Prediction for Computer Experiment Output Having Qualitative And Quantitative Input Variables

Gang Han

Ph.D student

Department of Statistics, The Ohio State University

September 10, 2006
Kickoff Workshop, SAMSI

With special thanks to Don Bartel and Tim Wright of the Joint Cornell-Hospital for Special Surgery Biomechanics Program

Outline

1. Introduction
   (a) Research Goal and A Motivating Example
   (b) Approaches for modeling computer experiments having one qualitative input variable
2. A Bayesian Model and Its Implementation
3. Competing Methods and Examples
4. Conclusion and Future Work

Introduction of Computer Experiment

- **Properties of Computer Experiments**
  1. Computer experiments complement physical experiments.
  2. Computer experiments yield deterministic answers but possibly have bias.
  3. Some difficulties
     - Some codes are time-consuming.
     - Some applications have large numbers of inputs.

Research Goals

1. Model the output of a computer experiment when the inputs are both quantitative and qualitative
2. Predict the unknown response
3. Applications such as calibration, optimization, etc.

A Motivating Example

Generic Knee Prosthesis

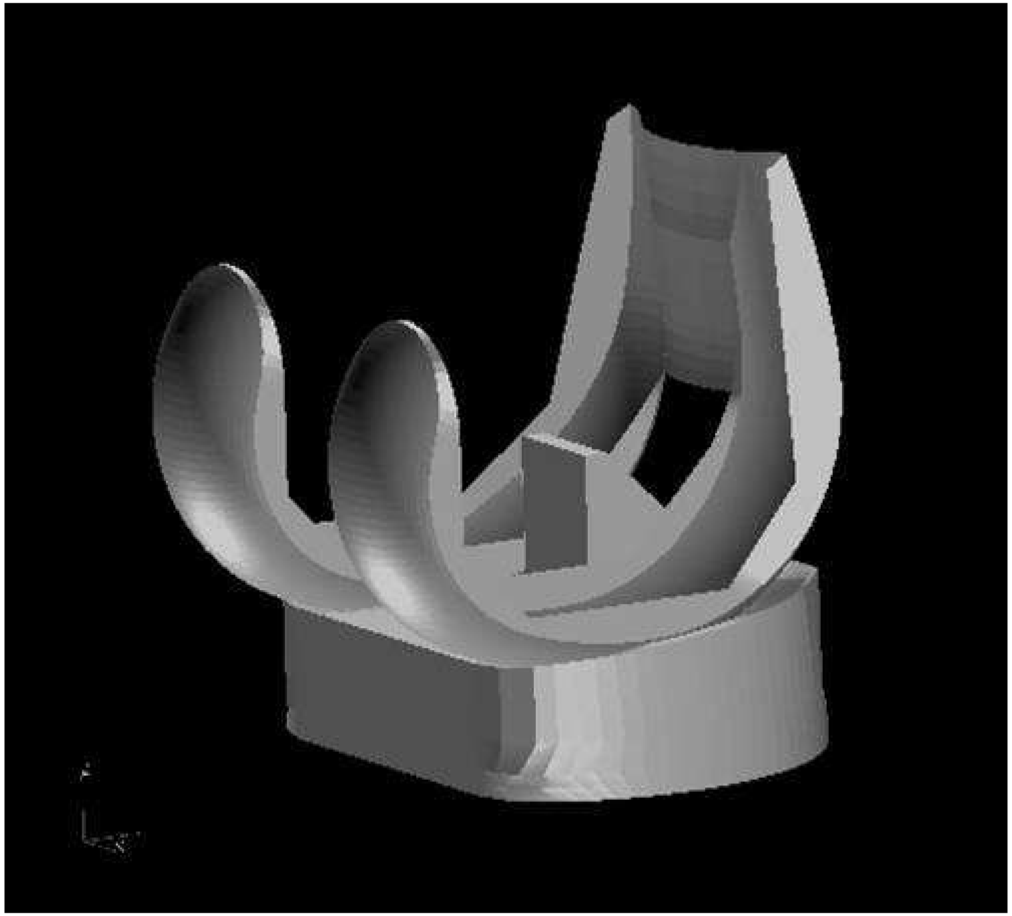

Figure 1: Appearance of the Knee Prosthesis

A Motivating Example

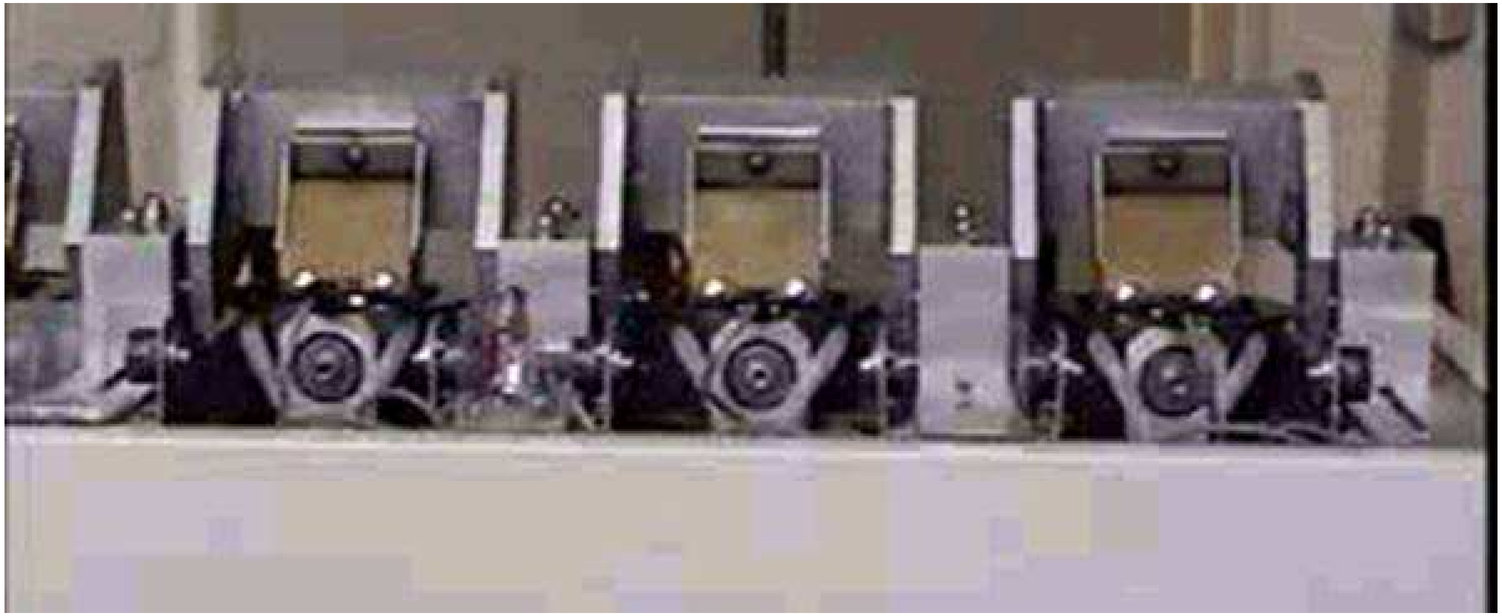

Figure 2: A 4-station *mechanical knee wear simulator machine (MKsim)*

A Motivating Example

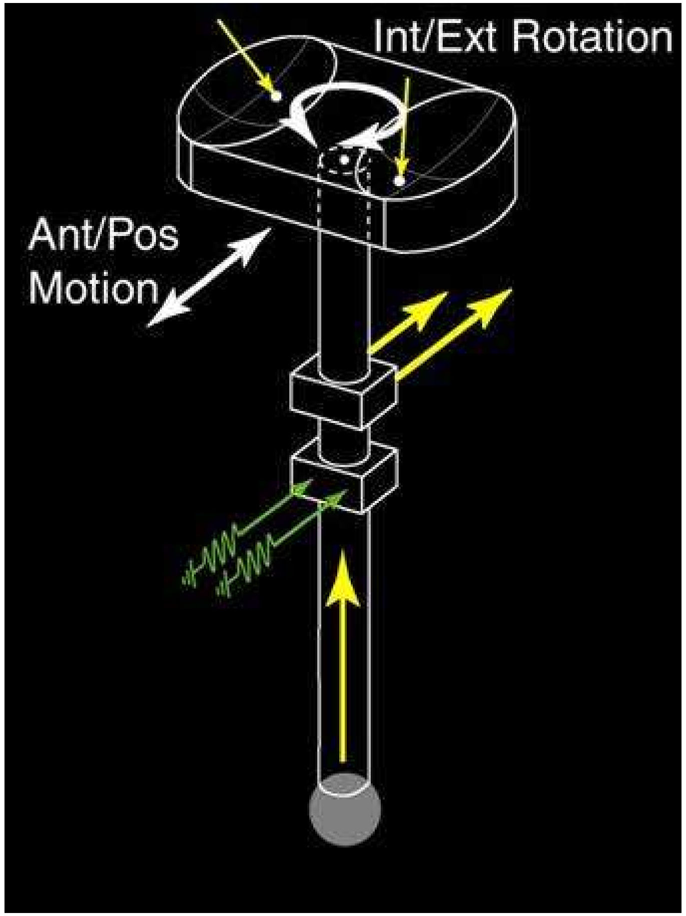

Figure 3: Movement of a mechanical knee simulator

A Motivating Example

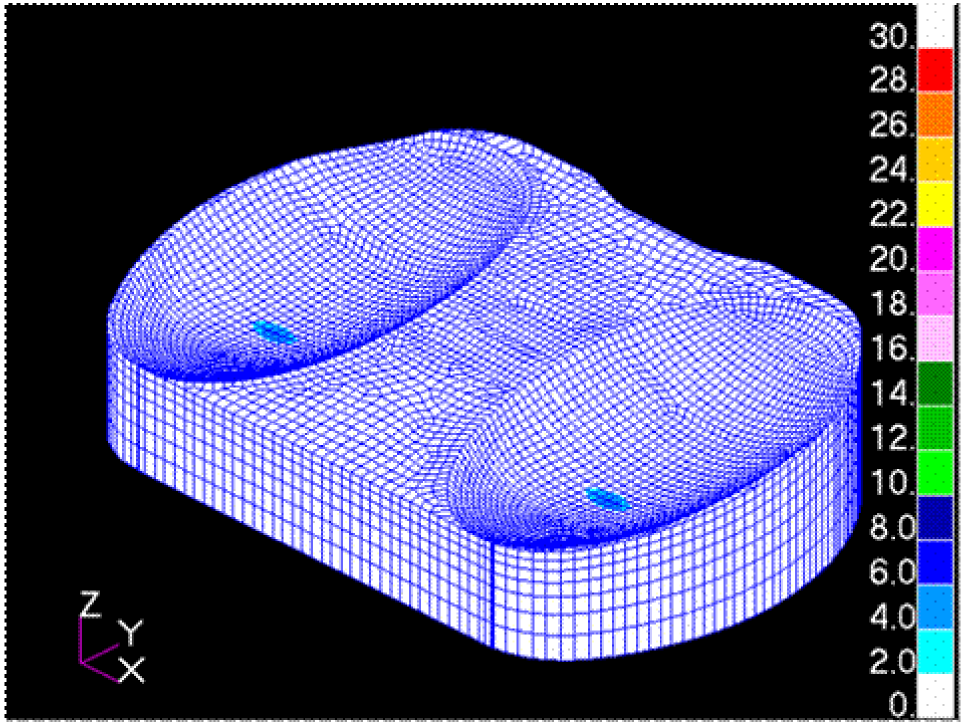

Figure 4: Mesh density used by the FEA computer program

A Motivating Example

This example has **mixed inputs**.

Some *qualitative input variables* are inherently qualitative or best treated as such

- *Mesh Density (FEA)*$\{$1,2,3$\}$
- *Initial AP(Anterior Posterior) position (FEA)*$\{$0,2,4,6$\}$

Some *quantitative input variables* are

- Spring length (FEA & MKSim)$\{$range:[24,48]$\}$
- Spring stiffness (FEA & MKSim)$\{$range:[14.5,43.5]$\}$
- Frequency (FEA)$\{$range:[0.3,1.4]$\}$

Some Approaches

Suppose the inputs are $(t, \boldsymbol{x})$, where $t \in \{1, 2, \ldots, T\}$ is qualitative and $\boldsymbol{x} \in [0, 1]^d$ is quantitative. The output is the real-valued function $Y(t, \boldsymbol{x})$. The training data is $\boldsymbol{Y}^n = (Y(t_1, \boldsymbol{x}_1), Y(t_2, \boldsymbol{x}_2), \ldots, Y(t_n, \boldsymbol{x}_n))$.

To illustrate the idea, we suppose that there is one qualitative variable with three categories, which means $t \in \{1, 2, 3\}$.

Some Approaches

Critical feature: $Y(t, \boldsymbol{x})$ should have valid covariance.

For any choice of $t_1, t_2, \ldots, t_n$ and the choice of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ the covariance function of $\boldsymbol{Y}^n$ is positive definite.

Approach 1

Idea: Use Gaussian process model with indicator variables in the covariance matrix. Let $I(t=1), I(t=2)$, and $I(t=3)$ be indicator functions, the response $y(t,\boldsymbol{x})$ can be modeled as a Gaussian Stochastic process with some mean β, and covariance

$$
\begin{aligned}
&Cov(Y(t_1,\boldsymbol{x}_1), Y(t_2,\boldsymbol{x}_2)) = \\
&\sigma^2 exp\{-\textstyle\sum_{j=1}^{d}\theta_j(x_{1,j}-x_{2,j})^2 \\
&-\alpha_1[I(t_1=1)-I(t_2=1)]^2 \\
&-\alpha_2[I(t_1=2)-I(t_2=2)]^2 \\
&-\alpha_3[I(t_1=3)-I(t_2=3)]^2\}.
\end{aligned}
$$

Approach 1

Critique:
1(-) The meaning of the parameters α_1, α_2, and α_3 are vague.

2(-) The parameterizations do NOT express all the possible correlation structures.
For example, if $t_1 = t_2$, the correlation structure will be the same no matter what the common value is.

Approach 2

Idea Model $Y(t, \boldsymbol{x})$ as a draw from a process constructed from linear combinations of independent processes $Z_0(\boldsymbol{x}), \ldots, Z_3(\boldsymbol{x})$.

The models below are based on linear combinations of independent Gaussian Processes, $Z_0(\cdot)$, $Z_1(\cdot)$, $Z_2(\cdot)$, and $Z_3(\cdot)$

$$
\begin{aligned}
Y(1, \boldsymbol{x}) &= \delta_{01} Z_0(\boldsymbol{x}) + \delta_{11} Z_1(\boldsymbol{x}) + \delta_{21} Z_2(\boldsymbol{x}) + \delta_{31} Z_3(\boldsymbol{x}) \\
Y(2, \boldsymbol{x}) &= \delta_{02} Z_0(\boldsymbol{x}) + \delta_{12} Z_1(\boldsymbol{x}) + \delta_{22} Z_2(\boldsymbol{x}) + \delta_{32} Z_3(\boldsymbol{x}) \\
Y(3, \boldsymbol{x}) &= \delta_{03} Z_0(\boldsymbol{x}) + \delta_{13} Z_1(\boldsymbol{x}) + \delta_{23} Z_2(\boldsymbol{x}) + \delta_{33} Z_3(\boldsymbol{x})
\end{aligned}
$$

Approach 2

Covariance structure for the comprehensive model

- Variance: $Var(Y(1,\boldsymbol{x})) = \delta_{01}^2\sigma_0^2 + \delta_{11}^2\sigma_1^2 + \delta_{21}^2\sigma_2^2 + \delta_{31}^2\sigma_3^2$.
- Covariace:
  - Same category, different inputs
    $Cov(Y(1,\boldsymbol{x}_1), Y(1,\boldsymbol{x}_2)) =$
    $\delta_{01}^2 Cov(Z_0(\boldsymbol{x}_1), Z_0(\boldsymbol{x}_2)) + \delta_{11}^2 Cov(Z_1(\boldsymbol{x}_1), Z_1(\boldsymbol{x}_2)) +$
    $\delta_{21}^2 Cov(Z_2(\boldsymbol{x}_1), Z_2(\boldsymbol{x}_2)) + \delta_{31}^2 Cov(Z_3(\boldsymbol{x}_1), Z_3(\boldsymbol{x}_2))$
  - Different categories, same inputs
    $Cov(Y(1,\boldsymbol{x}_1), Y(2,\boldsymbol{x}_1)) =$
    $\delta_{01}\delta_{02}\sigma_0^2 + \delta_{11}\delta_{12}\sigma_1^2 + \delta_{21}\delta_{22}\sigma_2^2 + \delta_{31}\delta_{32}\sigma_3^2$
  - Different categories, different inputs
    $Cov(Y(1,\boldsymbol{x}_1), Y(2,\boldsymbol{x}_2)) = \delta_{01}\delta_{02} Cov(Z_0(\boldsymbol{x}_1), Z_0(\boldsymbol{x}_2)) +$
    $\delta_{11}\delta_{12} Cov(Z_1(\boldsymbol{x}_1), Z_1(\boldsymbol{x}_2)) +$
    $\delta_{21}\delta_{22} Cov(Z_2(\boldsymbol{x}_1), Z_2(\boldsymbol{x}_2)) + \delta_{31}\delta_{32} Cov(Z_3(\boldsymbol{x}_1), Z_3(\boldsymbol{x}_2))$.

Approach 2

Special Case I: Given $\boldsymbol{\beta} = \{\beta_0, \beta_1, \beta_2, \beta_3\}$, $\boldsymbol{\delta} = \{\delta_{01}, \ldots, \delta_{33}\}$, $\boldsymbol{\sigma} = \{\sigma_0^2, \sigma_1^2, \sigma_2^2, \sigma_3^2\}$, and $\boldsymbol{\theta} = \{\theta_{01}, \ldots, \theta_{0d}, \ldots, \theta_{31}, \ldots, \theta_{3d}\}$,

$$\begin{aligned} Y(1, \boldsymbol{x}) &= \delta_{01} Z_0(\boldsymbol{x}) + \delta_{11} Z_1(\boldsymbol{x}), & (1) \\ Y(2, \boldsymbol{x}) &= \delta_{02} Z_0(\boldsymbol{x}) + \delta_{22} Z_2(\boldsymbol{x}), & (2) \\ Y(3, \boldsymbol{x}) &= \delta_{03} Z_0(\boldsymbol{x}) + \delta_{33} Z_3(\boldsymbol{x}). & (3) \end{aligned}$$

Approach 2

$Z_0(\cdot)$, $Z_1(\cdot)$, $Z_2(\cdot)$, and $Z_3(\cdot)$ are independent Gaussian Stochastic Processes.

$$Z_i(\cdot) \sim N(\beta_i, Cov(\cdot, \cdot)), \tag{4}$$

$$Cov(Z_i(\boldsymbol{x}_1), Z_i(\boldsymbol{x}_2)) = \sigma_i^2 exp\{-\sum_{j=1}^{d} \theta_{i,j}(x_{1,j} - x_{2,j})^2\}. \tag{5}$$

Approach 2

Critique of Special Case I:

1(+), $Y(1,\cdot)$, $Y(2,\cdot)$, and $Y(3,\cdot)$ are symmetric.

2(-), There are two many parameters. The model requires strong priors in $\boldsymbol{\delta}$ and $\boldsymbol{\sigma^2}$.

3(-), Z_0 is the only part for describing the correlations among all the processes, $Y(1,\cdot)$, $Y(2,\cdot)$, and $Y(3,\cdot)$.

Approach 2

Special Case II:

$$Y(1, \boldsymbol{x}) = \delta_{11} Z_1(\boldsymbol{x}) + \delta_{21} Z_2(\boldsymbol{x}) \quad (6)$$

$$Y(2, \boldsymbol{x}) = \delta_{22} Z_2(\boldsymbol{x}) + \delta_{32} Z_3(\boldsymbol{x}) \quad (7)$$

$$Y(3, \boldsymbol{x}) = \delta_{13} Z_1(\boldsymbol{x}) + \delta_{33} Z_3(\boldsymbol{x}) \quad (8)$$

Approach 2

Critique of Special Case II:

1(+), $Y(1,\cdot)$, $Y(2,\cdot)$, and $Y(3,\cdot)$ are symmetric.

2(-), There are two many parameters.

3(-), It is hard to expand this model when $T > 3$.

Approach 2

Special Case III:

We set $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$.

$$Y(1, \boldsymbol{x}) = \delta_{11} Z_1(\boldsymbol{x}) \tag{9}$$

$$Y(2, \boldsymbol{x}) = \delta_{12} Z_1(\boldsymbol{x}) + \delta_{22} Z_2(\boldsymbol{x}) \tag{10}$$

$$Y(3, \boldsymbol{x}) = \delta_{13} Z_1(\boldsymbol{x}) + \delta_{23} Z_2(\boldsymbol{x}) + \delta_{33} Z_3(\boldsymbol{x}) \tag{11}$$

Approach 2

Critique of Special Case III:

1(+), This model has the property that the number of the parameters δ equals to the number of the elements in the variance-covariance matrix we could control.

2(-), The variance of the process $Y(t, \boldsymbol{x})$ tends to be bigger as t increases. The model requires the processes corresponding to different t have orders.

Approach 2

Special Case IV:

We set $E(Z_i(\boldsymbol{x})) = 0$.

$$\begin{aligned} Y(1,\boldsymbol{x}) &= \beta_1 + Z_1(\boldsymbol{x}) \qquad (12) \\ Y(2,\boldsymbol{x}) &= \beta_2 + \delta_1 Z_1(\boldsymbol{x}) + Z_2(\boldsymbol{x}) \qquad (13) \\ Y(3,\boldsymbol{x}) &= \beta_3 + \delta_2(\delta_1 Z_1(\boldsymbol{x}) + Z_2(\boldsymbol{x})) + Z_3(\boldsymbol{x}) \qquad (14) \end{aligned}$$

Approach 2

Critique of Special Case IV:

1(+), By denoting the mean of each process $Y_i(t, \boldsymbol{x})$ by a parameter β_i, the model is parsimonious.

2(-), The model requires the processes corresponding to different t have orders.

A Bayesian Model

Idea of the model: The responses from different categories have ”similar” profiles in continuous variables.

Suppose the inputs are $(t, \boldsymbol{x})$, where $t \in \{1, 2, \ldots, T\}$ is qualitative and $\boldsymbol{x} \in [0, 1]^d$ is quantitative. The output is the real-valued function $y(t, \boldsymbol{x})$. The training data is $\boldsymbol{Y}^n = (Y(t_1, \boldsymbol{x}_1), Y(t_2, \boldsymbol{x}_2), \ldots, Y(t_n, \boldsymbol{x}_n))$.

A Hierarchical Qualitative Quantitative Variable (HQQV) Model

The parameters are $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_T)$, $\boldsymbol{\sigma^2} = (\sigma_1^2, \ldots, \sigma_T^2)$, $\boldsymbol{\rho} = (\boldsymbol{\rho}_1, ..., \boldsymbol{\rho}_T)$, and $\boldsymbol{\rho}_t = (\rho_{t,1}, \ldots, \rho_{t,d})$.

A Bayesian Model

Given these parameters,

$$Y(t,\boldsymbol{x}) = \beta_t + Z_t(\boldsymbol{x}). \tag{15}$$

$Z_1(\boldsymbol{x}), ..., Z_T(\boldsymbol{x})$ are *independent* stationary Gaussian Processes with mean 0 and

$$cov(Z_i(\boldsymbol{x}_1), Z_i(\boldsymbol{x}_2)|\boldsymbol{\rho}_i, \sigma_i^2) = \sigma_i^2 R(\boldsymbol{x}_1 - \boldsymbol{x}_2|\boldsymbol{\rho}_i), \tag{16}$$

where $\boldsymbol{h} = (h_1, ..., h_d)^T$ and $R(\boldsymbol{h}|\boldsymbol{\rho}_t) = \prod_{j=1}^{d} \rho_{tj}^{h_j^2}$.

$$\left[\boldsymbol{Y}^n|\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}\right] \sim N\left(\boldsymbol{F\beta}, \boldsymbol{\Sigma}_{yn}\right) \tag{17}$$

A Bayesian Model

- Parameter Model

$$[\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}] = [\boldsymbol{\beta}, \boldsymbol{\rho}][\boldsymbol{\sigma}^2] = [\boldsymbol{\beta}|\boldsymbol{\rho}][\boldsymbol{\sigma}^2][\boldsymbol{\rho}]. \tag{18}$$

  - $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_T$ are *independently* and *identically* distributed. For any $j \in \{1, \ldots, d\}$, $\rho_{1j}, \ldots, \rho_{Tj}$ are *independently and identically* distributed as $Beta(\alpha_j, \gamma_j)$. α_j and γ_j are chosen to be $\alpha_j = 1$, and $\gamma_j = 0.5$. ρ_{tj} has mean $2/3$ and variance 0.0889.
  - $\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2$ are *independently and identically* distributed. $\frac{1}{\sigma_i^2}$ is of $\Gamma(\alpha, \gamma)$, where α and γ are chosen to be $\alpha = 5$ and $\gamma = 0.2$. σ_i^2 has mean 1.25, mode 5/6, and variance 25/48.
  - $\beta_1, \ldots, \beta_T$ are *independently and identically* distributed with the standard uniform improper prior.

Notice that $Y(1, \boldsymbol{x}), \ldots, Y(T, \boldsymbol{x})$ in the HQQV model are conditionally independent, but unconditionally dependent on each other.

A Bayesian Predictor and Its Implementation

Goal: Predict $Y_0 = Y(t_0, \boldsymbol{x}_0)$ given $\boldsymbol{Y}^n$.

- The predictor

$$\begin{aligned} \widehat{Y}(i_0, \boldsymbol{x}_0) &= E\{Y(i_0, \boldsymbol{x}_0|\boldsymbol{Y}^n)\} \\ &= E\{E(Y(i_0, \boldsymbol{x}_0)|\boldsymbol{Y}^n, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\sigma}^2)\}. \end{aligned}$$

- Uncertainty of the prediction

$$\begin{aligned} Var(Y_0|\boldsymbol{Y}^n) &= var\{E(Y_0|\boldsymbol{Y}^n, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\sigma}^2)\} \\ &+ E\{var(Y_0|\boldsymbol{Y}^n, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\sigma}^2)\}. \end{aligned}$$

A Bayesian Predictor and Its Implementation

Idea We use the Monte Carlo Markov Chain method, the Metropolis Hastings Sampler, to draw the values of each parameter from the joint posterior distribution $[\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho} | \boldsymbol{Y}^n]$ and to predict $y(i_0, \boldsymbol{x}_0)$, the unknown output corresponding to the inputs $(i_0, \boldsymbol{x}_0)$.

A Bayesian Predictor and Its Implementation

In our case,

$$[\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho} | \boldsymbol{Y}^n] \propto [\boldsymbol{Y}^n | \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}][\boldsymbol{\beta}][\boldsymbol{\sigma}^2][\boldsymbol{\rho}]$$

and

$$\left[\begin{pmatrix} Y_0 \\ \boldsymbol{Y}^n \end{pmatrix} | \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho} \right] \sim N\left(\begin{pmatrix} u_0 \\ \boldsymbol{u}_n \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma_{y0}} & \boldsymbol{\Sigma_{0n}^T} \\ \boldsymbol{\Sigma_{0n}} & \boldsymbol{\Sigma_{yn}} \end{pmatrix} \right)$$

The conditional distribution is

$$[Y_0 | \boldsymbol{Y}^n, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\rho}] \sim N(u_{Y_0|\boldsymbol{Y}^n}, \boldsymbol{\Sigma}_{Y_0|\boldsymbol{Y}^n}),$$

where $u_{Y_0|\boldsymbol{Y}^n} = u_0 + \boldsymbol{\Sigma_{0n}^T} \boldsymbol{\Sigma_{yn}}^{-1} (\boldsymbol{Y}^n - \boldsymbol{u}_n)$, and $\boldsymbol{\Sigma_{Y_0|\boldsymbol{Y}^n}} = \boldsymbol{\Sigma_{y0}} - \boldsymbol{\Sigma_{0n}^T} \boldsymbol{\Sigma_{yn}}^{-1} \boldsymbol{\Sigma_{0n}}$.

Competing methods

1. A REML EBLUP model that fits the data in each category separately.
2. An autoregressive model proposed by Kennedy & O'Hagan(2000) (K&O AR). The response of the computer experiments $Z_{t+1}(\boldsymbol{x})$from code level $t+1$ can be modeled as

   $$Z_{t+1}(\boldsymbol{x}) = \rho_t \times Z_t(\boldsymbol{x}) + \delta_t(\boldsymbol{x}),$$

   where $\delta_t(\cdot) \sim N(0, c_t(\cdot,\cdot))$ is a stationary Gaussian Stochastic Process, and $c_t(x_1, x_2) = \sigma_t^2 \times e^{-b_i \times (x_1 - x_2)^2}$.
3. Fitting HQQV model for data in each category separately. We call this the Independence Bayesian model.

Example1

We describe first how simulation data are produced for comparison of the HQQV and alternative models. We generate a sequence of test data sets, each having a single qualitative variable with $T = 4$ categories and 2 continuous variables.

The parameters are $\beta_1, \ldots, \beta_4$, $\sigma_1^2, \sigma_2^2, \ldots, \sigma_4^2$, and $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_4$, where $\boldsymbol{\rho}_t = [\rho_{t1}\rho_{t2}]$, and ρ_{t1} and ρ_{t2} correspond to the first and the second continuous variable respectively.

Example1

Step 1 Set the parameters $\beta_1, \ldots, \beta_4$, $\sigma_1^2, \sigma_2^2, \ldots, \sigma_4^2$, and $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_4$ by generating them from prior distributions such that

- $\beta_1, \ldots, \beta_4$ are *independently* and *identically* distributed, with the standard non-informative prior.
- $\sigma_1^2, \ldots, \sigma_4^2$ are *independently* and *identically* distributed as Inverse $\Gamma(5, 0.2)$.
- $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_4$ are *independently* and *identically* distributed where, for any $j \in \{1, 2\}$, $\rho_{1j}, \ldots, \rho_{4j}$ are *independently* and *identically* distributed as $Beta(1, 0.5)$.

Example1

Step 2 Generate Test Surface

a)Make the two dimensional grid of (x_1, x_2) points by crossing $x_1 = [\frac{1}{8}, \frac{2}{8}, ..., 1]$ by $x_2 = [\frac{1}{8}, \frac{2}{8}, ..., 1]$. Generate $64(8 \times 8) y(\cdot)$ values at these grid points. As an example, the draws from surface j are from a 64-variate normal distribution with mean β_j and

$$Cov(Y(j, \boldsymbol{x}_{i1}), Y(j, \boldsymbol{x}_{i2})) = \sigma_j^2\, R(\boldsymbol{x}_{i1} - \boldsymbol{x}_{i2} | \boldsymbol{\rho}_j), \qquad (19)$$

and

$$R((h_1, \ldots, h_2) | \rho_1, \rho_2) = \prod_{k=1}^{2} \rho_k^{h_k^2} \quad (0 \le \rho_k \le 2). \qquad (20)$$

Example1

b) Use the Krigifier(Trosset, 1999)

$$Y(j, x_1, x_2) = \beta_j + r^T(x_1, x_2)R^{-1}(y^{64} - \beta_j \mathbf{1}_{64\times 1}) \qquad (21)$$

to interpolate these 64 values to $[0, 1]^2$. Generate 100 data points as true values from the krigifier. $x_1 \in 0.05, 0.15, \ldots, 0.95$ and $x_2 \in 0.05, 0.15, \ldots, 0.95$.

Example1

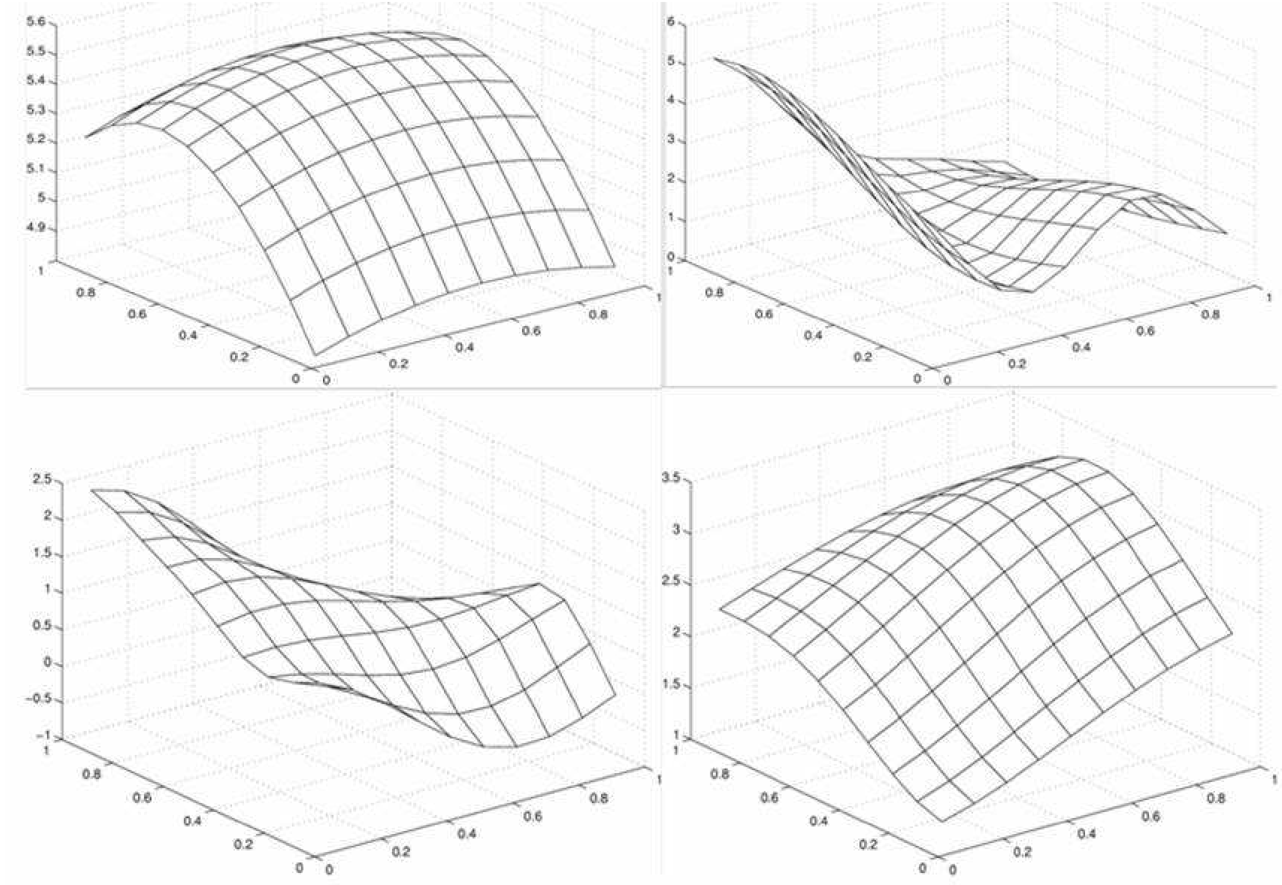

Figure 5: A draw of four surfaces from the $Y(t, \boldsymbol{x})$ Process of Example 1

Example1

This procedure ensures that the training data have several properties that satisfy the assumptions of the HQQV model.

First, all four surfaces are of similar shape because the parameters are drawn from a common distribution.

Second, we anticipate that the priors and posteriors will be similar because the likelihood function is made by the training data obtained from the assumed process.

Third, it is possible to assess the model reliability by calculating the prediction accuracy since multiple values can be drawn from formula (21) and compared with the predictions made by each statistical predictor.

Example1

Step 3 Generate Training Data

For each code level, select 8 points (x_1, x_2) in $[0, 1]^2$ according to a Maximin Latin Hypercube design (Johnson, Moore and Ylvisaker, 1990; Welch, 1985) and get their responses. We thus have a training dataset having $32(4 \times 8)$ values.

Step 4 Generate Evaluation Data

Predict $Y(t, \boldsymbol{x})$ at 100 ($x_1 \in 0.05, 0.15, \ldots, 0.95; x_2 \in 0.05, 0.15, \ldots, 0.95$) points for each surface using a statistical model and the training data in step 3.

Example1

Root Mean Squared Prediction Rrrors(RMSPE) over the 400 sample draws of 4 surfaces are shown in Table 1.

Example1

Models	HQQV Model	IB Model	REML EBLUP
Prediction Error 1	**0.1077**	0.1091	0.1151
Prediction Error 2	**0.1399**	0.1403	0.1462
Prediction Error 3	**0.1376**	0.1393	0.1933
Prediction Error 4	**0.1084**	0.1094	0.1122
Prediction Error 5	**0.1177**	0.1178	0.1250
Prediction Error 6	**0.1136**	0.1153	0.1226
Prediction Error 7	**0.1180**	0.1187	0.1330
Prediction Error 8	0.1354	**0.1285**	0.1399
Prediction Error 9	0.1364	**0.1330**	0.1417
Prediction Error 10	**0.1108**	0.1163	0.2139

Table 1: RMSPE for HQQV model, Independent Bayesian model, and Empirical EBLUP

Example1

Notice that the HQQV model has smaller prediction error than the other two models.

For problems as this example where the responses for different categories have similar trend/shape, modeling all the data together making more sense than doing that once for each category.

Example2

Suppose we have one qualitative and one quantitative input. The qualitative input identifies one of three quadratic curves, each a function of the quantitative input(x). We observe the first two curves at 21 equally spaced points from 0 to 20, and the third curve only on the points 10 to 20.

Goal: predict $y(3, x)$ over $x \in [0, 10]$.

Example2

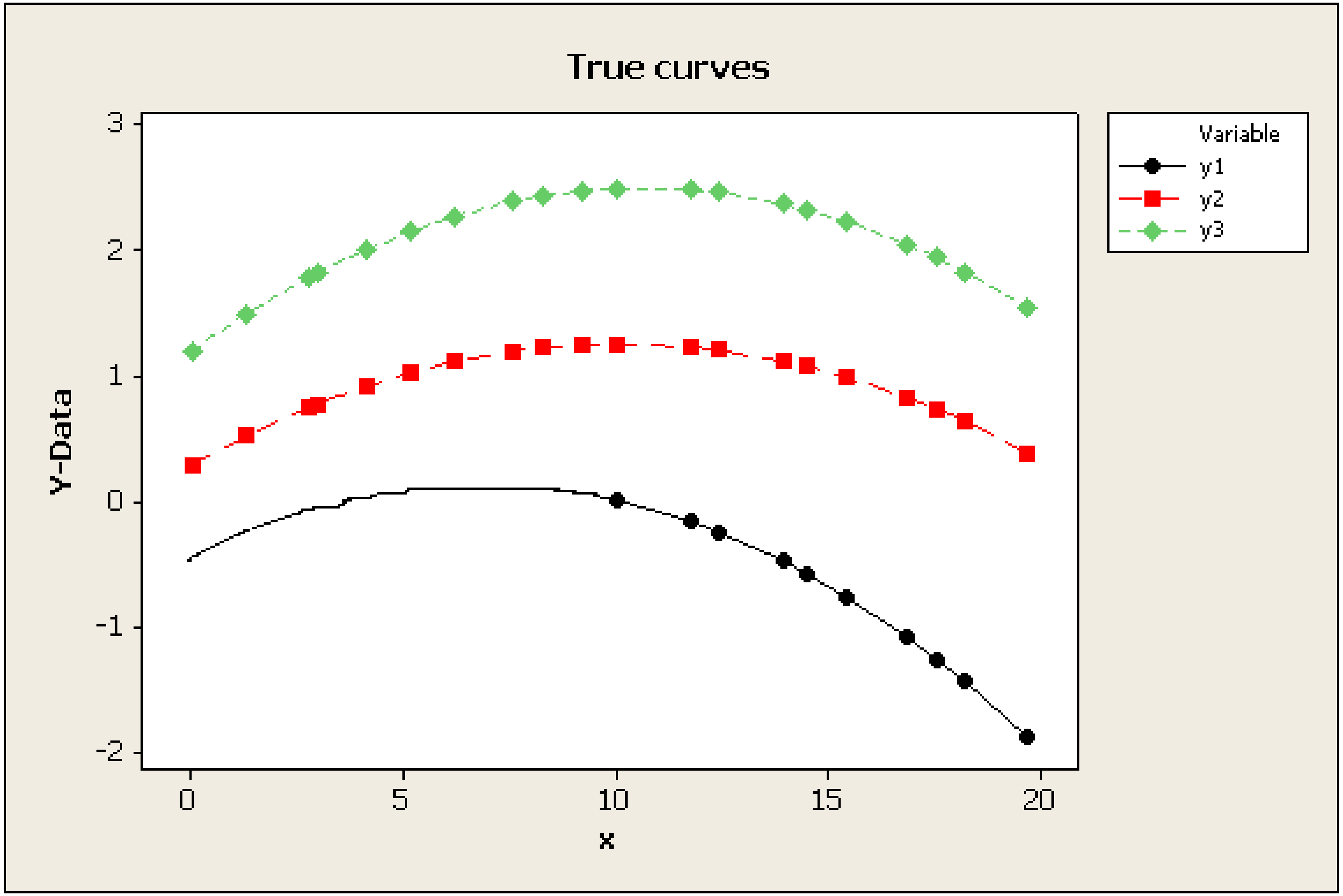

Figure 6: True Data

Example2

Goal: predict $y(3, x)$ over $x \in [0, 1]$.

Models	HQQV Model	K&O AR	IB	REML EBLUP
SSE	0.0463	??	0.4758	0.3185

Table 2: Root Mean Squared Prediction Errors from the HQQV model, K&O AR model, IB model and the REML EBLUP model

Models	HQQV Model	K&O AR	IB	REML EBLUP
SumVar	5.66	??	5.65	17.01

Table 3: Summed Prediction variance over 100 test points of the HQQV model, K&O AR model, IB model, and the REML EBLUP model

Conclusion and Future work

Conclusion

The HQQV model looks promising for the smaller prediction error than some competing methods when the responses from different code levels share certain properties.

Future work

- **Methodologies**
  - Extend the HQQV model to multiple qualitative inputs
  - Extend the HQQV model to functional outputs
- **Applications**
  - Design of Experiments for Calibration of computer codes having qualitative and quantitative inputs
  - Optimization for outputs having quantitative and qualitative inputs