Spatial Analysis to Quantify Numerical Model Bias and Dependence: How Many Climate Models Are There?

Mikyoung Jun  
Department of Statistics  
Texas A&M University  
mjun@stat.tamu.edu

May 8, 2007

This is joint work with Reto Knutti and Doug Nychka.
Outline

Introduction
   Climate models and questions of interest
   Data

Climate model biases for Mean state and Trend

Correlation between model biases
   Mean State
   Trend

Eigen-analysis
Outline

Introduction
Climate models and questions of interest
Data

Climate model biases for Mean state and Trend

Correlation between model biases
Mean State
Trend

Eigen-analysis
Intergovernmental Panel on Climate Change (IPCC)

- Established in 1988 by two United Nations organizations, the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP)
- Goal: assess the "risk of human-induced climate change"
- Global climate model: numerical model that gives climate output on grids
- We have 20+ climate models developed by various organizations over the world
<table>
<thead>
<tr>
<th>Group</th>
<th>Country</th>
<th>IPCC I.D.</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Beijing Climate Center</td>
<td>China</td>
<td>BCC-CM1</td>
</tr>
<tr>
<td>2</td>
<td>Canadian Center for Climate Modelling &amp; Analysis</td>
<td>Canada</td>
<td>CGCM3.1</td>
</tr>
<tr>
<td>3</td>
<td>Météo-France/ Centre National de Recherches Météorologiques</td>
<td>France</td>
<td>CNRM-CM3</td>
</tr>
<tr>
<td>4</td>
<td>CSIRO Atmospheric Research</td>
<td>Australia</td>
<td>CSIRO-Mk3.0</td>
</tr>
<tr>
<td>5</td>
<td>US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory</td>
<td>USA</td>
<td>GFDL-CM2.0</td>
</tr>
<tr>
<td>6</td>
<td>US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory</td>
<td>USA</td>
<td>GFDL-CM2.1</td>
</tr>
<tr>
<td>7</td>
<td>NASA/Goddard Institute for Space Studies</td>
<td>USA</td>
<td>GISS-AOM</td>
</tr>
<tr>
<td>8</td>
<td>NASA/Goddard Institute for Space Studies</td>
<td>USA</td>
<td>GISS-EH</td>
</tr>
<tr>
<td>9</td>
<td>NASA/Goddard Institute for Space Studies</td>
<td>USA</td>
<td>GISS-ER</td>
</tr>
<tr>
<td>10</td>
<td>LASG/Institute of Atmospheric Physics</td>
<td>China</td>
<td>FGOALS-g1.0</td>
</tr>
<tr>
<td>11</td>
<td>Institute for Numerical Mathematics</td>
<td>Russia</td>
<td>INM-CM3.0</td>
</tr>
<tr>
<td>12</td>
<td>Institut Pierre Simon Laplace</td>
<td>France</td>
<td>IPSL-CM4</td>
</tr>
<tr>
<td>13</td>
<td>Center for Climate System Research, National Institute of Environmental Studies, and Frontier Research Center for Global Change</td>
<td>Japan</td>
<td>MIROC3.2 (medres)</td>
</tr>
<tr>
<td>14</td>
<td>Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA, and Model and Data group</td>
<td>Germany/Korea</td>
<td>ECHO-G</td>
</tr>
<tr>
<td>15</td>
<td>Max Planck Institute for Meteorology</td>
<td>Germany</td>
<td>ECHAM5/MPI-OM</td>
</tr>
<tr>
<td>16</td>
<td>Meteorological Research Institute</td>
<td>Japan</td>
<td>MRI-CGCM2.3.2</td>
</tr>
<tr>
<td>17</td>
<td>National Center for Atmospheric Research</td>
<td>USA</td>
<td>CCSM3</td>
</tr>
<tr>
<td>18</td>
<td>National Center for Atmospheric Research</td>
<td>USA</td>
<td>PCM</td>
</tr>
<tr>
<td>19</td>
<td>Hadley Centre for Climate Prediction and Research/ Met Office</td>
<td>UK</td>
<td>UKMO-HadCM3</td>
</tr>
<tr>
<td>20</td>
<td>Hadley Centre for Climate Prediction and Research/ Met Office</td>
<td>UK</td>
<td>UKMO-HadGEM1</td>
</tr>
</tbody>
</table>
Intergovernmental Panel on Climate Change (IPCC)

Global Warming Projections

- CCSR/NIES
- CCCma
- CSIRO
- Hadley Centre
- GFDL
- MPIM
- NCAR PCM
- NCAR CSM

Temperature Anomaly (°C)

1900  1950  2000  2050  2100
Questions of interest

- How can we quantify the model bias?
- Relationship between bias in the mean state and bias in the trend?
- Are the climate model outputs random samples from a distribution symmetric around true climate?
- Especially, are the model biases correlated?
  ⇒ How many models do we really have?
Observation vs model output

- Monthly averages in 1970-1999 (unit: °C) of surface temperature
- Latitude range: 45° S - 72° N
- Observations: combined data set from CRU (Climate Research Unit, East Anglia) and the Hadley Centre (UK), a few missing data
- IPCC model outputs: 19 models, no missing data
- Discrepancy of grids: bilinear interpolation of the model outputs to the observation grid
- Naive imputation for the missing observations
Outline

Introduction
  Climate models and questions of interest
  Data

Climate model biases for Mean state and Trend

Correlation between model biases
  Mean State
  Trend

Eigen-analysis
Biases for Mean state

- $X(s, t)$: DJF (or JJA) averages of the observation at location $s$ and year $t$
- $Y_i(s, t)$: $i^{th}$ model output
- Difference: $D_i(s, t) = X(s, t) - Y_i(s, t)$
- $D_i(s, t) = b_i(s) + u_i(s, t)$
  - $b_i$ is the bias for the mean state
- Model $b_i$ based on $\bar{D}_i(s) = \frac{1}{30} \sum_{t=1}^{30} D_i(s, t)$
Averaged bias and RMS

Obs – average of 19 models (DJF)

RMS error (DJF)

Obs – average of 19 models (JJA)

RMS error (JJA)
Spatial model for the bias

- $b_i(s) = \mu_{0i} + \mu_{1i}L(s) + \mu_{2i}\mathbf{1}_{s \in \text{land}} + \mu_{3i}A(s) + a_i(s)$
- $a_i$: mean zero Gaussian random field
- Covariance structure of $a_i$:
  simple nonstationary covariance model valid on sphere
  - $a_i(s) = (\delta_i\mathbf{1}_{s \in \text{land}} + 1)Z_i(s)$
    ($\delta_i > 0$)
  - $\text{Cov}\{Z_i(s_1), Z_i(s_2)\} = \alpha_i M_{\nu_i+1}(d/\beta_i)$
    ($d$: chordal distance between $s_1$ and $s_2$)
  - Could apply more complex model for Nonstationarity (Jun and Stein 2007, Technometrics, To appear):
    - eg) $a_i(s) = \{\eta_i(L) \frac{\partial}{\partial L} + \psi_i(L) \frac{\partial}{\partial l}\}Z_i(s) + d_iZ_i(s)$. 
Biases for Trend

- $X(s, t)$: seasonal averages (DJF or JJA) of observation at location $s$ and year $t$
- $Y_i(s, t)$: $i^{th}$ model output
- Regress $X$ and $Y_i$ on $t - \bar{t}$, separately
- Bias is defined as the difference between slope from $X$ and slope from $Y_i$ (Trend 1)
- A lot of “noise” in the observation slope: smooth both observation and model output beforehand (Trend 2)
- Take spatial averages of observation and model output and then calculate the absolute difference between observation slope and model slope (Trend 3)
Biases for Mean state vs Trend

DJF (mean) JJA (mean)

DJF (trend 1) JJA (trend 1)

DJF (trend 2) JJA (trend 2)

DJF (trend 3) JJA (trend 3)

0.73 -0.34 0.73 0.53 0.7
0.27 -0.21 0.19 0.07 0.19
Outline

Introduction
  Climate models and questions of interest
  Data

Climate model biases for Mean state and Trend

Correlation between model biases
  Mean State
  Trend

Eigen-analysis
Quantify cross-correlation

\[ \sigma_{ij}(s) = \text{Cov}\{a_i(s), a_j(s)\} \]

\[ r_{ij}(s) = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}(s) \]

: correlation between \(i^{th}\) and \(j^{th}\) model biases at location \(s\)

Apply Gaussian kernel to \(\tilde{D}_{ij}(s) = \tilde{D}_i(s)\tilde{D}_j(s)\), where \(\tilde{D}_i(s)\) is “filtered \(\bar{D}_i(s)\)”

\[ \hat{\sigma}_{ij}(s) = \sum_{k=1}^{1656} K\left(\frac{|s, s_k|}{h}\right) \tilde{D}_{ij}(s_k) \cdot \left[ \sum_{k=1}^{1656} K\left(\frac{|s, s_k|}{h}\right) \right]^{-1} \]
Estimated correlation between biases

Figure gives $r_{5j}, j = 2, \ldots, 20$ (DJF, mean state)
Estimated correlation between biases

Figure gives $r_{ij}$ averaged over space (mean state)
Estimated correlation between biases

Figure gives $r_{ij}$ averaged over space (trend)
Issues with correlation for Trend

- Four ensemble members of model 2 (1,2,3,4)
- Two ensemble members of model 5 (5,6)
- Two ensemble members of model 6 (7,8)
Introduction
   Climate models and questions of interest
   Data

Climate model biases for Mean state and Trend

Correlation between model biases
   Mean State
   Trend

Eigen-analysis
How many eigenvalues needed?
First top 5 eigen values (mean state)
Multidimensional scaling

- With correlation matrices as similarity measure, we can form subgroups of climate models.