Potential Predictability of North American Surface Temperature. Part I: Information-Based versus Signal-To-Noise-Based Metrics

Y. TANG

State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou, China, and Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada

D. CHEN

State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou, China

X. YAN*

Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada

(Manuscript received 15 September 2012, in final form 25 September 2013)

ABSTRACT

In this study, the potential predictability of the North American (NA) surface air temperature was explored using information-based predictability framework and Ensemble-Based Predictions of Climate Changes and their Impacts (ENSEMBLES) multiple model ensembles. Emphasis was put on the comparison of predictability measured by information-based metrics and by the conventional signal-to-noise ratio (SNR)-based metrics. Furthermore, the potential predictability was optimally decomposed into different modes by maximizing the predictable information (equivalent to the maximum of SNR), from which the most predictable structure was extracted and analyzed.

It was found that the conventional SNR-based metrics underestimate the potential predictability, in particular in these areas where the predictable signals are relatively weak. The most predictable components of the NA surface air temperature can be characterized by the interannual variability mode and the long-term trend mode. The former is inherent to tropical Pacific sea surface temperature (SST) forcing such as El Niño–Southern Oscillation (ENSO), whereas the latter is closely associated with the global warming. The amplitude of the two modes has geographical variations in different seasons. On this basis, the possible physical mechanisms responsible for the predictable mode of interannual variability and its potential benefits to the improvement of seasonal climate prediction were discussed.

1. Introduction

The predictability of the surface air temperature (SAT) of North America (NA) at seasonal time scales has been widely investigated using both statistical models and dynamical models (e.g., Barnett and Preisendorfer 1987; Shabbar and Barnston 1996; Barnston and Smith 1996; Vautard et al. 1999; Smith and Livezey 1999). It was found that the seasonal climate prediction is more skillful in NA than in other extratropical regions, and significant prediction skills can be achieved for NA, especially when the tropical sea surface temperature (SST) anomalies are large such as at El Niño–Southern Oscillation (ENSO) events (e.g., Shukla et al. 2000; Derome et al. 2001; Kumar and Hoerling 1998).

An important aspect in the study of the seasonal climate predictability is to quantify the potential predictability and detect its source. This is often employed using signal-to-noise ratio (SNR)-based metrics (e.g., Kumar et al. 2003; Peng and Kumar 2005; Kang and Shukla 2006; Quan et al. 2006; Peng et al. 2011). With the analysis of SNR, it was found that the seasonal climate

* Current affiliation: Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, Virginia.

Corresponding author address: Y. Tang, Environmental Science and Engineering, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9 Canada.
E-mail: ytang@unbc.ca

DOI: 10.1175/JCLI-D-12-00654.1

predictability in NA relies mostly on the state of the tropical SST anomalies, and that the extratropical atmospheric seasonal variability over the Pacific–North American region has a pronounced response to the interannual variability of the tropical SST, especially in winter (e.g., Shabbar and Barnston 1996; Barnston and Smith 1996; Kumar and Hoerling 1998; Shukla 1998; Shukla et al. 2000; Schlosser and Kirtman 2005; Peng et al. 2011). The SNR-based predictability measures assume that the signal is always predictable and noise is certainly unpredictable. However, the assumption may not always hold. In some cases, the noise could produce or add additional predictability; for example, when noise varies from year to year having a physical basis, it can amplify the signal, or even develop into the signal by itself. A typical example is the role of the westerly wind burst (WWB) in ENSO predictability, where the former is the noise in nature but can improve ENSO prediction as an effective precursor (e.g., Gebbie et al. 2007). Indeed, some researches indicate that the SNR may not well characterize the potential predictability in some cases (e.g., Yang et al. 2012).

Recently, a new theoretical framework based on information theory has been developed and applied to systematically examine potential predictability at seasonal time scales (e.g, Schneider and Griffies 1999; Kleeman 2002; DelSole 2004, 2005; Tang et al. 2005, 2008; DelSole and Tippett 2007). The advantages and merits of this framework have been well evidenced in the study of ENSO, the Pacific–North American pattern (PNA), the Arctic Oscillation (AO), monsoons, and so on. For example, it has been argued that the information-based potential predictability can explain, to a large extent, the variation of overall prediction skill, especially the anomaly correlation skill (e.g., Tang et al. 2005, 2008). Hence, it is interesting to investigate the potential predictability of the NA seasonal climate using information-based measures, and to further compare it with the predictability by SNR-based measures.

Another issue of potential predictability study is to detect the most predictable component (PrC), which is significant in both a theoretical and practical sense. Theoretically, the temporal evolution and spatial structure of the PrC can shed light on possible physical and dynamical processes inherent to predictability source. Practically, on the other hand, PrC analysis (PrCA) may help detect good prediction targets and improve actual prediction skill. We will demonstrate these merits of PrCA in section 4, as an example, by analyzing NA SAT predictability.

In this paper, we will apply information-based metrics to explore the NA SAT potential predictability, including its temporal and spatial distribution and the most predictable component. This is achieved through analyzing ensemble hindcasts of Ensemble-Based Predictions of Climate Changes and their Impacts (ENSEMBLES; http://www.ecmwf.int/research/EU_projects/ENSEMBLES/exp_setup/index.html). Emphasis will be put on the comparison of potential predictability between information-based and SNR-based measures, and the PrCA using the maximization of predictable information (or equivalent to the maximum of SNR). The latter explores the most predictable pattern of the NA SAT, which has been little addressed in the literature. In addition, as a supplement to earlier studies that mostly focused on the winter season when the SST anomalies are quite large (e.g., Kumar and Hoerling 1998), this study explores the NA seasonal climate predictability for different seasons. The annual variation of the seasonal climate predictability over NA has not been fully demonstrated in the literature, especially in terms of the tier-1 coupled model predictions.

For a clarification, we define the potential predictability used in this study as the prediction skill of a ''perfect'' forecast system (model), which does not make use of observation to measure, whereas the actual prediction skill is the prediction skill evaluated using the observation. The predictability, used as a general expression, means both the potential predictability and actual prediction skill in this study.

This paper is structured as follows: section 2 briefly describes the model and data used. In section 3, the information-based potential predictability is quantitatively evaluated and further compared against the predictability by the conventional SNR measures. Section 4 extracts the most predictable components of NA SAT and analyzes its intrinsic connection to ENSO and PNA, followed by conclusions in section 5.

2. Data and methods

a. Data

The ENSEMBLES assembles ensemble forecasts from five global coupled atmosphere–ocean models, developed by the Met Office (UKMO), Météo France (MF), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Leibniz Institute of Marine Sciences at Kiel University [Leibniz-Institut für Meereswissenschaften (IFM)-GEOMAR], and the Euro-Mediterranean Centre for Climate Change [Centro Euro-Mediterraneo per i Cambiamenti Climatici–Istituto Nazionale di Geofisica e Vulcanologia (CMCC-INGV)] in Bologna. Each ensemble has nine runs of 7-month duration (except 14 in November) generated using different sets of ocean reanalysis generated from wind stress and SST perturbations, initiated separately from February, May, August, and November. The hindcast period of the ENSEMBLES covers the 46 years of

1960–2005. Compared with the previous Development of a European Multimodel Ensemble System for Seasonal to-Interannual Prediction (DEMETER; Palmer et al. 2004), the ENSEMBLES has many modifications such as increased resolution, better representation of subgrid physical processes, land and sea ice, and the inclusion of interannual variability in the greenhouse gas forcing. It has been reported that ENSEMBLES has better seasonal climate prediction skill than DEMETER, especially in the tropical Pacific. Further descriptions of the ENSEMBLES are available in Weisheimer et al. (2009) and Alessandri et al. (2011).

For verification purpose, the TS2.1 dataset of the monthly observed land surface air temperature from the Climatic Research Unit (CRU; http://www.cru.uea.ac.uk/cru/data/hrg.htm) was used. The CRU dataset has been widely used in climate community (e.g., Patz et al. 2002; Jones et al. 2012). An interpolation was performed to convert the data grid of $0.5° \times 0.5°$ to the model grid at the resolution of $2.5° \times 2.5°$.

b. Methods

1) Information-based potential predictability

Information-based potential predictability metrics are based on entropy, a quantitative measure of dispersion level (uncertainty). The entropy of a continuous distribution $p(x)$ is defined as (Cover and Thomas 2006)

$$H(x) = -\int p(x)\ln p(x)\,dx, \tag{1}$$

where the integral is understood to be a multiple integral over the domain of x. The entropy $H(x)$ is the average unpredictability in a random variable.

Suppose that the future state of a climate variable is predicted/modeled as a random variable denoted by ν with a climatological distribution $p(\nu)$. One ensemble prediction produces a forecast distribution, which is the conditional distribution $p(\nu\,|\,\Theta)$ given the initial condition Θ. The difference between the forecast entropy defined by the distribution $p(\nu\,|\,\Theta)$ and the climatological entropy defined by the distribution $p(\nu)$ is an indication of decrease of uncertainty (i.e., potential predictability). Based on the definition of the "difference" ("distance") of two entropies, there are two basic information-based measures, relative entropy (RE) and predictive information (PI), as defined below:

$$\mathrm{PI} = -\int p(\nu)\ln[p(\nu)]\,d\nu + \int p(\nu\,|\,\Theta)\ln[p(\nu\,|\,\Theta)]\,d\nu \quad \text{and} \tag{2}$$

$$\mathrm{RE} = \int p(\nu\,|\,\Theta)\ln\left[\frac{p(\nu\,|\,\Theta)}{p(\nu)}\right]d\nu. \tag{3}$$

Apparently, PI is simply defined by the arithmetical difference between climatology entropy and prediction entropy whereas the RE can be interpreted as a relative difference of both, referred to as Kullback–Leibler divergence. The difference between PI and RE is apparent for a special case when the climatology and prediction distribution are Gaussian, under which RE and PI can be expressed by the ensemble (or forecast) variance σ_p^2, the climatological variance $\sigma_{\mathbf{q}}^2$, and the ensemble (or forecast) mean μ_p and the climatological mean $\mu_{\mathbf{q}}$, namely (e.g., Kleeman 2002; DelSole and Tippett 2007)

$$\mathrm{PI} = \frac{1}{2}\ln\left(\frac{\sigma_{\mathbf{q}}^2}{\sigma_p^2}\right) \quad \text{and} \tag{4}$$

$$\mathrm{RE} = \frac{1}{2}\Bigg[\underbrace{\ln\left(\frac{\sigma_{\mathbf{q}}^2}{\sigma_p^2}\right) + \frac{\sigma_p^2}{\sigma_{\mathbf{q}}^2} - 1}_{\text{Dispersion}} + \underbrace{\frac{\mu_p^2}{\sigma_{\mathbf{q}}^2}}_{\text{Signal}}\Bigg] = \mathrm{PI} + \frac{1}{2}\left(\frac{\sigma_p^2}{\sigma_{\mathbf{q}}^2} - 1 + \frac{\mu_p^2}{\sigma_{\mathbf{q}}^2}\right). \tag{5}$$

The climatological variance is usually constant or slowly varying, thus PI by (4) is dominated by prediction variance (or ensemble spread). This builds a statistical foundation for ensemble spread that is often used as the potential predictability measure in ensemble weather forecast. However the PI ignores the role of ensemble mean that is often referred to as signal component in the RE by (5) and regarded as a response to slowly varying external forcing such as SST. For seasonal climate prediction, the signal component is a main source of predictability (e.g., Kumar et al. 2000, 2007). The relative contribution of the signal component and dispersion component may shed light on the source of potential predictability: internal dynamics or external forcing, since the former is often by characterized by ensemble spread and the latter is by ensemble mean.

A derivative of PI is predictive power (PP), defined by $\mathrm{PP} = 1 - \exp(-\mathrm{PI})$. Under the Gaussian assumption (Schneider and Griffies 1999),

$$\mathrm{PP} = 1 - \left(\frac{\sigma_p^2}{\sigma_{\mathbf{q}}^2}\right)^{1/2}. \tag{6}$$

The above metrics measure the potential predictability of individual prediction, as a function of initial condition and lead time, which offers a means of estimating the

uncertainty of a forecast using the dynamical model. This is similar to the so-called conditional signal-to-noise ratio that measures the predictability of individual prediction, but differs from the conventional SNR potential predictability measure by (13), also called unconditional SNR (Kumar 2007, 2009), which evaluates the overall predictability of the prediction system (simply called SNR hereafter unless otherwise indicated). Corresponding with the SNR-based overall predictability, the average of REs or PIs over all initial conditions reflects the average predictability and was proved to be equal to mutual information (MI), another quantity from information theory (DelSole 2004). In the context of predictability, MI is defined as (Cover and Thomas 2006; DelSole 2004)

$$\mathrm{MI} = \iint p(\nu,\Theta)\ln\left[\frac{p(\nu,\Theta)}{p(\nu)p(\Theta)}\right] d\nu d\Theta, \tag{7}$$

where $p(\nu,\Theta)$ is the joint probability distribution between ν and Θ. MI measures the statistical dependence between ν and Θ, and vanishes when ν and Θ are independent $[p(\nu,\Theta) = p(\nu)p(\Theta)]$; that is, since the prediction target is completely independent of the initial (boundary) conditions, it is regarded as unpredictable. It can be proven that the MI is equivalent to the average of RE or PI over all predictions (Cover and Thomas 2006).

When forecast and climatological distributions are Gaussian, the MI can be simply expressed by (e.g., Yang et al. 2012)

$$\mathrm{MI} = \frac{1}{2}\ln\sigma_{\mathbf{q}}^2 - \langle\ln\sigma_p^2\rangle. \tag{8}$$

2) Relationship between SNR-based metrics and MI-based metrics

For seasonal climate prediction, the total variance (i.e., climate variance) can be decomposed into signal variance and noise variance, if the signal S and noise N are assumed to be independent each other (e.g., Kumar and Hoerling 1998; Shukla et al. 2000; Rowell 1998), namely

$$\mathrm{Var}(T) = \mathrm{Var}(S) + \mathrm{Var}(N), \tag{9}$$

where

$$\mathrm{Var}(T) = \frac{1}{MK}\sum_{i=1}^{M}\sum_{j=1}^{K}(X_{i,j} - \overline{\overline{X}})^2, \tag{10}$$

$$\mathrm{Var}(S) = \frac{1}{M}\sum_{i=1}^{M}(\overline{X}_i - \overline{\overline{X}})^2, \quad \text{and} \tag{11}$$

$$\mathrm{Var}(N) = \frac{1}{MK}\sum_{i=1}^{M}\sum_{j=1}^{K}(X_{i,j} - \overline{X}_i)^2. \tag{12}$$

Note that $X_{i,j}$ is the jth member of the ensemble prediction starting from the ith initial condition. Also, K is the ensemble size and M is the total number of initial conditions (predictions), and

$$\overline{X}_i = \frac{1}{K}\sum_{j=1}^{K}X_{i,j} \quad \text{and} \quad \overline{\overline{X}} = \frac{1}{M}\sum_{i=1}^{M}\overline{X}_i.$$

Two common measures of potentially predictability are the signal-to-noise ratio and the signal-to-total variance (STR), that is,

$$\mathrm{SNR} = \frac{\mathrm{Var}(S)}{\mathrm{Var}(N)} \quad \text{and} \quad \mathrm{STR} = \frac{\mathrm{Var}(S)}{\mathrm{Var}(S) + \mathrm{Var}(N)}. \tag{13}$$

The inequality of arithmetic and geometric means[1] suggests that Eq. (8) could be written by (e.g., Yang et al. 2012)

$$\begin{aligned}\mathrm{MI} &= \frac{1}{2}(\ln\sigma_{\mathbf{q}}^2 - \langle\ln\sigma_p^2\rangle) \geq \frac{1}{2}(\ln\sigma_{\mathbf{q}}^2 - \ln\langle\sigma_p^2\rangle)\\ &= -\frac{1}{2}\ln\left(\frac{\langle\sigma_p^2\rangle}{\sigma_{\mathbf{q}}^2}\right) = -\frac{1}{2}\ln(1-\mathrm{STR}).\end{aligned} \tag{14}$$

Furthermore, we define a MI-based potential predictability as below:

$$\mathrm{AC}_{\mathrm{MI}} = \sqrt{1 - \exp(-2\mathrm{MI})}. \tag{15}$$

$$\text{Thus,} \quad \mathrm{AC}_{\mathrm{MI}} \geq \sqrt{\mathrm{STR}}. \tag{16}$$

The $\sqrt{\mathrm{STR}}$ is actually a perfect correlation skill, and also is equivalent to the correlation of the signal component to the prediction target itself (see the appendix).

The equality in (16) holds if and only if σ_p^2 is constant, as addressed in (14). As argued in DelSole (2004) and Yang et al. (2012), the $\mathrm{AC}_{\mathrm{MI}}$ measures the statistical dependence, linear or nonlinear, between the ensemble mean prediction μ_p and the hypothetical observation ν (an arbitrary ensemble member), whereas the STR only measures their linear correlation. When μ_p and ν are joint normally distributed, their statistical dependence reduces to linear correlation (DelSole 2004). When μ_p and ν are not joint normally distributed, MI naturally

[1] For any nonnegative real number N, $(a_1 + a_2 + \cdots + a_N)/N$ (the arithmetic mean) $\geq (a_1 a_2 \ldots a_N)^{1/N}$ (the geometric mean), and the two means are equal if and only if $a_1 = a_2 = \cdots = a_N$.

disagrees with SNR. The joint normally distributed variables have constant conditional variance. Note that prediction variance is also the conditional variance of ν given the ensemble mean μ_p. Thus, if the prediction variance is varied, μ_p and ν are definitely not joint normally distributed, making the SNR-based potential skill, a linear correlation between μ_p and ν, underestimate the nonlinear statistical dependence between μ_p and ν, which is a strict statistical definition of potential predictability. Thus, the MI-based potential predictability measures are better than SNR-based predictability in characterizing "true" potential predictability. When the climatology and prediction distribution are both Gaussian and the prediction variances are constant, the information-based measure is equivalent to the SNR-based potential measure.

It should be noted that the above conclusion should not be challenged by a possible fact that SNR-based skill might have a better relationship to actual skill than MI-based skill, simply because the actual skill is often measured by the linear correlation (or related quantity), which is inherent to the SNR-based skill.

3) PRINCIPAL PREDICTION COMPONENT ANALYSIS (PRCA) AND MAXIMUM SNR

As argued in Schneider and Griffies (1999), the predictive power PP of (6) is a positively oriented predictive index, defined by the difference between posterior (prediction) entropy and prior (climatology) entropy, thus measuring the decrease of uncertainties due to prediction.

The PrCA analysis is an approach to derive the most predictable component by maximizing PP, or minimizing $\sigma_p^2/\sigma_{\mathbf{q}}^2$. The minimization of $\sigma_p^2/\sigma_{\mathbf{q}}^2$ is equivalent to the maximization of SNR (i.e., $1-\sigma_p^2/\sigma_{\mathbf{q}}^2$). Theoretically, the signal and noise are independent when the ensemble size is infinite. However, the ensemble size is always finite in reality, thus the estimation of the signal is often contaminated by the noise. An optimal estimate for the largest potential predictability should maximize the SNR, from which the resultant signal component is the most predictable. In following discussions, the term of maximum SNR and PrCA are alternatively used given their complete equivalence for Gaussian systems. Both methods identify the spatial pattern, or weight matrix, providing an optimized filter to discriminate the signal and noise. The time series of PrCA reflects the temporal evolution of the dominant mode of the signal, whereas its spatial pattern, which is also referred to as the most predictable pattern, characterizes the spatial distribution of the dominant mode of signal.

Denoting by S and N the signal and noise, PrCA, or maximum of SNR, looks for a vector $\mathbf{q}$, which can maximize the ratio of the variance of signal and noise that are projected onto the vector (DelSole and Tippett 2007):

$$r_S = \mathbf{q}^{\mathrm{T}} S; \quad r_N = \mathbf{q}^{\mathrm{T}} N,$$

$$\mathrm{SNR} = \frac{\sigma_{r_S}^2}{\sigma_{r_N}^2},$$

$$\sigma_{r_S}^2 = \frac{1}{M} r_S r_S^{\mathrm{T}} = \frac{1}{M} \mathbf{q}^{\mathrm{T}} S S^{\mathrm{T}} \mathbf{q} = \mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_S \mathbf{q},$$

$$\sigma_{r_N}^2 = \frac{1}{KM} r_N r_N^{\mathrm{T}} = \frac{1}{KM} \mathbf{q}^{\mathrm{T}} N N^{\mathrm{T}} \mathbf{q} = \mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_N \mathbf{q}, \quad \text{and}$$

$$\mathrm{SNR} = \frac{\mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_S \mathbf{q}}{\mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_N \mathbf{q}} \rightarrow \mathrm{Maximum}. \tag{17}$$

Mathematically the optimization by (17) leads to a generalized eigenvalue/eigenvector problem based on the Raleigh quotient theorem (Venzke et al. 1999; Schneider and Griffies 1999):

$$\mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_S = \lambda \mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_N, \tag{18}$$

where $\boldsymbol{\Sigma}_S$ is the covariance matrix of signal and $\boldsymbol{\Sigma}_N$ is the covariance matrix of noise. The solution of (18) can be obtained by solving this eigenvalue equation:

$$\mathbf{q}^{\mathrm{T}} \boldsymbol{\Sigma}_S \boldsymbol{\Sigma}_N^{-1} = \lambda \mathbf{q}^{\mathrm{T}}. \tag{19}$$

Thus, the PrCA is also called the signal-to-noise maximizing EOF (MSN EOF) analysis, first introduced by Allen and Smith (1997), which derived the signal by removing the influences of noise, and widely used already in climate predictability study (e.g., Fukunaga 1990; Venzke et al. 1999; Sutton et al. 2000). Practically, the number of grid points is always much larger than the number of total samples in climate studies, thus $\boldsymbol{\Sigma}_N$ is usually not full-rank, leading to a solution of ill-conditioned inversions. There are two common methods to solve this issue: a truncated EOF approach and a whitening approach. Details of the two approaches can be found in Venzke et al. (1999) and Schneider and Griffies (1999).

c. *Analysis strategy*

In this study, the analyses were carried out for the period from January 1969 to December 2001 (33 yr), which is a common period with the second phase of the Canadian Historical Seasonal Forecasting Project (HFP2) ensemble products. The comparison of potential predictability between ENSEMBLES and HFP2 will be reported in Part II of this work (Y. Tang et al. 2013, unpublished manuscript). For consistency, we use this common period in

this work. However, the results and conclusions presented in this paper are not changed by using the entire period of ENSEMBLES experiments. The analysis area covers all land grids in NA (20°–80°N, 150°–60°W). The model climatology is obtained by averaging all its ensemble members. The multiple model ensemble (MME) is used for analysis. It has been argued that MME is usually better than single model ensemble (SME) since the uncertainties associated with the different model frameworks can offset each other and be relieved by large size of ensemble members (e.g., Krishnamurti et al. 1999, 2000; Palmer et al. 2004; Yan and Tang 2012). The MME mean was constructed here in the simplest way, namely by averaging each model ensemble mean (i.e., five means for ENSEMBLES). All the ensemble members of ENSEMBLES are pooled together with equal weights to form superensembles for the PrCA (i.e., 45 members for ENSEMBLES MME). To explore the seasonal dependence of predictability, we performed all analyses at each season. Unless otherwise indicated, the prediction is one season ahead using the seasonal mean as the prediction target in this study. For example, the mean over March–May (MAM) is used as the target time of spring prediction, initialized by the beginning of February. Similarly, the mean over June–August (JJA) is the target time of summer prediction, initialized by the beginning of May, and forth on for other two seasons. During the entire discussions of this paper, the calendar seasons mentioned or labeled in figures are the target time of prediction.

3. MI-based potential predictability

a. General features of MI-based skill

The MI-based potential predictability AC_{MI}, defined in (15), is shown in Fig. 1a. As can be seen in this figure, the AC_{MI} is location and season dependent. Relatively strong AC_{MI} occurs at winter and spring, followed by summer and fall. From (14) and (15), we can see that the variation in AC_{MI} is actually determined by the signal and noise variance. A further analysis of signal and noise variance, as shown in Fig. 2, indicates that the seasonal dependence of AC_{MI} is controlled by the signal component. The strong signal in winter and spring is due to two reasons. First, the tropical Pacific SSTA has a strong interannual variability in winter. Second, the response of NA climate to ENSO probably has a time lag from two weeks to one season (e.g., Jin and Hoskins 1995; Hall and Derome 2000; Mo et al. 1998; Kumar and Hoerling 2003). Unlike the signal, the noise displays a relatively weak seasonal dependence. This suggests that the noise is model dependent while the signal is mainly from external forcing (e.g., the tropical Pacific SSTA).

Spatially, AC_{MI} shows a larger amplitude over northwestern Canada and the southeastern United States in winter and spring, which is due to the teleconnection of NA climate to ENSO. In section 4, we will see that a strong PNA structure, represented in the 500-mb geopotential height (1 mb = 1 hPa), is an associated pattern of the most predictable structure of the NA SAT, leading to such a distribution of AC_{MI}. On the other hand, it has been well argued that although the PNA pattern is a natural internal mode of climate variability, it is also strongly influenced by the ENSO forcing. We will further discuss this issue in section 4.

The noise is nearly spatially uniform in winter, but shows some spatial variation in other seasons, with a strong amplitude distributed in Canada and in the northern United States, especially in spring as shown in Fig. 2. The relatively strong noise in northern NA may be due to large uncertainties in parameterizing physical processes of formation and dissolution of snow and ice. Compared with the signal, however, the noise has much less spatial variation, resulting in a signal-like distribution for AC_{MI}. Thus, the main features of predictability of NA SAT are dominated by the signal components, consistent with previous studies (e.g., Kumar et al. 2000, 2007).

It should be noted that the noise has a significant impact on the strength of potential predictability, although it has little contribution to the spatial distribution and seasonal dependence of potential predictability due to its relatively uniform variation in time and space. In fact, the noise variance from (12) is far larger than the signal counterpart from (11), as indicated by the color bar in Fig. 2. This also suggests that the internal variability is much larger than the external forcing over NA, probably explaining why the seasonal climate prediction skill over NA is quite limited (e.g., SAT prediction skill as shown below) compared with the predictability in the tropical regions.

b. SNR-based predictability

The SNR-based potential predictability, defined by $\sqrt{\mathrm{STR}}$, is also analyzed. The SNR (STR), measuring the relative role of the external forcing to the internal high frequent noise, has been widely used to quantify the potential predictability at seasonal or longer time scales. The result shows that SNR-based potential predictability is very similar to AC_{MI} in both spatial pattern and seasonal variations (not shown). This is not surprising since both measures are based on the signal and noise. However the MI-based potential predictability is higher than, or at least equal to, the SNR-based skill, as indicated by (16). Shown in the right panel of Fig. 1 is the subtraction of $\sqrt{\mathrm{STR}}$ from AC_{MI}, where the statistically

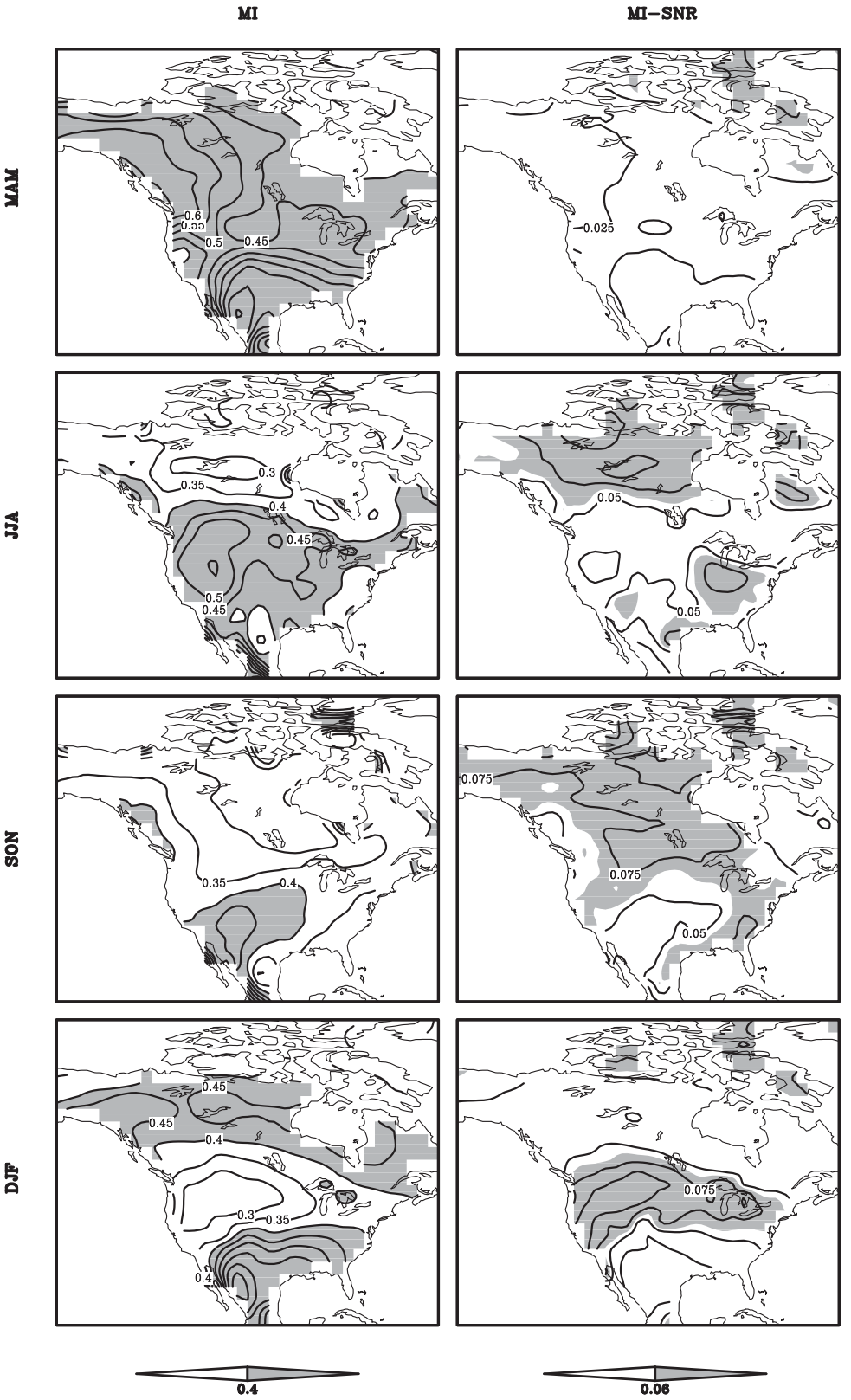

FIG. 1. (a) MI-based potential predictability AC_{MI} and its difference from the SNR-based measure $\sqrt{STR}$ (see text).

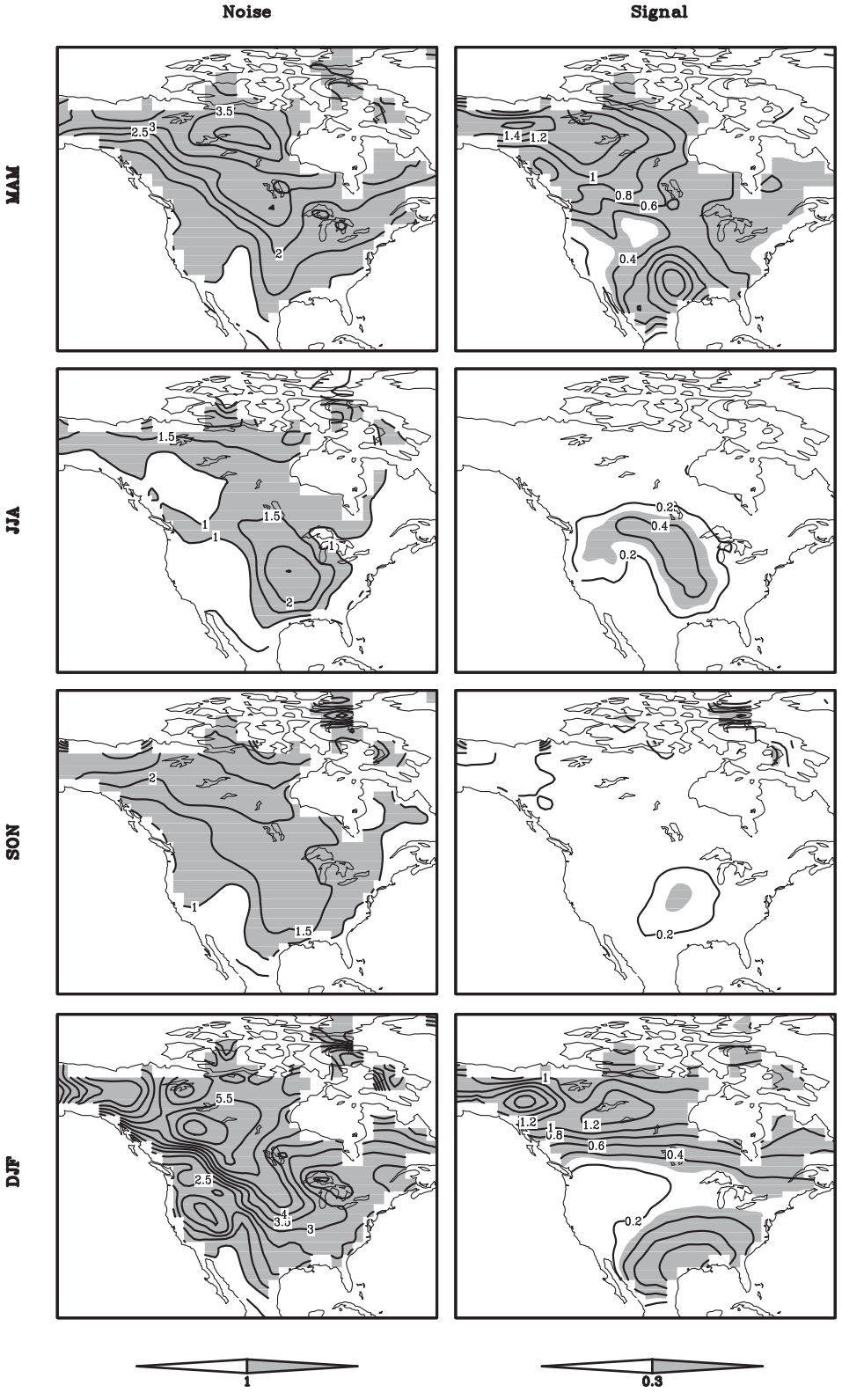

FIG. 2. Signal and noise distribution.

significant difference is shaded. The bootstrap experiment was here used to perform the significant test. The central idea of the test is to obtain an upper limit of the difference between AC_{MI} and $\sqrt{STR}$ due to sampling errors, beyond which the difference can be argued to be statistically significant. To implement this, we randomly took 80% of all ensemble predictions, from which $\sqrt{STR}$ was calculated, and also a surrogated AC_{MI} was obtained that used a constant prediction variance (i.e., the mean of the ensemble spreads). As argued above, the MI-based skill and SNR-based skill are both identical to each other if the prediction and climatological distributions are Gaussian, and the prediction variance is constant. Considering the fact that a Gaussian distribution is usually a good approximation for seasonal climate variables (e.g., Schneider and Griffies 1999; Tang et al. 2005), the difference between surrogated AC_{MI} and $\sqrt{STR}$ should be completely due to sampling error. This process was repeated 500 times, which can estimate an upper limit of sampling errors.

As shown in Fig. 1b, the difference between AC_{MI} and $\sqrt{STR}$ is significantly positive in some regions, indicating that the SNR-based metrics underestimate the potential predictability there. This underestimation is in particular apparent in the fall and summer. Comparing Figs. 1a and 1b reveals that the significant difference between AC_{MI} and $\sqrt{STR}$ often occurs at these regions where the potential predictability is relatively low. This is probably because these areas are more sensitive to the variation of noise (ensemble spread) due to their relatively weak signal strength.

Usually, as the initial perturbation grows with lead time, the ensemble spread tends to have more variations among individual predictions. Thus, one may expect that the SNR-based measure would more underestimate the potential predictability for longer climate prediction. To explore this, we calculated the potential predictability of two seasons ahead, as shown in Fig. 3. Comparing Fig. 3 with Fig. 1 reveals that the difference between AC_{MI} and $\sqrt{STR}$ indeed increases as lead time increases; that is, the SNR-based measure can lead to more severe underestimation for longer time prediction. Furthermore, it is found that the potential predictability is not much related to initial conditions at seasonal time scales, since Figs. 1 and 3 have different initial times of prediction but show a very similar potential predictability pattern. Again, the calendar seasons labeled in Fig. 3 are the target time of prediction. For example, the target season March–May in Fig. 3 corresponds with prediction starting from 1 November of the previous year (two seasons ahead) whereas the same target time in Fig. 1 corresponds the prediction from 1 February. The less impact of initial conditions on the potential predictability is due to the greater significance of external forcing to climate prediction at seasonal time scale or longer (Peng et al. 2011).

4. The most predictable pattern (PrCA) of the NA SAT

a. General feature of leading PrCA modes

Figures 4 and 5 show the spatial pattern and temporal variation of the first two most predictable component analysis modes, which explain totally over 50% of signal variance for all seasons as shown in Table 1. It was found that other PrCA modes beyond the first two modes have a very small and negligible contribution. It should be noted that, unlike in principal component analysis (PCA), the second mode can explain more variance than the first mode in PrCA as shown in Table 1 for fall and winter. This is because PrCA seeks the optimal modes based on predictable component (predictability) rather than on explained variance (variability).

As can be seen in Figs. 4 and 5, the two most predictable patterns characterize the predictability of interannual variability and long-term trend. When the first mode represents the predictability of interannual variability, the second mode mainly describes the long-term trend modulated by interannual variability, and vice versa. For example, the first PrCA mode characterizes the predictability of interannual variability in spring and winter and the predictability of long-term trend in summer and fall, whereas the second mode is just complementary with the first mode in characterizing the predictability of the two time scales. Since the prediction target is seasonal mean here, the intraseasonal variability has been filtered prior to PrCA analysis. One may expect that the PrCA also can include the predictability of intraseasonal variability if daily or weekly data are used.

Comparing Figs. 4 and 5 reveals that the PrCA mode of interannual variability shows spatially a dipole structure between northwestern Canada and the southeastern United States. Such a dipole structure is very significant in spring and winter, as shown in the PrCA mode 1, but less visible in summer and fall shown in PrCA mode 2, which is due to the seasonal variation of the strength of SST forcing, as argued in section 3a. We will further discuss the PrCA mode of interannual variability later. On the other hand, the predictability of long-term trend is inherent to the models' capability in capturing the signal of the global warming, which is nearly uniformly distributed in the whole domain over NA although its strength has some geographical variations in different seasons.

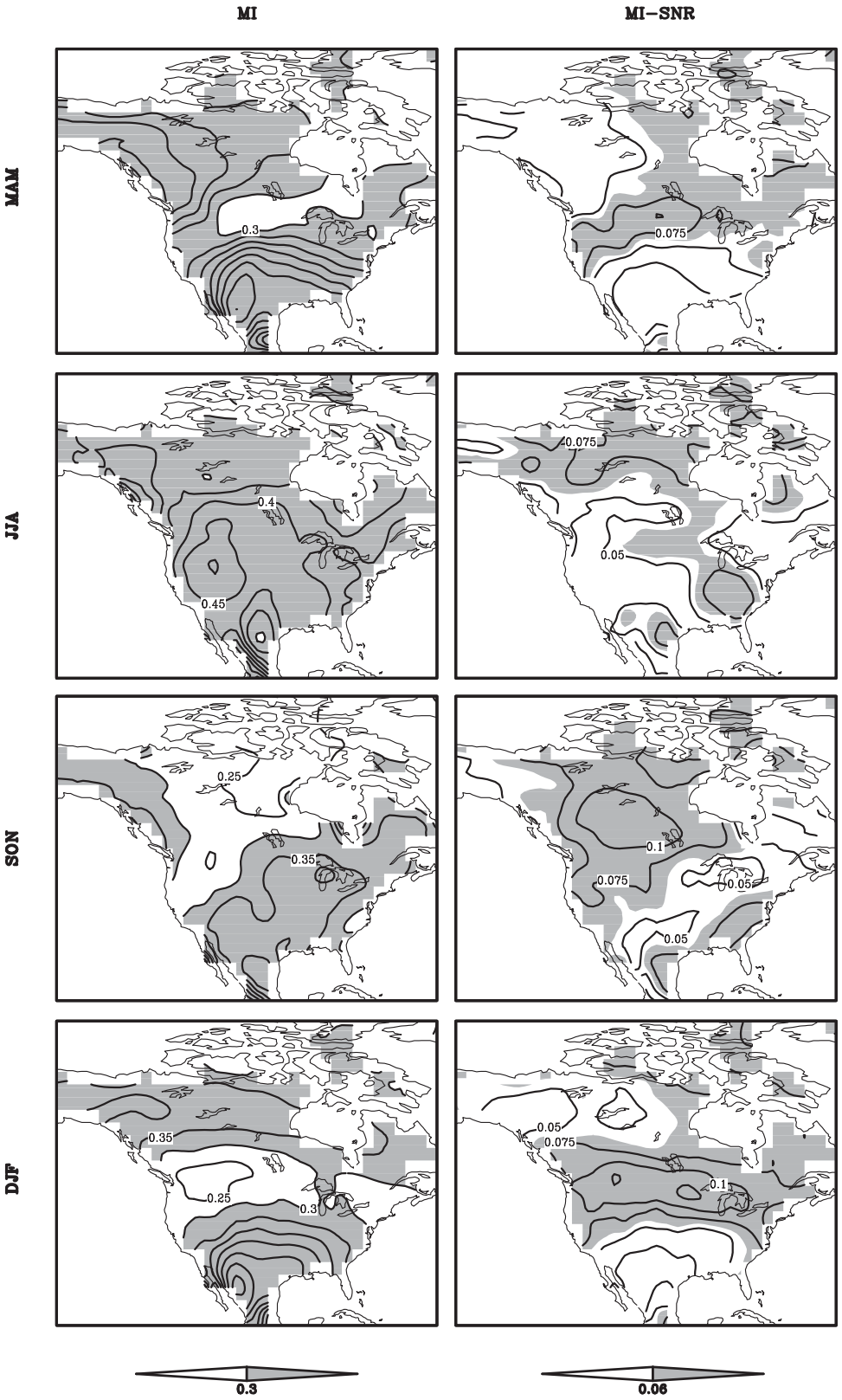

FIG. 3. As in Fig. 1, but for two-season prediction (the calendar season is the target time of prediction, such as MAM meaning the prediction starting from November).

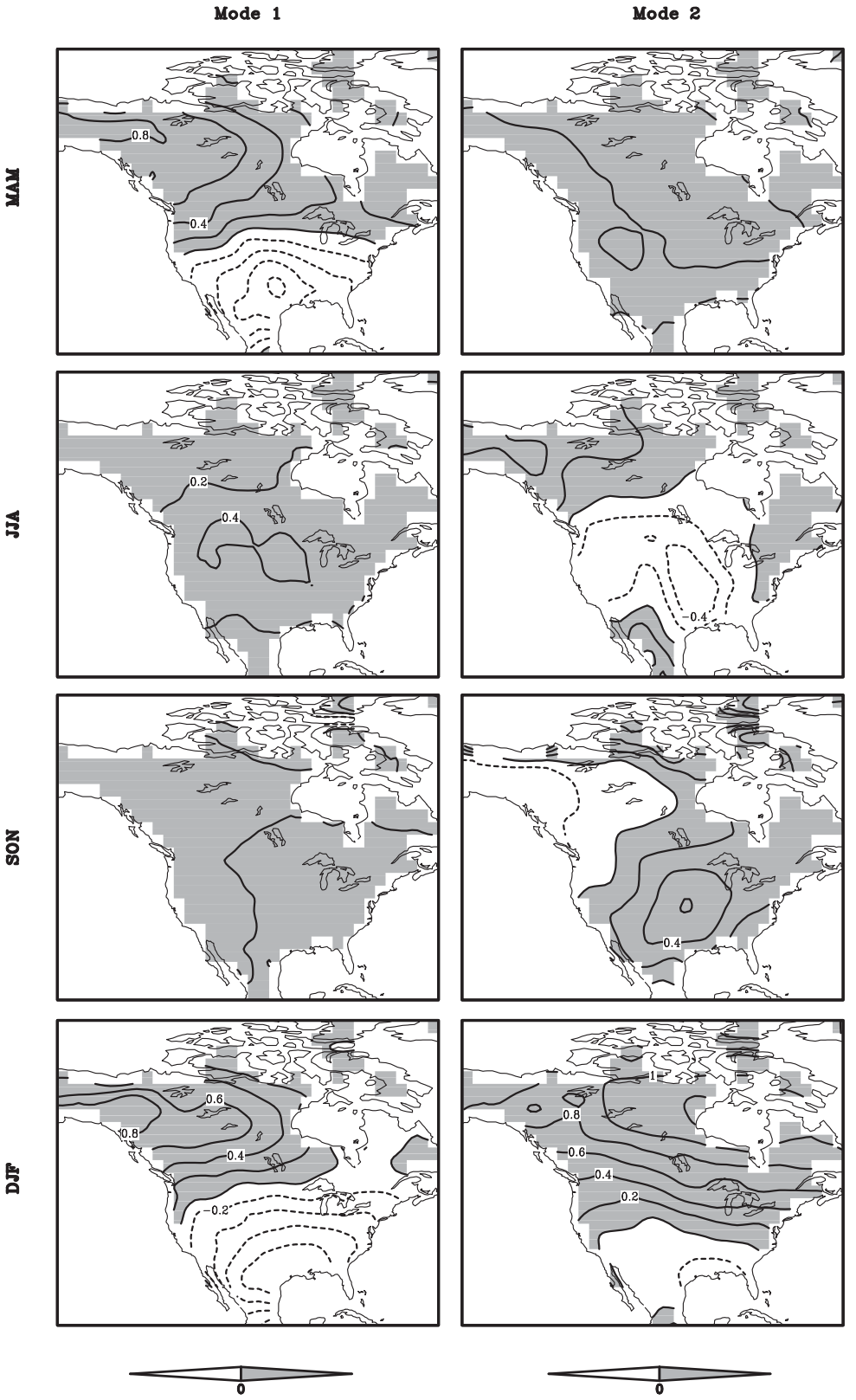

FIG. 4. The most predictable pattern of the NA SAT at the lead of one season for different seasons (prediction target).

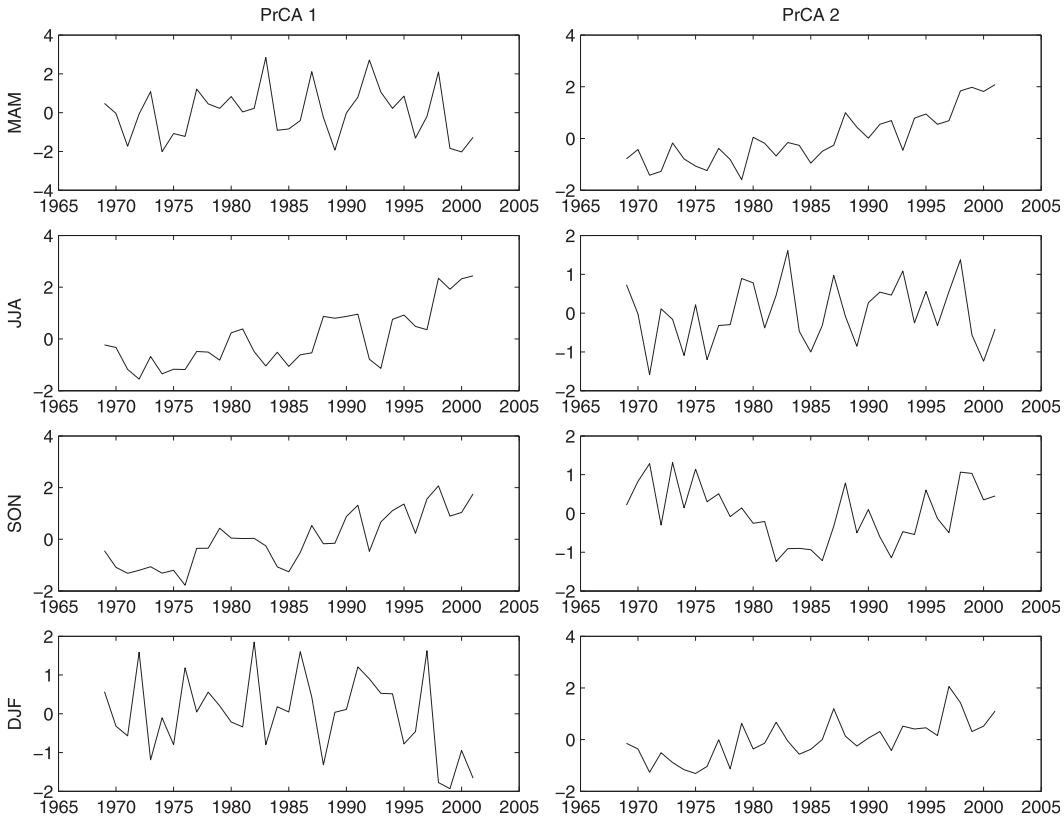

FIG. 5. The time series of PrCA modes 1 and 2.

Table 2 shows the prediction skill of the time series of the first two PrCA modes against the observed counterpart. The observed PrCA modes are obtained by projecting observed SAT onto the optimal filters derived from PrCA analysis. As can be seen in this table, the PrCA modes have high prediction skills for long-term trend for all seasons. However for interannual variability, the most predictable skill has significant seasonal dependence, with much higher skill in spring and winter than in summer and fall. In next section, we will discuss possible reasons for this seasonal dependence.

b. The PrCA of interannual variability

ENSO is the strongest interannual variability in the earth. It has been well recognized that the seasonal climate predictability of midlatitudes mainly originates from the ENSO predictability. This presumption is mainly supported by the fact that there exists a strong teleconnection between ENSO and the extratropical climate variability in both observations and modeling as discussed in the introduction. In this subsection, we will explore the direct link of the seasonal climate predictability to ENSO. Toward this, we derived the pattern of the tropical sea surface temperature anomalies (SSTAs) associated with the PrCA modes, by projecting the time series of the PrCA modes onto the tropical SSTA. This projection is equivalent to the regression coefficient of the time series of the PrCA mode to the SSTA.

Figure 6 is the SSTA spatial patterns associated with the first two PrCA modes of NA SAT, where the SSTA is from the ENSEMBLE MME mean. We also projected the PrCA modes onto the observed SSTA and obtained associated patterns similar to Fig. 6 (not shown). A comparison of Fig. 6 with Figs. 4 and 5 reveals that the PrCA mode of interannual variability is closely associated with ENSO. For example, the first PrCA mode of NA SAT

TABLE 1. Variance accounted for by each PrCA mode [September–November (SON) and December–February (DJF)].

	MAM	JJA	SON	DJF
Mode 1	55%	43%	26%	29%
Mode 2	24%	18%	30%	45%

TABLE 2. Prediction skill of each PrCA mode against observed counterpart.

	MAM	JJA	SON	DJF
Mode 1	0.73	0.75	0.71	0.63
Mode 2	0.77	0.47	0.40	0.67

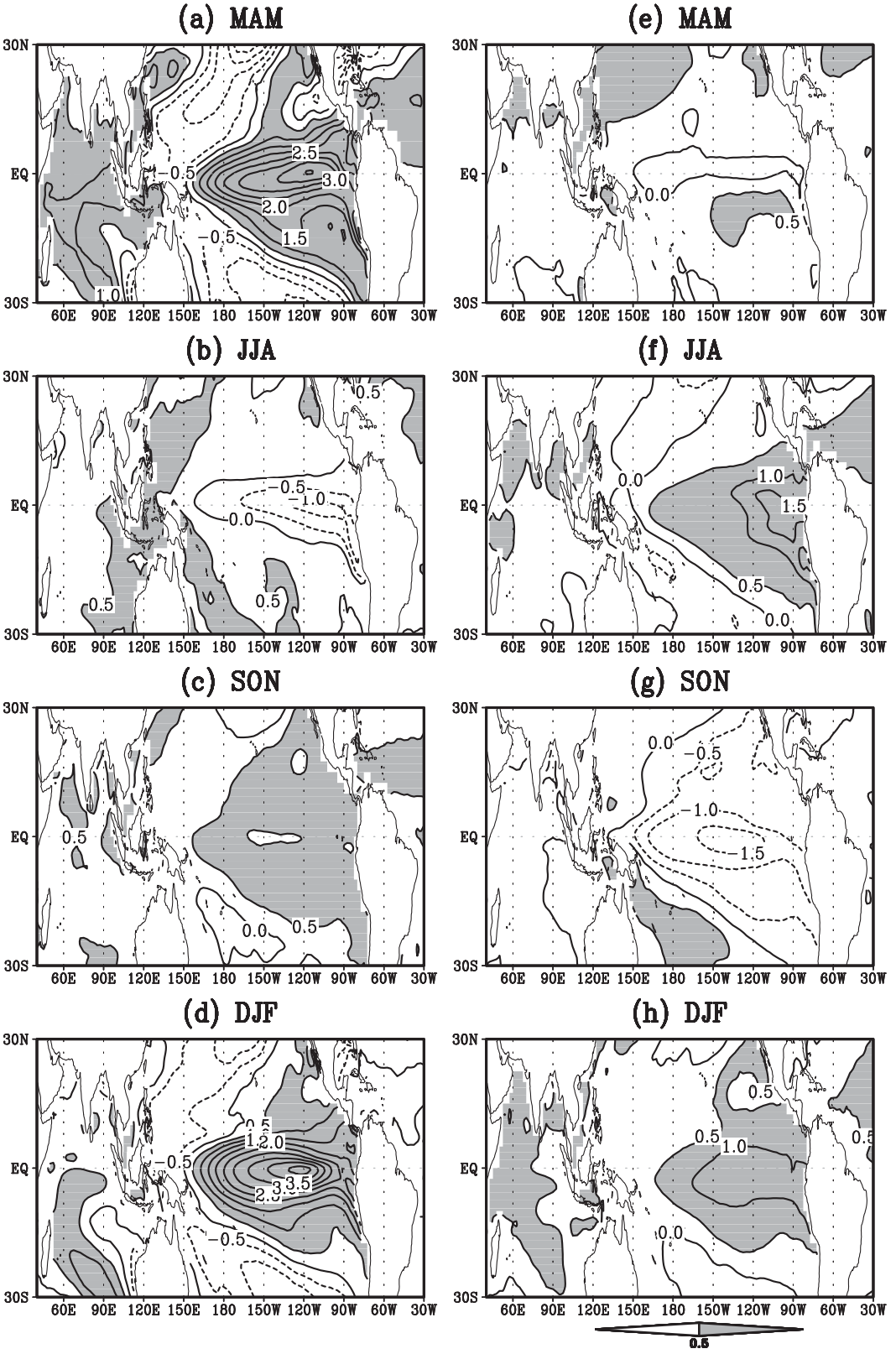

FIG. 6. The SST patterns associated with the PrCA mode 1 and mode 2 of SAT.

has strong interannual variability (dipole pattern) in spring and winter, which corresponds with strong El Niño events in the same seasons as shown in Fig. 6. A similar argument can be made for the association of the second PrCA modes of NA SAT with ENSO for summer and fall, when both PrCA and SSTA show relatively weak variability. Thus, the most predictable component of interannual variability of the NA SAT is clearly associated with the ENSO forcing.

As addressed above, the PrCA modes of interannual variability are characterized by a dipole structure. A further comparison between Figs. 4 and 6 reveals that the El Niño events correspond with warming in the northwestern Canada and cold in the southeastern United States. This is in particular obvious for mode 1 in winter and spring. On the other hand, the La Niña events lead to a cold climate in northwestern Canada and warmth in the southeastern United States as shown in mode 2 for fall.

To shed light on the link between the dipole structure of the PrCA of NA SAT with ENSO, we projected the time series of the first PrCA mode onto the ENSEMBLE MME 500-mb height ensemble mean, as shown in Fig. 7. As can be seen in this figure, a PNA-like pattern can be observed in the spring and winter, indicating that the dipole structure of NA SAT PrCA mode is highly associated with PNA-like pattern. For example, in winter and spring when the dipole is strong in Fig. 4, a strong positive PNA-like phase can be observed in Fig. 7. Likely, when the PNA-like pattern is very weak or absent in summer and fall, the dipole pattern of the PrCA mode of NA SAT almost disappears. Thus, the PNA bridges the ENSO and the NA SAT predictability. Such a teleconnection of NA climate predictability to ENSO is highly related to the teleconnection of the ENSO to NA climate variability, which has been well documented in literature (e.g., Trenberth et al. 1998; Quan et al. 2006; Hoerling and Kumar 2002). Figure 8 is the regression coefficient of the first principal component (PC) of the Pacific SSTA to the NA SAT, showing such a dipole pattern of NA climate variability teleconnected to ENSO.

The dipole structure of NA SAT can be easily interpreted by its associated atmospheric pattern PNA. The positive PNA phase features stronger than usual heights over the intermountain region of North America, denoting advection of maritime air into the western and northern parts of North America and some warming from subsidence directly under the high pressure ridge. Furthermore, the pressure gradient between the high ridge and the Aleutian low results in an increase in wind speed that steers more storms into the Pacific Northwest, thus reducing clouds cover and inducing higher daytime temperature. Also during the positive phase of the PNA, deepening troughs of the low pressure exist over the southern United States, bringing colder air down farther into the south from the northern latitudes, increasing the winter storm activity for these regions, and causing below average temperatures across the south-central and southeastern United States.

The above analysis offers a possible source of the most predictability component of NA SAT, namely that it originates from ENSO and is dominated by the PNA. Now, one question here is how much the ENSO and PNA variability can be explained in the prediction of the tropical Pacific SST and 500-mb height. A natural expectation is that the first PrCA mode of SSTA and 500-mb height would be respectively characterized by a typical ENSO and PNA pattern. To answer this, we conducted the PrCA using ENSEMBLES for SST and 500-mb height anomalies. Figure 9 is the first mode of PrCA of 500-mb height, showing a pattern very similar to Fig. 7, a strong PNA pattern in spring and fall. The first two modes of PrCA of SSTA are also very similar to Fig. 6 (not shown). Thus, a possible mechanism for why the most predictable component of NA SAT is a dipole structure can be proposed as shown in schematic Fig. 10, namely that the most predictable component of interannual variability over the NA SAT is dominated by PrCA1 of atmospheric and oceanic interannual variability.

c. Application of PrCA on the prediction

In this section, we will evaluate the benefits of PrCA to the improvement of actual ensemble prediction. It is straightforward to use PrCA leading modes to reconstruct the predicted field, which is achieved by multiplying the spatial modes by corresponding time series. This is the same as the approximation of the raw field using principal component analysis lead modes. In this study, since the variance contribution by individual modes become rather small after the first two modes, thus we used the first two PrCA modes to construct the prediction. To remove possible artificial skill, the PrCA was performed in the framework of cross validation. Namely, the forecast of one specific year was left intact as predictand, and the remaining forecasts were decomposed by PrCA to extract the most predictable components and their corresponding spatial patterns. A specification of the predictand was then obtained by projecting the predictand on the first two PrCA spatial modes and their time series. This process was iterated to generate the forecast field.

Shown in the left panels of Fig. 11 are the correlation skills by PrCA prediction against the observation. As can be seen in this figure, the PrCA prediction skill

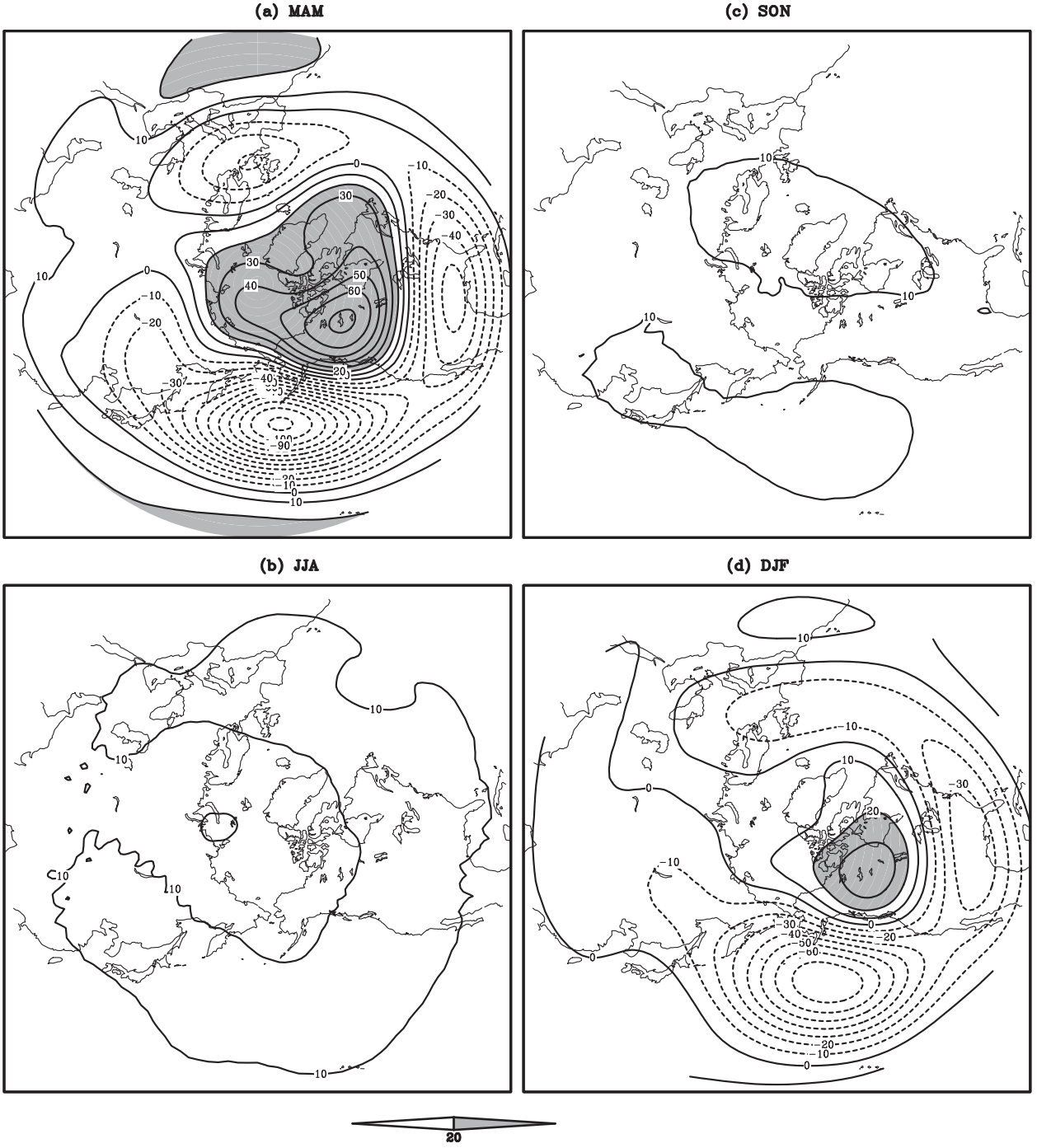

FIG. 7. The 500-mb height pattern associated with the PrCA SAT mode 1.

shows some seasonal dependence with the best skill in fall and winter, followed by spring and summer. This seasonal dependence is similar that of ensemble mean prediction but quite different from the potential predictability. The difference between potential predictability and actual prediction skill characterizes the impact of model bias on prediction skill, which we will discuss in future work.

The difference of correlation skill between PrCA prediction and ensemble mean prediction is shown in the right panel of Fig. 11. It is visible that the PrCA is better than the ensemble mean in some regions where

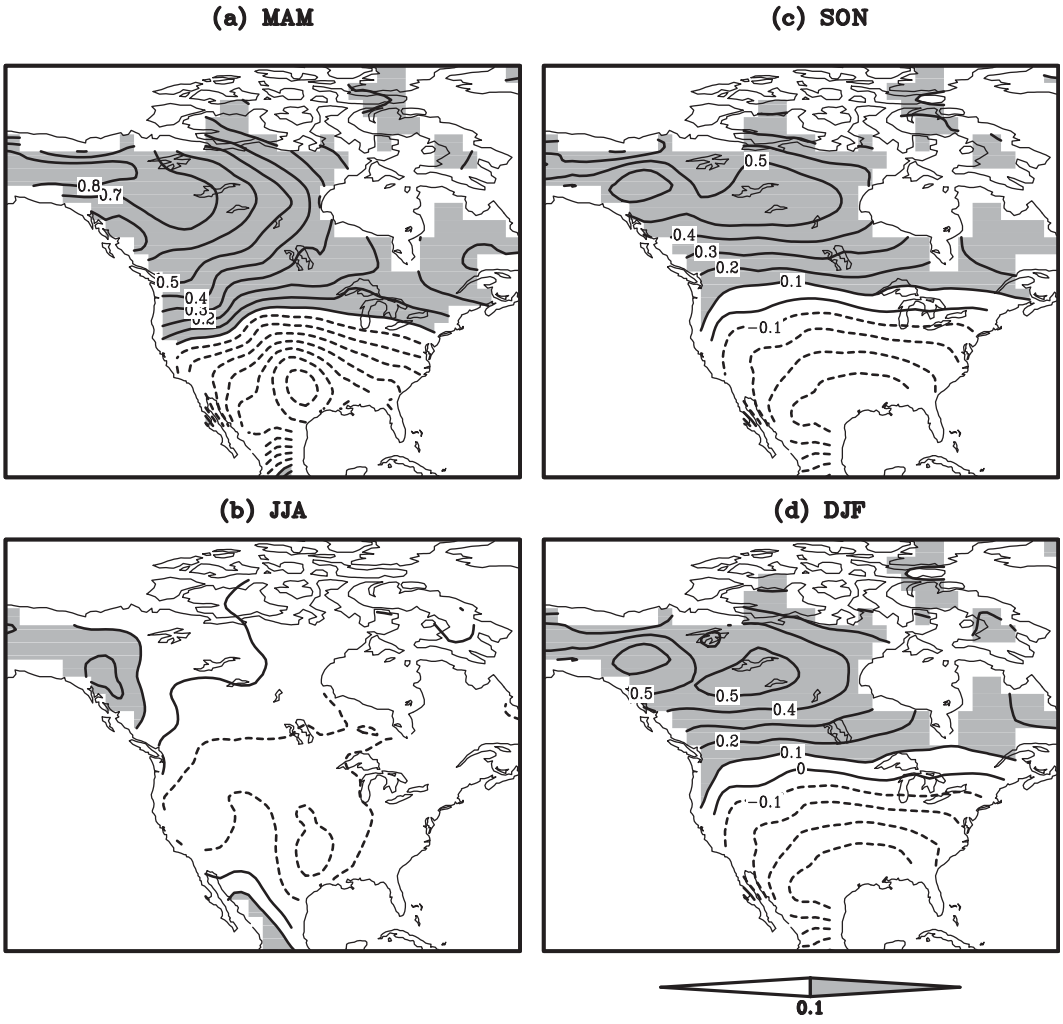

FIG. 8. Regression coefficient of NA SAT to the first PC of tropical SST.

the PrCA lead modes have large variance as shown in Fig. 4, which is particularly obvious in fall and summer. For these regions with small variance in PrCA spatial patterns, ensemble means seem better than PrCA prediction. Similar results can be obtained using RMSE (not shown). This may suggest a possible method to improve prediction skill for some regions, with the application of PrCA. A detailed discussion would be required for further application of PrCA into the practical prediction such as forecast validation and statistical significant test, which we will pursue in future work.

The advantage of PrCA can be more clearly presented by comparing Figs. 12 and 13, which show the first PrCA time series and PCA time series, both against their observed counterpart. The PCA was employed using the ensemble mean prediction, thus its first time series represents the main feature of prediction. Apparently, the PrCA time series has much better prediction skill than PCA counterpart for all seasons. This is not surprising because PrCA extracts the most predictable component whereas the PCA extracts the component of the most variance.

5. Conclusions and discussion

Many dynamical systems relevant to atmosphere and ocean have a fundamental limitation to their predictability caused by strong nonlinearity or stochastic forcing (i.e., the prediction uncertainty). Such uncertainty is intrinsic to physical system itself, which is basically independent of the representation of the physical system and characterized by the rate of separation of initially close states. In other words, the potential prediction uncertainty always exists even for a perfect model and cannot be removed or alleviated by the model improvement. Thus, the study of potential predictability can

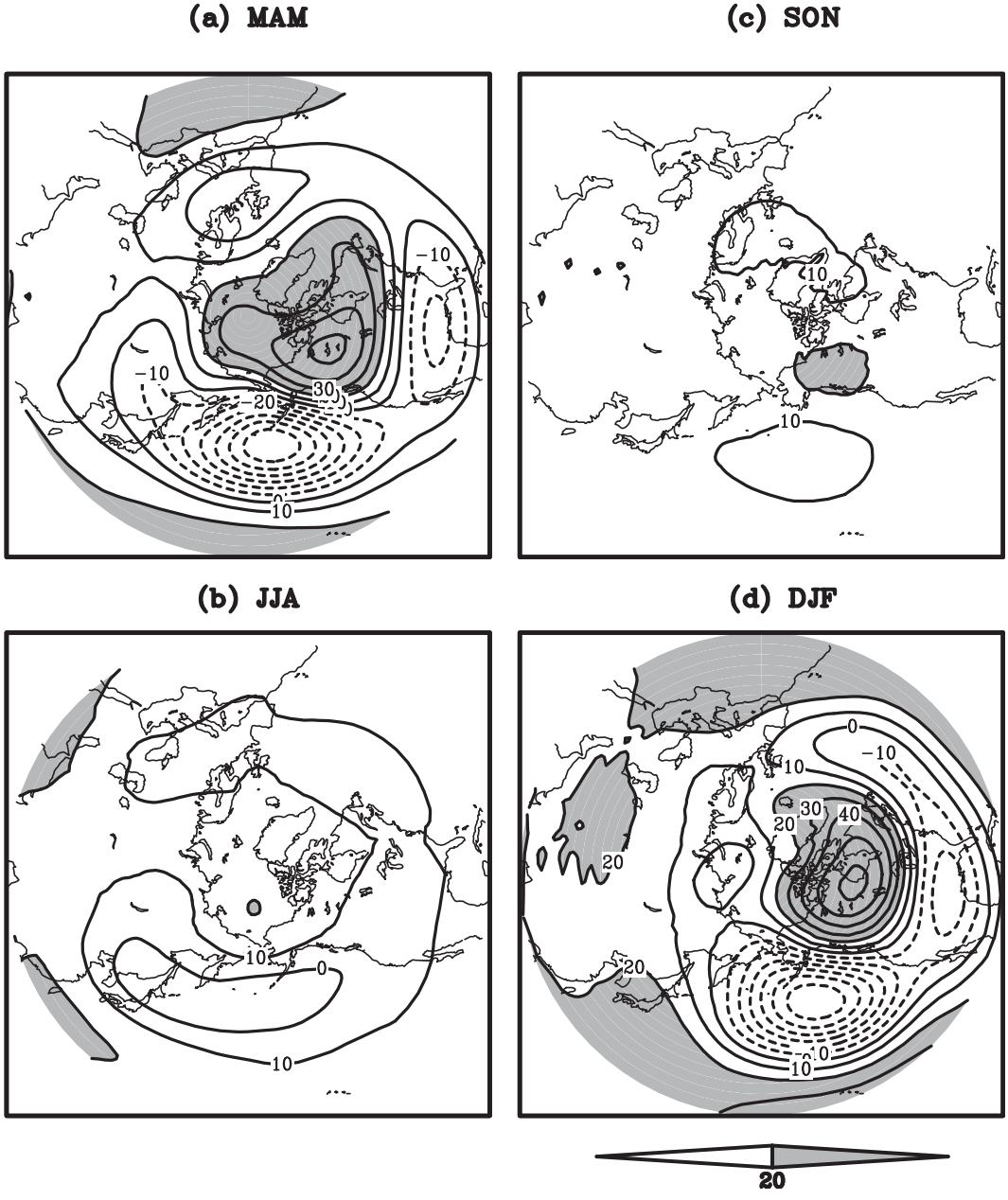

FIG. 9. The PrCA mode 1 for 500 mb.

probe the possible room of the improvement of prediction skill, and guide the direction of model development.

The accurate estimate of potential predictability is a challenging issue since its true value is actually unknown. Currently, the potential predictability is usually measured by the signal-to-noise ratio or the perfect model correlation skill (ensemble mean against ensemble member), both actually equivalent to each other as shown in the appendix. This approach raises some concerns. First, the assumption that the signal is predictable and noise is unpredictable is not always held. In some cases, the noise may produce or add additional predictability, for example, the role of the west wind burst (WWB) and high nonlinearity in ENSO predictability as mentioned in the introduction. Second, the correlation between ensemble mean and ensemble member actually

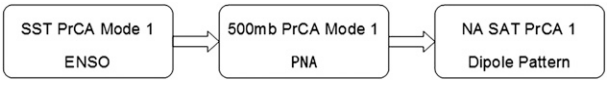

FIG. 10. Schematic diagram of teleconnection from ENSO to NA SAT predictability.

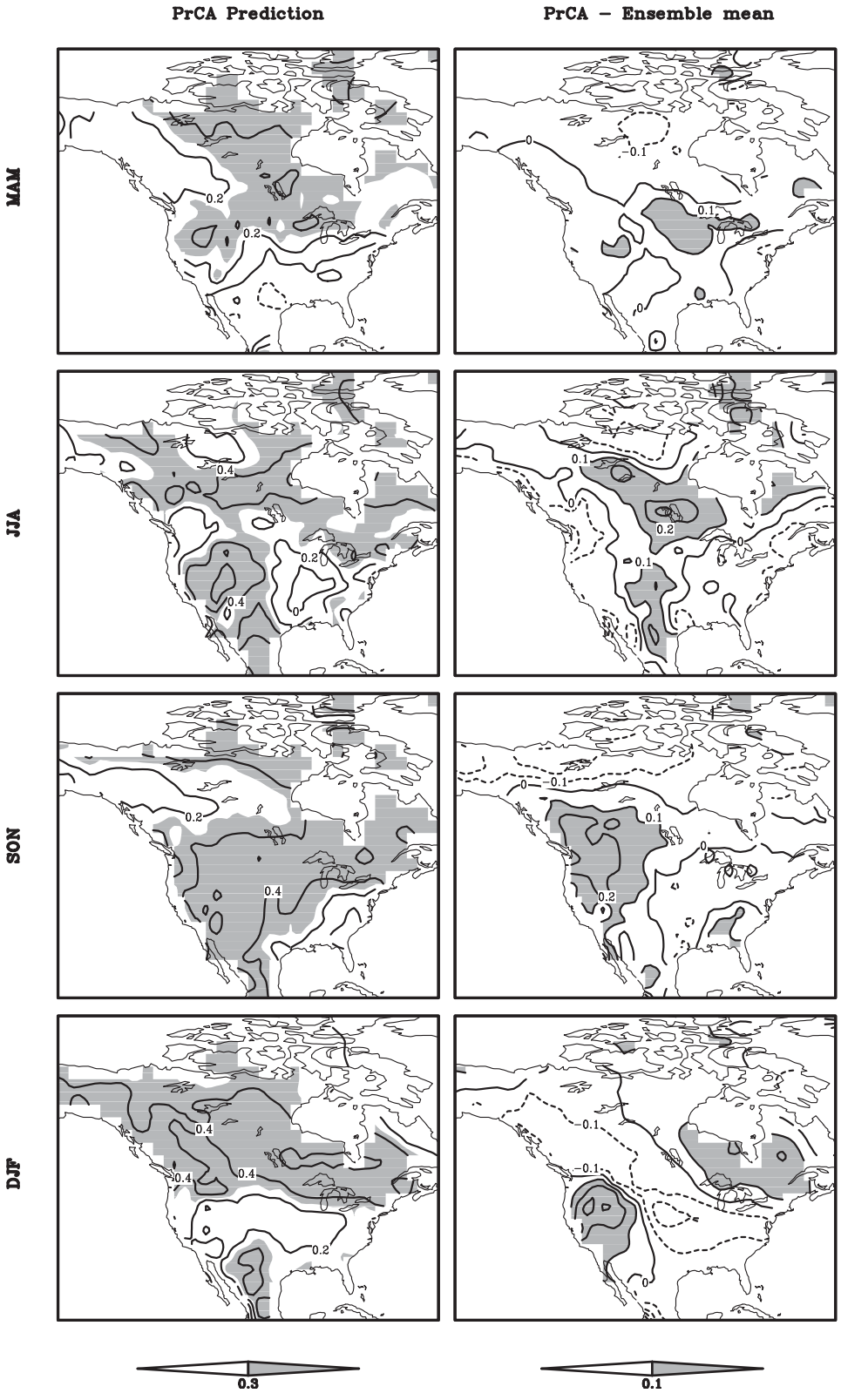

FIG. 11. Actual prediction correlation skill against observation by 1) the first two PrCA modes and 2) the difference between PrCA and ensemble mean.

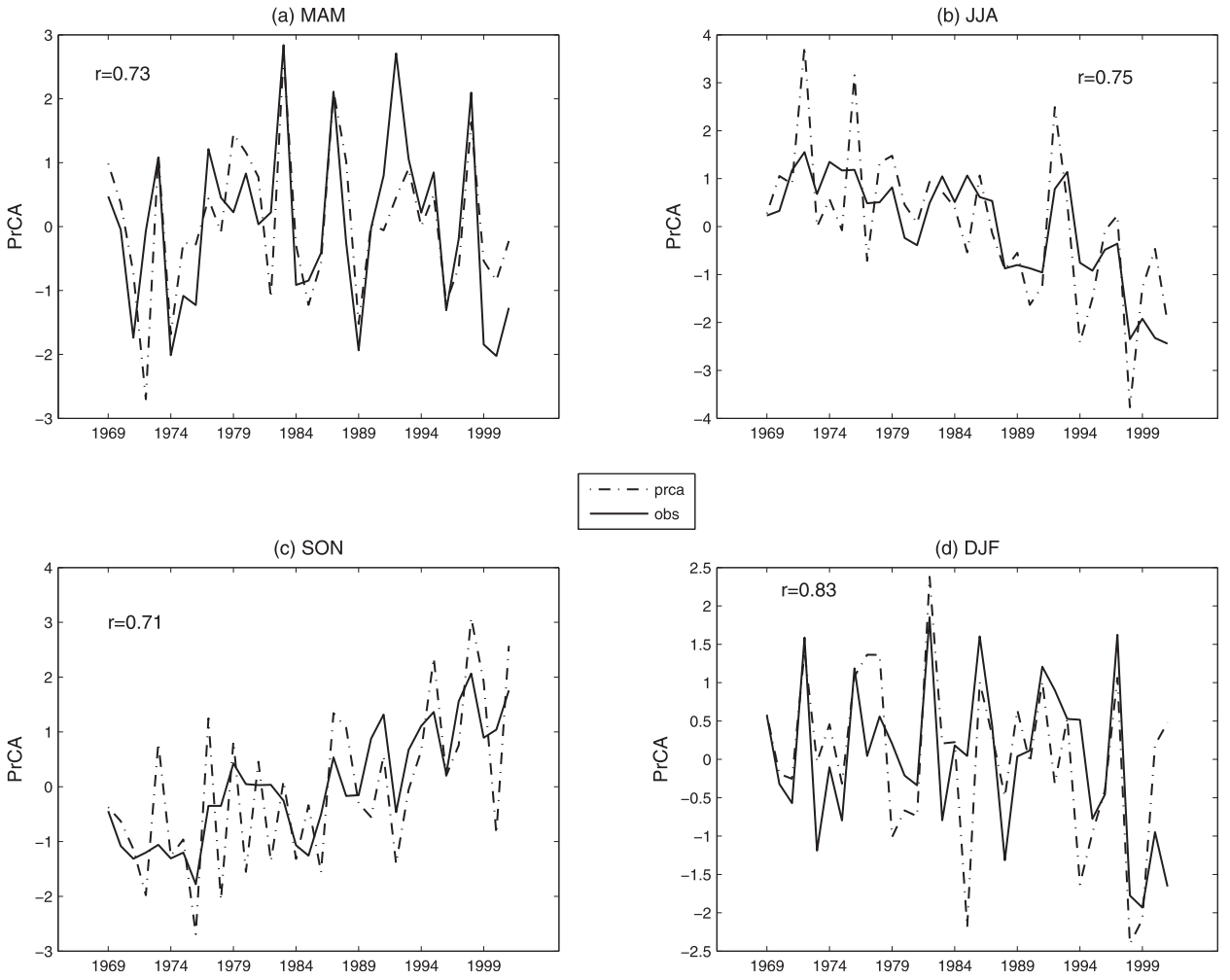

FIG. 12. The predicted time series of the first PrCA mode against the observation counterpart.

describes a linear relationship between initial conditions with predictions, which is statistically a specific case of the general definition of potential predictability. Intuitionally, the potential predictability should be quantified by the utility that can characterize the coherence between initial conditions and predictions. Thus, the SNR or perfect model skill actually underestimates the true potential predictability.

In this work, we applied the mutual information, an information-based metric, to study NA SAT potential predictability. As interpreted by (7), the mutual information well represents the coherence between initial conditions and predictions and thus can be expected to accurately measure the potential predictability. Using multiple ENSEMBLES ensemble products, we analyzed in detail the potential predictability by the MI-based measure and SNR-based measure for the NA SAT. Comparison between the two potential predictabilities reveals that the SNR often underestimates the potential predictability, and their significant difference often occurs at the region where the signal is relatively weak and noise dominates the SNR. It was also found that SNR can precisely measure the potential predictability if and only if the ensemble spread (the noise amplitude) is not varied with initial conditions. A large variation of ensemble spread with predictions will lead to severe underestimation of SNR to real predictability.

Another emphasis of this work is on the optimal decomposition of predictability and identifying the most predictable components. It was found that the first two most predictable components of NA SAT correspond to the predictability of the interannual variability and long-term trend. The amplitude of the two modes has geographical variations in different seasons. For the long-term trend, its most predictable component is spatially less varied, suggesting uniform warming over the whole NA domain, whereas for the interannual variability the most predictable component is spatially a dipole structure,

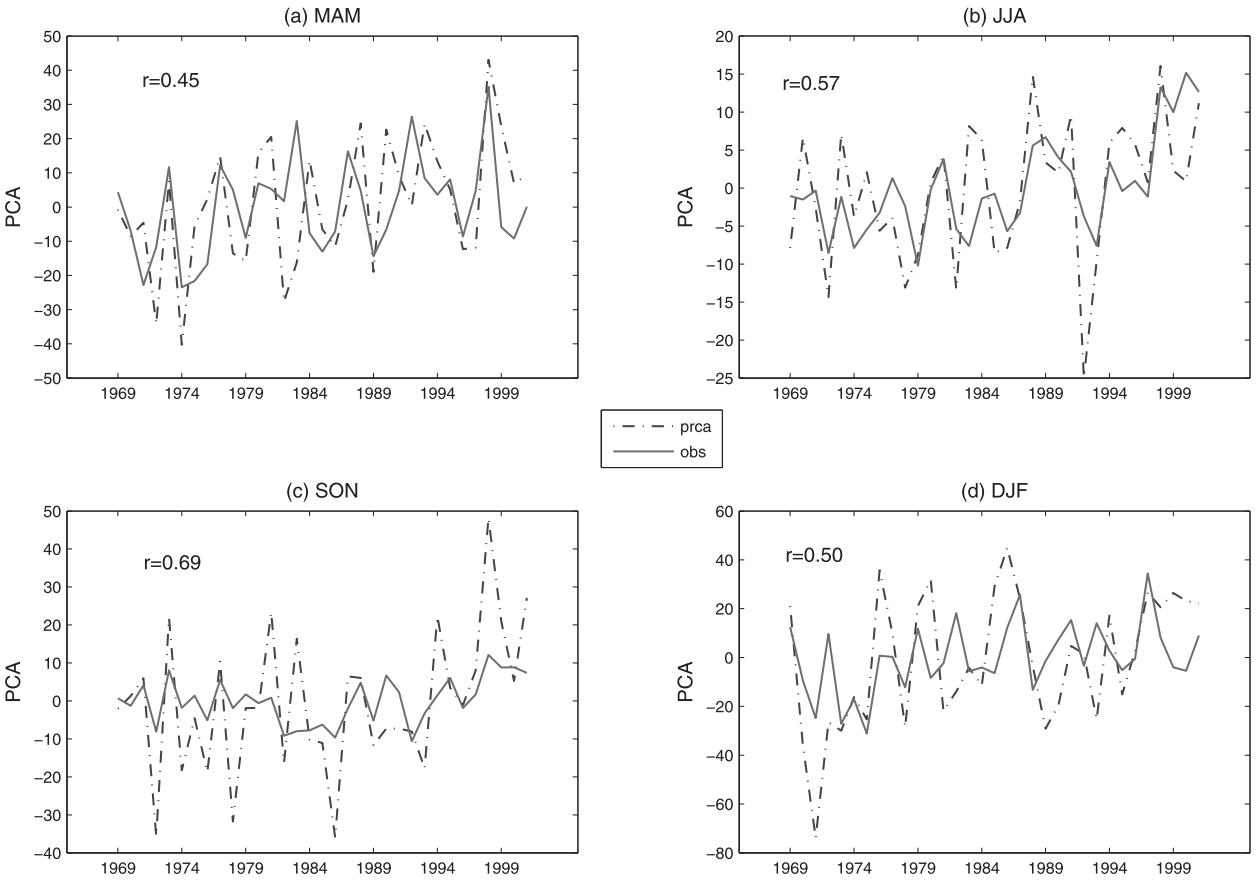

FIG. 13. As in Fig. 12, but for the prediction of the first PC.

characterized by the out-of-phase relationship between northwestern Canada and the southeastern United States. Such a dipole structure has a seasonal dependence, with the most significant pattern in spring and winter, followed by summer and fall. This is the most probably due to the seasonal variation of the strength of SST forcing. A further analysis of SST and 500-mb height reveals that the dipole structure of the most predictable pattern SAT is inherent to the most predictable patterns of SST and 500-mb height at the interannual time scale, which are characterized by ENSO and PNA. Thus, a physical interpretation of this dipole pattern can be proposed, namely that when El Niño occurs, a strong positive phase of PNA prevails in the middle to high latitude in spring and winter via the teleconnection, leading to warming in northwestern Canada and cooling in the southeastern United States. An opposite situation can be argued for La Niña.

It should be mentioned that all analyses are based on Gaussian assumptions in this work, which allow analytic solutions available for RE and MI. Fortunately, the Gaussian assumption is well held for climate prediction with a target of a monthly or seasonal mean. However, for shorter time scales the Gaussian assumption may be destroyed. In this case, the calculation of RE and MI is very challenging, which will be pursed in the near future. In addition, the number of ensemble samples is not very large in this analysis, which may result in some sampling errors in the comparison between MI-based and SNR-based measures.

Acknowledgments. This work was supported by Canada NSERC Discovery Grant, NSF of China (41276029) and National Key Basic Research Program of China (Grant 2013CB430302).

APPENDIX

The Equivalence of Perfect Correlation to the SNR

For seasonal climate prediction, it is well recognized that the prediction target $\mathbf{T}$[2] can be assumed to be the sum of two parts: a predictable component (signal S) and an unpredictable component or noise N, namely,

[2] Note that **T** is a vector here, for example, the SAT at a grid.

$$\mathbf{T} = S + N. \tag{A1}$$

Further assuming that the expectations of both signal and noise are 0, and the signal and noise are independent each other, we have

$$E(SN^{\mathrm{tr}}) = E(S^{\mathrm{tr}}N) = 0 \quad \text{and}$$
$$\mathrm{Var}(\mathbf{T}) = \mathrm{Var}(S) + \mathrm{Var}(N). \tag{A2}$$

The correlation R between S and T can be written as follows:

$$R^2 = \frac{E(S\mathbf{T}^{\mathrm{tr}})E(S\mathbf{T}^{\mathrm{tr}})}{E(SS^{\mathrm{tr}})E(\mathbf{T}\mathbf{T}^{\mathrm{tr}})} = \frac{E[S(S+N)^{\mathrm{tr}}]E[S(S+N)^{\mathrm{tr}}]}{E(SS^{\mathrm{tr}})E[(S+N)\,(S+N)^{\mathrm{tr}}]}$$
$$= \frac{\mathrm{E}(SS^{\mathrm{tr}})}{E(SS^{\mathrm{tr}}) + E(NN^{\mathrm{tr}})} = \frac{\mathrm{Var}(S)}{\mathrm{Var}(S) + \mathrm{Var}(N)} = \mathrm{STR}, \tag{A3}$$

where the superscript tr is the transpose. Thus $\sqrt{\mathrm{STR}}$ measures the potential correlation skill when the $\mathbf{T}$ is predicted by a perfect model. As the prediction target is completely composed of predictable signal components (i.e., the noise is zero), a perfect model will have a perfect correlation skill of 1. The difference between the potential correlation and the actual correlation quantifies the impact of model errors.

Furthermore, the $\sqrt{\mathrm{STR}}$ can also be interpreted as the correlation skill of a perfect prediction, which assumes that the observation is an arbitrary ensemble member. The perfect correlation skill ignores the imperfectness of model itself. To see this equality, we denote the ensemble mean μ as the prediction, thus the ''observation'' can be written by $\mu + \varepsilon$, where ε is a normally distributed white noise with the mean of zero and variance of σ_e^2.

The correlation of prediction against the observation can be written as follows:

$$\mathrm{AC}_{\mathrm{pef}} = \frac{E[\mu(\mu+\varepsilon)]}{\{E(\mu^2)E[(\mu+\varepsilon)^2]\}^{1/2}}$$
$$= \frac{E(\mu^2)}{\sqrt{E(\mu^2)E[\mu^2 + 2\mu\varepsilon + \varepsilon^2]}} = \frac{\sqrt{E(\mu^2)}}{\sqrt{E(\mu^2) + E(\varepsilon^2)}}. \tag{A4}$$

Comparison of (13) and (9) reveals the equality of $\mathrm{AC}_{\mathrm{pef}}$ and $\sqrt{\mathrm{STR}}$.

REFERENCES

Alessandri, A., A. Borrelli, A. Navarra, A. Arribas, M. Déqué, P. Rogel, and A. Weisheimer, 2011: Evaluation of probabilistic quality and value of the ENSEMBLES multimodel seasonal forecasts: Comparison with DEMETER. *Mon. Wea. Rev.,* **139,** 581–607.

Allen, M. R., and L. A. Smith, 1997: Optimal filtering in singular spectrum analysis. *Phys. Lett.,* **234,** 419–428.

Barnett, T. P., and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115,** 1825–1850.

Barnston, A. G., and T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J. Climate,* **9,** 2660–2697.

Cover, T. M., and J. A. Thomas, 2006: *Elements of Information Theory.* Wiley, 776 pp.

DelSole, T., 2004: Predictability and information theory. Part I: Measures of predictability. *J. Atmos. Sci.,* **61,** 2425–2440.

——, 2005: Predictability and information theory. Part II: Imperfect forecasts. *J. Atmos. Sci.,* **62,** 3368–3381.

——, and M. K. Tippett, 2007: Predictability: Recent insights from information theory. *Rev. Geophys.,* **45,** RG4002, doi:10.1029/2006RG000202.

Derome, J., and Coauthors, 2001: Seasonal predictions based on two dynamical models. *Atmos.–Ocean,* **39,** 485–501.

Fukunaga, K., 1990: *Introduction to Statistical Pattern Recognition.* 2nd ed. Academic Press, 592 pp.

Gebbie, G., I. Eisenman, A. Wittenberg, and E. Tziperman, 2007: Modulation of westerly wind bursts by sea surface temperature: A semistochastic feedback for ENSO. *J. Atmos. Sci.,* **64,** 3281–3295.

Hall, N. M., and J. Derome, 2000: Transience, nonlinearity, and eddy feedback in the remote response to El Niño. *J. Atmos. Sci.,* **57,** 3992–4007.

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

Jin, F., and B. J. Hoskins, 1995: The direct response to tropical heating in a baroclinic atmosphere. *J. Atmos. Sci.,* **52,** 307–319.

Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice, 2012: Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.,* **117,** D05127, doi:10.1029/2011JD017139.

Kang, I.-S., and J. Shukla, 2006: Dynamical seasonal prediction and predictability of the monsoon. *The Asian Monsoon,* B. Wang, Ed., 585–612.

Kleeman, R., 2002: Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.,* **59,** 2057–2072.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Willford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate prediction forecasts from multimodel superensemble. *Science,* **285,** 1548–1550.

——, ——, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate,* **13,** 4196–4216.

Kumar, A., 2007: On the interpretation and utility of skill information for seasonal climate predictions. *Mon. Wea. Rev.,* **135,** 1974–1984.

——, 2009: Finite samples and uncertainty estimates for skill measures for seasonal predictions. *Mon. Wea. Rev.,* **137,** 2622–2631.

——, and M. P. Hoerling, 1998: Annual cycle of Pacific–North American seasonal predictability associated with different phases of ENSO. *J. Climate,* **11,** 3295–3308.

——, and ——, 2003: The nature and causes for the delayed atmospheric response to El Niño. *J. Climate,* **16,** 1391–1403.

——, 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.

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

——, B. Jha, Q. Zhang, and L. Bounoua, 2007: A new methodology for estimating the unpredictable component of seasonal atmospheric variability. *J. Climate,* **20,** 3888–3901.

Mo, R. P., J. Fyfe, and J. Derome, 1998: Phase-locked and asymmetric correlations of the wintertime atmospheric patterns with the ENSO. *Atmos.–Ocean,* **36,** 213–239.

Palmer, T., and Coauthors, 2004: Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

Patz, J. A., and Coauthors, 2002: Regional warming and malaria resurgence. *Nature,* **420,** 627–628.

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.

——, ——, and W. Wang, 2011: An analysis of seasonal predictability in coupled model forecasts. *Climate Dyn.,* **36,** 637–648, doi:10.1007/s00382-009-0711-8.

Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate,* **19,** 3279–3293.

Rowell, D., 1998: Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J. Climate,* **11,** 109–120.

Schlosser, C. A., and B. P. Kirtman, 2005: Predictable skill and its association to sea surface temperature variations in an ensemble climate simulation. *J. Geophys. Res.,* **110,** D19107, doi:10.1029/2005JD005835.

Schneider, T., and S. M. Griffies, 1999: A conceptual framework for predictability studies. *J. Climate,* **12,** 3133–3155.

Shabbar, A., and A. G. Barnston, 1996: Skill of seasonal climate forecasts in Canada using canonical correlation analysis. *Mon. Wea. Rev.,* **124,** 2370–2385.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science,* **282,** 728–731.

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

Smith, T. M., and R. E. Livezey, 1999: GCM systematic error correction and specification of the seasonal mean Pacific–North America region atmosphere from global SSTs. *J. Climate,* **12,** 273–288.

Sutton, R. T., S. P. Jewson, and D. P. Rowell, 2000: The elements of climate variability in the tropical Atlantic region. *J. Climate,* **13,** 3261–3284.

Tang, Y., R. Kleeman, and A. Moore, 2005: Reliability of ENSO dynamical predictions. *J. Atmos. Sci.,* **62,** 1770–1791.

——, ——, and ——, 2008: Comparison of information-based measures of forecast uncertainty in ensemble ENSO prediction. *J. Climate,* **21,** 230–247.

Trenberth, K. E., G. W. Branstrator, 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.

Vautard, R., G. Plaut, R. Wang, and G. Brunet, 1999: Seasonal prediction of North American surface air temperatures using space–time principal components. *J. Climate,* **12,** 380–394.

Venzke, S., M. R. Allen, R. T. Sutton, and D. P. Rowell, 1999: The atmospheric response over the North Atlantic to decadal changes in sea surface temperature. *J. Climate,* **12,** 2562–2584.

Weisheimer, A., and Coauthors, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.,* **36,** L21711, doi:10.1029/2009GL040896.

Yan, X., and Y. Tang, 2012: An analysis of multi-model ensemble for seasonal climate predictions, *Quart. J. Roy. Meteor. Soc.,* **139,** 1179–1198, doi:10.1002/qj.2020.

Yang, D., Y. Tang, Y. Zhang, and X. Yang, 2012: Information-based potential predictability of the Asian summer monsoon in a coupled model. *J. Geophys. Res.,* **117,** D03119, doi:10.1029/2011JD016775.