A New Methodology for Estimating the Unpredictable Component of Seasonal Atmospheric Variability

ARUN KUMAR

Climate Prediction Center, NCEP/NOAA, Camp Springs, Maryland

BHASKAR JHA AND QIN ZHANG

RSIS, Climate Prediction Center, NCEP/NOAA, Camp Springs, Maryland

LAHOUARI BOUNOUA

Biospheric Sciences Branch, Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 19 January 2006, in final form 20 September 2006)

ABSTRACT

Predictability limits for seasonal atmospheric climate variability depend on the fraction of variability that is due to factors external to the atmosphere (e.g., boundary conditions) and the fraction that is internal. From the analysis of observed data alone, however, separation of the total seasonal atmospheric variance into its external and internal components remains a difficult and controversial issue. In this paper a simple procedure for estimating atmospheric internal variability is outlined. This procedure is based on the expected value of the mean square error between the observed and the general circulation model simulated (or predicted) seasonal mean anomaly. The end result is a spatial map for the estimate of the observed seasonal atmospheric internal (or unpredictable) variability. As improved general circulation models become available, mean square error estimated from the new generation of general circulation models can be easily included in the procedure proposed herein, bringing the estimate for the internal variability closer to its true estimate.

1. Introduction

Predictability of seasonal climate anomalies can arise from two possible sources: 1) boundary conditions external to the atmosphere [e.g., sea surface temperatures (SSTs)] and 2) atmospheric initial conditions. Within the paradigm of seasonal atmospheric predictability due to external boundary conditions, the potential for skillful predictions depends on the fraction of the atmospheric seasonal mean variability that is related to the anomalous boundary conditions and the fraction that is internal to the atmosphere. The influence of the anomalous boundary conditions on the atmospheric seasonal variance can be further augmented by the influence of atmospheric initial conditions, and in general, this influence depends on the separation between the initial condition and the target forecast season (i.e., the forecast lead time).

Based on observational data alone, however, the separation of seasonal atmospheric variance into its external and internal components, as well as determining the influence of atmospheric initial conditions on seasonal mean variability, remain difficult and controversial tasks. The difficulty arises because for the individual realizations of observed seasonal mean atmospheric anomalies, the estimation of boundary forced and internal components of the atmospheric variance, as well as the influence of initial conditions on them, cannot be made. Alternate approaches for the estimation of seasonal internal variability based on the daily atmospheric variability have been proposed (Madden 1976; Shea and Madden 1990). These methods rely on the analysis of the autocorrelation of daily observations to infer the variance of monthly and seasonal time averages and their comparison with the corresponding ob-

Corresponding author address: Dr. Arun Kumar, Climate Prediction Center, 5200 Auth Rd., Rm. 800, Camp Springs, MD 20746.
E-mail: arun.kumar@noaa.gov

DOI: 10.1175/JCLI4216.1

served interannual variability. Such methods also rely on various assumptions (e.g., changes in boundary conditions have no influence on the characteristics of daily variability). Such assumptions could lead to erroneous estimates for the internal and external variability of seasonal means (Shukla 1983; Trenberth 1984; Zwiers 1987).

An alternate approach for estimating seasonal climate predictability is the use of atmospheric general circulation models (AGCMs). For example, for decomposition of the seasonal mean atmospheric variability into its internal and external components, long multiple realizations of AGCM simulations starting from different atmospheric initial conditions, but forced with identical evolution for the observed boundary conditions (the so called AMIP simulations), are made. The ensemble mean of the AGCM-simulated anomaly is the atmospheric response to the observed boundary forcing, whereas the departure from the ensemble mean is the component of seasonal mean that is internal to the atmosphere, making it possible to estimate the atmospheric external and internal variances (Barnett 1995; Harzallah and Sadourny 1995; Kumar and Hoerling 1995). A similar setup can also be used for estimating the influence of atmospheric initial conditions on seasonal means; except for this case, AGCM simulations, in contrast to the long AMIP integrations, start from observed initial conditions and are of short duration (Branković and Palmer 2000; Shukla et al. 2000).

Although ensemble AGCM simulations can be used to decompose seasonal mean atmospheric variability into its external and internal components, such a procedure leads to an estimate that is an AGCM's rendition of observed atmospheric variability and could be biased by the AGCM errors. Indeed, AGCM-based estimates of external and internal components of seasonal mean variability show a large range of variations from one AGCM to another (Shukla et al. 2000; Kumar et al. 2000).

To lessen the influence of AGCM biases, in this paper an approach for estimating the upper bound for the observed internal variability is outlined, based on the aggregation of simulations from many different AGCMs. This procedure provides a local measure for the seasonal mean atmospheric internal variability. The estimate depends on the selection of the least-biased AGCM among the collection of AGCM simulations one has. Further, as models improve, the procedure described in this paper can incorporate ensemble simulations from the next generation models, and estimates for the observed atmospheric internal variability obtained herein can be easily updated. The procedure for estimating the atmospheric internal variability is described in section 2, and results are presented in section 3. As our estimates for the internal and external variance in section 3 are based on the specification of SST boundary conditions alone, a discussion in section 4 includes a review of factors that are not included in our analysis but may influence estimates of external and internal variance (e.g., the influence of atmospheric initial conditions, coupled air–sea interactions, etc.). Concluding remarks are presented in section 5.

2. Analysis procedure and data

Within the paradigm of seasonal atmospheric predictability due to the observed interannual variability in SSTs, the observed seasonal mean anomalies O_j for the season "j" can be written as a sum of the atmospheric response μ_{oj} to SSTs and the component $\varepsilon_{\mathrm{oj}}$ due to the atmospheric internal variability:

$$O_j = \mu_{\mathrm{oj}} + \varepsilon_{\mathrm{oj}}. \tag{1}$$

From the perspective of SSTs as the external forcing, the atmospheric response is considered to be potentially predictable, while the internal variability represents the unpredictable component of the observed variability. For ensemble simulations from the AGCM, the simulated *ensemble mean* anomaly M_j can be written as the sum of the atmospheric response μ_{mj} and the component $\varepsilon_{\mathrm{mj}}$ due to atmospheric internal variability:

$$M_j = \mu_{\mathrm{mj}} + \varepsilon_{\mathrm{mj}}. \tag{2}$$

Because of the AGCM biases, the AGCM-simulated seasonal mean atmospheric response to the boundary forcing need not be the same as its observed counterpart. Further, as the atmospheric internal variability for the AGCM is based on the ensemble mean, it has smaller amplitude than its observed counterpart (Kumar and Hoerling 1995). In fact, for large ensemble sizes, the internal variability component in (2) approaches zero. The corresponding estimates for observed (σ_o^2) and model-simulated variance of ensemble mean (σ_m^2) are given by

$$\sigma_o^2 = \sigma_{oe}^2 + \sigma_{\mathrm{oi}}^2 \tag{3}$$

and

$$\sigma_m^2 = \sigma_{\mathrm{me}}^2 + \sigma_{\mathrm{mi}}^2, \tag{4}$$

where subscript i and e on the right-hand side of Eqs. (3) and (4) refer to internal and external variance, respectively. The internal variance for the ensemble mean of AGCM simulations is related to the internal variance of a single AGCM simulation by a multiplicative factor of $1/n$, where n is the number of realizations in the ensemble (Kumar and Hoerling 1995). As men-

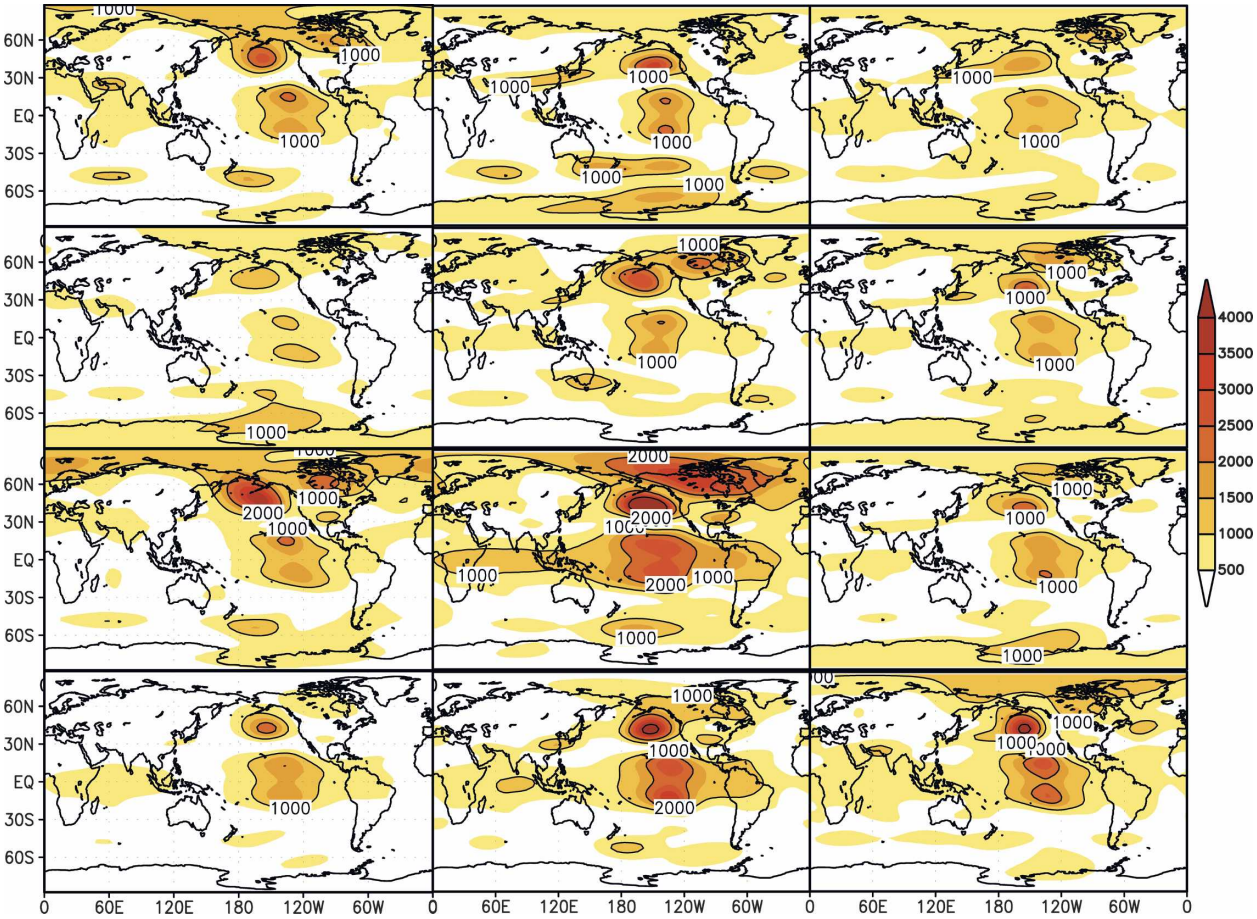

FIG. 1. Ensemble mean variance (m^2) for DJF seasonal mean 200-mb heights simulated by different AGCMs. AGCM simulations are forced by the observed SSTs and are from 1951 to 2000, and ensemble sizes for different AGCMs range from 7 to 24. The ensemble mean variance is indicative of SST-forced atmospheric variability.

tioned earlier, for large ensembles the internal variability of the ensemble mean of AGCM simulations approaches zero.

For a season j, the mean-square error (MSE) of the ensemble mean prediction is

$$\mathrm{MSE}_j = (M_j - O_j)^2, \tag{5}$$

and its expected value, under the assumption that the observed and the modeled internal variability are uncorrelated, is given by

$$\mathrm{MSE} = \sigma_{oi}^2 + \sigma_{mi}^2 + \langle(\mu_{mj} - \mu_{oj})^2\rangle. \tag{6}$$

The expected value of MSE is the sum of three terms: the observed internal variability, the internal variability of the ensemble mean of AGCM simulations, and a term that is the error in the *model-simulated atmospheric response relative to the observed response to SSTs*. For large ensembles and an AGCM with unbiased atmospheric response, MSE equals the observed internal variability and is the lower bound for the MSE in the paradigm of boundary-forced seasonal predictions. For other cases (e.g., small ensembles and errors in the AGCM's atmospheric response to SSTs), the MSE for the ensemble mean of AGCM simulations is constrained to always be larger than the observed internal variability. We use this property of MSE *to estimate the internal variability of the observed seasonal means.*

In the present analysis, ensemble mean simulations from 11 different AGCMs are used. All the AGCM simulations are the AMIP-type long integrations for 1951–2000 and are forced by the evolution of the observed SSTs. Depending on the model biases and ensemble size for each AGCM, the spatial distribution of MSE for each AGCM differs. At each geographical location, we next find the *minimum value of MSE* irrespective of which AGCM it came from, and the spatial

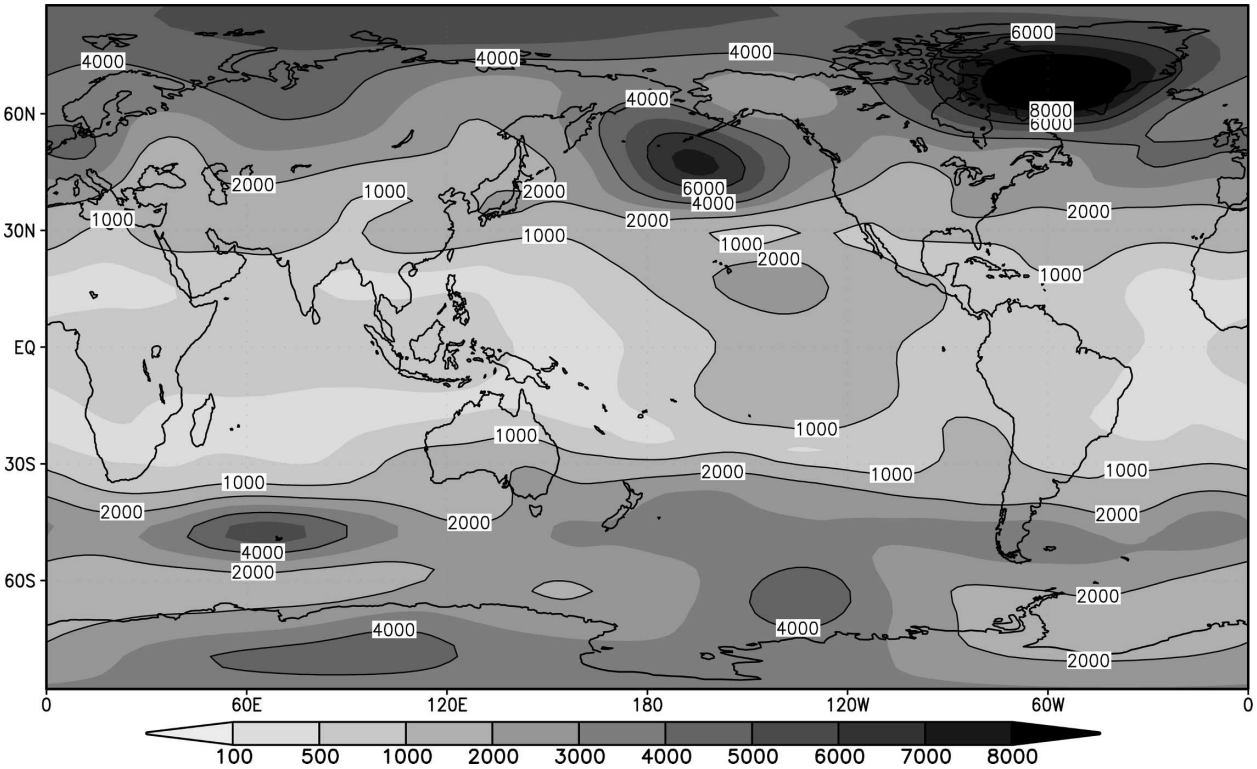

FIG. 2. Variance of the observed (or reanalysis; m^2) DJF seasonal mean 200-mb heights. The variance is computed from the observed seasonal means from 1951 to 2000.

map of the minimum value of MSE is our best estimate for the observed atmospheric internal variability.

To illustrate the application of the proposed procedure, we focus on the December–February (DJF) 200-mb seasonal mean heights. The National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis forms the basis for the observed anomalies (hereafter referred to as the observations), and model-simulated heights are the ensemble mean height anomalies. The minimum ensemble size for AGCM simulations is 7 while the maximum ensemble size is 24. We also utilize the average of ensemble mean anomalies from all 11 AGCMs and consider the average (or superensemble, hereafter referred to as the twelfth AGCM) as an additional specification of the seasonal mean anomaly. The length of all the AGCM simulations is 49 DJFs from 1951 to 2000. The observed and model-simulated 200-mb height anomalies are the departures from their respective climatologies.

3. Results

Shown in Fig. 1 is the ensemble mean variance of 200-mb height anomalies for all 12 AGCMs in our analysis. This variance corresponds to model variance in (4) and is dominated by the SST-forced atmospheric variability. There is a remarkable degree of consistency in the spatial pattern of variance across different models. The spatial structure is also consistent with the expected spatial structure of the atmospheric response to the tropical SST anomalies related to El Nino–Southern Oscillation (ENSO; Trenberth et al. 1998). The amplitude of external variance differs across AGCMs because of differing model biases influencing the atmospheric response to imposed SSTs, as well as the influence of different ensemble sizes leading to differing contributions of the internal variability [i.e., the second term on the right-hand side of Eq. (6)].

The total variance of 200-mb observed DJF seasonal mean 200-mb heights [as in Eq. (3)] is shown in Fig. 2 and is the variance whose internal variability component we seek to estimate. Because of the contribution of the atmospheric internal variability to the seasonal means, the observed variance, particularly in the extratropical latitudes, is much larger than the variance for different AGCMs shown in Fig. 1 (where the process of ensemble averaging leads to a considerable reduction in the contribution of the model-simulated atmospheric internal variability).

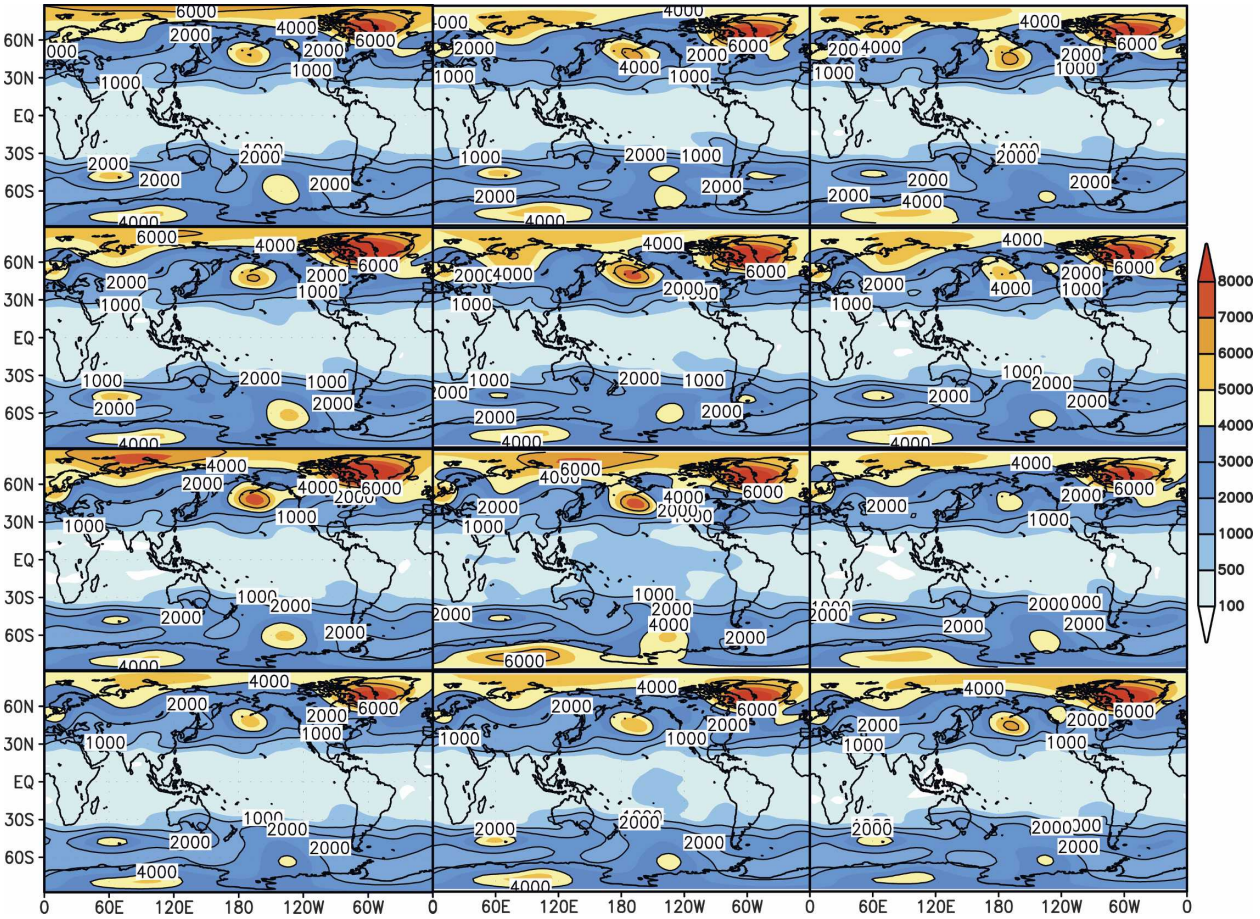

FIG. 3. MSE (m^2) of the seasonally averaged DJF 200-mb heights for different AGCM simulations. MSE is computed over the 1951–2000 period.

The MSE for the ensemble mean AGCM-simulated anomaly for all 12 AGCMs is shown in Fig. 3. For different AGCMs, such plots correspond to (6) and are different estimates for the observed seasonal internal variability of 200-mb heights. Once again, the estimates for different AGCMs differ because of differing contributions from the AGCMs' internal variability [the second term on the right-hand side of (6)] and because of differences in AGCMs' characteristic responses to SSTs [the third term on the right-hand side of (6)].

At each grid point, the AGCM for which the combination of the error in the atmospheric response to SSTs and the contribution of the AGCM-simulated internal variability adds least to the estimate of the MSE is the value of MSE closest to the observed internal variability. From 12 different estimates of the observed internal variability, the gridpoint-by-gridpoint minimum value of MSE across 12 AGCMs is shown in Fig. 4 and is our best estimate for *the internal variability of the observed 200-mb seasonal mean heights.* Also shown in Fig. 4 (bottom panel) are the regions over which the minimum value of MSE is significant at the 95% confidence level based on the Monte Carlo approach (see appendix for details and further discussion). The estimate of internal variability in Fig. 4, by design of the experimental setup of AGCM simulations, is the atmospheric variability that is not related to the interannual variability of observed SSTs and cannot be predicted from the specification of SSTs alone.

One can also estimate the SST-forced external variability of the observed 200-mb seasonal mean heights by simply subtracting the total variance in Fig. 2 from the estimate of internal variability in Fig. 4. This is shown in Fig. 5 (top panel). This estimate of external variability is also compared with the variance of 200-mb observed heights that is linearly associated with Niño-3.4 SST variability (Fig. 5, bottom panel). In the Northern Hemisphere there is a remarkable degree of spatial resemblance between the two estimates. In the Southern Hemisphere, the estimate of SST-related atmospheric variability shows a zonal band of external variability between 30° and 60°S that is not present in the

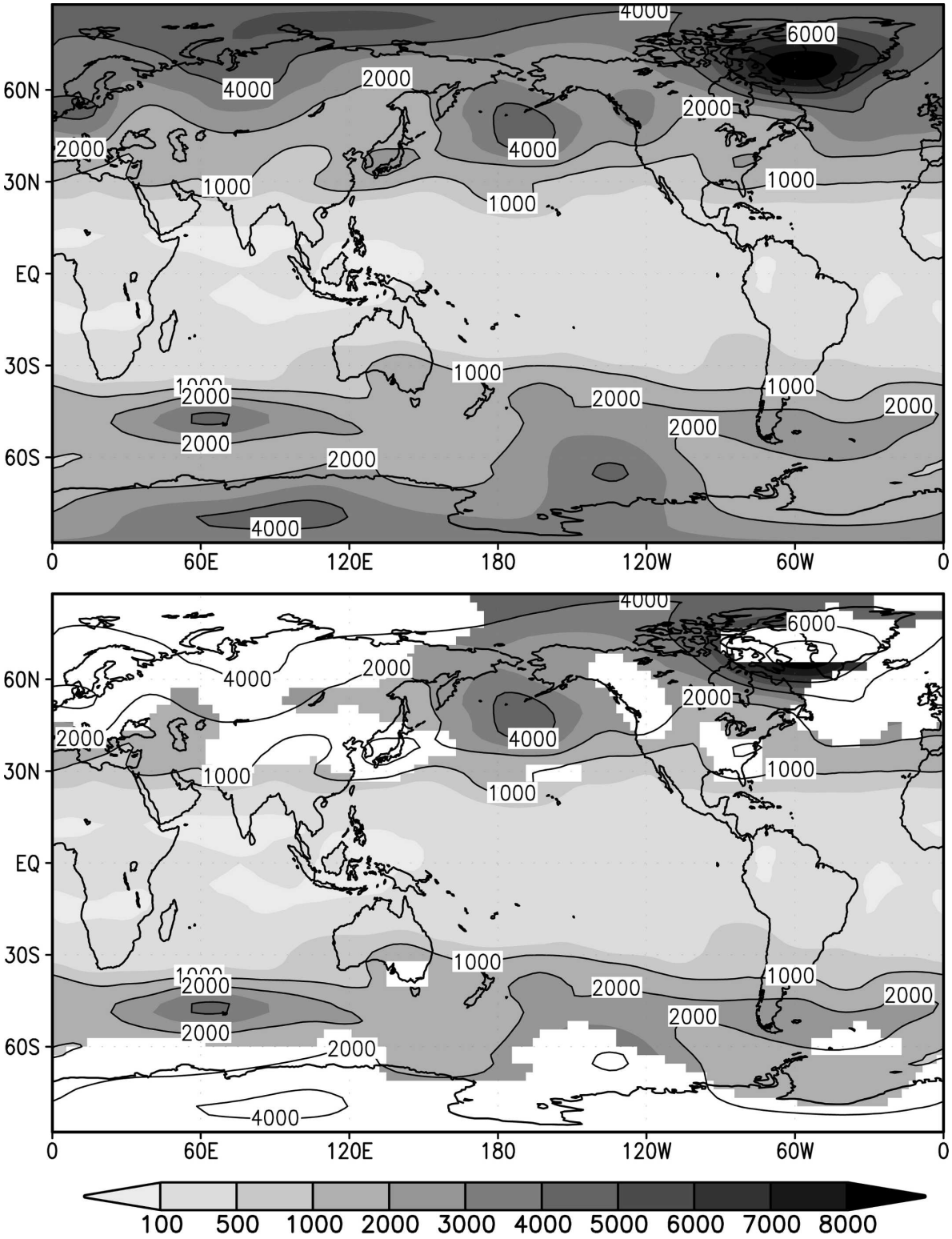

FIG. 4. (top) The best estimate of the internal variance of observed 200-mb DJF seasonal mean heights (m^2). The best estimate is obtained from the gridpoint-by-gridpoint minimum value of MSE for different AGCMs in Fig. 3. (bottom) A replot of the top panel with shaded regions indicating where the estimated MSE is significant at the 95% level based on the Monte Carlo approach.

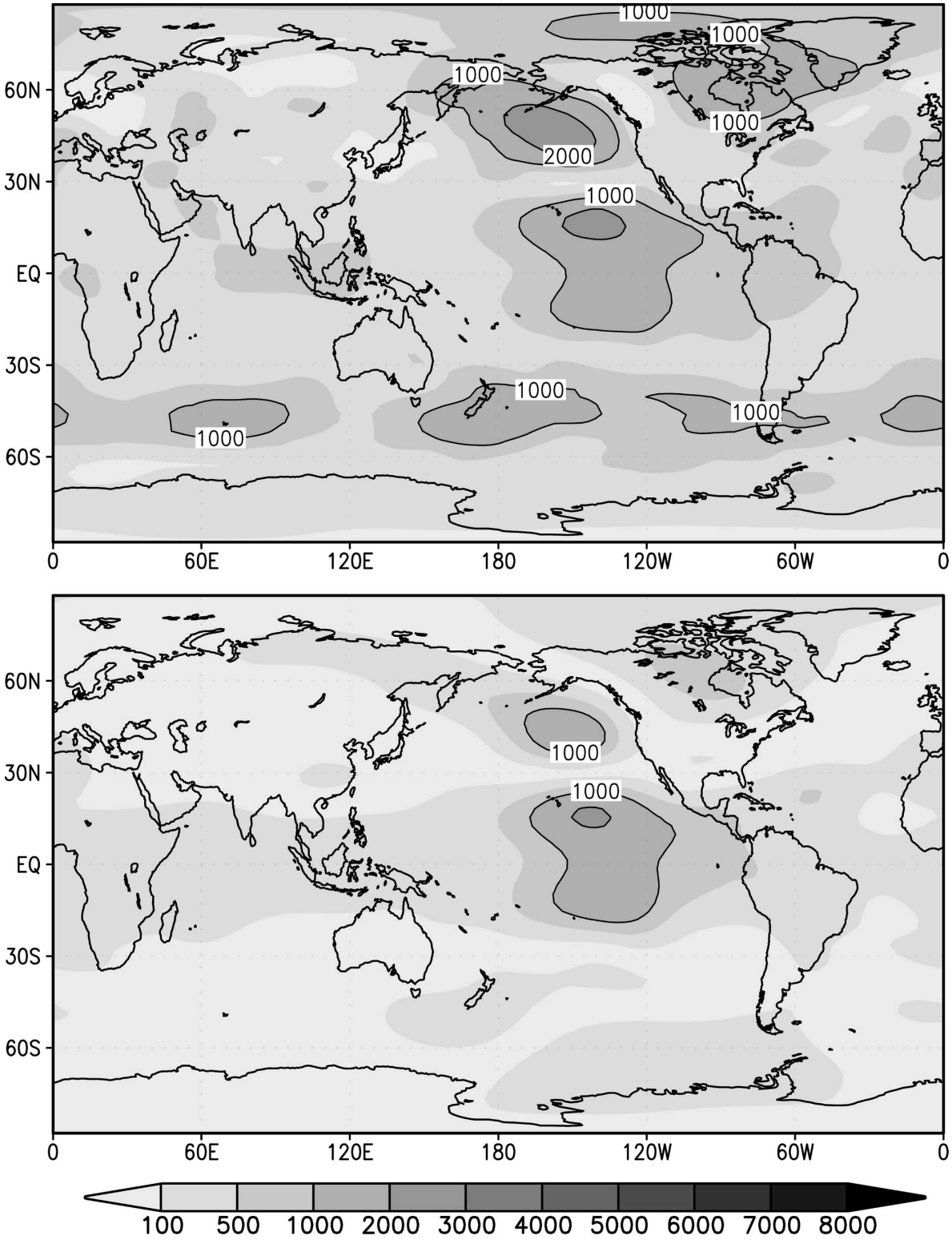

FIG. 5. (top) The best estimate of the external variance of observed 200-mb DJF seasonal mean heights (m^2) obtained by subtracting the best estimate of internal variance (Fig. 4) from the total variance (Fig. 2). (bottom) Observed 200-mb DJF seasonal mean variance linearly related to the Niño-3.4 SST index.

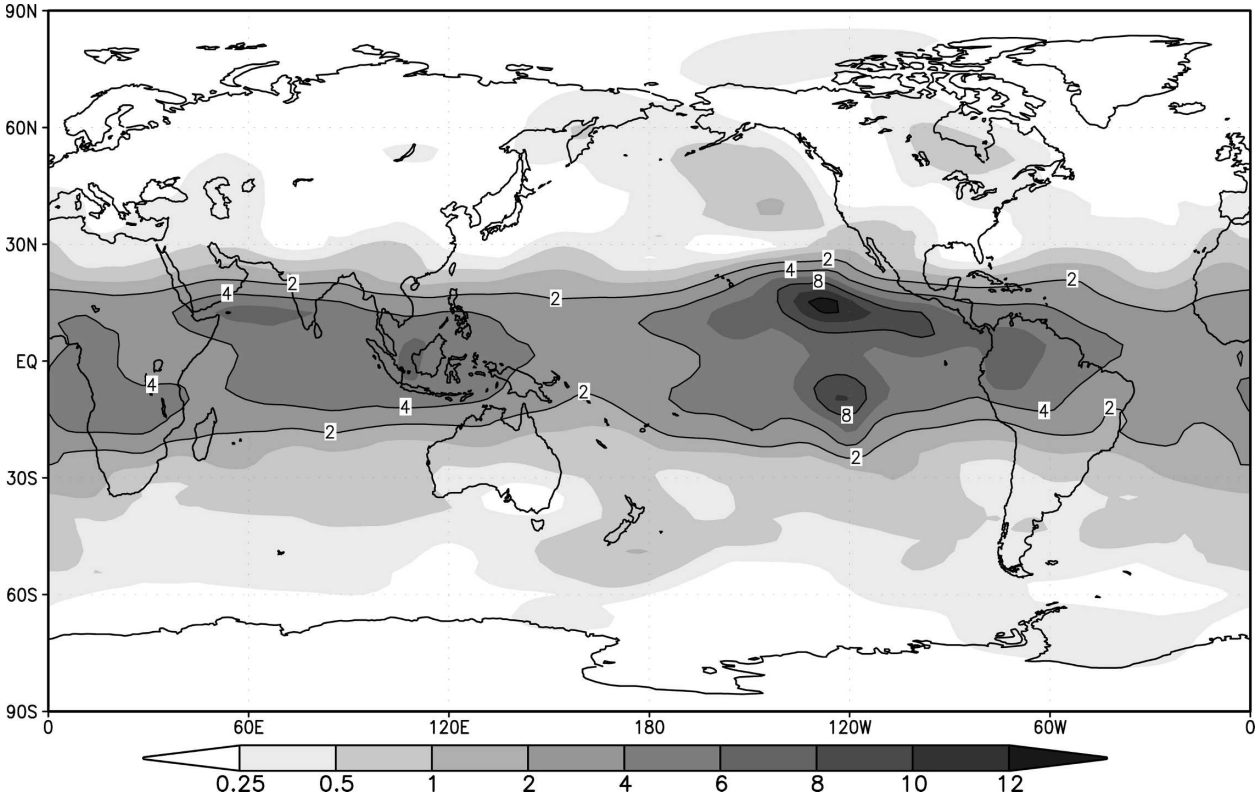

FIG. 6. SN ratio estimate for the DJF seasonal mean 200-mb heights computed as the ratio of external-to-internal std dev in Fig. 5 top and Fig. 4, respectively. Higher ratios imply higher potential predictability.

linear estimate based on the Niño-3.4 SSTs. This is related to the recent trend in the Southern Hemisphere annular mode (Marshall 2003; Marshall et al. 2004; Renwick 2004), which to some extent can be replicated in the AGCM simulations forced with the global SSTs, however is not present in the atmospheric response to Niño-3.4 SSTs.

The estimate of SST-forced variability obtained in the present analysis is slightly larger than the atmospheric variability based on the Niño-3.4 SST index alone. This is to be expected, as the AGCM simulations take into account the linear and nonlinear atmospheric influence of global SSTs. However, the point to note is that the difference is small in comparison with the *magnitude of atmospheric variability that is unrelated to SSTs* (Fig. 4) and will not result in a substantial change in the signal-to-noise (SN) ratio. The small difference in the estimate of the external variability based on the AGCM simulations and that obtained linearly based on the Niño-3.4 SST index also implies the dominance of ENSO on the predictability of seasonal means (see also Hoerling and Kumar 2002).

Shown in Fig. 6 is the spatial pattern of signal-to-noise ratio obtained from our estimates of the observed internal and external variability and computed as the ratio of standard deviation of external and internal variance. The SN ratio is largest in the tropical latitudes, decreases in the extratropical latitudes, and conforms to a long history of the analysis of atmospheric potential predictability due to SSTs (Kumar and Hoerling 1995; Stern and Miyakoda 1995; Rowell 1998; Zwiers et al. 2000; Peng et al. 2000; Straus et al. 2003; Kumar et al. 2003).

The SN ratio has direct relevance for the other measures of seasonal prediction skill (Kumar and Hoerling 2000; Sardeshmukh et al. 2000). In general, the higher the SN ratio, the higher the skill for the seasonal prediction. The theoretical relationship between the SN ratio and the expected value of anomaly correlation (AC) for the atmospheric response as the prediction is shown in Fig. 7 (top panel). Similar relationships can be derived for any skill metric (Kumar et al. 2001). The relationship between the AC and the SN ratio is consistent with our a priori expectations. For example, the expected value of AC is higher for higher SN ratios, and asymptotes to one for a large SN ratio.

Also shown in Fig. 7 (bottom panel) is the scatterplot of SN ratio and the maximum value of the anomaly

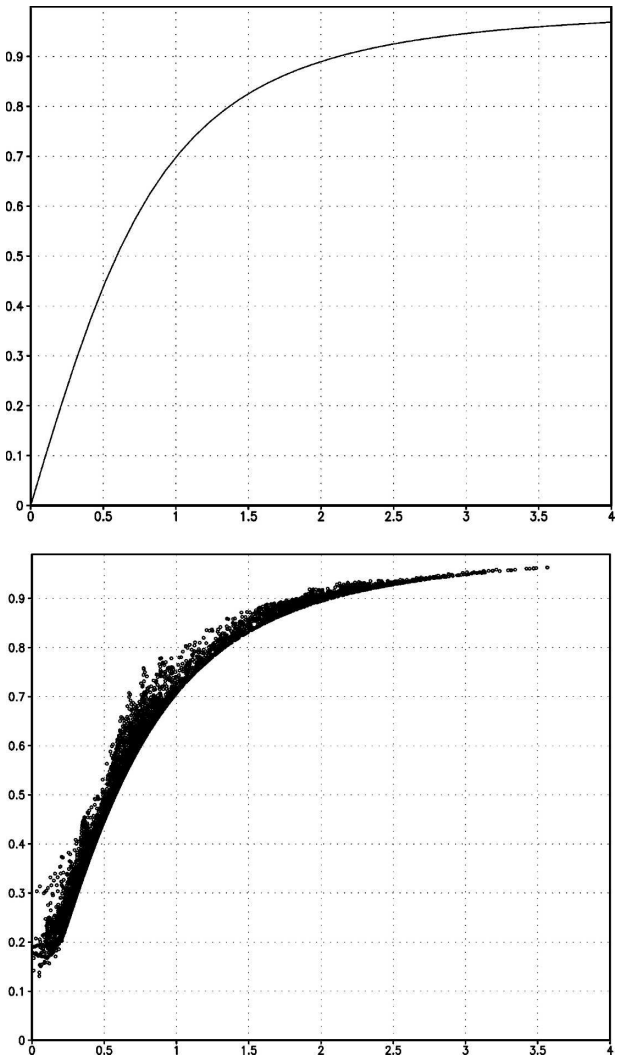

FIG. 7. (top) Theoretical relationship between SN (x axis) and the corresponding expected value of AC (y axis). (bottom) The same relationship obtained as a scatterplot between the SN ratio in Fig. 6 and the maximum value of the AC.

correlation (which is defined as the point-by-point maximum value of anomaly correlations from 12 anomaly correlations that are obtained as the temporal correlation between the AGCM-simulated and the observed anomalies over the 49 DJFs). A close resemblance between the scatterplot in Fig. 7 (bottom panel) and the theoretical relationship in Fig. 7 (top panel) is striking and reconfirms the theoretical relationship between the SN ratio and the expected value of predictive skill.

4. Discussion

In the analysis above, the decomposition of the observed seasonal variance was based on long AMIP simulations in which the only externally specified forcing was SSTs. Accordingly, the decomposition of observed variance into external and internal components and the corresponding SN ratio and predictability estimates relate to an estimate of the potential predictability of seasonal mean atmospheric anomalies due to SSTs alone. There are at least three factors that are missing from the experimental setup of AMIP simulations that could also have an influence on the decomposition of seasonal mean variance (and our estimate of potential predictability). These are discussed next.

a. Influence of atmospheric initial conditions

The outstanding scientific question about the influence of atmospheric initial conditions on the seasonal predictability is quantifying the role of atmospheric initial conditions in the seasonal mean variability of the subsequent season. In the AMIP simulations, the atmospheric variability is far removed from the initial conditions from which the integrations began, and the statistical characteristics of seasonal mean anomalies are consistent with observed SST forcing alone.

One can also envision an ensemble of AGCM predictions from the observed atmospheric initial states, with different AGCM integrations starting from slightly perturbed initial conditions. This is a common practice for the medium-range weather predictions and can also be extended to seasonal predictions. In this setup of model integrations, it is possible that the memory of atmospheric conditions may persist and influence the atmospheric seasonal variability.

There have been at least two coordinated efforts to quantify seasonal atmospheric predictability with AGCM integrations starting from the observed initial conditions and forced with the observed SSTs (Branković and Palmer 2000; Shukla et al. 2000). The MSE error of heights and corresponding predictability estimates obtained from these runs do not seem to differ dramatically from the ones obtained here. As an example, the root-mean-square error (RMSE) of 500-mb heights over the Pacific–North America (PNA) region in Table 2 of Shukla et al. (2000) ranges between 29 and 31 for different AGCMs. One interpretation for the "narrow range" of RMSE is that it represents the observed internal variability. For the 200-mb heights, the RMSE in our case varies between 40 and 45 for 12 different AGCMs. Following an alternate approach, Phelps et al. (2004) and Peng and Kumar (2005) analyzed differences in seasonal mean atmospheric variability for a target season with different lead times from the initial conditions. Their results indicate that the variability of seasonal means approaches its climatological value within a month of the start of AGCM

integrations, implying that memory from the initial conditions is lost rapidly and has little influence on constraining the envelope of seasonal atmospheric variability.

Another variant of the AGCM integrations is the use of predicted SSTs (e.g., persisting the observed SST anomalies at the start of integrations into the forecast period; Derome et al. 2001; Frederiksen et al. 2001) together with the start of integrations from the observed initial conditions. However, it is unlikely that the predictability estimates based on the use of predicted SSTs will be higher than the "potential predictability" estimates when observed SSTs are specified.

b. Influence of coupled air–sea interactions

AGCM simulations with specified observed SSTs lack another crucial factor (i.e., the coupled ocean–atmosphere evolution). It is possible that the inclusion of coupled evolution may lead to a higher estimate of the variance that is external (or predictable), thereby reducing the MSE of the prediction and leading to a smaller estimate for the internal (or unpredictable) component of the observed variability. Model integrations that are most relevant for this formalism are the coupled seasonal forecasts from the observed ocean and atmospheric initial conditions, as long coupled model simulations that can also replicate the observed history of SSTs are not feasible.

The above described coupled model integrations are similar to AGCM integrations with observed atmospheric initial conditions discussed in section 4a, the difference being that the SSTs are now predicted and the evolution of SST among different realizations need not be the same as the observed history of SST. As a consequence, while on the one hand predictability estimates might improve because of the inclusion of coupled ocean–atmosphere evolution, on the other hand the predictability estimate might degrade because SST is no longer predicted perfectly.

The availability of the Development of a European Multimodel Ensemble System for Seasonal to Interannual Prediction (DEMETER) dataset (Palmer et al. 2004) provides an opportunity to assess how the inclusion of coupled ocean–atmosphere evolution changes the component of observed atmospheric variability that is unpredictable, and how it differs from the estimate of the unpredictable component obtained from the AMIP simulations (Fig. 4). Shown in Fig. 8 is the estimate of observed atmospheric internal (top panel) and external (bottom panel) variability based on coupled forecasts from the DEMETER project. In this analysis we used 6 out of 7 models from the DEMETER project, and the analysis period is for DJF 1974–2000. One of the DEMETER models was not used and this allowed the analysis to start from 1974 (instead of starting from 1980), thereby increasing the sample size and adding six additional years to the analysis. Otherwise, the analysis procedure is the same as for obtaining Figs. 4, 5 (top panels).

The estimate of observed internal and external variability with the inclusion of coupled evolution (Fig. 8) is similar to the one obtained based on the AMIP simulations (Figs. 4, 5, top panels). In the tropical and Northern Hemisphere extratropical latitudes, the external variability for DEMETER (Fig. 8, bottom panel) is higher than its AMIP counterpart (Fig. 5, top panel) because of the tendency for large ENSO events to be concentrated during this period. This is consistent with the fact that the linear ENSO signal for the 1974–2000 period (not shown) is also higher than the corresponding signal for the 1951–2000 period (Fig. 5, bottom panel). In the Southern Hemisphere extratropical latitudes, on the other hand, the estimate of external variability based on DEMETER is closer to its counterpart based on ENSO (Fig. 5, bottom panel) and is consistent with the fact that beyond the core region of ENSO SST variability in the equatorial tropical eastern Pacific, SST predictions based on a coupled model are not skillful in maintaining SST trends and also result in the loss of corresponding atmospheric response.

The analysis of atmospheric variability based on DEMETER implies that the influence of the coupled evolution on the seasonal mean atmospheric variability that is internal to the atmosphere may not be significant, and therefore, predictability estimates obtained from the AMIP simulations may also represent what is achievable in the short lead coupled forecasts. An additional point to note is that DEMETER runs are initialized with the observed initial conditions and are consistent with the discussion in the previous section, implying that the influence of initial conditions on the predictability of seasonal means is small.

c. Influence of boundary conditions other than SSTs

Another factor that could have an influence on the predictability of seasonal climate variability is the initialization of land boundary conditions or interannual changes in vegetation amount (Bounoua et al. 2000). On a conceptual level, the assessment of the influence of land initial boundary conditions on seasonal predictability follows the same approach as the one discussed for the assessment of coupled ocean–atmosphere integrations (i.e., an ensemble of short model integrations from perturbed land–atmosphere initial conditions).

A crucial question related to the possible influence of land boundary conditions on the seasonal atmospheric variability is their potential to constrain the tropo-

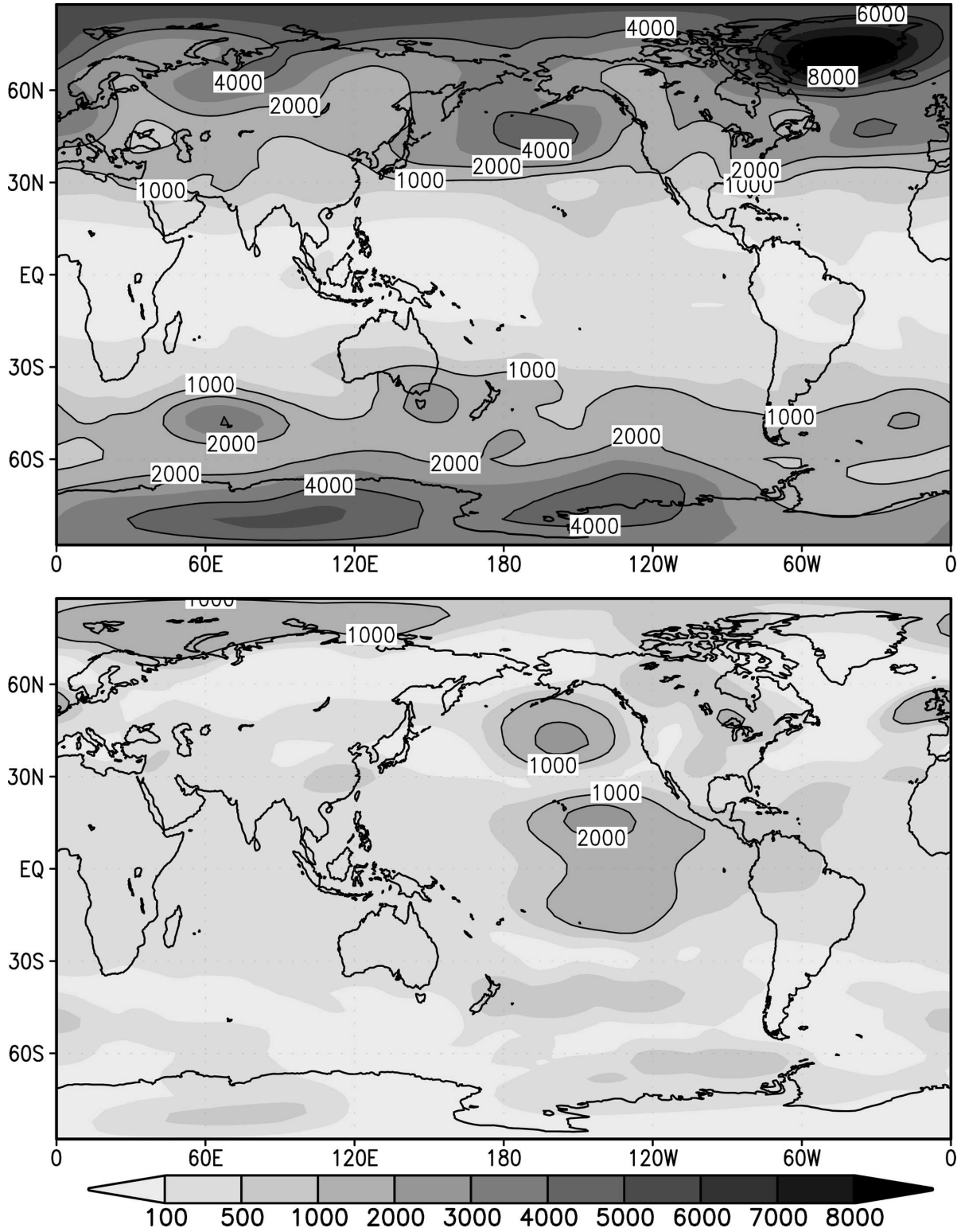

FIG. 8. (top) The best estimate of the internal variance of observed 200-mb DJF seasonal mean heights obtained from the coupled models participating in the DEMETER project. This estimate is the same as the one shown in Fig. 4 based on AMIP simulations. (bottom) The best estimate of the external variance of observed 200-mb DJF seasonal mean heights obtained by subtracting the best estimate of internal variance in the top panel from the total observed variance (Fig. 2).

spheric atmospheric variability. For example, if the influence of anomalous land boundary conditions (e.g., snow and soil moisture) remains confined to the lower troposphere, the time scale of their influence would be similar to their decay time scale in isolation. On the other hand, if the influence of anomalous land boundary conditions penetrates to the upper troposphere and is able to constrain large-scale circulation, predictability associated with the land initial conditions is expected to be larger (and longer lasting). Published research so far points to the former, that is, the influence of land boundary conditions is confined to the lower troposphere alone. For example, Kumar and Yang (2003) compared the vertical extent of the influence of extratropical snow and tropical SST variability over North America. Their results demonstrate that although the remote influence of tropical SST variability on the seasonal mean variability over NA extends throughout the troposphere, the extent of the influence of snow anomalies was confined to the lower troposphere.

Also consistent with above results are the estimates of lag correlations between the soil moisture, surface temperature, and rainfall anomalies. Almost all observational and model results suggest the larger influence of soil moisture anomalies on the surface temperature on 1–2-month time scales (which results from the direct influence of soil moisture anomalies on the partitioning of sensible and latent heat fluxes) and a much weaker influence on the rainfall anomalies (which are more a reflection of large-scale circulation features; Wu and Dickinson 2004; Kanamitsu et al. 2003; Wang and Kumar 1998; Huang et al. 1996).

5. Concluding remarks

The estimate for the internal variability of DJF 200-mb observed seasonal mean heights in Fig. 4 is our best estimate of the observed internal (or unpredictable) component of variability based on the current generation of AGCMs (and data available to us) that is not related to the observed history of SSTs. Similarly, Fig. 8 replicates the best estimate of the predictability of seasonal atmospheric climate anomalies based on the DEMETER dataset that includes the influence of observed ocean, atmosphere, and land initial states, as well as the influence of realistic coupling between different components of the earth's system. Simulations from the future generation of model integrations and the corresponding spatial map for MSE can be used to update the spatial map of the unpredictable component of variability in Fig. 4 (and in Fig. 8). At the geographical locations, where the MSE for models is higher than the current estimate, this will not lead to any update in the estimate of the unpredictable component of variability. Only at the geographical location where the estimate of MSE is lower than the estimate of variability in Fig. 4 (Fig. 8) will a lower (higher) estimate of atmospheric internal variability seasonal climate predictability be found. It remains to be seen how much of the internal variability from estimates based on the current generation of the model simulations, and shown in Figs. 4, 8, can be moved to the variance that is predictable because of the improved models, higher resolution, improved initial conditions, etc. Based on the unique properties of MSE, in this paper we provide a methodology that could be used to document such improvements.

Acknowledgments. Support provided by NOAA's Office of Global Programs "Climate Dynamics and Experimental Prediction" program is gratefully acknowledged. DEMETER data were provided by the European Centre for Medium-Range Weather Forecasts. Data for 11 AMIP simulations were provided by the International Research Institute for Climate and Society (IRI), Geophysical Fluid Dynamics Laboratory (GFDL), Global Modeling and Assimilation Office (GMAO), Climate Diagnostics Center (CDC), and Experimental Climate Prediction Center (ECPC), which are gratefully acknowledged. Constructive comments from two anonymous reviewers and Drs. Wanqiu Wang and Peitao Peng helped greatly to improve the manuscript.

APPENDIX

Test of Significance

The statistical significance for the internal variability in Fig. 4 (and therefore for the external variability in Fig. 5, top panel) is tested based on the Monte Carlo approach. Recall that the internal variability estimate in Fig. 4 is based on the following steps:

- computing the mean-square error between the observed and model-simulated ensemble mean 200-mb heights
- repeating the above step for 12 different models providing 12 different estimates for the MSE
- finding the point-by-point minimum value of MSE resulting in the estimate of internal variability shown in Fig. 4.

To test what the probability is that this MSE can be obtained by chance alone, steps 1–3 were repeated after first randomizing observed and model-simulated time series. This randomization was done for the observed

and the modeled time series independently. For each randomized sample, the minimum value of MSE, which is equivalent to Fig. 4, is computed. Following this procedure, we generated 10 000 estimates of the MSE. From these estimates of the MSE, the percent of samples for which the MSE was smaller than the one in Fig. 4 is obtained. Based on this procedure, regions where the MSE in Fig. 4 is smaller than other estimates of the MSE (based on 10 000 samples) 95% of the time are highlighted. In other words, regions where the MSE shown in Fig. 4 is highly likely to be different from what could be obtained by chance alone are highlighted.

The randomization for the observed and the modeled time series could be achieved in two different ways (i.e., with and without replacement). Randomization without replacement was chosen since reshuffling the years between 1951 and 2000 preserves the total variance of the observed and modeled time series. This constraint is not guaranteed for randomization with replacement (e.g., just by chance years with low anomalies can be repeated, leading to small variance for the observed anomalies). This has the consequence that for lots of samples, the MSE could be less than the one in Fig. 4 because variability for the randomized sample itself is lower.

From a comparison of the shaded region in Fig. 4 (bottom panel) with Fig. 6, it is immediately apparent that in almost all regions where the signal-to-noise ratio exceeds 0.25, the MSE is also significant and cannot be obtained by chance alone. Recall that at a gridpoint level, the MSE in Fig. 4 is the minimum of that obtained from 12 different models. The MSE for that particular model is minimized because of the covarying observed and the ensemble mean model-simulated signal. This can be easily inferred from expanding Eq. (6). For regions where the signal-to-noise ratio exceeds 0.25, any randomization of the observed or the model-simulated time series by making the covariability term close to zero leads to an increase in the estimated value of MSE.

REFERENCES

Barnett, T. P., 1995: Monte Carlo climate forecasting. *J. Climate,* **8,** 1005–1022.

Bounoua, L., G. J. Collatz, S. O. Los, P. J. Sellers, D. A. Dazlich, C. J. Tucker, and D. A. Randall, 2000: Sensitivity of climate to changes in NDVI. *J. Climate,* **13,** 2277–2292.

Branković, Č., and T. N. Palmer, 2000: Seasonal skill and predictability of ECMWF PROVOST ensembles. *Quart. J. Roy. Meteor. Soc.,* **126,** 2035–2067.

Derome, J., G. Brunet, A. Plante, N. Gagnon, G. J. Boer, F. W. Zwiers, S. Lambert, and H. Ritchie, 2001: Seasonal predictions based on two dynamical models. *Atmos.–Ocean,* **39,** 485–501.

Frederiksen, C. S., H. Zhang, R. C. Balgovind, N. Nicholls, W. Drosdowsky, and L. Chambers, 2001: Dynamical seasonal forecasts during the 1997/98 ENSO using persisted SST anomalies. *J. Climate,* **14,** 2675–2695.

Harzallah, A., and R. Sadourny, 1995: Internal versus SST-forced atmospheric variability as simulated by an atmospheric general circulation model. *J. Climate,* **8,** 474–495.

Hoerling, M. P., and A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *J. Climate,* **15,** 2184–2203.

Huang, J., H. M. van den Dool, and K. P. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *J. Climate,* **9,** 1350–1362.

Kanamitsu, M., C.-H. Lu, J. Schemm, and W. Ebisuzaki, 2003: The predictability of soil moisture and near-surface temperature in hindcasts of the NCEP seasonal forecast model. *J. Climate,* **16,** 510–521.

Kumar, A., and M. P. Hoerling, 1995: Prospects and limitations of seasonal atmospheric GCM predictions. *Bull. Amer. Meteor. Soc.,* **76,** 335–345.

——, and ——, 2000: Analysis of a conceptual model of seasonal climate variability and implications for seasonal prediction. *Bull. Amer. Meteor. Soc.,* **81,** 255–264.

——, and F. Yang, 2003: Comparative influence of snow and SST variability on extratropical climate in northern winter. *J. Climate,* **16,** 2248–2261.

——, A. G. Barnston, P. Peng, M. P. Hoerling, and L. Goddard, 2000: Changes in the spread of the variability of the seasonal mean atmospheric states associated with ENSO. *J. Climate,* **13,** 3139–3151.

——, ——, and M. P. Hoerling, 2001: Seasonal predictions, probabilistic verifications, and ensemble size. *J. Climate,* **14,** 1671–1676.

——, S. D. Schubert, and M. S. Suarez, 2003: Variability and predictability of 200-mb seasonal mean heights during summer and winter. *J. Geophys. Res.,* **108,** 4169, doi:10.1029/2002JD002728.

Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.,* **104,** 942–952.

Marshall, G. J., 2003: Trends in the southern annular mode from observations and reanalyses. *J. Climate,* **16,** 4134–4143.

——, P. A. Stott, J. Turner, W. M. Connolley, J. C. King, and T. A. Lachlan-Cope, 2004: Causes of exceptional atmospheric circulation in the Southern Hemisphere. *Geophys. Res. Lett.,* **31,** L14205, doi:10.1029/2004GL019952.

Palmer, T. N., and Coauthors, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

Peng, P., and A. Kumar, 2005: A large ensemble analysis of the influence of tropical SSTs on seasonal atmospheric variability. *J. Climate,* **18,** 1068–1085.

——, ——, A. G. Barnston, and L. Goddard, 2000: Simulation skills of the SST-forced global climate variability of the NCEP–MRF9 and the Scripps–MPI ECHAM3 models. *J. Climate,* **13,** 3657–3679.

Phelps, M. W., A. Kumar, and J. J. O'Brien, 2004: Potential predictability in the NCEP CPC dynamical seasonal forecast system. *J. Climate,* **17,** 3775–3785.

Renwick, J. A., 2004: Trends in the Southern Hemisphere polar vortex in NCEP and ECMWF reanalyses. *Geophys. Res. Lett.,* **31,** L07209, doi:10.1029/2003GL019302.

Rowell, D. P., 1998: Assessing potential seasonal predictability

with an ensemble of multidecadal GCM simulations. *J. Climate,* **11,** 109–120.

Sardeshmukh, P. D., G. L. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate,* **13,** 4268–4286.

Shea, D. J., and R. A. Madden, 1990: Potential for long-range prediction of monthly mean surface temperatures over North America. *J. Climate,* **3,** 1444–1451.

Shukla, J., 1983: Comments on "Natural variability and predictability." *Mon. Wea. Rev.,* **111,** 581–585.

——, and Coauthors, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.,* **81,** 2593–2606.

Stern, W., and K. Miyakoda, 1995: Feasibility of seasonal forecasts inferred from multiple GCM simulations. *J. Climate,* **8,** 1071–1085.

Straus, D., J. Shukla, D. Paolino, S. Schubert, M. Suarez, P. Pegion, and A. Kumar, 2003: Predictability of the seasonal mean atmospheric circulation during autumn, winter, and spring. *J. Climate,* **16,** 3629–3649.

Trenberth, K. E., 1984: Some effects of finite sample size and persistence on meteorological statistics. Part II: Potential predictability. *Mon. Wea. Rev.,* **112,** 2359–2379.

——, G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.,* **103** (C7), 14 291–14 324.

Wang, W., and A. Kumar, 1998: A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *J. Geophys. Res.,* **103** (D22), 28 637–28 646.

Wu, W., and R. E. Dickinson, 2004: Time scales of layered soil moisture memory in the context of land–atmosphere interaction. *J. Climate,* **17,** 2752–2764.

Zwiers, F. W., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Mon. Wea. Rev.,* **115,** 2957–2974.

——, X. L. Wang, and J. Sheng, 2000: Effects of specifying bottom boundary conditions in an ensemble of atmospheric GCM simulations. *J. Geophys. Res.,* **105,** 7295–7316.