Using COMSOL to Estimate the Heat Losses of Composite Panels Undergoing Repairs Using Bayesian Inference

A. F. Emery and K. C. Johnson
University of Washington, Seattle, WA 98195,
Corresponding Author: emery@u.washington.edu

Abstract: The usual approach for estimating parameters of mathematical models of engineering systems is based upon the least squares method (LS) in which a system response is compared to a model and the sum of the squared differences between the model and the measured response is minimized. The method as usually employed is based upon the model sensitivities, S, often determined by finite differencing. Using the PDE features of COMSOL allows us to solve directly for S. The results are maximum *a posterior* probable values along with estimates of confidence intervals. For other than Gaussian distributed noise and when more detailed statistical information is desired, recourse must be made to Bayesian inference. Unfortunately this approach is computationally expensive and is realistic when reasonably large numbers of parameters are to be found only if the complex responses can be quickly sampled. One solution is the use of sparse grids. Such sparse grids require the prior specification of a realistic parameter space for sampling. This is best based upon an initial guess of the parameter values obtained from the least squares approach.

Keywords: Parameter estimation, Sparse Grids, Bayesian Inference, Least Squares

1. Introduction

The standard approach to estimating parameters is the Least Squares analysis in which the estimated parameters are those that minimize the norm of the residuals defined as the difference between the predictions of the model and the data. If desired, the residuals can be weighted but usually the weights are taken to be unity. The least squares approach can be shown to be to same as the maximum likelihood method based upon the fundamental assumptions that the model is correct and that any deviation of the data from the model is due to normally distributed zero mean errors. In the likelihood approach, the weights, if used, are based upon the statistical properties of the errors. For non linear problems, the solution is an iterative one based upon sensitivities computed at each estimate of the parameters. Letting the model be represented by $M(\Theta)$ where Θ represents a vector of the parameters sought, the estimated parameters, $\hat{\Theta}$, are those that minimize the functional $L(\hat{\Theta})$ weighted by Σ^{-1}

$$L(\hat{\Theta}) = r^T(\hat{\Theta})\Sigma^{-1}r(\hat{\Theta}) \qquad (1a)$$

where the residuals r are defined by

$$r(\hat{\Theta}) \equiv D - M(\hat{\Theta}) \qquad (1b)$$

Linearizing Eq. 1b about an estimate $\hat{\Theta}_i$ gives

$$r(\hat{\Theta}_i) = D - M(\hat{\Theta}) - \frac{dM}{d\Theta}\Big|_{\hat{\Theta}_i}(\hat{\Theta}_i - \Theta) \qquad (1c)$$

and $\hat{\Theta}_{i+1}$ is given by

$$\hat{\Theta}_{i+1} - \hat{\Theta}_i = (A_i^T\Sigma^{-1}A_i + \beta I)^{-1} A_i^T\Sigma^{-1}[D - M(\hat{\Theta}_i)] \qquad (2a)$$

where $A_i = \partial M/\partial\Theta|\hat{\Theta}_i$. For N measurements and d parameters, Θ is a $[d\ x\ 1]$ column vector, A is a $[N\ x\ d]$ matrix and β, the Levenberg-Marquardt parameter [1], is used for ill conditioned problems. Upon convergence of the iterations, the estimate $\hat{\Theta}$ satisfies

$$E[\hat{\Theta}] = \Theta \qquad (2b)$$

$$cov[\hat{\Theta}] = (A^T\Sigma^{-1}A)^{-1} \qquad (2c)$$

A serious question is how to compute $\partial M/\partial\Theta$. The simplest method is to use finite differencing, i.e., $\partial M/\partial\Theta = (M(\Theta + \delta\Theta) - M(\Theta - \delta\Theta))/2\delta\Theta$. Questions arise about the size of $\delta\Theta$ and the accuracy of the 1st order differencing. A better way is to solve the sensitivity equations directly. Consider a one dimensional transient thermal problem for which the temperature, T, and the sensitivity to the convective heat transfer coefficient, $u = \partial T/\partial h$, satisfy

$$\rho c\frac{\partial T}{\partial t} = k\nabla^2(T) \qquad (3a)$$

$$-k\frac{\partial T}{\partial x} = h(T - T_\infty) \text{ on S} \qquad (3b)$$

$$\rho c\frac{\partial u}{\partial t} = k\nabla^2(u) \qquad (3c)$$

$$-k\frac{\partial u}{\partial x} = (T - T_\infty) + hu \text{ on S} \qquad (3d)$$

where S is the boundary subject to convective heat transfer. By using the PDE solver in COMSOL it is possible to solve directly for u. There is no significant difference in computing time when solving for u compared to finite differencing, but the accuracy with which T and u are computed are comparable.

2. The Problem

A composite panel, Figure 1, was heated with an electric blanket applied to its upper surface and cooled by free convection from the bottom surface and by forced convection from the stringer. Figure 2 is a cross section of the panel. The goal was to estimate the heat losses from the upper and lower surfaces from temperatures measured with thermocouples attached to the surface and located as shown and from infrared thermograms of the upper surface.

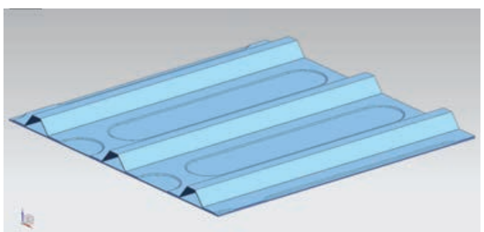


Figure 1. The Panel (upside down) (the distance along the stringer is x, perpendicular to the stringer is y, and vertical is z)

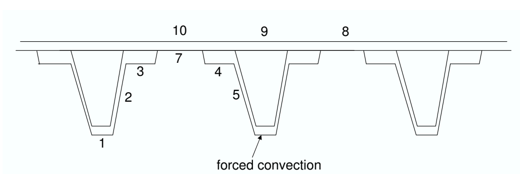


Figure 2: Schematic of the panel (along the y axis at the mid x value) (showing typical locations of thermocouples placed at the numbered points)

The panel was initially at room temperature and transiently heated by a constant current applied to the heating blanket. The heat losses, characterized by the parameter vector $\Theta = \{h_t, h_u, h_s, h_b\}$ where the h_i represent the convective heat transfer coefficients for the panel top and underside surfaces, and stringer side and bottom surfaces respectively, were estimated from the temperatures measured during heating as shown in Figure 3 plus the temperatures on the surface of the electric blanket measured with infrared thermograms.

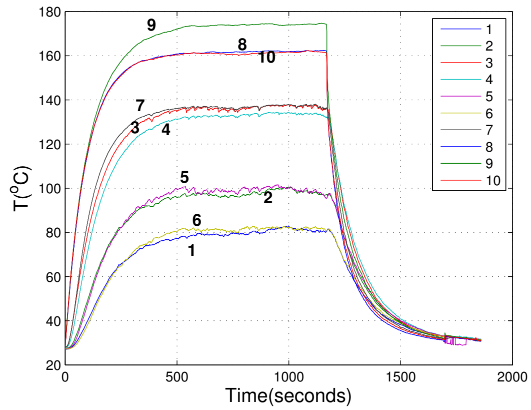


Figure 3: Typical Time History of Panel Temperatures (numbers denote the thermocouples shown on Figure 2)

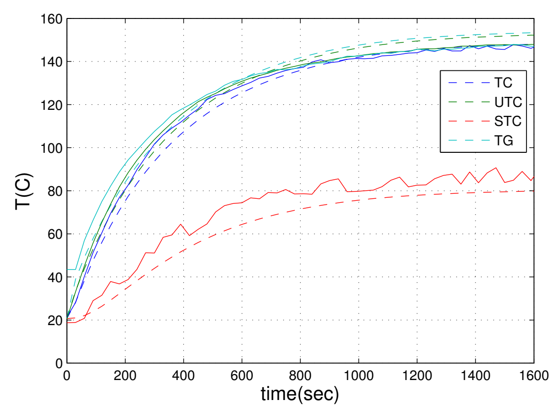


Figure 4: Comparison of Model (dashed) and Data (solid) (TC=panel top, UTC=under panel, STC=stringer, TG thermogram)

The panel temperatures needed in Eq. 2 for the model $M(\Theta)$ were computed using Comsol 3.5 [2]. Figure 4 shows the agreement. Unfortunately, the agreement is not particularly good for several reasons: 1) it is difficult to separate the effects. Heat lost from the region of the panel that does not have a stringer under it is proportional to $h_t + h_u$ and several different combinations of h_t and h_u that have the same sum will produce comparable results; 2) Because several of the sensitivities, Figure 5a, are comparable, the matrix $A^T\Sigma^{-1}A$ is very ill conditioned. As a consequence, a map of the norm of the residuals, Figure 5b, displays many local minima and a very flat surface. Consequently the least squares algorithm leads to multiple solutions; 3) because of the flatness of the surface, small differences in computing the sensitivities will lead to dramatic movement through the 4 dimensional space of parameters; 4) the physical properties, density, specific heat, and conductivity, are temperature dependent and poorly quantified.

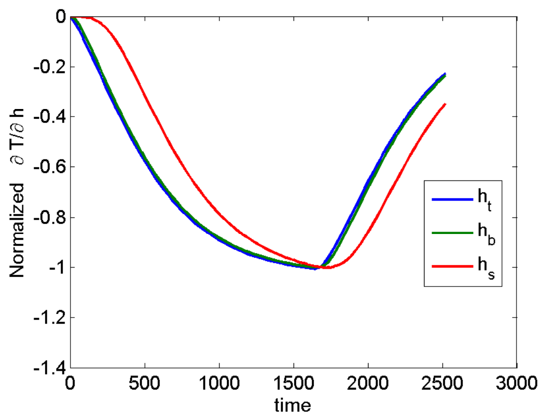


Figure 5a: Sensitivities

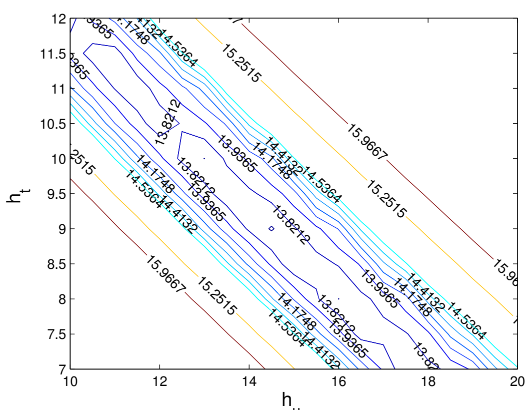


Figure 5b: Contours of L showing local minima and flatness

3. Computing Sensitivities

With such a flat surface, small errors in computing the sensitivities, $\partial T/\partial h$, gave rise to very erratic progress during the iterative solution of Eq. 2. Figures 6 compare the sensitivities to h_b computed for several different thermocouple positions (136 is at the bottom of the stringer and 65 is on the upper surface centered above the stringer) . Figure 6a is the solution for $\partial T/\partial h_b$ computed from the sensitivity equation using the PDE solver, Figure 6b is based upon finite differencing. Note that while the sensitivities at thermocouple 136 are comparable, those for the other locations based on finite differencing are dramatically wrong.

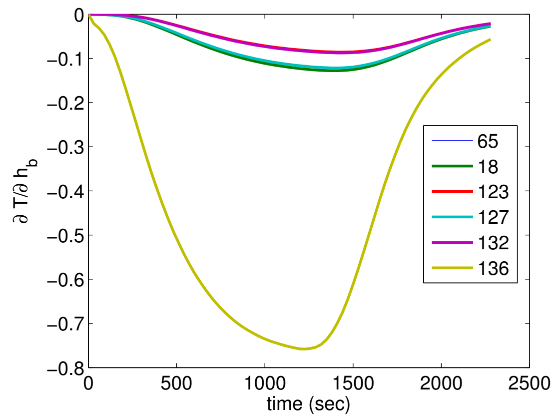


Figure 6a: Time History of $\partial T/\partial h_b$

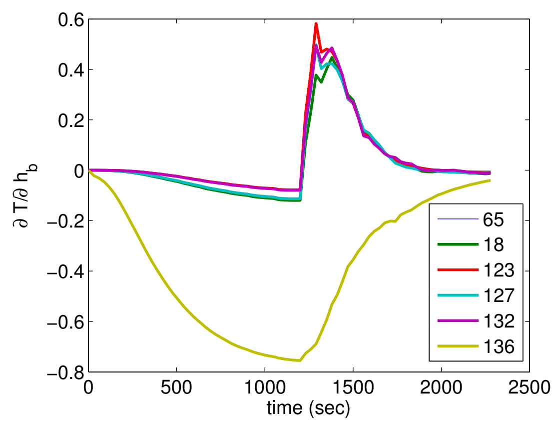


Figure 6b: Time History of $\partial T/\partial h_b$ based on Finite Differences

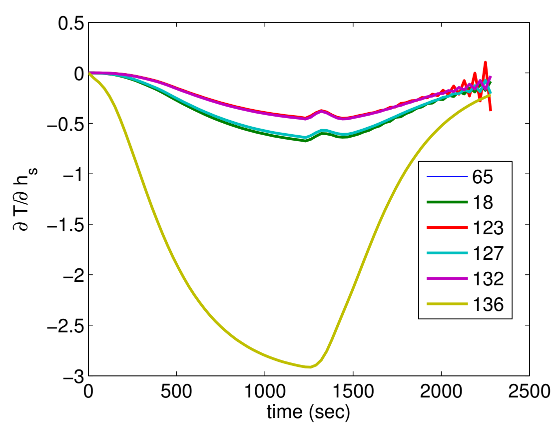


Figure 7: Time History of $\partial T/\partial h_s$ based on Finite Differences using *strict*

If the time integration is set to *strict* then there is negligible difference between the two methods. However, using the strict option can lead to unusual results. Figure 7 shows the results for computing $\partial T/\partial h_s$ using the strict option and we see unexpected results near the time that the heating is turned off and at the end of the experiment. It is not clear why the finite difference result for the sensitivities to h_s and h_b differ.

4. Using Bayesian Inference

An approach based upon hierarchical Bayesian inference [3] allows us to incorporate knowledge about the reasonable values of the parameters sought and information about any other parameters involved in the model. If the model response can be computed by interpolating a response surface, then this approach becomes feasible. Sparse grids can provide a reasonable method for constructing and interpolating the response surface.

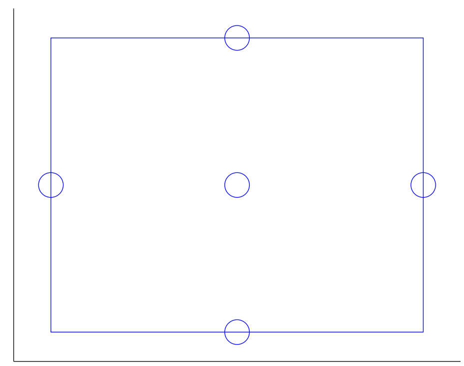
The Bayesian approach consists of determining the joint probability distribution from Bayes relation. Consider the case of two parameters to be estimated, Θ_1, Θ_2, from a set of data D that is presumed to be contaminated with a normally distributed error of standard deviation σ. The joint pdf is given by Eq. 4a and the marginal pdf for Θ_1 is obtained by integrating over all other parameters under the assumption of know prior distributions for all of the parameters.

$$p(\Theta_1, \Theta_2|D) \propto p(D|\Theta_1, \Theta_2)p(\Theta_1, \Theta_2) \quad (4a)$$

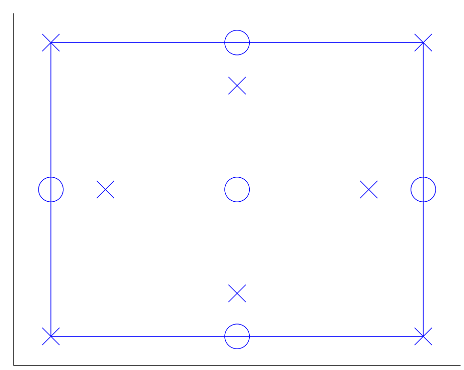
$$p(\Theta_1|D) = \int p(\Theta_1, \Theta_2|D)d\Theta_2 \quad (4b)$$

For simple models, the evaluation of the right hand side of Eq. 4b for the integration usually does not entail substantial computational costs. However, for complex models the computational cost is often too high to be acceptable. Using 7 Gaussian integration points requires evaluating the model at 7^4=2401 points. On the other hand, a sparse grid of level 4 for 4 parameters requires only 441 sample points.

Smolyak [4] developed a method that yields a high level of quality when interpolating and integrating functions. It consists of combining different orders of one dimensional interpolation on nested spaces, i.e., sets of sample points for which the points of each set are intermediate to the points of the preceding set. The interpolating function is a set of linear combinations of products of the univariate polynomials. Each one dimensional interpolation is exact on certain nested spaces. Consider three sample points in each of the x and y directions, Figure 8a. This will allow us to represent the univariate functions, $f(x) = [1, x, x^2]$, $f(y) = [1, y, y^2]$. Smolyak's method represents $f(x, y)$ by products of these univariate representations, i.e., $f(x, y) = f(x)f(y)$. In this case we obtain
$f(x, y) = [1, x, y, x^2, xy, y^2, x^2y, xy^2, x^2y^2]$.



8a) Level 1 Points (O)



8b) Level 2 Additional Points (X)
Figure 8. Level 1 and Level 2 Sparse Grid Points

Obviously with the 5 sample points the product cannot represent all 9 terms and some of the terms are eliminated, giving $f(x, y) = [1, x, y, x^2, y^2]$, that is we filter the product solution and the terms xy, x^2y, y^2x are not represented.

A sparse grid can be characterized by its 'level', essentially the order of the complete polynomial that can be interpolated. When using nested grids, each level contains the contributions of the lower levels and each level lacks some terms needed to complete the polynomial. Level 2 is formed by adding eight (8) additional points as shown in Figure 8b. The univariate function representations would be $f(x) = [1, x, x^2, x^3, x^4]$, $f(y) = [1, y, y^2, y^3, y^4]$, giving a product involving 25 terms. And again the 13 sample points would result in a representation that lacks the terms x^3y, xy^3 to make a complete 4th order polynomial.

A sparse grid was defined for the 4 different values of h, with each ranging from 0.1 $h(LS)$ to 3 $h(LS)$ where $h(LS)$ was the estimate obtained from the least squares method, Eq. 3. Response surfaces were obtained for the temperature, T, and each of the sensitivities, $\partial T/\partial h_i$. The accuracy of the sparse grid interpolation was judged by comparing the interpolated values with the exact ones over a tensor grid. Figure 9 depicts the error at a time of 1100 seconds, just before the heating was turned off for a sparse grid of level 4 (441 sparse grid points).

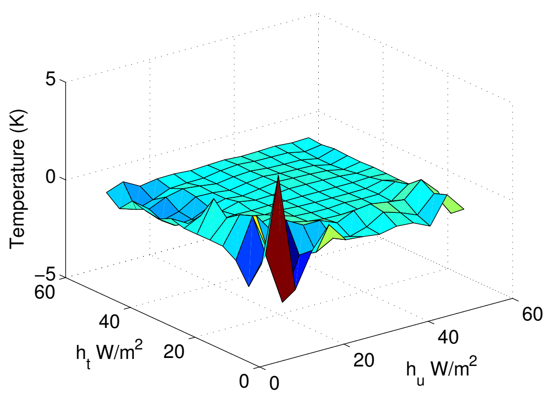


Figure 9a. Error in Interpolating T
Using the Clenshaw-Curtis grid, Level 4

The large error in the vicinity of $h_t = h_u = 0$ is due rapidly increasing temperature as shown in Figure 9b. Figure 9c displays the error in the sensitivity $\partial T/\partial h_u$, showing comparable behavior.

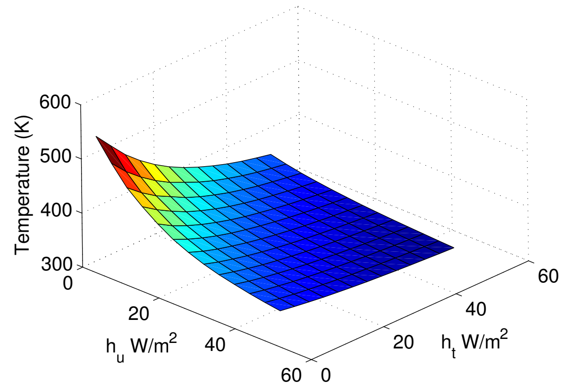


Figure 9b. Temperature Response Surface

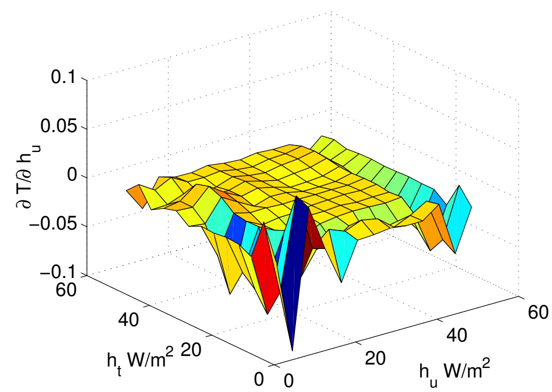


Figure 9c. Error in $\partial T/\partial h_u$
Using the Clenshaw-Curtis grid, Level 4

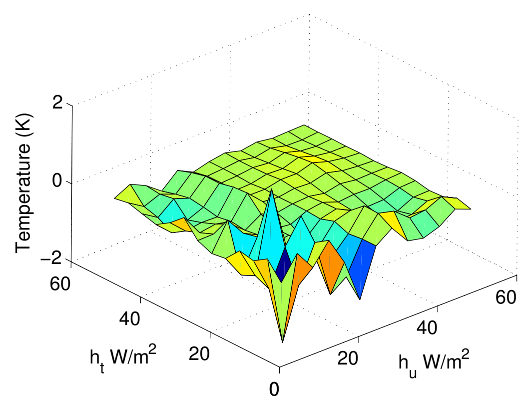


Figure 10a. Error in Interpolating T
Using the Chebyshev grid, Level 4

An additional option when using sparse grids is the choice of the grid. The most common are those based on Chebyshev polynomials, the Clenshaw-Curtis grid, and the Patterson grid. Figure 10 shows the error associated with the Chebyshev grid and in general the errors are about 1/2 of those of the Clenshaw-Curtis grid. However, the Chebyshev grid shows a tendency for ripple like errors as one would expect since it is based on sinusoidal min-max interpolants.

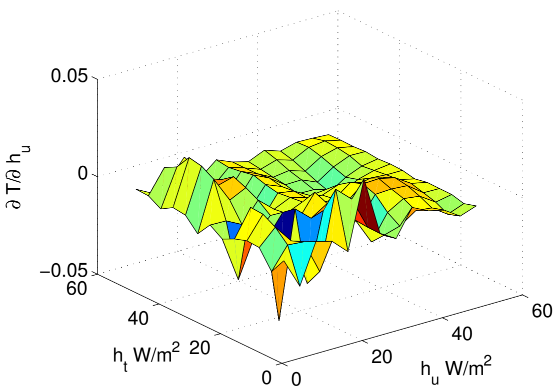


Figure 10b. Error in $\partial T/\partial h_u$
Using the Chebyshev grid, Level 4

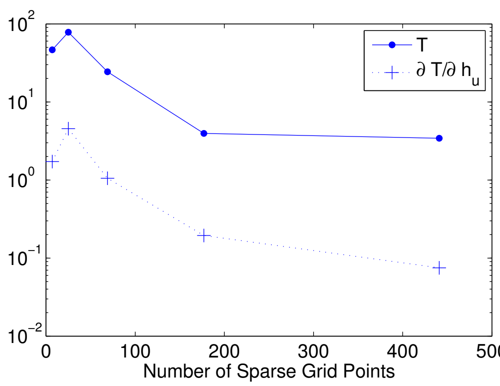


Figure 11. Convergence as a function of level

A difficult question to resolve is the level of sparse grid needed. Figure 11 shows the convergence of the errors as a function of the number of sparse grid points for interpolating the temperatures and the sensitivities at a time of 1100 sec. It would appear that adequate convergence has been achieved at level 4. The errors displayed in Figures 9 and 10 are over the entire parameter space for a specific time. Unfortunately this is not the entire story. One also needs to be concerned about the accuracy of the interpolation over the entire time of the experiment. Figure 12 shows the effect of different levels. It is clear that adequate convergence has been reached by level 4 as long as the interpolation is restricted to the heating period. The interpolation is not capable of capturing the correct response over the entire period. The solution, of course is to fit each period, the heating and the cooling, independently.

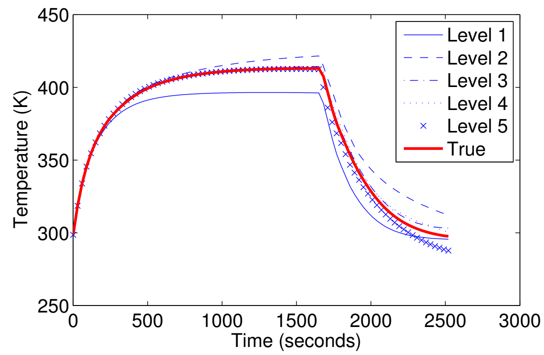


Figure 12a. Temperature as a function of time

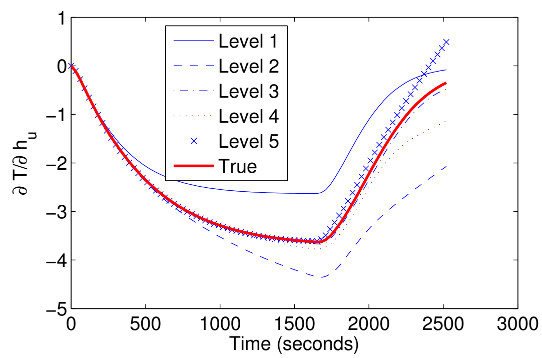


Figure 12b. $\partial T/\partial h_u$ as a function of time

5. Computing the pdf

Given an adequate level of interpolation accuracy, the marginal probability densities for the different parameters are obtained from Eq. 4b by Gaussian integration. The only problem is that if the parameter space sampled by the sparse grid is too large, the interpolation may not be sufficiently accurate to give good results. In addition, if the integration points have to cover too wide a parameter range, values of h far from the least squares estimates give such large values of $T(model) - T(data)$ that are used in evaluating the likelihood that substantial numerical problems will be encountered. Figure 13 shows a typical pdf based on 40 Gaussian quadrature points. For Figure 13a the quadrature points spanned a range $\pm$ 4 standard deviations derived from the least square estimates. For Figure 13b, the integration spanned $\pm$ 0.5 standard deviations.

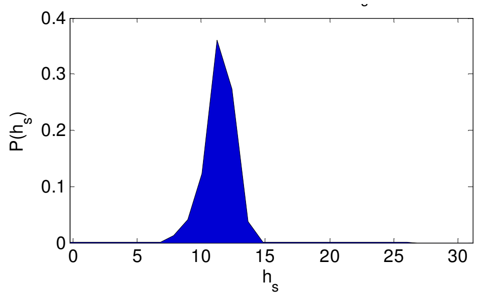


Figure 13a. Estimated pdf

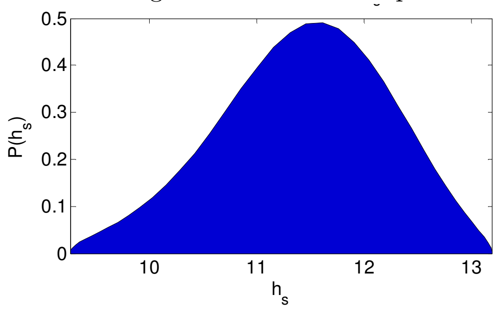


Figure 13b. Focused range of h

6. Conclusions

Calculating the marginal probability distribution from Eq. 3 requires a fairly dense sampling of the integrand near the point of maximum probability and, if using a sparse grid a reasonable level of accuracy. For the problems considered here, and true in general, the response surface shows some strong variations near the vertices of the parameter space, Figure 9b. Consequently it is important to choose the smallest parameter ranges possible. Although the maximum a-posteriori probability may not coincide with the estimated value of from the least squares analysis, $\hat{\theta}$, we have found that defining the range of each parameter to be $\hat{\theta} \pm n\ \sigma(\hat{\theta})$ where $\sigma(\hat{\theta})$ is given by Eq. 2c and n is a small number, usually 4 or 5, suffices. The advantage of the sparse grid (or of any other interpolation) is that locating the peak probability and the effect of assuming different priors in Eq. 4 can be examined with minimal computational expense. For the panel, computing the time history for a given set of parameters required in the order of 65 seconds compared to 0.8 seconds when using the sparse grid of level 5. Further details of both the sparse grid and Bayesian inference methods are available in [5].

7. Acknowledgments

A portion of this research was conducted under the sponsorship the Washington State Technology Center, contract RTD09 UW GS09, the United States Federal Aviation Administration Cooperative Agreement 08-C-AM-UW, and the National Science Foundation, Grant 0626533. We wish to thank Professor N. R. Aluru and Doctoral candidate, N. Agarwal for the initial version of the sparse grid basis program.

References

[1] Press, W. H., Flannery, B. P., Teukoslky, S. A. and Vetterling, W. T., 1986, *Numerical Recipes,*, Cambridge Univ. Press, Cambridge, UK

[2] COMSOL Version 3.5a, *Heat Transfer Module User's Guide,*COMSOL Multiphysics, 2008

[3] O'Hagan, A and Forster, J, 2004, *Bayesian Inference, Kendall's Advanced Theory of Statistics,* Vol. **2b**, Oxford Univ. Press, Oxford, UK

[4] Smolyak, S., 1963, "Quadrature and Interpolation Formulas for Tenser Product of Certain Classes of Functions," *Soviet Math. Dokl.,* 4, pp 240-243

[5] Emery, A. F. and Johnson, K. C., 20112, "Practical Considerations when Using Sparse Grids with Bayesian Inference for Parameter Estimation," *Proc. 2011 International Conference on Inverse Problems in Engineering*, May 4-6, Orlando, Florida, USA