

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

CLAUDIA TEBALDI

*Institute for the Study of Society and Environment, National Center for Atmospheric Research, Boulder, Colorado*

RICHARD L. SMITH

*Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina*

DOUG NYCHKA

*Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, Colorado*

LINDA O. MEARNS

*Institute for the Study of Society and Environment, National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 25 January 2004, in final form 11 November 2004)

### ABSTRACT

A Bayesian statistical model is proposed that combines information from a multimodel ensemble of atmosphere–ocean general circulation models (AOGCMs) and observations to determine probability distributions of future temperature change on a regional scale. The posterior distributions derived from the statistical assumptions incorporate the criteria of bias and convergence in the relative weights implicitly assigned to the ensemble members. This approach can be considered an extension and elaboration of the reliability ensemble averaging method. For illustration, the authors consider the output of mean surface temperature from nine AOGCMs, run under the A2 emission scenario from the Synthesis Report on Emission Scenarios (SRES), for boreal winter and summer, aggregated over 22 land regions and into two 30-yr averages representative of current and future climate conditions. The shapes of the final probability density functions of temperature change vary widely, from unimodal curves for regions where model results agree (or outlying projections are discounted) to multimodal curves where models that cannot be discounted on the basis of bias give diverging projections. Besides the basic statistical model, the authors consider including correlation between present and future temperature responses, and test alternative forms of probability distributions for the model error terms. It is suggested that a probabilistic approach, particularly in the form of a Bayesian model, is a useful platform from which to synthesize the information from an ensemble of simulations. The probability distributions of temperature change reveal features such as multimodality and long tails that could not otherwise be easily discerned. Furthermore, the Bayesian model can serve as an interdisciplinary tool through which climate modelers, climatologists, and statisticians can work more closely. For example, climate modelers, through their expert judgment, could contribute to the formulations of prior distributions in the statistical model.

### 1. Introduction

Numerical experiments based on atmosphere–ocean general circulation models (AOGCMs) are one of the primary tools used in deriving projections for future climate change. However, the strengths and weaknesses that individual AOGCMs display have led authors of

model evaluation studies to state that “no single model can be considered ‘best’ and it is important to utilize results from a range of coupled models” (McAvaney et al. 2001). In this paper, we propose a probabilistic approach to the synthesis of climate projections from different AOGCMs, in order to produce probabilistic forecasts of climate change.

Determining probabilities of future global temperature change has flourished as a research topic in recent years (Schneider 2001; Wigley and Raper 2001; Forest et al. 2002; Allen et al. 2000). This work has been based for the most part on the analysis of energy balance or

---

*Corresponding author address:* Claudia Tebaldi, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000.  
E-mail: tebaldi@ucar.edu

reduced climate system models that facilitate many different model integrations and hence allow the estimation of a distribution of climate projections. However, the focus on low-dimensional models prevents a straightforward extension of this work to regional climate change analyses, which are the indispensable input for impacts research and decision making. In this study we draw on more conventional AOGCM experiments in order to address specifically regional climate change. We recognize, though, that probabilistic forecasts at the level of resolution of the typical AOGCM is a feat of enormous complexity. The high dimensionality of the datasets, the scarcity of observations in many regions of the globe, and the limited length of the observational records have required ad hoc solutions in detection and attribution studies, and it is hard to structure a full statistical approach able to encompass these methods and add the further dimension of the super-ensemble dataset. At present, we offer a first step toward a formal statistical treatment of the problem, by examining regionally averaged temperature signals. This way, we greatly simplify the dimensionality of the problem, while still addressing the need for regionally differentiated analyses.

#### *a. Model bias and model convergence*

Recently, Giorgi and Mearns's reliability ensemble average (REA) method (Giorgi and Mearns 2002) quantified the two criteria of *bias* and *convergence* for multimodel evaluation, and produced estimates of regional climate change and model reliabilities through a weighted average of the individual AOGCM results. The REA weights contain a measure of model bias with respect to current climate and a measure of model convergence, defined as the deviation of the individual projection of change with respect to the central tendency of the ensemble. Thus, models with small bias and whose projections agree with the consensus receive larger weights. Models with lesser skill in reproducing the observed conditions and appearing as outliers with respect to the majority of the ensemble members receive less weight. In a subsequent note Nychka and Tebaldi (2003) show that the REA method is in fact a conventional statistical estimate for the center of a distribution that departs from a Gaussian shape by having heavier tails. Thus, although the REA estimates were proposed by Giorgi and Mearns as a reasonable quantification of heuristic criteria, there exists a formal statistical model that can justify them as an optimal procedure. The research in this paper is motivated by this statistical insight. Here we start from the same data used by Giorgi and Mearns (i.e., 30-yr regional climate averages representative of current and future conditions) and propose statistical models that recover the REA framework but are flexible in the definition of bias and convergence and are easily generalizable to richer data structures. The main results of our analysis consist of

probability distributions of temperature change at regional scales that reflect a relative weighing of the different AOGCMs according to the two criteria.

#### *b. A Bayesian approach to climate change uncertainty*

There is interest in the geosciences in moving from single-value predictions to probabilistic forecasts, in so far as presenting a probability distribution of future climate is a more flexible quantification for drawing inferences and facilitating decision making (Reilly et al. 2001; Webster 2003; Dessai and Hulme 2003). Bayesian methods are not the only option when the goal is a probabilistic representation of uncertainty, but are a natural way to do so in the context of climate change projections.

The choice of prior distributions is crucial, when adopting a Bayesian approach, representing the step in the analysis more open to subjective evaluations. Within the climate community, different viewpoints have been expressed regarding the use of expert judgement in quantifying uncertainty. For example, Wigley and Raper (2001) explicitly chose prior distributions to be consistent with their own expert judgement of the uncertainty of key parameters such as climate sensitivity, and Reilly et al. (2001) argued that such assessments should be part of the Intergovernmental Panel on Climate Change (IPCC) process for quantifying uncertainty in climate projections. Opposing this view, Allen et al. (2001) argued that "no method of assigning probabilities to a 100-year climate forecast is sufficiently widely accepted and documented in the refereed literature to pass the extensive IPCC review process." (They suggested that there might be better prospects of success with a 50-yr forecast, since over this time frame there is better agreement among models with respect to both key physical parameters and the sensitivity to different emissions scenarios.) Following this philosophy, Forest et al. (2002) stated "an objective means of quantifying uncertainty . . . is clearly desirable" and argued that this was achievable by choosing parameters "that produce simulations consistent with 20th-century climate change." However, their analysis did not use multimodel ensembles. We view the present approach as an extension of the philosophy reflected in the last two quotations, where we adopt uninformative prior distributions but use both model-generated and observational data to calculate meaningful posterior distributions. Another point in favor of uninformative prior distributions is that they typically lead to parameter estimates similar to non-Bayesian approaches such as maximum likelihood. However, Bayesian methods are more flexible when combining different sources of uncertainty, such as those derived from present-day and future climate model runs, and we regard this as a practical justification for adopting a Bayesian approach.

### c. Outline

Section 2 contains the description of the basic statistical model and its extensions. Prior distributions for the parameters of interest, and distributional assumptions for the data, conditional on these parameters, are described. We then present analytical approximations to the posterior distribution in order to gain insight into the nature of the statistical model and its results. In section 3 the model is applied to the same suite of AOGCM experiments as in Giorgi and Mearns (2002), and some findings from the posterior distributions are presented. In section 4 we conclude with a discussion of our model assumptions and what we consider promising directions for extending this work. The Markov chain Monte Carlo (MCMC) method used to estimate the posterior distributions is described in the appendix.

## 2. Statistical models for AOGCM projections

Adopting the Bayesian viewpoint the uncertain quantities of interest become the parameters of the statistical model and are treated as random variables. A *prior probability distribution* for them is specified independently of the data at hand. The *likelihood* component of the statistical model specifies the conditional distribution of the data, given the model parameters. Through Bayes's theorem prior and likelihood are combined into the *posterior distribution* of the parameters, given the data. Formally, let  $\Theta$  be the vector of model parameters, and  $p(\Theta)$  their prior distribution. The data  $D$ , under the assumptions formulated in the statistical model, has likelihood  $p(D|\Theta)$ . Bayes's theorem states that

$$p(\Theta|D) \propto p(\Theta) \cdot p(D|\Theta),$$

where  $p(\Theta|D)$  is the posterior distribution of the parameters and forms the basis of any formal statistical inference about them.

When the complexity of  $p(\Theta|D)$  precludes a closed-form solution, which is the case in our application, an empirical estimate of the posterior distribution can be obtained through MCMC simulation. MCMC techniques are efficient ways of simulating samples from the posterior distribution, bypassing the need of computing it analytically, and inference can be drawn through smoothed histograms and numerical summaries based on the MCMC samples.

### a. The model for a single region

In this section we present the analytical form of our basic statistical model. We list first the distributional assumptions for the data (likelihood), then the priors for all the model parameters. We then present conditional approximations to the posterior that will help to interpret our results. Throughout, let  $X_i$  and  $Y_i$  denote

the temperature simulated by AOGCM  $i$ , seasonally and regionally averaged and aggregated into two 30-yr means representative of current and future climate, respectively.

#### 1) LIKELIHOODS

We assume Gaussian distributions for  $X_i, Y_i$ :

$$X_i \sim N(\mu, \lambda_i^{-1}), \quad (1)$$

$$Y_i \sim N(\nu, (\theta\lambda_i)^{-1}), \quad (2)$$

where the notation  $N(\mu, \lambda^{-1})$  indicates a Gaussian distribution with mean  $\mu$  and variance  $1/\lambda$ . Here  $\mu$  and  $\nu$  represent the true values of present and future temperature in a specific region and season. A key parameter of interest will be  $\Delta T \equiv \nu - \mu$ , representing the expected temperature change. The parameter  $\lambda_i$ , reciprocal of the variance, is referred to as the *precision* of the distribution of  $X_i$ . To allow for the possibility that  $Y_i$  has different precision from  $X_i$ , we parameterize its distribution by the product  $\theta\lambda_i$  where  $\theta$  is an additional parameter, common to all AOGCMs. A more general model might allow for  $\theta$  to be different in different models, but that is not possible in the present setup given the limited number of data points.

The assumptions underlying (1) and (2) are that the AOGCM responses have a symmetric distribution, whose center is the "true value" of temperature, but with an individual variability, to be regarded as a measure of how well each AOGCM approximates the climate response to the given set of natural and anthropogenic forcings. The assumption of a symmetric distribution around the true value of temperature for the suite of multimodel responses has been implicitly supported by the Coupled Model Intercomparison Project (CMIP) studies (Meehl et al. 2000) where better validation properties have been demonstrated for the mean of a superensemble rather than the individual members. The additional assumption in our model that the single AOGCM's realizations are centered around the true value could be easily modified in the presence of additional data. For example, if single-model ensembles were available, then an AOGCM-specific random effect could be incorporated.

We model the likelihood of the observations of current climate as

$$X_0 \sim N(\mu, \lambda_0). \quad (3)$$

Here  $\mu$  is the same as in (1), but  $\lambda_0$  is of a different nature from  $\lambda_1, \dots, \lambda_9$ . While the latter are measures of model-specific precision, and depend on the numerical approximations, parameterizations, grid resolutions of each AOGCM,  $\lambda_0$  is a function of the natural variability specific to the season, region and time average applied to the observations. In our model we fix the value of the parameter  $\lambda_0$ , using estimates of regional natural vari-

ability from Giorgi and Mearns (2002). The parameter  $\lambda_0$  could be treated as a random variable as well if our data contained a long record of observations that we could use for its estimation.

## 2) PRIOR DISTRIBUTIONS

The model described by (1)–(3) is formulated as a function of the parameters  $\mu, \nu, \theta, \lambda_1, \dots, \lambda_9$ . Following the arguments presented in section 1b, we choose uninformative priors as follows.

- Precision parameters  $\lambda_i, i = 1, \dots, 9$  have Gamma prior densities [indicated by  $\text{Ga}(a, b)$ ], of the form

$$\frac{b^a}{\Gamma(a)} \lambda_i^{a-1} e^{-b\lambda}$$

with  $a, b$  known and chosen so that the distribution will have a large variance over the positive real line. Similarly,  $\theta \sim \text{Ga}(c, d)$ , with  $c, d$  known. These are standard prior choices for the precision parameters of Gaussian distributions. In particular they are *conjugate*, in the sense of having the same general form as the likelihood considered as a function of  $\lambda_i$ , thus ensuring the computational tractability of the model. We choose  $a = b = c = d = 0.001$ , that translate into Gamma distributions with mean 1 and variance 1000. By being extremely diffuse the distributions have the noninformative quality that we require in our analysis.

- True climate means,  $\mu$  and  $\nu$  for present and future temperature, respectively, have uniform prior densities on the real line. Even if these priors are improper (i.e., do not integrate to 1), the form of the likelihood model ensures that the posterior is a proper density function. Alternative analyses in which the prior distribution is restricted to a finite, but sufficiently large, interval, do not in practice produce different results.

## 3) POSTERIOR DISTRIBUTIONS

We apply Bayes's theorem to the likelihood and priors specified above. The resulting joint posterior density for the parameters  $\mu, \nu, \theta, \lambda_1, \dots, \lambda_9$  is given by, up to a normalizing constant,

$$\prod_{i=1}^9 \lambda_i^{a-1} e^{-b\lambda_i} \times \lambda_i \theta^{1/2} \exp\left\{-\frac{\lambda_i}{2}[(X_i - \mu)^2 + \theta(Y_i - \nu)^2]\right\} \times \theta^{c-1} e^{-d\theta} \times \exp\left\{-\frac{\lambda_0}{2}(X_0 - \mu)^2\right\}. \quad (4)$$

The distribution in Eq. (4) is not a member of any known parametric family and we cannot draw inference from its analytical form. The same is true for the marginal posterior distributions of the individual parameters, that we cannot compute from (4) through closed-

form integrals. Therefore, MCMC simulation is used to generate a large number of sample values from (4) for all parameters, and approximate all the summaries of interest from sample statistics. We implement the MCMC simulation through a Gibbs sampler, and details of the algorithm are given in the appendix.

Here we gain some insight as to how the posterior distribution synthesizes the data and the prior assumptions by fixing some groups of parameters and considering the conditional posterior for the others.

For example, the distribution of  $\mu$  fixing all other parameters is a Gaussian distribution with mean

$$\tilde{\mu} = \left( \sum_{i=0}^9 \lambda_i X_i \right) / \left( \sum_{i=0}^9 \lambda_i \right) \quad (5)$$

and variance

$$\left( \sum_{i=0}^9 \lambda_i \right)^{-1}. \quad (6)$$

In a similar fashion, the conditional distribution of  $\nu$  is Gaussian with mean

$$\tilde{\nu} = \left( \sum_{i=1}^9 \lambda_i Y_i \right) / \left( \sum_{i=1}^9 \lambda_i \right) \quad (7)$$

and variance

$$\left( \theta \sum_{i=1}^9 \lambda_i \right)^{-1}. \quad (8)$$

The forms (5) and (7) are analogous to the REA results, being weighted means of the nine AOGCMs and the observation, with weights  $\lambda_1, \dots, \lambda_9, \lambda_0$ , respectively. As in the case of the REA method, these weights will depend on the data, but a fundamental difference is that in our analysis they are random quantities, and thus we account for the uncertainty in their estimation. Such uncertainty will inflate the width of the posterior distributions of  $\nu, \mu$ , and thus also  $\Delta T$ .

An approximation to the mean of the posterior distribution of the  $\lambda_i$ s, for  $i = 1, \dots, 9$ , is

$$E(\lambda_i | \{X_0, \dots, X_9, Y_1, \dots, Y_9\}) \approx \frac{a + 1}{b + \frac{1}{2}[(X_i - \tilde{\mu})^2 + \theta(Y_i - \tilde{\nu})^2]}. \quad (9)$$

The form of (9) shows that the individual weight  $\lambda_i$  is large provided both  $|X_i - \tilde{\mu}|$  and  $|Y_i - \tilde{\nu}|$  are small. These two quantities correspond to the *bias* and *convergence* criteria, respectively. Here  $|Y_i - \tilde{\nu}|$  measures the distance of the  $i$ th model future response from the overall average response, and so has characteristics similar to the convergence measure in Giorgi and Mearns (2002). The important difference from REA for this model is that the distance is based on the future

projection ( $Y_i$ ) rather than the temperature change ( $Y_i - X_i$ ). As for the bias term, notice that in the limit, if we let  $\lambda_0 \rightarrow \infty$  (we model the observation  $X_0$  as an extremely precise estimate of the true temperature  $\mu$ ),  $\bar{\mu} \rightarrow X_0$ , and the bias term becomes in the limit  $|X_i - X_0|$ , the same definition of bias as in the REA analysis. The form of (9) also reveals how by choosing  $a = b = 0.001$  we ensure that the contribution of the prior assumption to (9) is negligible.

*b. Introducing correlation between present and future climate responses within each AOGCM and robustifying the model*

The model presented in section 2a reproduces the basic features of the REA method. In this section we modify our model in two important directions. First, we relax the assumption of independence between  $X_i$  and  $Y_i$ , for a given climate model. We do so by linking  $X_i$  and  $Y_i$  through a linear regression equation; this is equivalent to assuming that  $(X_i, Y_i)$  are jointly normal, given the model parameters, with an unknown correlation coefficient that is left free to vary, a priori, between  $-1$  and  $+1$ .

Thus, adopting a slightly different notation than in section 2a, we express the statistical assumptions for  $X_i$  and  $Y_i$  as follows. Let

$$X_i = \mu + \eta_i, \quad (10)$$

where  $\eta_i \sim N(0, \lambda_i^{-1})$ , making (10) equivalent to (1). Then, given  $X_i$ ,  $\mu$ , and  $\nu$ , let

$$Y_i = \nu + \beta_x(X_i - \mu) + \xi_i/\sqrt{\theta}, \quad (11)$$

where  $\xi_i \sim N(0, \lambda_i^{-1})$ .

In this model we use the priors listed in section 2a for the parameters  $\mu$ ,  $\nu$ ,  $\lambda_1, \dots, \lambda_9$ ,  $\theta$ , and we choose a uniform prior on the real line for  $\beta_x$ .

In Eq. (11)  $\beta_x$  introduces a direct (if positive) or inverse (if negative) relation between  $X_i - \mu$  and  $Y_i - \nu$ . By using a common parameter we force the sign of the correlation to be the same for all AOGCMs, a constraint that is required because only two data points (current and future temperature response) per model are available for a given region and season. The value of  $\beta_x$  has also implications in terms of the correlation between  $Y_i - X_i$ , the signal of temperature change produced by the  $i$ th AOGCM, and the quantity  $X_i - \mu$ , the model bias for current temperature. A value of  $\beta_x$  equal to 1 translates into conditional independence of these two quantities, while values greater than or less than 1 would imply positive or negative correlation between them. (In the REA analysis, Giorgi and Mearns treat these two quantities as independent, after computing empirical estimates of their correlations and finding that they are for the most part low in absolute value and not statistically significant.) But the most important feature of the model with correlation has to do with the

form of the posterior distribution for the  $\lambda_i$  parameters. By an approximation similar to what we use in section 2a for the basic model, we derive that the posterior mean has now the approximate form:

$$E(\lambda_i | \{X_0, \dots, X_9, Y_1, \dots, Y_9\}) \approx \frac{a + 1}{b + \frac{1}{2} \{(X_i - \bar{\mu})^2 + \theta[Y_i - \bar{\nu} - \beta_x(X_i - \bar{\mu})]^2\}}. \quad (12)$$

If  $\beta_x \approx 1$  the form in (12) uses a convergence term now similar to the REA method's, measuring the distance between  $\Delta T_i \equiv Y_i - X_i$  and the consensus estimate of temperature change  $\bar{\nu} - \bar{\mu}$ . We will compare the results from the basic model and this modified version, in order to assess how differently the two definitions of convergence reflect on the final posterior distributions of regional temperature changes.

A second concern is the sensitivity of our model to outlying data points. This is a consequence of hypothesizing normal distributions for  $X_i$  and  $Y_i$ , given the model parameters. Traditionally, robustness is achieved by choosing distributions with heavier tails than the Gaussian, and Student's  $t$  distributions with few degrees of freedom are often a convenient choice because of flexibility (the degrees of freedom parameter allowing to vary the degree of their departure from the Gaussian) and computational tractability. Thus, we model  $\eta_i$  and  $\xi_i$  in Eqs. (10) and (11) above as independently distributed according to a Student's  $t$  distribution, with  $\phi$  degrees of freedom. We perform separate analyses for  $\phi$  varying in the set  $\{1, 2, 4, 8, 16, 32, 64\}$ , the lower end of the range being associated with heavier tail distributions, the higher being equivalent to a Gaussian model. We are not a priori labeling some of the AOGCMs as outliers, but we examine the difference in the posterior distributions that ensues as a function of the statistical model assumed for the error terms. Varying the degrees of freedom can accommodate extreme projections to a larger or smaller degree and is equivalent to the convergence criterion being less or more stringent, respectively. This achieves at least qualitatively the same effect as the exponents in the REA's weights, by which the relative importance of the bias and convergence terms could be modified.

### 3. The analysis: Results from the MCMC simulation

#### a. Data

We analyze temperature responses under the A2 SRES emission scenario (Nakicenovic et al. 2000), just a portion of the dataset analyzed by the REA method in Giorgi and Mearns (2002). Output from nine different AOGCMs (see Table 1 and references therein) con-

TABLE 1. The nine AOGCMs whose output constitutes the data in our analysis

AOGCM	Reference	Climate sensitivity
CCSR-NIES/version 2	Emori et al. (1999)	4.53
MRI/version 2	Noda et al. (1999)	1.25
CCC/GCM2	Flato and Boer (2001)	3.59
CSIRO/Mk2	Gordon and O'Farrell (1997)	3.50
NCAR/CSM	Dai et al. (2001)	2.29
DOE-NCAR/PCM	Washington et al. (2000)	2.35
GFDL/R30-c	Knutson et al. (1999)	2.87
MPI-DMI/ECHAM4-OPYC	Stendel et al. (2000)	3.11
UKMO/HADCM3	Gordon et al. (2000)	3.38

sists of present and future average surface temperatures, aggregated over the 22 regions shown in Fig. 1, two seasons [December–January–February (DJF) and June–July–August (JJA)] and two 30-yr periods (1961–90 and 2071–2100). Thus, we will separately treat  $(X_i, Y_i)$ ,  $i = 1, \dots, 9$ , for each region and season. Observed temperature means for the 22 regions and the 1961–90 period ( $X_0$ ) were also taken from Giorgi and Mearns (2002), together with estimates of regional natural variability, interpretable as the standard deviation of the three-decadal average that  $X_0$  represents. We use these estimates (listed in Table 2) to determine the values of the parameter  $\lambda_0$ . (The dataset can be downloaded from [www.cgd.ucar.edu/~nychka/REA](http://www.cgd.ucar.edu/~nychka/REA).)

### b. Results for a group of representative regions

We present posterior distributions derived from the basic model (section 2a) and the model introducing correlation between  $X_i$  and  $Y_i$  (section 2b), both models assuming Gaussian distributions for the likelihoods. We

TABLE 2. Natural variability ( $^{\circ}\text{C}$ ) of observed temperature. These are taken from Giorgi and Mearns (2002). They were estimated by computing 30-yr moving averages of observed, detrended, regional mean temperatures over the twentieth century and taking the difference between the maximum and minimum values.

Region	DJF	JJA
NAU	0.40	0.50
SAU	0.50	0.25
AMZ	0.30	0.25
SSA	0.35	0.75
CAM	0.65	0.35
WNA	0.75	0.50
CNA	1.40	0.92
ENA	1.30	0.80
ALA	1.75	0.80
GRL	0.85	0.50
MED	0.40	0.60
NEU	1.00	0.90
WAF	0.45	0.50
EAF	0.42	0.50
SAF	0.40	0.50
SAH	0.45	0.75
SEA	0.40	0.25
EAS	0.85	0.75
SAS	0.40	0.25
CAS	1.00	0.75
TIB	0.85	0.75
NAS	1.00	0.70

choose to focus on a representative group of six regions [Alaska (ALA), east North America (ENA), southern South America (SSA), northern Europe (NEU), East Africa (EAF), northern Australia (NAU)], for both the winter (DJF) and summer (JJA) seasons.

### 1) TEMPERATURE CHANGE

Figures 2–4 show posterior distributions of temperature change  $\Delta T \equiv \nu - \mu$  in the six regions, for the

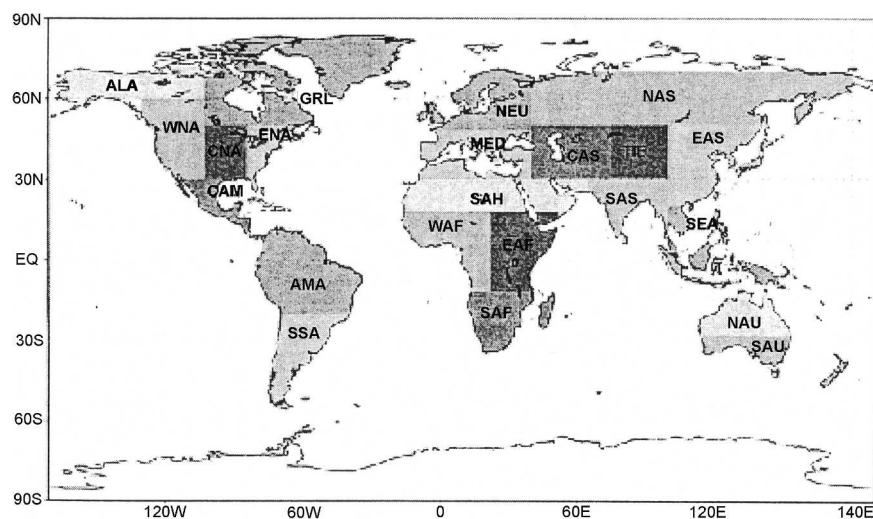


FIG. 1. The 22 regions into which the landmasses were discretized for both the REA analysis and our analysis.

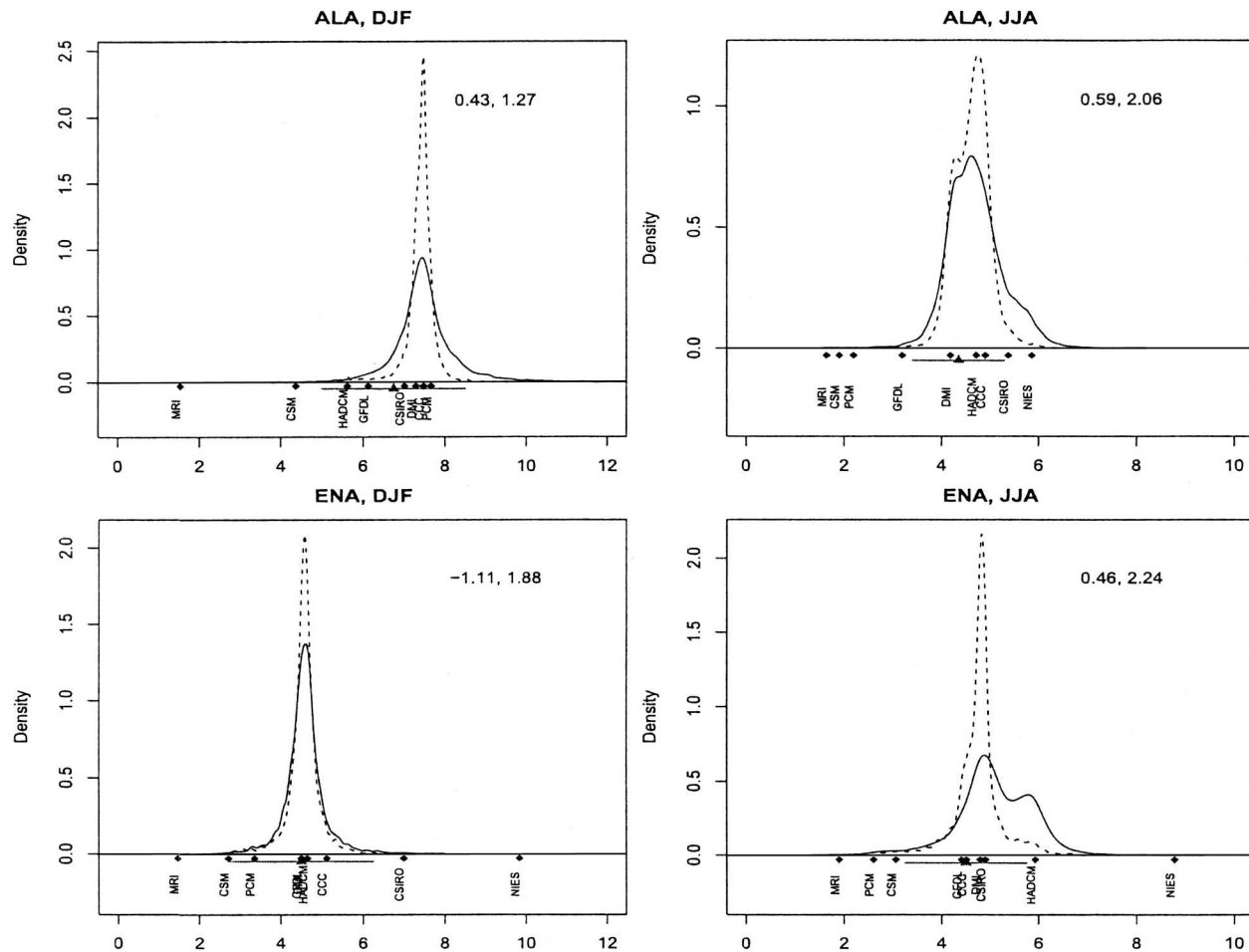


FIG. 2. Posterior distributions of  $\Delta T \equiv \nu - \mu$  for (top) ALA and (bottom) ENA regions for the (left) winter (DJF) and (right) summer (JJA) seasons. Densities derived from the basic model (solid lines) and from the model with correlation between  $Y_i$  and  $X_i$  (dashed lines). The points along the base of the densities mark the nine AOGCMs temperature change predictions. The triangle and segment indicate the REA estimate of mean change  $\pm$  a measure of natural variability. The two numbers in the upper-right corner of each panel are the limits of the 95% posterior probability region for  $\beta_x$ .

winter and summer season, for the basic model (solid lines) and for the model with correlation (dashed lines). For reference, the nine models' individual responses  $Y_i - X_i$ ,  $i = 1, \dots, 9$  are plotted along the  $x$  axis (as diamonds), together with Giorgi and Mearns's REA estimates of  $\Delta T$  plus or minus two standard deviations (as a triangle superimposed on a segment). From the relative position of the diamonds one can qualitatively assess a measure of convergence for each model (defined in terms of the individual AOGCM's temperature change response), and discriminate between models that behave as outliers and models that reinforce each other by predicting similar values of temperature change. The bias measure is not immediately recoverable from this representation, so in Table 3 we list values of the bias for each of the nine AOGCMs in the six regions for both summer and winter. A comparison of the densities in Figs. 2–4 with the bias values listed in the table reveals how models having relatively smaller

bias receive relatively larger weight. Models that perform well with respect to both criteria are the ones where the probability density function is concentrated. The REA results are consistent overall with mean and range of the dashed curves. For most of the regions (especially in winter) the solid and dashed lines are in agreement, with the solid lines showing a slightly wider probability distribution. This is expected because of the introduction, in (11) of the covariate  $X_i$ , which has the effect of making the distribution of  $Y_i$  more concentrated about  $\nu$ . As a consequence the distribution of the difference  $\nu - \mu$  is also tighter. For these region–season combinations, and for a majority of those not shown, it is the case that the two models, with or without correlation, result in almost identical posterior distributions of temperature change. They represent cases in which the nine AOGCMs maintain the same behavior, when judged in terms of their absolute projections of future climate, or in terms of their temperature change signals.

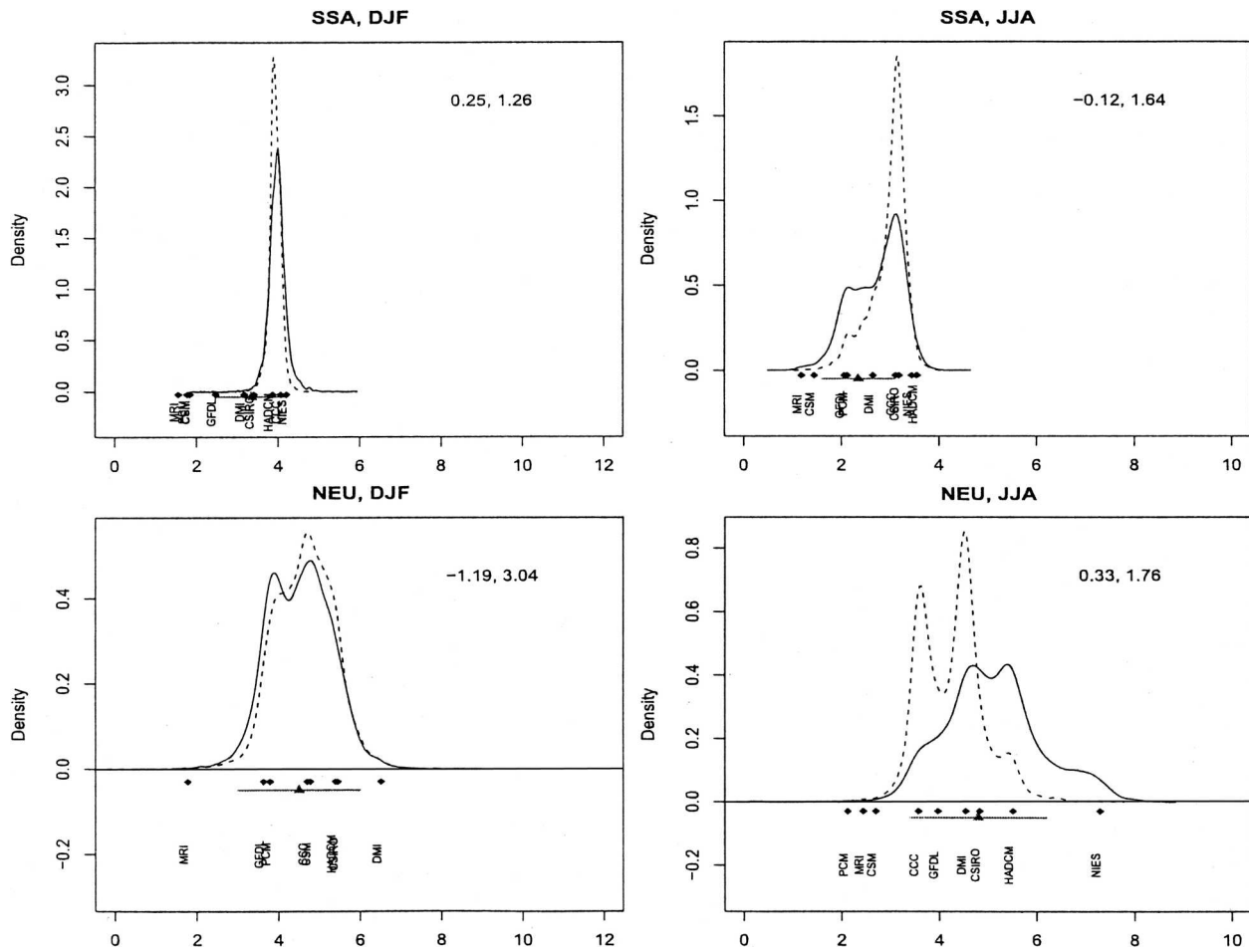


FIG. 3. Same as Fig. 2, but for (top) SSA and (bottom) NEU regions.

Those that can be defined as extreme are so in both respects, those that agree with each other, and form the consensus, do so in both respects.

For a few region–season combinations (we show three representatives: EAF, NEU, and NAUs in summer) the posterior densities are dramatically different for the two statistical models. This is the consequence of applying two different measures of convergence, and AOGCM results that are “outliers” with respect to one but are not so with respect to the other. In these instances, it matters which criteria of convergence we apply, and so the REA method is in agreement only with the dashed curves.

Since the introduction of the parameter  $\beta_x$  heavily influences the outcome of the analysis for some region–season combinations, it deserves a careful consideration. For most of the regions its posterior concentrates to the right of zero, indicating a significant positive correlation between present and future temperature responses of all AOGCMs. Figure 5 shows the range of distributions of  $\beta_x$  for the six regions and two seasons. It is also interesting to notice that most of the mass of these posterior

densities is concentrated around 1, supporting (as explained in section 2b) the REA finding of weak correlation between the temperature *change* signal and bias.

The shapes of the densities in Figs. 2–4 vary significantly. Regular, unimodal curves are estimated in regions where there is obvious agreement among models, or where outlying models are downweighted due to large biases. Multimodal curves characterize regions where AOGCMs give disparate predictions, none of which can be discounted on the basis of model bias. These are also most often the cases when the results are sensitive to whether the basic model or the one with correlation is adopted. The value of presenting results in terms of a posterior distribution is obvious in the case of multimodal curves. For these densities the mean and median would not be good summaries of the distribution. Although we do not necessarily see these multimodal results as having an obvious physical interpretation, they definitely highlight the problematic nature of predictions in specific areas of the globe. For these regions our analysis gives more insight than traditional measures of central tendency and confidence intervals,



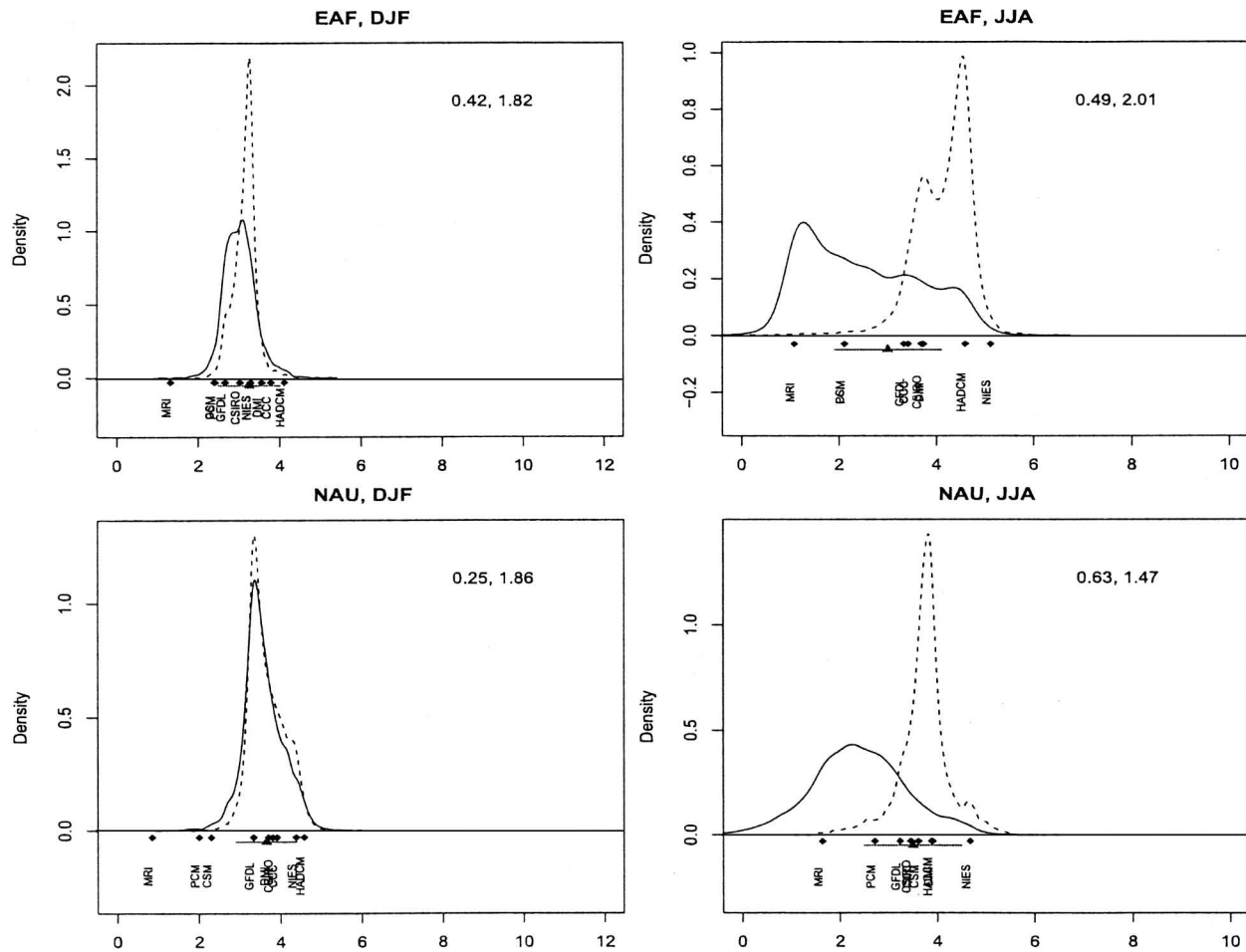


FIG. 4. Same as Fig. 2, but for (top) EAF and (bottom) NAU regions.

and considering the sensitivity of the results to the statistical model assumptions is extremely relevant.

## 2) PRECISION PARAMETERS

Figures 6 and 7 summarize the posterior distributions for the precision parameters  $\lambda_i$ , in the form of box plots

(shown on a logarithmic scale, because of their high degree of skewness). In our analysis  $\lambda_i$ s are the analog to the model weights in the REA's final estimates. However,  $\lambda_i$  is for us a random variable, and the scoring of the nine AOGCMs should be assessed here through the relative position of the nine box plots, rather than by comparing point estimates. Comparing these distri-

TABLE 3. Model bias, for the nine AOGCMs temperature response in the six regions in winter (DJF) and summer (JJA). Biases are computed as the deviation of the single AOGCM's response,  $X_i$ , from the mean of the posterior distribution for  $\mu$  derived by our analysis.

	CCC	CSIRO	CSM	DMI	GFDL	MRI	NIES	PCM	HADCM
ALA (DJF)	-0.23	5.83	2.95	0.98	4.29	11.88	2.15	-2.40	-2.62
ALA (JJA)	0.50	1.37	-3.57	-0.32	2.26	1.82	0.57	-4.83	-0.64
ENA (DJF)	1.15	-1.37	0.18	3.03	-0.17	4.44	-1.32	-1.56	0.10
ENA (JJA)	-1.03	1.21	-2.48	1.04	7.26	8.31	1.16	-1.01	-0.62
SSA (DJF)	-0.00	0.69	-1.82	3.16	3.46	4.85	-1.20	-0.97	0.09
SSA (JJA)	-1.03	-0.48	-0.98	0.72	0.74	1.28	-0.39	-1.74	-2.01
NEU (DJF)	0.47	0.34	-1.43	0.30	-1.04	0.57	-6.30	0.03	-3.25
NEU (JJA)	0.10	2.18	-3.65	1.05	3.52	5.96	-0.63	-2.38	-0.83
EAF (DJF)	-4.76	-1.06	-1.39	0.91	0.30	1.70	-0.28	-1.90	0.10
EAF (JJA)	-3.21	-1.58	-2.49	1.12	-1.48	0.57	-1.29	-3.49	-0.05
NAU (DJF)	-1.31	-1.58	-1.16	2.17	0.23	3.32	0.60	-2.05	-0.71
NAU (JJA)	-3.31	-3.04	-5.15	1.62	-1.70	1.57	-0.95	-5.58	-2.15

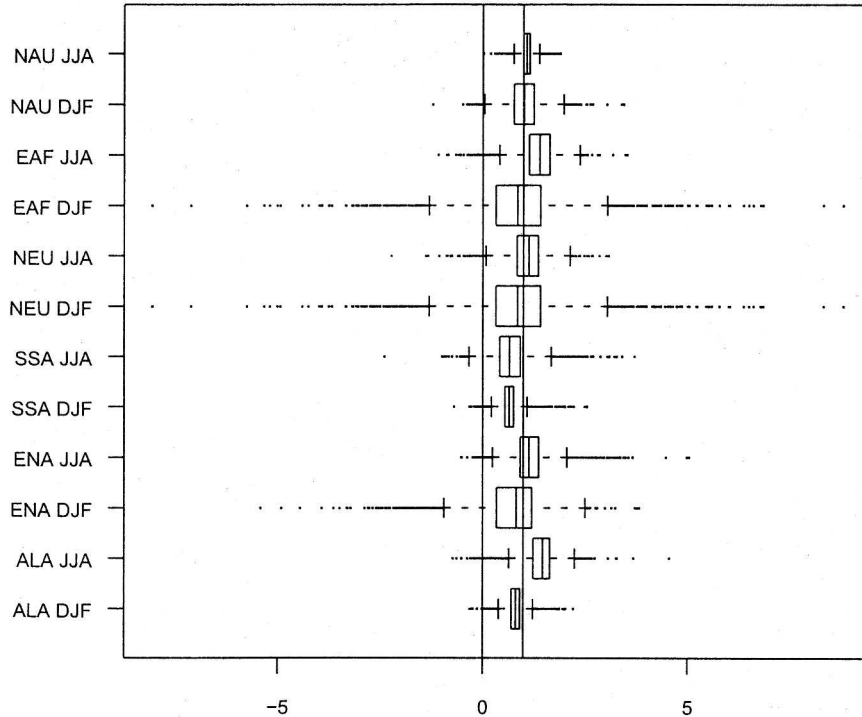


FIG. 5. Posterior distributions of  $\beta_x$ , the regression parameter that introduces correlation between  $X_i$  and  $Y_i$ , common to all nine AOGCMs, for six chosen regions for the winter and summer seasons. For reference, we draw a vertical line at 0, which is useful to assess the significance of the parameter magnitude, and a vertical line at 1 to assess the consistency of our results to the REA analysis' assumption of independence between  $Y_i - X_i$  and  $X_i - \mu$ .

butions across models for a single region and season may suggest an ordering of the nine AOGCMs: large  $\lambda_i$ s indicate that the distribution of the AOGCM response is more tightly concentrated around the true climate response. Thus, posterior distributions that are relatively shifted to the right indicate better performance by the AOGCMs associated to them than distributions shifted to the left. However, the large overlap among these distributions indicates that there is substantial uncertainty in the relative weighing of the models and cautions against placing too much faith in their ranking. Further substantiating this finding, we have calculated the posterior means for all  $\lambda_i$ s and standardized them to percentages. They are listed in Tables 4 and 5. The tables show clearly that the nine AOGCMs are weighted differently in different regions and seasons. This suggests differential skill in reproducing regional present-day climate and a different degree of consensus among models for different regional signals of temperature change, again pointing to the need of evaluating an AOGCM over a collection of regions and types of climate, before placing too much or too little confidence in it. In this perspective, then, Table 6 presents our summary measure of the nine models' relative weighting, by listing the median rank of each model over the 22 regions, separately for the winter

and summer seasons. This ranking is also in the same spirit as the overall reliability index produced by REA. (All the results presented here are from the model with correlation, where the posterior means of the  $\lambda_i$  parameters have closer resemblance to the REA weights.)

### 3) INFLATION–DEFLATION PARAMETER

The parameter  $\theta$  represents the inflation–deflation factor in the AOGCMs' precision when comparing simulations of present-day to future climate. When adopting the model with correlation the posterior for  $\theta$  is always concentrated over a range of values greater than 1, as a consequence of the tightening effect of the form (11) described earlier. As for the posterior distributions derived from the basic model, we present them for each of the six regions in Fig. 8, in the form of box plots. For some of the regions the range is concentrated over values less than 1, suggesting a deterioration in the precision of the nine AOGCMs. Figure 8 indicates that each region tells a different story, and so does each season. For ENA, the simulations (i.e., their agreement) deteriorate in the future, and this holds true for both seasons. For NAU, especially in the summer, and for SSA in both seasons the contrary appears to be true.

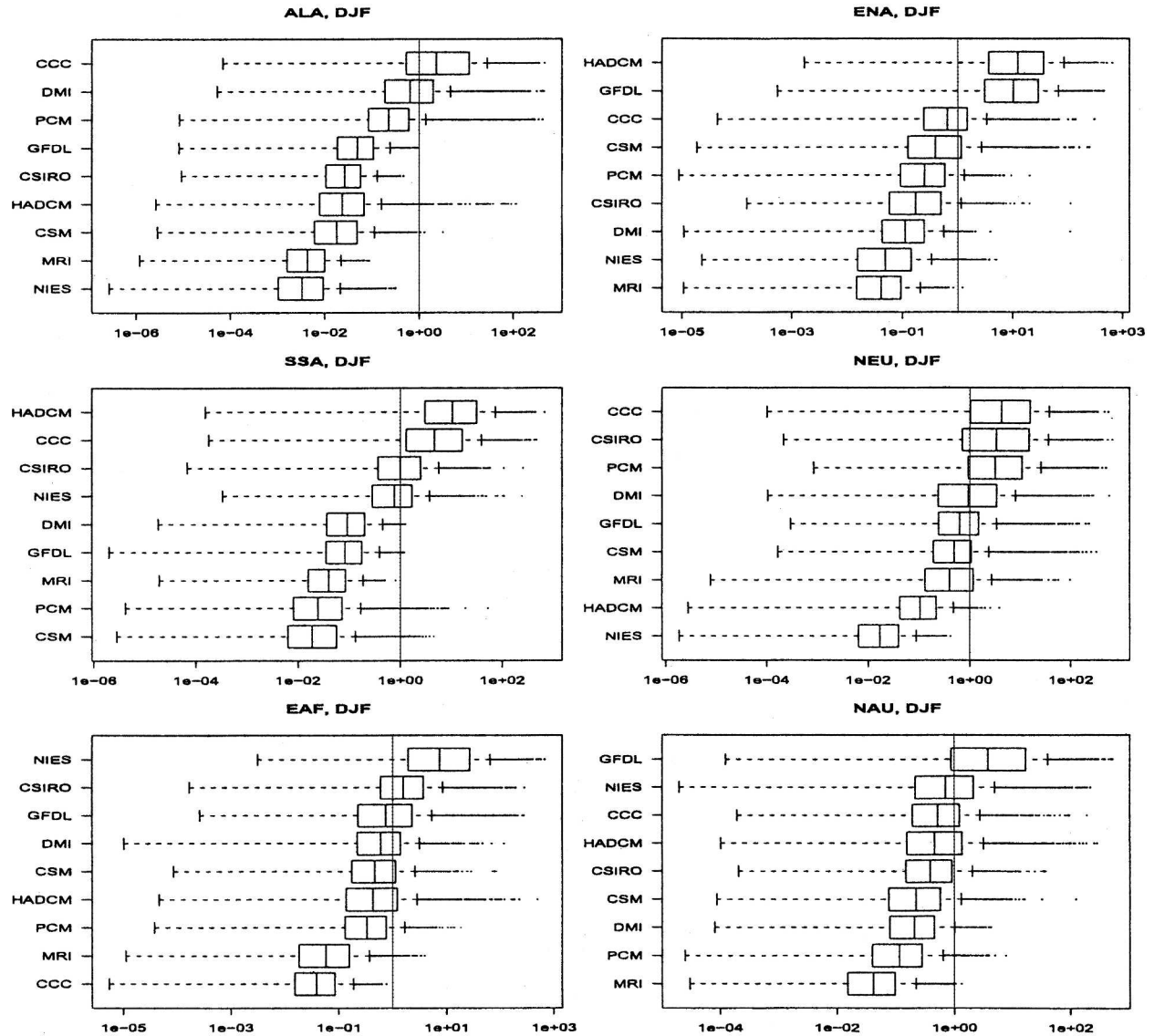


FIG. 6. Posterior distributions of  $\lambda_i$ , the model-specific precision parameter, for six chosen regions for the winter season.

For ALA the behavior is different in winter (where a loss of precision is indicated) than in summer. For the two other regions, EAF and NEU, the inference is not as clear-cut. We cannot offer any insight into the reason why this is so, our expectation being that the projections in the future, in the absence of the regression structure (11), would be at least as variable around the central mean as the current temperatures simulations are, likely more so. Considering in detail the properties of the AOGCMs for these regions may help to interpret these statistical results. However, we point out that the assumption of a common parameter for all AOGCMs may be too strict, and these results may change if a richer dataset, with single-model ensembles, allowed us to estimate model-specific  $\theta$  factors.

#### 4) INTRODUCING HEAVY-TAIL DISTRIBUTIONS

We estimated alternative statistical models, with heavier-tail characteristics, by adopting Student's  $t$  distributions for  $X_i$  and  $Y_i$ , as explained in section 2b.

Different values of degrees of freedom ( $\phi$  in the notation of section 2b) were tested, from 1 (corresponding to heavy-tail distributions, more accommodating of outliers) to 64 (corresponding to a distribution that is nearly indistinguishable from a Gaussian). By comparing the posterior densities so derived, we can assess how robust the final results are to varying the assumptions on the tails of the distributions of  $X_i$ s and  $Y_i$ s. It is the case for all regions that the overall range and shape of the temperature change distributions are insensitive

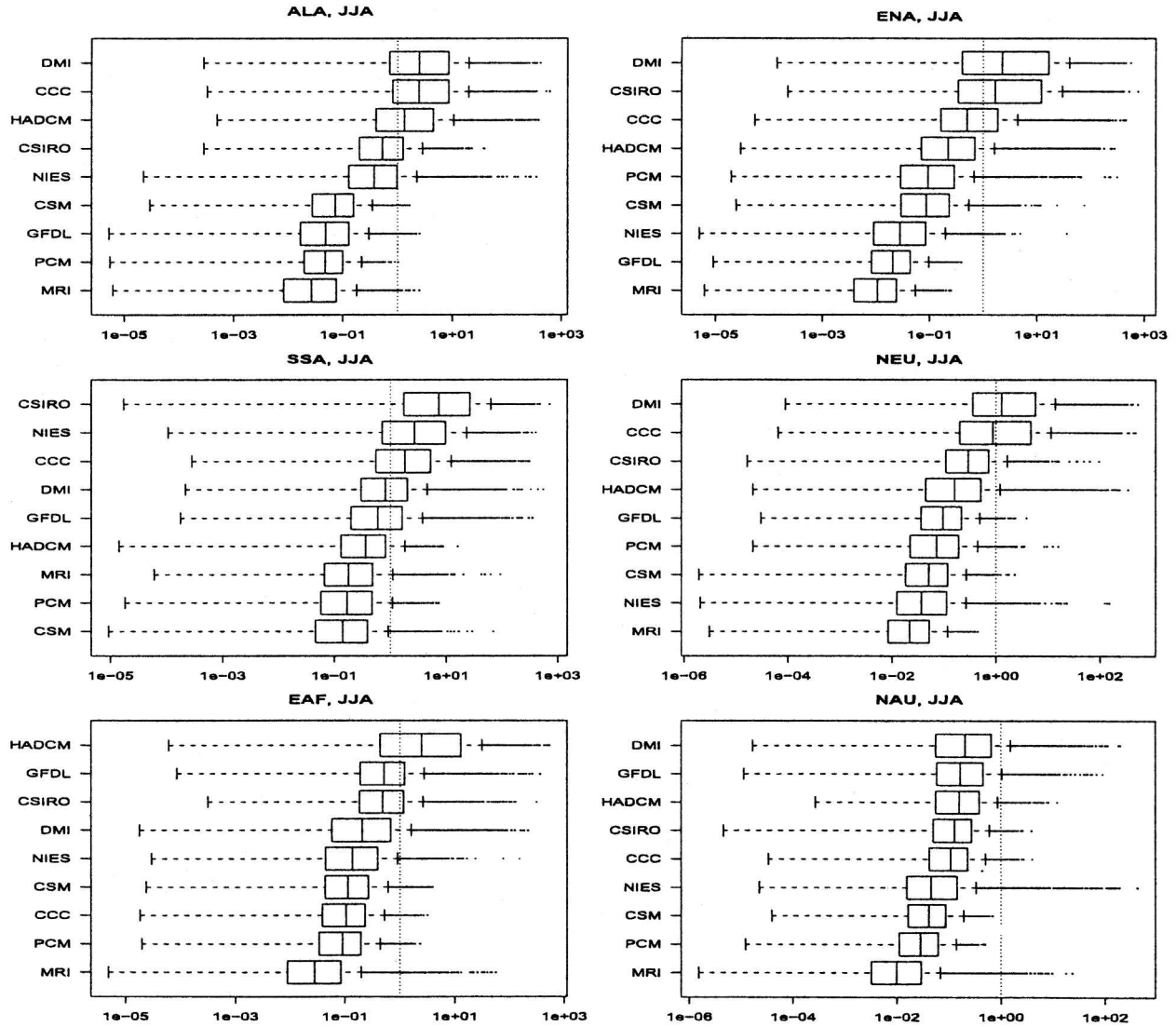


FIG. 7. Same as Fig. 6, but for the summer season.

to the degrees of freedom, and this holds true when testing either the basic model, or the model with correlation. Only for a small number of regions using one or two degrees of freedom produce posterior distribu-

tions of climate change that spread their mass toward the more isolated values among the nine AOGCM responses. The difference from the basic Gaussian model is hardly detectable in a graph comparing the posterior

TABLE 4. Relative weighting of the nine AOGCMs across six regions chosen as examples. The values are computed as  $100 \times (\lambda_i^* / \sum_{i=1}^9 \lambda_i^*)$ , where the 9  $\lambda_i^*$ s are the means of the posterior distributions derived by MCMC simulation. Winter temperature change analysis.

	CCC	CSIRO	CSM	DMI	GFDL	MRI	NIES	PCM	HADCM
ALA	63.80	0.19	0.19	21.71	0.35	0.03	0.04	12.28	1.40
ENA	2.84	0.74	2.99	0.34	42.26	0.11	0.24	0.75	49.73
SSA	36.08	4.40	0.13	0.30	0.26	0.13	3.20	0.22	55.28
NEU	28.21	30.93	3.08	8.71	3.46	2.15	0.05	23.12	0.29
EAF	0.17	10.50	2.60	3.17	11.29	0.39	65.49	1.58	4.82
NAU	4.81	2.96	2.09	1.29	66.55	0.29	13.18	0.86	7.97

TABLE 5. Same as in Table 4, but for summer temperature change analysis.

	CCC	CSIRO	CSM	DMI	GFDL	MRI	NIES	PCM	HADCM
ALA	34.88	3.12	0.34	34.41	0.33	0.21	4.59	0.22	21.91
ENA	14.07	34.38	0.58	43.57	0.08	0.04	0.22	1.78	5.27
SSA	12.35	47.90	0.85	5.87	6.69	1.03	23.21	0.80	1.29
NEU	40.50	2.92	0.39	43.88	0.69	0.16	1.09	0.70	9.67
EAF	0.84	7.93	1.03	7.98	9.52	0.94	1.98	0.70	69.08
NAU	4.59	5.67	1.72	31.09	15.68	2.04	29.30	1.23	8.67

densities from different model formulations. We conclude that the posterior estimates are not substantially affected by the distributional assumptions on the tails for  $X_i$  and  $Y_i$ .

#### 4. Discussion, extensions, and conclusions

##### a. Overview of temperature change distributions

Figures 9 and 10 look at differences in the estimates of temperature change and uncertainty among regions, in the form of a series of box plots, sorted by their median values. This representation is useful for assessing different magnitudes of warming, and different degrees of uncertainty (variability) across regions and between seasons. Notice that all distributions are limited to the positive range, making the case for global warming. Warming in winter is on average higher than in summer, for all regions, and the high latitudes of the Northern Hemisphere are the regions with a more pronounced winter climate change. These all are by now undisputed results from many different studies of climate change (Cubasch et al. 2001). The variability of the distributions is widely different among regions, supporting the notion that for some regions the signal of climate change is stronger and less uncertain than for others. Regions such as ENA, SSA, eastern Asia (EAS), the Mediterranean (MED) in winter and West Africa (WAF) and southern Australia (SAU) in summer show extremely tight distributions predicting the number of degrees of warming with relative certainty. On the other hand, regions like northern Asia (NAS) and central Asia (CAS) in winter, EAF and west North America (WNA) in summer, and the Amazons (AMZ), central North America (CNA), and NEU in both seasons show a wide range of uncertainty in the degrees of warming. As we already mentioned, the width and shape of these distributions may be taken as a signal of the degree of agreement among AOGCMs over the temperature projection in the regions. Problematic re-

gions, characterized by multimodal or simply diffuse distributions may suggest areas that merit special attention by the climate modeling community.

##### b. Advantages of a Bayesian approach

Recent work (Raisanen and Palmer 2001; Giorgi and Mearns 2003) has addressed the need for probability forecasts, but offered only an assessment of the probability of exceeding thresholds, identified as the (weighted) fraction of the ensemble members by which the thresholds in question were exceeded. Moreover, in the case where each ensemble member is equally weighted, issues of model validation are ignored. A differential weighting of the members on the basis of the REA method is more appropriate but depends on a subjective formulation that is difficult to evaluate. In contrast, we think that the Bayesian approach is not only flexible but facilitates an open debate on the assumptions that generate probabilistic forecasts.

The results in this paper demonstrate how the quantitative information from a multimodel experiment can be organized in a coherent statistical framework based on a limited number of explicit assumptions. Our Bayesian analysis yields posterior probability distributions of all the uncertain quantities of interest. These posterior distributions for regional temperature change and for a suite of other parameters provide a wealth of information about AOGCM reliability and temporal (present to future) correlations. We deem this distributional representation more useful than a point estimate with error bars, because important features such as multimodality or long tails become evident. Also, the uncertainty quantified by the posterior distributions is an important product of this analysis and it is not easily recovered using non-Bayesian methods. We note here that the width of the posterior distributions is also a reflection of the limited number of data points that we used in estimating the parameters of the statistical

TABLE 6. Median rank for each model over the 22 regions for winter and summer temperature change predictions. For each region and season the models were ranked from first [1] to last [9] on the basis of the sorted values (from largest to smallest) of the posterior means of  $\lambda_i$ ,  $i = 1, \dots, 9$ . Then, the median value of the ranking for each model over the 22 regions was computed. If the median rank for model  $i$  is  $k$ , model  $i$  was ranked  $k$ th or better in at least half the regions.

Season	CCC	CSIRO	CSM	DMI	GFDL	MRI	NIES	PCM	HADCM
DJF	6.5	5.0	5.0	5.0	3.0	9.0	6.0	5.0	3.0
JJA	3.5	3.0	6.0	3.0	6.0	9.0	5.0	6.5	3.0

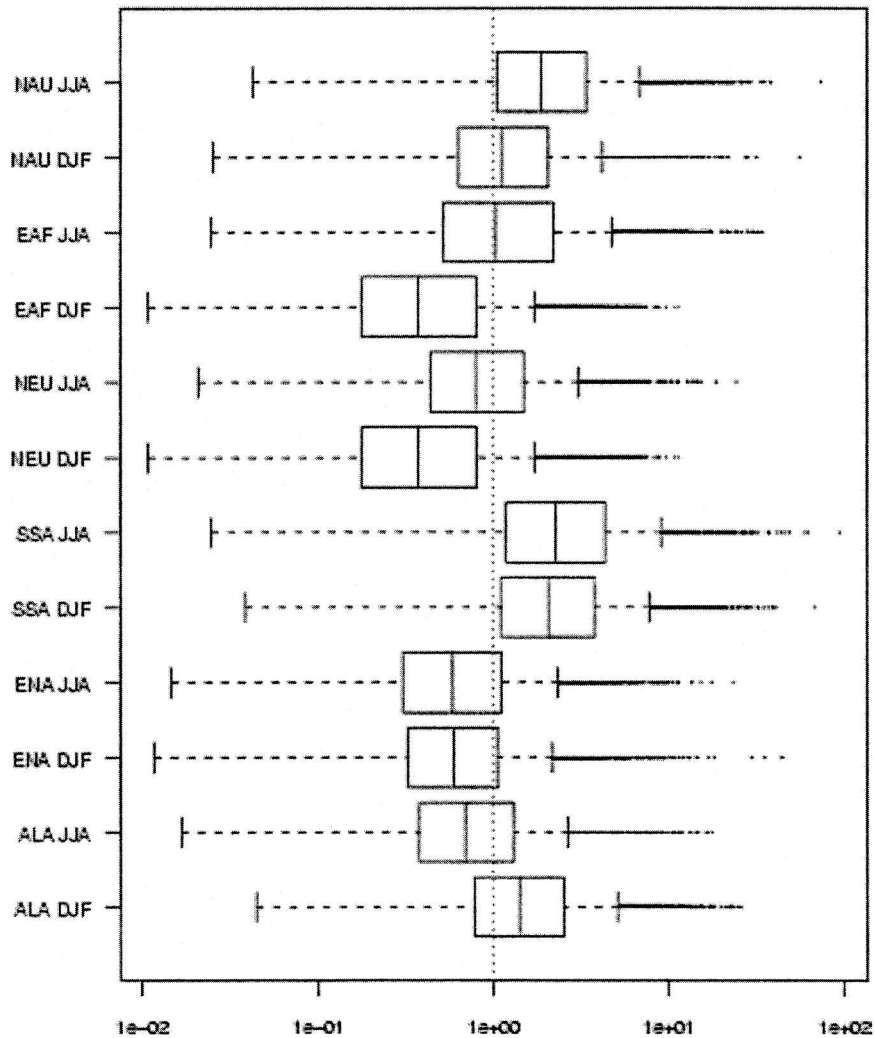


FIG. 8. Posterior distributions of  $\theta$ , the inflation/deflation factor for the precision parameters when simulating future climate, common to all nine AOGCMs for six chosen regions for the winter and summer seasons.

model, particularly for the AOGCM-specific parameters  $\lambda_1, \dots, \lambda_9$ . Conversely, it is worth underlining that even if we start from extremely diffuse priors, we are able to estimate informative posterior distributions for all parameters.

### c. Sensitivity analysis of the statistical assumptions

The posterior distributions of regional temperature change in many region–season combinations may differ both in variance and shape, depending on the statistical model adopted, that is, when introducing the correlation structure between present and future model projections. These sensitivities suggest the need to enrich the experimental setting, thus formulating the statistical model as free as possible from constraints. For example, imposing a common correlation parameter  $\beta_x$

among models is a strong assumption that could be relaxed if more information were available. Similarly, single-model ensembles would make it possible to estimate the internal variability of the AOGCM and allows for the separation of intramodel versus intermodel variability. In addition, climatological information could be incorporated into the prior. The range of our posteriors for the present and future temperature means would remain the same in the presence of proper, but still uninformative priors (i.e., either uniform distributions limited to a physically reasonable range, or Gaussian distributions concentrating most of their mass over such ranges). However, a modeling group may have confidence that results in a particular region are positively biased, or relatively more accurate than in other regions, on the basis of separate experiments. This kind of information could be applied to the formulation of a

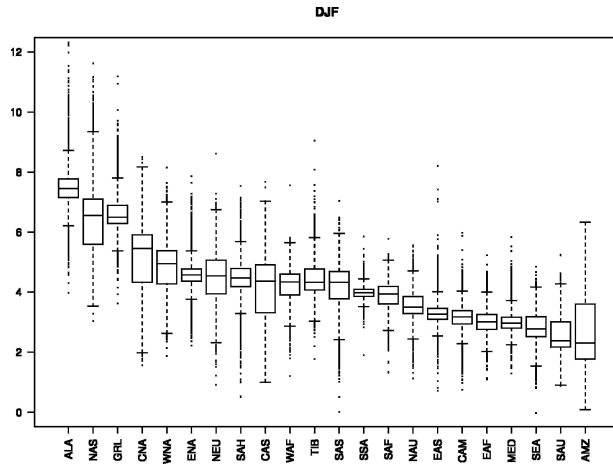


FIG. 9. Posterior distributions of  $\Delta T \equiv \nu - \mu$  for all 22 regions for the winter season.

different prior for that AOGCM's precision parameter. In our opinion, analyzing the robustness of results to model assumptions is as valuable as analysis of the results themselves, highlighting the weak parts of the statistical formulation, and pointing to the need of closer communication between modelers, climatologists, and statisticians in order to circumscribe the range of sensible assumptions. In this regard the Bayesian model presented here may be viewed as a device to foster more effective collaborations.

A natural extension to the current region-specific model is a model that introduces AOGCM-specific correlation between regions and thus borrows strength from all the regional temperature signals in order to estimate model variability and biases. As we discussed in section 1, working with higher-resolution AOGCM output, both spatially and temporally, will require a much more complex effort. Richer information on AOGCM performance as linked to specific regional/

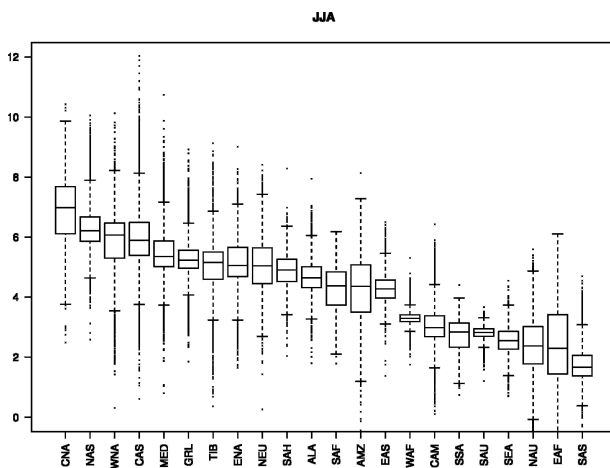


FIG. 10. Same as Fig. 9, but for the summer season.

seasonal climate reproductions will have to be incorporated in the likelihood of the AOGCM's responses and/or the priors on the precision parameters. In addition, other climate responses may be considered, jointly with temperature. Precipitation is of course the obvious candidate, and a bivariate distribution of temperature and precipitation is a natural extension.

#### d. About reliability criteria and statistical modeling

The criteria of bias and convergence established by Giorgi and Mearns (2002, 2003) and adopted in this work are commonly discussed in the climate change literature as relevant to evaluating climate change projections. However, one can question whether these two criteria are the most relevant for evaluating the reliability of projections of climate change. The ability of a model to reproduce the current climate is usually regarded as a necessary, but not sufficient condition for considering the model's response to future forcings as reliable (McAvaney et al. 2001). The convergence criterion has been advocated by some authors (Raisanen 1997; Giorgi and Francisco 2000) and can be theoretically derived by assuming that the observed AOGCMs represent a random sample from a superpopulation of AOGCMs. A pitfall in applying model convergence is that models may produce similar responses due to similarities in model structures, not because the models are converging on the "true" response. Furthermore, one may argue that extreme projections could be the result of a model incorporating essential feedback mechanisms that the majority of other models ignore. Yet, agreement among many models has been viewed as strengthening the likelihood of climate change (Cubasch et al. 2001; Giorgi et al. 2001a,b). As we noted in section 2b both the form of the REA weights (with the use of two different exponents for the bias and convergence terms) and the specification of the likelihood in our statistical model allow the final result to depend to a lesser degree on the convergence criterion. Besides changing the likelihood assumptions one could impose a prior distribution for the  $\theta$  parameter that concentrates most of its mass on values less than 1. In this way, one posits a lower level of confidence in the precision of the future simulation than of the present, implicitly accepting a less stringent criterion of convergence for the future trajectories. The form of (9) shows that a value of  $\theta$  less than 1 would down weight large deviations in the convergence term  $(Y_i - \nu)^2$ , thus achieving a similar effect to the exponent in the REA weight.

From a more general perspective, however, our goal was to extend the REA method, not to reevaluate its criteria. It is expected that other criteria, and more complex combinations of criteria will evolve over time as this statistical work is refined. Much work is being devoted to more sophisticated and extensive methods of model performance evaluation and intermodel comparison (Hegerl et al. 2000; Meehl et al. 2000) and fu-

ture analyses may modify or extend the two criteria of bias and convergence to account for them.

In conclusion, we view our results as an illustration of the power of bringing statistical modeling to experiments where quantifying uncertainty is an intrinsic concern. Moreover, we hope this work may foster more deliberate analysis that incorporates scientific knowledge of climate modeling at global and regional scales.

*Acknowledgments.* This research was supported through the National Center for Atmospheric Research Initiative on Weather and Climate Impact Assessment Science, which is funded by the National Science Foundation. Additional support was provided through NSF Grants DMS-0084375 and DMS-9815344. We are indebted to two anonymous reviewers and the editor for thoughtful and constructive criticisms that have greatly informed and added value to the paper in its final form. Also, we wish to thank Filippo Giorgi, Kevin Trenberth, Tom Wigley, and Art Dempster for helpful and cogent remarks on intermediate drafts.

## APPENDIX

### Prior-Posterior Update through MCMC Simulation

The joint posterior distributions derived from the models in section 2 are not members of any known parametric family. However, the distributional forms (Gaussian, uniform, and gamma) chosen for the likelihoods and priors are conjugate, thus allowing for closed-form derivation of all full conditional distributions (the distributions of each parameter, as a function of the remaining parameters assuming fixed deterministic values). We list here such distributions, for the robust model that includes a correlation between  $X_i$  and  $Y_i$  in the form of regression equation (11). The variables  $s_i$ ,  $t_i$   $i = 1, \dots, 9$  are introduced here as an auxiliary randomization device, in order to efficiently simulate from the Student's  $t$  distributions within the Gibbs sampler. They are not essential parts of the statistical model. Fixing  $s_i = t_i = 1$ ,  $\beta_x = 0$  allows the recovery of the full conditionals for  $\lambda_1, \dots, \lambda_9$ ,  $\mu$ ,  $\nu$ , and  $\theta$  of the basic univariate model as a special case:

$$\lambda_i | \dots \sim \text{Ga} \left( a + 1, b + \frac{s_i}{2} (X_i - \mu)^2 + \frac{\theta t_i}{2} [Y_i - \nu - \beta_x (X_i - \mu)]^2 \right), \quad (\text{A1})$$

$$s_i | \dots \sim \text{Ga} \left( \frac{\phi + 1}{2}, \frac{\phi + \lambda_i (X_i - \mu)^2}{2} \right), \quad (\text{A2})$$

$$t_i | \dots \sim \text{Ga} \left( \frac{\phi + 1}{2}, \frac{\phi + \theta \lambda_i [Y_i - \nu - \beta_x (X_i - \mu)]^2}{2} \right), \quad (\text{A3})$$

$$\mu | \dots \sim N \left( \tilde{\mu}, \left( \sum s_i \lambda_i + \theta \beta_x^2 \sum t_i \lambda_i + \lambda_0 \right)^{-1} \right), \quad (\text{A4})$$

$$\nu | \dots \sim N \left( \tilde{\nu}, \left( \theta \sum t_i \lambda_i \right)^{-1} \right), \quad (\text{A5})$$

$$\beta_x | \dots \sim N \left( \tilde{\beta}_x, \left[ \theta \sum t_i \lambda_i (X_i - \mu)^2 \right]^{-1} \right) \quad (\text{A6})$$

$$\theta | \dots \sim \text{Ga} \left( c + \frac{N}{2}, d + \frac{1}{2} \sum t_i \lambda_i [Y_i - \nu - \beta_x (X_i - \mu)]^2 \right). \quad (\text{A7})$$

Above, we have used the following shorthand notation:

$$\tilde{\mu} = \frac{\sum s_i \lambda_i X_i - \theta \beta_x \sum \lambda_i t_i (Y_i - \nu - \beta_x X_i) + \lambda_0 X_0}{\sum s_i \lambda_i + \theta \beta_x^2 \sum \lambda_i t_i + \lambda_0}, \quad (\text{A8})$$

$$\tilde{\nu} = \frac{\sum t_i \lambda_i \{Y_i - \beta_x (X_i - \mu)\}}{\sum t_i \lambda_i}, \quad (\text{A9})$$

$$\tilde{\beta}_x = \frac{\sum t_i \lambda_i (Y_i - \nu)(X_i - \mu)}{\sum t_i \lambda_i (X_i - \mu)^2}. \quad (\text{A10})$$

The Gibbs sampler can be easily coded so as to simulate iteratively from this sequence of full conditional distributions.

After a series of random drawings during which the MCMC process forgets about the arbitrary set of initial values for the parameters (the burn-in period), the values sampled at each iteration represent a draw from the joint posterior distribution of interest, and any summary statistic can be computed to a degree of approximation, a direct function of the number of sampled values available and the inverse function of the correlation between successive samples. To minimize the latter, we save only one iteration result every 50, after running the sampler for a total of 500 000 iterations, and discarding the first half as a burn-in period. These many iterations are probably not needed for this particular application but by performing them we are eliminating any possibility of bias resulting from too few MCMC iterations. The convergence of the Markov chain to its stationary distribution (the joint posterior of interest) is verified by standard diagnostic tools (Best et al. 1995). [A self-contained version of the MCMC algorithm, implemented in the free software package R (R Development Core Team 2004), is available at [www.cgd.ucar.edu/~nychka/REA](http://www.cgd.ucar.edu/~nychka/REA).]



## REFERENCES

- Allen, M. R., P. A. Stott, J. F. B. Mitchell, R. Schnur, and T. L. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617–620.
- Allen, M., S. Raper, and J. Mitchell, 2001: Uncertainty in the IPCC's Third Assessment Report. *Science*, **293**, 430–433.
- Best, N. G., M. K. Cowles, and S. K. Vines, cited 1995: CODA Convergence Diagnosis and Output Analysis software for Gibbs Sampler output: Version 0.3. [Available online at <http://www.mrc-bsu.cam.ac.uk/bugs/classic/coda04/readme.shtml>.]
- Cubasch, U., and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 525–582.
- Dai, A., T. M. L. Wigley, B. A. Boville, J. T. Kiehl, and L. E. Buja, 2001: Climates of the twentieth and twenty-first centuries simulated by the NCAR Climate System Model. *J. Climate*, **14**, 485–519.
- Dessai, S., and M. Hulme, 2003: Does climate policy need probabilities? Tyndall Centre Working Paper 34, 43 pp. [Available online at <http://www.tyndall.ac.uk/publications/publications.shtml>.]
- Emori, S., T. Nozawa, A. Abe-Ouchi, A. Numaguti, M. Kimoto, and T. Nakajima, 1999: Coupled ocean–atmosphere model experiments of future climate change with an explicit representation of sulfate aerosol scattering. *J. Meteor. Soc. Japan*, **77**, 1299–1307.
- Flato, G. M., and G. J. Boer, 2001: Warming asymmetry in climate change simulations. *Geophys. Res. Lett.*, **28**, 195–198.
- Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, **295**, 113–117.
- Giorgi, F., and R. Francisco, 2000: Evaluating uncertainties in the prediction of regional climate change. *Geophys. Res. Lett.*, **27**, 1295–1298.
- , 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**, 1141–1158.
- , and —, 2003: Probability of regional climate change based on the Reliability Ensemble Averaging (REA) method. *Geophys. Res. Lett.*, **30**, 1629, doi:10.1029/2003GL017130.
- , and Coauthors, 2001a: Regional climate information: Evaluation and projections. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 583–638.
- , and Coauthors, 2001b: Emerging patterns of simulated regional climatic changes for the 21st century due to anthropogenic forcings. *Geophys. Res. Lett.*, **28**, 3317–3321.
- Gordon, C., C. Cooper, C. A. Senior, H. T. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transport in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.*, **16**, 147–168.
- Gordon, H. B., and S. P. O'Farrell, 1997: Transient climate change in the CSIRO coupled model with dynamic sea ice. *Mon. Wea. Rev.*, **125**, 875–907.
- 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 Dyn.*, **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**, 30 981–30 996.
- McAvaney, B. J., and Coauthors, 2001: Model evaluation. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 471–524.
- Meehl, G. A., G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer, 2000: The Coupled Model Intercomparison Project (CMIP). *Bull. Amer. Meteor. Soc.*, **81**, 313–318.
- Nakicenovic, N., and Coauthors, 2000: *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 599 pp.
- Noda, A., K. Yamaguchi, S. Yamaki, and S. Yukimoto, 1999: Relationship between natural variability and CO<sub>2</sub>-induced warming pattern: MRI AOGCM experiment. Preprints, *10th Symp. on Global Change Studies*, Dallas, TX, Amer. Meteor. Soc., 359–362.
- Nychka, D., and C. Tebaldi, 2003: Comments 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.
- R Development Core Team, cited 2004: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Available online at <http://www.R-project.org/>.]
- Raisanen, J., 1997: Objective comparison of patterns of CO<sub>2</sub> induced climate change in coupled GCM experiments. *Climate Dyn.*, **13**, 197–211.
- , and T. N. Palmer, 2001: A probability and decision model analysis of a multimodel ensemble of climate change simulations. *J. Climate*, **14**, 3212–3226.
- Reilly, J., P. H. Stone, C. E. Forest, M. D. Webster, H. D. Jacoby, and R. G. Prinn, 2001: Uncertainty in climate change assessments. *Science*, **293**, 430–433.
- Schneider, S. H., 2001: What is dangerous climate change? *Nature*, **411**, 17–19.
- Stendel, M., T. Schmidt, E. Roeckner, and U. Cubasch, 2002: The climate of the 21st century: Transient simulations with a coupled atmosphere–ocean general circulation model. Danmarks Klimacenter Rep. 02-1, 52 pp.
- Washington, W. M., and Coauthors, 2000: Parallel Climate Model (PCM) control and transient simulations. *Climate Dyn.*, **16**, 755–774.
- 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.