A Better Way to Predict Reliability from Lifetest Data Using a Regression Technique

John Scarpulla, David C. Eng, and Denise L. Leung

TRW Inc., One Space Park Drive, Redondo Beach, CA 90278
john.scarpulla@trw.com

Abstract

In order to predict reliability of GaAs integrated circuits, the lognormal-Arrhenius model is often used. An extrapolation to ordinary usage temperatures is often justified with this model. However, lifetest data is often incomplete (censored), making a simple regression analysis difficult. A regression technique is introduced here for the analysis of lifetest data. It solves these difficulties, is rigorous and reproducible, and is easy to implement. It also handles the setting of confidence limits on lifetime at the desired probability level. Also addressed is the important problem of specifying the test conditions against which product wafers are to be screened.

INTRODUCTION

Many reliability mechanisms in GaAs technologies are thermally activated and follow an Arrhenius law, such as metal gate "sinking" or interdiffusion in HEMTs, and gain degradation in HBTs. The usual way to assess lifetime is to perform stressing experiments at three (or more) extremely high temperatures. The activation energy of the degradation mechanism is then established from the Arrhenius line, which is then extrapolated to actual usage temperatures to determine the expected lifetime. A JEDEC specification has even been written to cover the experimental details [1]. The specification fails to address some commonly encountered problems, however.

A typical Arrhenius plot resembles Fig. 1, where MTF's (mean times to failure) have been determined at each of three temperature (235 °C, 250 °C, and 265 °C). At the usage temperature of 90 °C the lifetime is extrapolated to be very long (2×10^{10} hours in this case) and all is well.

However, we have noticed that this extrapolated lifetime estimate is extremely sensitive to the mean failure times. Changing one of the three MTFs by only 10% can result in a 50X change in extrapolated lifetime! This is shown in Fig. 1 where the lower line results when the lowest temperature MTF is reduced by 10%. This has caused us to doubt conventional wisdom when interpreting lifetest data, and prompts the following questions: What is the statistical confidence and variability associated with the estimate? When extrapolating to temperatures far away from the test temperatures, it is intuitive that a statistical risk is incurred. How is this risk quantified? What is the procedure in case the sample sizes are different at different temperatures, or if the data is censored? What does one do if interested not in MTF but in the time to reach, say 0.1% failure? How does one design a screen for circuits or wafers to guarantee a certain reliability?

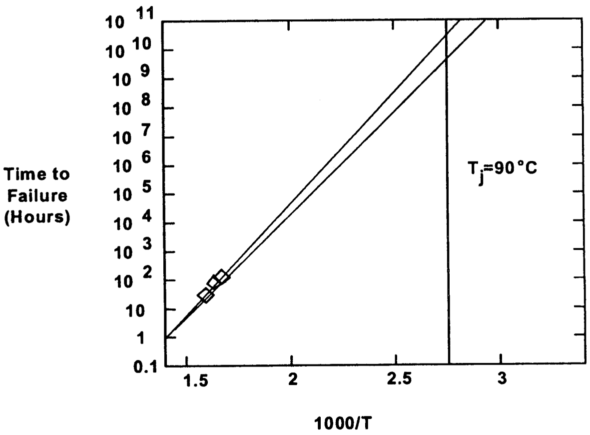

Fig. 1 Arrhenius line as generated using the MTF's per JEDEC specification [1] (upper line). If the lower temperature MTF reduces by only 10%, the predicted MTF at 90°C changes by 50 × (lower lines).

CONVENTIONAL REGRESSION TECHNIQUE

Before attempting to answer these questions, the standard methodology is reviewed. The standard regression technique for performing life analysis is given by Nelson [2]. Unfortunately, it is required that there be equal numbers of data points at each temperature, that all temperature experiments be complete (that is, 100% of the samples must experience failure), and that there be no left or right censoring of the data (no missing early or late failures, respectively). Unfortunately, these requirements are seldom met in practical life tests. Nor are these problems addressed by the JEDEC specification. To get around these difficulties, the hard-pressed engineer will simply find (or estimate) three MTF's, one for each temperature, and then perform a three-point regression fit (with only one degree of freedom) for the Arrhenius line as shown in Fig. 1.

NEW MULTIPLE-REGRESSION TECHNIQUE

The statistically pure and correct approach to dealing with real lifetest data is to perform an analysis using the maximum likelihood method [2], requiring specialized software and a nonlinear optimizer. Engineers, however, prefer regression-based methods, because the regression technique is simple, algebraic, and deterministic. All of the questions posed above are addressed by the multiple-regression technique presented here.

It is assumed that the physical model is Arrhenius, and the failure distributions are lognormal and homoscedastic (lognormal spreads are temperature-independent). This means that the following Arrhenius equation describes the mean time to failure, t_{50}:

$$t_{50} = A\exp\left(\frac{E_A}{kT}\right) \tag{1}$$

where A is the pre-exponential constant, E_A is the activation energy, k is Boltzmann's constant, and T is the absolute junction temperature. The actual times to failure are assumed to be distributed lognormally about this mean with lognormal shape parameter σ. The trick used here is to fit the lifetest data to the regression equation:

$$\ln(t_f) = a + \frac{E_A}{kT} + \sigma Z(P) \tag{2}$$

where t_f are the failure times, a is a constant equal to ln(A), P are the cumulative probabilities for each failure time, calculated using median ranks [3], and the function $Z(P)$ is the inverse normal cumulative distribution function (z-score for probability P). For P = 50% for the MTF, $Z(P) = 0$, and the above equation is the Arrhenius model, with the ln of the MTF linearly proportional to the inverse temperature. Making the variable transformations $y = \ln(t_f)$, $x_1 = 1/kT$, and $x_2 = Z(P)$, it is possible to use the standard multiple linear regression form in two variables

$$y = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon \tag{3}$$

where ε is the error term. The determination of the regression parameters is straightforward and is discussed in many textbooks [4, for example].

In this way, all the data collected in the lifetest experiment participates in the fit, not just the three means. Furthermore, any number of samples and combinations of temperatures can be analyzed, and incomplete or censored data is easily handled. The multiple regression gives the needed constants $b_0 = \ln(A)$, $b_1 = E_A$ and $b_2 = \sigma$.

As an example, Fig. 2 shows the fit for the mean time to failure, t_{50} using this multiple regression technique. The data points (right censored) are from a process qualification of TRW's 2 µm HBT process described elsewhere [5]. The failure times are defined when the gain degradation of the amplifier used as the SEC ("standard evaluation circuit") exceeds 1.0 dB. The constants from the fit were: pre-exponential constant $A = 6.017 \times 10^{-8}$ hours, activation energy E_A = 1.17 eV, and lognormal shape parameter σ = 0.466. The residuals of the fit are also available from the regression analysis, and aid in assessments of whether the assumed Arrenhius lognormal model is reasonable [4].

The time to reach the failure percentage, P is

$$t_P(T) = \exp[y_P(T)] = A\exp\left(\frac{E_A}{kT} + \sigma Z(P)\right) \tag{4}$$

where, for example, to find the time to reach 0.1% failures, the cumulative standard normal distribution function $Z(0.1\%) = -3.09$.

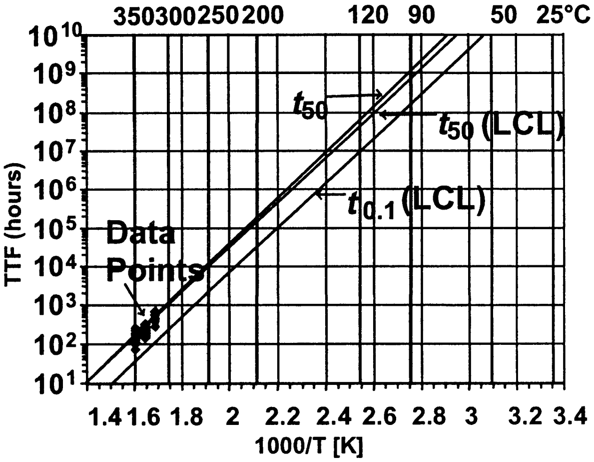

Fig. 2 Example of a regression fit using a set of right-censored three temperature lifetest data. Also shown are the lower confidence limits (LCL) for the MTF and for the time to 0.1% failure. The data is from a 2µm HBT process [5].

CONFIDENCE LIMITS

Associated with any statistical estimates such as those just described for t_{50} and $t_{0.1}$, is a degree of uncertainty or randomness. This uncertainty decreases with larger sample sizes. The uncertainty increases if the regression line is extrapolated far away from the data points and more risk is incurred. Unfortunately, the Arrhenius line is always extrapolated down to the usage temperatures of the GaAs devices which are far from the lifetest temperatures. The degree of risk or uncertainty in this extrapolation is quantified by providing confidence limits.

The lower confidence limit for lifetime (at any desired percentile P, and confidence level γ) is a hyperbola given by

$$t_{LCL,P,\gamma} = \exp\left(y_{LCL,P,\gamma}\right)$$

where (5)

$$y_{LCL,P,\gamma} = y_P - Z(\gamma)\sigma\sqrt{\frac{[Z(P)]^2}{2v} + S}$$

and, S is the sum of the squared errors [4]. In this expression, y_P refers to eq. (4) for the desired percentile, $Z(\gamma)$ and $Z(P)$ are the z-scores for the cumulative normal distribution at γ and P respectively, and v is the degrees of freedom (number of lifetest failures less three). Equation (5) resembles closely the confidence limit proposed by Nelson [2], which is an approximation. Figure 2 shows the 95% lower confidence limit on the MTF , and on the time to reach 0.1% failures. The lower confidence limit quantifies just how much uncertainty lies in the extrapolations to usage temperatures.

SCREENING FOR RELIABILITY

Suppose now that a multiple temperature lifetest has been performed, and a certain baseline reliability has been established. The question now arises as to just how to compare a new lot or a new wafer against this existing reliability data. More importantly, it is desired to design a one temperature stress test for a fixed time to determine whether the lifetime of the present wafer can be expected to be sufficiently good compared to the previous baseline. (At TRW this screen is referred to as a WALT, or wafer acceptance life test.) The screen consists of sampling test devices or circuits from a wafer and stressing them at elevated times and temperatures. We would like this screening test to have as small a sample size as possible and still give an indication of whether the wafer has good reliability comparable to the previous baseline.

To do this, let there be m samples submitted to the screen allowing up to c failures. From binomial theory, the confidence γ for a reliability level of at least $(1 - P_f)$ is given by the cumulative binomial distribution:

$$1-\gamma = \sum_{x=0}^{c} \frac{m!}{(m-x)!\,x!} P_f^x \left(1-P_f\right)^{m-x} \qquad (6)$$

where P_f is the probability of failure. For example if the sample size is $m = 5$, with $c = 1$ failure allowed, then for a confidence of 95%, the reliability is 34% (probability of failure is 66%). This means that having done this screen, it is possible to state with 95% confidence that the probability of failure is less than 66% (or the reliability is greater than 34%). This is, in fact, a poor reliability. However, because the test is at a very highly accelerated stress level, and because the baseline process regression parameters are known, this may well translate into excellent reliability at actual usage conditions. Table 1 shows some other failure probabilities associated with different sampling plans.

	Number of allowed failures, c		Number of allowed failures, c	
	0	1	0	1
Sample size, m	P_f, probability of failure with 95% confidence is at most:		P_W, underlying wafer failure probability of screen to reject P_S = 2% of good wafers is at most:	
5	45%	66%	0.41%	4.7%
10	26%	40%	0.2%	2.2%
20	14%	22%	0.1%	1.1%

Table 1. Possible sampling plans for screening wafers.

In every sampling plan, there is a tradeoff between the "consumer's risk" and the "producer's risk". This means that if the sampling plan is too strict, the producer takes the risk of scrapping wafers and losing revenue. On the other hand, if the sampling plan is too lenient, the consumer takes the risk of accepting unreliable wafers. The two risks must be balanced. We have designed our screens a-priori that will reject 1/50 good wafers or 2%. The underlying wafer failure rate corresponding to a 2% screen rejection rate is again calculated from the cumulative binomial distribution as 4.7% for this example. The right side of Table 1 shows this probability for other sampling plans.

RELIABILITY AND FIT LEVELS

It is now possible to calculate the reliability of wafers subjected to this screen. We continue with the example sampling plan $m = 5$ and $c = 1$. First, the probability of failure as a result of performing the screen is equal to or less than $P_f = 66\%$ with $\gamma = 95\%$ confidence. It is possible to translate this down to usage conditions. The way to do this is to establish a temperature acceleration factor in going from the selected screening temperature T_s to the usage temperature T_u. This acceleration factor is

$$AF = \exp\left[y_{LCL,Pw,\gamma}(T_u) - y_{LCL,Pw,\gamma}(T_s) \right] \quad (7)$$

and takes into account the statistical uncertainty in the extrapolation to usage temperatures. Note the use of the underlying wafer failure probability P_w here, although the screen can only determine that the failure probability is less than P_f = 66% with γ = 95% confidence. The acceleration factor has the following interpretation: 1 hour at the stress temperature is equivalent to AF hours at the usage temperature. The probability of failure under usage conditions is then given by

$$P_u = \Phi\left(Z\left(P_f\right) - \frac{\ln\left(\frac{AF \times t_s}{t_u}\right)}{\sigma} \right) \quad (8)$$

where Φ is the cumulative normal distribution function. Eq. (8) is depicted in Fig. 3. In this way, the probability of failure of all screened wafers is guaranteed to be at most P_u.

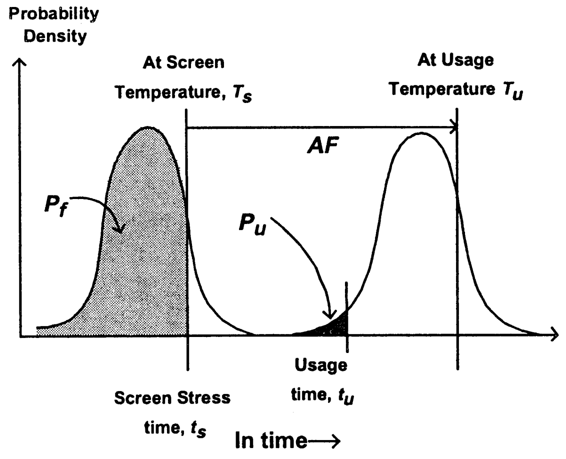

Fig. 3 Illustration of the probability calculation used to translate from the screen (stress) temperature to the usage temperature. The distributions are lognormal, and the shaded areas represent the failure probabilities.

To relate this probability P_u to the average failure rate (expressed in FITs), we use the relationship [3]:

$$AFR = \frac{-\ln\left(1 - P_u\right)}{t_u} \times 10^9 \quad (9)$$

where the usage time t_u is in hours. The AFR is a way to determine an approximate constant failure rate, useful in reliability simulations that use exponential failure distributions with constant failure rates. The MIL-HDBK 217 method [6] is one example of this approach.

A proposed screen was calculated based upon the three temperature lifetest qualification data for our 2 μm HBT process example (Fig. 2). The proposed screen for each new wafer is:

- sample 5 devices per wafer (m = 5)
- allow 0 or 1 failures (c = 1)
- allow 2% reject rate for good wafers
- failure definition: 1.0 dB gain degradation
- stress time, t_s = 250 hours
- stress junction temperature, T_s = 320 °C

The resulting AFR for the wafer screen as proposed here is a guaranteed 0.818 FITs with 95% confidence for a usage or mission time of 25 years, and an operating temperature of 150 °C. The reliability at the usage conditions is at least $1 - P_u$ = 99.982% with 95% confidence.

CONCLUSION

A method for fitting lifetest data has been presented which uses a multiple regression technique on incomplete data. The life data may be right or left censored. The technique provides the parameters A, E_A, and σ which are the basis for reliability predictions. The design of screen to guarantee reliability was described using the results of the regression technique.

REFERENCES

[1] JEDEC "Guidelines for GaAs MMIC and FET Life Testing" JEP118, Electronic Industries Assoc., Wash. DC, 1993.

[2] Wayne Nelson Accelerated Testing Wiley: New York, 1990.

[3] P.A. Tobias and D.C. Trindade Applied Reliability (2nd ed.) Van Nostrand Reinhold: New York, 1995.

[4] J. Neter, M.H. Kutner, C. Nachtsheim, and W. Wasserman Applied Linear Statistical Models (4th ed.) McGraw-Hill: New York, 1996.

[5] F.M. Yamada, et. al "High-Reliability GaAs HBT Monolithic Microwave Amplifier", IEEE RFIC Symposium, Digest of Tech. Papers, Denver, CO June 8-11, 1997, p. 193.

[6] MIL-HDBK-217, Rev. F, Notice 2, U.S. Dept. of Defense, Wash. DC, Feb. 1995.