|t_1,\ldots,t_n) = h_{\beta|\alpha}(\beta)\, h_{\alpha}(\alpha) \tag{16}$$

where, by some algebra,

$$h_{\beta|\alpha}(\beta) = \beta^{n\alpha-1} S^{n\alpha} e^{-\beta S}/\Gamma(n\alpha) \tag{17}$$

and

$$h_{\alpha}(\alpha) = p^{\alpha-1}\Gamma(n\alpha)/\{k_{\alpha} S^{n\alpha}[\Gamma(\alpha)]^{n}\}. \tag{18}$$

The function in Equation (17) is recognized as a gamma density function, while the function in Equation (18) is a density that is not a member of a well-studied family of densities. In fact, $h_{\alpha}(\alpha)$ cannot be integrated analytically to find the corresponding cumulative distribution function. Thus, the acceptance-rejection method of sampling is used to generate values from this density.

Acceptance-Rejection Sampling

The acceptance-rejection method of random deviate generation is applicable when the values are generated from a density function that is bounded from above. This condition holds for the density in Equation (18). Although it is not always an efficient algorithm to use, it can be employed when only the density function is known. Thus, acceptance-rejection methods can be used in a wider variety of situations than many other methods. Briefly, acceptance-rejection generation of a random deviate is accomplished using the steps:

1. Transform the random variable X (i.e. α in (18)) to another random variable Y so that the range of values of the random variable Y is [0,1]. Since the random variable $\alpha > 0$, the transformation that is used is $Y = \alpha/(1 + \alpha)$. Denote the density of the transformed random variable by $h^*(y)$. It is assumed that $h^*(y)$ is bounded from above.

2. Calculate the maximum value of the density function of the transformed random variable Y. Call this value w.

3. Generate two uniform random deviates, u_1 and u_2.

4. If $u_2 \leq h^*(u_1)/w$, then let the generated random deviate, y, be equal to u_1, and the reverse transformation, $\alpha = Y/(1-Y)$, gives the value of the random deviate α. Otherwise, return to Step 3.

The potential inefficiency of the acceptance-rejection methodology stems from the test in Step 4. If the inequality does not hold true, no random deviate is generated. Thus, particularly if the density of the transformed random variable is very peaked, many uniform random deviates may be needed in order to generate one value of Y. However, the cost of generating uniform random deviates is very low compared to the cost of function evaluation, and thus acceptance-rejection procedures may be more efficient if the alternative method requires difficult function evaluations.

Results for the Gamma Probability Model

Figures 17 to 19 show the 90% Bayesian uncertainty bounds for the gamma probability model resulting from the acceptance-rejection sampling procedure for plant-centered, grid, and overall data, respectively. The bounds in Figure 17 are quite narrow reflecting the greater number of observations for plant-centered data (n=43), and the empirical derived step function is entirely contained within the bounds. The bounds in Figure 18 for the grid data are much wider than those for the plant-centered data reflecting the smaller sample size (n=13). Again these bounds provide good coverage of the empirical step function in the graph. With n=63, the bounds in Figure 19 are narrow, but the deficiency of the gamma probability model for the overall data is evidenced by empirical step function being on, or outside of, the bounds.

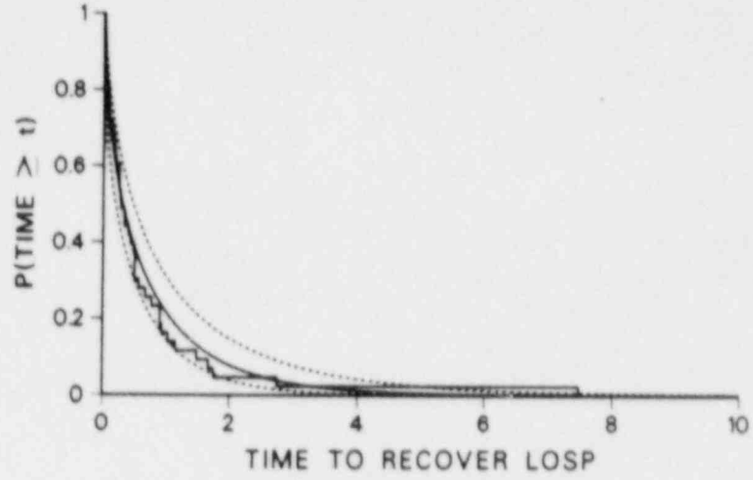

Figure 17. Ninety Percent Uncertainty Bounds for the Gamma Model for PC

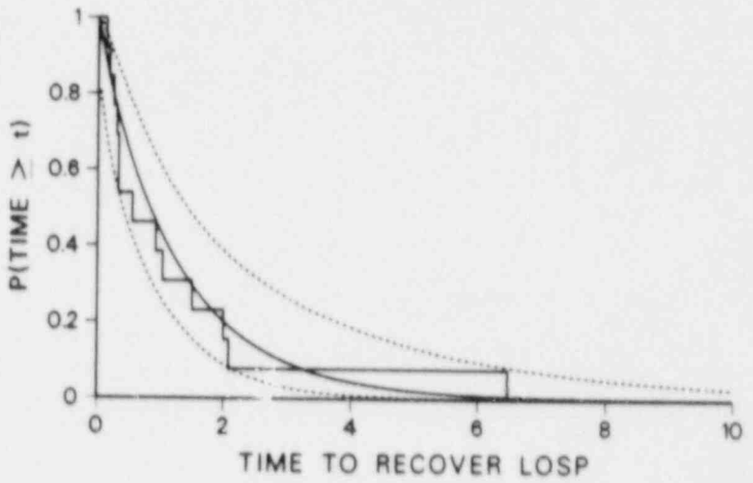

Figure 18. Ninety Percent Uncertainty Bounds for the Gamma Model for Grid

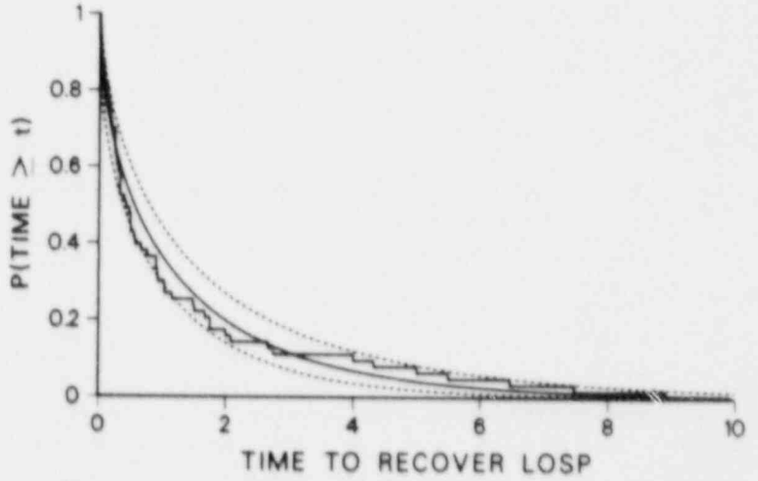

Figure 19. Ninety Percent Uncertainty Bounds for the Gamma Model for All

Interpreting the Uncertainty Bounds

The Bayesian method of generating the uncertainty bounds shown in Figures 17 to 19 lends itself to an empirical interpretation. The process of generating pairs of values of α and β may be viewed as one that generates complementary cumulative probability functions or curves for the time to recovery of LOSP. Each pair of values, (α, β), is mapped into a distinct curve. Thus a Monte Carlo simulation of N pairs of values of α and β is a Monte Carlo simulation of N recovery curves.

No single curve generated in the simulation forms an uncertainty bound, rather the uncertainty bound is found from an envelope of curves. For example, the 90% uncertainty bound is found by identifying the .05×Nth smallest curve and the .95×Nth largest curve at each value of the time variable. Thus it is possible that one curve could provide part of the upper uncertainty bound at one value of the time variable and part of the lower uncertainty bound at a different value of the time variable.

This method of constructing uncertainty bounds is based on the axioms of probability and thus has a strong foundation. The gamma probability model was used here because the data are fit better by gamma than by either the exponential or lognormal models. Bounds based on the Weibull model are now considered.

Bounds for the Weibull Probability Model

Bayesian analysis of the uncertainty associated with the Weibull probability model proceeds in the same way as the analysis just presented for the gamma probability model. Again, a diffuse joint prior distribution is employed for the two uncertain parameters. In this case the joint prior is $1/(\lambda\beta)$. The likelihood function analogous to that given in Equation (11) for a sample of n observations given by $t_1, \ldots, t_n$ is

$$L(t_1, \ldots, t_n) = (\lambda^{\beta}\beta)^n P^{\beta-1} e^{-\lambda^{\beta} S}, \tag{19}$$

where $S = \Sigma\, t_i^{\beta}$ and $P = \Pi\, t_i$. Multiplying the joint prior distribution by the likelihood function and a normalizing constant, k, yields the posterior distribution for λ and β given by

$$h(\lambda, \beta | t_1, \ldots, t_n) = k\lambda^{n\beta-1}\beta^{n-1}P^{\beta-1}e^{-\lambda^{\beta} S}. \tag{20}$$

Equation (20) can be factored into the unconditional or marginal density function of β, $h_{\beta}(\beta)$, and the conditional density function of λ given β, $h_{\lambda|\beta}(\lambda)$. These densities are

$$h_{\beta}(\beta) = k\Gamma(n)\beta^{n-1}P^{\beta-1}/S^n \tag{21}$$

and

$$h_{\lambda|\beta}(\lambda) = \lambda^{n\beta-1} S^n e^{-\lambda^\beta S}/\Gamma(n). \qquad (22)$$

The uncertainty bounds for the Weibull probability model are generated using the same Monte Carlo scheme that was used to generate the uncertainty bounds for the gamma probability model. Sampling is done first from $h_\beta(\beta)$ using acceptance-rejection methods. The generated value of β is used to specify the conditional distribution of λ, and a value of λ is selected using the inverse function method of random deviate generation.

Results for the Weibull Probability Model

Figures 20 to 22 show the 90% uncertainty bounds for the Weibull probability model resulting from the acceptance-rejection sampling procedure for plant-centered, grid, and overall data, respectively. The results shown in Figures 20 and 21 are very similar to those for the gamma probability model in Figures 17 and 18, respectively. The results shown in Figure 22 may be slightly better than the corresponding results shown in Figure 19 for the overall data. This is consistent with the posterior probabilities given earlier for the gamma and Weibull models for the overall data.

7. A COMPOSITE MODEL FOR TIME TO RECOVERY OF LOSP

The analyses in the previous sections provide support for modeling the plant-centered data and the grid data with either gamma or Weibull probability models. Thus, it is appropriate to recall the last paragraph of Section 3, which indicated that probability models appropriate for portions of the data will usually not be appropriate for the entire composite data set. That is, the composite data set represents a mixture of distributions. An alternative to modeling the entire data with a single probability model, as has been done in the previous sections, is to use a composite probability model that combines the various entities of plant-centered, grid, and weather into a single model in order to predict the time to recovery, given a LOSP. The composite model is formed by pooling data from plant-centered, grid-centered, and weather sources. The question that needs to be addressed is how to construct a composite recovery curve from models for each of these three types of power loss. A solution to this problem is to form a mixture distribution that is a weighted average of each of the three types of failures. Such a model can be expressed as

$$G(x) = p_1 G_1(x) + p_2 G_2(x) + p_3 G_3(x) \qquad (23)$$

where the $G_i(x)$ are complementary cumulative probability functions representing the three types of failures, and the p_i are weights that reflect the occurrence rates of the three types of power loss. Of course, $p_1 + p_2 + p_3 = 1$.

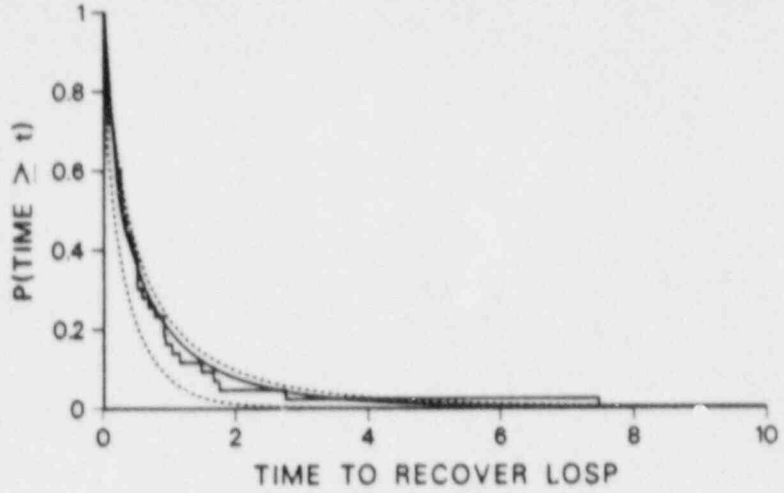

Figure 20. Ninety Percent Uncertainty Bounds for the Weibull Model for PC

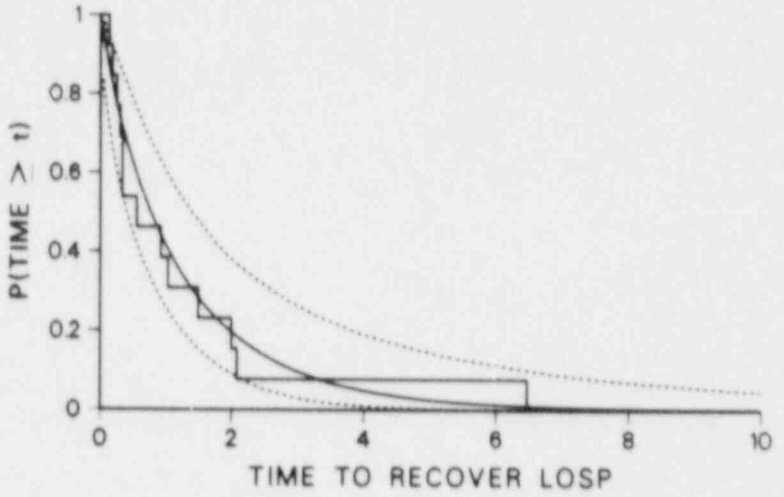

Figure 21. Ninety Percent Uncertainty Bounds for the Weibull Model for Grid

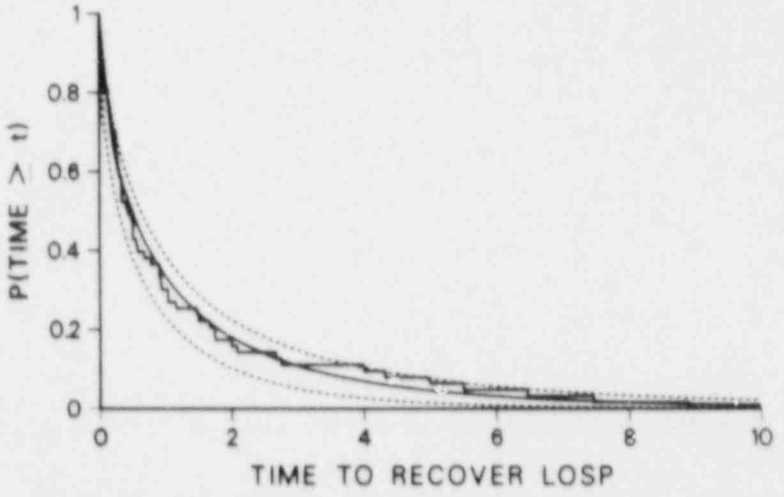

Figure 22. Ninety Percent Uncertainty Bounds for the Weibull Model for All

The model in Equation (23) contains the parameters p_1, p_2, and p_3, which may be estimated by the observed relative frequencies of each of the three types of power loss. There is also uncertainty in these estimates so that the total uncertainty in the curve in Equation (23) is a function of the uncertainties about the p_i as well as the uncertainties about the $G_i(x)$.

Dirichlet Distribution

The information about the uncertainty in the p_is may be analyzed by Bayesian methods using a noninformative prior distribution and a multinomial likelihood function. The resulting posterior distribution is known as a Dirichlet distribution. The Dirichlet distribution for two random variables, p_1 and p_2, will be used since $p_3 = 1 - p_1 - p_2$. This Dirichlet density is given by

$$f(p_1,p_2|n_1,n_2,n_3) = (\Gamma(n_1+n_2+n_3)/[\Gamma(n_1)\Gamma(n_2)\Gamma(n_3)])\, p_1^{n_1-1}\, p_2^{n_2-1}\, (1-p_1-p_2)^{n_3-1} \quad (24)$$

where n_i is the frequency of observations in the ith category. The likelihood function for frequencies of mutually exclusive events is the multinomial likelihood function given by

$$L(n_1,n_2,n_3|p_1,p_2) = (n_1+n_2+n_3)!/[n_1!n_2!n_3!]p_1^{n_1}\, p_2^{n_2}\, p_3^{n_3}. \quad (25)$$

When Equation (25) is combined with the noninformative prior distribution

$$w(p_1,p_2) = 1/[(p_1\, p_2)(1 - p_1 - p_2)] \quad (26)$$

via Bayes' Theorem, the resulting posterior density is that given in Equation (24).

Uncertainty for the Composite Model

The expression of uncertainty about the function in Equation (23) is accomplished in the same manner as was used for the individual models discussed in the previous sections. The difference here is that not only will the uncertain parameters of the distributions $G_i(x)$ be sampled, but in addition, values of p_1 and p_2 will be sampled from the posterior distribution given in Equation (24). For the demonstration of this technique, the gamma family of distributions is used to represent the distribution of recovery time for each type of power loss. The technique is to sample from the collection of eight uncertain parameters (three alphas, three betas, and two values of the proportions, p_i) and construct a curve for each sampled set of values.

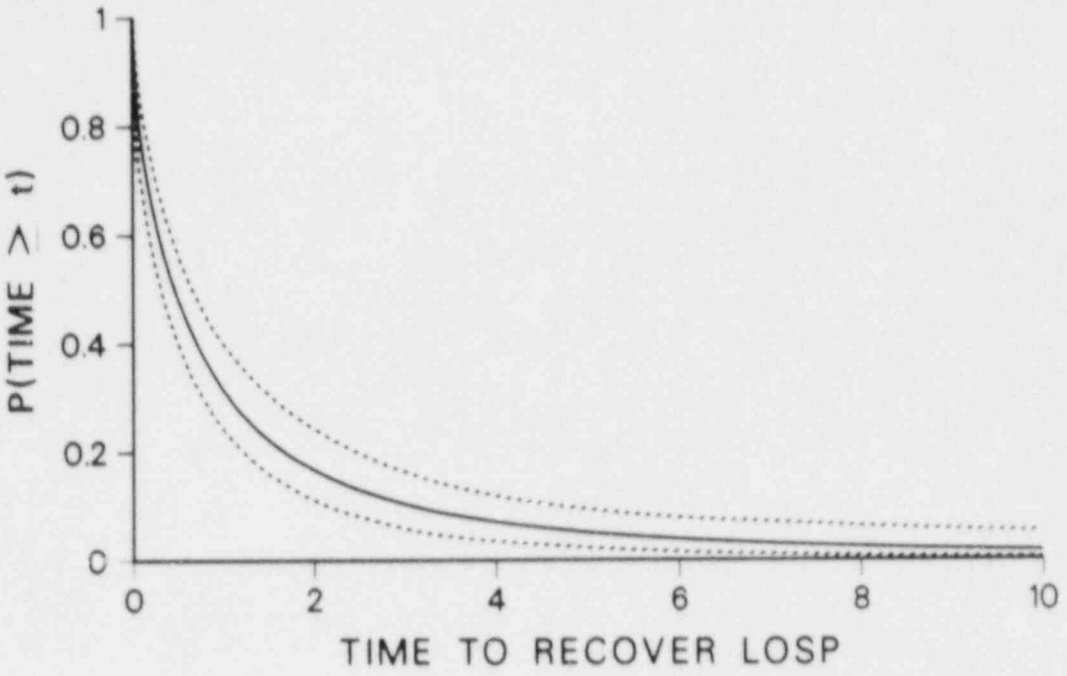

Figure 23. Uncertainty Bounds for G(x) from Equation (23) Based on Using All of the Plant-Centered Data

This process is repeated many times so that a collection of curves is accumulated. The uncertainty bounds are derived at various values of the recovery time in the identical manner that was used for the individual curves calculated in the preceding sections.

Sampling from the Dirichlet distribution is most easily facilitated by the acceptance-rejection method. This case requires the sampling of pairs of values of p_1 and p_2 rather than a single parameter value. One pair of values of p_1 and p_2 is formed by generating three uniformly distributed random numbers. Two of the random numbers are the potential values of p_1 and p_2, while the third number is compared to the height of the density at p_1, p_2 divided by the maximum value of the density. If the third number is less than the density divided by the maximum density, then the pair is accepted. Otherwise the pair is rejected and three new random numbers are created and the process is repeated.

The uncertainty results obtained by Monte Carlo estimation of G(x) are shown in Figure 23. The results shown in Figure 23 are based on using all the PC data (n_1=43), the grid data (n_2=13), and the weather data (n_3=7). The three curves shown in Figure 23 represent the .05, .50 and .95 quantiles for G(x) at each value of time.

8. MODIFYING THE COMPOSITE MODEL FOR SPECIFIC PLANTS

The composite model in Equation (23) is very general and flexible and is easily modified to reflect conditions at specific plants. There are several ways in which the model can be modified:

1. The functions in Equation (23) are completely general and do not all have to be the same. That is, one distribution such as the Weibull could be used to represent $G_1(x)$ for the plant-centered term, another distribution such as the gamma could be used to represent $G_2(x)$ for the grid term, while yet another distribution could be used for the weather term $G_3(x)$.

2. A specific $G_i(x)$ can be modified for a particular plant. For example, it will be shown in this section that the plant-centered data can be split into three separate groups based on switchyard design with each switchyard design having its own model.

3. The weights p_i can be modified to reflect conditions at a particular plant. For example, it will be shown in this section how the weather weight, p_3, might be modified for a specific plant.

Modifying the Distribution for Plant-Centered LOSP

The composite model in Equation (23) is generic in the sense that all data from all plants have been combined to form a single model. That model may or may not be appropriate for a specific plant. For example, the data previously given for time to recovery of LOSP for plant-centered data was subdivided in NUREG-1032 into three plant switchyard design groups (see NUREG-1032 for a detailed explanation of the design groups). The data are given in Table 6 and the corresponding boxplots shown in Figure 24.

Based on the three switchyard groupings in Table 6, three separate Weibull probability models as in Equation (4) have been fit to the data. The parameter estimates for λ and β determined from Equations (5) and (6) are given in Table 7. The resulting fitted probability models are shown in Figure 25 along with the Weibull fit for the entire set of plant-centered data previously shown in Figure 14. Figure 25 shows that Group I2 has the closest agreement with the model based on the overall data set. The three groupings given in NUREG-1032 were made with respect to what is physically valid for the plant under consideration. The grouping of these data will now be examined by considering a statistical analysis of the fitted models shown in Figure 25.

Table 6. Plant-Centered Data from Table 1 Categorized by Switchyard Configuration

Plant-Centered Group I1, n=14

0.002	0.003	0.013	0.017	0.080
0.150	0.167	0.183	0.250	0.270
0.330	0.430	0.480	0.500	

Plant-Centered Group I2, n=13

0.003	0.020	0.070	0.083	0.130
0.250	0.250	0.280	0.334	0.500
0.670	1.030	1.480		

Plant-Centered Group I3, n=16

0.004	0.015	0.067	0.200	0.400
0.500	0.570	0.767*	0.900	0.900
0.930	1.150	1.667*	1.750	2.750
7.467*				

Table 7. Maximum Likelihood Estimates for the Parameters λ and β for the Weibull Distributions Fit to the Switchyard Type in Table 6

	Group I1	Group I2	Group I3
$\hat{\lambda}$:	5.2721	2.7897	0.9663
$\hat{\beta}$:	0.8248	0.8331	0.7294

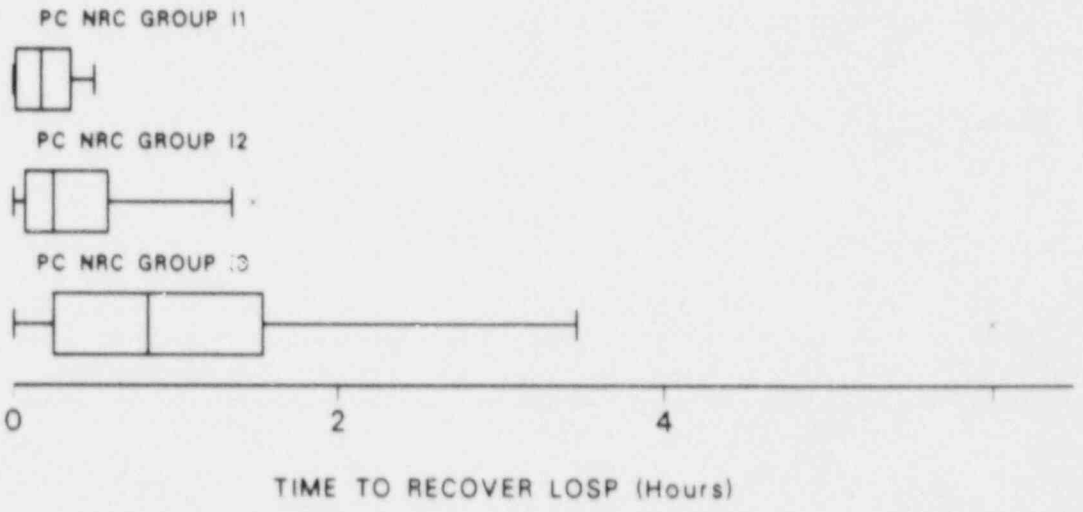

Figure 24. Boxplots of Three Groupings of the Plant Centered Data

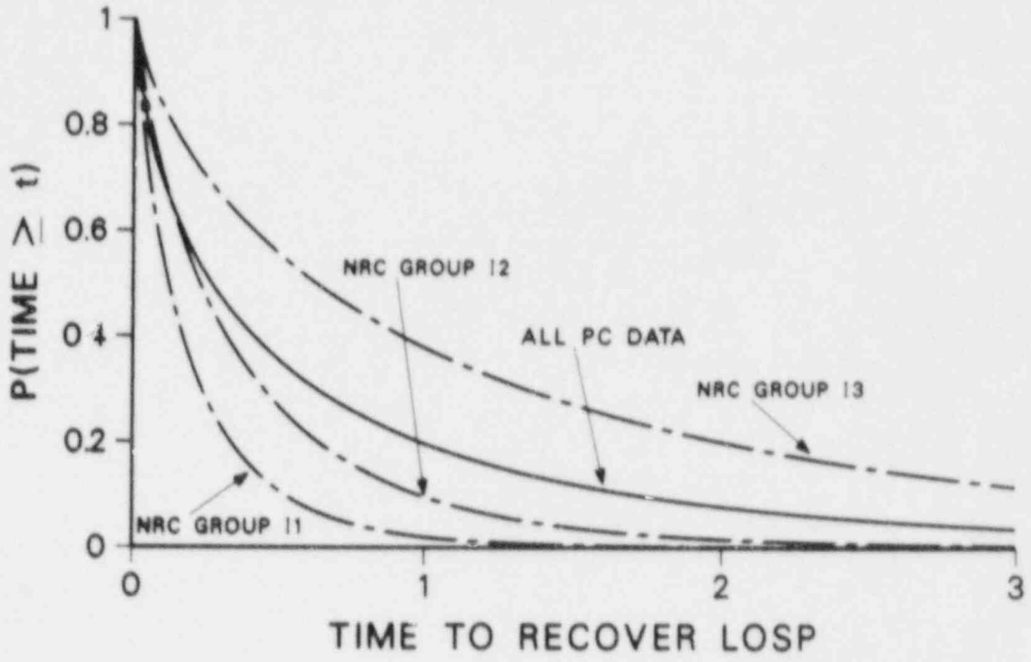

Figure 25. Fitted Weibull Probability Models for the Three Plant Centered Groupings Along With the Fit for All Plant Centered Data Previously Shown in Figure 14

Tests for Pooling Data

The question that arises with respect to the plant-centered data is whether the three plant design groups of data should be modeled as one data set or three distinct data sets. A test of the homogeneity of the distributions of the data provides useful information to the analyst in making this decision.

The method that has been employed to ascertain whether pooling of the three groups of data is appropriate is a generalized likelihood ratio test (GLRT). The GLRT is a method of constructing tests of competing hypotheses. The hypotheses must be structured so that one hypothesis contains a subset of the possibilities contained in the other hypothesis. For example, assume that the data in the ith group are distributed according to the gamma distribution with parameters α_i and β_i for i=1, 2, 3. The test of the homogeneity of the distributions of the data is a test of the null hypothesis:

$$H_0: (\alpha_1 = \alpha_2 = \alpha_3) \text{ and } (\beta_1 = \beta_2 = \beta_3) \quad (27)$$

against the alternative hypothesis that at least two of the α_i and/or at least two of the β_i are different.

The GLRT statistic is constructed by finding the maximum of the likelihood function under the null hypothesis and the maximum

of the likelihood function without the constraints imposed by the null hypothesis. A function of the ratio of these two maxima forms the test statistic. Under some mild regularity conditions about the parent distribution of the data, the probability distribution of the test statistic is the chi-squared probability distribution with ν degrees of freedom where ν is the number of parameters constrained by the null hypothesis. For the null hypothesis given in Equation (27), $\nu = 4$, since two of the values of α_i and two of the values of β_i are constrained by the null hypothesis. Let T denote the test statistic and $f(x|\alpha,\beta)$ represent the density function of the data. Then, for the three group problem, T is given by:

$$T = -2 \frac{\max_{\alpha,\beta} \sum_{i=1}^{3} \sum_{j=1}^{n_i} \log f(t_{ij}|\alpha,\beta)}{\sum_{i=1}^{3} \max_{\alpha_i,\beta_j} \sum_{j=1}^{n_i} \log f(t_{ij}|\alpha_i,\beta_j)} \qquad (28)$$

where t_{ij} is the jth observation among the n_i observations from the ith group of observations.

The GLRT test described above has the following practical implications from the potential conclusions of the test. If the null hypothesis is rejected, the analyst should conclude that the distribution of the recovery time depends upon the type of electrical switchyard. If the null hypothesis is not rejected, the analyst should not conclude that the distribution of recovery times is dependent upon the electrical switchyard configuration.

The GLRT requires that a family of probability distributions, such as the gamma or Weibull distributions, be specified to calculate the statistic T given in Equation (28). In order to ameliorate this assumption, parallel analyses have been made assuming gamma probability distributions and Weibull distributions. For each family of distributions, tests of the following hypotheses were performed:

1. All three groups are homogeneous.

2. The first and second groups are homogeneous.

3. The first and third groups are homogeneous.

4. The second and third groups are homogeneous.

The results of each test using both the gamma and Weibull family of distributions are given in Table 8.

Table 8. Results of Hypotheses Tests for Pooling Data

Tests of Homogeneity
Gamma Distributions

Question	ν	T	p-value
1	4	14.80	.0051
2	2	1.96	.3745
3	2	11.85	.0027
4	2	5.64	.0595

Tests of Homogeneity
Weibull Distributions

Question	ν	T	p-value
1	4	12.22	.0158
2	2	1.75	.4173
3	2	10.01	.0067
4	2	4.78	.0916

The results obtained from the gamma and Weibull distributions are very similar, indicating that the findings are not dependent on the choice of the family for these two distributions. The p-value associated with the test of Question 1 is small enough to warrant rejection of the null hypothesis in Equation (27). Since the null hypothesis of one model for all three switchyard configurations has been rejected, the next step in the analysis is to proceed with testing the hypotheses in questions 2 through 4. Tests of questions 2 and 4 show that the groups 1 and 2 are not significantly different, and that groups 2 and 3 are not significantly different; however, groups 1 and 3 are significantly different. This is one of the types of problems that can be encountered when making comparisons of this type, in that it is not clear whether to treat group two by itself or to pool these data with either group one or group three. Thus, this would be a good place to supplement the statistical analysis with engineering judgment with respect to what is physically valid for the plants under consideration.

One impact of modeling the plant-centered data based on switchyard configuration is to increase the width of the uncertainty bounds because of the decreased sample size used to fit the probability model for the plant-centered term in Equation (23). The results shown in Figure 26 are based on using only the PC data from Group I2. As one might suspect, the bounds in Figure 26 based on the smaller sample size are wider than those in Figure 23 for the larger sample size. Note that the models shown in Figures 23 and 26 both use a point value of $p_i = 43/63$ for the plant-centered component.

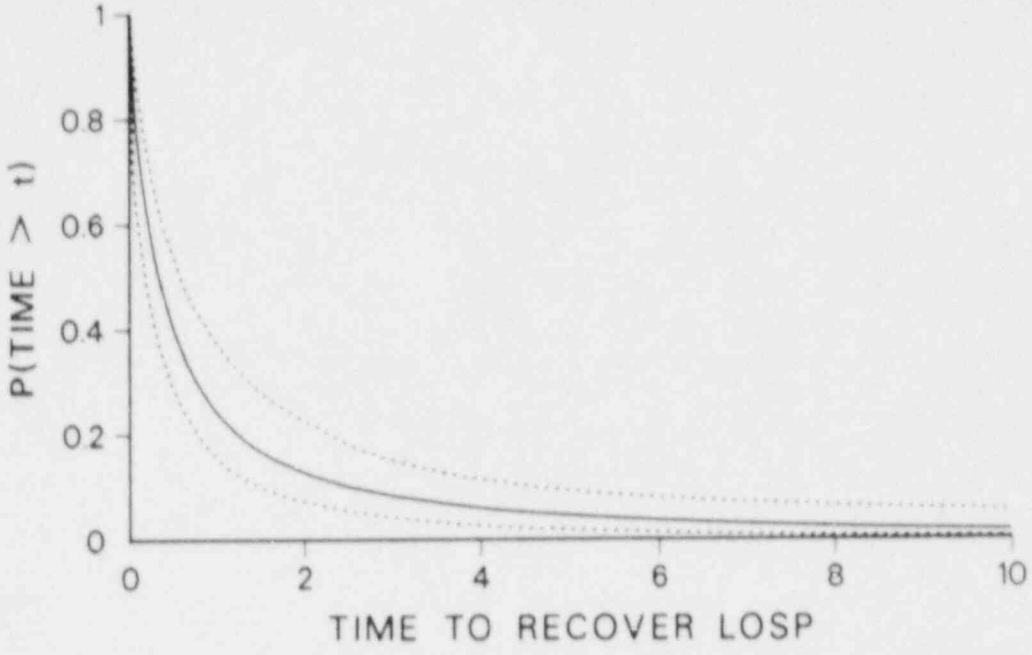

Figure 26. Uncertainty Bounds for G(x) from Equation (23) Based on Using Only the Plant Centered Data in Group I2

Modifying the Weight for the Weather Term in the Composite Model

The weights p_j appearing in the composite model are used to determine the relative influence of each of the three sources of LOSP. Based on the 63 incidents of LOSP, the weights are 43/63 for plant-centered losses, 13/63 for grid losses, and 7/63 for weather-related losses. Thus, if any one weight is adjusted, the remaining weights must also be adjusted, since the weights sum to 1.0. In this subsection the influence of weather is modified by adjusting the weight p_3. This modification is intended to show how to adjust the weights, but not necessarily as the proper way to modify the weather weight for a specific plant.

In NUREG-1032, proportionality factors are presented that are used in modeling the frequency of various types of weather such as snow and ice, winds from 75 to 125 m.p.h., and tornados. These proportionality factors can be combined in a linear fashion with the weather conditions at a specific site to form a weather hazard index for the site. Based on information supplied by the NRC, these indexes range from .00020 to .14508 covering 79 plants. All but two of these indexes are .01310 or less. The median of all indexes is .00510, while the arithmetic mean for the indexes, excluding the two largest observations, is .00566.

An index is simply a scale free measure that can be used to compare the relative effect of weather at different sites. If the weather hazard index is normalized by dividing each index by either the median or truncated mean, the relative comparisons remain unchanged. However, the division does provide a number that can be used to illustrate how to modify the weight p_3 of the

weather term in the composite model. Division by the truncated mean gives weather hazard ratios ranging from .03531 to 2.31300 (excluding the two largest observations). This weather hazard ratio (R) can then be used to modify the effective weight, p_i, in the composite model. That is, the p_i in (23) would be estimated by $p_1/(p_1 + p_2 + p_3R)$, $p_2/(p_1 + p_2 + p_3R)$, and $p_3R/(p_1 + p_2 + p_3R)$ where p_3 is the weight associated with the weather component. Thus, a value of R=1 would result in the composite model as shown in Equation (23), while values of $R > 1$ have the effect of increasing the contribution attributable to weather, and values of $R < 1$ have the opposite effect.

Figure 23 shows the composite model utilizing all plant-centered data and a weather hazard ratio of R=1. Figure 26 shows the composite model utilizing only the Group I2 plant-centered data and a weather hazard ratio of R=1. Figures 27 and 28 show two of the many variations possible with the composite model. Figure 27 illustrates a "best case" situation in that an I1 plant configuration was used (see Figure 25 for reference) with a weather hazard ratio of R = .035. Figure 28 illustrates a "worst case" situation with an I3 configuration and a value of R = 2.313.

Modifying the Grid Component in the Composite Model

The preceding discussion has shown how modifications could be made to the composite model for plant-centered and weather-related LOSP. The data base for grid-related LOSP consists of only 13 values, and 7 of these values come from just one plant. Thus, some plant specific modifications may be warranted. For example, if the seven values from one plant are dropped from the data (perhaps because the problem no longer exists), the basic grid model would not be appreciably changed, but there would be greater uncertainty associated with the grid component. However, the weight, p_i, in the composite model would be lowered from 13/63 = .206 to 6/56 = .107. If information were available similar to that used to split the plant-centered data based on switchyard configuration, or similar to the weather hazard index, then it could be used to modify the grid component of the composite model in a manner similar to what has been demonstrated. With such adjustments, the composite model would be used to produce the uncertainty distribution for $P(T \geq t)$ for a site-specific probabilistic risk assessment.

The Expected Time to Recovery for the Composite Model

Without modifying the weights for the composite model in Equation (23), the expected value of the time to recovery of LOSP is 1.21 hours when the $G_i(x)$ are modeled with gamma distributions, and 1.20 hours when Weibull distributions are used in the model. When the $G_i(x)$ are made plant specific, the respective expected values for I1, I2, and I3 switchyard configurations are 0.90, 1.03 and 1.62 hours for the gamma model and 0.96, 1.03, and 1.63 hours for the Weibull model.

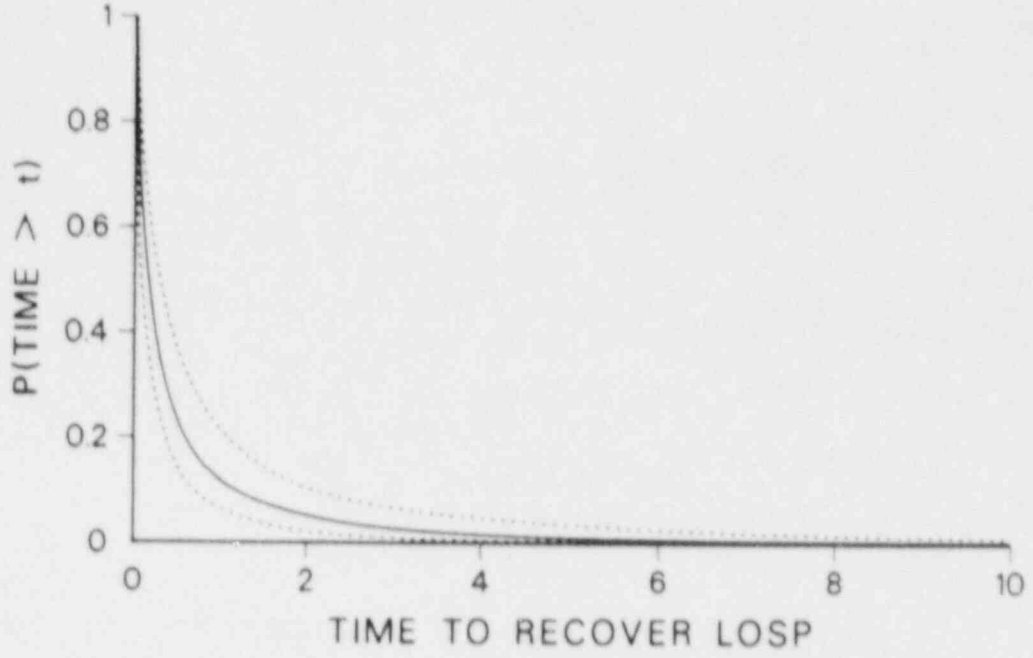

Figure 27. Uncertainty Bounds From the Composite Model With an I1 Plant Configuration and a Weather Hazard Ratio of R = 0.035.

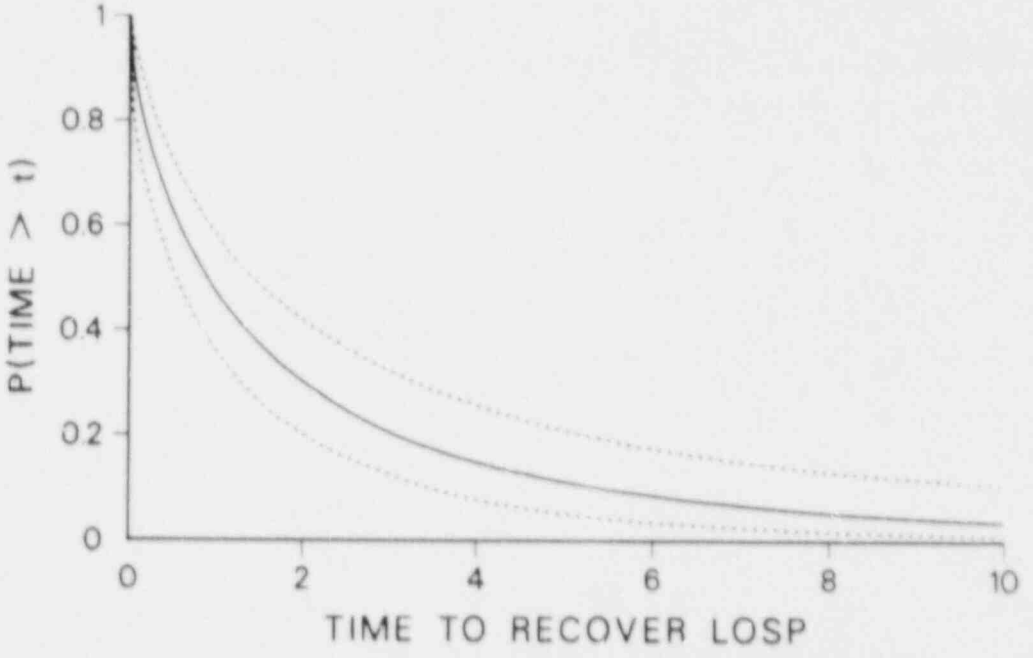

Figure 28. Uncertainty Bounds From the Composite Model With an I3 Plant Configuration and a Weather Hazard Ratio of R = 2.313.

9. A BAYESIAN ANALYSIS OF THE FREQUENCY OF THE INITIATING EVENTS

The modeling thus far in this report has concerned the time to recovery of LOSP, given that a LOSP incident has occurred. In this section the frequency of the initiating events leading to LOSP is modeled under two different assumptions based on Bayesian methods. The first model assumes that all nuclear power generating stations share a common incidence rate for LOSP. Under this model, the data for incidents leading to LOSP are pooled together and the posterior distribution is found for the mean incidence rate for all plants. The second approach models the incidence rate for each plant relative to the incidence rates of all other plants, and the posterior distribution is found for the incidence rate for each plant.

Bayesian Analysis Assuming a Common Incidence Rate

When all nuclear power generating stations are assumed to share a common incidence rate, the procedures outlined in Martz and Waller (1982, pages 254-5) can be followed. With these procedures the number of LOSP incidents n, occurring in time t, is modeled by a Poisson distribution.

$$f(n|\lambda) = P(n \text{ occurrences in total time } t|\lambda)$$

$$= e^{-\lambda t}(\lambda t)^n/n! \qquad \lambda > 0, \quad n = 0, 1, 2, \ldots \tag{29}$$

where the parameter λ is the intensity of the process.

Martz and Waller (pages 286-7) show that a noninformative prior for λ of $1/\lambda^{1/2}$ yields a posterior distribution for λ given n of

$$p(\lambda|n) = \frac{p(\lambda)\ f(n|\lambda)}{\int_0^\infty p(\lambda)\ f(n|\lambda)\ d\lambda}$$

$$= \frac{(1\ /\ \lambda^{1/2})\ e^{-\lambda t}\ (\lambda t)^n/n!}{\int_0^\infty (1\ /\ \lambda^{1/2})\ e^{-\lambda t}\ (\lambda t)^n/n!\ d\lambda}$$

$$= \frac{t^{n+1/2}\ \lambda^{n-1/2}\ e^{-\lambda t}}{\Gamma(n + 1/2)}, \qquad \lambda > 0 \tag{30}$$

Given n, $2\lambda t$ follows a $\chi^2(2n+1)$ distribution with

$$E(\lambda|n) = (2n + 1)/2t \tag{31}$$

Table 9. Results of Using Equations 29 to 31 for Initiating Event Frequencies

	Weather & Grid	Plant Centered
$E(\lambda\|n)$	.0284	.0864
$Var(\lambda\|n)$	3.941E-5	1.717E-4
Lower 5% Bound	.0189	.0661
Upper 5% Bound	.0394	.1090

$$Var(\lambda|n) = (2n + 1)/2t^2 \qquad (32)$$

A two-sided Bayes probability interval for λ may be derived from the $\chi^2(2n+1)$ distribution as

$$[\chi^2_{\alpha/2}(2n+1)/2t,\ \chi^2_{1-\alpha/2}(2n+1)/2t] \qquad (33)$$

The data presented in Table 1 showed 43 incidents of LOSP that were plant-centered, 13 incidents that were grid-related, and 7 incidents that were weather-related. For the analysis presented in this section the grid and weather incidents are pooled together since they both influence the grid. Plant operating data through December 1987 shows 43 plant-centered incidents in 503.29 operating years, and 20 weather and grid incidents in 721.26 calender years. Using these numbers with Equations (31) to (33) gives the results shown in Table 9.

The total frequency of LOSP is the sum of the weather/grid losses and the plant-centered losses. Thus, the uncertainty in the initiating event frequency is modeled by finding the distribution of the random variable $Z = \lambda_1 + \lambda_2$, where λ_1 is the failure rate for weather/grid losses, and λ_2 is the failure rate for plant-centered losses. Since the random variable $2\lambda_i t$ has a $\chi^2(2n+1)$ distribution, it is necessary to make a change of variable to find the distribution of λ_i and then in turn to find the distribution of Z.

The chi-squared distribution with ν degrees of freedom has the density

$$f(y) = \frac{y^{\nu/2-1}\ e^{-y/2}}{\Gamma(\nu/2)\ 2^{\nu/2}}$$

Let $y = 2\lambda t$ and $dy = 2t\, d\lambda$, thus

$$f(\lambda) = \frac{(2\lambda t)^{\nu/2-1}\, e^{-2\lambda t/2}\, 2t}{\Gamma(\nu/2)\, 2^{\nu/2}} = \frac{\lambda^{\nu/2-1}\, t^{\nu/2}\, e^{-\lambda t}}{\Gamma(\nu/2)} \quad (34)$$

This is a gamma density as appeared in Equation (3), with $\alpha = \nu/2$ and $\beta = t$. The density of the random variable $Z = \lambda_1 + \lambda_2$ is given as follows:

$$g(z) = \int_0^z f_{\lambda_1}(z - \lambda_2)\, f_{\lambda_2}(\lambda_2)\, d\lambda_2 \quad (35)$$

and the distribution function is

$$G(z^*) = \int_0^{z^*} g(z)\, dz \quad (36)$$

$$= \frac{t_1^{\nu_1/2}\, t_2^{\nu_2/2}}{\Gamma(\nu_1/2)\, \Gamma(\nu_2/2)} \int_0^{z^*} \int_0^{z} (z-\lambda_2)^{\nu_1/2-1}\, e^{-(z-\lambda_2)t_1}\, \lambda^{\nu_2/2-1}\, d\lambda_2\, dz$$

The integrand in Equation (36) is a confluent hypergeometric function and would have to be evaluated numerically. While this could be done, the numerical evaluation is quite difficult in this case because of the particular values of the parameters for this problem. Therefore, it is somewhat easier in this case to use a Monte Carlo procedure to approximate the distribution of the sum of two gamma random variables. For the plant-centered frequency, the gamma distribution has parameters of $\alpha = 43.5$ and $\beta = 503.29$, while the weather/grid frequency has parameters $\alpha = 20.5$ and $\beta = 721.26$.

A Monte Carlo procedure with n = 1000 was used to estimate the distribution function for the frequency of the initiating events leading to LOSP. The graph of the estimated distribution function is given in Figure 29. The summary statistics from that simulation are given in Table 10. The sample mean is very close to the sum of the expected values of the weather and grid component, and the plant-centered component given above, that is .0284 + .0864 = .1149. Note also that the sample mean is consistent with the quote from NUREG-1032 given in the introduction as "... LOSP has occurred about once per 10 site years."

Table 10. Summary Statistics for the Monte Carlo Derivation of the Distribution for Frequency of Loss of Off-Site Power

$\bar{X}$	X_{min}	$X_{.05}$	$X_{.50}$	$X_{.95}$	X_{max}
.1146	.0684	.0929	.1130	.1389	.1734

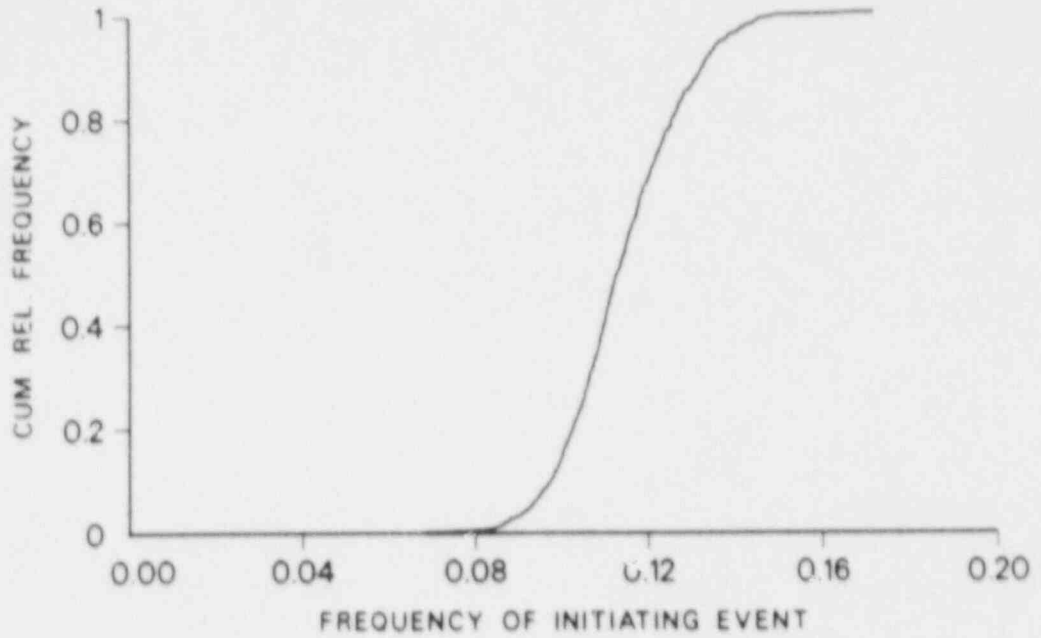

Figure 29. Estimated Cumulative Distribution Function for the Frequency of Initiating Events Leading to LOSP

Bayesian Analysis Assuming Individual Incidence Rates

An alternative to the model just presented, which assumes that all power generating stations share a common incidence rate, is to model the incidence rate for each plant individually. Under this model the incidence rates for individual plants are assumed to belong to a superpopulation of incidence rates, thus the incidence rate for an individual plant is assumed to have been randomly selected from this superpopulation of incidence rates. This approach was used by Hora and Iman (1987) to model the failure rate of the solid rocket boosters on the space shuttle. In that application the failure rate for predecessor solid rocket systems, as well as the failure rate for the space shuttle were assumed to belong to a superpopulation of failure rates.

Data are available for n+1 plants. The value λ_i, $i = 0, 1, \ldots, n$ is the "true" incidence rate/year for the ith plant, which has operated for t_i years with s_i incidents of LOSP. Thus, s_i/t_i

provides an estimate of λ_i. With this notation the plant of interest is i=0.

For this analysis the gamma distribution is used as a model for the superpopulation, although many other distributions could be used. The probability density function for the two-parameter gamma distribution was previously given in Equation (3) as

$$h(\lambda) = f_\gamma(\lambda|\alpha,\beta) = \beta^\alpha[\Gamma(\alpha)]^{-1}\ \lambda^{\alpha-1}\ e^{-\beta\lambda} \qquad \lambda \geq 0,\ \alpha,\beta > 0 \quad (37)$$

The parameters α and β are unknown, and as in Section 6, the noninformative prior

$$p(\alpha,\beta) \propto 1/(\alpha\beta) \qquad \alpha,\beta > 0 \quad (38)$$

is used. The likelihood function of the datum (s_i,t_i) is Poisson

$$L(s_i,t_i|\lambda_i) = (\lambda_i t_i)^{s_i}\ e^{-\lambda_i t_i}/s_i! \quad (39)$$

The joint probability of $\alpha,\beta,\lambda_0,\lambda_1,\ldots,\lambda_n,s_0,s_1,\ldots,s_n$ is

$$p(\alpha,\beta)\ \prod_{i=0}^{n}\ [h(\lambda_i)\ L(s_i,t_i|\lambda_i)] \quad (40)$$

The posterior density of $\alpha,\beta,\lambda_0,\lambda_1,\ldots,\lambda_n$ is

$$p^*(\alpha,\beta,\lambda_0,\ldots,\lambda_n) =$$

$$\frac{p(\alpha,\beta)\ \prod_{i=0}^{n}\ [h(\lambda_i)\ L(s_i,t_i|\lambda_i)]}{\int_0^\infty\int_0^\infty\int_0^\infty\cdots\int_0^\infty\ p(\alpha,\beta)\ \prod_{i=0}^{n}\ [h(\lambda_i)\ L(s_i,t_i|\lambda_i)]\ d\alpha\ d\beta\ d\lambda_0\ldots d\lambda_n} \quad (41)$$

The denominator of (41) is a numerical constant that depends on the s_i, t_i but not on α, β, λ_0, ..., λ_n. Denoting this constant by 1/k, (41) becomes

$$k\ p(\alpha,\beta) \prod_{i=0}^{p} [h(\lambda_i)\ L(s_i,t_i|\lambda_i)] \tag{42}$$

$$= k\ \frac{1}{\alpha\beta} \prod_{i=0}^{n} \left[\frac{\beta^{\alpha}\ \lambda_i^{\alpha-1}\ e^{-\beta\lambda_i}}{\Gamma(\alpha)} \cdot \frac{(\lambda_i t_i)^{s_i}\ e^{-\lambda_i t_i}}{s_i!} \right] \tag{43}$$

Combining terms gives

$$p^*(\alpha,\beta,\lambda_0,\ldots,\lambda_n) =$$

$$= \frac{k}{\alpha\beta} \cdot \frac{\beta^{(n+1)\alpha}}{[\Gamma(\alpha)]^{n+1}} \cdot \prod_{i=0}^{n} \frac{\lambda_i^{\alpha+s_i-1}\ e^{-\lambda_i(\beta+t_i)}\ t_i^{s_i}}{s_i!} \tag{44}$$

In order to obtain the posterior density of λ_0, Equation (44) must be integrated with respect to $\lambda_1,\ldots,\lambda_n,\alpha,\beta$. Noting that

$$\int_0^{\infty} \frac{\lambda_i^{\alpha+s_i-1}\ e^{-\lambda_i(\beta+t_i)}\ t_i^{s_i}}{s_i!}\ d\lambda_i = \frac{\Gamma(\alpha+s_i)\ t_i^{s_i}}{(\beta+t_i)^{\alpha+s_i}\ s_i!} \tag{45}$$

allows Equation (44) to be simplified to

$$p^*(\alpha,\beta,\lambda_0) =$$

$$= \frac{k}{\alpha\beta} \frac{\beta^{(n+1)\alpha}}{[\Gamma(\alpha)]^{n+1}} \frac{t_0^{s_0}\ \lambda_0^{\alpha+s_0-1}\ e^{-\lambda_0(\beta+t_0)}}{s_0!} \prod_{i=1}^{n} \frac{\Gamma(\alpha+s_i)\ t_i^{s_i}}{(\beta+t_i)^{\alpha+s_i}\ s_i!} \tag{46}$$

It remains to integrate $p^*(\alpha,\beta,\lambda_0)$ numerically with respect to α and β and to find k. The value k is found numerically by the relation,

$$\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} p^*(\alpha,\beta,\lambda_0) = 1 \tag{47}$$

Thus,

$$\frac{1}{k} = \int_0^\infty \int_0^\infty \frac{\beta^{(n+1)\alpha}}{\alpha\beta[\Gamma(\alpha)]^{n+1}} \prod_{i=0}^{n} \frac{\Gamma(\alpha+s_i)\, t_i^{s_i}}{(\beta+t_i)^{\alpha+s_i}\, s_i!}\, d\alpha\, d\beta \qquad (48)$$

Thus, substituting the numerical value of k from Equation (48) into Equation (46) and integrating Equation (46) with respect to α and β produces the posterior density of λ_0.

The posterior cumulative distribution function of λ_0 is

$$F(\lambda_0) =$$

$$k \int_0^\infty \int_0^\infty \frac{\beta^{(n+1)\alpha}}{\alpha\beta[\Gamma(\alpha)]^{n+1}} F_\gamma(\lambda_0|\alpha+s_0, \beta+t_0) \prod_{i=0}^{n} \frac{\Gamma(\alpha+s_i)\, t_i^{s_i}}{(\beta+t_i)^{\alpha+s_i}\, s_i!} d\alpha d\beta \qquad (49)$$

where $F_\gamma(\lambda_0|\alpha+s_0, \beta+t_0)$ is the cumulative gamma distribution function,

$$F_\gamma(\lambda_0|\alpha+s_0, \beta+t_0) = \int_0^{\lambda_0} \frac{y^{\alpha+s_0-1}\, (\beta+t_0)^{\alpha+s_0}\, e^{-(\beta+t_0)y}}{\Gamma(\alpha+s_0)}\, dy. \qquad (50)$$

The Bayesian analysis presented in the first part of this section modeled the plant-centered and grid/weather data separately and used a Monte Carlo procedure to find the distribution of the sum of the two posterior random variables. A similar approach is applied here with respect to Equation (49). Three typical cases have been selected to demonstrate the technique. These cases are as follows.

	Plant-Centered		Grid/Weather	
Case	s_0	t_0	s_0	t_0
1	0	11.00	0	13.51
2	0	2.27	0	4.00
3	1	9.81	0	13.09

Cases 1 and 2 have no recorded incidents of LOSP, with Case 1 having a much longer operating history. In Case 3 the plant has been in operation for a time similar to Case 1, however there was been one occurrence of a plant-centered LOSP. The distribution functions for these three cases are shown in Figures 30 to 32, respectively. The distribution for the common incidence model

Table 11. Summary Statistics for the Monte Carlo Derivation of the Distribution for Frequency of LOSP Using the Bayesian Analysis for Individual Incidence Rates

	$\hat{\chi}$	χ_{min}	$\chi_{.05}$	$\chi_{.50}$	$\chi_{.95}$	χ_{max}
Previous Results Using a Common Incidence Rate	.1147	.0684	.0929	.1136	.1389	.1734
Case	Results Using Individual Incidence Rates					
1	.0915	.0000	.0207	.0900	.1686	.3505
2	.1093	.0001	.0276	.0974	.2265	.5514
3	.1111	.0000	.0453	.1010	.2106	.4514

given in Figure 29 is repeated in Figure 30 for ease of comparing the differences in the two modeling approaches. Figure 30 makes it clear that the model assuming individual incidence rates produces the greater uncertainty. Table 11 has been constructed to shown the summary statistics in comparison to the mean results shown previously in Table 10. The results in Table 11 emphasize the greater uncertainty for the individual distributions, compared to the mean distribution given by the previous model. This greater uncertainty is due to the shorter operating history for individual cases compared to combining all operating histories. As shown in Figure 31, Case 2, with a shorter operating history than Case 1, has a much longer tail in the upper part of the distribution, indicating greater uncertainty than in Case 1, even though both have no reported incidents of LOSP. Moreover, Table 11 shows the .95 quantile to be .1686 for Case 1 and .2265 for Case 2. The means in Cases 1 and 2, however, are both lower than the mean of .1147 given for the mean model. Cases 1 and 3 are shown jointly in Figure 32. In spite of the similar operating histories for these cases, the uncertainty in Case 3 is similar to that for Case 2 because of the one recorded incident of LOSP. Additionally, this causes the mean to be higher in Case 3 than in Case 1.

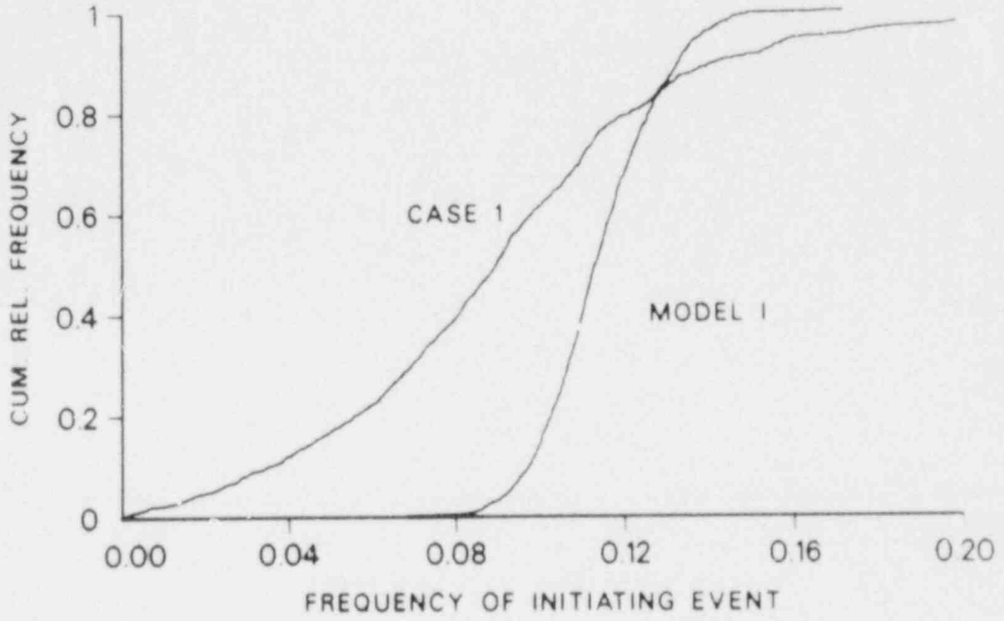

Figure 30. Estimated Cumulative Distribution Function for the Frequency of Initiating Events Leading to LOSP for Case 1 Compared to the Mean Model

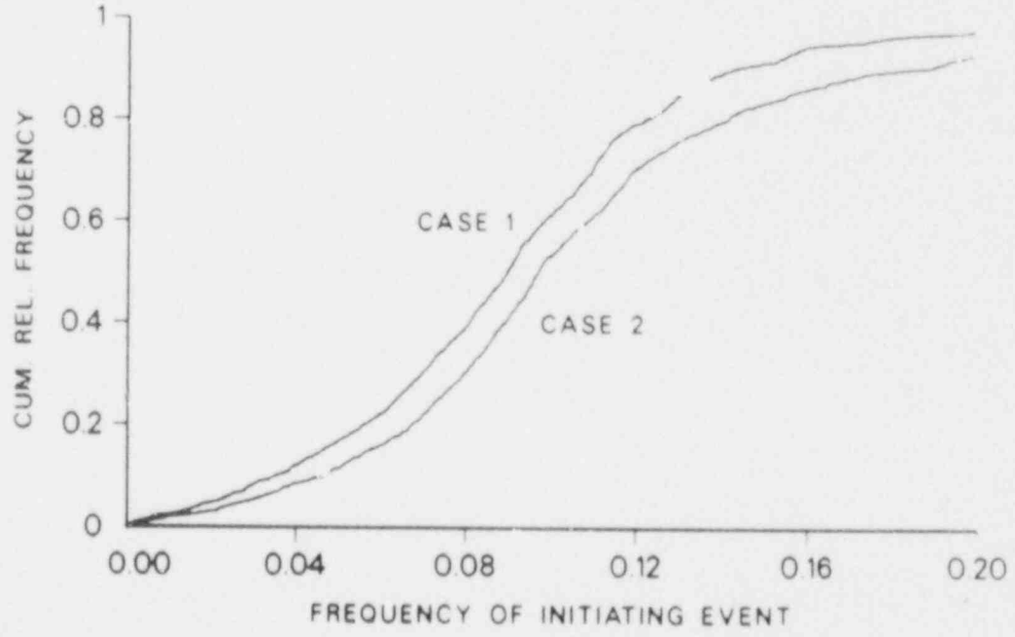

Figure 31. Estimated Cumulative Distribution Function for the Frequency of Initiating Events Leading to LOSP for Case 2 Compared to Case 1

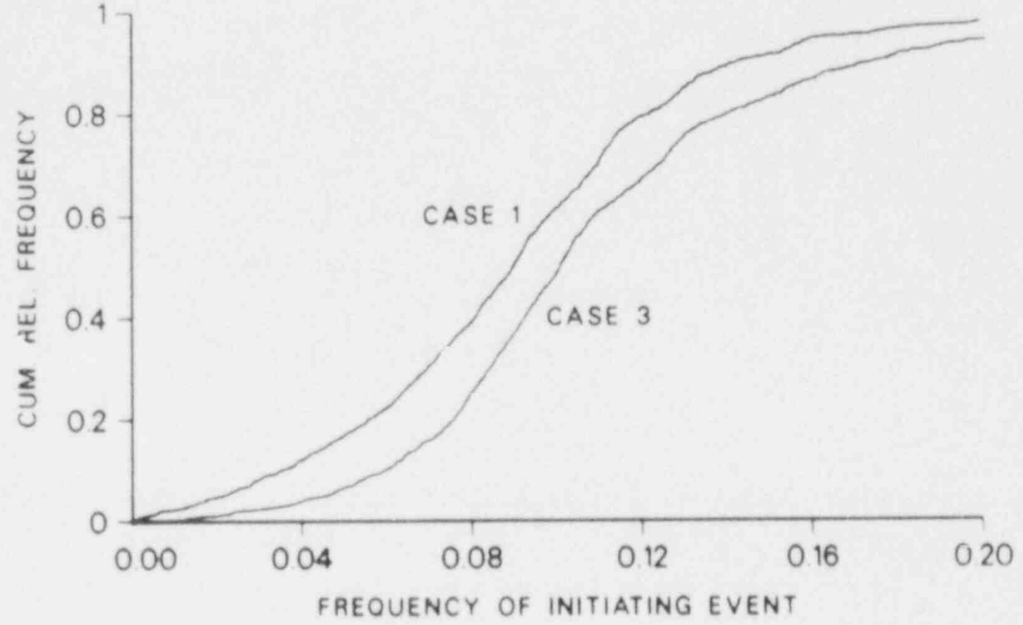

Figure 32. Estimated Cumulative Distribution Function for the Frequency of Initiating Events Leading to LOSP for Case 3 Compared to Case 1

REFERENCES

Barnowsky, P. W. (1985). "Evaluation of Station Blackout Accidents at Nuclear Power Plants," Draft Version, NUREG-1032, Nuclear Regulatory Commission.

Conover, W. J. (1980). **Practical Nonparametric Statistics, 2nd Ed.**, John Wiley & Sons, Inc., New York.

Hora, S. C. and Iman, R. L. (1987). "Bayesian Analysis of Learning in Risk Analyses," **Technometrics**, 29(2), 221-228.

Jeffreys, H. (1961). **Theory of Probability, 3rd Ed.**, Clarendon Press, Oxford.

Lawless, J. F. (1982). **Statistical Models and Methods for Lifetime Data**, John Wiley & Sons, Inc., New York.

Martz, H. F. and Waller, R. A. (1982). **Bayesian Reliability Analysis**, John Wiley & Sons, Inc., New York.

DISTRIBUTION:

U.S. Government Printing Office
Receiving Branch (Attn: NRC Stock)
8610 Cherry Lane
Laurel, MD 20707
250 copies for RG, 1S

US NRC (53 copies total)
Office of Nuclear Regulatory Research
Division of Risk Analysis and Operations
Washington, DC 20555
B. Agrawal
P. Barnowsky (5)
G. Burdick
F. Coffman
J. Craig (3)
M. Cunningham (5)
W. Farmer
J. Flack (5)
J. Johnson (5)
C. Johnston
L. Lancaster
T. Margulies
J. Murphy (5)
A. Potapous
D. Rasmuson (5)
F. Rosa (5)
A. Rubin (5)
A. Serkiz (5)

Los Alamos National Laboratories (5)
Group S1, MS 606
Los Alamos, NM 87545
Attn: H. F. Martz
R. J. Beckman
M. E. Johnson
M. D. McKay
R. A. Waller

David C. Aldrich
SAIC
1710 Goodridge Drive
P. O. Box 1303
McLean, VA 22102

S. Alexander
Vendor Inspection Branch
NRR/DRIS, M/S EWS 346
US Nuclear Regulatory Commission
Washington, DC 20555

George Apostolakis
UCLA
5532 Boelter Hall
Los Angeles, CA 90024-1597

A. Bayer
INR-Kernforschungszentrum Karlsruhe
D-7500 Karlsruhe 1
Postfach 3640
WEST GERMANY

R. Battle (2)
Oak Ridge National Laboratory
Building 3500
P. O. Box X
Oak Ridge, TN 37830

Robert Bertucio
Energy Incorporated Services
1851 So. Central Place
Suite 201
Kent, WA 98031

Dennis Bley
Pickard, Lowe & Garrick
2260 University Drive
Newport Beach, CA 92660

Gary Boyd
Safety & Reliability Optimization
Services
624 Glen Willow Dr.
Knoxville, TN 37922

Carla Brofferio
Comitato Nazionale per l'Energia Nucleare
Viale Regina Margherita, 125
Casella Postale N. 2358
I-00100 Roma A.D.
ITALY

Robert Budnitz
Future Resources Associates
2000 Center Street
Suite 418
Berkeley, CA 94707

Klaus Burkart
Institut fur Neutronenphysik and
Reaktortechnik (INR)
Kernforschungszentrum Karlsruhe G.m.b.H.
Postfach 3640
D-7500 Karlsruhe 1
WEST GERMANY

Pietro Cagnetti
PAS
Comitato Nazionale per l'Energia Nucleare
Centro di Studi Nucleari della Casaccia
Via Anguillarese km 1+300
I-00060 Roma
ITALY

David Campbell
JBF Assoc. Inc.
1000 Tech Park Center
Knoxville, TN 37932

James E. Campbell
Intera Environmental Consultants Inc.
6850 Austin Center Blvd.
Suite 300
Austin, TX 78731

M. Childers (2)
Nuclear Utilities Group on
Station Blackout
1200 17th Street NW, Suite 800
Washington, DC 20036

Michael L. Corradini
University of Wisconsin-Madison
Nuclear Engineering Dept.
145 Engr. Res. Building
1500 Johnson Drive
Madison, WI 53706

W. J. Conover (10)
College of Business Administration
Texas Tech University
Lubbock, TX 79409

George Crane
1570 E. Hobble Creek Dr.
Springville, Utah 84663

Malcolm Crick
National Radiological Protection Board
Chilton
Didcot
Oxon. OX110RQ
United Kingdom

Pamela Doctor (2)
Battelle Northwest
P. O. Box 999
Richland, WA 99352

Darryl Downing (6)
P. O. Box Y. Bldg. 9207A, MS2
ORNL
Oak Ridge, TN 37831

Mary Drouin (10)
SAIC
505 Marquette NW
Suite 1200
Albuquerque, NM 87102

James Dukelow
Westinghouse Hanford Co.
Mgr., Preclosure Safety
CDC-1/234/3000 Area
P.O. Box 1970
Richland, WA 99352

Friedmar Fischer
Kernforschungszentrum Karlsruhe G.m.b.H.
Postfach 3640
D-7500 Karlsruhe 1, WEST GERMANY

Karl Fleming
Pickard, Lowe & Garrick
2260 University Drive
Newport Beach, CA 92660

Ralph R. Fulwood
Engineering Technology Division
BNL, Bldg. 130
Upton, NY 11973

J. Gaertner
Electric Power Research Institute
3412 Hillview Ave.
Palo Alto, CA 94303

R. H. Gardner
Environmental Sciences Division
ORNL
Oak Ridge, TN 37830

Richard Gunst (5)
Department of Statistics
Southern Methodist University
Dallas, TX 75275

Allan Gutjahr
Department of Mathematics
NMIMT
Socorro, NM 87801

William V. Harper (2)
Performance Analysis Department
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201

Bernard Harris
Department of Statistics
University of Wisconsin
1210 W. Dayton St.
Madison, WI 53711

Michael Haynes
United Kingdom Atomic Energy Authority
Safety & Reliability Directorate
Wigshaw Lane
Culcheth
Warrington WA3 4NE
UNITED KINGDOM

F. O. Hoffman
Health and Safety Research Division
ORNL
Oak Ridge, TN 37830

Stephen C. Hora (10)
Division of Business & Economics
University of Hawaii at Hilo
1400 Kapiolani St.
Hilo, HI 96720

Frank W. Horsch
Kernforschungszentrum Karlsruhe G.m.b.H.
Postfach 3640
D-7500 Karlsruhe 1
WEST GERMANY

Toshinori Iijima
Division of Reactor Safety Evaluation
Reactor Safety Research Center
Japan Atomic Energy Research Institute
Tokai Research Establishment
Tokai-mura
Naka-gun
Ibaraki-ken 319-11
JAPAN

Geoffrey D. Kaiser
Consulting Division
NUS Corporation
910 Clopper Road
Gaithersburg, MD 20878

Stan Kaplan
Pickard, Lowe & Garrick
2260 University Drive
Newport Beach, CA 92660

Samuel C. Kao (2)
Applied Mathematics, 515
Brookhaven National Laboratory
Upton, NY 11973

William E. Kastenberg
Mechnical, Aerospace and Nuclear
Engineering Department
School of Engineering & Applied
Science
University of California, Los Angeles
Los Angeles, CA 90024

Alan Kolaczkowski (10)
SAIC
505 Marquette NW
Suite 1200
Albuquerque, NM 87102

Jim Kolanowski (2)
Commonwealth Edison Co.
35 1st National West
Chicago, IL 60690

Jerald F. Lawless
Dept. of Statistics and
Actuarial Science
University of Waterloo
Waterloo, Ontario
CANADA N21 3G1

E. V. Lofgren
Science Applications Int. Corp.
1710 Goodrich Drive
McLean, VA 22102

Steve Maloney
Devonrue
180 Linden St.
Boston, MA 02111

Daniel Manesse
Institut de Protection et de
Surete Nucleaire (IPSN)
Commissariat a l'Energie Atomique
Centre d'Etudes Nucleaires de
Fontenay-aux-Roses
Boite Postale 6
F-92260 Fontenay-aux-Roses
FRANCE

David S. Margoles
Mathematics and Statistics Division
Lawrence Livermore Laboratory
P. O. Box 808 (L-316)
Livermore, CA 94550

M. McGarry
Nuclear Utilities Group on
Station Blackout
1200 17th Street NW, Suite 800
Washington, DC 20036

Shan Nair
Research Division
Central Electricity Generating Board
Berkeley Nuclear Laboratories
Berkeley
Gloucestershire GL13 9PB
UNITED KINGDOM

William Nixon
United Kingdom Atomic Energy Authority
Safety & Reliability Directorate
Wigshaw Lane
Culcheth
Warrington WA3 4NE
UNITED KINGDOM

Gareth Parry
NUS Corporation
910 Clopper Rd.
Gaithersburg, MD 20878

Blake Putney
Science Applications Int. Corp.
5150 El Camino Real, Suite C31
Los Alto, CA 94022

Norman C. Rasmussen
Department of Nuclear Engineering
Massachusetts Institute of Technology
77 Mass Avenue
Cambridge, MA 02139

Ilkka Savolainen
Technical Research Centre of Finland
Nuclear Engineering Laboratory
P. O. Box 169
SF-00181 Helsinki 18
FINLAND

Sebastiano Serra
ENEL-DCO
Ente Nazionale per l'Energia Elettrica
Via G.B. Martini, 3
Casella Postale N. 386
I-00186 Roma
ITALY

Juan Bagues Somonte
Consejo De Seguridad Nuclear
Sor Angela de la Cruz, 3
Madrid-28020
SPAIN

David A. Stanners
Commission on European Communities
Joint Research Center
Ispra Establishment
21020 Ispra (Varese)
ITALY

John R. D. Stoute
Health Physics Division
Energieonderzoek Centrum Nederland (ECN)
Westerduinweg, 3
Postbus 1
NL-1755 Petten ZG
NETHERLANDS

Søren Thykier-Nielsen
Health Physics Department
Risø National Laboratory
Postbox 49
DK-4000 Roskilde
DENMARK

Alfred Torri
Pickard, Lowe and Garrick, Inc.
191 Calle Magdalena
Suite 290
Encinitas, CA 92024

Ulf Tveten
Institute for Energy Technology
Postboks 40
N-2007 Kjeller
NORWAY

Stephen C. Unwin (3)
Department of Nuclear Energy
Division of Safety and Risk Evaluation
Brookhaven National Laboratory
Bldg. 130, Brookhaven Lane
Upton, NY 11973

William E. Vesely
Science Applications Int. Corp.
2929 Kenny Rd., Suite 245
Columbus, OH 43221

I. B. Wall
Electric Power Research Institute
3412 Hillview Avenue
P. O. Box 10412
Palo Alto, CA 94303

Keith Woodard
Pickard, Lowe & Garrick, Inc.
1200 18th Street, NW
Suite 612
Washington, DC 20036

H. Wykoff
EPRI
3412 Hillview Ave.
Palo Alto, CA 94303

R. Wyrick
Institute for Nuclear
Power Operations
1100 Circle 75 Parkway, Suite 1500
Atlanta, GA 30339

Eric R. Ziegel
Standard Oil Company (Indiana)
Amoco Research Center
P. O. Box 400
Naperville, IL 60566

2561 G. W. Smith
2614 M. J. Hannah
3141 S. A. Landenberger (5)
3151 W. L. Garner
6400 D. J. McCloskey
6410 N. R. Ortiz
6412 A. L. Camp
6412 W. R. Cramond
6412 D. M. Kunsman
6412 K. J. Maloney
6412 A. C. Payne
6412 T. T. Sype
6412 T. A. Wheeler
6412 D. W. Whitehead
6413 E. D. Gorham-Bergeron
6413 R. J. Breeding
6413 J. J. Gregory
6413 F. T. Harper
6415 F. E. Haskin
6415 M. K. Carmel
6415 S. E. Dingman
6415 J. C. Helton
6415 R. L. Iman (15)
6415 H. N. Jow
6415 M. J. Shortencarier
6415 J. L. Sprung
6415 G. D. Wyss
6416 R. M. Cranwell
6416 I. J. Hall
6418 J. E. Kelly
6447 M. P. Bohn
8024 P. W. Dean

NRC FORM 335 (2-84) NRCM 1102, 3201, 3202

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

SEE INSTRUCTIONS ON THE REVERSE

1. REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)
NUREG/CR-5032
SAND87-2428

2. TITLE AND SUBTITLE
MODELING TIME TO RECOVERY AND INITIATING EVENT FREQUENCY FOR LOSS OF OFF-SITE POWER INCIDENTS AT NUCLEAR POWER PLANTS

3. LEAVE BLANK

4. DATE REPORT COMPLETED
MONTH: December YEAR: 1987

5. AUTHOR(S)
R. L. Iman and S. C. Hora

6. DATE REPORT ISSUED
MONTH: January YEAR: 1988

7. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Sandia National Laboratories
Albuquerque, NM 87185-5800

8. PROJECT/TASK/WORK UNIT NUMBER

9. FIN OR GRANT NUMBER
A1393

10. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Division of Risk Analysis and Operations
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

11a. TYPE OF REPORT

b. PERIOD COVERED (Inclusive dates)

12. SUPPLEMENTARY NOTES

13. ABSTRACT (200 words or less)

Industry data representing the time to recovery of loss of off-site power at nuclear power plants for 63 incidents caused by plant-centered losses, grid losses, or severe weather losses are fit with exponential, lognormal, gamma and Weibull probability models. A Bayesian analysis is used to compare the adequacy of each of these models and to provide uncertainty bounds on each of the fitted models. A composite model that combines the probability models fitted to each of the three sources of data is presented as a method for predicting the time to recovery of loss of off-site power. The composite model is very general and can be made site specific by making adjustments on the models used, such as might occur due to the type of switchyard configuration or type of grid, and by adjusting the weights on the individual models, such as might occur with weather conditions existing at a particular plant. Adjustments in the composite model are shown for different models used for switchyard configuration and for different weights due to weather. Bayesian approaches are also presented for modeling the frequency of initiating events leading to loss of off-site power. One Bayesian model assumes that all plants share a common incidence rate for loss of off-site power, while the other Bayesian approach models the incidence rate for each plant relative to the incidence rates of all other plants. Combining the Bayesian models for the frequency of the initiating events with the composite Bayesian model for recovery provides the necessary vehicle for a complete model that incorporates uncertainty into a probabilistic risk assessment.

14. DOCUMENT ANALYSIS - a. KEYWORDS/DESCRIPTORS

b. IDENTIFIERS/OPEN-ENDED TERMS

15. AVAILABILITY STATEMENT
Unlimited

16. SECURITY CLASSIFICATION
(This page) Unclassified
(This report) Unclassified

17. NUMBER OF PAGES
60

18. PRICE

☆ U.S. GOVERNMENT PRINTING OFFICE: 1988—773-049/61035