Design and Analysis for Multifidelity Computer Experiments

Ying Hung
Department of Statistics and Biostatistics
Rutgers University
Overview

• Introduction to computer experiments
  – Design issues
  – Modeling issues

• Analysis for multifidelity computer experiments

• Improvements based on variable selection
Introduction to computer experiments

- First computer experiments were conducted at Los Alamos National Laboratory to study the behavior of nuclear weapons.
- Computer experiments are becoming popular because many physical experiments are difficult or impossible to perform.
Properties of computer experiments

• Computer experiments refer to those experiments that are performed in computers using physical models and finite element analysis.

• Deterministic outputs (no random error)
  ➢ No replicates required
  ➢ Interpolation

• Large number of variables
• Time-consuming, expensive
Experimental design for computer experiments

- Latin hypercube design (LHD).
  - McKay, Beckman, Conover (1979).
  - Easy to construct.
  - One-dimensional balance.

<table>
<thead>
<tr>
<th>run</th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Computer experiments modeling

• Universal kriging \[ Y(x) = \mu(x) + Z(x), \]

  • \[ \mu(x) = \sum_{i=0}^{m} \mu_i v_i(x), \]

  • \[ Z(x): \text{a weak stationary stochastic process with mean 0 and covariance function } \sigma^2 \psi. \]

  • \( v_i \)'s: known functions, \( \mu_i \)'s: unknown parameters.

  • \( \text{cov}\{Y(x + h), Y(x)\} = \sigma^2 \psi(h), \) where the correlation function \( \psi(h) \) is a positive semidefinite function with \( \psi(0) = 1 \) and \( \psi(-h) = \psi(h). \)

• Ordinary kriging

  \[ Y(x) = \mu_0 + Z(x) \]
GP example

true y and 4 predicted lines
Analysis for Multifidelity Computer Experiments

• Multifidelity computer experiments
  – Physical experiments and computer experiments
  – Computer experiments with different accuracy

• An example in electronic packaging

• Objective:
  – Study effect of initial PWB warpage on low cycle fatigue reliability of solder bumps based on:
    • computer experiments: Finite Element Modeling (FEM)
    • physical experiments: accelerated thermal cycling test
  – How to calibrate?
Finite Element Modeling

Purpose: To Study How Initial PWB (Printed wiring board) Warpage Affects Solder Bump Fatigue Reliability

- PWB warpage was measured at eutectic temp. and used as initial geometric input to FEM

Max. warpage across PWB = 2158.9 micron

Warpage Measurement of Sample 2 at 183°C

Meshed PWB with 35×35mm PBGA
Study of How PWB Warpage Affects Solder Bump Fatigue

- Case Studies for:
  - Sn-Pb (Tin-Lead) and Sn-Ag-Cu (Lead-Free) Solder Bumps on
    - Two Packages (256-bump 27×27-mm PBGA and 352-bump 35×35-mm PBGA)
      - Each package placed at three different locations:

- PWB samples can have different initial warpage or can be flat
  - PWBA warpage can be either convex or concave as shown below:

- Total 42 cases for each package [including 2 types of solder, 3 chip locations, 3 PWB samples, 2 warpage shapes, and ideal PWB (2 solder types plus 3 package locations) w/o warpage]
## Factors studied in FEM

- **Factors:**

<table>
<thead>
<tr>
<th>$w_{\text{max}}$</th>
<th>maximum initial PWB warpage at 25°C (mm)</th>
<th>2105.3, 3076.6, 3824.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_{\text{shape}}$</td>
<td>warpage shape</td>
<td>+1: Convex up; -1 Concave up</td>
</tr>
<tr>
<td>$d_p$</td>
<td>package dimension (mm)</td>
<td>27 by 27, 35 by 35</td>
</tr>
<tr>
<td>$l_p$</td>
<td>location of package (mm)</td>
<td>Center, 60-30, Outmost</td>
</tr>
<tr>
<td>$m_s$</td>
<td>solder bump material</td>
<td>Sn-Pb, Lead-free</td>
</tr>
</tbody>
</table>

$N_f = \text{fatigue life estimation of solder bumps (cycles)}$
Data from computer experiments

- FEM data (84 runs)

<table>
<thead>
<tr>
<th></th>
<th>27by27mm</th>
<th>Initial Warp</th>
<th>-3824</th>
<th>-3076.6</th>
<th>-2105.3</th>
<th>0</th>
<th>2105.3</th>
<th>3076.6</th>
<th>3824</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>at 25C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sn-Pb</td>
<td>Center</td>
<td></td>
<td>1356</td>
<td>1523</td>
<td>1737</td>
<td>1981</td>
<td>1755</td>
<td>1556</td>
<td>1393</td>
</tr>
<tr>
<td></td>
<td>60-30</td>
<td></td>
<td>1438</td>
<td>1618</td>
<td>1823</td>
<td>2005</td>
<td>1846</td>
<td>1644</td>
<td>1481</td>
</tr>
<tr>
<td></td>
<td>Outmost</td>
<td></td>
<td>1477</td>
<td>1652</td>
<td>1861</td>
<td>2034</td>
<td>1884</td>
<td>1691</td>
<td>1518</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3076.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2105.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2105.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3076.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lead-Free</td>
<td>Center</td>
<td></td>
<td>1618</td>
<td>1807</td>
<td>2042</td>
<td>2215</td>
<td>2061</td>
<td>1845</td>
<td>1660</td>
</tr>
<tr>
<td></td>
<td>60-30</td>
<td></td>
<td>1709</td>
<td>1905</td>
<td>2144</td>
<td>2248</td>
<td>2167</td>
<td>1946</td>
<td>1751</td>
</tr>
<tr>
<td></td>
<td>Outmost</td>
<td></td>
<td>1756</td>
<td>1950</td>
<td>2189</td>
<td>2280</td>
<td>2214</td>
<td>1990</td>
<td>1804</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>35by35mm</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sn-Pb</td>
<td>Center</td>
<td></td>
<td>1429</td>
<td>1620</td>
<td>1838</td>
<td>2098</td>
<td>1865</td>
<td>1656</td>
<td>1482</td>
</tr>
<tr>
<td></td>
<td>60-30</td>
<td></td>
<td>1512</td>
<td>1706</td>
<td>1933</td>
<td>2125</td>
<td>1969</td>
<td>1739</td>
<td>1558</td>
</tr>
<tr>
<td></td>
<td>Outmost</td>
<td></td>
<td>1549</td>
<td>1751</td>
<td>1975</td>
<td>2161</td>
<td>1998</td>
<td>1786</td>
<td>1599</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3076.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2105.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2105.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3076.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lead-Free</td>
<td>Center</td>
<td></td>
<td>1715</td>
<td>1905</td>
<td>2158</td>
<td>2343</td>
<td>2183</td>
<td>1947</td>
<td>1759</td>
</tr>
<tr>
<td></td>
<td>60-30</td>
<td></td>
<td>1841</td>
<td>2006</td>
<td>2251</td>
<td>2380</td>
<td>2278</td>
<td>2054</td>
<td>1885</td>
</tr>
<tr>
<td></td>
<td>Outmost</td>
<td></td>
<td>1882</td>
<td>2063</td>
<td>2304</td>
<td>2415</td>
<td>2341</td>
<td>2097</td>
<td>1929</td>
</tr>
</tbody>
</table>
Objective: To verify and correlate 3-D finite element simulation results.

PWB with 35×35 mm PBGA at Location 2

PWB with 35×35 mm PBGA at Location 4

Accelerated Thermal Cycling Test

Standard Thermal Cycling Profile
FEM Simulation vs Experimental Study

Experimental Data

FEM Simulations

Fatigue life (Cycle)

maximum PWB warpage

-4000 -2000 0 2000 4000

-4000 -2000 0 2000 4000

1400 1600 2000 2400

1400 1600 2000 2400
Analysis for the two types of data

- Model fitting base on FEM and experimental data:
  - Step 1: Fit kriging model using only simulation data

\[
\hat{N}_k(\bar{x}) = 1101.6 + \varphi(\bar{x})^T \Psi^{-1}(N_{out} - 1101.6I)
\]

where \(N_{out}\) = FEM output data.

- Step 2: Calibrate fitted model in Step 1 with experimental data

\[
\hat{N}_f(\bar{x}) = 1830.3 - 540w_{\text{max}} + \hat{N}_k(\bar{x})
\]

\[
= 2931.9 - 540w_{\text{max}} + \varphi(\bar{x})^T \Psi^{-1}(N_{out} - 1101.6I)
\]

where \(\hat{N}_f(\bar{x})\) = fatigue life prediction

\(w_{\text{max}}\) = maximum initial PWB warpage at 25°C
Calibration Base on Experimental Data

Experimental Data – Kriging Prediction

$w_{\text{max}}$  PWB warpage (Standardized)
Improvement based on variable selection

• Universal kriging \( Y(x) = \mu(x) + Z(x) \),

  \( \mu(x) = \sum_{i=0}^{m} \mu_i v_i(x) \),

• \( Z(x) \): a weak stationary stochastic process with mean 0 and covariance function \( \sigma^2 \psi \).

• \( v_i \)'s: known functions, \( \mu_i \)'s: unknown parameters.

• \( cov\{Y(x + h), Y(x)\} = \sigma^2 \psi(h) \), where the correlation function \( \psi(h) \) is a positive semidefinite function with \( \psi(0) = 1 \) and \( \psi(-h) = \psi(h) \).

• Ordinary kriging

  \( Y(x) = \mu_0 + Z(x) \)
Problems with GP model

- Problems with the ordinary kriging model
  - The prediction can be poor if there are some strong trends.
  - It is not easy to understand the effects of the factors by just looking at the predictor.
  - Predictor not robust to the misspecification in the correlation parameters.

- It has been noted that the prediction accuracy and model efficiency of a GP model can be improved by identifying important variables (Welch et al. 1992, Cressie 1993, Martin and Simpson 2005, Gramacy and Lee 2008, Joseph et al. 2008, Stein 2008, Kaufman et al. 2013).
Drawbacks with existing approaches

• Selections are perform based on specific types of a model with convenient by questionable assumptions.
  ➢ A GP model \( y(x) = \mu(x) + Z(x), \quad \mu(x) = \sum_{k=1}^{p} \beta_k x_k = f(x)'\beta, \)
  ➢ Blind kriging selects important variable based only on the mean function of GP models.
  ➢ Linkletter et al. (2006) introduced a variable selection procedure only for the correlation function.

• Computationally intensive
Bayesian variable selection for kriging

- A unified approach that can perform variable selection in a general GP model is attractive but nontrivial. Because the mean function and the correlation structure are not independent. The same variable can appear in either one part or both parts of the model to contribute the effect(s).
- Idea: Using a hierarchical Bayes formulation to connect different effects of the same variables in kriging models.
- Introduce a latent variable into kriging model to indicate if a particular variable is active or not. For those active variables, they can have effect in the mean function and/or in the correlation function.
Bayesian variable selection for kriging

- kriging model:

\[ y(x) = \mu(x) + Z(x), \quad \mu(x) = \sum_{k=1}^{p} \beta_k x_k = f(x)' \beta, \]

- Define a binary vector \( \gamma = (\gamma_1, \ldots, \gamma_p)' \). Such a vector is used to indicate if a particular variable is active or not.

- Priors: \[
\pi(\beta_k | \gamma_k) = (1 - \gamma_k)\delta(0) + \gamma_k DE(0, \tau_k), \\
\pi(\theta_k | \gamma_k) = (1 - \gamma_k)\delta(0) + \gamma_k Exp(\lambda_k), \\
P(\gamma) \propto q^{\|\gamma\|}(1 - q)^{p-\|\gamma\|}, \\
\sigma^2 \propto (\sigma^2)^{-\nu_0/2-1} \exp(-1/(2\sigma^2)).\]
Bayesian variable selection for kriging

- This approach is flexible but obtaining the posterior can be computationally difficult because it involves high-dimensional integration.
- With some mild assumptions, we can approximate the posterior by

\[ P(\gamma | y) \approx C(y)(\sqrt{\sigma^2 w})|\gamma| \times \exp \left( -\frac{1}{2} \min_{\beta, \theta} L_\rho(\beta, \theta) \right) . \]

\[ L_\rho(\beta, \theta) = \log |\Phi(\theta)| + \frac{(y - F\gamma \beta \gamma)' \Phi^{-1}(\theta)(y - F\gamma \beta \gamma) + \rho_1 \sum_{k \in \gamma} |\beta_k| + \rho_2 \sum_{k \in \gamma} \theta_k}{\sigma^2} . \]

- The approximation leads to a double penalized likelihood estimation problem
- Estimation: Coordinate descent algorithm
Summary

• Illustrates how to analyze multifidelity Computer Experiments using a real example.

• Analysis of computer experiments mainly based on GP models, in particular, ordinary kriging model.

• Proposed a Bayesian variable selection framework to improve prediction accuracy and model efficiency.
Dinner is ready!

Thank you!