BAYESIAN MODELING OF UNCERTAINTY IN ENSEMBLES OF CLIMATE MODELS

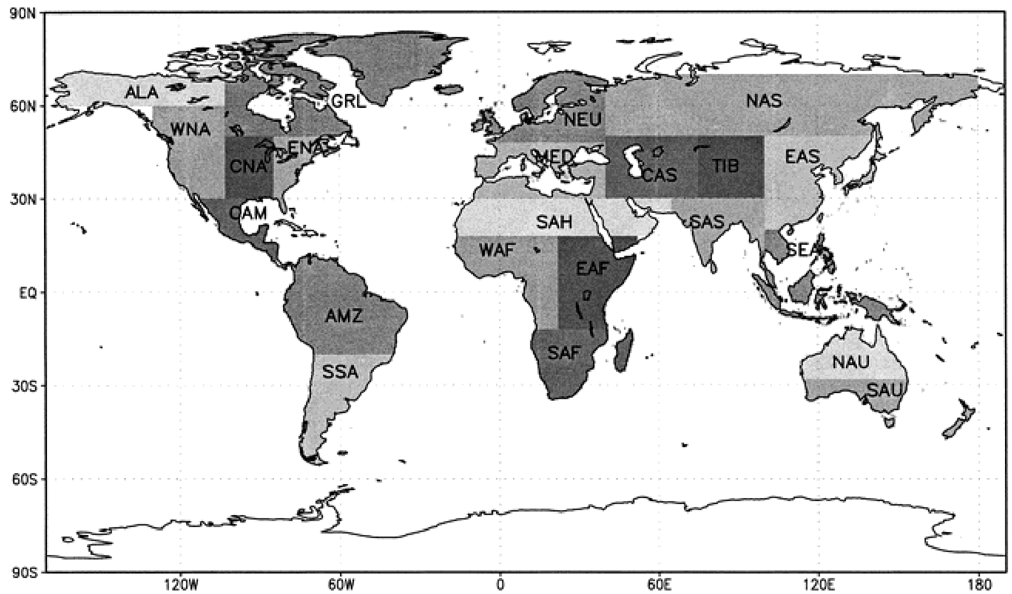
Richard L. Smith
(University of North Carolina, Chapel Hill, visiting NCAR)
Claudia Tebaldi, Doug Nychka and Linda Mearns
(National Center for Atmospheric Research, Boulder, CO)

Hawaii International Conference on Statistics, Mathematics and Related Fields
January 18, 2006

Supported by NSF, NOAA and NCAR Assessment Initiative

Model	Comment	Sensitivity
CCC	Canada	3.59
CSIRO	Australia	3.50
CSM	NCAR, USA	2.29
DMI	Max Planck (Germany)	3.11
GFDL	Princeton, USA	2.87
MRI	Japan	1.25
NIES	Japan	4.53
PCM	Parallel Climate Model (USA)	2.35
HADCM	Hadley Centre, U.K.	3.38

Climate models used in this study, and their climate sensitivities in ^{o}K.



Regions used in this study

The Giorgi-Mearns Approach

- M climate models
- X_j — projection of current climate by model j
- Y_j — projection of future climate by model j
- X_0 — current true climate (uncertainty ϵ)

If $Var(Y_j) = \sigma^2/\lambda_j$, consider

$$\tilde{Y} = \frac{\sum_{j=1}^{M} \lambda_j Y_j}{\sum_{j=1}^{M} \lambda_j}. \tag{1}$$

Giorgi and Mearns called λ_j the "reliability" of model j and formula (1) the "reliability ensemble estimator" or REA.

To estimate the reliabilities, Giorgi and Mearns proposed

$$\lambda_j = (\lambda_{B,j}^m \lambda_{D,j}^n)^{1/mn} \tag{2}$$

where

$$\lambda_{B,j} = \min\left(1, \frac{\epsilon}{|X_j - X_0|}\right), \quad \lambda_{D,j} = \min\left(1, \frac{\epsilon}{|Y_j - \tilde{Y}|}\right), \tag{3}$$

where $|X_j - X_0|$ is the "bias" of model j, $|Y_j - \tilde{Y}|$ the "convergence" of model j, and the parameters m and n control the relative importance given to these two quantities (Giorgi and Mearns suggested $m = n = 1$).

Iterate to convergence...

Bayesian (Univariate) Approach

$$X_0 \sim N[\mu, \lambda_0^{-1}], \quad (\lambda_0 \text{ known}) \tag{4}$$
$$X_j \sim N[\mu, \lambda_j^{-1}], \tag{5}$$
$$Y_j \mid X_j \sim N[\nu + \beta(X_j - \mu), (\theta\lambda_j)^{-1}], \tag{6}$$

where parameters μ, ν, β, θ, λ_j have prior distributions

$$\mu, \nu, \beta \sim U(-\infty, \infty), \tag{7}$$
$$\theta \sim G[a, b], \tag{8}$$
$$\lambda_1, ..., \lambda_M \sim G[a_\lambda, b_\lambda], \tag{9}$$
$$a_\lambda, b_\lambda \sim G[a^*, b^*]. \tag{10}$$

Here the hyperparameters a, b, a^*, b^* are chosen so that each of θ, a_λ, b_λ has a proper but diffuse prior. In practice we set $a = b = a^* = b^* = 0.01$.

Define

$$\tilde{\mu} = \frac{\lambda_0 X_0 + \sum \lambda_j X_j - \theta\beta \sum \lambda_j (Y_j - \nu - \beta X_j)}{\lambda_0 + \sum \lambda_j + \theta\beta^2 \sum \lambda_j}, \tag{11}$$

$$\tilde{\nu} = \frac{\sum \lambda_j \{Y_j - \beta(X_j - \mu)\}}{\sum \lambda_j}, \tag{12}$$

$$\tilde{\beta} = \frac{\sum \lambda_j (Y_j - \nu)(X_j - \mu)}{\sum \lambda_j (X_j - \mu)^2}. \tag{13}$$

Gibbs sampler updates:

$$\begin{aligned}
\mu \mid \text{rest} &\sim N\left[\tilde{\mu}, \frac{1}{\lambda_0 + \sum \lambda_j + \theta\beta^2 \sum \lambda_j}\right], && (14)\\
\nu \mid \text{rest} &\sim N\left[\tilde{\nu}, \frac{1}{\theta \sum \lambda_j}\right], && (15)\\
\beta \mid \text{rest} &\sim N\left[\tilde{\beta}, \frac{1}{\theta \sum \lambda_j (X_j - \mu)^2}\right], && (16)\\
\lambda_j \mid \text{rest} &\sim G\Big[a + 1, \\
&\quad b + \frac{1}{2}(X_j - \mu)^2 + \frac{\theta}{2}\{Y_j - \nu - \beta(X_j - \mu)\}^2\Big], && (17)\\
\theta \mid \text{rest} &\sim G\left[a + \frac{M}{2}, b + \frac{1}{2}\sum \lambda_j \{Y_j - \nu - \beta(X_j - \mu)\}^2\right]. && (18)
\end{aligned}$$

For the parameters a_λ, b_λ, use Metropolis

Cross-Validation in the Univariate Model

How could we validate this approach to calculating a predictive distribution for the true temperature difference $Y_0 - X_0$?

Method 1: Wait 100 years, measure the true Y_0, calibrate against predictive density.

Disadvantage of this method?

Method 2: *Population of Models* approach

Think of the given ensemble of 9 models as a random sample from a hypothetically infinite population of models.

Suppose someone came along with a new model, for which the projected climate variables were $X^\dagger, Y^\dagger$. Conditionally on the hyperparameters $\mu,\ \nu,\ \beta,\ \theta,\ a_\lambda$ and b_λ, the distribution of $Y^\dagger - X^\dagger$ is derived from

(i) $\lambda^\dagger \sim G[a_\lambda, b_\lambda]$,

(ii) conditionally on $\lambda^\dagger$, $Y^\dagger - X^\dagger \sim N\left[\nu - \mu, \frac{(\beta-1)^2+\theta^{-1}}{\lambda^\dagger}\right]$.

so by mixing this conditional predictive distribution over the posterior distribution of $(\mu, \nu, \beta, \theta, a_\lambda, b_\lambda)$, we obtain a full posterior predictive distribution for $Y^\dagger - X^\dagger$.

Use this as the basis for cross-validation

Cross-Validation Algorithm

1. For each $j \in \{1, ..., M\}$, rerun the REA.GM procedure without model j.

2. Using parameters at nth MCMC iteration, say $a_\lambda^{(n)}, b_\lambda^{(n)}, \nu^{(n)}$, $\mu^{(n)}, \beta^{(n)}, \theta^{(n)}$, draw a random $\lambda_{j,n} \sim G[a_\lambda^{(n)}, b_\lambda^{(n)}]$ and calculate

$$U_j^{(n)} = \Phi \left\{ \frac{Y_j - X_j - \nu^{(n)} + \mu^{(n)}}{\sqrt{\{(\beta_x^{(n)} - 1)^2 + \theta^{(n)^{-1}}\}/\lambda_{j,n}}} \right\}$$

3. Let U_j be the mean of $U_j^{(n)}$ over all n

4. Recompute steps 1–3 for each region, so we have a set of test statistics $U_{ij},\ i = 1, ..., R,\ j = 1, ..., M$.

5. Plot the U_{ij}'s, apply tests of fit, etc.

Multivariate Approach

Assume we have current and future climate model projections, X_{ij} and Y_{ij}, models $j = 1, ..., M$, regions $i = 1, ..., R$. X_{i0} is the current observed mean temperature in region i, variance $\frac{1}{\lambda_{0i}}$.

Model and prior densities:

$$\begin{aligned}
X_{i0} &\sim N[\mu_0 + \zeta_i, \lambda_{0i}^{-1}], && (19)\\
X_{ij} &\sim N[\mu_0 + \zeta_i + \alpha_j, (\eta_{ij}\phi_i\lambda_j)^{-1}], && (20)\\
Y_{ij} \mid X_{ij} &\sim N[\nu_0 + \zeta_i' + \alpha_j' + \beta_i(X_{ij} - \mu_0 - \zeta_i - \alpha_j), \\
&\quad (\eta_{ij}\theta_i\lambda_j)^{-1}], && (21)\\
\mu_0, \nu_0, \zeta_i, \zeta_i', \beta_i, \beta_0 &\sim U(-\infty, \infty), && (22)\\
\theta_i, \phi_i, \psi_0, \theta_0, c, \alpha_\lambda, \beta_\lambda &\sim G[a, b], && (23)\\
\lambda_j \mid a_\lambda, b_\lambda &\sim G[a_\lambda, b_\lambda], && (24)\\
\eta_{ij} \mid c &\sim G[c, c], && (25)\\
\alpha_j \mid \psi_0 &\sim N[0, \psi_0^{-1}], && (26)\\
\alpha_j' \mid \alpha_j, \beta_0, \theta_0, \psi_0 &\sim N[\beta_0\alpha_j, (\theta_0\psi_0)^{-1}]. && (27)
\end{aligned}$$

Features of this Model

- Allows overall model biases α_j and α'_j, *but* these are treated differently from the region effects ζ_i and ζ'_i

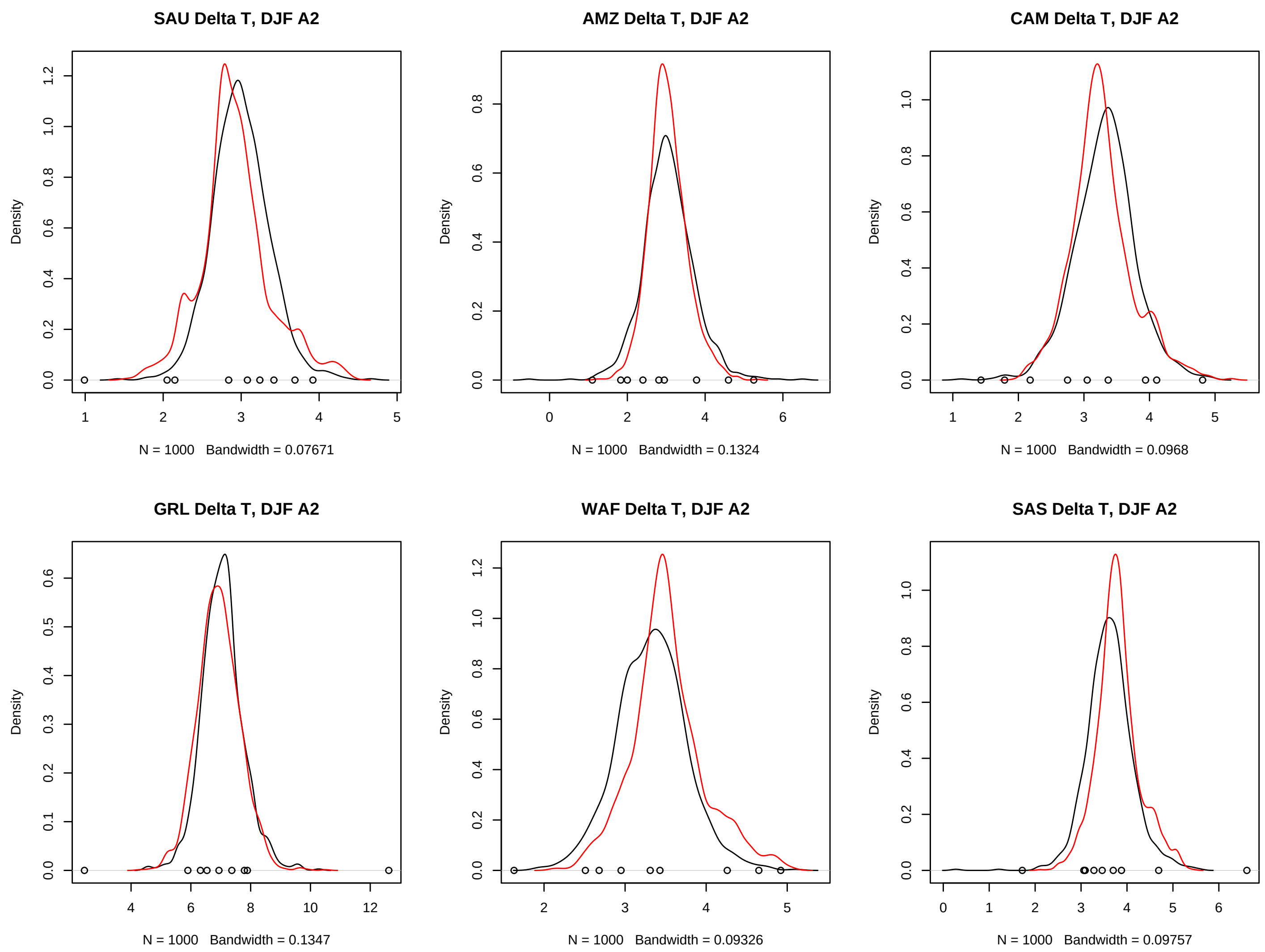
- Variances: Introduce interaction term η_{ij} governed by parameter c
  - $c \to \infty$ corresponds to $\eta_{ij} \equiv 1$ — no interaction
  - $c \to 0$ implies no constraint on variances — equivalent to univariate approach
  - Actual approach allows calculation of a posterior distribution for c — interaction finds its onw level

Implementation

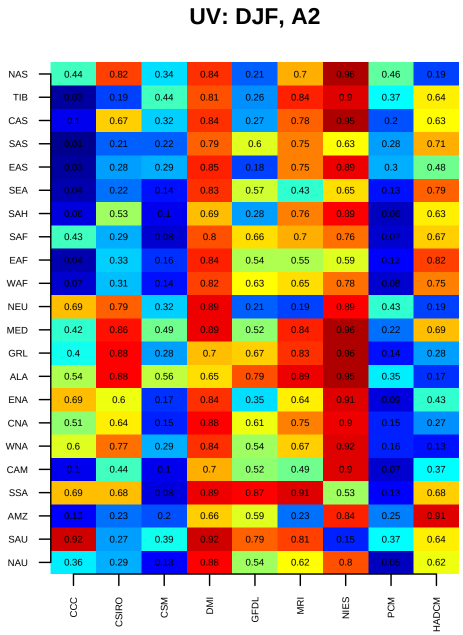
- MCMC for posterior densities
- Cross-validation statistics U_{ij}

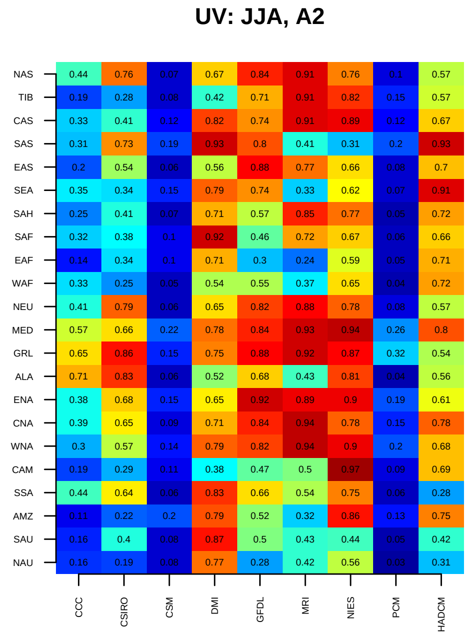
Both may be calculated by generalizing the univariate approach

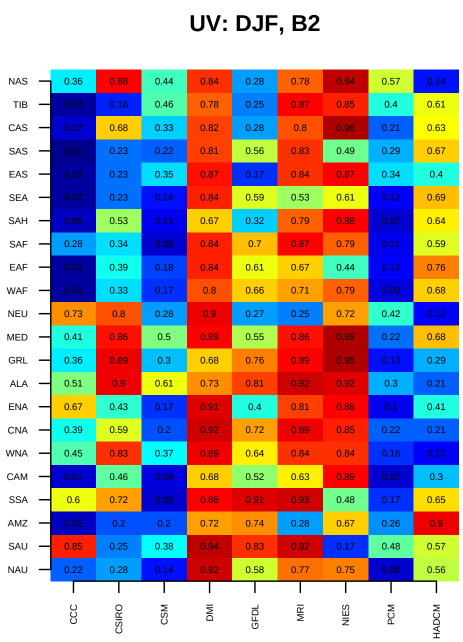
Results...

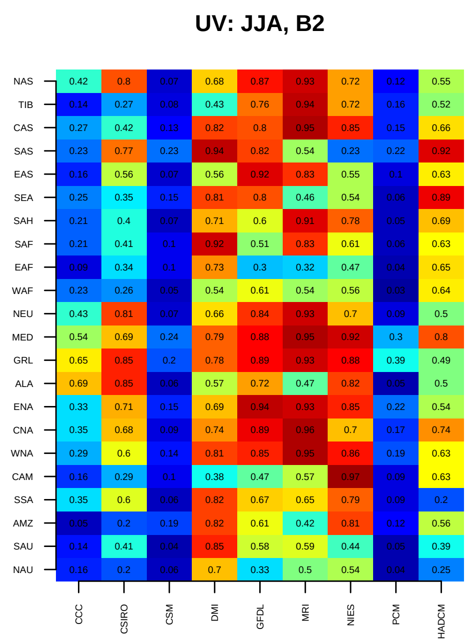


Posterior density of temperature change in 6 regions
(black — UV model; red — MV model to be discussed later)

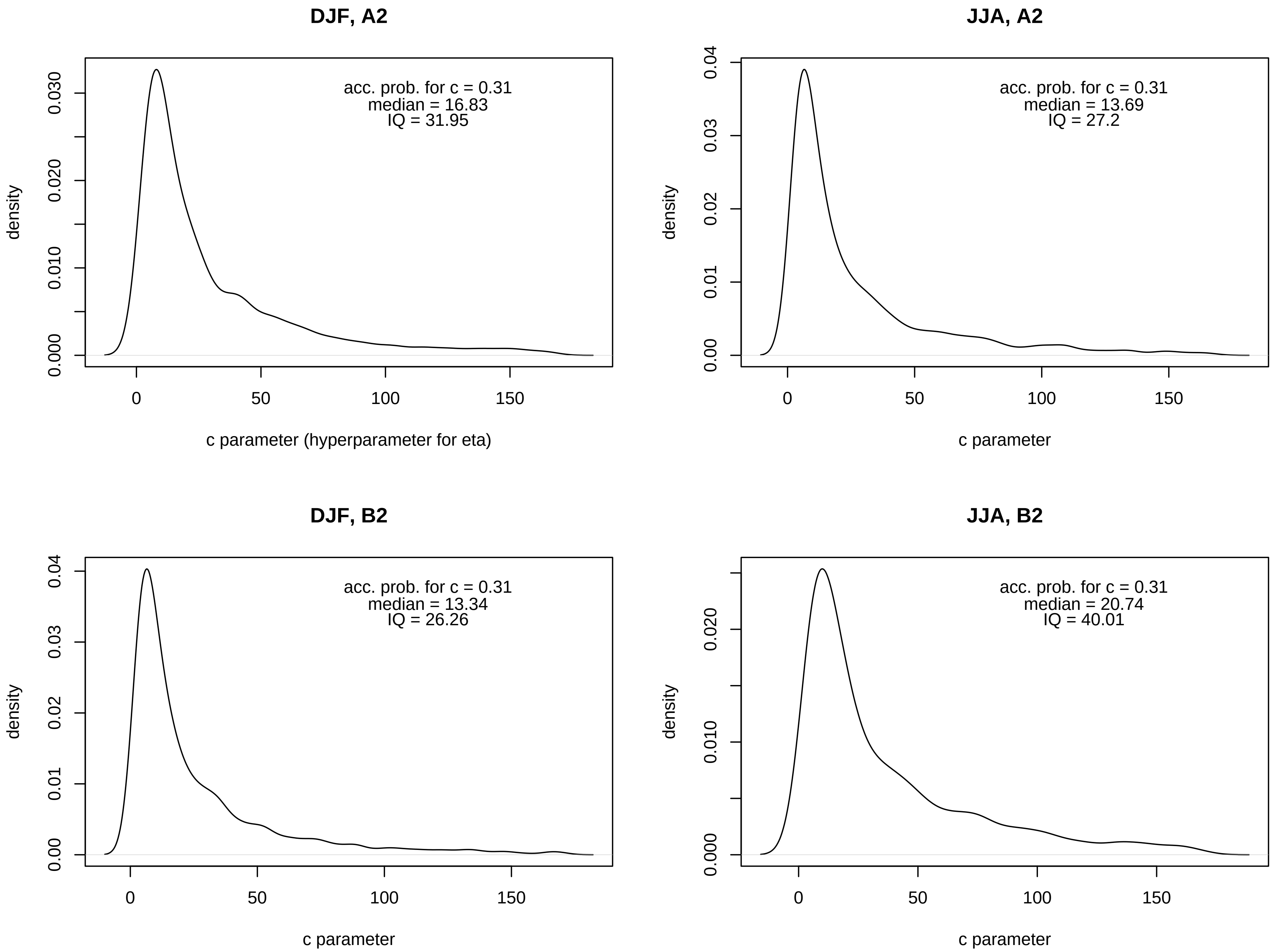




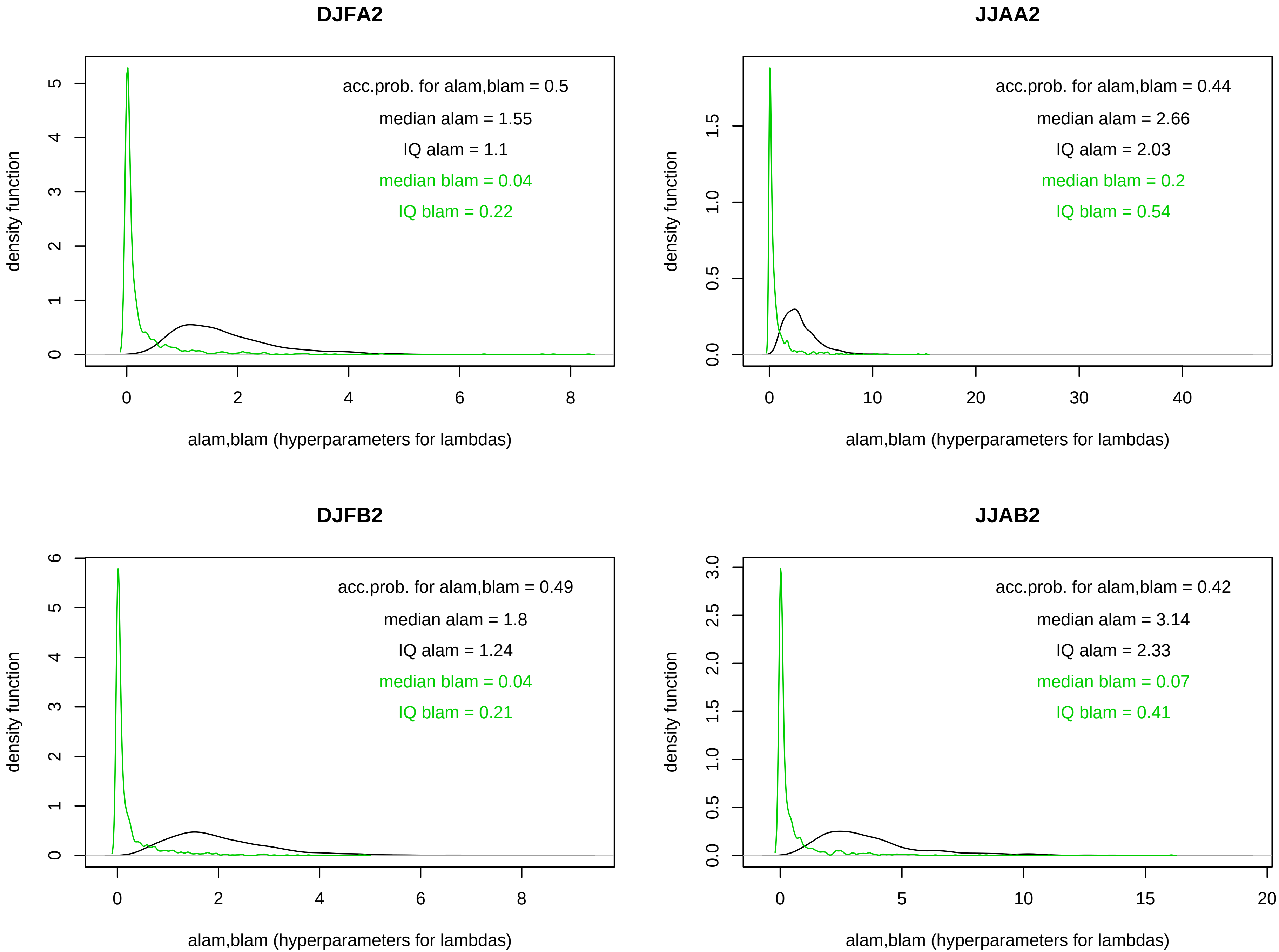




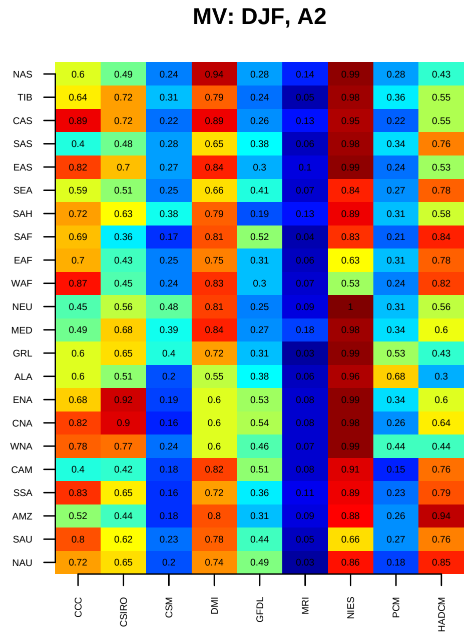
U-statistic plots — Univariate model

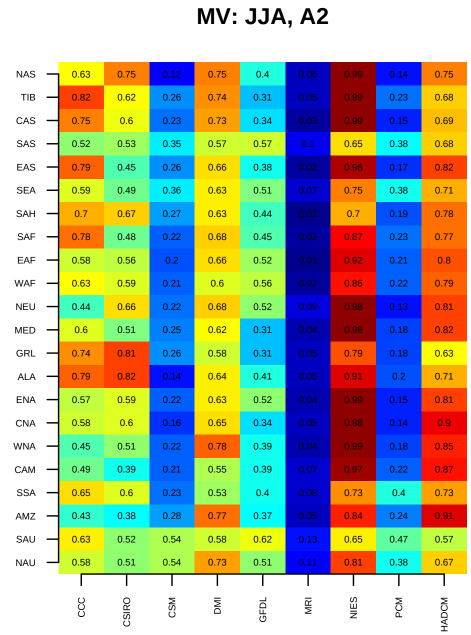


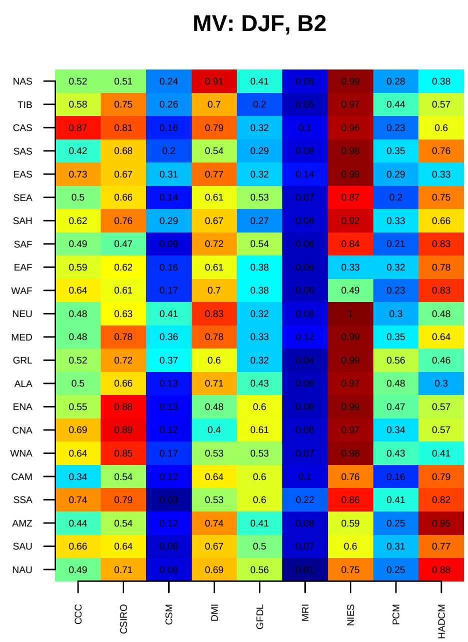
Posterior density of c

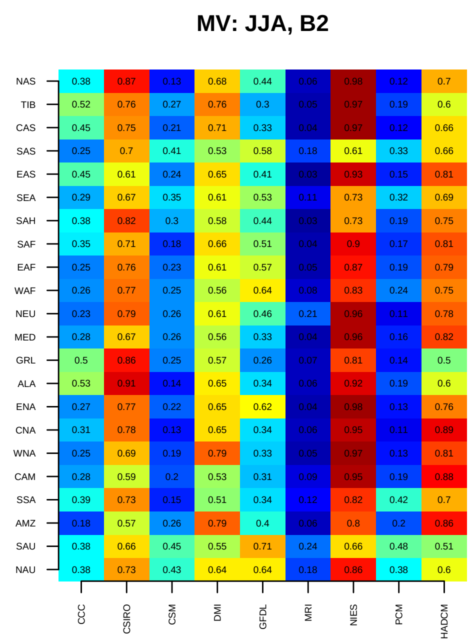


Posterior density of a_λ and b_λ









U-statistic plots — Multivariate model

Goodness of Fit of the U Statistics

For each region, 9 values of U_{ij}

Based on these 9 values, carry out a *two-sided* Kolmogorov-Smirnov test of size .05. Critical values by simulation.

Note total number of rejections in 22 regions.

	Univariate	Multivariate
DJFA2	1	3
DJFB2	2	3
JJAA2	0	3
JJAB2	1	3

All rejections in *lower* tail of K-S statistic.

Similar results for Cramér-von Mises,

	Univariate	Multivariate
DJFA2	3	2
DJFB2	1	3
JJAA2	0	4
JJAB2	2	3

Anderson-Darling,

	Univariate	Multivariate
DJFA2	3	3
DJFB2	3	3
JJAA2	1	3
JJAB2	2	5

Correlation test (analogous to Shapiro-Wilk)

	Univariate	Multivariate
DJFA2	0	3
DJFB2	1	1
JJAA2	0	2*
JJAB2	0	1

* only one in the whole set of tests where rejection indicated that empirical CDF was *too far* from theoretical CDF.

Conclusions: The fit to the uniform CDF seems *too good* (more so in the case of the multivariate method). Possible artifact of correlation effects.

Are the multivariate model predictive distributions "tighter" than those based on the univariate model?

Compare IQR, I15R (difference between 85th and 15th percentiles) and I5R (difference between 95th and 5th percentiles)

In each case, compute the ratio of IQR/I15R/I5R for predictive distributions based on the univariate (numerator) and multivariate (denominator) methods. We also noted (in parentheses) the number of times in each comparison of 22 regions that the multivariate method gave the smaller IQR/I15R/I5R.

	IQR	I15R	I5R
DJFA2	1.11 (13)	1.09 (12)	1.12 (15)
DJFB2	1.04 (13)	1.04 (14)	1.05 (12)
JJAA2	1.05 (13)	1.04 (14)	1.00 (14)
JJAB2	1.10 (15)	1.08 (16)	1.08 (14)

Overall, predictive distributions based on the multivariate method are tighter than those based on the univariate method. However, individual ratios vary widely.

The following regions have consistently tighter predictions under the univariate method: SAU, SAH, SEA

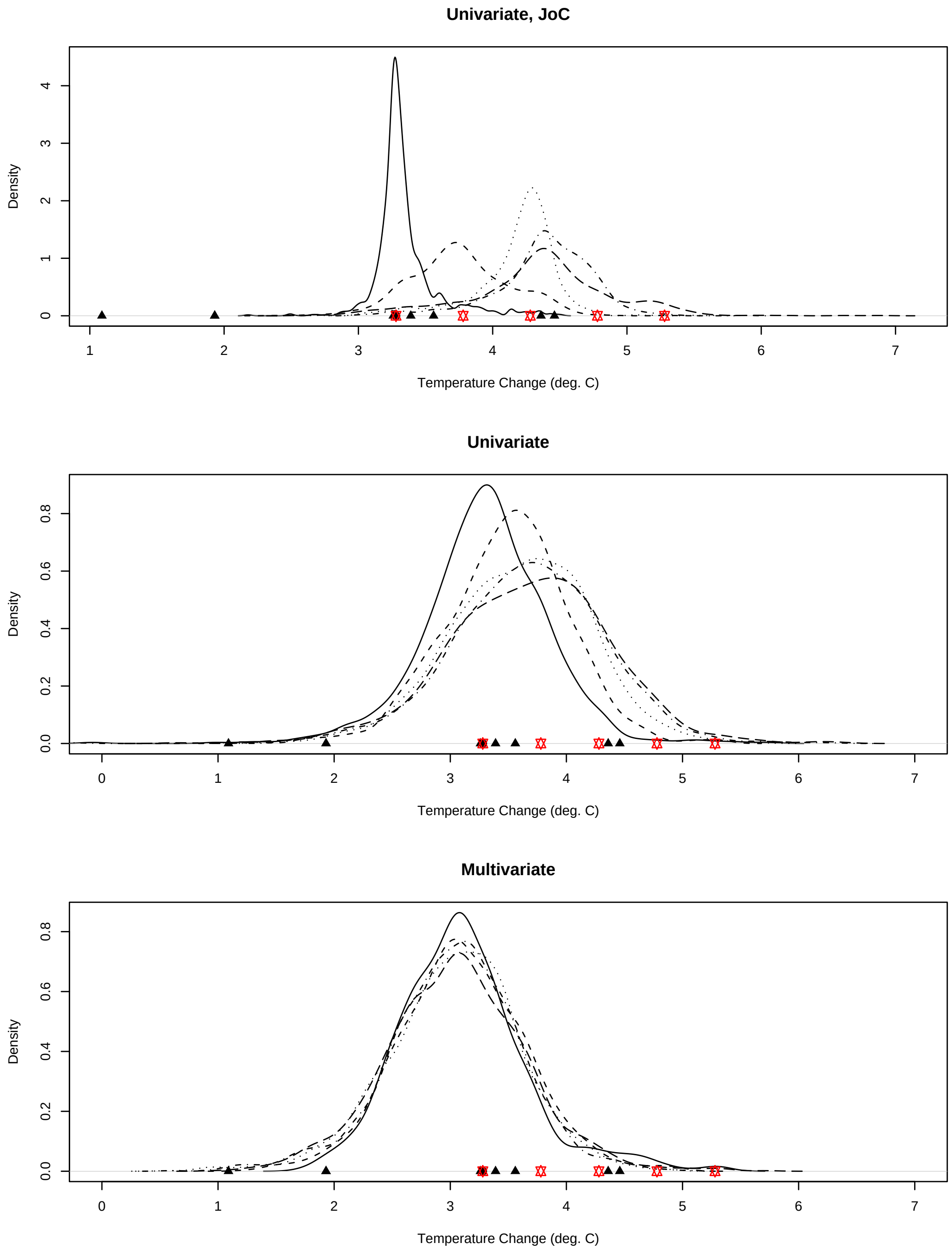
The following regions have consistently tighter predictions under the multivariate method: AMZ, WNA, CNA, ENA, NEU, EAS, CAS, TIB, NAS

Sensitivity Analysis

In our earlier analysis (*J. Climate 2005*), we noted that sometimes the posterior predictive distribution is very tight, but also highly sensitive to small perturbations of the data.

The analyses posed in this paper do not seem to suffer nearly so badly.

As an example, consider the WAF region under JJA-A2. We moved just one of the data points into five different values (represented by red dots) and recomputed the posterior densities using each of the three methods.



Sensitivity of different statistical approaches (WAF, JJA, A2)

Finally, the actual predictive distributions...

We use the multivariate approach to calculate predictive distribution functions for the temperature change in each of 22 regions, two seasons and two emissions scenarios.

	Q.025	Q.25	Mean	Q.75	Q.975
NAU	2.40	3.00	3.30	3.50	4.30
SAU	2.10	2.70	2.90	3.10	4.00
AMZ	2.10	2.70	3.00	3.30	4.20
SSA	2.00	2.60	2.90	3.10	4.00
CAM	2.40	3.00	3.30	3.50	4.40
WNA	3.00	3.80	4.20	4.50	5.40
CNA	3.00	3.70	4.10	4.40	5.40
ENA	3.10	3.90	4.30	4.70	5.60
ALA	5.00	6.30	7.00	7.70	9.20
GRL	5.60	6.50	6.90	7.40	8.30
MED	2.30	3.00	3.30	3.60	4.40
NEU	3.70	4.60	5.00	5.50	6.50
WAF	2.70	3.30	3.50	3.70	4.60
EAF	2.30	2.90	3.20	3.40	4.30
SAF	2.40	3.00	3.30	3.50	4.30
SAH	3.00	3.70	4.00	4.20	5.10
SEA	1.90	2.50	2.70	3.00	3.80
EAS	3.40	4.00	4.30	4.70	5.50
SAS	2.90	3.50	3.80	4.00	4.90
CAS	3.10	3.80	4.10	4.40	5.30
TIB	3.90	4.60	4.90	5.20	6.00
NAS	4.60	5.80	6.40	7.00	8.10

Predictive distribution for temperature change: DJF, A2, Multivariate model.

	Q.025	Q.25	Mean	Q.75	Q.975
NAU	1.70	2.10	2.30	2.50	3.00
SAU	1.40	1.80	2.00	2.20	2.70
AMZ	1.40	1.90	2.10	2.30	2.80
SSA	0.90	1.40	1.70	1.90	2.50
CAM	1.60	2.00	2.30	2.50	2.90
WNA	2.20	2.80	3.00	3.30	3.80
CNA	2.20	2.70	3.00	3.30	3.90
ENA	2.20	2.90	3.20	3.40	4.10
ALA	3.80	4.80	5.30	5.80	7.00
GRL	4.20	4.90	5.20	5.50	6.30
MED	1.70	2.10	2.40	2.60	3.10
NEU	2.50	3.30	3.60	4.00	4.80
WAF	1.80	2.20	2.40	2.60	3.10
EAF	1.60	2.00	2.20	2.40	2.90
SAF	1.70	2.10	2.30	2.50	2.90
SAH	2.10	2.60	2.80	3.00	3.50
SEA	1.30	1.70	1.90	2.10	2.50
EAS	2.50	3.00	3.30	3.60	4.10
SAS	2.00	2.40	2.60	2.80	3.30
CAS	2.30	2.70	3.00	3.20	3.80
TIB	2.90	3.40	3.60	3.90	4.40
NAS	3.70	4.50	4.90	5.30	6.30

Predictive distribution for temperature change: DJF, B2, Multivariate model.

	Q.025	Q.25	Mean	Q.75	Q.975
NAU	2.10	2.90	3.30	3.70	4.60
SAU	1.70	2.40	2.80	3.10	3.90
AMZ	2.90	3.60	4.00	4.40	5.20
SSA	1.80	2.40	2.70	3.00	3.70
CAM	2.60	3.10	3.50	3.80	4.50
WNA	3.60	4.30	4.70	5.00	5.70
CNA	3.70	4.50	4.80	5.20	6.00
ENA	3.40	4.10	4.50	4.80	5.60
ALA	2.80	3.50	3.80	4.10	4.90
GRL	3.10	4.10	4.50	4.90	5.90
MED	3.90	4.60	5.00	5.30	6.10
NEU	3.10	3.70	4.00	4.40	5.00
WAF	2.30	2.90	3.20	3.50	4.10
EAF	2.50	3.10	3.40	3.70	4.30
SAF	2.90	3.50	3.80	4.10	4.70
SAH	3.30	3.90	4.20	4.50	5.10
SEA	1.60	2.30	2.60	2.90	3.60
EAS	2.90	3.60	3.90	4.20	4.80
SAS	1.30	2.00	2.40	2.70	3.50
CAS	4.20	4.90	5.30	5.60	6.40
TIB	3.60	4.30	4.60	5.00	5.70
NAS	3.80	4.50	4.90	5.30	6.10

Predictive distribution for temperature change: JJA, A2, Multivariate model.

	Q.025	Q.25	Mean	Q.75	Q.975
NAU	1.20	1.90	2.20	2.50	3.30
SAU	0.90	1.50	1.80	2.10	2.80
AMZ	1.90	2.50	2.80	3.00	3.80
SSA	1.00	1.60	1.90	2.20	2.90
CAM	1.70	2.30	2.50	2.80	3.50
WNA	2.50	3.10	3.50	3.70	4.40
CNA	2.60	3.20	3.50	3.80	4.50
ENA	2.40	3.00	3.30	3.50	4.30
ALA	1.90	2.60	2.90	3.20	3.90
GRL	2.20	3.00	3.50	3.90	4.90
MED	2.80	3.40	3.70	3.90	4.60
NEU	2.10	2.70	3.00	3.30	4.00
WAF	1.20	1.80	2.10	2.40	3.10
EAF	1.50	2.00	2.30	2.60	3.30
SAF	1.80	2.40	2.60	2.90	3.60
SAH	2.20	2.70	3.00	3.20	3.90
SEA	0.90	1.60	1.80	2.10	2.80
EAS	2.10	2.60	2.90	3.10	3.80
SAS	0.70	1.30	1.60	1.90	2.60
CAS	3.00	3.60	3.90	4.10	4.90
TIB	2.50	3.10	3.40	3.70	4.40
NAS	2.60	3.30	3.60	3.90	4.70

Predictive distribution for temperature change: JJA, B2, Multivariate model.