Bayesian Approaches to the Analysis of Computer Model Output

Online Seminar Series May 20, 2014

Mark Berliner
Department of Statistics
Ohio State University
mb@stat.osu.edu

Collaborators: Radu Herbei, Ralph Milliff, Chris Wikle

Research Support: NSF, ONR
Outline

• Model error and data discussion: focus on space-time models
• Statistical models for data
• Approaches to modeling model error
  1) small to moderate scales
  2) massive scales
• Two examples of 1) and 2)
Model Error and Data

• Many sources of error:
  approximation versus true model;
  un-modeled effects;
  parameterizations; numerical error; ...

• **Observations are critical:**
  assess and moderate the impact of error;
  calibrate models;
  interpolation and short-horizon prediction
Statistical Models for Data

- Observational data: \( Y \)
- Process (state variable) of interest: \( X \)
- **Data model:** \( s = \text{spatial location}, t = \text{time} \)
  \[
  Y(s,t) = X(s,t) + e(s,t)
  \]
  where \( \{e\} \) are measurement errors
- Usually assume \( \{e\} \) all have expectation = 0
- In probability language
  \[
  Y(s,t) \mid X(s,t), \sigma^2 \sim \text{Gau}(X(s,t), \sigma^2)
  \]

For example
Approaches: 1) Small to moderate scales

• Process model:

Nature: \( G(\{ X, F, Q \}(s,t)) = 0 \)

is approximated by

\[ g(\{\tilde{X}, f, \theta, \eta \}(s,t)) = 0 \]

where \( \{\eta\}(s,t) \) is modeled error process

• \( \{\varepsilon\}(s,t) \) is true model error

\[ \{X\}(s,t) - \{\tilde{X}\}(s,t) = \{\varepsilon\}(s,t) \]
Approaches: 2) Massive scales

- HP computing models; climate system models
- Summary/dim. reduc.: \( X = S(\text{true}), \overline{X} = S(\text{Output}) \)
- Observational data model \( Y = X + e \)
- Computer model data model \( \overline{X} = X + b(\epsilon, \ldots) + \xi \)

where \( b(\epsilon) \) denotes bias and \( \xi \) denotes other sources of variation (e.g., ensemble “sampling”)

- Process model prior on \( X \)
- Notes:
  - Not putting \( \eta \) into the numerical model
  - \( b \) and \( \xi \) vary with \( S \)
  - need a model for \( b \)
A. Kennedy & O’Hagan

- Observational data model: for each observation
  \[ Y = X(c, \theta) + e \]
  where \( c \) are model inputs, \( \theta \) is unknown
- Computer model data model
  \[ X(c, \theta) = \tilde{X}(c, \theta) + \delta(c) \]
  where \( \delta(\cdot) \) denotes discrepancy (GRF)
- Hence
  \[ Y = \tilde{X}(c, \theta) + \delta(c) + e \]
- Bayesian inference for \( \theta \)
- Similar to Approach 2.

B. Special issues for emulators of large models

C. Dimension reduction (EOF, PCA, ...) bridge scales
## Rough Comparison of Probability Models

<table>
<thead>
<tr>
<th>Outline above</th>
<th>Kennedy &amp; O’Hagan</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbf{\cdot} \quad Y = X(\theta) + e \quad P(Y \mid X, \theta)$</td>
<td>$\mathbf{\cdot} \quad Y = X(\theta) + e \quad P(Y \mid X, \theta)$</td>
</tr>
<tr>
<td>$\mathbf{\cdot} \quad \tilde{X} = X + b + \xi \quad P(\tilde{X} \mid X,b,\theta)$</td>
<td>$\mathbf{\cdot} \quad X(\theta) = \tilde{x}(\theta) + \delta \quad P(X \mid \delta,\theta)$ [ t versus $\theta$ ]</td>
</tr>
<tr>
<td>$\diamond Prior \ P(X, b, \theta)$</td>
<td>$\diamond Prior \ P(\delta, \theta)$</td>
</tr>
<tr>
<td>$\diamond Posterior: \ P(X,b,\theta \mid Y, \tilde{X})$</td>
<td>$\diamond Posterior: \ P(\delta,\theta \mid Y,)$</td>
</tr>
</tbody>
</table>
Simulation study:
Box indicates domain; Dots indicate locations where data are available; Tracer X is oxygen concentration

True model

$$\frac{\partial X}{\partial t} = \frac{\partial}{\partial x} \left( \kappa^x \frac{\partial X}{\partial x} \right) + \frac{\partial}{\partial y} \left( \kappa^y \frac{\partial X}{\partial y} \right) - u \frac{\partial X}{\partial x} - v \frac{\partial X}{\partial y} - \lambda X$$

with spatially varying diffusivities

$$\kappa^x = ax + b \quad \text{and} \quad \kappa^y = cy^2 + dy + h$$
• Discretize model: 703 sites. Data simulated for 335 sites for $T = 12$ months, prediction to $T = 20$

• Generate $X$ and then generate data: $(x, y) = s$

$Y(s_i, t_j) = X(s_i, t_j) + e(s_i, t_j)$

• Process model prior

$$\frac{\partial \tilde{X}}{\partial t} = k^x \frac{\partial^2 \tilde{X}}{\partial x^2} + k^y \frac{\partial^2 \tilde{X}}{\partial y^2} - u \frac{\partial \tilde{X}}{\partial x} - v \frac{\partial \tilde{X}}{\partial y} - \lambda \tilde{X} + \eta$$

• Assessment: Average MSE

$$MSE(t) = \frac{1}{|\text{sites}|} \sum_{x, y} (X(x, y, t) - \tilde{X}(x, y, t))^2$$
E1: No model error;

E2: \( \eta = A \frac{\partial \hat{X}}{\partial x} + B \frac{\partial \hat{X}}{\partial y} \)

E3: \( \eta(t) = A(t) \frac{\partial \hat{X}}{\partial x} + B(t) \frac{\partial \hat{X}}{\partial y} ; \)

\( A, B \) - linear (quadratic) functions of \( t \);

E4: \( \eta(t) = A(t) \frac{\partial \hat{X}}{\partial x} + B(t) \frac{\partial \hat{X}}{\partial y} \)

\( A, B \) - AR(1) models;
E1: No model error;

E2: $\eta = A \frac{\partial \tilde{X}}{\partial x} + B \frac{\partial \tilde{X}}{\partial y}$

E3: $\eta(t) = A(t) \frac{\partial \tilde{X}}{\partial x} + B(t) \frac{\partial \tilde{X}}{\partial y}$; $A, B$ - linear (quadratic) functions of $t$;

E4: $\eta(t) = A(t) \frac{\partial \tilde{X}}{\partial x} + B(t) \frac{\partial \tilde{X}}{\partial y}$; $A, B$ - AR(1) models;
(1 - \frac{MSE(t)}{MSE(t)}) \times 100

Not bad: E3, E4: modeled errors were temporal; actual error was spatial. E2: some spatial error.
Steady Flow of Glaciers and Ice Sheets

- Flow: gravity moderated by drag (base & sides) & stuff
- Simple models: flow from geometry

Data

- Program for Arctic Climate Regional Assessments
- Radarsat Antarctic Mapping Project
- Surface topography (laser altimetry)
- Basal topography (radar altimetry)
- Velocity data (interferometry)
North East Ice Stream, Greenland
Physical Modeling: Surface: s, Thickness: H, Velocity: u

- Basal Stress: \( \tau = -\rho g H \frac{ds}{dx} ( + \ "stuff") \)
- Velocities: \( u = u_b + b_0 H \tau^n \) where \( u_b = k \tau^p + (\rho g H)^{-q} \)

Our Model

- Random Basal Stress: \( \tau = -\rho g H \frac{ds}{dx} + \eta \)
  where \( \eta \) is a "corrector process"
Approaches: 2) Massive scales
Low dimensional summaries $S$ of true $X$ and model output $\tilde{X}$

- Observational data model
  $$Y = X + e$$

and computer model data model

  $$\tilde{X} = X + b(\varepsilon) + \xi$$

- Process model prior on $X$ and $b$
Climate Projection Example
Berliner and Kim (2008) *J Climate*

- $X_t$ hemispheric- & monthly-averaged surface temp's
- Two climate system models $\tilde{X}_t$: PCM (n=4), CCSM (n=1) for 2002-2197
- Keys
  1) Model j, ensemble k: $\tilde{X}_{tjk} = X_t + b_{j(t)} + \xi_{tjk}$
  2) Prior process model: For each hemisphere,
     \[
     X_t = \mu_t + \alpha_t (X_{t-1} - \mu_{t-1}) + \eta_t
     \]
     where parameters depend on CO$_2$ and SOI
  3) $b$'s: prior mean 0 and slowly varying in time
Summaries of Posterior Distribution
Mean of NH Temp (Climate)  NH Temp (Weather)
Ocean (the Med.) Modeling Example
Berliner et al. (2014)

• Process: profiles of temperature $X(z,t)$
• 16 vertical levels from 0m to 300m
• $t=1,...,60$ days
• Bayes wind-model gives ensembles of boundary conditions for massive ocean models:
  1) Ocean Parallelized (OPA)
  2) Nucleus for European Modeling of the Ocean (NEMO)
Observations

Initial – Boundary conditions

Bayes

BHM Winds

Model 1

Model 2

BHM Oceans

BHM Ocean Post. Dist.

Models

- **Model output data model:**

\[
\tilde{X}_{\text{OPA} j} = X + b_{\text{OPA}} + \xi_{j(O)} \quad j=1,\ldots,10
\]

\[
\tilde{X}_{\text{NEMO} j} = X + b_{\text{NEMO}} + \xi_{j(N)}
\]

- **Prior for X:**

Prior mean = Analysis field from a data assimilation

Prior covariance = Analysis error covariance matrix or ... (unavailable in this example so we developed something based on climatology).

- **Prior for b’s:** prior mean 0 and slowly varying in time.

- **More work:** covariances of \( \xi \)'s; ...
Thank You!