SAND Report 2008-0632A, AIAA Paper AIAA-2008-5944

Model Calibration under Uncertainty: Matching Distribution Information

Laura P. Swiler[*], Brian M. Adams[†], and Michael S. Eldred[‡]
Sandia National Laboratories[§], Albuquerque NM 87185

We develop an approach for estimating model parameters which result in the "best distribution fit" between experimental and simulation data. Best distribution fit means matching moments of experimental data to those of a simulation (and possibly matching a full probability distribution). This approach extends typical nonlinear least squares methods which identify parameters maximizing agreement between experimental points and computational simulation results. Several analytic formulations for the distribution matching problem are provided, along with results for solving test problems and comparisons of this parameter estimation technique with a deterministic least squares approach.

I. Introduction

Nonlinear models are frequently used to model physical phenomena, including engineering applications. In this paper, we refer to a nonlinear model very broadly: the output of the model is a nonlinear function of the parameters [Draper98]. Thus, nonlinear models can include systems of partial differential equations (PDEs). Some examples include CFD (computational fluid dynamics), groundwater flow, heat transport, etc. Nonlinear models also include functional approximations of uncertain data via regression or response surface models. In most cases, we have some type of simulation model which is a nonlinear model, so we use the term nonlinear model and simulation model interchangeably in this paper. In addition to a simulation model, we assume that we have experimental data which may be used to calibrate the model. The calibration often requires the solution of an optimization problem to determine the optimal parameter settings for the simulation model.

In this paper, we are concerned with identifying model parameters which result in a "best fit" between experimental data and simulation results in a nondeterministic context. That is, instead of matching point estimates we are concerned with matching moments (e.g., mean or variance) between experimental and simulation data where the variability in the model output is due to parametric uncertainty. We develop an approach extending the typical nonlinear least squares formulation to allow for this distribution matching. Note that "parameters" may be parameters in an approximation model such as a regression model, or physics modeling parameters which are used in physical simulation models such as PDEs. We distinguish data from parameters: data are physical data which are input either to a regression or physical simulation. For example, in groundwater flow modeling, hydraulic conductivity is a parameter and data may include measured flow rates from well tests. In this paper, we denote parameters that will be calibrated as $\boldsymbol{\theta}$, and the independent input data (e.g., state variables, configuration data, boundary conditions) as $\boldsymbol{x}$. We also assume that there are uncertain variables, denoted by $\mathbf{u}$, that represent inherent variability or lack of exact knowledge influencing the simulation, but which we cannot observe. The effect of the uncertain variables is reflected in both the output variability of the nonlinear model and the experimental data, but we can only explicitly account for these uncertain variables in the simulation model. Thus, our simulation model, f,

[*] Principal Member of Technical Staff, Optimization and Uncertainty Estimation Dept., P.O. Box 5800, Sandia National Laboratories, MS 1318.
[†] Senior Member of Technical Staff, Optimization and Uncertainty Estimation Dept., P.O. Box 5800, Sandia National Laboratories, MS 1318.
[‡] Principal Member of Technical Staff, Optimization and Uncertainty Estimation Dept., P.O. Box 5800, Sandia National Laboratories, MS 1318, Associate Fellow AIAA.
[§] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

is a function of ($\boldsymbol{x}$, $\boldsymbol{\theta}$, $\mathbf{u}$) but the experimental observations y are only a function of $\boldsymbol{x}$: we observe $\boldsymbol{x}$ and $y(\boldsymbol{x})$. Given these assumptions, we want to use a calibration framework in a specific way, to aggregate or "integrate out" the effects of the uncertain variables at the simulation model level by approximating their effects with statistical metrics such as moments, and matching such moments to those of the experimental data.

There are several motivations for matching distribution information between the model and the experimental data instead of directly matching point values. Often experimental data is of poor quality, has very few data points, or may not be directly relevant to the particular simulation model under investigation. In this case, we may just want to match means and/or variances. Another case is where modelers just want to obtain a ballpark estimate based on bounds matching: that is, analysts want the model upper and lower bound estimates (accounting for model uncertainty) to match the upper and lower bounds on the data. One might have several sets of experiments and want to match the mean or moments of each to a simulation model. Finally, the model may have some uncertain variables (some of which may be "calibrated" by choosing parameter values to describe them such as distribution means, and other uncertain variables will not be calibrated) and we want to incorporate this uncertainty in the response.

The outline of this paper is as follows: we present a brief overview of nonlinear least squares optimization techniques in Section II, and provide several formulations for matching moments or distributions in Section III. In Section IV we outline a number of approaches which can be used to solve these formulations, with examples and results in Section V, followed by the conclusions in Section VI.

II. Nonlinear Least Squares Optimization Methods

Nonlinear regression extends linear regression for use with a much larger and more general class of functions [Bates88]. Unlike linear regression, there are few limitations on the way parameters may enter into the functional part of a nonlinear regression model. The way in which the unknown parameters in the nonlinear function are estimated, however, is conceptually the same as in linear least squares regression.

The nonlinear model of the response y as a function of the n-dimensional inputs $\boldsymbol{x}$ is given as:

$$y = f(\boldsymbol{x};\boldsymbol{\theta}) + \varepsilon \tag{1}$$

where f is the nonlinear model, $\boldsymbol{\theta}$ is a vector of parameters to be calibrated, ε is a random error term, and we assume that $\mathbf{E}[\varepsilon] = 0$ and $\mathbf{V}ar[\varepsilon] = \sigma^2$ and the error terms are independent and identically distributed (iid). Usually y is a function of $\boldsymbol{x}$ but this dependence is often implicit and $y(\boldsymbol{x})$ simply written as y. Note that for nonlinear functions, the derivative of f with respect to the parameters $\boldsymbol{\theta}$ depends on at least one of the parameters of the vector $\boldsymbol{\theta}$. Given observations of the response y corresponding to the independent variables $\boldsymbol{x}$, the goal of nonlinear regression is to find the optimal values of $\boldsymbol{\theta}$ to minimize the error sum of squares function $S(\boldsymbol{\theta})$, also referred to as SSE:

$$S(\boldsymbol{\theta}) = \sum_{i=1}^{n} [(y_i - f(\boldsymbol{x}_i;\boldsymbol{\theta})]^2 = \sum_{i=1}^{n} [R_i(\theta)]^2 \tag{2}$$

where $R_i(\theta)$ are the residual terms. Nonlinear regression employs an optimization algorithm to find the least squares estimator $\hat{\boldsymbol{\theta}}$ of the true minimum $\boldsymbol{\theta}^*$; a process that is often difficult [Seber03]. Derivative-based nonlinear least squares optimization algorithms exploit the structure of such a sum of squares objective function. If $S(\boldsymbol{\theta})$ is differentiated twice, terms of residual $R_i(\theta)$, $R_i''(\theta)$, and $[R_i'(\theta)]^2$ result. By assuming that the residuals $R_i(\theta)$ are close to zero near the solution, the Hessian matrix of second derivatives of $S(\boldsymbol{\theta})$ can be approximated using only first derivatives of $R_i(\theta)$.

An algorithm that is particularly well-suited to the small-residual case and the above formulation is the Gauss-Newton algorithm. This formulation and algorithm combination typically requires the user to explicitly formulate each term in the least squares (e.g., n terms for n data points) along with the gradients for each term. This may be very expensive for computationally intensive evaluations of f, and even more so when finite differences are used, since the number of necessary calculations also will increase as the number of parameters increases. Additionally, the approximation of gradients in the presence of errors in the problem is problematic; the gradient approximation may have larger errors than the objective function approximation [Borggaard]. Despite these limitations, gradient-based methods are usually much more efficient than derivative-free approaches. Three gradient-based methods tailored for nonlinear least squares optimization problems are available in a software tool called DAKOTA [Eldred06]. They are a standard Gauss-Newton algorithm, NLSSOL, and NL2SOL. NLSSOL uses a sequential

quadratic programming formulation (SQP), and exploits the structure of the least squares objective function through the periodic use of Gauss-Newton Hessian approximations to accelerate the SQP algorithm. NL2SOL uses a trust-region method (and thus can be viewed as a generalization of the Levenberg-Marquardt algorithm) and adaptively chooses between two Hessian approximations, the Gauss-Newton approximation alone and the Gauss-Newton approximation plus a quasi-Newton approximation to the rest of the Hessian. This enables good performance on problems with large residuals at the optimal parameter values, but even on small-residual problems, the latter Hessian approximation can be useful when the initial parameter iterate is far from the solution. On problems that are not over-parameterized (i.e., that do not involve more optimization variables than the data support), NL2SOL usually exhibits fast convergence. In the case of high cost and/or questionable accuracy in computing gradients, global optimization algorithms not requiring derivatives, also available in DAKOTA, may be necessary.

III. Matching Distribution Information

Often we are interested in running a simulation model repeatedly to obtain information about the spread and variability in its response. In particular, uncertainty propagation refers to the process where, given a characterization of uncertain model input values **u**, a method such as Monte Carlo sampling is used to obtain model output values which reflect the input uncertainty. One may be interested in moments, bounds, or a full distribution characterization of these output values. In this paper, we assume the simulation model is deterministic: for a particular set of input parameter values (***x*, θ, u**), the model always returns the same output.

If we sample over the uncertain input values, e.g. draw values from the probability distributions describing **u**, then for a particular set of data values ***x*** and calibration parameter values **θ**, we can obtain a distribution of response values *y*. This yields more information from the simulation than the set of pairs $(y_i, f(\mathbf{x}_i;\boldsymbol{\theta}))$ that are typically matched in regression analysis. For example, the computational model *f* depends on a realization $\boldsymbol{u}_j$ of the uncertain variables: $f(\boldsymbol{x}_i;\boldsymbol{\theta};\boldsymbol{u}_j)$, and sampling over values $\boldsymbol{u}_j$ yields a distribution of simulation values associated with each experimental point $y(\boldsymbol{x}_i)$. In this work, we aggregate the effects of the uncertainty by calculating moments or other statistical measures of *f* and calibrating to these instead of explicitly fitting the experimental data (which implicitly depends on **u**) to each of the simulation points $f(\boldsymbol{x}_i;\boldsymbol{\theta};\boldsymbol{u}_j)$.

In the following formulations, the vector **θ** always denotes calibration parameters, ***x*** denotes input data such as configuration data, state variables, boundary conditions, etc., and $\boldsymbol{u}_j$ represents a realization of the stochastic inputs **u**. It is possible for the characteristics of the uncertain input distributions to depend on **θ** (with an abuse of notation, **u(θ)**). Our approach can find the optimal values of calibration parameters $\boldsymbol{\theta}_A$ entering directly into the model, or those $\boldsymbol{\theta}_I$ which influence the characterization of **u,** as shown in Figure 1.

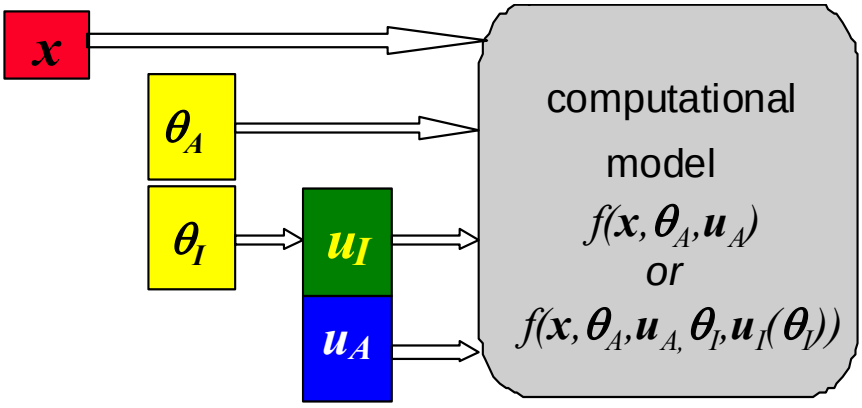

Figure 1. Depiction of simulation dependence on configuration variables *x*, calibration parameters θ, and uncertain variables u

This paper illustrates several variations on the standard nonlinear least squares (NLLS) formulation given in equation (2), each accounting for statistical behavior of the model/experimental response. In the first case, one has repeated values taken in a single experiment (e.g., one experimental configuration) and wants to obtain parameter

values so that the mean of the simulation matches the mean of the experimental results as shown in Figure 2. Note that the experimental and simulation data may not follow the same distribution in the case of matching means. The objective function for the NLLS formulation in this case is:

$$S(\boldsymbol{\theta}) = (\mu_{\exp} - \mu_{sim})^2 = \left(\frac{\sum_{i=1}^{n} y_i}{n} - \frac{\sum_{j=1}^{m} f(\mathbf{x};\boldsymbol{\theta};\boldsymbol{u}_j)}{m} \right)^2 \tag{3}$$

where $\mu_{\exp}$ is the experimental mean, μ_{sim} is the mean of the simulations taken over realizations $\boldsymbol{u}_j$ of $\boldsymbol{U}$, and the number of experimental data points (n) may be different from the number of simulation runs (m).

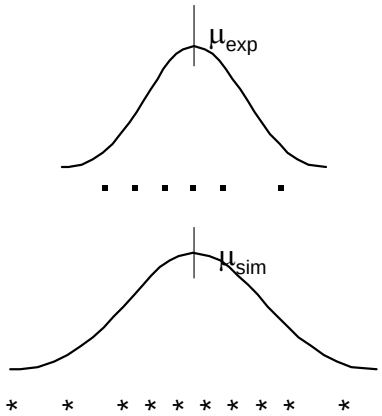

Figure 2. Matching simulation and experimental means

Note that we could choose to match both means and variances or standard deviations (and higher order moments as well). Figure 3 shows the case of matching means and standard deviations. In this case, the NLLS formulation becomes: $S(\boldsymbol{\theta}) = (\mu_{\exp} - \mu_{sim})^2 + (\sigma_{\exp} - \sigma_{sim})^2$ where $\sigma_{\exp}$ and σ_{sim} are the standard deviation of the experimental data and simulation results, respectively. In this case, one might want to weight the residual terms in accordance with the importance of matching particular moments. Usually it is more critical to match the mean or lower moments, so these would be more heavily weighted. One non-heuristic approach to weighting mean and variance is to instead match specific moment projections formulation using non-generalized reliability indices. A reliability index can be thought of as the number of standard deviations a particular output value z^* is from the mean. The reliability indices are calculated as: $\beta = \frac{z^* - \bar{z}}{\bar{\sigma}}$, where $\bar{z}$ is the mean of the experimental response and $\bar{\sigma}$ is the standard deviation of the experimental response. Instead of explicitly matching the mean and standard deviation, the problem is reformulated to match several reliability indices corresponding to experimental data. Doing this implicitly matches the simulation mean and standard deviation to the experimental mean and standard deviation, because the simulation $\boldsymbol{\beta}$ terms are calculated as $\beta_{sim} = \frac{z^*_{sim} - \bar{z}_{sim}}{\bar{\sigma}_{sim}}$. A strength of this formulation is that various response values used to calculate the $\boldsymbol{\beta}$ terms implicitly determine weights for the mean and standard deviation. This approach may be more appropriate than arbitrarily choosing a weighting scheme such as 2:1 for mean: standard deviation. While this may appear to be CDF matching, it is not, since only the underlying weighted moments are being estimated. Later, generalized reliability indices $\boldsymbol{\beta}^*$ are estimated; this is a form of CDF matching.

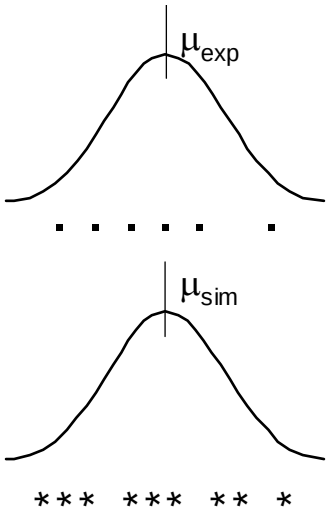

Figure 3. Matching simulation and experimental means and variances

The next case is more interesting: matching the means between simulation and experimental data at several experimental configurations, where each configuration contains replicated points and allows a different variance at each experimental configuration. This is similar to nonlinear weighted least squares, with the difference that we are matching means and not the actual data at these different configurations. In this situation, the NLLS objective function is:

$$S(\boldsymbol{\theta}) = \sum_{k=1}^{K} (\mu_{\exp(k)} - \mu_{sim(k)})^2 = \sum_{k=1}^{K} \left(\frac{\sum_{i=1}^{n_k} y_{ik}}{n} - \frac{\sum_{j=1}^{m_k} f(\mathbf{x};\boldsymbol{\theta};\boldsymbol{u}_j)_k}{m} \right)^2 \qquad (4)$$

where each experimental configuration k has its own mean, and both the number of experimental data points and simulation data points may vary across the K configurations. Note also that K could represent a number of response functions which one is trying to simultaneously match, instead of a number of experimental configurations. Figure 4 shows this case for K=3 configurations or responses.

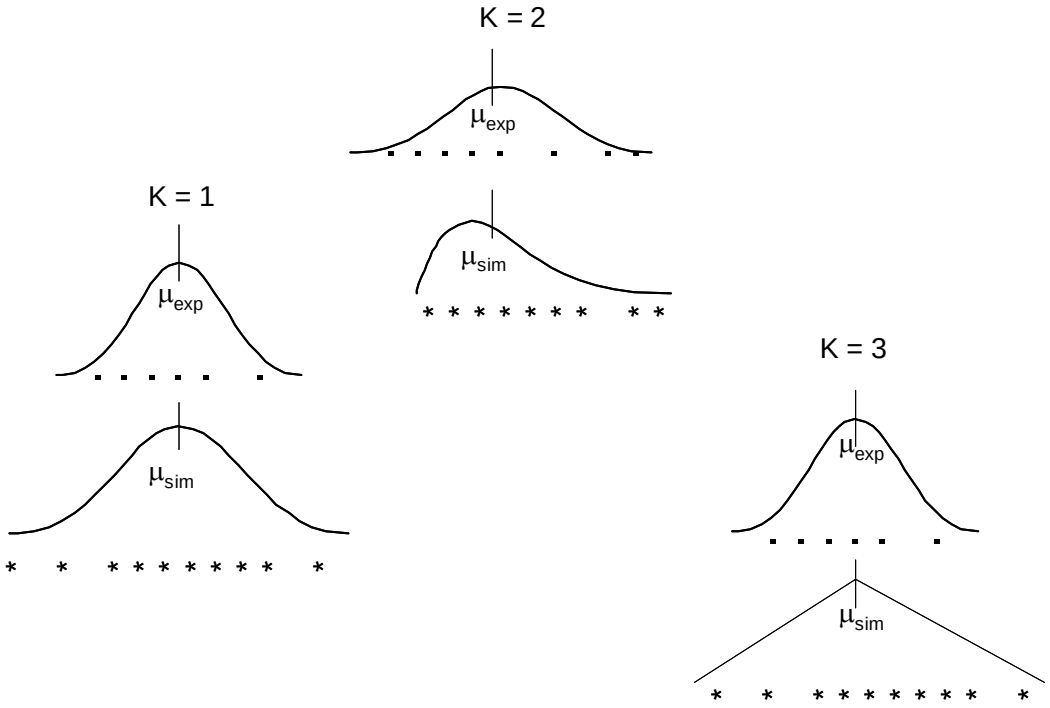

Figure 4. Matching several simulation and experimental means

Finally, we match entire probability distributions by matching discrete values from the experimental and simulation cumulative distribution functions (CDFs), although other distance metrics between probability distributions could be used. For example, in Figure 5 we show an example where we match the response levels at 5 fixed percentile values p_k (= 0.2, 0.4, 0.6, 0.8, and 1.0) from a CDF of the experimental data to a CDF of the simulation data. This type of parameter optimization can be very powerful: we are choosing values of input parameters which yield output results from simulation runs to match a generic experimental distribution. This NLLS formulation can be given in terms of experimental response levels Y_k and model response f_k as:

$$S(\boldsymbol{\theta}) = \sum_{k=1}^{K} (CDF_{\exp(k)} - CDF_{sim(k)})^2 = \sum_{k=1}^{K} \left(\left[Y_k : \Pr(y \le Y_k) = p_k \right] - \left[f_k : \Pr(f(\mathbf{x};\boldsymbol{\theta};\boldsymbol{U}) \le f_k) = p_k \right] \right)^2 \quad (5)$$

where the CDF of both the experimental data and the simulation data is discretized to K levels. For each level, simple counting, distribution fitting, or interpolation may be used to calculate the empirical response level corresponding to the probability levels p_k. Alternately, one may fix response levels Y_k of interest and use probabilities in the least squares criterion as shown in equation 6, where simple counting, distribution fitting, and/or interpolation may be used to calculate the probability that both the experimental data and simulation data are less than a certain value Y_k:

$$S(\boldsymbol{\theta}) = \sum_{k=1}^{K} (CDF_{\exp(k)} - CDF_{sim(k)})^2 = \sum_{k=1}^{K} \left(\Pr(y \le Y_k) - \Pr(f(\mathbf{x};\boldsymbol{\theta};\boldsymbol{U}) \le Y_k) \right)^2 \quad (6)$$

For example, if Y_k = 52.5, and the probability that the experimental data was less than or equal to 52.5 is 0.8, then the NLLS optimization would seek values of $\boldsymbol{\theta}$ such that the probability that the simulation model yields a value less than or equal to 52.5 is also approximately 0.8. Note that the formulations in Equation 5 and 6 may be expanded to include the matching of several CDFs, similar to the multiple mean responses in Equation 4. Also, the formulations may be modified to allow for different measures of a stochastic response. For example, instead of matching percentiles along a CDF, one may match generalized reliability indices, which are just a nonlinear transformation of probabilities that often results in a more linear dependence on $\boldsymbol{\theta}$.

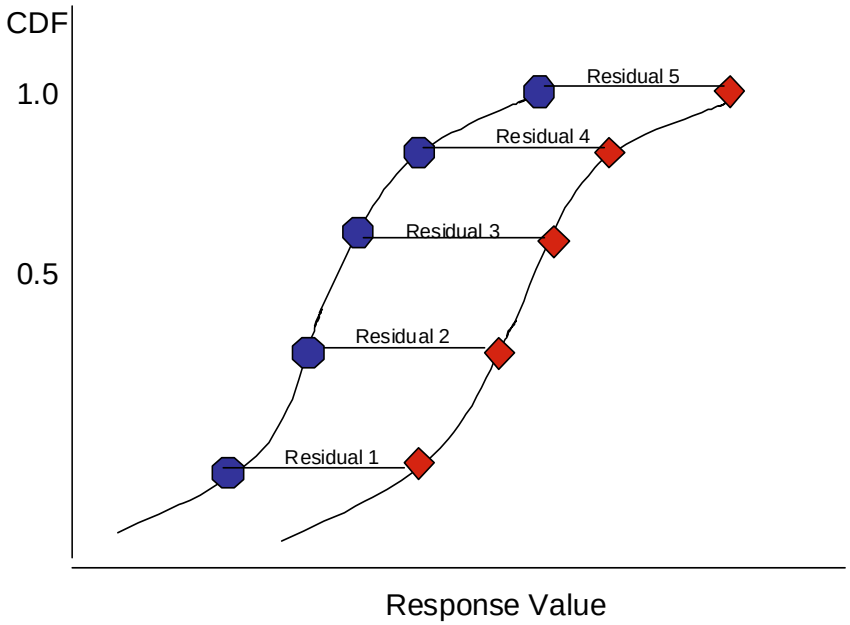

Figure 5. Matching response levels at percentile values values between the simulation (in red) and experiment (in blue)

In this section, we introduced modified nonlinear least squares formulations that allow matching stochastic quantities such as moments, percentiles, and response levels. Note that these formulations focus on the stochastic response of the simulation introduced by parameter uncertainty in the model. It is also possible to obtain confidence intervals on the calibrated parameter values, based on the variability in the experimental data and on certain assumptions about the linearity of the model and the distribution of experimental error. There are a variety of methods that one can use to obtain individual or joint confidence intervals (CIs) on parameters obtained in a NLLS formulation, include Bonferroni intervals, the linear approximation method, the F-test method, and the log-likelihood method. These methods along with application results are well documented in [Donaldson87], [Rooney01], and [Vugrin06]. The issue of generating confidence intervals around calibrated parameters is not addressed in this paper, but the CI formulations could extend naturally to the calibration framework being proposed.

IV. Solution Approaches: NLLS under uncertainty

The formulations in Section III are all examples of nonlinear least squares methods under uncertainty. A variety of methods may be used to solve such optimization under uncertainty (OUU) problems. For this paper, we will primarily focus on nested approaches, consisting of an outer loop nonlinear least squares optimization method (such as Gauss-Newton) and an inner loop uncertainty quantification (UQ) method that iterates on the simulation and provides statistical information (such as mean, variance, or percentile values) to the outer loop for incorporation in the NLLS objective function. Graphically, the flow of information is shown in Figure 6.

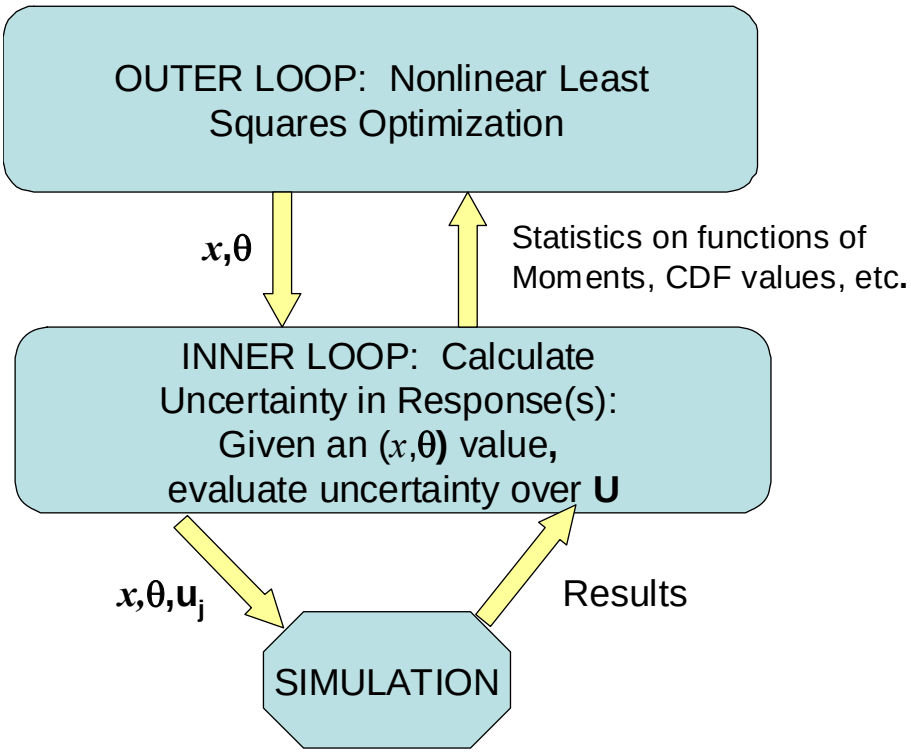

Figure 6. NLLS under Uncertainty: Nested solution approach

There are several ways one may perform UQ to calculate statistical information on the inner loop. These include sampling methods, reliability methods, and stochastic expansion methods, which are outlined in the following sections. A powerful alternative to the strictly nested paradigm is trust region-managed surrogate-based calibration under uncertainty, where surrogate (meta-) models for the inner loop statistics as a function of the calibration parameters θ are iteratively constructed and updated as optimization progress is made in converging to the optimal least squares regressor $\hat{\boldsymbol{\theta}}$. Other methods include sequential or uni-level reliability optimization methods, which more closely couple the optimization and uncertainty quantification by iterating between these and updating the optimization goals based on the most recent probabilistic assessment results. [Eldred and Bichon, 2006].

A. Sampling methods

The most common method of calculating uncertainty with simulations is to specify distributions on the uncertain input values, and then repeatedly sample from those distributions, running the model with the sampled values to build up a distribution of the outputs. This is classical propagation of uncertainty. The sampling techniques can be a variety of Monte Carlo methods, including stratified sampling (Latin hypercube sampling) which spread the samples over the space [Swiler and Wyss, 2004], or quasi-Monte Carlo sampling, which is a way of generating sequences that approximate a uniform distribution [Knuth, 1997; Gentle, 2004]. Typically m samples are taken from distributions on inputs, and each one is run through the simulation model, resulting in m realizations of the outputs.

B. Reliability Methods

Reliability methods are based on probabilistic approaches that compute approximate response function statistics based on specified uncertain variable distributions. These response statistics include response mean, response standard deviation, and cumulative or complementary cumulative distribution functions (CDF/CCDF). These methods are often more efficient at computing statistics in the tails of the response distributions (events with low probability) than sampling-based approaches in which the number of samples required to resolve a low probability can be prohibitive. A full discussion of the theory behind analytic reliability methods will not be presented here; instead see the discussion in Section 6.3 of the DAKOTA User's Manual [Eldred et al.] and in [Haldar and Mahadevan].

The mean value reliability method (also called First-order Second-Moment method) is based on a first-order Taylor series approximation of the response function linearized at the mean values of the random variables. The mean value method calculates the mean and variance of the output. More advanced reliability methods based on most-probable point analysis, answer the fundamental question: Given a set of uncertain input variables, **u**, and a scalar response function, *f*, what is the probability that the response function is below or above a certain level, $\bar{z}$? This probability calculation involves a multi-dimensional integral over an irregularly shaped domain of interest. The reliability methods all involve transforming the user-specified uncertain variables to a space of independent standard Gaussian random variables, possessing a zero mean and unit variance. The region of interest is also mapped to the transformed space. In the transformed space, probability contours are circular and the multi-dimensional integrals can be approximated by simple functions of a single parameter, β, called the reliability index. β is the minimum Euclidean distance from the origin in the transformed space to the response surface, determined by $\bar{z}$. This point is also known as the most probable point (MPP) and is found by solving a nonlinear equality-constrained optimization problem. The distance from the MPP to the origin has the meaning of the number of input standard deviations separating the mean response from a particular response threshold. There are many reliability methods: most of them involve different ways to approximate the limit state, solve for the MPP, or integrate to calculate the probabilities.

Note that although analytic reliability methods may require fewer function evaluations than sampling methods, they do require finite difference or analytic gradients of the output variable with respect to the uncertain variables. Hessian information, if available, tends to help the performance of these methods. Finally, finding each point in a cumulative distribution function requires performing a separate optimization calculation.

C. Stochastic Expansion Methods

Stochastic expansion methods approximately represent random response functions in terms of finite-dimensional series expansions. We focus on two stochastic expansion methods: polynomial chaos expansions (PCE) and stochastic collocation methods (SC). PCE is a type of stochastic response surface method [Ghanem and Red-Horse]. These methods approximate outputs of the uncertain system through series expansions of standard random variables. In PCE, the series expansion is often based on Hermite polynomials which are functions of Gaussian random variables, although other orthogonal polynomials which are functions of other distribution types may be used. Eldred, Webster, and Constantine [2008] provide an overview of the various types of random input distributions that can be modeled and the corresponding orthogonal polynomials, as well as describe solution approaches. The goal of a PCE analysis is to determine the unknown coefficients of the orthogonal polynomials in the series expansion. Usually, these coefficients can be calculated from a limited number of model simulations.

Conceptually, the propagation of input uncertainty through a model using PCE consists of the following steps: (1) input uncertainties are expressed in terms of a set of unit Gaussian random variables, (2) a functional form such as Hermite polynomials is assumed for selected outputs, and (3) the parameters of the functional approximation are determined. An important distinguishing feature of the PCE methodology is that the solution series expansions are expressed as random processes, not merely as statistics as with many other nondeterministic methods. Thus, one obtains a stochastic approximation of the response which can be sampled thousands of times (by sampling the underlying Gaussian random variables) to obtain statistics of interest. Note that PCE can be most efficient when there is additional information from the problem physics which dictates the most appropriate type of approximating polynomials. For this paper, we used 2 variants of stochastic expansion methods: a PCE approach and a stochastic collocation approach. Both methods used 3rd order quadrature to sample the points upon which the expansion is based. In the stochastic collocation approach, Lagrange polynomial interpolants are used instead of orthogonal polynomials. Finally, once the expansion coefficients were determined for PCE or SC, the expansion was sampled 100,000 times to determine the statistics of interest (such as the mean or percentile values).

D. Software Implementation

To implement the examples discussed in the following section, we use the DAKOTA toolkit [DAKOTA – Eldred et al. 06]. DAKOTA (Design Analysis Kit for Optimization and Terascale Applications) is a software tool which provides an integrated suite of optimization methods, data sampling methods, and response surface approximation methods. It is a publicly-available software package that can readily be employed for various optimization studies, parameter estimation, uncertainty analysis, etc. For the results presented in this paper, we employed DAKOTA version 4.1+ on a 64-bit Intel microprocessor-based computer workstation running the Red Hat Enterprise Linux 4 operating system. This version of DAKOTA is available to the public, under the restrictions of the GNU General Public License, from http://www.cs.sandia.gov/DAKOTA.

DAKOTA is designed to explore advanced solution methodologies, such as those proposed here, coupling NLLS algorithms to uncertainty analysis capabilities. DAKOTA contains algorithms for uncertainty quantification with sampling, reliability, and PCE/stochastic finite element methods and for parameter estimation with nonlinear least squares methods. It also contains general gradient and non-gradient based optimization algorithms and parameter study capabilities. The individual algorithms may be linked together within "strategies" such as surrogate-based optimization or optimization under uncertainty. We specifically draw on the following capabilities of DAKOTA:

1. Algorithm "nesting" to specify an outer and inner loop formulation as depicted in Figure 6. This facilitates sending the calibration parameters θ from the outer to inner loop and sending statistical measures from the inner loop back to the outer loop to be incorporated in the outer loop objective function.
2. Three NLLS methods (Gauss-Newton, NLSSOL and NL2SOL) for nonlinear least squares optimization and three uncertainty quantification method classes (sampling, reliability methods, and stochastic expansion methods) for the inner loop calculation of statistical measures of simulation responses.
3. Option to utilize analytic or generate finite difference gradients.
4. Design variable insertion: the capability to calibrate parameters which specify or govern distributions (such as means, variances, and bounds).
5. In some cases, the ability to compute analytic derivatives of the statistics of interest with respect to the calibration parameters, without resorting to additional evaluations of the computational model.

V. Examples and Results

For Examples A, B, and C, we explore a problem from the reliability optimization community [Sues01] to demonstrate the formulations outlined in Section III. The problem involves a cantilever beam, depicted in Figure 7. This problem is often formulated as a design problem where the goal is to minimize the weight (or, equivalently, the cross-sectional area) of the beam subject to displacement and stress reliability constraints. We modify this formulation, since we are interested in obtaining values (estimated parameters) for the width and thickness of the beam (w and t) which result in the simulated displacement "matching" the experimental displacement in a variety of ways (e.g., matching means, matching CDFs, etc.)

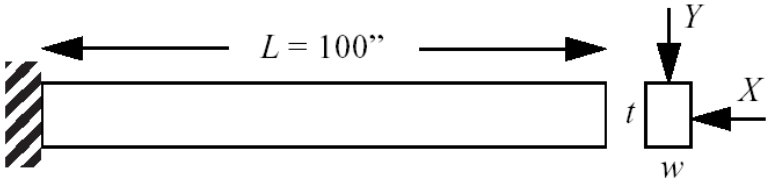

Figure 7. Cantilever Beam Schematic

The variables characterizing the beam problem are shown in Table 1. The random variables $\boldsymbol{u}$ in the problem include the yield stress R of the beam material, the Young's modulus E of the material, and the horizontal and vertical loads, X and Y, which are modeled with normal distributions. Typically R provides an uncertain constraint on the calculated stress, but will not be considered in the modified problem formulation, where we are matching distributions of displacement. The length of the beam is constant: $L=100$ inches. Note that for the common formulation used in reliability optimization, the displacement is subject to a constraint [DAKOTA06], but for the purposes of this study, it is not.

Three response functions are typically calculated as part of the simulation: area, stress, and displacement. Area is simply $w*t$. Stress and displacement (assuming simplifying assumptions such as isotropic material properties of the beam and elastic displacement) are given by:

$$Stress = \frac{600}{wt^2}Y + \frac{600}{w^2 t}X$$

$$Displacement \quad D = \frac{4L^3}{Ewt}\sqrt{\left(\frac{Y}{t^2}\right)^2 + \left(\frac{X}{w^2}\right)^2}$$

Table 1. Variables characterizing the Cantilever Beam Problem

	Variable	Description	Mean or Nominal Value	Standard Deviation
constant	L	length	100 inches	N/A
design	w	width	4 inches	N/A
$\boldsymbol{\theta}$	t	thickness	4 inches	N/A
uncertain	X	horizontal load	500 lbs	100
$\boldsymbol{u}$	Y	load	1000 lbs	100
	R	yield stress	40000 psi	2000
	E	elastic modulus	2.9E7 psi	1.45E6

A. Example of matching means

For this first example, we demonstrate a model calibration where the outer loop nonlinear least squares optimization seeks values of the parameters w and t which result in the mean of the simulation model for the displacement of the beam approximately matching the experimental data sample mean for the displacement of the beam. The inner loop involves sampling or other uncertainty analysis over three uncertain variables: E, X, and Y. For this exercise, we simulated 20 experimental data points by holding the design variables fixed at w* = 2.5 inches and t* = 2.5 inches, and sampling over the uncertain variables. The mean of this particular sample is 3.993 inches. The displacement of the nominal beam given the properties in Table 1 above (with $w = t = 4$ inches) is 0.605 inches. Thus, when using the nominal values as an initial iterate, we expect the calibration to result in smaller values for w and t to match with the greater displacement observed in the simulated experiment.

Table 2 shows the results of matching the experimental and model means, using three NLLS optimization methods and several uncertainty quantification methods. The NL2SOL algorithm consistently performs well. While in a few instances Gauss Newton's performance rivaled that of NL2SOL, in general Gauss Newton and NLSSOL did not perform as well; therefore, the remaining results are presented using NL2SOL.

The calibration results in Table 2 illustrate a typical tension between accuracy and efficiency (computational cost) when performing uncertainty quantification. For example, the mean value reliability method requires significantly fewer function evaluations than sampling. This is to be expected: each inner loop mean value UQ evaluation builds a first-order approximation of response uncertainty, requiring only a single function and gradient evaluation with the calibrated parameter values set from the outer loop, and the uncertain parameters set to their mean values. The efficiency of model calibration under uncertainty also depends on the method for computing required derivatives of statistics with respect to calibration parameters. Numerically differencing across UQ analyses is required when using sampling and is demonstrated for mean value, but can be costly (we employ central finite differences here, with a finite difference relative step size of 1% with respect to each calibration parameter). In contrast, mean value and stochastic expansion methods allow analytic calculation of calibration derivatives, substantially reducing outer loop cost. So, for example, the mean value method with analytic derivatives requires the fewest function evaluations. (For inner loop UQ methods themselves needing derivatives, e.g., mean value and stochastic expansion methods, we use analytic derivatives from the cantilever simulator.)

Regarding accuracy, there is a difference between the inner loop UQ method assessment of the response mean (evidenced by the very small residual errors in the mean estimate, between 0 and 1.0E-6) and the actual response mean obtained from more resolved LHS sampling. The reported LHS-verified mean was calculated by performing a 10,000 sample LHS study with the estimated optimal parameter values. This verification of the mean estimate demonstrates that the mean value method is least accurate for this problem, followed by (under-resolved) sampling, and the stochastic expansion methods, which are most accurate. The mean value method estimated means around 4.02 compared to the experimental mean of 3.993.

Many combinations of w and t will yield a simulation mean comparable to the experimental mean of 3.993. This non-uniqueness or lack of identifiability of parameter values is a common issue in inverse problems, and matching higher-level statistics such as moments may exacerbate the problem. Regularization approaches may be used to address this [Aster et al., 2005], for example, Tikhonov regularization where a term is added to the sum of squared residuals to improve problem conditioning. The term added may penalize deviations from some expected parameter

range (prior knowledge), favor smooth solutions (penalize large derivatives), or be a factor of the response and/or input variable variance. For example, a Cholesky factorization of the inverse covariance matrix of the data is often used as the Tikhonov matrix. In ridge regression, a bias term is added to the response data to reduce the variance of the estimated parameters. We have not used regularization approaches in this paper and identify this as an area of future research (specifically, what type of regularization terms should be used when optimizing statistical measures?)

Table 2. Results for Matching Simulation Mean with Experimental Mean of 3.993

NLLS Method (Outer Loop)	UQ Method (Inner Loop)	$\hat{t}$	$\hat{w}$	Residual Norm Value	LHS-verified Residual Norm	Num Fn Evaluations
NL2SOL	sampling (20)	2.116	2.795	1.03E-10	4.52E-3	840
Gauss Newton	sampling (20)	3.237	2.224	1.00E-6	4.76E-4	1700
NLSSOL	sampling (20)	2.437	2.530	2.22E-15	2.78E-3	2120
NL2SOL	mean value, numerical deriv.	2.110	2.791	7.33E-11	3.48E-2	42
Gauss Newton	mean value, numerical deriv.	3.243	2.220	1.19E-6	1.51E-2	85
NLSSOL	mean value, numerical deriv.	2.428	2.530	3.11E-15	2.77E-2	106
NL2SOL	mean value, analytic deriv.	2.110	2.791	0.0	3.48E-2	20
NL2SOL	PCE	2.117	2.795	4.44E-16	4.76E-6	270
NL2SOL	stochastic collocation	2.117	2.795	4.44E-16	7.19E-7	270

B. Example with means and standard deviations

The next example demonstrates a model calibration where the outer loop nonlinear least squares optimization is choosing the values of the parameters w and t which result in the mean and standard deviation of the simulation model approximately matching the mean and standard deviation of the experimental data for the displacement of the beam. In this example, we included 2 residual terms (the difference between experimental and simulation means, and the difference between experimental and simulation standard deviations) in a least squares objective function with equal weights. If instead for example we wanted to place more priority on matching the means, we could use weighted least squares. Table 3 shows the results. Note that the experimental mean and standard deviation are 3.993 and 0.4142, respectively. We see that matching both mean and standard deviation requires more function evaluations than matching only the mean. NL2SOL coupled with the mean value reliability method appears to be the most efficient for this formulation, but it is the least accurate. The stochastic expansion methods (PCE and stochastic collocation) offer a good balance in terms of efficiency and accuracy.

The last two rows of Table 3 show an alternative approach for the inner loop UQ. We use the concept of reliability indices as explained in Section III. For this particular problem, we took 5 experimental data points corresponding to the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentile values. We then optimized with respect to matching the simulation data so that the simulation reliability indices would approximately correspond to the experimental reliability indices at these percentile values. This formulation implicitly matches the mean and standard deviation. In this particular example, the performance of the calibration method using reliability indices is comparable to using PCE or SC to calculate the first two moments directly in the inner loop.

Table 3. Results for Matching Mean to 3.993 and Standard Deviation to 0.414, using NL2SOL.

UQ Method (Inner Loop)	$\hat{t}$	$\hat{w}$	Residual Norm values (mean, std.dev)	LHS-verified Residual norm values	Num Fn Evaluations
sampling (20)	2.552	2.469	6.13E-03, 5.24E-02	3.86E-3, 5.30E-3	1640
mean value, numerical deriv.	2.020	2.910	1.48E-8, 1.92E-10	3.54E-2, 8.38E-2	49
PCE	2.328	2.602	4.69E-12, 1.15E-12	9.48E-6, 2.95E-4	351
stochastic collocation	2.328	2.602	4.76E-12, 1.16E-12	9.48E-6, 2.88E-4	351
PCE specifying reliability indices	2.328	2.603	5.94E-5, 3.68E-5	5.00E-5, 3.31E-4	324
stochastic collocation specifying reliability indices	2.328	2.603	5.95E-5, 3.68E-5	5.00E-5, 3.25E-4	324

C. Example with CDFs

Next we show an example matching percentiles from a cumulative distribution. For the experimental data, we take the original 20 samples, sort them in ascending order, and determine response levels (displacements) corresponding to the 0.05, 0.25, 0.50, 0.75, and 0.95 percentiles (so that the probability of the displacement of the experimental data P(Data $\leq$ 3.65) = 0.25, etc.). In this example, there are 5 terms in the least squares objective function, corresponding to the five percentiles of the CDF we are trying to match. The experimental response values at these percentiles are: 3.42, 3.65, 3.86, 4.33, and 4.66.

Table 4 shows a wide range of results for this calibration problem. In general, matching percentile values requires significantly more function evaluations than matching the first and second moments, but should offer more power in parameter estimation since the output PDF is better characterized. The various solution approaches obtain different values for the optimal parameters w and t. Some parameter values result in better matches to the middle percentiles and are very inaccurate in the tail (for example, the 20 sample LHS result has a residual difference of 0.24 for the 95th percentile value), while others (mean value and stochastic expansion) have the most error matching the 50th percentile value. There is a tradeoff in matching the various percentiles: some can be matched closely at the expense of others. If the user wanted to force matching at tails, for example, one could weight the residuals to obtain better matches at the percentiles of interest.

UQ performed with 20 sample LHS suffered from problems resolving tail probabilities. Increasing the sample size from 20 to 100 greatly increases the accuracy of the result, but at the cost of significantly more function evaluations due to its nesting within the model calibration. In this problem, we see that the Advanced Mean Value (AMV) reliability method, which performs a single Taylor series approximation in the space of the original uncertain variables ("x-space") centered at the uncertain variable means, and searches for the MPP using this approximation, yields similar results to the AMV+ method, which iteratively updates the Taylor series approximation at each MPP prediction until the MPP converges. In other cases AMV+ may significantly outperform AMV, although in general it will require more function evaluations to converge. Finally, the stochastic expansion methods perform well.

Table 4. Results for Matching Experimental and Simulation Percentile Values from the CDF using NL2SOL: Experiment percentile values of 3.42, 3.65, 3.86, 4.33, and 4.66.

UQ Method (Inner Loop)	$\hat{t}$	$\hat{w}$	Residual Norm values (one row per 0.05, 0.25, 0.5, 0.75, and 0.95 percentile)	LHS-verified residual norm values of percentiles	Num Fn Evaluations
sampling (20)	2.000	2.958	5th: 0.1582 25th: 0.1213 50th: 0.02687 75th: 0.0832 95th: 0.2452 TOTAL: 0.3279	5th: 0.2298 25th:0.0211 50th: 0.0980 75th: 0.0123 95th: 0.1826 TOTAL: 0.3104	1440
sampling (100)	2.444	2.534	5th: 0.0889 25th: 0.0489 50th: 0.0712 75th: 0.1333 95th: 0.0886 TOTAL: 0.2024	5th: 0.0924 25th: 0.0298 50th: 0.0885 75th: 0.1002 95th: 0.0140 TOTAL: 0.1658	6200
MPP x-Taylor mean (AMV)	4.300	2.008	5th: 0.0799 25th: 0.0658 50th: 0.1231 75th: 0.0760 95th: 0.0315 TOTAL: 0.1807	5th: 0.1232 25th: 0.0373 50th: 0.1180 75th: 0.0363 95th: 0.0866 TOTAL: 0.1983	385
MPP x-Taylor mpp (AMV+)	4.305	2.007	5th: 0.0799 25th: 0.0659 50th: 0.1231 75th: 0.0760 95th: 0.0315 TOTAL: 0.1807	5th: 0.1233 25th: 0.0373 50th: 0.1179 75th: 0.0363 95th: 0.0866 TOTAL: 0.1983	550
PCE, specifying percentiles	2.395	2.560	5th: 0.0903 25th: 0.0046 50th: 0.1063 75th: 0.0769 95th: 0.0077 TOTAL: 0.1660	5th: 0.0826 25th: 0.0444 50th: 0.1078 75th: 0.0817 95th: 0.0093 TOTAL: 0.1648	1809
stochastic collocation, specifying percentiles	2.408	2.552	5th: 0.0895 25th: 0.0475 50th: 0.1065 75th: 0.0776 95th: 0.0071 TOTAL: 0.1663	5th: 0.0827 25th: 0.0445 50th: 0.1079 75th: 0.0818 95th: 0.0075 TOTAL: 0.1649	2349

D. Example with design variable insertion

This bi-level example illustrates the case where outer loop design variables parameterize inner loop uncertain variable distributions. For example, one might wish to determine the mean and variance of a normal input variable distribution to minimize the least squares error in the output. This is called *design variable insertion* into the distribution parameters of the uncertainty analysis. We consider a steel column example from reliability-based structural optimization [Kuschel97], which is typically considered with three design variables $\theta = [b,d,h]$ as shown in Table 5.

Table 5. Design Variables for the Steel Column Example

Variable	symbol	design range	reference value	units
mean of flange breadth	*b*	[200, 400]	200.00	mm
mean of flange thickness	*d*	[10, 30]	17.50	mm
mean of height of steel profile	*h*	[100, 500]	100.00	mm

The reference values correspond to the design originally reported to have optimal reliability. The components of the uncertain vector $u = \left[F_s, P_1, P_{2,} P_3, B, D, H, F_0, E\right]$, enter into the response function of interest *f* as follows:

$$f(\boldsymbol{\theta}, \boldsymbol{u}(\boldsymbol{\theta})) = F_s - P\left(\frac{1}{A_s} + \frac{F_0}{M_s} \cdot \frac{E_b}{E_b - P}\right),$$

where $P = (P_1 + P_2 + P_3)$, $A_S = 2BD, M_S = BDH, M_i = 0.5BDH^2$, and $E_b = \frac{\pi^2 EM_i}{s^2}$, and the uncertain parameters have characteristics shown in Table 6, some of which depend on θ (inserted variables):

Table 6. Uncertain Variables for the Steel Column Example

variable	symbol	distribution	mean/standard deviation	unit
yield stress	F_s	lognormal	400/35	MPa
dead weight load	P_1	normal	500000/50000	N
variable load	P_2	Gumbel	600000/90000	N
variable load	P_3	Gumbel	600000/90000	N
flange breadth	*B*	lognormal	*b*/3	mm
flange thickness	*D*	lognormal	*d*/2	mm
height of profile	*H*	lognormal	*h*/5	mm
initial deflection	F_0	normal	30/10	mm
Young's modulus	*E*	Weibull	21000/4200	MPa

While any of the matching conditions could be applied to this example, we focus on the case of matching CDF response levels corresponding to specified probability percentiles 0.05, 0.25, 0.50, 0.75, and 0.95, as in Example C. Simulated experimental data were generated by setting *b*, *d*, and *h*, to the reference values and then evaluating *f* over 40 Latin hypercube samples of the uncertain vector *u*. As before, binning the sample evaluations yields response levels [76.974, 115.02, 169.64, 191.93, 227.36] corresponding to the percentiles of interest. The goal of the least squares estimation is then to find θ such that the stochastic simulation output *f*, given ***u***, matches these CDF data *y*. For this problem we apply NL2SOL with initial iterate θ_0 = [300, 20, 300].

The nested uncertainty analysis is performed using mean value or most-probable point (MPP) analytic reliability methods enhanced by a first-order Taylor series surrogate model constructed either (1) once at the uncertain variables means and then used to find all requested MPPs (Advanced Mean Value method, AMV), or (2) initially at the uncertain variable means and iteratively updated at each MPP prediction until each MPP converges (AMV+) [Eldred2006]. An advantage of MPP methods is that analytic derivatives of the UQ response levels *f* with respect to θ can be computed without performing finite differencing across the nested analysis, resulting in improved efficiency. Second-order enhancements (AMV^2 and AMV^2+) to the MPP searches were explored, but did not appreciably impact the optimization results and are not reported here.

Table 7 shows the results for this example, including verification of the residuals using 10,000 sample LHS analysis. It is not surprising that we cannot always precisely recover the optimal parameters *b* and *d* since the parameters *B* and *D* are somewhat confounded in the response function *f*, and there are only 5 CDF percentiles to match. With mean value, analytic derivatives are not available, so the least squares solver must rely on finite difference derivatives across the entire nested UQ analysis, typically at considerably greater cost. In this case, however, the mean value approximation resolved the probability sufficiently for optimization, with overall lower cost. This highlights the importance of selecting a UQ method appropriate for the optimization problem at hand: with two design optima occurring at the design boundary, the optimizer works well given a reasonable, though

perhaps crude approximation to the trend. For a less well-behaved problem, the more advanced MPP methods might prove more efficient.

Table 7. Results for the Steel Column Example, matching percentiles with design variable insertion

UQ algorithm	*b**	*d**	*h**	solver residual norm	LHS-verified residual norm	function evals
mean value	200.0	17.07	100.0	19.32	20.87	91
u-space AMV	228.7	14.42	220.4	24.07	28.77	145
u-space AMV+	200.0	17.03	100.0	18.63	21.45	399

The panels in Figure 8 show the experimental CDF, the CDF corresponding to the optimization initial iterate, and the optimal CDFs determined by minimization with the NL2SOL algorithm using each of the MPP search methods in u-space. These graphically demonstrate the progress an optimizer makes in calibrating to a CDF. While more costly, the use of iterative model updating in AMV+ helps resolve tail probabilities slightly more accurately (see right figure).

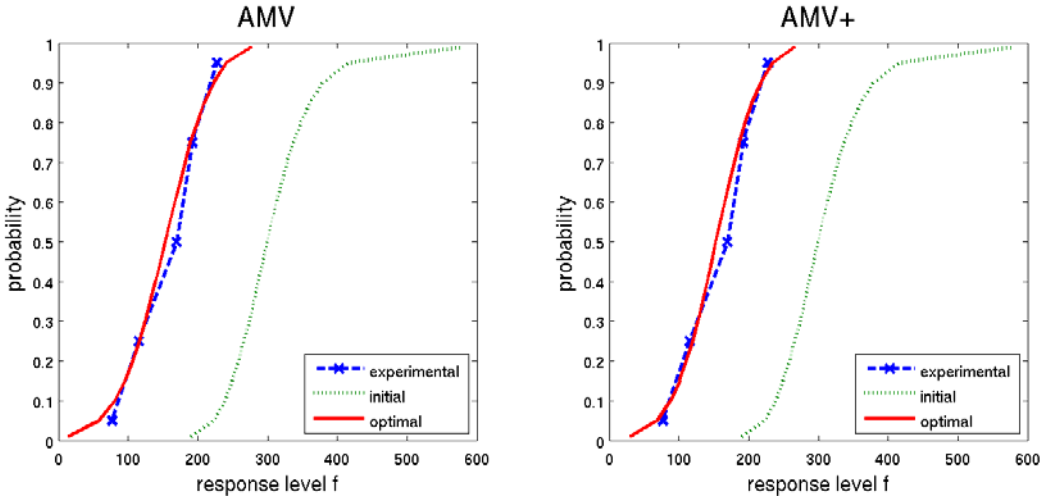

Figure 8. Initial and optimal calibrated simulation CDFs matching the experimental CDF for the Steel Column example.

VI. Conclusions

This paper has presented several formulations to estimate model parameters which result in the "best distribution fit" between experimental and simulation data, including matching the experimental mean with the simulation mean, matching both mean and standard deviation, and matching percentile values. Our results show reliability methods or stochastic expansion methods such as PCE are more effective than sampling in the propagation of uncertainty in the inner loop. This is partly due to the fact that reliability and stochastic expansion methods tend to be more accurate and efficient than sampling methods for uncertainty propagation, and partly due to the availability of analytic derivatives of the statistics for these methods. Specifically, reliability methods support CDF sensitivities and PCE/SC support moment sensitivities. NL2SOL generally was more efficient than NLSSOL or Gauss Newton on the problems in our case studies. The optimal model parameters obtained when only matching moments, especially matching only the experimental and simulation mean, may be non-unique, a common situation in inverse problems. Best distribution fit problems may require regularization approaches to address the lack of identifiability in parameter estimation. Finally, we have shown that finding optimal parameters to match percentiles may result in some percentiles being closely matched and others (such as tail probabilities) not being as well matched. Weighting the residual terms of interest can help address this. In terms of future work, we would like to consider treating the uncertainty as epistemic instead of aleatory, and matching interval bounds. We would also like to investigate methods which are not nested, such as sequential methods to handle optimization of calibration parameters under uncertainty.

Acknowledgments

The authors thank David Gay and Kay Vugrin for their help and discussion of nonlinear least squares optimization methods and confidence interval estimation. The authors thank the Advanced Simulation and Computing (ASC) Program of the National Nuclear Security Administration (NNSA) at Sandia National Laboratories for funding this research.

References

Aster, R. C., Borchers, B., and C. H. Thurber. *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, 2005.

Bates, D. B. and D. G. Watts. *Nonlinear Regression Analysis and its Applications.* John Wiley & Sons, 1988.

Donaldson, J. R. and R. B. Schnabel, "Computational Experience with Confidence Regions and Confidence Intervals for Nonlinear Least Squares", *Technometrics*, Vol. 29, No. 1, pp. 67-82, 1987.

Draper, N. R. and H. Smith, *Applied Regression Analysis*, John Wiley & Sons, NY, NY 1998.

Eldred, M.S. and Bichon, B.J., "Second-Order Reliability Formulations in DAKOTA/UQ," paper AIAA-2006-1828 in *Proceedings of the 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (8th AIAA Non-Deterministic Approaches Conference), Newport, Rhode Island, May 1 - 4, 2006.

Eldred, M.S., Brown, S.L., Adams, B.M., Dunlavy, D.M., Gay, D.M., Swiler, L.P., Giunta, A.A., Hart, W.E., Watson, J.-P., Eddy, J.P., Griffin, J.D., Hough, P.D., Kolda, T.G., Martinez-Canales, M.L. and Williams, P.J., "DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.0 Users Manual," Sandia Technical Report SAND2006-6337, October 2006.

Eldred, M.S., Webster, C.G., and Constantine, P., "Evaluation of Non-Intrusive Approaches for Wiener-Askey Generalized Polynomial Chaos," paper AIAA-2008-1892 in *Proceedings of the 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (10th AIAA Non-Deterministic Approaches Conference), Schaumburg, IL, April 7-10, 2008.

Ghanem, R. and J. R. Red-Horse. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element technique. Physica D 133: 137-144, 1999.

Haldar, A., and S. Mahadevan, *Probability, Reliability, and Statistical Methods in Engineering Design*, John Wiley and Sons, 2000.

Gentle, J.E. *Random Number Generation and Monte Carlo Methods*. Springer Statistics and Computing Series, 2004.

Knuth, D. E. *The Art of Computer Programming, Vol 2: Seminumerical Algorithms*. 3rd edition, 1997.

Oberkampf, W. L. and M. F. Barone, "Measures of Agreement between Computation and Experiment: Validation Metrics," SAND Report 2005-4302, Sandia National Laboratories, 2005.

Rooney, W. C. and L. T. Biegler, "Design for Model Parameter Uncertainty Using Nonlinear Confidence Regions", *Process Systems Engineering*, Vol. 47, pp. 1794—1804, 2001.

Seber, G. A. F. and C. J. Wild, *Nonlinear Regression*, Wiley & Sons, 2003.

Sues, R., M. Aminpour, and Y. Shin. "Reliability-based multidisciplinary optimization for aerospace systems." In *Proc. 42rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference,* No.AIAA-2001-1521, Seattle, WA, April 16-19 2001.

Swiler, L. P. and A. A. Giunta. "Aleatory and Epistemic Uncertainty Quantification for Engineering Applications", Proceedings of the Joint Statistical Meetings, American Statistical Association, August 2007.

Swiler, L.P. and Wyss, G.D., "A User's Guide to Sandia's Latin Hypercube Sampling Software: LHS UNIX Library Standalone Version," Sandia Technical Report SAND2004-2439, July 2004.

Vugrin, K. W, L. P. Swiler, R. M. Roberts, N. Stuckey-Mack, and S.P. Sullivan. "Confidence Region Estimation Techniques for Nonlinear Regression: Three Case Studies." *Water Resources Research*, Vol. 43, WO3423, 2006.