Material Identification of Solder Joint in Microelectronics Package Using Bayesian Approach

J.H. Gang\textsuperscript{a}, J.H. Choi\textsuperscript{b}, D. An\textsuperscript{a} and J.W. Joob\textsuperscript{b}

\textsuperscript{a}School of Aerospace and Mechanical Engineering, Korea Aerospace University \\
kangjinx@naver.com, jhchoi@kau.ac.kr, skal34@nate.com

\textsuperscript{b}Dept. of Mechanical Engineering, Chungbuk National University \\
jinwon@chungbuk.ac.kr

Abstract: In this study, a method of computer model calibration is applied to quantify the uncertainties arising in the material characterization of the solder joint in the microelectronics package subject to a thermal cycle. The uncertainties include inherent experimental error and insufficient number of experiments. In the case of costly computation, surrogate model is also employed to approximate the original response function using a finite number of analyses, which adds another uncertainty. In this study, all these uncertainties are addressed by using a Bayesian calibration approach. A special specimen that characterizes the solder property due to the shear deformation is prepared, from which the Moiré fringe is measured by running a thermal cycle. Viscoplastic finite element analysis procedure is constructed for the specimen based on the Anand model. Gaussian process model known as Kriging is employed to approximate the original FEA model. Posterior distribution for the unknown Anand parameters is formulated from the likelihood function for joint full-field displacements of computation and experiment. Markov Chain Monte Carlo (MCMC) method is employed to simulate posterior distribution. As a result, the displacements and stresses are predicted in the form of confidence interval. The results show that the proposed approach can be a useful tool in the estimation of the unknown material parameters in a probabilistic manner by effectively accounting for the uncertainties due to the experimental and computational models.

Keywords: Anand Model, Bayesian Calibration, Kriging, Likelihood Function, Markov Chain Monte Carlo Technique, Material Identification, Solder Joint, Viscoplasticity.

1. Introduction

One of the most important concerns of microelectronic packaging technology such as ball grid array is to estimate the deformation behavior of solder joint. The reason is that the solder joint failure occurs as a result of the crack growth due to the excessive deformation under the repeated thermal loading, which is induced by the difference of thermal expansion coefficient of the components. Once we have the information, reliability of the solder joint in service can be estimated at the design stage. Finite element analysis is widely used to evaluate the deformation behavior as a means to save cost and time. In the analysis, characterizing material behavior of the solder, which exhibits strong nonlinearity with respect to the temperature and time, is utmost important for the sake of correct prediction. Numerous material models have been proposed in this respect. While the simplest attempt was to employ temperature dependent
elasto-plastic model, viscoplasticity and creep model are much more widely accepted in the analysis in order to account for the transient inelastic behavior at the elevated temperatures (Bodner, 2002).

In the analysis, material parameter values for the associated model of the solder are usually taken from the published literatures. The problem, however, is that the values are different with each other, which makes it hard to choose. Besides, the values are mostly from the tensile test of bulk specimen, whereas the solder may undergo shear stress predominantly. To overcome this difficulty, better practice is to measure the real deformation of the solder joint, and calibrate the material parameters so that the FEA results match those by the measurements. In this study, a special specimen of solder joint is fabricated for this purpose that can measure the shear deformation predominantly. In view of the solder materials, Tin-lead eutectic solder has been widely used in electronic systems, because it has low electrical resistance and low melting temperature. The leaded products, however, are not allowed due to the environmental reasons. Nevertheless, in this study, conventional solder of Sn37Pb is still considered to illustrate feasibility of our approach. Numerous researchers have tested this eutectic solder extensively and many constitutive properties are available in the literatures (Darveaux et al., 1995; Wang et al., 2001; Ye et al., 2002; Liu et al., 2008), in which the values vary substantially from each other in spite of the same material composition.

In a deterministic sense, calibration of the material parameters is just to obtain a single set of values that minimizes the difference of the results between the computer model and experiment. This is however awkward to be applied in the design due to the uncertainties experienced during the procedure, which include inherent experimental error and insufficient number of experiments. Furthermore, in the case of costly computation, which is the case in this study, surrogate model is also employed to approximate the original response function using a finite number of analyses. This adds another uncertainty. In this study, all these uncertainties are addressed by using a Bayesian calibration approach. Before the main applications, a mathematical example is considered to illustrate the concept and the procedure of the computer model calibration. The method is then applied to the inverse problem to determine properties of solder joint in the special specimen.

2. Procedure for Computer Model Calibration

In this section, overall procedure is addressed that accounts for the uncertainties in the calibration of the computer model with the experimental data. The computer code has two types of input variables: a set of input variables \( \mathbf{x} \) can be manipulated and measured in the field. Another set of parameters \( \mathbf{t} \) are used to adjust the computer model in order to represent the physical system more closely. Then, the experimental observations of the response variable are assumed to be given by

\[
y^{e} (\mathbf{x}) = y^{f} (\mathbf{x} | \mathbf{t}) + \varepsilon
\]

where \( y^{e} \) and \( y^{f} \) are the computer model and experimental output, respectively, and \( \varepsilon \) is the white-noise error with distribution \( N(0, \sigma^2) \). The parameters \( \mathbf{t} \), which often denote the physical constants, are the uncertain to be determined along with the variance \( \sigma^2 \) as a result of the calibration. The notation \( y^{f} (\mathbf{x} | \mathbf{t}) \) describes that the response mean of the computer model is a function of \( \mathbf{x} \) conditional on the given parameters \( \mathbf{t} \). In the original procedure (Loeppky et al., 2006), one more term called bias or discrepancy of the computer model is considered in the formulation, which is attributed to a poor or incorrect formulation of underlying physics. In this study, this is not considered for simplicity by assuming that the computer
model reflects the reality appropriately, i.e., the calibrated computer model gets closer toward the reality as the number of experimental data increases.

Suppose we have \( n \) experimental observations \( y^e = \{y_1^e, ..., y_n^e\} \) at a set of points \( D^e = \{x_1^e, ..., x_n^e\} \). Then the likelihood of the observation error at the points \( D^e \) is defined as

\[
L(y^e | \mathbf{t}, \sigma) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left( y^e - y^e (\mathbf{x}, \mathbf{t}) \right)' \left( y^e - y^e (\mathbf{x}, \mathbf{t}) \right) \right\}
\]

which is namely the simultaneous probability to get \( y^e \) given the parameters \( \mathbf{t}, \sigma \). In the Bayesian approach, prior information for the parameters is multiplied to the likelihood to obtain the posterior distribution of the parameters. In this study, the prior is not considered and only the likelihood is used, since this is conceptually and numerically more simple. Then Eq. (2) is the posterior distribution of the unknown parameters \( \mathbf{t}, \sigma \) conditional on the observed data \( y^e \).

If the computer model is computationally expensive as is the case of the FEA, surrogate model is a viable choice, which approximates the original model using the computer experiments at a finite number of points. This, however, introduces additional uncertainty, called code uncertainty. Let us denote \( \mathbf{X} = (\mathbf{x}, \mathbf{t}) \) and suppose we have evaluated computer outputs \( y^c = \{y_1^c, ..., y_m^c\} \) at the \( m \) number of DOE points \( D^c = \{X_1^c, ..., X_m^c\} \). The points of \( D^c \) need not be the same as those of \( D^e \). Common practice is to employ Latin Hypercube Sampling (LHS) for generating DOE points. Then the original model is approximately expressed via Gaussian random function (Gelman et al., 2004), which is also known as Kriging, as

\[
\hat{y}(\mathbf{X}) = f(\mathbf{X})\mathbf{\beta} + Z(\mathbf{X}), \quad Z \sim N(0, \sigma_c R), \quad R = R(\mathbf{X}; \mathbf{X}')
\]

where ^ denotes the surrogate representation, and \( R \) is a correlation function between \( \mathbf{X} \) and \( \mathbf{X}' \) which is typically represented by

\[
R(\mathbf{X}, \mathbf{X}') = \exp \left\{ -\theta \|\mathbf{X} - \mathbf{X}'\|^2 \right\}
\]

where \( \theta \) is correlation parameter that controls the smoothness of the function. Owing to the surrogate model, the Kriging parameters \( \mathbf{\beta} \), standard deviation \( \sigma_c \) and correlation parameter \( \theta \) are added to the unknown uncertainties. Then the likelihood for the joint experimental and computational data is defined as

\[
L(\mathbf{y} | \mathbf{\beta}, \sigma^2, \mathbf{t}) \propto |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{F}\mathbf{\beta})' \Sigma^{-1} (\mathbf{y} - \mathbf{F}\mathbf{\beta}) \right\}
\]

In this expression, \( \sigma^2 = \{\sigma_c^2, \sigma_t^2\} \) is a collection of variances, and

\[
\mathbf{y} = \begin{bmatrix} y^e \\ y^c \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{F}^e \\ \mathbf{F}^c \end{bmatrix} = \begin{bmatrix} \mathbf{F}(D^e) \\ \mathbf{F}(D^c) \end{bmatrix},
\]

are \( (m + n) \) dimensional vector and \( (m + n) \times p \) matrix, respectively. Here, \( D_t^e = \{(x_1^e, t), ..., (x_p^e, t)\} \) denotes the augmented set of points including \( \mathbf{t} \) in order to relate with points at \( D^e \). The subscript \( \mathbf{t} \) represents the dependency of \( D^e \) with respect to \( \mathbf{t} \). \( \Sigma \) is the \( (m + n) \times (m + n) \) covariance matrix

\[
\Sigma = \begin{bmatrix} \sigma_c^2 \mathbf{R}^c & \sigma_c^2 \mathbf{R}^{ ce} \\ \sigma_t^2 \mathbf{R}^e & \sigma_t^2 \mathbf{R}^{ te} + \sigma_c^2 \mathbf{I}_n \end{bmatrix} = \sigma_c^2 \mathbf{R} + \begin{bmatrix} 0 & 0 \\ 0 & \sigma_t^2 \mathbf{I}_n \end{bmatrix}
\]

where

\[
\mathbf{R}^{ ce} = \mathbf{R}(D^e, D^c), \quad \mathbf{R}^{ te} = \mathbf{R}(D_t^e, D^c) \quad \text{and} \quad \mathbf{R} = \mathbf{R}(D, D)
\]
are \( m \times n, n \times n \) and \((m+n) \times (m+n)\) matrices, respectively, and \( D = \{ D^c, D^e \} \) denotes the entire set of points. Note that the likelihood function depends on \( \theta \) through the set of points \( D^c \). Ignoring the priors, the likelihood as given in Eq. (8) becomes the posterior distribution of the parameters conditional on the data \( y \).

In general, the posterior distribution is not given in a form of predefined standard distribution functions which makes it hard to exploit existing formula. A convenient but general approach is to use MCMC (Markov Chain Monte Carlo) simulation, which is a modern computational method to draw random sequence of parameters that samples the given distribution (Andrieu et al., 2003). From the generated samples, we can estimate the statistical properties of the parameters including the mean and confidence interval. The advantages of MCMC are that it is computationally more efficient than the grid approximation or importance sampling, it does not require normalizing constant of target distribution, and it is useful when the target probability distribution is given as complex or implicit in terms of parameters, which is difficult to integrate in traditional way.

Once we have obtained the distribution of the parameters, we can proceed to draw the posterior predictive distribution of \( y^p \) at a new set of untried points \( D^p \), which may represent our degree of confidence supported by the provided data. The predictive distribution follows multivariate normal distribution, in which the mean is given by

\[
E\left( y^p | y, \beta, \sigma^2, \theta \right) = F^p \beta + \Sigma^p \Sigma^{-1} (y - F \beta) \tag{9}
\]

where

\[
F^p = F(D^p), \quad \Sigma^p = \sigma^2 \left[ R^c, R^e \right] = \sigma^2 R^p \tag{10}
\]

\[
R^c = R(D^c), \quad \text{and} \quad R^e = R(D^e, D) \tag{11}
\]

and the covariance is given by

\[
cov\left( y^p | y, \beta, \sigma^2, \theta \right) = \Sigma^p - \Sigma^p \Sigma^{-1} \Sigma^p \tag{12}
\]

where

\[
\Sigma^p = \sigma^2 R^p \quad \text{and} \quad R^p = R(D^p, D^p) \tag{13}
\]

The overall procedure is summarized as follows.

1. Establish computer model \( y^c(x | \theta) \) where \( \theta \) are unknown parameters to be determined. Perform experiments to obtain observed data, which are related by the computer model through Eq. (1) where \( \xi \sim N(0, \sigma^2) \).

2. In the case of costly computation for the computer model, generate DOE points using LHS technique and evaluate response at the points in order to construct surrogate approximation.

3. Draw samples of the unknown parameters \( \beta, \sigma^2, \theta \) using MCMC technique from the posterior distribution given by Eq. (5).

4. Generate samples of posterior predictive distribution using Eq. (9) and (12) from the drawn samples.
3. Mathematical example

A mathematical example is considered to demonstrate the presented method, which was used by (Loeppky et al., 2006). Suppose that the true physical model is given by the equation

\[ y(x) = 1.5 + 3.5 \exp(-1.7x); \quad x \in [0, 3] \]  

(14)

Let us assume that we have measured responses fictitiously for this physical model with the error \( \epsilon \sim N(0, 0.36) \). Fictitious experimental data \( \mathbf{y}^e = \{y_1^e, \ldots, y_n^e\}' \) are randomly generated at the points \( \mathbf{D}^e = \{\mathbf{x}_1^e, \ldots, \mathbf{x}_n^e\}' \) which are equally spaced with the interval 0.3. The result is given as star symbols in Figure 1. Assuming that we have no prior knowledge on this model, computer model is posed by a single parameter in the following form:

\[ y^e(x \mid t) = 5\exp(-tx) \]  

(15)

The objective is to best estimate the unknown parameter \( t \) conditional on the experimental data. Before applying the proposed method, the problem is solved for the case of single parameter function using the

![Figure 1. Fictitious experimental data and results of calibration](image1)

![Figure 2. Posterior distributions of the two parameters and their MCMC history](image2)
classical regression technique for comparison purpose. Take log to Eq. (15) to get
\[ w(x) = \beta x \]  
(16)
where \( w(x) = \log \left( \frac{y^c}{5} \right) \) and \( \beta = -t \). Then the problem becomes classical linear regression. Following the regression procedure (Hines et al., 2003), one obtains
\[ \beta = (F'F)^{-1} F'w \]  
and
\[ \sigma^2 = \frac{1}{n-1} (F\beta - w)' (F\beta - w) \]  
(17)
where \( F = [x_1^c, \ldots, x_n^c]' \), \( w = \log \left( \frac{y^c}{5} \right) \) and \( \sigma^2 \) is the variance of the regression model. The computed values are \( t = 0.5293 \) and \( \sigma = 0.3270 \), and the regression curve is plotted in Figure 1. As expected, the curve deviates somewhat from the main stream of the data due to the wrong assumption of the trial function.

| Table I. Mean and confidence bands of the two parameters |
|------------|------------|------------|
|            | 2.5%       | Median     | 97.5%      |
| \( t \)    | 0.4781     | 0.6557     | 0.9147     |
| \( \sigma \) | 0.4542     | 0.6855     | 1.1284     |

Figure 3. Superposition of original and surrogate model

Figure 4. Posterior prediction of surrogate model as a result of calibration
Material Identification of Solder Joint in Microelectronics Package Using Bayesian Approach

In the proposed method, the unknown parameters $t$ and $\sigma$ are obtained in the form of the posterior pdf using Eq. (2). In order to sample the parameters that follow this joint pdf, 50,000 iterations of MCMC simulation is implemented. The traces of the MCMC history as well as the posterior distribution are given in Figure 2 for the two parameters. The median values and the confidence intervals are also given in Table I. Using the obtained samples, one can generate the same number of functions over the range $[0, 3]$, which represents the posterior predictive distribution. The median and predictive intervals of the function are extracted from this, which are plotted in Figure 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2.5%</th>
<th>Median</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>-0.0735</td>
<td>0.2414</td>
<td>0.5308</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.3739</td>
<td>0.5594</td>
<td>0.8616</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.8070</td>
<td>1.8070</td>
<td>1.8088</td>
</tr>
<tr>
<td>$\sigma_\epsilon$</td>
<td>1.9055</td>
<td>1.9196</td>
<td>1.9198</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.3110</td>
<td>0.4731</td>
<td>0.6968</td>
</tr>
</tbody>
</table>

In the proposed method, the unknown parameters $t$ and $\sigma$ are obtained in the form of the posterior pdf using Eq. (2). In order to sample the parameters that follow this joint pdf, 50,000 iterations of MCMC simulation is implemented. The traces of the MCMC history as well as the posterior distribution are given in Figure 2 for the two parameters. The median values and the confidence intervals are also given in Table I. Using the obtained samples, one can generate the same number of functions over the range $[0, 3]$, which represents the posterior predictive distribution. The median and predictive intervals of the function are extracted from this, which are plotted in Figure 1.

Though the example involves a simple mathematics, the computer model is now assumed as costly to evaluate for illustration purpose. By treating $x$ and $t$ in (14) as the two independent input variables, surrogate model is constructed using the computed values at the 5 LHS points. The results are given in Figure 3 which shows superposed surfaces of the original and surrogate model. Posterior distributions of the unknown parameters, which are $t, \sigma, \beta, \sigma_\epsilon$, and $\theta$, are obtained as given in the Table II. The mean and predictive interval on the true function is given in Figure 4, which represents our knowledge based on the given data.

4. Bayesian Calibration of the Special Specimen

A special specimen is designed and fabricated to estimate material property of solder for the purpose to induce the shear deformation predominantly as shown in Figure 5. The solder material is Sn37Pb. The upper and lower materials are SUS630 and copper, respectively. Temperature independent material properties are given in Table III. The specimen undergoes a thermal cycle as shown in Figure 6. Due to the different coefficient of thermal expansion (CTE), the solder in between the two metals undergoes shear deformation. Fringe patterns are measured by Moiré interferometry (Joo et al., 2005), as shown in Figure 7. Finite element analysis procedure is created to conduct viscoplastic analysis. Anand model is employed for characterizing the solder material as shown in Table IV. It is remarked that the parameter values from the literatures differ significantly from each other as is illustrated in table which are the case of Sn40Pb. Initial analysis is carried out using the Sn37Pb values by (Liu et al., 2008), and the results are compared with those by the experiments in Figure 8. In the figure, lines with the solid circles and squares denote the displacement at A and B respectively as shown in Figure 8. The four points at each curve denote the displacements at the points (a) ~ (d) in the Figure 6, which are the beginning of the dwell at 75°C, 100°C, 125°C and the end of cooling to 25°C, respectively. As shown in the figure, the two does not agree, which means that the material property taken from the literature differ from the actual ones of our experiment.

Table II. Mean and confidence bands of the parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2.5%</th>
<th>Median</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>-0.0735</td>
<td>0.2414</td>
<td>0.5308</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.3739</td>
<td>0.5594</td>
<td>0.8616</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.8070</td>
<td>1.8070</td>
<td>1.8088</td>
</tr>
<tr>
<td>$\sigma_\epsilon$</td>
<td>1.9055</td>
<td>1.9196</td>
<td>1.9198</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.3110</td>
<td>0.4731</td>
<td>0.6968</td>
</tr>
</tbody>
</table>
Table III. Material properties of the special specimen

<table>
<thead>
<tr>
<th>Material</th>
<th>Young’s modulus (GPa)</th>
<th>Poisson’s ratio</th>
<th>CTE (ppm / °C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUS630</td>
<td>200</td>
<td>0.29</td>
<td>10.8</td>
</tr>
<tr>
<td>Sn37Pb</td>
<td>32</td>
<td>0.35</td>
<td>23.5</td>
</tr>
<tr>
<td>Copper</td>
<td>130</td>
<td>0.34</td>
<td>16.3</td>
</tr>
</tbody>
</table>

Figure 5. Schematic of the special specimen and finite element model

Figure 6. Illustration of temperature profile used in experiment

Figure 7. Representative fringe patterns of the special specimen during a thermal cycle
Bayesian approach is implemented to find out the true values of material parameters. Four constants $s_0$, $Q/R$, $\zeta$ and $m$ are taken from the Table V as the unknowns to be determined. The reason to choose these is that after a sensitivity study, these are found to affect the analysis result most significantly. By treating four constants as the independent input variables, surrogate model is constructed using the computed values at the 50 LHS points. After implementing MCMC, the posterior distributions of the four parameters are obtained in Figure 9. The trace of the MCMC history is also given to show that they have converged to the given distribution. The values of the mean and the confidence bands are also given in Table V. Once the distributions of the parameters are obtained, the posterior predictive displacements using these values are determined in the form of distribution. The results at the four time points are given in Figure 10 along with their confidence bands. The mean displacements show close agreement with the experimental results.
Table V. Mean, median and confidence bands of the four parameters and measurement error

<table>
<thead>
<tr>
<th></th>
<th>2.5 %</th>
<th>Median</th>
<th>Mean</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$</td>
<td>9.8372</td>
<td>12.9200</td>
<td>12.9060</td>
<td>15.7330</td>
</tr>
<tr>
<td>$Q/R$</td>
<td>8677.0000</td>
<td>10586.0000</td>
<td>10555.0000</td>
<td>12337.0000</td>
</tr>
<tr>
<td>$\xi$</td>
<td>1.3065</td>
<td>2.0424</td>
<td>2.1217</td>
<td>3.4745</td>
</tr>
<tr>
<td>$m$</td>
<td>0.0035</td>
<td>0.0598</td>
<td>0.0834</td>
<td>0.3142</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.4716</td>
<td>1.2355</td>
<td>1.2382</td>
<td>1.9531</td>
</tr>
</tbody>
</table>

Figure 9. Posterior distributions of the four parameters and their MCMC history

Figure 10. Confidence bands of posterior predictive distribution
5. Conclusions

In this study, computer model calibration is implemented to estimate unknown material parameters of the solder with viscoplastic behavior. Special specimen is prepared to this end, and displacements are measured by Moiré interferometry through a thermal cycle. The uncertainties arising from the experiments and surrogated computer model are taken into account. The posterior distributions of the unknown material parameters are determined by using MCMC iteration. The displacements at the interested time points are then predicted in the form of confidence interval. The results show that the proposed approach can be a useful tool in the estimation of the unknown material parameters in a probabilistic manner by effectively accounting for the uncertainties due to the experimental and computational models.

The work provided in this study was concentrated on the maximum V displacement in SUS630 and copper because the displacements are affected directly by behavior of solder. More important values, however, are the maximum strains or maximum stresses in the solder. Therefore, another attempt must be conducted to obtain the maximum strains in the solder and compared with the numerical predictions. In addition, the estimated material parameters by the proposed approach should be applied to the actual microelectronic packages such as WB-PBGA and FC-PBGA and get the agreement within the confidence level. Then the approach is proved to be truly valid for inverse estimation of constitutive solder properties.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2007-0056530).

References

