Microelectronics Reliability 54 (2014) 349–353

Contents lists available at ScienceDirect

Microelectronics Reliability

journal homepage: www.elsevier.com/locate/microrel

Estimating activation energies for multi-mode failures

Dario Nappa *, Gergana I. Drandova

TriQuint Semiconductor, 500 W Renner Road, Richardson, TX 75080, United Sates

ARTICLE INFO

Article history:
Received 8 July 2013
Received in revised form 25 September 2013
Accepted 27 September 2013
Available online 17 November 2013

ABSTRACT

In Accelerated Life Time Modeling, the goal is to estimate the activation energy and failure time distribution. Existing methods assume data sets come from just one mechanism of failure. However, in many applications, more failure modes can be involved and few data are available; hence, we have to develop a method to identify the number of failure modes and assign observations to the appropriate failure mode. We developed a methodology based on Finite Mixture models and Bayesian Model selection to identify multiple failure modes. The approach provides the probability for each observation being associated with a certain failure mode, and provides good estimates for the activation energy of each mode.

1. Introduction

In Accelerated Life Time Modeling (ALTM), we investigate the Life Time (LT) distribution of a given product and how LT changes with a stress factor. Temperature is an example of such a stress factor and, in most cases, increasing the temperature results in an increase in the failure rate. In ALTM, the LT follows a certain distribution with parameters changing with temperature. For example, the failure time can follow a Weibull, Lognormal, or Gamma distribution. These distributions have two parameters and one of the parameters, the scale parameter, changes with temperature according to the Arrhenius law

$$\text{scale} = Ae^{\frac{E_a}{k_B T}}, \tag{1}$$

where k_B is Boltzmann's constant, T is temperature (stress factor), and E_a is the activation energy (higher activation energies correspond to failure mechanisms with stronger temperature dependence).

The activation energy can be estimated using Proportional Hazard Models [1,2] or the Maximum Likelihood Method (MLE) [3]. In this article we use the MLE method. Eq. (1) assumes that we have just one failure mechanism with one value for the activation energy. However, in many cases units' failures can be driven by two or more mechanisms. For example, one failure mechanism can be observed through device degradation (e.g. current drop, resistance increase, etc.), while another mechanism can be associated with an abrupt change in a parameter (e.g. sudden appearance of leakage or catastrophic burnout). For some units we observe failures (uncensored data) and in most cases we can assign a mechanism of failure. For other units we do not observe failures through the end of the test (censored data) or it is not easy to state what the mechanism of failure is.

Multi-failure mode situations are sometimes handled by reliability engineers by discarding the data for secondary failure modes. However, by removing observations, we decrease the accuracy and increase the bias which can lead to the erroneous life time distributions and activation energy estimates for the primary failure mode as well. Therefore, the ability to determine how many different failure mechanisms are involved and, furthermore, the ability to assign each unit to a mechanism of failure is extremely important. It is needed for correctly estimating life time distributions and activation energies, making life time predictions, knowing which failure mode to attack through process changes first and failure root cause analysis.

In the general case of multi-temperature accelerated life test data where multiple failure mechanisms operate and the failure times are possibly censored, we would like to know the answers to the following:

(1) How many possible mechanisms of failure (groups) exist?
(2) For each observation: what is the probability of being from a specific group?
(3) For each group: what is the best statistical distribution modeling the data (for example, Normal, Lognormal, or Weibull)?
(4) What are the estimates for the activation energy for each failure mode?

There are methodologies [4] and software packages (Minitab, Reliasoft) providing multimode reliability analysis with input from the user on failure mode for each sample. However, there is not a methodology which analyzes the data to determine the failure mode for each individual device and provide activation energies for all failure modes at once. In this work, we provide a methodol-

* Corresponding author. Tel.: +1 214 677 6908.
E-mail address: dario.nappa@gmail.com (D. Nappa).

0026-2714/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.microrel.2013.09.025

ogy for answering to the above questions. In Section 2, we provide the general theory about ALTM; in Section 3, we introduce the current approach for ALTM and in Section 4, we introduce the proposed model based on Finite Mixture Models. Finally, in Section 5, we provide examples with synthetic data.

2. Accelerated life testing models

ALTM relates [1,5,6] the lifetime distribution to explanatory variables E (temperature, pressure, voltage, etc.). The distribution can be defined by the survival function. Let us assume that the explanatory variable X is constant with time (constant stress) and we will define T_x as the failure time when $X = x$. The survival function is defined as

$$S_x(t) = P(T_x \geqslant t), \quad t > 0, \quad x \in E, \tag{2}$$

and provides the probability for a unit to survive for a time equal or greater than t when the stress factor has a value $X = x$.

A simple equation often used to relate how t changes with X is given [6] by

$$S_x(t) = S_0\{r(x) * t\}, \tag{3}$$

where S_0 is called the baseline survival function and does not depend on x. For any fixed x, the value $r(x)$ can be interpreted as the acceleration constant of the survival function or the time scale. Typically, the survival times are assumed to be from some class of parametric distributions (Weibull, Lognormal, etc.).

A parametric model assumes that life times' variation can follow a certain distribution whose scale changes with the stressing factor. For example, when assuming lifetime is changing according to a Weibull distribution, this means that

$$P(T = t \mid X = x) = \frac{k}{\tau}\left(\frac{t}{\tau}\right)^{k-1} e^{-(t/\tau)^k}, \tag{4}$$

where the scale parameter changes with X. For example, when X is the temperature it is assumed (Arrhenius model) that τ changes log-linearly with $1/X$ as shown in Eq. (1).

To estimate the unknown parameters various units are stressed at fixed values of X producing a data set with failure times $t_1, t_2, \ldots, t_n$ and stress conditions $x_1, x_2, \ldots, x_n$. For each unit we have a censoring indicator c_i assuming value of zero when the unit failed at time t_i and value of one when the unit was still working (censored observation).

To estimate the parameters, various approaches can be used [2]. In our investigation, we focus on a parametric approach by assuming that we know the distribution of life times. Then, for uncensored observations, we have a probability density function

$$P(i) = f(t_i \mid \theta_i), \tag{5}$$

and for censored observations we have a survival function $P(i) = 1 - F(t_i \mid \theta_i)$, where $F(t_i \mid \theta_i)$ is the cumulative distribution function.

In Eq. (5), we have to estimate the unknown parameters

$$\theta_i = \theta_i(x_i). \tag{6}$$

Given a dataset of n observations, we can estimate the parameters using the Maximum-Likelihood method [3]. This means maximizing

$$L = \prod_{i=1}^{n} f(t_i \mid \theta_i)^{1-c_i} * (1 - F(t_i \mid \theta_i))^{c_i}. \tag{7}$$

The above methodology is quite easy to implement and software is available. However, in some cases, the data set is more complex and we have to develop a new method. In the next section, we propose using Finite Mixture Models and Bayesian Model selection.

3. Finite mixtures modeling

In some applications in the engineering field, several failure mechanisms may be present. For example, some failures can be driven by defects, contamination, etc., while other failures can be driven by wear-out. Then, we would like to identify how many failure modes are present and find the failure distribution for each. The Finite Mixture Models approach [7] is an obvious choice for modeling multiple failure modes.

Finite Mixture Models (FMM) assume that each observation can be from multiple groups that are mutually exclusive. The probability of an observation can be expressed as

$$P\ (\text{fail time} \leqslant t) = \sum_{g=1}^{G} P(g) * P\ (\text{fail time} \leqslant t; g). \tag{8}$$

Eq. (8) tells us that the probability of a part failing by a given time t is the sum of the probabilities over all groups $g = 1 \cdots G$. For each group g, we have a probability $P(g)$ for the failure mechanism to occur and a probability $P(t; g)$ for having an observed failure time less than or equal to t. To fit the above model to the data, we have to estimate $P(g)$ and, for each group, we have to estimate the unknown parameters: $\theta = \{\theta_1, \theta_2, \ldots, \theta_G\}$.

In FMM we estimate $P(g)$ and θ using the MLE method. In most cases the Likelihood is very complex and we search for the MLE using the Expectation Maximization method [7]. The Expectation Maximization (EM) method uses an iterative procedure for computing the MLE for problems with incomplete data. In our case, the incomplete data is the lack of an assignment of each observation to a particular group. The name EM highlights the two steps performed in each iteration:

E-step: Given current estimates for $P(g)$ and θ, calculate the probability $P(t; g)$ of each observation being part of a group g.
M-step: Given $P(t; g)$, calculate $P(g)$ and θ such as to maximize a conditional likelihood.

For example, let us assume we investigate a mixture of two distributions and before step s we have initial estimates $\theta_{1,s-1}\theta_{2,s-1}$ for their parameters, and $P_{1,s-1}$ and $P_{2,s-1}$ for the groups' probabilities.

In the sth E-step we calculate the probability that an observation belongs to a group g using the formula

$$P_s(i; g) = \frac{P(i \mid \theta_{g,s-1}) * P_{g,s-1}}{P(i \mid \theta_{1,s-1}) * P_{1,s-1} + P(i \mid \theta_{2,s-1}) * P_{2,s-1}}$$
$$i = 1, 2, \ldots, n \quad g = 1, 2 \tag{9}$$

where $P(i \mid \theta_{g,s-1}) * P_{g,s-1}$ is the probability to have observation i from group g and using the parameters θ estimated in step $s - 1$.

The M-step requires [7] the maximization of the conditional expectation of the likelihood of the parameters θ given the data D. For our data modeling this translates into new estimates for group probabilities

$$P_{g,s} = \frac{1}{n}\sum_{i=1}^{n} P_s(i; g). \tag{10}$$

The parameters $\theta_{g,s-1}$ are estimated by solving the equations

$$\sum_{g=1}^{G}\sum_{i=1}^{n} P_k(i; g)\frac{\vartheta \log P(i \mid \theta_{g,s-1})}{\vartheta \theta_g} = 0. \tag{11}$$

The EM procedure can be used recursively up to when there is not much change in the estimated parameters. Ref. [7] provides information on how to verify when convergence is achieved.