Information-based potential predictability of the Asian summer monsoon in a coupled model

Dejian Yang,[1,2] Youmin Tang,[2] Yaocun Zhang,[1] and Xiuqun Yang[1]

Received 26 August 2011; revised 30 November 2011; accepted 15 December 2011; published 15 February 2012.

[1] In this study, we applied two information-based measures, relative entropy (RE) and mutual information (MI), to explore the potential predictability of the Asian summer monsoon (ASM) at seasonal time scales using the hindcasts of the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS). Several important issues related to ASM predictability were addressed, including the dominant precursors of forecast skill and the degree of confidence that can be placed in an individual forecast. We found that the MI-based average potential predictability can explain, to a large extent, the variation in overall prediction skill, especially anomaly correlation skill. Compared with the conventional signal-to-noise ratio approach, the MI-based measures can characterize more potential prediction utility. As a potential predictability measure for an individual prediction, RE has a good relationship with C, the contribution of an individual forecast to overall anomaly correlation skill, and a poor relationship with the absolute error (AE). Further analyses revealed that RE is highly related to sea surface temperature (SST) and SST-ASM correlation patterns in the model resemble the typical El Niño-Southern Oscillation (ENSO) structure. The different impacts of ENSO on the South Asian summer monsoon and the East Asian summer monsoon, the two main components of the ASM, well explain the different potential predictability features in the two regions, in particular, in terms of their interannual variability. Thus, ENSO is the main source of ASM seasonal predictability.

Citation: 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.

1. Introduction

[2] The Asian summer monsoon (ASM) is a major climate system that is important in forming and maintaining the unique Asian monsoon climate. Its interannual fluctuation and variability is often associated with floods, droughts, and other natural disasters during summer and further critically influences economies and societies in this region. Thus, improving predictions of the ASM climate several months in advance is extremely important for decision making and risk management [e.g., *Webster et al.*, 1998] (also see the reviews given by *Wang* [2006]).

[3] The study of monsoon prediction and predictability has a long history. In the early 1980s, *Charney and Shukla* [1981] found that a large part of monsoon variability is due to external (boundary) forcings such as anomalous sea surface temperature (SST), albedo, and soil moisture and could be predictable at a longer range than the weather forecast scale. With the availability of high-quality observational data sets and state-of-the-art climate models, there have been significant advances in recent years in understanding the basic external and internal processes that determine the temporal and spatial structures of monsoon variability [e.g., *Webster and Yang*, 1992; *Palmer*, 1994; *Ju and Slingo*, 1995; *Webster et al.*, 1998; *Wang et al.*, 2000; *Krishnamurthy and Shukla*, 2000; *Sperber et al.*, 2001; *Molteni et al.*, 2003; *Wang et al.*, 2004; *Goswami et al.*, 2006; *Zhou et al.*, 2009; *Wang et al.*, 2008a; *Yang et al.*, 2008; *Zhou and Zou*, 2010] (also see the reviews given by *Wang* [2006]) and in developing strategies to predict monsoon variability such as turning to tier-1 coupled climate model predictions from tier-2 forced atmospheric model predictions and using multiple model ensembles instead of single model ensembles [e.g., *Wang et al.*, 2005; *Wu and Kirtman*, 2005; *Wu et al.*, 2006; *Krishnamurti et al.*, 2006; *Kang and Shukla*, 2006]. In terms of the potential predictability estimate for the ASM, most previous research has been confined to exploring the (average) signal-to-noise ratio in a perfect model framework, which is thought to be able to measure the relative contribution of the potentially predictable external variability to monsoon variability [e.g., *Goswami*, 1998; *Cherchi and Navarra*, 2003; *Kang et al.*, 2004; *Kang and Shukla*, 2006; *S.-S. Lee et al.*, 2011]. In general, climate predictability studies can be roughly classified into two categories: diagnostic

[1]School of Atmospheric Sciences, Nanjing University, Nanjing, China.
[2]Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada.

0148-0227/12/2011JD016775

predictability studies and prognostic predictability studies. The former usually give a summary and an overall estimate of the extent to which a climate system is predictable and thus focus on the average predictability [e.g., *Rowell*, 1998; *Yang et al.*, 1998; *Kang and Shukla*, 2006], whereas the latter consider the variation of the predictability from one particular prediction to another and focus on forecasting the predictability of a certain prediction [e.g., *Palmer*, 2000; *Tang et al.*, 2007]. ASM predictability studies have so far mainly addressed the diagnostic average predictability. A detailed exploration of ASM statistical predictability, especially its prognostic predictability, should therefore be undertaken.

[4] An important aspect of a prognostic predictability study is an exploration of the ability of potential predictability measures in estimating a priori individual prediction skills. Specifically, if a relationship can be identified between a potential predictability measure and an actual model prediction skill, a practical means for estimating the confidence level of an individual prediction of the given model could be developed. For example, in numerical weather prediction, ensemble spread is usually used as an indicator of individual forecast skill [e.g., *Buizza and Palmer*, 1998]. However, few connections have been found between this spread and forecast skill in several dynamical models [e.g., *Kumar et al.*, 2000; *Tang et al.*, 2005, 2008b]. Instead, in some studies, an alternate indicator of forecast skill is the leading eigenmode amplitude of the forecast initial conditions [*Kleeman and Moore*, 1999; *Tang et al.*, 2005], which essentially represents the contribution of persistence to prediction skill [*von Storch and Xu*, 1990; *von Storch and Baumhefner*, 1991].

[5] Recently, a new theoretical framework based on information theory has been developed and applied to evaluate potential climate predictability systematically [e.g., *Leung and North*, 1990; *Schneider and Griffies*, 1999; *Kleeman*, 2002; *DelSole*, 2004, 2005; *Tang et al.*, 2005, 2007, 2008a; *DelSole and Tippett*, 2007]. It has been argued that relative entropy (RE), defined as the difference between the prediction probability density functions (PDFs) and the climatological PDF, can well explain why the two reliability measures discussed above are central to predictability studies [e.g., *Kleeman*, 2002; *Tang et al.*, 2005]. The RE is a prognostic measure of individual predictability, whereas mutual information (MI), the average of RE over all initial conditions, is a diagnostic measure of average predictability.

[6] In this study, we evaluate both the prognostic and diagnostic potential predictability of the ASM using these two information-based measures. This evaluation is implemented through an analysis of ensemble hindcasts of the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS). Our emphasis is on a comparison between potential and actual measures of predictability. Herein, we define potential predictability as the prediction skill of a "perfect" forecast system (model), which is evaluated without using observations whereas the actual prediction skill is the prediction skill evaluated using observations. The predictability, used as a general expression, means both the potential predictability and actual prediction skill in this study.

[7] The remainder of the paper is structured as follows: Section 2 briefly describes the model and data used. In section 3, two information-based potential predictability measures, RE and MI, are introduced and interpreted. Section 4 investigates and compares overall potential predictability and actual prediction skill. Section 5 explores the relationship between relative entropy and individual model forecast skills. Section 6 further explores the source of predictability with a focus on the dominant role of SST forcing. A summary and discussion is provided in section 7.

2. Model and Data

2.1. Model

[8] The NCEP CFS is a fully coupled ocean-land-atmosphere dynamical seasonal prediction system, which became the NCEP operational model in August, 2004 [*Saha et al.*, 2006]. The atmospheric component of the CFS is a coarser-resolution version of the NCEP operational Global Forecast system (GFS) [*Moorthi et al.*, 2001]. This component uses a spectral triangular truncation of 62 waves in the horizontal direction and finite differencing in 64 vertical sigma layers, with the top of the model at 0.2 hPa. The ocean component is the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version 3 (MOM3) [*Pacanowski and Griffies*, 1998]. MOM3 has a quasi-global domain from 74°S to 64°N, with a zonal resolution of 1° and a meridional resolution of 1/3° between 10°S and 10°N, gradually increasing to become 1° fixedly poleward at 30°S and 30°N. There are 40 layers in the vertical direction, with 27 layers in the upper 400 m, and the bottom depth is approximately 4.5 km. The vertical resolution is 10 m from the surface to a depth of 240 m and gradually increases to approximately 511 m at the bottom. For the land model, the two-layer model described by *Mahrt and Pan* [1984] is used.

[9] The atmospheric and oceanic components are coupled without flux adjustment. The two components exchange daily averaged quantities like heat and momentum fluxes once a day. Full atmosphere–ocean interaction is confined to 65°S to 50°N. Between 74° and 65°S and between 64° and 50°N, the weighted average of the observed climatology and modeled SST is used in the way that the weights vary linearly with latitude such that the SSTs at 74°S and 64°N equal observed climatologies and the SSTs at 65°S and 50°N equal modeled climatologies. Poleward of 74°S and 64°N, the SSTs are taken from the observed climatology (see *Saha et al.* [2006] for details).

2.2. Data

[10] In this work, we analyze the CFS hindcasts starting in all 12 calendar months during the period from 1981 to 2008 (available from http://www.cdacb.in/html/spim_data.aspx). Each prediction is a nine-month integration and has an ensemble of 15 members generated from 15 different initial conditions. The atmospheric initial conditions were taken from the NCEP/Department of Energy (DOE) Atmospheric Model Intercomparison Project II Reanalysis (R2) [*Kanamitsu et al.*, 2002], and the ocean initial conditions were from the NCEP Global Ocean Data Assimilation System (GODAS). Following *Yang et al.* [2008], for the target season JJA (June–July–August), the ensemble prediction using the initial conditions on days 9–13, 19–23, the last two days of May and the first three days of June is called the 0-month lead (LM0) prediction. Similarly, the ensemble prediction using the initial conditions on days 9–13, 19–23,

and the last two days of April and the first three days of May is called the 1-month lead (LM1) prediction. Predictions for longer leads are similarly named. The longest lead prediction is the 6-month lead (LM6) prediction, starting on days 9–13, 19–23, and the last two days of November and the first three days of December in the previous year. In total, there are 7 lead times used in this study: LM0 to LM6.

[11] The verification data sets used in this study include the Climate Prediction Center merged analysis of precipitation (CMAP) [*Xie and Arkin*, 1996], winds from the NCEP/DOE R2 [*Kanamitsu et al.*, 2002], and the NOAA optimally interpolated (OI) SST analysis [*Reynolds et al.*, 2002].

3. Information-Based Measures of Potential Predictability

3.1. Relative Entropy

[12] 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 that is the conditional distribution $p(\nu|i)$ given the initial condition i. The climatological distribution is also the unconditional distribution, and we thus have

$$p(\nu) = \int p(\nu|i)\,p(i)di, \tag{1}$$

where $p(i)$ is the probability distribution for the initial condition i. The extent to which the forecast distribution $p(\nu|i)$ differs from the climatological distribution $p(\nu)$ is an indication of potential predictability. Various statistical tests are commonly used to examine the differences between the two distributions, e.g., the t- and F tests for Gaussian distributions and the Kolmogorov-Smirnov test for general distributions [*Anderson and Stern*, 1996; *Sardeshmukh et al.*, 2000]. The relative entropy, RE, or the Kullback–Leibler distance, is a quantitative measure of the difference between two distributions from information theory [*Cover and Thomas*, 1991]. In the context of predictability, RE is defined as

$$RE = \int p(\nu|i)\ln\frac{p(\nu|i)}{p(\nu)}d\nu. \tag{2}$$

In terms of information theory, the quantity RE measures the informational inefficiency of using the climatological distribution $p(\nu)$ rather than the forecast distribution $p(\nu|i)$, and $RE \geq 0$, with the equality holding if and only if $p(\nu|i) = p(\nu)$ [*Cover and Thomas*, 1991]. In Bayesian terminology, the climatological distribution is a *prior* distribution that can be usually derived from long-term historical observations. An ensemble prediction augments this prior information, and the additional information measured by RE is a natural measure of the utility or usefulness of this prediction and thus implies the potential predictability. In practice, $p(\nu|i)$ and $p(\nu)$ can be estimated directly from samples or alternatively approximated using kernel density estimation.

[13] If the two distributions are Gaussian, equation (2) reduces to an expression including the ensemble (or forecast, i.e., conditional) variance $\sigma^2_{\nu|i}$, the climatological variance σ^2_ν, and the difference $\mu_{\nu|i} - \mu_\nu$ between the ensemble (or forecast, i.e., conditional) mean and the climatological mean [*Kleeman*, 2002]:

$$RE = \frac{1}{2}\left[\ln\left(\frac{\sigma^2_\nu}{\sigma^2_{\nu|i}}\right) + \frac{\sigma^2_{\nu|i}}{\sigma^2_\nu} + \frac{(\mu_{\nu|i} - \mu_\nu)^2}{\sigma^2_\nu} - 1\right]. \tag{3}$$

The first two terms on the right-hand side of equation (3) are determined by the prediction and climatological variances and represent the contribution of the dispersion or spread of the ensemble to the RE. The third term on the right-hand side of equation (3) is governed by the ensemble mean and measures the contribution of the signal size to the RE. In addition, $\sigma^2_{\nu|i}/\sigma^2_\nu$ and $(\mu_{\nu|i} - \mu_\nu)^2/\sigma^2_\nu$ can be respectively understood as the normalized ensemble variance (NEV) and the square of the normalized ensemble mean shift (SNEMS) [*Tippett et al.*, 2004].

[14] It has been reported that the assumption of a Gaussian distribution is a good approximation for seasonal mean variables [e.g., *Schneider and Griffies*, 1999; *Kleeman*, 2002; *Tang et al.*, 2005], so we calculate the RE using equation (3) in this study. Here, the ensemble mean $\mu_{\nu|i}$ and the ensemble variance $\sigma^2_{\nu|i}$ are estimated from the 15 members of the ensemble. The climatological mean μ_ν and the climatological variance σ^2_ν are estimated using a lead-time-dependent approach similar to that used by *Peng et al.* [2011]; i.e., each lead time has a respective climatological mean and variance. For a particular lead time, the climatological mean and variance are estimated from all the members at this lead time (sample size: 28×15).

[15] In this study, the model prediction is the anomaly with respect to the model climatology, so the climatological mean μ_ν in equation (3) actually approaches zero. Additionally, the observed anomaly with respect to the observed climatology is used in forecast verification.

3.2. Mutual Information

[16] RE is a predictability measure for individual predictions. The average of RE over all initial conditions reflects the average predictability and has been proven to be equal to the mutual information (MI), another quantity from information theory [*DelSole*, 2004]. In the context of predictability, MI is defined as [*DelSole*, 2004]

$$MI = \iint p(\nu, i)\ln\left[\frac{p(\nu, i)}{p(\nu)p(i)}\right]d\nu di, \tag{4}$$

where $p(\nu, i)$ is the joint probability distribution between ν and i. The MI measures the statistical dependence between ν and i and vanishes when ν and i are independent ($p(\nu, i) = p(\nu)p(i)$). The equality of the MI and the average RE indicates that predictability can be measured in two equivalent ways: by the difference between the forecast and climatological distributions and by the degree of statistical dependence between the initial condition i and the future state ν [*DelSole and Tippett*, 2007]. If the future state ν is on average unpredictable, individual forecasts should have probability distributions identical to the climatological distribution, i.e., $p(\nu|i) = p(\nu)$ and $RE = 0$ for all predictions. This is equivalent to independence between i and ν. Therefore, independence indicates unpredictability and dependence implies predictability. The MI is invariant with respect to nonlinear, invertible (nonsingular) transformations of state [*DelSole*,

2004]. Thus, the MI between ν and i equals the MI between ν and $\mu_{\nu|i}$. The latter is probably more straightforward in understanding MI-based predictability because the dependence between ν and ensemble mean $\mu_{\nu|i}$ can be interpreted as the dependence between observation (ν) and prediction ($\mu_{\nu|i}$) under the assumption of a perfect model.

[17] When the forecast and climatological distributions are Gaussian, the MI can be expressed as [*DelSole and Tippett*, 2007] (also see Appendix A)

$$MI = \frac{1}{2}\left(\ln \sigma_{\nu}^2 - \left\langle \ln \sigma_{\nu|i}^2 \right\rangle\right). \tag{5}$$

Equation (5) is the formula we use to calculate MI in this study. *Joe* [1989] and *DelSole* [2005] showed that the transformations $\sqrt{1-\exp(-2MI)}$ and $1-\exp(-2MI)$ produce "potential" skill scores, which exhibit proper limiting behavior: they have values between 0 and 1, and the minimum (maximum) value 0 (1) occurs when MI vanishes (approaches infinity). Here, "potential" indicates that they are perfect model measures. In this study, we use the two potential skills to represent the MI. Furthermore, if the forecast and climatological distributions are Gaussian, and the forecast variance is constant, the two potential skills above respectively reduce to two other conventional potential skills: the potential anomaly correlation (AC_p) and the potential mean square skill score ($MSSS_p$) [e.g., *DelSole*, 2004; *DelSole and Tippett*, 2007].

[18] The actual AC and MSSS make use of actual observations and quantify the actual ability of the model to reproduce reality, as expressed by [e.g., *Murphy and Epstein*, 1989]

$$AC_a = \frac{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^p - \overline{x^p}\right)\left(x_j^o - \overline{x^o}\right)}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^p - \overline{x^p}\right)^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^o - \overline{x^o}\right)^2}}, \tag{6}$$

and

$$MSSS_a = 1 - \frac{MSE}{MSE_{c\,\lim}} = 1 - \frac{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^p - x_j^o\right)^2}{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^o - \overline{x^o}\right)^2}, \tag{7}$$

where x_j^p denotes the forecast using ensemble mean, x_j^o denotes the corresponding observation, $\overline{x^p}$ ($\overline{x^o}$) denotes the mean of the forecasts (observations), and N is the number of samples.

[19] The potential skills AC_p and $MSSS_p$ have the same formulas as AC_a and $MSSS_a$ but use one arbitrary realization of the forecast distribution as the hypothetical observation instead of the real observation and are thus often called perfect model skills [e.g., *Rowell*, 1998] or theoretical skills [e.g., *Newman et al.*, 2003; *Kang and Shukla*, 2006]. It has been argued that these two conventional potential skill measures are functions of the so-called signal-to-noise ratio (SNR) [e.g., *Sardeshmukh et al.*, 2000; *Newman et al.*, 2003] (also see Appendix A). Thus, the two conventional potential skills are also often called SNR-based potential skills. If the model is perfect, that is, the actual observation can be considered to be one arbitrary realization of the model forecast distribution, AC_a should be equal to AC_p within the tolerances of the sampling and observational errors.

[20] If both the forecast and climatological distributions are Gaussian, and the forecast variance (ensemble spread) is constant, the potential skills based on the MI reduce to the conventional ones [*DelSole*, 2004; *DelSole and Tippett*, 2007] (also see Appendix A), as expressed by

$$AC_p = \sqrt{1-\exp(-2MI)}, \tag{8}$$

and

$$MSSS_p = {AC_p}^2 = 1-\exp(-2MI). \tag{9}$$

In the following discussions, the skills based on either the MI or the SNR that do not make use of actual observations are formally called *potential predictability*, whereas the skills that make use of actual observations through equations (6) and (7) are called *actual skills*. In the next section, we explore the ability of the potential predictability in explaining the variation of actual skills. To avoid overfitting, we calculated the actual AC in a cross-validated mode using a leave-three-out method, i.e., the anomaly is obtained using the climatology that is calculated excluding the data for the current year and the two adjacent years. However, we did not find a substantial reduction of AC after the cross-validation scheme was used, compared with the skill scores without cross-validation.

4. Overall Potential Predictability and Actual Skill

[21] In this section, we first explore the variation of RE from year to year and then examine in detail the MI-based potential predictability and actual skill as functions of lead time.

[22] For simplicity, we only examine the prediction of two monsoon circulation indices at LM0. The first index is the Webster-Yang monsoon index (WYMI) [*Webster and Yang*, 1992], defined as the vertical zonal wind shear between 850 and 200 hPa (u850–u200) averaged over the South Asian region (0°–20°N, 40°–110°E), which measures the broad-scale South Asian summer monsoon (SASM) circulation anomalies. The second index is the East Asian summer monsoon index (EASMI), which is a reversed *Wang and Fan* [1999] index and is defined by the 850 hPa zonal winds averaged over a given domain (22.5°–32.5°N, 110°–140°E) minus that averaged over an offset domain (5°–15°N, 90°–130°E) [*Wang et al.*, 2008b; *Wu et al.*, 2009]. Note that the East Asian summer monsoon (EASM) system discussed here actually consists of two traditional monsoon subsystems: the East Asian summer monsoon and the Western North Pacific summer monsoon. The SASM and the EASM are the two major components of the grand ASM. They behave quite independently, although they share several fundamental characteristics [e.g., *Yang and Lau*, 2006; *Lau and Wang*, 2006].

[23] Figure 1 shows the RE for the WYMI and EASMI predictions at LM0 from 1981 to 2008, where large variations in RE are evident from one particular summer to another for both indices, indicating the large interannual

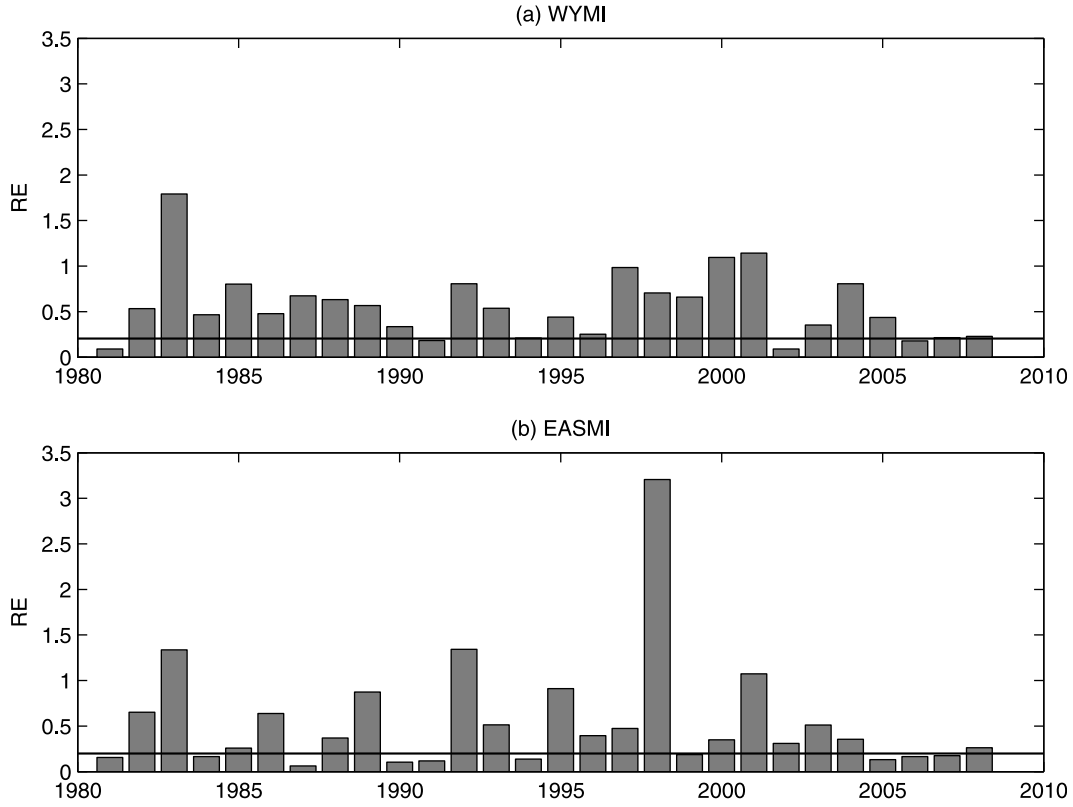

Figure 1. Relative entropy RE as a function of year from 1981 to 2008 at the 0-month lead time (LM0) for the JJA predictions of (a) the WYMI and (b) the EASMI. Solid lines in Figures 1a and 1b show the 95% confidence levels using the Monte Carlo test.

variability of the ASM potential predictability. When the forecast and climatological distributions are identical, the RE is equal to zero, and there is no extra information from the prediction, according to its definition. However, although the two distributions are identical, sampling error introduced due to the finite ensemble size may lead to a nonzero RE. Therefore, the statistical significance level should exceed the extent of uncertainty due to the finite ensemble size. We quantified the extent of uncertainty with a Monte Carlo method used by *Tang et al.* [2007] and *Tippett et al.* [2004]. Fifteen samples were randomly selected from the entire set of 15×28 samples to form a new random forecast distribution, which was used to calculate its RE with respect to the climatological distribution according to equation (3). This process was repeated 1,000 times, and the sorted results indicated the likelihood that the RE exceeds a given value by chance. Values above the 95th percentile were considered significant. Among the 28 predictions, 22 predictions (78%) of the WYMI had a significant RE, and 18 predictions (64%) of the EASMI had a significant RE.

[24] An interesting feature in Figure 1 is that the RE has relatively very large values only in a few years (e.g., the RE greater than 1.0 only in 1983, 1997, 2000, and 2001 for the WYMI and in 1983, 1992, 1998, and 2001 for the EASMI), and many predictions have relatively small RE values. For the years with very large RE values, the ASM predictions, denoted by the two circulation indices, have much additional information beyond the prior information given by the climatological distribution, thereby having high potential prediction utility. The suitability of the RE for measuring predictability is further demonstrated by the grouped AC_a values, classified by the median value RE_0 (as the threshold) of all RE values. Given the small sample size (28) at each lead time, we pooled the predictions for all the lead times to group them for the AC_a skills. We found that the AC_a skills were 0.52 (0.14) and 0.72 (0.17) for $RE > RE_0 (RE < RE_0)$ for the WYMI and the EASMI, respectively. Apparently, the predictions with larger RE values have much higher correlation skills than those with smaller RE values. A bootstrap resampling test showed that at a 95% confidence level, the grouped AC_a skills, classified by $RE > RE_0 (RE < RE_0)$, are statistically significantly larger (smaller) than the overall AC_a skills obtained using all the samples (0.39 for the WYMI and 0.58 for the EASMI). Previous work on monsoon predictability has mainly focused on the overall predictability for the entire period [e.g., *Yang et al.*, 2008; *S.-S. Lee et al.*, 2011]. Here, the RE can quantify the predictability of individual predictions and capture its temporal variation, which have important practical significance. A comprehensive analysis of the usefulness of RE for the a priori estimation of individual prediction skill is provided in section 5.

[25] Further comparison revealed that all the years with large RE values were accompanied by El Niño-Southern Oscillation (ENSO) events (warm events: 1983, 1992, 1997, and 1998; cold events: 2000 and 2001), implying that the ENSO has a significant impact on the seasonal predictability of the ASM. Furthermore, among the years with large RE

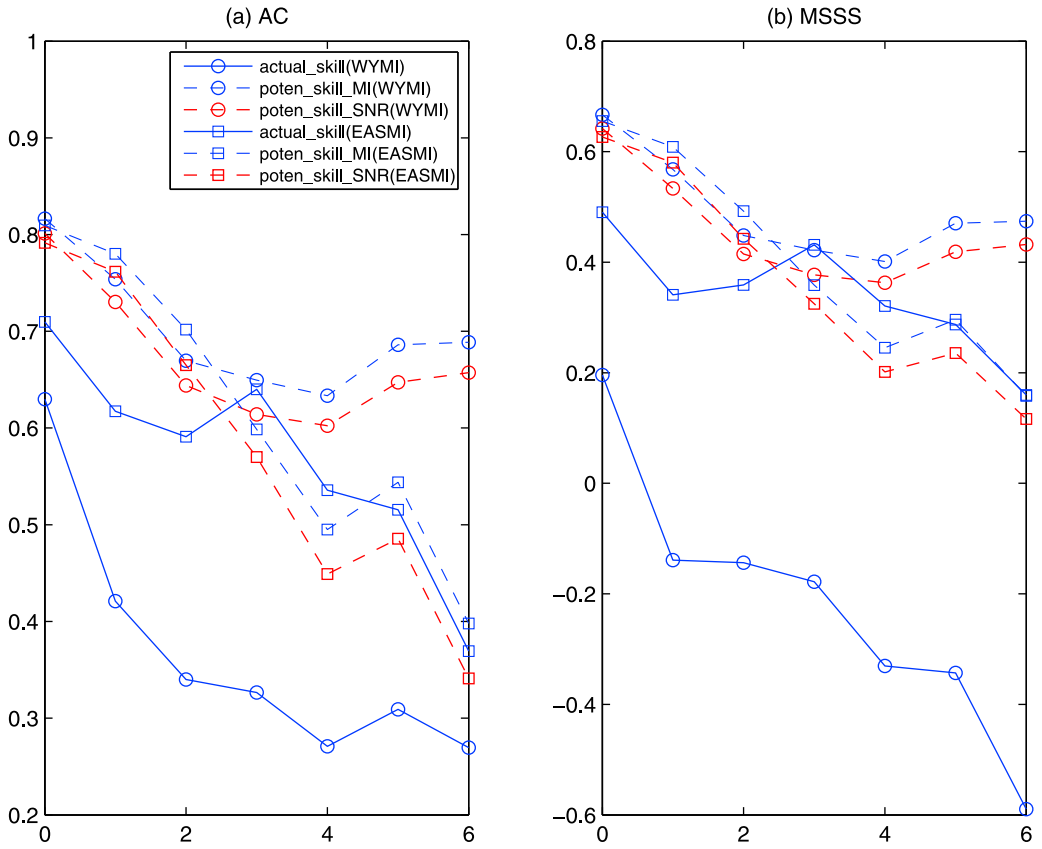

Figure 2. Potential predictability and actual skill of (a) the AC and (b) the MSSS. The actual skill, the MI-based potential predictability, and the SNR-based potential predictability are denoted by blue solid lines, blue dashed lines, and red dashed lines, respectively.

values, two are ENSO-developing years (1997 and 2000) and two are ENSO-decaying years (1983 and 2001) for the WYMI, whereas for the EASMI, all four years (1983, 1992, 1998, and 2001) are ENSO-decaying years. This pattern suggests that the impacts of the ENSO on ASM seasonal predictability and their mechanisms differ for the SASM and the EASM. A more detailed discussion on the role of the SST in ASM seasonal predictability is presented in section 6.

[26] Figure 2 shows the MI-based potential predictability and actual skill as functions of lead time for the WYMI and the EASMI. The most prominent feature here is the decrease of skill with lead time, which is in good agreement with the general conclusion that predictability declines with lead time in chaotic or stochastic systems. The temporal evolution is not strictly monotonic for either the potential or actual skills, and there are some fluctuations at longer lead times. For example, the potential AC has local minimum values at LM4 for both the WYMI and the EASMI, and the actual AC appears as a local minimum at LM4 for the WYMI and at LM2 for the EASMI. This non-monotonic evolution in skills is probably due to two causes. First, the finite sample size used for skill evaluation may cause sampling errors. Second, the non-monotonic nature may be related to the underlying physics, for example, the "Spring Prediction Barrier" [e.g., *Wang et al.*, 2008a; *Webster and Yang*, 1992].

[27] The SNR-based potential predictability, defined in Appendix A, is also shown in Figure 2, where the SNR was estimated using the method proposed by *Rowell* [1998] (hereinafter referred to as R98). A prominent merit of the R98 method is its bias correction, which takes into account the impact of finite sample size on the estimation of signal variance. As shown in Figure 2, the MI-based potential predictability is higher than the SNR-based potential predictability, as indicated by the theoretical relationships of equations (A11) and (A12). As argued in Appendix A, the MI-based skill and SNR-based skill are identical if the prediction and climatological distributions are Gaussian, and the prediction variance is constant. Considering the fact that Gaussian distributions are generally good approximations for seasonal climate variables, the difference between the MI- and SNR- based skills shown in Figure 2 should be the consequence of the varied prediction variance (ensemble spread). However, their difference here is not large, suggesting that the prediction variance does not differ much from prediction to prediction for the two monsoon indices. To examine this difference, we performed a significance test: for a selected lead time, the median value of all 28 forecast variances was chosen as the reference variance, and the F test was then used to determine whether or not each prediction variance was statistically significantly different from the reference variance. The results show that at a 95% confidence level, only 4% and 6%, respectively, of the WYMI and EASMI predictions have significantly different variances.

[28] One interesting question arising here is how to understand the MI-SNR discrepancy when the prediction variance

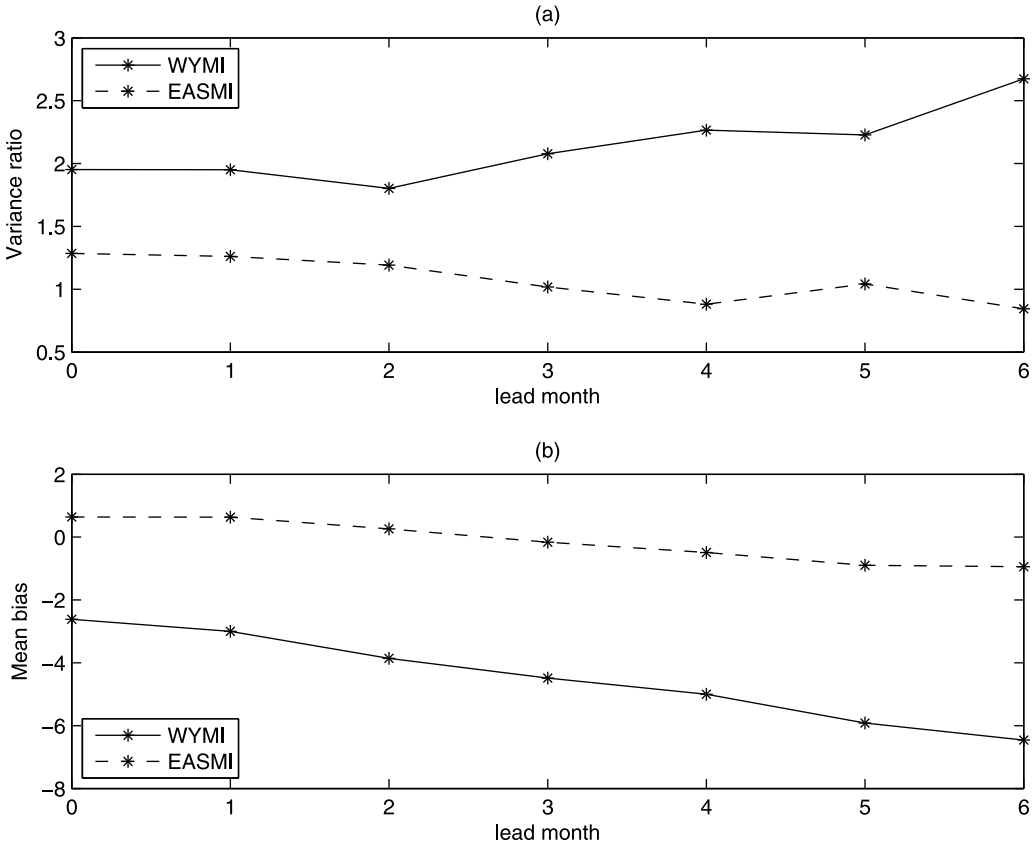

Figure 3. (a) Ratio of the modeled climatological variance to the observed counterpart and (b) bias (difference, unit: m/s) between the modeled climatological mean and the observed counterpart.

has a significant variation, as suggested by equations (A11) and (A12). As discussed in section 3.2, the MI-based potential predictability measures the statistical dependence, linear or nonlinear, between the ensemble mean prediction $\mu_{\nu|i}$ and the hypothetical observation ν, whereas the SNR-based potential predictability only measures their linear correlation. When $\mu_{\nu|i}$ and ν are joint normally distributed, their statistical dependence reduces to a linear correlation. When $\mu_{\nu|i}$ and ν are not joint normally distributed, MI naturally disagrees with SNR. Joint normally distributed variables have a constant conditional variance. Note that the prediction variance is also the conditional variance of ν given the ensemble mean $\mu_{\nu|i}$. Thus, if the prediction variance is not constant, $\mu_{\nu|i}$ and ν are definitely not joint normally distributed, causing the SNR-based potential predictability, a linear correlation between $\mu_{\nu|i}$ and ν, to underestimate the nonlinear statistical dependence between $\mu_{\nu|i}$ (or i) and ν, which is a strict statistical definition of potential predictability.

[29] It should be noted that the above conclusion should not be challenged by the possible fact that the SNR-based potential predictability may have a better relationship with the actual skill than the MI-based potential predictability simply because the actual skill is often measured by a linear correlation (or a related quantity) that is inherent to the SNR-based potential predictability. Thus, a more challenging issue is how to design new metrics to measure the actual forecast skill that could realize the MI-based extra potential predictive information beyond the SNR. In principle, the MI between the ensemble mean prediction $\mu_{\nu|i}$ and the actual observation O would potentially provide the ability to realize the MI-based potential predictability [e.g., *DelSole*, 2005]. However, the effective estimation of the MI in this context is difficult. Further discussion of this topic is beyond the scope of this study, and we leave it for future work in this field.

[30] Despite the same evolution trend with lead time, there remains a discrepancy between the actual skill and the potential predictability based on both the MI and the SNR. For instance, for the WYMI, the former is much smaller than the latter. Model imperfection is the main source of this kind of discrepancy. Thus, a comparison between the actual skill and the potential predictability in Figure 2 indicates that the model imperfection in the WYMI prediction is more severe than that in the EASMI prediction. Below, we depict the model imperfection in another way. Following the discussion in the previous section, model perfection here refers to the actual observation being statistically indistinguishable from members of the forecast ensemble. Therefore, if the model is perfect, the variances of all the members and observations, which are the modeled and observed climatological variances, respectively, are statistically identical. Figure 3a shows the ratio of the modeled variance to the observed counterpart. Compared with the EASMI, the variance ratios for the WYMI are far from unity at all lead times. In addition, an F test was performed. At a confidence level of 95%, the differences between the modeled and observed

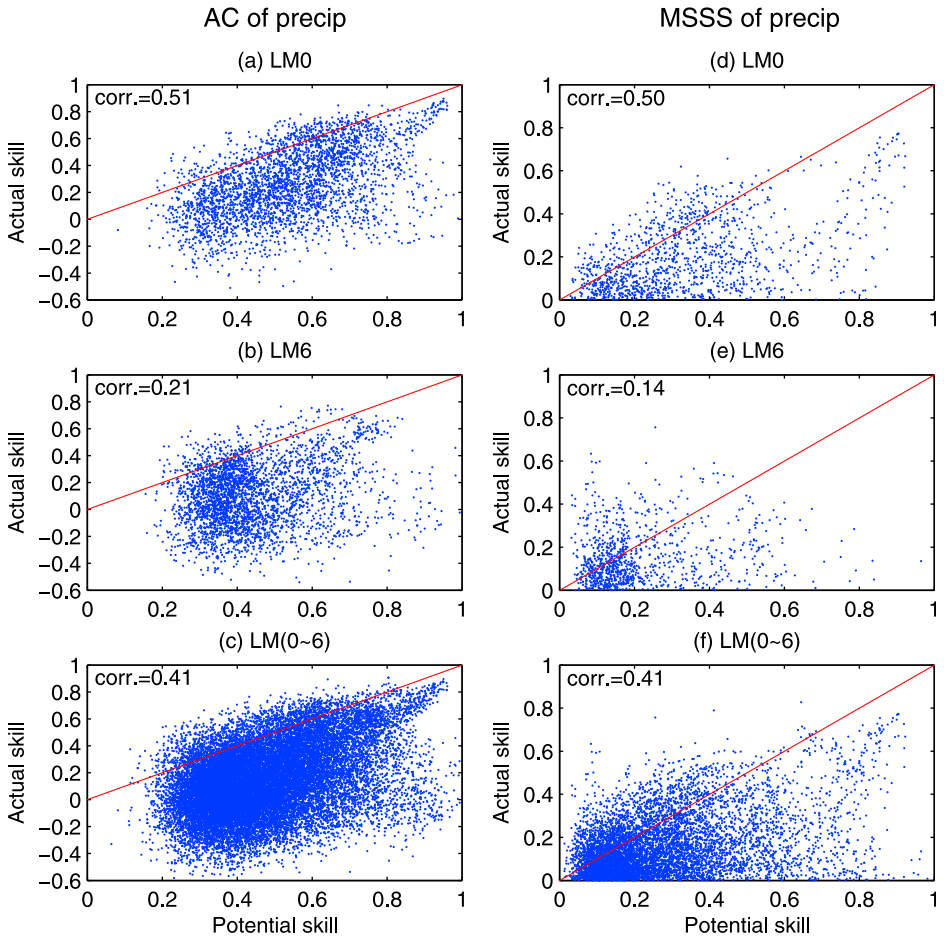

Figure 4. Scatterplots of the MI-based potential predictability versus the actual skills for precipitation prediction within the Asian monsoon domain (30°S–50°N, 40°–180°E). Each dot represents the skill at one grid. For the MSSS, only the actual skills with nonnegative values are presented. The dots on the straight line connecting (0,0) and (1,1) have actual skills equal to the potential predictability. (a and d) The 0-month lead (LM0) predictions. (b and e) The 6-month lead (LM6) predictions. (c and f) Predictions at all lead times (LM0–LM6).

variances for the EASMI are not statistically significant at all lead times. For the WYMI, conversely, the variance differences are statistically significant. In other words, the model imperfection in the WYMI prediction is more severe than that in the EASMI prediction. If the model is perfect, there should also be no significant difference in the means. However, because the prediction target is the anomaly, there is always no mean difference for both indices (it is automatically zero). Thus, in terms of the first moment, the model is perfect. Evidently, this cannot exclude the existence of mean bias in the raw data, as shown in Figure 3b. Student's t-test was performed, and the results indicate that the biases for the WYMI are statistically significant at the confidence level of 95% at all lead times, whereas for the EASMI, there are no significant biases.

[31] It is worth noting that model imperfection can cause the potential predictability to be greater or smaller than the actual skill, depending on its distribution and feature. If model error only exists in the signal component (the ensemble mean) as random error, the potential predictability is greater than the actual skill, as proven by *Sardeshmukh et al.* [2000] and *Kang and Shukla* [2006]. In Figure 2, the smaller actual skill for the WYMI is a good example of this case. However, in other cases, the potential predictability may be smaller than the actual skill, such as the EASMI predictability at the lead times of LM3 and LM4, as shown in Figure 2. This is most likely due to the model imperfection in estimating the noise variance. For example, if the ensemble is over-dispersive and prediction variance is overestimated, the potential predictability can be underestimated compared with actual skill [e.g., *Batté and Déqué*, 2011].

[32] It is also interesting to explore the potential and actual skills in a grid-by-grid manner for the ASM domain. Here, the Asian monsoon region is considered to cover the area (30°S–50°N, 40°–180°E), which includes both the Asian lands and the nearby waters of the Indian Ocean and the western Pacific. Figures 4 and 5 present scatterplots of the MI-based potential predictability against the actual skills for the grids within the ASM domain, measured using two important monsoon variables: precipitation and 850 hPa zonal wind (u850). In Figures 4 and 5, each dot represents

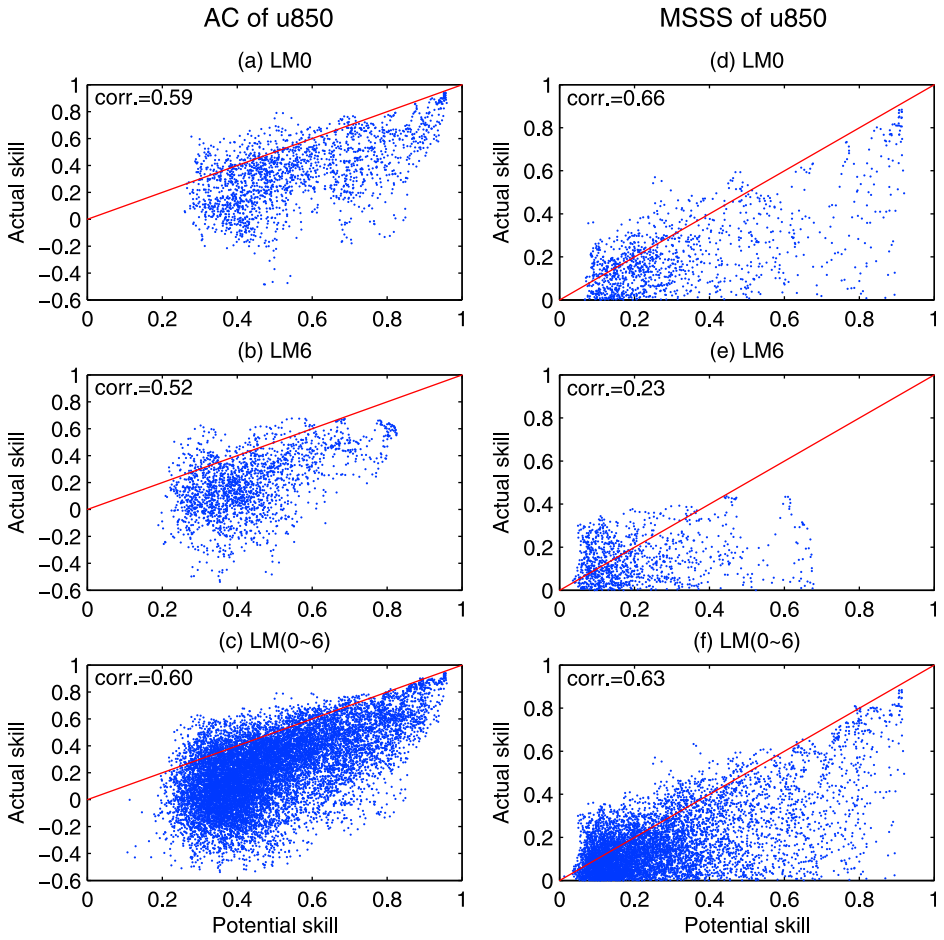

Figure 5. Same as Figure 4 but for u850 prediction.

the skill at one grid. The dots on the straight line indicate the grids where the actual skill equals the potential predictability. The dots above the line represent actual skill larger than potential predictability, and vice versa. The potential predictability is higher than the actual skills for most grids, but there are a few opposite cases, implying the complex impact of model error on the potential predictability, as mentioned above. However, there is still a good correlation between the potential and actual AC skills, as indicated in the upper left corners of Figures 4a–4c and 5a–5c. Their correlation coefficients are statistically significant at the confidence level of 95%, suggesting that the AC-based potential predictability dominates the variation of the actual skill within the ASM domain. For the MSSS, in contrast, the scatterplots using all grids display no significant correlations between the potential and actual skills (figures not shown). This difference indicates that model error severely contaminates the underlying relationship between the potential and actual MSSSs. Compared with the AC-based measure, the relationship between the potential and actual MSSSs is more sensitive to model error, probably due to the unboundedness of the MSSS. In contrast, the correlation has a lower bound of -1, preventing the actual AC from becoming too small even when there is very large model error. Instead, if we choose only those grids with nonnegative values of the actual MSSS, there exists a significant relationship between the potential and actual skills, as shown in Figures 4d–4f and 5d–5f. The threshold of zero was not chosen arbitrarily. Usually, a larger threshold indicates a better relationship between the potential and actual skills. However, as indicated in Appendix A, the skill of a perfect model cannot be smaller than zero. Therefore, a threshold of zero is appropriate to remove the grids with severe model error but still retain enough grids.

[33] In addition, the relationship between the potential predictability and the actual skill is lead-time dependent. As shown in Figures 4 and 5, their relationship becomes weaker as the lead time increases, implying that the model error becomes more severe with increasing lead time. This result is consistent with that of *S.-S. Lee et al.* [2011].

[34] An analysis similar to that presented in Figures 4 and 5 was also performed using the SNR-based potential predictability. The results showed that the SNR-based potential predictability has a good relationship with the actual skill, as good as the MI-based potential predictability for u850 (figure not shown). This similarity is because the ensemble spread does not differ much from prediction to prediction for u850, as mentioned above. For precipitation, however, the correlation between the SNR-based potential predictability and the actual skill is about 0.1 higher than the MI-based counterpart due to the relatively large variability in the ensemble spread (figure not shown).

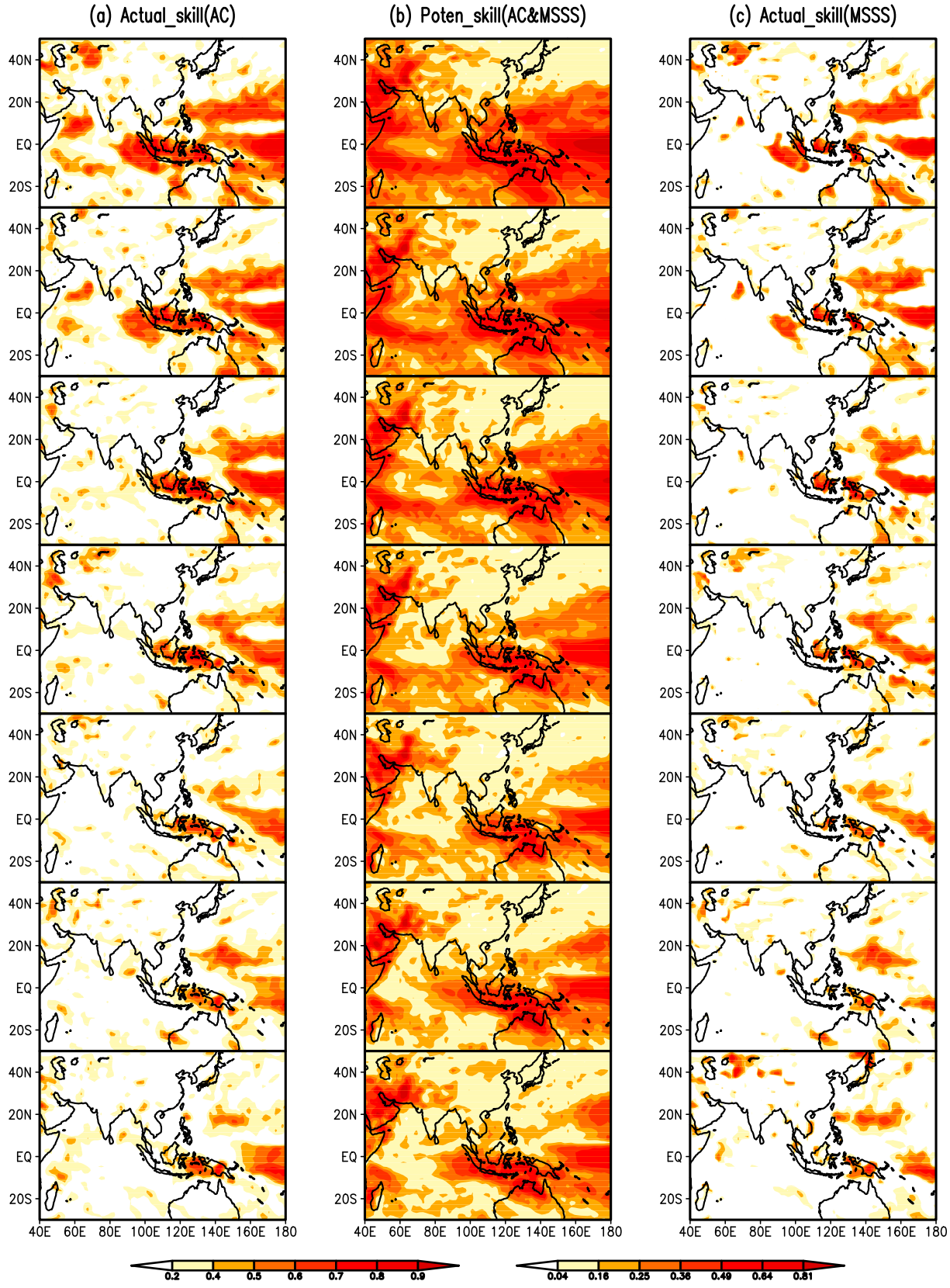

Figure 6. Spatial distributions of the MI-based potential predictability and the actual skills for precipitation prediction. In each column, seven subplots, from top to bottom, correspond to the seven lead times, from LM0 to LM6. The left color bar is for the AC and the right bar is for the MSSS. Note that the scale of the right color bar is the square of that of the left one. For Figure 6b, when discussing the AC (MSSS), the left (right) color bar is used.

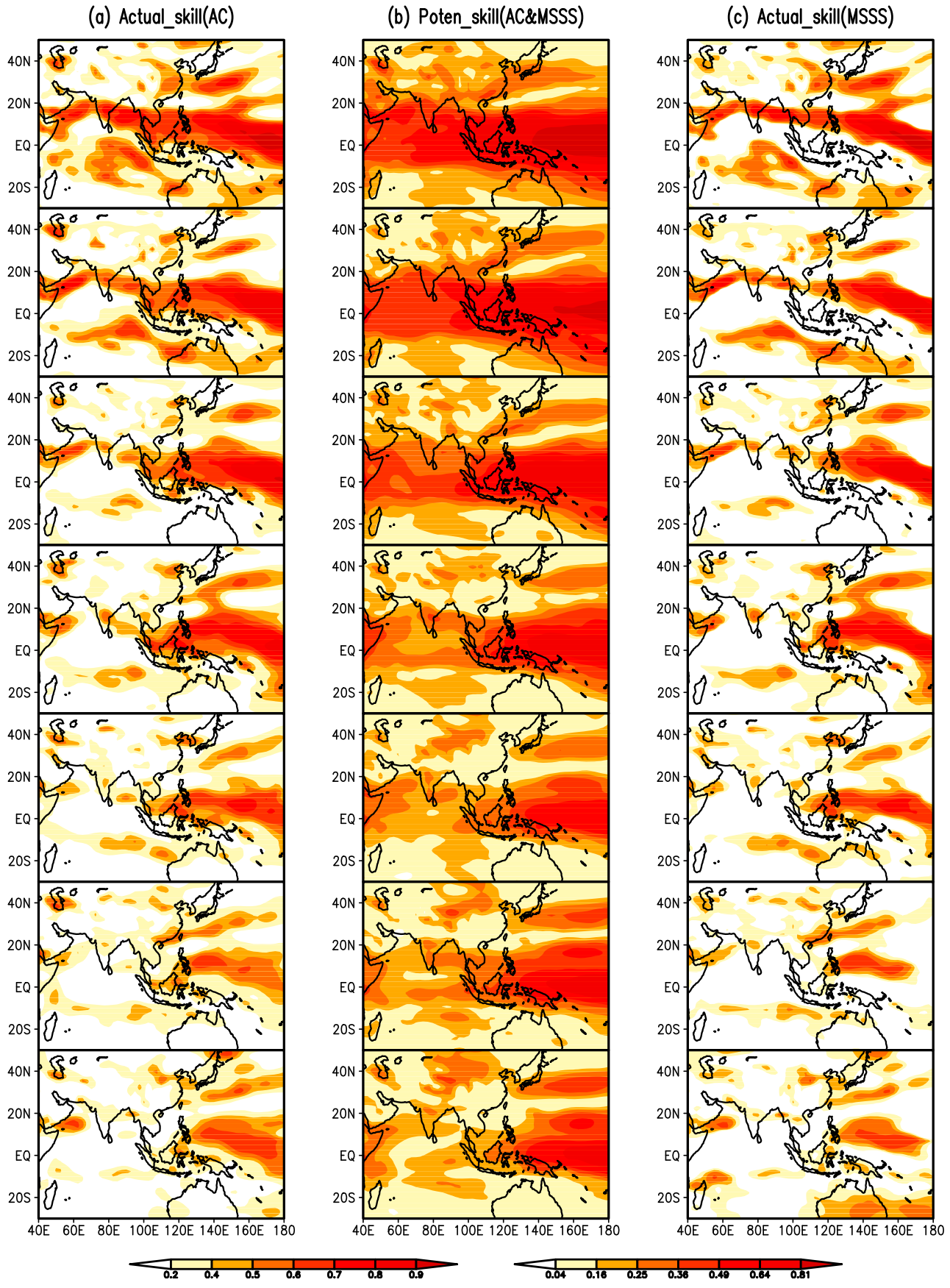

Figure 7. Same as Figure 6 but for u850 prediction.

[35] The spatial distributions of the AC and MSSS for the precipitation and u850 are illustrated in Figures 6 and 7. Figures 6a and 7a (Figures 6c and 7c) display the actual AC (MSSS) skills. Figures 6b and 7b display the potential predictability of both the AC and MSSS, which can be represented on the same plot due to their equivalence, as shown by the first equality in equation (9). There are inconsistencies between the MI-based potential predictability and the actual skill due to model error and (or) the nonlinear property of the MI-based potential predictability, as argued earlier.

[36] A striking feature in Figure 6 is the large predictability occurring in the tropical region (20°S–20°N) for both the potential and actual skills, particularly in the tropical western Pacific. A high predictability in the tropical Pacific is a common feature in climate models, mainly as the result of a strong response (signal) to the SST forcing, especially the SST forcing related to ENSO, and the relatively weak unpredictable atmospheric internal variability there [e.g., *Charney and Shukla*, 1981; *Stern and Miyakoda*, 1995; *Shukla*, 1998; *Rowell*, 1998; *Hoerling and Kumar*, 2002; *Palmer et al.*, 2004]. Another similarity between the potential and actual skills is their decreasing magnitudes with lead time. While the potential predictability still shows some visible patterns for long lead times, actual skills almost vanish after LM2. The difference between the potential and actual skills mainly occurs in the tropical Indian Ocean, the Middle East and Northern Australia, where a relatively high potential predictability exists against a weak actual skill.

[37] In Figure 7, the potential predictability of the u850 prediction at LM0 is zonally distributed within 20°S–20°N, with the large values mainly located in the tropical western Pacific Ocean and the small values mainly located in the extratropical region. The actual AC and MSSS display spatial distributions similar to those of the potential predictability, i.e., high values mainly in the tropical region and low values in the extratropical region. Similar to the precipitation, the similarity between the actual and potential skills decreases with lead time, with the best similarity at LM0.

[38] We also performed a spatial analysis for the SNR-based potential predictability. The SNR-based spatial pattern for u850 (figure not shown) is very similar to Figure 7 because the ensemble spread of the u850 prediction changes little, as noted above. For precipitation, the SNR-based potential predictability shows a spatial distribution similar to that in Figure 6 but shows large differences in two regions: the Middle East and Northern Australia, as illustrated in Figure 8. A comparison between Figures 7b and 8 reveals that the MI-based measure gives a higher potential predictability than the SNR-based measure, suggesting that the MI-based measure provides extra potential predictive information beyond the conventional SNR in the two regions.

5. The Relationship Between Relative Entropy and Individual Prediction Skills

[39] In the previous section, we analyzed the average potential predictability of the ASM system. In this section, we focus on the potential predictability of an individual prediction. From the viewpoint of practical prediction, it is more interesting to seek the relationship between the potential predictability and the actual skill for an individual prediction, offering an a priori means of estimating the confidence level of the individual prediction.

[40] Figures 9a and 9c present scatterplots of the RE versus the absolute error (AE) for the WYMI and EASMI predictions, respectively, for all lead times. Here, the AE is defined as the absolute value of the difference between each prediction and the corresponding observation. As can be seen, there is a weak positive correlation between the RE and the AE for the WYMI, with a correlation coefficient of 0.19, and no significant correlation between the RE and the AE was detected for the EASMI. The positive correlation between the RE and the AE seems a striking counterexample to the widespread perception that high potential predictability is associated with a high forecast skill (and thus a small AE). This correlation is most likely due to the contamination of the model error. As discussed above, the model error is more severe in the WYMI predictions than in the EASMI predictions. To further examine the possible impact of model error on the relationship between RE and AE, we constructed scatterplots of RE versus AE using a perfect model scenario, i.e., by randomly selecting a member from the set of 15 ensemble members as the hypothetical observation for the calculation of AE, as shown in Figures 10a and 10c. A comparison of Figures 9a and 10a reveals that the unexpected positive correlation between RE and AE for the WYMI is actually due to model error because such a positive relationship is absent in Figure 10a. However, even in a perfect model scenario, there is no large negative correlation between RE and AE, indicating that the RE is not a good indicator of AE-based actual skill. The inefficiency of the RE in estimating the RMSE-based skill was also found in ENSO prediction [e.g., *Tang et al.*, 2008a].

[41] To further explore the relationship between the RE and individual prediction skill, we examined the contribution (C) of each prediction to the overall AC skill, as defined by *Tang et al.* [2005]:

$$C = \frac{1}{AC_a} \frac{\frac{1}{N}\left(x_j^p - \overline{x^p}\right)\left(x_j^o - \overline{x^o}\right)}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^p - \overline{x^p}\right)^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_j^o - \overline{x^o}\right)^2}} \times 100\%. \tag{10}$$

Figures 9b and 9d are scatterplots of RE versus C for the WYMI and the EASMI, respectively, for all lead times. Here, there is a significant positive correlation between them for both the WYMI and the EASMI, with the EASMI showing a better correlation than the WYMI. Similarly to the perfect model scenario depicted in Figures 10a and 10c, scatterplots of RE versus the "perfect" C are also shown in Figures 10b and 10d. On comparing Figure 9b (Figure 9d) and Figure 10b (Figure 10d), we find that the model error indeed also has some impact on the RE-C relationship.

[42] From equation (3), the RE is determined by two items: the NEV and the SNEMS. The former actually reflects the ensemble spread. In a perfect model scenario, it is not difficult to ascribe the relationship between the RE and the AE to that between the NEV and the AE (the SNEMS

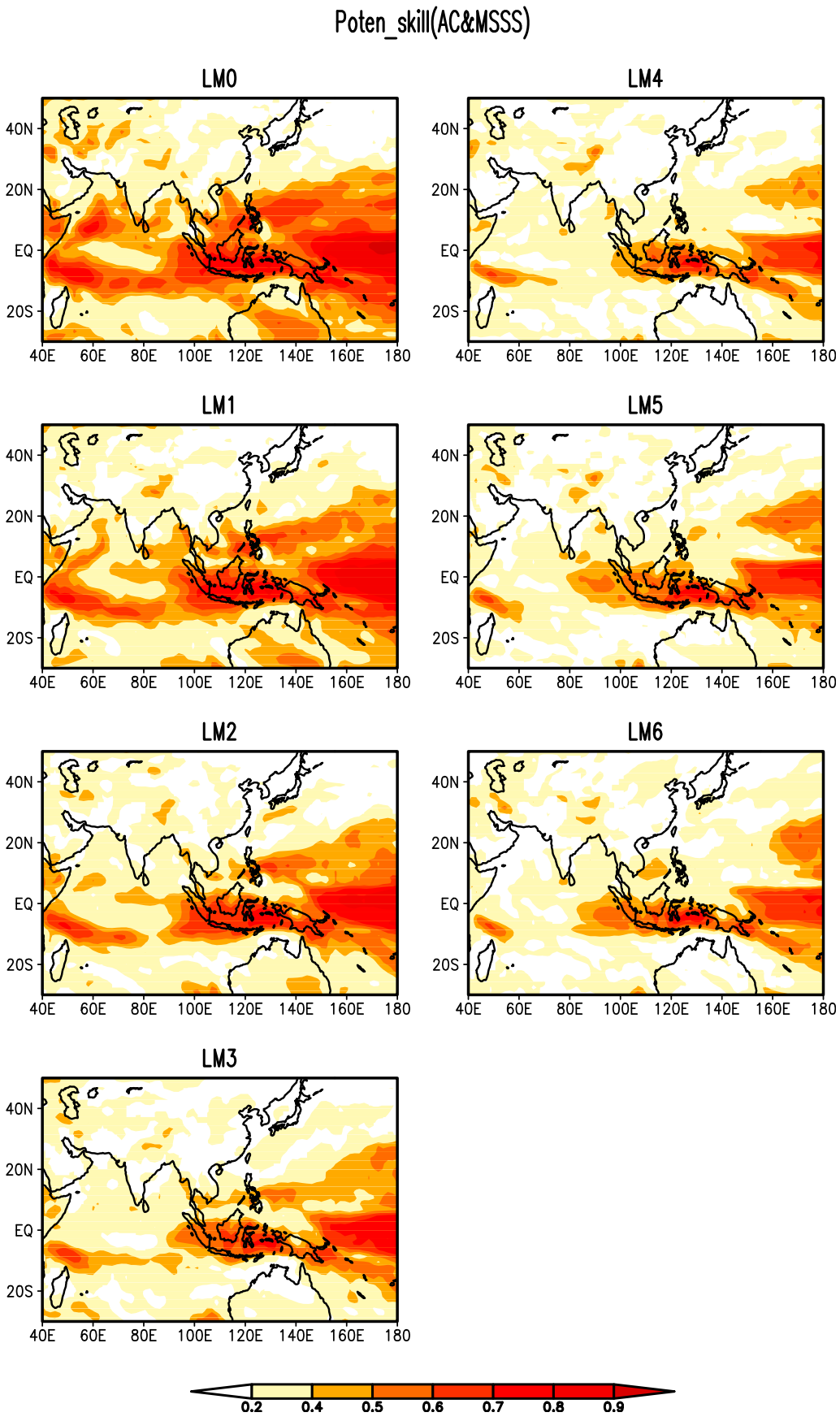

Figure 8. Same as Figure 6b but for SNR-based potential predictability. The color bar is for AC.

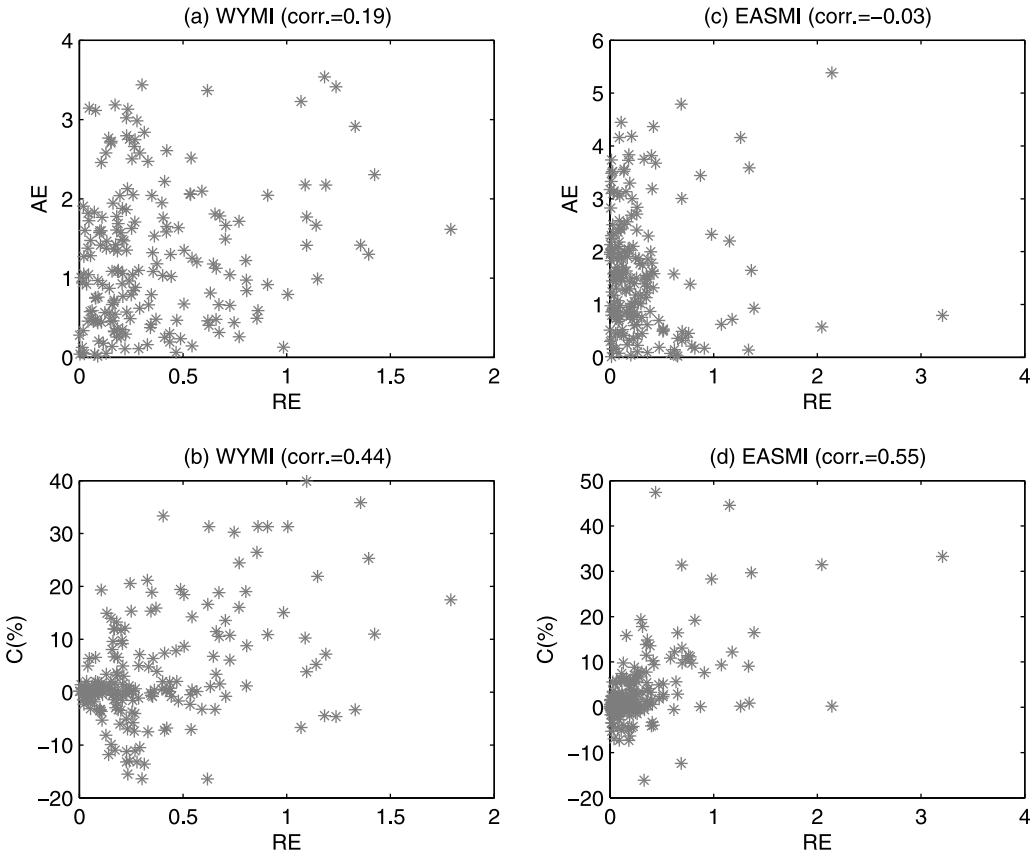

Figure 9. (a and c) Scatterplots of RE versus AE and (b and d) scatterplots of RE versus C.

can impact the AE only via its impact on the spread). The scatterplots, Figures 11a and 11c, reveal that there is no significant correlation between the NEV and the AE. To exclude the possible impact of model error on the NEV-AE relationship, scatterplots between the NEV and the "perfect" AE were constructed, similarly to the above method (Figures 12a and 12c). As can be seen, even in a perfect model scenario, the correlation between the NEV and the AE is also small although it is reasonably positive. This suggests that the weak correlation between the NEV and the AE is not predominantly due to model error; instead, it is probably an inherent property of the ensemble climate forecast system. The weak NEV-AE correlation is consistent with the spread-skill relationship found in ensemble weather forecasts [e.g., *Houtekamer*, 1993; *Whitaker and Loughe*, 1998]. Under the assumption of a perfect ensemble, *Houtekamer* [1993] derived an analytical relationship between the ensemble spread and forecast accuracy, which is a function only of the variability of the ensemble spread (β). When β is very small, the spread is considered to be approximately constant, and the prediction error is a random number drawn from a fixed distribution, so the correlation between the ensemble spread and forecast accuracy approaches zero [*Whitaker and Loughe*, 1998]. In ensemble climate predictions, the variability in the ensemble spread is usually relatively small compared with that for the weather forecasts [e.g., *Kumar et al.*, 2000; *Tippett et al.*, 2004; *Wu and Kirtman*, 2006; *Tang et al.*, 2007, 2008b] (also see section 4), leading to a weak relationship between the ensemble spread and the AE for ensemble climate predictions.

[43] *Tang et al.* [2008b] found that the ensemble mean square is a good indicator of C. The SNEMS contains the ensemble mean square. Figures 11b and 11d are scatterplots between the SNEMS and C. Here, there is a significant positive correlation between the SNEMS and C for both the WYMI and the EASMI, which accounts for the significant correlation between the RE and C. However, as indicated in Figures 12b and 12d, the counterparts of Figures 11b and 11d with the perfect model scenario, the model error still somehow biases the SNEMS-C relationship. As mentioned above, the variation in the ensemble spread in climate prediction is usually relatively small. If we assume the spread is constant, the RE-C correlation instantly reduces to the SNEMS-C correlation. *Tang et al.* [2007, 2008b] presented a relationship between the theoretical correlation (denoted ρ) of RE-C (or SNEMS-C) and the model actual AC skill (r) as follows, based on the assumption of a perfect model, with Gaussian forecast and climatological distributions and a constant ensemble spread (see Appendix B):

$$\rho = \sqrt{\frac{2r^2}{r^2+1}}. \tag{11}$$

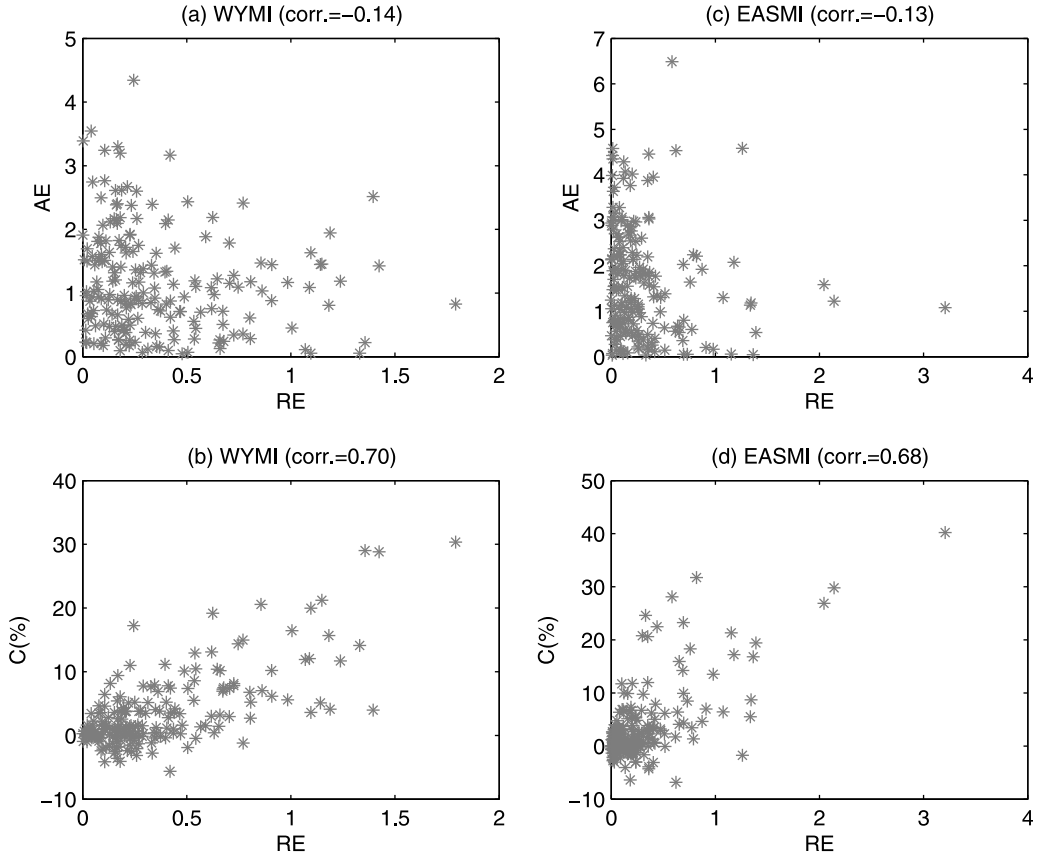

Figure 10. (a and c) Scatterplots of RE versus AE and (b and d) scatterplots of RE versus C using a perfect model scenario.

Equation (11) builds the connection between the RE-C correlation and the AC skill. It is interesting to examine equation (11) using the actual ensemble predictions. Under the perfect model scenario, the AC values (r) are 0.71 and 0.66, respectively, for the WYMI and the EASMI, which leads to theoretical RE-C correlation (ρ) values of 0.82 and 0.78, close to the actual values of 0.70 and 0.68 shown in Figures 10b and 10d. Furthermore, we used the actual AC skill in equation (11) and obtained theoretical ρ values of 0.51 and 0.71 for the WYMI and the EASMI. Again, these values are not far away from the actual RE-C correlations of 0.44 and 0.55 shown in Figures 9b and 9d. The good agreement between the theoretical RE-C and the actual RE-C correlations may offer an effective means for estimating the reliability of individual ASM predictions.

6. The Role of SST Forcing

[44] It is of interest to explore the possible sources of information-based potential predictability. It has been argued in the literature that the physical basis for seasonal climate prediction lies in slowly varying lower boundary forcing, especially the SST forcing [e.g., *Charney and Shukla*, 1981; *Shukla*, 1998; *Wang et al.*, 2007; *J.-Y. Lee et al.*, 2011]. In this section, we explore SST forcing and its impact on potential predictability by analyzing the RE, the SST and ensemble prediction itself.

[45] First, we calculated the correlation between the RE and the square of the SST anomaly (SSTA2) of the prediction initial time at each grid over the global ocean for all lead times. The reason for using the SSTA2, not the SSTA itself, is that the RE is actually independent of the sign of the ensemble mean, as indicated in equation (3), whereas the SSTA usually induces both positive and negative atmospheric responses. For simplicity, we present only two representative correlation maps here, as shown in Figure 13: the correlation between the RE of the prediction starting in May (LM0) and the observed May SSTA2, and the correlation between the RE of the prediction starting in January (LM4) and the observed January SSTA2. For the WYMI at LM0, significant correlations mainly appear in the tropical central-eastern Pacific, the western North Pacific, the subtropical South and North Pacific. At LM4, the correlations display a pattern similar to that at LM0 except that significant correlations also exist in the western Indian Ocean. For the EASMI at LM0, significant correlations appear in the tropical central-eastern Pacific, the subtropical South and North Pacific, and the eastern Indian Ocean. At LM4, significant correlations mainly appear in the tropical central-eastern Pacific and the western Indian Ocean. The stronger correlation for LM4 than for LM0 is probably due to two reasons:

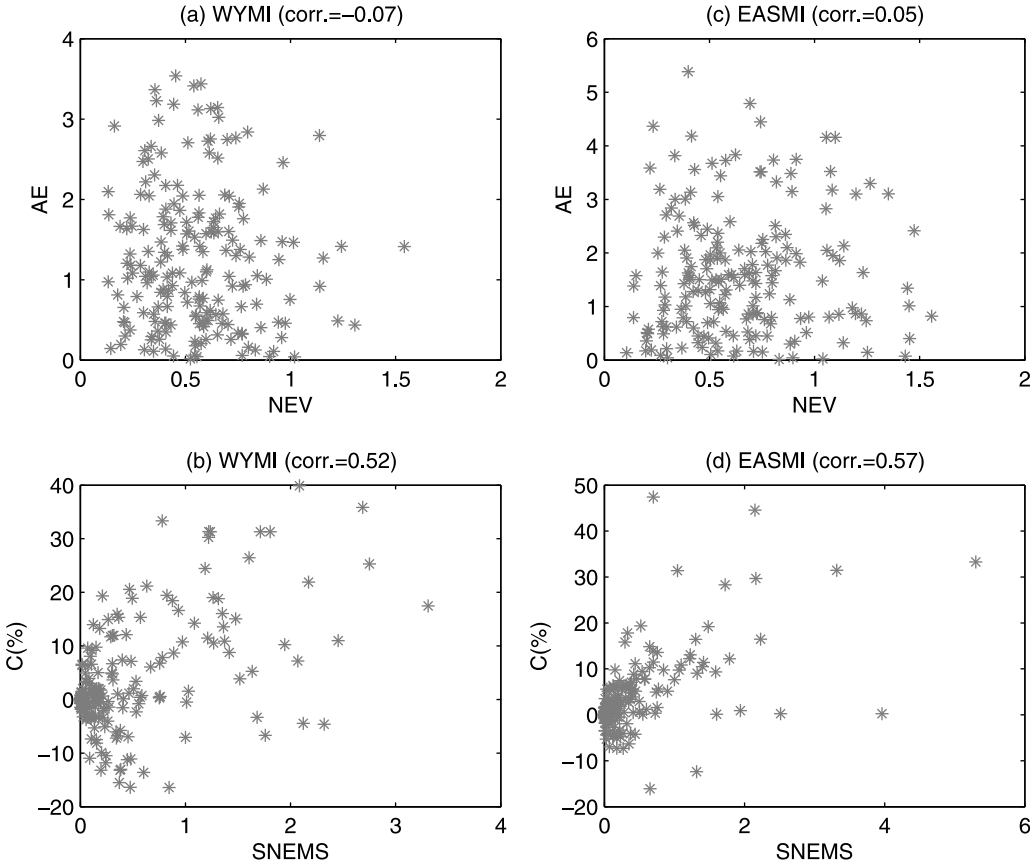

Figure 11. (a and c) Scatterplots of NEV versus AE and (b and d) scatterplots of SNEMS versus C.

1) There is a much stronger SST signal in winter (January) than in later spring (May) in the tropical Pacific ocean, as documented in the observational studies [e.g., *Rasmusson and Carpenter*, 1982]; and 2) the atmospheric response over the EASM region to the ENSO has a time lag of around 6 months, as reported in the literature [e.g., *Wang et al.*, 2000, 2008b].

[46] Because the RE is a function of both the NEV and the SNEMS, the NEV-SSTA2 correlation and the SNEMS-SSTA2 correlation were also calculated. It was found that the spatial pattern of the SNEMS-SSTA2 correlation (figure not shown) is almost the same as that of the RE-SSTA2 correlation, which is obviously the result of weak spread variability in the ensemble climate prediction. This similarity indicates that the SSTA2, if it has any effect, mainly induces variation in the ensemble mean amplitude. Therefore, the significant correlations in Figure 13 are mainly due to the contribution of the SST forcing to the amplitude of the ensemble mean. In reality, however, the ensemble mean amplitude usually varies accompanied by a change of the ensemble mean sign because the SSTA can lead to both positive and negative atmospheric responses. Thus, it is probably more informative to directly correlate the global SSTA with the model ensemble mean. As shown in Figure 14, all the correlation patterns resemble the typical ENSO pattern, implying that the ensemble mean is mainly an ENSO-forced component; in other words, the ENSO is a major source of ASM seasonal predictability. Other sources, such as those due to the internal dynamics of the atmospheric system itself, may also provide some predictability for the ASM [e.g., *Fu et al.*, 2007; *Kleeman*, 2008].

[47] From the perspective of model evaluation, the modeled SST-monsoon relationship and its observed counterpart are often compared [e.g., *Yang et al.*, 2008; *S.-S. Lee et al.*, 2011]. We also calculated the correlation between the observed JJA monsoon indices and the observed May SST and the correlation between the observed JJA monsoon indices and the observed January SST, as shown in Figure 15. Comparing Figure 15 with Figure 14 reveals that the modeled SST-monsoon relationship is much stronger than the observed counterpart, especially for the WYMI. This is consistent with the result of *Yang et al.* [2008]. Besides the model error, one possible reason for the discrepancy between the modeled and observed SST-monsoon relationships is that the model ensemble mean filters out a large portion of the fast, unpredictable component of the coupled system that is present in the actual observations. SST forcing is mainly related to the slowly varying, predictable component. Thus, the existence of the unpredictable component in the observations degrades the observed SST-monsoon relationship compared with the modeled counterpart. To reproduce the observed relationship in the model, it is more reasonable to use an individual ensemble member that contains the unpredictable component, rather than the

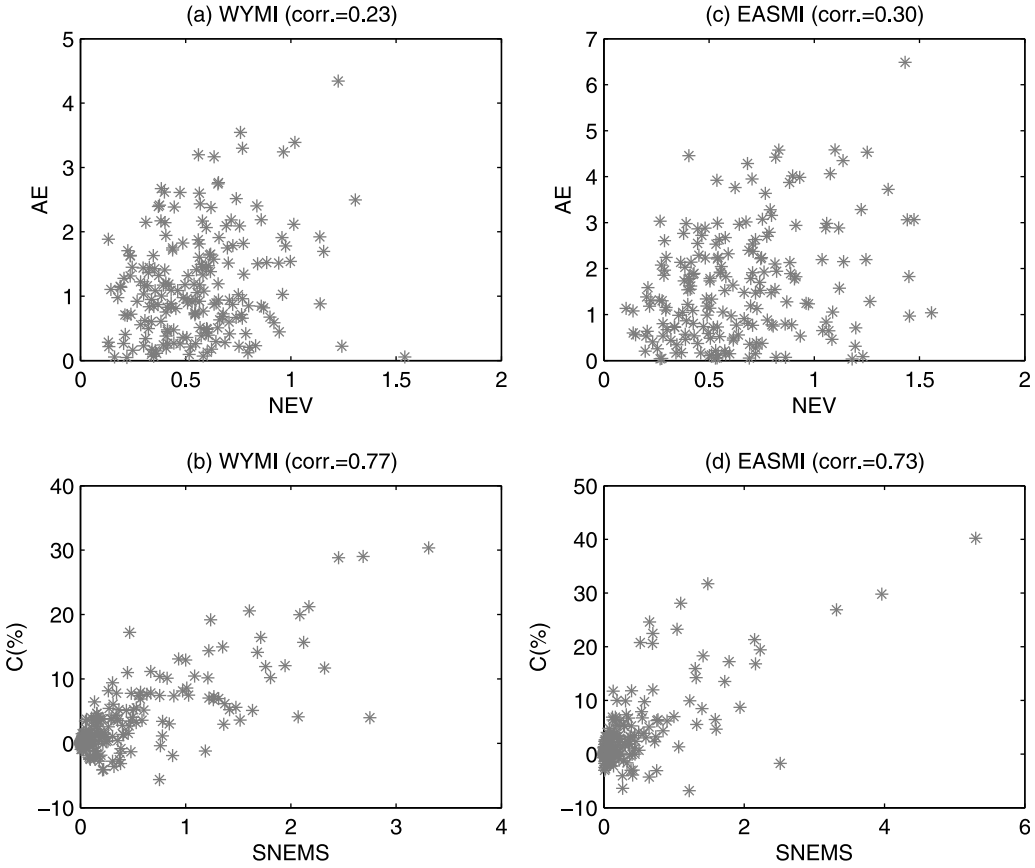

Figure 12. (a and c) Scatterplots of NEV versus AE and (b and d) scatterplots of SNEMS versus C using a perfect model scenario.

ensemble mean, to represent the modeled monsoon. To test this hypothesis, we performed this experiment and repeated the calculations until all the 15 members were used. The average of the fifteen correlations is shown in Figure 16. Here, the SST-monsoon correlation is much weaker than that in Figure 14 and more similar to the observed counterpart in Figure 15.

[48] A further examination of Figure 15 suggests that the ENSO may influence the SASM and the EASM in different ways. Comparison of the correlation coefficients for LM4 and LM0 in Figure 15 reveals that the lag correlation is lower than the "simultaneous" correlation for the WYMI, but the opposite conclusion is drawn for the EASMI. As noted by *Yang and Lau* [2006], the ENSO can directly impact the SASM by Walker circulation, and thus, the response of monsoon circulation to the ENSO is almost simultaneous. This result also confirms previous finding that the EASM has a 2-season (6-month) lagged response to the ENSO forcing during an ENSO-decaying summer [e.g., *Wang et al.*, 2000, 2008b]. The lagged response of the EASM to the ENSO can be accounted for by two possible mechanisms: monsoon-warm pool interaction [*Wang et al.*, 2000] and the Indian Ocean capacitor effect [*Xie et al.*, 2009]. *Wang et al.* [2000] pointed out that the local warm pool ocean-monsoon interaction maintains the Philippine Sea anomalous anticyclone, which is formed during the peak phase of El Niño and influences the EASM during the following summer. *Xie et al.* [2009] reported that the persistent Indian Ocean warming after the peak phase of El Niño triggers a baroclinic Kelvin wave into the western Pacific, which induces an Ekman divergence mechanism in the western North Pacific and further influences the EASM.

7. Summary and Discussion

[49] In this study, two information-based measures were used to explore the potential predictability of the Asian summer monsoon (ASM): relative entropy (RE), which measures individual potential predictability, and mutual information (MI), the average of the RE over all initial conditions, which measures the average potential predictability. The Webster-Yang monsoon index (WYMI) that measures broad-scale South Asian summer monsoon (SASM) circulation anomalies, the East Asian summer monsoon (EASM) index (EASMI), and the ASM regional precipitation and 850 hPa zonal wind, were chosen as the targets of this study.

[50] For the average predictability, an emphasis was placed on a comparison between the potential predictability based on the MI without using observations and the actual skill with observations along with the relationship between the MI-based potential predictability and another widely used potential predictability measure based on the signal-to-

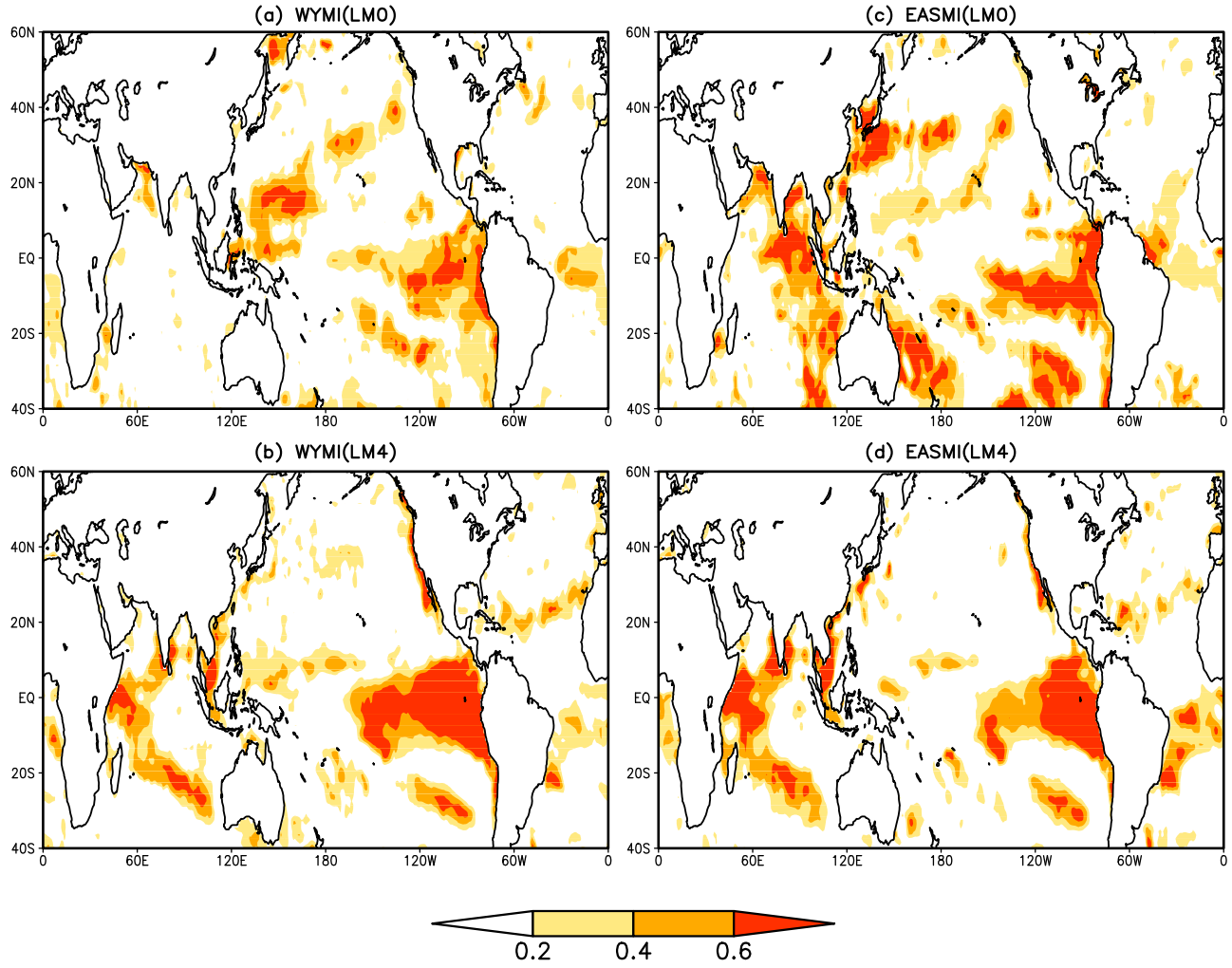

Figure 13. Spatial pattern of the correlation between the RE of predictions of the JJA ASM indices and the square of the observed initial SSTA.

noise ratio (SNR). We found that although model error contaminates the theoretical relationship, the MI-based potential predictability has a significant relationship with its actual counterpart for the seasonal prediction of the ASM in terms of the spatial distribution and lead-time variation of the predictability. Thus, the MI-based potential predictability explains, to a large extent, the variation of overall actual skill. In addition, the potential predictability of the EASMI is more realistic than that of the WYMI, in terms of its comparison with the actual skills, suggesting that model imperfection has a more severe impact on the SASM than on the EASM.

[51] A theoretical analysis and practical comparison between the MI- and SNR-based potential predictability revealed that the former can measure more potential prediction utility than the latter. The MI-based predictability measures the statistical dependence, linear or nonlinear, between the ensemble mean (a prediction) and an ensemble member (a hypothetical observation), whereas the SNR-based predictability only measures a linear relationship between the prediction and the hypothetical observation. When the prediction and climatological distributions are Gaussian, and the ensemble spread is constant with predictions, the two measures are equivalent. When the ensemble spread is not constant, the SNR-based predictability often underestimates the potential predictability.

[52] For individual predictability, the WYMI and EASMI show different interannual RE variabilities, which is mainly due to the differing impacts of the ENSO on the SASM and the EASM. The SASM responds to the ENSO almost simultaneously, whereas the EASM shows a lagged response to the ENSO. Furthermore, we examined the relationship between the RE and individual actual prediction skills for the WYMI and EASMI predictions, which is very interesting from a practical prediction viewpoint because the identified relationship offers a means of estimating a priori the confidence of an individual prediction. It was found that the RE is not a good predictor of the absolute error (AE) but is a good predictor of C, the contribution of an individual forecast to the overall anomaly correlation skill. With a comprehensive analysis and examination of each item of the RE and based on two ideal theoretical frameworks, we found the underlying

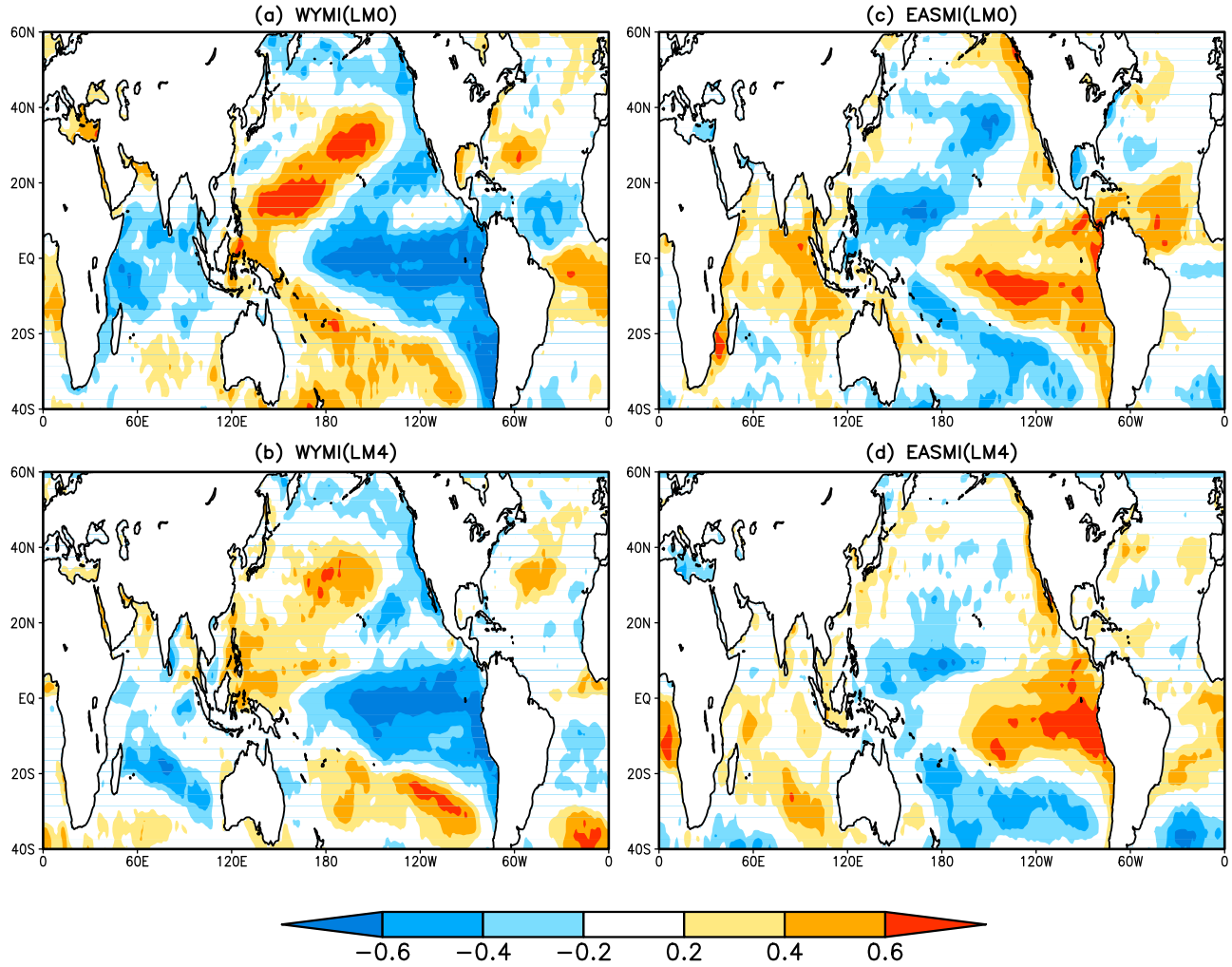

Figure 14. Spatial pattern of the correlation between the ensemble mean of the predictions of the JJA ASM indices and the observed initial SSTA.

reasons for the existence of a weak RE-AE correlation and a good RE-C correlation. The small variability of the ensemble spread in the ensemble climate forecast is the main factor responsible for the weak RE-AE correlation, although model error also has some impact, and the large average predictability yields the good RE-C correlation.

[53] An interesting issue is our understanding of the relationship between the potential predictability and the actual skill for individual predictions. Note that even for a perfect model, the observation is only one arbitrary realization and not the mean of the forecast distribution, which just reflects the chaotic or stochastic nature of the climate system and thus inevitably introduces randomness into the individual skill. This randomness undoubtedly worsens the relationship between the potential predictability and the actual skill. On the spread-accuracy relationship, as *Mason and Stephenson* [2008] argued, in a perfect model framework, it is the variance of the errors, not the error itself that should increase with increasing ensemble variance (spread). The weak NEV-AE relationship shown in Figures 12a and 12c provides good evidence for this assertion. Note that randomness also exists in the C. However, the relatively large signal component (ensemble mean) existing in both the SNEMS and the C guarantees a good SNEMS-C relationship, as shown in Figures 12b and 12d.

[54] The dominant role of SST forcing in ASM seasonal predictability was explored. We found that the RE has a good correlation with the square of the SSTA in several regions across the global oceans. Further analysis showed that SST forcing impacts ASM seasonal predictability, mainly by changing the amplitude of the ensemble mean. The SST-ASM correlation patterns in the model resemble the typical ENSO pattern, suggesting that the ENSO is the main source of ASM seasonal predictability. In addition, a possible explanation is provided for the much stronger SST-monsoon correlation in the model than in the observation, namely, that the model ensemble mean filters out a large portion of the unpredictable internal components of the coupled system that are contained in the actual observations.

[55] It should be noted that the potential predictability studied in this work is based on the perfect model assumption; thus, the conclusions and findings can be expected to

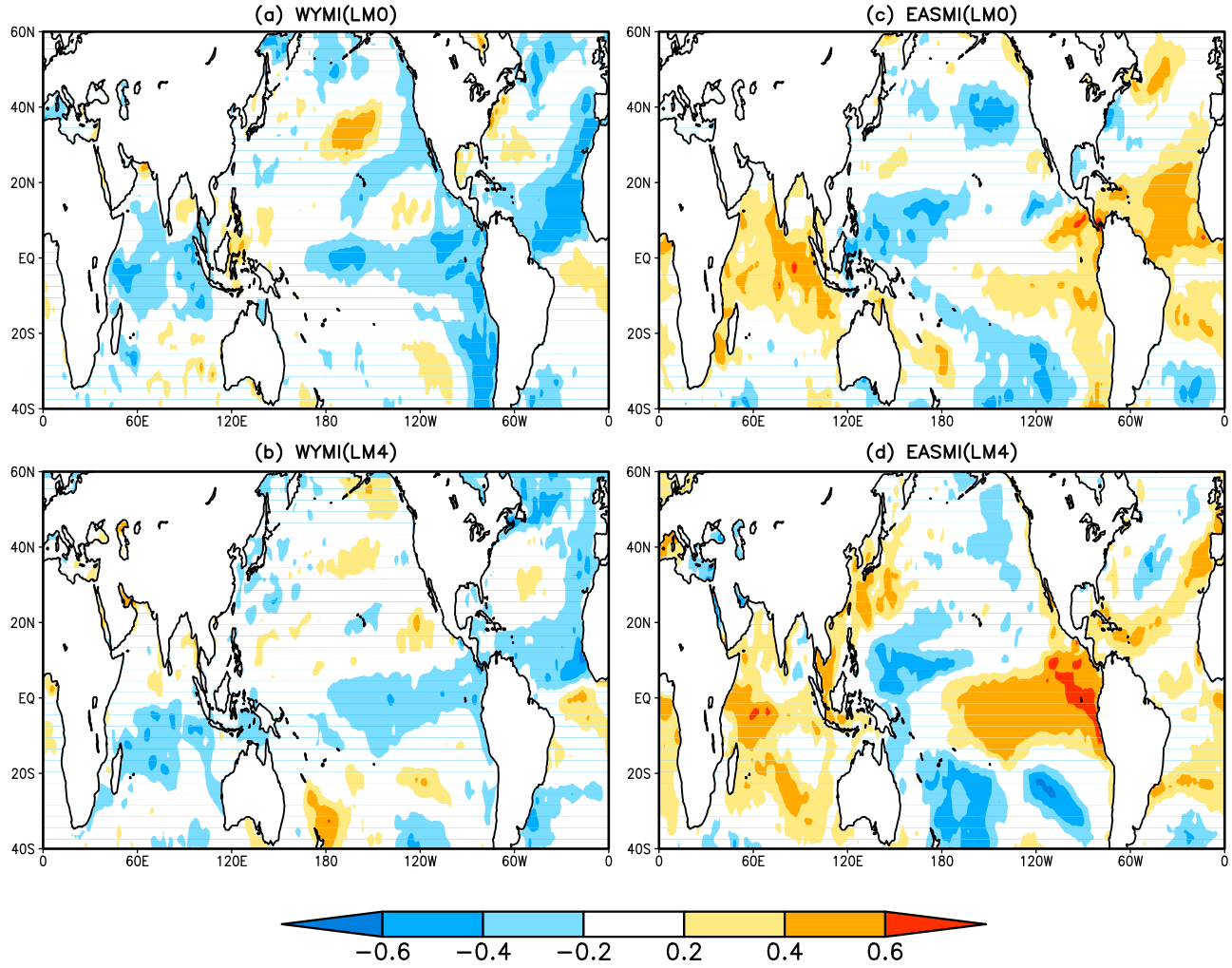

Figure 15. Spatial pattern of the correlation between the observed JJA ASM indices and the observed initial SSTA.

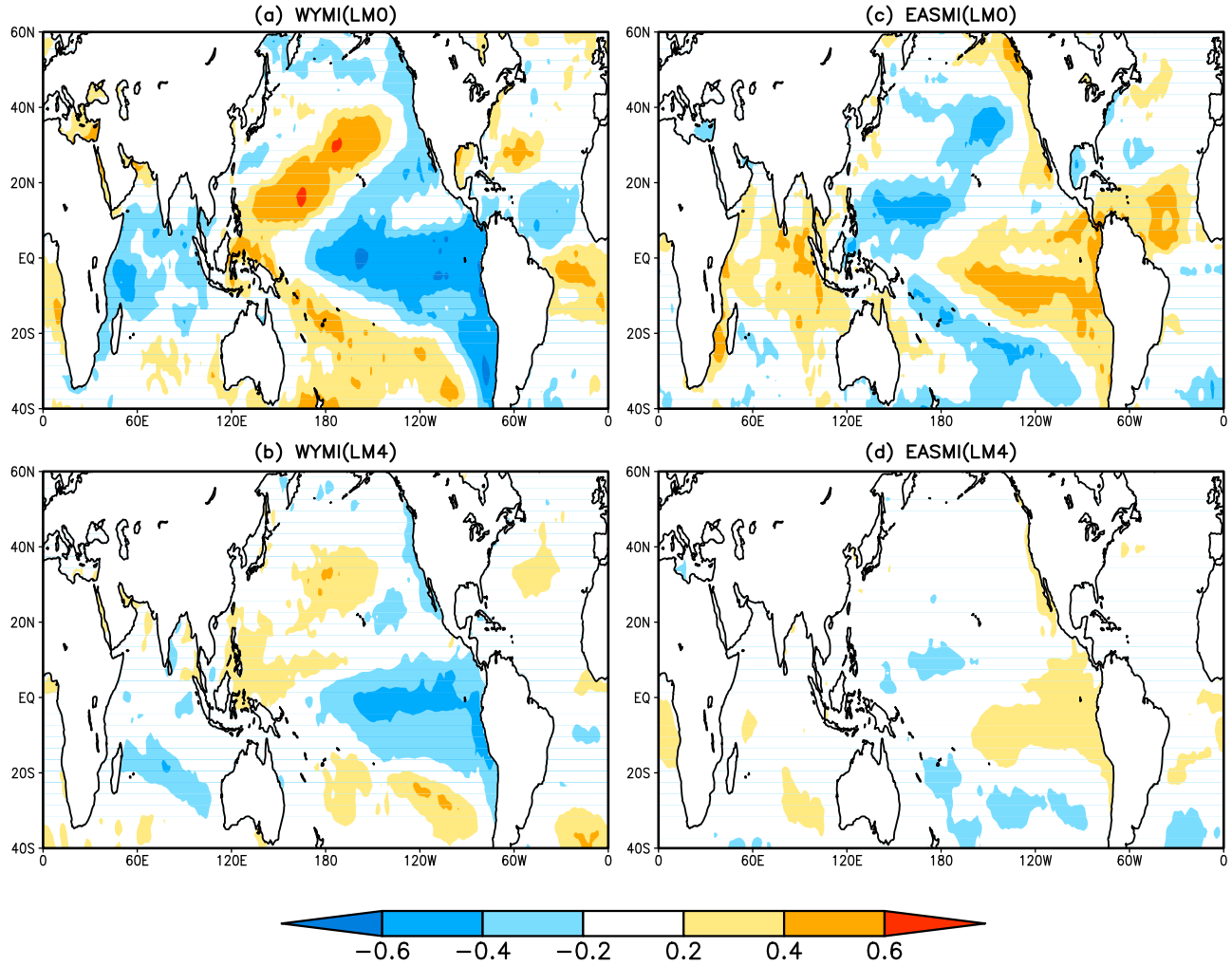

Figure 16. Spatial pattern of the correlation between the forecast of the JJA ASM indices using individual ensemble members and the observed initial SSTA. Shown here is the average of the fifteen individual correlations.

be slightly model dependent. For verification, we performed a similar analysis using stream 2 hindcasts of 5 state-of-the-art coupled models from the ENSEMBLES project [*Weisheimer et al.*, 2009]. We found that the main qualitative conclusions drawn in this study still held for the ENSEMBLES models.

Appendix A: The Theoretical Relationship of MI With Overall Skill

[56] According to *DelSole and Tippett* [2007] and *Tippett et al.* [2010],

$$\sigma_\nu^2 = \langle \sigma_{\nu|i}^2 \rangle + \left\langle (\mu_{\nu|i} - \mu_\nu)^2 \right\rangle, \tag{A1}$$

where $\langle \ldots \rangle$ denotes the mean (expectation) over all initial conditions. (A1) can be interpreted as a decomposition of the climatological variance (also called the total variance, in some literature) into a sum of noise variance and signal variance. Note that (A1) holds for any kind of probability distribution. Using (A1), averaging RE over all initial conditions makes the last three terms on the right-hand side of equation (3) vanish, giving

$$MI = \langle RE \rangle = \frac{1}{2}\left(\ln \sigma_\nu^2 - \langle \ln \sigma_{\nu|i}^2 \rangle\right). \tag{A2}$$

Without loss of generality, on the other hand, we assume that the climatological mean μ_ν is zero, and we can then write the potential AC and MSSS as

$$AC_p = \frac{\langle \mu_{\nu|i}, \nu \rangle}{\sqrt{\langle \mu_{\nu|i}^2 \rangle}\sqrt{\langle \nu^2 \rangle}} = \frac{\langle \mu_{\nu|i}^2 \rangle}{\sqrt{\langle \mu_{\nu|i}^2 \rangle}\sqrt{\langle \mu_{\nu|i}^2 \rangle + \langle \sigma_{\nu|i}^2 \rangle}}$$
$$= \frac{\sqrt{\langle \mu_{\nu|i}^2 \rangle}}{\sqrt{\langle \mu_{\nu|i}^2 \rangle + \langle \sigma_{\nu|i}^2 \rangle}}, \tag{A3}$$

$$MSSS_p = 1 - \frac{\left\langle (\mu_{\nu|i} - \nu)^2 \right\rangle}{\langle \nu^2 \rangle} = 1 - \frac{\langle \sigma_{\nu|i}^2 \rangle}{\langle \mu_{\nu|i}^2 \rangle + \langle \sigma_{\nu|i}^2 \rangle}$$
$$= \frac{\langle \mu_{\nu|i}^2 \rangle}{\langle \mu_{\nu|i}^2 \rangle + \langle \sigma_{\nu|i}^2 \rangle}, \tag{A4}$$

where we have already applied (A1) and

$$\langle \mu_{\nu|i}, \nu \rangle = \int E[\mu_{\nu|i}\nu|i]p(i)di = \int \mu_{\nu|i}E[\nu|i]p(i)di = \int \mu_{\nu|i}^2 p(i)di = \langle \mu_{\nu|i}^2 \rangle, \quad \text{(A5)}$$

$$\left\langle (\mu_{\nu|i} - \nu)^2 \right\rangle = \int E[(\mu_{\nu|i} - \nu)^2|i]\, p(i)di = \int \sigma_{\nu|i}^2 p(i)di = \langle \sigma_{\nu|i}^2 \rangle. \quad \text{(A6)}$$

Note that in (A5) and (A6), $E[\ldots\ |i]$ is the conditional expectation given the initial condition i, and we used the rule that the total expectation is the expectation of the conditional expectations.

[57] In general, $STR = \langle \mu_{\nu|i}^2 \rangle / (\langle \mu_{\nu|i}^2 \rangle + \langle \sigma_{\nu|i}^2 \rangle)$ is called the signal (variance)-to-total (variance) ratio, and $SNR = \langle \mu_{\nu|i}^2 \rangle / \langle \sigma_{\nu|i}^2 \rangle$ is called the signal (variance)-to-noise (variance) ratio. Thus, we have

$$STR = \frac{SNR}{SNR + 1}, \quad \text{(A7)}$$

$$AC_p = \sqrt{STR} = \sqrt{\frac{SNR}{SNR + 1}}, \quad \text{(A8)}$$

and

$$MSSS_p = STR = \frac{SNR}{SNR + 1}. \quad \text{(A9)}$$

Note that (A8) and (A9) hold without demanding a Gaussian assumption, as can be seen in (A3) to (A6).

[58] If the forecast and climatological distributions are Gaussian, we have

$$MI = \frac{1}{2}\left(\ln \sigma_\nu^2 - \left\langle \ln \sigma_{\nu|i}^2 \right\rangle\right) \geq \frac{1}{2}\left(\ln \sigma_\nu^2 - \ln\left\langle \sigma_{\nu|i}^2 \right\rangle\right) = -\frac{1}{2}\ln\left(\frac{\left\langle \sigma_{\nu|i}^2 \right\rangle}{\sigma_\nu^2}\right) = -\frac{1}{2}\ln(1 - STR). \quad \text{(A10)}$$

The inequality in (A10) is due to the fact that the arithmetic mean is larger than or equal to the geometric mean or, more strictly, is a result of Jensen's inequality from information theory. Therefore, we have

$$AC_p \leq \sqrt{1 - \exp(-2MI)}, \quad \text{(A11)}$$

$$MSSS_p \leq 1 - \exp(-2MI). \quad \text{(A12)}$$

The equalities in (A11), (A12), and (A10) hold if and only if $\sigma_{\nu|i}^2$ is constant. The conditions that the forecast and climatological distributions are both Gaussian, and that the forecast variance $\sigma_{\nu|i}^2$ is constant, are equivalent to the condition that i and ν are joint normally distributed [*DelSole and Tippett*, 2007]. If i and ν are joint normally distributed, the probability distributions $p(i)$, $p(\nu)$ and $p(\nu|i)$ are all Gaussian distributions, and we obtain [*DelSole and Tippett*, 2007; *Tippett et al.*, 2010]

$$\mu_{\nu|i} = \rho_0 \frac{\sigma_\nu}{\sigma_i}(i - \langle i \rangle), \quad \text{(A13)}$$

$$\sigma_{\nu|i}^2 = \left(1 - \rho_0^2\right)\sigma_\nu^2 = const, \quad \text{(A14)}$$

and

$$MI = -\frac{1}{2}\ln\left(1 - \rho_0^2\right), \quad \text{(A15)}$$

where ρ_0 is the linear correlation between the initial state i and the future state ν. As mentioned in section 3.2, MI measures the statistical dependence between i and ν. As can be seen from (A15), the statistical dependence reduces to a linear correlation ρ_0 if the two variables are joint normally distributed. Because the conditional mean $\mu_{\nu|i}$ is a linear function of the initial state i (see (A13)), ρ_0 is also the linear correlation between $\mu_{\nu|i}$ and ν, which is the potential anomaly correlation skill AC_p in (A3). Note that if i and ν are not joint normally distributed, ρ_0 is usually different from AC_p.

Appendix B: The Theoretical Relationship Between RE and C

[59] If the forecast and climatological distributions are Gaussian, and the forecast variance is constant, the RE-C correlation reduces to the SNEMS-C correlation. The SNEMS is proportional to $\mu_{\nu|i}^2$, where $\mu_{\nu|i}$ is the ensemble mean. If the model is perfect, the actual observation O can be described as the model random variable ν. Thus, C is proportional to $\mu_{\nu|i}\nu$. The square of the correlation between the RE and the correlation contribution C is

$$\begin{aligned}
\rho^2(RE, C) &= \rho^2(SNEMS, C) \\
&= \frac{\langle (SNEMS - \langle SNEMS \rangle)(C - \langle C \rangle) \rangle^2}{\left\langle (SNEMS - \langle SNEMS \rangle)^2 \right\rangle \left\langle (C - \langle C \rangle)^2 \right\rangle} \\
&= \frac{\left\langle \left(\mu_{\nu|i}^2 - \left\langle \mu_{\nu|i}^2 \right\rangle\right)\left(\mu_{\nu|i}\nu - \left\langle \mu_{\nu|i}^2 \right\rangle\right) \right\rangle^2}{\left\langle \left(\mu_{\nu|i}^2 - \left\langle \mu_{\nu|i}^2 \right\rangle\right)^2 \right\rangle \left\langle \left(\mu_{\nu|i}\nu - \left\langle \mu_{\nu|i}^2 \right\rangle\right)^2 \right\rangle} \\
&= \frac{\left\langle \left(\mu_{\nu|i}^2 - \left\langle \mu_{\nu|i}^2 \right\rangle\right)^2 \right\rangle^2}{\left\langle \left(\mu_{\nu|i}^2 - \left\langle \mu_{\nu|i}^2 \right\rangle\right)^2 \right\rangle \left\langle \mu_{\nu|i}^2\nu^2 - 2\mu_{\nu|i}\nu + \left\langle \mu_{\nu|i}^2 \right\rangle^2 \right\rangle} \\
&= \frac{\left\langle \left(\mu_{\nu|i}^2 - \left\langle \mu_{\nu|i}^2 \right\rangle\right)^2 \right\rangle}{\left\langle \mu_{\nu|i}^2\nu^2 - 2\mu_{\nu|i}\nu + \left\langle \mu_{\nu|i}^2 \right\rangle^2 \right\rangle} \\
&= \frac{\left\langle \mu_{\nu|i}^4 \right\rangle - \left\langle \mu_{\nu|i}^2 \right\rangle^2}{\left\langle \mu_{\nu|i}^2\nu^2 \right\rangle - 2\langle \mu_{\nu|i}\nu \rangle + \left\langle \mu_{\nu|i}^2 \right\rangle^2} \\
&= \frac{\left\langle \mu_{\nu|i}^4 \right\rangle - \left\langle \mu_{\nu|i}^2 \right\rangle^2}{\left\langle \mu_{\nu|i}^4 \right\rangle - \left\langle \mu_{\nu|i}^2 \right\rangle^2 + \left\langle \mu_{\nu|i}^2 \right\rangle \sigma_{\nu|i}^2} \\
&= \frac{2\left\langle \mu_{\nu|i}^2 \right\rangle^2}{2\left\langle \mu_{\nu|i}^2 \right\rangle^2 + \left\langle \mu_{\nu|i}^2 \right\rangle \sigma_{\nu|i}^2} \\
&= \frac{2SNR}{2SNR + 1} = \frac{2STR}{1 + STR},
\end{aligned}$$

where we previously used

$$\left\langle \mu_{\nu|i}^2 \nu^2 \right\rangle = \int E[\mu_{\nu|i}^2 \nu^2|i]p(i)di = \int \mu_{\nu|i}^2 E[\nu^2|i]p(i)di = \int \left(\mu_{\nu|i}^4 + \mu_{\nu|i}^2 \sigma_{\nu|i}^2\right)p(i)di = \left\langle \mu_{\nu|i}^4 \right\rangle + \left\langle \mu_{\nu|i}^2 \right\rangle \sigma_{\nu|i}^2,$$

and the fact that $\langle \mu_{\nu l i}^4 \rangle = 3\langle \mu_{\nu l i}^2 \rangle^2$ for normally distributed variable.

[60] According to Appendix A,

$$\rho^2 = \frac{2r^2}{1+r^2},$$

where r is the overall AC skill.

[61] **Acknowledgments.** This work was supported by an NSERC Discovery Grant to Y. Tang. DY is grateful to the China Scholarship Council and Nanjing University for supporting his visit at the University of Northern British Columbia. This study was jointly supported by the National Basic Research Program of China (973 Program) (2012CB955901), the National Natural Science Foundation of China (41130963 and 40730953), the National Public Benefit Research Foundation of China (GYHY201006019), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (20100091110003). The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (contract 505539) whose support is gratefully acknowledged.

References

Anderson, J. L., and W. F. Stern (1996), Evaluating the potential predictive utility of ensemble forecasts, *J. Clim.*, *9*, 260–269, doi:10.1175/1520-0442(1996)009<0260:ETPPUO>2.0.CO;2.

Batté, L., and M. Déqué (2011), Seasonal predictions of precipitation over Africa using coupled ocean–atmosphere general circulation models: Skill of the ENSEMBLES project multimodel ensemble forecasts, *Tellus, Ser. A*, *63*, 283–299, doi:10.1111/j.1600-0870.2010.00493.x.

Buizza, R., and T. N. Palmer (1998), Impact of ensemble size on ensemble prediction, *Mon. Weather Rev.*, *126*, 2503–2518, doi:10.1175/1520-0493(1998)126<2503:IOESOE>2.0.CO;2.

Charney, J. G., and J. Shukla (1981), Predictability of monsoons, in *Monsoon Dynamics*, edited by J. Lighthill and R. P. Pearce, pp. 99–110, Cambridge Univ. Press, Cambridge, U. K., doi:10.1017/CBO9780511897580.009.

Cherchi, C., and A. Navarra (2003), Reproducibility and predictability of Asian summer monsoon in the ECHAM4-GCM, *Clim. Dyn.*, *20*, 365–379.

Cover, T. M., and J. A. Thomas (1991), *Elements of Information Theory*, 576 pp., John Wiley, Hoboken, N. J., doi:10.1002/0471200611.

DelSole, T. (2004), Predictability and information theory. Part I: Measures of predictability, *J. Atmos. Sci.*, *61*, 2425–2440, doi:10.1175/1520-0469(2004)061<2425:PAITPI>2.0.CO;2.

DelSole, T. (2005), Predictability and information theory. Part II: Imperfect forecasts, *J. Atmos. Sci.*, *62*, 3368–3381, doi:10.1175/JAS3522.1.

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

Fu, X., B. Wang, D. E. Waliser, and L. Tao (2007), Impact of atmosphere–ocean coupling on the predictability of monsoon intraseasonal oscillations, *J. Atmos. Sci.*, *64*, 157–174, doi:10.1175/JAS3830.1.

Goswami, B. N. (1998), Interannual variations of Indian summer monsoon in a GCM: External conditions versus internal feedbacks, *J. Clim.*, *11*, 501–522, doi:10.1175/1520-0442(1998)011<0501:IVOISM>2.0.CO;2.

Goswami, B. N., G. Wu, and T. Yasunari (2006), The annual cycle, intraseasonal oscillations and roadblock to seasonal predictability of the Asian summer monsoon, *J. Clim.*, *19*, 5078–5099, doi:10.1175/JCLI3901.1.

Hoerling, M. P., and A. Kumar (2002), Atmospheric response patterns associated with tropical forcing, *J. Clim.*, *15*, 2184–2203, doi:10.1175/1520-0442(2002)015<2184:ARPAWT>2.0.CO;2.

Houtekamer, P. L. (1993), Global and local skill forecasts, *Mon. Weather Rev.*, *121*, 1834–1846, doi:10.1175/1520-0493(1993)121<1834:GALSF>2.0.CO;2.

Joe, H. (1989), Relative entropy measures of multivariate dependence, *J. Am. Stat. Assoc.*, *84*, 157–164, doi:10.2307/2289859.

Ju, J., and J. M. Slingo (1995), The Asian summer monsoon and ENSO, *Q. J. R. Meteorol. Soc.*, *121*, 1133–1168, doi:10.1002/qj.49712152509.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP–DOE AMIP-II reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kang, I.-S., and J. Shukla (2006), Dynamic seasonal prediction and predictability, in *The Asian Monsoon*, edited by B. Wang, chap. 15, pp. 585–612, Springer, New York, doi:10.1007/3-540-37722-0_15.

Kang, I.-S., J.-Y. Lee, and C.-K. Park (2004), Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction, *J. Clim.*, *17*, 834–844, doi:10.1175/1520-0442(2004)017<0834:PPOSMP>2.0.CO;2.

Kleeman, R. (2002), Measuring dynamical prediction utility using relative entropy, *J. Atmos. Sci.*, *59*, 2057–2072, doi:10.1175/1520-0469(2002)059<2057:MDPUUR>2.0.CO;2.

Kleeman, R. (2008), Limits, variability and general behavior of statistical predictability of the mid-latitude atmosphere, *J. Atmos. Sci.*, *65*, 263–275, doi:10.1175/2007JAS2234.1.

Kleeman, R., and A. M. Moore (1999), A new method for determining the reliability of dynamical ENSO predictions, *Mon. Weather Rev.*, *127*, 694–705, doi:10.1175/1520-0493(1999)127<0694:ANMFDT>2.0.CO;2.

Krishnamurthy, V., and J. Shukla (2000), Intraseasonal and interannual variability of rainfall over India, *J. Clim.*, *13*, 4366–4377, doi:10.1175/1520-0442(2000)013<0001:IAIVOR>2.0.CO;2.

Krishnamurti, T. N., T. S. V. V. Kumar, and A. K. Mitra (2006), Seasonal climate prediction of Indian summer monsoon, in *The Asian Monsoon*, edited by B. Wang, chap. 14, pp. 553–583, Springer, New York, doi:10.1007/3-540-37722-0_14.

Kumar, A., 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. Clim.*, *13*, 3139–3151, doi:10.1175/1520-0442(2000)013<3139:CITSOT>2.0.CO;2.

Lau, N.-C., and B. Wang (2006), Interaction between the Asian monsoon and the EL Niño/Southern Oscillation, in *The Asian Monsoon*, edited by B. Wang, chap. 12, pp. 479–512, Springer, New York, doi:10.1007/3-540-37722-0_12.

Lee, J.-Y., et al. (2011), How predictable is the Northern Hemisphere summer upper-tropospheric circulation?, *Clim. Dyn.*, *37*, 1189–1203, doi:10.1007/s00382-010-0909-9, in press.

Lee, S.-S., et al. (2011), Deficiencies and possibilities for long-lead coupled climate prediction of the western North Pacific-East Asian summer monsoon, *Clim. Dyn.*, *36*, 1173–1188, doi:10.1007/s00382-010-0832-0.

Leung, L.-Y., and G. R. North (1990), Information theory and climate prediction, *J. Clim.*, *3*, 5–14, doi:10.1175/1520-0442(1990)003<0005:ITACP>2.0.CO;2.

Mahrt, L., and H.-L. Pan (1984), A two-layer model of soil hydrology, *Boundary Layer Meteorol.*, *29*, 1–20, doi:10.1007/BF00119116.

Mason, S. J., and D. B. Stephenson (2008), How do we know whether seasonal climate forecasts are any good, in *Seasonal Climate: Forecasting and Managing Risk*, edited by A. Troccoli et al., chap. 10, pp. 259–290, Springer, Dordrecht, Netherlands, doi:10.1007/978-1-4020-6992-5_10.

Molteni, F. S., S. Corti, L. Ferranti, and J. M. Slingo (2003), Predictability experiments for the Asian summer monsoon: Impact of SST anomalies on interannual and intraseasonal variability, *J. Clim.*, *16*, 4001–4021, doi:10.1175/1520-0442(2003)016<4001:PEFTAS>2.0.CO;2.

Moorthi, S., H.-L. Pan, and P. Caplan (2001), Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system, *Tech. Proced. Bull., 484*, 14 pp., Off. of Meteorol., Natl. Weather Serv., Silver Spring, Md. [Available online at http://www.nws.noaa.gov/om/tpb/484.htm]

Murphy, A. H., and E. S. Epstein (1989), Skill scores and correlation coefficients in model verification, *Mon. Weather Rev.*, *117*, 572–582, doi:10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2.

Newman, M., P. D. Sardeshmukh, C. R. Winkler, and J. S. Whitaker (2003), A study of subseasonal predictability, *Mon. Weather Rev.*, *131*, 1715–1732, doi:10.1175//2558.1.

Pacanowski, R. C., and S. M. Griffies (1998), MOM3.0 manual, Geophys. Fluid Dyn. Lab., NOAA, Princeton, N. J.

Palmer, T. N. (1994), Chaos and predictability in forecasting the monsoons, *Proc. Indian Natl. Sci. Acad.*, *60A*, 57–66.

Palmer, T. N. (2000), Predicting uncertainty in forecasts of weather and climate, *Rep. Prog. Phys.*, *63*, 71–116, doi:10.1088/0034-4885/63/2/201.

Palmer, T. N., et al. (2004), Development of a European multi-model ensemble system for seasonal to interannual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*, 853–872, doi:10.1175/BAMS-85-6-853.

Peng, P., A. Kumar, and W. Wang (2011), An analysis of seasonal predictability in coupled model forecasts, *Clim. Dyn.*, *36*, 637–648, doi:10.1007/s00382-009-0711-8.

Rasmusson, E. M., and T. H. Carpenter (1982), Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/ El Niño, *Mon. Weather Rev.*, *110*, 354–384, doi:10.1175/1520-0493(1982)110<0354:VITSST>2.0.CO;2.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang (2002), An improved in situ and satellite SST analysis for climate, *J. Clim.*, *15*, 1609–1625, doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

Rowell, D. P. (1998), Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Clim.*, *11*, 109–120, doi:10.1175/1520-0442(1998)011<0109:APSPWA>2.0.CO;2.

Saha, S., et al. (2006), The NCEP climate forecast system, *J. Clim.*, *19*, 3483–3517, doi:10.1175/JCLI3812.1.

Sardeshmukh, P. D., G. P. Compo, and C. Penland (2000), Changes of probability associated with El Niño, *J. Clim.*, *13*, 4268–4286, doi:10.1175/1520-0442(2000)013<4268:COPAWE>2.0.CO;2.

Schneider, T., and S. Griffies (1999), A conceptual framework for predictability studies, *J. Clim.*, *12*, 3133–3155, doi:10.1175/1520-0442(1999)012<3133:ACFFPS>2.0.CO;2.

Shukla, J. (1998), Predictability in the midst of chaos: A scientific basis for climate forecasting, *Science*, *282*, 728–731, doi:10.1126/science.282.5389.728.

Sperber, K. R., et al. (2001), Dynamical seasonal predictability of the Asian summer monsoon, *Mon. Weather Rev.*, *129*, 2226–2248, doi:10.1175/1520-0493(2001)129<2226:DSPOTA>2.0.CO;2.

Stern, W. F., and K. Miyakoda (1995), Feasibility of seasonal forecasts inferred from multiple GCM simulations, *J. Clim.*, *8*, 1071–1085, doi:10.1175/1520-0442(1995)008<1071:FOSFIF>2.0.CO;2.

Tang, Y., R. Kleeman, and A. M. Moore (2005), On the reliability of ENSO dynamical predictions, *J. Atmos. Sci.*, *62*, 1770–1791, doi:10.1175/JAS3445.1.

Tang, Y., H. Lin, J. Derome, and M. K. Tippett (2007), A predictability measure applied to seasonal predictions of the Arctic Oscillation, *J. Clim.*, *20*, 4733–4750, doi:10.1175/JCLI4276.1.

Tang, Y., R. Kleeman, and A. M. Moore (2008a), Comparison of information-based measures of forecast uncertainty in ensemble ENSO prediction, *J. Clim.*, *21*, 230–247, doi:10.1175/2007JCLI1719.1.

Tang, Y., H. Lin, and A. M. Moore (2008b), Measuring the potential predictability of ensemble climate predictions, *J. Geophys. Res.*, *113*, D04108, doi:10.1029/2007JD008804.

Tippett, M. K., R. Kleeman, and Y. Tang (2004), Measuring the potential utility of seasonal climate predictions, *Geophys. Res. Lett.*, *31*, L22201, doi:10.1029/2004GL021575.

Tippett, M. K., A. G. Barnston, and T. DelSole (2010), Comments on "finite samples and uncertainty estimates for skill measures for seasonal prediction," *Mon. Weather Rev.*, *138*, 1487–1493, doi:10.1175/2009MWR3214.1.

von Storch, H., and D. P. Baumhefner (1991), Principal oscillation pattern analysis of the tropical 30- to 60-day oscillation. Part II: The prediction of equatorial velocity potential and its skill, *Clim. Dyn.*, *5*, 1–12.

von Storch, H., and J. S. Xu (1990), Principal oscillation pattern analysis of the tropical 30- to 60-day oscillation. Part I: Definition of an index and its prediction, *Clim. Dyn.*, *4*, 175–190.

Wang, B. (2006), *The Asian Monsoon*, Springer, New York.

Wang, B., and Z. Fan (1999), Choice of South Asian summer monsoon indices, *Bull. Am. Meteorol. Soc.*, *80*, 629–638, doi:10.1175/1520-0477(1999)080<0629:COSASM>2.0.CO;2.

Wang, B., R. Wu, and X. Fu (2000), Pacific–East Asian teleconnection: How does ENSO affect East Asian climate?, *J. Clim.*, *13*, 1517–1536, doi:10.1175/1520-0442(2000)013<1517:PEATHD>2.0.CO;2.

Wang, B., I.-S. Kang, and J.-Y. Lee (2004), Ensemble simulations of Asian–Australian monsoon variability by 11 AGCMs, *J. Clim.*, *17*, 803–818, doi:10.1175/1520-0442(2004)017<0803:ESOAMV>2.0.CO;2.

Wang, B., et al. (2005), Fundamental challenge in simulation and prediction of summer monsoon rainfall, *Geophys. Res. Lett.*, *32*, L15711, doi:10.1029/2005GL022734.

Wang, B., et al. (2007), Coupled predictability of seasonal tropical precipitation, *CLIVAR Exch.*, *12*, 17–18.

Wang, B., et al. (2008a), How accurately do coupled climate models predict the Asian-Australian monsoon interannual variability?, *Clim. Dyn.*, *30*, 605–619, doi:10.1007/s00382-007-0310-5.

Wang, B., et al. (2008b), How to measure the strength of the east Asian summer monsoon?, *J. Clim.*, *21*, 4449–4463, doi:10.1175/2008JCLI2183.1.

Webster, P. J., and S. Yang (1992), Monsoon and ENSO: Selectively interactive systems, *Q. J. R. Meteorol. Soc.*, *118*, 877–926, doi:10.1002/qj.49711850705.

Webster, P. J., et al. (1998), Monsoons: Processes, predictability, and prospects for prediction, *J. Geophys. Res.*, *103*, 14,451–14,510.

Weisheimer, A., et al. (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.

Whitaker, J. S., and A. F. Loughe (1998), The relationship between ensemble spread and ensemble mean skill, *Mon. Weather Rev.*, *126*, 3292–3302, doi:10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2.

Wu, R., and B. P. Kirtman (2005), Roles of Indian and Pacific Ocean air-sea coupling in tropical atmospheric variability, *Clim. Dyn.*, *25*, 155–170, doi:10.1007/s00382-005-0003-x.

Wu, R., and B. P. Kirtman (2006), Changes in spread and predictability associated with ENSO in an ensemble coupled GCM, *J. Clim.*, *19*, 4378–4396.

Wu, R., B. P. Kirtman, and K. Pegion (2006), Local air-sea relationship in observations and model simulations, *J. Clim.*, *19*, 4914–4932, doi:10.1175/JCLI3904.1.

Wu, Z., B. Wang, J. Li, and F. F. Jin (2009), An empirical seasonal prediction model of the east Asian summer monsoon using ENSO and NAO, *J. Geophys. Res.*, *114*, D18120, doi:10.1029/2009JD011733.

Xie, P., and P. A. Arkin (1996), Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, *J. Clim.*, *9*, 840–858, doi:10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2.

Xie, S. P., et al. (2009), Indian ocean capacitor effect on Indo-western Pacific climate during the summer following El Niño, *J. Clim.*, *22*, 730–747, doi:10.1175/2008JCLI2544.1.

Yang, S., and W. K.-M. Lau (2006), Interannual variability of the Asian monsoon, in *The Asian Monsoon*, edited by B. Wang, chap. 6, pp. 259–293, Springer, New York, doi:10.1007/3-540-37722-0_6.

Yang, S., et al. (2008), Simulations and seasonal prediction of the Asian summer monsoon in the NCEP climate forecast system, *J. Clim.*, *21*, 3755–3775, doi:10.1175/2008JCLI1961.1.

Yang, X.-Q., J. L. Anderson, and W. F. Stern (1998), Reproducible forced modes in AGCM ensemble integrations and potential predictability of atmospheric seasonal variations in the extratropics, *J. Clim.*, *11*, 2942–2959, doi:10.1175/1520-0442(1998)011<2942:RFMIAE>2.0.CO;2.

Zhou, T., and L. Zou (2010), Understanding the predictability of East Asian summer monsoon from the reproduction of land-sea thermal contrast change in AMIP-type simulation, *J. Clim.*, *23*, 6009–6026, doi:10.1175/2010JCLI3546.1.

Zhou, T., et al. (2009), The CLIVAR C20C project: Which components of the Asian-Australian monsoon circulation variations are forced and reproducible?, *Clim. Dyn.*, *33*, 1051–1068, doi:10.1007/s00382-008-0501-8.

Y. Tang, Environmental Science and Engineering, University of Northern British Columbia, Prince George, BC V2N 4Z9, Canada. (ytang@unbc.ca)

D. Yang, X. Yang, and Y. Zhang, School of Atmospheric Sciences, Nanjing University, Nanjing 210093, China.