5.17

BAYESIAN CLIMATE CHANGE ASSESSMENT USING MULTI-AOGCM ENSEMBLES: GLOBAL AND REGIONAL SURFACE TEMPERATURES

Seung-Ki Min* and Andreas Hense

Meteorological Institute, University of Bonn, Germany

1. INTRODUCTION

There have been increasing efforts to apply Bayesian statistics to climate change detection and attribution in order to consider errors in the signal (or model response) as well as natural climate variability (Min et al., 2004, hereinafter referred to as M04, 2005a; Lee et al., 2005; Schnur and Hasselmann, 2005; IDAG, 2005 and references therein). However, uncertainties from inter-model differences could not be assessed reasonably due to lack of enough samples of model simulations.

In this study, we use multi-model ensembles archived for the 4th Assessment Report (AR4) of Intergovernmental Panel on Climate Change (IPCC) in order to test the sensitivity of Bayesian detection and attribution of climate change to inter-model uncertainties. The Bayesian decision method by M04 is applied to global and regional mean surface air temperatures (SATs) from single-model ensembles (SMEs) with the ECHO-G model (Legutke and Voss, 1999; Min et al., 2005b,c,d) and multi-model ensembles (MMEs) with IPCC AR4 models (Min and Hense, 2005a, hereinafter referred to as MH05a) and their results are compared.

In total, six scenarios are considered as a possible explanation of observed SAT changes over the 20th century: CTL (control), N (natural forcing; solar and volcanic), G (greenhouse-gas), S (sulfate aerosol), ANTHRO (anthropogenic forcing), and ALL (natural plus anthropogenic forcing). Given the scenarios, the Bayesian assessment is done for observed SATs for the whole 20th century and its first (1900-1949) and second half (1950-1999) separately. Parameters necessary to define the scenarios (means and covariance matrices, see section 2.1) are estimated from SMEs or MMEs and its effect on Bayesian decisions is examined as well as the effect of varying priors.

2. METHODOLOGY

2.1 *Bayesian Decision Method*

Here we provide a brief explanation (see M04 for details). Given a set of N possible scenarios (m_i, i = 1, …, N) and the observational data ($\mathbf{d}$), an appropriate question on climate change detection and attribution will be "How probable is the scenario m_i given the observation $\mathbf{d}$?" This can be expressed as a conditional probability $\mathrm{P}(m_i|\mathbf{d})$ which is the posterior probability of the scenario given the observation. Using Bayes' rule, this can be evaluated from the prior probability $\mathrm{P}(m_i)$ which represents a subjective belief in the scenario, and likelihood function $l(\mathbf{d}|m_i)$ which characterizes the observational probability given the scenario:

$$P(m_i \mid \mathbf{d}) = \frac{l(\mathbf{d} \mid m_i) P(m_i)}{\sum_{j=1}^{N} l(\mathbf{d} \mid m_j) P(m_j)}. \quad (1)$$

Assuming multivariate Gaussian distributions for the detection variables of the scenario m_i and the observations $\mathbf{d}$, the likelihood function can be expressed as

$$l(\mathbf{d} \mid m_i) = \frac{1}{\sqrt{(2\pi)^q}} \sqrt{\frac{\det \mathbf{A}_i^{-1}}{\det \boldsymbol{\Sigma}_i \det \boldsymbol{\Sigma}_0}} \exp\left(-\frac{1}{2}\Lambda_i\right), \quad (2)$$

where q is the dimension of the data vector $\mathbf{d}$, $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_i$ are the covariance matrices of the observation $\mathbf{d}$ and the scenario m_i respectively. $\mathbf{A}_i$ is a linear combination of these covariance matrices, and Λ_i is a generalized distance measure between the observation and scenario (for more details see M04).

According to the Bayesian decision theory, posterior probability in Eq. (1) can be used as

* *Corresponding author address*: Seung-Ki Min, Meteorological Institute, University of Bonn, Auf dem Hügel 20, 53121 Bonn, Germany
E-mail: skmin@uni-bonn.de

a decision function (Duda and Hart, 1973; Berger, 1985), and we select the scenario with maximum posterior so that the theoretical error becomes a minimum (Duda and Hart, 1973; M04). In the special case of identical priors, the likelihood ratio or Bayes factor, which represents observational evidences for the scenario concerned, becomes a decision function. Kass and Raftery (1995) suggested descriptive scales of Bayes factors. If the logarithm of Bayes factor is larger than 1, 2.5, or 5, the observations represent ‘substantial’, ‘strong’, or ‘decisive’ evidences in favor of the scenario concerned against a reference scenario (CTL here). Most of recent Bayesian studies apply these scales to interpret their analysis results (M04; Min et al., 2005a; Lee et al., 2005; Schnur and Hasselmann, 2005).

2.2 *Legendre Series Expansions*

We treat one time series of SATs as the realization of a multivariate random variable. Therefore we need to reduce dimension to avoid singular or near-singular covariance matrices in Eq. (2). We propose the use of Legendre polynomials (LP) by which the time series of 20th century SAT anomalies are expanded. The amplitude of overall warming (*scale*, 0th degree: LP0), linear trend (*trend*, 1st degree: LP1), and decadal time-scale variations can be represented effectively (MH05a). Figure 1 shows an example of Legendre decomposition of global mean SATs over 1900-1999 for Legendre degrees 0-12 from CRU observations and two model simulations with natural and anthropogenic forcing (see section 3).

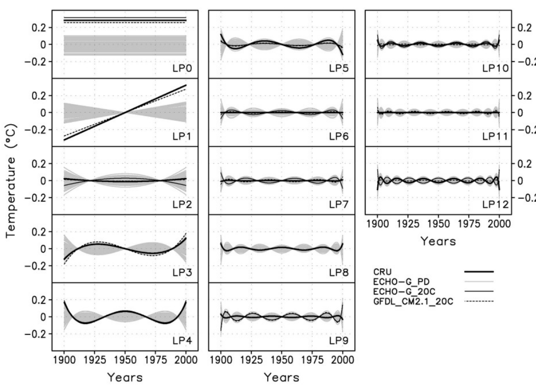

FIG. 1. Reconstructed time series of global mean SATs for 1900-1999 from Legendre series expansions for degrees from 0 to 12: CRU observations (thick solid), selected two 20C3M simulations with different AOGCMs (thin solid and dashed), and 91 samples from ECHO-G_PD. After MH05a.

Additional analysis of power spectra of global and regional mean SATs from MME control runs shows that models do well in simulating the internal variability on decadal time scales. Hence we can apply LP truncation up to the 12th degree which corresponds to about 20 year period for 1900-1999 and 10 years for 1900-1949 and 1950-1999 (MH05a; Min and Hense, 2005b, hereinafter referred to as MH05b).

3. DATA AND MODEL SIMULATIONS

As the observational dataset we use Climate Research Unit (CRU) data (HadCRUT2v for global and CRUTEM2v for regional SATs) for 1900-1999 (Jones and Moberg, 2003). Regional domains are applied following Stott (2003). There are six continental scale regions (see caption of Fig. 3) with each region composed of two or three subregions (Fig. 5). Combing Legendre coefficients of area-averaged SAT for the subregions constructs space-time vector for continental scale regions for model simulations and observations (MH05b).

There are two sources for the model dataset. The first comes from SME simulations with ECHO-G. These consist of an existing 1000-year present-day control run (ECHO-G_PD; Min et al., 2005b,c) and newly performed historical simulations for 1860-2000 under different external forcing factors consistent with the forced scenarios: G, S, N, ANTHRO, and ALL (For details about the simulations and applied forcing, see MH05a and Min et al., 2005d). From ECHO-G ensemble simulations, we obtain at least three non-overlapping samples of 100-year (1900-1999) global/regional mean SATs for five forced scenarios. From the ECHO-G_PD, 91 time series of 100-year SATs are sampled for the CTL scenario using a moving window of 100-year length with a shift of 10 years.

The second source for model data is the IPCC AR4 archive (http://www-pcmdi.llnl.gov /ipcc/about_ipcc.php). We extracted monthly mean SATs from 20C3M simulations and preindustrial control simulations from total 22 models. Detailed model information can be found at the archive. Overall 48, 25, and 80 non-overlapping 100-year global mean SATs could be extracted for MME_ALL (12 models with all forcing), MME_ANTH (12 models with anthropogenic forcing only), and MME_PI (22 models with no forcing) respectively. Another sampling for MME_PI is done with overlapping 100-year moving windows with a 10-year shift where data from individual models are kept separate. This produces 644 samples.

For each 100-year long sample, Legendre coefficients are obtained for the whole period and the first and second 50 years which corresponds to observational periods of 1900-1999, 1900-1949, and 1950-1999. Then the coefficients for three periods are used to estimate means and covariance matrices for likelihood calculation in Eq. (2). For the mean estimation, we use SMEs only for G, S, and N since we don't have available MMEs for the scenarios while for the other scenarios SMEs or MMEs are selectively used. Covariance matrices from forced scenarios are assumed to be identical to that of CTL which is estimated from SMEs or MMEs. We refer to these two kinds of settings with SMEs and MMEs as SINGLE and MULTI experiment respectively.

We treat all model simulations as independent samples. For every 100-year SAT sample the anomaly is calculated with respect to the first 20 years to be compared with the 20th century simulations. Taking different reference periods for anomaly calculation does not change main results. All model data are interpolated linearly onto the observational grid of 5° × 5° and masked with observational coverage on a month-by-month basis prior to analysis.

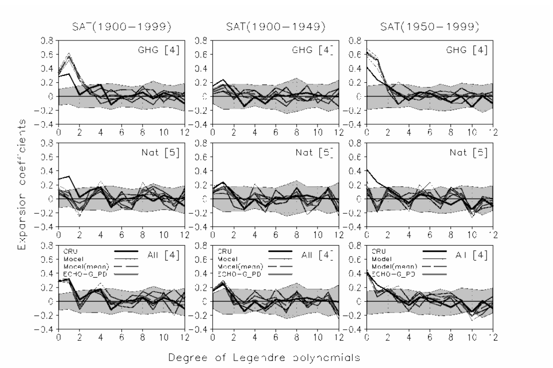

FIG. 2. Legendre expansion coefficients for global mean SATs for the period of 1900-1999, 1900-1949, and 1950-1999 from ECHO-G (thin) and CRU observations (thick). Gray shading represents the range from ECHO-G_PD. After MH05a.

4. DETECTION VARIABLES

Figure 2 shows the Legendre coefficients for global mean SATs from SMEs and observations. For the whole 20th century and its second half, model runs with greenhouse-gas forcing show too large warming (see LP0 and LP1 for *scale* and *trend*), those with natural forcing cannot simulate the warming reasonably. Model with all forcing exhibits a best consistency with observations, suggesting that observed global mean SAT changes are explainable with natural and anthropogenic forcing together. The observed and simulated coefficients in 1900-1949 indicate an important role of natural forcing. Coefficients from MME_ALL and MME_PI show similar patterns (MH05a).

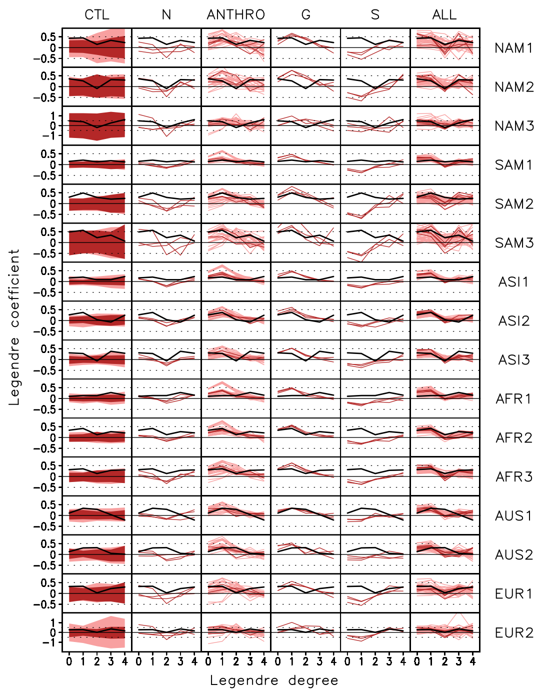

FIG. 3. Legendre expansion coefficients for regional annual mean SATs for 1900-1999 over 16 subregions from six continental scale regions: North America (NAM), Asia (ASI), South America (SAM), Africa (AFR), Australia (AUS), and Europe (EUR). Thick lines are CRU observations and light [dark] grey lines represent results from SMEs [MMEs]. After MH05b.

Figure 3 displays the Legendre coefficients of SATs for 16 subregions during 1900-1999. In general we can find the observed warming in most subregions, although the warming amplitude varies from region to region relative to the internal variability range. Model simulations with different forcing factors show that the coefficients from ALL and ANTHRO runs are closer to observational values than the other runs. In some regions, positive coefficients of LP4 are dominant in N and ALL runs, which are also found in the observations. This represents a possible role of natural forcing since LP4 contains the early warming near 1940s (Fig. 1).

5. BAYESIAN DECISION RESULTS

5.1 *Global Mean Temperatures*

Distribution of posterior probabilities and Bayesian decision for global mean SATs from the MULTI experiment is shown in Fig. 4. ALL scenario is decided for 1900-1999 and 1950-1999 whereas SAT change for 1900-1949 is classified into N or ALL scenarios. This is well consistent with previous results (Mitchell et al., 2001; IDAG, 2005 and references therein). Figure 4 also shows that the Bayesian decisions are mostly insensitive to inter-model variability and prior probability. The results from SINGLE are very similar, meaning that introducing larger inter-model uncertainty does not change the Bayesian assessment for global mean SATs (MH05a).

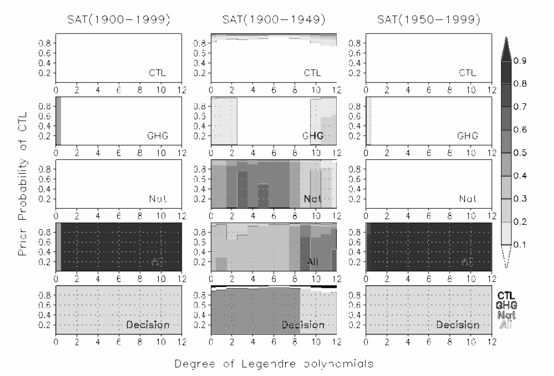

FIG. 4. Distributions of posterior probabilities for the four scenarios of CTL, G, N, and ALL (upper four panels) and corresponding Bayesian decisions (bottom panels) for observed global mean SATs for 1900-1999, 1900-1949, and 1950-1999 in case of varying prior probabilities from the MULTI experiment. Ordinate depicts the prior probability of CTL and abscissa represents Legendre degrees retained at 0 to 12. After MH05a.

5.2 *Regional Mean Temperatures*

Figure 5 represents the decision results for six continental scale regions from MULTI. Comparing with the global mean results, we consider here two additional scenarios of S and ANTHRO, and analyze space-time vectors by combining the Legendre coefficients for two or three subregions. Decision results for 1900-1999 (Fig. 5a) show that NAM, SAM, AFR, and EUR have dominant ALL signals (green) while ASI and AUS prefer ANTHRO scenario (blue). In 1900-1949, the N signal over SAM and ALL signal over AFR are dominant (Fig. 5b) unlike global mean result (Fig. 4). Result for 1950-1999 resembles that for 1900-1999 except that AFR is dominated by G signal, which indicates an underestimated warming trend simulated by MMEs.

Generally regional-scale decisions are insensitive to the priors and temporal truncations. In comparison with the results from SINGLE, some different decisions can be found, e.g., AHTHRO signal is stronger over AFR for all three periods and over SAM in 1950-1999. According to patterns of the Bayes factors (not shown), the strength of ALL and ANTHRO signals over those regions are very similar (MH05b). Hence the effect of inter-model uncertainties appears to be weak even for regional-scale climate change assessment.

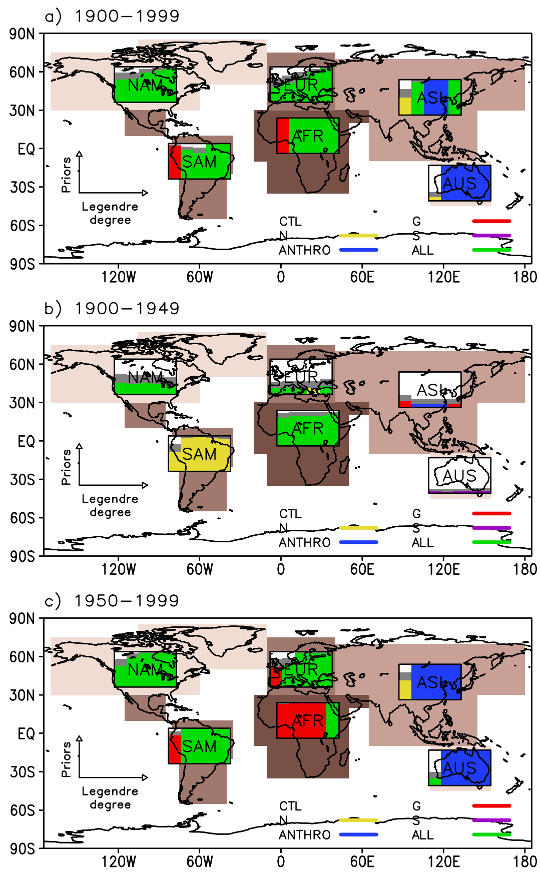

FIG. 5. Bayesian decisions for regional mean SATAs for a) 1900-1999, b) 1900-1949, and c) 1950-1999 in case of varying priors and Legendre degrees retained at 0 to 4 from the MULTI experiment. After MH05b.

6. CONCLUSION

A Bayesian approach is applied to the observed global and regional SAT changes using MMEs of the IPCC AR4 simulations and SMEs with the ECHO-G model. A Bayesian decision method is used as a tool for classifying observations into six scenarios (CTL, N, G, S, ANTHRO, and ALL) which are used to explain observed SAT changes. Observed and simulated area mean SATs are decomposed into temporal components of overall mean, linear trend, and decadal variabilities through Legendre series expansions. The coefficients are used as detection variables. Parameters (means and covariance matrices for likelihood calculation) for defining each scenario are estimated from SMEs or MMEs, by which sensitivity of Bayesian decision results to inter-model uncertainties is examined.

Main finding is that Bayesian assessment of climate change for global and regional SAT changes provides observational evidences for natural plus anthropogenic signals (i.e. ALL scenario), corroborating previous results (Mitchell et al., 2001; IDAG, 2005). The results are largely insensitive to inter-model uncertainties and prior probability in the global and regional mean SATs.

Acknowledgements. This work was supported by the German Research Foundation (DFG) with grant He1916/8. ECHO-G simulations were performed with SX-6 supercomputers at DKRZ, Germany and KISTI, Korea. We acknowledge the international modeling groups for providing their data for analysis.

REFERENCES

Berger, J. O., 1985: *Statistical decision theory and Bayesian analysis*, 2nd eds. Springer, New York, Berlin, 617 p.

Duda, R. O., and P. E. Hart, 1973: *Pattern classification and scene analysis*. John Wiley, Hoboken, N. J., 482 pp.

IDAG (International ad hoc Detection and Attribution Group), 2005: Detecting and attributing external influences on the climate system: A review of recent advances, *J. Climate*, **18**, 1291-1314.

Jones, P. D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001, *J. Climate*, **16**, 206– 223.

Kass, R. E., and A. E. Raftery, 1995: Bayes factors, *J. Am. Stat. Assoc.*, **90**, 773-795.

Lee, T., F. Zwiers, G. Hegerl, X. Zhang, and M. Tsao, 2005: A Bayesian approach to climate change detection and attribution assessment. *J. Climate*, **18**, 2429-2440.

Legutke, S, and R. Voss, 1999: The Hamburg atmosphere-ocean coupled circulation model ECHO-G. *Technical report* No. 18, German Climate Computre Centre (DKRZ), Hamburg, Germany, 62 pp.

Min, S.-K., and A. Hense, 2005a: A Bayesian assessment of climate change using multi-model ensembles. Part I: Global mean surface temperature. *J. Cilmate*, accepted.

———, and ———, 2005b: A Bayesian assessment of climate change using multi-model ensembles. Part II: Regional and seasonal mean surface temperatures. *J. Climate*, to be submitted.

———, ———, and W.-T. Kwon, 2005a: Regional-scale climate change detection using a Bayesian decision method. *Geophys. Res. Lett.*, **32**, L03706, doi:10.1029/2004GL021028.

———, ———, H. Paeth, and W.-T. Kwon, 2004: A Bayesian decision method for climate change signal analysis. *Meteorol. Z.*, **13**, 421-436.

———, S. Legutke, A. Hense, and W.-T. Kwon, 2005b: Internal variability in a 1000-year control simulation with the coupled climate model ECHO-G – I. Near-surface temperature, precipitation and sea level pressure. *Tellus*, **57A**, 605-621.

———, ———, ———, and ———, 2005c: Internal variability in a 1000-year control simulation with the coupled climate model ECHO-G – II. El Niño Southern Oscillation and North Atlantic Oscillation. *Tellus*, **57A**, 622-640.

———, ———, ———, U. Cubasch, W.-T. Kwon, J.-H. Oh, and U. Schlese, 2005d: East Asian climate change in the 21st century as simulated by the coupled climate model ECHO-G under IPCC SRES scenarios. *J. Meteorol. Soc. Jpn.*, accepted.

Mitchell, J. F. B., D. J. Karoly, G. C. Hegerl, F. W. Zwiers, M. R. Allen, and J. Marengo, 2001: Detection of climate change and attribution of causes. In: *Climate Change 2001: The Scientific Basis*. Houghton, J. T. et al., (eds.) Cambridge University Press, UK.

Schnur, R., and K. Hasselmann, 2005: Optimal filtering for Bayesian detection and attribution of climate change. *Climate Dyn.*, **24**, 45-55.

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, **30**, 1728, doi:10.1029/2003GL017324.