Bayesian Modeling and Computation in Complex Geophysical Problems

Mark Berliner
Ohio State University

NCAR Workshop: Petascale Computing 5-7 May, 2008

Outline

- Bayesian Modeling: Selected Features
- Computation: Monte Carlo (MC)
  1. Basic MC
  2. Importance Sampling, Particle Filters
  3. Markov Chain MC
- Examples of Multiscale Models
- Discussion

Bayesian Modeling: Selected Features

- **Bayesian Analysis: treating uncertainty and knowledge**
  - Combine observations & other information sources formally
  - Uncertainty management is paramount
  - Inputs and outputs are probability distributions
- **Mechanism: probability theory**
- **Challenges: (I) formulation of models; (II) computation.**
- **Two general arenas**
  1. Stochastic Dynamic Modeling: developing probability models for a complex system (within & across space-time scales; coarse graining; stochastic parameters & parameterizations)
  2. Forecasting: learning about and predicting an unobserved trajectory of a dynamical system (NWP; data assimilation)

Modeling device: Bayesian Hierarchical Models (BHM)

- **HM: Sequences of conditional probability models**

$$\mathbf{p}(\mathbf{x},\mathbf{y},\mathbf{z}) = \mathbf{p}(\mathbf{x} \mid \mathbf{y},\mathbf{z})\, \mathbf{p}(\mathbf{y} \mid \mathbf{z})\, \mathbf{p}(\mathbf{z})$$

- **Skeleton BHM, Observations y; Processes x; Parameters $\boldsymbol{\theta}$**

  1. **Data Model** $\quad \mathbf{q}(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta})$
  2. **Prior Process Model** $\quad \mathbf{p}(\mathbf{x} \mid \boldsymbol{\theta})$
  3. **Prior Parameter Model** $\quad \mathbf{p}(\boldsymbol{\theta})$

- **Bayes' Theorem gives Posterior Distribution:**

$$\begin{aligned}
\mathbf{p}(\mathbf{x},\boldsymbol{\theta} \mid \mathbf{y}) &\propto \mathbf{q}(\mathbf{y} \mid \mathbf{x},\boldsymbol{\theta})\mathbf{p}(\mathbf{x} \mid \boldsymbol{\theta})\mathbf{p}(\boldsymbol{\theta}) \\
&= \mathbf{q}(\mathbf{y} \mid \mathbf{x},\boldsymbol{\theta})\mathbf{p}(\mathbf{x} \mid \boldsymbol{\theta})\mathbf{p}(\boldsymbol{\theta})/\mathbf{q}(\mathbf{y})
\end{aligned}$$

$$\textbf{where} \quad \mathbf{q}(\mathbf{y}) = \int \mathbf{q}(\mathbf{y} \mid \mathbf{x},\boldsymbol{\theta})\mathbf{p}(\mathbf{x} \mid \boldsymbol{\theta})\mathbf{p}(\boldsymbol{\theta})\mathbf{d}\mathbf{x}\mathbf{d}\boldsymbol{\theta}$$

What Does This Buy Us?

- Combining information: "Physical-statistical modeling" (Berliner, JGR, 2003)
- Quantifying and dealing with uncertainty!!
- SGS parameterization? (ECMWF & "stoch-physics")

A. Operational impact of chaos: treat things as random.

B. From a deterministic physical model, $\mathcal{D}(\mathbf{x}, \boldsymbol{\theta}_{\mathrm{m}}) = 0$
to a stochastic model $\mathbf{p}(\mathbf{X} \mid \boldsymbol{\theta})$

- "Approximate physics ($\mathcal{D}$) applied approximately (discretize $\mathcal{D}$) and unsurely ($\boldsymbol{\theta}$; forcings unknown)"
- (Berliner, Milliff, Wikle, 2003, JGR) Air-sea interaction:

$$(\nabla^2 - \frac{1}{r^2})\frac{\partial \psi}{\partial t} = -\mathbf{J}(\psi, \nabla^2\psi) - \beta\frac{\partial \psi}{\partial x} + \frac{1}{\rho H}\mathrm{curl}_z \boldsymbol{\tau}(\mathbf{W}) - \gamma\nabla^2\psi + a_h\nabla^4\psi.$$

I see $\mathbf{p}(\psi_{t+1} | \psi_t, \boldsymbol{\theta}, \text{winds, boundary \& initial con.})$

C. Parameterization: Physical variables $\mathbf{X}, \mathbf{Z} = (\mathbf{Z_r}, \mathbf{Z_u})$

- **Discrete Time Physical Model:**
  - $\mathbf{X_{t+1}} = \mathbf{h}(\mathbf{X_t}, \mathbf{Z_{t+1}})$
  - $\mathbf{Z_{t+1}} = \mathbf{g}(\mathbf{X_t}, \mathbf{Z_t})$
- **Numerical, parameterized model**
  - $\mathbf{x_{t+1}} = \tilde{\mathbf{h}}(\mathbf{x_t}, \mathbf{z_{r,t+1}}, \mathbf{z_{u,t+1}})$ **and** $\mathbf{z_{u,t+1}} \approx \mathbf{F}(\mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta})$ **give**
  - $\mathbf{x_{t+1}} = \mathbf{G}(\mathbf{x_t}, \mathbf{z_{r,t+1}}, \mathbf{F}(\mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta}))$

1. **Stochastic-Bayesian Parameterization**

$$\mathbf{p}(\mathbf{x_{t+1}} \mid \mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta}) = \int \mathbf{p}(\mathbf{x_{t+1}} \mid \mathbf{x_t}, \mathbf{z_{u,t+1}}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta})\mathbf{p}(\mathbf{z_{u,t+1}} \mid \mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta})\mathbf{d}\mathbf{z_{u,t+1}}$$

2. **On-the-fly Stochastic-Bayesian Parameterization**

$$\begin{aligned}\mathbf{p}(\mathbf{x_{t+1}} \mid \mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta}, \mathbf{Y}) \;=\; & \int \mathbf{p}(\mathbf{x_{t+1}} \mid \mathbf{x_t}, \mathbf{z_{u,t+1}}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta}, \mathbf{Y}) \\ & \mathbf{p}(\mathbf{z_{u,t+1}} \mid \mathbf{x_t}, \mathbf{z_{r,t+1}}, \boldsymbol{\theta}, \mathbf{Y})\mathbf{d}\mathbf{z_{u,t+1}}\end{aligned}$$

- **Both Bayesian parameterizations are built using observations, model explorations, etc.**

Bayesian Computation and Monte Carlo

- **Bayes' Theorem gives <u>Posterior Distribution</u>:**

$$\mathbf{p}(\mathbf{x}, \boldsymbol{\theta} \,|\, \mathbf{y}) = \mathbf{q}(\mathbf{y} \,|\, \mathbf{x}, \boldsymbol{\theta})\mathbf{p}(\mathbf{x} \,|\, \boldsymbol{\theta})\mathbf{p}(\boldsymbol{\theta})/\mathbf{q}(\mathbf{y})$$

$$\textbf{where} \quad \mathbf{q}(\mathbf{y}) = \int \mathbf{q}(\mathbf{y} \,|\, \mathbf{x}, \boldsymbol{\theta})\mathbf{p}(\mathbf{x} \,|\, \boldsymbol{\theta})\mathbf{p}(\boldsymbol{\theta})\mathbf{dxd}\boldsymbol{\theta}$$

- **If q(y) is intractable, turn to Monte Carlo.**

$\boxed{\textbf{1. Monte Carlo (MC)}}$

- **Sample or ensemble $\mathbf{x}^1, \ldots, \mathbf{x}^M$ from $\mathbf{p}(\mathbf{x} \mid \mathbf{y})$ (Suppress $\boldsymbol{\theta}$)**

- **Estimate expectations: (notation: E( ) same as $< >$)**

$$\mathbf{E}(\mathbf{h}(\mathbf{X}) \mid \mathbf{y}) = \int \mathbf{h}(\mathbf{x})\mathbf{p}(\mathbf{x} \mid \mathbf{y})\mathbf{dx} \quad \textbf{by} \quad \hat{\mathbf{E}}(\mathbf{h} \mid \mathbf{y}) = \frac{\mathbf{1}}{\mathbf{M}} \sum \mathbf{h}(\mathbf{x}^i)$$

- **That is, approximate $\mathbf{p}(\mathbf{x} \mid \mathbf{y})$ by discrete, uniform distribution on the sample: $\widehat{\mathbf{Pr}}(\mathbf{X} = \mathbf{x}^i) = \mathbf{1}/\mathbf{M}$**

2. Importance Sampling (ISMC)

- **Direct sampling from $p(x \mid y)$ very hard or not possible**
- **Sample $x^1, \ldots, x^M$ from g**
- **Estimate**

$$E(h(X) \mid y) = \int h(x)\frac{p(x \mid y)}{g(x)}g(x)dx \quad \text{by} \quad \frac{1}{M}\sum h(x^i)\frac{p(x^i \mid y)}{g(x^i)}$$

- **Usual alternative:**
  - **Define normalized ISMC weights $\alpha_i = w(x^i) \,/ \sum w(x^j)$ where $w(x^i) = p(x^i \mid y)/g(x^i)$**
  - **Estimation: $\hat{E}(h \mid y) = \sum \alpha_i h(x^i)$**
  - **Approximate $p(x \mid y)$ by discrete distribution $\{x^i, \alpha_i : i = 1, \ldots, M\}$: $\widehat{Pr}(X = x^i) = \alpha_i$**
  - **Key: the normalizer q(y) of $p(x \mid y)$ cancels in the α's so we only need $p(x \mid y)$ up to proportionality.**

Notes on ISMC

- **Particle Filter: Evolve or generate $\mathbf{x}_t^i$ over time.**
  - **Sample $\{\mathbf{x}_{t-1}^i, \boldsymbol{\alpha}_{i,t-1} : \mathbf{i} = 1, \ldots, \mathrm{M}\}$ representing $\mathbf{p}(\mathbf{x}_{t-1} \mid \mathbf{y}_{t-1})$**
  - **Generate $\mathbf{x}_t^i \sim \mathbf{p}(\mathbf{x}_t \mid \mathbf{x}_{t-1}^i)$**
    $\{\mathbf{x}_t^i, \boldsymbol{\alpha}_{i,t-1} : \mathbf{i} = 1, \ldots, \mathrm{M}\}$ represents $\mathbf{p}(\mathbf{x}_t \mid \mathbf{y}_{t-1})$
    (Forecast Step)
  - **Bayes' Theorem converts to $\{\mathbf{x}_t^i, \boldsymbol{\alpha}_{i,t} : \mathbf{i} = 1, \ldots, \mathbf{m}\}$ representing $\mathbf{p}(\mathbf{x}_t \mid \mathbf{y}_t)$ where**
    $$\boldsymbol{\alpha}_{i,t} \propto \mathbf{q}(\mathbf{y}_t \mid \mathbf{x}_t^i)\boldsymbol{\alpha}_{i,t-1}$$
    (Analysis Step)
- **What we can do with an ensemble depends on how it was made.**
- **In high dimensions α's are poorly behaved:**
  They concentrate on a few (or one!) ensemble members

3. Markov Chain Monte Carlo (MCMC)

- **Finding normalizer q(y) vrs finding "partition function" in Statistical Mechanics**
- **MCMC: Develop a stationary (ergodic) Markov chain with limiting distribution coinciding with the target posterior $p(x \mid y)$.**
  - **After a "burn-in" (like "spin-up") period, realizations from the chain form an ensemble from $p(x \mid y)$ (approximately).**
  - **Ensemble members are dependent, but MC estimation works**

Metropolis-Hastings

- **State of chain at iterate i : $\mathbf{x}^{\mathbf{i}}$**
  - **generate $\tilde{\mathbf{x}}$ from some proposal distribution $\mathbf{g}(\tilde{\mathbf{x}} \mid \mathbf{x}^{\mathbf{i}})$**
  - **generate independent U = Uniform(0,1) RV**
  - **set $\mathbf{x}^{\mathbf{i+1}} = \tilde{\mathbf{x}}$ if**

$$\mathbf{U} < \frac{\mathbf{p}(\tilde{\mathbf{x}} \mid \mathbf{y})}{\mathbf{p}(\mathbf{x}^{\mathbf{i}} \mid \mathbf{y})} \frac{\mathbf{g}(\mathbf{x}^{\mathbf{i}} \mid \tilde{\mathbf{x}})}{\mathbf{g}(\tilde{\mathbf{x}} \mid \mathbf{x}^{\mathbf{i}})}$$

  - **set $\mathbf{x}^{\mathbf{i+1}} = \mathbf{x}^{\mathbf{i}}$, otherwise.**
- **Key: Again, normalizer $\mathbf{q}(\mathbf{y})$ cancels.**

Gibbs Sampler

- $\mathbf{x}$ is a K-dimensional vector, $(\mathbf{x_1}, \ldots, \mathbf{x_K})$
- Derive "full conditionals" $\mathbf{p}(\mathbf{x_k} \mid \mathbf{x_1}, \ldots, \mathbf{x_{k-1}}, , \mathbf{x_{k+1}}, \ldots, \mathbf{x_K}$ (and $\mathbf{y}$))
  - state of chain at iterate i: $(\mathbf{x_1^i}, \ldots, \mathbf{x_K^i})$
  - generate $\mathbf{x_1^{i+1}}$ from $\mathbf{p}(\mathbf{x_1} \mid \mathbf{x_2^i}, \mathbf{x_3^i}, \ldots, \mathbf{x_K^i})$
  - generate $\mathbf{x_2^{i+1}}$ from $\mathbf{p}(\mathbf{x_2} \mid \mathbf{x_1^{i+1}}, \mathbf{x_3^i}, \ldots, \mathbf{x_K^i})$

$$\vdots$$

  - generate $\mathbf{x_K^{i+1}}$ from $\mathbf{p}(\mathbf{x_K} \mid \mathbf{x_1^{i+1}}, \ldots, \mathbf{x_{K-1}^{i+1}})$

Others

- **Metropolis-within-Gibbs: replace intractable full conditionals by Metropolis steps.**
- **Using stochastic differential equation**

$$\mathbf{du(t) = b(u)dt + \boldsymbol{\sigma}(u)dW(t)}$$

where $\{\mathbf{W(t) : t \geq 0}\}$

  - **(Theory & assumptions) $\mathbf{U(t)}$ has a density function $\mathbf{p(u, t)}$**
  - **$\mathbf{p(u, t)}$ is solution Fokker-Planck Equation**
  - **Stationary solutions:**

$$\mathbf{0.50\, \frac{\partial^2}{\partial u^2}(\boldsymbol{\sigma}^2\, p) = \frac{\partial}{\partial u}(b\, p)}$$

  - **Pick b and $\boldsymbol{\sigma}$ so that p is the target posterior**
- **ISMC-MCMC**

Key Points

- **Monitoring convergence**
- **Output Analysis: Using output to summarize the target posterior.**
- **Tensions:**
  - **Multiple runs vrs one long run**
    Wasted burnin periods vrs "mixing"
  - **Output from a run are dependent**

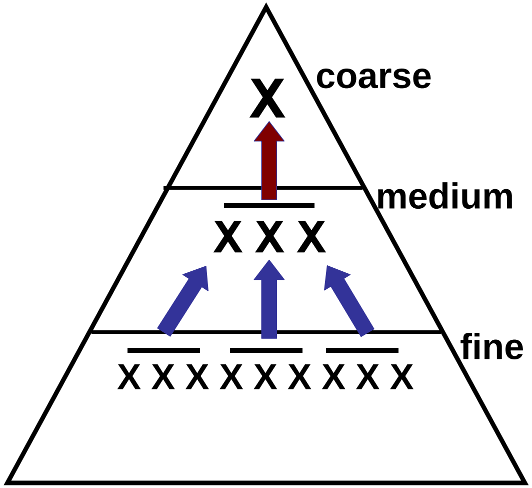
$$\mathbf{Var}(\bar{\mathbf{x}}^{\mathbf{i}}) = \frac{\mathbf{v}^2}{\mathbf{M}}(\mathbf{1} + \sum \mathbf{c}(\mathbf{i})\boldsymbol{\rho}_{\mathbf{i}})$$

- **Embarassingly Parallel? Obvious: Multiple runs, but?**

Two Notions of Multiscale Modeling

1. Space-Time Filtering

- "Hierarchical" process prior: $\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \vec{\mathbf{X}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{m}} \mid \vec{\mathbf{X}}_{\mathbf{f}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{f}})$
- Up-down scaling: $\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}}, \vec{\mathbf{X}}_{\mathbf{f}} \mid \vec{\mathbf{X}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{m}})$
- Terra incognita: $\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{m}} \mid \vec{\mathbf{X}}_{\mathbf{f}}, \vec{\mathbf{X}}_{\mathbf{c}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{f}}\,\vec{\mathbf{X}}_{\mathbf{c}})$

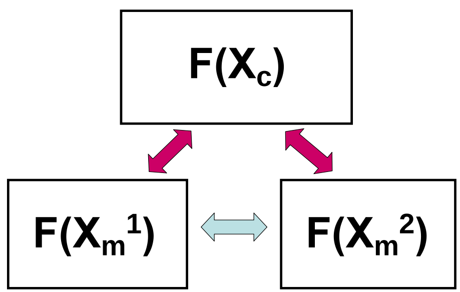


Example

- Data Model: $\mathbf{q}(\vec{\mathbf{Y}}_{\mathbf{c}} \mid \vec{\mathbf{X}}_{\mathbf{c}})\,\mathbf{q}(\vec{\mathbf{Y}}^{\mathbf{1}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})\,\mathbf{q}(\vec{\mathbf{Y}}^{\mathbf{2}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})$
- Process Prior: $\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}}, \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}}) \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})$

Full Conditionals for Gibbs Sampler:

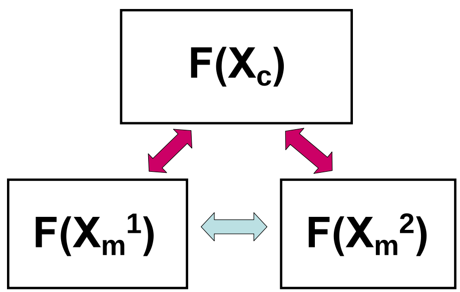
- $\mathbf{F}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \textbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}_{\mathbf{c}} \mid \vec{\mathbf{X}}_{\mathbf{c}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}}, \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})$
- $\mathbf{F}(\vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}} \mid \textbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}^{\mathbf{2}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}}, \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})$
- $\mathbf{F}(\vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}} \mid \textbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}^{\mathbf{1}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}_{\mathbf{c}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}}, \vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}^{\mathbf{2}}_{\mathbf{m}} \mid \vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})\,\mathbf{p}(\vec{\mathbf{X}}^{\mathbf{1}}_{\mathbf{m}})$
- Note how all levels intertwine: challenge to parallel code



Example Cont'd: Potentials for parallel codes

(1) Run bottom nodes holding $\vec{X}_c$ fixed
(i.e., we don't have to update every variable every time, though not doing so may slow convergence/)

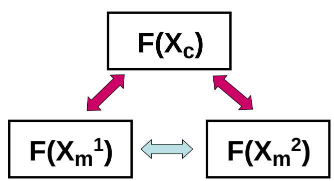
- Master swaps across scales occasionally.
- Many scales: Management system



Example Cont'd: Potentials for parallel codes

(2) Partial Conditionals for Gibbs Sampler with ISMC:

- $\mathbf{F}(\vec{\mathbf{X}}_\mathbf{c} \mid \mathbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}_\mathbf{c} \mid \vec{\mathbf{X}}_\mathbf{c})\, \mathbf{p}(\vec{\mathbf{X}}_\mathbf{c} \mid \vec{\mathbf{X}}^1_\mathbf{m}, \vec{\mathbf{X}}^2_\mathbf{m})$
- $\mathbf{F}_\mathbf{p}(\vec{\mathbf{X}}^2_\mathbf{m} \mid \mathbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}^2_\mathbf{m} \mid \vec{\mathbf{X}}^2_\mathbf{m})\, \mathbf{p}(\vec{\mathbf{X}}^2_\mathbf{m} \mid \vec{\mathbf{X}}^1_\mathbf{m}) \quad \{\mathbf{p}(\vec{\mathbf{X}}_\mathbf{c} \mid \vec{\mathbf{X}}^1_\mathbf{m}, \vec{\mathbf{X}}^2_\mathbf{m})\}$
- $\mathbf{F}_\mathbf{p}(\vec{\mathbf{X}}^1_\mathbf{m} \mid \mathbf{rest}\ ) \propto \mathbf{q}(\vec{\mathbf{Y}}^1_\mathbf{m} \mid \vec{\mathbf{X}}^1_\mathbf{m})\, \mathbf{p}(\vec{\mathbf{X}}^1_\mathbf{m}) \quad \{\mathbf{p}(\vec{\mathbf{X}}_\mathbf{c} \mid \vec{\mathbf{X}}^1_\mathbf{m}, \vec{\mathbf{X}}^2_\mathbf{m})\, \mathbf{p}(\vec{\mathbf{X}}^2_\mathbf{m} \mid \vec{\mathbf{X}}^1_\mathbf{m})\}$
  - **Ignoring terms in brackets if results are simple**
  - **But, those terms form required IS weights**
  - **Speed versus memory**



2. Parameterization of Scales

- **Build or "parameterize" scales into dynamic model for X**
  Example (Berliner & Kim, 2008, J Clim)
- **X: monthly surface temperatures**
- **Time series models (AR) with <u>time varying parameters</u>**

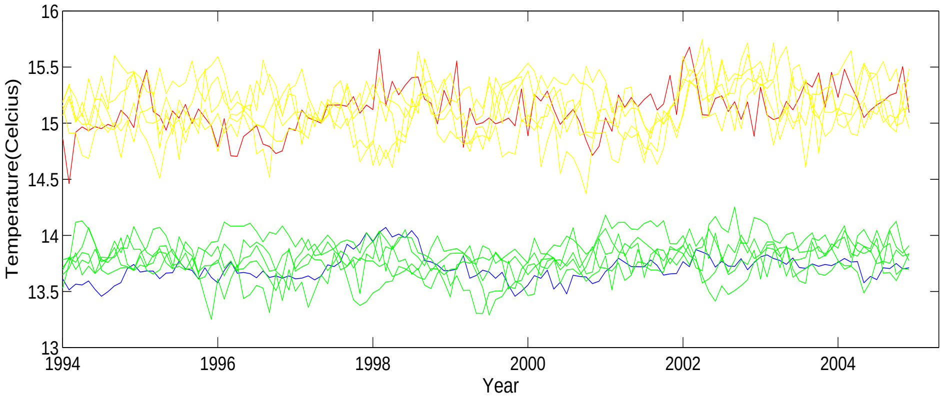
$$\mathbf{X_t = \boldsymbol{\mu}_{i(t)} + \boldsymbol{\beta}_{j(t)}(X_{t-1} - \boldsymbol{\mu}_{i(t-1)}) + e_{(t)}}$$

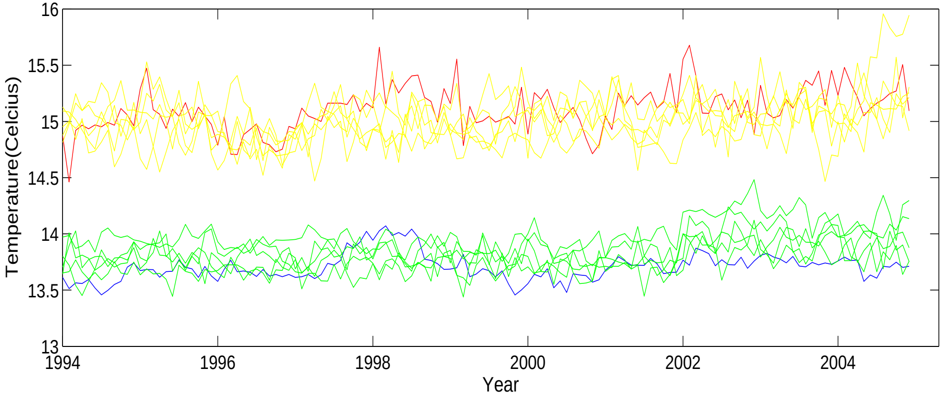
- **$\boldsymbol{\mu}_{i(t)}$ slowly vary (climate scale); $\boldsymbol{\beta}_{j(t)}$ vary moderately (another climate scale); e_t vary quickly ("weather"), but their variances slowly vary (climate scale):**
  - $\boldsymbol{\mu}_i = \mathbf{a + b\ CO_{2i} + noise}$
  - $\boldsymbol{\beta}_j = \mathbf{c + d\ SOI_j + noise}$

- **Computational challenge: Model selection**
  With what rates should the $\boldsymbol{\mu}_{\mathrm{i(t)}}$, $\boldsymbol{\beta}_{\mathrm{j(t)}}$,
  and variances of the $\mathrm{e_t}$ evolve?
  (1000's of combinations)

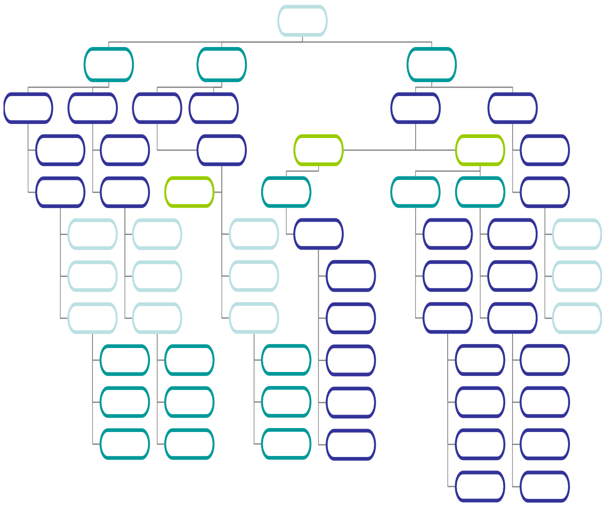
- **Decadal Prediction**

  - **Build model using observations up to 1994**
  - **Forecast for the following 10 years using a stochastic model for SOI**
  - **Next Graphic: show NH and SH observed temp's and ensembles from our predictive distributions**
    (First panel: $\boldsymbol{\mu}_{\mathrm{i(t)}}$ varying every 8 years,
    $\boldsymbol{\beta}_{\mathrm{j(t)}}$ varying every 2 years
    second panel: $\boldsymbol{\mu}_{\mathrm{i(t)}}$ varying every 8 years,
    $\boldsymbol{\beta}_{\mathrm{j(t)}}$ varying every 4 years)

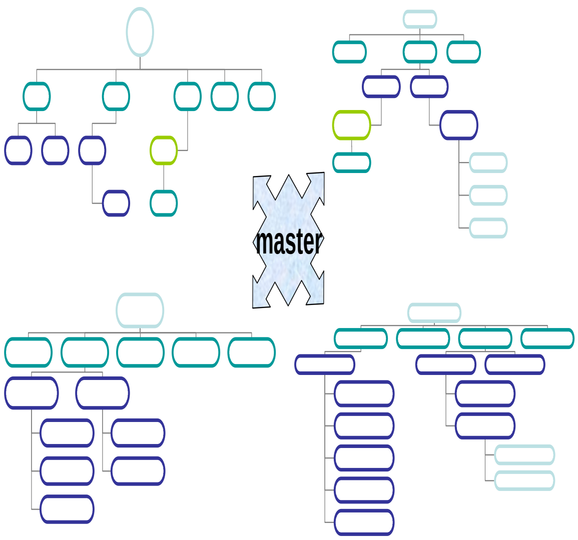

16
15.5
15
14.5
14
13.5
13
Temperature(Celcius)
1994
1996
1998
2000
2002
2004
Year


16
15.5
15
14.5
14
13.5
13
Temperature(Celcius)
1994
1996
1998
2000
2002
2004
Year

1. Bayesian Networks



2. Competing Networks



Discussion

- **Joe's talk**