

Model Validation Methodology: From Validation Experiments to Systems Level Application

Richard G. Hills[†] and Ian H. Leslie[‡]
New Mexico State University, Las Cruces, NM

Abstract— Our increased dependence on computer models leads to the natural question – How do we increase the rigor in validating models against experimental data? Models have traditionally been tested against experimental data through simple comparisons such as x-y plots, scatter plots, or contour plots. While such qualitative comparisons are appropriate for model building, the use of such comparisons for model validation naturally leads to the questions – When is the agreement between experimental measurements and model predictions sufficient, and how should we quantify this agreement? Unfortunately, defining rigorous metrics for such comparisons is difficult since there are uncertainties in the validation experiment measurements and in the model parameters. Because of these uncertainties, we expect there to be differences between experimental observations and model predictions, even for perfect models. In addition, when models predict multivariate data (time histories or spatial distributions for example), the differences between the experimental observations and model predictions can be highly correlated. Furthermore, we often measure one quantity from a validation experiment, but desire to predict another quantity for the target application of our model. Finally, complex multi physics models often require a suite of validation experiments to test the model over the range of parameters and physics addressed by the anticipated target application.

Here we present an approach to the development of rigorous model validation methodology that accounts for measurement and model parameter uncertainty. Specifically, we present methodology, based on first order sensitivity analysis, to 1) evaluate whether the validation experiments “cover” the physics of the target application, 2) combine the validation data from a suite of experiments to best represent the decision variables for the target application, 3) evaluate the uncertainty associated with the combined data, and 4) define a validation metric based on this combination of data that accounts for uncertainty in the experimental measurements and the model parameters. We present examples of the methodology based on thermal diffusion.

Index Terms— Prediction methods, Sensitivity, Software verification and validation, Uncertain systems.

I. INTRODUCTION

THIS paper describes some of the methodology developed in [1], [2], [3], and [4] to test model predictions against experimental observations using statistical techniques. The focus of the present work is to further develop the relation between component, or unit level validation experiments, and the system level target application, originally developed in [2]. Specifically, we investigate the relationship between the decision variables that are important to the target application and the measurements obtained from the suite of supporting validation experiments.

A decision variable for the target application is a predicted quantity critical to the success of the target application, i.e., a variable that we use to decide whether the target system was successful at meeting its design goal. Examples include temperature in a temperature sensitive critical component, depth of penetration of a projectile, or the probability of detonation of an explosive device. Note that a decision variable for a target application may or may not correspond to the quantities measured during validation experiments. For example, we may measure the velocity of a foam decomposition front due to thermal loading in a one-dimensional experiment whereas we may really be interested in the thermal protection afforded by the foam on a component in a system. While measurement of front speed is a good indicator of the predictive ability of a model and can be measured non-evasively using radiography, it is not a direct measure of the thermal environment of the component. However, this measure may be very

[†] Associate Dean of Engineering Research and Professor of Mechanical Engineering, Box 30001, Las Cruces, NM 88003, (505-646-3422), email: rhills@nmsu.edu.

[‡] Associate Professor of Mechanical Engineering, email: ileslie@nmsu.edu.

appropriate if the location of the front (or the arrival time of the front at the component) significantly impacts the thermal environment of the component.

Fig. 1 illustrates the basic problem we face. We have some target application for which we wish to test a model. The model for this application is composed of a system of sub-models which each represent some subset of the physics for the target application. For example, the physics for a target application may include conduction heat transfer, thermal contact resistance, convection, radiation heat transfer, and phase changes. We represent the system level target application by the graphic labeled System Level in Fig. 1.

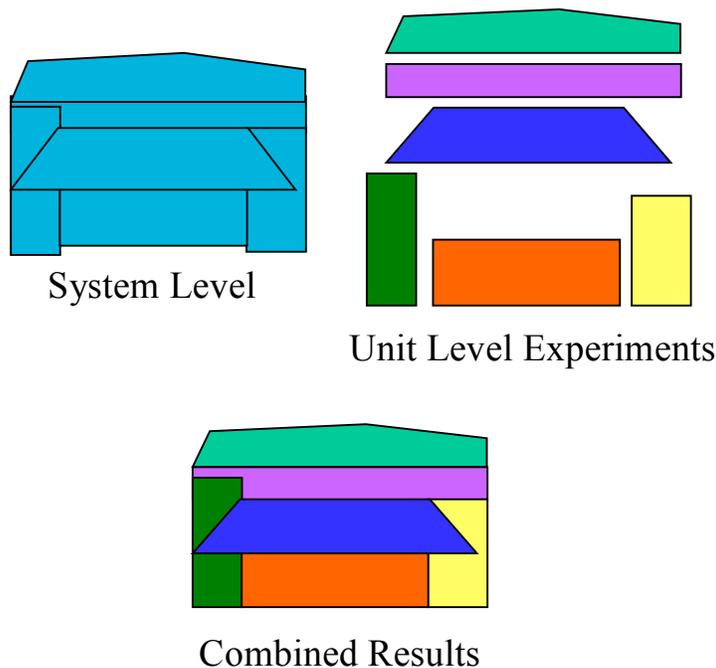


Fig. 1. Unit to System Level Validation

We often cannot perform model validation at the system level due to expense, environmental impact, safety, or simply the inability to run sufficiently controlled experiments at the system level. In these cases, we perform validation experiments on the sub-system level, which can be done under more carefully controlled and monitored conditions. Here we call these sub-system experiments Unit Level (or validation) Experiments. For a unit to system approach to work, we must have some assurance that testing at the unit level fully represents the physics important at the system level. In terms of the Fig. 1, the unit level experiments are represented by the building blocks that combine to represent the target application or system (Combined Results in Fig. 1).

From Fig. 1, we see that the individual building blocks or unit level validation experiments may not be independent. There may be overlap between several of the unit experiments. For example, it is very difficult to design unit level validation experiments for radiation heat transfer or thermal convection that do not also contain heat conduction. Thus, we may have a suite of unit level validation experiments that all contain some form of heat conduction. While the overlap in physics between validation experiments is normal, we do want our unit level experiments to cover all of the physics deemed relevant to our target application. In terms of Fig. 1, we do not want any holes in the graphic labeled Combined Results.

A second issue is to decide how the data from the unit level experiments should be combined to best reflect the validity of the model for our target application. If the performance of the model for the target application is especially sensitive to a particular sub-physics, say thermal radiation, then we expect that we may want the results from the thermal radiation validation experiments to be weighted more heavily in our assessment of the model. We also want to know what level of uncertainty we can tolerate in the validation experiments (both in the measurements and in the model parameters) so that the results are useful for a target application.

Fig. 2 illustrates issues related to uncertainty. We begin with the component of the figure labeled Unit

Experiments. Note that we show probability density for both the uncertainty in the model parameters α used to model the validation experiment or experiments, and the uncertainty in the measurements γ . Once we define a procedure to combine these experiments (actually to combine data from the experiments) to best reflect the target application, we would then like to estimate what the corresponding uncertainty is in the target application, based on our models for the system physics and the uncertainties in the measurements and the parameters. We wish this uncertainty to be less than the acceptable uncertainty for the target application. If, in-fact, the uncertainty is larger for our combined results than is acceptable for the target application, then our validation experiments are not of sufficient to resolve the target application. The system level target application is shown in Fig. 2. We denote the critical target application predicted decision variable by d . This variable may be a scalar or a vector. Note that we have included the uncertainty in the system model parameters and shown the effect of the uncertainty in d (a scalar in this case). We have also shown the uncertainty in the model parameters for the unit experiments and the measurements, and included the effect in the combined results for the decision variable d . We intentionally show the uncertainty in the model parameters for the unit/validation experiments as less than for the application. Generally, we can more carefully control (less uncertainty) these parameters for the model validation experiments than we can for the application. This is desirable because we want the combined results from the unit experiments to have less uncertainty (narrower PDF) than the acceptable level of uncertainty for the eventual application of the systems level model.

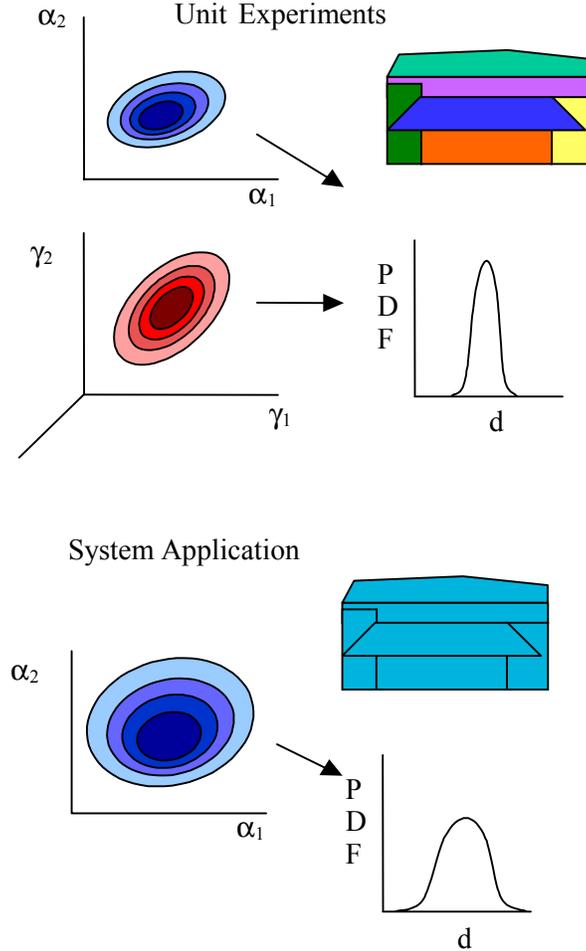


Fig. 2 Uncertainty in Unit to System Level Validation.

II. THEORY

We begin the theoretical development of unit to system validation methodology with a list of assumptions required for the present approach.

1. We assume that it is practical to perform first order sensitivity analysis of the system level application and each of the unit level validation experiments for the important model parameters (i.e., those whose effect on the application is significant).

2. We assume that a representation of first order dependencies between various physical phenomena is adequate at the unit level to represent the target application. If we have zero first order sensitivity, but non-zero second order sensitivity, then the present analysis will not properly account for this dependence.

3. We assume that we have adequate characterization of the uncertainty in the experimental measurements for the validation experiments. Specifically, we assume that we know the form of the distribution, and have estimated values of all of the statistical parameters for the measurements except the statistic describing central tendency of the distribution (i.e., mean, median, mode). The distribution assumptions are required late in the analysis for the development of a validation metric. These assumptions are not required for the first order sensitivity analysis, which addresses coverage of the target application by the unit level experiments and estimates covariance of the reconstructed decision variable.

Assume the following form for the model of the target application:

$$\mathbf{d} = \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a), \quad (1)$$

where \mathbf{d} is a vector of decision variables, \mathbf{G} is a model for the target application, $\boldsymbol{\alpha}$ represents the model parameters that represent the important physics that one wishes to capture in the target application, and $\boldsymbol{\alpha}_a$ represents the model parameters that contain uncertainty. $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_a$ may be vectors of the same parameters, different parameters, or some of each. The lengths of these two vectors need not be the same. Note that components in $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_a$ may represent the same parameter. In this case, we can take $\boldsymbol{\alpha}$ as an expected value and $\boldsymbol{\alpha}_a$ as a perturbation from the expected value due to uncertainty. In this case, the expected value of $\boldsymbol{\alpha}_a$ is zero. $\boldsymbol{\alpha}$ is the important model parameters and/or calibration parameters and/or important model quantities. We can add other arguments/independent variables (such as time or position) to the list of variables in (1) without loss of generality. However, because our sensitivity analysis is only in terms of the variable listed in (1), there is no practical need at this point to do this. The application of this methodology to a time dependent model will be demonstrated in the last example provided here.

We write a similar expression for the unit level validation measurements.

$$\boldsymbol{\gamma} = \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v), \quad (2)$$

where $\boldsymbol{\gamma}$ represents the prediction vector of validation measurements from a suite of experiments, \mathbf{F} a vector of models for the validation experiments, and $\boldsymbol{\alpha}_v$ the uncertain model parameters. We intentionally use different symbols for $\boldsymbol{\alpha}_v$ and $\boldsymbol{\alpha}_a$ since we allow different uncertainties to exist for either one. We emphasize that these measurements can be from suites of experiments that involve different experimental apparatus. Each suite may test the model for different physics. However, we intentionally use the same parameter vector for the $\boldsymbol{\alpha}$ in (2) as we did in (1). The vector of parameters $\boldsymbol{\alpha}$ in (1) represents the vector of all parameters of significance to the target application. If our validation experiments span the physics of the target application, then we should expect that all of the model parameters of importance to the target application should also be important in the suite of validation experiments. In addition, we may expect that individual validation experiments may not be sensitive to the full set of model parameters. In this case, the parameter will appear as a dummy parameter in that particular model for the experiment.

The vector of model parameters is important to the present development since it provides the linkage to relate the validation experiments to the target application. The choice of which parameters to include in this set will require

judgment.

We now perform a first order sensitivity analysis.

$$\Delta\boldsymbol{\gamma} \approx \nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \Delta\mathbf{a} + \nabla_{\boldsymbol{\alpha}_v} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \Delta\boldsymbol{\alpha}_v \quad (3)$$

$$\Delta\mathbf{d} \approx \nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a) \Delta\mathbf{a} + \nabla_{\boldsymbol{\alpha}_a} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a) \Delta\boldsymbol{\alpha}_a \quad (4)$$

Next we take a weighted combination of the predicted measurements to best represent the target application.

$$\begin{aligned} & \mathbf{A}^T \nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \Delta\mathbf{a} + \mathbf{A}^T \nabla_{\boldsymbol{\alpha}_v} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \Delta\boldsymbol{\alpha}_v \\ &= \nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a) \Delta\mathbf{a} + \nabla_{\boldsymbol{\alpha}_a} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a) \Delta\boldsymbol{\alpha}_a \end{aligned} \quad (5)$$

We can use (5) to define the weights for the measurements such that the sensitivity of the decision variable to the important model parameters is the same as the weighted combination of the measurements to the parameters. This effectively combines the validation experiments so that they represent the sensitivities of the target application to these parameters, to first order. For the present analysis, we take $\boldsymbol{\alpha}_v$ and $\boldsymbol{\alpha}_a$ as expected values and $\Delta\boldsymbol{\alpha}_v$ and $\Delta\boldsymbol{\alpha}_a$ as differences from the expected values. Taking the expected value of (5) leads to

$$(\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v))^T \mathbf{A} = (\nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a))^T, \quad (6)$$

since the expected values of each component $\Delta\boldsymbol{\alpha}_v$ and $\Delta\boldsymbol{\alpha}_a$ were defined to be zero. Note that we evaluate the gradient terms at fixed parameter values. For the case that the columns of the gradient on the left hand side of (6) do not span the column space of the right side, the validation experiments do not span the space of the target application decision variables. This suggests that the sensitivity matrices for the validation experiments cannot represent the sensitivities of the target application. For the case that there are as many parameters as unknowns and the rank of the left hand gradient matrix equal to the number of unknowns, a unique \mathbf{A} exists.

$$\mathbf{A} = \left((\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v))^T \right)^{-1} (\nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a))^T \quad (7)$$

The remaining case is when the left matrix does span the column space of the RHS, but we have more measurements than model parameters. In this case, the system is under constrained and we have the opportunity for additional constraints. We choose the \mathbf{A} that satisfies (6) that minimizes the uncertainty in the sensitivity of the reconstructed decision variable to uncertainties in the experimental data, as measured by the variance. Using Lagrange multipliers to minimize the variance subject to the constraint given by (6) gives:

$$\min L = \mathbf{A}^T \text{cov}(\mathbf{F} - \boldsymbol{\gamma}) \mathbf{A} + \boldsymbol{\lambda}^T \left((\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v))^T \mathbf{A} - (\nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a))^T \right)$$

The first term represents the covariance for the reconstructed decision variable and the second represents the constraint. Choosing \mathbf{A} to minimize (8) results in [4]:

$$\mathbf{A} = (\text{cov}(\mathbf{F} - \boldsymbol{\gamma}))^{-1} (\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v)) \left((\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v))^T (\text{cov}(\mathbf{F} - \boldsymbol{\gamma}))^{-1} (\nabla_{\mathbf{a}} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v)) \right)^{-1} (\nabla_{\mathbf{a}} \mathbf{G}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_a))^T, \quad (9)$$

where the covariance in the differences between the model predictions and the experimental observations is approximated by (see [1]):

$$\text{cov}(\mathbf{F} - \boldsymbol{\gamma}) = \text{cov}(\boldsymbol{\gamma}) + \nabla_{\boldsymbol{\alpha}_v} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \text{cov}(\boldsymbol{\alpha}_v) \nabla_{\boldsymbol{\alpha}_v}^T \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \quad (10)$$

Given the \mathbf{A} defined by (9), we can evaluate the uncertainty in the reconstructed decision variable as follows. Take the resulting \mathbf{A}^T times (3), and use the results in (4) and (5) to get:

$$\Delta \mathbf{d} \approx \begin{bmatrix} \mathbf{A}^T & -\mathbf{A}^T \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v) & \nabla_{\alpha_a} \mathbf{G}(\alpha, \alpha_a) \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{\gamma} \\ \Delta \alpha_v \\ \Delta \alpha_a \end{bmatrix} \quad (11)$$

Assume that uncertainties in the experimental measurements are independent of the uncertainties in the validation model parameters, which in turn, are independent of the uncertainties in the target application model parameters. In this case, we can estimate the total covariance in the decision variable for the reconstructed model as follows:

$$\text{cov}(\mathbf{d}) = \mathbf{A}^T \text{cov}(\boldsymbol{\gamma}) \mathbf{A} + \mathbf{A}^T \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v) \text{cov}(\alpha_v) (\mathbf{A}^T \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v))^T + \nabla_{\alpha_a} \mathbf{G}(\alpha, \alpha_a) \text{cov}(\alpha_a) (\nabla_{\alpha_a} \mathbf{G}(\alpha, \alpha_a))^T \quad (12)$$

Equation (12) defines a measure of uncertainty in the reconstructed decision variable vector; given the covariance the model parameters for the validation experiments, the covariance in the model parameters for the target application, and the uncertainty in the validation measurements. Note that the last term in (12) represents the covariance in the target application model prediction of the decision variable due to the uncertainties considered here. Thus, the covariance in the reconstructed variable will be greater than that due to only target application parameter covariance, due the added uncertainties in the validation measurements and in the model parameters for the validation models.

We can now use our linear combination of data to define a validation metric appropriate for this weighting. Consider a set of validation measurements $\boldsymbol{\gamma}$, the corresponding covariance matrix $\text{cov}(\boldsymbol{\gamma})$, the model predictions of the measurements $\mathbf{F}(\alpha, \alpha_v)$, and covariance matrix of the uncertain model parameters $\text{cov}(\alpha_v)$. We are interested in developing a metric for the weighted linear combination of the differences where the weighting is given by \mathbf{A} .

$$\Delta \mathbf{d} = \mathbf{A}^T [\boldsymbol{\gamma} - \mathbf{F}(\alpha, \alpha_v)] \quad (13)$$

The covariance matrix for this linear combination of differences is given by (see (12)):

$$\text{cov}(\mathbf{d}) = \mathbf{A}^T \text{cov}(\boldsymbol{\gamma}) \mathbf{A} + \mathbf{A}^T \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v) \text{cov}(\alpha_v) (\mathbf{A}^T \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v))^T \quad (14)$$

Note that we have not included the uncertainty for the target application model parameters. This is largely a philosophical choice as we are interested in measuring consistency between this weighted combination of the differences between the validation experiment measurements and model predictions without the added uncertainty of the target application. Collecting terms on \mathbf{A} in (14) gives:

$$\text{cov}(\mathbf{d}) = \mathbf{A}^T \left(\text{cov}(\boldsymbol{\gamma}) + \nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v) \text{cov}(\alpha_v) (\nabla_{\alpha_v} \mathbf{F}(\alpha, \alpha_v))^T \right) \mathbf{A} \quad (15)$$

At this point, we ask whether the differences between the weighted combination of measurements and model predictions (i.e., (13)) are significant relevant to the uncertainty in the weighted combination of differences as represented by (15). Before we can define significance, we need to know the functional form for the probability distribution for the uncertainty in $\Delta \mathbf{d}$.

If the measurements and the model predictions are normally distributed, then a linear combination of these differences will be normally distributed with the covariance defined by (15). For the case of normally distributed differences, we use the following statistic (see [2] and [4]):

$$r^2 = \Delta \mathbf{d}^T \text{cov}^+(\mathbf{d}) \Delta \mathbf{d} \quad (16)$$

The + superscript indicates a pseudoinverse. $\text{cov}(\mathbf{d})$ will not be of full rank for the case of more measurements than important model parameters, $\boldsymbol{\alpha}$. In this case, $\text{cov}(\mathbf{d})$ will generally have a rank equal to the number of important parameters for the target application. As such, we take the inverse of $\text{cov}(\mathbf{d})$ in just that subspace spanned by $\text{cov}(\mathbf{d})$, and set the remaining contributions to zero. This is accomplished through the use of the pseudoinverse of $\text{cov}(\mathbf{d})$. Specifically, we use singular value decomposition [5] and remove those directions for which the singular values are zero. Our metric thus only measures differences in the direction of importance to the target application decision variable.

For normally distributed $\Delta\mathbf{d}$, the r^2 statistic is distributed as $\chi^2(n^+)$ distribution with n^+ degrees of freedom where n^+ is the rank of the covariance matrix. To differentiate the case of the rank of the covariance matrix different from its dimension, we use the symbol n^+ to represent degrees of freedom. Given a value for r^2 from our measurements, we can evaluate the cumulative probability (significance) that a set of measurements give an r^2 value larger than observed, given that the model is valid. This is expressed in (17).

$$P(\chi^2(n^+) > r^2) \quad (17)$$

If the significance is small, we question the validity of the model (or the validity of the probability models used to evaluate this statistic). If the measurements are not normally distributed, then the $\chi^2(n^+)$ is not an appropriate choice. Reference [4] presents a Maximum Likelihood approach that is well suited for non-normal distributions and models that are non-linear in parameters.

III. EXAMPLE APPLICATIONS

Here we present two examples based on thermal conduction. Other examples of this methodology are provided in [4].

A. Example 1: Steady Heat Conduction

Consider the one-dimensional thermal heat conduction problem illustrated in Fig. 3. For this example, we assume that the validation experiments and the target application have the same geometry and materials and there are no uncertainties in the geometry or materials. For validation Exp. 1, we assume that only internal temperature measurements are taken and that there is no measure of flux on the surface, nor internal generation.

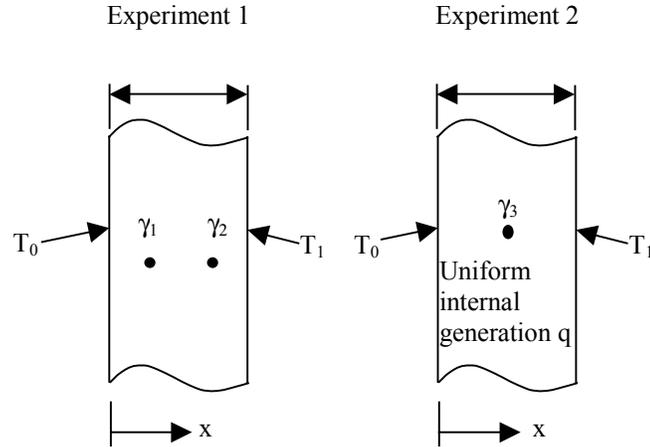


Fig. 3 Example 1: Heat Conduction Validation Experiment

We design validation Exp. 2 to have uniform internal generation with a single temperature measurement made at $x = x_v$. Assume our decision variable for the target application is the flux at the $x = 1$ surface. The target application is shown in Fig. 4. Note that the target application contains uniform internal generation. The important variables for

this application are listed below:

<u>Experiment 1:</u>	<u>Experiment 2:</u>	<u>Application:</u>
T_0, T_1	T_0, T_1, q	T_0, T_1, q

Thus for this example, we have:

$$\boldsymbol{\alpha} = \begin{bmatrix} T_0 \\ T_1 \\ q \end{bmatrix} \quad (18)$$

We see that while our target application and the validation experiments have the same geometries, the important dependent variables (temperatures for the validation experiments and surface flux for the target application) are different. Reference [4] presents an example application of this methodology using shock propagation data where the geometries are different.

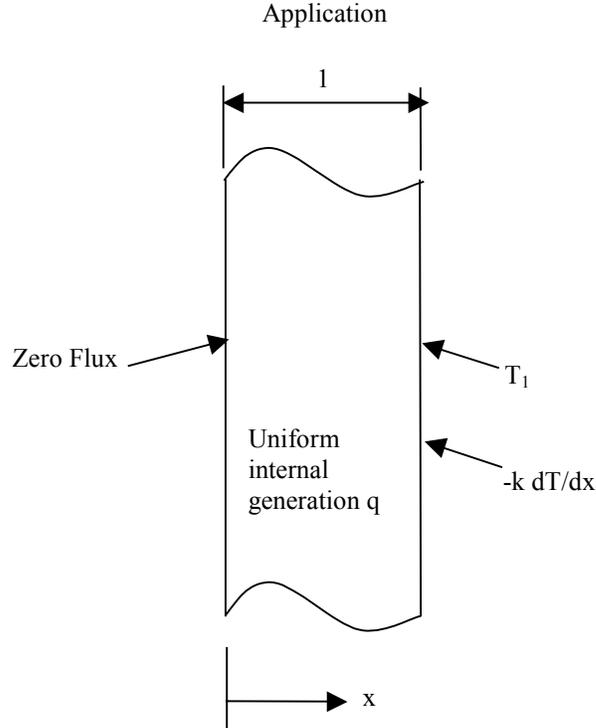


Fig. 4 Example 1: Target Application

We assume that the important model parameters are the boundary conditions at $x = 0$ and $x = 1$, and the internal generation q . In other words, we are assuming that we expect our model to predict thermal response as a result of the boundary conditions and internal generation. We also assume that the covariance matrix for the validation measurements is given by:

$$\text{cov}(\boldsymbol{\gamma}) = \sigma_m^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \sigma_m^2 \mathbf{I} \quad (19)$$

The mathematical models for the two validation experiments are:

Experiment 1:

$$\frac{d^2T}{dx^2} = 0$$

$$T(0) = T_0 = \alpha_1$$

$$T(1) = T_1 = \alpha_2$$

$$\gamma_1 = T(0.25)$$

$$\gamma_2 = T(0.75)$$

Experiment: 2

$$\frac{d^2T}{dx^2} = q = \alpha_3$$

$$T(0) = T_0 = \alpha_1$$

$$T(1) = T_1 = \alpha_2$$

$$\gamma_3 = T(x_v)$$

(20a,b)

(21a,b)

(21c,d)

(22a,b)

(22c)

The corresponding model for our target application is:

$$\frac{d^2T}{dx^2} = q = \alpha_3 \quad (23)$$

$$\frac{dT(0)}{dx} = 0 \quad (24a)$$

$$T(1) = T_1 = \alpha_2 \quad (24b)$$

$$d = -k \frac{dT(1)}{dx} \quad (24c)$$

Since our models are simple, we obtain closed form solutions through simple integration and application of the boundary conditions.

Experiment 1:

$$T = \alpha_1(1-x) + \alpha_2 x;$$

Experiment: 2

$$T = \alpha_1(1-x) + \alpha_2 x + \frac{\alpha_3}{2}(x^2 - x) \quad (26a,b)$$

The model predictions for the measurements are thus given by:

Experiment 1:

$$\gamma_1 = 0.75\alpha_1 + 0.25\alpha_2;$$

$$\gamma_2 = 0.25\alpha_1 + 0.75\alpha_2$$

Experiment: 2

$$\gamma_3 = \alpha_1(1-x_v) + \alpha_2 x_v + \frac{\alpha_3}{2}(x_v^2 - x_v) \quad (27a,b)$$

(28)

The solution for our target application is:

$$T = \alpha_2 + \frac{\alpha_3}{2}(x^2 - 1) \quad (29)$$

The decision variable is found by evaluating flux at $x = 1$ from the expression in (29).

$$d = -k\alpha_3 \quad (30)$$

Taking derivations of (29) and (30) with respect to the important model parameters gives:

$$\begin{array}{ll} \textbf{Experiment:} & \textbf{Application:} \\ \nabla_{\mathbf{a}} \mathbf{F} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.25 & 0.75 & 0 \\ 1-x_v & x_v & (x_v^2-x_v)/2 \end{bmatrix} & \nabla_{\mathbf{a}} \mathbf{G} = [0 \ 0 \ -k] \end{array} \quad (31a,b)$$

We will now consider three cases.

1) Case 1: Experiment 1 Only

In this case we used data from only Experiment 1. Thus we use only the first two rows of (31a).

$$\nabla_{\mathbf{a}} \mathbf{F} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.25 & 0.75 & 0 \end{bmatrix} \quad \nabla_{\mathbf{a}} \mathbf{G} = [0 \ 0 \ -k] \quad (32a,b)$$

With the matrix \mathbf{A} replaced by the vector \mathbf{a} , substitution of (32a) and (32b) into (6) gives:

$$\begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \\ 0 & 0 \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix} \quad (33)$$

Note that while the two columns of the matrix are independent, they cannot represent the right hand side, as there is no linear combination of these columns that can represent the right hand side for a non-zero k . In the context of the present method, the first validation experiment, by itself, cannot represent the target application. This is not surprising since we considered internal generation as an important model parameter for the target application, which is not present in the first validation experiment. Because a solution to (33) does not exist, we cannot represent the sensitivity of the target application to the important model parameters without additional validation experiments.

2) Case 2: Experiment 2 Only

In this case, we use only the last row of (31a) corresponding to Exp. 2.

$$\nabla_{\mathbf{a}} \mathbf{F} = [1-x_v \ x_v \ (x_v^2-x_v)/2]; \quad \nabla_{\mathbf{a}} \mathbf{G} = [0 \ 0 \ -k] \quad (34a,b)$$

so,

$$\begin{bmatrix} 1-x_v \\ x_v \\ \frac{x_v^2-x_v}{2} \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix} \quad (35)$$

Note that there is no x_v that gives zeros in the first two elements of the left hand side. Thus, we cannot find an \mathbf{a} that satisfies (35). In the context of the present method, the second validation experiment by itself, cannot represent the target application. This is a bit more surprising since the second application contains the same physics (conduction and internal generation) as the target application. Apparently, we cannot reconstruct the decision variable, heat flux, without the additional internal measurements. This suggests that we may need to use both validation experiments to

resolve the target application decision variable (or perhaps more internal measurements from the second experiment).

3) *Case 3: Experiments 1 and 2*

In this case, we use all of (31a).

$$\begin{array}{cc} \textbf{Experiment:} & \textbf{Application:} \\ \nabla_{\mathbf{a}}\mathbf{F} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.25 & 0.75 & 0 \\ 1-x_v & x_v & (x_v^2-x_v)/2 \end{bmatrix}; & \nabla_{\mathbf{a}}\mathbf{G} = [0 \ 0 \ -k] \end{array} \quad (36a,b)$$

so,

$$\begin{bmatrix} 0.75 & 0.25 & 1-x_v \\ 0.25 & 0.75 & x_v \\ 0 & 0 & (x_v^2-x_v)/2 \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix} \quad (37)$$

Do the columns of the matrix span the right hand side of (37)? The answer depends on the value for x_v . For example, if we take $x_v = 0$, (37) gives:

$$\begin{bmatrix} 0.75 & 0.25 & 1 \\ 0.25 & 0.75 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix} \quad (38)$$

Clearly, no linear combination of the columns of the coefficient matrix can reproduce the right hand side. If $x_v = 1$, then all the elements in the last row of the coefficient matrix are also zero, leading to the same results. Taking measurements on the boundaries of the second validation experiment does not help us resolve the effect of internal generation in the target application. In contrast, there is a solution to (38) for interior x_v . For example, taking the measurement from the center of the conducting slab, $x_v = 0.5$, results in:

$$\begin{bmatrix} 0.75 & 0.25 & 0.5 \\ 0.25 & 0.75 & 0.5 \\ 0 & 0 & -0.125 \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix}, \quad (39)$$

which does have a unique solution. Thus this set of experiments can represent the sensitivities of the target application to the important model parameters. Solving for \mathbf{a} gives:

$$\mathbf{a} = k \begin{bmatrix} -4 \\ -4 \\ 8 \end{bmatrix} \quad (40)$$

Substitution of (19) and (40) into (15) gives:

$$\sigma_d^2 = \sigma_m^2 \mathbf{a}^T \mathbf{I} \mathbf{a} = 96 \sigma_m^2 k^2 \quad (41)$$

While these measurements can be weighted to represent the target application, we see that our estimate of the corresponding decision variable is very sensitive to the measurement error. While the suite of experiments spans the target application, the experimental design resulted in the representation of the decision variable that is very sensitive to small errors in the measurements.

Can we improve this by taking the internal temperature measurement from Exp. 2 at some other location x_v ? The x_v that minimizes (41), given (37), can be shown to be $x_v = 0.5$. So x_v is already optimum. However, we may be able to reduce the sensitivity to the measurements noise by using additional measurements.

Consider two measurements taken in Exp. 2 at $x = 0.5 - \delta$ and $x = 0.5 + \delta$. Adding the appropriate column to (37) to reflect the additional measurement gives:

$$\begin{bmatrix} 0.75 & 0.25 & 0.5 + \delta & 0.5 - \delta \\ 0.25 & 0.75 & 0.5 - \delta & 0.5 + \delta \\ 0 & 0 & \frac{(0.5 - \delta)^2 + \delta - 0.5}{2} & \frac{(0.5 + \delta)^2 - \delta - 0.5}{2} \end{bmatrix} \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ -k \end{bmatrix} \quad (42)$$

Note that since we have 4 measurements, but are only attempting to represent the sensitivity of the system to 3 parameters, (42) has one free variable. There are an infinity of solutions. We choose the solution that minimizes the target variable variance,

$$\sigma_d^2 = \sigma_m^2 \mathbf{a}^T \mathbf{I} \mathbf{a}, \quad (43)$$

subject to the constraints given by (42) using the Lagrange multiplier approach discussed earlier. For example, if we take $\delta = 0.25$, we find:

$$\mathbf{a} = k \begin{bmatrix} -5.333 \\ -5.333 \\ 5.333 \\ 5.333 \end{bmatrix} \quad (44)$$

Substitution of (44) into (43) gives:

$$\sigma_d^2 = \sigma_m^2 \mathbf{a}^T \mathbf{I} \mathbf{a} = 113.8 \sigma_m^2 k^2 \quad (45)$$

Searching through all possible δ to find the one that minimizes (45) gives $\delta = 0$. This is equivalent to taking two independent measurements at the same location, the center of the slab. This can be accomplished by installing two thermal couples along the centerline. The corresponding weighting of the 4 measurements becomes:

$$\mathbf{a} = k \begin{bmatrix} -4 \\ -4 \\ 4 \\ 4 \end{bmatrix} \quad (46)$$

with

$$\sigma_d^2 = \sigma_m^2 \mathbf{a}^T \mathbf{I} \mathbf{a} = 64 \sigma_m^2 k^2 \quad (47)$$

We see that the use of two measurements in Exp. 2 does improve the ability to resolve the decision variable somewhat, but the weighting of the measurements to represent the decision variable is still very sensitive to the measurements uncertainty.

This example shows the power of the present approach. The approach not only shows how to weight the measurements to resolve the first order sensitivity of the decision variables for the target application to the important model parameters, but also relates uncertainty in the validation measurements to the corresponding uncertainty in the reconstruction of the decision variable. We now present an example of transient heat conduction.

B. Example 2: Transient Heat Conduction

The last example considers transient heat conduction in a slab where we have measurements at multiple locations and times, and where we have different uncertainties for the model parameters of the target application, for the validation experiments, and for the measurements. In this case the geometry of the target application and the validation experiment are the same. We consider most of the thermal parameters to be uncertain for the application, but a subset of the parameters to be uncertain for the validation experiment. The model equations for this example are given below:

<u>Experiment:</u>	<u>Application:</u>	
$\frac{\partial T}{\partial t} = \frac{k}{\rho C_p} \frac{\partial^2 T}{\partial x^2}$	$\frac{\partial T}{\partial t} = \frac{k}{\rho C_p} \frac{\partial^2 T}{\partial x^2}$	(48a,b)

T(x, 0) = 0	T(x, 0) = 0	(49a,b)
-------------	-------------	---------

T(0, t) = T ₀	T(0, t) = T ₀	(49c,d)
--------------------------	--------------------------	---------

T(1, t) = T ₁	T(1, t) = T ₁	(49e,f)
--------------------------	--------------------------	---------

$\gamma_1 = T(0.25, t_j), j = 1, n$	$d = -k \frac{dT(1, t_a)}{dx}$	(50a,b)
-------------------------------------	--------------------------------	---------

$\gamma_2 = T(0.75, t_j), j = 1, n$		(50c)
-------------------------------------	--	-------

The ratio of the thermal conductivity k to the density-specific heat product ρC_p is considered a single parameter and is known as the thermal diffusivity. The important and uncertain model parameters are listed below:

<u>Experiment:</u>	<u>Application:</u>
Important: T ₀ , T ₁ , k/ρC _p	Important: T ₀ , T ₁ , k/ρC _p
Uncertain: k/ρC _p	Uncertain: T ₀ , T ₁ , k/ρC _p

In this case, we define the α 's as follows:

$$\alpha_1 = T_0, \quad \alpha_2 = T_1, \quad \alpha_3 = k/\rho C_p \quad (51)$$

The unit and system level uncertain parameters are:

$$\alpha_{v1} = k, \quad \alpha_{a1} = T_1, \quad \alpha_{a2} = T_2, \quad \alpha_{a3} = k/\rho C_p \quad (52)$$

Due to the simple form of our model, we use a closed form solution [6] to (48) and (49).

$$T(x, t) = T_1(1-x) + T_2x + \sum_{n=1}^{\infty} A_n \exp\left(-\frac{k}{\rho C_p} (n\pi)^2 t\right) \sin(n\pi x) \quad (53)$$

where

$$A_n = -\frac{2}{n\pi} [T_1 - T_2(-1)^n] \quad (54)$$

Our decision variable is given by (see (50b))

$$d(t) = -k \left(T_2 - T_1 + \sum_{n=1}^{\infty} A_n n\pi \exp\left(-\frac{k}{\rho C_p} (n\pi)^2 t\right) \cos(n\pi x) \right) \quad (55)$$

The uncertainties in the model parameters, and the measurements, are defined in Table I.

Parameter	Mean Value	Standard Deviation
Validation Experiment		
$k/\rho C_p$	1.0	0.05
γ		0.25
Application		
T_1	10.0	2.0
T_2	20.0	2.0
$k/\rho C_p$	1.0	0.1

In this example, we ignore the uncertainty in k but only consider the uncertainty in the thermal diffusivity $k/\rho C_p$. Here we take $k = 1$.

Rather than evaluate the sensitivity derivatives by differentiating (53) and (55) directly, we use finite differences to estimate these derivatives. This is simply a convenience, but also is typical of the method used to evaluate more complex models that must be solved numerically. A first order finite differences approximation is used with a step size equal to 1% of the mean value of the parameters shown in Table I.

For purposes of demonstration, we assume that our measurement times do not correspond to the target application decision variable times. The measurement times and locations, and the sensitivity coefficients for these measurements are given in Table II.

Time	x	$\frac{\partial F}{\partial T_1}$	$\frac{\partial F}{\partial T_2}$	$\frac{\partial F}{\partial \left(\frac{k}{\rho C_p}\right)}$
0.10	0.25	0.5761	0.08834	4.711
0.25	0.25	0.7118	0.2118	2.790
0.50	0.25	0.7468	0.2468	0.4677
0.75	0.25	0.7497	0.2497	0.05877
1.00	0.25	0.7410	0.2500	0.006565
0.10	0.75	0.08834	0.5766	5.186
0.25	0.75	0.2118	0.7118	2.793
0.50	0.75	0.2468	0.7468	0.4677
0.75	0.75	0.2497	0.7497	0.05877
1.00	0.75	0.2500	0.7500	0.006565

The finite difference approximations to the sensitivity coefficients for the target application are listed in Table III.

TABLE III
SENSITIVITY COEFFICIENTS FOR TARGET APPLICATION MODEL

Time	$\frac{\partial \mathbf{F}}{\partial T_1}$	$\frac{\partial \mathbf{F}}{\partial T_2}$	$\frac{\partial \mathbf{F}}{\partial \left(\frac{k}{\rho C_p} \right)}$
0.10	0.5761	0.08834	4.711
0.25	0.7118	0.2118	2.790
0.50	0.7468	0.2468	0.4677
0.75	0.7497	0.2497	0.05877
1.00	0.7410	0.2500	0.006565
0.10	0.08834	0.5766	5.186
0.25	0.2118	0.7118	2.793
0.50	0.2468	0.7468	0.4677
0.75	0.2497	0.7497	0.05877
1.00	0.2500	0.7500	0.006565

Given these sensitivities, we use (9) to evaluate the weighting matrix \mathbf{A} . The standard deviation of the reconstructed decision variable can now be evaluated from (12). The standard deviation of the reconstructed decision variable \mathbf{d} , for each measurement time (i.e., each component in \mathbf{d}) is given by the square root of the diagonal elements in the corresponding covariance matrix, $\text{cov}(\mathbf{d})$. We can look at the diagonal elements in each of the matrices that make up $\text{cov}(\mathbf{d})$. These results are given in Table IV. The second column shows the contribution due to measurement uncertainty (first term on RHS of (12)), the third column is the contribution due to parameter uncertainty in the unit level model (second term on the RHS of (12)), the fourth column is the contribution due to uncertainty in the target application model (last term on RHS of (12)), and the last column gives the total uncertainty. The results of Table IV indicate that most of the sensitivity in reconstructing the decision variable originates from uncertainties in the model parameters in the target application. This is not unexpected since the uncertainties in the model parameters for the target application are larger than for the validation experiments, as indicated by the values listed in Table I.

TABLE IV
DISTRIBUTION OF UNCERTAINTY IN DECISION VARIABLE

Time	$\sigma_{\mathbf{d}\text{-meas}}$	$\sigma_{\mathbf{d}\text{-v}}$	$\sigma_{\mathbf{d}\text{-a}}$	$\sigma_{\mathbf{d}}$
0.125	0.949	1.106	4.84	5.06
0.250	0.579	0.621	3.47	3.57
0.375	0.381	0.269	3.10	3.14
0.500	0.329	0.104	3.02	3.04
0.625	0.320	0.038	3.01	3.02
0.750	0.318	0.013	3.00	3.02
0.875	0.318	0.004	3.00	3.02
1.000	0.318	0.001	3.00	3.02
10.00	0.318	0.000	3.00	3.02

C. Validation

For the previous analysis, we did not need to assume a form for the probability density functions (other than the means and variances exist). However, the statistical inference for model validity will require that we make additional assumptions concerning the underlying distributions. Here we assume that the uncertainties for these parameters can all be modeled as independent normal distributions. An example on non-normal distributions are provided in [4].

Note from Table I that we show more uncertainty in the thermal conductivity for the target application than we do for the unit level validation experiments. We expect our validation experiments to be better controlled than our target applications.

To generate a set of simulated measurement data, we use the following procedure:

1. Randomly generate a set of validation model parameters using the probability distribution for the parameters. For this case, we have one model parameter, the ratio $k/\rho C_p$, with the statistics given in Table I. We randomly generate this parameter because there is uncertainty as to the true value of this parameter for the validation experiment.
2. Given this realization of the parameter, we use the unit level model to generate a set of predicted measurements.
3. We then add random noise to the predicted measurements to represent the measurement noise, assuming a normal distribution for each of the measurements with the statistics given in Table I.

This procedure generates a set of measurements that we might obtain if the model were valid. To make the analysis more interesting, we repeated the above procedure multiple times until a low probability set of measurements was obtained. We did this to demonstrate the methodology when the measurements were near the region of non-acceptance for a model. The resulting measurements are shown in the last column of Table V. We also show the model predictions (using the mean value of the parameter and no measurement noise) in the table for the two measurement locations and the 5 measurement times.

TABLE V
TEMPERATURE PREDICTIONS AND MEASUREMENTS

Time	x	T_{pred}	T_{meas}
0.10	0.25	0.5761	0.08834
0.25	0.25	0.7118	0.2118
0.50	0.25	0.7468	0.2468
0.75	0.25	0.7497	0.2497
1.00	0.25	0.7410	0.2500
0.10	0.75	0.08834	0.5766
0.25	0.75	0.2118	0.7118
0.50	0.75	0.2468	0.7468
0.75	0.75	0.2497	0.7497
1.00	0.75	0.2500	0.7500

We can now apply the validation metric given by (16) to obtain:

$$r^2 = (\mathbf{F}(\boldsymbol{\alpha}, \langle \boldsymbol{\alpha}_v \rangle) - \boldsymbol{\gamma})^T \text{cov}^+ (\mathbf{F} - \boldsymbol{\gamma}) (\mathbf{F}(\boldsymbol{\alpha}, \langle \boldsymbol{\alpha}_v \rangle) - \boldsymbol{\gamma}), \quad (56)$$

where

$$\text{cov}(\mathbf{F} - \boldsymbol{\gamma}) = \text{cov}(\boldsymbol{\gamma}) + \nabla_{\boldsymbol{\alpha}_v} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v) \text{cov}(\boldsymbol{\alpha}_v) (\nabla_{\boldsymbol{\alpha}_v} \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_v))^T. \quad (57)$$

From the data provided in the tables, we find:

$$r^2 = 9.38 \quad (58)$$

The covariance matrix has a rank of 3 since we have only 3 independent parameters. The significance of obtaining

this value of r^2 with 3 degrees of freedom is (see [2], [4]):

$$r^2 = 9.38; P(\chi^2(3) > r^2) = 0.025 \quad (59)$$

Assuming that this model is valid, the probability that this model would give the above r^2 or larger is only 2.5%. This result suggests that we have good evidence to reject the model as valid. In other words, we would expect a valid model to produce this large or larger value of r^2 only 2.5% of the time. References [2], [3], and [4] provide additional examples of applications of similar metrics to data for real, but more complex applications.

IV. DISCUSSION AND CONCLUSIONS

A method has been presented to weight the validation measurements to better reflect a target application of a model. This method is based on a first order sensitivity analysis of the models for the validation experiments and for the target application. The method weights the measurements so that they reflect the sensitivity of the target application to the important model parameters. We call this the representation method since we use the validation experimental measurements to represent the target application decision variable. The representative approach has the effect of not only throwing out the directions in the validation space that are not important to the target application, but also weighting the remaining directions based on their importance. The implementation of the method into a validation metric, as presented here, does require that the validation experiments cover the target application. More specifically, the validation experiments collectively must be sensitive to the same parameters as is the target application. For example, if transient heat conduction is important to a target application decision variable, then the target application decision variable will be dependent on the thermal diffusivity and this parameter should be chosen as an important model parameter. If the validation experiments are not dependent on thermal diffusivity, then the validation experiments do not reflect the target application and weights cannot be developed.

The representative method also allows us to evaluate the sensitivity of the reconstructed target application decision variable to the validation measurements. If the decision variable is not dependent on the same model parameters as the validation experiments or suite of experiments, then the target application is not resolved by the validation experiments and additional experiments must be developed. In contrast, if the decision variables are overly sensitive to small differences in the measurements, then the decision variable will be overly sensitive to noise in the measurements. In this case, we must make sure that the experiments are designed so that uncertainty is sufficiently small.

In the case of more validation measurements than important model parameters, we have the luxury of weighting these measurements so that the decision variable is not as sensitive to measurement noise, while still representing the target application. In doing so, we still must require that the validation experiments provide adequate coverage of the target application. For example, if we have several temperature measurements at the same distance from a boundary in a one-dimensional heat-conducting slab, then the present methodology will take an average of these multiple measurements. This averaging reduces the effect of noise in the measurements.

A critical requirement for the implementation of this method is that it requires that we identify the model parameters that are important to the target application. This provides the link between the validation experiments and the target application. Failure to do so can result in the representation of our target application being incomplete. Methods of this type cannot replace engineering judgment. However, these methods can more rigorously quantify the results of this judgment, which will greatly facilitate consensus building concerning the validity of the model.

Finally, we would like to emphasize that our approach to model validation is different from the more common one which: 1) directly compares experimental data to model predictions, 2) develops a model for uncertainty based on the differences between the predictions and the observations directly, and then 3) evaluates the statistical significance of these differences relative to this model of uncertainty. In this more common approach, the sources of the uncertainty in the differences are not as important as the ability to characterize the uncertainty directly from the data. Our approach is different in that we try to understand and model the sources of uncertainty, and then propagate these sources through the model. While this approach is more difficult (because one must understand more about the

system), we suggest that this approach is more suitable for cases where we plan to use the model in a predictive sense. For example, we may know that there are manufacturing uncertainties in the thermal properties of a material. Thus we can develop models for the uncertainty in these properties through multiple samples from multiple manufactures. We can then propagate this model for uncertainty through the predictive model, using techniques such as sensitivity analysis or Monte Carlo analysis, to provide a model for prediction uncertainty due to the “modeled” parameter uncertainty. If we change the geometry, but not the material (and sources of materials) of an item being tested, we can revise the model for the new geometry, using our existing model for parameter uncertainty. If we have properly accounted for all of the significant sources of uncertainty, and if our model is valid, then we should have some confidence that the model predictions, and the corresponding estimate of uncertainty in the predictions, are reasonable. In contrast, if we simply look at the statistics of the differences between experimental observation and model prediction for the original geometry, then we should not expect that the model for uncertainty would be valid for a different geometry.

ACKNOWLEDGMENT

We wish to thank Tim Trucano and Kevin Dowding of Sandia National Laboratories, for their insight and guidance during the development of this model validation methodology.

REFERENCES

- [1] R. G. Hills and T. G. Trucano, “Statistical validation of engineering and scientific models: Background,” Sandia National Laboratories, SAND99-1256, 1999.
- [2] R. G. Hills and T. G. Trucano, “Statistical validation of engineering and scientific models with application to CTH,” Sandia National Laboratories, SAND2001-0312, 2001.
- [3] R. G. Hills and T. G. Trucano, “Statistical validation of engineering and scientific models: A maximum likelihood based metric,” Sandia National Laboratories, SAND2001-1783, 2002a.
- [4] R. G. Hills, I. H. Leslie, and T. G. Trucano, “Statistical validation of engineering and scientific models: Validation experiments to Application,” Sandia National Laboratories, 2002b, submitted for SAND publication.
- [5] G. Strang, *Linear Algebra and Its Applications*. New York: Academic Press, 1976, pp. 135.
- [6] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*. Oxford University Press, 1978, ch. 3.