Medium Voltage Cable Failure Trends

Research Status Report

1002256

Medium Voltage Cable Failure Trends

Research Status Report

1002256

Technical Update, December 2003

EPRI Project Manager

R. Keefe

EPRI • 3412 Hillview Avenue, Palo Alto, California 94304 • PO Box 10412, Palo Alto, California 94303 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

VMN Group LLC

This is an EPRI Level 2 report. A Level 2 report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.

ORDERING INFORMATION

Requests for copies of this report should be directed to the EPRI Distribution Center, 207 Coggins Drive, P.O. Box 23205, Pleasant Hill, CA 94523, (800) 313-3774.

CITATIONS

This document was prepared by

VMN Group LLC
26789 Shady Oaks Court
Los Altos Hills, CA 94022

Principal Investigators

C. Feinstein, Ph.D.

P. Morris, Ph.D.

This document describes research sponsored by EPRI.

The publication is a corporate document that should be cited in the literature in the following manner:

Medium Voltage Cable Failure Trends: Research Status Report, EPRI, Palo Alto, CA: 2003. 1002256.

ABSTRACT

The primary objective of this report is to present a method whereby utilities can apply their existing data on cable inventory and create a useful hazard function. Further, this report provides a default hazard function that can be used by utilities that have insufficient data. The purpose of the hazard function is to forecast the failures of the cable inventory over time. Therefore, another objective of this project is to provide a method for forecasting failures as a consequence of cable inventory and hazard functions.

This project is closely related to a companion project, the Cable Reliability Management Strategies project. The management strategy model described in the companion report, *Cable Reliability Management Strategies: Research Status Report,* EPRI, Palo Alto, CA: 2003 {1002257}, determines an optimal policy for underground cable repair and replacement in distribution systems. The hazard functions in the present report provide key inputs for the replacement strategy models developed in the Cable Reliability Management Project.

CONTENTS

1 INTRODUCTION

This report presents a method that can be used by utilities to estimate the hazard function of an underground cable system. The hazard function is a mathematical representation of the dynamics of failure of an asset, such as an underground cable. The hazard function is estimated from observed failure data.

There are four additional parts to this report. In the next part, Chapter 2, we determine which failure models are appropriate to describe cable failure. We have used piecewise linear hazard functions in earlier studies. In this report, we will extend the class of hazard functions to include Weibull, lognormal, normal, exponential and extreme distributions. In Chapter 3, we provide parameter estimation procedures. All the hazard functions are *parametric*, which means that a small number of parameters (e.g., mean and variance) are sufficient to characterize the entire distribution. In Chapter 4, we discuss how to select a hazard function and offer a default for those that have either insufficient data or experience in analyzing hazards. In Chapter 5, we use the hazard function to forecast failures of the cable inventory.

Failure and Repair

The context of the present research, specifying the hazard function for underground cable, is failure and repair of such underground cable. We addressed the general methodology issue in our report *Reliability of Electric Utility Distribution Systems: EPRI White Paper (1000424).* The most popular models in the literature represent the failure and repair of equipment as a Markov process (Hillier (1986)). These Markov models provide a useful perspective on the key issues, so we will provide a brief summary here.

The Markov process identifies various asset states (e.g., “failed”, “operating normally”) that characterize the condition of equipment. The transitions from state to state are governed by probability distributions. The typical assumptions made include specification of constant transition rates. This defines a *stationary* Markov model, the solution of which, for an arbitrary number of states, is well understood (Barlow (1965), Billinton (1983)).

Hazard Rates

An essential idea underlying the applicability of the stationary Markov model is the assumption that the *hazard rate* is constant. The *reliability* of an asset is defined as the probability that it survives (does not fail) at least up until some arbitrary time. The hazard rate is the conditional probability that the asset fails in the next instant of time given that it survived until the present. If we represent the cumulative probability distribution of the lifetime of as asset by the function F(t), such that the probability that the lifetime is less than t is F(t), then the hazard rate h(t) is given by the equation

$$h(t) = dF(t)/dt\,[1 - F(t)]^{-1} \tag{1.1}$$

where dF(t)/dt = f(t), the probability density of the lifetime of the asset.

The concept of the hazard rate is interesting for several reasons. First, the hazard rate can be empirically observed and is most often expressed as a so-called *bathtub curve.* (See figure 1-1.) The nature of the hazard rate is that it tends to start out relatively large and decreases during the *burn-in* period, remains constant for an arbitrary time, during the *steady-state* period, and then increase, during the *burnout* period. Second, the burnout period reflects the effect of aging. Hence, we are motivated to consider the behavior of the hazard rate as a fundamental modeling issue for the study of aging assets. Third, the stationary Markov model applies in the steady-state period, since the hazard rate is constant, reliability is exponential (see chapter 2 for a discussion of the exponential distribution), and the reciprocal of the hazard rate is the mean time to failure. It is important to note that these specifications are generally not valid for aging distribution assets.

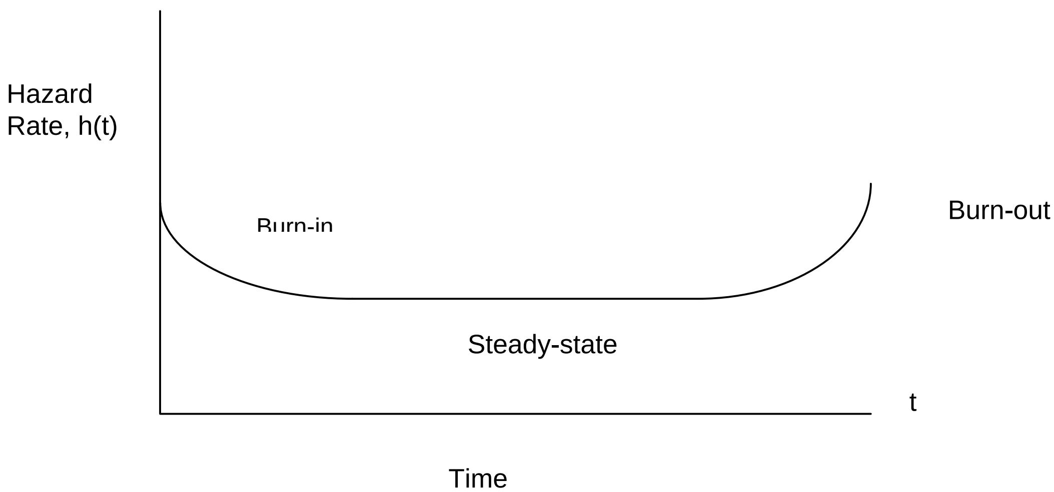


Figure 1-1
Hazard Rate "Bathtub" Curve

For aging assets, it is natural to address the effect of hazard rate directly. We use a nonstationary Markov approach that permits variable transition rates that change over time. The present report provides a method that takes observed data, converts that data into a hazard function, and then uses the hazard function to specify the nonstationary Markov transition rates. This method is described below (see chapter 5) and has been successfully applied in several case studies. The essential characteristic of the method is that it applies nonstationary hazard rates directly, as a function of age and other components of the current state of the asset. That dependence appears to be fundamental to capturing failure and repair of aging assets, and combines the increasing hazard rate and Markov modeling concepts. We discuss this point a bit further immediately below.

As noted above, the hazard rate is theoretically a conditional probability. We relax that definition somewhat and consider the hazard rate as an empirically determined failure rate, typically expressed in failures per foot of installed cable per year. Observed values are on the order of 10^{-5} failures/foot/year, with some variation. In practice, we will interpret that empirical rate as the probability that a particular foot of cable experiences a failure in a particular year.

The importance of the hazard rate is based on (at least) three ideas. First, referring to equation (1.1), if the hazard rate is empirically observed, then the life distribution can be found by solving the differential equation. We note that

$$F(t) = 1 - \exp\left(\int_0^t h(x)\,dx\right). \qquad (1.2)$$

If the hazard rate were constant, which is the behavior during the steady-state period, between burn-in and burnout, as presented in figure 1-1, then (1.2) specifies an exponential distribution. Thus, the hazard rate determines the life distribution. We discuss various life distributions and hazard rates in chapter 2.

Second, the hazard rate summarizes the data that is directly available at a utility. The typical data is a direct observation of the number of failures of underground cable in a year. At some utilities, failure data is organized by both type of cable and age of cable. The most direct way to summarize that data is to fit a hazard rate to it. We discuss methods for fitting data and estimating hazard rates in chapter 3.

Third, the hazard rate can be used to estimate the nonstationary transition probabilities that can be used as parameters in a Markov model of failure of underground cable. Because hazard rates are conditional probabilities, if we assume that non-overlapping time periods are independent, it is straightforward to express the probability that no failure occurs in the period t+1 to t+n years as the product of n terms $(1-h(t+1))(1-h(t+2))\ldots(1-h(t+n))$, where $h(x)$ is the hazard rate in year x. Thus, the hazard rate completely determines the nonstationary Markov model parameters. We discuss this further in chapter 5.

Description of the Report

This report contains four additional chapters. The next chapter, chapter 2, describes the hazard functions that we consider in the report. We discuss appropriateness of each function. In chapter 3, we present methods that can be used to estimate the parameters of the hazard functions using data that might be available at a utility. In chapter 4, we discuss how hazard functions can be selected. In chapter 5, we show how the hazard function can be used to forecast failures.

2 CABLE FAILURE MODELS

We distinguish between failure models that are based on statistical life distributions—exponential, normal, lognornal, Weibull, extreme value—and failure models that are purely empirical—the bathtub curve of figure 1-1 and the piecewise linear hazard rate. The statistical life distributions typically respond to a particular phenomenon or failure mechanism. We begin with those distributions.

Exponential

The exponential distribution on time to failure is described by the density function $f(t) = \lambda\exp(-\lambda t)$, the cumulative distribution $F(t) = 1 - \exp(-\lambda t)$; and the hazard rate $h(t) = \lambda$. There is one parameter in the exponential model, λ, which is the reciprocal of the average lifetime of the equipment modeled. This parameter, like all parameters in the other hazard functions, must be specified using either data or expert judgment.

The physical process that is modeled by the exponential can be thought of as the arrival of a peak stress or overload that the equipment cannot support. When the peak stress arrives, the equipment fails. The arrival process is a Poisson process, such that the probability that k peak stresses occur in the period (0, t) is given by the Poisson probability

$$p(k) = \exp(-\lambda t)\,(\lambda t)^k / k!\ , \quad k = 0, 1, 2, \ldots \tag{2.1}$$

where λ is the constant rate of occurrence of peak stresses.

Thus, the equipment will not fail in the interval (0, t) if no peak stresses arrive, which occurs with probability $p(0) = \exp(-\lambda t)$. Hence, $F(t)$, which is the probability that the equipment fails before t, is $1 - p(0)$ which is equal to $1 - \exp(-\lambda t)$, as noted above.

Normal

The normal distribution is described by the density function $f(t) = (2\pi)^{-1/2}\,(1/\sigma)\exp(-(t-\mu)^2/ 2\sigma^2)$, which is valid for $-4 < t < 4$. The graph of this density function is the familiar so-called bell-shaped curve. The distribution has two parameters, the mean time to failure μ and the variance of the time to failure σ^2. The hazard rate is a monotonically increasing convex function. See figure 2-1.

The normal hazard is used to represent the failure of equipment when the number of shocks required to cause failure is greater than one. This is an extension of the exponential distribution. (The mathematical details of this assumption are beyond the scope of this report. Briefly, the main idea here is that another distribution, the gamma distribution, is used to represent the failure time of equipment that fails after receiving r shocks; the gamma is approximated by the normal; the shocks arrive following a Poisson distribution with parameter λ; and the mean and variance of the approximating normal are given by the gamma distribution parameters $\mu = r/\lambda$, $\sigma^2 = r/\lambda^2$.)

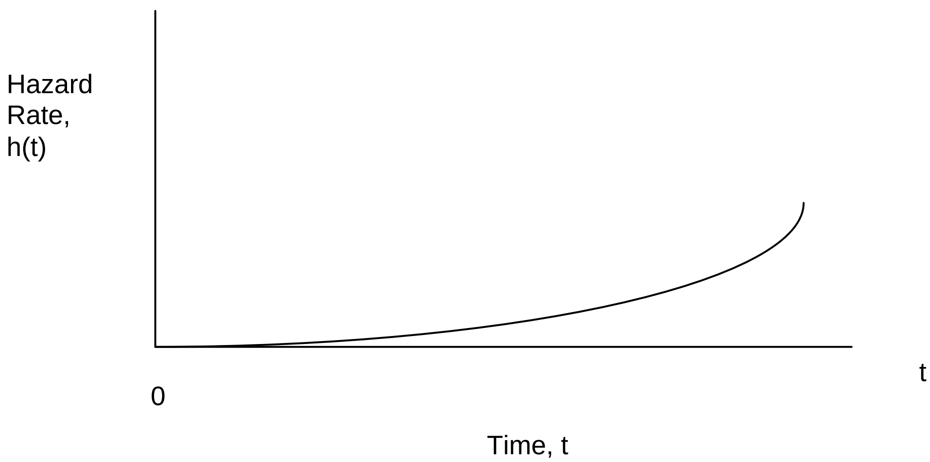


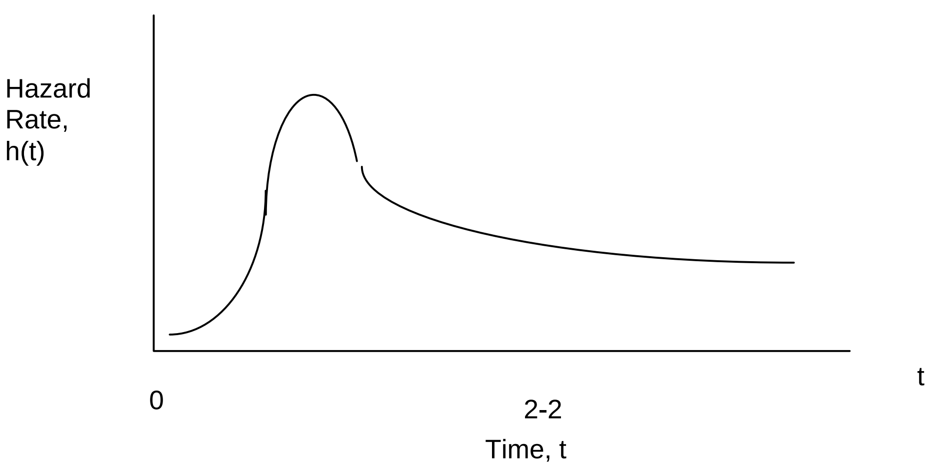
Figure 2-1
Normal Hazard Rate

Lognormal

The lognormal distribution governs a variable the logarithm of which is normally distributed. The distribution is skewed to the right and has a density function that is specified by two parameters, μ, σ, where $f(t) = (2\pi)^{-1/2} (1/\sigma t) \exp(-(\ln(t)-\mu)^2 / 2\sigma^2)$. The mean time to failure is $\mu_{\ln} = \exp(\mu + \sigma^2/2)$ and the variance of the time to failure is $\sigma^2_{\ln} = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$.

The usefulness of the lognormal distribution arises from a form of the Central Limit Theorem that states that the product of n independent random variables is lognormally distributed (for large n). The failure process modeled is multiplicative, such that the effect of succeeding shocks on the equipment is proportional to the level of the effect of all preceding shocks. Another application of the lognormal is with respect to repair times, rather than failure times. This is because the hazard function is not monotone, and possesses an interior maximum. The lognormal hazard is shown in figure 2-2.

Figure 2-2
Lognormal Hazard Rate



Weibull

The Weibull distribution is given by the density function $f(t) = (\alpha/\beta)\,(t/\beta)^{\alpha-1}\exp[-(t/\beta)^{\alpha}]$, valid for $t \geq 0$. The cumulative is $F(t) = 1 - \exp[-(t/\beta)^{\alpha}]$. The hazard rate is $h(t) = (\alpha/\beta)\,(t/\beta)^{\alpha-1}$. The parameters α and β are known as the shape parameter and the scaling parameter respectively. In log-log coordinates, the hazard function is a straight line. Thus, the hazard rate is increasing if $\alpha > 1$, constant if $\alpha = 1$ (and the Weibull becomes an exponential distribution), and decreasing if $\alpha < 1$. See figure 2-3.

The phenomenon modeled by the Weibull is associated with the theory of extreme values. The Weibull is the distribution of the minimum value of a collection of independent observations from a gamma distribution. If a system consists of a collection of components, each of which has a lifetime governed by a gamma distribution (which may be approximated by a normal), and if the system fails when any of its components fail, then the time to failure of the system is the minimum time to failure of any of the components. The Weibull is also robust in that the components may each have somewhat different gamma (normal) parameters, and the minimum is still distributed by the Weibull.

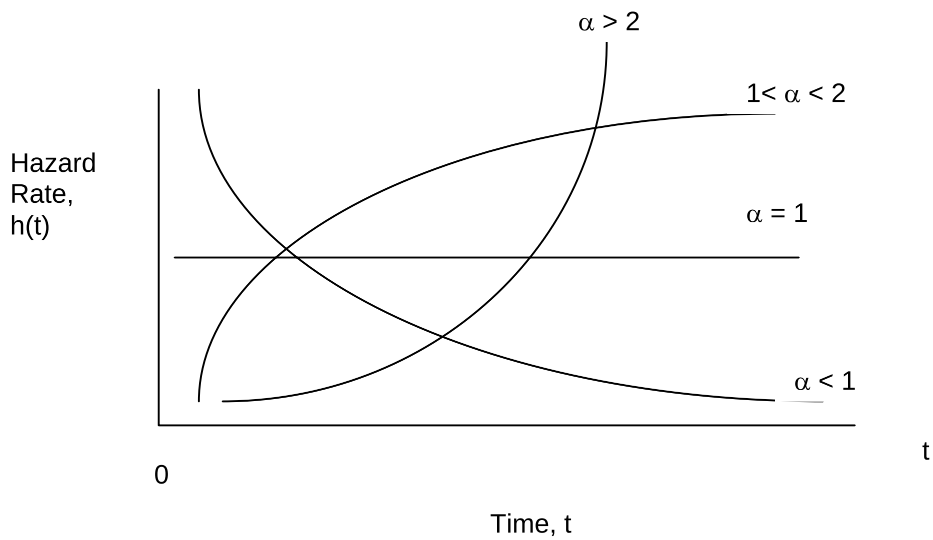


Figure 2-3
Weibull Hazard Rate

Extreme Value Distributions

The Weibull is an example of an extreme value distribution. It is also called a type III asymptotic distribution of minimum values. Also useful are other extreme value distributions, including the type I asymptotic distributions of the maximum and the minimum values. We shall not pursue these distributions in this report. The Weibull appears to provide sufficient modeling capability for capturing the distribution of the minimum. The classic reference for discussion of extreme value distributions is Gumbel (1958). Abernethy (1996) provides a good reference for understanding Weibull distributions.

The behavior of the extreme value hazard functions is shown in figure 2-4. The hazard rate of the asymptotic distribution of the minimum value grows exponentially. The hazard rate of the asymptotic distribution of the maximum value approaches a constant as t approaches infinity.

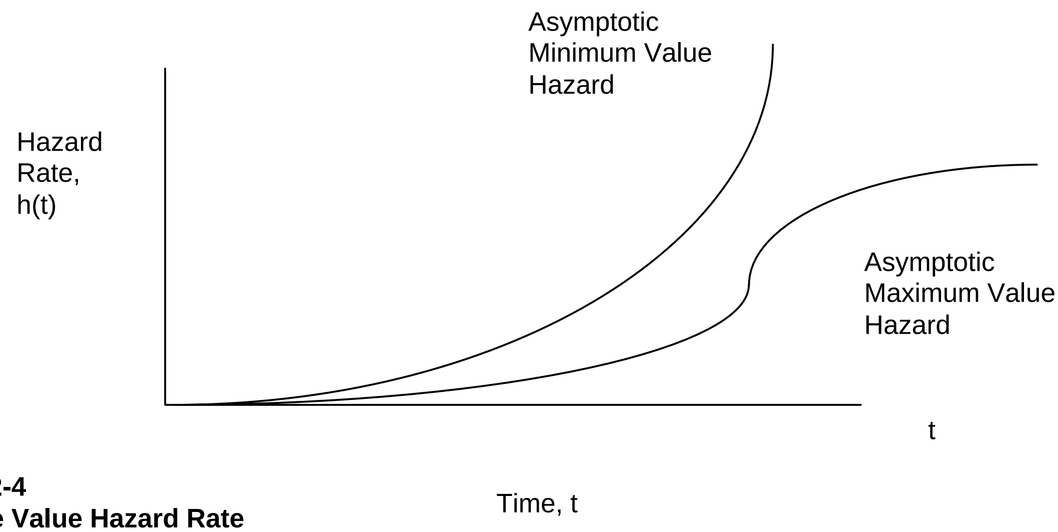


Figure 2-4
Extreme Value Hazard Rate

Empirical Hazard Functions

The empirical hazard functions are not based on any underlying life distribution. That is, instead of deriving the hazard function h(t) from the life distribution F(t) using equation (1.1), the hazard function can be used to find F(t) using equation (1.2). Therefore, the hazard function may be estimated directly from available data.

Bathtub Curve

The bathtub curve, shown in figure 1-1, is an empirically observed representation of failure for most equipment. The choices for the burn-in and burnout periods are arbitrary. Clearly, the normal, Weibull, or extreme distributions can provide models for the burnout period. Exponential burnout, corresponding to the type I asymptotic distribution of the minimum as shown in figure 2-3, is a popular form. For underground cable repair/replace studies, we have not found it important to represent the burn-in period. Therefore, we tend not to use the complete bathtub curve model. However, the burnout behavior of the bathtub curve is generally applicable. It is only a matter of choosing the rate of burnout in order to apply the model.

Piecewise Linear Model

A simple approximation to any of the hazard functions described above is the piecewise linear function shown in figure 2-5. The hazard rate is constant at the steady-state rate h_{ss} until the onset of the burnout period, which begins at age T. After the onset of burnout, the hazard function grows linearly with slope m. Thus, for t > T, $h(t) = h_{ss} + m(t-T)$. Therefore, the function is specified by three parameters, h_{ss}, T, and m. We have applied the piecewise linear model in several previous studies.

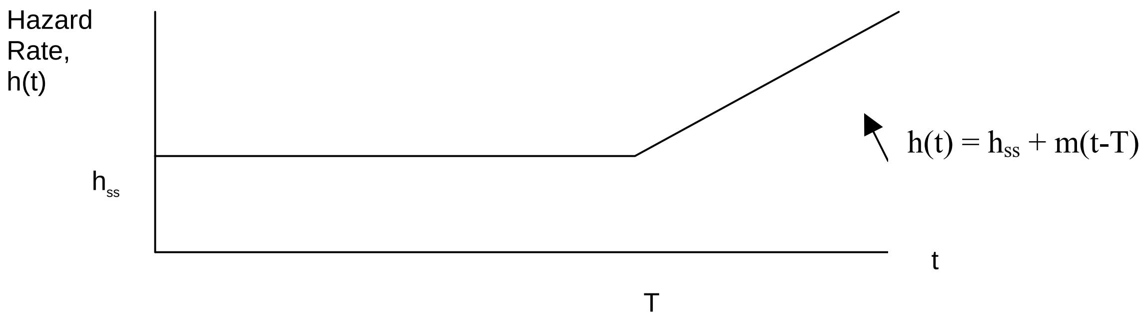


Figure 2-5
Piecewise Linear Hazard Function

Discussion

There are two fundamental ways to interpret the hazard function, as noted above. It can be thought of as a logical consequence of an underlying life distribution, or it can be thought of as the determinant of an empirical life distribution. The mathematical analysis of underground cable does not change depending on which of these interpretations is adopted, but the empirical analysis of utility data can be very different. The difference arises in the role of judgment in the analysis of the data.

The usual procedure for fitting a hazard function derived from a life distribution is to estimate the best values of the parameters by minimizing the sum of the squared errors in the estimate. Thus, a single set of parameter estimates is used for the entire range of data. However, in practice, because of a general paucity of data, it often turns out that the data are not scattered sufficiently over the range of lifetimes to provide a balanced weighting of both the steady-state period and the burnout period. The consequence of this is that the parameter estimates are more responsive to the denser region of the data, typically the earlier or shorter lifetimes, than to the sparser or longer lifetimes. What happens then is that the fit is better where the data is denser. And because the hazard function is completely determined by the parameters, the burnout period may not be accurately represented by the data. In other words, one is stuck with whatever the parameters predict for burnout, but the parameters are determined by the steady-state behavior. This is a weakness that can only be overcome by judgment, whereby the analyst adjusts the hazard function to fit better the actual data, overriding the estimated values of the parameters. We have done exactly this in our application of the Weibull to some observed data.

This difficulty tends not to arise in empirically fitted bathtub and piecewise linear hazards. In the piecewise linear case, the analyst typically assigns a value to T, the onset of burnout, and lets the data drive the best estimates of h_{ss} and m, the steady-state hazard and the slope of the burnout period. As we shall see, some form of judgment is required in all cases when the data is extremely sparse.

3 FAILURE MODEL PARAMETER ESTIMATION PROCEDURES

Introduction

In this chapter, we solve the estimation problem: given a collection of data describing the failure of underground cable, estimate the hazard function. An example of a data set provided by a utility is the following.

Year Installed	Year Failed	Number of Failures	Feet Installed	Age (t)	Hazard Rate h(t)
1997	1998	1	33208	1	3.01132E-05
1989	1997	1	54674	8	1.82902E-05
1978	1997	1	4160	19	0.000240385
1978	1999	1	4160	21	0.000240385
1978	2000	1	4160	22	0.000240385
1985	2000	1	13329	15	7.50244E-05
1975	1998	1	3416	23	0.00029274
1990	1997	3	7477	7	0.00040123
1990	1999	1	7477	9	0.000133743
1991	1998	1	33020	7	3.02847E-05
1991	1999	1	33020	8	3.02847E-05
1994	1999	1	5375	5	0.000186047
1994	1996	1	26891	2	3.71872E-05
1994	1997	1	26891	3	3.71872E-05
1995	1996	1	56311	1	1.77585E-05
1995	1997	3	56311	2	5.32756E-05
1995	1997	1	56311	2	1.77585E-05
1995	1999	1	56311	4	1.77585E-05
1996	1997	2	17712	1	0.000112918
1996	1998	1	29117	2	3.43442E-05
1996	1999	1	29117	3	3.43442E-05
1997	1998	1	34240	1	2.92056E-05
1997	1999	2	34240	2	5.84112E-05
1997	2000	1	33208	3	3.01132E-05

Table 3-1. Failure Data

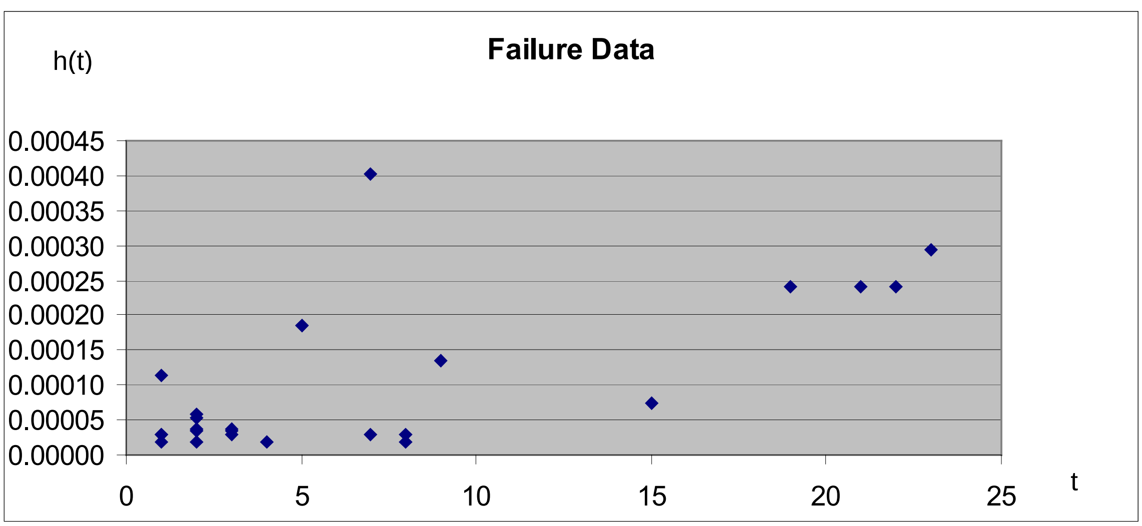
This table lists the times to failure of an inventory of underground cable. The year the cable was installed is in the first column. The second column is the year in which a failure was observed. The difference between these two times is the age of the cable when it failed. That is shown in column five. The number of failures observed for that cable is listed in column three. The number of feet installed is listed in column four. The ratio of the number of failures observed (column three) to the number of feet installed (column four) is the hazard rate, expressed in failures/foot/year as a function of age. The hazard rate is shown in column six. The last two columns of table 3-1 comprise the data required for estimating the hazard function. For example, the first row of the table corresponds to the year 1998. In that year, 1 failure occurred in cable that was 1 year old (the cable was installed in 1997). There were 33,208 feet installed in 1997. Therefore, the hazard rate was 1 failure /33208 feet of cable installed = 3.01132 x 10^{-5}, as noted in column six.

In order to create table 3-1, what must be known is the year the cable was installed, the number of feet of cable installed in that year, and the number of failures of that vintage cable in any year. It is important to note that the data in table 3-1 is not separated by vintage. By combining all vintages of cable into one data set, we make the underlying assumption that vintage does not affect hazard. If that assumption is not true, then the data should be collected separately by vintage. For example, if it is believed that cable installed in the 1970s differs in behavior from cable installed after 1979, and all cable installed after 1979 behaves in the same way, then the data should be separated into two groups. For such data, two hazard functions would be separately estimated. An interesting question would be to ask whether the functions found in that way are really different. This question could be answered using statistical analysis. In the limit, data could be kept separately for each year cable is installed. Regardless of how many different data sets there may be, the hazard functions can be estimated in the same way.

Estimating the Hazard Function

A reasonable first step towards estimating the hazard function is to plot the data in table 3-1. This is shown in figure 3-1.

Figure 3-1 Observed Failure Data



It may be reasonable to discern a burnout period that begins at approximately fifteen years in the data shown in figure 3-1. Reasonable choices for hazard functions appear to be normal, Weibull, and piecewise linear. We discuss fitting these hazard functions to this data set.

Piecewise Linear Hazard Function

There are several ways to fit a piecewise linear hazard function to hazard data. The most straightforward is to use each data point to define an estimated value of the hazard as a function of the parameters (h_{ss}, m, T) such that for each time t, the estimated value is

$$h_{est}(t;\, h_{ss},\, m,\, T) = h_{ss} + \max\{0,\, m(t\text{-}T)\} \tag{3.1}$$

(where the second term is zero if $t \leq T$ and m(t-T) if $T > t$). Now, define the error, e, at age t, as the difference between the observed hazard h(t) given by the data and the estimated value, or

$$e(t,\, h(t);\, h_{ss},\, m,\, T) = h(t) - h_{est}(t;\, h_{ss},\, m,\, T). \tag{3.2}$$

The parameters (h_{ss}, m, T) are selected to minimize the sum of the squared errors over the entire data set, or

$$\min Q(h_{ss},\, m,\, T) = \Sigma_t\, e^2((t,\, h(t);\, h_{ss},\, m,\, T). \tag{3.3}$$

It is technically somewhat easier to solve the problem (3.3) if T is specified by expert judgment. We illustrate the solution for the case in which T is selected to be 15. The parameters of the piecewise linear hazard are $h_{ss} = 7.4825 \times 10^{-5}$ and $m = 2.7329 \times 10^{-5}$. See figure 3-2.

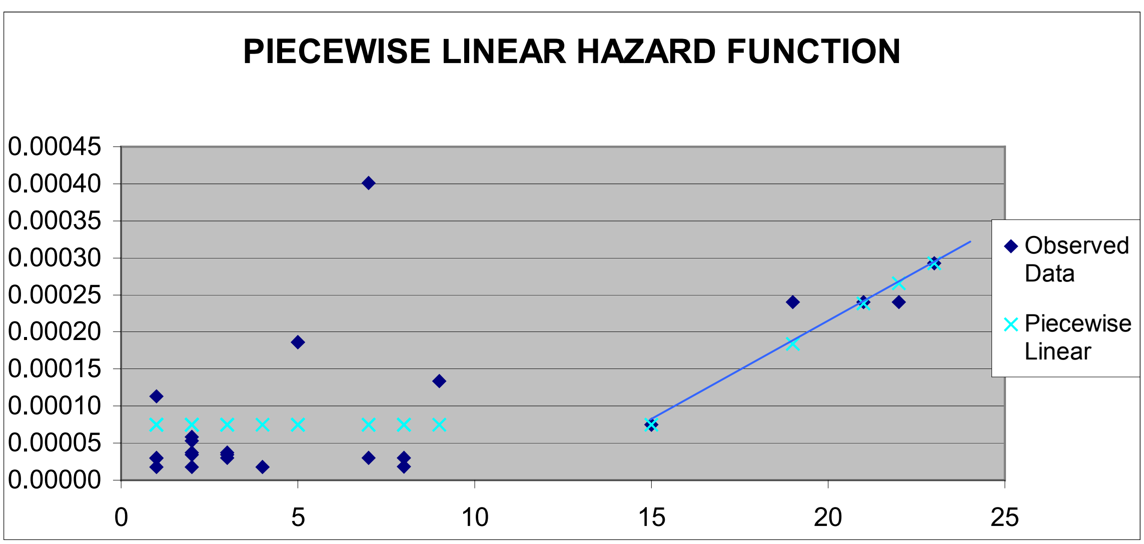


Figure 3-2 Piecewise Linear Hazard Function Based on Observed Data

The consequences of this estimation method for predicting failures and specifying policy will be made apparent below and in the companion report *Cable Reliability Management Strategies.*

An alternate approach to fitting a piecewise linear hazard function is useful in the case in which only the aggregate number of failures is known. Thus, instead of the data in table 3-1, what is known is (a) the total number of failures in a given year for the entire inventory (i.e., failures are not known by age of failed cable) and (b) the age distribution of the inventory (i.e., the number of feet installed by age).

Then, the three parameters of the piecewise linear hazard can be estimated by following five steps.

1. Record the inventory (feet) of cable by age, v(t).

2. Record the total number of failures in year y, f(y).

3. Use expert judgment to estimate the onset of burnout, T.

4. Use expert judgment to estimate the slope hazard function in the burnout period, m.

One approach we have found useful for this estimation process is to consider how long it takes for the failure rate to double from the onset of burnout. Let this doubling time be represented by d. Then the slope is given by $m = h_{ss}/d$. This follows from the fact that

$$h(T+d) = 2h_{ss} = h_{ss} + m([T+d] - T) = h_{ss} + md. \tag{3.4}$$

5. Equate the observed number of failures, f(y), to the number predicted by the hazard rate h(t), by setting

$$f(y) = \Sigma_t\, h(t)v(t), \tag{3.5}$$

where the right hand side of the equation is the expected sum of failures over the entire inventory. Because of the results of the first four steps, the hazard function is specified by only one parameter, h_{ss}. Therefore, equation (3.5) provides a solution for $h_{ss,}$

$$h_{ss} = f(y)/\ [\Sigma_{t \le T}\, v(t) + \Sigma_{t > T}\, v(t)(t-T)/d\ ] \tag{3.6}$$

Example. In this simple example, we assume there are only three vintages of cable, 50,000 feet that is 5 years old, 75,000 feet that is 14 years old, and 100,000 feet that is 21 years old. In the notation we have been using, v(5) = 50,000; v(14) = 75,000; and v(21) = 100,000. This is the solution to step 1. We also assume that there were 4 failures in year 2002. Thus, f(2002) = 4. This is the solution to step 2. For step 3, we assume that burnout begins at T = 20 years. For step 4, we assume that the hazard rate doubles in four years. Thus, $m=h_{ss}/4$. Therefore, the hazard function is $h(t) = h_{ss}$, $t \le 20$ and $h(t) = h_{ss}(1+ [t-20]/4)$, $t > 20$. Hence, the hazard function depends on one unknown parameter, h_{ss}. In step 5, we have the equation f(2002) = 4 = h(5)v(5) + h(14)v(14) + h(21)v(21) = $50{,}000h_{ss} + 75{,}000h_{ss} + 100{,}000(1.25h_{ss})$. Thus, $h_{ss} = 4/250{,}000 = 1.6 \times 10^{-5}$ failures/foot/year.

There are other ways to fit a piecewise linear hazard, depending on what data is available, but we have found these two approaches particularly useful. Using data augmented by judgment appears to be a reasonable way to proceed.

Weibull Hazard Function

Using a method similar to that for fitting the piecewise linear function to the data in table 3-1, it is possible to select the two parameters of the Weibull, α and β, to minimize the sum of the squared errors in an expression analogous to (3.3), where the error is defined by

$$e(t, h(t); \alpha, \beta) = h(t) - (\alpha/\beta)\,(t/\beta)^{\alpha-1}. \qquad (3.7)$$

The second term in the expression is the estimated hazard rate using the Weibull with the to-be-selected parameters α and β. The error defined by 3.7 is nonlinear in the parameters α and β. Therefore, it is common practice to take logarithms of the hazard and define the logarithmic error (not the logarithm of the error as defined in (3.7))

$$e_l(t, h(t); \alpha, \beta) = \ln h(t) - [(\alpha\text{-}1)\ln t + \delta] \qquad (3.8)$$

where

$$\delta = \ln \alpha - \alpha \ln \beta. \qquad (3.9)$$

The importance of this transformation, obtained by taking the logarithm of the data and the logarithm of the Weibull hazard function, is that the logarithmic error e_l is linear in the unknowns α and δ. Once α and δ are known, then β is found using (3.9). Selecting α and δ to minimize the sum of the squared logarithmic errors

$$\min Q_l(\alpha, \beta) = \Sigma_t\, e_l{}^2((t, h(t); \alpha, \beta) \qquad (3.10)$$

is a well-understood problem. It is referred to as linear regression or least-squares fit. Commercially available software packages including Excel have easy-to-use least-squares subroutines. (The Excel command is LINEST.)

We illustrate the results of the least-squares fit to the logarithmic error in figure 3-3. The parameters of the Weibull are $\alpha = 1.60093$, $\delta = -10.55765$, and $\beta = 981.082$.

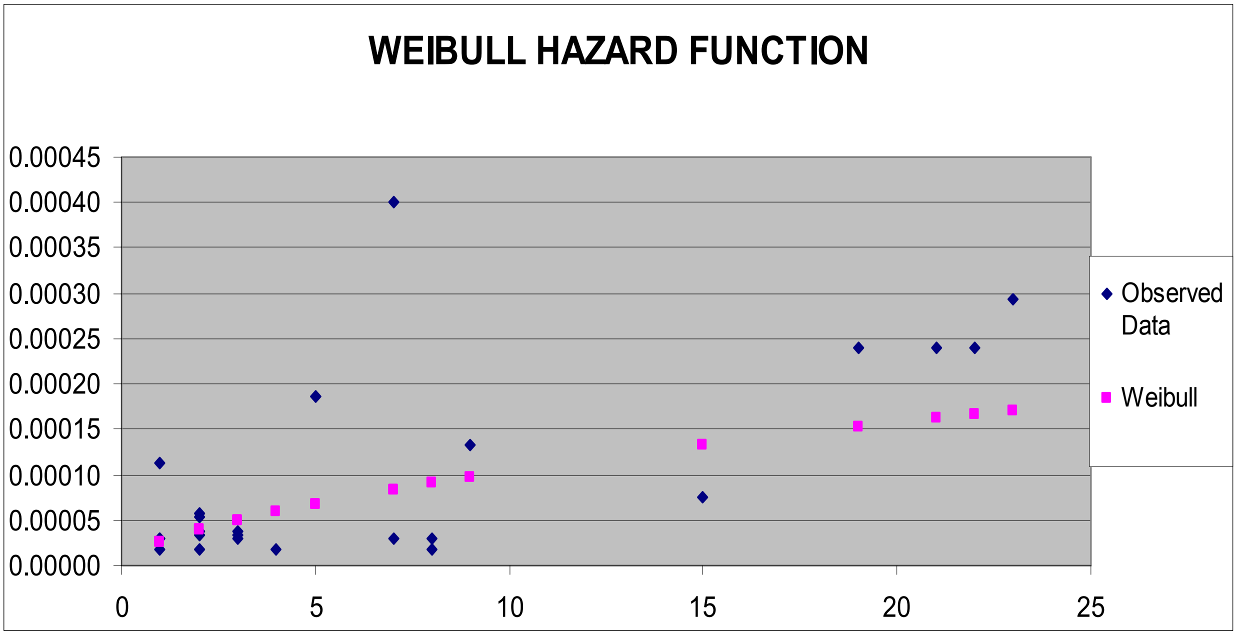


Figure 3-3 Weibull Hazard Function Based on Observed Data

If only summary failure data are available, as discussed in the piecewise linear case, so that we know the total number of failures in year y, f(y), and the total inventory by age {v(t): t=1,2,3,…}, then it is possible to estimate a Weibull hazard, but only after one of the parameters has been selected by expert judgment. This is because there is only one equation, equation (3.5) above, but two unknowns, the parameters α and β. In this case, the application of expert judgment is essential to determining the hazard function.

The Importance of Expert Judgment

The Weibull fit obtained in figure 3-3 illustrates the point made above about the density of the data driving the parameter specification. It is clear that for the larger number of data points for the shorter lifetimes, the fitted function attempts to balance the errors, while for the fewer data points at the longer lifetimes, the fitted function is uniformly lower. The uniformity of the fitting error—values of e_l (equation (3.8)) are always positive for longer lifetimes, compared with both positive and negative errors for shorter lifetimes—suggests that the fitting method could introduce bias. This is another situation that suggests the importance of judgment.

In this case, it is straightforward to test various settings of the parameters α and β. The most convenient way to do this testing is graphically. In the present situation, the more important region to represent accurately is the burnout period because that is what typically determines the optimal repair/replace policy for underground cable inventory as that inventory ages. Therefore, we developed a manual fit that appears to track the data more accurately than the fit obtained by minimizing the sum of squares of the logarithmic error. The manually fit hazard function does not agree with the data very well in the earlier period. The setting of the parameters that we found that seemed most appropriate overall, while emphasizing the agreement with the data in the burnout period, are $\alpha = 4$ and $\beta = 110$. We hasten to say that there can be any number of settings that appear to be reasonable. See figure 3-4.

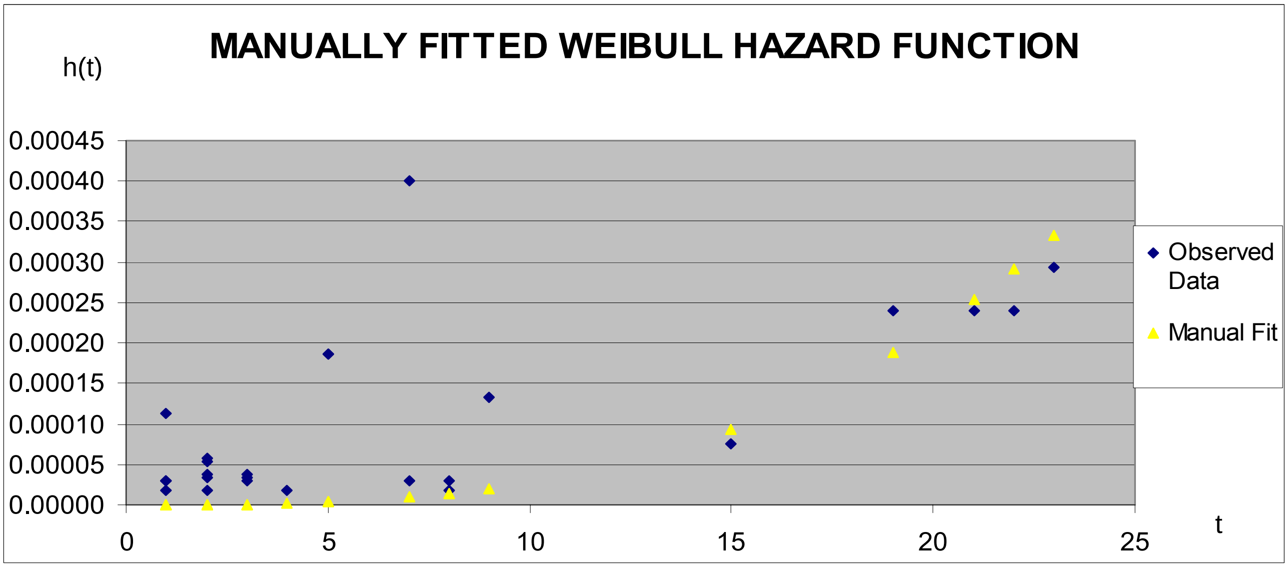


Figure 3-4 Weibull Hazard Function Based on Observed Data Fit Manually

This is one example of how expert judgment can be brought into play in the hazard function estimation process. Another instance of the importance of expert judgment arises when we use the estimated hazard function to predict actual failures. In this aspect of the work, it is essential that the hazard rates at longer lifetimes that are extrapolated from the fitted data provide a reasonable forecast of behavior. We discuss below how such considerations constrain the parameter settings of the hazard functions and how expert judgment is again required to specify the parameters even when sufficient data are present.

It is also important to recognize that one of the reasons that this problem arises is that there is somewhat less flexibility in a two-parameter analytic function such as a Weibull than there is in an empirical function like a bathtub curve or a piecewise linear hazard. In the former, the steady-state or earlier behavior is tightly linked to the burnout behavior. In the latter, the various regions are almost completely decoupled. In particular, the slope of the burnout is generally completely independent of the steady state hazard for a piecewise linear function or a bathtub curve, but that slope is completely determined by the parameter settings of a Weibull, which are often selected with great dependence on the earlier data, as we have seen in the example above.

Finally, if even summary data is unavailable, then expert judgment appears to be the only way that a hazard function can be specified.

Lognormal Hazard Function

It is possible to fit the data in table 3-1 to a lognormal hazard function. However, inspection of figures 2-2, the lognormal hazard, and figure 3-1, the observed data, reveals that the curve-fitting process would attempt to find a lognormal fit such that the data would correspond to the rising part of the hazard function only, or the left-most portion of the curve in figure 2-2. Therefore, this model is not interesting and is indeed misleading because there is no reason to believe that the data would exhibit non-monotonicity as time increased. Further, even if we restricted attention to the left-most part of the lognormal hazard, no particular advantage would be gained because the Weibull would be able to capture the nonlinear burnout period equally well. We therefore did not consider the lognormal as an appropriate model for underground cable.

Other Models, Including the Bathtub Curve

Although there are many possible life distributions, and therefore many possible hazard functions, the Weibull hazard is perhaps the most popular. There seems to be no compelling reason to adopt the exponential hazard (no burnout), the normal hazard (nonlinear burnout, as in the Weibull), the lognormal, or any other model. The underlying life distribution of the normal hazard indicates a concentration of lifetimes about the mean life. We have no reason to believe that this is appropriate for underground cable. The lognormal, as we noted above, seems inappropriate for cables. Therefore we propose to adopt either the Weibull hazard, the piecewise linear hazard, or the bathtub curve (figure 1-1). The main advantage of the bathtub curve is that it combines the nonlinear burnout period of the Weibull with the decoupled burnout period of the piecewise linear.

Fitting a bathtub curve is an exercise in combining expert judgment with data analysis. The steady-state hazard, as in the piecewise linear case, is simply the horizontal function $h(t) = hss$, for $0 < t < T$. The burnout period ($t > T$) can be represented by any nonlinear monotonic function. Therefore, Weibull burnout can be applied to a bathtub curve, and the data during the burnout period can be used to estimate the Weibull parameter α, which applies only over the burnout period. Therefore, estimating a bathtub curve requires splitting the data into two parts, the steady-state part and the burnout. We ignore burn-in, which appears not to apply to cable. Thus, one sets T using expert judgment. Based on that specification, the data prior to T can be used to estimate hss by a simple averaging process. Beyond T, expert judgment is required to specify the shape of the burnout hazard rate. Once that shape is specified, then the parameters are selected to minimize the sum of squared errors over the burnout period, in a manner virtually identical to the examples above. If Weibull burnout is assumed, then equations (3.8)-(3.10) apply to the data in the burnout period. Other functional forms will require other error formulations, but the ideas involved in parameter estimation are identical.

4 SELECTING A HAZARD FUNCTION

The question posed in this chapter is how to select a hazard function.

In the previous chapters, we identified several hazard functions and presented methods whereby data and expert judgment can be used to estimate the parameters of a hazard function. We ask, in this chapter, whether there is a general method that can be applied to selecting a hazard function. We will also propose a default hazard function that can be used if no better information is available.

Data and Judgment

There are several dimensions to the discussion about the relative role of data and judgment in selecting and fitting a hazard function. It is reasonable to select a hazard function that captures the behavior of the data that is available. But it is also reasonable to apply expert judgment in cases in which the data yields results that do not appear reasonable.

For example, we discussed above the relative merits of the results of the Weibull estimation using only the data compared with the results of the estimation that is based on judgment and the recognition that it is important to represent the burnout period with reasonable accuracy.

The conclusion is that it is impossible to eliminate the need for judgment. But the challenge is to apply judgment in a way consistent with engineering practice. This is a theme in many of the models that we have developed for EPRI, including the project prioritization methodology (P^2), the aging assets methodology, the load forecasting methodology (LoadDynamics), and the fuel inventory model (UFIM). In each case, the role of judgment is to augment the information contained in the available data to make a statement that more accurately represents the beliefs of the decision-makers.

Relationship to Policy Models

This research is part of a larger effort to understand and manage underground cable inventories. The hazard function is an essential input to a model that determines the least cost policy for repairing and replacing underground cable inventories. (This work is described in the companion report *Cable Reliability Management Strategies: Research Status Report,* EPRI, Palo Alto, CA: 2003 {1002257}.) Therefore, the hazard function should be estimated in such a way that it can provide useful input to the model it supports.

For example, using the data in Table 3-1, we were able to fit a Weibull hazard, with parameters $\alpha = 1.60093$ and $\beta = 981.082$, as noted above. But, as we noted above, we recognized that this data-driven estimate did not capture the burnout period very well. To overcome the bias induced by the parameter estimation process, we developed a manual fit with parameters $\alpha = 4$ and $\beta = 110$. However, when we used this hazard function as input to the policy model, we discovered that for cable that is more than thirty-five years old, the Weibull hazard forecasts what appear to be an unreasonable amount of failures. The forecast is deemed unreasonable based on expert judgments from our previous case studies. Therefore, we adjusted again, this time because of the requirements of the policy model, and used, for purposes of analysis and illustration, the parameters $\alpha = 5$ and $\beta = 95$. These parameters provided a fit that responded to both the data and the judgment that future failure rates were constrained. This result is shown in figure 4-1. Compare with figure 3-4.

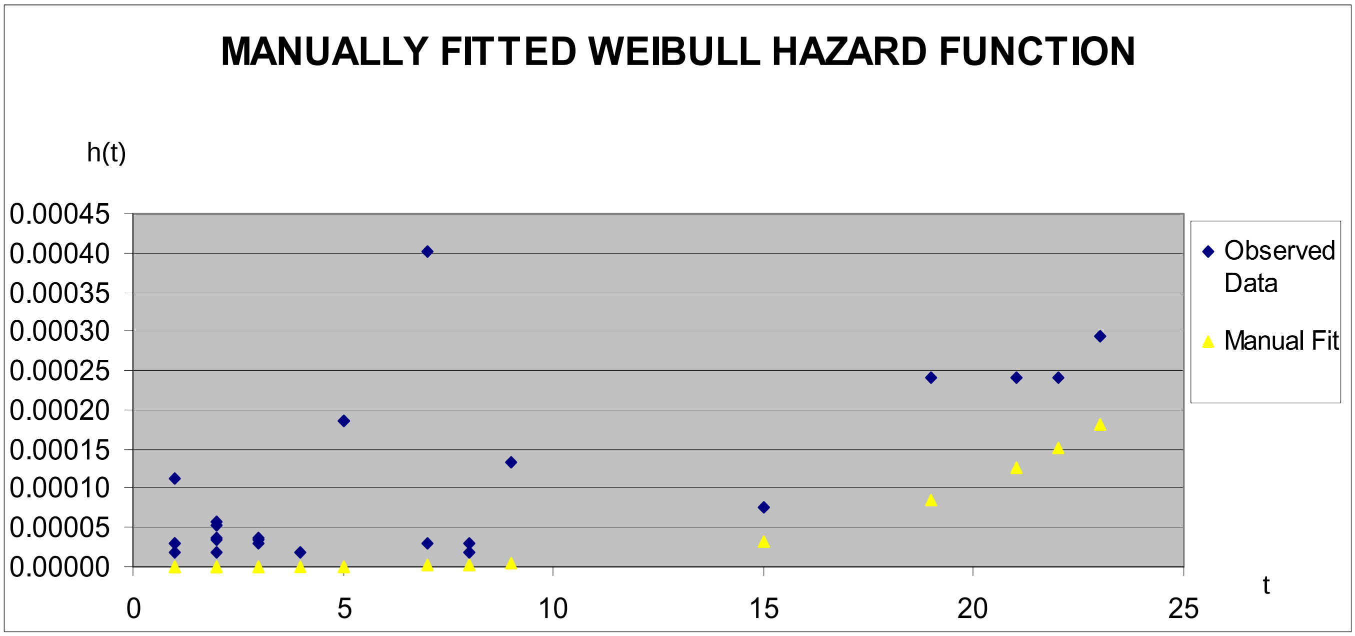


Figure 4-1 Weibull Hazard Function Based on Observed Data Fit Manually In Order to Provide Reasonable Forecasts of Future Failure Rates

Default Hazard Function

We have neither enough experience nor enough data to suggest a definitive default hazard function that will not change as additional utility data is observed. Nevertheless, we can offer the following guidance.

1. The simplest hazard function to estimate without data for support is the piecewise linear function.

2. Expert judgment can be used to specify the three parameters, h_{ss}, T, and m, the steady-state hazard rate, the time of onset of burnout, and the slope of the burnout period.

3. The steady-state hazard rate, h_{ss}, is on the order of 10^{-5} failures/year/foot.

4. The onset of burnout occurs approximately at T = 25 years.

5. The slope of the burnout period, m, is approximately 0.2, which indicates that the failure rate doubles in five years.

6. These parameter estimates can be thought of as our best estimates of industry-wide averages—although they are based on very small samples at present—and therefore can be systematically varied by a given utility, when that utility believes that the average does not describe the unique conditions experienced.

Conclusion

It is not possible to give a general method for selecting a hazard function, even if we restrict our attention to underground cables. We believe that the most robust hazard functions are Weibull, piecewise linear, and the bathtub curve (with a power function or something similar to describe burnout). We suggest that failure data should be plotted first, providing something like figure 3-1. The hazard function should be selected so that it captures the behavior of the data. The parameters of the hazard function should be specified, first, by least-squares error analysis of the data. But that specification should be reviewed by experts in order to be sure that the estimation process is not biased during the burnout period and that the hazard function provides reasonable failure rates in the future.

5 FORECASTING CABLE FAILURES

The hazard function provides two capabilities that are essential to further policy analysis. First, we may use the hazard function as an input to a policy specification model. This requires a conversion of the information in the hazard function into a form that the model can apply. The model is a stochastic dynamic program (Hillier (1986)) and the hazard function must be used to specify transition rates. Second, the hazard function can be used to forecast cable system failures. This forecast can be used to validate the hazard function over time.

Converting Hazard Functions into Failure Transition Rates

We noted in Chapter 4 that one of the criteria for selecting a hazard function is that the hazard function provides reasonable behavior in a policy analysis model. The policy analysis model requires that the probability that a section of cable experiences a failure in each of the analysis years be specified. This probability is known as a transition rate, because it is the probability that a section of cable makes a transition from the operating state to the failed state in a given year. The method discussed here is also presented in the companion report *Cable Reliability Management Strategies: Research Status Report,* EPRI, Palo Alto, CA: 2003 (1002257).

The hazard function provides the value h(t), the failure rate (failures/year) per foot of installed cable that is t years old. We interpret this number as the arrival rate of failures that are governed by the Poisson distribution. This is the same assumption that is made in deriving the exponential hazard function. The Poisson probability distribution is given in equation (2.1), above. Thus, we identify h(t) with the Poisson parameter λ. What we are assuming is that the arrival process of the failures is time-varying Poisson. Further, h(t) is measured per foot of cable. We are interested in the failure behavior of a segment of cable that may be as much at 0.3 miles long. We make the simplifying assumption, typically made when applying a Poisson process, that each foot of cable fails independently. Hence the failure rate per cable segment is lh(t), where l is the length of a segment. Therefore, the parameter of the Poisson in equation (2.1) becomes $\lambda = lh(t)$. It is worth noting that the probability that no failure occurs in the interval (0, 1), or a period that is one year long, in a cable segment of length l that is age t is

$$p(0; t) = e^{-\lambda} = e^{-lh(t)}. \tag{5.1}$$

(The zero in the argument of p corresponds to k = 0 in the notation of equation (2.1), which means that no event occurred.) Now, in the policy model we have developed for the study of underground cable, we wish to analyze failures over arbitrarily longer time periods. In particular, in this study, we wish to address policies covering periods of 5-year intervals. We make the simplifying assumption that the behavior in each year is independent of the behavior in any other year. Therefore, the probability that no failure occurs in a five-year period in a cable segment of length l that is age t at the beginning of the five-year period is

$$p\,(0; t, 5) = e^{-\lambda} = e^{-l[h(t) + h(t+1) + h(t+2) + h(t+3) + h(t+4)]}. \tag{5.2}$$

(The third argument, 5, in the expression p(0; t, 5) denotes the number of years in an analysis period.) Equation (5.2) expresses the probability as a product of the probabilities given in equation (5.1). The formulation (5.2) captures the time-varying effect of the hazard function h(t). The probability of a transition to the failed state at any time in the five-year period is then the complement of p(0; t, 5), or letting q(1; t, 5) denote the probability of at least one failure in a five year period for cable that is t years old at the beginning of the period,

$$q(1; t, 5) = 1- p(0; t, 5) = 1- e^{-l[h(t) + h(t+1) + h(t+2) + h(t+3) + h(t+4)]}. \quad (5.3)$$

We also consider the effect of multiple past failures on present hazard rate. The simplest approach to treating this phenomenon is to multiply the hazard rate by a rate, say m, which we select to capture the increase in hazard rate because of a single prior failure. This may be estimated if data is available, or assessed based on expert judgment. Now if the number of past failures is greater than one, say f, and this number is known, then the failure rate multiplier is m^f. (The current default value for m is 2, based on a small sample of data and judgment.) This factor is included in the specification of q(1; t, 5) in the obvious way:

$$q(1; t, 5) = 1- p(0; t, 5) = 1- e^{-l\ (m\hat{\ }f)\ [h(t) + h(t+1) + h(t+2) + h(t+3) + h(t+4)]} \quad (5.4)$$

This completes the specification of the segment transition probabilities for use in the policy analysis model. Clearly, if h(t) grows rapidly, then q(1; t, 5) rapidly approaches 1 (which means virtually certain failure in any year) for large t. Such behavior may not accurately reflect reality. The remedy, as noted above, is to apply expert judgment to the specification of the hazard function as illustrated in the sequence of figures 3-3, 3-4, and 4-1 above.

Forecasting Failures

Given a hazard function and a specification of the inventory of installed cable, it is possible to forecast the number of failures of that cable inventory.

The forecast is accomplished by treating the hazard function as the growth rates in a discrete time dynamic system (Luenberger (1979)). The equations that govern the cable population dynamics can be specified as follows.

Let $v_f(t)$ = feet of cable that failed f times by age t. This is the cable inventory. If the failure history is not known, we set f to zero.

Let h(t) = hazard rate per foot of cable at age t.

Let r = repeat failure rate. This is a parameter of the model that can be determined by data or specified by expert judgment. It is the rate at which multiple failures can occur in a failed segment in a single year. If it is known that, say, 100 feet failed in a year, then 100r of them will experience a second failure in that year. We had used the estimate r = 0.10 in a previous study.

Let m =hazard multiplier, as noted above. If cable has previously failed f times, then the multiplier is mf. We have used the estimate m=2 in a previous study. (Note that this use of the letter m is not the same as that used to indicate the slope of a piecewise linear hazard. We trust that this will not cause confusion.)

For simplicity, we assume that we do not count past failures beyond three. Therefore, we need expressions for $\{v_f(t+1):\ f = 0, 1, 2, 3\}$, based on the values for t.

It is simplest to express the equations in vector-matrix form. Let A(t) be the 4 x 4 transition matrix

$$A(t) = \begin{bmatrix} 1-h(t) & 0 & 0 & 0 \\ (1-r)h(t) & 1-mh(t) & 0 & 0 \\ rh(t) & (1-r)mh(t) & 1-m^2h(t) & 0 \\ 0 & rmh(t) & m^2h(t) & 1 \end{bmatrix} \qquad (5.5)$$

The matrix A(t) can be interpreted as the matrix of transition probabilities from the number of failures associated with the column at time t to the number of failures associated with the row at time t. (It may be worth noting that the expression 1 – h(t), the entry in the first row, first column, is the probability that a foot of cable that had not failed up to t will not fail by t+1; and this is the first-order Taylor expansion of the probability discussed above, $e^{-h(t)}$, in equation (5.1)). Thus, for example, (1-r)mh(t) is the entry in column 2, row 3. It is the fraction of cable that has experienced one failure by time t (column 2) that will experience two failures by time t + 1 (row 3). Then, letting $\underline{v}(k)$ denote the (column) vector $[v_o(t)\ \ v_1(t)\ \ v_2(t)\ \ v_3(t)\ ]^T$, we have the linear system

$$\underline{v}(t+1) = A(t)\,\underline{v}(t) \qquad (5.6)$$

that determines the population dynamics.

For example, the first equation is

$$v_o(t+1) = (1-h(t))v_o(t) \qquad (5.6a)$$

which indicates that the number of feet that have not failed by age t+1 is a fraction of those that have not failed by age t. The remainder, $h(t)v_o(t)$, will have failed either once $((1-r)h(t)v_o(t))$ or twice $(rh(t)v_o(t))$ by year t+1. These two terms are found in the second and third equations, as defined by the second and third rows of the matrix A(t), which specify $v_1(t+1)$ and $v_2(t+1)$, respectively. The fourth equation is

$$v_3(t+1) = rmh(t)\,v_1(t) + m^2h(t)\,v_2(t) + v_3(t) \qquad (5.6d)$$

which indicates that the number of feet that have failed three times by age t+1 is the number of feet that had failed once by age t and that failed twice more in the next year plus the number of feet that had failed twice by age t and failed (at least) once in the next year plus the number of feet that had failed three times by age t. Recall that for simplicity we are not counting past failures beyond three.

It is straightforward to show that the number of failures in any year is

$$N(t) = h(t)\,(1+r)\,(v_o(t) + mv_1(t) + m^2v_2(t) + m^3v_3(t)) \tag{5.7}$$

Example. In order to illustrate the forecasting process as transparently as possible, we take a very simple example. As in the example in chapter 3, suppose that there are only three vintages of cable, 50,000 feet that is 5 years old, 75,000 feet that is 14 years old, and 100,000 feet that is 21 years old. We assume that the failure history is not known. In the notation used above, $v_o(5) = 50{,}000$; $v_o(14) = 75{,}000$; and $v_o(21) = 100{,}000$. All other inventory variables are zero.

We will use the piecewise linear hazard function determined in that example, $h(t) = h_{ss}$, $t < 20$ and $h(t) = h_{ss}(1+ [t-20]/4)$, $t \geq 20$, with $h_{ss} = 4/250{,}000 = 1.6 \times 10^{-5}$ failures/foot/year.

Further, we assume that the segment length l = 500 feet, the failure rate multiplier m = 2, and the repeat failure rate r = 0.1.

Using these data, the transition matrix A(t), as given in equation (5.5), for cable age $t \leq 20$, the steady-state period, is

$$\mathbf{A(t) =} \begin{bmatrix} 0.999984 & 0 & 0 & 0 \\ 0.0000144 & 0.999968 & 0 & 0 \\ 0.0000016 & 0.0000288 & 0.999936 & 0 \\ 0 & 0.0000032 & 0.000064 & 1 \end{bmatrix}$$

Table 5-1. Steady-state transition matrix

Applying this transition matrix for five years, beginning with the initial vector of cable of age 5, the vector $[v_o(t)\ \ v_1(t)\ \ v_2(t)\ \ v_3(t)\ ]^T = [50{,}000\quad 0\quad 0\quad 0\ ]$, the forecast for the number of feet that experience additional failures is the sequence of vectors (shown as column vectors) as follows. Bear in mind that we are allocating the feet of cable among the various classes that are defined by failure history. Each element in the table below is a number of feet. All columns sum to 50,000 feet, the initial inventory. The column corresponding to forecast year 0 is the initial vector of feet of cable age 5.

	Forecast Year					
	0	1	2	3	4	5
v0	50000	49999.20	49998.40	49997.60	49996.80	49996.00
v1	0	0.72	1.44	2.16	2.88	3.60
v2	0	0.08	0.16	0.24	0.32	0.40
v3	0	0.00E+00	7.42E-06	2.23E-05	4.45E-05	7.42E-05

Table 5-2. Forecasting the behavior of 5 year old cable over a five year horizon.

This table indicates, for example, that by the fifth forecast year, 3.60 feet of cable that was five years old at the beginning of the forecasting horizon have experienced one failure. 0.40 feet experienced two failures, and a negligible amount of feet experienced three or more failures.

Applying equation (5.7), the number of failures forecast in each year is 0.88. (There is a negligible amount of inventory with multiple failures. The number of failures, then, is approximately 50,000 feet x $1.6\text{x}10^{-5}$ failures/year/foot x 1.1 = 0.88 failures for each year.)

Similarly, applying this transition matrix for five years, beginning with the initial vector of cable of age 14, the vector $[v_o(t)\ \ v_1(t)\ \ v_2(t)\ \ v_3(t)\ ]T = [75{,}000\quad 0\quad 0\quad 0\ ]$, the forecast for the number of feet that experience additional failures is the sequence of vectors (shown as column vectors) as follows. The number of failures in each year behaves similarly to that experienced by the 50,000 feet of five-year old cable. That number is 1.32 failures in each year.

	Forecast Year					
	0	1	2	3	4	5
v0	75000	74998.80	74997.60	74996.40	74995.20	74994.00
v1	0	1.08	2.16	3.24	4.32	5.40
v2	0	0.12	0.24	0.36	0.48	0.60
v3	0	0.00E+00	1.11E-05	3.34E-05	6.68E-05	1.11E-04

Table 5-3. Forecasting the behavior of 14 year old cable over a five year horizon.

For the cable that was 21 years old at the beginning of the forecasting horizon, the matrix in table 5-1 does not apply, because this cable is experiencing burnout. For example, the transition matrix in for the first forecast year, based on the fact that the cable is 21 years old, is

$$A(21) = \begin{bmatrix} 0.99998 & 0 & 0 & 0 \\ 0.000018 & 0.99996 & 0 & 0 \\ 0.000002 & 0.000036 & 0.99992 & 0 \\ 0 & 0.000004 & 0.00008 & 1 \end{bmatrix}$$

Table 5-4. Burnout period transition matrix (age 21).

Applying the sequence of transition matrices beginning with the one in table 5-4 for year 21, beginning with the initial vector of cable of age 21, the vector $[v_o(t) \;\; v_1(t) \;\; v_2(t) \;\; v_3(t)\,]^T = [100{,}000 \quad 0 \quad 0 \quad 0\,]$, the forecast for the number of feet that experience additional failures is the sequence of vectors (shown as column vectors) as follows.

	Forecast Year					
	0	1	2	3	4	5
v0	100000	99998.00	99995.60	99992.80	99989.60	99986.00
v1	0	1.80	3.96	6.48	9.36	12.60
v2	0	0.20	0.44	0.72	1.04	1.40
v3	0	0.00E+00	2.78E-05	9.93E-05	2.33E-04	4.50E-04

Table 5-5. Forecasting the behavior of 21 year old cable over a five year horizon.

Applying equation (5.7) to these forecasts, the number of failures in each year of the forecasting period increases, because the hazard rate is increasing during burnout, and is given by the values 2.2, 2.64, 3.08, 3.52, and 3.96. The dominant effect here is the growth rate in the hazard. The approximate number of failures in the first forecast year is $100{,}000 \times 1.25 \times 1.6 \times 10^{-5} \times 1.1 = 2.2$ for the 21 year old cable. In the next forecast year, the cable is 22 years old, two years into the burnout period, so the number of failures is approximately$100{,}000 \times 1.50 \times 1.6 \times 10^{-5} \times 1.1 = 2.64$. The remaining numbers are found in a similar manner.

The final forecast of the number of failures for the entire inventory in this example is the sequence {4.4, 4.84, 5.28, 5.72, 6.16}. (Note that 4.4 = 0.88 + 1.32 + 2.20.) Clearly, forecasts like these can be used to test the hazard function over time. This concludes the example.

6 CONCLUSIONS

In this report, we have discussed various models of failure of underground cable. The failure behavior of cable is captured by a hazard function, which is defined as the probability that a section of cable will experience a failure in the next instant of time given that it has survived until that time. The hazard function can also be interpreted as the failure rate of underground cable inventory, expressed as failures/year/foot of cable installed. This interpretation permits us to estimate the hazard function from observed data. We presented several straightforward methods of estimating the parameters of hazard functions. The methods combine least-squares error analysis and expert judgment.

It is possible to suggest a default hazard function for use if no data is available. We suggest that a piecewise linear hazard function is a reasonable place to begin. We offer some guidance with respect to setting the three parameters of that function based on our limited experience with underground cable.

We indicated how the hazard function can be used to supply transition rates for a policy analysis model. We showed how the hazard function can be used to forecast failures over an arbitrary forecasting horizon for any inventory of cable. These two applications can serve as a means of testing whether the hazard function is a reasonable representation of the failure phenomenon.

A
References

Abernethy, Robert B. (1996). *The New Weibull Handbook* (Second Edition), Dr. Robert B. Abernethy Publisher, North Palm Beach, Florida

Barlow, R.E., Proschan, F. (1965). *Mathematical Theory of Reliability.* Wiley: New York.

Billinton, R., Allan, R.N. (1983). *Reliability Evaluation of Engineering Systems: Concepts and Techniques.* Plenum: New York.

Gumbel, E.J. (1958). *Statistics of Extremes*. Columbia University Press, NY.

Hillier, F.S., Lieberman, G.J. (1986). *Introduction to Operations Research.* Holden-Day, Oakland, CA.

Luenberger, D.G. (1979) *Dynamic Systems*. Addison-Wesley: New York.

About EPRI

EPRI creates science and technology solutions for the global energy and energy services industry. U.S. electric utilities established the Electric Power Research Institute in 1973 as a nonprofit research consortium for the benefit of utility members, their customers, and society. Now known simply as EPRI, the company provides a wide range of innovative products and services to more than 1000 energy-related organizations in 40 countries. EPRI's multidisciplinary team of scientists and engineers draws on a worldwide network of technical and business expertise to help solve today's toughest energy and environmental problems.

EPRI. Electrify the World

1002256

Printed on recycled paper in the United States of America

EPRI • 3412 Hillview Avenue, Palo Alto, California 94304 • PO Box 10412, Palo Alto, California 94303 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com