Environmental Stress Screening Tutorial

Bruce Peterson
Accolade Engineering Solutions
15520 Rockfield Blvd., Suite H
Irvine, CA 92618
949-597-8378
www.accoladeeng.com

1.0 Introduction

The purpose of this paper is to provide an overview of the technology of environmental stress screening. In the sections that follow, several areas of environmental stress screening technology will be discussed. These areas will include:

1. Description of why environmental stress screening is performed

2. Show the results from surveys which rank the most effective screening stimuli

3. Describe dominant product failure modes and the best stimuli for precipitating them

4. Describe the most common mathematical models for quantifying stress screening programs (also called the empirical stress screen equations)

5. Show the results of these mathematical models for the competing stress screening profiles

6. Describe the benefits of screens with high stress levels

7. Provide and overview of thermal chamber technology

8. Describe the developmental aspects of an effective environmental stress screening process

9. Present some conclusions

This document is not intended to provide a thorough treatment of product profiling as part of stress screen development. Every product should be considered a unique specimen with its own requirements for an effective screen. The specific steps toward product qualification and screen development are best determined by an engineer with relevant background and experience.

2.0 References

2.1 *'Environmental Stress Screening Guidelines for Assemblies'*, Institute of Environmental Sciences, 1990

2.2 Hobbs, G.K., HALT, HASS, *Precipitation and Detection Screens*.

2.3 IPC-SM-785

2.4 IPC-9701

2.5 IPC-A-610C

2.6 Jensen, F. and Petersen, N., *'Burn In'*, John Wiley and Sons, New York, 1982

2.7 Kapur, K. and Lamberson, L., *'Reliability in Engineering Design'*, John Wiley and Sons, New York, 1977

2.8 Kececioglu, D., Sun, F.,*'Environmental Stress Screening'*, Prentice Hall, 1995

2.9 Kececioglu, D., Sun, F., *'Environmental Stress Screening : Its Quantification, Optimization and Management'*, Prentice Hall, 1995

2.10 Kececioglu, D., *'Reliability Engineering Handbook, Volume 2'*, Prentice Hall, 1991

2.11 Lindgren, T., *'Optimizing ESS Effectiveness using Weibull Techniques'*, Proceedings from the Institute of Environmental Sciences, 1986

2.12 Montgomery, D., *'Introduction to Statistical Quality Control'*, John Wiley and Sons, New York, 1991

2.13 O'Connor, P., *'Practical Reliability Engineering'*, John Wiley and Sons Ltd., England, 1991

2.14 Proceedings of IEEE 1997 Workshop on Accelerated Stress Testing

2.15 Saari, A.E., Schafer, R.E., VanDenBerg, S.J., *Stress Screening of Electronic Hardware'*, Rome Air Development Center, 1982

2.16 Schlagheck, J., *Environmental Stress Screening*

2.17 Tustin, W. and Mercado R., *Random Vibration in Perspective*

2.18 Walpole, R. and Myers, R., *'Probability and Statistics for Engineers and Scientists'*, MacMillan Publishing Company, New York, 1985

3.0 Definitions

Bathtub Curve – A methodology for quantifying reliability when the hazard rate is plotted against time. The bathtub curve has three distinct regions. The first region, with decreasing hazard rate is the infant mortality region, the second region represents the constant hazard rate of the useful life and the third region is the increasing hazard rate which represents product wearout.

Burn-In Test – A test in which finished product is subjected to elevated temperatures over a period of time to accelerate failure modes which follow the Arrhenius reaction rate model.

Creep – The time-dependent visco-plastic deformation as a function of applied stress.

Coffin-Manson Model – A predictive model which relates the number of cycles to failure to the applied plastic strain.

Corrective Action - To find the fundamental cause for a defect and eliminate the cause. The corrective action must be proven to be the actual cause by appropriate experiments as corrective action can frequently be inappropriate if great care is not exercised in determining the true root cause and then verifying that the defect has been totally eliminated.

Cyclic Differential Expansion – Expansion differences due to the differences in coefficients of thermal expansion and cyclic temperature changes during operational use or temperature cycling tests.

Cyclic Temperature Range or Swing – Temperature amplitude between maximum and minimum temperatures occurring in operational service cycles or temperature cycling tests.

Destruct Limit - The stress level beyond which the product will suffer permanent damage and not function properly.

Detection Screen - A screen intended to provide an environment wherein patent defects can be detected.

Environmental Stress Screening – A screening process in which a product is subjected to environmentally generated stresses to precipitate latent product defects. The environmental stresses may be any combination of temperature, vibration or humidity.

Highly Accelerated Stress Screen – A screening process like ESS with stress levels typically beyond product operating ranges but within the product destruct ranges.

Infant Mortality – A failure during production screens (burn-in, ESS or HASS), initial functional testing or early service life where failures are associated with manufacturing process defects.

Latent - A defect which is undetectable in its current state with the tests to be used.

Maximum Cyclic Strain Range – The total stain range experienced after complete stress relaxation during exposure to cyclically induced thermal of mechanical deformations.

NAVMAT P-9492 Random Vibration Profile – The NAVMAT profile is a random vibration profile with a starting frequency of 20Hz and a vibration magnitude of 0.01 G^2/Hz. The vibration magnitude increases at the rate of 3dB/Octave to 80Hz where the vibration magnitude is 0.04 G^2/Hz. The vibration magnitude is maintained until the frequency reaches 350Hz where the vibration magnitude decreases at the rate of 3dB/Octave until the frequency reaches 2000Hz.

Operating Limit - The stress level beyond which the product will not operate properly, but below which it will even after repeated cycles above the operating limit. Some stresses will have upper and lower limits and others will not.

Patent - A defect which is detectable with the tests to be used

Precipitate - To change a defect from latent to patent.

Precipitation Screen - A screen intended to precipitate latent defects to patent.

Proof of Screen - A procedure for determining if the selected screening regimen will substantially degrade products run through the screen regimen.

Random Vibration – A vibration profile in which an assembly is subjected to a range of frequencies concurrently. Random vibration magnitude is expressed as G^2/Hz.

Sine Vibration – A vibration profile in which an assembly is subjected to a single frequency of cyclic displacement. A variation of Sine vibration is swept Sine. Swept Sine vibration applies Sine vibration profiles over a selected range of frequencies. Sine and Swept Sine vibration magnitude is expressed as Grms.

Solder Attachment – The collection of solder joints associated with a component.

Step Stress Approach - An approach wherein a given stress is adjusted in discrete steps until an operational or destruct limit is reached.

Stress Relation – The time-dependent decrease in stress due to visco-plastic deformation as a function of applied displacement.

Thermal Cycling – Exposure of assemblies to cyclic temperature changes where the rate of temperature rate of change averages less than 30°C/minute between 10% and 90% of the total temperature range.

Thermal Shock – Exposure of assemblies to cyclic temperature changes where the rate of temperature rate of change averages more than 30°C/minute between 10% and 90% of the total temperature range.

Vibration – A periodic, typically elastic, motion of a structure in alternately opposite directions from the position of equilibrium. Vibration can be single axis or multiple axis as well as Sine or Random vibration.

4.0 Background

Environmental Stress Screening (ESS) is a process used by factories to precipitate process related defects from latent to patent for detection by a product verification test. For most processes, the product verification tests are electrical tests but may include other forms of testing which are non-electrical. To conduct an effective screen, the product must be capable of surviving the high stimulation levels needed to accelerate the failure mechanism of assembly related defects. The participation of design and reliability engineering is to determine the limits of environmental stimulation which the product can endure before its performance is permanently degraded. A mechanically "weak" design may be changed to improve its margins with respect to a specific form of environmental stimulus. A by-product of this activity a more rugged product which may enjoy a higher demonstrated MTBF. The form of ESS chosen by the factory is dependent on the failure mechanisms for the relevant field failures. An ESS program is faulty when it does not expose the locus of faults seen by the customer. The ESS process is a dynamic process which must change as product failure behavior changes. For this reason, it is not appropriate to "spec" an ESS regimen and leave the regimen unchanged throughout product life.

Again, the purpose of any environmental stress screening (ESS) regimen is to expose the hidden defects that were introduced during the manufacturing process. More succinctly, manufacturing defects are precipitated from latent to patent. ESS, however, is not designed to find deficiencies in product design although in many cases it does expose design deficiencies. Rather than design an ESS program to find design weaknesses, an important ingredient of product qualification must be the undertaking of environmental testing to ensure that the design is robust enough to meet its design goals.

As the electronics industry has matured, component technology and assembly techniques have changed profoundly. The first products to be subjected to environmental stress screening in the form of "burn-in" was products designed with vacuum tube technology. For these products, high temperature burn-in was the best stimuli for precipitating latent defects. Early transistor and integrated circuit technology defects were also efficiently stimulated to failure using high temperature burn-in. During this time, most product defects were component related defects. Today we see a much different failure behavior for electronic products. Components have become so reliable that most product defects are related to the assembly process. As the product fault spectrum has changed, the screening stimuli for rapid defect precipitation must also change to ensure that latent defects are efficiently stimulated to an observable failure.

In 1990 Motorola conducted a study of burn-in effectiveness. They found that after burn-in only 0.000658% of the units failed an electrical test. Motorola's conclusion was that burn-in, prior to usage, does nothing to remove many failures but may cause failures due to additional handling (See table 1 below).

Product Family	Quantity Tested	Electrical Rejects	Comments
LS	38,855	1	Mech. Surface
Fast	51,830	3	1- assembly, 1 – WFAB, 1 – undetermined
MECL 10K	33,078	4	Oxide pinholes
MECL 10H	62,573	5	3- assembly, 2- undetermined
MG CMOS	26,408	1	Photo resist
Total	212,744	14	0.000658%

Table 1 – Summary of Electrical Burn-in Results

"The reliability of integrated circuits has improved considerably over the past five years. As a result, Burn-in, prior to usage, does not remove many failures. On the contrary it may cause failures due to additional handling" – Motorola Handbook 1990

In addition, two studies were conducted in 1984 to determine which environmental stimuli were considered the most effective. These studies were performed by the Institute of Environmental Sciences and the French ESS Task Teams. The results of the IES survey is shown in figure 1 and the results of the French ESS task team is shown in figure 2. In both of these studies, thermal cycling was cited as the most effective environmental stimulus.

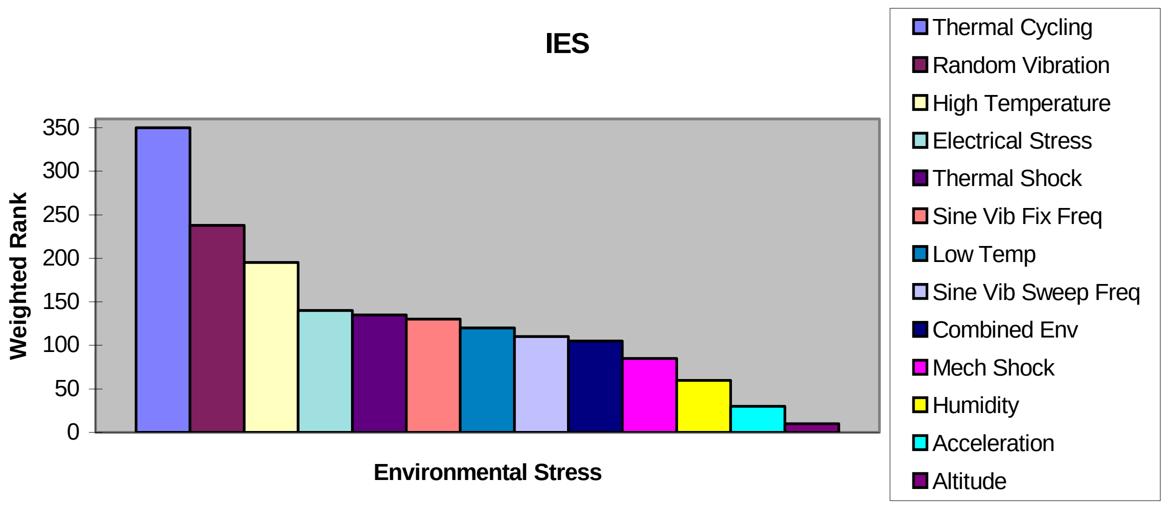

Figure 1 – IES Survey of Screening Effectiveness

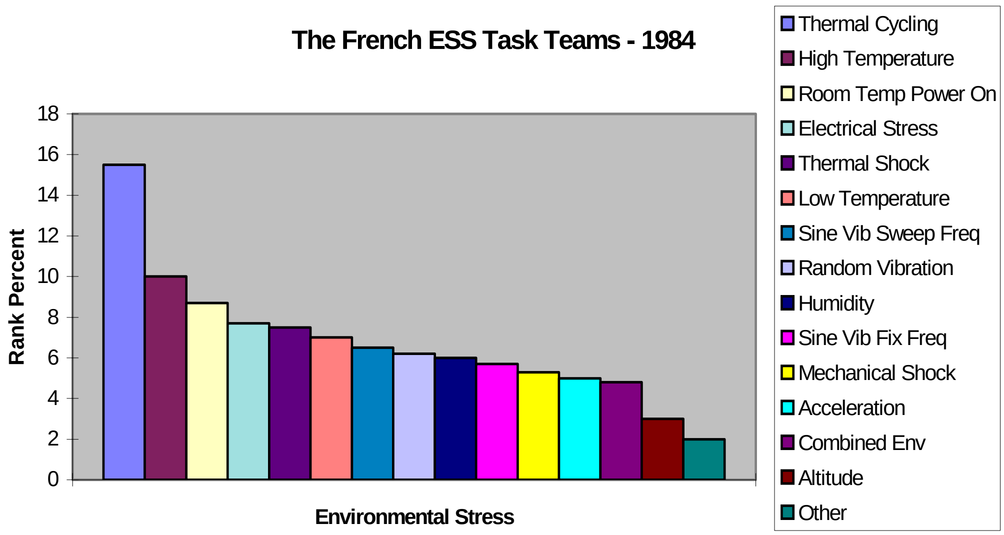

Figure 2 – French ESS Task Team ESS Effectiveness Survey

Additionally, the IES study cited workmanship as the dominant failure mechanism for unscreened development systems. To help organizations with no prior ESS experience, the IES developed a set of baseline thermal screening parameters. Products that have not participated in a HALT study (to be described later), should make use of the IES recommendations and perform a proof of screen (also described later) for initial screening regimens. The IES recommendations are shown in table 2 below.

Characteristic	Level of Assembly		
	PWA	Unit	System
Temperature range of hardware	-50°C/+75°C to -65°C/+100°C	-40°C/+70°C to -55°C/+85°C	-40°C/+60°C to -55°C/+70°C
Temperature rate of change of hardware	10°C/min to 20°C/min	5°C/min to 20°C/min	5°C/min to 10°C/min
Soak time of hardware at temperature extremes - if unmonitored	5 minutes	5 minutes	5 minutes
- if monitored	Long enough to perform functional testing		
Equipment Conditions	Unpowered	Powered, monitored	Powered, monitored

Table 2 - Baseline Regimen for Organizations Lacking ESS Experience

Studies have also shown that two stimuli applied concurrently during the stress screen are more effective than applying the same environmental stimuli in multiple screens. Power cycling, for example, during the cold exposure interval of the environmental stress screen is a stronger screen than performing the same power cycling at room temperature.

Although constant temperature burn-in is still very popular, studies over the last several years have clearly shown that constant temperature burn-in is an inefficient stimuli for precipitating the class of defects commonly found on current electronic products. The use of a constant temperature burn-in is more appropriate for detection screens or product mission environmental demonstrations. Table 3 below clearly shows that structural defects (the most common class of defects today) are best stimulated by thermal cycling and thermal shock. For this reason, many companies, have abandoned constant temperature burn-in as a defect precipitation screen. The IES survey confirms this trend. The form of stimulus now used by most organizations includes thermal cycling while the most capable screening processes combine environments such as thermal cycling along with random vibration and power cycling. The baseline parameters for a thermal cycling screen is 5-20°C/minute ramp rate, a 100-120°C temperature swing and soak times only long enough to ensure temperature stabilization or to perform functional testing. The IES survey also confirms this trend.

Screening Test	IC Failure Mechanisms									
	Substrate mounting defects	Bulk silicon defects	Substrate surface defects	Bonding and Wire	Partricle contamination + extraneous material	Seal Defects	Package Defects	External lead defects	Thermal Mismatch	Electrical stability
Internal visual exam	●		●	●	●		●			
External visual exam						●	●	●		
Stabilization bake		●	●	●						●
Thermal Cycling	●		●	●		●	●		●	
Thermal Shock	●		●	●		●	●		●	
Centrifuge	●			●	●		●			
Shock	●			●	●		●			
Vibration	●			●	●		●			
X-Ray	●			●	●	●	●			
Burn-in	●	●	●	●						●
Leakage Tests						●				

Table 3 – IC Failure Mechanism and the best screens for precipitation or detection

5.0 Solder joint failure mechanisms and environmental stimuli

In thermal shock screening, the stresses imposed on a solder joint are a result of warpage of the surface mount assembly and the differences in the coefficient of thermal expansion between the materials which comprise the surface mount assembly. During rapid temperature change, the warpages result in tensile and shear stresses on the solder joint. Even assemblies with matched coefficients of thermal expansion will exhibit solder joint failures when subjected to thermal shock. In figure 3 below, a stress/strain chart shows the characteristics of solder joints during thermal cycling. As the solder joints are brought to the highest temperature, the leaded and leadless devices will develop internal stresses and will also exhibit some elastic yielding. Over time, the solder joints will creep to provide stress relaxation. The leaded components also have different stress relief characteristics as the lead in addition to the solder joint contribute to the stress relieving process. After the solder joint completely stress relieves and the solder joint is brought to the lowest temperature, the stress and stress relaxation processes occur in the same manner as for the high temperature case described previously. To provide the optimum level of stress relation of the solder joints during the thermal shock screen, a dwell at each temperature extreme must be provided. The solder joint can relax in minutes at high temperature and days at very low temperatures. Generally, a dwell time of five to fifteen minutes is prescribed for most screens.

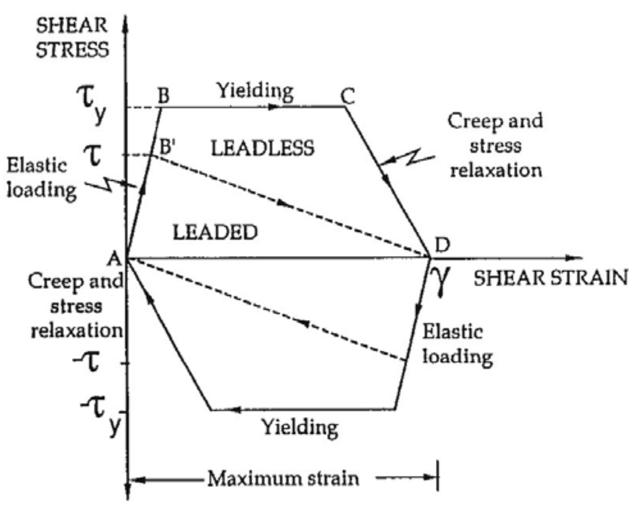

Figure 3 – Stress Strain Curves for Solder Joints

Another study was performed to determine the optimum temperature range for the temperature cycling screens. The temperature range is the magnitude of the total temperature change from the lowest temperature to the highest temperature. The IES recommendations (table 2) show temperature ranges from 100°C to 165°C. Figure 4 below shows that the number of lifecycle loads was minimized when the temperature range was 120°C. In addition, the same figure shows the affect of unequal dwell times.

Most practitioners of stress screens would believe that unequal dwell times are needed to allow proper solder joint relaxation in the hot and cold temperature exposures. In an unequal dwell time scenario, the cold temperature exposure would have a longer dwell time due to the reduced plasticity of the solder at cold temperatures. As figure 4 shows, the duration of the dwell times and the equality of the dwell times for hot temperature and cold temperature exposure had little affect on the part life cycle loads

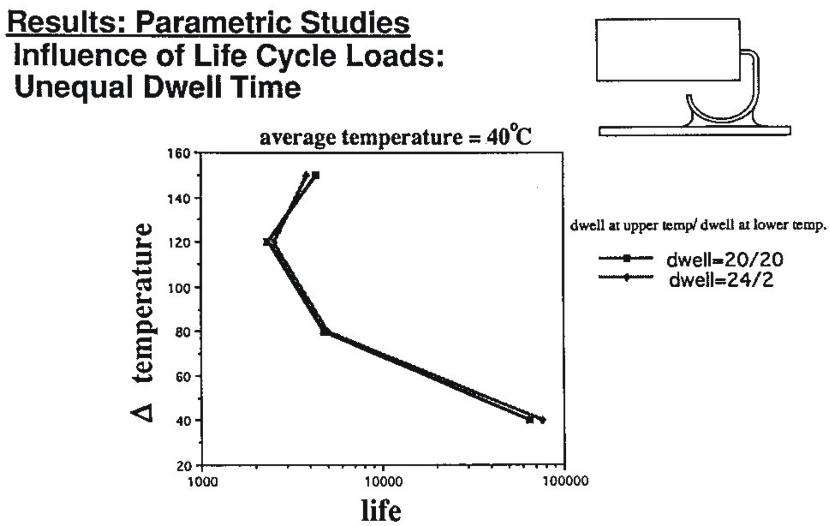

Figure 4 – Solder Joint Lifetime as a function of thermal cycling temperature range

6.0 Introduction of Mathematical Models

To help develop screening parameters which are able to accelerate product failure modes and also provide an empirical measure of screening strength, a set of mathematical equations have been developed. The most ubiquitous model for constant temperature burn-in is the Arrhenius reaction rate model and the most accepted models for empirical screening strength are the RADC models. RADC models are available for constant temperature burn-in as well as thermal cycling.

6.1 The Arrhenius Reaction Rate Model

The Arrhenius Reaction Rate Model relies on the premise that defects, like many chemical reactions, have an energy threshold that must be exceeded before the reaction can occur. Also, like chemical reactions, when the applied energy level is increased, the speed of the reaction is also increased. For constant temperature burn-in, the magnitude of the temperature is expected to have a direct affect on the speed of defect precipitation and yields a result called the acceleration factor.

The Arrhenius acceleration factor (AF) equation is:

$$AF := \exp\left[\frac{E_A}{K}\cdot\left(\frac{1}{T_o}-\frac{1}{T_a}\right)\right]$$

Where:

AF = acceleration factor
E_A = activation energy in eV (0.7eV is common value for CMOS ICs)
K = Boltzmann's Constant = 8.617385×10^{-5} eV/°K
T_o = Operating Temperature in °K (use 298°K – same as 25°C)
T_a = Accelerated Temperature °K

Figure 5 below shows the levels of acceleration as a function of applied temperature for several different activation energies. For figure 5, the operating temperature is assumed to be 25°C. From the graphs, it is clear that to gain any significant acceleration level at nominal activation energies, temperatures above 85°C must be used. Unfortunately, most commercial electronic products have operating temperatures below 70°C which prevent functional testing during the stress screen unless the temperature is lowered during segments of the screen to accommodate the functional test. Due to the lack of high acceleration levels during burn-in, the screening duration must be increased to allow sufficient time to precipitate all relevant defects. Burn-in screening times on the order of days is common. In addition, constant temperature screens are less capable of providing the level of stress needed to accelerate structural latent defects. The solder joint stress-relaxation scenario discussed previously would only be active during the initial heating phase of the constant temperature burn-in.

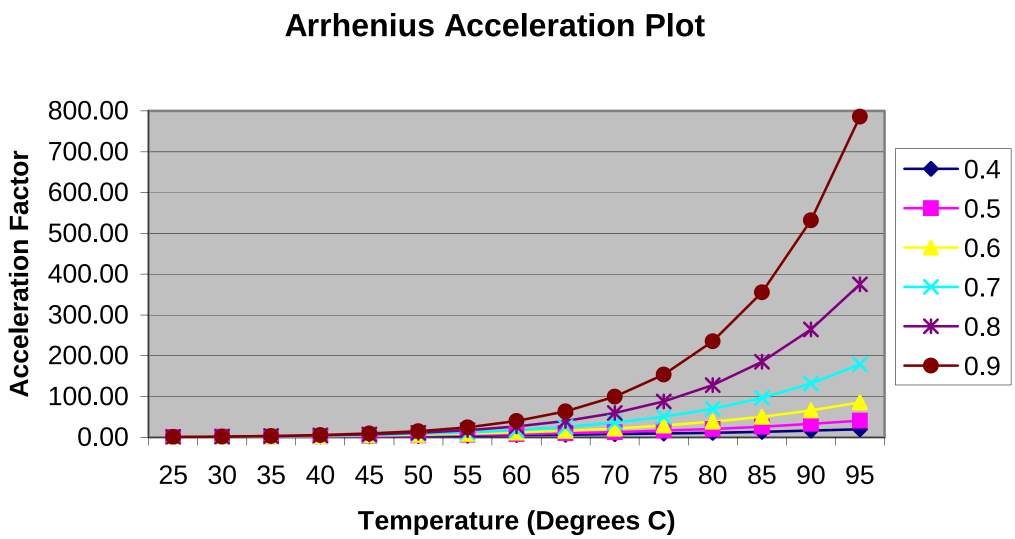

Figure 5 – Arrhenuis Acceleration Plot as Function of Temperature

6.2 The RADC models

The RADC equations are empirical stress screen equations which are used to compare the screening strengths of different thermal screening profiles. These equations by themselves to not infer a screening duration or an acceleration (aging) factor but are useful for manipulating screening parameters to gain the greatest screening strength.

6.2.1 The RADC constant temperature model

The RADC equation for constant temperature burn-in is:

$$SS := 1 - \exp\left[-0.0017\left(T_R + 0.6\right)^{0.6} \cdot T_b\right]$$

Where:

SS = screening strength
T_R = Temperature Range above ambient (ambient is 25°C)
T_b = Burn-in Time in hours

Figure 6 shows a plot of screening strength as a function of temperature for several screening temperatures. Like the plots in the Arrhenius acceleration plot, the figure 6 plots show that higher temperatures and longer screening durations (measured in days) are required to achieve the minimum levels of screening strength.

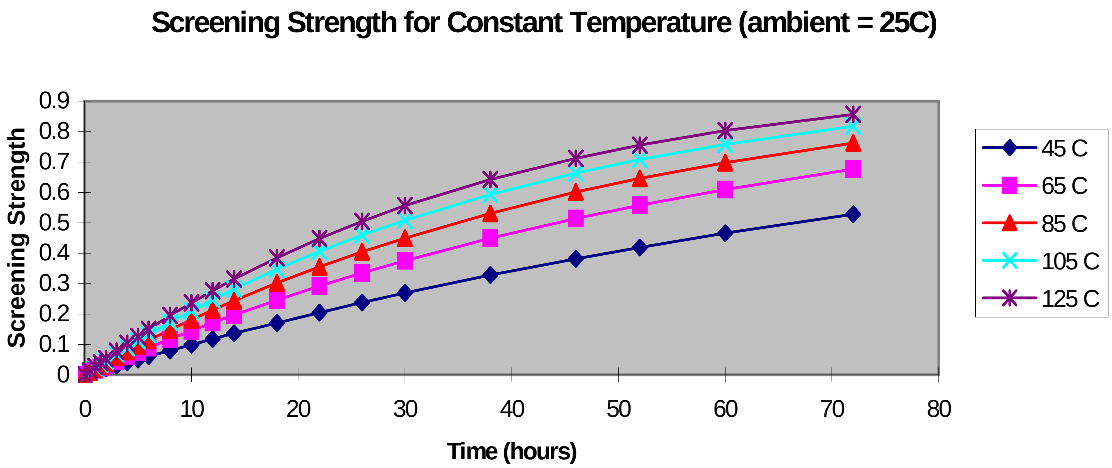

Figure 6 – RADC model for screening strength of constant temperature screen

6.2.2 The RADC thermal cycling model

The RADC equation for thermal cycling is:

$$SS := 1 - \exp\left[-0.0017\left(T_R + 0.6\right)^{0.6} \cdot \left(\ln(e + dt)\right)^{3} \cdot N_{cy}\right]$$

Where:

SS = screening strength
T_R = Temperature Range (highest temp - lowest temp) in °C
e = natural log base = 2.7183
dt = temperature rate of change in °C/minute
N_{CY} = Number of thermal cycles

Figure 7 shows a plot of screening strength as a function of thermal cycles for several different rates of temperature change. The plot of the thermal cycling screen strength clearly shows that the best stress screen performance occurs when the temperature rate of change exceeds 10°C/minute. Unlike the constant temperature screen, the thermal cycling shows the ability to quickly gain high screening strengths in as little as 2 to 10 thermal cycles.

One of the most important characteristics of the thermal cycling screening strength plots is that when the temperature rate of change is increased, the number of thermal cycles to achieve a desired screening strength is reduced. Figure 8 shows a study that was performed to experimentally determine the relationship between the number of thermal cycles required for a screen and the temperature rate of change.

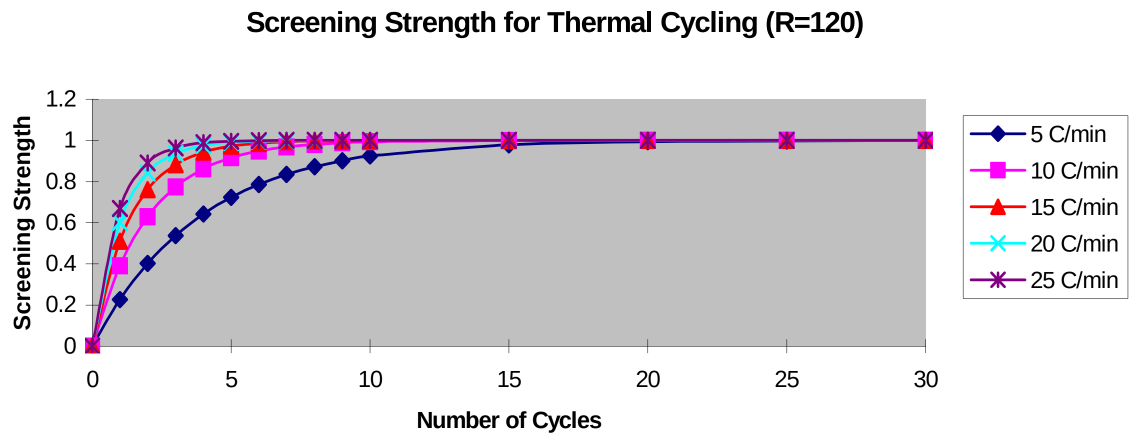

Figure 7 – Screening strength as a function of number of thermal cycles

Referencing the dark band in the plot we can see how significantly the temperature rate of change affects overall screening time. At 5°C/minute, 400 thermal cycles are needed for the screen while at 30°C/minute, 2 thermal cycles are needed for the screen.

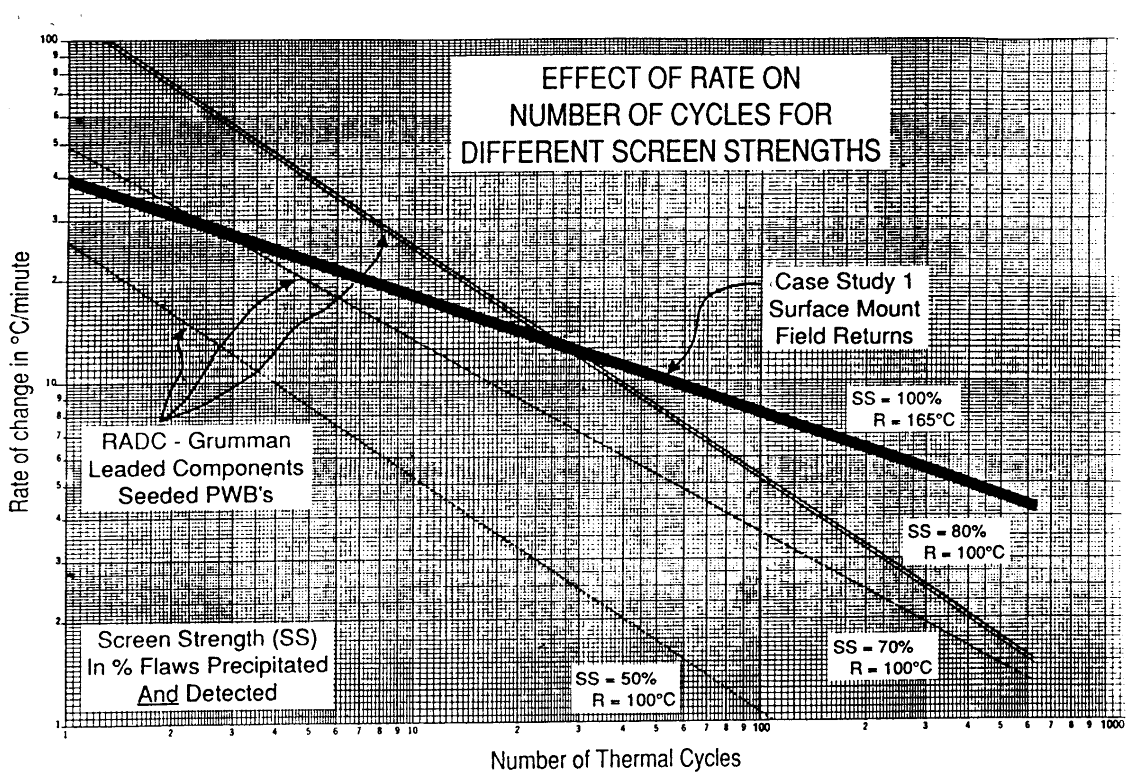

Figure 8 – Number of thermal cycles as a function of rate of temperature change

To more clearly illustrate this point, table 4 shows the overall screening time for given temperature rates of change with 5 minutes dwells at each temperature extreme. At 5°C/minute the screen duration will be 23,200 minutes or 387 hours while at 40°C/minute the screen duration will be 16 minutes or 0.27 hours. From a standpoint of equipment cost and process efficiency, faster thermal rates of change are most desired.

Temperature Rate of Change (°C/minute)	Screening Time (minutes)
5	23,200
10	2,040
20	154
30	27
40	16

Table 4 - Screening time and temperature rate of change

Going back to the original RADC screening strength plot for thermal cycling, a more appropriate view of the thermal cycling performance is to view the same data with the thermal cycles converted to hours. Figure 9 shows a plot of screening strength as a function of time for several different rates of temperature change. In this plot, a dwell of 5 minutes is provided at the end of each thermal ramp to allow for solder stress relaxation. Again we see that higher thermal rates of change are desired with an average of 10°C/minute to15°C/minute temperature rate of change being the minimum threshold.

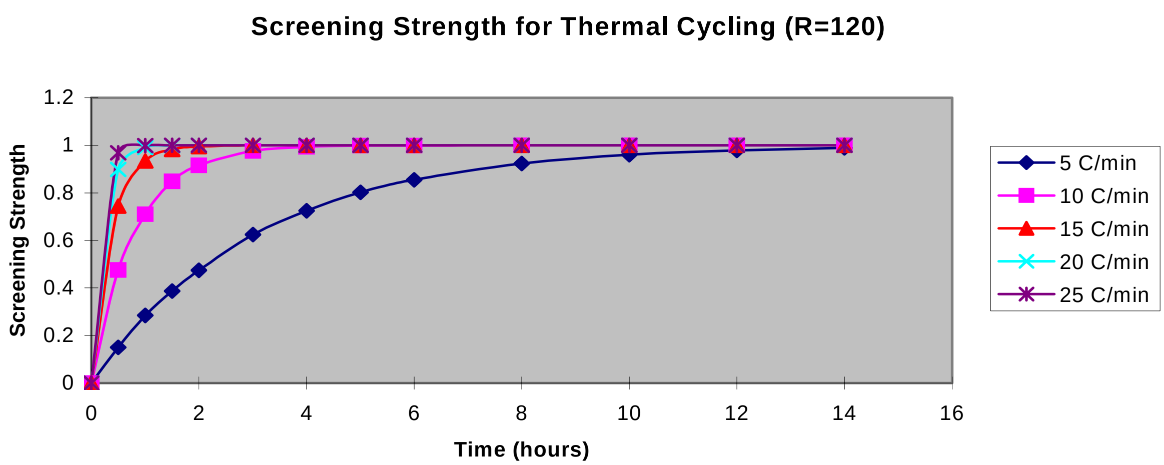

Figure 9 – Screening strength as a function of time

7.0 Product Profiling

Prior to the implementation of the production screening process, the product profiling phase of ESS should be executed by the designers engineering. The product profiling phase may also be called the HALT (Highly Accelerated Life Test). The goal of the product profiling phase is to use environmental stimuli to ensure that new products are designed with generous margins of strength to:

- improve reliability (MTBF)

- allow higher stress levels during manufacturing screens to substantially reduce the screening duration

- learn about product failure behavior to determine the type of environmental stimuli which may be used during manufacturing screens

In the absence of the product profiling phase, the implementers or the production screen will need to be conservative in the choice of parameters for the screening equipment and stay within the IES guidelines. A proof of screen (described later) will be used to verify the baseline screening parameters do not damage good hardware. An effective screen will result but the screen will not be optimized for the shorted duration and lowest possible screening cost.

7.1 Documentation

A product profiling plan must be completed prior to the start of the profiling process and approved by the responsible authorities. The end use environment of the product must be considered during the product profiling plan. Product profiling plans, therefore, are not universal and each product or type of product must have unique profiling plans.

7.2 Selection of Stimuli

The product profiling plan will enumerate the type(s) of stimuli that will be used during the product profiling phase. The type(s) of stimuli will be determined by the cognizant reliability engineer and design engineer. The stimulus environments used by manufacturing on existing products should be used as a minimum to determine if existing screening equipment will satisfy the requirements for a new product. Some stimuli may include, but not limited to, temperature, random vibration, voltage, frequency, power cycling and humidity.

7.3 Determination of Product Limits

For each of the environments specified in the product profiling plan, engineering must determine the operating and destruct limits. Some types of stimuli such as voltage or temperature have upper and lower operating limits as well as upper and lower destruct limits. These limits can best be determined by using the step-stress-test approach. During the step-stress-test approach, each environment is incremented or decremented in discrete steps between product exposure and test. The time that a product is exposed to an environment should be the same for each environment. The test which is performed to determine proper product operation must be the same test after each exposure period. The step size for each environment is determined by the type of stimuli and expected range of stimuli. Temperature, for example, may be incremented in large steps until the expected operating or destruct limit is near; the step size can then be reduced to find the closest number which represents that limit under investigation.

After the operating and destruct limits are found for a single applied stimuli, the process can be repeated for combined stimuli where only a single stimuli is allowed to change between each iteration; all other stimuli are held at a constant level.

Any product failures during the step-stress-test procedure must be analyzed to root cause to completely understand the nature of the failure. The appropriate product changes are best determined when a thorough understanding of the failure mechanism is gained. All failure data and root cause analyses must be documented to aid manufacturing in selecting a baseline screening stimuli or to gain knowledge of possible product weaknesses prior to start of high volume production.

7.4 Product Changes to Improve Margins

When the measured operating and destruct limits are too close to each other, too low or when the operating limits are too close to the guaranteed operating limits, changes to the product must occur to improve design margins. Other design problems may be detected by analyzing the measurement data. During vibration, for example, swept sine stimuli may be used to perform a resonant search for various structures in a product. Poor damping at the resonant frequency or overlapping resonances between stacked structures are indicators of possible product reliability problems.

The nature of any detected design problem or potential reliability problem must be communicated to the responsible design engineer so that the appropriate changes can be made. In most cases, the changes are simple with minimal cost impact if the changes are made early in the design process. After the changes are incorporated into the design, the profiling process must be repeated to ensure that the design was actually improved and that another parameter was not compromised. The engineers involved can determine if the entire profiling process must be repeated or only a portion of it. All actions must be documented.

8.0 Screen Development

The screen development phase of ESS implementation is executed by manufacturing. The goal of the manufacturing ESS or HASS implementation is to perform a screen which will precipitate relevant latent defects in the most cost effective manner and leave enough life in the product to fulfill its mission. Relevant defects are defects which correspond to the locus of manufacturing detected failures and field detected failures.

8.1 Selection of Stimuli

For new technology for which no failure history exists, manufacturing must rely on the recommendations of engineering for the baseline screening stimuli. The report from the product profiling (HALT) phase should be studied to make the best selection. In some cases, studies from other organizations may help select the stimuli which has the best probability of finding the expected types of defects for a particular technology. Studies conducted by the IES is one such resource.

For products which utilize technology for which failure experience is available, one can select the appropriate stimuli from an evaluation of the root cause analysis of the failures. The physics of the failure analysis will indicate which screening stimuli is best for rapid precipitation. In many cases more than one type of stimuli is equally effective at defect precipitation; in such cases, the stimuli which encompasses the majority of defect mechanisms should be chosen. Since no single stimuli can precipitate all defect types, more than one stimuli may be necessary. Combined environments has been proven to be more effective at defect precipitation than the same environments executed in tandem. Thermal cycling and power cycling, for example, are environments which are commonly combined.

For most screens, the selection of stimuli and the end use environment have no relevance. The goal of an effective screening program is to find latent defects which would fail in the field if screening was not performed. The method we choose to use to precipitate the defect from latent to patent does not matter as long as we precipitate it quickly, cost effectively and without compromising the remaining life in the product. The field environment, however, should be considered when it is abnormally abusive. The stimulation levels used for such products should be no less than the expected end use stimuli.

8.2 Magnitude of Precipitation Stimuli

The magnitude of the environmental stimuli is determined by the operating and destruct limits measured during the product profiling (HALT) phase. The type of stimuli will also determine the baseline screening magnitude. The sections below discuss the baseline levels for each type of precipitation stimuli

8.2.1 Voltage

The applied voltage should be 5% less than the destruct limit

8.2.2 Vibration

The applied vibration level should range from the midpoint of between the operating limit and destruct limit to 25% less then the destruct limit. The lower of the two numbers should be used initially and gradually increased to the higher of the two numbers. The shaker type used to determine the product limits should also be the shaker type used during the manufacturing screen. Using different shaker types can lead to serious problems because the vibration intensity levels (measured as G_{rms}) used during random vibration is profoundly different than the same numeric G_{rms} value used in sine or swept sine environments. The correct measure of vibration intensity for random vibration is power spectral density (PSD). The use of G_{rms} values for random vibration is to allow comparison of narrow band random vibration levels to sine of swept sine vibration levels. Be sure to measure the vibration levels on the product on multiple axes.

8.2.3 Thermal Cycling

The applied thermal levels (upper and lower) should range from the midpoint of between the operating limit and destruct limit to 10°C less then the destruct limit. The lower of the two numbers should be used initially and gradually increased to the higher of the two numbers. The thermal rate of change should be as high as possible within the limits of the technology or to the maximum rate of the chamber. Be sure to measure the temperature levels and rate of change on the product.

8.2.4 Power Cycling

Ten power cycles should be selected as a starting point for power cycling. Power cycling may involve the cycling of AC power to an entire system or the cycling of DC power to a subassembly or circuit board. Power cycling may also include a test between power cycles to determine if a failure has occurred.

8.3 Duration of Applied Stimuli

The duration of applied stimuli is based on the observed time to defect precipitation for targeted defect types. The Weibull techniques are used to determine the optimum screening duration (see section 8.8). Studies should be conducted periodically (perhaps monthly) whereby extended screens are performed to search for new failure modes which occur at times beyond the current screening duration. The screening duration may need to be adjusted to capture multiple failure modes which do not all precipitate within the same time intervals.

In all cases, the magnitude of the applied stimuli has a significant affect on the screening duration. A general rule is that doubling the applied stimuli can reduce the screening duration by a factor of ten. For example, a 120°C range thermal cycling screen at 5°C takes 400 cycles (about 23,200 minutes) to precipitate a failure which is best precipitated by thermal rates of change. The same screen at 10°C will take about 60 cycles (2,040 minutes) to precipitate the same defect. For thermal cycling chambers with high rates of change (>40°C) the number of cycles can be as few as two. Higher screening levels, therefore, result in lower capital equipment costs. The significance of product profiling and mechanically robust designs from engineering is easily recognized.

For power cycling we do not have any method to increase the intensity of its stimuli alone to reduce the number of power cycles. If we are using power cycling to precipitate electromigration defects, we can combine high temperature and power cycling to accelerate electromigration defect precipitation. By measuring the number of power cycles to failure precipitation, the optimum number of power cycles can be determined.

8.4 Proof of Screen

Once the desired screening stimuli and baseline parameters have been selected we must ensure that the screening regimen has not taken too much life out of the product. The purpose of ESS is to accelerate the product aging process to find latent defects which would become patent and be experienced by the customer if the screening was not performed. The screen, therefore, takes life out of the product. To determine if enough life is remaining in the product to fulfill its mission, a proof of screen is conducted.

The proof of screen involves the application of the screening regimen and test repeatedly without failure. The number of iterations should be about twenty but no less than ten. If a failure is found during the proof of screen, the cause must be investigated to determine if the cause was due to excessive fatigue damage from the screen. The screening parameters must be adjusted if the magnitude of the stimulation is too great. After the screening parameters have been adjusted, a proof of screen must be performed again. If the screening parameters are substantially lower than the original limits provided during the product profile (HALT), engineering should be contacted to investigate the cause.

As product failure behavior changes or as the product evolves through its life, the proof of screen must be repeated to ensure the validity of the currently selected screening parameters. A new product profile (HALT) may be in order if the product has undergone many changes.

8.5 Testing During the Precipitation Screen

To improve the probability of detecting precipitated defects, a test executed during the screen is desirable. Since the screening levels are usually beyond the operational levels of the product, it may not always be possible to perform a test during the precipitation screen. Often we may find that at least some areas of the product can and should be tested during the precipitation screen.

Some specimens undergoing the precipitation screen may not be electrically or mechanically complete to perform a test unless additional hardware is attached. In such cases, especially at a subassembly level, the requirement for test during the precipitation screen can be relaxed. In general, the higher the level of assembly the greater the need to perform test during the precipitation screen.

8.6 Detection Screens

Although a defect may have been exposed during the precipitation screen, the defect may not be observable when the stimuli is removed. Defects such as these are commonly called intermittent defects and they can be difficult to diagnose. The use of a detection screen can be used to solve this problem. The main characteristics of the detection screen is stimuli within the operational limits of the product while a thorough test is performed. The magnitude of the environmental stimuli for the detection screen is determined by the operating limits measured during the product profiling (HALT) phase. The sections below discuss the baseline levels for each type of detection stimuli

8.6.1 Voltage

The applied voltage should be 5% less than the operational limit.

8.6.2 Vibration

The applied vibration level should range from nearly zero to 20% less then the operational limit. The type of shaker used for detection is less critical than the type of shaker used for precipitation but one should recognize that the farther you deviate from the precipitation environment the less likely the failure can be duplicated or observed. Be sure to measure the vibration levels on the product on multiple axes.

8.6.3 Thermal Cycling

The applied thermal levels (upper and lower) should lie under 10°C of the upper and lower operating limits. The thermal rate of change should be as high as possible but may not be practical due to the equipment limitations for the detection environment. Be sure to measure the temperature levels and rate of change on the product.

8.6.4 Elevated Temperature

The use of a burn-in environment as a detection screen can be very effective. The burn-in detection screen is useful to observe a previously precipitated but intermittent failure and to demonstrate proper environmental performance. When an elevated temperature screen is used the temperature level should lie under 10°C of the upper operating limits and the unit should undergo testing during the initial temperature rise to the temperature setpoint.

8.7 Root Cause Analysis

To determine the physical nature of the failures precipitated during screening, a thorough root cause analysis program must be implemented. The root cause analysis can determine if product failure behavior is changing or if the screening parameters are too destructive. In both cases, the screening regimen should be changed to adapt to the changing failure behavior of the product. Another obvious by-product of root cause analysis is to implement corrective action to prevent future repeat occurrences of the failure. The level of analysis during root cause analysis is very important to determine the appropriate corrective action. Component or solder failures can fall into one of several failure modes. The goal of root cause analysis is to determine the precise failure mechanism of the defect.

8.8 Optimizing the ESS Process

The IES baseline ESS regimen provides an extremely conservative process that does gain the levels of efficiency that are most desired. Practitioners of ESS must optimize the screening duration to conclude when all latent defects have been precipitated.

The method used for ESS screening duration optimization was first described in the book “burn-in” by Jensen and Petersen. The essence of the method described by Jensen and Petersen is that a group of products contain units without defects and units with defects. The units without defects are called the strong main population. The units with defects are called the weak sub-population. During the observation period, starting from time zero, we tend to see failure behavior that represents a decreasing failure rate. The decreasing failure rate is commonly referred to as the infant mortality region of the bathtub curve. This region of the bathtub curve is the region during which the weak sub-population of units with latent defects are exhausted. As time progresses, the failure rate approaches a constant level which is called the “useful region” or “unit operating life” of the bathtub curve. This region of the bathtub curve is the region where units from a strong main population will live most of their life. This constant level in the bathtub curve or constant failure rate is not at a magnitude of zero owing to the non-infinite MTBF value of products. The constant failure rate implies that failures from the strong main population have equal probabilities of failing over all observation times. The infant mortality region of the bathtub curve is actually a superposition of failures from the weak sub-population and the strong main population. The mathematical techniques presented by Jensen and Petersen are designed to determine at which time the members of the weak sub-population have been exhausted. The additional ingredient of adding an environmental stress does not alter the computation technique since it just compresses the variable time.

This method uses the time to failure history of monitored units in ESS to fit the data to a Weibull distribution to create a mathematical model of the product failure behavior. After the mathematical model of the failure behavior is derived, Bayesian methods are used to apply probabilities to each observed failure. Failures can be from a weak sub-population or a strong main population. The weak sub-population includes all units which have the latent defects we wish to precipitate. The strong main population includes all the defect free products that will fail at times in accordance with the exponential distribution model. The exponential distribution models the constant failure rate region of the bathtub curve and is actually a special case of the Weibull distribution. The trend of the Bayesian probabilities is that early failures are more likely to belong to the weak sub-population. Late failures are more likely to belong to the strong main population. Through the manipulations of these probabilities we can find a decision point at which the weak sub-population has been exhausted. This time will be the time at which the environmental stress screening process should be terminated. The details of all the mathematical processes are beyond the scope of this paper.

8.9 Corrective Action

The last necessary ingredient of a successful screening program is corrective action. This stage completes the entire screening process loop which includes the precipitation screen, the detection screen, root cause analysis and corrective action. The loss of any if these steps results in an open loop process which is not capable of correcting its process errors. Without corrective action, the customer still enjoys the benefits of a reliable product but the manufacturer suffers with the expense of the repair, re-screen and re-test operations. For the manufacturer to achieve a low cost screening process, screening must evolve into a sampling operation as soon as possible. If 100% of all product is screened, the equipment costs could be overwhelming. We must remember that screening adds cost to a product as the capital equipment costs are amortized over all products produced.

The appropriate corrective action is determined by the analysis of the root cause analysis data. Since the wrong corrective action can be nearly as ineffective as no corrective action, the accuracy of root cause analysis is paramount. In addition to appropriate corrective action, we must have timely corrective action. The time between the end of the assembly process to the implementation of corrective action must occur quickly to maintain a sampling ESS process while protecting the customer. This is one of the reasons that short screens are desired.

Very often during final assembly screening operations, defects on sub-assemblies built outside of the facility are detected. These subassemblies may have been built at a manufacturing site a great distance away and they may have been built days or weeks ago. Although the customer is protected from receiving a product which is likely to fail, the ability to quickly determine the root cause and take corrective action may be impacted due to distance in time or geography. In such cases, the screening programs at these remote sites (if they exist) should be reviewed to determine their ability to precipitate the defects experienced in the final assembly operations. The salient point here is that the screening program for a particular manufacturing site should be designed to detect the latent defects that result from *their* process so that speedy corrective action can take place.

8.10 Re-Screening Product

Since precipitation screens take life out of a product, care must be exercised to ensure that the product is not re-screened too many times. A general rule is that the maximum number of precipitation screens a unit is exposed to should not exceed 10% of the proof of screen iterations. Since detection screens use stimulus levels which are much lower than precipitation screens, the acceleration levels are lower as well. Detection screens can be repeated as many times as desired.

9.0 Thermal Screening Equipment

From the previous sections of this paper the basics for the implementation of an effective environmental stress screening program was described. The most significant aspects of the stress screening program is the type of stimulus and the screening parameters. From many studies, thermal cycling is universally accepted as the most effective stimulus. Other studies have shown that rapid rate of change of the temperature will provide the shortest and most economical screening processes.

The question that now remains is to determine the type of thermal cycling equipment to use. Thermal cycling chambers can use mechanical refrigeration or liquid nitrogen to perform the cooling function. Heating in all types of thermal cycling chambers is accomplished through electric heating coils. In thermal cycling chamber designs, the main considerations are the heating capacity, the cooling capacity and the air- flow. The heating capacity tends to be less of a problem as it is one of the least expensive resources in the chamber construction. The air-flow is more important as the ability of a product to track chamber air temperature is largely determined by the velocity of the air-flow. A study performed by engineers at Hughes Aircraft showed that for thermal screening at greater than 10°C/minute required air-flow greater than 1,000 linear feet/minute. The ESS chamber, therefore, must have a high performance circulation system. The last section of chamber technology involves the type and size of the refrigeration system. For small lab applications where linear rates of change or ultra-fast rate changes (>40°C/minute) are needed, liquid nitrogen is the cooling method most preferred. The use of nitrogen, however, can be problematic. Liquid nitrogen is hazardous to handle, cannot be stored for more than a few weeks, requires proper venting from the facility to the outside environment and can be expensive to use in large chambers. For most thermal chamber applications, mechanical refrigeration is used. The refrigeration consists of one of more compressors in a single stage or cascade configuration and either water cooled or air cooled condensers. The single stage compressor system allows cooling to about -20°C to -30°C. The cascade or dual stage compressor systems allow cooling to -70°C. The choice of air cooled or water cooled is determined by the size of the compressor pumps. Small lab type systems with compressors under 5hp are generally air cooled (residential air conditioning systems are air cooled). High performance compressors that are 10hp to 30hp are usually water cooled. These systems can be air cooled but their efficiency can be impacted and the quantity of expensive refrigerant that is needed in the system is greater.

Figure 10 shows a drawing of a conventional single zone chamber layout. Inside the chamber are the products to be screened (UUT), the chamber walls (liner), the heating coils, the circulation fan and the evaporator. All these units constitute the load to the refrigeration system. For high performance compressors (30hp), the evaporator alone will be 100lbs of aluminum coils and fins. In some cases, the load of the chamber components can exceed the load represented by the UUT. In chambers of this design, a cascade 30hp x 30hp system should not be expected to perform cooling at rates greater than 10°C/minute. Certainly, extra compressors can be added to improve performance but the cost of the system and the cost to operate the system may become excessive.

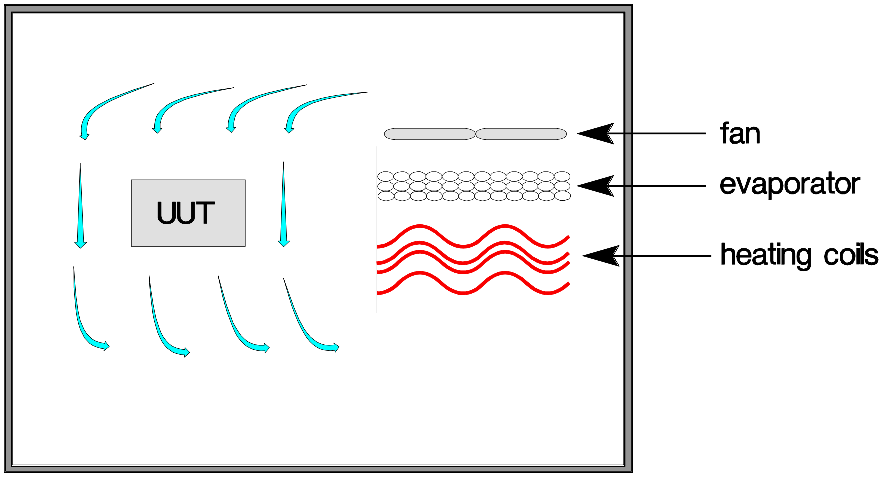

Figure 10 – Single zone chamber layout

An alternative to a single zone thermal cycling chamber is the use of a two zone chamber. Figure 11 shows a drawing of a two zone chamber. In this design, the chamber has one zone that is maintained at the hot temperature and the other zone maintained at the low temperature. In this arrangement, the product is moved between these two zones with the use of an elevator. When the product is transferred into the cold zone, the load the refrigeration sees is the floor and ceiling of the transfer basket and the UUT. Since the heavy refrigeration coils are never heated, the refrigeration coils do not contribute to the refrigeration load. These chambers are more expensive than single zone chambers with the same heating and cooling system but an air-to-air chamber with a cascade 30hp x 30hp system can achieve rates of temperature change exceeding 20°C/minute. The air-to-air thermal chambers are the most efficient design of thermal chambers with mechanical refrigeration which can achieve high thermal rates of change.

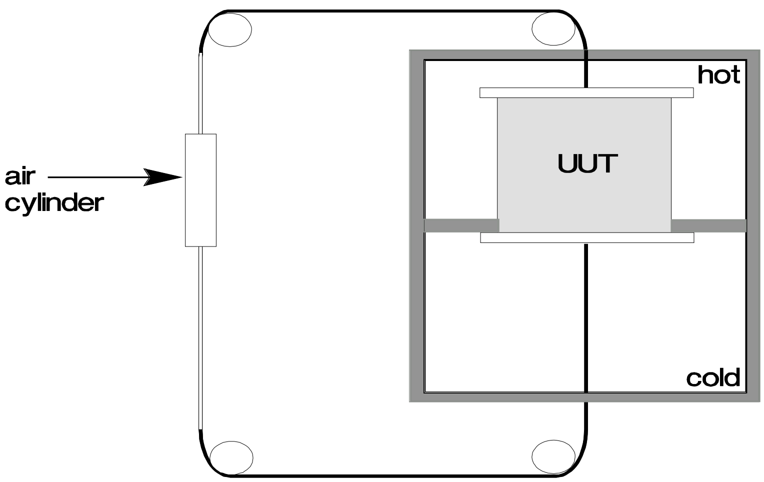

Figure 11 – 2 Zone air-to-air chamber

10.0 Conclusions

The need to perform product environmental screening to help ensure quality and reliability is undisputed. The methods to achieve an efficient and capable screening process can employ many forms.

The fault spectrum of modern electronic assemblies has shifted from component related defects to assembly related defects. The nature of environmental stress screening programs must ensure that they properly stimulate the types of defects that are expected to be found in production and in the field.

From many studies and the availability of stress screen equations, we are aware of the benefits of thermal cycling over all other forms of environmental stress. We are also aware of the need to subject the products in the screen to high thermal rates of change to quickly and cost effectively stimulates latent defects to patent defects. In addition, the combined environments of thermal cycling with random vibration provide the highest probability of precipitating the most common defect types.

To minimize facility costs and operating costs, the use of two zone air-to-air chambers are an attractive choice for average rates of change between 20°C/minute to 30°C/minute. Performing power cycling during the thermal screen allows additional fault modes to be exercised and is more effective than performing the same screens in tandem.

If a superficial functional test can be performed during the thermal screen and time to failure tracked, we have an opportunity to optimize the screening duration. Using the Weibull techniques to establish the screening duration allows the screening time to adapt to the product failure behavior. This statistically tractable method provides the best guarantee that products of consistently high quality will be delivered to customers. The use of fixed screening durations, however, precludes the possibility of lowering the screening cost or, worse yet, not allowing the duration of the screen to be extended to ensure that all latent defects are precipitated.

After the precipitation screen, a lower level environmental stimulus (within operating limits) should be applied during a comprehensive functional test to serve as a detection screen.