Substation Transformer Asset Management and Testing Methodology

1012505

Substation Transformer Asset Management and Testing Methodology

1012505

Technical Update, December 2006

EPRI Project Manager
J. Bloom

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ 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 Technical Update report. A Technical Update 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 EPRI Orders and Conferences, 1355 Willow Way, Suite 278, Concord, CA 94520. Toll-free number: 800.313.3774, press 2, or internally x5379; voice: 925.609.9169; fax: 925.609.1310.

CITATIONS

This report was prepared by

VMN Group LLC
200 Cervantes Road
Redwood City, CA 94062

Principal Investigators
Charles D. Feinstein, Ph.D.
Peter A. Morris, Ph.D.

This document describes research sponsored by the Electric Power Research Institute (EPRI).

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

Substation Transformer Asset Management and Testing Methodology, EPRI, Palo Alto, CA: 2006. 1012505.

PRODUCT DESCRIPTION

Increasing pressure from both customers and regulators to maintain and enhance service reliability, while at the same time controlling costs, has put many utilities' power delivery businesses in a classic dilemma of conflicting objectives. For that reason, asset management has become an increasingly important aspect of corporate business strategies. A significant focus of EPRI's asset management research in recent years has been to develop a rational basis for selecting repair or replacement options for specific classes of equipment by balancing the risks of equipment failure against the costs of continued maintenance or capital replacement.

This report discusses methods for making decisions about aging assets in electric transmission and distribution systems. This report focuses on substation transformers. EPRI plans to continue this research in the future in order to improve the specificity, precision, and scope of these guidelines and to extend them to other classes of assets.

Background

For many utilities, particularly those focusing on the power delivery business, their substation transformer inventory represents a substantial asset. Managing this inventory entails significant costs and directly affects the reliability of electric service. Thus utilities need cost-effective strategies for maintaining their transformer assets.

Objective

The purpose of this report is to provide analytical tools for developing economical strategies for managing transformers.

Approach

EPRI has developed a decision framework that enables utilities to generate business cases for asset management policies. This framework takes a life-cycle costing approach that enables corporate financial managers and regulators to assess the multi-year financial impacts of maintaining specific classes of power delivery infrastructure assets, such as substation transformers.

The analytical tools presented in this report share a basic framework for decision-making that specifies the evolution of the condition of the asset population over time, the various decision alternatives that are available, including testing, and the basic data needed to support the decision model.

The decision framework represents the dynamic process of transformer deterioration mathematically, using a set of equations that provide a forecast of future deterioration. The dynamic equations describe the evolution of transformer condition probabilistically – given the current transformer state, there will be a probability distribution of states in which the transformer might be observed the next time it is inspected. The model of transformer condition developed in this report depends upon what data is available for decision-making.

Results

This report describes a set of analytical models for dealing with the complexities of transformer management decisions. This framework is valuable for several reasons:

- It systematically and logically captures the interrelationships among the factors that influence the cost effectiveness of transformer management policies.
- It identifies the key information needed for making good decisions.
- It provides an objective way to choose among decision alternatives.
- It enables calculating the cost and performance consequences of transformer management policies.
- It evaluates the benefits of different test protocols.

EPRI Perspective

EPRI has been developing methods for power delivery asset management since 1992. Methodology, software, and equipment failure data have been under development for several years to aid companies in developing economic asset management strategies – strategies that meet customer needs for reliability and power quality at least cost. The objective of this project is to use the information, tools, and experience that has been developed in EPRI's asset management research to deliver general guidelines and strategies for managing specific equipment categories, in this case substation transformers, based on the knowledge assembled in previous years' work.

This report, which describes the methodology for asset management of substation transformers, has two companion reports:

EPRI report 1012503, *Equipment Failure Model and Data for Substation Transformers* (2006), provides the default settings of parameters that users will find in the present implementation of the methodology.

EPRI report 1012504, *Transformer Population Model with Testing* (2006), provides prototype software implementing the current version of the dynamic decision model formulated and discussed in this report. The Users Guide included in that report provides a series of screen shots to give the reader sufficient guidance to use the model.

Keywords

Asset management
Distribution systems
Substations
Transformers
Aging assets
Equipment repair and replacement
Reliability

CONTENTS

1 INTRODUCTION 1-1
Basic framework for decision-making 1-2
Objective 1-2
Dynamic Behavior: Deterioration and Hazard Rates 1-2
Unobservable Conditions and Stressors 1-5
Empirical and Judgmental Data 1-6
Diagnostic Tests 1-7
Decision Alternatives and Inventory Policies 1-8
Specifying a Case for Analysis 1-9
Contents of this report 1-11

2 REPRESENTING TRANSFORMER CONDITION AND MANAGEMENT DECISIONS 2-1
State Variables for Transformers 2-1
State Dynamics 2-3
Inspection and testing 2-5
Stressors 2-6
Policy Specification 2-6

3 MODELING TRANSFORMER DYNAMICS AND MANAGEMENT POLICIES 3-1
Optimal Policy Model Based on State Variable Dynamics 3-1
Condition Dynamics and Testing 3-2
Dynamic Decision Modeling 3-4
Costs and Objective Function 3-5
Data Required to Support the Optimal Policy Model 3-6

4 ESTIMATING HAZARD FUNCTIONS 4-1
Overview of the Estimation Method 4-1
Transformer Failure Models 4-3

Exponential 4-3
Normal 4-4
Lognormal 4-5
Weibull 4-5
Extreme Value Distributions 4-6
Empirical Hazard Functions 4-7
Bathtub Curve 4-7
Piecewise Linear Model 4-8
Discussion 4-8
Estimating the Hazard Function Based on Transformer Age 4-9
Estimating a Piecewise Linear Hazard Function 4-10
Estimating a Weibull Hazard Function 4-11
The Role of Expert Judgment 4-11
Conclusion 4-12

5 HAZARD FUNCTION ESTIMATES: AN EXAMPLE 5-1
Estimating the Hazard Function 5-2
Piecewise Linear Hazard Function 5-3
Weibull Hazard Function 5-4

6 CONCLUSIONS 6-1

LIST OF FIGURES

Figure 1-1 Estimated Hazard Functions, Based on Observed Failure Data....1-5
Figure 4-1 Observed Hazard Data for an Inventory of Transformers....4-2
Figure 4-2 Hazard Rate "Bathtub" Curve....4-3
Figure 4-3 Normal Hazard Rate....4-4
Figure 4-4 Lognormal Hazard Rate....4-5
Figure 4-5 Weibull Hazard Rate....4-6
Figure 4-6 Extreme Value Hazard Rate....4-7
Figure 4-7 Piecewise Linear Hazard Function....4-8
Figure 5-1 Observed Failure Data....5-2
Figure 5-2 Piecewise Linear Hazard Function for *T*=27....5-4
Figure 5-3 Weibull Hazard Function Based on Observed Data, Fit with Logarithmic Transformation....5-6

LIST OF TABLES

Table 2-1 Policy Specification as a Function of Observable State (Age) and Transformer Test Outcome 2-8
Table 5-1 Failure Data 5-1
Table 5-2 Hazard Observations During Burnout, for T=30 5-3
Table 5-3 The Effect of Varying Onset of Burnout (T) Upon Parameter Estimates and Goodness-of-fit (J) 5-4
Table 5-4 Logarithmic Transformation of Hazard Data 5-5

1 INTRODUCTION

Utilities today face the twin challenges of satisfying increasingly high standards for reliability and service quality while at the same time reducing costs and improving earnings. To meet the challenges, utilities are adopting *asset management* as their framework for allocating capital and operation/maintenance budgets. Simply stated, asset management consists of decision-making processes that have the goal of deriving the most value from utility assets within the available budget.

For many utilities, particularly those focusing on the power delivery business, their substation transformer inventory represents a substantial asset. Managing this inventory entails significant costs and directly affects the reliability of electric service, since transformer failures can lead to lengthy and widespread outages. Thus utilities need cost-effective strategies for maintaining their power delivery assets. The purpose of this report is to provide analytical tools, information, and guidance on developing economical strategies for managing an inventory of transformers.

Among the asset management decisions that arise in dealing with transformers are the following:

- Determining what level of maintenance to apply to a transformer over its lifetime;
- Determining how frequently to inspect or test transformers and what types of inspections or tests to perform (e.g., visual checks for oil leaks, assessment of general condition [corrosion of external metal work, tanks, pipe work, radiators, etc.], annual infrared surveys, dissolved gas oil sampling, and others);
- Determining which transformers to repair and what repairs to apply to individual transformer components, such as the radiator or the pumps or the bushings;
- Determining whether and when to overhaul a transformer, thereby rejuvenating it;
- Determining when to replace a transformer.

Of course, these issues are closely related, so that deciding, for instance, to use a certain level of maintenance can affect the frequency of inspections, tests, repairs, overhauls and replacements. Thus, balancing the many factors that influence the overall cost and reliability impacts of transformer management strategies forms a complex, multi-faceted decision problem.

This report describes a set of analytical models for dealing with these complexities. This framework is valuable for several reasons:

- It systematically and logically captures the interrelationships among the factors that influence the cost effectiveness of transformer management policies.
- It identifies the key information needed for making good decisions.

- It provides an objective way to choose among decision alternatives.
- It enables calculating the cost and performance consequences of transformer management policies.

Basic framework for decision-making

The analytical tools presented in this report share a basic framework for decision-making that specifies the evolution of the condition of the transformer inventory over time, the various decision alternatives that are available, and the basic data needed to support the decision model. This section summarizes that framework. (Previous EPRI reports also discuss the basic methodology, including 1002086 [*Guidelines for Intelligent Asset Replacement, Vol.1*], 1002087 [*Guidelines for Intelligent Asset Replacement, Vol.2, Wood Poles*], 1002257 [*Transformer Reliability Management Strategies*], 1010740 [*Guidelines for Intelligent Asset Replacement, Vol.3, Underground Distribution: Underground Cable*] and 1012500 [*Guidelines for Intelligent Asset Replacement, Vol.4, Wood Poles (Expanded)*].) The following are the elements of the decision framework:

- Objective
- State definitions and dynamics
- Test definitions and accuracy
- Decision alternatives and policies
- Data requirements

Objective

The objective in this decision framework is to minimize the lifecycle cost of maintaining the transformer inventory, subject to serviceability requirements. The lifecycle cost comprises the total of the installation, inspection, repair, overhaul, and replacement costs throughout a multi-year time horizon, all taken on a present value basis. The serviceability requirement means that transformers that fail or do not meet the minimum standards (e.g. for rated capacity) must be replaced, repaired, or overhauled.

Dynamic Behavior: Deterioration and Hazard Rates

The dynamic behavior of transformers is based on the idea of deterioration. Over time, the condition of an individual transformer may deteriorate as it experiences thermal loading, through faults, environmental damage (including salt contact or effects of weather), mechanical wear-out, improper maintenance, and other factors. The rate of deterioration depends on the transformer type, the vintage and manufacturer, the operating environment, and other factors, and it also can be influenced by some remedial maintenance activities. The decision framework utilized in this report represents the dynamic process of transformer deterioration mathematically, using a set of equations that provide a forecast of future transformer conditions. The forecast depends on the present condition of the transformer. In the present methodology, although it is recognized that a single transformer can be thought of as an interacting set of components ((a) the transformer, the bushings, the lightning arrestors; (b) the cooling system, the radiator, the pumps, the fans; and (c) the tap changer), a single, discrete variable is used to characterize the aggregate condition of the transformer. Future extensions of this methodology could consider these components separately.

In general, information summarizing what is known about a transformer is called its *state*. The condition of the transformer is an unobservable part of the state. The state may include other information as well, including the transformer's age or the time since the transformer was last overhauled, and the results of any tests performed on the transformer. The importance of the state variable is that it contains information about the transformer that is sufficient to forecast its future behavior.

In this report, the observable state of a transformer has two components. The first component represents its age. The actual state variable is the *Service Age*, which is the chronological age of the transformer modified by any maintenance actions taken that affect the age. For example, the effect of overhauling a transformer is to make it behave similarly to a transformer with less age that has not been overhauled. The second component of the observable state is whether or not a transformer has been overhauled. These state components have been chosen because, based on prior experience, they provide a reasonable basis for deciding among the various courses of action available for transformers. It is assumed that that these two components of the state are known at any time for all transformers.

Generally speaking, with current technology the condition of individual transformers may be monitored frequently. However, such monitoring, or testing (in the terminology used in this methodology) is not costless. The virtue of testing is that information about the unobservable condition may be discovered, and this information can suggest the appropriate action to take in order to minimize the costs of operating the transformer. Therefore, whether or not to test is a *decision*, a part of the overall asset management policy for transformers.

Expanding on this idea, the decision framework developed in this report represents the decisions regarding testing, overhaul, maintenance, and replacement as depending on the state of a transformer . The specification of a decision for each transformer state is called a *policy*, so the decision framework develops a *state-dependent policy* that *minimizes the lifecycle cost* of maintaining the transformer population.

In general, as noted above, the evolution of the transformer state cannot be predicted with certainty; that is, the deterioration of a transformer is subject to random influences. This fact implies that the dynamic equations must describe the state evolution probabilistically – given the current state, there will be a probability distribution of states in which the transformer might be observed at some future time. The random evolution of the state results from three factors:

- The transformer component degradation processes proceed randomly. Furthermore, even if the state of a transformer were known with certainty at some point in time, over time these random influences would destroy that knowledge.
- The state does not completely capture the current condition of a transformer because some aspects of the condition cannot be observed directly. For instance, the effect of past overloads may not have any apparent manifestation and testing methods for the ability of the

transformer to handle overloads may not detect any effects until the degradation is far advanced. This incomplete information results in uncertainty about the current condition that in turn creates uncertainty about the future evolution of the state. If degradation is present but not detected, the probability of failure at some future moment must include the probability that degradation at some level was present, but not detected, at the present time.

- The state does not necessarily represent all aspects of a transformer's history relevant to predicting its future evolution. That is, choosing a representation of the state is a design decision within this analytical framework. Usually, the amount of information that the state can represent is limited, for two reasons. First, the amount of information in a utility's records is limited for reasons of cost or practicality. For instance, a utility may record only the most recent inspection of a transformer and not its entire history. Second, a complex definition of the state can make the dynamic equations mathematically intractable. The more information that is encoded in the state, the higher the dimension that is needed in the dynamic equations, a phenomenon called the "curse of dimensionality." For instance, even if one could record the entire history of a transformer, computing a policy based on this information would be impossibly complicated. With limited information encoded in the state, two transformers with the same state might in fact have different histories, and differences in their future condition would appear to be random because the information needed to predict them are not available.

The deterioration of a transformer's condition is represented mathematically by a probability distribution called the *hazard function*. The value of the hazard function at a particular service age t, called the *hazard rate*, is the probability that a transformer that has survived to that age does not survive to age $t+1$. That is, the hazard rate at age t is the probability that it will fail in the next year. Figure 1-1 illustrates the general shape of the hazard functions used to model transformer deterioration.

The figure indicates the relationship between the hazard function and the observed data that the function summarizes. The hazard in Figure 1-1 is expressed as the failure rate of transformers installed as a function of the service age of the transformers. For example, the hazard rate of transformers of service age forty is approximately four percent, hence one would expect that in any year, four out of every 100 transformers that begin the year at service age twenty would fail during that year.

One way to capture the information in the data is to fit the parameters of a hypothesized hazard function to the data. The figure shows the result of fitting two hazard functions, a piecewise-linear fit, using three parameters—the steady-state hazard rate, the age at onset of burnout, and the burnout rate—and a Weibull function. In this case, both fitted models are relatively close, and they both can only approximate the actual data. Other models that are popular include the exponential hazard function, the lognormal hazard function, and the Normal hazard function. The methodology described in this report can apply any hazard function that the analyst chooses.

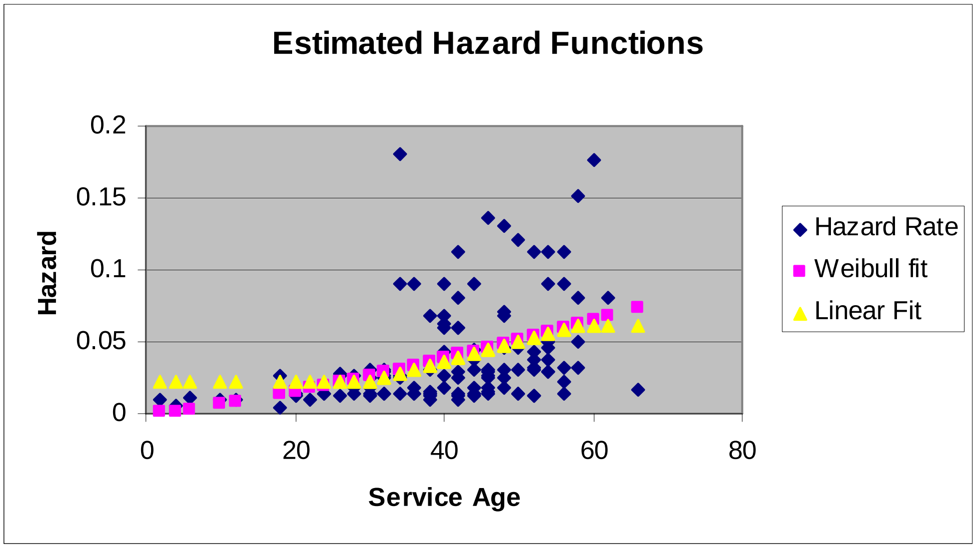

Figure 1-1
Estimated Hazard Functions, Based on Observed Failure Data

In this example, the hazard function exhibits the following plausible behavior. When a transformer is relatively new, the likelihood that it will fail is small and constant. As it ages, the transformer reaches a point at which the failure probability increases more rapidly. This increase continues for the foreseeable future.

Clearly, the hazard function depends directly on the age of the transformer, but it also can depend on other components, observable and unobservable, of the transformer's state. Chapters 4 and 5 of this report discuss how to develop good estimates of the failure behavior of transformers based on observed data.

Unobservable Conditions and Stressors

In addition to observed data, the decision framework permits the analyst to specify two other considerations that affect the failure of transformers, the dynamics of unobservable condition and the presence of stressors.

Unobservable conditions are levels of deterioration that increase the risk of failure or the cost of maintenance of the transformer that cannot be directly detected. At present, although we recognize that it is reasonable to represent the transformer as three separate subsystems ((1) the transformer, the bushings, the lightning arrestors; (2) the cooling system, the radiator, the pumps, the fans; and (3) the tap changer), the present implementation of the methodology applies a single, discrete summary condition to characterize the entire transformer. The alternatives are: {*Acceptable* (green), *Monitor* or *Watch* (blue), *Marginal* (yellow), and *Unacceptable* (red)}. In principle, the analyst may specify unobservable conditions arbitrarily, subject to the ability to specify the effect of the unobservable conditions on the hazard function and to specify the likelihood that a transformer is in any of the condition states as a function of the service age of the transformer (and any other observable attributes that may be relevant; the methodology recognizes that condition dynamics depend on the maintenance level applied to the transformer).

The main idea here is that the unobservable conditions change over time, and when they change, the hazard rates change. How the unobservable conditions affect the hazard rate can be specified in many ways. The current implementation relies on the ability of the analyst to specify hazard rate multipliers (see EPRI report 1012503).

Stressors are external influences that may increase the likelihood that a transformer will fail. The definition of a stressor is arbitrary. The present implementation of the decision framework identifies two kinds of stressors, environmental and utilization (operational) stressors. Generally, environmental stressors represent factors that are beyond the control of the utility, such as weather or temperature conditions, while utilization stressors represent factors that are potentially within its control, such as transformer loading. Detailed definitions of these stressors are given in EPRI report 1012503. The analyst must be able to specify the effect of the presence of stressors on the hazard rate. The present implementation expresses this dependency as hazard rate multipliers (see EPRI report 1012503).

Empirical and Judgmental Data

Much of the flexibility and realism of the decision framework used in this report stems from the contributions of expert engineers to its formulation. This expertise is reflected, in particular, in the definitions of the degradation states of transformers, the relationship between degraded conditions and hazard rates, the impacts of stressors on transformer degradation and on hazard rates, and the specifications of diagnostic test protocols. These factors represent parameters that must be quantified for use in the decision model. Ideally, one would like to use empirically observed data to estimate statistically the values of the parameters, such as, for instance, the hazard function illustrated in figure 1-1. However, in many cases, the data that are available are not sufficiently rich to provide a basis for good statistical estimates for many of these parameters.

There are several reasons why existing data may not be sufficient. First, because asset management is a fairly new concept, many utilities did not have sufficient reasons to collect detailed information to track equipment condition. Even when they did collect data, it was often for other purposes and therefore was not available through integrated databases that facilitated to sorts of analysis required for asset management decision-making. Many utilities are just now converting their legacy data into integrated databases and instituting systematic processes for collecting and validating such data. Second, equipment replacement decisions occur when the equipment reaches the end of its life, and as is apparent from the discussion of hazard functions, equipment behaves quite differently as it approaches that stage than when it is comparatively young. Thus, even though data collection processes are improving, there will be a considerable lag before information tracking the full lifetime of equipment becomes available. Third, gathering certain kinds of data is very costly and potentially destructive. Furthermore, for the purposes of statistical estimation, it would be necessary to do this data-gathering for an adequately sized sample of multiple transformers to give a certain level of statistical precision to the estimates. It would be difficult for most utilities to justify the expense just to provide data for estimating the parameters of a decision model. This motivates efforts at EPRI to create a database of industry-wide data as it becomes available. Finally, the dimensionality of the data increases cost and complexity of the data collection task. Dimensionality means the number of factors that influence a given parameter. For example, consider the hazard rate. It depends on at least the following factors: the transformer’s age, the number of prior overhauls, the condition of the oil and the core, the presence and degree of one or more stressors, etc. Even with a small number of distinct levels for each of these factors, it is easy to see that dozens of hazard rate estimates would be needed to properly account for all of them. Utility data collection processes are usually not set up to differentiate among all of these factors, and even if they were, the required sample sizes would make data collection extremely expensive.

Given the limitations of statistical data analysis to estimate key relationships among parameters of the decision model, use of expert judgment provides an important element of the analysis. Among the parameters estimated by experts are the differential effects on the hazard rates of the following factors:

- condition states of the transformer
- environmental and operating stressors
- level of maintenance (indirectly, through the dynamics of condition)

In addition, the following parameters have also been estimated by experts:

- the proportion of the transformer population in each condition state as a function of age and level of maintenance
- the accuracy of test protocols in revealing the true condition of the transformer.

Default values of the parameters have been provided by EPRI Solutions in EPRI report 1012503, *Equipment Failure Model and Data for Substation Transformers*, 2006.

Diagnostic Tests

Diagnostic tests can play an important role in managing substation transformer assets. Because transformer condition is not directly observable, diagnostic tests can provide some information about their condition. However, three issues limit the usefulness of tests. First, tests are generally not completely accurate; that is, for instance, a transformer that is not degraded (“green” condition) may give a test result indicating some level of degradation and vice versa. It should be noted, however, that even an inaccurate test may provide information useful for asset management decisions (see Raiffa, H., *Decision Analysis*). In principle, it should be possible to quantify the level of test accuracy (a subject of on-going research at EPRI), but at present such estimates rely on expert judgment. Second, the relationship between test outcomes and hazard rates is not well established; that is, how much more likely is a 30-year-old transformer with moderate degradation (in “blue” condition) to fail in the next year than a transformer of the same age with no degradation (in “green” condition)? Again, in principle, it should be possible to quantify this relationship (also a subject of on-going research at EPRI), but at present such estimates rely on expert judgment. Finally, testing can be costly and possibly destructive. In fact, these three issues are linked together, and the decision of whether to test or not is basically a question of whether the value of the information, even if somewhat inaccurate, outweighs the cost of obtaining it. (A more complete discussion of this issue, with particular attention to the value of information may be found in Raiffa, H., *Decision Analysis* and Schlaifer, R.O., *Analysis of Decisions Under Uncertainty*.)

The decision framework used in the report models testing using several overall principles. First, whether or not to test is a decision, a part of the overall asset management policy for transformers. That is, the value of test information depends on the actions one takes based on the test and on the cost savings attainable over acting without doing the tests. Said another way, a test has value only if its outcome would lead to making a different decision, and even then, unless the value exceeds the cost of the test, one should not use the test. Thus, testing is included among the decision alternatives considered by the framework. Second, the decision framework separates specification of the test outcome from the action taken based upon it. This principle stands in contrast to much of the current industry practice, which is to state the test outcome as a recommendation for action, such as, "replace immediately" or "retest in 3 years." Separating test outcome from recommended action permits optimizing the asset management policy as a function of test outcome, rather than prespecifying the decision. Third, test outcomes are specified in terms of the condition state they diagnose; that is, test outcomes are reported as, for instance, "the transformer is in marginal condition." This specification was chosen because it is fairly easy to understand, it applies across many kinds of tests, and it conforms to typical industry practice of rating transformer condition on a four-point qualitative scale. Alternatively, one could specify the test outcome as the specific numerical measurement produced by the test; however, it is not clear that this would improve the overall decision model, and it would make it much more complicated. One consequence of this assumption is that test accuracy can be specified as the probability that the test says the condition is X when in fact the true condition is Y, where X and Y represent condition states. Fourth, the decision framework represents test protocols rather than individual tests. A test protocol is a group of related tests applied to an individual transformer either at the same time or in succession over a relatively short period of time (say within 2 or 3 years). This representation conforms to current industry practice that takes advantage of the different kinds of information available from different tests (e.g. oil condition--DGA, power factor, infrared hotspot detection, visual inspection for mechanical degradation and leaks, etc.) It also allows representing some of the contingent test policies utilities are using, such as retesting within two years to determine how rapidly degradation is progressing. Finally, through the use of test protocols to represent groups of related tests, the decision framework assumes that no other dependence exists among tests done on the same transformer at different times; that is, the model does not remember the outcome of a previous test bank. This assumption greatly simplifies the state dynamic equations by not requiring them to represent the unobservable state. It is a reasonable assumption under two conditions: (i) test intervals are fairly long relative to the speed of degradation, or (ii) the action taken as a result of the test changes the transformer's condition so that the previous condition is no longer relevant. However, the representation of test information over time remains an area of possible extension of the decision model, which appears to be an important consideration for power system transformers.

Decision Alternatives and Inventory Policies

At any point in time, the methodology assumes that the observable state of a transformer is known, the unobservable state of the transformer is specified by a probability distribution (which is also known), and the asset manager must make a decision about the transformer. The methodology permits the choice of the following alternatives:

- *no action (do nothing)*, which does not change the state of the transformer; doing nothing means that the behavior of the transformer is governed by the hazard function given by the current state of the transformer—which means that the current state is sufficiently satisfactory so that no intervention need be taken;

- *overhaul* the transformer, which changes the state of the transformer; when a transformer is overhauled, its effective service age is reduced, which means an overhauled transformer behaves as if it were a younger (and not-yet-overhauled) transformer; therefore, its condition is changed favorably; the analyst defines both the age reduction due to overhaul and the probability distribution of the condition state at the reduced age; also, after overhaul, the maintenance level applied to the transformer can be changed;

- *test* the transformer, which means that further information about the unobservable condition is needed prior to deciding what to do with the transformer; the outcome of the test is a claim about the unobservable condition; testing is not perfect, so that the test outcome revises the probability distribution on the unobservable condition, hence the hazard rate is changed, which in turn changes the subsequent decision; the decision is then based on knowledge about the state and the test result;

- *replace* the transformer, which means that the present state of the transformer was such that the risk of failure was too great to accept (with respect to cost-effectiveness); the transformer can be replaced with a like transformer or a different transformer type; the hazard function for that replacement type must be specified if the replacement type differs from the transformer under analysis.

By convention, decisions occur at the beginning of each year of the planning period. In the present implementation, the time between decisions is not a single year. Instead, decisions occur at the beginning of a time period that lasts five years. This is done for two reasons. First, five-year intervals make the computations required for a solution tractable. Second, as a consequence of the fact that there are five years between decisions, we may treat successive tests as independent. That is, the outcome of any test does not depend on the outcome of the previous test. This has two important consequences. Not only does this make the description of the asset state simpler, because it is not necessary to remember the results of previous tests, but also it relieves the analyst of the burden of specifying the conditional behavior of a sequence of tests.

A *policy* is a complete specification of what decisions to make in each interval of the planning period (or, in the present implementation, at the beginning of each five-year interval) as a function of the state of the transformer. The methodology determines the least cost policy over the indefinite future that makes the same decision for a transformer state regardless of what time in the planning period the decision is taken. This is called a *stationary policy*. The least cost stationary policy is the solution to the asset management problem.

Specifying a Case for Analysis

In addition to specifying the hazard function, the methodology discussed in this report is based on a collection of user inputs that describe the behavior of an inventory of transformers, the economics of transformers, and the decisions that can be made with respect to the inventory of transformers. The inputs are listed and described below.

- Inventory. The number of installed transformers as a function of transformer age and overhaul status (which is binary: either the transformer has been overhauled or it has not yet been overhauled).

- Hazard Function. The hazard function for the transformer inventory must be specified. In the present implementation, the hazard function is specified as a piecewise-linear function, which is discussed below in Chapters 4 and 5. Often, data is available to help estimate the hazard function.

- Unobservable states must be defined. The present implementation has four distinct unobservable states. The effect of the unobservable state on the hazard function is given as a multiplier, which must be specified. The probability distribution on the occurrence of the state as a function of the transformer age is required.

- Stressors can be defined as needed. The present implementation has four levels of stressors: {average environment and utilization, severe environment, severe utilization, severe environment and severe utilization (both active)}. The effect of the presence of stressors on hazard rates is given as a multiplier, which must be specified.

- Tests. Specified tests report on the level of an unobservable condition. The accuracy of the test must be specified as a likelihood function. (See Chapter 4 for further discussion of test likelihoods.)

- Economic data are required.

    - Failure Cost. The cost of a failure, not including the cost of replacing any failed equipment. The failure cost measures the impact of an outage both to the utility and the customer. The failure cost typically includes the direct cost to the utility of a failure, plus any indirect costs that may arise from a failure, as well as the imputed cost to the customers served when an outage occurs. Therefore, the failure cost can vary greatly by transformer type. It is reasonable to expect that this parameter will be very influential in determining the optimal policy. The analyst will almost surely want to determine the sensitivity of the optimal policy to changes in failure cost.

    - Replacement cost, overhaul cost, testing cost (per test). These are costs associated with their respective decisions. They are each expressed as single amounts.

    - Maintenance costs. The methodology identifies three levels of maintenance: {Good, Average, Poor (run-to-failure)}. The annual cost of each, per transformer, is required.

- Discount rate. The methodology minimizes the net present value of the inventory policy.

- Decision Consequences. In general, what a decision does to the state variable must be given.
  - Replace. Provides a new transformer, service age = 0, not-yet-overhauled, unobservable condition = *acceptable* (green).
  - Overhaul. Reduces the service age of the transformer by a fixed amount, specified by the user. Therefore, the condition distribution, and hence the hazard function, changes.
  - Test. Provides further information about the unobservable states. The probability distribution on the unobservable states is revised conditional on the test outcome. The analyst specifies the test accuracy.
  - Do Nothing. No change in the state.

This completes the list of inputs.

Contents of this report

This introductory chapter discusses the motivation for developing decision models for managing transformer assets and provides an overview of the technical approach.

Chapter 2 describes the basic components of decision models for managing transformer assets: how to represent the condition of transformer and the outcome of decisions (do nothing, overhaul, test, and replace) applied to transformer.

Chapter 3 describes how to represent mathematically the dynamics of transformer degradation and the effects of decisions applied to transformers. In this Chapter, the dynamic decision model is formulated. The decision model describes how the transformer state evolves in response to various degradation factors and how decisions affect the transformer state.

Chapter 4 discusses estimating the hazard functions used in the dynamic decision model. Procedures are given for estimating piecewise linear and Weibull hazard functions.

Chapter 5 presents an example that illustrates the methods discussed in Chapter 4 that can be applied to estimate hazard rates for transformers.

Chapter 6 provides brief statement of conclusions.

This report, which describes the methodology for asset management of substation transformers, has two companion reports:

EPRI report 1012503, *Equipment Failure Model and Data for Substation Transformers* (2006), provides the default settings of parameters that users will find in the present implementation of the methodology.

EPRI report 1012504, *Transformer Population Model with Testing* (2006), provides prototype software implementing the current version of the dynamic decision model formulated and discussed in this report. The Users Guide included in that report provides a series of screen shots to give the reader sufficient guidance to use the model.

2
REPRESENTING TRANSFORMER CONDITION AND MANAGEMENT DECISIONS

This Chapter generally describes the basic components of decision models for managing substation transformer assets: how to represent the condition of transformers and the outcome of replacement, maintenance, and testing decisions applied to transformers.

State Variables for Transformers

The *state* of an asset such as a transformer is the collection of information about the asset that is sufficient to predict the asset's condition in the next time interval. Choosing the definition of states is a design decision within this analysis framework, and there may be a number of useful, alternative definitions. There are four considerations that apply to the definition of the state:

1. First, the state is descriptive. It contains sufficient information about the asset at any time to describe the relevant aspects of its condition at that time. For instance, the transformer state might include information about the presence of contaminants in the oil or about the past maintenance applied to the transformer.
2. Second, the definition of the state is not unique. The choice of state variable depends on several things, including what information is available about the asset's condition, what information is needed to predict the future evolution of the its condition, and how the evolution of condition is represented. That is, given two different mathematical models of asset condition, the same asset can have different state representations.
3. Third, the state is dynamic. The state of an asset such as a transformer changes over time, as external conditions change or as the asset responds, as it ages, to the external conditions it encounters. For instance, a transformer that does not reveal evidence of mechanical degradation in the current state has a non-zero probability of experiencing such degradation the next time it is inspected or tested.
4. Fourth, the entire state of an asset need not be directly observable. In that case, tests and other observations may permit inferring the occurrence of unobservable components of the state as a consequence of what is observed. For example, since insulation degradation is usually not directly observable, a partial discharge test might allow an inspector to infer its presence.

Some asset characteristics do not change with time, such as its location, size, manufacturer, and year installed. These constant characteristics do not need to be represented in the state; rather, a separate decision model is formulated for each subcategory. This separation reduces the complexity of the state description and the dynamic equations without severely limiting the kinds of testing/maintenance/replacement policies that can be considered. Henceforth, it is assumed that the state description and dynamic equations refer to a single type of transformer, defined by its location, capacity, voltage, and other installation conditions. If information is not available about these characteristics, then transformers of all types can be considered together in a single set of dynamic equations.

The simplest, and essential, state variable for a transformer is its *service age*. Although, as discussed below, transformer management policies depend upon other factors, especially on condition, nevertheless the transformer's age represents an important determinant of the management policy. Age serves as a surrogate for the damage that accumulates over time, and since condition can be only indirectly observed, age provides a prior expectation of condition that can be tested, as discussed below. The methodology uses service age, or *effective* age, rather than simple chronological age, in order to capture the effect of overhauling a transformer, which can be most easily represented as a reduction in the effective age of the transformer. The effective age of the transformer is the age of a transformer than has not yet been overhauled that behaves in an identical manner to the transformer in question. The service age of a transformer is essential information because the probability of failure in any year depends on the service age at the beginning of the year.

If more is known about the transformers, then the state can include further information, such as the overhaul history of the transformer. These are observable components of the state, and we restrict the observable state to the pair (service age, overhaul status). Thus, we will develop policies that will indicate what action to take if, for example, a transformer is presently age 35 years and has not yet been overhauled. The methodology does not currently capture the time since overhaul, although a future extension could capture it if it proves useful to modeling the evolution of condition after overhaul.

The unobservable component of the state represents the condition or level of deterioration that cannot be directly detected. The present implementation of the methodology applies a single, discrete summary condition to characterize the entire transformer: *Acceptable* (green), *Monitor* or *Watch* (blue), *Marginal* (yellow), and *Unacceptable* (red). This representation of condition was chosen in consultation with utility experts on transformers and is reasonably consistent with current condition indicators used in utilities.

The fundamental modeling assumption associated with the state definition is that *all assets in the same state at a given time evolve the same way in the next time interval*. This assumption does *not* mean that the future state of all transformers sharing the same initial state will be the same. The future state for any transformer is the result of degradation processes that evolve probabilistically. Hence, the current state specifies the *probability distribution* of the future state of a transformer. Thus, the fundamental assumption means that the probability distribution of the future state is identical for all transformers having the same initial state.

This fundamental assumption has implications for the way the states are defined. Characteristics of transformers that would cause differences in the future evolution of their condition must be captured in the state definition. Conversely, limiting the amount of information contained in the state (which is frequently necessary to make the dynamic equations tractable mathematically) limits the ability of the model to differentiate future condition evolution among transformers. That limitation is not necessarily a bad thing, since it may reflect the unavailability or costliness of data needed to differentiate condition states. However, it does imply that sound engineering judgment needs to be applied in formulating the model of asset condition evolution, to balance the degree of specificity with the availability of data and the tractability of the calculations.

Managing the population of transformers requires making decisions about what to do with individual transformers, such as: do nothing, test, overhaul, or replace. The state must contain sufficient information to determine the evolution of the state that results from making a decision about a particular . For instance, a state-dependent decision for a transformer might take the form "if a transformer is less than 30 years old and has not yet been overhauled and a dissolved gas analysis indicates that the transformer is in "watch" condition, then overhaul the transformer." In addition, the cost of that decision may be a function of the present state. For instance, the cost of overhauling might depend on the age of the transformer. However, such dependency has not been implemented in the present version of the methodology, in part in order to make the data requirements as simple as possible.

The definition of the state thus has fundamental implications for the economic analysis of alternative maintenance policies. The fundamental assumption implies that all transformers with the same state will be managed the same way; that is, the transformer management policy is *state-dependent*. Hence, given the dynamic equations governing the evolution of the state and given a transformer management decision policy that is a function of the state, the lifecycle costs of the management policy can be calculated. Therefore, in principle, it is possible to find the management policy that has least cost by determining the state dynamics associated with any policy. That idea is fundamental to the methodology presented in this guidebook.

State Dynamics

While it is sometimes useful to develop a maintenance policy for each transformer individually, this report considers asset management of an entire fleet of transformers. Thus, the dynamic equations represent the evolution of the entire transformer inventory. In particular, define an array to represent the population of transformers as a function of the state variables. Let

$X(s)_k$ = transformer population that is in state s at time k; call the population in a particular state a *cohort*.

The subscript k indexes chronological time, measured in 5-year intervals, during the planning period, which is can be represented as the interval $\lfloor 0, T_f \rfloor$, for some arbitrary number of years, T_f, called the *planning horizon*. The state s, as discussed above, can be any collection of information that is sufficient to describe the transformer and to determine the state evolution in the next period.

The dynamic equations mathematically describe how the transformer population evolves over time. The next chapter discusses the dynamic equations in detail. Generally, there are four important aspects of the formulation. First, the distribution of the transformer population among the states in the next interval is a function of the current distribution. For instance, in the simple case that the state is the transformer's age, and transformers do not fail, the population at age t in interval k equals the population at age $t+1$ in interval $k+1$

$$X(t+1)_{k+1} = X(t)_k .$$

Second, the migration of transformers among the states usually occurs probabilistically. Again, let the state *s* be the age *t*. Now, if transformers fail at age t at the rate $h(t)$, then the surviving population of age $t+1$ in interval $k+1$ is

$$X(t+1)_{k+1} = [1-h(t)]X(t)_k$$

The function $h(t)$ is the *hazard rate* for transformers. The hazard rate is determined empirically, using some generally applied functional forms. We discuss the estimation process in detail, below (see Chapter 4).

In general, the probability that a transformer in one state evolves to another state in the next time period is called the *state transition probability*.

Third, the dynamic equations represent the evolution of the entire transformer population rather than that of an individual transformer. The population dynamics take advantage of the law of large numbers, which states, roughly speaking, that the average behavior of a population that consists of a large number of independent, identical random variables is approximately the average or expected behavior of an individual member of that population. In this case, if the failure probability of an individual transformer at a given age is $h(t)$, then $h(t)$ is the fraction of the population at that age that will fail and the remaining fraction $1-h(t)$ will survive. Using the population dynamics allows one to compute the state-dependent asset management policy for the entire population at once, whereas using the individual asset model would require many runs.

Fourth, the migration of transformers among the states can depend on some action taken with respect to a particular transformer in a particular state. For instance transformers that are to be replaced as a consequence of the state-dependent optimal policy or that have failed may be replaced with new transformers, so that the population of new (that is, age 0) transformers at the beginning of the next interval is the total of the transformers of all ages that are replaced in the current period. Suppose that the optimal policy indicates that all transformers of age t^* or greater must be replaced; then

$$X(0)_{k+1} = \sum_{t^*>t>0} h(t)X(t)_k + \sum_{t\geq t^*} X(t)_k$$

where the first term represents replacement of failed transformers and the second represents scheduled replacements of over-age transformers. (Another formulation, as noted above, could include a decision to replace or overhaul a failed transformer, if the number of past failures does not exceed some threshold. The next chapter discusses the formulation of the mathematical model in greater detail.)

A *decision rule* specifies the action to be taken in a particular state at a given time in the analysis period. Let

$d(s)_k$ represent the decision made at time k for a transformer that is in state s.

In general, the dynamic equations describe how the transformer population migrates among the states as a function of the state transition probabilities and state-dependent decision rules. A *policy* is a complete specification of the decision rules used for all possible states s at all times k during the analysis period.

Two kinds of costs can be incurred in implementing a policy, the costs of the decisions themselves and the costs incurred by transformers in a particular state, called the *occupancy* costs. For example, decision costs could include the costs of transformer inspection and maintenance. Occupancy costs could arise from handling trouble reports that occur between tests or inspections, which may increase with transformer age or changes in transformer condition. The dynamic equations determine the costs of implementing a policy.

Inspection and testing

The purpose of inspection and testing is to determine the condition (that is, the state) of the transformer and thereby decide what maintenance actions to take. In general, the purpose of inspection and testing is to assess the unobservable components of the state of a transformer. In the present formulation of the methodology, the unobservable component of the state of a transformer is a single, summary condition variable that can take on one of four values: {*Acceptable* (green), *Monitor* or *Watch* (blue), *Marginal* (yellow), and *Unacceptable* (red)}. The outcome of a test or a test protocol is assumed to be a claim about the true condition of the transformer. The test is imperfect, hence the consequence of the test is to revise the probability distributions on the transformer condition. Prior to inspection and testing, the probability distribution on the unobservable conditions is governed by an age-dependent dynamic process that is specified as part of the input to the methodology.

As the probability distribution of the unobservable components of the state changes, the future behavior of the transformer changes. In particular, as the transformer condition deteriorates, failure becomes more likely. Therefore, knowledge of the transformer condition can motivate overhaul or replacement of the transformer. It is the interplay between test results, dynamically changing conditions, and the costs and consequences of decisions that guides the specification of the optimal transformer management policy.

There are several modeling and analysis issues that are important considerations with respect to testing. The state description selected must be consistent with the actual test outcomes so that any particular outcome can be expressed as a claim or conclusion about the unobservable state of the transformer. The typical test result will be a number, perhaps a temperature, chemical analysis, or a map of hot spots, or a phase angle, depending on the test. These direct measurements suggest something about the condition of the transformer. The modeling issue is how best to relate the observed measurement to the state description. In the present implementation, the unobservable state of the transformer is assumed to be in one of several qualitative conditions with some probability. For example, one condition that a transformer can be found in is *Marginal*. For a test that measures degradation, a range of test outcomes must be specified such that if the actual outcome is in that range, then the conclusion of the test is that the transformer is in the condition *Marginal.* Test outcome ranges must be specified such that any possible test outcome corresponds to a single transformer condition.

The analysis issue is to specify the accuracy of the test. The accuracy of the test is given by its *likelihood function*. The likelihood function of a test is the probability that the test indicates that the transformer is in condition X when it is truly in condition Y. This is a conditional probability. The conditional probability of the occurrence of an event X given that the event or condition Y is known is written p{X|Y}. We use the notation {“X”} to indicate the event that the test result indicates that the transformer state is X. The notation is adopted to indicate that the test is making a claim or a statement about the condition. Thus, the likelihood function of the test is the collection of probabilities p{“X”|Y} for all possible unobservable transformer states X (and Y). In particular, the likelihood p{“X”|X}, which is the probability that the test says the transformer is in state X when the transformer is actually in state X, is the accuracy of the test.

The final analysis issue is to specify the revised state of the transformer given a test result. This is the conditional probability p{Y|“X”}, which answers the question “How likely is it that the transformer is in state Y if the test said it was in state X?”. This revised probability is found by an application of Bayes’ Theorem, which is a well-known result. (See virtually any text on probability; e.g., Ross, S. *A First Course In Probability, 4th Ed., Macmillan: New York, 1988.*)

In the present implementation, the analyst is asked to specify the likelihood function for each test. As we discuss below, there are expert default values that are available. These default values have been implemented in the current version of the methodology.

The costs of inspection, testing, overhaul, and replacement vary by utility and must be specified as part of the input data for the analysis.

Stressors

The behavior of transformers also depends on additional considerations that we collect under the heading of *stressors*. There are two classes of stressors, *environmental* (weather, temperature, …) and *utilization* (electrical loading, duty cycles, …). Stressors are interesting because it is reasonable to suppose that the presence of a stressor can affect the hazard rate (the probability that a transformer will fail), the accuracy of a test, and the occurrence of unobservable conditions.

In the present methodology, the stressors are expressed in the simplest way: the stressor is either present or it is not. The effects of stressors on the hazard rate, test accuracy, and the arrival rate of unobservable conditions are specified as part of the inputs to the methodology.

Policy Specification

As noted above, a policy, which is the solution to the transformer management problem, is a specification of the decision made with respect to transformers in each state s at each time k, for all states and all times. A policy that makes the same decision for a given state for all times is called a *stationary* policy. The optimal stationary policy is that set of decisions that minimizes the expected present value of the cost of following the policy for the indefinite future.

For example, a simple policy based on the simple state $s = t$, where t is the service age of the transformer, is to replace a transformer when it becomes ten years old and do nothing, not even test, with any transformer that is less than ten years old. (Such a policy is almost surely not cost effective, unless transformers are very likely to fail when they age beyond ten years, and the cost of a failure of a single transformer is approximately the same as the cost to replace a transformer – neither of which is likely to be the case.) In the present notation, this is the policy

$d(s)_k$ = *replace*, when $s \geq 10$ for all k, and

$d(s)_k$ = *do nothing*, when $s < 10$ for all k .

Notice that in this simple case, the state variable for the transformer is its age, which means that nothing else need be known for policy analysis. The decision policy is based only on the age of the transformer. The decision alternatives are the mutually exclusive but not collectively exhaustive actions *replace* and *do nothing*. The policy is stationary, so that the choice is the same for any period k . Note also that this specification of the policy, combined with the appropriate costs, would be sufficient to evaluate the expected present value of the cost of this policy applied to a population of transformers if one additional aspect of transformer behavior were known – the fraction of transformers of any age (less than ten years old) that will fail in the next year. This fraction is specified by the hazard rate $h(t)$.

In the cases discussed in the rest of this report, the structure of the problem is virtually the same as suggested in this simple example, but the state variables and the decisions can become far more complex.

As an example of a more complex policy, consider the interaction between transformer age and the outcome of a test or test protocol on the unobservable condition of the transformer. When the unobservable condition falls below a given level for a transformer that is sufficiently old, then the transformer is replaced. Thus, a test that detects, with some probability, the presence of condition degradation, will provide information that is sufficient to determine policy.

Thus, the observable transformer state in this situation is the service age of the transformer and the *outcome* of the degradation detection test, represented by the pair {service age, test outcome}. In the present formulation, we suppose that the outcome of the test can be one of four conditions: {No Degradation Detected, Mild Degradation Detected, Moderate Degradation Detected, and Severe Degradation Detected}. Note that the test outcomes are similar to the actual conditions.

The test outcomes are claims about the actual condition, which is unobservable. If the test were perfectly accurate, then when the test makes a claim, the claim is identical to the actual condition. But in most real cases, the test is not perfect and so the test outcome does not perfectly indicate the actual condition. Instead, the test outcome revises the probability distribution on the actual condition, as discussed above.

The decision alternatives are {Do Nothing, Overhaul, Replace}. The policy specification can be presented in a matrix that indicates which alternative to adopt as a function of the observed state (see Table 2-1; the colors in the table have nothing to do with the color labels used to describe the conditions). The table describes policy regions. For example, if the test result is None = No Degradation Detected, then the optimal policy is to do nothing, regardless of age. (We have suppressed the quotation marks around the test result.) If the test result is Mild = Mild Degradation Detected, the optimal policy is to overhaul a transformer that is no more than 25 years old and replace transformers that are at least 26 years old. If the test result is Moderate = Moderate Degradation Detected, the optimal policy is to overhaul transformers no older than 10 years and replace those that are at least 11 years. The same policy applies if the test result is Severe = Severe Degradation Detected. Of course, this policy is only an illustration and does not necessarily correspond to a real policy based on analysis.

Table 2-1
Policy Specification as a Function of Observable State (Age) and Transformer Test Outcome

Degradation Condition Test Result	Transformer Age							
	0-5	5-10	11-15	16-20	21-25	26-30	31-40	>41
None "Green"	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing
Mild "Blue"	Overhaul	Overhaul	Overhaul	Overhaul	Overhaul	Replace	Replace	Replace
Moderate "Yellow"	Overhaul	Overhaul	Replace	Replace	Replace	Replace	Replace	Replace
Severe "Red"	Overhaul	Overhaul	Replace	Replace	Replace	Replace	Replace	Replace

3

MODELING TRANSFORMER DYNAMICS AND MANAGEMENT POLICIES

This chapter describes how to represent mathematically the dynamics of transformer behavior with respect to failure and degradation, and the effects of maintenance, testing, and replacement decisions applied to an inventory of transformers. The dynamic behavior of the transformer inventory is captured by the changes in the state of the transformer and the behavior of the unobservable conditions. The model development links the dynamic behavior and the decisions made to the costs of the policy specified by the decision rules. These costs are minimized by the optimal policy.

Optimal Policy Model Based on State Variable Dynamics

The purpose of this model is to specify the optimal stationary policy for managing a population of transformers. A policy is the specification of the decision made for transformers in each state s at each time k for all states and all times. A stationary policy makes the same decision for a given state for all times. The optimal stationary policy is that set of decisions that minimizes the expected present value of the cost of following the policy over the indefinite future.

In this model, the observable state of the transformer consists of the pair (t,o), where t is the age of the transformer, and o indicates whether the transformer has been overhauled. The reason that it is important to know whether the transformer has been overhauled is to specify the applicable maintenance policy.

Define the one-period *hazard rate*,

$h(t,o)$ = the probability that a transformer that has survived until age t and has overhaul status o does not survive to age $t+1$. Note that the dependence of the hazard rate on service age is an empirical question. We discuss how to estimate hazard functions in Chapters 4 and 5.

Let

$X(t,o)_k$ = transformer population at time k of service age t and that have overhaul status o. The subscript k indexes chronological time, measured in 5-year intervals, during the planning period. By our conventions, the population is identified at the beginning of interval k.

Notice that the pair (t,o), which is the state variable s for this model, is sufficient to identify a part, or cohort, of the transformer population. All members of the transformer cohort with the same state behave dynamically in the same way. Therefore, at any time k , the transformer population can be represented as a matrix such that each element of the matrix is the number of transformers in state (t,o) at time k, which we denote $\left[X(t,o)_k\right]$ with columns representing ages $t > 0$ and the rows identifying whether or not the transformers have been overhauled. The notation o = 0 indicates that the transformer has not yet been overhauled, and the notation o = 1 indicates that the transformer has been overhauled.

The state dynamics of transformers are as follows. A transformer enters the (*service age, overhaul status*) state (t, o) at the beginning of interval k if at the beginning of the previous interval (k-1) it was in state (t-1, o) and during the interval k-1 it did not fail and its overhaul status did not change. Further, a transformer can enter the state (t,1) at the beginning of interval k if at the beginning of the previous interval (k-1) it was in the state (t-1+t_o, 0) and it was overhauled (which changes the overhaul state from 0 to 1, and overhauling reduces the service age by t_o).

That is,

$$X(t,0)_k = X(t-1,0)_{k-1}(1-h(t-1,0)) \tag{3-1a}$$

and

$$X(t,1)_k = X(t-1,1)_{k-1}(1-h(t-1,1)) + X(t-1+t_o,0)_{k-1} \tag{3-1b}$$

if transformers in state (t-1+t_o, 0) are overhauled under the optimal policy.

A policy for transformer management indicates when transformer should be replaced. The policy is specified by the stationary (not time-dependent) decision rule $d(s)$. A typical decision rule specifies that a transformer should be replaced if its age is at least t^* regardless of whether it has been overhauled. For such a rule, we find that the number of transformers that are replaced at the end of interval k-1 is

$$\sum_{t\geq t^*}\sum_{o} X(t,o)_k = X(0,0)_k \tag{3-2}$$

which we denote as $X(0,0)_k$ because these are new transformers that have not been overhauled and are put in service at the beginning of interval k. One way of thinking about such replacements is that the policy prevents a transformer that would be in state (t^*, o) (or worse) in interval k from being retained in service for that interval.

Condition Dynamics and Testing

At any time, a transformer can be in one of several unobservable conditions. These conditions, as defined in Chapter 2 (see also EPRI report 1012503), are levels of transformer degradation. The hazard function depends on the condition of the transformer.

Let c_j denote the j^{th} condition.

Let $p\{c_j \mid t,o\}$ denote the probability that a transformer of age *t* and overhaul status *o* is in condition c_j. These probabilities are input parameters to the model. It is important to note that in this formulation, the condition of the transformer is both dynamic and uncertain (in addition to being unobservable). The dependency of condition on transformer age and overhaul status is typically specified by expert judgment rather than recorded data. Thus, the condition specification is a forecast of an uncertain variable.

Let $h(t,o \mid c_j)$ denote the hazard rate conditional on the transformer being in condition c_j. These conditional hazard rates are inputs to the model.

Then the unconditional hazard rate, the hazard rate used in (3-1a, b), is given by

$$h(t,o) = \sum_{c_j} p\{c_j \mid t,o\} h(t,o \mid c_j) . \qquad (3\text{-}3)$$

Therefore, the hazard rate of a transformer depends on the probability distribution on the (unobservable) condition of the transformer.

Although it is not possible to observe the condition directly, the transformer can be tested for the presence of various degraded conditions. The test reports that the transformer is in a particular condition, but the test is not perfectly accurate. As we noted in the Introduction (p. 1-8 ff.) and Chapter 2 (pp. 2-5 ff.), we describe the accuracy of the test by the likelihood that the test will report that the transformer is in condition c_j when it is indeed in that condition, or the probability $p\{"c_j" \mid c_j\}$, which is written to suggest that the test *says* that the condition is c_j. In order to specify completely the behavior of the test, the likelihood function must be given. The likelihood function indicates the chances that the test outcome, in whatever units it is actually observed, will be translated into a conclusion about the condition of the transformer, as a function of the actual transformer condition. There are two interesting aspects to this question. First, the actual, or directly observed, test outcome must be translated into a statement about the actual transformer condition. Second, the uncertainty associated with the test outcome must be specified. This uncertainty can be thought of as the degree to which the test can make a false identification. We write the likelihood that the test concludes that the transformer is in condition c_l when the actual condition of the transformer is c_j as the probability $p\{"c_l" \mid c_j\}$. Note that when the subscripts are equal, $j=l$, the likelihood is interpreted as the test accuracy, as stated above. All likelihoods must be specified as inputs to the model.

As a result of the test outcome, the probability distribution on condition is revised. The probability distribution of the state $p\{c_j \mid t,o\}$ is called the *prior* distribution, since it represents the presumed condition of the transformer at the given age prior to performing the test. The revision is accomplished by application of Bayes' Theorem, which determines the (posterior) probability that the transformer is in state c_j given that the test says that it is in any state c_l:

$$p\{c_j \mid "c_l"\} = p\{"c_l" \mid c_j\} p(c_j\} / p\{"c_l"\} . \qquad (3\text{-}4)$$

where $p\{"c_l"\} = \sum_{c_j} p\{"c_l" \mid c_j\} p(c_j\}$ is the probability of test outcome $"c_l"$.

This equation, valid for all states (*t,o*), permits the hazard rate to be updated as a result of the test. Therefore, the test-outcome-based hazard rate can be found:

$$h(t,o\,|\,"c_l") = \sum_{c_j} p\{c_j\,|\,"c_l",t,o\}h(t,o\,|\,c_j) \qquad (3\text{-}5)$$

Dynamic Decision Modeling

These definitions and equations are sufficient to describe the dynamics of the decision model. At the beginning of any period, *k*, a decision is made. The state of the transformer at the beginning of the period is the pair (*t,o*) and the unobservable condition is given by the prior probability distribution $p\{c_j\,|\,t,o\}$. Based on the state of the transformer, a decision is made.

Let this decision be denoted $d(t,o,"c")$.

Notice that the decision function has three arguments, the age of the transformer, the overhaul status, and the test outcome. At the beginning of the period, the test outcome is unknown, so the value of this argument is null. The result of the decision can be either to replace the transformer, to maintain it in some way, to test it, or to do nothing. (In the present implementation, the maintenance decision is restricted to *overhaul*. The consequences of the *overhaul* decision are modeled by changing the service age of the transformer. This changes the probability distribution on the unobservable condition. This, in turn, modifies the hazard rate for the transformer through (3-3). Therefore, *overhauling* a transformer reduces the probability that it will fail in the next period.) If the decision were to test, then the test outcome *"*c_l*"* is revealed, and the third argument in the decision function is known. After testing, the possible outcomes of the decision function are to replace the transformer, to maintain it in some way (*overhaul*), or to do nothing.

If the decision is to do nothing (before or after testing), then the transformer that is in state (*t,o*) in interval *k* either fails and is replaced or does not fail and enters state (*t*+1*, o*) in interval *k*+1. This behavior is described by (3-1a).

The simplest way to represent the effect of the overhaul decision is to identify a transformer of less age that has not yet been overhauled and behaves, going forward in time, identically to an overhauled transformer. This identification can be done in at least two ways. Either a specific age change is given as in input to the analysis (e.g., an overhauled transformer behaves as if it were ten years younger) or changes to the observable and unobservable conditions are given as inputs (e.g., the probability distribution of the condition of an overhauled transformer is modified such that probability that the transformer is in the worst condition state goes to zero, and the overhaul status is set to one). These considerations introduce the concept of the *effective age* of a transformer. The actual age of the transformer does not change, but if the transformer is overhauled, it is, in effect, made younger. Therefore, if the decision is to overhaul (before or after testing), then the transformer that is in state (*t,*0) in interval *k* is changed to a transformer that is in state (t-t_o,1) in interval *k*, where t_o is the reduction in effective age that is the result of an overhaul. (The effect of changing the probability distribution on the unobservable state is represented as a change in transformer age. Thus, *overhaul* does what it claims to do in this model.) The overhauled transformer either fails or does not fail, and is then either replaced or enters state (t-t_o+1, 1) in interval k+1. This behavior is described by (3-1a).

If the decision is to replace (before or after testing), then the transformer state is changed to (0,0). This behavior is described by (3-2).

If the decision is to test, then the result of the test is observed, the hazard function is updated [as described by (3-3), (3-4), and (3-5)], and the decision to replace, overhaul, or do nothing is specified by $d(t,o,"c")$. Equations (3-1) and (3-2) apply.

This describes the decisions made in a single period. At the end of the period, which is equivalent to the beginning of the next period, the transformer state (t,o) is known and the unobservable condition probabilities are known, so the identical decision process can be applied.

Costs and Objective Function

The costs associated with the transformer inventory are the replacement cost, the failure cost, the cost of testing, the maintenance (overhaul) cost, and operating costs that are associated with the unobservable states. The last category of cost is presently not implemented. In general, such a cost is based on the probability of the transformer occupying any of the unobservable conditions.

Let R = the cost of replacing a transformer. Then the cost of replacements in interval k is

$$R_k = RX(0,0)_k \tag{3-6a}$$

Let F = the cost of a transformer failure. Then the cost of failures in interval k is

$$F_k = \sum_t \sum_f \sum_{c_l} FX(t,o)_k h(t,o\,|"c_l")p\{"c_l"\}. \tag{3-6b}$$

Note that if there is no test result, because testing was not chosen, then the conditional outcome is null and the hazard function is given by (3-3) instead of (3-5).

Let = the cost of testing a transformer . Then the cost of testing in interval k is

$$T_k = \sum_t \sum_o \tau X(t,o)_k d(t,o,\phi) \tag{3-6c}$$

where the third argument in the decision function, which is the test outcome, is null and the summation is over all pairs (t,o) such that the value of the decision function $d(t,o,\phi) = TEST$.

Let V = the cost of maintenance (overhaul). Overhauling can be chosen before or after a test is done. There are two terms in the cost of overhauling. Then the cost of overhauling in interval k is

$$V_k = \sum_t \sum_o VX(t,o)_k d(t,o,\phi) + \sum_t \sum_o \sum_{c_l} VX(t,o)_k\, p\{"c_l"\}d(t,o,"c_l") \tag{3-6d}$$

where the summations are over all pairs (*t,o*) such that $d(t,o,\phi) = d(t,o,"c_l") = overhaul$.

Let O_{c_j} = the operating cost of a transformer if it is in (unobservable) condition c_j. Then the operating cost in interval *k* is

$$O_k = \sum_t \sum_o \sum_c \sum_{c_j} O_{c_j} X(t,o)_k \, p\{c_j \mid t,o,"c"\} p\{"c"\} \qquad \text{(3-6e)}$$

where the probability distribution on the unobservable condition is modified by the test outcome if testing were chosen.

Then, the present worth of the policy that makes decisions $d(t,o,"c")$ is, for annual interest rate *r*

$$PV_\infty(r,d) = \sum_{k=1}^{\infty} (1+r)^{-k} [R_k + F_k + T_k + V_k + O_k]. \qquad \text{(3-6f)}$$

The decision problem is to choose the decision rule $d(t,o,"c")$ that minimizes the present value of the policy given by (3-6f). The solution to this problem has been implemented.

Data Required to Support the Optimal Policy Model

The essential description of the behavior of transformers is the hazard rate function *h*(*t,o*). Many utilities have data that can be used to estimate these hazard rates. We discuss methods for estimating hazard functions in the next chapter.

The consequences of decisions such as overhauling must be specified with respect to the state variables. This can be done using actual data or provided by expert assessment.

The unobservable conditions are described by probability distributions $p\{c_j \mid t,o\}$. These are usually provided by expert assessment.

Testing accuracy, given by likelihood functions $p\{"c_j" \mid c_j\}$, are usually provided by expert assessment.

The economic data required are the following.

- The cost of replacing a transformer.
- The cost of failure of a transformer.
- The cost of testing a transformer.
- The cost of rejuvenating a transformer.
- The operating cost if a transformer is in a particular condition.

These costs can be found from utility records or provided by expert assessment.

4
ESTIMATING HAZARD FUNCTIONS

The previous chapter presented the dynamic decision model that determines the least cost policy to manage a population of transformers. This chapter turns to the issues of data. The model described above requires estimates of parameters in order to perform an analysis. Specifying model parameters usually requires a combination of data analysis and expert judgment, with judgment playing an essential role when data is sparse or unavailable.

There is an important relationship between models and data: it is inefficient and misleading to collect data prior to specifying a model. Ideally, modeling requirements should drive data gathering, so that one does not incur the expense of gathering unnecessary or irrelevant data. In reality, of course, utilities usually have existing data sets that were gathered for any number of reasons. Therefore, in practice, modeling is often formulated to conform to existing data or to available judgment; however, this is not the ideal situation. Even in this situation, though, a model can tell an analyst what data to look for within the existing data sets and how to analyze that data.

As discussed in the previous chapter, the models require three general types of data:

1. Data to describe the dynamic behavior of the transformer.
2. Data to describe the failure behavior of the transformer.
3. Data to describe the economic behavior of the transformer.

The data requirements vary according to the modeling assumptions and cases, which in turn depend on the state description of the transformer and the policy alternatives that are to be investigated, as discussed in Chapter 3. The data requirements for describing the dynamic behavior and the economic behavior of transformers were discussed in the preceding chapter. This chapter focuses on the issue of describing the failure behavior of transformers.

Describing the failure behavior of transformers is essential when applying the dynamic decision model. As noted above, we seek a hazard function, the probability that a transformer will fail in the next year, given that it has survived until the end of the previous year. We show in this chapter how event data can be used to estimate the hazard function.

Overview of the Estimation Method

In developing hazard functions from empirical data, it is important to keep in mind two properties of hazard functions that one expects the estimates to obey. First, the hazard rate is an increasing function of age; that is, transformers are more likely to fail as they get older. Second, the hazard function is relatively smooth; that is, the hazard rates for age t and $t+1$ are not too different from each other.

In practice, purely empirical estimates of failure rates often fail to behave this way, due primarily to natural, random fluctuations in the data (see for instance, Figure 4-1). These fluctuations may become particularly severe when the underlying data is sparse for a particular age or another explanatory variable, because each individual observation carries more weight when there are relatively few of them. The noise in the estimated hazard rates should not be allowed to drive the treatment/replacement decision. Therefore, some means is needed to filter these naturally occurring fluctuations.

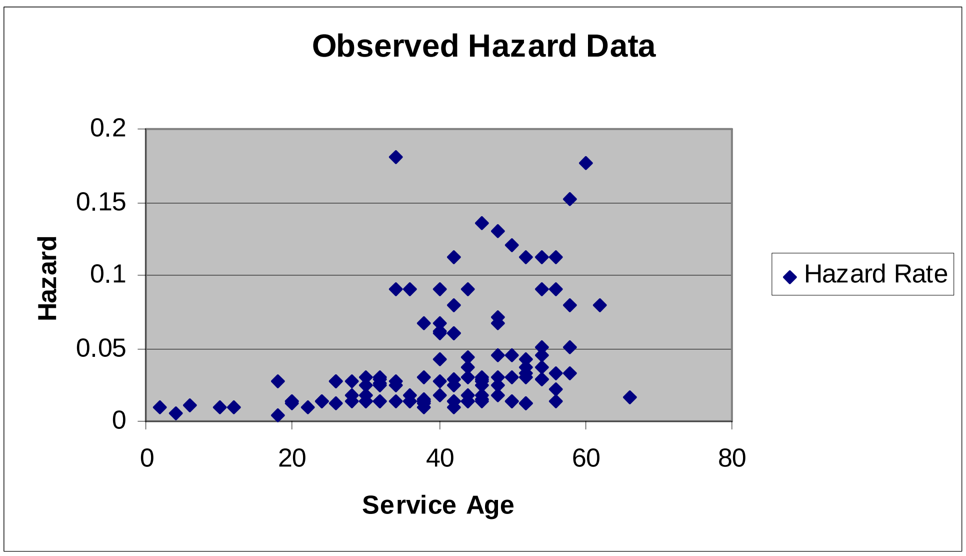

Figure 4-1
Observed Hazard Data for an Inventory of Transformers

The approach we use to control empirically observed variations is to develop a *model* of the hazard function. An empirical model summarizes the information contained in observed data in a way that is consistent with expectations about the behavior of equipment. In many applications, the observed hazard rate is expressed as a so-called *bathtub curve* (see Figure 4-2). 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. The burnout period reflects the effect of aging.

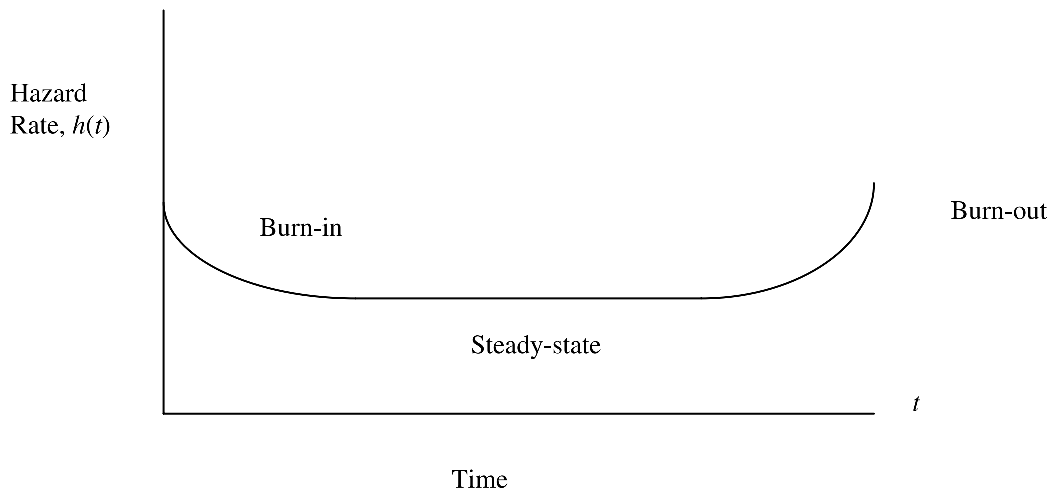

Figure 4-2
Hazard Rate "Bathtub" Curve

Other models specify a functional form for a hazard function that depends on several parameters that can be estimated from the hazard data. Many applications in statistical analysis commonly use a linear regression model, in which a straight line represents the observed data. However, the linear regression model does not represent hazard data very well. We propose alternative approaches, including the use of the piecewise linear model and the Weibull model. These approaches have properties that more closely represent the expected and observed behavior of transformer failure data.

Transformer Failure Models

We distinguish between failure models that are based on statistical life distributions—exponential, normal, lognormal, Weibull, extreme value—and failure models that are purely empirical—the bathtub curve of Figure 4-2 and the piecewise linear hazard rate. The statistical life distributions typically respond to a particular phenomenon or failure mechanism. We begin the discussion 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

$$p(k) = \exp(-\lambda t)\ (\lambda t)^k / k!\ , \quad k = 0, 1, 2, \ldots \qquad (4\text{-}1)$$

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.

For asset management purposes, the exponential failure model is not very useful, since such failures are essentially unpredictable, the equipment shows no deterioration with time, and therefore, run-to-failure is the optimal policy.

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 -• < t < •. 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 4-3.

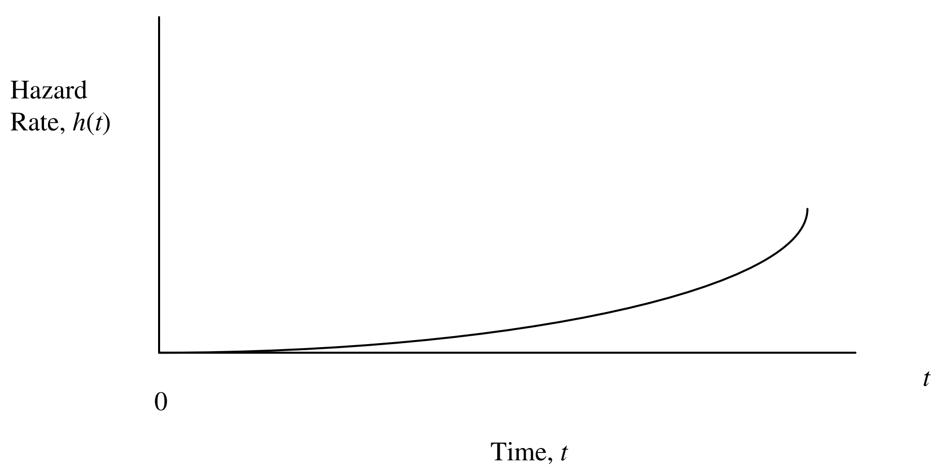

Figure 4-3
Normal Hazard Rate

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 $\mu = r/\lambda$, $\sigma^2 = r/\lambda^2$.)

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 4-4.

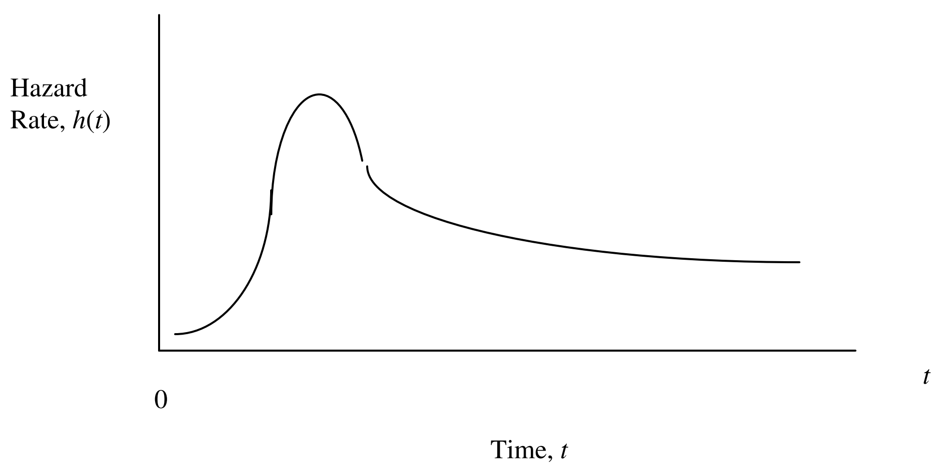

Figure 4-4
Lognormal Hazard Rate

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 α

> 1, constant if $\alpha = 1$ (and the Weibull becomes an exponential distribution), and decreasing if $\alpha < 1$. See Figure 4-5.

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.

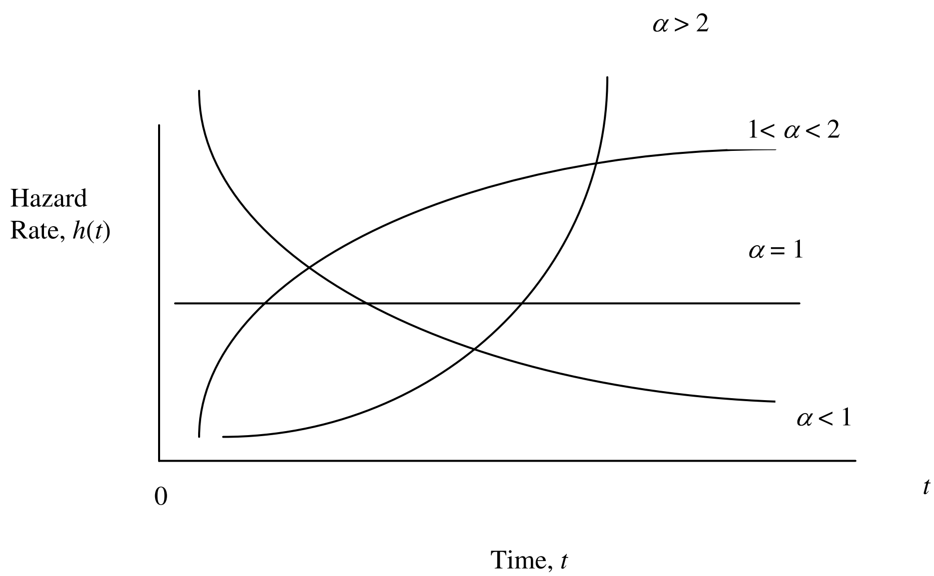

Figure 4-5
Weibull Hazard Rate

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 (*Statistics of Extre*mes, Columbia University Press, 1958). Abernethy (*The New Weibull Handbook, 2nd Edition,* Dr. Robert B. Abernethy Publisher, North Palm Beach, FL, 1996) provides a good reference for understanding Weibull distributions.

The behavior of the extreme value hazard functions is shown in Figure 4-6. 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.

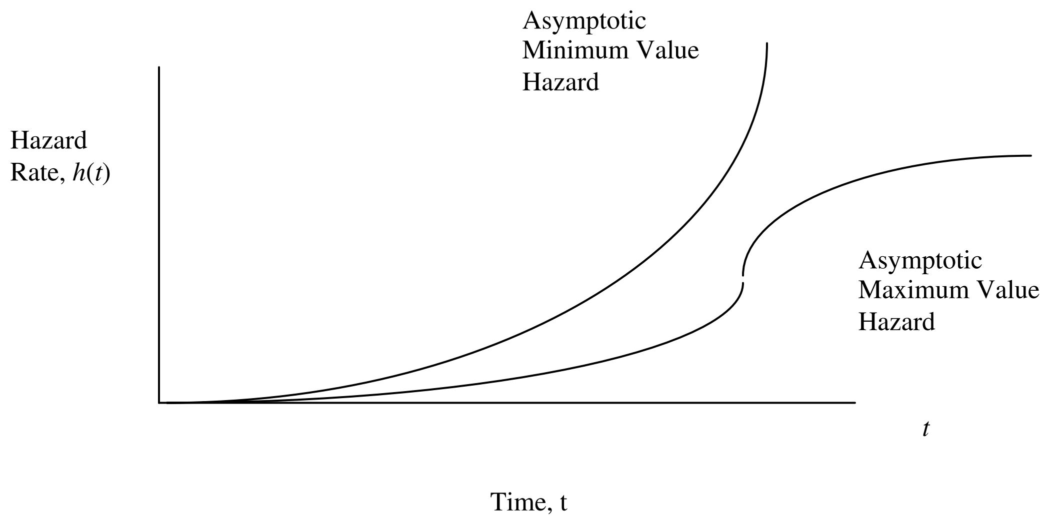

Figure 4-6
Extreme Value Hazard Rate

Empirical Hazard Functions

The empirical hazard functions are not based on any underlying life distribution. (This is a mathematical modeling distinction which may not be useful in practice.) 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

$$h(t) = dF(t)/dt[1 - F(t)]^{-1} \qquad (4\text{-}2)$$

where $dF(t)/dt = f(t)$, the probability density of the lifetime of the asset. This implies that the hazard rate is dependent on what is assumed about the life distribution, $F(t)$. The life distributions—exponential, normal, lognormal, Weibull, extreme—are what characterized the hazard functions discussed above, and (4-2) determined $h(t)$. Alternatively, the empirical hazard functions assume a model for the hazard function directly, find the best fit of the data to that model, and then find the life distribution by solving the differential equation (4-2), the solution of which is

$$F(t) = 1 - \exp(\int_0^t h(x)dx). \qquad (4\text{-}3)$$

Bathtub Curve

The bathtub curve, shown in Figure 4-1, 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 4-6, is a popular form. For transformer repair/replace studies, which focus on end-of-life behavior, 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 to transformers. 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 4-7. 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 $t > T$, $h(t) = h_{ss} + m(t\text{-}T)$. Therefore, the function is specified by three parameters, h_{ss}, *T*, and *m*.

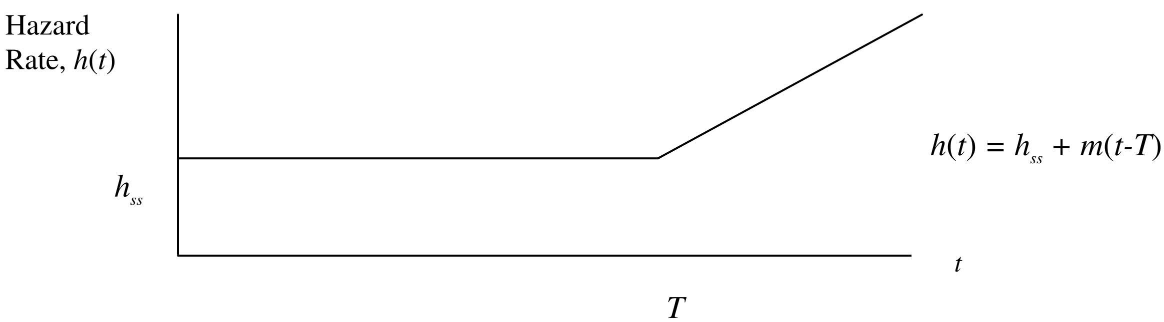

Figure 4-7
Piecewise Linear Hazard Function

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 transformers 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.

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.

Estimating the Hazard Function Based on Transformer Age

The simplest state description of a transformer is its age. In this case, if age is all that is known about the condition of a transformer, the state has a very elementary dynamic behavior: the transformer ages every year until it is taken out of service. Hence, the basic data for this case is the failure likelihood of a transformer as a function of its age, that is, a hazard function.

Estimating hazard rates from empirical data in this case uses the following procedure. Identify two time scales in the data, t denoting the age of the transformer and k denoting the calendar time (e.g., the year 2004). The hazard rate for transformers of age t observed at time k equals the ratio of the number of transformers in service of age t at time k that failed between years k and $k+1$ divided by the number of transformers of age t that are in service at time k. Assuming that the failure process is a stationary one – that is, the hazard rate does not depend on k – then a single snapshot of the transformer inventory at a single time k would be sufficient to estimate the hazard function, assuming there are sufficient transformers of all ages t in the observed snapshot. Generally, however, data can be collected over time, for various times k.

An example of an observed set of hazard rates was given in Figure 4-1. The data presented in the Figure are estimated exactly as described in the paragraph above. Notice that the hazard rates for different ages show considerable variation, not a smooth progression. This variation results from the circumstances that applied in this particular set of underlying failure data, and in particular, the large swings at later ages (20 and beyond) result from the relatively small numbers of transformers that survive. Therefore, rather than allow these sample-specific fluctuations to drive the analysis, it is preferable to develop a smoother functional representation of the hazard function. One might try to fit a straight line to the hazard data using least squares linear regression. However, the data observed in Figure 4-1 suggests that the hazard function can be better approximated by the two-parameter piecewise linear hazard function described above, shown in Figure 4-7. Another reasonable candidate that summarizes the information in the observed data set might be the Weibull hazard, as shown in Figure 4-5, with parameter > 1.

In the next Chapter, we will provide an example. In the present section, we will describe the steps necessary to estimate a piecewise linear hazard function and a Weibull hazard function from observed data. In the present implementation of the methodology, only the piecewise linear hazard function is implemented. We present the Weibull estimation methods for completeness.

Estimating a Piecewise Linear Hazard Function

The estimation problem is the following: given a collection of data describing the failure of transformers, estimate the three parameters, (h_{ss}, *m, T*) of a piecewise linear hazard function. We describe the solution in the following steps.

Step 0. Create the data set. The data set will have only two variables for each year *k*, the age (*t*) of the transformers and the fraction of installed transformers of that age that failed in the given year. That fraction may be expressed as failures per year. We may denote this hazard rate as the observed hazard at age *t*, or $h_{obs}(t)$. As a result of this step, the collection of values ($t, h_{obs}(t)$) are specified. At this point, it is always useful to plot the data, to find a Figure similar to Figure 4-1.

Step 1. Estimate the age at onset of burnout, *T*. This is estimated by expert judgment and will vary in the analysis until a suitable value is found. A reasonable starting point is approximately 25 years, although the specific data set will suggest a value.

Step 2. Estimate the steady-state value h_{ss} by taking the average of all observed hazards for age $\leq T$.

Step 3. Subtract h_{ss} from all observed hazards for age greater than *T*. Subtract *T* from all observed ages greater than *T*. These steps create a partial data set that contains all the points observed after onset of burnout. The subtractions express these values as departures from the (*age, hazard*) pair (T, h_{ss}).

Step 4. Estimate the slope *m* by fitting a straight line to the points with ages greater than *T*. The straight line must go through the point (0,0) in order to estimate *m* correctly. (This estimation may be accomplished by using the Excel LINEST command with the restriction that the intercept of the estimated linear function is 0.) This means that the slope found will be the best value of *m* for the piecewise linear function that goes through the point (T, h_{ss}).

Step 5. Use the parameter estimates to determine the estimated hazard as a function of age. That is, find

$$h_{est}(t;h_{ss},m,T) = h_{ss} + \max\{0, m(t-T\} \qquad (4\text{-}4)$$

(where the second term is equal to zero if $t \leq T$ and is equal to $m(t\text{-}T)$ if $T > t$).

Define the estimation error

$$e(t) = h_{est}(t;h_{ss},m,T) - h_{obs}(t) \qquad (4\text{-}5)$$

which is the difference between the estimated value of the hazard and the actual observed data. Compute the sum of the squares of the errors, the value $J(T)$, which can be considered a function of the estimated value of *T*,

$$J(T) = \sum_t e(t)^2 \qquad (4\text{-}6)$$

$J(T)$ is a measure of the goodness-of-fit of the piecewise linear hazard.

Step 6. Return to step 1. Vary systematically the estimate of T. Select T that minimizes $J(T)$.

Estimating a Weibull Hazard Function

Again, the estimation problem is the following: given a collection of data describing the failure of transformers, estimate the two parameters, (,), of a Weibull hazard function. We describe the solution in the following steps. The basis of this procedure is that the Weibull hazard function, $h(t) = (\alpha/\beta)\,(t/\beta)^{\alpha-1}$, can be expressed as *log* $h(t)$ = (-1)*log t* + , where =*log* – *log* . Thus, the logarithm of the hazard rate is a linear function of the logarithm of the transformer age. Therefore, the parameters (,) can be estimated by the ordinary linear least-squares procedure.

Step 0. Create the data set. This is identical to Step 0 for the piecewise linear hazard function.

Step 1. Apply a logarithmic transformation to the data. Take the logarithm of the age t and the logarithm of the observed hazard $h_{obs}(t)$. This creates the data set, the pairs (*log t, log* $h_{obs}(t)$).

Step 2. Estimate the parameters -1 and using the linear parameter estimation procedure. (In Excel, this procedure is LINEST.) Find the parameter = exp[(log –)/].

Step 3. Use the parameter estimates to determine the estimated hazard as a function of age. That is, find

$$h_{est}(t;\alpha,\beta) = (\alpha/\beta)(t/\beta)^{\alpha-1} \qquad (4\text{-}7)$$

Define the estimation error

$$e(t) = h_{est}(t;\alpha,\beta) - h_{obs}(t) \qquad (4\text{-}8)$$

which is the difference between the estimated value of the hazard and the actual observed data. Compute the sum of the squares of the errors, the value J(,),

$$J(\alpha,\beta) = \sum_t e(t)^2 \qquad (4\text{-}9)$$

J(,) is a measure of the goodness-of-fit of the Weibull hazard. (If J is divided by the number of observations in the data set to create an average, such average values of J can be compared to determine relative goodness of fit.)

The Role of Expert Judgment

The full state of a transformer influences the hazard rate. This includes both the overhaul status of the transformer as well as the dependence of the hazard rate upon the unobservable condition.

It is reasonable to expect that utilities may not have overhaul status-dependent failure data available for each transformer. It is also reasonable to expect that utilities will not have sufficient data on hand to create forecasts for the dynamic process whereby transformer condition changes. Nor will data exist that indicates the effect of condition on hazard rate. Therefore, this effect, too, must be specified in other ways.

The approach taken to specifying these additional effects is expert elicitation. The process of expert elicitation results in estimates that are not supported by data and substitutes for data analysis when data is not present. The next chapter discusses the role of expert elicitation in the methodology.

In particular, the following issues are addressed by expert elicitation.

- The effect of overhaul status on condition dynamics
- The effect of transformer condition on the hazard rate, which includes the definition of the unobservable conditions, the specification of the dynamic process whereby conditions change, and the effect of condition on the parameters of the hazard function
- The effect of stressors on the hazard rate.

Conclusion

This chapter presented methods for estimating piecewise linear and Weibull hazard functions using available event data, that a utility might be expected to have. The methods are based on parameter estimation. The chapter also indicated that the hazard rates also depend upon the other components of the state of the transformer, such as the failure history and the condition of the transformer (the unobservable state). The expert elicitation that develops these dependencies is discussed in the next chapter. The chapter following the next one presents, in some detail, an example of the procedures for a sample data set.

5

HAZARD FUNCTION ESTIMATES: AN EXAMPLE

In this Chapter, we solve the estimation problem: given a collection of data describing the failure of transformers, estimate the hazard function. An example of a data set provided by a utility is the following. The data are ordered by increasing year installed, and decreasing age, *t*.

Table 5-1
Failure Data

Year Installed	Year Failed	Number of Failures	Transformers Operating	Age (t)	Hazard Rate h(t)
1951	1998	1	16	47	0.063
1953	1997	2	25	44	0.080
1953	1997	1	23	44	0.043
1953	1999	2	22	46	0.091
1960	2000	2	30	40	0.067
1964	2000	1	6	36	0.167
1965	1998	1	21	33	0.048
1965	1997	3	20	32	0.150
1966	1999	1	14	33	0.071
1966	1998	1	13	32	0.077
1969	1999	1	21	30	0.048
1969	1999	1	20	30	0.050
1969	1996	1	19	27	0.053
1970	1997	1	29	27	0.034
1970	1996	1	28	26	0.036
1970	1997	3	27	27	0.111
1970	1997	1	24	27	0.042
1971	1999	1	36	28	0.028
1971	1997	1	35	26	0.029
1971	1998	1	34	27	0.029
1972	1999	1	26	27	0.038
1972	1998	1	25	26	0.040
1972	1999	2	24	27	0.083
1972	2000	1	22	28	0.045
1979	2002	1	25	23	0.040
1983	2003	1	18	20	0.056
1995	2004	1	12	9	0.083

This table lists the times to failure of an inventory of transformers. The year the transformer 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 transformer when it failed. That is shown in column five. The number of failures observed for all the transformers installed in a given year is listed in column three. The number of operating transformers that were installed in the given year is listed in column four. The ratio of the number of failures observed (column three) to the number of operating transformers (column four) is the hazard rate, shown in column six. Note that the transformers that have failed decreases the number of operating transformers. The last two columns of Table 5-1 comprise the data required for estimating the hazard function.

In order to create Table 5-1, what must be known is the year the transformer was installed and the number of failures of that vintage transformer in any year. It is important to note that the data in Table 5-1 is not separated by vintage. By combining all vintages of transformers 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 transformers installed prior to the 1970s differ in behavior from transformers installed after 1969, and all transformers installed after 1969 behave 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 a transformer 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

Figure 5-1 presents the data in Table 5-1 graphically. Creating such a plot is a natural first step towards estimating the hazard function.

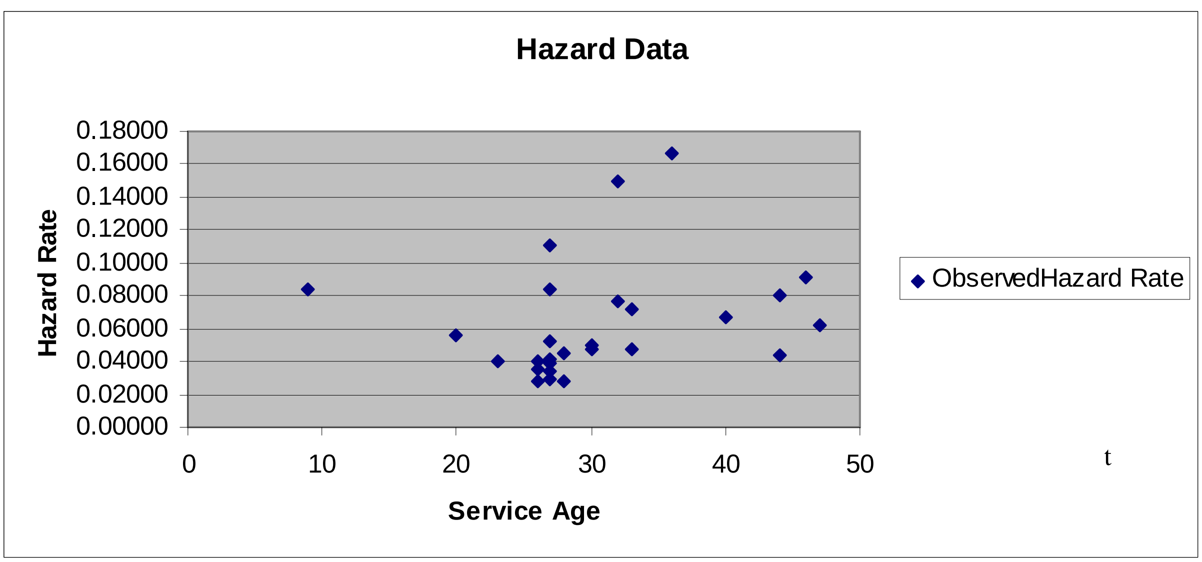

Figure 5-1
Observed Failure Data

It may be reasonable to discern a burnout period that begins at approximately thirty years in the data shown in Figure 5-1. Appropriate choices for hazard functions appear to be piecewise linear, normal, and Weibull. We restrict our attention to the piecewise linear and Weibull models.

Piecewise Linear Hazard Function

We will apply the procedure specified in Chapter 4 to fit a piecewise linear hazard function to the data in Table 5-1.

Step 0. This has been accomplished by creating Table 5-1.

Step 1. The onset of burnout is uncertain in this data set. It may begin as early as 25 years or as late as 35. We will investigate value of T in the range $25 \leq T \leq 35$.

Step 2. The results of the computations are presented in Table 5-2, below. As T varies over the range specified in step 1, several values of h_{ss} are found. The value of h_{ss} increases as T increases.

Step 3. For example, when $T = 30$ there are ten data points that have t > 30. These values are found in the first ten rows of Table 5-1. Subtracting 30 from each of the ages and subtracting 0.04759, the value of h_{ss} for $T = 30$, from each of the observed hazards yields the data set shown in Table 5-2 (net increasing hazard is the difference between the steady state hazard rate and the observed hazard rate).

Table 5-2
Hazard Observations During Burnout, for T=30

Net Increasing Hazard	Time beyond Steady-State
0.014911726	17
0.032411726	14
-0.004110013	14
0.043320817	16
0.019078393	10
0.119078393	6
3.07735E-05	3
0.102411726	2
0.023840297	3
0.029334803	2

Step 4. The slope of the straight line that fits these ten points with minimum sum of squared errors and that also goes through the point (0,0) is $m = .00235$.

Step 5. The sum of squared errors when $T = 30$ is $J = .03187$.

Step 6. Varying T and repeating steps 1-5 yields the results in Table 5-3.

Table 5-3
The Effect of Varying Onset of Burnout (T) Upon Parameter Estimates and Goodness-of-fit (J)

T	hss	m	t2-T	J
26	0.03476	0.00295	11.78	0.03509
27	0.04954	0.00202	24.53	0.02949
28	0.04738	0.00226	21.01	0.03052
29	0.04738	0.00233	21.39	0.03118
30	0.04759	0.00235	20.24	0.03187
31	0.04759	0.00235	20.2	0.03276

The minimum sum of squared errors occurs for $T = 27$. However, the sum of squared errors varies very slowly with respect to changes in T over the interval (27, 31). Any of those values appear to be reasonable. An additional consideration is the slope of the burnout period. The fourth column of Table 5-3 identifies the doubling time, the time it takes for the steady-state hazard rate to double after onset of burnout. This is a parameter of the hazard function that some experts believe is an important characterization of the hazard rate. For the piecewise linear model, the doubling time is the ratio of the steady-state hazard rate to the slope of the burnout period. This value varies over the entire interval of T values. Indeed, there is relatively small variation with respect to changes in T over the interval (28, 31). This variation may motivate the choice of a value of T other than the minimizer $T = 27$. For illustrative purposes, we present the result for $T = 27$ in Figure 5-2.

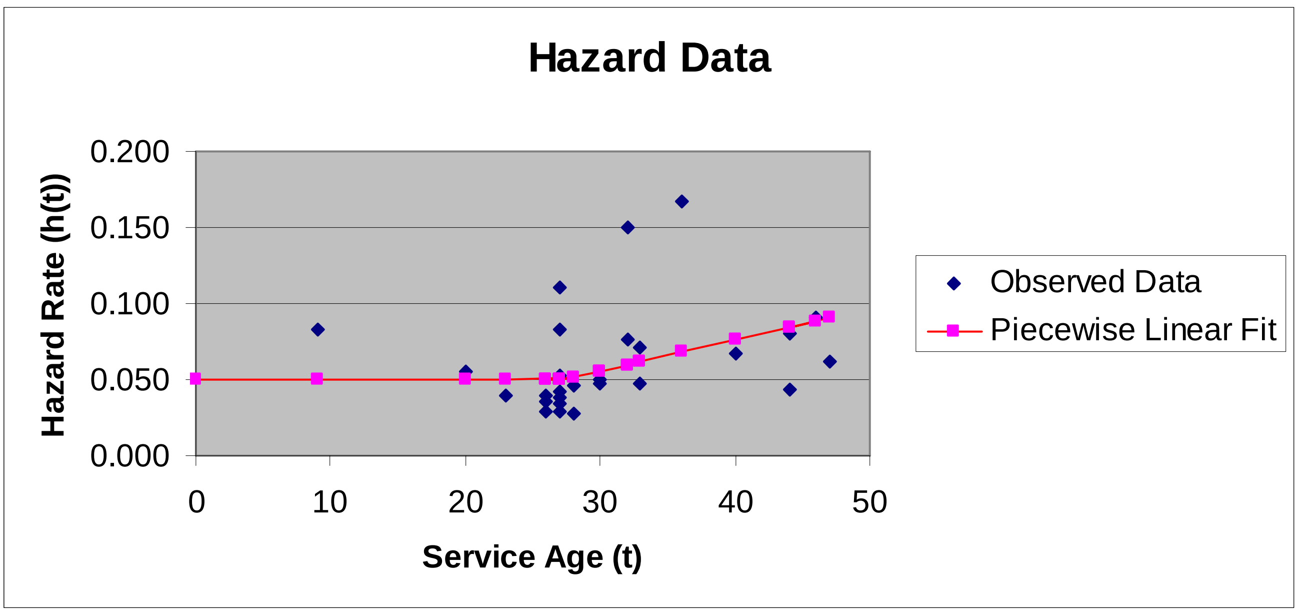

Figure 5-2
Piecewise Linear Hazard Function for *T*=27

Weibull Hazard Function

We will apply the procedure specified in Chapter 4 to fit a Weibull hazard function to the data in Table 5-1.

Step 0. This has been accomplished by creating Table 5-1.

Step 1. The logarithmic transformation of the data in Table 5-1 is shown in Table 5-4. The last two columns of the table contain the data that are used to fit the logarithm of the Weibull hazard.

Step 2. The estimates of the parameters are = 1.2294 and = 23.1682. The sum of squares of the errors of the fit is J = .032031, which is somewhat greater than the value of J for the piecewise linear hazard function at T=27. The Weibull hazard is shown in Figure 5-3.

Table 5-4
Logarithmic Transformation of Hazard Data

Year Installed	Year Failed	Number of Failures	Transformers Operating	Age (t)	Hazard Rate $h(t)$	Log $h(t)$	Log t
1951	1998	1	16	47	0.063	-2.772588722	3.850147602
1953	1997	2	25	44	0.080	-2.525728644	3.784189634
1953	1997	1	23	44	0.043	-3.135494216	3.784189634
1953	1999	2	22	46	0.091	-2.397895273	3.828641396
1960	2000	2	30	40	0.067	-2.708050201	3.688879454
1964	2000	1	6	36	0.167	-1.791759469	3.583518938
1965	1998	1	21	33	0.048	-3.044522438	3.496507561
1965	1997	3	20	32	0.150	-1.897119985	3.465735903
1966	1999	1	14	33	0.071	-2.63905733	3.496507561
1966	1998	1	13	32	0.077	-2.564949357	3.465735903
1969	1999	1	21	30	0.048	-3.044522438	3.401197382
1969	1999	1	20	30	0.050	-2.995732274	3.401197382
1969	1996	1	19	27	0.053	-2.944438979	3.295836866
1970	1997	1	29	27	0.034	-3.36729583	3.295836866
1970	1996	1	28	26	0.036	-3.33220451	3.258096538
1970	1997	3	27	27	0.111	-2.197224577	3.295836866
1970	1997	1	24	27	0.042	-3.17805383	3.295836866
1971	1999	1	36	28	0.028	-3.583518938	3.33220451
1971	1997	1	35	26	0.029	-3.555348061	3.258096538
1971	1998	1	34	27	0.029	-3.526360525	3.295836866
1972	1999	1	26	27	0.038	-3.258096538	3.295836866
1972	1998	1	25	26	0.040	-3.218875825	3.258096538
1972	1999	2	24	27	0.083	-2.48490665	3.295836866
1972	2000	1	22	28	0.045	-3.091042453	3.33220451
1979	2002	1	25	23	0.040	-3.218875825	3.135494216
1983	2003	1	18	20	0.056	-2.890371758	2.995732274
1995	2004	1	12	9	0.083	-2.48490665	2.197224577

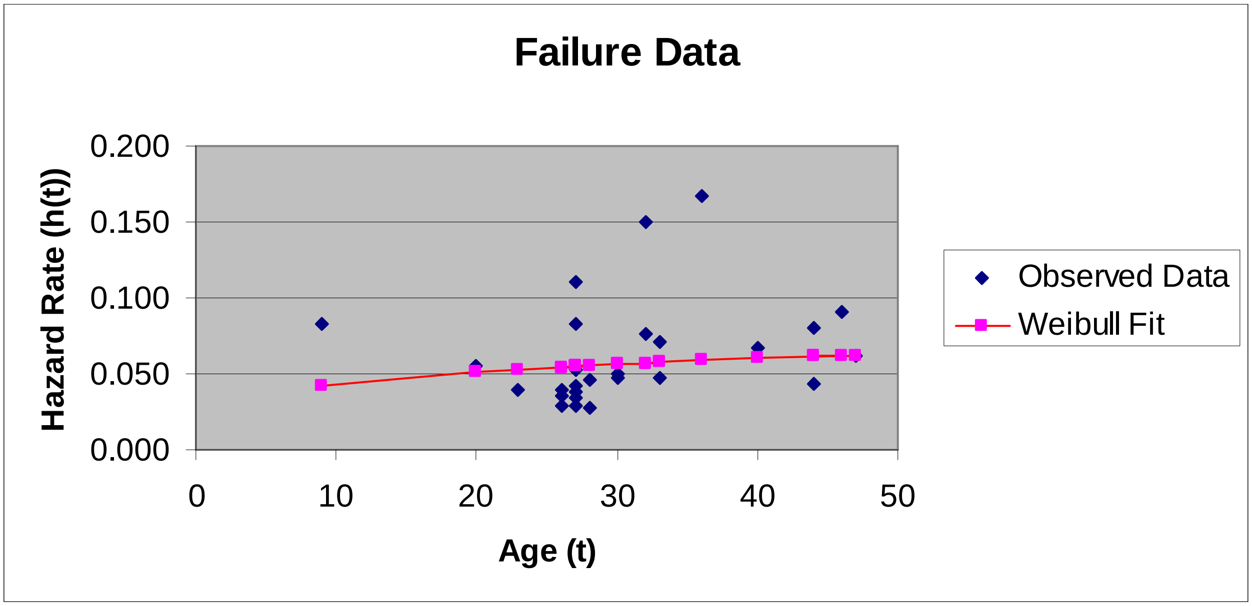

Figure 5-3
Weibull Hazard Function Based on Observed Data, Fit with Logarithmic Transformation

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 transformers. The lognormal seems inappropriate for transformers because of its interior maximum hazard. Therefore we propose to adopt either the piecewise linear hazard, the Weibull hazard, or the bathtub curve (Figure 4-2). The simplicity of the piecewise linear hazard and the intuitive content of the three parameters that determine it, motivated its use in the present implementation of the methodology to the exclusion of the others.

6 CONCLUSIONS

This report describes a methodology that can be used to specify the least-cost policy to manage an inventory of transformers. The input data requirements have been discussed, and default values are provided. We presently await user experience with the methodology, and we will modify the structure of the system as user experience suggests.

This report, which describes the methodology for asset management of substation transformers, has two companion reports:

EPRI report 1012503, *Equipment Failure Model and Data for Substation Transformers* (2006), provides the default settings of parameters that users will find in the present implementation of the methodology.

EPRI report 1012504, *Transformer Population Model with Testing* (2006), provides prototype software implementing the current version of the dynamic decision model formulated and discussed in this report. The Users Guide included in that report provides a series of screen shots to give the reader sufficient guidance to use the model.

Future work on this topic will focus on enriching the model of substation transformer dynamic behavior, refining the estimates of failure and deterioration rates, and gaining practical experience in applying the methodology in utility situations.

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI)

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, California, and Charlotte, North Carolina, was established in 1973 as an independent, nonprofit center for public interest energy and environmental research. EPRI brings together members, participants, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of electricity generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents nearly 15% of EPRI's total research, development, and demonstration program.

Together…Shaping the Future of Electricity

Printed on recycled paper in the United States of America

1012505

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com