Reproducible Forced Modes in AGCM Ensemble Integrations and Potential Predictability of Atmospheric Seasonal Variations in the Extratropics

XIU-QUN YANG*

Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey

JEFFREY L. ANDERSON AND WILLIAM F. STERN

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey

(Manuscript received 27 August 1997, in final form 24 November 1997)

ABSTRACT

An approach to assess the potential predictability of the extratropical atmospheric seasonal variations in an ensemble of atmospheric general circulation model (AGCM) integrations has been proposed in this study by isolating reproducible forced modes and examining their contributions to the local ensemble mean. The analyses are based on the monthly mean output of an eight-member ensemble of 10-yr Atmospheric Model Intercomparison Project integrations with a T42L18 AGCM.

An EOF decomposition applied to the ensemble anomalies shows that there exist some forced modes that are less affected by the internal process and thus appear to be highly reproducible. By reconstructing the ensemble in terms of the more reproducible forced modes and by developing a quantitative measure, the potential predictability index (PPI), which combines the reproducibility with the local variance contribution, the local ensemble mean over some selective geographic areas in the extratropics was shown to result primarily from reproducible forced modes rather than internal chaotic fluctuations. Over those regions the ensemble mean is potentially predictable. Extratropical potentially predictable regions are found mainly over North America and part of the Asian monsoon regions. Interestingly, the potential predictability over some preferred areas such as Indian monsoon areas and central Africa occasionally results primarily from non-ENSO-related boundary forcing, although ENSO forcing generally dominates over most of the preferred areas.

The quantitative analysis of the extratropical potential predictability with PPI has shown that the preferred geographic areas have obvious seasonality. For the 850-hPa temperature, for example, potentially predictable regions during spring and winter are confined to Alaska, northwest Canada, and the southeast United States, the traditional PNA region, while during summer and fall they are favored over the middle part of North America. It has also been shown that the boreal summer season (June–August) possesses the largest potentially predictable area, which seems to indicate that it is a favored season for the extratropical potential predictability. On the contrary, boreal winter (December–February) appears to have a minimum area of extratropical potential predictability.

The results have been compared with the more traditional statistical tests for potential predictability and with observations from the National Centers for Environmental Prediction reanalysis, which indicates that the PPI analysis proposed here is successful in revealing extratropical potential predictability determined by the external forcing.

1. Introduction

It is generally believed that information from the atmospheric initial state is more important than external forcing for predicting the evolution of anomalous atmospheric circulation features on timescales up to about 1 month. As the length of prediction extends beyond a month or season, the external forcing, especially the slowly varying anomalous lower boundary forcing, has much more impact on atmospheric development than the initial conditions. Although snow cover and soil moisture may also play a significant role in the boundary forcing of the atmosphere (Shukla 1984), SST anomalies are commonly assumed to be the major component of the lower boundary forcing for the atmosphere. A variety of previous studies have been devoted to understanding the response of the atmosphere to SST anomalies either in the Tropics or in the midlatitudes (e.g., Lau 1985; Latif et al. 1990; Gates 1992; Lau and Nath 1990, 1994; Smith 1995; Harzallah and Sadourny 1995). Atmospheric GCM simulations of the seasonal

*Permanent affiliation: Department of Atmospheric Sciences, Nanjing University, Nanjing, China.

Corresponding author address: Dr. Xiu-Qun Yang, Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Forrestal Campus, US Route 1, P.O. Box 308, Princeton, NJ 08542.
E-mail: xqy@gfdl.gov

response to the ENSO component of SST anomalies generally exhibit global circulation patterns resembling those observed; this can be interpreted as evidence for the feasibility of seasonal prediction (Palmer and Anderson 1994).

The tropical atmosphere is primarily driven thermally (Matsuno 1966; Gill 1980) and has variability on interannual timescales that is dominated by the lower boundary anomalies, especially the dominant ENSO-related SST anomalies (Charney and Shukla 1981). Unlike the Tropics, much of the extratropical atmospheric variation is dominated by ''chaotic'' internal variability, which is presumably not predictable on seasonal timescales, although SST anomalies may have significant impact on the extratropical atmospheric variations via Rossby wave propagation or other mechanisms. It is of interest to isolate those regions and seasons for which the effects of the external forcing are able to dominate the effects of internal variability in the extratropical atmosphere. Isolating and understanding the dynamics through which tropical (or extratropical) SSTs are able to create some predictable response in the extratropics is also of interest. Hypothetically, one could hope to isolate the externally forced response even in regions where that response is insignificant compared to internally generated atmospheric noise.

Strictly speaking, an assessment of the predictability of seasonal atmospheric variations should be based on a coupled ocean–atmosphere system. As a preliminary step, it is important to understand the predictability of the atmosphere forced by prescribed SSTs. A number of studies have investigated this issue using atmospheric GCM (AGCM) integrations (e.g., Chervin 1986; Brankovic et al. 1994; Ebisuzaki 1995; Stern and Miyakoda 1995; Kumar et al. 1996; Deque 1997), and many approaches have been proposed to distinguish the externally forced variations from the internal chaotic variations. Following the work of Madden (1976, 1981) and Shea and Madden (1990), many earlier studies employed variance analysis methods in which the AGCM with anomalous boundary forcing is compared with the hypothetical situation in which the AGCM is forced by climatological mean SSTs; significant differences between the variances of these two cases are interpreted as indicating potential predictability (e.g., Chervin 1986).

More recently, ensemble integrations generated from multiple GCM simulations have been conducted by a number of groups, many in association with the Atmospheric Model Intercomparison Project (AMIP). The potential predictability of seasonal atmospheric variations has been reexamined based on these ensemble integrations. Frequently, a ratio of internal variance to external variance (noise-to-signal) or a related quantity is used to measure the impact of external forcing. For example, the reproducibility, a description of the spread among ensemble members, is defined to measure potential predictability by Stern and Miyakoda (1995). They found a high degree of reproducibility over most tropical areas but found that most areas in the extratropics were not reproducible. Using the same ensemble integrations, Anderson and Stern (1996) presented a different method, using the Kolmogorov–Smirnov and Kuiper's statistical test for comparing two discrete distributions, and found some extratropical areas where the ensemble forecast distribution is significantly different from an appropriate climatological distribution. Both of these measures were defined so that the measured potential predictability was a function not only of geographical location but also of time.

There are a variety of ways to use ensemble AGCM integrations (Murphy 1989, 1990; Mureau et al. 1993; Ferranti et al. 1994; Anderson and Stern 1996; Anderson 1996), but the most common is to use the ensemble mean as a substitute for a single traditional forecast (Brankovic et al. 1990; Milton 1990; Tracton and Kalnay 1993). It is important to determine if the ensemble mean is predictable even in the presence of large spread among the ensemble members. In a perfect model sense, a small spread among ensemble members does imply good predictability. However, a large spread or a low degree of reproducibility as defined in Stern and Miyakoda(1995) may not imply lack of predictability. It is possible that over some geographic regions in the extratropics the unpredictable internal noise has little influence on the ensemble mean, even though the noise may be large. If the ensemble mean is controlled by externally forced modes, potential predictability of the mean can exist even in the presence of large amounts of noise. One way to assess this in the context of reproducibility would be to apply a standard t test to determine if the ensemble means for different years were significantly different from the model climatological distribution. Results from such an approach are generally quite similar to the Kuiper's results of Anderson and Stern (1996).

Unlike previous studies based on statistical tests or direct measurement of noise-to-signal ratio, the present study focuses on the ensemble behavior of a class of orthogonal modes closely related to the external forcing anomalies. Although derived statistically, there is some hope that these modes can provide insight into the dynamics responsible for extratropical potential predictability in AGCM ensemble integrations. The time-mean potential predictability of the extratropical atmospheric seasonal variations is investigated by examining the reproducibility of these modes. In this paper a particular geographic location is said to have potential predictability when the model ensemble mean at the location is dominated by a highly reproducible component, which comes from the reproducible forced modes. A potential predictability index, which combines the local reproducibility and local variance contribution is defined to quantitatively describe the potential predictability. The major results from this study are compared with those from both the reproducibility analysis of

Stern and Miyakoda (1995) and the potential predictive utility analysis with the Kolmogorov–Smirnov or Kuiper's statistic test of Anderson and Stern (1996). The results are also validated with observations from the National Centers for Environmental Prediction (NCEP) reanalysis.

The model and ensemble integrations used in this study are described in section 2. In section 3, an EOF decomposition of the ensemble mean is used to detect the existence of reproducible forced modes. Based on these forced modes, the ensemble is filtered to exclude both that part of the chaotic component that is independent of the ensemble mean and modes that are predominantly influenced by internal processes. The geographic distribution of reproducibility based on the filtered ensemble is presented. In section 4, the potential predictability index (PPI) is proposed to quantitatively describe the extratropical potential predictability, and the preferred geographical areas and seasons for the extratropical seasonal potential predictability are examined. The extratropical potential predictability inferred from the PPI analysis approach is compared with that obtained from other statistical tests and with observations in section 5.

2. Model and ensemble integrations

The AGCM used to generate ensemble integrations has a spectral triangular truncation at wavenumber 42 (T42) and 18 vertical sigma levels (L18) (Gordon and Stern 1974, 1982). The physical parameterizations are documented in Stern and Miyakoda (1995). Eight multiyear integrations forced by observed AMIP SSTs (Gates 1992) were generated with the AGCM. Each ensemble member was integrated for 10 yr from 1 January 1979 through 31 December 1988. Initial conditions for the ensemble were taken from analyses for 17 December 1978 through 21 January 1979, spaced 5 days apart. Each of these analyses was then used as an initial condition as if it were the analysis for 1 January 1979. Stern and Miyakoda (1995) provided a detailed analysis of the reproducibility of these ensemble integrations and evaluated the model's simulation of the atmosphere's climatology. Anderson and Stern (1996) applied a variety of statistical tests to investigate the potential predictive utility of these ensembles. The present study applies a different method to identify patterns of significant time-mean extratropical potential predictability.

Of interest here is the degree to which the specified, observed SST forcing can influence the seasonal mean state of the AGCM, particularly in the extratropics. The first year of the ensemble integration was discarded in order to eliminate most direct effects of the initial conditions. The analysis in the following is based on monthly mean output of the eight-member ensemble integration for 9 yr from January 1980 through December 1988. Results are presented for seasonal means; the four boreal seasons are referred to as MAM (March–May), JJA (June–August), SON (September–November), and DJF (December–February). For any given season, say MAM, there are a total of 27 individual months from each ensemble integration. Two representative variables, the 850-hPa temperature anomaly (T850), and 300-hPa geopotential height anomaly (Z300), are selected to document the AGCM's seasonal variations in the lower and upper troposphere. The following sections describe the seasonal variations by finding reproducible forced modes and evaluating their contributions to the extratropical potential predictability.

3. Reproducible forced modes in ensemble

a. Definition of reproducibility

Assume that X_{ij} is a monthly average quantity of the ith ensemble member of an ensemble with size M ($M = 8$ here) for the jth month among N months ($N = 27$ months here as noted above) of a particular season. It is important to keep in mind that the definitions below are all for a single season, say MAM. The ensemble monthly mean is defined as

$$\langle X\rangle_j = \frac{1}{M}\sum_{i=1}^{M} X_{ij}, \tag{3.1}$$

which is a function of time, and the ensemble seasonal mean is

$$\overline{X} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} X_{ij}. \tag{3.2}$$

Three variances, σ_{I}, σ_{E}, and σ_{S} can be defined:

$$\sigma_{\mathrm{I}}^2 = \frac{1}{MN}\sum_{j=1}^{N}\sum_{i=1}^{M} (X_{ij} - \langle X\rangle_j)^2, \tag{3.3}$$

$$\sigma_{\mathrm{E}}^2 = \frac{1}{N}\sum_{j=1}^{N} (\langle X\rangle_j - \overline{X})^2, \tag{3.4}$$

$$\sigma_{\mathrm{S}}^2 = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} (X_{ij} - \overline{X})^2, \tag{3.5}$$

where σ_{I}, a measure of the ensemble spread, reflects the internal ''chaotic'' variability of the atmosphere, whereas σ_{E} measures variability that may be induced by the external forcing, sometimes referred to as the external variability. These variances satisfy the relationship $\sigma_{\mathrm{S}}^2 = \sigma_{\mathrm{I}}^2 + \sigma_{\mathrm{E}}^2$. As revealed in later sections of this study, σ_{E} almost always includes some amount of internal variability that is not directly attributable to externally imposed forcing.

Traditionally, the potential predictability of some field in an ensemble of AGCM integrations is assessed using the ratio of external to internal variability, $\sigma_{\mathrm{E}} : \sigma_{\mathrm{I}}$, or some closely related quantity. If the internal variability does not dominate the external variability, one can say that the seasonal atmospheric variation is potentially predictable.

Stern and Miyakoda (1995) assessed the feasibility

of seasonal forecasts inferred from an ensemble of multiple AGCM simulations by defining the reproducibility for an ensemble. In the present study, a *time-mean reproducibility* is defined as

$$R = 1 - \frac{\sigma_I}{\sigma_S}, \tag{3.6}$$

which will be referred to as simply the reproducibility in the following. When the internal chaotic variability (σ_I) is very small, R will be close to 1 and the ensemble is highly reproducible, indicating that the seasonal variation is largely controlled by external forcing. This definition of reproducibility is somewhat different from that used by Stern and Miyakoda, who initially defined their reproducibility as $\sigma_I(j)/\sigma_S$ with the ensemble variance at a time, $\sigma_I(j)$, and therefore R defined as a function of time. To measure cumulative reproducibility over the ensemble integration period, Stern and Miyakoda defined a time-averaged reproducibility, which is qualitatively consistent with $1 - R$ as defined in (3.6); note that this means that high reproducibility in Stern and Miyakoda was associated with small values of their metric.

Reproducibility can be used to determine if the distribution of individual predictions within the ensemble shows some systematic reproducible behavior or whether it is unpredictable. For example, Fig. 1 shows the ensemble integrations for T850 at selected grid points in the Tropics (0°, 125°W) and the extratropics (60°N, 125°W) during MAM. It is evident that in the Tropics the ensemble members are close together, whereas those for the extratropics have relatively large scatter. Figure 7a shows the geographic distribution of reproducibility as defined in (3.6) and confirms that large values of R are primarily confined to the Tropics consistent with the results of Stern and Miyakoda (1995).

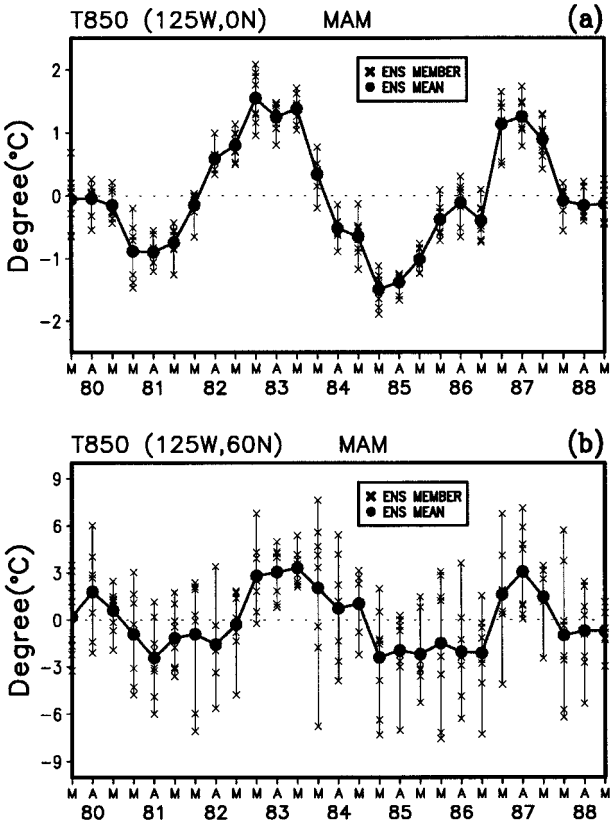

FIG. 1. Anomalies of ensemble integrations from ensemble-mean seasonal climate for 850-hPa temperature during MAM at a selected grid point in the Tropics (0°, 125°W) (upper panel) and the extratropics (60°N, 125°W) (lower panel). Circles denote the ensemble-mean anomaly, $\langle X\rangle_j - \overline{X}$, and crosses denote individual ensemble member anomalies, $X_{ij} - \overline{X}$.

b. Reproducible forced modes

The features appearing in Figs. 1 and later in Fig. 7a indicate that the tropical seasonal atmospheric variation is dominated by the external forcing, whereas in the extratropics the internal variation is so large that the influence of the external forcing is difficult to detect. This section presents a method for extracting that part of the variability in the AGCM that is highly reproducible, even in extratropical regions that are dominated by internal variability.

Monthly mean anomalies of the *i*th ensemble member for the *j*th month may be considered to be composed of three parts:

$$X_{ij} - \overline{X} = Y^I_{ij} + Y^{IE}_{ij} + Y^E_{ij}, \tag{3.7}$$

where Y^I_{ij} is that part of the internal variation that is not reproducible and has no contribution to the ensemble mean ($\langle Y^I\rangle_j = 0$), Y^E_{ij} represents the highly reproducible part, which is a direct response to the external forcing, and Y^{IE}_{ij} is the part with low reproducibility that also contributes to the ensemble mean. Here Y^{IE}_{ij} may be induced by interaction between external forcing and internal processes but is dominated by the effects of internal variability. The ensemble-averaged anomaly results only from the latter two parts; that is,

$$\langle X\rangle_j - \overline{X} = \langle Y^{IE}\rangle_j + \langle Y^E\rangle_j. \tag{3.8}$$

The second term on the right side of (3.8) (the reproducible part) is the key to potential predictability. If this component dominates the complete ensemble mean, potential predictability will exist. However, in the presence of ''noise'' from the other two components, Y^I_{ij} and Y^{IE}_{ij}, it becomes problematic to extract this reproducible component.

For a particular geographic location, it is difficult to detect the reproducible component locally. In this study, reproducible global modes are derived from the ensemble and the local reproducible component is then computed by assessing the local contribution from the global modes. In principle, the reproducible modes are fundamentally forced modes resulting from some kind of external forcing, and the existence of reproducible

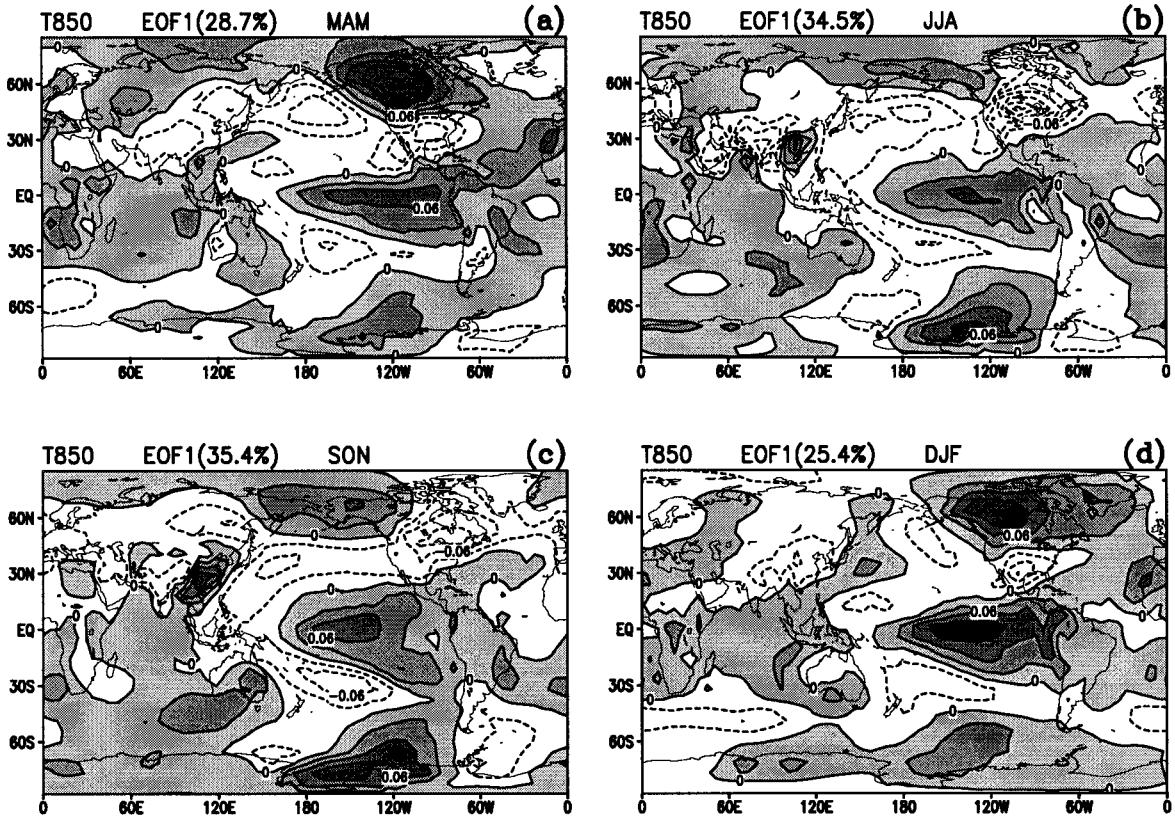

FIG. 2. Spatial patterns of the leading EOF of ensemble-averaged 850-hPa temperature anomalies for (a) MAM, (b) JJA, (c) SON, and (d) DJF.

modes depends on whether the external forcing can win a competition with the internal chaotic processes.

The reproducibility of a given spatial pattern for a given mode can be assessed by examining the time series of the projection of the individual ensemble monthly means onto the spatial pattern. Then the reproducibility of the time series corresponding to that spatial pattern is assessed. High reproducibility indicates that the pattern is less affected by atmospheric internal variability. In this study, the spatial patterns of interest are computed by performing an EOF decomposition of the ensemble-averaged anomalies:

$$\langle X \rangle_j - \overline{X} = \sum_{k=1}^{N} a_{jk} P_k(\mathbf{r}), \tag{3.9}$$

where a_{jk} is the principal component at the jth month corresponding to the kth EOF spatial pattern, $P_k(\mathbf{r})$, where $\mathbf{r}$ is a vector spanning the phase space of the AGCM (a vector composed of the value of a field at each model grid point for our purposes).

Figure 2 shows the spatial pattern of the leading EOF of the ensemble-averaged T850 anomaly for the four different seasons. The spatial structure in the Tropics is characterized by the classical ''ENSO pattern'' with large amplitude in the eastern equatorial Pacific and little seasonal dependence. The spatial patterns in the extratropics, especially over North America and Asia, exhibit obvious seasonality. In the synopsis that follows, the signs of the anomalies are those that occur in conjunction with a positive anomaly in the east tropical Pacific. Over North America in winter (Fig. 2d) a large positive center in north Canada accompanies a negative center in the southeast United States, whereas in summer (Fig. 2b) a negative center occupies the central part of North America with weak positive areas shifted to Alaska through eastern Siberia. Over Asia, seasonality is particularly evident in southeast Asia, with a positive center in summer (Fig. 2b) but negative center in winter (Fig. 2d). The spatial patterns for MAM and SON (Figs. 2a and 2b) seem to be intermediate to those for DJF and JJA.

Similar characteristics can be observed in the spatial patterns of the leading EOF for the ensemble-averaged Z300 anomaly field (not shown). The spatial pattern over the eastern equatorial Pacific appears to be roughly symmetric about the equator and is seen throughout the year; this might be attributed to the Rossby wave response to the ENSO-related SST forcing. In the extratropics of the Northern Hemisphere, the Pacific–North American (PNA) pattern is dominant in winter as well as in spring, whereas a zonally elongated pattern spans the central part of North America in summer and in fall.

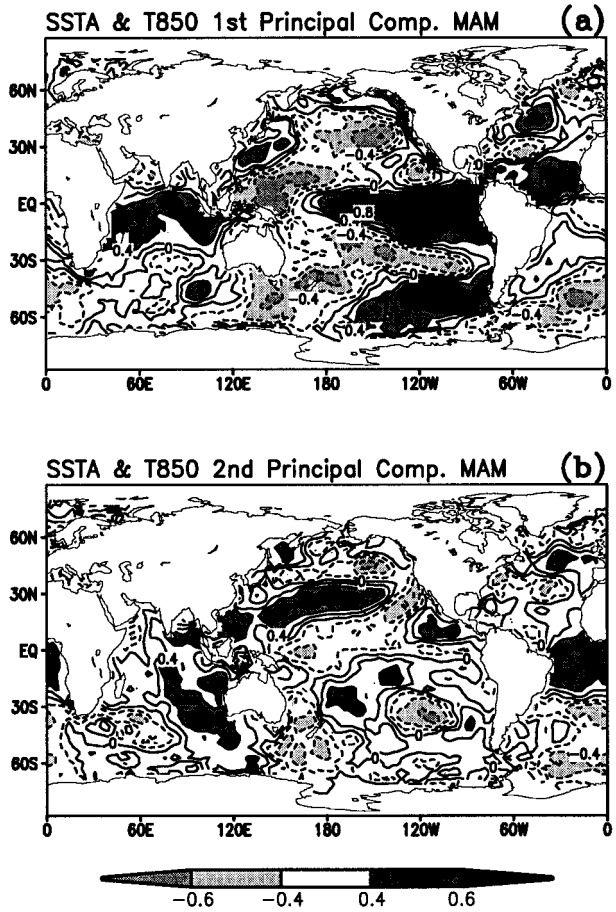

FIG. 3. Distribution of simultaneous correlation between global SST anomaly and (a) the first principal component and (b) the second principal component of ensemble averaged 850-hPa temperature anomaly during MAM. Contour interval is 0.2. Regions of correlations greater than 0.4 are heavily shaded and regions less than −0.4 are lightly shaded.

However, only in JJA and SON is a wave train pattern seen in the Southern Hemisphere.

Overall it is evident that the first EOF of the ensemble mean is fundamentally dominated by the ENSO-like pattern as revealed in many previous studies (e.g., Wallace and Gutzler 1981). This can be further confirmed by Fig. 3a, which shows the spatial distribution of the temporal correlation of the first principal component of the ensemble-averaged T850 with global SST anomalies, again revealing a typical El Niño pattern.

The relation of the other EOFs to the external forcing appears to be somewhat complicated. This is partly because there is no globally dominant external forcing signal except for ENSO, and partly because more internal processes are involved in the higher EOFs. However, it is still possible that some of the higher-order EOFs reflect the effect of non-ENSO external forcing. For example, Fig. 3b shows the temporal correlation of the second principal component of ensemble-averaged T850 with global SST anomalies during MAM. Larger values of the correlation are found in midlatitudes (around 30°N) of the North Pacific, rather than in the equatorial Pacific as in Fig. 3a.

Projecting the original ensemble integrations onto the kth ensemble-average EOF pattern can reveal to what extent the EOF is affected by internal processes versus external forcing. In terms of the spatial patterns defined in (3.9), one can define an ensemble of *projected principal components*, b_{ijk}, for the kth ensemble-average EOF by

$$b_{ijk} = \sum_{\mathbf{r}} (X_{ij} - \overline{X}) P_k(\mathbf{r}). \tag{3.10}$$

The projected principal components (b_{ijk}, $k = 1, \ldots, N$) do not form a complete set for the ith ensemble member. However, the ensemble mean of all of the projected principal components does form a complete set for the ensemble mean (hereafter the usage of complete set is in the sense of ensemble mean), since the ensemble average of b_{ijk} is the same as the principal component of the original ensemble mean—that is, $\langle b \rangle_{jk} = a_{jk}$. The projected components indicate the contribution that each ensemble member can make to the ensemble mean.

Figure 4 illustrates the projected principal component ensemble for the first ensemble-average EOF for T850 for the four seasons. Consistent with the spatial pattern in Fig. 2, the first principal component time series of the ensemble mean exhibits the ENSO signals with peaks in 1982–83 and 1986–87 for all seasons. The individual projected principal components display very little scatter demonstrating that this first EOF is highly reproducible with values of reproducibility exceeding 0.59 for all seasons [here the reproducibility for the projected principal components was computed with a definition analogous to (3.6)]. Note that the reproducibility found for an eight-member ensemble of red noise having the same 1-month autocorrelation is about 0.1. This reveals that the first mode is predominantly a response to external forcing and is mostly unaffected by the presence of internal atmospheric noise.

Generally, the value of reproducibility of the projected principal component ensembles decreases with increasing EOF order for all the seasons, as shown in Fig. 5. However, the distribution of reproducibility as a function of EOF order has obvious seasonal dependence. The largest number of relatively reproducible EOFs was found in summer, with the fewest found in winter. This may simply indicate that the internal variability is strongest in winter but weakest in summer.

It is of particular interest that, besides the first EOF, relatively high reproducibility is also found for some higher-order EOFs, which are generated by non-ENSO-related boundary forcing such as midlatitude SST anomalies (shown in Fig. 3b). As discussed in the next section, those relatively reproducible modes controlled by non-ENSO-related boundary forcing are as important in contributing to the ensemble mean over some local regions as are the ENSO-related modes.

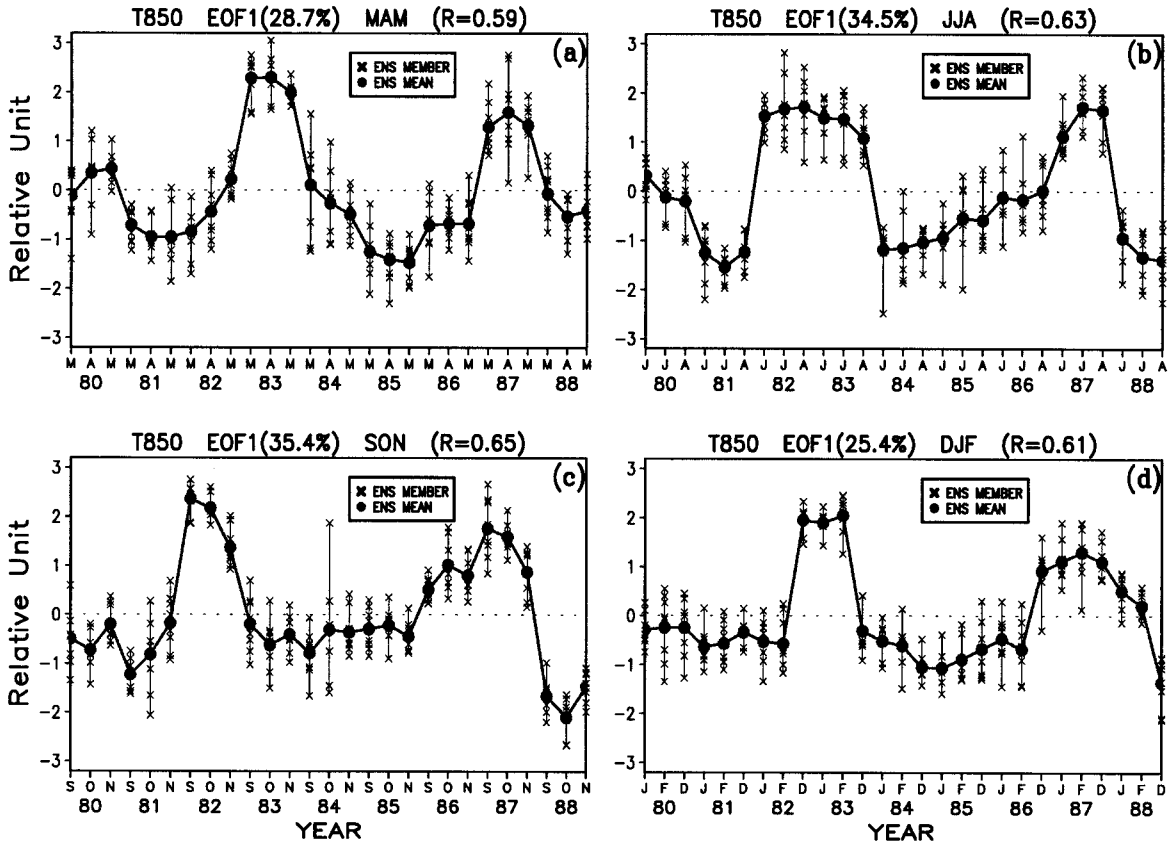

FIG. 4. Projection of ensemble members on the first principal component of the ensemble mean, b_{ij1}, for 850-hPa temperature anomalies for (a) MAM, (b) JJA, (c) SON, and (d) DJF. Circles denote the ensemble mean of the projections and crosses mark the projections of the individual members.

c. *Reconstruction of ensemble*

The results above indicate that the ENSO-forced EOF is highly reproducible in all seasons and that some additional predominantly forced EOFs also exist. Since reproducibility is found to decrease with increasing EOF order, one can easily reconstruct the ensemble based only on those modes that are relatively reproducible. This reconstructed ensemble should be useful for isolating the locally reproducible components of the ensemble and for further understanding the sources of potential predictability in the extratropics.

Using the projected principal components defined in (3.10), the reconstructed anomaly of a monthly mean quantity for the *i*th ensemble member and the *j*th month can be written in terms of the first *K* reproducible modes as

$$Z_{ij}^{K} = \sum_{k=1}^{K} b_{ijk} P_k(\mathbf{r}). \tag{3.11}$$

As discussed in section 3b, two types of atmospheric internal variability can exist. One appears to be purely noise and has nothing to do with the ensemble mean. The other is a result of internal processes interactively mixed with effects of external forcing. Reconstructing an ensemble in terms of the EOFs of the ensemble mean has two advantages. First, the reconstruction can eliminate chaotic variation, which is independent of the original ensemble mean. If the reconstruction is based on a complete set of *N* EOFs (i.e., $K = N$), the reconstructed ensemble-averaged anomaly is exactly the same as that for the original ensemble mean—that is, $\langle Z^N \rangle_j = \langle X \rangle_j - \overline{X}$. However, the reconstructed ensemble mean does not include the noise:

$$Y_{ij}^{\mathrm{I}} = X_{ij} - \overline{X} - Z_{ij}^{N}, \tag{3.12}$$

from (3.7) that is independent of the ensemble mean. Second, by excluding higher-order EOFs that are primarily a result of internal chaotic processes—that is, eliminating the unreproducible component Y_{ij}^{IE} in (3.7), one can isolate the local reproducible component, the term Y_{ij}^{E} in (3.7).

Suppose that the first *L* EOFs are relatively reproducible. Using (3.11) together with (3.7) and (3.12), the two components, which could both be significant to the ensemble mean, can be calculated: the reproducible component,

$$Y_{ij}^{\mathrm{E}} = Z_{ij}^{L}, \tag{3.13}$$

and the unreproducible component,

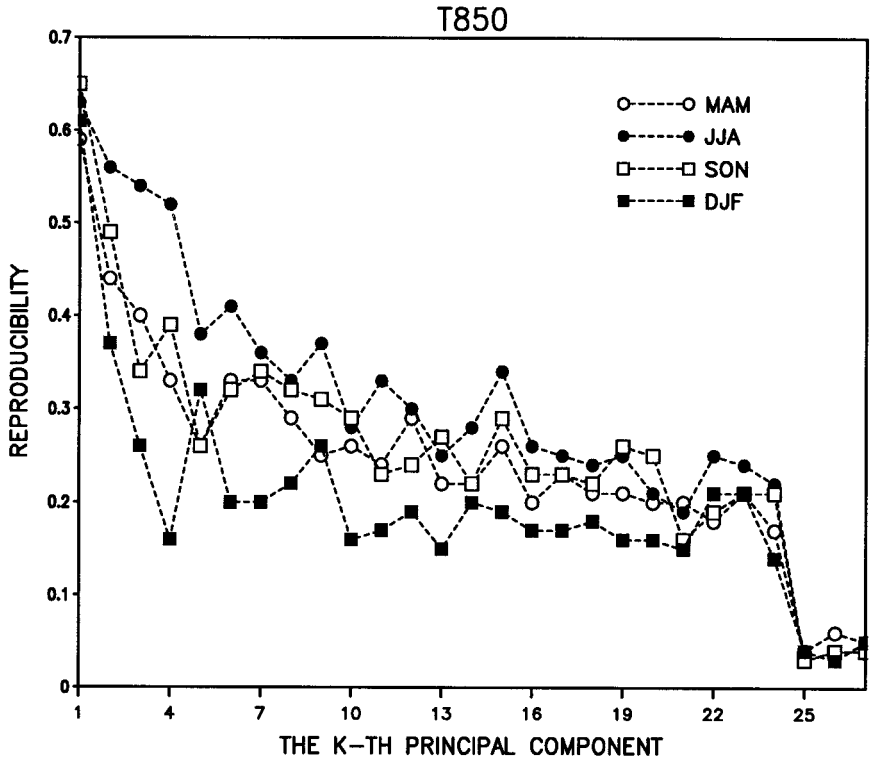

FIG. 5. The reproducibility of the projected principal components, plotted as a function of EOF order.

$$Y_{ij}^{\mathrm{IE}} = Z_{ij}^{N} - Z_{ij}^{L}. \tag{3.14}$$

Retaining the first L EOFs is not optimal, but the results are not particularly sensitive to this truncation (see, e.g., Fig. 10). The determination of the number of retained EOFs, L, is an important issue to be discussed in the next section.

Both advantages discussed above can be seen in Fig. 6, which presents a comparison of the original and reconstructed ensembles of T850 anomalies at a selected grid point (60°N, 125°W) during MAM. A notable reduction of noise in the reconstructed ensemble in terms of a complete set of 27 EOF modes (Z_{ij}^{N}) is observed in Fig. 6b relative to the original ensemble (Fig. 6a). Note that the ensemble means of the two panels are the same. The bottom panel shows the reconstructed ensemble in terms of only the first EOF mode, $Y_{ij}^{\mathrm{E}} = Z_{ij}^{1}$. The highly reproducible time series in Fig. 6c, largely a response to ENSO, dominates the original ensemble mean (Fig. 6a). At least at this grid point, the highly reproducible externally forced component contributes significantly to the local ensemble mean.

d. Geographic distribution of reproducibility

One goal of the method developed here is to find the preferred geographic regions with potential predictability. The reconstruction of the ensemble in terms of reproducible modes can isolate part of the externally forced signal. However, the selection of the number of modes to be retained in the reconstruction seems somewhat arbitrary. The transition from reproducible modes to unreproducible higher-order modes appears to be gradual. In addition, retaining too many modes leads to a reconstruction that contains noise, which obscures the reproducible signal, whereas retaining too few modes can eliminate useful reproducible information.

As an example, the reproducibility of the reconstructed ensemble in terms of all of the EOFs (i.e., Z_{ij}^{N}), denoted as $R^{N} = 1 - \sigma_{\mathrm{I}}^{N}/\sigma_{\mathrm{S}}^{N}$, is examined first. Such a reconstruction reduces noise without altering the ensemble mean. The geographic distribution of high reproducibility based on the untruncated reconstructed ensemble indicates where the original ensemble mean has potential predictability. Figure 7b presents the distribution for MAM. Comparing with the raw reproducibility shown in Fig. 7a, increased areas with high reproducibility can be observed in the extratropics, especially over North America. In areas with high reproducibility in Fig. 7, the reproducible forced modes dominate the ensemble mean implying that the ensemble mean is locally potentially predictable.

It is encouraging to find some extratropical regions where the ensemble mean has significant reproducible components. However, applying a more truncated EOF filter may be able to extract additional signal. Figure 8 shows the percentage of grid points with reproducibility

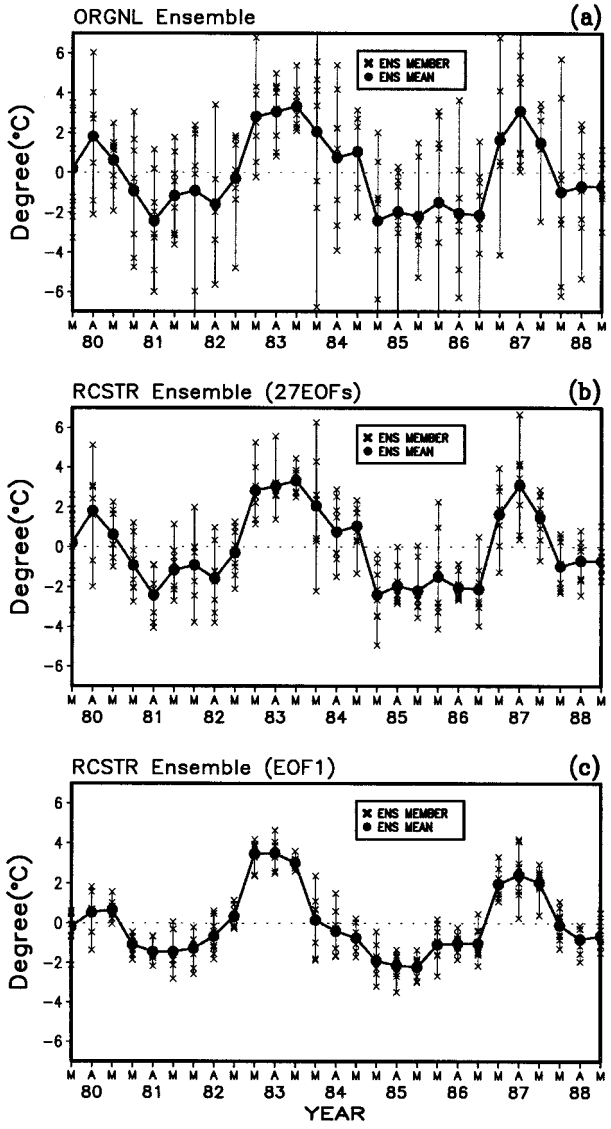

FIG. 6. Time series of (a) original ensemble anomaly, $X_{ij} - \overline{X}$; (b) reconstructed ensemble with a complete set of EOF modes (27 modes), Z_{ij}^{27}; and (c) reconstructed ensemble in terms of the leading EOF mode, Z_{ij}^{1}, for 850-hPa temperature during MAM at (60°N, 125°W). Circles denote ensemble mean and crosses denote individual ensemble members.

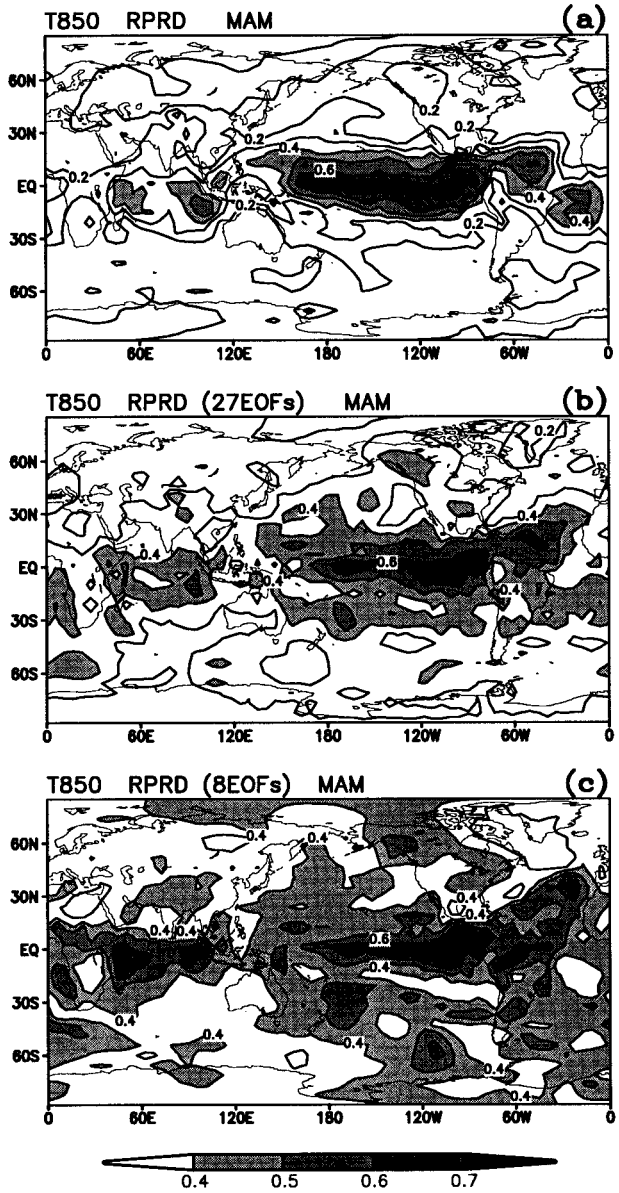

FIG. 7. Global distribution of reproducibility for 850-hPa temperature during MAM (a) for the original ensemble, (b) for the ensemble reconstructed in terms of a complete set of 27 EOF modes (b), and (c) for the ensemble reconstructed in terms of the first eight EOF modes. Contour interval is 0.1 with values larger than 0.4 shaded.

greater than 0.4 as a function of the number of EOFs retained in the reconstructed ensemble. Generally, this percentage decreases as the number of retained EOFs is increased. Figure 7c presents the geographic distribution of reproducibility for T850 during MAM, calculated from the reconstructed ensemble in terms of a truncated sets EOFs. The number of retained modes was empirically selected in an attempt to retain as many relatively potentially predictable features of the complete ensemble mean as possible. The result is that much larger areas with significant reproducibility can be found in the extratropics than for the complete reconstruction shown in Fig. 7b. Of particular interest is the relatively high degree of reproducibility was found over many areas of North America and even Asia. These reproducible regions in the truncated reconstructed ensemble indicate that the lower boundary forcing is able to generate a forced component that contributes to the ensemble mean in many areas of the extratropics.

The reproducibility of the reconstructed ensemble has considerable seasonal dependence, as shown in Fig. 8. In general, summer (JJA) is a preferred season, which possesses the largest area with high reproducibility,

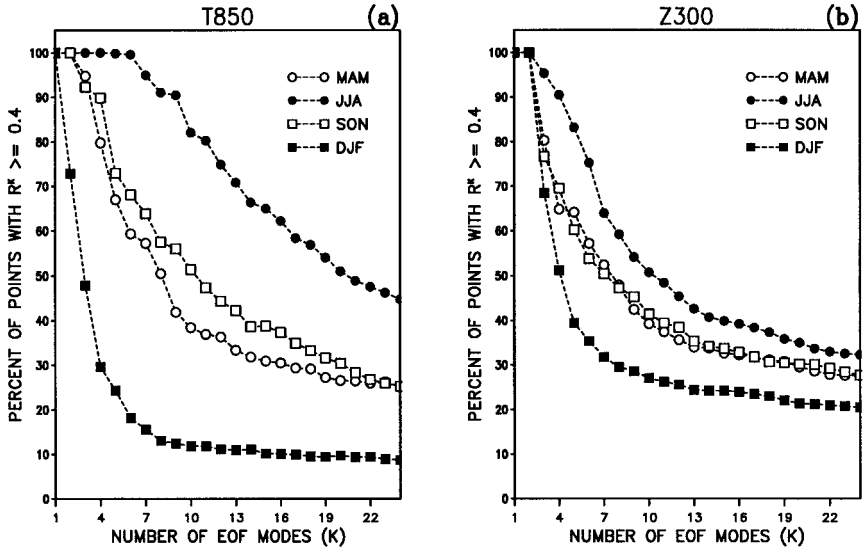

FIG. 8. Percentage of grid points with reproducibility no less than 0.4 as a function of number, K, of EOF modes included in the reconstructed ensemble for 850-hPa temperature anomalies (a) and 300-hPa geopotential height anomalies (b).

whereas winter (DJF) has the smallest area. The transition seasons are in between. The seasonal dependence is dominated by the Northern Hemisphere, which has the largest preferred area. Also, the seasonal difference is larger for T850 than for Z300. One might speculate that winter is the season during which the externally forced signal is most difficult to detect because chaotic processes in the atmosphere dominate the forced components, whereas summer seems to be, at least for T850, the most favorable for the external forcing to generate significant responses.

4. Potential predictability of local ensemble mean

According to the analysis in the last section, two primary prerequisites for the existence of potential predictability of the ensemble mean are (i) the local variance of the ensemble mean is dominated by forced modes, and (ii) these forced modes are highly reproducible. The simultaneous occurrence of these two criteria implies that the ensemble mean at the particular geographic location is potentially predictable in a time-mean sense. This section presents a quantitative measure of potential predictability based on these criteria.

a. Potential predictability index

As seen in section 3, it is problematic to use reproducibility of reconstructed ensembles in terms of the ensemble-mean EOFs to examine the local potential predictability because one does not know how many EOFs to include in the truncated reconstruction [selection of L in (3.13)]. Here, the *potential predictability index* (hereafter referred to as PPI), a combination of the principal modes' reproducibility and their variance contribution to the ensemble mean is proposed as a practical tool for assessing potential predictability in this context.

Assume that a reconstructed ensemble in terms of the first K EOFs is represented by Z_{ij}^K. The reproducibility of the reconstructed ensemble, denoted as $R^K = 1 - \sigma_I^K/\sigma_S^K$, can be calculated according to (3.1)–(3.6). Let $F^K = \sigma_E^K/\sigma_E$ be the ratio of the local variance explained by the truncated reconstruction to the original ensemble mean's variance (note that $F^K = 1$ when $K = N$). The PPI at a particular grid point is defined as

$$\text{PPI} = F^K R^K. \tag{4.1}$$

Since $0 \le R^K \le 1$, and $0 \le F^K \le 1$, the PPI satisfies $0 \le \text{PPI} \le 1$. Large PPI at a geographic location corresponds not only to high reproducibility but also to a high percentage of variance contribution to the ensemble mean, thus indicating potential predictability in the time mean. In addition, as discussed below, PPI analysis can help to determine the appropriate truncation for the reconstructed ensemble in order to extract the reproducible component (Y_{ij}^E).

b. Geographic distribution and seasonality of PPI

One advantage of PPI analysis is that the spatial pattern of potentially predictable areas is relatively insensitive to the EOF truncation as shown in Fig. 9. The pattern of high PPI regions is relatively consistent for truncations ranging from 1 through 20 modes. As noted in the next subsection, this is because the first EOF mode (i.e., ENSO-forced mode) dominates over most areas although other forced modes may come into play over some local regions.

Another advantage of PPI analysis is that one can

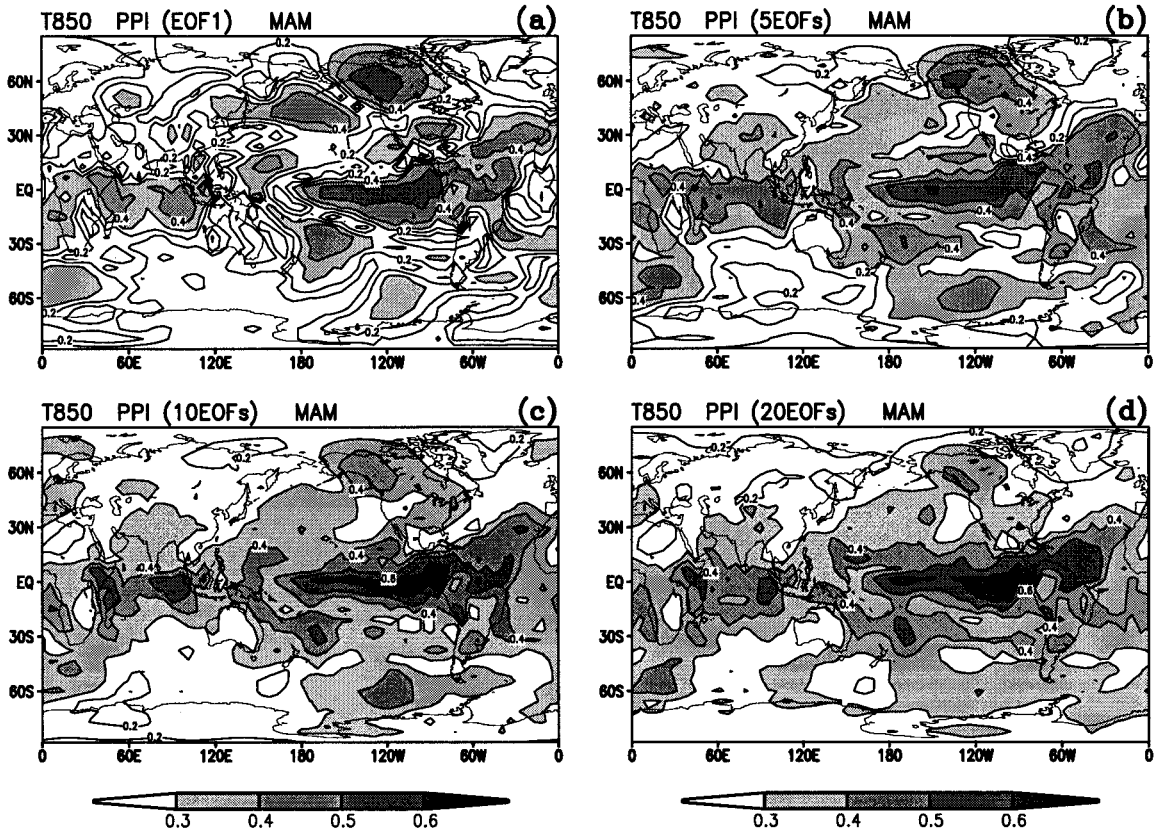

FIG. 9. Global distribution of PPI for 850-hPa temperature during MAM, for different EOF truncations: (a) 1 EOF, (b) 5 EOFs, (c) 10 EOFs, and (d) 20 EOFs.

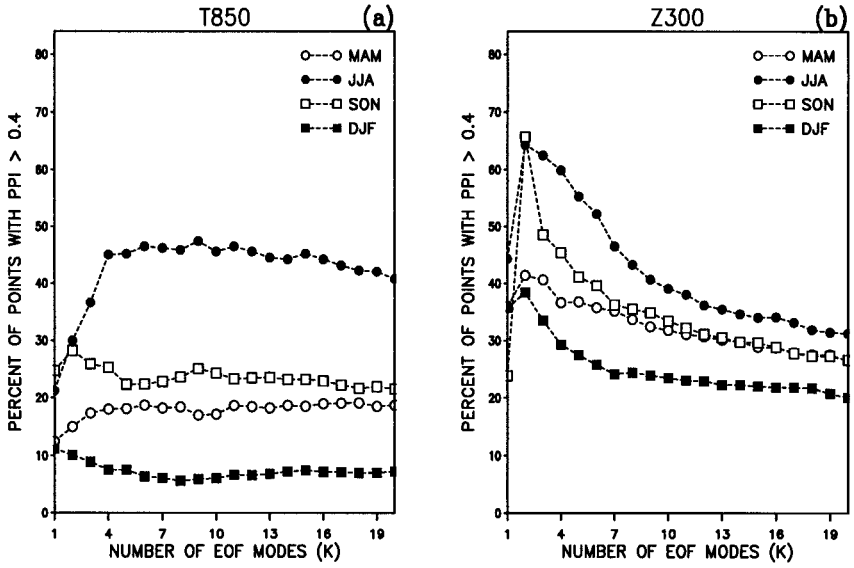

FIG. 10. Percentage of grid points with PPI larger than 0.4 as a function of number, *K*, of EOF modes included in the reconstructed ensemble for 850-hPa temperature anomalies (a) and for 300-hPa geopotential height anomalies (b).

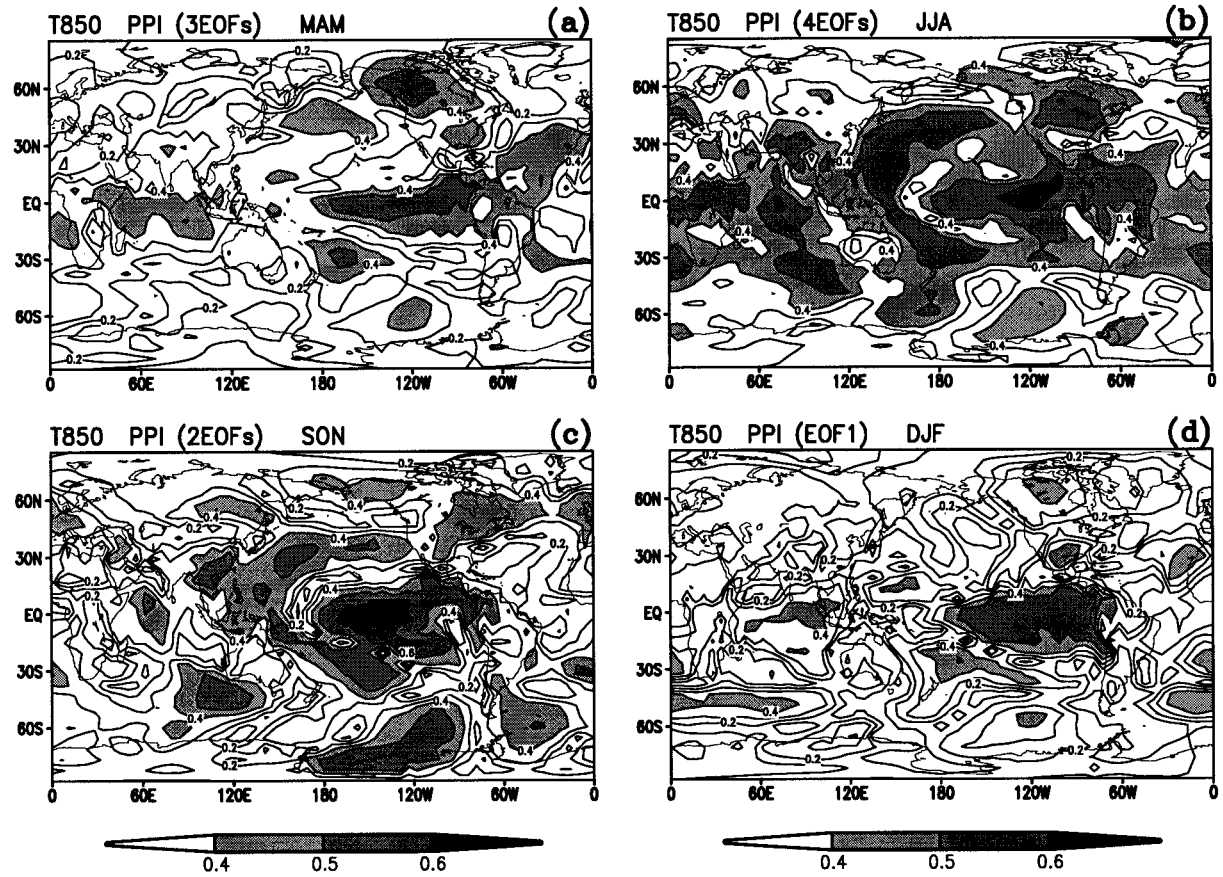

FIG. 11. Global distribution of PPI for EOF truncation, L, corresponding to the peak in Fig. 10 for 850-hPa temperature during (a) MAM ($L = 3$), (b) JJA ($L = 4$), (c) SON ($L = 2$), and (d) DJF ($L = 1$). Contour interval is 0.1 with regions greater than 0.4 shaded.

find an EOF truncation that maximizes the geographic area with a given level of potential predictability. Figure 10 plots the percentage of grid points with PPI larger than 0.4 as a function of the number of EOFs for T850 and Z300. Generally, the area with PPI exceeding a threshold becomes approximately constant when the number of EOFs included in the reconstructed ensemble becomes large. The PPI distribution for a complete set of EOF modes (i.e., $K = N$) has the same geographic area as for the distribution of reproducibility from the untruncated reconstruction as shown in Fig. 7b.

Figure 10 shows that there is a maximum area as a function of the number of retained EOFs for each season. For example, for T850 (Fig. 10a), retaining the first four EOFs ($L = 4$) produces the maximum PPI area during JJA, whereas retaining only a single EOF is optimal for DJF ($L = 1$), three EOFs is optimal for MAM, and two for SON. For Z300 (Fig. 10b), the first two EOFs are optimal for all seasons ($L = 2$). The results here are relatively insensitive to the empirical PPI threshold value; results for other values greater than the 0.4 used here gave essentially identical results.

Figure 11 shows the PPI distribution obtained from the reconstructed ensemble truncated at the peaks from Fig. 10 for T850. The most significant areas with large PPI in the extratropics are over North America. Like the reproducibility examined in section 3 (Fig. 8), the extratropical areas of high PPI have a large seasonal dependence. During MAM and DJF, the preferred areas coincide with the traditional PNA regions with the major center over north Canada and a relatively small center over the southeast United States (Figs. 11a,d). During JJA and SON (Figs. 11b,c), the significant areas are shifted to the middle part of North America. Areas that have potential predictability can be found over Asia for all seasons except DJF. For example, large PPI was found over the Indian monsoon region during JJA (Fig. 11b), as well as over southeast China and northeast Asia during SON (Fig. 11c). A similar PPI distribution can be observed in the Z300 field (not shown).

The PPI distributions in Fig. 11 suggest that the extratropical potential predictability is lowest during winter, especially for T850. It can be seen from Fig. 10a that JJA has the largest degree of potential predictability. For Z300 (Fig. 10b), the seasonality tends to be smaller although the seasonal variation is qualitatively similar.

c. Contribution from ENSO and non-ENSO forcing

The PPI analysis above reveals the preferred geographic areas where the ensemble mean is potentially predictable. These areas are attributed to several leading EOF reproducible modes. Dynamically, the reproducible modes may be generated by different boundary forcing. As mentioned in section 3 (Fig. 3b), non-ENSO boundary forcing can play a role in this process. For the T850 field, the correlation of the principal component time series with SST shows that all modes except the first are not strongly related to the ENSO SST signal. These higher-order reproducible modes may come from non-ENSO boundary forcing such as midlatitude SST anomalies as shown in Fig. 3b.

One can use the PPI distributions from different truncations to evaluate the contributions from ENSO and non-ENSO forcing. For T850 during MAM, Figs. 12a,b show the PPI distribution separately for contributions from EOF1 (ENSO forcing) and EOF2 (leading non-ENSO forcing). In comparison with the maximal area PPI (truncated at three modes) shown in Fig. 12c, ENSO forcing plays a dominant role in generating potential predictability of the ensemble mean over most extratropical areas, especially over North America. Figure 12b shows that the leading non-ENSO forcing, such as SST anomalies in the subtropical western Pacific, makes a significant contribution to potential predictability over the Indian monsoon areas, western Europe, and the eastern United States where no direct ENSO effect was found.

The contribution from non-ENSO forcing is concentrated in central Africa and southern South America during JJA, and southeast Asia and northern Canada during SON (not shown). Very little contribution from non-ENSO forcing was found during DJF.

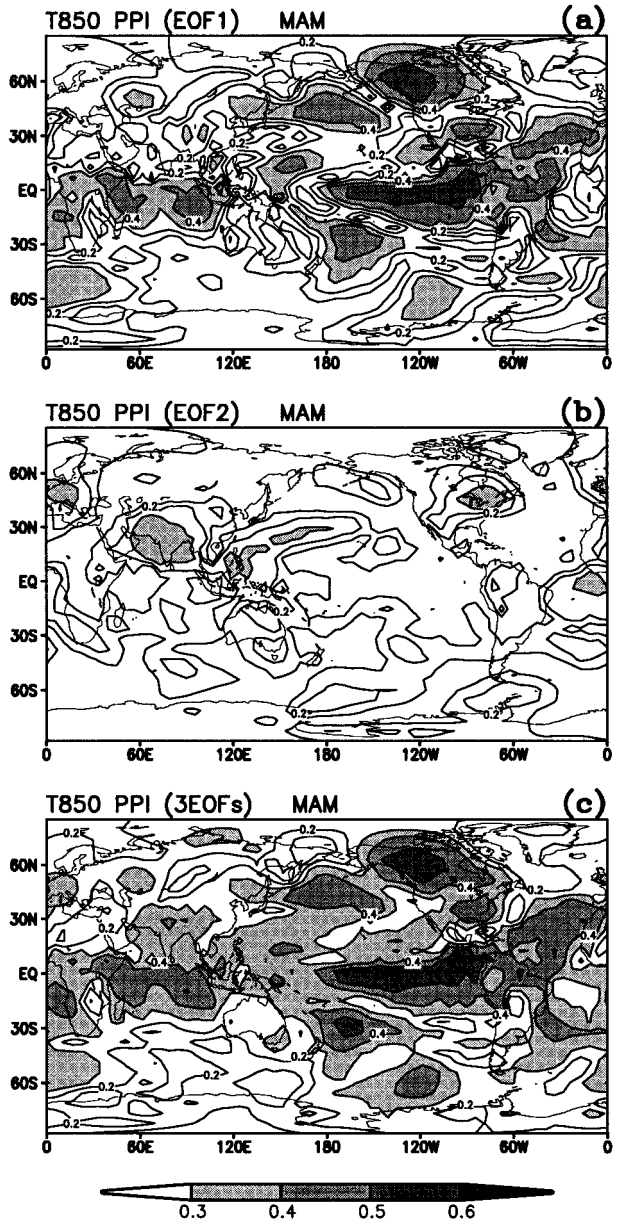

FIG. 12. Global distribution of PPI for 850-hPa temperature during MAM, calculated from the reconstructed ensemble in terms of (a) only the first EOF, (b) only the second EOF, and (c) the first three EOFs. Contour interval is 0.1 with regions of greater than 0.4 shaded.

5. Comparison to potential predictability statistics and observations

By investigating the reproducible forced modes and extracting the local reproducible component (i.e., forced signal), regions in the extratropics where the local ensemble mean is potentially predictable can be identified. It is interesting to explore how the PPI defined in the previous section relates to previous definitions of potential predictability. In addition, the model results can be compared to observations in light of the PPI results. In this section, both of these issues are explored to gain further insight into the relevance of the PPI approach.

a. Compared with statistical tests

The PPI distribution displays spatial patterns with centers where the time-mean ensemble mean is strongly controlled by external forcing. One can compare these time-mean results to measures of the time-dependent potential predictability. A statistical test for potential predictability should measure whether the ensemble mean is significantly different from the model climatology when there exists a notable external forcing anomaly. As an example, Fig. 13a shows the spatial distribution of confidence of Student's *t*-test, which indicates if the ensemble mean in MAM of 1983 is significantly different from the climatological mean. The higher confidence centers are roughly consistent with the larger PPI centers shown in Fig. 11a, suggesting that much of the time-mean predictability is a result of forcing from strong ENSO events.

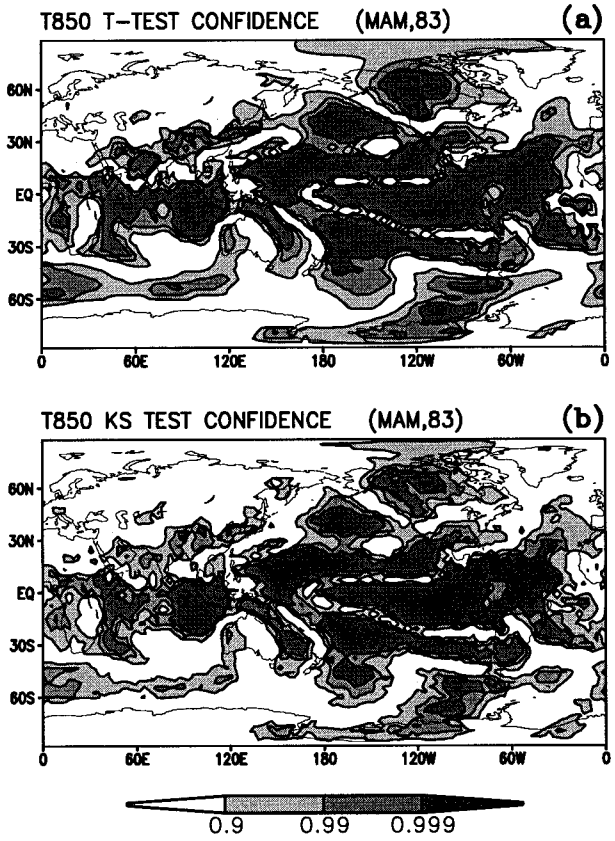

FIG. 13. Global distribution of confidence of Student's *t*-test (a) and KS test significance (b) for 850-hPa temperature during MAM of 1983. Light (dark) shaded areas indicate values greater than 0.9 (0.99).

Additional statistical tests such as the Kolmogorov–Smirnov (KS) test and Kuiper's test (Anderson and Stern 1996) evaluate whether two distributions are different. Such tests can be more powerful tools than the Student's *t*-test or *F* test to evaluate the time-dependent potential predictability. Figure 13b presents the spatial distribution of confidence of a KS test comparing the ensemble distribution for MAM 1983 to the climatological distribution. Over most areas the difference between two distributions (ensemble distribution in a particular time and climatological distribution) is closely related to the ensemble-mean difference as shown in Fig. 13a.

Both statistical tests for time-varying potential predictability exhibit similar patterns (Figs. 13a,b) to those for the time mean from PPI analysis (Fig. 11a). According to PPI analysis, over those preferred extratropical areas, the local ensemble mean is dominated by a distinguishable, systematic externally forced component. This should be the reason why over those areas the local ensemble distribution relative to an anomalous external background is always found to differ from the climatology. Thus, results from the statistical tests can be related to global forced modes according to the PPI analysis.

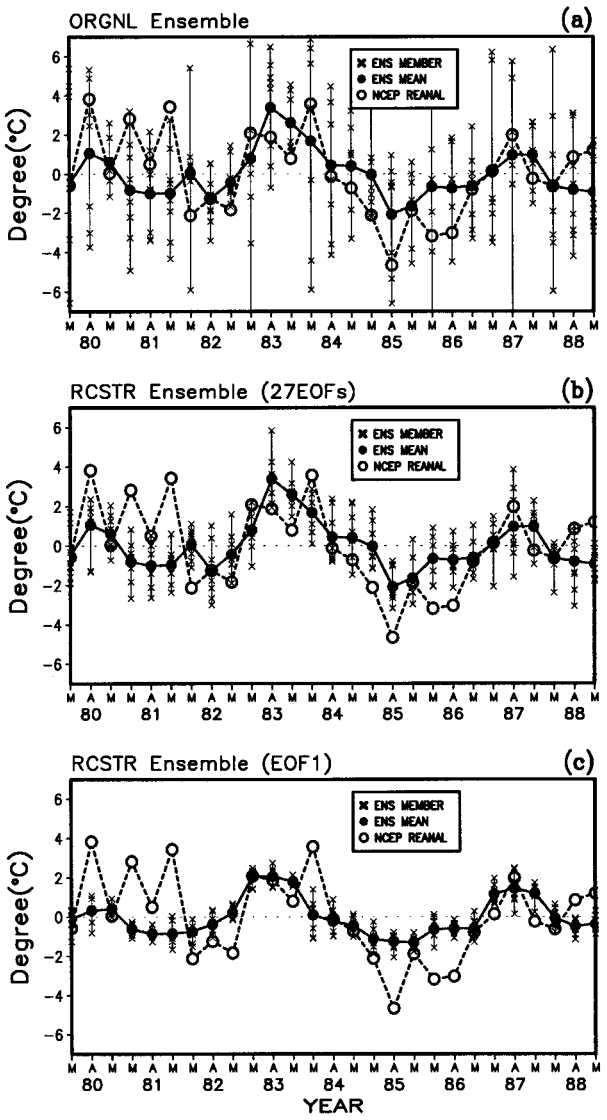

FIG. 14. Time series of (a) original ensemble anomaly, $X_{ij} - \overline{X}$; (b) reconstructed ensemble with a complete set of EOF modes (27 modes), Z_{ij}^{27}; and (c) reconstructed ensemble in terms of the leading EOF mode, Z_{ij}^{1}, for 850-hPa temperature during MAM at (60°N, 150°W) compared with NCEP reanalysis. Closed circles denote ensemble mean, crosses denote individual ensemble members, and open circles mark the NCEP reanalysis.

b. Comparison with NCEP reanalysis field

The PPI analysis has identified extratropical regions with enhanced time-mean potential predictability. Over those regions, the ensemble mean is controlled by external forcing rather than by internal chaotic process. Since the external forcing is specified from the real world, if the model realistically simulates the real world, the AGCM ensemble mean should be similar to the observations in areas with high PPI. In this study, the external forcing includes only the SST so the potential predictability here is referred to as the SST-induced signal.

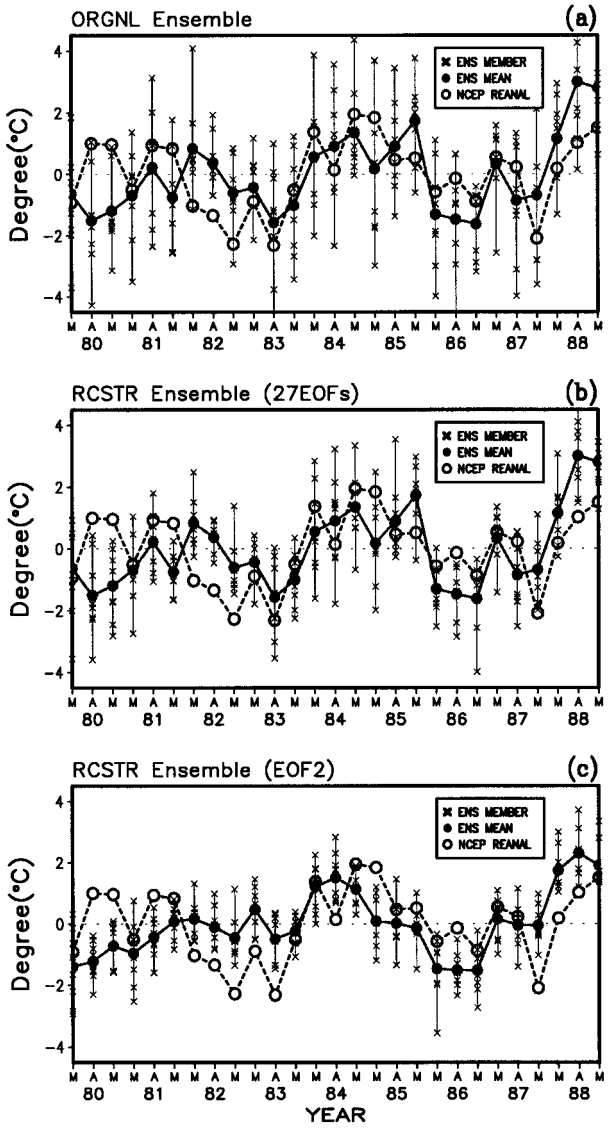

FIG. 15. Same as Fig. 14 except for at (30°N, 76°E).

Figures 14 and 15 display the NCEP reanalysis and the ensemble behavior of the T850 anomalies during MAM for the original ensemble (top panel), reconstructed ensemble in terms of the complete set of EOF modes (middle panel), and reconstructed ensemble in terms of a single EOF mode (bottom panel, EOF1 in Fig. 14 and EOF2 in Fig. 15) at two selected grid points over regions where potential predictability was found. Figure 14a shows that the original ensemble mean agrees well with NCEP reanalysis at this point in the Gulf of Alaska, with warm anomalies during two warm event years (1983 and 1987) and cold anomalies during a cold event (1985), despite the large ensemble spread. Figure 14b shows that the reconstruction in terms of a complete set of 27 EOF modes filters a considerable part of the noise that has nothing to do with the ensemble mean. Figure 14c demonstrates that the ensemble mean is largely controlled by the first EOF, which is a reproducible ENSO-related forced mode. At this grid point, the ensemble mean is predictable, it is determined by the ENSO-related SST anomaly, and the internal chaotic processes have limited impact on the ensemble mean.

Figure 15 presents results for a grid point (30°N, 76°E) over the Indian monsoon region. The simulated local ensemble mean is also consistent with the NCEP reanalysis. However, the model ensemble mean is dominated by the second EOF mode, which is related to non-ENSO boundary forcing as shown in Figs. 3b and 12b. In this case, the first EOF mode (ENSO mode) makes a relatively minor contribution to the ensemble mean.

Figure 16 shows the spatial distribution of temporal correlation between the model ensemble mean and the NCEP reanalysis of T850 anomalies in MAM. Although the correlations are not generally large, over those regions (such as most of the Tropics, Alaska, the southeastern United States, and part of the Indian monsoon region) where good potential predictability was found with PPI analysis, simulated T850 anomalies agree well with the observations. To a first approximation, the model ensemble mean correlates with the observations only over regions with large PPI values.

The fact that the simulated ensemble-mean anomalies over selected regions such as the southeastern United States, Alaska coast, and even Indian monsoon areas where there exists potential predictability have been found to be reasonably consistent with the observations is encouraging. This confirms not only that the ensemble average forecasts are effective, but that the current model has some ability to simulate SST effects on atmospheric seasonal variations.

6. Discussion and conclusions

An approach to assess potential predictability of the extratropical atmospheric seasonal variations produced in ensemble integrations forced with prescribed observed SSTs has been developed in this study. EOF decomposition was used to show that there exist some modes that are less affected by the internal chaotic noise of an AGCM and thus appear to be highly reproducible. These reproducible modes are fundamentally forced modes that are closely related to both ENSO and also non-ENSO-related forcing. By reconstructing the ensemble in terms of a complete set of EOF modes, some part of the internal variability, which appears to be independent of the local ensemble mean, can be excluded from the ensemble without altering the ensemble mean. Reconstructing the ensemble in terms of a subset of the more reproducible modes (forced modes) can further reduce the impacts of internal atmospheric noise and isolate the externally forced response even in regions where that response is insignificant compared to internally generated atmospheric noise.

The PPI, which combines the reproducibility with the

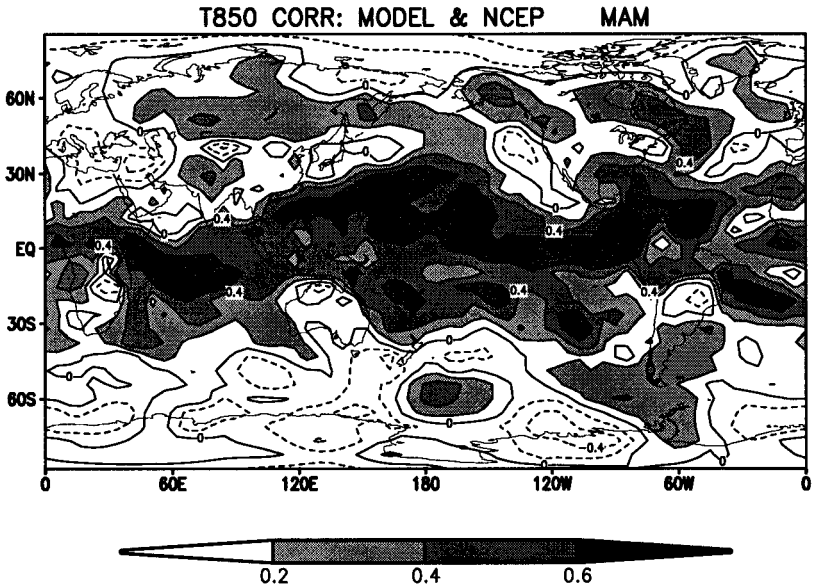

FIG. 16. Spatial distribution of temporal correlation between model and NCEP reanalysis of 850-hPa temperature in MAM; contour interval is 0.2 with values greater than 0.2 shaded.

local variance contribution, was defined to isolate potentially predictable signals over the largest possible extratropical area. Over the regions with larger PPI we say that the ensemble mean whose major part as a distinguished reproducible component represents the SST effect is potentially predictable. Extratropical potential predictability is found primarily over North America and part of the Asian monsoon region. Although ENSO-related SST effects are dominant, especially over North America, non-ENSO SST anomalies were found to be of considerable importance in some seasons over some regions such as the Indian monsoon areas in MAM, central Africa in JJA, and Southeast Asia in SON.

Quantitative analysis of the extratropical potential predictability with PPI has shown that the preferred geographic areas have obvious seasonality. For the 850-hPa temperature, for example, potentially predictable regions during spring and winter are confined to Alaska, northwest Canada, and the southeastern United States—that is, the traditional PNA region, whereas, during summer and fall they are confined to the middle part of North America. It has also been shown that the JJA season possesses the largest potentially predictable area while DJF has the least. The MAM and SON seasons are intermediate.

The seasonality of extratropical potential predictability indicates the influence of the annual cycle on the relative impact of anomalous lower-boundary forcing. However, as argued by Brankovic et al. (1994), there are two potentially conflicting effects of the annual cycle on extratropical predictability. Dynamical teleconnections between the Tropics and extratropics are stronger in winter when potential vorticity gradients and hence Rossby wave dynamics are more intense, suggesting the extratropics has larger predictability conferred through tropical SSTs in that season. On the other hand, the internal chaotic dynamics of the atmosphere is weaker in summer, suggesting a relatively stronger role for extratropical lower-boundary forcing and hence enhanced potential predictability. The model used in this study found that wintertime was not a preferred season for extratropical potential predictability, consistent with the European Centre for Medium-Range Weather Forecasts (ECMWF) model (Brankovic et al. 1994). However, the model used here suggests the preferred season is boreal summer, whereas the ECMWF model favors the spring season. Some possible explanations for this difference are the use of a perfect model framework in the present study and the fact that this study depends solely on boundary forcing while the ECMWF studies make use of the initial atmospheric conditions.

The PPI analysis method seems to be effective in uncovering the preferred geographic regions and favored season for the extratropical potential predictability. It provides considerable useful information in the extratropics, which cannot be described with the direct measure of reproducibility defined by Stern and Miyakoda (1995) due to large chaotic variability. The results from the PPI analysis are consistent with those from the statistical tests for time-varying potential predictability proposed by Anderson and Stern (1996). However, unlike the statistical tests, the PPI analysis may give additional insight into the dynamics leading to the potential predictability. According to this analysis, whether the ensemble average is potentially predictable is determined by whether there exist highly reproducible forced modes that dominate the local ensemble mean. A key point is that there must be some sort of repro-

ducible forced modes that are not distorted by internal dynamics. Fortunately, in the current model, the ENSO-forced modes satisfy these conditions for all seasons and non-ENSO-forced modes do so for all seasons except for DJF. It is encouraging to note that for those geographic regions with high potential predictability, the simulated ensemble mean agrees well with observations.

The results obtained from this study have some implications for seasonal predictions. Due to the large chaotic variability, the prediction of extratropical seasonal variation must be based on an ensemble of forecasts rather than on individual forecasts. Also, despite the larger ensemble scatter in the extratropics useful information is available from the external forcing over some selected regions. In the current model, for example, the ensemble average simulation is able to extract useful information over North America. Since the SST forcing is prescribed from observations, the results from the forced ensemble are an approximate upper bound on the AGCM's capability to forecast seasonal variations.

There are a number of caveats to the conclusions presented here. First, the external forcing used in the AGCM integrations is only the SST. If other lower boundary forcing such as soil moisture, snow cover, and sea ice anomalies are included, potential predictability might be enhanced to an unknown degree. The potentially predictable regions in the present study were mainly limited to North America because the PNA region is unusually sensitive to the SST. Second, the results could be influenced by systematic errors in the AGCM, which could cause the detected regions of potential predictability to be misplaced or distorted. However, this model does appear to be relatively successful in simulating climatology and anomaly patterns over the PNA region. The impact of model systematic error on potential predictability should be a focus of future research. Third, this analysis is based on an ensemble of AGCM simulations, not forecasts. The impact of initial conditions would have an influence on the behavior of forecasts; in general, this should lead to a decrease in potential predictability (and predictability) with lead time. Fourth, due to interaction among air–sea, air–land, and air–ice systems, the change of one type of lower boundary condition would influence another. The final evaluation of potential predictability should be based on results from a fully coupled model of the complete climate system. Finally, as stated in the previous section, the use of the ensemble average forecast is the most straightforward application of ensemble forecasts. As long as the ensemble distribution remains quasi-normal, the use of the ensemble mean is straightforward. However, if the ensemble distribution is not normal, the ensemble average forecast may not be a particularly good choice for prediction. Because of a variety of uncertainties such as initial conditions and model physical parameterizations, a more reasonable seasonal prediction with GCMs would be based on probabilistic forecasts. How to produce and evaluate the probabilistic forecasts from model ensemble forecasts thus become another important issue (Anderson 1996). It seems likely that seasonal prediction with GCMs could be significantly improved with further understanding of the issues discussed above.

Acknowledgments. The authors are grateful to N.-C. Lau and John Lanzante for their valuable reviews of earlier versions of this paper. We also appreciate two anonymous reviewers for their constructive comments on this paper.

REFERENCES

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate,* **9,** 1518–1530.

——, and W. F. Stern, 1996: Evaluating the potential predictive utility of ensemble forecasts. *J. Climate,* **9,** 260–269.

Brankovic, C., T. N. Palmer, F. Molteni, S. Tibaldi, and U. Cubasch, 1990: Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. *Quart. J. Roy. Meteor. Soc.,* **116,** 867–912.

——, ——, and L. Ferranti, 1994: Predictability of seasonal atmospheric variations. *J. Climate,* **7,** 217–237.

Charney, J. G., and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics,* J. Lighthill and R. Pearce, Eds., Cambridge University Press, 735 pp.

Chervin, R. M., 1986: Interannual variability and seasonal climate predictability. *J. Atmos. Sci.,* **43,** 233–251.

Deque, M., 1997: Ensemble size for numerical seasonal forecasts. *Tellus,* **49A,** 74–86.

Ebisuzaki, W., 1995: The potential predictability in a 14-year GCM simulation. *J. Climate,* **8,** 2749–2761.

Ferranti, L., F. Molteni, C. Brankovic, and T. N. Palmer, 1994: Diagnosis of extratropical variability in seasonal integrations of the ECMWF model. *J. Climate,* **7,** 849–868.

Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.,* **73,** 1962–1970.

Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.,* **106,** 447–462.

Gordon, C. T., and W. F. Stern, 1974: Spectral modeling at GFDL. The GARP Programme on Numerical Experimentation. *Int. Symp. on Spectral Methods in Numerical Weather Prediction,* Copenhagen, Denmark, WMO, 46–80.

——, and ——, 1982: A description of the GFDL global spectral model. *Mon. Wea. Rev.,* **110,** 625–644.

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.

Kumar, A., M. Hoerling, M. Ji, A. Leetmaa, and P. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. *J. Climate,* **9,** 115–129.

Latif, M., J. Biercamp, H. von Storch, M. J. McPhaden, and E. Kirk, 1990: Simulation of ENSO related surface wind anomalies with an atmospheric GCM forced by observed SST. *J. Climate,* **3,** 509–521.

Lau, N.-C., 1985: Modeling the seasonal dependence of the atmospheric response to observed El Niño in 1962–75. *Mon. Wea. Rev.,* **113,** 1970–1996.

——, and M. J. Nath, 1990: A general circulation model study of the atmospheric response to extratropical SST anomalies observed in 1950–79. *J. Climate,* **3,** 965–989.

——, and ——, 1994: A modeling study of the relative roles of tropical and extratropical SST anomalies in the variability of the global atmosphere–ocean system. *J. Climate,* **7,** 1184–1207.

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

——, 1981: A quantitative approach to long-range prediction. *J. Geophys. Res.*, **86,** 9817–9825.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44,** 25–43.

Milton, S. F., 1990: Practical extended-range forecasting using dynamical models. *Meteor. Mag.*, **119,** 221–233.

Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble prediction using dynamically conditioned perturbations. *Quart. J. Roy. Meteor. Soc.*, **119,** 299–323.

Murphy, J. M., 1989: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114,** 463–493.

——, 1990: Assessment of the practical utility of extended range ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **116,** 89–125.

Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.*, **120,** 755–793.

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., 1984: Predictability of time averages. Part II: The influence of the boundary forcing. *Problems and Prospects in Long and Medium Range Weather Forecasting*, D. M. Burridge and E. Kallen, Eds., Springer-Verlag, 155–206.

Smith, I. N., 1995: A GCM simulation of global climate interannual variability: 1950–1988. *J. Climate*, **8,** 709–718.

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

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8,** 379–398.

Wallace, M. J., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109,** 784–812.