Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles

Claudia Tebaldi [1], Richard L. Smith [2] , Doug Nychka[3] and Linda O. Mearns [4]

December 29, 2003

Technical Description of the Gibbs Sampler

Univariate model

For the simplest model, with no correlation structure between X's and Y's and Gaussian likelihoods, we can write the joint density of θ, μ, ν, λ_i, Y_i $(1 \leq i \leq 9)$ and X_i $(0 \leq i \leq 9)$, as proportional to:

$$\prod_{i=1}^{N} \left[\lambda_i^{a-1} e^{-b\lambda_i} \cdot \lambda_i \theta^{1/2} \exp\left\{ -\frac{\lambda_i}{2} ((X_i - \mu)^2 + \theta(Y_i - \nu)^2) \right\} \right] \cdot \theta^{c-1} e^{-d\theta} \cdot \exp\left\{ -\frac{\lambda_0}{2} (X_0 - \mu)^2 \right\}. \tag{1}$$

From (1) consider the joint distribution of ν and θ, conditional on all other parameters being known. These are easily derived as

$$\nu | \theta, Y_1, \ldots, Y_N, \lambda_1, \ldots, \lambda_N \sim N\left(\tilde{\nu}, \left(\theta \sum_1^N \lambda_i \right)^{-1} \right) \tag{2}$$

$$\theta | Y_1, \ldots, Y_N, \lambda_1, \ldots, \lambda_N \sim \Gamma \left\{ c + \frac{N-1}{2}, d + \frac{1}{2} \sum_1^N \lambda_i (Y_i - \tilde{\nu}^2) \right\} \tag{3}$$

Also, by integrating out μ and ν but conditioning on θ in (1) we can write the posterior distribution of the precision parameters $\lambda_1, \ldots, \lambda_9$:

$$p(\lambda_1, \ldots, \lambda_N | \theta, X_0, \ldots, X_N, Y_1, \ldots, Y_N) \propto \left(\sum_0^N \lambda_i \right)^{-1/2} \left(\sum_1^N \lambda_i \right)^{-1/2} \cdot$$
$$\cdot \prod_{i=1}^{N} \left[\lambda_i^{a-1} e^{-b\lambda_i} \cdot \lambda_i \theta^{1/2} \exp\left\{ -\frac{\lambda_i}{2} ((X_i - \tilde{\mu})^2 + \theta(Y_i - \tilde{\nu})^2) \right\} \right]. \tag{4}$$

From (4), ignoring the first two factors, one can recognize that the posterior density of λ_i given θ is

$$\Gamma \left\{ a + 1, b + \frac{1}{2} ((X_i - \tilde{\mu})^2 + \theta(Y_i - \tilde{\nu})^2) \right\} \tag{5}$$

[1]Project Scientist, National Center for Atmospheric, Boulder, CO
[2]Professor, University of North Carolina, Chapel Hill, NC
[3]Senior Scientist, National Center for Atmospheric, Boulder, CO
[4]Senior Scientist, National Center for Atmospheric, Boulder, CO

Robust model with correlation between present and future AOGCM response

There is a simple way of characterizing Student-t distribution through the introduction of an auxiliary randomization. Rather than modeling X_i and Y_i directly as in (**??**) and (**??**), we write

$$X_i = \mu + s_i \epsilon_i \tag{6}$$

and

$$Y_i = \nu + \beta(X_i - \mu) + t_i \epsilon'_i / \sqrt{\theta} \tag{7}$$

where s_i, t_i are independently distributed as χ^2 with ϕ degrees of freedom, and $\epsilon_i, \epsilon'_i, \theta$ have the same distributions specified for the basic univariate model. This modeling choice exploits the fact that if a random variable x has a standard normal distribution, y has a $\Gamma(\phi/2, 1/2)$ (also known as a χ^2 distribution with ϕ degrees of freedom), and x and y are mutually independent, then

$$z = \frac{x}{(y/\phi)^{1/2}}$$

has a Student-t density with ϕ degrees of freedom. A feature of this modeling choice is that by imposing degenerate prior distributions for $\beta = 0$, $t_i = 1$ and $s_i = 1$ this model reduces to the basic univariate model of Section **??**. Note that in this modified version of the model we have to fix the value of $\theta = 1$ to make the parameters identifiable.

We write the joint density of the random variables in this model, as proportional to:

$$\theta^{c+N/2-1} e^{-d\theta} \frac{\left(\frac{\phi}{2}\right)^{N\phi}}{\left\{\Gamma\left(\frac{\phi}{2}\right)\right\}^{2N}} e^{-\frac{\lambda_0}{2}(X_0-\mu)^2} \cdot \tag{8}$$

$$\cdot \prod_{i=1}^{N} \left[\lambda_i^a e^{-b\lambda_i} (s_i t_i)^{(\phi-1)/2} e^{-\phi(s_i+t_i)/2} \cdot \exp\left[-\frac{\lambda_i}{2} \{ s_i (X_i - \mu)^2 + \theta t_i (Y_i - \nu - \beta(X_i - \mu))^2 \} \right] \right].$$

From it, a series of full conditional distributions, on which the Gibbs sampler's iterations are based, is easily derived.

We conducted several separate analyses by varying the degrees of freedom ϕ of the Student-t distribution over the set $\{2, 4, 8, 16, 32, 64, 200\}$. Low values of the degrees of freedom accentuate the heavy tail nature of the distribution, accommodating for larger outliers. The higher the value of the degrees of freedom the closer the approximation to a Gaussian assumption. Although in principle one could estimate the degrees of freedom along with the other parts of the model we found that this added complexity was difficult to handle given the small sample sizes of the model output data. As an simpler alternative we choose to investigate the sensitivity of the conclusions to this parameter.

Gibbs sampler implementation

The Gibbs sampler can be coded so as to simulate iteratively from the following sequence of full conditional distributions (the ... at the right of the conditioning sign refer to all the random variables in the model, apart from the parameter to be drawn, from which the definition "full conditional"):

$$\lambda_i|\ldots \sim \Gamma\left(a+1, b+\frac{s_i}{2}(X_i-\mu)^2+\frac{\theta t_i}{2}\{Y_i-\nu-\beta(X_i-\mu)\}^2\right), \quad (9)$$

$$s_i|\ldots \sim \Gamma\left(\frac{\phi+1}{2}, \frac{\phi+\lambda_i(X_i-\mu)^2}{2}\right), \quad (10)$$

$$t_i|\ldots \sim \Gamma\left(\frac{\phi+1}{2}, \frac{\phi+\theta\lambda_i\{Y_i-\nu-\beta(X_i-\mu)\}^2}{2}\right), \quad (11)$$

$$\mu|\ldots \sim N\left(\tilde{\mu}, \left(\sum s_i\lambda_i+\theta\beta^2\sum t_i\lambda_i+\lambda_0\right)^{-1}\right), \quad (12)$$

$$\nu|\ldots \sim N\left(\tilde{\nu}, \left(\theta\sum t_i\lambda_i\right)^{-1}\right), \quad (13)$$

$$\beta|\ldots \sim N\left(\tilde{\beta}, \left(\theta\sum t_i\lambda_i(X_i-\mu)^2\right)^{-1}\right), \quad (14)$$

$$\theta|\ldots \sim \Gamma\left(c+\frac{N}{2}, d+\frac{1}{2}\sum t_i\lambda_i\{Y_i-\nu-\beta(X_i-\mu)\}^2\right). \quad (15)$$

In what precede, we have used the following shorthand notation:

$$\tilde{\mu} = \frac{\sum s_i\lambda_i X_i - \theta\beta\sum\lambda_i t_i(Y_i-\nu-\beta X_i)+\lambda_0 X_0}{\sum s_i\lambda_i+\theta\beta^2\sum\lambda_i t_i+\lambda_0}, \quad (16)$$

$$\tilde{\nu} = \frac{\sum t_i\lambda_i\{Y_i-\beta(X_i-\mu)}{\sum t_i\lambda_i}, \quad (17)$$

$$\tilde{\beta} = \frac{\sum t_i\lambda_i(Y_i-\nu)(X_i-\mu)}{\sum t_i\lambda_i(X_i-\mu)^2}. \quad (18)$$

We have shown here the sampling distributions for the robust model that includes correlation between X_i and Y_i in the form of a regression law. We can fix $s_i = t_i = 1$, $\beta = 0$ and sequence only through draws of $\lambda_1, \ldots, \lambda_N$, μ, ν and θ in order to simulate from the posterior distributions of the basic univariate Gaussian model presented in Section **??**.

The actual implementation of the Gibbs sampler consists of cycling through the following sequence of random variate generations:

1. Given the current values of all the remaining parameters, sample a value of μ from the full conditional distribution $[\mu|\ldots]$.

2. Given the new value of μ, and the current values of the remaining parameters, sample a value of ν from the full conditional distribution $[\nu|\ldots]$.

3. Given the new values of μ, ν, and the current values of the remaining parameters, sample a value of λ_i from the full conditional distribution $[\lambda_i|\ldots]$. Repeat for $i = 1, \ldots, 9$.

4. Given the new values of $\mu, \nu, \lambda_1, \ldots, \lambda_9$, and the current values of the remaining parameters, sample a value of θ from the full conditional $[\theta|\ldots]$.

5. Given the new values of $\mu, \nu, \lambda_1, \ldots, \lambda_9, \theta$, and the current values of the remaining parameters, sample a value of β from the full conditional $[\beta|\ldots]$.

6. Given the new values of $\mu, \nu, \lambda_1, \ldots, \lambda_9, \theta, \beta$, and the current values of the remaining parameters, sample a value of s_i from the full conditional $[s_i|\ldots]$. Repeat for $i = 1, \ldots, 9$.

7. Given the new values of $\mu, \nu, \lambda_1, \ldots, \lambda_9, \theta, \beta, s_1, \ldots, s_9$, sample a value of t_i from the full conditional $[t_i|\ldots]$. Repeat for $i = 1, \ldots, 9$.

8. Repeat from the top.

Markov Chain Monte Carlo implementation

For all models' estimation, we ran the sampler for a total of $500,000$ iterations, discarding the first half of the simulated values and saving only one draw every 50. Thus, we base our conclusions on a total of $5,000$ values for each parameter, representing a sample from its posterior distribution. The convergence of the Markov chain to its stationary distribution was verified by standard diagnostic tools (Best et al. 1995).

References

Allen, M.R., Stott, P.A., Mitchell, J.F.B., Schnur, R. and T.L. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617-620.

Best N.G., M. K. Cowles, and S. K. Vines, 1995: *CODA Convergence Diagnosis and Output Analysis software for Gibbs Sampler output: Version 0.3.*

Bernardo, J. M., and A. F. Smith, 1994: *Bayesian Theory.*, John Wiley & Sons.

Cubasch, U., G.A. Meehl, G.J. Boer, R.J.Stouffer, M. Dix, A. Noda, C.A. Senior, S. Raper, and K.S. Yap, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 525-582.

Forest, C.E., Stone, P.H., Sokolov, A.P., Allen, M.R. and M.D. Webster, 2002: Quantifying Uncertainties in Climate System Properties with the Use of Recent Climate Observations. *Science.*, **295**, 113-117.

Gilks, W.R., S. Richardson, and D. J. Spiegelhalter, 1996: *Markov Chain Monte Carlo Methods in Practice.*,Chapman & Hall.

Giorgi F., and R. Francisco, 2000: Evaluating uncertainties in the prediction of regional climate change. *Geophysical Research Letters*, **27**, 1295-1298.

Giorgi F., and L.O. Mearns, 2002: Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method. *J. Climate,* **15**, n. 10, 1141-1158.

Giorgi F., and L.O. Mearns, 2003: Probability of Regional Climate Change Calculated using the Reliability Ensemble Averaging (REA) Method. *Geophysical Research Letters*, **30**, n.12, 311-314.

Giorgi F., and co-authors, 2001a: Regional climate information: Evaluation and projections. In J. T. Houghton et al. (eds.) *Climate Change 2001: The Scientific Basis.* Contribution of Working Group I to the Third Assessment Report of the Intergovenmental Panel on Climate Change. Chapter 10. Cambridge: Cambridge University Press. pp. 583–638.

Giorgi, F., and co-authors, 2001b: Emerging patterns of simulated regional climatic changes f or the 21st century due to anthropogenic forcings. *Geophys. Res. Lett.* **28**, 3317-3321.

Hegerl, G.C., P.A. Stott, M.R. Allen, J.F.B. Mitchell, S.F.B. Tett and U. Cubasch, 2000: Optimal detection and attribution of climate change: sensitivity of results to climate model differences. *Climate Dynamics,* **16**, 737-754.

Knutson, T. R., T. L. Delworth, K. W. Dixon, and R. J. Stouffer, 1999: Model assessment of regional surface temperature trends (1949-97). *J. Geophys. Res.,* **104**, 30981-30996.

McAvaney, B.J, C.Covey, S. Joussaume, V. Kattsov, A. Kitoh, W. Ogana, A.J. Pittman, A.J. Weaver, R.A. Wood, and Z.-C. Zhao, 2001: Model evaluation. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 471-524.

Mearns, L.O., M. Hulme, T. R. Carter, R. Leemans, M.Lal, and P. Whetton, 2001: Climate scenario development. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 739-768.

Meehl, G.A., G.J. Boer, C. Covey, M. Latif, and R.J.Stouffer, 2000: The Coupled Model Intercomparison Project (CMIP). *Bull. Am. Met. Soc.*, **81**, 313-318.

Nakicenovic, N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grbler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Raihi, A. Roehrl, H-H. Rogner, A. Sankovski, M. Schlesinger, P.

Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, Z. Dadi, 2000: IPCC Special Report on Emissions Scenarios, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 599 pp.

Nychka, D., and C. Tebaldi, 2003: Comment on 'Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method'. *J. Climate,* **16**, 883-884.

Raisanen, J., 1997: Objective comparison of patterns of CO_2 induced climate change in coupled GCM experiments. *Climate Dynamics*, **13**, 197-211.

Raisanen, J and T. N. Palmer, 2001: A Probability and Decision Model Analysis of a Multi-model Ensemble of Climate Change Simulations. *J. Climate,* **14**, 3212-3226.

Reilly, J., P. H. Stone, C. E. Forest, M. D. Webster, H. D. Jacoby, R. G. Prinn, 2001, Uncertainty in climate change assessments. *Science*, **293**, 5529,430-433.

Schneider, S. H., 2001: What is dangerous climate change?. *Nature*, **411**, 17-19.

Webster, M., 2003: Communicating climate change uncertainty to policy-makers and the public. *Climatic Change* **61**, 1-8.

Wigley, T. M. L., and S. C. B. Raper, 2001: Interpretation of high projections for global-mean warming. *Science*, **293**, 451-454.