Int J Adv Manuf Technol (2005)
DOI 10.1007/s00170-004-2181-9

ORIGINAL ARTICLE

Nathan Gnanasambandam · Anthony Primavera · K. Srihari

The reliability prediction of electronic packages – an expert systems approach

Received: 21 January 2003 / Accepted: 15 March 2004 / Published online: 25 May 2005

Abstract The exponential growth of the electronics packaging industry has fueled the availability of a variety of area array packages. The reliability of these packages, as characterized by their capacity to withstand the IPC- (formerly Institute of Interconnecting and Packaging Electronic Circuits) prescribed swings in temperature, differentiates one from the other. With design cycles shrinking and competition surging, the capability to make instant package selection decisions by leveraging prior empirical data could pose as a potential alternative for exhaustive experimentation. By employing expert systems techniques, this research developed suitable models that accurately depict field conditions in order to assist in delineating trends in package reliability data.

The fatigue behavior of the solder joints subjected to accelerated thermal cycling is often used as an indicator of the reliability of electronic packages in field conditions. Design for reliability (DFR) could be pursued if the thermal fatigue behavior can be predicted in the design phase of a product. The finite element method (FEM) and accelerated testing such as air-to-air thermal cycling (AATC) have been used extensively to study second-level package reliability. Factors like incorrect assumptions or unknown material properties involved in the development of the FEM models are the cause of deviations between actual and predicted values The mathematical complexity and the time needed for model development further aggravate the situation.

The focal point of this research was to develop a generic method that could be used to predict the second-level solder joint reliability of area array packages from the analysis of empirical data. While package characteristics play an important role in identifying the similarities between various subsets of packages, the role of assembly parameters is crucial in terms of their impact on reliability. Weights, in terms of the parameters' impact on reliability, are computed by examining each individual experiment. Based upon identified trends and the separation of qualitative and quantitative impact of the contributing parameters, regression models may be developed to capture the second-level reliability behavior of the package. These models make it possible to predict reliability and potentially save time and resources for an end-user.

N. Gnanasambandam · K. Srihari (✉)
Department of Systems Science and Industrial Engineering,
State University of New York,
Binghamton, NY 13902-6000, USA
E-mail: srihari@binghamton.edu

A. Primavera
Surface Mount Technology Laboratory,
Universal Instruments Corporation,
Binghamton, NY 13902-0825, USA

Keywords Accelerated testing · Electronics manufacturing · Expert system · Knowledge-based systems · Reliability prediction

1 Introduction

Evaluating the reliability of electronic packages that have been assembled on to printed circuit boards (PCBs) through accelerated methods generates huge volumes of data. These data sets could be effectively leveraged to understand their behavior vis-à-vis second-level, or solder joint, reliability. The estimated reliability of an individual electronic component leads to the identification of critical issues relating to the entire PCB assembly, thereby influencing many financial and marketing decisions that relate to the product. Hence, the investment of both time and resources in a well-structured reliability enhancement program is adequately justified by the prospect of huge tangible as well as intangible losses in the event such research is overlooked at the product development stage. Losses may range from the cost associated with returned or reworked goods to a drop in customer goodwill. Therefore, it is imperative to build suitable models or conduct extensive tests that provide an upper bound on the expected second-level (solder joint) reliability.

Many prediction mechanisms and mathematical models have been proposed to parameterize the behavior of the solder joint under thermal cycling [1–3]. Air-to-air thermal cycling (AATC) of the test assembly, especially in the case of chip scale packages

(CSPs) and ball grid arrays (BGAs), is an irreplaceable baseline that is used to compare the accuracy of the aforementioned models. However, using thermal cycling as a measure of the accelerated test performance is fraught with many difficulties, such as the time involved and the gamut of experiments that need to be conducted to reach a statistically significant conclusion. The rapid development and evolution of electronic packages has resulted in a wide variety of packages being available, making the choice of a reliable package increasingly difficult for a design engineer. The prolonged process of testing may have to be repeated for every new package. However, this may not be the most efficient alternative to get a quick estimate of reliability and its associated trends.

The finite element method (FEM) has been exhaustively used to capture the dynamics of solder joint failure through fatigue. The volume of the solder joint is divided into many elements, for which constitutive equations have been derived for inelastic strain energies caused by cyclic thermal stress conditions that eventually lead to fatigue. More information about these studies could be found in [1, 4]. Though these models are cost-effective, they may be of varying complexities involving several assumptions. One common assumption is that the material properties of the solder joint do not vary during thermal cycling. The ranges considered on strain or strain energy may sometimes be insufficient. This is coupled with the use of complex differential equations that govern the mechanics of each considered element. The calculations involved need to consider effects from neighboring solder bumps and the interactions that may influence failure conditions. This translates into a problem with huge computational complexity that takes several hours to solve on high-end workstations. Further, several workstation models may neglect micro-structural growth as well as crack growth and propagation. The model-building process is also theoretically complicated and requires several man-hours of domain experts before the model is set on a platform to run. Modeling is iterative in nature and needs continuous monitoring by experienced analysts who have the capability to interpret the results and fine tune the model [5].

Other modeling techniques that have been documented in literature use failure data from cycling experiments performed on several component types to extract the correlation between inelastic strain energy and cycles-to-failure per unit cross-sectional area of the solder joint [6–9]. This innovative procedure has been compared with traditional FEM models and is known to produce similar or close results. But the main drawback is that the model building process mentioned above can only be done by domain experts. Though the model may be a fast turnaround design tool for known components, it may by itself not be adaptive. Hence, it may require additional work and experimentation for use with newer components. However, the question remains how to quickly assess the reliability of an electronic assembly based on the data available from prior experimentation with the help of observed similarities in terms of the component characteristics and the parameters of the experiment. The approach described in this paper uses data-mining-assisted trend- extraction algorithms on cataloged data and generates reliability prediction using adaptive expert system architecture [10].

2 Problem statement and research objective

The explosive growth of the electronics packaging industry has made a whole gamut of area array packages available in the market. As a result, engineers across different rungs in this field are faced with the significant challenge of estimating the reliability of electronics packages, and selecting those that best suit their needs. One alternative is to exhaustively test every possible package. However, this approach may be inappropriate for some circumstances. On the other hand, an extensive variety of packages that consistently exhibit similarity to one another, both in terms of construction and reliability behavior, have been evaluated over many years at the facility where this research was conducted. The objective of this research effort was to utilize the comprehensive data bank that was available on the second-level reliability of area array packages in conjunction with analytical techniques to form the knowledge base of a cost-effective, "intelligent", predictive tool that could potentially save time for a large number of engineers in this field. To this end, an adaptive expert system was built to assist decision-making by predicting the second-level solder joint reliability for engineers in the field of design, manufacturing, reliability testing, and failure analysis. This expert system can:

1. Support input and storage of component characteristics, experimental and test conditions, failure data, and test status;
2. Conduct failure analysis and estimate distribution parameters;
3. Predict the reliability of a test component (provided the test sample has similar attributes to the samples that exist in the database);
4. Identify deficiencies the available data set, thereby assisting with the planning of subsequent experiments;
5. Automatically adapt the knowledge base on reliability trends for new components;
6. Provide a mechanism to compare the relative importance of experiment/component parameters;
7. Make available a suitable testing ground for "what-if scenarios" (e.g., what is the effect of a 1.57-mm (62-milliinch) substrate versus a 2.36-mm (93-milliinch) substrate on Package A's reliability? How does an increase in substrate pad diameter affect reliability?).

3 Research methodology

The expert system developed in this research works in conjunction with several assisting algorithms that discover and enhance reliability-related knowledge by using heuristic techniques on the experimental data. Knowledge that is extracted is fed back into the system, so that the system can update its configuration and use the latest information for further analysis. This imparts intelligence to the expert system besides making it adaptive to changes in the knowledge base. For any given component (assuming a suitable "database match" exists), the system isolates a similar component, weighs the important parameters that affect reliability, identifies a similar experiment for the matching component, and uses it as a baseline for scaling the reliability of the requested parameter. The expert system described in this paper provides a generic methodology that can be adapted to other systems with very few changes. However, as far as this research is concerned, the application area is focused on the domain of electronics packaging.

The reliability prediction methodology has been programmed into a user-friendly software module. This module is part of a larger expert system that addresses various other aspects including assembly yield and board routing. This paper focuses on the prediction sequence and hence does not discuss the software engineering aspects of the program. Also, this paper uses the terms accelerated test performance and reliability interchangeably, but all references relate to the former.

3.1 Reliability testing

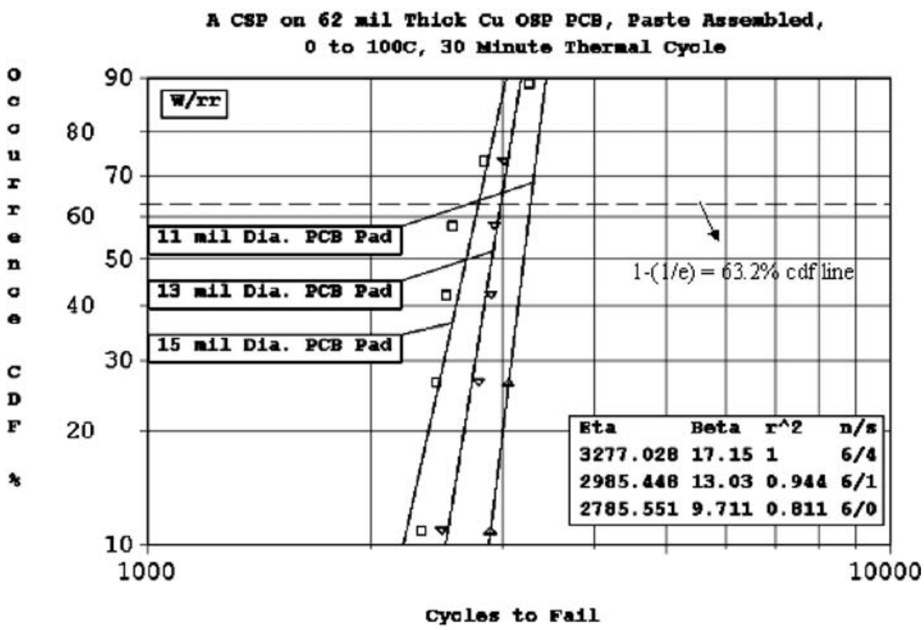
In order to estimate the reliability of products, samples are typically tested in laboratory environments under prescribed conditions [1]. In most cases, it may take a time for the product to fail. To accelerate the test failure, samples are often subjected to temperature/mechanical stress regimes similar to actual field conditions, however on an abbreviated time scale. Depending on the area of application, methodologies or scaling factors that correlate laboratory and actual life estimates may be available.

The first step in creating a reliability program is to identify the appropriate test environments. Careful consideration has to be given to the actual operating conditions of the product. Harsher operating conditions may necessitate more strenuous tests. In accelerated testing, a predetermined number of products are tested while the electrical connectivity of the assembly is monitored by means of computerized data logging equipment. The sample size of the test is often small because the testing is destructive in nature and expensive in terms of the test equipment. As and when failures occur, the time to failure (in cycles) is noted. The test is continued until a predetermined number of all products fail or until a set number of accelerated cycles have been completed. In either case, there need to be enough failures to utilize mathematical methods such as maximum likelihood estimators (MLE), least-squares method, Weibayes method, and the Kaplan-Meier method to obtain a statistical estimate of the characteristic life of the product [11].

Failure data can often be fitted to a suitable probability distribution to estimate the product life. Failure mechanisms may be classified into time-zero or early failures (attributed to manufacturing defects) and fatigue (wear out) failure.

Additionally, literature suggests a third mechanism, which is the transitioning of a solder system from a ductile fatigue failure to an interfacial solder to the attachment pad [1]. Results should be screened rigorously to eliminate early manufacturing defects, metallurgical non-solder-fatigue failure, or mixed-mode failure [11, 12]. For this study, only fatigue failures were included in the analysis. While the log-normal and Weibull distributions fit fatigue and wear-out mechanisms well, this research uses the Weibull distribution throughout [11]. The distribution parameters are calculated using the failure points, and a graph of the cumulative distribution function versus time (in cycles) is plotted. From the graph, the mean life or characteristic life may be computed. For the Weibull distribution, the 63rd percentile (cdf = 63.2%) denotes the characteristic life (eta) of the assembly [11]. The R^2 parameter denotes the closeness of fit and is required to be as close as possible to 100%. The shape parameter (beta) typically has a value greater than 2 for fatigue failure and indicates that the failure rate increases rapidly as time increases. Figure 1 is an example of a Weibull plot for a component

Fig. 1. Reliability plot of a CSP after thermal cycling



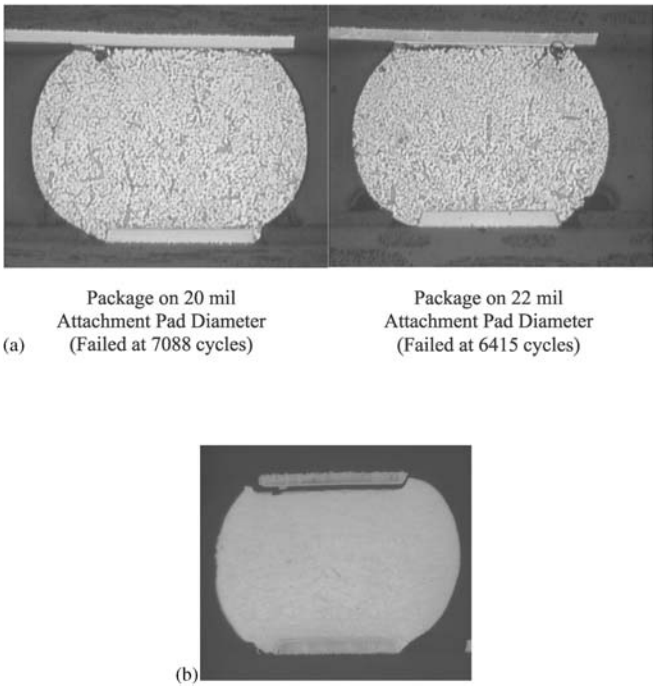


Fig. 2. a Examples of fatigue failure for different PCB pad sizes **b** Example of a non-fatigue failure

that underwent reliability testing to evaluate the effect of pad diameter.

3.2 Reliability of electronic packages

Understandably, the reliability of electronics packages depends on the length of time that the entire component, including the solder, can maintain electrical connectivity to the substrate pad when subjected to testing conditions. The package design, assembly parameters, severity of the test, and material properties of the solder together determine the life of the solder joint. The electrical connection may be broken due to various failure mechanisms such as fatigue, delamination, and embrittlement of the solder. Some examples of typical failures are shown in Fig. 2. The focus of this research is however on failure due to fatigue. Figure 2a shows the impact of attachment pad size on reliability, whereas Fig. 2b shows a non-fatigue failure. Time zero defects are ignored in this study because they may be due to manufacturing defects or handling issues.

This research emphasizes the development of a generic method that could be used for any AATC testing if sufficient data exists. In particular, this research interpreted the extensive information from second-level reliability testing of area array devices to predict second-level solder joint reliability. The packages under consideration were subjected to a variety of IPC-prescribed thermal regimes such as 0 to 100 °C and -40 to 125 °C to study the impact on their reliability [13]. The systematic compilation of package characteristics, assembly conditions, and testing parameters for every group of failure cycles forms the basis for this analytical technique described in detail in Sect. 4.

3.3 Generic approach

The approach adopted in this research was applied to predicting the accelerated thermal cycling performance that may be related to the second-level reliability of assembled electronic packages. However, the generic nature of the approach makes this method viable for predicting the reliability of other materials or assemblies. Data from tests and experiments are the primary driver in this research and hence this method is most suited to studies that have extensive data available. Though dormant for many years, data mining methodology has come to the limelight because of advancements in computer hardware and software [14, 15]. This research uses data mining to enhance the analytical ability of expert systems and to discover new knowledge from prior experimentation. In principle, this method could be adjusted minimally and used for the prediction of reliability of say, bolts or rivets, provided data is available.

4 Reliability prediction theory

From the point of the test board design to the eventual analysis of results, data constantly changes form because of processing at every stage. More importantly, since the amount of data is staggering, one can decide to build a data warehouse model (a highly advanced data storage mechanism favoring the efficiency of analytical methods) or a relational model to tackle storage issues. Since the relational model is more appropriate and easier to develop, it has been given priority over the cost-prohibitive and time-consuming warehousing model [16]. The prediction sequence (discussed Sect. 4.5) is preceded by several carefully planned stages that primarily deal with experimental design or the handling of data. The sections below provide a detailed description of the different steps adopted in this implementation.

4.1 Design of experiments (DOE)

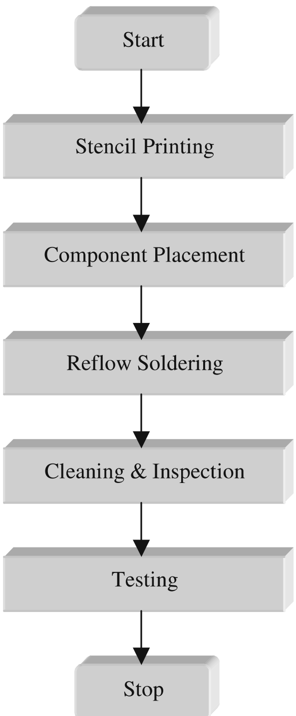
Usually, second-level reliability testing for a package is a time-consuming process that follows the assembly process sequence illustrated in Fig. 3 [17]. Since assembly parameters and board characteristics play a vital role in determining the reliability of the end product, the board design and DOE must encompass the whole gamut of possibilities or experimental levels for each component to bring out the influence of these parameters. This facilitates the identification of a trend or relationship that could be linear or polynomial, single or multivariate. Often, sample size limitations and cost considerations conflict with the above guideline. However, the prediction depends heavily on the existence and completeness of data from experimentation. Hence, in most cases, a trade-off must be reached between the costly and destructive nature of reliability-related experimentation, and the need to select a sufficient sample size (usually around 32).

A number of factors impact reliability and they may include component or PCB characteristics, process parameters, workmanship issues, equipment used, and field conditions. While it may not be possible to gauge the combined impact of all these factors, designed experiments may be run under controlled conditions with a manageable matrix of significant parameters. Based on prior experimentation and literature, a set of factors such as attachment pad size, substrate thickness, pad thickness, and finish were chosen [18]. Full factorial experiments with four to five levels were then run on specific classes of packages. After the assembly, thermal cycling conditions such as the ramp rates and the dwell times were also included in the experiment because of their known impact on cycles-to-failure. Additionally, component characterization that was conducted prior to experimentation yielded a statistical perspective of several design parameters. Consequently, some significant design parameters, process parameters, and testing conditions were included in the framework of this study. The impact of equipment and workmanship on reliability was ignored, primarily because most of the assembly was carried out in a single facility.

4.2 Collection of data

This process involves the collection of data regarding components, assembly conditions, and failure. The task of package characterization precedes the assembly sequence (Fig. 3) and deserves a thorough treatment as this information is used by the expert system to identify a baseline to the prediction mechanism. All possible process variables need to be tracked during the assembly sequence because prior knowledge of key factors governing reliability is not known. For example, the board characteristics such as thickness and pad diameter may play a more dominant role than other parameters such as pad finish in determining reliability. The assembled boards are then put into reliability chambers for AATC. Test regimes (e.g., 0 to 100 °C, −40 to 125 °C) are outlined in JEDEC test specifications, which exist for testing electronics packages for different field conditions [13]. New reliability test results that continuously keep arriving need to be linked to assembly information and maintained in spreadsheets or databases so that further consolidation and refinement of the data can be performed.

Fig. 3. Surface mount assembly sequence



4.3 Preliminary cleaning and storage

Second-level reliability data is voluminous and, due to the sensitivity of the test equipment, often electrical noise-prone in nature. Consequently, in this scenario, classification is key to the organization of data. This step involves careful scrutiny of the data in the spreadsheets and the temporary databases in order to remove noise (i.e., false electrical test data) or redundant information so that it can be moved to permanent storage and analyzed. Categorization enables reuse and saves valuable database space. Experiment numbers need to be allotted for every unique combination of component type, assembly type, and test type. All experiments are potentially comprised of multiple types of failure and hence the definition of each failure type needs to be clearly delineated. For all datasets in this research, solder joint failure was detected by measuring resistance. In most cases, electrical failure was the first intermittent failure (300 ohms in-situ for a minimum duration of 200 ns) monitored by an event detector. This is a slight modification of the IPC-SM-785 test specification which suggests 1000 ohms for a duration of 1 ms. The study of reliability and solder joint failures in particular is characterized by the bathtub curve shown in Fig. 4 [12]. Infant mortality failures occur early in the life cycle and may be attributed to poor quality assemblies or poor component quality. Wear-out failures, on the other hand, occur due to the build-up of stress. While considering failures, the occasional infant mortality or random failures need to be carefully isolated from fatigue failures.

4.4 Reliability analysis

After the data is“cleaned”, it needs to be analyzed to determine the statistical characteristic life of the package, which is used

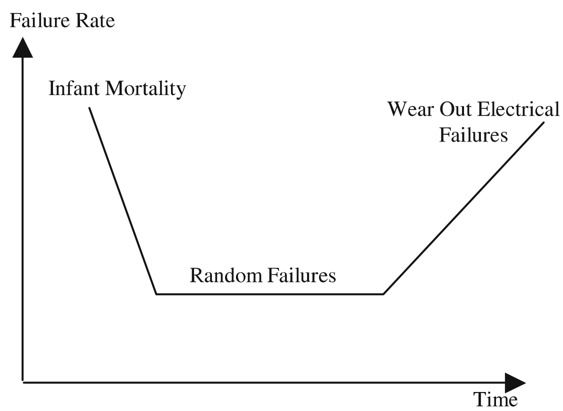


Fig. 4. The bathtub reliability curve

for further reliability prediction. Some literature suggests the use of failure free time and the three parameter Weibull distribution as opposed to log-normal or the two-parameter Weibull because they obtained a better fit for early wear out failure across a large database [9]. However, the primary objective in this research was not to determine failure free time, but rather to use the characteristic lives obtained from several hundred experiments to formulate a trend prediction algorithm. Hence, all characteristic lives ("eta" – N_{63}) were obtained using the two-parameter Weibull distribution. The fit (R^2) in most cases was above 95% with the possibility to rule out experiments that did not meet specific criteria such as sample size or fit.

4.5 Prediction mechanism

At this stage, all the prerequisite steps towards the actual prediction have been completed. This section concentrates on the processing sequence (Fig. 5) adopted to predict reliability from the extensive database and describes how each segment conforms to the expert system architecture.

4.5.1 Identification of matching component

By comparing the requested user input relating to test component characteristics with the existing database entries, an appropriate component needs to be isolated. All experiments performed on this match (or other close matches) may serve as the search domain for trend extraction. This step facilitates the data mining activity by providing a list of components characteristically similar to the selected input component. Regression modeling using data sets relating to existing components is automatically accomplished [19]. Although the most common use of the algorithm (discussed in this paper) is the sensitivity analysis of the thermal cycling performance of characterized packages, the user may often be interested in predicting the reliability of a new or unknown component. As long as the new component exhibits similarities to existing packages, a metric for similarity can be defined. The metric used in this approach involves a penalty-based ranking scheme that is explained below. For each parameter, the deviation of the test component from each component in the database is calculated one at a time. Some parameters may be more important and may have a dominant effect on reliability. These parameters are differentiated from one another by weights. The weighted deviation of every parameter is ascertained as follows.

$$P = \frac{\sum w_i \frac{|u_i - d_i|}{u_i}}{\sum w_i} \qquad (1)$$

where P is the cumulative penalty, w_i is the weight of the parameter i, u_i and d_i are the values in the user input and database respectively. This is repeated for every package in the database and the cumulative penalty is stored temporarily. An ordered listing of all available components is then made available to the user. The deviations are arranged in increasing order (of P) thereby resulting in a decreasing order of similarity. The user may then exercise discretion to choose a specific component. Based on the number of packages available, the module may come up with several matches that the user can subsequently review and narrow down.

The deviations calculated above relate to a specific class of components. For example, a new CSP may be compared to other CSPs rather than BGAs or peripheral leaded quad flat packs (QFPs). Another important aspect in this stage is that variables may be qualitative or quantitative in nature [10]. Qualitative variables are difficult to handle, but the penalty calculation is concerned only with a binary situation such as "exact match"/"no match" between the test and database records. If an exact match is obtained, then deviation is at a minimum (or zero); it is a constant positive quantity (say 1) otherwise. The weighting step (Sect. 4.5.3) also deals with qualitative variables, but in a different fashion.

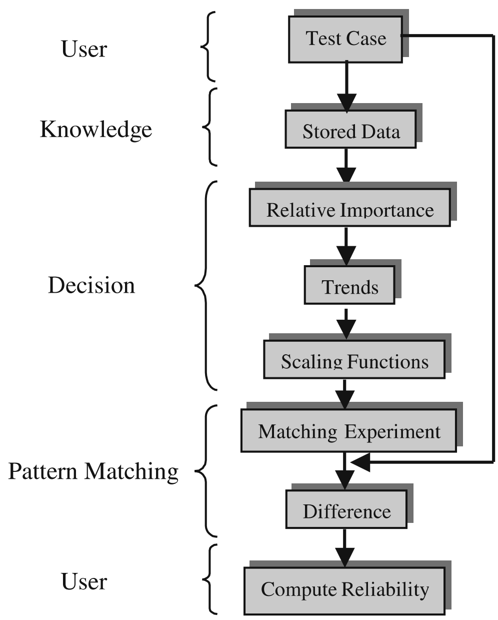


Fig. 5. Expert system view of the overall sequence

4.5.2 Screening of variables

This approach tracks several variables during various stages of manufacturing because of their potential impact on reliability. The expert system would then scan through the data and quantify the extent to which each variable affects reliability. In order to narrow down the variables for the first iteration of expert system construction, knowledge from prior experimentation could be used. It is sound judgment to try the prediction mechanism on a few parameters first so that it could be extended later in an iterative fashion. Table 1 depicts a list of some important parameters that were chosen on the basis of knowledge gained through experience.

4.5.3 Weighting

"Weighting" deals with the quantification of the relative importance of various parameters. In order to perform this operation,

Table 1. List of parameters that affect reliability

Assembly parameters	Thermal cycling parameters
Quantitative • Motherboard thickness • Pad diameter Qualitative • Paste/flux process • Pad finish • Solder metallurgy • Reflow atmosphere	Quantitative • Start temperature • Final temperature • Ramp • Dwell • Cycle time

variables need to be classified into qualitative or quantitative types. For quantitative variables, the percentage change in quantity is related to the percentage change in reliability. Changes are computed between the user's case and the database experiment, as shown by the example described in Table 2. Weights are computed across all experiments belonging to each separate class of components. The average percentage change may then be computed for every parameter. This identifies the parameter that has a greater impact on reliability.

In the case of qualitative parameters, percentage changes cannot be computed. However, one can note the average reliability change for a particular transition in a qualitative parameter. The user's case and database experiment need to be compared on a given parameter, such as pad finish, to determine transitions that change reliability to relatively greater extents. By considering all the records in the database, a look-up table may be dynamically built. This aids in weighting the parameter for a specific user request.

4.5.4 Identification of matching experiment

Based on the component match and the input from the user regarding assembly and test conditions, an experiment needs to be chosen as the baseline. This experiment is the combination of various parameters that have the closest resemblance to the one required by the user. The combination of parameters may not exactly match the search requirements and, as expected, they may be off on some parameters. In order to bridge the gap between the search result and the user's case, "scaling functions" (described later in Sect. 4.5.5) are used.

There may be multiple experiments that come reasonably close to the search criteria and often, the number of experiments obtained is related to the search order. The order shown below has been used preferentially while searching through the database in this research. It can be altered on a case-by-case basis depending on the weights computed before.

Motherboard thickness → pad diameter → thermal cycle → paste/flux process → pad finish

This order may seem rather arbitrary but it is based on experiments that were performed which suggest, for example, that pad diameter and motherboard thickness may have a greater impact on reliability than pad finish. However, the search order is fine-tuned based on the exact compilation of the entire dataset. Consequently, the adaptive nature of this methodology can be utilized in future iterations of expert system development.

At the end of the match step, a list of experiments that exactly match/closely match the search criteria may be identified. The user may choose to use one of the search results to base the prediction on. Scaling is performed with respect to the nearest match identified by the module and approved by the user. In some cases, the search may not yield any results. This may be indicative of the lack of datasets that closely match the required criteria.

Table 2. A pair of datasets (all values refer to board thickness)

Value in mm	Accelerated thermal cycles (reliability)
1.57 mm (62 mils)	7000
0.79 mm (31 mils)	16 000
Change in value in mm	**Change in reliability (ratio)**
0.79 mm (31 mils) decrease	2.29 increase

(For the same package)

4.5.5 Development of scaling functions

This is the heart of the expert system where trends are extracted from qualitative as well as quantitative parameters. An example of a trend may be, within a certain range for area array components, an increase in PCB pad diameter resulting in a decrease in AATC solder joint reliability. Such trends constitute the new knowledge that is discovered by the expert system. However, the challenge resides in quantifying the importance of each piece of information that is discovered and utilizing it towards prediction.

Based on the knowledge that has been gathered and stored during the weighting process, trends on quantitative variables and their effect on reliability is examined first. Further, the stored information has to be revisited and examined for statistical significance. The trends relate to a change in a quantitative parameter and the corresponding change in reliability. This analysis is carried out one variable at a time, keeping all the other variables constant. Table 1 shows how one data point relating to a "change-factor" is obtained. This is repeatedly performed to identify several points that can be modeled using regression. Simple linear regression (in one variable) is carried out and the least squares best-fit polynomial is computed. The quantitative difference between the user's scenario and the database match in any given parameter is used to predict the change in reliability. When the user's case exactly matches the database record for the parameter being considered, scaling need not be performed. This is used to predict the change in reliability for the quantitative difference in one parameter caused by the dissimilarity between the user's case and database experiment. In cases where an exact match is obtained, the scaling function need not be used at all.

Regression in one variable may lead to some interactions between parameters being ignored, but what complicates the situation is that they might turn out to be more than just binary or ternary interactions. Such a situation is difficult to deal with and may possibly involve a dynamic approach that models a combination of the important variables (as determined in the weighting phase) using multiple regression. For this experimentation, this was viewed seen more as an improvement and was not considered in the initial stages of the expert system development.

The aforementioned scaling functions, developed using regression, are relevant only to quantitative variables that lend themselves to mathematical manipulation. For qualitative variables, the different cases seen in the database can be adapted to form semantic rules that may assist in prediction. As pointed out before, data sets need to be considered in pairs to see a change in the quantitative parameter and the corresponding reliability change. Likewise, a pair of data sets could help in formulating a rule for an arbitrary transition in the state of a qualitative parameter. Obviously, the state of the qualitative parameter becomes important here and hence keeping all the other parameters constant between the pair of datasets helps to remove ambiguity. Then, it can be stated that the change in the qualitative parameter uniquely causes the observed effect on reliability. Interactions with other variables are mostly ignored because they seem to have a small impact on the prediction. In some cases, the interactions may be the reason that causes a deviation between the actual and predicted value of reliability. Additional research is needed to automatically factor in the more prominent interactions. However, for qualitative variables, rules may be formed and stored based on many instances of actual transitions from one state to another. This is illustrated in Table 3.

In the context of this paper, the rules or models that help calculate the change in reliability for a change or transition in a qualitative or quantitative parameter are called a "scaling functions." The reason they are called scaling functions is because they help scale the gap between the user's case and the experiment identified by the module. But they may not directly point out whether the reliability would increase or decrease. To address this issue, certain semantic conclusions could be made when calculating the weights for different parameters. For example, if on an average, the increase in a parameter such as pad diameter causes a decrease in reliability then it may be worthwhile to note that it exhibits this trend over a particular range of pad diameters. It is very likely that the same trend may be followed in the areas surrounding the domain where empirical evidence is present, however there is a certain degree of statistical uncertainty when performing extrapolation well beyond the known dataset. Therefore, while extrapolation may be a possibility, one has to exercise caution and limit it to areas within close proximity to experimental cases. Appropriate flagging may be necessary to warn the user that the case may be beyond the predictive capability of the data based approach discussed in this paper. In some cases, a mixed trend may be observed and hence the proximity to available data (or a range of data points) enables the differentiation of random trends from genuine cases such as inflexion points.

Table 3. Basis for rule formation

An example of reliability change with pad finish		
From	To	Average change
HASL	OSP	0.08
OSP	NiPd	0.1
OSP	ImmAu	0.08

(All other parameters are kept the same)

4.5.6 Comparing the test scenario and the experiment match

When the user is conducting sensitivity analyses on different parameters that affect reliability, often the values provided by the user for specific parameters may closely match experimental records stored in the database. Understandably, in some situations, the test case (or the user scenario) may not coincide exactly with the experimental match determined by this algorithm. There may be a few parameters where there is a difference between the two cases and these need to be addressed using the scaling functions. The difference in quantity (quantitative variable) or the transition (qualitative variable) needs to be substituted into the regression models or semantic inferences to determine how much the reliability might change. In case there is no transition or change in value, the scaling operation need not be performed for that parameter.

4.5.7 Prediction and interpretation

After scaling functions have been developed, they could be applied on specific variables that require scaling. The change in the parameter is substituted into the scaling function to determine the change in reliability. Further, the trend exhibited by the parameter would determine the direction of scaling. In the case of a qualitative variable, the knowledge base (look-up table) is checked for the relevant transition. If there is empirical evidence, then the average change in reliability for that transition could be applied to the test scenario. Here, care has to be taken to see if the database evidence points directly to the required transition or its converse. For example, in the case of pad finish, the database may have a transition from organic solder protect (OSP) to nickel-gold (NiAu) when the requirement is quite the opposite. This piece of information could still be used provided the module takes into consideration that the transition is the opposite of what is required. The approach followed in this research uses this kind of information intelligently and scales appropriately as warranted by the situation.

The change in reliability and the direction of scaling are combined to obtain a scaling factor for each variable. The scaling factor is multiplied by the basis reliability or the reliability of the basis experiment (experiment match). The result is the expected scaled reliability and varies from parameter to parameter. However, it is intuitive to expect that some variables will have a more pronounced effect. This is factored in through the weights that were calculated before. The weights are first normalized using a "min-max" scheme that standardizes them and facilitates their comparison across multiple scenarios [10].

One option is to consider the weighted, scaled reliability rather than the individual (for each parameter) scaled reliability. However, this is just a matter of personal preference. The scaled reliability does give an idea of the range in which reliability is expected to vary and what factors may be more important. To this extent, using an approach that scales reliability allows for the interpretation of what the reliability of a package may be and also the bounds of its variability. Of course, determining the confidence level of such a prediction was considered but that is outside the scope of this discussion. By altering one or more of the parameters, the user can run "what-if" scenarios and understand the sensitivity of each parameter affecting reliability. This capability may directly translate to proactive design that improves failure-free product life.

5 Expert system architecture

The prediction of reliability is a sequential process that conveniently fits the expert system framework shown in Fig. 6. By analyzing the different functionalities that are achieved by the above steps, it is possible to realize some of the salient features of expert systems and their relevance to this research effort. Some aspects of expert systems that are called into play in this adaptive expert system are mentioned below.

5.1 User interface (UI)

The user interface is the bridge between the user and the logic embedded in the system. This paper ties together the method used to predict reliability and its good fit with the expert system framework. Hence, this paper does not dwell on the details of the UI although a comprehensive user interface was built.

5.2 Storage

Storage relates to both temporary and permanent storage of information relating to the following:

1. Input/output details;
2. Historical/empirical data;
3. Knowledge base: weights, impact on reliability, semantic inferences, rules;
4. Results of intermediate calculations.

5.3 Knowledge management

This is concerned with the logic that keeps the knowledge base updated upon the addition of a new experiment or changes to previous experiments. As a result, the management of knowledge becomes very adaptive and the expert system could be customized to specific types of components. It performs the following functions:

1. Flushes previously calculated weights/rules from storage;
2. Recalculates/reformulates weights or rules on a dynamic basis.

5.4 Search methods

The functions handled by this section are described below:

1. Concerns the logic/algorithms that is used for searching through the database for specific trends;
2. Combines experiential knowledge with logic to form heuristic search methods.

5.5 Knowledge interpretation and control

The section deals with the inference mechanism of the expert system. The functions are described below:

1. Combines the use of search methods and knowledge management techniques to generate inferences that pertain to solder joint reliability;
2. Makes the predictions and displays them to the user using the interface;
3. Exercises overall control of the program.

6 Case studies

The following scenario describes how thermal cycling performance is predicted for a new component. The example considered is a 0.8-mm pitch CSP on BT substrate with eutectic Sn/Pb solder bumps. This component is to be assembled on a 1-mm (40-mil) thick motherboard using a Sn/Pb paste process on an attachment pad whose diameter is 32 mm

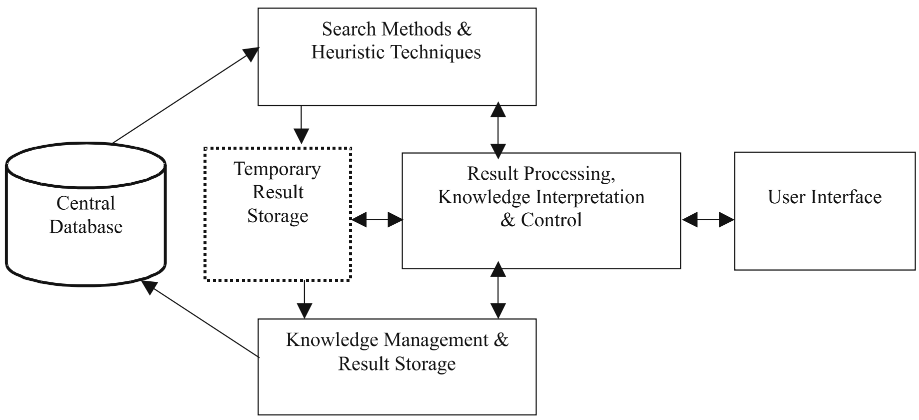


Fig. 6. Expert system architecture and information flow

(125 milliinch) with OSP finish and reflowed in air. The die size is 5 mm^2 (200 mils^2). The user is interested in predicting the performance of this hypothetical component that is subjected to a 0–100 °C, 20-minute thermal cycle with a five-minute ramp and dwell.

The module searches the database of 288 components of various geometries and assembly processes and evaluates the similarity of each component in the database from the perspective of component geometry, assembly process, and testing conditions. The module then suggests a component and an experiment that was performed on a similar component along with a few other choices. The user makes a judgment as to which component and assembly to base the prediction upon.

The module then refers to its knowledge base to see how components have been behaving with respect to various factors that influence reliability. Trends that are extracted are stored and rules are stored. For example, a rule may state that for a given subset of components, pad finish did not have a significant impact on reliability. By using the rules and the quantitative deviation of the test sample from actual empirical results, the module scales up or down each parameter in its list. All parameters may not have an equal effect on reliability, e.g., parameters such as board thickness may affect reliability drastically. Therefore, the module uses its domain of basis experiments, constructs a regression model, and predicts the thermal cycling performance. As the model depends on empirical data for it to work, sufficient datasets or experiments need to exist for the prediction to be accurate.

7 Discussion

The approach used to deduce solder joint reliability is based on data mining methods that use statistics and empirical data to provide insight into the reliability of a new component. Using the method described above, it accomplishes the prediction of reliability and the window in which it varies. This gives an idea as to what the reliability of a package would tend to be, what variables affect it, and how prominent the effect of each parameter may be, based on actual testing results. The method discussed above ranks the parameters considered and provides a normalized weight (between 0 and 1) that is indicative of the individual parameter's importance.

The use of empirical data greatly reduces the time involved to predict the reliability of an assembled package. Package design engineers could obtain quick estimates of reliability and an idea of what parameters are crucial. This aids DFR and sensitivity analyses. The other alternatives are to use FEM and thermal cycling that are either too rigorous to model or not possible to conduct at the design stage. Further, FEM or cycling methods often consume enormous amounts of time, which prevents them from being successfully implemented in a DFR program. This is where the proposed approach proactively assists the design engineer by conveniently estimating and predicting the expected thermal cycling performance.

8 Conclusions

This research yielded the following conclusions:

1. A generic method that can be used to predict the reliability of solder joints has been identified, tested, and validated. The technique that is discussed is based on accelerated reliability evaluation through LLTS testing.
2. Although the reliability prediction mechanism used in this paper is focused on electronics packaging, it can be changed to suit a different domain.
3. This method serves as a personal computer based reliability (thermal cycling performance) prediction mechanism based on empirical evidence, ideally suited for use during the package design stage. It helps to maintain the feedback between reliability testing and the design stage and thus completes a vital link in the program to produce reliable packages. This method potentially reduces testing for packages that are similar to the ones existing in the database and hence can drive costs down.
4. In order to accomplish its goals, this research has drawn upon and combined various concepts from fields such as data mining, statistics, and computing.
5. The methodology described above predicts the reliability and bounds on the variability associated with the prediction. It also calculates the relative importance of various parameters considered in a preliminary matrix.
6. The methodology discussed is this paper can be improved. The exact nature of the algorithms and the details of its implementation have not been discussed. However, the paper outlines some of the general aspects that are pivotal in designing such a predictive tool.
7. The architecture of the entire tool is in the form of an expert system that is adaptive to changes.

References

1. Lau JH (1995) Ball grid array technology. McGraw-Hill, New York
2. Zhang X, Lee R, Pao Y (2000) A damage evolution model for thermal fatigue analysis of solder joints. ASME J Electron Packag 122(3):200–206
3. Clech J-P, Noctor DM, Manock JC, Lynott GW, Bader FE, Augis JA (1993) A comprehensive surface mount reliability (CSMR) model covering several generations of packaging and assembly technology. In: Proceedings 43rd ECTC, pp 62–70
4. Lau JH, Pao Y-H (1996) Solder joint reliability of BGA, CSP, flip-chip, and fine pitch SMT assemblies. McGraw-Hill, New York
5. Luger GF, Stubblefield WA (1998) Artificial intelligence – structures and strategies for complex problem solving. Addison-Wesley, MA
6. Clech J-P (1998) Flip-chip/CSP assembly reliability and solder volume effects proceedings. In: Surface Mount International Conference, pp 315–324
7. Clech J-P (1996) Solder reliability solutions: a PC-based design-for-reliability tool proceedings. In: Surface Mount International Conference, San Jose, CA, vol I, pp 136–151
8. Clech J-P (1996) Solder reliability solutions: from LCCCs to area array assemblies. In: Proceedings NEPCON West, Anaheim, CA, pp 1665–1680
9. Clech J-P, Noctor DM, Manock JC, Lynott GW, Bader FE (1994) Surface mount assembly failure statistics and failure free time. In: Proceedings 44th ECTC, Washington, DC, pp 487–497

10. Han J, Kamber M (2001) Data mining: concepts and techniques. Kaufmann, New York
11. Abernethy RB (2000) The new Weibull handbook, 4th edn. Abernethy, North Palm Beach, FL
12. Lewis EE (1995) Introduction to reliability engineering. Wiley, New York
13. IPC Standard IPC- SM-785 (2002) Guidelines for accelerated reliability test of surface mount solder attachments. www.ipc.org, December 2002
14. Apte C, Weiss SM (1997) Data mining with decision trees and decision rules. Future Gener Comput Syst 13(2):211–229
15. Hosking J, Pednault E, Sudan E (1997) A statistical perspective on data mining. Future Gener Comput Syst 13(2):117–134
16. Elmasri R, Navathe S (2000) Fundamentals of database systems, 3rd edn. Addison-Wesley, Menlo Park, CA
17. Prasad R (1997) Surface mount technology–principles and practice, 2nd edn. Kluwer, New York
18. Primavera A (2000) The influence of PCB parameters on CSP assembly and reliability. Advanced Packaging, September 2000, pp 29–52
19. Montgomery DC, Peck EA (1982) Introduction to linear regression analysis. Wiley, New York