Bayesian Hierarchical Models: A Brief Introduction

Christopher K. Wikle
Department of Statistics
University of Missouri-Columbia

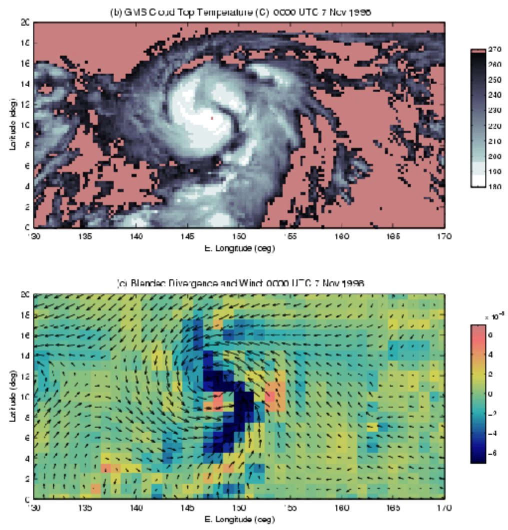


IMAGE TOY2006 Workshop III: February 27, 2006

Outline of Talk

- Motivation: physical/statistical models
- Bayesian Statistics
  - Simple Example
- Hierarchical Bayesian Models
  - Simple Example (cont.)
- Examples
  - Advection-Diffusion: physical space
    - Application: invasive species
  - Advection-Diffusion: spectral space
    - Application: tropical winds

Two types of geophysical models?

- Physical:
  - Laws of classical physics
  - Uncertainties: e.g., representativeness; neglecting higher order terms; choice of space-time averaging (subgrid-scale parameterizations); linking systems (e.g., atmosphere/ocean coupling), etc.
- Statistical:
  - Descriptive
  - Uncertainties: inefficient in complex settings; data not representative of full dynamics; application under system changes; estimation

Simple Illustration

PDE:

$$\frac{\partial u}{\partial t} = \mathcal{M}(u, w, \gamma),$$

where $\mathcal{M}$ is a function (e.g., spatial derivatives) of the variable of interest, u, other potential variables, w, and parameters γ.

Numerical Model:
Simple finite difference representations suggest an approximate difference equation model,

$$\mathbf{u}_{t+\Delta_t} = \mathbf{M}(\mathbf{u}_t, \mathbf{w}_t, \boldsymbol{\gamma}, \delta),$$

where Δ_t is the selected time step, $\mathbf{M}$ is based on $\mathcal{M}$, and δ involves the spatial grid sizes.

Statistical Model:

$$\mathbf{u}_{t+\Delta_t} = \mathbf{F}(\mathbf{u}_t, \mathbf{w}_t, \boldsymbol{\theta}) + \mathbf{e}_t,$$

where $\mathbf{F}$ is an unknown operator and $\boldsymbol{\theta}$ are unknown parameters. The additive error term $\mathbf{e}_t$ represents model and discretization errors.

"Physical/Statistical Models"

Better to consider a "spectrum of models" (e.g., Berliner, 2003)

$$\textbf{Deterministic} \quad \Longleftrightarrow \quad \textbf{Stochastic}$$

e.g., Hybrid Physical/Statistical Model:

$$\mathbf{u}_{t+\Delta_t} = \tilde{\mathbf{M}}(\mathbf{u}_t, \mathbf{w}_t, \boldsymbol{\theta}) + \mathbf{e}_t,$$

where $\tilde{\mathrm{M}}$ is an approximation to the "true" discretized model physics (M) and $\boldsymbol{\theta}$ are unknown, but **random** parameters. The additive error term $\mathbf{e}_t$ represents additional model errors.

Physical/Statistical Model Estimation

Uncertainties in the Physical/Statistical Model Framework:

- Data/Model agreement
- model representativeness
- parameter uncertainty

How do we accommodate these sources of uncertainty in a coherent, probabilistic framework?

Bayesian Approach:

- Natural for combining information sources while managing their uncertainties.
  - Multiple data sources
  - Uncertain model
  - Stochastic parameters
  - Expert opinion
- Obtain predictive distributions of quantities of interest, conditioned on observations

Distributional Notation

We use the short-hand notation $[\quad]$ for *probability distribution*.

Given random variables A and B, represent:

- **Joint** distribution of A and B: $[A, B]$
- **Conditional** distribution of A given B: $[A|B]$
- **Marginal** distribution of A: $[A]$

Also, note the following are equivalent:

$$Y|\mu, \sigma^2 \sim N(\mu, \sigma^2)$$

$$Y = \mu + \epsilon, \quad \epsilon \sim N(0, \sigma^2)$$

where "$\sim$" means "is distributed as".

Bayesian Viewpoint

Bayes' Theorem:

$$[Y|Z] = \frac{[Y,Z]}{[Z]} = \frac{[Z|Y][Y]}{\int [Z|Y][Y]dY}$$

For random variables with known distributions, this is a "fact" from probability theory. It is *not* controversial!

Bayesian Statistics:

Modeling unknown distributions and updating those models based on data. Might be controversial!

Bayesian Statistics

$$[Y|Z] = \frac{[Z|Y][Y]}{\int [Z|Y][Y]dY} \propto [Z|Y][Y]$$

- Want to make inference about Y but we only observe Z
- We **update** our uncertainty about Y after observing Z
- $[Y]$ reflects our **prior** knowledge of Y
- $[Z|Y]$ is the **"likelihood"** or **"data model"**
- $[Y|Z]$ is the **posterior distribution**

Bayesian Modeling:

Treat all unknowns *as if* they are random and evaluate probabilistically

$$[\text{Process}|\text{Data}] \propto [\text{Data}|\text{Process}][\text{Process}]$$

Simple Example I

- Temperature observations: $D_T(s_1), D_T(s_2), \ldots, D_T(s_n)$
- These are observations of the true (unknown) temperature process: T, at ”locations” (or replicates) $s_1, \ldots, s_n$; measurement error variance σ^2
- We have prior information about T, say T_0 from a “numerical model” but we don’t believe our model is perfect (τ^2).

Probabilistically, we may believe:

$$[\text{Data}|\text{Process}] : \qquad D_T(s_i)|T, \sigma^2 \sim N(T, \sigma^2)$$

$$[\text{Process}] : \qquad T|T_0, \tau^2 \sim N(T_0, \tau^2)$$

We want the posterior: [Process|Data], i.e.,

$$[T|D_T(s_1),\ldots,D_T(s_n)] \propto [D_T(s_1),\ldots,D_T(s_n)|T,\sigma^2][T|T_0,\tau^2]$$

where for simplicity, we assume T_0, τ^2, σ^2 are known. Then,

$$T|D_T(s_1),\ldots,D_T(s_n) \sim N(w\bar{D}_T + (1-w)T_0, \frac{\sigma^2\tau^2}{\sigma^2+n\tau^2})$$

where

$$\bar{D}_T = \frac{1}{n}\sum_{i=1}^{n} D_T(s_i) \quad \text{and} \quad w = \frac{\tau^2}{\tau^2+\sigma^2/n}.$$

Note: this is a weighted average of the sample mean and the prior mean, where the weights are related to our certainty about the observations and prior.

Basic Bayes is Not New to Atmospheric Science

- Epstein, early 1960's used Bayesian decision making in applied meteorology
- Weather Modification Research in 1970's (e.g., Simpson et al. 1973, *JAS*)
- Data Assimilation can be viewed from Bayesian viewpoint (e.g., Lorenc 1986, *Q. J. Roy. Met. Soc.*)
  - Kalman filter is Bayesian (e.g., Meinhold and Singpurwalla, 1983)
  - Ensemble Kalman Filters (Sequential Monte Carlo); e.g., Evensen and van Leeuwen, 2000, *MWR*
- Climate Change (e.g., Solow 1988; *J. Clim.*; Leroy 1998, *J. Clim.*)
- Satellite Retrieval (many!)

Hierarchical Bayes **is** new to meteorology!

Hierarchical Bayesian Viewpoint

Cornerstone of hierarchical modeling is **conditional thinking**.

Joint distribution can be represented as products of conditionals. e.g.,

$$[A, B, C] = [A|B, C][B|C][C]$$

NOTE: Our choice for this decomposition is based on what we know about the process, and what assumptions we are willing and able to make for simplification:

e.g., conditional independence: [A|B,C]=[A|B].

Often easier to express conditional models than full joint models.

Separate unknowns into two groups:

- process variables: actual physical quantities of interest (e.g., temperature, pressure, wind)
- model parameters: quantities introduced in model development (e.g., propagator matrices, measurement error and sub-grid scale variances, unknown physical constants, etc.)

Basic Hierarchical Model

1. [data|process, parameters]
2. [process|parameters]
3. [parameters]

The posterior [process,parameters|data] is proportional to the product of these three distributions!

Simple Example I (cont.)

Say that in Example I, we believe that the numerical model value of T is biased. We do not know exactly the bias, but suspect it is related to known factors $\mathbf{x} \equiv (x_1, x_2, \ldots, x_k)'$ (e.g., MOS). Then, our hierarchical model has the following stages:

- Data Model:

  $$D_T(s_i)|T, \sigma^2 \sim N(T, \sigma^2)$$

- Process Model:

  $$T|T_0, \mathbf{x}, \boldsymbol{\beta}, \tau^2 \sim N(T_0 + \mathbf{x}'\boldsymbol{\beta}, \tau^2)$$

- Parameter Model:

  $$\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma} \sim N(\boldsymbol{\beta}_0, \boldsymbol{\Sigma})$$

We then seek the posterior distribution: $[T, \boldsymbol{\beta}|D_T(s_1), \ldots, D_T(s_n)]$

Complicated Processes

Usually, we can't find analytically the normalizing constant in Bayes' theorem. Thus we can't easily get the posterior distribution.

- Numerical integration (o.k. for low dimensions)
- Use Monte Carlo methods to sample from the posterior.
  - Markov Chain Monte Carlo (MCMC)
    - Sample from a Markov Chain that has the same ergodic distribution as the posterior
    - Don't need to know normalizing constant
    - e.g., Metropolis (1950s), Gibbs sampler (1980s) algorithms
    - revolutionized Bayesian statistics in the early 1990's
  - Can be very computationally intensive
- Importance Sampling Monte Carlo (ISMC) (tomorrow's talk)

Simple Example II: Advection/Diffusion

Assume an advection-diffusion equation is fairly representative of the true physics of the system of interest:

$$\frac{\partial u}{\partial t} = -u_0 \frac{\partial u}{\partial x} + A \frac{\partial^2 u}{\partial x^2}$$

where our uncertainty is primarily related to the unknown parameters u_0, A.

Finite Difference Approximation:

$$\frac{u_{t+1}(x) - u_t(x)}{\delta_t} \approx -u_0 \left[\frac{u_t(x+1) - u_t(x-1)}{\delta_x}\right] + A \left[\frac{u_t(x+1) - 2u_t(x) + u_t(x-1)}{\delta_x^2}\right]$$

Difference Equation Representation

Collecting terms in the finite difference representation:

$$u_t(x) = \theta_1 u_{t-1}(x) + \theta_2 u_{t-1}(x+1) + \theta_3 u_{t-1}(x-1)$$

where θ_i are functions of δ_x, δ_t, A, and u_0.

In vector form, for locations $s_1, \ldots, s_n$ $(s_j = (x_j, y_j))$:

$$\mathbf{u}_t = \mathbf{M}\mathbf{u}_{t-1} + \mathbf{M}_b \mathbf{u}^b_{t-1}$$

where $\mathbf{M}$, $\mathbf{M}_b$ are functions of $\theta_1, \ldots, \theta_3$; $\mathbf{u}^b$ corresponds to the boundary process.

ADE Based Hierarchical Model: Physical Space

Data Model: Observations $\mathbf{Z}_t$

$$\mathbf{Z}_t | \mathbf{u}_t, \sigma_\epsilon^2 \;\sim\; N(\mathbf{H}_t \mathbf{u}_t, \sigma_\epsilon^2 \mathbf{I}), \quad t = 1, \ldots, T$$

where $\mathbf{Z}_t$ is $m_t \times 1$, $\mathbf{u}_t$ is $n \times 1$, and $\mathbf{H}_t$ is an $m_t \times n$ matrix that maps data to prediction locations.

Process Model:

$$\mathbf{u}_t | \mathbf{u}_{t-1}, \boldsymbol{\theta}, \boldsymbol{\Sigma}(\gamma) \;\sim\; N(\mathbf{M}(\boldsymbol{\theta}) \mathbf{u}_{t-1} + \mathbf{M}_b(\boldsymbol{\theta}) \mathbf{u}_{t-1}^b, \boldsymbol{\Sigma}(\gamma)),$$

for $t = 1, \ldots, T$. (For now, assume boundary process $\mathbf{u}_t^b$ is known; note: random boundaries fit nicely in the hierarchical framework!) Also, need distribution for *initial condition* : $\mathbf{u}_0 \;\sim\; N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$

ADE Hierarchical Model: Parameter Models

Priors for θ_i could be based on relationships suggested by the finite difference:

$$\theta_i \sim N(\tilde{\theta}_i, \sigma_i^2),$$

where

$$\tilde{\theta}_1 = 1 - \frac{2\delta_t A}{\delta_x^2}$$

$$\tilde{\theta}_2 = \frac{\delta_t A}{\delta_x^2} - \frac{\delta_t u_0}{\delta_x}$$

$$\tilde{\theta}_3 = \frac{\delta_t A}{\delta_x^2} + \frac{\delta_t u_0}{\delta_x}$$

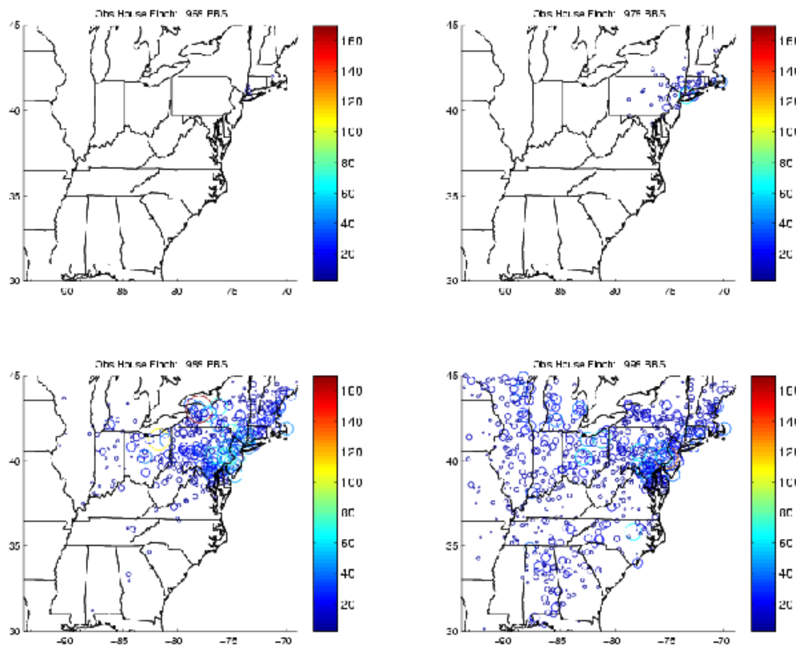
Note: One could also put a distribution on A and u_0 directly and/or allow them to vary spatially (e.g., forthcoming Ecology example).

Note: One may need to constrain these distributions to ensure stability. However, not as important as in numerical solutions to PDEs because data is available to naturally constrain. Also, in some cases, it is nice to have the flexibility!

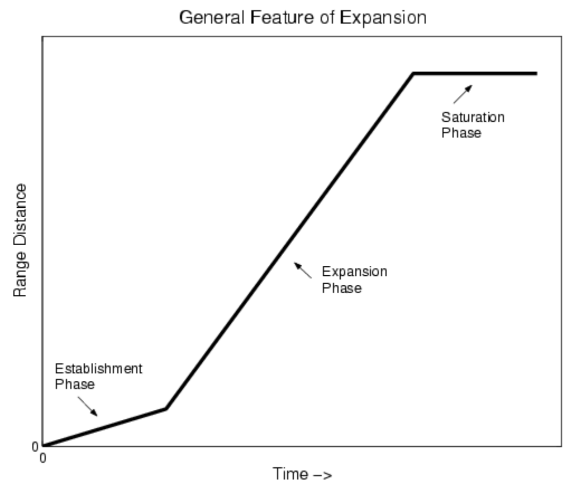
Reaction-Diffusion Application: (Wikle 2003; Ecology)

Spread of invasive species on landscape scale:

Breeding Bird Survey (BBS) counts for house finch (*Carpodacus mexicanus*) [substantial observer error and bias!]



Typical Invasions



Invasive species phases:

- Introduction
- Establishment
- Range Expansion
- Saturation

Ecological Models (dispersal and growth)

Ecology Example: Process Model Motivation

Skellam's (1951) Model (Diffusion plus Malthusian growth):

$$\frac{\partial u}{\partial t} = \delta \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + \alpha u.$$

Over large spatial scales, constant diffusion rate is not realistic. Spread will not be spatially homogeneous! e.g., a better model would be:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\delta(x, y) \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(\delta(x, y) \frac{\partial u}{\partial y} \right) + \alpha u,$$

- No general analytical solution
- Although this can be solved numerically, we still don't expect it to be exactly appropriate
- Growth process naive
- How do we estimate the spatial process $\delta(x, y)$?
- How do we account for the measurement/sampling errors (e.g., BBS data)?

Ecology Example: Hierarchical Model (Sketch)

- **Data Model:** $Z_t(\mathbf{s}_i)$ observed BBS count

$$Z_t(\mathbf{s}_i)|\lambda_t(\mathbf{s}_i) \sim iid\ Poisson(\lambda_t(\mathbf{s}_i))$$

- **Process Models:**

$$\log(\lambda_t(\mathbf{s}_i)) = \mu_t + \mathbf{h}'_{it}\mathbf{u}_t + \eta_t(\mathbf{s}_i), \quad \eta_t(\mathbf{s}_i) \sim iid\ N(0, \sigma^2_\eta)$$

$$\mu_t = \mu_{t-1} + \epsilon_t, \quad \epsilon_t \sim iid\ N(0, \sigma^2_\epsilon)$$

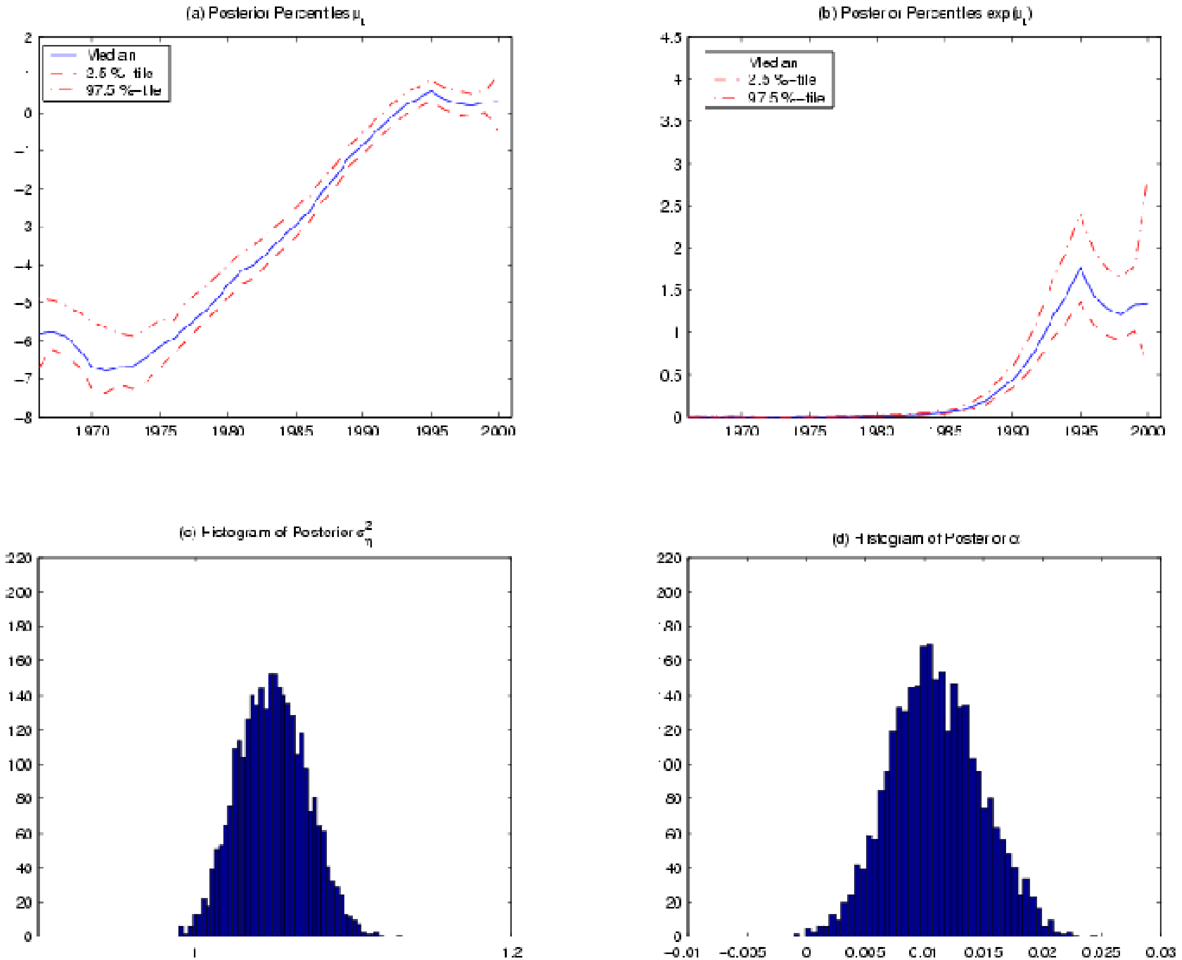
$$\mathbf{u}_t = \mathbf{M}(\boldsymbol{\delta}, \alpha)\mathbf{u}_{t-1} + \boldsymbol{\gamma}_t, \quad \boldsymbol{\gamma}_t \sim iid\ N(\mathbf{0}, \boldsymbol{\Sigma}(\theta_\gamma))$$

- **Parameter Models:**

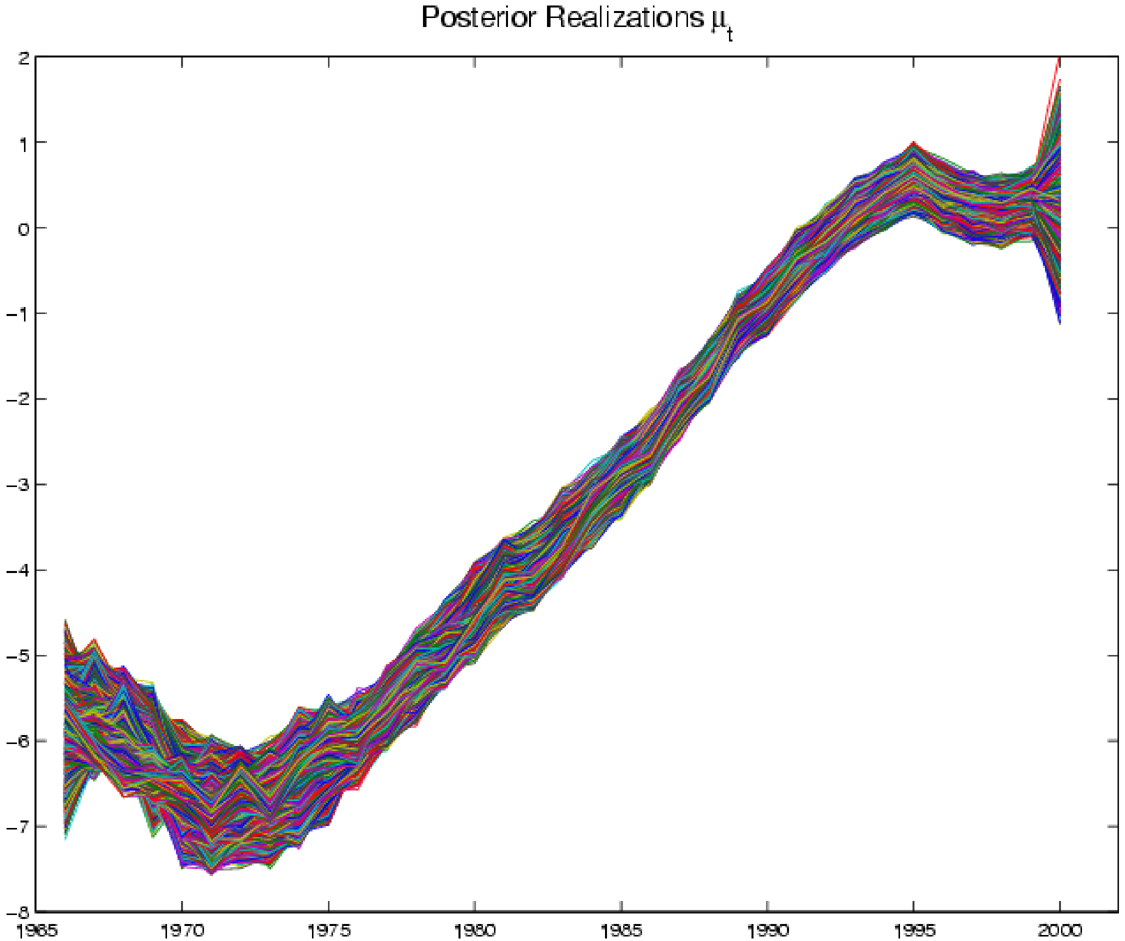
$$\boldsymbol{\delta}|\boldsymbol{\beta}, \sigma^2_\delta, \mathbf{R}_\delta \sim N(\boldsymbol{\Phi}\boldsymbol{\beta}, \sigma^2_\delta \mathbf{R}_\delta),$$
$$\alpha \sim N(\tilde{\alpha}_0, \tilde{\sigma}^2_\alpha)$$

Others: $\mathbf{u}_0$, $\boldsymbol{\beta}$, variances

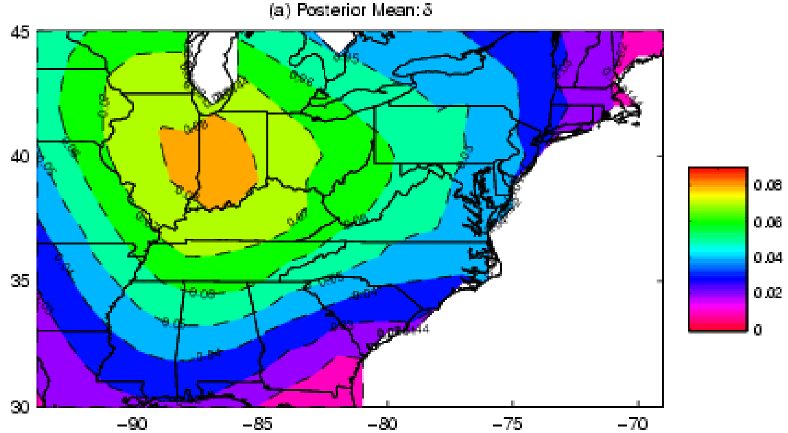
Ecological Example Posteriors: Growth

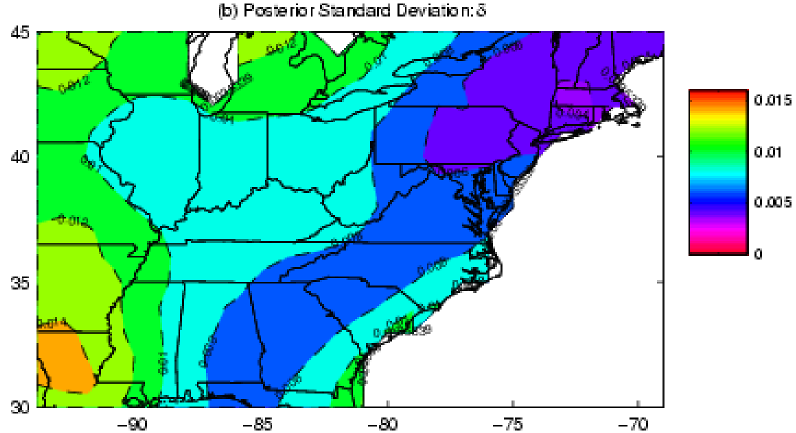


Ecological Example: (log) Growth Ensembles from Posterior

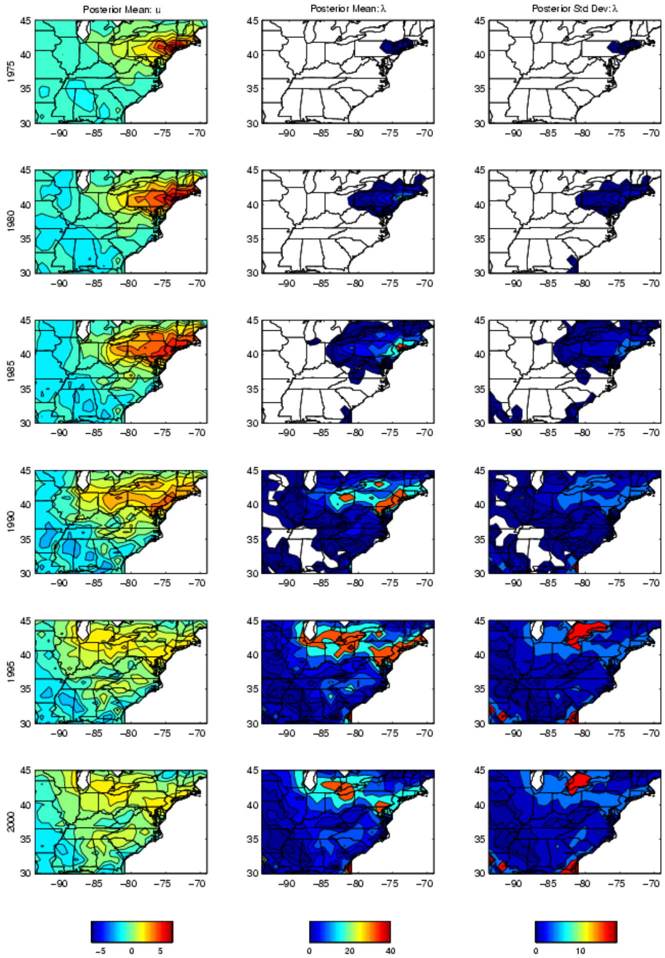


Ecological Example Posterior: Diffusion Parameter





Ecological Example Results: Process



Advection/Diffusion Example Revisited: Spectral

Again, consider the advection/diffusion equation:

$$\frac{\partial u}{\partial t} = -u_0 \frac{\partial u}{\partial x} + A \frac{\partial^2 u}{\partial x^2}$$

Assume solutions are a superposition of wave modes of the form:

$$u_t(x) = \sum_j [a_{1j}(t) \cos(\omega_j x) + a_{2j}(t) \sin(\omega_j x)]$$

where $\omega_j = 2\pi j / D_x$ is the spatial frequency of a wave with wave number j over domain D_x.

Thus, for all spatial locations of interest, $\mathbf{u}_t = \mathbf{\Phi} \boldsymbol{a}_t$ where $\mathbf{\Phi}$ is made up of the Fourier basis functions, and $\boldsymbol{a}_t$ is the collection of all wave-mode coefficients.

ADE Spectral (cont.)

The deterministic solution gives formulae for $a_{1j}(t), a_{2j}(t)$ (exponentially decaying sinusoids in time):

$$\begin{aligned} a_{1j}(t) &= \exp(-A\omega_j^2 t)\sin(u_0\omega_j t) \\ a_{2j}(t) &= \exp(-A\omega_j^2 t)\cos(u_0\omega_j t) \end{aligned}$$

Note time evolution:

$$\begin{bmatrix} e^{-A\omega_j^2(t+\delta)}\sin\{\omega_j(t+\delta)\} \\ e^{-A\omega_j^2(t+\delta)}\cos\{\omega_j(t+\delta)\} \end{bmatrix} = \mathbf{G}_j \begin{bmatrix} e^{-A\omega_j^2 t}\sin\{\omega_j t\} \\ e^{-A\omega_j^2 t}\cos\{\omega_j t\} \end{bmatrix}$$

where

$$\mathbf{G}_j = \begin{bmatrix} e^{-A\omega_j^2\delta}\cos\{\omega_j\delta\} & e^{-A\omega_j^2\delta}\sin\{\omega_j\delta\} \\ -e^{-A\omega_j^2\delta}\sin\{\omega_j\delta\} & e^{-A\omega_j^2\delta}\cos\{\omega_j\delta\} \end{bmatrix}$$

Thus (deterministic) linear wave theory suggests:

$$\boldsymbol{a}_j(t+\delta) = \mathbf{G}_j \boldsymbol{a}_j(t)$$

However, we don't expect the true process to behave *exactly* as the linear wave theory suggests!

- Let $a_{1j}(t), a_{2j}(t)$ be *stochastic*
- Add noise term $\boldsymbol{\eta}_j(t)$ to account for uncertainty
- Let the propagator be $\mathbf{M}_j$ with *prior mean* $\mathbf{G}_j$

$$\boldsymbol{a}_{t+\delta} = \mathbf{M}\boldsymbol{a}_t + \boldsymbol{\eta}_{t+\delta}$$

where $\boldsymbol{a}_t \equiv [\boldsymbol{a}_1(t)' \ \ldots \ \mathbf{a}_J(t)']'$ and $\mathbf{M}$ is block diagonal with blocks $\mathbf{M}_j$, $j = 1, \ldots, J$, where J is number of wave modes.

Hierarchical Spectral Spatio-Temporal Model

- Stage 1:

$$\mathbf{Z}_t = \mathbf{H}_t \mathbf{u}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I})$$

- Stage 2: (truncated modes)

$$\mathbf{u}_t = \boldsymbol{\Phi} \boldsymbol{a}_t + \boldsymbol{\gamma}_t, \quad \boldsymbol{\gamma}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\gamma)$$

- Stage 3:

$$\boldsymbol{a}_t = \mathbf{M} \boldsymbol{a}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$$

- Stage 4:

$$\text{vec}(\mathbf{M}_j) \sim N(\text{vec}(\mathbf{G}_i), \boldsymbol{\Sigma}_m)$$

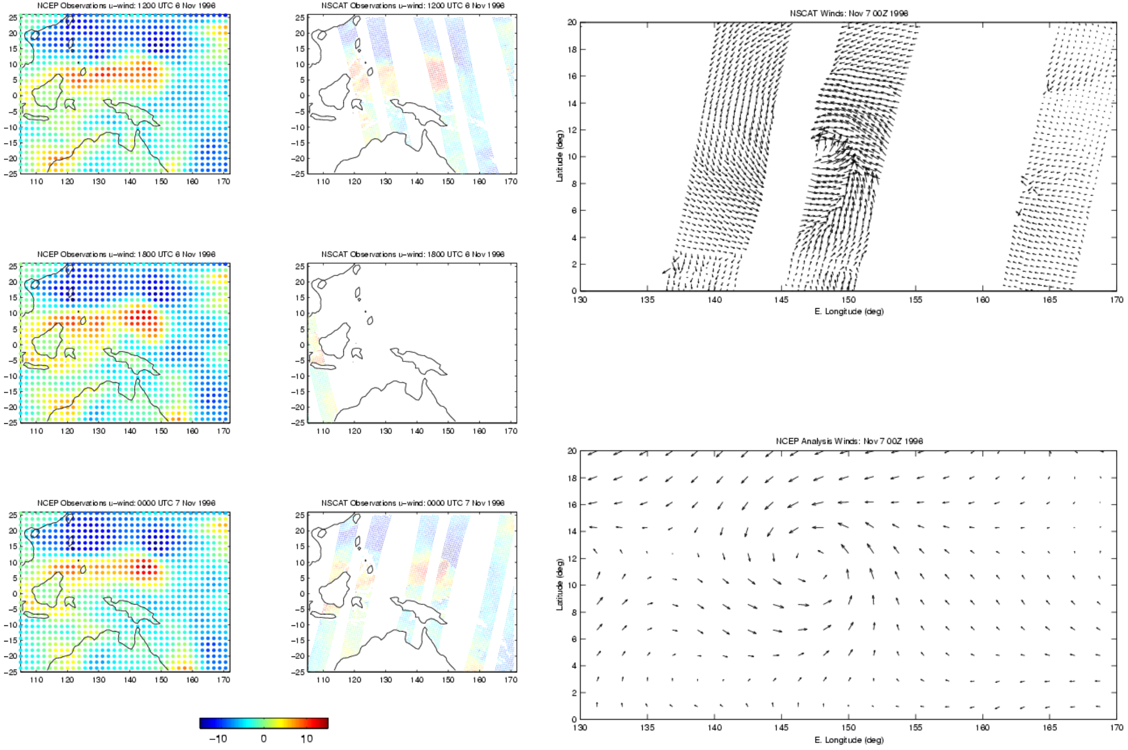
$$\boldsymbol{\Sigma}_\eta^{-1}(j) \sim W((\nu_j S_j)^{-1}, \nu_j)$$

$$\sigma_\epsilon^2 \sim IG(q_\epsilon, r_\epsilon)$$

$\boldsymbol{\Sigma}_\gamma(\theta)$is a stationary spatial cov matrix

Spectral Application: Tropical Ocean Winds

Problem: Blending tropical surface winds given high-resolution satellite scatterometer observations and low-resolution assimilated model output. (Wikle, Milliff, Nychka, Berliner, 2001; *JASA*)



Wind Problem: Process Model Motivation I

Consider the linear shallow water equations (on an equatorial beta-plane):

$$\frac{\partial u}{\partial t} - \beta_0 y v + g\frac{\partial h}{\partial x} = 0$$
$$\frac{\partial v}{\partial t} + \beta_0 y u + g\frac{\partial h}{\partial y} = 0$$
$$\frac{\partial h}{\partial t} + \bar{h}(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) = 0$$

- u, v: east-west, north-south wind components at location (x,y)
- h: deviation of the fluid depth about its mean
- $\bar{h}$: mean fluid depth
- β_0: constant related to rotation of earth
- g: gravitational acceleration

This linear system can be solved analytically, giving a series of traveling waves (*equatorial normal modes*) - **Observed in tropics!**

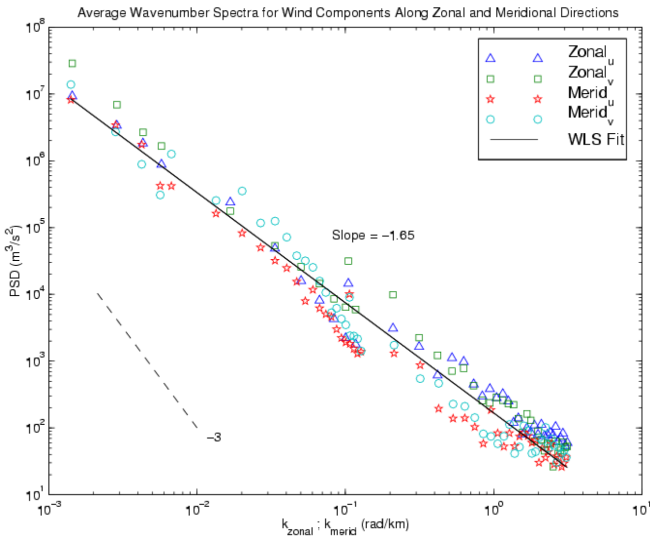
Wind Problem: Process/Parameter Model Motivation II

Empirical results suggest turbulent scaling behavior for near surface tropical winds (e.g., Wikle, Milliff and Large, 1999; *JAS*)

In particular:

$$S_v(\omega) \propto \frac{\sigma_v^2}{|\omega|^\kappa}$$

where $\kappa = 5/3$, and ω is the spatial frequency.



Wind Model: Hierarchical Model Sketch (v-Component)

- **Data Model:** Change of support

$$[\mathbf{Z}_{st}{}' \;\; \mathbf{Z}_{at}{}']' = \mathbf{H}_t \mathbf{v}_t + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$$

- **Process Model:**

$$\mathbf{v}_t = \boldsymbol{\mu} + \boldsymbol{\Phi} \boldsymbol{a}_t + \boldsymbol{\Psi} \mathbf{b}_t,$$

$\boldsymbol{\mu}$ - mean vector reflecting "climatological" winds

$\boldsymbol{\Phi}$ - matrix containing "important" normal mode basis functions (importance determined from empirical studies, e.g., Wheeler and Kiladis, 1999)

$\boldsymbol{a}_t$ - time-varying equatorial mode spectral coefficients

$\boldsymbol{\Psi}$ - matrix containing multiresolution (wavelet) basis functions (representing small/meso-scale variability)

$\mathbf{b}_t$ - time-varying multiresolution coefficients

$$\boldsymbol{\mu}|\boldsymbol{\beta}, \boldsymbol{\Sigma}_\mu \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_\mu)$$

$\boldsymbol{\beta}$ - "regression" coefficients for climatological covariates $\mathbf{X}$

$$\boldsymbol{a}_t = \mathbf{M}(\boldsymbol{\theta})\boldsymbol{a}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$$

$\boldsymbol{\theta}$ - informative priors based on equatorial wave theory and empirical studies

$$\boldsymbol{b}_t = \mathbf{M}_b(\boldsymbol{\theta}_b)\boldsymbol{b}_{t-1} + \boldsymbol{\gamma}_t, \quad \boldsymbol{\gamma}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\gamma)$$

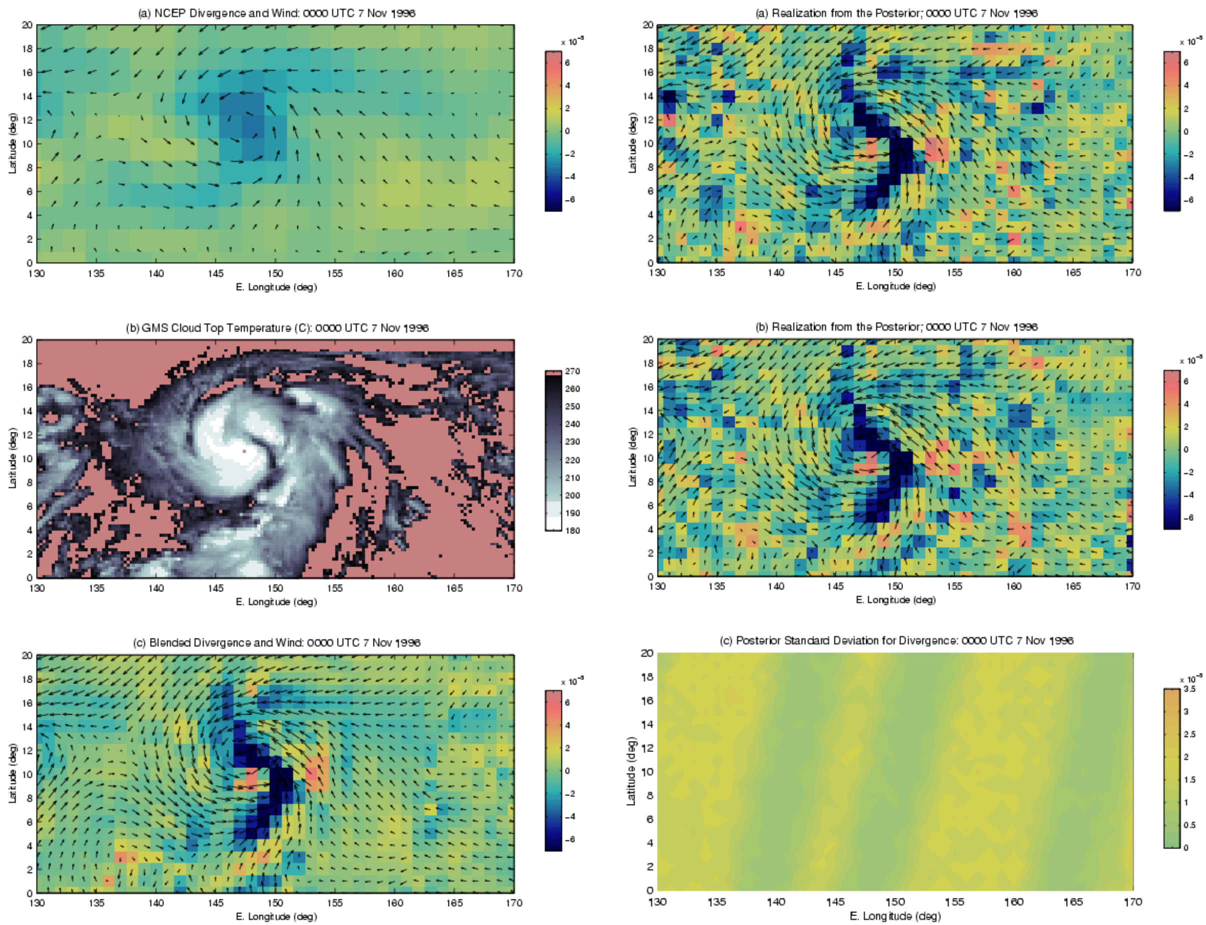
$\mathbf{M}_b(\boldsymbol{\theta})$ - diagonal propagator (AR(1))

$\boldsymbol{\Sigma}_\gamma$ - diagonal, but informative priors to suggest empirical power-law scaling relationships

- **Parameter Models:**

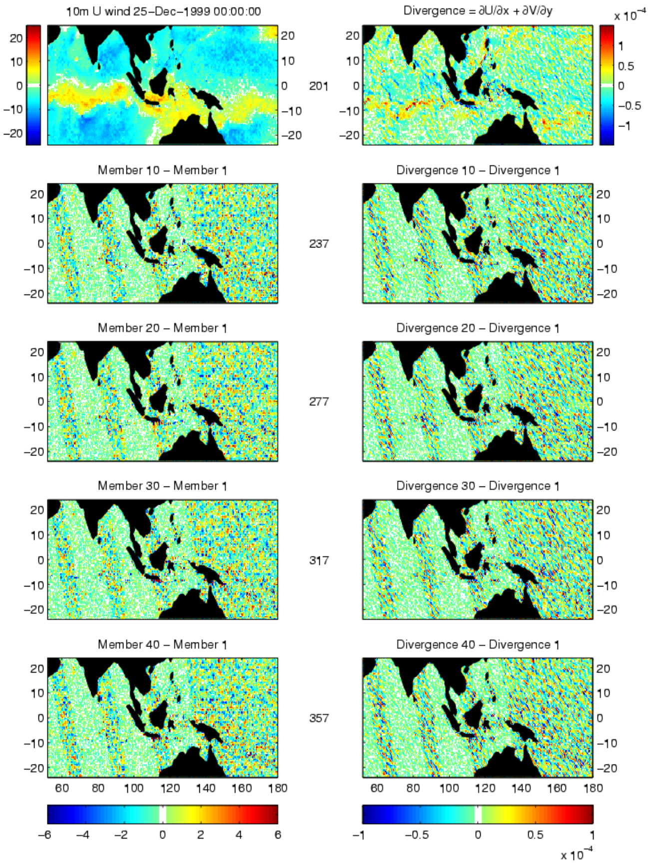
  Parameter distributional choices reflect theoretical and empirical science related to equatorial normal modes and spectral scaling relationships.

Wind Problem Results: Tropical Cyclone Dale



Wind Blending: “Operational”

Tim Hoar, IMAGe, NCAR has made this model “operational” on 1/2 deg grid



Conclusion

- Hierarchical Bayesian methods allow one to quantify uncertainty in all aspects of the problem (data, process, parameters) and distributional output reflects the quantification of uncertainty
- Hierarchical Bayesian methods allow one to decompose the problem into a series of simpler conditional models
  - Can accommodate data of different sources/resolution/alignment
  - Can accommodate complicated spatio-temporal dependence
  - Can accommodate physics (e.g., shallow water PDE; reaction-diffusion)
  - Can accommodate empirical results (e.g., turbulent scaling laws)
  - Can accommodate stochastic parameterizations!
- Downside: Complicated and computationally intense

Tomorrow's Talk: Preview

Realistic Implementation of BHM Parameterization

- Qualitative Dynamics
  - Long Lead Forecast of SST
  - Nowcast of Radar Reflectivities
- Coupling Systems
  - Air/Sea Interaction
- Stochastic Parameterization in Climate Models
  - Convective Initiation in Mesoscale Forecast Model