# The Nature and Role of *True Value* in Monte Carlo Uncertainty Analyses with Application to BEAU Methodology

Paul Sermer and Keith Weaver

Reactor and Radiation Physics Department Nuclear Safety Solutions Ltd. 700 University Ave., 11<sup>th</sup> Floor, Toronto, Ontario, M5G 1X6 paul.sermer@nuclearsafetysolutions.com keith.weaver@nuclearsafetysolutions.com

## Summary

The current best estimate and uncertainty analysis (BEAU) methodology for LOCA safety analysis requires as input probability distributions for the relevant process variables as well as their associated uncertainties. In particular, among these process variables are maximum channel and bundle powers including their uncertainties. Upon deeper consideration it becomes apparent that it is not possible to satisfy the requirement in a straightforward manner. The reason for this is that the mathematical notion of the variables in the compliance with channel/bundle power licence limit analysis (the CU approach) are conceptually different. A crucial difference between the two approaches is that the CU analysis considers the error in the simulated variable, estimated by a Monte Carlo procedure, while BEAU analysis does not. The Monte Carlo procedure of the current BEAU analysis lacks a formal description of its fundamentals and treatment of the uncertainties, while the CU methodology provides a rigorous mathematical framework for the simulation of variables, its meaning and the associated uncertainty analyses. Using this mathematical framework, it is shown how to understand the BEAU simulation, what the input process variables should be and how the input errors should be treated.

The result of a BEAU analysis is an extreme value, such as a *maximum* fuel sheath temperature, and hence the crucial statistical properties of such a result depend on the understanding of the extreme value statistics. The current BEAU analysis may suffer from the lack of consideration for the extreme value statistics. In this paper some examples are provided to demonstrate the effect of the extreme value statistics on the final result of the desired computation. It is also argued that the CU approach applied in BEAU analysis leads to estimates for the simulated variables with well defined statistical properties. Moreover, such an approach may lead to more favourable results than the current BEAU analysis in a sense that the former may produce results with larger operating margins.

### I. INTRODUCTION

Many results in nuclear safety analysis rely on the ability to calculate an extreme value of something, say T, such as a maximum fuel temperature in LOCA BEAU analysis [1]. It is relatively straightforward to find a conservative extreme value for T. But to operate a reactor under practical licensing constraints one needs to find extreme values and at the same time be adequately confident that they are only as extreme as is useful for the application. The need for the answer to lie within some band of conservatism, and the associated level of confidence vary from application to application. Consequently, there is a need for a general method to calculate extreme values, but which can be tailored to meet the needs of different applications. In addition to LOCA BEAU analysis, such applications include compliance with maximum channel power licence limits, implementation of low flow channel power limits which are maximum channel power dependent, compliance with minimum margin to fuel constraint, computation of NOP trip set points, etc.

Such a general approach has been developed and is called the compliance uncertainty (CU) approach (based on the work on compliance with channel power licence limits [2]). It is described in some detail below.

The main advantages of the CU approach are:

- The result is a probability distribution for the error in the estimated extreme value T (considered to be a random variable), with well-defined statistical properties. This allows one to calculate directly the upper bound on T for some associated probability 1  $\alpha$ . In contrast, the use of a probability distribution for T in which the error term has not been computed explicitly is basically an empirical approach and has no such statistical assurance.
- It can deliver estimates of the extreme value which are restrictive only to the degree required or desired because the error term is handled explicitly.
- The nature of the estimated extreme values of T can be related to the characteristics of the function  $\mathcal{F}$  by which the values of T are generated. The way in which these two functions (i.e.  $\mathcal{F}$ and T) are related is complex and it will be seen that as  $\mathcal{F}$  changes in nature the changes in the estimates of T can be considerable. More significantly, however, these changes in the estimated T can be qualitatively different from those seen in a more traditional approach (where the errors are included only implicitly). This can lead to situations where the predictions of a more traditional approach can be shown by the CU method to be non-conservative.

In the present paper, we will

• outline the mathematical formalism underlying the CU approach, with particular attention to the concept of "true value" and the application of this concept,

- describe how a Monte Carlo calculation using this approach is set up,
- show explicitly how the CU and traditional (BEAU) approaches differ,
- provide some analysis of the extreme values for a given computation  $\mathcal F$

#### II. THE CONCEPT OF A TRUE VALUE IN THE CU METHODOLOGY

The "true value" of some parameter, say  $P_i^{true}$ , is a somewhat metaphysical concept, in that a true value is generally unknowable and unobservable. We can speak of a measured value, say  $P_i$ , and an associated error  $\varepsilon_i$ , and by virtue of this one can infer a true value, provided the definitions are kept adequately rigorous. The mathematical expression of this statement is

$$P_i = P_i^{true} + \varepsilon_i, \tag{1}$$

and this expression can be viewed as a definition of error. Equation (1) is sometimes written as

$$P_i^{true} = P_i + \varepsilon_i,$$

but this is incorrect as a conceptual definition since the true value exists regardless of the error. (This equation is used only numerically to derive confidence limits on the true value.) Furthermore, since the error and the true value must be considered as variables over which we have no control, the observed value,  $P_i$ , is dependent on them and thus the mathematical formalism of Equation (1) is correct.

From Equation (1), it is evident that getting to grips with the true value requires an understanding of both the observed value and the error. Consider a general model in which a parameter  $T^{true}$  is a function,  $\mathcal{F}$ , of *n* variables  $P_1^{true}$ ,  $P_2^{true}$ ,  $P_3^{true}$ , ...,  $P_n^{true}$ , i.e.,

$$\boldsymbol{T}^{true} = \mathcal{F}(P_1^{true}, P_2^{true}, P_3^{true}, \dots, P_n^{true}).$$
(2)

Given measured (or computed) values  $P_1, P_2, P_3, ..., P_n$ , an approximate values T of  $T^{true}$  is defined as

$$T = \mathcal{F}(P_1, P_2, P_3, ..., P_n),$$
(3)

where T now includes an error, say  $\tau$ , which can expressed as

$$T = T^{true} + \tau. \tag{4}$$

In this expression,  $\tau$  can be considered as a random variable and will possess some probability distribution.

Following the standard approach for statistical testing, for a small positive number,  $\alpha$  (such as 0.02), we can say that the probability that  $\tau$  has a value greater than or equal to a value  $-\tau_{\alpha}$  is 1 -  $\alpha$ , or

$$P\{\tau \ge -\tau_{\alpha}\} = 1 - \alpha. \tag{5}$$

(The *minus* sign in front of  $\tau_{\alpha}$  is used for convenience only since  $\tau_{\alpha}$  is, typically, a positive number.) Using this value of  $\tau_{\alpha}$  an upper confidence bound,  $T^{\alpha}$ , on  $T^{true}$  is defined as

$$\boldsymbol{T}^{\boldsymbol{\alpha}} = \boldsymbol{T} + \boldsymbol{\tau}_{\boldsymbol{\alpha}}. \tag{6}$$

The important property of  $T^{\alpha}$  is that

$$P\{\boldsymbol{T}^{\alpha} \geq \boldsymbol{T}^{true}\} = 1 - \alpha.$$
<sup>(7)</sup>

What Equation (7) says is that if we prescribe some limit, L, then based on a choice of  $\alpha$ and our derived knowledge of the error associated with T, we can state with a confidence of  $1 - \alpha$ that when  $T^{\alpha}$  is less than or equal to L,  $T^{true}$  is also less than or equal to L. In the language of statistical testing,  $T^{\alpha}$  is the test statistic, the Null Hypothesis is " $T^{true}$  is greater than L", and the Alternative Hypothesis is " $T^{true}$  is less than or equal to L". So if the value of the test statistic  $T^{\alpha}$ is less than or equal to L, we reject the Null Hypothesis and conclude that the Alternative Hypothesis is true to a confidence level of  $1 - \alpha$  (or significance level of  $\alpha$ ).

Note the difference between  $T^{\alpha}$  and T. In more traditional approaches, T incorporates the error in some non-specific way (for example applying errors to  $P_i$ s in (3), in the sense that the focus is on the observed or computed value T. T is then presented as an empirical distribution which has no defined properties, and in order to make statements about any conservatism which T may include one is forced simply to specify arbitrarily some (upper) percentile of this distribution as a means of estimating where the desired extreme value may lie. In the CU approach, on the other hand, the error is treated explicitly and a variable  $T^{\alpha}$  is constructed which allows inferences of statistical significance to be made about the parameter  $T^{true}$ .

Note particularly that the values of T were obtained using some calculation model  $\mathcal{F}$  - see (3). Because of this, the value of T will depend on the functional nature of  $\mathcal{F}$ . For a particular function  $\mathcal{F}_1$ , the statistical test applied to T would produce some outcome A. For a different func-

tion  $\mathcal{F}_2$ , some different outcome *B* could arise. It is possible to make these distinctions only because the CU approach treats the error explicitly and allows a statistical confidence of any desired level of stringency to be applied. More traditional approaches which consider the error implicitly allow no such distinctions to be made. Consequently, the CU approach makes it possible to choose an extreme value which has a proximity to the given limit appropriate to the constraints of the situation being considered. In other words, it is possible to select an extreme value in relation to a limit such that the degree of conservatism one wishes to incorporate in the selected extreme value can be specified (or minimized).

#### III. MONTE CARLO CALCULATIONAL METHOD

Because of equation (7) and the above discussion, it becomes clear that the determination of  $T^{\alpha}$  should be the goal of a required computational procedure. This appears consistent, in principle, with more traditional analyses, such as the current BEAU analysis, in which an upper 95% confidence limit on T is obtained to serve as an upper bound on a true value. However, this confidence limit is not  $T^{\alpha}$  (for  $\alpha = 0.05$ ) as defined in (6) and (7). In fact, we will demonstrate that this 95% confidence limit on T may be much larger than  $T^{\alpha}$ . Using the above mathematical formalism, it is clear that in order to compute  $T^{\alpha}$  we need to determine the probability distribution for the error  $\tau$  (see (4)) and hence the *compliance uncertainty*  $\tau_{\alpha}$  given by (5) - this is the essence of the socalled CU approach as mentioned earlier.

The probability distribution for  $\tau$  can be estimated by a Monte Carlo approach (cf. [2] and [3]) using actual reactor operating data. (For the sake of simplicity we will refer to T in (2) and (3) as the *fuel temperature*.) The rationale behind the approach is the following. Let us assume that for some reactor state the *true* process variables are  $P^{true} = [P_1^{true}, P_2^{true}, P_3^{true}, ..., P_n^{true}]$ , which are not known to us. If we apply the available computer codes and measurements, for the same state, we will obtain the estimated process variables  $P = [P_1, P_2, P_3, ..., P_n]$ . We conceive of the *true* process variables,  $P^{true}$ , as being statistically realized, along with random computational or measurement errors, as the computed process variables P, which are known to us. Because computational and measurement errors have random components, the set P is not a unique realization of the *true* process variables  $P^{true}$ . Rather, there are infinitely many other realizations of the *true* process variables each of which could have been the one computed, but happened not to be. Let the different realizations, or hypothetical process variables, be denoted by  $P^{(j)}$ , j = 1, 2, 3, .... We are interested in the probability distribution of the error  $T^{(j)} - T^{true} = \mathcal{F}(P^{(j)}) - \mathcal{F}(P^{true})$ , cf. (2) (3) and (4). The issue is to estimate this error without knowing  $P^{true}$ . Essentially, there is only one way to solve this problem (see [3]):

Although the process variable estimates P are not the *true* ones, let us consider a fictitious reactor in which they were the *true* ones. We assume that this fictitious reactor is not too different from the actual one with the process variables  $P^{true}$ . Furthermore, let us assume that the shape of the probability distribution

$$\boldsymbol{T}^{(j)} - \boldsymbol{T} = \mathcal{F}(\boldsymbol{P}^{(j)}) - \mathcal{F}(\boldsymbol{P}),$$

in the fictitious reactor is the same, or very nearly the same, as the shape of the probability distribution

$$T - T^{true} = \mathcal{F}(P) - \mathcal{F}(P^{true}),$$

in the real reactor. Note that we are not assuming that P and  $P^{true}$  are equal; they are certainly not. We are only assuming that the way in which random errors enter the computation does not vary rapidly as a function of  $P^{true}$ , so that P can serve as a reasonable surrogate.

Now, in the fictitious reactor, the distribution of  $T^{(j)} - T$  is within our ability to calculate. The formulation of the Monte Carlo procedure is as follows. For a (real) reactor state, let P be the set of computed or measured process variables and  $T = \mathcal{F}(P)$ . We randomly select

$$\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n],$$

(cf. Equation (1)), from the known error distributions and compute

$$\mathbf{T}^{(1)} = \mathcal{F}(\mathbf{P} + \mathbf{\varepsilon}),$$
$$\mathbf{\tau}^{(1)} = \mathbf{T}^{(1)} - \mathbf{T}.$$

 $\tau^{(1)}$ , of course, is only one possible random realization of  $\tau$ . For a different set of random values of  $\epsilon$ , we compute

$$T^{(2)} = \mathcal{F}(\boldsymbol{P} + \boldsymbol{\varepsilon}),$$
$$\tau^{(2)} = T^{(2)} - T.$$

 $\tau^{(2)}$  is another possible random realization of the error  $\tau$ . This process is repeated for j = 1, 2, 3, 4 ... N, where N is a large number, such as 10,000 or 100,000. The values  $\tau^{(1)}$ ,  $\tau^{(2)}$ ,  $\tau^{(3)}$ , ..., determine the probability distribution for  $\tau$ . Finally, the compliance uncertainty  $\tau_{\alpha}$  is computed from (5) and the (well defined) upper bound  $T^{\alpha}$  on  $T^{true}$  is computed from (6).

#### IV. UNDERSTANDING OF THE TRADITIONAL BEAU COMPUTATION

Using the above terminology, the traditional BEAU Monte Carlo methodology can be described as follows. At every simulation stage j = 1, 2, 3, ..., randomly obtain

$$\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \varepsilon_3, ..., \varepsilon_n],$$

from the known error distributions, and also randomly obtain  $P = [P_1, P_2, P_3, ..., P_n]$  from the given probability distributions. (Note that in the previous Section P is not chosen randomly but, rather is given for the specific reactor state.) A new random realization of the fuel temperature is

$$T^{(j)} = \mathcal{F}(P + \varepsilon), \qquad j = 1, 2, 3, ...$$

The set of values  $T^{(1)}$ ,  $T^{(2)}$ ,  $T^{(3)}$ , ..., (in fact, the same as in the previous Section) determine the probability distribution  $T^{sim}$ . The upper 95<sup>th</sup> percentile of this distribution, denoted by  $T_{95}$ , is deemed to be an "upper bound" on the *true* value  $T^{true}$ .

The above statement on the nature of  $T_{95}$  does not provide a quantitative relationship to the *true* value  $T^{true}$ , similar to (7), since there is no error reference in the above computation.  $T_{95}$ is simply greater than 95% of the approximate values and hence "somewhat likely" to be greater than  $T^{true}$ . A quantification of the "somewhat likely" requires knowledge of the error. In the context of our understanding of the Monte Carlo procedure described in the previous Section,  $T^{sim}$  is in fact an approximation of the fuel temperature T in the fictitious reactor, which itself is an approximation of the fuel temperature  $T^{true}$  in the actual reactor. Thus, as an approximation of the fuel temperature in the actual reactor,  $T^{sim}$  is a "worse" approximation to  $T^{true}$  than T is. Consequently, it is reasonable to expect  $T_{95}$  to be significantly different (probably larger) than the rigorously defined upper bound  $T^{\alpha}$  in (6).

In order to appreciate that the two described approaches are indeed different, we applied the CU and the traditional methodologies to compute estimates of maximum channel power. Using the described notation, the process variables  $\boldsymbol{P} = [P_1, P_2, P_3, ..., P_n]$  would be the channel powers and the computation  $\mathcal{F}$  would be the *maximum* function. The result,  $\mathcal{F}(\boldsymbol{P})$ , i.e.,  $max(\boldsymbol{P})$ , is denoted by  $\boldsymbol{S}_{max}$  for the CU approach and  $\boldsymbol{Q}_{max}^{sim}$  for the traditional BEAU approach (these play the role of  $\boldsymbol{T}$  above). In the CU approach,  $\boldsymbol{S}_{max}^{\alpha}$  is computed with  $\alpha = 2\%$ . The probability density functions for  $\boldsymbol{S}_{max}^{\alpha}$  and  $\boldsymbol{Q}_{max}^{sim}$  are presented in the figure below. These were computed using actual reactor operational data over a three year period - almost 2000 sates for all Darlington reactor Units. The 95<sup>th</sup> percentiles  $\boldsymbol{S}_{max}^{\alpha,95}$  and  $\boldsymbol{Q}_{max}^{sim,95}$ , respectively, are indicated in the figure as well. We observe that  $\boldsymbol{S}_{max}^{\alpha,95}$  is much smaller than  $\boldsymbol{Q}_{max}^{sim,95}$  demonstrating that the traditional BEAU approach is unnecessarily pessimistic in this case. That is,  $\boldsymbol{Q}_{max}^{sim}$  would suggest that there is a significant period of time when the true maximum channel power exceeds the licence limit (7200 kW), while this is not the case as the  $S_{max}^{\alpha}$  would indicate.



# V. PROPERTIES OF THE BEAU MONTE CARLO METHODOLOGIES WITH RESPECT TO EXTREME VALUE STATISTICS

The result of a BEAU analysis is an extreme value, such as a *maximum* fuel sheath temperature, and hence the crucial statistical properties of such results depend on the understanding of the extreme value statistics. We observe that the traditional BEAU Monte Carlo method (described in the previous Section) does not possess a formal mathematical framework and no useful statistical properties of  $T^{sim}$ , or  $T_{95}$  are given (some analysis to that effect is presented below). The socalled CU approach described in Section II. does possess a formal mathematical framework and important statistical properties can be demonstrated - such as property (7). (Recall that the fundamental distinction between the two approaches is that the goal of the Monte Carlo procedure in the CU method is to compute the error of the final result, while the traditional BEAU method has no reference to such an error.) The seeming disadvantage<sup>1</sup> of the CU approach is that it is computationally more expensive in that all (or possibly only a subset) of reactor channels are involved, while in the current implementation of the BEAU method only one reactor channel is considered. It is to be understood, however, that the latter approach may be deficient due to the lack of consideration for the extreme value statistics. In this Section we will provide some examples to demonstrate the effect of the extreme value statistics on the final result of the desired computation.

A "more" rigorous approach to the current BEAU method is to treat the input process variables in a similar manner as with the maximum channel power, namely, use  $S_{max}^{\alpha}$  without additional error sampling. Thus, assume that

$$\boldsymbol{P}^{\alpha} = [P_1^{\alpha}, P_2^{\alpha}, P_3^{\alpha}, \dots, P_n^{\alpha}],$$

are the input process variables that are guaranteed to exceed their corresponding *true* values with a probability  $1-\alpha$ , (cf. (7)), i.e.,

$$P\{P_i^{\alpha} \ge P_i^{true}\} = 1 - \alpha, \tag{8}$$

for all i = 1, 2, 3, ..., n, and some small  $\alpha$ , such as 0.02. (Note that for some variables the *less than* sign may be appropriate. For the sake of simplicity of the presentation and without loss of generality, we will not distinguish such cases.) Now, for the given computation  $\mathcal{F}$  (see (3)), we define

$$\boldsymbol{U}^{\alpha} = \mathcal{F}(\boldsymbol{P}^{\alpha}) \tag{9}$$

and the distribution for  $U^{\alpha}$  (an estimate of  $T^{true}$  in (2)) may be obtained using a Monte Carlo simulation by sampling from the known distributions for  $P^{\alpha}$ . Once such a distribution for  $U^{\alpha}$  is computed, the final result is the 95<sup>th</sup> percentile, or  $U_{95}^{\alpha}$ . We will refer to the just described approach as the *updated* BEAU Monte Carlo method.

The statistical properties of this *updated* BEAU methodology depend on the form of the function  $\mathcal{F}$ . For a general form of this function, an analytical determination of the statistical properties is too intractable to pursue.<sup>2</sup> However, if  $\mathcal{F}$  were monotone increasing then equation (8) would imply

$$P\{\mathcal{F}(P_i^{\alpha}) \geq \mathcal{F}(P_i^{true})\} = 1 - \alpha,$$

<sup>1.</sup> The disadvantage can be mitigated by using parallel computation. For example, McMaster University's SharcNet supercomputer with the current 128 nodes can be utilized for this purpose.

<sup>2.</sup> In case of the CU approach, the form of the function  $\mathcal{F}$  is implicitly included in the numerical method. The uncertainty  $\tau_{\alpha}$ , the result of the computation, depends on the form of  $\mathcal{F}$  and hence the property (7) is true for any  $\mathcal{F}$ .

$$P\{\boldsymbol{U}^{\alpha} \geq \boldsymbol{T}^{true}\} = 1 - \alpha$$

This is the desired property (7). Thus, in case of a monotone increasing function  $\mathcal{F}$ , we have two rigorously computed estimates of "maximum fuel temperatures",  $U^{\alpha}$  and  $T^{\alpha}$  defined, respectively, by (9) and (6). As argued before, it is "cheaper" to compute  $U^{\alpha}$ . However, it may or may not be a *better* estimate than  $T^{\alpha}$ . (We will consider  $U^{\alpha}$  to be a *better* estimate then  $T^{\alpha}$  if  $U_{95}^{\alpha} < T_{95}^{\alpha}$ .) Again, this depends on the specific form of the function  $\mathcal{F}$ . (We note that the actual function  $\mathcal{F}$  is likely not monotone in all its variables and hence the CU method is needed if rigorous results with known statistical properties are required. We include the monotone case as an exercise to gain deeper understanding of the issues involved.)

In order to demonstrate the above ideas, we have arbitrarily selected five different monotone increasing functions  $\mathcal{F}$  and computed the corresponding estimates  $U^{\alpha}$  and  $T^{\alpha}$  by the *updated* BEAU method and the CU method, respectively. The computed probability density functions (PDFs) for each selected function  $\mathcal{F}$  (denoted by  $\mathbf{T}_1, ..., \mathbf{T}_5$ ) are depicted in the graphs below together with the graphs of the functions  $\mathbf{T}_i = \mathbf{T}_i(\mathbf{S})$  on the left hand side. The first two cases are examples of convex functions for which the CU approach is "better" (as defined above). The third case is a concave function and the CU approach is "worse". The last two cases are examples of neither convex nor concave functions and the obtained results are essentially unpredictable. In the fifths case the CU approach is "worse", while in the fourth case it is "better". Interestingly enough, in the latter case, the result would be opposite for a percentile larger than 95%. Suffice it to say that the form of the function  $\mathcal{F}$  is crucial in determining the final result.

The effect of the form of the function  $\mathcal{F}$  is somewhat easier to understand in case of the *updated* BEAU approach. The *mode*, or, the peak of the distribution for  $U^{\alpha}$  corresponds to, approximately, 6900 kW for all five cases. No such, or similar observation can be made about the position or the form of the distribution for  $T^{\alpha}$ . The reason is that the function  $\mathcal{F}$  significantly alters the "fuel temperature" distribution across the reactor channels which alters the extreme value distribution of the result (cf. Reference X), i.e., the maximum "fuel temperature", and hence the error  $\tau_{\alpha}$ . Thus it is difficult to predict  $T^{\alpha}$  a priori.



- 11 -





#### REFERENCES

- [1] Internal Ontario Power Generation and Bruce Power documentation on LLOCA BEAU analysis.
- [2] P. Sermer, C. Olive and F.M. Hoppe, Efficient Compliance with Prescribed Bounds on Operational Parameters by Means of Hypothesis Testing using Reactor Data, Proceedings of the CNS 2000 Annual Conference, Toronto, June 11 - 14, 2000.
- [3] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge 1986, pp. 529 532.