Extensive Investigation of Calibrated Accelerated Life Testing (CALT) in Comparison with Classical Accelerated Life Testing (ALT)

B. Sal, M. Altun
Istanbul Technical University, Istanbul, Turkey

ABSTRACT: Various accelerated reliability test methodologies have been recently proposed with an aim of reducing sample sizes. Dramatic decrease in failure rates, for electronic products targeted in this study, is the main motivation for these methodologies; performing conventional accelerated life tests (ALT) is getting extremely time consuming and costly. Calibrated accelerated life testing (CALT) is one of the well accepted methodologies that aims to use fewer samples than those used in ALT. We investigate ALT and CALT with different failure rates. We show that as opposed to the conventional view, considering required test time CALT does not always overwhelm ALT; it highly depends on failure rates, acceleration factors, and stress levels. . We also compare these methodologies by considering accuracy that is the main drawback of using CALT. In this study, we define an accurate test method choice between CALT and ALT by failure rates, acceleration factor, warranty times and mean-time-to-failures.

1 INTRODUCTION

In reliability, various test methodologies are used due to finding failure rates by testing products. Accelerated Life Testing (ALT) and Calibrated Accelerated Life Testing (CALT) are mainly used test methods. Also Highly Accelerated Life Testing (HALT) is used before these testing methods to determine absurd stress levels which can be used for both test methodologies (Paschkewitz, 2009, Donahoe et al. 2008).

In controversial ALT method, various stress profiles (temperatures, voltages, pressures, vibrations etc.) under the absurd stress levels, are applied to product with effect of acceleration factor with short amount of time to determine potential failure modes that happened generally its warranty time (Donahoe et al. 2008). Additionally, stress profiles are chosen by using accelerated stress models (Arrhenius, Peck etc.) (O'Connor & Kleyner, 2011). Sample size is defined in respect of complicated calculations (Guo et al. 2013).

CALT is firstly introduced as a solution for controversial ALT method's high amount of sample size and also has similar mechanism with ALT. In CALT, there are three stress profiles which defines products life time by using stress line. First stress level is selected considering %10-15 below absurd stress level which is defined by the help of HALT. After that, second stress level is chosen %10-%15 below the first stress level. Moreover, with these two points first stress-life plot can be plotted. By the help of plotted stress-life plot, third stress level is identified. Remaining time is another factor to determine third stress level. Each stress level contains two samples (Paschkewitz, 2009).

Although CALT uses much fewer samples than classic ALT, it is accuracy is controversial and we investigate its accuracy for the first time in literature.

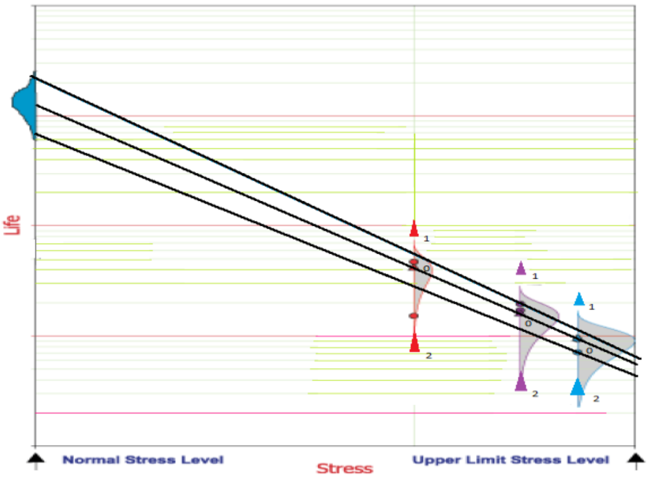

Figure 1: CALT process, Life-Stress Plot's Changing (Paschkewitz, 2009)

CALT's life-stress plot can be drawn in many ways, according to products failures in different stress levels. Figure 1 shows that, if 0 point is estimated plot, there can be another points like 0-1 and 0-2 that can change the whole stress-life plot. So that, two samples in every stress level does not seem accurate in this part. Besides, in CALT method, whole mechanism has to be run until it fails. For example, for %0.1 failure rate which is defined for electronic products, mean-time-to-failure could be as 1000 years or more. So that, it takes so much time until it fails even with effect of acceleration factor.
In the Section 2.2, we compare ALT and CALT for different failure rates, warranty times, acceleration factors and accuracy. We compare result for estimated cases and intensify with problem in Section 2.3 and Section 2.4 respectively. For instance, with defined failure rate and estimated accuracy, we define smallest sample size and minimum total test time between CALT and ALT.

2 COMPARISON OF ALT AND CALT

In this section, there is an overview of performance parameters that affects CALT and ALT. Furthermore, for both test methodologies different failure rates, acceleration factors, warranty times and mean-time-to-failures are applied. Elements which are used for comparison, are sometimes calculated by proven formulas and sometimes are chosen according to studies (Guo et al. 2013, Reliasoft Corp., 2012).

2.1 *OVERVIEW OF PERFORMANCE PARAMETERS*

2.1.1 *BOUNDS RATIO*

Bounds ratio is defined as ratio between two bounds of estimated two-sided B_x life. By the help of bounds ratio, accuracy can be determined (Guo et al. 2013).
As mentioned before, we defined a bounds ratio for ALT. However, bounds ratio may be calculated with defined sample size and confidence level (Reliasoft Corp., 2012). Bounds ratio for six samples on CALT is higher than that defined for ALT. Since few sample size cause higher bounds ratio (Guo et al. 2013). Additionally, high bounds ratio lead to uncertainty for test results.
We use MATLAB code 'wblrnd' to define bounds ratio for both test methods by defining their eta, beta and bunds ratio values. Moreover, in Figure 2 and Figure 3, we show that, for six sample size, bounds ratio is very high.

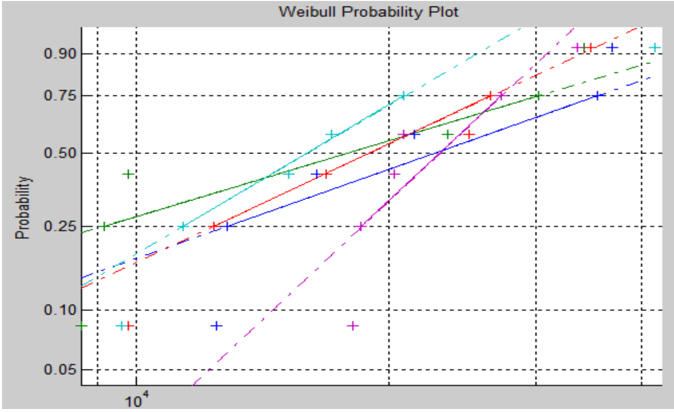

Figure 2: Bounds ratio for 6 sample size, CALT

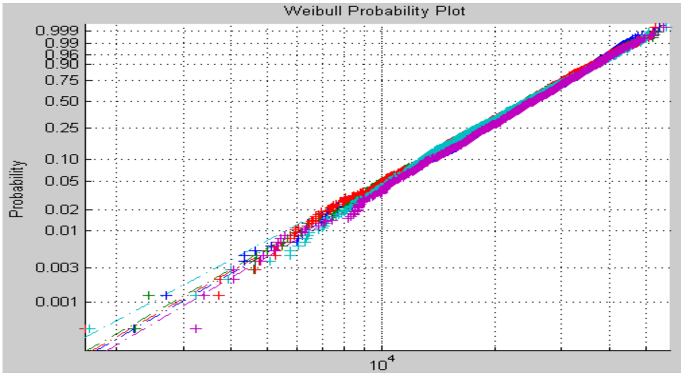

Figure 3: Bounds ratio for 1211 sample size, ALT

We define same bounds ratio for both test methodologies and appoint samples for CALT on bounds ratio two by two using MATLAB code, 'normrnd'. We repeat same assignment a hundred times to make sure its accuracy. Also, values remain outside of the defined ratio are considered false.

2.1.2 *CONFIDENCE INTERVAL AND CONFIDENCE LEVEL*

Confidence interval is a term that define estimated interval for population (Cox & Hinkley, 1974). Confidence level is, if there are different and iterated test which leads to confidence interval construction between these tests, the ratio among intervals contains true values of parameters is equal to confidence level. Furthermore it is absolute with confidence intervals (Kendall & Stuart, 1973). For example, with known population and confidence interval for each population, %90 confidence level says that, true population is underlying in %90 confidence interval.

2.1.3 *UNIT AND TOTAL TEST TIMES*

Unit test time defines, the test time for 1 unit that includes parameter calculation for ALT and parameter estimation for CALT (Guo et al. 2013, Paschkewitz,

2009). We have considered 10 hours of working of a product for 1 week and we multiply it by 52 for 1 year working time also we multiply 1 year test time by warranty time in ALT and mean-time-to-failure for CALT. Additionally, we divide it by acceleration factor.
For total test time, we multiply unit test time by known sample size for each methods.

2.1.4 *MEAN-TIME-TO-FAILURE ACCORDING TO FAILURE RATE*

Mean-time-to-failure according to failure rate calculations has made by equation 1/λ=MTTF where λ is failure rate (O'Connor & Kleyner, 2011). This means, we need to define mean-time-to-failure according to failure rate instead of defining same mean-time-to-failure for all failure rates. So that, we have shown that mean-time-to-failure for lowest failure rate. However, we used mean-time-to-failure by failure rate values for each case.

2.2 *CALCULATION OF PARAMETERS*

2.2.1 *ALT*

ALT method has formulas for different factors of test. Additionally, these formulas are proven by comprehensive studies (O'Connor & Kleyner, 2011). So that our equations will be;

1. Reliability, R(t)= exp(-WT/MTTF)
   a. $R(t)= \exp(-t/n)^{\beta}$
   b. β= Beta, n= Eta, t= WT (hours)
   c. WT= Warranty Time, MTTF= Mean-Time-to-Failure
2. $AF=\exp[(Ea/k)*(1/T_{field}-1/T_{test})]$
   a. E_a= Activation Energy, T_{field}= Field Temperature, T_{test}= Test Temperature, AF= Acceleration Factor.
   b. $n_1 /AF= n_2$
   c. $P_1= 1-\exp(-t/n_1)^{\beta}$
   d. $P_2= 1-\exp(-t/n_2)^{\beta}$
   e. P_1 and P_2 Probability of Failure values, t= Estimated Time (hours), k= Boltzmann constant.
3. Bounds Ratio= Upper Limit/Lower Limit
   a. $\ln T_p + z*std(\ln T_p)$= Upper Limit
   b. $\ln T_p - z*std(\ln T_p)$= Lower Limit
   c. T_p= Standard Deviation, z= Normal Distribution Parameter.
4. Sample Size= $(z*A*BR)^c$
   a. A= Average variance coefficient, z= Normal Distribution Parameter, c= Distribution Parameter. Average Coefficient of Variance Theorem (Meeker & Escobar, 1998)
5. Unit Time= (YT*WT)/AF
   a. YT= Yearly Working Time (hours)
6. Total Time= Sample Size × Unit Time

2.2.2 *CALT*

Recommended sample size for CALT is 6 (Paschkewitz, 2009). For time for one unit and total time calculations;

1. Sample Size= 6
   a. Recommended sample size for CALT is 6, however, sample size can be increased in order to increase accuracy.
2. Unit Time= (YT*MTTF)/AF
   a. MTTF= Mean-Time-to-Failure
3. Total Time= Sample Size × Unit Time

These calculations in Section 2.2.1 and 2.2.2 are used for every defined values. However, in CALT method six samples have to be used for every failure rates (Guo et al. 2013). Other defined values are calculated by these formulas.

2.3 *COMPARISON OF RESULTS*

We compared two test methodologies by previous given formulas with different failure rates, acceleration factors, warranty times and mean-time-to-failures.

Firstly we choose %10, %1 and %0.1 failure rates for each methods. Other values defined as;

- Mean-Time-to-Failure: 30 years
- Acceleration Factor: 10
- Warranty Time: 3 years
- Bounds Ratio~5

Table 1. %10 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	96	~%100	155.28	14,906.88
CALT	6	~%97	1465.104	9282.6019

Table 2. %1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	1211	~%100	155.28	188,044.08
CALT	6	~%50	14,651.04	92,826.019

Table 3. %0.1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	15,554	~%100	155.28	2,415,225.112
CALT	6	~%14	14,6510.4	928,260.19

We change warranty time to one year to define effects on each test methods.

- Mean-Time-to-Failure: 30 years
- Acceleration Factor: 10
- Warranty Time: 1 year
- Bounds Ratio~5

Table 4. %10 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	96	~%100	51.76	4,968.96
CALT	6	~%97	1465.104	9282.6019

Table 5. %1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	1211	~%100	51.76	62,681.36
CALT	6	~%50	14,651.04	92,826.019

Table 6. %0.1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	15,554	~%100	51.76	805,075.04
CALT	6	~%14	146,510.4	928,260.19

We also change mean-time-to-failure from 30 years to 10 years. Furthermore, that changing affects not only parameters but mean-time-to-failure by failure rate values.

- Mean-Time-to-Failure: 10 years
- Acceleration Factor: 10
- Warranty Time: 3 years
- Bounds Ratio~5

Table 7. %10 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	96	~%100	155.28	14,906.88
CALT	6	~%97	488.368	2930.208

Table 8. %1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	1211	~%100	155.28	188,044.08
CALT	6	~%50	4883.68	29,302.08

Table 9. %0.1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	15,554	~%100	155.28	2,415,225.112
CALT	6	~%14	48,836.8	293,020.8

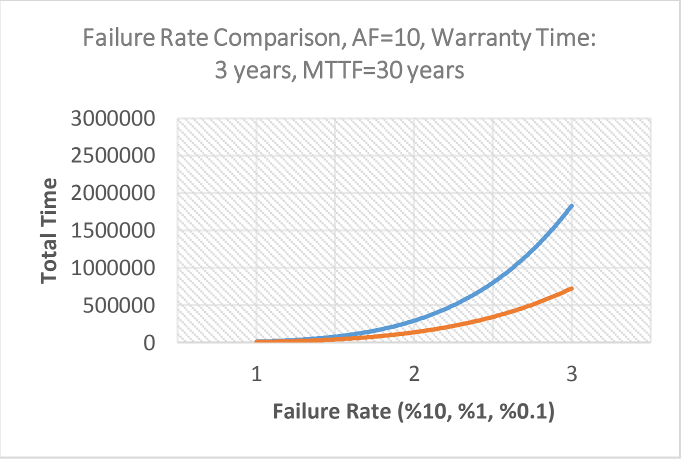

Figure 4: AF=10, WT=3 years, MTTF=30 years

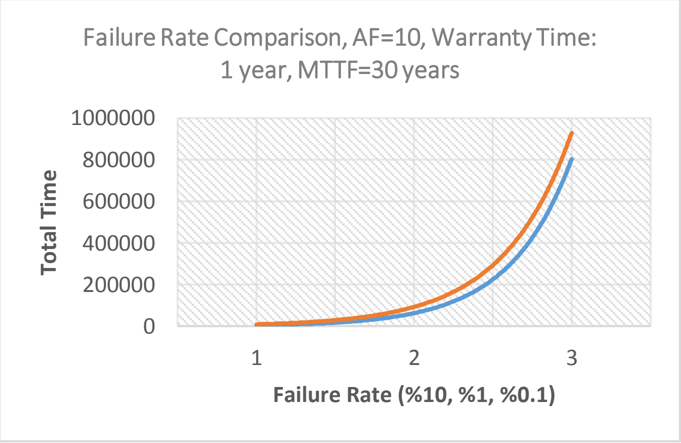

Figure 5: AF=10, WT=1 year, MTTF=30 years

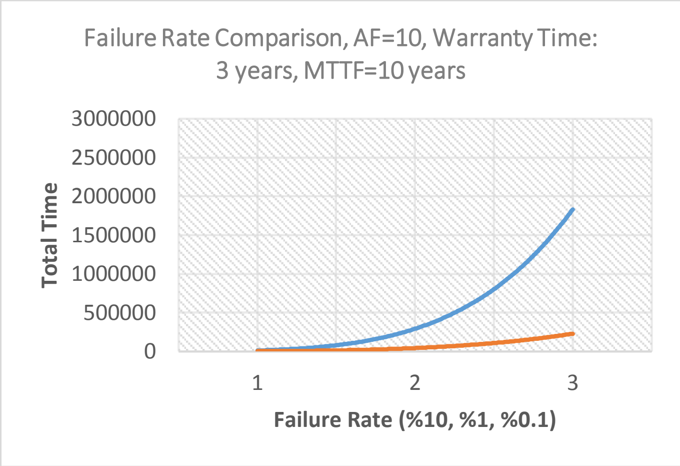

Figure 6: AF=10, WT=3 years, MTTF=10 years

It is clear to say that, failure rate increment increases total time of accelerated tests but time for one sample remains the same. For defining bounds ratio for each methods there is an accuracy effect. As mentioned in Section 2.2, bounds ratio requirement for lower failure rates for CALT is using much more samples. Even though CALT uses six samples for every failure rates accuracy decrease by failure reduction. As a result CALT has a positive effect on sample size and total time as well as it has negative effect on accuracy which can be bad for reliability for product.

Changing of warranty time is a factor that reduces total time for ALT but not for CALT since CALT uses mean-time-to-failure for its total time calculations. However, changing of mean-time-to-failure has a direct effect on CALT by reducing time for 1 sample as well as reducing total test time.

We define different values for acceleration factor to see its effect on both methods with sample size, accuracy and total test time. Also, we use different acceleration factor on different failure rates as we define before. Determined acceleration factor values for each test as;

- Mean-Time-to-Failure: 30 years
- Acceleration Factor: 20
- Warranty Time: 3 years
- Bounds Ratio~5

Table 10. %10 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	28	~%100	77.64	2173.92
CALT	6	~%100	732.552	4395.312

Table 11. %1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	234	~%100	77.64	18,167.76
CALT	6	~%97	7325.52	43,953.12

Table 12. %0.1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	2676	~%100	77.64	207,764.64
CALT	6	~%50	73,255.2	439,531.2

- Mean-Time-to-Failure: 30 years
- Acceleration Factor: 30
- Warranty Time: 3 years
- Bounds Ratio~5

Table 13. %10 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	20	~%100	51.76	1035.2
CALT	6	~%100	488.368	2930.208

Table 14. %1 Failure Rate Comparison

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	96	~%100	51.76	4968.96
CALT	6	~%98	4883.68	29302.08

Table 15. %0.1 Failure Rate,

Test Method	Sample Size	Accuracy	Time for 1 sample	Total Time
ALT	956	~%100	51.76	49,482.56
CALT	6	~%72	48,836.8	293,020.8

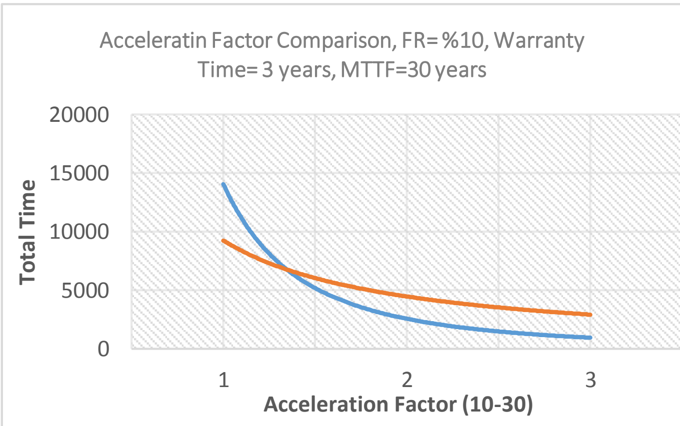

Figure 7: FR=%10, WT=3 years, MTTF=30 years

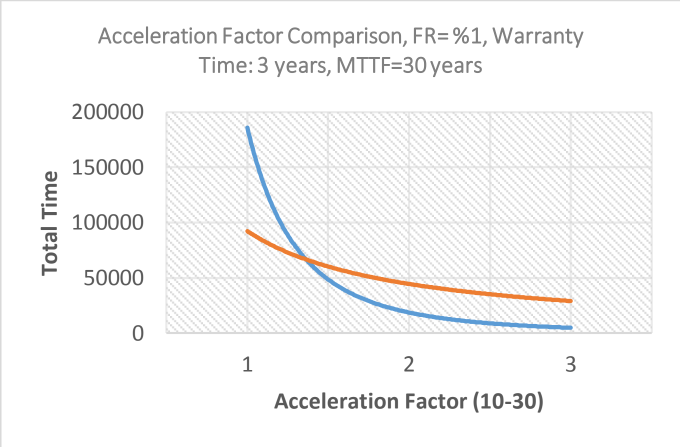

Figure 8: FR=%1, WT=3 years, MTTF=30 years

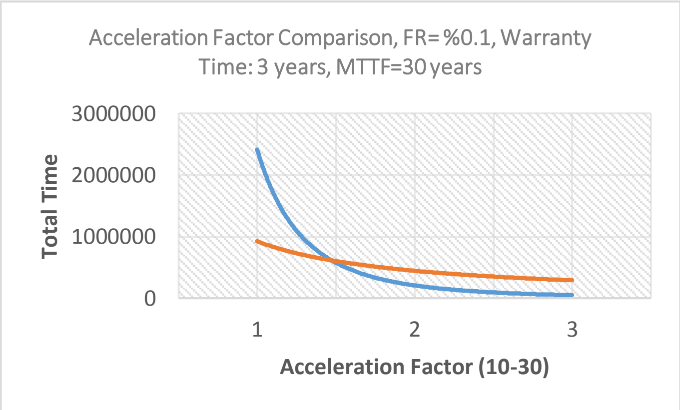

Figure 9: FR=%0.1, WT=3 years, MTTF=30 years

We compared acceleration factors by failure rates. For each defined acceleration factor values there are three failure rates which are %10, %1 and %0.10. Besides, we did not change the warranty time and mean-time-to-failure values to see only acceleration factor effect during the comparison.

With the increment of acceleration factor in both methodologies unit test times decreased because detractive effect of acceleration factor. However, in ALT method sample size did not remain same. Since, as mentioned Section 2.2.1, probability of failures which affect sample size with Average Coefficient of Varience theorem, has an ifluence.

In CALT method, unit and test times decrease while sample size remains same because of recommended sample size in this method. Moreover, CALT's accuracy is increasing since acceleration factor affects standard deviation as well as distribution of probability of failure.

We define 2 different problems for showing ALT and CALT choice according to wanted values of some product.

2.3.1 *3D GRAPHS*

We determine a formula with total time, acceleration factor and failure rate for both methodologies to take 3D graphs about them. Furthermore, we used total time calculations that mentioned before. Mean-Time-to-Failure values for CALT calculations and Warranty Time values for ALT calculations. We preferred to show total time as a colored z-axis and as can be seen from the graphs, x-axis for acceleration factor and y-axis for warranty time/mean-time-to-failure.

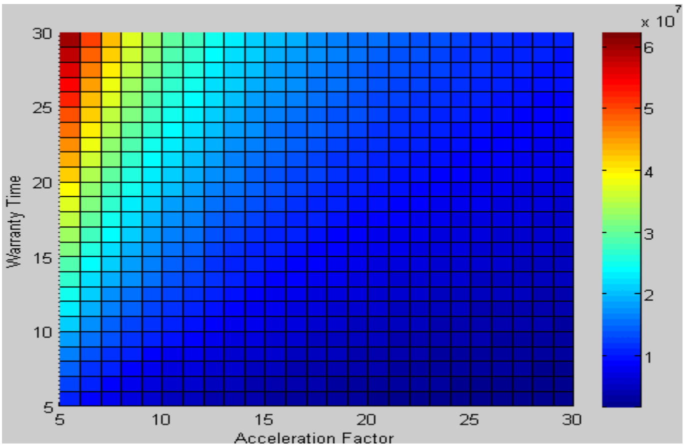

Figure 10: ALT 3D Graph (AF=5-30, WT=5-30)

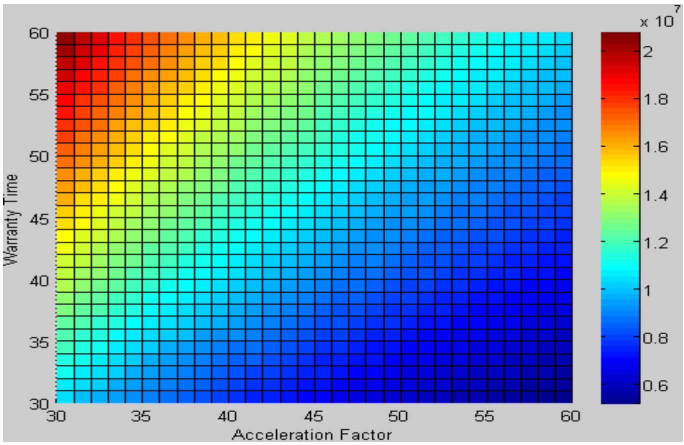

Figure 11: ALT 3D Graph (AF=30-60, WT=30-60)

We used warranty time values from 5 year to 60 years. However warranty time values above 30 years and acceleration factor values above 40 is not accurate nor realistic.

These graphs show a simple explanation of our equation of total test time. Since, with minimum acceleration factor and maximum warranty time, we get very high points that covers all the points, we divide graphs into two groups with AF=5-30, WT=5-30 and AF=30-60, WT=30-60.

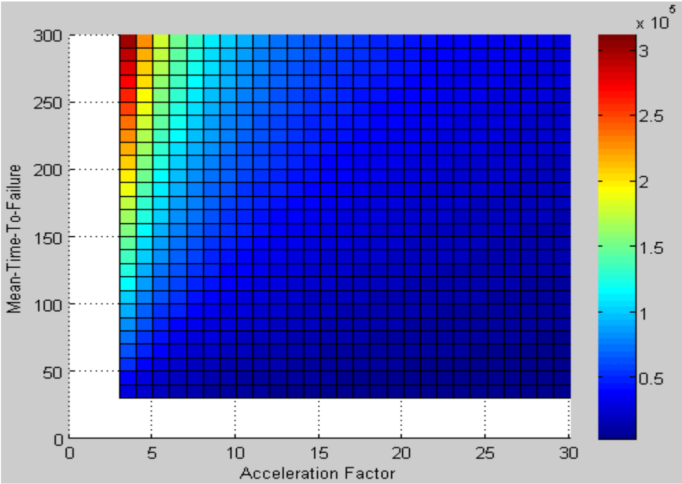

Figure 12: CALT 3D Graph (AF=3-30, MTTF by failure rate=30-300)

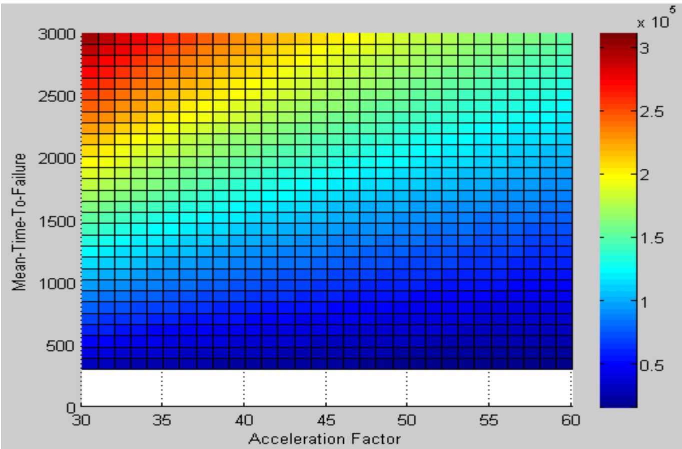

Figure 13: CALT 3D Graph (AF=30-60, MTTF by failure rate=300-3000)

We used mean-time-to-failure by failure rate values from 30 to 3000 years. Mean-time-to-failure values below 30 years are not seem accurate, nevertheless. We have used these values just for calculations.
Since, with minimum acceleration factor and maximum warranty time, we get very high points that covers all the points, we divide graphs into two groups with AF=5-30, WT=5-30 and AF=30-60, WT=30-60.
Calculations on CALT method's total test time are similar to ALT since both used same equations for unit test time. However CALT seems that takes less time than ALT because of sample size is 6. However, graphs must be interpreted from values that we found. By the help of those values it can be understood.

2.4 *PROBLEM*

We define a case for which accelerated test methodology will be used according to failure rate and accuracy.
Example 1: A reliability engineer wants to do accelerated test on a product. She knows the beta parameter of the Weibull distribution that 3. She also knows that her product's mean-time-to failure is 30 years. From the previous tests acceleration factor is known as 15. Also determined warranty time for product is 3 years. She wants %1 failure rate and more than %90 accuracy for test. As a result, she wants to define;

- Required minimum sample size,
- Required minimum test time,
- Accelerated test methodology for her test.

Solution for Example 1: In here, we can use identified formulas and estimated values from Section 2.2.1 and 2.2.2 also we can derive a formula with acceleration factor and total test time by the help of given and estimated formulas. So that minimum and suggested sample size for CALT is n=6.
If we can use those formulas to find sample size in ALT. Our minimum sample size for this test will be;
n= 662, as a result CALT has required minimum sample size. However, CALT's accuracy for desired values is %50 and it does not fit for engineer's design. To fulfill the requirements CALT's sample size should be at least 25 for %90 accuracy. So minimum sample size will be 25 in CALT.
Required minimum test time associates with sample size. Thus, with the help of 3D graphs in Section 2.3.1 we can see for 3 years of warranty time, 30 years of mean-time-to-failure and acceleration factor value of 15 our total test time value will be so much below from 2 × 10^7 for ALT and seems double times of 1 × 10^5 for CALT. With the all necessary calculations and equations for ALT method our total time will be;

TT=(4 × 10^8) × (AF^-3.314),
TT=50,640 hours, as can be seen its accuracy from the graph for ALT.

Nevertheless, CALT's formula has to multiply with 25/6 since the formulas for 6 sample size in CALT and for CALT our total test time it will be as;

TT= (25/6) × (10^6) × $(AF)^{-1.053}$
TT= 240,638 hours.

As a conclusion, even though CALT has much fewer sample size than ALT for desired accuracy value, ALT takes less time than CALT for total test time. Therefore, ALT methodology have to be chosen for her test.

3 CONCLUSION

In this paper, we extensively investigate CALT and we use classic ALT to compare with each other. We compare two test methods in different acceleration factor, mean-time-to-failure, warranty time, failure rates and accuracy points. Moreover, we go through all the datas we found and we explain the formulas that we use for finding parameters of each test methods.

CALT's sample size requires no equation and firstly it is introduced with high accuracy in every failure rates. Besides, it is also introduced with simple plot that requires only three point to maintain product's stress-life plot. We show that it is not clear to draw stress-life plot without knowing its trending. Trending should be identified with more points while points represent samples.

We find that although CALT uses much fewer sample size than ALT for test, its accuracy is not good enough. For CALT's accuracy increment, sample size should be increased. However, since sample size is associates with total test time CALT method can not overwhelm ALT method on different levels.

We also discover that, since ALT's sample size is decreasing by increment of acceleration factor with much higher trending than CALT. ALT can be the best test method from breaking point which is 40 in this matter. However, acceleration factor should be kept below 50 for realistic results.

ACKNOWLEDGEMENTS

This work is cooperated with Arcelik A.S. and supported by TUBITAK (The Scientific and Technological Council of Turkey) University-Industry Collaboration Grant Program #5130034.

REFERENCES

Cox D.R. & Hinkley D.V., 1974, Theoretical Statistics, *Chapman & Hall*, 49-209.

Donahoe, D., Zhao, K., Murray, S., Ray, R. M. 2008. Accelerated Life Testing, *Encyclopedia of Quantitative Risk Analysis and Assessment*.

Guo, H., Pohl, E., Gerokostopoulos, A., 2013, Determining the Right Sample Size for Your Test: Theory and Application, *Annual RELIABILITY and MAINTAINABILITY Symposium.*

Kendall, M.G. & Stuart, D.G., 1973, The Advanced Theory of Statistics. Vol 2: Inference and Relationship, Section 20.4.

Meeker, W. & Escobar, L., 1998, Statistical Methods for Reliability Data. New York: John Wiley & Sons, Inc.

O'Connor, P. & Kleyner, A. (5th) 2012. Practical Reliability Engineering. West Sussex: John Wiley & Sons.

Paschkewitz, J., 2009, Calibrated Accelerated Life Test with Time Varying Stress Profiles to Verify Product Reliability Requirements, *International Applied Reliability Symposium*.

ReliaSoft Corp., 2012, *Life Data Analysis Reference*, available online at: http://www.reliawiki.com/index.php/ReliaSoft_Books .