SAND2016-7339C

A Framework and Taxonomy for the Design and Analysis of Margins

Justin Newcomer, PhD, Department of Statistical Sciences, Sandia National Laboratories
Rene Bierbaum, Department of Reliability and Electrical Systems, Sandia National Laboratories

Key Words: Margin Testing, Quantification of Margins and Uncertainties, Product Qualification

SUMMARY & CONCLUSIONS

There are many statistical challenges in the design and analysis of margin testing for product qualification. To further complicate issues, there are multiple types of margins that can be considered and there are often competing experimental designs to evaluate the various types of margin. There are two major variants of margin that must be addressed for engineered components: performance margin and design margin. They can be differentiated by the specific regions of the requirements space that they address. Performance margin are evaluated within the region where all inputs and environments are within requirements, and it expresses the difference between actual performance and the required performance of the system or component. Design margin expresses the difference between the maximum (or minimum) inputs and environments where the component continues to operate as intended (i.e. all performance requirements are still met), and the required inputs and conditions. The model

$$\text{Performance} = f(\text{Inputs, Environments}) + \varepsilon \qquad (1)$$

can be used to help frame the overall set of margin questions. The interdependence of inputs, environments, and outputs should be considered during the course of development in order to identify a complete test program that addresses both performance margin and design margin questions. Statistical methods can be utilized to produce a holistic and efficient program, both for qualitative activities that are designed to reveal margin limiters and for activities where margin quantification is desired. This paper discusses a holistic framework and taxonomy for margin testing and identifies key statistical challenges that may arise in developing such a program.

1 HISTORY AND BACKGROUND

The formulation of this margin framework evolved from the lessons learned the authors have gained applying these concepts at Sandia National Laboratories. The concept of margin has an interesting history at Sandia. Consideration of margin was an integral part of the engineering practices underlying the legacy nuclear weapon stockpile. Performance parameters were defined as part of the product specification, and component performance relative to requirements was monitored during development and production. Conservative design practices were used and environmental over-testing was a key part of development and qualification motivating design changes to address any observed weaknesses. While margin was not always explicitly acknowledged or quantified and a formal framework for evaluating margins didn't exist, testing across (and beyond) requirements helped to identify margin limiters of all types that could be addressed prior to fielding of the product. The result is that very few margin insufficiency problems have been observed in the legacy stockpile components designed by Sandia. The conservative design practices and robust test and analysis processes thirty years ago were effective in removing margin limiters during development and qualification. This isn't automatic for today's components; it must be earned through deliberate efforts during design, development, qualification, and production which has motivated our effort to develop a holistic framework and taxonomy for margin at Sandia.

The original framework for evaluating margins in the nuclear weapon stockpile, introduced in a 2003 joint Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL) report [1] to assess and certify the performance of the stockpiled weapon systems, focused solely on quantification of performance margins which largely remains the case for margin analyses that support the stockpiled systems today. As attention has shifted to the development and qualification of new complex systems, however, more attention is being placed on quantification of margins to environmental and input requirements. These margin analyses are motivated by uncertain and changing requirements that can arrive late in the development process requiring significant work to understand the impact of the changing requirements if margins were not previously quantified. The quantification of environmental and input margins provides valuable insight about the robustness of the components and system to their requirements.

Through the evaluation of performance margins in legacy systems, an increased awareness of characterizing input and environmental margins on new development projects, and continued efforts to improve the statistical methodologies for margin analyses at Sandia applications, we have come to better understand and appreciate the importance and value of such analyses. The process of quantifying margins helps the component and system designers understand the key performance parameters, their requirements and thresholds, how robust they are to the required input levels and environmental conditions, and how they interface with other components in a system framework. Hence, these methodologies are particularly relevant for development programs and can be used to support product qualification, inform product acceptance criteria in production, and provide a baseline for future surveillance sampling strategies.

However, a key obstacle is that the word "margin" has different meanings to different people, and there are several types of "margin" that should explicitly be evaluated during the design process. In this paper, an integrated and comprehensive taxonomy of margin will be illustrated to help guide in collecting a complete set of evidence addressing all variants of margin. By gathering this evidence early in the design and development phase, areas of low (or negative) margin can be identified and addressed prior to fielding and monitored more closely during production and surveillance.

In addition, by viewing margin in an integrated fashion, there is an opportunity to make the testing and analysis more efficient compared to the more discrete concepts of margin used in the past.

2 AN INTEGRATED FRAMEWORK FOR MARGIN

In introducing this framework, it is helpful to describe the requirements space and show how margins are evaluated with respect to this space. There are two general classes of margin:

1. Performance margins
2. Design margins

Performance margins relate to the desired outputs or outcomes when the component functions as intended. Design margins relate to the external conditions that are applied to the component: required inputs for function or environments to which the component may be exposed.

Components must meet all performance requirements across the space defined by the limits of the input and environmental requirements. This is depicted notionally by the blue box in Figure 1 below.

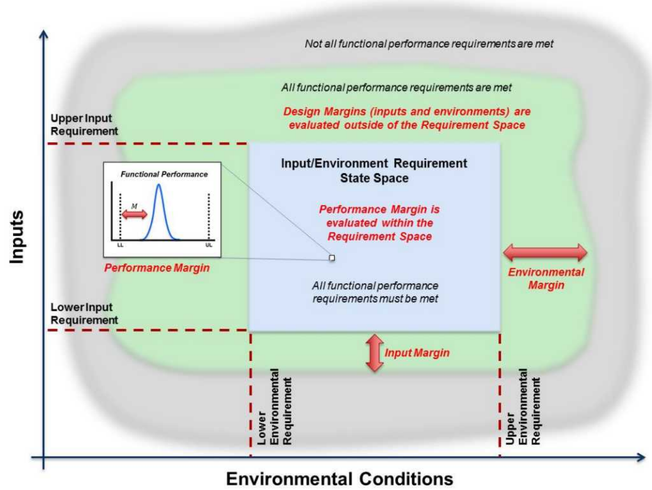

Figure 1. Performance and Design Margins Relative to Requirement State Space.

If either the inputs or environmental conditions exceed their requirements (the green and gray regions outside of the blue box), there is no requirement for proper function. However, it is desirable for many reasons that the component is still capable of proper function beyond the requirements for inputs and environments (inside the red region), at least to some extent. Requirements for inputs and environments may change during development, or even after a component is fielded and there are often uncertainties associated with the specification of input and environmental requirements. Further, the ability to operate beyond the input and environment requirements early in life may provide some benefit if age-related performance trends occur later in life.

The main distinction between performance margins and design margins is the portion of the requirements space in which margins are evaluated. Performance margins are evaluated within this input/environment requirement space (the blue box in Figure 1). On the other hand, to understand how robust the component designs are to the input and environmental requirements (design margins), we evaluate how the component performs beyond the input/environment requirement space. The region where the component still performs as intended, but beyond the input/environment requirements, is depicted by the green region in Figure 1. The design (input and environmental) margins are evaluated by understanding the point beyond the requirements space at which the component no longer performs as required (depicted by the gray space in Figure 1).

In an idealistic sense, the ultimate goal would be to fully characterize the entire space shown in Figure 1 and identify the points where the minimum input, environmental, and performance margins exist. Resource constraints and test limitations however limit one's ability to test freely across this space. Therefore a major statistical challenge is how to optimize testing across this space to provide the most information on all three margins. To further complicate this challenge, there are often multiple performance requirements that must be met within the blue box and there are often multiple inputs and environments that define these regions. In reality this is a complex multi-dimensional space and Figure 1 provides a simplified 2-dimensional slice of the real requirements space that is used for illustration.

2.1 An Integrated Model for Margin

As discussed above, this two-dimensional representation is a greatly simplified view. In practice, there are numerous functional performance requirements, input requirements, and environmental requirements.

The general model for considering different types of margin is given in (1) and reproduced below.

$$\text{Performance} = f(\text{Inputs, Environments}) + \varepsilon$$

By *Performance* we mean the specific functions that a component or system performs. This could be either a particular signal produced by the component (e.g., a firing set produces a high current pulse to fire detonators) or the particular function a component performs (e.g., safety contacts are driven to a closed position), upon receipt of proper inputs and within the given environmental conditions. Often in the latter case, variables are defined that help to characterize the go/no-go action (e.g., "time-to-closure") but success or failure is directly observable in a component-level test. In the former case, success may not be directly observable when functioning the component by itself, one doesn't know for sure if the firing set would have fired the detonators if the detonators are not present in the test so "success" must be defined by specifying requirements (upper, lower, or both) on the performance parameter from which we infer success or failure. Setting requirements is thus an extremely important part of margin evaluation. Margin to a requirement is only meaningful if that requirement is also meaningful, but further discussion of this is beyond the scope of this paper.

By *Inputs*, we mean electrical inputs such as firing set arming signals, use control human-intent inputs, safety unique signals, and so forth that a particular component must receive to perform its function. These inputs may also be specific environmental conditions such as launch acceleration or reentry deceleration that are required to occur in order for a component to perform its function. Thus when we talk about an "input", it can be either an electrical input or an environmental input required for the component to function. Note that many "inputs" for components are actually a performance "output" from another component. These require specialized types of margin analyses that allow system designers to evaluate how components interact and interface within the systems context. We denote these as "interface margins".

By *Environments*, we mean the environmental conditions (temperature, radiation, vibration, shock, chemical/biological, etc.) to which a component is exposed throughout its lifetime. This includes conditions prior to use (transportation, handling, dormant or field storage) as well as conditions at time of function (ejection shock, reentry vibration, etc.). Implicit in this are also the conditions of service life and operational life. Service life refers to the length of time (in days, months, years, etc.) that the component or system must continue to function as intended. Operational life refers to the number of operations (safety switch closures, capacitor discharges, etc.) that the system or component must be able to perform as intended.

Finally, the error (ε) represented in the equation is a compilation of component characteristics such as the unit-to-unit differences driven by production variation and random variation in materials or parts.

There are typically two basic questions that are asked about this relationship between Performance, Inputs, and Environments. Both of them are referred to as "margin", but they take a slightly different form; (1) "Given that the [Inputs, Environments] space is constrained to be within specified requirements for each element, how does the observed Performance compare to the required Performance?" and (2) "What is the entire [Inputs, Environments] space over which the resulting Performance of a component are still considered acceptable (e.g., within the performance requirements)?"

These two questions lead to the characterization of performance and design margins respectively. Each of these two questions is important, but they are fundamentally different in terms of the information they give about the component and also in terms of how they are answered.

Note that each of these two questions reference (a) requirements for inputs and environments and (b) requirements for performance. Given that the inputs and environments can be anywhere within their entire required range, we conceptually define performance margin as the distance between the best estimate of the minimum observed performance and the performance that is needed for successful function. Here the minimum observed performance notes that the performance may, and likely will, change as the inputs and environments vary. Therefore, the goal is to identify the worst case [Input, Environment] condition within their requirements that give the minimum performance margin.

In the second question, a simplistic way to characterize design margin is to consider each input/environment one at a time and observe how much beyond the requirement that input or environment can be while still meeting all performance requirements. Currently this is how design margin is often evaluated, and given the interdependence of performance, inputs, and environments, it highlights the fact that design margin for a component, and more so for a system, is a misnomer. We often only characterize design margin for a single input or condition, not the integrated collection of inputs and environments. Idealistically we want to find the complex space of [Inputs, Environments] where performance requirements are just met and compare it to the space defined by the requirements imposed on [Inputs, Environments]. The notional difference between these two spaces is what truly constitutes design margin for a component. However in practice it is very difficult to characterize. This will be discussed in more detail in a later section.

3 STATISTICAL CHALLENGES IN EVALUATING MARGINS

In this section we discuss several statistical methods and challenges with evaluating margins across and beyond the requirements state space. As discussed before, there are often numerous constraints (test time, test capability, availability of hardware, cost, etc.) that limit one's ability to test the requirements space completely.

3.1 Evaluating Performance Margins and Design Margins

Within the input and environmental requirement space, performance margins are characterized using the statistical framework presented in [2] which is driven by answering the statement, "*Are we **YY**% confident that **at-least XX**% of the unit population will yield a performance output **greater than** the requirement **R**?*" of for design (input and environmental) margins, "*Are we **YY**% confident that at-least **XX**% of the population failures will be in an environment **more severe*** *than the maximum required environment **R**?"*

The main difference between these two questions will be in the evaluation of success or failure relative to the requirement R and the type of data collected to characterize margin. In the quantification of performance margins the requirement R is generally assumed to be a fixed value with no uncertainty (e.g. a performance requirement specified in a component or system requirements document). It represents a threshold between success and failure. The data collected is the functional performance response of the component at specified input and environmental conditions within the input and environmental requirement space. This data characterizes successful component or system performance (success data). On the other hand, in the quantification of design margins we test the component beyond the required input or environmental levels (R) until the component fails to perform as required. The data collected are the input or environmental levels where the component fails (failure data). The two figures below show the statistical framework for both performance and environmental margins (a figure for input margins is omitted because it is similar to the environmental margin figure).

Performance margin is depicted notionally in Figure 2. $\hat{Q}_{XX}$ is the best estimate of performance that $XX\%$ of the units will achieve (are above) and $\hat{Q}_{XX,YY}$ is the $YY\%$ confidence bound on $\hat{Q}_{XX}$ (i.e. the point at which we are $YY\%$ confident that at-least $XX\%$ of the unit population will yield a performance response greater than). Here the requirement is the minimum performance required to ensure successful system function. The performance margin is the difference between the best estimate of performance ($\hat{Q}_{XX}$) and the required performance (R).

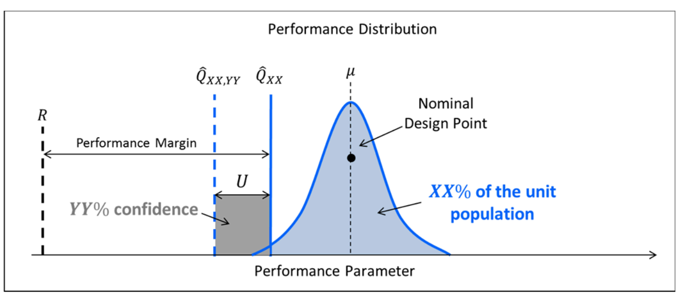

Figure 2. Depiction of Component Performance Margin.

Design margin is shown notionally in Figure 3. Here, $\hat{Q}_{XX}$ is the best estimate of the environmental severity at which $XX\%$ of the units will begin to fail (the environmental threshold) and $\hat{Q}_{XX,YY}$ is the $YY\%$ confidence bound on $\hat{Q}_{XX}$ (i.e. the point at which we are $YY\%$ confident that at-least $XX\%$ of the unit population will fail above). Here the design margin is the difference between the best estimate of the failure threshold ($\hat{Q}_{XX}$) and the maximum environmental severity that the system is required to function under within the component's environmental requirements (R).

This curve is shown as a cumulative probability distribution to highlight the fact that this is a failure distribution as a function of the environmental severity. A similar graphic could be shown for input margins since conceptually the same type of evaluation would be performed (i.e. inputs would be varied until the input level where $XX\%$ of component failures was determined).

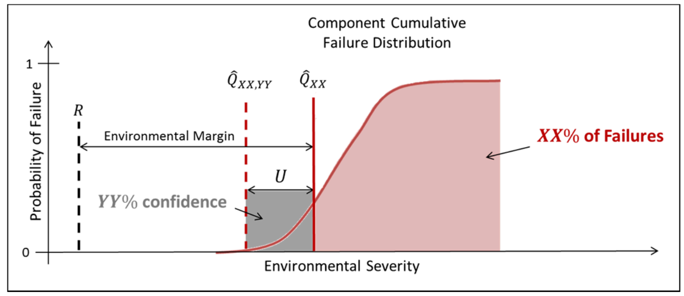

Figure 3. Depiction of Component Design (Environmental) Margin.

There are several statistical challenges that arise with the evaluation of both performance and design margins. These are discussed in the sections below

3.2 Assessing Distributional Assumptions

One of the main statistical challenges in assessing both performance and design margins is the need to make distributional assumptions. In both margin cases, estimates of a meaningful percentile and a statistical confidence bound must be made to guide design and qualification decisions. Early in the development process however there is often limited data and the learning process of how performance depends on inputs and environments is still evolving. Therefore, one must often make distributional assumptions that are not easily verifiable (i.e. due to limited data there is not sufficient evidence to rule out a range of statistical distributions).

The current state of statistical tests for evaluating parametric distributional assumptions is inadequate and often provides uninterpretable and misleading information. Such tests merely provide evidence that a particular distributional choice is insufficient, provided there is sufficient data to support this evidence. The tests do not provide a measure of the adequacy of a given distribution and distributional form uncertainty is not measured. Further, percentile estimates, as is the focus here, are extremely sensitive to the distributional choice.

Therefore, the step in this process for evaluating performance and design margins lacks adequate tools for evaluating the goodness of the estimates $\hat{Q}_{XX}$ and $\hat{Q}_{XX,YY}$. This induces a large amount of unquantifiable uncertainty into the design and qualification process. Development of actionable guidance and implementation tools is a current focus area at Sandia and a necessary first step toward adoption of methods for better percentile estimation and hence characterization of performance and design margins.

3.3 Interactions of Inputs and Environments

The general model for performance shown in (1) may tempt one to consider testing one state at a time to build a full understanding of output. However it is more complicated than this in that there are typically interactions between Inputs and Environments. A simple example is the environment of temperature, which typically will affect the inputs applied (e.g., at hot temperature, the input voltage to a safety switch provided by a weapon controller will be higher) as well as the performance of the component (e.g., the resistance of the resistors in the safety switch increases with temperature). The specific case of performance margins is illustrated in Figure 4.

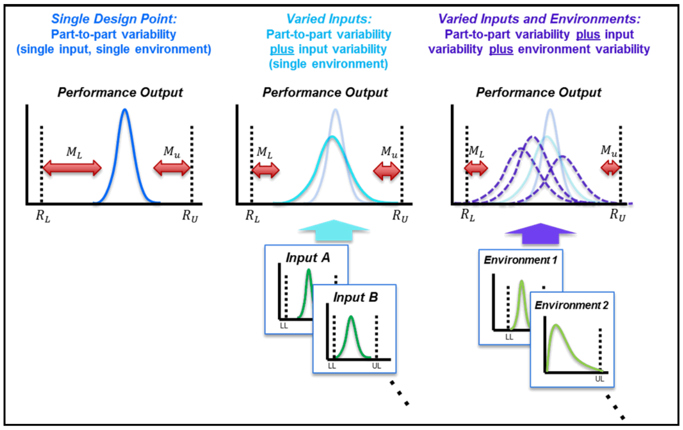

Figure 4: Performance Margin as a Function of Inputs and Environments

From a point design, one can sometimes directly determine the performance margin. However the challenge for performance margin is that most outputs are dependent on both inputs and environments, as well as dependent upon the inherent part-to-part variability for a component. And of course most components have multiple performance parameters that are considered important in inferring successful function. Each may behave differently with respect to inputs and environments.

Performance margin is typically characterized by testing with inputs at the maximum and minimum values and with environments set to the minimum or maximum required environment, at least to the extent that one can, to see the effect on the performance parameter(s) of interest. The distribution of outputs is then compared to the requirement to determine margin. Margin testing of this type can be complemented by computational tools such as electrical modeling, which allow for modeling of output variability given the variability and any known, modelable environmental dependence (such as temperature or radiation exposure) of the constituent parts. Expert judgment in recognizing relationships between inputs and environmental conditions (e.g., cold temperature tends to result in low/slow inputs) can make evaluation (both testing and simulation) more efficient.

Design of Experiments is another tool that can make state space exploration more efficient in identifying minimum performance margin conditions. However, it is important to remember that worst-case margin may not necessarily be at the extremes of inputs and/or conditions; it is prudent to try to cover the state space with sufficient testing to identify the true performance margin.

3.4 Adequacy of Requirements

Evaluation of performance margin is very dependent upon good requirements definition (both for specified inputs/conditions and for the required outputs), and mistakes or excessive conservatism in requirements can result in either too optimistic or too pessimistic a view of a component's performance margin.

One other very important point is differentiating between the "assessed performance margin" (the relationship of an output to its specified performance requirement) and "true performance margin" (the relationship of an actual output to the point where the component or system actually fails to function). Most performance margin questions in reality require consideration of the relationship between two separate distributions (a component and the interface with its succeeding component), not the relationship of a single distribution to a deterministic limit.

The former of course is what we would like to characterize; however that can rarely be done because the "true performance margin" itself depends upon the distribution of performance of any adjoining components (which almost certainly varies by condition). Although it provides a conservative answer, it is almost always much easier to use a fixed requirement.

3.5 Challenges for Design Margins

Determination of design margins has challenges similar to those described above for performance margins. One of the main additional challenges with the evaluation of design margin is that there are two types of failures that must be considered in the evaluation of the cumulative failure distribution shown in Figure 3; catastrophic failures (e.g. the component shattered when exposed to a severe shock and produced no output) and margin failures (e.g. the component functioned but produced a degraded performance output due to a thermal exposure). This complexity is shown notionally in Figure 5 where the boundary between the green and gray regions must consider both types of failures.

There are two common approaches for dealing with and combining these two types of failures. First, one could treat both as binary go/no-go events by reducing the performance data to binary go (within performance requirements) and no-go (outside performance requirements) data. This however does not utilize the full information contained in the data and may lead one to overlook potential issues that could manifest if small changes in the performance output occur. The second approach is to model the probability of failure due to the environmental or input severity (P_f) by evaluating the probability of not having a catastrophic failure and the probability of meeting performance requirements given a catastrophic failure did not occur,

$$P_f = 1 - P(X = 1) \cdot P(R_L < Y < R_U), \qquad (2)$$

where $X = 1$ if there is no catastrophic failure and $X = 0$ otherwise, and Y is the random variable for the given performance parameter of interest. This allows one to more completely evaluate the design margins, however there may be limited (or for high reliability components) often no observable catastrophic failures. This limited data poses increased uncertainty in the estimates of P_f and requires one to make additional assumptions on the distribution of catastrophic failures.

Again, ideally when evaluating design margins, we would like to find the minimum design across the entire n-space of inputs, environments. In Figure 5 this is depicted showing the minimum input and environmental margins for this simplified 2-dimensional case.

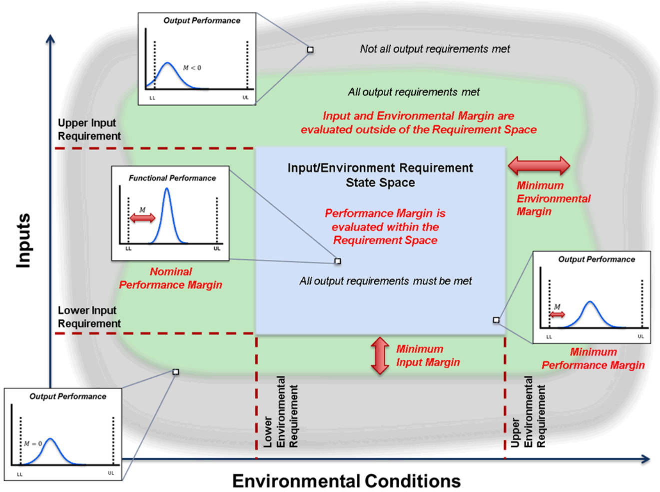

Figure 5. Evaluation of Performance Margins and Design Margins.

Note that this general taxonomy of performance margins and design margins can be used for all types of performance questions, be they reliability, safety, use control, or yield.

The point here is not to cause despair but rather to highlight the critical importance of thinking about margin holistically, considering both performance margin and design margin during development and recognizing that the problem is multi-dimensional. Thinking about margin questions in an integrated way opens up opportunities to increase efficiency in testing and simulation. Careful selection of key performance parameters is essential as well and should be done early in development.

4 TESTING CONSIDERATIONS FOR MARGIN CHARACTERIZATION

Performance margin is typically characterized by testing or simulation across the range of inputs and environments that are within requirements. Design margin is typically evaluated by over-testing or testing-to-failure of various kinds. Examples for environments include applying a mechanical environment such as shock until a component fails, or lowering the temperature until a component's output drops below the level needed for function. Accelerated testing is a commonly used tool and it should be noted that design margin activities early in design is typically more qualitative, with the intent to identify margin limiters so they can be removed, rather than provide some quantitative margin estimate. A good example of this is Highly Accelerated Life Testing (HALT), where environments beyond requirements are applied to induce failure. Computational modeling can also play a role in characterizing design margin.

There is no "one size fits all" approach to determining test quantities for characterization of either performance margins or design margins. Considerations relating to each type of margin should result in definition of a diverse series of tests, each with their own specific objective, plan, and test approach. It is highly recommended that statisticians be engaged to help formulate an appropriate, limited-scope objective and then develop a plan to help answer it, with an understanding of the uncertainty in the answer. For all margin evaluations, either performance or design, it is essential to keep in mind the integrated view, where margin for a particular condition or input is likely dependent upon the specific state of other conditions and/or inputs

REFERENCES

1. B.T. Goodwin, R.J. Juzaitis, National Certification Methodology for the Nuclear Weapons Stockpile, UCRL-TR-223486. Lawrence Livermore National Laboratory, Livermore, CA, 2006.
2. J.T. Newcomer, A New Approach to QMU for Physical Simulation Data, SAND2012-7912, Sandia National Laboratories, Albuquerque, NM, September 2012

BIOGRAPHIES

Name1
Address 1
Address 2
Address 3
City, State or Province, Postal Code, Country

e-mail:

Text.

Name2
Address 1
Address 2
Address 3
City, State or Province, Postal Code, Country

e-mail:

Text.