Accelerated stress testing in a time-driven product development process

Yuan Lu\textsuperscript{a}, Han Tong Loh\textsuperscript{a}, Aarnout Cornelis Brombacher\textsuperscript{b,}\textsuperscript{*}, Elke den Ouden\textsuperscript{c}

\textsuperscript{a}Centre for Robust Design, National University of Singapore, Singapore
\textsuperscript{b}Eindhoven University of Technology, Faculty of Technology Management, Product and Process Quality Section, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
\textsuperscript{c}Centre for Manufacturing Technology Asia Pacific, Philips Singapore Pte Ltd, Singapore

Abstract

In order to compete in the market, companies have to produce the right products with a shorter time to market and at lower costs than before. Shorter time to market requires the product development process (PDP) to change the way of working from the classical ‘wait and react’ to anticipating and preventing problems as early as possible in the development process. This requires a new, and different role for the quality and reliability tests used. While in a classical PDP products could be tested when available from (pilot) production, a modern, time-driven development process requires optimisation long before larger series of products are available. Accelerated stress testing (AST) is a classical solution for the implementation of tests where product failures need to be activated faster (and cheaper) in a well-controlled environment at the early stage of the PDP. This paper reviews the classical AST strategy and some most recent AST strategies. It demonstrates that these accelerated test strategies are mainly based on generic lists of failure mechanisms and have only very limited relation with the actual failure rate curve of products. The theoretical background of the four-phase roller coaster failure rate curve is addressed and from this an alternative AST strategy is developed based on the relevant phases of the roller coaster curve using a concept called stressor-susceptibility analysis. A discussion, on the application of the proposed AST strategies and their impacts on the four-phase roller-coaster curves, is given at the end. (C) 2000 Elsevier Science B.V. All rights reserved.

Keywords: Reliability testing; Time-driven product development process; Validation test; Analysis test; Accelerated stress testing; HAST; MEOST; Stressor-susceptibility analysis

1. Introduction

1.1. New demands for PDP

In order to manufacture products for a competitive market, four main business drivers, function, quality, time and cost, become the cardinal concerns for the manufacturers. Increasing demands from the market requires more companies to adopt a systematic strategy to produce a better product in a faster and cheaper way. It is better to be the first to market with a good product than to be the last to market with the best product. Therefore, the product development processes (PDP) of companies, active in areas with strong demands on costs, quality and with a high degree of innovation, are faced with a number of, often conflicting, demands.

In order to meet these requirements different companies adopt different strategies [1]. Some
companies try to improve their time to market by shortening existing development processes. Although initially this may lead to short-term successes due to strong motivation of the people involved, in the long run these classical time-squeezed PDPs can suffer from a lot of problems (and therefore delay) in the back-end of the process.

Another approach, often applied in companies driven strongly by quality standards like ISO9000, is to stick rigidly to procedures. Although this prevents a lot of failures in the back-end of the process it will lead to lack of flexibility and lack of capability to introduce new technology [1].

A third approach to reduce time to market is the use of sophisticated tools in the design process such as QFD, FMEA, FTA, etc. Without these tools embedded as part of the PDP, however, the authors have observed that, without adequate input data and without feedback loops on the results of these tools in the PDP, the added value is questionable.

At the same time, the importance of product quality has been changed from ‘nice to have’ to a prerequisite during the last 10 years [1]. If a product does not have a competitive quality, it does not make sense to produce it for the market. The requirements on product quality have also changed over years. The end users have a higher expectation on the product performance of both functionality and quality. There is also a trend to extend warranty periods for consumer products.

All of these new trends in the market have added new demands to the PDP. It is supposed to be able to make changes as early as possible to meet the shorter time to market and low cost requirements. It has to move from the reactive way of working to the proactive way, i.e. anticipate and resolve problems early. Concurrent engineering is widely promoted under this circumstance [2]. Although many people perceive Concurrent Engineering as ‘just doing things in parallel’, the main aim of concurrent engineering is to identify and resolve risks early in the PDP, in a phase of maximal flexibility where design changes cause minimal delay.

1.2. New demands for quality and reliability testing

These changes in the PDP have also a major impact on the role of testing in the development processes. A classical test satisfies, when a product prototype becomes available, the ability of the product to meet its specifications. In a classical, sequential, development process functional aspects are validated in the design phase, aspects of production are validated in the (pre-) production phase and aspects of customer use are validated at the moment when a larger quantity of products becomes available for the first time. This causes, however, a direct conflict with the concurrent engineering development process because in such a process potential problems, found normally during product tests, should have been identified already during the earlier phases of product development.

In a concurrent engineering process therefore it is possible to identify two classes of tests: analysis tests and validation tests. An analysis test is a test where potential critical aspects of a product are tested early in product development. The test is ‘failure oriented’ and aimed at confirming or rejecting risks that may or may not happen in a future product. The test product bears only limited resemblance with the final product; it is just made to identify and resolve risks. A second class of tests, later in the PDP, consists of validation tests. These tests are ‘success oriented’; they aim at confirming the already predicted (good) performance of the product.

In this respect, classical reliability testing, which is under the previous definition usually a validation test, serves to validate the product quality performance against its specifications. It is usually conducted at the late stage of the PDP where a product prototype with a certain level of final product details is available. The intention is to have as few failures as possible. The disadvantage of a validation test is that it is always reactive and never proactive. If a serious product reliability issue arises from the test, it could be too expensive and too late to introduce the necessary design changes. A common estimate is that a design change made one stage later in the design process will be a factor ten more expensive. In the highly competitive market today where many new technologies are occurring, fast time-to-market profitability is strongly encouraged and customers expect reliable products with a competitive price, there are multiple-fold risks that the PDP have to take into account. The validation test is no longer suitable to assist the PDP
to tackle all these risks. For example, the length of a product development cycle is originally half a year with 6–8 weeks for reliability testing. Due to the increased market pressure, the length of the development cycle has to be reduced to 4 months. Then, to perform the original 6–8 weeks reliability testing is no longer feasible to meet the reduced development cycle time. There is a strong need to supplement validation tests by something giving information earlier in the PDP.

Testing the product at the earliest possible stage in the PDP is desirable because design improvements are least costly and times consuming when the design is still not definitively defined and fewer agreements with third parties (suppliers of tools, materials, etc.) have been settled [3]. An analysis test is recommended under this condition to be performed at the very early stage of the PDP where a product prototype with limited final product information is available. It is meant to invoke realistic product failures early in the PDP and to detect as many failures as possible. Then, control loops and follow-up actions could be taken to prevent and control those potential failure mechanisms. Finally, product design could be optimised.

For a time-driven PDP, both validation tests and analysis tests are required to optimise the product design through monitoring realistic and potential product failure mechanisms and to validate product design against specification to confirm the product’s quality and reliability. An obvious challenge is how to do analysis tests. It requires implementing analysis tests at the early stage of the PDP to stimulate the most realistic product failures that may happen a few years from now in the field. A classical solution is to use accelerated stress testing (AST) to find reliability failures in products already during product development. AST aims to invoke product failures faster and cheaper in house according to its strategy. A natural question that arises is whether the available AST strategies are effective and efficient enough to support the role of analysis tests in the time-driven PDP.

In the following sections, a review of the classical AST strategy is given first. Most recent AST strategies, Highly Accelerated Stress Test (HAST), Multiple Environmental Over Stress Test (MEOST) and Random Multiple Environmental Over Stress Test (RMEOST) are discussed. All the available AST strategies are shown to be irrelevant to realistic product reliability performance in the field. It is mentioned that a Maturity Index on Reliability (MIR) [1,7] analysis can help a company obtain reliable failure mechanism information from product field performance. The theoretical background of the four-phase roller coaster failure rate curve is discussed. The knowledge of the product reliability performance, four-phase roller coaster failures rate curve [4] and stress-susceptibility relationship [4,11] can be used to derive relevant test strategies for the time-driven PDP.

2. Classical accelerated stress testing strategies

2.1. Generic list

The classical AST tests products according to a commonly accepted test standard/generic list against the constant failure rate model. For example, products are tested according to IEC standard under the condition of vibration, shock, temperature, or transient for mechanical failure mechanisms. A typical example of using a generic list is Arrhenius law. Already in the early days of reliability analysis, people realised that for failure mechanisms dominated by chemical reactions there is indeed an option of accelerated testing. The reaction speed of chemical reactions is described by the so-called Arrhenius equation [5].

The use of Arrhenius equation in reliability engineering has had considerable impact. By elevating the temperature it becomes possible to carry out tests not requiring millions of products and test hours (product × hours) but only several hundreds. Fig. 1 [6] below gives a typical example of a relation between temperature and number of products × hours required in a test.

In this case, testing of components with failure rates in the order of magnitude of $10^{-6}$ at room temperature would require millions of product × hours. From an industrial perspective this means most likely a combination of a large number of products and a long period of testing time.
Fig. 1. Accelerated testing assuming activation energy of 0.6 eV [6].

Testing at an elevated temperature of 120 K higher than room temperature would reduce the $10^6$ product $\times$ hours to $10^2$ product $\times$ hours. This would result in tests that can be done in-house under controlled conditions involving much reduced numbers of products and testing hours.

It is obvious that the latter is far more attractive than the former. This could be one explanation why accelerated life-test has become very popular in industry. The use of the constant failure rate model in combination with applying Arrhenius law results in testing of components only using a small number of components at an elevated temperature. However, the use of Arrhenius law is valid only where the failure mechanisms of the products are determined by the components only and the failure mechanisms in the components are of a thermo-chemical nature. Researches have shown that this is in many practical situations not the case [7,8].

There are some known risks of using the generic list to conduct AST. It could be possible that unrelated failure mechanisms might be activated, irrelevant stresses might be used, or test results might be interpreted wrongly without the knowledge of the genuine failure mechanisms that may happen in the field. In other words, classical AST only tests for the constant failure rate with poor correlation to the actual product field performance.

2.2. More recent accelerated stress testing strategies

There are some very recently developed AST strategies, i.e. Highly Accelerated Stress Test (HAST) [9], Multiple Environment Over Stress Test (MEOST) [10] and Random Multiple Environment Over Stress Test (RMEOST).

- HAST: HAST tests products under a condition where a single stress is increased step by step from one level to another until failure occurs. From the test results, rating or derating strategies could be determined. A risk of carrying out this test is that irrelevant failure mechanisms may be activated. A challenge to HAST is to assure that the relevant failure mechanisms are to be stimulated. A drawback of applying HAST is that the possible interactions among different stresses which are valid in the field may be ignored.

- MEOST: MEOST tests products under a condition where several combinations of stresses are increased step by step until failure occurs. To perform MEOST requires the knowledge of relevant failure mechanisms in the field, stresses, and interaction among different stresses. It also requires many in-house experiments.

- RMEOST: RMEOST tests products under a condition where several combinations of the stresses are randomly applied. It requires knowledge of failure mechanisms and stresses. This test method has shown its advantage in the identification of the localised failures when limited information in failure mechanisms and their interactions is available. Product passes and failures are counted to determine rating or derating strategies. However, RMEOST requires even more experiments than MEOST.

From authors’ experiences, the use of HAST is very common in consumer industry. MEOST is still in the experiment stage. However, generic lists instead of the realistic failure mechanisms are used to conduct these tests.

It is observed that all the available AST strategies use generic lists and there is lack of distinct knowledge of failure behaviour of (sub-) populations of products. The use of such AST strategies in analysis tests would result in low predictability during the early stage of the PDP because irrelevant failure mechanisms may be tested under irrelevant stress profiles and test results may be mis-used to predict the product field performance.
The most important step is to identify the most relevant field failure mechanisms from field. This results in three questions.

1. How to obtain genuine product failure information from the field?
2. What is the genuine product reliability performance in the field?
3. What do we want to predict by using analysis tests?

Alternative strategies can be developed provided these three questions are answered.

3. Alternative AST strategy

It has been mentioned that the classical AST strategy does not correspond with the failure mechanisms from the field. An alternative to the classical AST strategy that arises naturally is to test the product against the failure mechanisms identified from the analysis of the physics of the field failures. This strategy is termed Physics of Failure strategy. It requires the knowledge of the relevant failure mechanisms, product susceptibility, product specifications, the interaction between different failure mechanisms, etc. The Maturity Index on Reliability (MIR) concept [1,7] and a stress-susceptibility concept [4,11] can help companies determine the relevant failure mechanisms from the field. A MIR analysis can help companies to obtain reliable product quality and reliability information from the field while a stressor-susceptibility analysis can derive the relevant failure mechanisms from the reliable information.

3.1. Modern reliability analysis

In many companies reliability analysis is performed as part of product development or marketing. Reliability figures are generated, for example, as part of the commercial specification of a product. Although industry uses a wide range of reliability analysis methods [11], most methods use standard component reliability data as input. In a situation where product reliability is predominantly determined by the reliability behaviour of components and the component reliability models reflect the behaviour of the respective components in the field, constant component failure rate model and parts count method can be used without serious risks. Researches, by Wong [8] as well as Brombacher [7], have shown that in several branches of electronics industry, especially in the areas with a high degree of technological innovation, the requirements mentioned above are not fulfilled. A roller coaster failure rate (Fig. 2) has been developed to replace the constant failure rate to model the product behaviour generally in the field.

In order to explain the roller coaster curve one of the authors has introduced the so-called ‘stressor-susceptibility’ concept [4,11]. This concept is based on the analysis of physical failure mechanisms in products. A stressor is defined as a physical stress influencing the quality and reliability of products while susceptibility of a product to a certain failure mechanism is defined as a probability function indicating the probability that the product will fail after a certain time under a given set of stressors. Although mathematically quite similar to load-strength analysis or stress-strength analysis [12,13] there are some differences:

- Stressor-susceptibility analysis uses four different phases instead of three phases to describe the failure rate or hazard rate curve of products (Fig. 3).
- Stressor-susceptibility concentrates strongly on the behaviour of (weak, extreme) sub-populations within a large batch of products.
As mentioned above the stressor-susceptibility concept of the failure rate curve uses four different classes of failures not all of which may occur (Fig. 4):

1. **Hidden 0-hour failures**: Products that arrive out-of-(customer)spec at the customer. These products have either slipped through final tests, have been damaged during transport or are used in an unanticipated manner. Although, theoretically, these failures should all be observed at the moment of commissioning of the product, complex functionality or delay in customer reporting can cause delay in observing and reporting a failure.

2. **Early wear-out**: For high-volume consumer products it is not unlikely that there are considerable differences between either any two items of a product or between how any two customers will use the same product. In some cases this can lead to situations where a distinct sub-population of products shows different reliability behaviour than the main population with respect to wear-out. Examples are products that are produced with internal flaws. These flaws can cause a far faster wear-out than the main population. In the failure rate curve these sub-populations can appear as one or more ‘humps’. These sub-populations are quite difficult to test during production because on the product level they initially perform according to specifications.

3. **Random failures**: Products are designed to be used against anticipated (‘normal’) user conditions. It is, however, difficult to anticipate and to design against all events to which a product can be subjected. External events with a strong ‘random’ character, such as lightning and mechanical shocks, can cause product failure at any moment in time. Although many of these events have a comparatively low probability, these rare events can always happen. In those cases where the likelihood of occurrence for these events is constant in time and constant over the product population the effect will be a constant failure rate.

4. **Systematic wear-out**: Many products, particularly mechanical products ‘also certain categories of electronic products’ show some form of degradation over time. Well-known time effects are corrosion of metals and increased brittleness of plastics. Although the level of degradation will be different for every product in a large population there will be a moment in time when the first product fails due to degradation and a moment when the last product has failed. At the moment in time where these failures start to dominate the failure rate curve it will lead to an increasing failure rate.

It is important, however, to emphasise that phase 1 and phase 2 failures occur in distinct sub-populations of products (or product users) and phases 3 and 4 are relevant for the entire product population. By deriving the genuine relevant failure mechanism information from field through MIR analysis and stressor-susceptibility analysis, there are two
common strategies that can be used to accelerate the stress in AST.

(a) Increase the probability of the extreme stress
One strategy is to increase the probability of the occurrence of the failure by increasing the frequency of the real but extreme stress condition of the field or by increasing the operation cycles of the product under test conditions given that the failure mechanisms remain the same for both the test and the field. Consider a light bulb. Its extreme stress occurs when it is switched on and off. Assume that this light bulb is designed for an operation life of 5 years and it is to be switched on/off twice every day under normal usage. If, instead, it is switched on and off thousands of times each day, then a failure can be expected to happen in a few days instead of months. The probability of the occurrence of failure is increased. This strategy requires the knowledge of the real but extreme stresses and the frequency of the occurrence of the extreme stresses. The advantage of using this strategy is that products are tested under real but extreme operating conditions and that translation from testing conditions to actual field conditions is comparatively easy. When knowledge on the frequency of occurrence of a dominant failure mechanism and the corresponding extreme stress levels are known the acceleration factor can be derived from the ratio between frequency of occurrence of events in the field and the frequency of occurrence during tests. However, an obvious difficulty appears as how to obtain realistic stochastic models of extreme stresses (Fig. 5).
(b) Increase the level of extreme stress
The other strategy is to increase the severity of the real but extreme stress given the failure mechanisms remain the same at both field and test. For example, high temperature or high humidity can weaken an adhesive mechanical bond in the field. If during the in-house test, the level of temperature or humidity is highly increased but the relevant failure mechanism, weakening an adhesive mechanical bond, still remains the same, the occurrence of such a failure would be increased. This strategy also requires the knowledge of the stresses and their relevant ranges to elicit the same failure mechanisms. It is very easy to perform in practice. However, a strong understanding on the stress severity is necessary to maintain the link between the test and reality (Fig. 6).

Based on the idea of the stress-susceptibility interaction [4], theoretically there is still a third strategy that could be used. It is to decrease the product strength so that normal stress acts like extreme stress. In this way, the probability of having a failure is also increased. However, such a test is very difficult to design and perform in practice. This test method is still in the theoretical research stage only.

A summary of the different AST strategies discussed here is given in Fig. 7.

4. Four-phase roller coaster curve and AST strategies
In this section, how to correlate the AST strategies with the four-phase roller coaster curve is discussed. In order to have a clear understanding of the common AST strategies used currently and their future trend, five high-volume consumer products manufacturers were invited for a discussion with some of the authors. The participants represent a variety of functional backgrounds. They are from Development, Quality Assurance and Production Departments. The main questions directed to them were

- What are the AST strategies used in your company? What is the basis of using these strategies?
- Do you use HAST, MEOST and RMEOST? How do you choose stressors for the test?

From their responses, it is learnt that the common AST strategy used by them is mainly based on generic lists. The generic lists are generated from various industry standards or company internal
standards. HAST is widely used by them while MEOST is still at the experimentation stage. However, the stressors for the test are identified from the generic lists not the real failure mechanisms. From the discussion in Section 2, we know that the use of generic lists in AST has its disadvantages as well as advantages. Then more questions were directed to them to find out how generic lists detect the realistic product failures effectively.

- What phases of the four-phase roller coaster failure rate curve are relevant to your products?
- What phases of the four-phase roller coaster failure rate curve do you want to test?
- What phases of the four-phase roller coaster failure rate curve are you testing?
- Do you have the knowledge of the statistical behaviour of the entire product population?
- Do you have the knowledge of the statistical behaviour of the relevant product sub-populations?

For all the companies, the responses were that phase 1 and 2 of the four-phase roller coaster curve are relevant. But the standard tests they apply are only able to measure phases 3. For consumer products, phases 3 and 4 are less relevant now than before. Due to improved technologies, product quality has improved and faces fewer problems in phase 3. There are not too many phase 3 and 4 problems. At the same time, they are also economically less important due to rate of technology development, as new products come into the market to replace old ones before they reach phase 3 and 4. Phases 1 and 2 have become relatively more and more important. From previous section, we have known that only certain sub-populations of products contribute to phases 1 and 2. However, the companies’ knowledge of the sub-populations is limited. Further discussion was held to identify areas of improvement for the correlation between the AST strategies and the relevant phases of their products. The common steps that the companies are willing to commit resources for improvement are

- Identify sub-populations from phase 1 and 2.
- Identify relevant realistic failure mechanisms.
- Define alternative AST strategies based on the realistic failure mechanisms.

The discussion results are summarised in Table 1. From this case study results and discussion in previous sections, a general approach to correlate the AST strategies to the four-phase roller coaster curve can be developed.

As discussed earlier only certain sub-populations behave according to phases 1 and 2. The classical AST strategies will not be able to detect them. To identify these two phases, the most important step

<table>
<thead>
<tr>
<th>Questions</th>
<th>Companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>AST strategies used</td>
<td>A</td>
</tr>
<tr>
<td>Based on what information</td>
<td>Generic lists</td>
</tr>
<tr>
<td>Use of HAST, MEOST, and RMEOST</td>
<td>Standards</td>
</tr>
<tr>
<td>How stresses are chosen</td>
<td>Generic lists</td>
</tr>
<tr>
<td>What phases of the four-phase roller coaster failure rate are relevant to your product</td>
<td>1 + 2</td>
</tr>
<tr>
<td>Which phases of the four-phase roller coaster failure rate do you want to test</td>
<td>3</td>
</tr>
<tr>
<td>Which phases of the four-phase roller coaster failure rate do you test</td>
<td>Yes</td>
</tr>
<tr>
<td>Do you have the knowledge of the statistical behaviours of the entire product population</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1
Case study results
to go is to identify the relevant sub-populations from reliable field information. Then, the root cause of these sub-populations has to be found. From there, appropriate AST strategies can be developed to test the relevant failure mechanisms. Though phases 1 and 2 are both early product failures, they are different intrinsically. Phase 1 failures happen when the product does not fulfil customer specifications, or not caught by the test program or just fail after tests are performed. The idea is to pick out this sub-population under the test program, in which the relevant stress is applied before sending them to the customer. Phase 2 failures happen when a sub-population that is made of weak materials, comes from an unstable production line, and is misused by customer. It is very difficult to test them in production since the failure can only be triggered when high stresses are applied. The appropriate approach is to build special analysis prototypes for this sub-population and test them by either increasing the stress levels or by increasing the likelihood of the extreme stresses at the early PDP.

Phases 3 and 4 product failures are related to the entire population. Phase 3 failures are mainly due to customer use. Operating condition that mimics extreme customer behaviours are simulated on prototypes to increase the likelihood of detecting product failures in the early PDP. Then, design improvement can be done and reliability of this product can be improved. Phase 4 failures are mainly due to the natural wear-out of design and material and due to customer use. Analysis prototypes that represent the entire population are to be tested under the elevated identified relevant stress levels at the early PDP. It would lead to an optimisation of the product design.

The ideas given in previous section can help to identify the relevant sub-populations and relevant failure mechanisms and to design AST strategies. Early design changes can be made to improve product performance and prevent quality and reliability problems later.

5. Conclusion

In this paper, we have discussed the new trends in reliability testing in a time-driven PDP. A fast and balance reliability-testing scheme is required to invoke realistic product failures at the early stage of the PDP. AST is a classical solution to the requirements. Currently available AST strategies do not take into account the four-phase roller coaster failure rate curve, but only the constant failure rate. From the authors’ experiences, early failures are especially relevant for most consumer products. Systematic strategies for testing all the phases of the failure rate curve can be derived based on MIR analysis, the knowledge of the four-phased roller coaster curve and the stress-susceptibility concept.

References