Part I:

Multivariate Bayesian Analysis of Atmosphere-Ocean General Circulation Models

GSP-NCAR, May 2nd 2005

Bayes Model for Climate Projections

Synthesizing temperature and precipitation climate projections from the outputs of several AOGCMs' with a hierarchical Bayesian approach.

In collaboration with: Stephan Sain - CU at Denver
Claudia Tebaldi - NCAR
Jerry Meehl - NCAR
Linda Mearns - NCAR
Tom Wigley - NCAR
Doug Nychka - NCAR

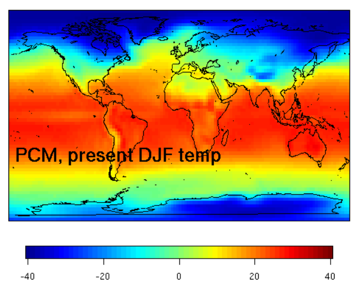
Study Climate with AOGCMs

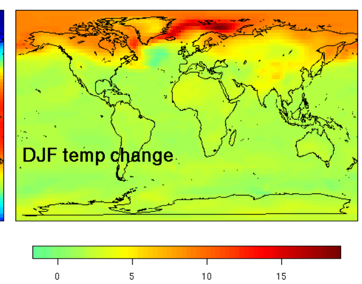
AOGCM: Atmosphere-Ocean General Circulation Models

Numerical models that calculate the detailed large-scale motions of the atmosphere and the ocean explicitly from hydrodynamical equations.

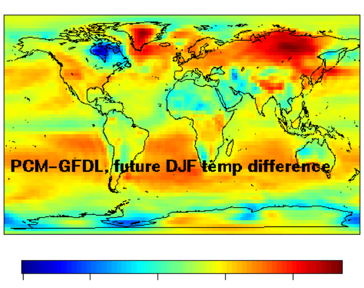
Study Climate with AOGCMs

AOGCM: Atmosphere-Ocean General Circulation Models





Models Do Not Agree



Example: Atmospheric Model

Input
- External forcings (radiation, volcanos, . . .)
- Anthropogenic forcings (GHG, aerosols, . . .)
- Initial conditions

↓

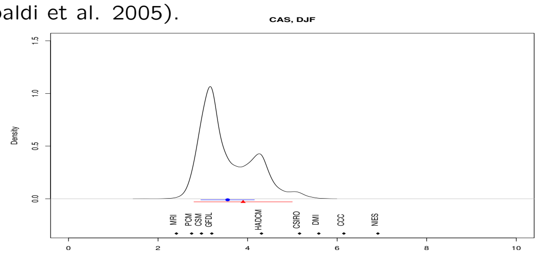
- Flow dynamics, PDEs
- Large scale discretization and simplifications
- Parametrization

↓

Output
- "Mean" temperatures and precipitation
- Pressure, wind, . . .

Models Do Not Agree

- Variability of global temperature increase across 16 models. MAGICC/SCENGEN program (Wigley, 2003).
- Probabilistic description of regional climate changes. (Tebaldi et al. 2005).



Data

Data provided for the Fourth Assessment Report of IPCC:

- Several models (CNRM, CSIRO, GFDL, GISSR, HADCM, INMCM, IPSL, MIROC, MPI, MRI, PCM, . . .)
- $2.8^\circ \times 2.8^\circ$ resolution (8192 data points)
- A2 scenario (“business as usual”)
- Temperature, . . . , monthly and seasonal averages over years 1961–1990 and 2070–2099
- NCEP reanalysis as “observations”

Statistical Model

For models $i = 1, \dots, N$, stack the gridded output into vectors:

$$\mathbf{X}_i = \text{simulated present climate}_i$$
$$\mathbf{Y}_i = \text{simulated future climate}_i$$

Objective:

Probabilistic description of modeled climate change

Statistical Model

Data level:

$$\begin{aligned} \mathbf{D}_i = \mathbf{Y}_i - \mathbf{X}_i &= \text{simulated climate change} \\ = \mathbf{Y}_i - \mathbf{X}_i &= \text{large scale structure} + \text{small scale structure} \\ = \mathbf{Y}_i - \mathbf{X}_i &= \mathbf{M}\boldsymbol{\theta}_i + \varepsilon_i, \qquad \text{with } \varepsilon_i \sim \mathcal{N}(\mathbf{0}, \phi_i \boldsymbol{\Sigma}) \end{aligned}$$

Process level:

$$\boldsymbol{\theta}_i \sim \mathcal{N}(\boldsymbol{\mu}, \psi_i \mathbf{I})$$

Prior level:

$$\boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}) \qquad \mathbf{M}$$
$$\phi_i \sim \mathrm{I}\Gamma(\xi_1, \xi_2) \qquad \boldsymbol{\Sigma}$$
$$\psi_i \sim \mathrm{I}\Gamma(\xi_3, \xi_4)$$

Basis Functions

We need a “rich” truncated basis set.

Current candidate:

- Harmonic functions on the sphere
- Indicators for continents, sea ice, . . .
- Patterns of current climate from NCEP reanalysis

Covariance Matrices

Examples of positive definite functions on the sphere:

1. representation with an infinite series of Legendre polynoms
$$c(h; \sigma, \tau) = \sigma\Big(1 - 2\tau\cos(h) + \tau^2\Big)^{-3/2}$$
2. restriction of a positive definite function on $\mathbb{R}^3$ to the sphere
$$c(h; \sigma, \tau) = \sigma\exp\Big(-\tau\sin(h/2)\Big)$$

We only parameterize the scale σ of the covariance matrices.

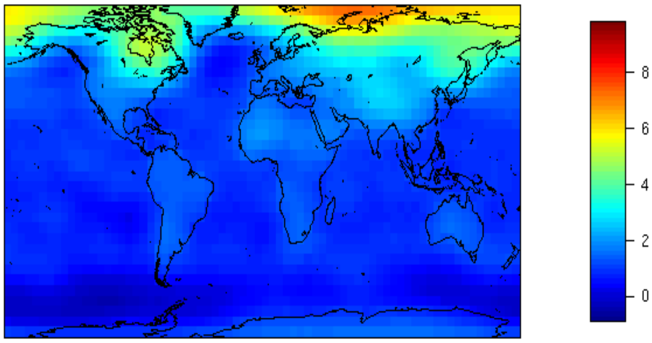
The “range” τ is choosen according an “empirical Bayes” approach.

Gibbs Sampler

- Full conditionals for the parameters are available
- Gibbs sampler programmed in R
- Run 10000 iterations
  (5000 burn-in, keep every 10th, takes ≈ 2 hours)
- Assessing convergence with:
  trace plots, different starting values, . . .

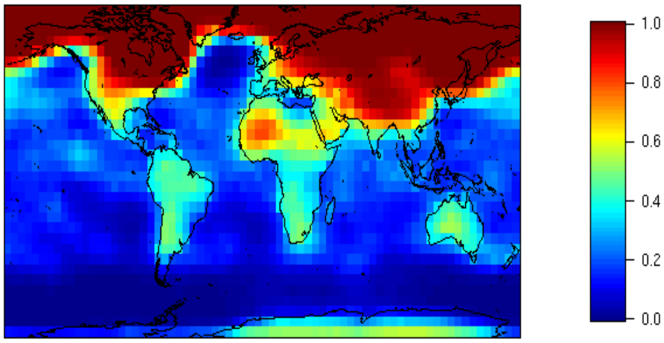
Posterior Climate Change

Posterior 20% quantile of temperature change

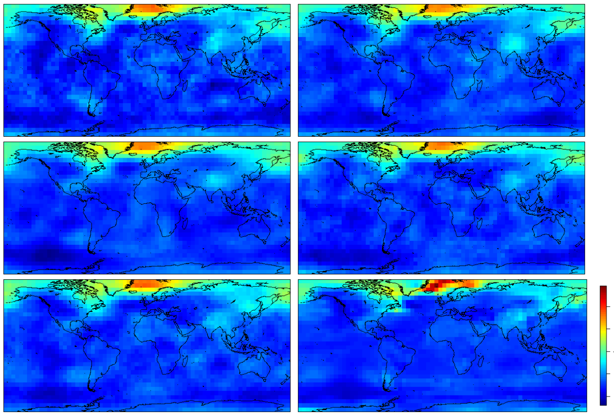


Posterior Climate Change

Probability that we observe at least a 2°K temperature increase



Posterior Model Realisations



Discussion and Further Work

- Promising approach (statistically and climatologically)
- Improve model for precipitation
- Generalize covariance parameterization
- Use “current” climate for better priors
- Extend to multivariate setting
- Implement ensemble runs

Part II:

KriSp:

a Package for Interpolation of Large Datasets Using Covariance Tapering

May 2nd, GSP NCAR

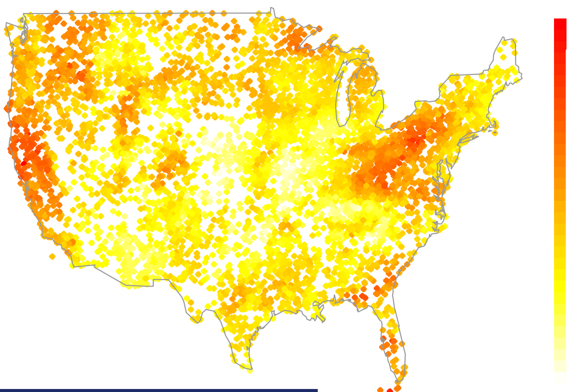
Tapering and Kriging

Introduce a sparseness structure in the covariance via tapering to gain computational advantages in large kriging problems constraint to maintaining asymptotic optimality.

In collaboration with Marc Genton and Doug Nychka.

Motivation

Precipitation anomaly in April 1948



Best Linear Unbiased Predictor

Suppose a spatial process of the form

$$Z(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}, \qquad \mathsf{E}(Z) = \mathbf{0}, \quad \mathsf{Cov}(Z) = \Sigma$$

with observations $\mathbf{Z} = \big(Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)\big)^\mathsf{T}$, $\mathbf{x}_i \in \mathcal{D} \subset \mathbb{R}^d$.

The kriging estimator (BLUP) is

$$\widehat{Z}(\mathbf{x}_0) = \mathbf{c}^\mathsf{T} \Sigma^{-1} \mathbf{Z}$$

with $\mathbf{c}_i = \mathsf{Cov}\big(Z(\mathbf{x}_0), Z(\mathbf{x}_i)\big)$.

Motivation

Prediction at a large number of points involves either

- solving one large linear system,
- solving many tiny linear systems

or, being smarter and

- "approximating" the covariance and solving "efficiently" one large but sparse linear system.

Synopsis

Completed:

- Submitted paper
- Beta-version of package with comprehensive tutorial

In progress:

- Apply to other statistical problems, like microarrays

In future:

- Nonstationarity, MLE, . . .