UPPER SUBCRITICAL CALCULATIONS BASED ON CORRELATED DATA

Vladimir Sobes, Bradley T. Rearden, Don E. Mueller, William J. Marshall, John M. Scaglione, Michael E. Dunn ·
Oak Ridge National Laboratory
P.O. Box 2008, Oak Ridge, TN, 37831
sobesv@ornl.gov, reardenb@ornl.gov, muellerde@onrl.gov, marshallwj@ornl.gov,
scaglionejm@ornl.gov, dunnme@ornl.gov

ABSTRACT

The American National Standards Institute and American Nuclear Society standard for *Validation of Neutron Transport Methods for Nuclear Criticality Safety Calculations* defines the upper subcritical limit (USL) as "a limit on the calculated *k*-effective value established to ensure that conditions calculated to be subcritical will actually be subcritical." Often, USL calculations are based on statistical techniques that infer information about a nuclear system of interest from a set of known/well-characterized similar systems. The work in this paper is part of an active area of research to investigate the way traditional trending analysis is used in the nuclear industry, and in particular, the research is assessing the impact of the underlying assumption that the experimental data being analyzed for USL calculations are statistically independent. In contrast, the multiple experiments typically used for USL calculations can be correlated because they are often performed at the same facilities using the same materials and measurement techniques. This paper addresses this issue by providing a set of statistical inference methods to calculate the bias and bias uncertainty based on the underlying assumption that the experimental data are correlated. Methods to quantify these correlations are the subject of a companion paper and will not be discussed here.

The newly proposed USL methodology is based on the assumption that the integral experiments selected for use in the establishment of the USL are sufficiently applicable and that experimental correlations are known. Under the assumption of uncorrelated data, the new methods collapse directly to familiar USL equations currently used. We will demonstrate our proposed methods on real data and compare them to calculations of currently used methods such as USLSTATS and NUREG/CR-6698. Lastly, we will also demonstrate the effect experiment correlations can have on USL calculations.

KEYWORDS
subcritical limit, correlation, bias, uncertainty

· This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

1. INTRODUCTION

The American National Standards Institute (ANSI) and the American Nuclear Society (ANS) standard for *Validation of Neutron Transport Methods for Nuclear Criticality Safety Calculations* [1] defines the upper subcritical limit (USL) as “a limit on the calculated *k*-effective value established to ensure that conditions calculated to be subcritical will actually be subcritical.” Often, USL calculations are based on statistical techniques that infer information about a nuclear system of interest (application) from a set of known/well-characterized similar systems (experiments).

In determining a USL, it is common to statistically determine a calculation’s *bias* and *bias uncertainty*. *Bias* is “the systematic difference between calculated results and experimental data.” [1] The *bias uncertainty* is “the uncertainty that accounts for the combined effects of uncertainties in the benchmarks, the calculational models of the benchmarks, and the calculational methods.” [1] Both the bias and the bias uncertainty are often inferred through statistical methods applied to the selected set of experiments.

Trending analysis has been the workhorse of USL calculations in the nuclear industry for many years. Currently, ORNL is reexamining the way traditional trending analysis is used in the nuclear industry. In particular, the motivation for this work is to assess the impact of the underlying assumption that the experimental data being analyzed for USL calculations are statistically independent. It is important to note that the multiple experiments typically used for USL calculations can be correlated because they are often performed at the same facilities using the same materials and measurement techniques. This paper addresses this issue by providing a set of statistical inference methods to calculate the bias and bias uncertainty based on the underlying assumption that the experimental data are correlated.

Recent developments in the field of integral experiment correlations are allowing researchers to determine the statistical correlations between those experiments. The Sampler Sequence within SCALE 6.2 developed at Oak Ridge National Laboratory is an example of one of those developments. [2] All of the statistical methodologies in this paper are based on the assumption that the integral experiments selected for use in the establishment of the USL are sufficiently applicable and that experimental correlations are known. The establishment of applicability and the quantification of correlations are the subjects of numerous other publications and are outside the scope of this work.

The procedures for calculating USLs at nuclear facilities which operate under the United States Department of Energy (DOE) are governed by ANSI/ANS Standards. The USL calculations in the private sector are governed by the United States Nuclear Regulatory Commission under 10CFR50.68. [3] In this paper, we are concerned with meeting the statistical requirements laid out in 10CFR50.68. The requirements can be summarized in this quotation:

The estimated ratio of neutron production to neutron absorption and leakage (k-effective) ... must not exceed 0.95, at a 95 percent probability, 95 percent confidence level.

Lastly, even though the primary application of the research presented in this paper is for USL calculations, there is nothing in the mathematical derivations that limits the application of the proposed methods to the nuclear industry. Therefore, this work can simultaneously serve as a general reference on determining statistical limits from correlated experimental data.

2. MODEL DEFINITION

For the remainder of this paper we will assume the following nomenclature:

p = Number of independent variables describing the benchmark experiments. Scalar value.
n = Number of benchmark experiments that depend on the independent variables. Scalar value. $p \leq n$.
X = Independent variables for all benchmarks. Examples: Energy of Average Lethargy of Neutrons Causing Fission (EALF), number of fueled rods, reflector thickness, correlations due to nuclear data uncertainties, etc. Matrix $n \times p$.
Y = Dependent variables. Example: k_{eff} for benchmark experiments. Column vector $n \times 1$.
β = Regression parameters. Column vector $p \times 1$.
Σ = Known correlation matrix for the dependent variable. Matrix $n \times n$.
σ = The square root of the variance of each dependent variable. Such that $\sigma^2 \Sigma$ is the covariance matrix of Y.

It should be noted that the methods presented here enable the simultaneous use of multiple independent variables. For example, one can perform simultaneous regression on k_{eff} versus the number of fueled rods and reflector thickness. However, as the name implies, the independent variables must be independent of each other.

The assumption necessary for all the methods proposed in the rest of this paper is this

$$Y = X\beta + E \,, \tag{1}$$

$$E \sim N(0, \sigma^2 \Sigma) \,. \tag{2}$$

That is, Y is linear in the vector of regression parameters, and the true error, E, has a multivariate normal distribution with zero mean and covariance matrix $\sigma^2 \Sigma$.

Other than accounting for experimental correlations, this model is no different from the typical model that is assumed in most traditional regression methods for calculating USLs. We must bring to light, however, what this model actually means. It means that we are going to assume that all of the experiments actually come from one multivariate normal distribution with mean $X\beta$ and covariance matrix $\sigma^2 \Sigma$. Furthermore, in order for us to be able to infer anything about our application system, we must assume that the application system also comes from the same distribution. This requirement is met where the same computer code system, nuclear data libraries, modeling techniques, etc., are used in calculating k_{eff} for the benchmark experiments used in establishing the USL and the analysis of the application system.

3. RESULTS

In theory, there are many ways of calculating a statistical estimator of the bias based on a set of experiments; we present the least squares and maximum likelihood estimator. This is also the type of estimator used in traditional trending analysis for USL calculations in the nuclear engineering field.

We begin with the model already discussed in the Section 2, that is,

$$Y = X\beta + E \,, \tag{3}$$

$$E \sim N(0, \sigma^2 \Sigma)\,. \tag{4}$$

Given a set of coordinate pairs *(X_i, Y_i)* we seek a vector β that will minimize the linear algebra expression,

$$(Y - X\beta)'(\sigma^2\Sigma)^{-1}(Y - X\beta)\,, \tag{5}$$

where the notation, A', indicates the transpose of A. Since σ^2 is a positive real constant, minimizing the above expression is equivalent to minimizing

$$(Y - X\beta)'\Sigma^{-1}(Y - X\beta)\,. \tag{6}$$

We proceed by taking the derivative of the above expression with respect to β and setting it to zero.

$$0 = -X'\Sigma^{-1}(Y - X\beta) - (Y - X\beta)'\Sigma^{-1}X\,. \tag{7}$$

Since, Σ, is a positive semi-definite correlation matrix, the above equation is of the form

$$A' + A = 0\,, \tag{8}$$

and $A = 0$ is the only solution. Therefore, the solution is

$$X'\Sigma^{-1}X\beta = X'\Sigma^{-1}Y\,, \tag{9}$$

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y\,. \tag{10}$$

Using the estimator of β, the unbiased χ^2 estimator of σ^2 can be written as,

$$s^2 = \left(Y - X\hat{\beta}\right)'\Sigma^{-1}\left(Y - X\hat{\beta}\,\right). \tag{11}$$

Furthermore, the two estimators are independent and are distributed as

$$\beta \sim N\left(\hat{\beta}, \sigma^2 \left(X'\Sigma^{-1}X\right)^{-1}\right), \tag{12}$$

$$\frac{(n-p)s^2}{\sigma^2} \sim \chi^2_{n-p}\,. \tag{13}$$

From the above distributions, it is obvious that the uncertainty on the bias, β, for this estimator is $\sigma^2(X'\Sigma^{-1}X)^{-1}$. With σ^2 unknown, we can naturally state that the true mean of our distribution, $\hat{Y}$, will not be less than,

$$X\beta - t_{n-p}^{\alpha/2}\left(diag(Xs^2(X'\Sigma^{-1}X)^{-1}X')\right)^{1/2}, \tag{14}$$

$100\alpha/2$ percent of the time, where $t_{n-p}^{\alpha/2}$ is the upper $100\alpha/2$ percent point of the *t* distribution with *n-p* degrees of freedom, and $diag(X(s^2(X'\Sigma^{-1}X)^{-1})X')$ is simply the estimator of the variance of β propagated to the value of $X\beta$. The *diag* operator turns the main diagonal of a square matrix that it is acting on into a column vector.

It is well known in the statistics field that in order to produce a simultaneous interval from this point-wise interval, the value of $t_{n-p}^{\alpha/2}$ must be replaced by $\left(pF_{p,n-p}^{\alpha}\right)^{1/2}$, where $F_{p,n-p}^{\alpha}$ is the upper 100α percent point of the *F* distribution with *p* degrees of freedom in the numerator and *n-p* degrees of freedom in the denominator. Therefore we can state that the true mean $\hat{Y}$ will be less than the simultaneous limit,

$$X\beta - \left(pF_{p,n-p}^{\alpha}\right)^{1/2}\left(diag(Xs^2(X'\Sigma^{-1}X)^{-1}X')\right)^{1/2}, \tag{15}$$

$100\alpha/2$ percent of the time.

We have also assumed in our model that the population is distributed around the true mean $\hat{Y}$ as $N(0,\sigma^2\Sigma)$. Therefore *100(1-P)%* of the population will be below $\hat{Y} - Z(P)\sigma$, where we have defined *Z(P)* as the *100P%* lower percent point of the standard normal distribution. Once again, σ is unknown. However, we do know the distribution of the estimator s^2, and therefore we can bind σ with a certain degree of confidence. We now state that $100\alpha/2$ percent of the time,

$$\sigma \leq \left(s^2 \frac{n-p}{\chi^2_{\alpha/2,n-p}}\right)^{1/2}, \tag{16}$$

where, $\chi^2_{\alpha/2,n-p}$, is the lower $\alpha/2$ percentile of the χ^2 distributions with *n-p* degrees of freedom.

At last, we can state that the true mean $\hat{Y}$ will not be in the calculated limit $100(\alpha/2)\%$ of the time, and 100*(1-P)%* of the population will be below $\left(\hat{Y} - Z(P)\left(s^2(n-p)/\chi^2_{\alpha/2,n-p}\right)^{1/2}\right)$ $100(\alpha/2)\%$ of the time. Therefore, we can say that 100*P%* of the population will be above the calculated USL,

$$USL = X\beta - \left(pF_{p,n-p}^{\alpha}\right)^{1/2}\left(diag(Xs^2(X'\Sigma^{-1}X)^{-1}X')\right)^{1/2} + Z(P)\left(\frac{s^2(n-p)}{\chi^2_{\alpha/2,n-p}}\right)^{1/2}, \tag{17}$$

at least $100(1-\alpha)\%$ of the time.

Both the bias and the USL calculations collapse to traditional trending analysis equations if experiment correlations are ignored (i.e., the correlation matrix Σ is set to unity (identity matrix)). Therefore, our proposed methods can be considered to be just a generalized version of the traditional trending analysis.

In the next section we will compare our proposed method for calculating the USL to methods currently used in the industry. The first two methods presented are from the code USLSTATS. [4] The two methods are called USL1 and USL2 and are described in full in Ref. [4]. The other method presented is the *single-sided simultaneous tolerance band* that we have calculated using the methodology presented in NUREG/CR-6698. [5]

In order to facilitate a fair comparison between all of the methods, we will not use any administrative margin for all of the methods presented. Further, all of the calculations are done for 95% confidence and 95% probability. Lastly, all of the methods follow the explicit guidance of ISG-10 NRC:

No credit should be taken for positive bias, because this would result in making changes in a non-conservative direction without having a clear understanding of those changes. [6]

4. APPLICATION

This section demonstrates our proposed methodology on data from a previous nuclear criticality safety validation. In Fig. 1 the calculated k_{eff} is divided by the experimental k_{eff} (C/E) plotted versus *EALF* from 30 critical experiments that were chosen in the validation to be a representative data set for the application. Fig. 1 also shows the 1 standard deviation error bars for the *C/E* data.

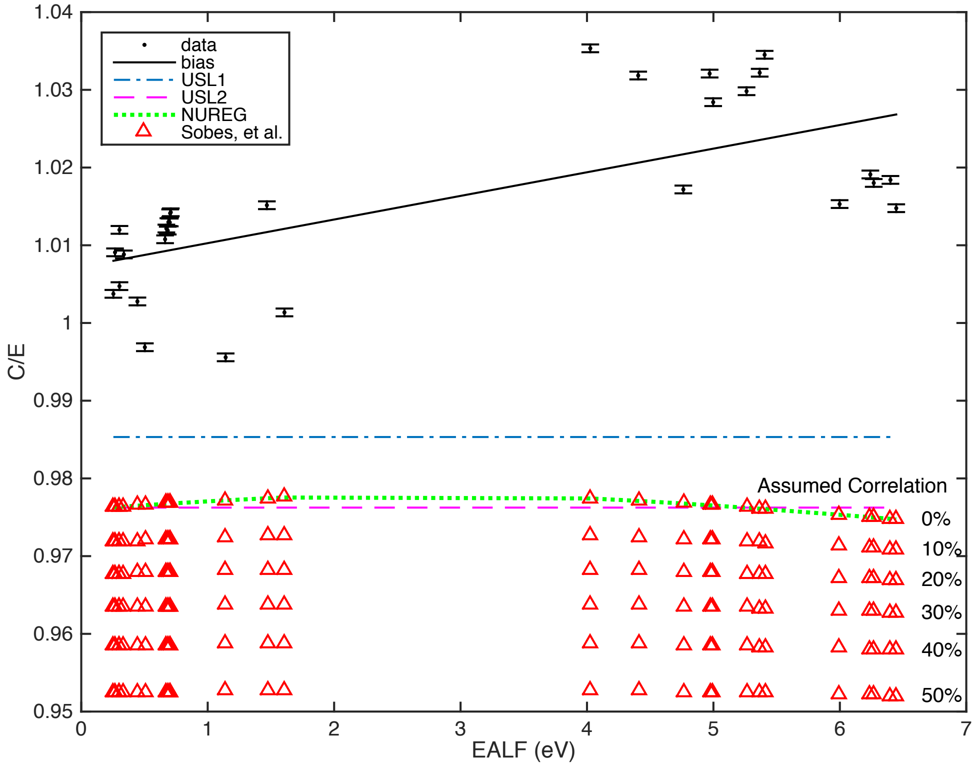

Figure 1. Plot of the data for the previous validation study along with several USL calculations, all at 95% confidence and 95% probability. The top set of Sobes et. al. data corresponds to assuming that the data are uncorrelated and the sets below increase the experiment correlations in increments of 10%.

On the same figure several traditional trending analysis limits currently used in the industry are plotted. Both USL1 and USL2 are plotted without administrative margins. In the middle of Fig. 1, the *single-sided simultaneous tolerance band* has been calculated using the methodology presented in NUREG/CR-6698. Since no credit has been taken for positive bias, which the experimental data have in this case, all of the USL calculations have no slope.

Lastly, the USL calculations have been plotted based on our methodology. The top set of triangles corresponds to the calculation of our method when we assume that the experimental data are uncorrelated. Notice that our method coincides almost exactly with the NUREG/CR-6698 calculation if we assume that the data are uncorrelated.

The exact amount of shared uncertainty between this set of experiments has not yet been quantified. The area of quantification of correlated uncertainty is currently an active research area. Even the range of possible values for uncertainty shared between similar experiments is a much-debated topic, with results varying from negligibly small correlation coefficients to values as high as 80%, [2] depending on the assumptions involved. Therefore, the analysis in this work should not be considered an assessment or review of the previous validation study. Rather, the work illustrates the possible impact experimental correlations could have on a USL calculation if the experimental correlations are known. As noted previously, the amount of shared uncertainty between the set of experiments examined in this paper has not been quantified.

If we assume, for demonstration purposes, that all experiments are correlated to each other and that 10% of the uncertainty in each experiment is shared among all experiments, we calculate the second-from-the-top set of triangles. The result is a USL that is on average 450 percent mille k_{eff} (pcm) lower than if the experiments are assumed to be uncorrelated. Note that all of the other methods presented in Fig. 1 have zero administrative margins. As we increase the correlation level between the integral experiments, we get progressively farther and farther away from the bias (Fig. 1.). This makes intuitive sense: as the level of the correlation of the integral experiments increases, there is less and less independent information from which to determine the USL.

5. CONCLUSION

This paper addresses a long-standing issue in the nuclear community of how to carry out USL calculations when the experimental data are known to be correlated. We have done so by providing a methodology to estimate the bias and calculate the USL based on correlated data. We have stated all of the assumptions of our model clearly, two of which are that the experiments are applicable for USL determination and that the experiment correlation matrix is known.

In the preceding section, we have demonstrated the application of our methodology on real data and compared our methods to both USLSTATS calculations and the methodology from NUREG/CR-6698. Most importantly, we have shown, in a quantifiable way, the effect that experiment correlations have on USL calculations.

ACKNOWLEDGMENTS

This work was supported by the DOE Nuclear Criticality Safety Program, funded and managed by the National Nuclear Security Administration for DOE.

REFERENCES

1. "Validation of Neutron Transport Methods for Nuclear Criticality Safety Calculations," ANS-8.24-2007 (R 2012), American Nuclear Society (2007).
2. W. J. Marshall and B. T. Rearden, "Determination of Experimental Correlations Using the Sampler Sequence within SCALE 6.2," *ICNC 2015*, Charlotte, North Carolina, September 13–17, 2015.
3. United States Nuclear Regulatory Comission. "Part 50—Domestic Licensing of Production and Utilization Facilities," United States Nuclear Regulatory Comission. http://www.nrc.gov/reading-rm/doc-collections/cfr/part050/ (2014).
4. J. J. Lichtenwalter, S. M. Bowman et. al., "Criticality Benchmark Guide for Light-Water-Reactor Fuel in Transportation and Storage Packages," NUREG/CR-6361, U.S. Nuclear Regulatory Commission (1997).
5. J. C. Dean, R. W. Tayloe, Jr., D. Morey, "Guide for Validation of Nuclear Criticality Safety Calculatioal Methodology," NUREG/CR-6698, US Nuclear Regulatory Commission (2001).
6. "Notice of Availability of Interim Staff Guidance Documents for Fuel Cycle Facilities," FCSS ISG-10, Rev. 0 (2005).