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

Outline

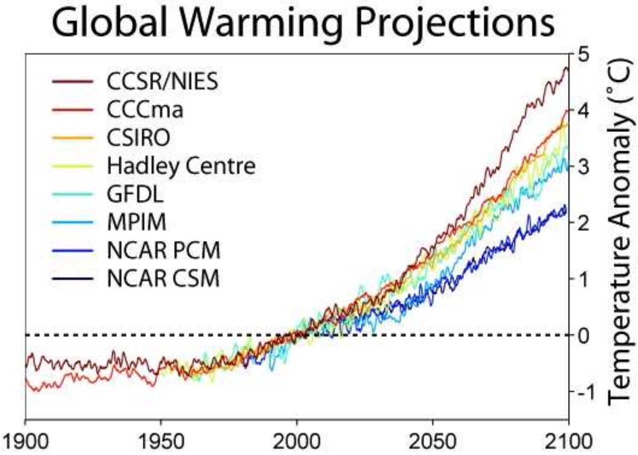
Intergovernmental Panel on Climate Change (IPCC)

- 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

Intergovernmental Panel on Climate Change (IPCC)

	Group	Country	IPCC I.D.	Resolution
1	Beijing Climate Center	China	BCC-CM1	192×96
2	Canadian Center for Climate Modelling & Analysis	Canada	CGCM3.1	96×48
3	Météo-France/ Centre National de Recherches Météorologiques	France	CNRM-CM3	128×64
4	CSIRO Atmospheric Research	Australia	CSIRO-Mk3.0	192×96
5	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory	USA	GFDL-CM2.0	144×90
6	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory	USA	GFDL-CM2.1	144×90
7	NASA/Goddard Institute for Space Studies	USA	GISS-AOM	90×60
8	NASA/Goddard Institute for Space Studies	USA	GISS-EH	72×46
9	NASA/Goddard Institute for Space Studies	USA	GISS-ER	72×46
10	LASG/Institute of Atmospheric Physics	China	FGOALS-g1.0	128×60
11	Institute for Numerical Mathematics	Russia	INM-CM3.0	72×45
12	Institut Pierre Simon Laplace	France	IPSL-CM4	96×72
13	Center for Climate System Research, National Institute of Environmental Studies, and Frontier Research Center for Global Change	Japan	MIROC3.2 (medres)	128×64
14	Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA, and Model and Data group	Germany/ Korea	ECHO-G	96×48
15	Max Planck Institute for Meteorology	Germany	ECHAM5/MPI-OM	192×96
16	Meteorological Research Institute	Japan	MRI-CGCM2.3.2	128×64
17	National Center for Atmospheric Research	USA	CCSM3	256×128
18	National Center for Atmospheric Research	USA	PCM	128×64
19	Hadley Centre for Climate Prediction and Research/ Met Office	UK	UKMO-HadCM3	95×73
20	Hadley Centre for Climate Prediction and Research/ Met Office	UK	UKMO-HadGEM1	192×145

Intergovernmental Panel on Climate Change (IPCC)



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?
  - $\Rightarrow$ How many models do we really have?

Observation vs model output

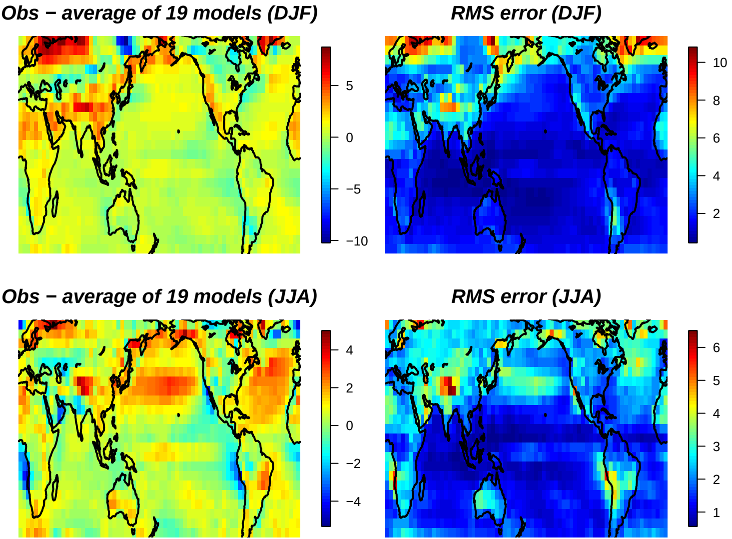
- Monthly averages in 1970-1999 (unit: $^\circ$ C) of surface temperature
- Latitude range: 45° S - 72° N
- Observations: combined data set from CRU (Climate Research Unit, East Anglia) and the Hadley Centre (UK), a few missing data
- IPCC model outputs: 19 models, no missing data
- Discrepancy of grids: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

Outline

Biases for Mean state

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

Averaged bias and RMS



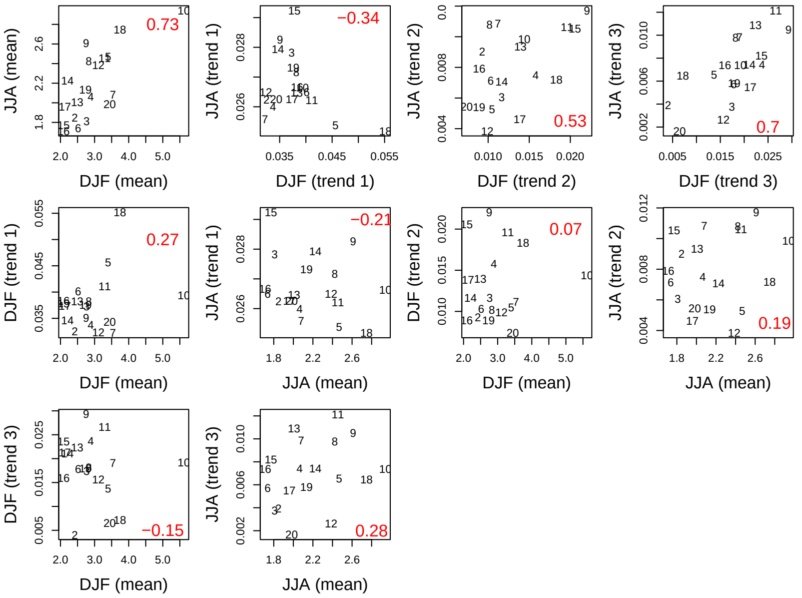
Spatial model for the bias

- $b_i(\mathbf{s}) = \mu_{0i} + \mu_{1i}L(\mathbf{s}) + \mu_{2i}\mathbf{1}_{(\mathbf{s}\in\text{land})} + \mu_{3i}A(\mathbf{s}) \quad + \quad a_i(\mathbf{s})$
- a_i: mean zero Gaussian random field
- Covariance structure of a_i:
  simple nonstationary covariance model valid on sphere
  - $$a_i(\mathbf{s}) = (\delta_i \mathbf{1}_{(\mathbf{s}\in\text{land})} + 1)Z_i(\mathbf{s})$$
    ($\delta_i > 0$)
  - $\text{Cov}\{Z_i(\mathbf{s}_1), Z_i(\mathbf{s}_2)\} = \alpha_i \mathcal{M}_{\nu_i+1}(d/\beta_i)$
    (d: chordal distance between $\mathbf{s}_1$ and $\mathbf{s}_2$)
  - Could apply more complex model for Nonstationarity (Jun and Stein 2007, *Technometrics*, To appear):
    $$\text{eg)}\ \ a_i(\mathbf{s}) = \{\eta_i(L)\frac{\partial}{\partial L} + \psi_i(L)\frac{\partial}{\partial l}\}Z_i(\mathbf{s}) + d_i Z_i(\mathbf{s}).$$

Biases for Trend

- $X(\mathbf{s}, t)$: seasonal averages (DJF or JJA) of observation at location **s** and year t
- $Y_i(\mathbf{s}, t)$: 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



Outline

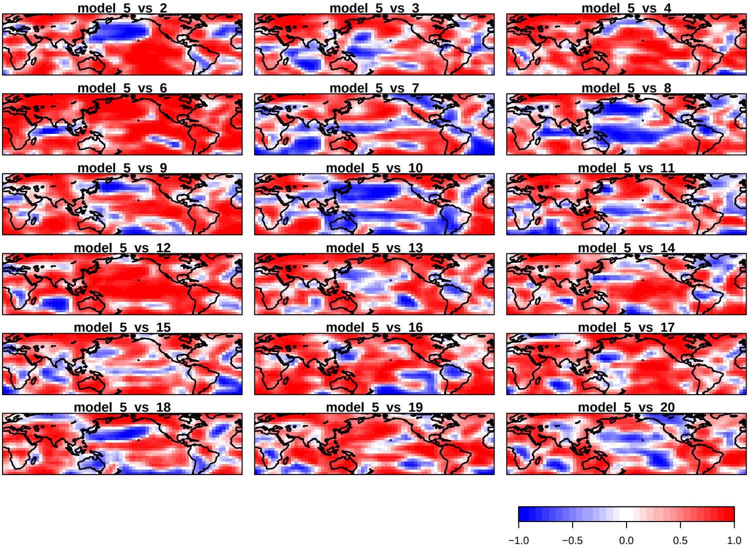
Quantify cross-correlation

- $\sigma_{ij}(\mathbf{s})$=Cov$\{a_i(\mathbf{s}), a_j(\mathbf{s})\}$
- $r_{ij}(\mathbf{s}) = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}(\mathbf{s})$
  : correlation between i^{th} and j^{th} model biases at location $\mathbf{s}$
- Apply Gaussian kernel to $\tilde{D}_{ij}(\mathbf{s}) = \tilde{D}_i(\mathbf{s})\tilde{D}_j(\mathbf{s})$, where $\tilde{D}_i(\mathbf{s})$ is "filtered $\bar{D}_i(\mathbf{s})$"
-

$$\hat{\sigma}_{ij}(\mathbf{s}) = \sum_{k=1}^{1656} K\left(\frac{|\mathbf{s}, \mathbf{s}_k|}{h}\right)\tilde{\tilde{D}}_{ij}(\mathbf{s}_k) \cdot \left[\sum_{k=1}^{1656} K\left(\frac{|\mathbf{s}, \mathbf{s}_k|}{h}\right)\right]^{-1}$$

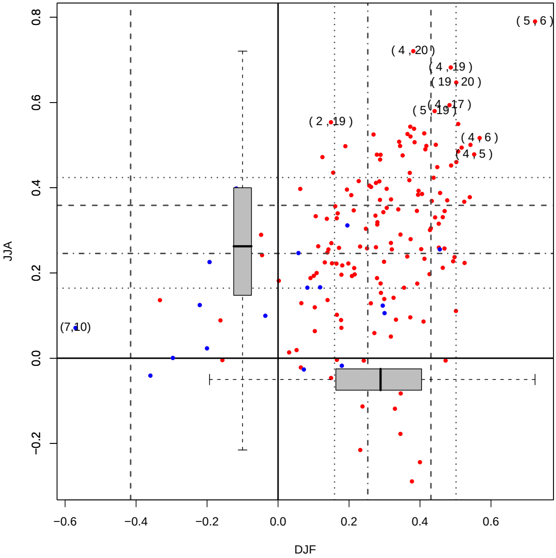
Estimated correlation between biases

Figure gives $r_{5j}, j = 2, \cdots, 20$ (DJF, mean state)



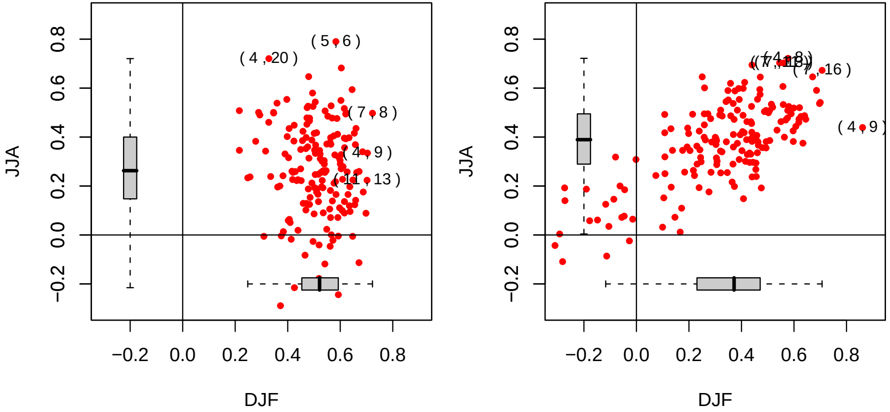
Estimated correlation between biases

Figure gives r_{ij} averaged over space (mean state)



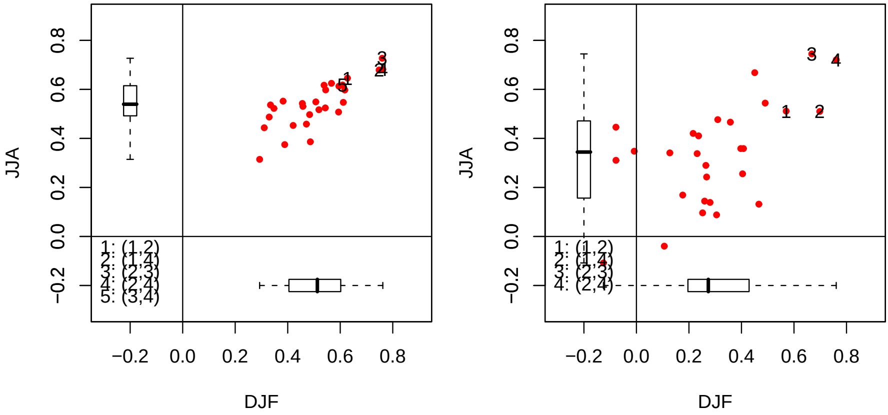
Estimated correlation between biases

Figure gives r_{ij} averaged over space (trend)



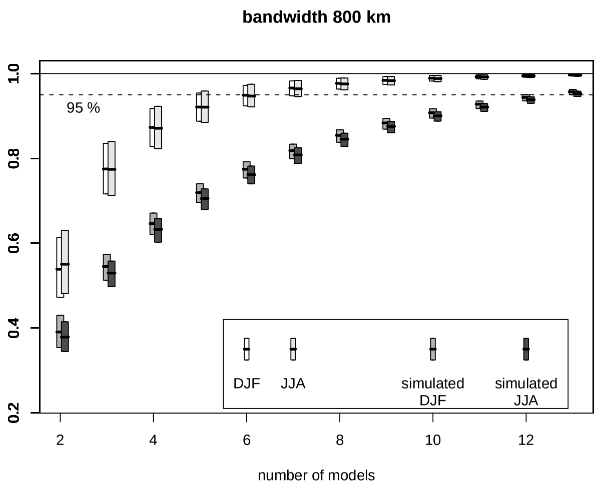
Issues with correlation for Trend

- 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)

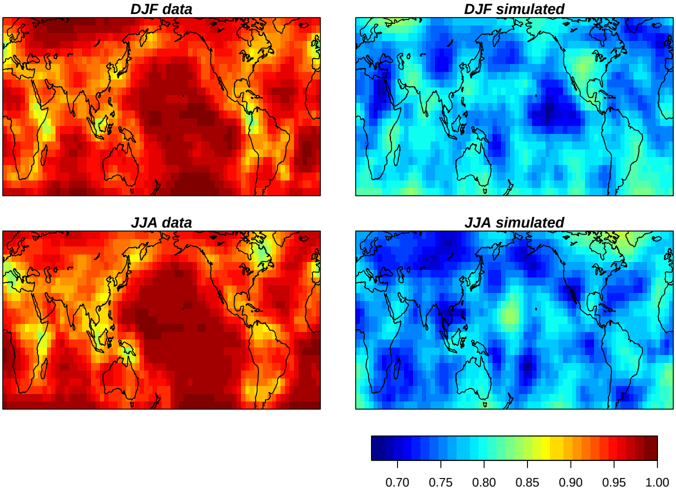


Outline

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