Proceedings of InterPack '01
Pacific Rim International Electronic Packaging Conference
July 2001, Kauai, Hawaii

IPACK2001-15516

Dealing with Uncertainties and Variations in Thermal Design

Brent A. Cullimore
C&R Technologies, Inc.
9 Red Fox Lane
Littleton Colorado, 80127-5710
303 971 0292, 303 742 1540 (FAX), brent@crtech.com

ABSTRACT

The major influence on the reliability of electronics is temperature, yet thermal/fluid modeling is plagued with uncertainties and unknowns. Nonetheless, if appropriate values of these unknown parameters are available for any specific electronics package, then its temperature response can be accurately predicted using modern thermal/fluid analysis tools.

Traditionally, uncertainties are dealt with by a combination of testing, safety factors or margins, and worst-case design scenarios. Analyses are performed iteratively in a repetitive "point design evaluation" mode. Computer-based design simulation tools have emphasized increasing detail and fidelity to physical phenomena, seemingly ignoring the fact that the inputs to these simulations are highly uncertain.

This paper describes both current and future methods of dealing with uncertainties in thermal engineering. It introduces advanced tools and alternative methodologies that can automate not only the quantification of reliability, but can also help synthesize designs on the basis of reliability. It advocates using rapid gains in computer speed not to increase the degree of detail in a model, but to help the engineer find a robust design by automating high-level design tasks.

INTRODUCTION

The sources of uncertainty are numerous for thermal/fluid analyses of electronic equipment. Contact conductances (e.g., between circuit board and chassis, or heat sink mounts) are notoriously difficult to predict ahead of time, and exhibit wide as-built variations. Other examples of uncertain parameters include film coefficients (especially for natural convection), component dissipation levels and/or duty cycles, heat sink by-pass ratios, filter resistances and blockage, environmental conditions, usage scenarios, and unit-to-unit variations in fan performance.

Traditional Treatment of Uncertainty and Variation

Variation can be classified in three categories:

1. uncertainties in performance parameters (contact conductances, film coefficients, dissipation levels, effective thermal capacitances and conductances of complex parts, etc.)
2. environmental or usage uncertainties (ambient temperature and humidity, duty cycle, etc.) as well as degradations over the maintenance life of the product
3. unit-to-unit (manufacturing) variations (bonding, fan performance, filter resistance, etc.)

Each type of variation is traditionally handled using different approaches. Because of differences between organizations, products, etc., the following attempt to describe "typical" approaches is necessarily a generalization.

During preliminary design phases, minima and maxima are established for uncertain performance parameters. These limits are based on experience and perhaps tests of prior similar designs. For the most part, environmental and usage uncertainties are similarly imposed as design requirements: minimum and maximum expected levels of ambient temperature, humidity, etc. Worst case hot and cold scenarios are then stacked up to assure that the electronics do not exceed operating temperature specifications, or that condensation of

humid air is precluded, etc. For example, a worst-case hot environment is often combined with a the lowest estimate of bond conductance and the highest estimate of component dissipation rates.

Subsequent tests might reduce at least the performance uncertainties if not the variations in usage and ambient environment, but more often than not the results of these tests are applied not to the current unit (which has already been designed), but to the design and analysis of future equipment. Values for performance uncertainties are calculated from limited test data, primarily for use in calibrating a thermal/fluid model of the equipment. Unfortunately, because of the system-level interactions of radiation and fluid flow, it rarely makes sense to perform thermal tests at low levels of assembly, and this means that the thermal/fluid model to be calibrated contains several (perhaps 5 to 30) simultaneous unknowns. Also, some unknowns (e.g., film coefficients) will vary over a range of test conditions (e.g., fan speeds). Most often, multiple unknowns are addressed serially: the parameter judged to be the most influential is corrected first, then fixed while the second parameter is subsequently adjusted, ad nauseam.

Unfortunately, even if all uncertainty could be removed in performance parameters such as contact conductances and film coefficients, variations and uncertainties remain in environment, usage, manufacturing, and installation. Often, these uncertainties are treated using design margins, safety factors, or other conservatisms based on experience or tradition. For example, a common upper limit for silicon-based transistors is 125°C. This temperature represents the "knee in the curve" past which failure becomes increasingly likely. To overcome environmental uncertainties as well as unit-to-unit variations, a margin is applied: the operational temperature limits might be set to 115°C, for example, allowing 10°C margin. Ideally, such margins are based on tests and analysis of the specific unit. Realistically, however, it is based on company or customer standards: generalized traditions that may or may not be appropriate for the current package. To overcome uncertainties in performance parameters and in the analysis itself, additional margin (perhaps 5 to 15°C) is typically applied during preliminary (pre-test) design phases. Again, these margins are based on standards or traditions, and not on the particulars of the current application.

The Need for a New Method

Overdesign is expensive. Underdesign is both risky *and* expensive (damaged reputation). Different disciplines (thermal, electrical, structural, etc.) communicate worst-case *derived* requirements to each other rather than attempting to couple the design analyses. For example, dissipation levels in batteries are difficult to quantify, so extremes are used as design cases. This leads to designs that are heavier and more costly than they need to be, *and in some cases does not even result in a safer or more reliable design* if a more risky technology had to be selected in order to meet strenuous derived requirements.

Even within the single discipline of thermal management, overdesign exists due to stack-up of margins and worst-case scenarios until the design case is unrealistic and will likely never occur. Additional margin is then added to cover *uncertainties* in thermal modeling, environment, and component sizing. Only when meeting an extreme stack-up of margins and uncertainties becomes impossible does a renegotiation of adequate margin begin, and such renegotiations are seldom based on any mathematical rigor or true knowledge of the underlying risk.

AUTOMATED MODEL CALIBRATION TECHNIQUES

As was noted above, if there were no uncertainties in input performance parameters, then temperature responses can be accurately predicted using modern thermal/fluid analysis tools. This fact is often exploited to calibrate thermal/fluid models: to reduce or eliminate uncertainties by backing out the values of those uncertainties that generate the best comparison with test data. In many ways, thermal/fluid analyses become elaborate extrapolations of known (tested) points.

As was also noted above, however, the means for calibrating (or "correlating") these models is primitive: repetitive analysis runs made varying one parameter at a time. A better calibration would result by varying all parameters simultaneously. Furthermore, the basis of comparing test data with predictions is rarely based on any mathematical procedure; often, a visual comparison is made between plotted data and predictions.

Automated techniques are available for finding the best-fit estimates of performance uncertainties, although they are as yet rarely employed. These automated techniques not only eliminate a laborious and commonly disliked task, they also get better results than the traditional manual, visual methods. Therefore it is not because of user hesitation that these techniques are not more commonly employed. Rather, it is because these techniques are not well known and because few thermal/fluid modeling tools have these techniques built-in and software packages external to the modeling tools are cumbersome, slow, and expensive. These deployment issues will be addressed at the end of this paper. For now, the basic techniques will be described.

Calibration as an Optimization Process

The task of an automatic calibration algorithm is to find the values of N uncertain parameters ("*calibration variables*") such that the difference between test data and simulation predictions is minimized: find the best-fit values of those parameters. The key to automating this task is to recognize that this is the same algorithm as that followed by a generalized optimization algorithm (also referred to as nonlinear programming, or NLP). To use such an algorithm for calibration, one needs to simply

redefine nomenclature, replacing “design variables” with “calibration variables” and making the “objective to minimize” a measure of the error between predictions and test data. In other words, a wealth of previous research and software solutions is already available that can be readily exploited.

Figure 1 illustrates how an optimization algorithm applies to a calibration task. The algorithm seeks the values of calibration variables (two are depicted in Figure 1 for visualization purposes) such that the objective function (O), being a measure of the difference between test (T) and predictions (P), is minimized. There are many ways to describe such an objective function, with the simplest being a least-squares fit over all comparison points (i):

$$O = \sqrt{\sum_i (T_i - P_i)^2}$$

With this method, there are usually no constraints (such as the one depicted in Figure 1).

In the case of clean (noise-free) data, a better fit (albeit more costly and more difficult to set up[1]) can be found using minimized maximum error, or “MINIMAX” techniques:

$$O = MAX_i |T_i - P_i|$$

Note that a key feature of using an optimization algorithm is that all calibration parameters are changed simultaneously, rather than the traditional serial approach, which corresponds to a series of optimizations each applied to a single parameter. Working in the N-dimensional space of all N calibration parameters at once is not only more efficient, it yields a better correlation.

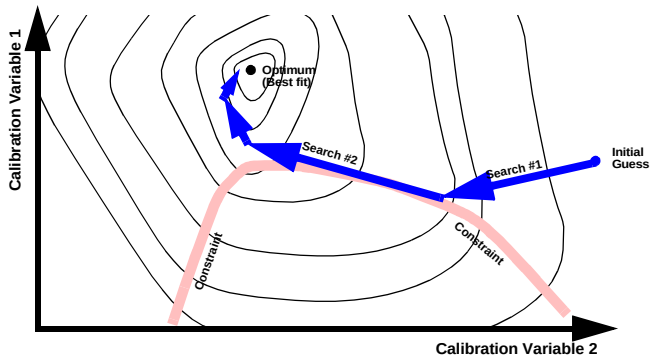


Figure 1. Calibration as an Optimization (Hill Climbing or Valley Descending) Process

1. Such approaches are heavily constrained owing to the non straightforward approach required to avoid introduction of discontinuities in the objective function, as discussed in Cullimore, 2001.

Figure 2 shows how a traditional point-design simulation can be executed iteratively to implement such an optimization-based approach. Typically, between 30 and 300 executions of the simulation are required to perform most calibration tasks. Depending on the implementation (see last section in this paper), the cost of such runs can be substantially less than this number (30 to 300) times the cost of a single stand-alone simulation run. In any case, fewer runs are required than in a traditional “manual, serial, visual” technique, and all of them are automated.

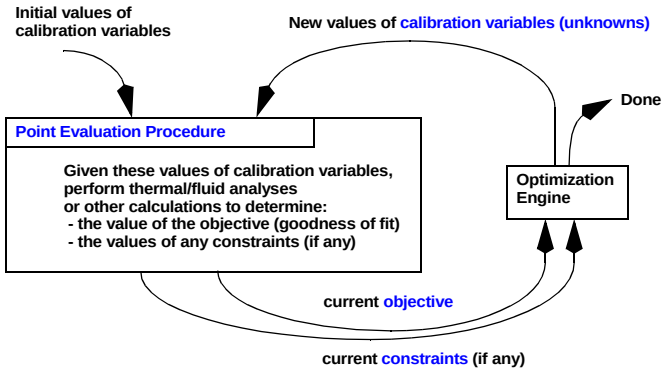


Figure 2. Thermal/fluid Point Simulations Executed as an Iterative Subprocess of Calibration

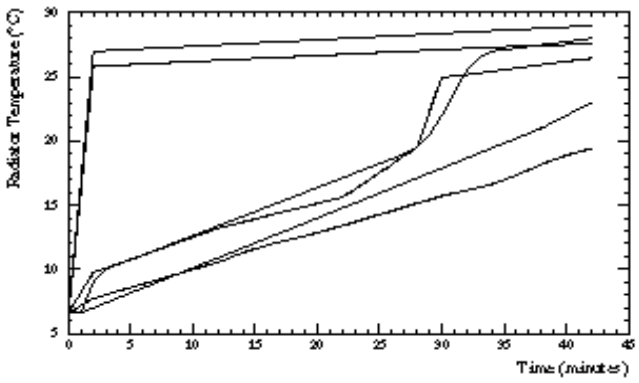
Figure 3 illustrates the application of such automated calibration tools to the case of a transient in a two-phase heat exchanger. While the details of this comparison are beyond the scope of this paper, note that four correlation parameters were chosen, and a maximum error of less than 3°C was achieved with 42 iterations of the transient simulation (42 transient runs, each using a different set of values for the four calibration parameters).

Problems and Responsibilities

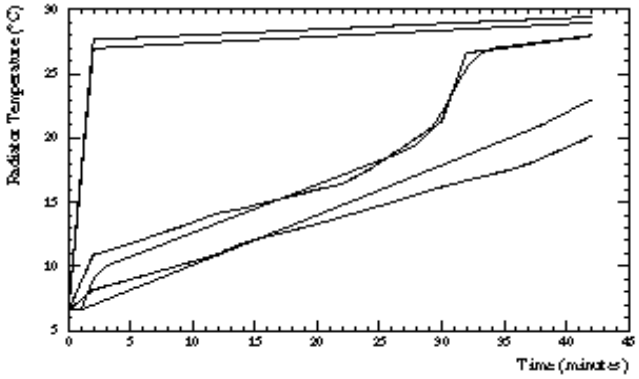
Despite all of the benefits of automated thermal/fluid model calibration techniques, analyst responsibility is not eliminated so much as shifted.

First, the choice of which parameters to declare as uncertain, and within which bounds, is critical. Failure to include a critical parameter or sufficient variation in a parameter can yield a false fit, yet too many parameters with bounds that are too liberal is inefficient.

Second, as was noted above, many different definitions of “best fit” can be mathematically specified. For example, a weighted least-squares is possible assigning more value to good correlation at critical components, or at critical simulation times, etc.



Before (best guess, vendor data)



After (calibrated to test)

Figure 3. Example of Automated Calibration of Four Uncertainties in a Heat Exchanger Transient

Third, unlike linear algebra it is possible to have more unknowns (calibration variables) than equations (data points) and still yield a useful solution. However, it is increasingly possible in those cases (few or inappropriate test data points compared to a large number of uncertainties declared) to yield multiple solutions: multiple local minima in the objective function. Therefore, the user retains the responsibility of assuring that the test data contains enough information to "pin down" the values of all uncertainties applied. For example, no amount of steady-state test data will yield a value for the effective thermal mass of some component: at least one transient temperature trace is required.

With practice, however engineers quickly gain intuition regarding the selection of uncertain parameters, their ranges, and appropriate procedures and subprocedures for extracting their values from test data using automated techniques.

Implementing Automated Calibration Techniques

Recognizing the fact that model calibration is a task that can be posed to a general optimization engine is only half of the solution. The other half is to have these capabilities work conveniently with the thermal/fluid analyzer that is used to generate the performance predictions. Because the implementation of automated calibration techniques parallels the implementation of reliability analysis techniques that will be described next, the discussion of such deployment issues is deferred until the end of the paper.

RELIABILITY OF THE THERMAL DESIGN

Of course, determining unknowns by test can never completely eliminate all sources of uncertainty. Uncertainties remain in manufacturing and installation tolerancing, environmental and usage variation, degradations, etc.

As an alternative to stacking up worst-case scenarios and using rigid margins or safety factors, the engineer could combine these factors statistically to yield information about the degree of confidence ("reliability") in a particular point design. At the very least, the appropriateness of the traditional margins and methods can be assessed. In other words, the engineer could generate not just a single performance predictions but also a distribution of performance predictions with associated probabilities of occurrence, as shown graphically in Figure 4.

Random Variables and Their Distributions

To perform a reliability assessment, the analyst starts by identifying which parameters (dimensions, properties, boundary conditions, etc.) are uncertain. These *random variables* will be allowed to vary over a prescribed range, and any one value of such a random variable has a given probability of occurrence, at least in comparison to other values. This variation is called a *probability distribution*.

The simplest type of distribution is a *uniform* one: the random variable may assume any value with equal probability between a lower limit and an upper limit (depicted at the left of Figure 5). This is an important class of distributions because it represents an easy transition from the current margin-based approach of worst-case high and low values. The margin-based approach to handling uncertainty is excessively conservative, corresponding to two delta (spike) distribution functions at the upper and lower limits, whereas the uniform distribution acknowledges that values in between are at least as likely to occur as the extremes. Nonetheless, the uniform distribution is very simplistic: in most distributions values near the extremes are much less likely to occur than values near the middle.

The most common type of nontrivial distribution is the *normal* or *Gaussian* distribution. It is a symmetric distribution that can be completely described by a mean value and a standard deviation. Many times, an engineer will know the

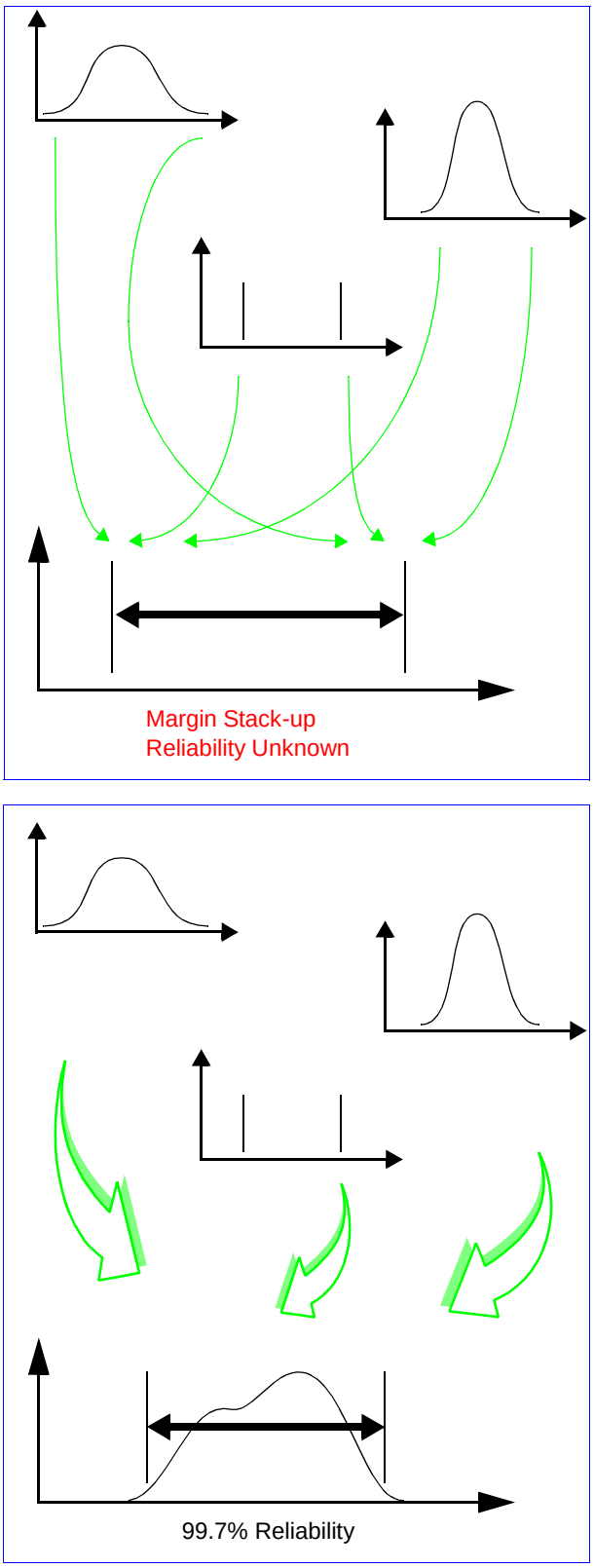


Figure 4. Assessing the Appropriateness of Derived Requirements and Margins

nominal value of a parameter along with an upper and/or lower limit. Frequently these upper and lower limits correspond to a known number of standard deviations (usually two or three) off the mean. Sometimes, a Gaussian distribution is appropriate, but a theoretical range between negative and positive infinity is nonphysical or would cause numerical problems: a truncated normal distribution is required. Another possibility is a triangular (witch's hat) distribution, useful when all that is known is a most likely value plus a lower and upper bound.

In fact, there are many types of distributions possible (e.g., log normal, Weibull, Chi-square, exponential, etc.), each suited for a different purpose. It even possible that a distribution function is produced from test or manufacturing data or from a previous analysis. Therefore, the ability to treat arbitrary probability distribution functions is important.

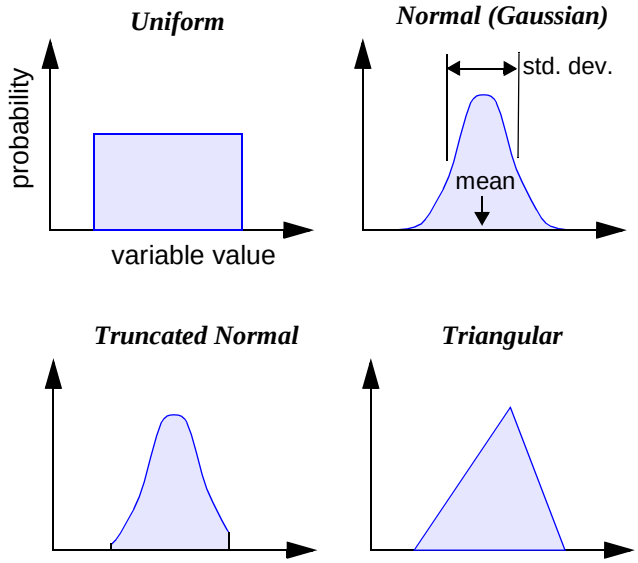


Figure 5. Various Types of Distributions for Random Variables

Reliability Estimation Methods

Given a thermal/fluid model with random variables and defined failure limits, there are several means of estimating the chances that these limits will not be exceeded. For example, the following three choices employ three very different statistical analysis techniques:

1. Monte Carlo Sampling. This method selects the values of uncertain variables randomly according to their probability distribution functions. As an example, for a uniform distribution any value within the valid range is selected using a uniform random number generator. For normal distributions, random values are selected, but

values near the center (the mean) will be generated more frequently than those at the extremes. The Monte Carlo approach requires many samples (on the order of 1000 analyses) and is therefore expensive. However, it yields the most information, and its accuracy can be easily extended with cumulative runs.

2. Descriptive Sampling. Unlike Monte Carlo methods, this approach has a known cost: the user specifies the number of samples to be made (based on what they can afford). This number becomes the resolution with which the distributions in the random variables are subdivided. For example, if the number of samples selected is 100, then each input profile will be divided into 100 regions of equal probability. Once the distributions of the random variables have been subdivided, only one value from each subdivision (the center of the corresponding region in the cumulative distribution function) is sampled, since each of these values is as probable as any of the others. Although five to ten times faster than Monte Carlo sampling for the same accuracy, drawbacks to this approach include lack of accuracy estimation or confidence: have enough samples been taken? Also, accuracy is only marginally improved via cumulative runs.

3. Gradient Method. This method is very fast, requiring only N+1 "samples", where N is the number of random variables. In fact, this technique is not a sampling technique at all. Rather, it estimates reliability by measuring gradients in the responses with respect to the random variables, and by assuming (but not requiring) that all distributions (both input and response) are Gaussian. It further assumes that the mean of the responses can be predicted using the mean values of the random variables, and that response variations from that point are linear with respect to changes in inputs. This method cannot estimate overall reliability, and cannot handle variable failure limits (reliability constraints). The accuracy is not cumulative. However, because it is so inexpensive, it often plays an important role in reliability-based design synthesis.

A comparison of these three techniques is made in Table 1. Experience has shown that Monte Carlo Sampling can only be afforded for the simplest problems, and that Descriptive Sampling is more commonly used. Because of its speed, the Gradient Method is very useful for advanced applications such as reliability-based optimization: the synthesis of a design based on reliability considerations, or the calculation of tolerable variation in a design (selection of tolerances).

Figure 6 illustrates the iterative nature of the sampling or gradient perturbations. The parallels with Figure 2 are intentional and important with respect to implementation. In both cases, what is traditionally performed by thermal/fluid analysts, a "point design simulation" of a device under steady-state and/or transient conditions, becomes merely a subprocess of a larger analysis whether performing optimization (sizing, selection), calibration, or reliability estimation.

Table 1: Comparison of Three Reliability Estimation Techniques

Method	Monte Carlo sampling	Descriptive sampling	Gradient method
Speed	Slow (~1000 analyses)	Intermediate (~100 analyses)	Fast (~10 analyses)
Convergence Detected?	Yes	No	No
Fixed Execution Cost?	No	Yes	Yes
Overall Reliability?	Yes	Yes	No
Cumulative?	Yes	Somewhat	No
Applicability	Unlimited	Unlimited	Limited. Assumes: - Gaussian variables - Continuous, linear responses - Fixed failure limits

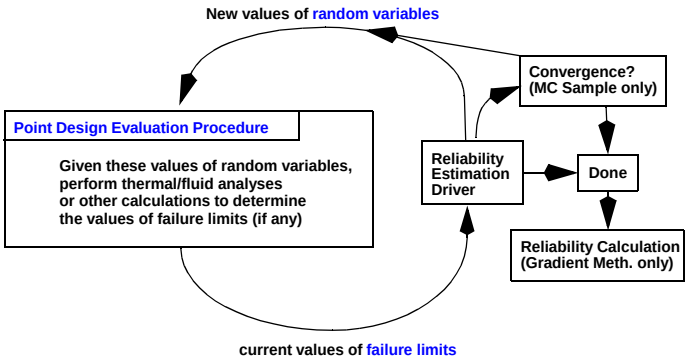


Figure 6. Thermal/fluid Point Design Simulations Executed as an Iterative Subprocess of Reliability Estimation

Figure 7 presents a sample of postprocessing of reliability estimations: histograms, or relative probability of occurrence as a function of a parameter under study (e.g., a failure limit such as peak operating temperature). Because of the number of samples required for non gradient-based techniques, keeping the results is critical such that new statistics can be performed on old data. For example, it is possible to use hindsight in the postprocessing phase and ask new questions, such as revising failure limits (e.g., "What are the chances that the temperature will exceed 50°C instead of 60°C?").

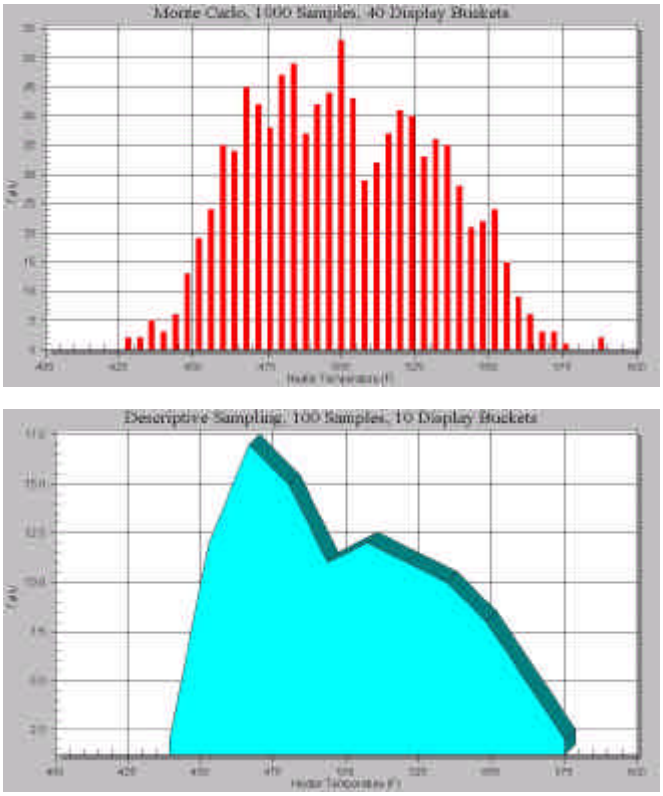


Figure 7. Examples of Histograms used to Postprocess Reliability Studies

IMPLEMENTATION

How does a thermal design engineer exploit the availability of these advanced techniques using their favorite thermal/fluid analyzer? Comparing Figure 2 and Figure 6 shows that both model calibration techniques and reliability estimation techniques are related not only because both deal with uncertainty and variation, but because both involve a higher level of analysis beyond a traditional "point design simulation." Most engineering analysis software is set up to solve a deterministic set of equations, either steady state or transient, given a fixed set of inputs. In other words, these programs provide predictions of how a single point design performs under specific environments. Model calibration and reliability estimation, therefore, can both be achieved by using or creating a software tool that can perform multiple iterative point design evaluations. This section describes three such approaches.

The first option uses an in-house development approach. First, engineers can write their own optimization engine or purchase one commercially. Reliability estimation "engines" are comparatively easy to write, but might still be purchased commercially. Then, a means of executing the thermal/fluid analyzer iteratively must then be achieved, perhaps via an API (application programmer interface) if available, or perhaps simply by modifying and rewriting text input files and reading text output files. A script can then be generated to iteratively run the thermal/fluid analyzer, either (1) driving the uncertain inputs with the optimization engine such that a best match is achieved between simulation predictions and test data or (2) perturbing the random variables and sampling outputs or measuring gradients for reliability estimation. This option is cost effective only if software development labor is inexpensive or if an organization is large enough to recoup the investment of the development of a general-purpose utility. Otherwise, considerable effort will be spent rewriting the software every time a new calibration task arises.

As the second option, engineers can acquire a general purpose MDO (multidisciplinary optimization) environment, most of which also feature statistical analysis tools. Examples of such software include Engineous' iSIGHT®, Phoenix Integration's ModelCenter®, MSC Software's RDCS, Synapse' Pointer®, VR&D's VisualDOC®, and Samtech's BossQuattro. To varying degrees, these programs enable the engineer to set up their favorite thermal/fluid simulation code as part of the evaluation of any one set of unknown or random inputs. The advantages are that these thermal/fluid simulation codes need not "know" that they are being used in such an iterative fashion: little to no modifications of the simulation codes and models are required. This approach also has the advantage of providing an infrastructure that reduces the time to create a new calibration or reliability estimation task. However, disadvantages of the MDO approach include the cost or acquiring and learning such codes, and the relatively slow speeds resulting from inefficiencies in running the simulation code in such a disconnected fashion. Nonetheless, such an approach is clearly better than the current "manual" and "serial" method of calibrating models.

A third choice is to use a thermal/fluid analyzer that already has these advanced features built-in (Cullimore, 1998). This avoids the overhead associated with the first choice, and the additional costs associated with the second choice, and is much faster to execute than either of those choices for various reasons.[2] However, choices are limited for two reasons. First and most important, few thermal/fluid analysts are aware of such capabilities, and hence they demand additional detailed phenomenological modeling power rather than more help with design and calibration tasks. Forgivably, commercial vendors listen to them, and the demand for high-level decision support tools is therefore slack. Second, once analysts discover these gains in productivity and different means of approaching uncertainties, sizing tasks, etc., software suppliers will find it

2. In addition to avoiding interprocess communication and overhead associated with starting and restarting programs, a built-in capability can exploit the advantage that previous steady state solutions (which usually comprise the majority of calibration and reliability assessment tasks) in the search were close to the current solution, and can jump quickly to incremental answers.

difficult to accommodate these requests without significant changes in their software. To accommodate high-level analyses such as model calibration and reliability estimation, the software must first become fully parametric instead of expecting single-valued (“hard-wired”) design and environment specifications. There is hope, however: structural analysis and CAD software have increasingly emphasized such capabilities in their new releases over the last five years. It is hoped that thermal/fluid analysis tools can follow these examples and catch up once the user community has been educated and the demand for new capabilities is established.

CONCLUSIONS

The uncertainties that abound in thermal/fluid modeling of electronics can be treated by a combination of calibration to available test, and to statistical evaluation of the influence of the remaining uncertainties. Automated techniques are available to assist in both tasks. Traditional steady state or transient “point design evaluation” analyses then become a subprocess of a larger system that provides decision support at a higher lever.

REFERENCES

Cullimore et al, SINDA/FLUINT Version 4.4 User’s Manual, PDF available at www.crtech.com, May 2001

Cullimore, B., Optimization, Data Correlation, and Parametric Analysis Features in SINDA/FLUINT Version 4.0, SAE 98-1547, July 1998.