Relationship between predictability and forecast skill of ENSO on various time scales

Yanjie Cheng,[1,2] Youmin Tang,[1] and Dake Chen[3,4]

Received 27 April 2011; revised 15 September 2011; accepted 16 September 2011; published 7 December 2011.

[1] Ensemble predictions are performed using the LDEO5 model for the period from 1856 to 2003 based on a well developed El Niño–Southern Oscillation (ENSO) ensemble system. Information-based and ensemble-based potential predictability measures of ENSO are explored using ensemble predictions and the recently developed framework of predictability. Relationships of these potential predictability measures and actual predictability measures are investigated on multiple time scales from interannual to decadal. Results show that among three information-based potential predictability measures, relative entropy (RE) is better than predictive information (PI) and predictive power (PP) in quantifying correlation-based prediction skill, whereas PI and PP are better indicators in estimating mean square error (MSE)-based prediction skill. The primary reason for these relationships is analyzed and the control factors of the potential predictability measures are identified. It is found that RE is dominated by the signal component, but the dispersion component has a comparable contribution during weak ENSO periods.

Citation: Cheng, Y., Y. Tang, and D. Chen (2011), Relationship between predictability and forecast skill of ENSO on various time scales, *J. Geophys. Res.*, *116*, C12006, doi:10.1029/2011JC007249.

1. Introduction

[2] Predictability is the extent to which events can be predicted [e.g., *DelSole*, 2004]. Generally, there are two types of predictability measures: one consists of actual measures that make use of observations, and the other consists of potential predictability measures that do not make use of observations. The actual predictability quantifies the prediction accuracy of models against observations, measured using either deterministic or probabilistic scores, whereas the potential predictability addresses the quantification of predictable components, often referred to as the ideal skill or the upper limit of skill. A common strategy in potential predictability study is to assume model to be perfect, under which the impact of initial uncertainty or stochastic (chaotic) noise, which is intrinsic to the system itself, can be explored. For example, *Goswami and Shukla* [1991] studied the potential predictability of ENSO by exploring the initial perturbation growth in the ZC model [*Zebiak and Cane*, 1987] in such a manner. An arbitrary model integration is usually used as the "observation" to measure the skill of another integration which has a subtly different initial condition generated by optimal perturbation. The skill of the perfect model is then used to gain insights into the potential predictability. Traditionally there are two methods for constructing optimal initial perturbations: singular vector analysis (SV) or breeding vector analysis (BV). Both methods have been widely used to study ENSO predictability and produce ensemble predictions [e.g., *Chen et al.*, 1997; *Xue et al.*, 1997a, 1997b; *Fan et al.*, 2000; *Tang et al.*, 2006; *Cheng et al.*, 2010a; *Toth and Kalnay*, 1993; *Cai et al.*, 2003; *Tang and Deng*, 2010a, 2010b].

[3] Studies of potential predictability often use ensemble predictions, which assemble the evolutions of initial perturbations (errors). If the ensemble size were large enough, the forecast probability density function (pdf) could be estimated and the prediction uncertainty could be precisely measured. However, a very large ensemble size is presently not practical due to its vast computational expenses. Instead, several useful statistic metrics are developed to measure the potential predictability, including signal characterized by ensemble mean (EM), noise represented by ensemble spread (ES) and the ratio of signal over noise [e.g., *Buizza and Palmer*, 1998; *Moore and Kleeman*, 1998; *Scherrer et al.*, 2004]. A large ES generally suggests a relatively low predictability in ensemble weather forecasts. However, these ensemble potential measures have often met with challenges and limitations in quantifying ENSO and climate prediction skill [e.g., *Kumar et al.*, 2000; *Tang et al.*, 2005, 2007, 2008a, 2008b]. For example, using 20 year retrospective predictions of two hybrid ENSO models, *Tang et al.* [2008b] found that ES is not a good predictor in quantifying climate

[1]Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada.
[2]National Climate Center, China Meteorological Administration, Beijing, China.
[3]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA.
[4]State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou, China.

0148-0227/11/2011JC007249

prediction skill in comparison with the ensemble mean square (EM2). An important task in predictability study is to quantitatively estimate actual skill by potential predictability measures, by which the degree of confidence to be placed in an individual forecast can be assessed [*Moore and Kleeman*, 1998; *Tang et al.*, 2008a].

[4] Recently, new ideas from information theory have been applied to ENSO and seasonal climate predictability [e.g., *DelSole and Tippett*, 2007]. Several information-based potential measures have been used to qualify the potential predictability, such as relative entropy (RE), predictive information (PI), predictive power (PP), and mutual information (MI) [*Schneider and Griffies*, 1999; *Kleeman*, 2002, 2008; *Tippett et al.*, 2004; *Tang et al.*, 2005, 2008a; *DelSole*, 2004, 2005; *DelSole and Tippett*, 2007, 2008]. Like the ensemble-based potential prediction skill metrics, information-based skill metrics also have the property that they are measures of predictability without the use of observations in the framework of perfect model assumption. Information-based measures have several important characteristics. First, for a normally distributed, stationary, Markov process, predictability declines monotonically with the length of the forecast [*Kleeman*, 2002; *DelSole*, 2004; *Tang et al.*, 2008a]. Second, at a given lead time, the averaged RE and PI over a long time period should be identical, equal to MI [*DelSole and Tippett*, 2007], which indicates the overall prediction skill of a forecast system. *Tang et al.* [2008a] first examined these characteristics of information-based potential predictability measures for ENSO retrospective predictions using two realistic hybrid models. Their results show that the aforementioned characteristics of information-based measures generally held well in their ENSO models. However, their study was limited to the 18 year period from 1981 to 1998, which covers only a few ENSO cycles, thus basically precluding statistically robust conclusions. In addition, it has been recognized that the ENSO variability and its prediction skill have striking decadal/interdecadal variations [e.g., *Zhang et al.*, 1997; *Kleeman et al.*, 1999; *An and Wang*, 2000; *Chen et al.*, 2004; *Tang et al.*, 2008c; *An*, 2004, 2009]. One might be able to shed light on the mechanism of such decadal/interdecadal variability by exploring the decadal/interdecadal variations of potential predictability. Obviously, the previous analysis based on an 18 year period is unable to achieve this goal.

[5] *Chen et al.* [2004] used KAPLAN sea surface temperature anomaly (SSTA) reanalysis data [*Kaplan et al.*, 1998] and the latest version of the ZC model (LDEO5 hereafter) to perform a 148 year hindcast experiment for the period of 1856–2003, with the SSTA being the only data used for model initialization. *Tang et al.* [2008c] further analyzed the interdecadal variation in ENSO prediction skill from 1881 to 2000 using multiple models. These retrospective ENSO predictions allowed us to achieve a robust and stable estimate of ENSO potential predictability, and to investigate the decadal/interdecadal variation of the predictability. Here we explore both information-based and ensemble-based potential predictability measures for ENSO using long-term retrospective ensemble predictions from 1856 to 2003 with the LDEO5 model. Relationships between actual prediction skill measures and potential predictability measures on various time scales from years to decades are investigated. Some theoretical properties of the information-based predictability measures are examined using the 148 year LDEO5 ensemble predictions. With such a long-term retrospective ensemble forecast, new findings and understanding in ENSO predictability can be expected.

[6] This paper is structured as follows: Section 2 briefly describes the LDEO5 model and the method of ensemble construction. Section 3 gives the definitions of the actual skill measures, the ensemble-based potential predictability measures, and the information-based potential predictability measures. Section 4 examines the information-based potential predictability measures using the 148 year retrospective ensemble prediction, in comparisons with previous findings of *Tang et al.* [2008a]. Section 5 discusses the relationship between information-based potential predictability measures and actual prediction skill on different time scales, followed by a summary in section 6.

2. Prediction Model and Ensemble Scheme

2.1. The LDEO5 Model

[7] The model used in this study is the ZC model [*Zebiak and Cane*, 1987], which has been widely applied to ENSO simulation and prediction. LDEO5 is the latest version of the ZC model [*Chen et al.*, 2004]. The atmosphere dynamics follows *Gill* [1980] using steady state, linear shallow-water equations, with the circulation forced by a heating anomaly which depends on SSTA and moisture convergence. The ocean dynamics uses a reduced-gravity model, with ocean currents generated by surface winds. An oceanic mixed layer is included to account for the three-dimensional thermodynamic processes that control SSTA. The model domain covers the tropical Pacific Ocean (124°E-80°W; 28.75°S-28.75°N), with a time step of 10 days. The model grid for ocean dynamics is 2° longitude × 0.5° latitude, while that for SST physics and the atmospheric model is 5.625° longitude × 2° latitude.

[8] The observational SSTA data set used in this study for model initialization is from the reconstructed analysis of *Kaplan et al.* [1998] for the period from January 1856 to December 2003. We solely rely on SSTA because it is the only oceanic data available for initializing long-term retrospective predictions over 100 years [e.g., *Chen et al.*, 2004; *Tang et al.*, 2008a]. Note that in the coupled initialization procedure of the LDEO forecast system, assimilating SSTA data is not simply putting a constraint on the ocean model at the surface; the SSTA data translate to surface wind field and subsurface ocean memory. There are two model output statistics (MOS) schemes to correct systematic model biases in LDEO5, one for SSTA and the other for thermocline depth and wind field. These bias correction schemes are interactive, in the sense that corrections are calculated at each time step from projections of predetermined error modes, thus affecting the coupling cycle of the model [*Chen et al.*, 2000, 2004]. The implementation of these schemes makes the mode behavior more realistic and largely eliminates the incompatibility between the model and observations.

2.2. The Strategy of Ensemble Construction

[9] The strategy of ensemble construction in this study aims at two major sources of uncertainties in ENSO prediction: the errors in initial conditions and the external stochastic atmospheric noise during the forecast period [e.g.,

Moore and Kleeman, 1998]. Thus, a joint perturbation, composed of the leading singular vector of SST (SV1_sst) perturbation of the initial conditions and the leading stochastic optimal perturbation of winds (SO1_wind) during the whole forecast period, was applied to construct ensemble predictions. The SV1_sst and SO1_wind represent the optimal growth of perturbation due to uncertainties in initial SSTA and atmospheric transients, respectively. They were obtained by perturbing the tangent linear model (TLM) of the LDEO5 model as outlined in our previous work [*Cheng et al.*, 2010a, 2010b]. It was found that this joint perturbation is able to provide reliable and skillful ENSO probabilistic forecasts for the LDEO5 model, as validated by the reliability term and the resolution term of the Brier score (BS). The details of the ensemble construction by this joint perturbation and resultant probabilistic prediction are given by *Cheng et al.* [2010b].

[10] The model was initialized by assimilating SSTA every month for the 1856–2003 period [*Chen et al.*, 2004]. Thus a total of 148 years × 12 months/year (=1776) forecast initial conditions were obtained. From each initial time, an ensemble forecast was performed with an ensemble size (M) of 100 and a lead time of 24 months. Therefore, a total of 1776 × 100 × 24 (4262400) forecasts are available for our predictability analyses.

3. Prediction Skill Metrics

3.1. Metrics of Ensemble Mean Prediction Skill

[11] Correlation-based skill and MSE-based skill are used to measure deterministic prediction skill. The overall skill of ensemble mean predictions over the 148 years is measured by anomaly correlation (R) and root mean square error (RMSE) [e.g., *Scherrer et al.*, 2004]. To evaluate the skill of an individual prediction, the mean square error of individual prediction (MSEIP) and anomaly correlation of individual prediction (CIP) are estimated for all lead times up to 24 months, as defined by *Moore and Kleeman* [1998] and *Tang et al.* [2008a, 2008b]

$$MSEIP(i) = \frac{1}{100}\sum_{m=1}^{M=100}\left[\frac{1}{24}\sum_{t=1}^{t=24}\left(T^p(i,t,m) - T^o(i,t)\right)^2\right] \quad (1)$$

$$CIP(i) = \frac{1}{100}\sum_{m=1}^{M=100}\left\{\frac{\sum_{t=1}^{t=24}\left[T_t^p(i,t,m)T_t^o(i,t)\right]}{\sqrt{\sum_{i=1}^{t=24}\left(T_i^p(i,t,m)\right)^2}\sqrt{\sum_{i=1}^{t=24}\left(T_i^o(i,t)\right)^2}}\right\} \quad (2)$$

where T is the index of NINO3.4 SSTA (averaged over 5°N to 5°S, 170°W to 120°W), t is the lead time of the prediction from 1 to 24 months, M is the ensemble size (100 here), T^p is the predicted NINO3.4 SSTA, and T^o is the observed counterpart. Subscript i denotes the initial time of prediction (i = 1, …, N), and N is the number of samples used over 148 years in this study. A total of 148 × 12 (N = 1776) forecasts, initialized from January 1856 to December 2003, were started at each month and carried on for 24 months using the LDEO5 model. The seasonal cycle was always removed from forecasts and observations prior to measuring prediction skill.

3.2. Ensemble-Based Measures of Potential Predictability

[12] Ensemble mean (EM or μ_p), ensemble spread (ES or σ_p), and ensemble ratio (ER or λ_p) are common ensemble-based measures of potential predictability that do not make use of observations. They are defined as

$$\mu_p(i,t) = \frac{1}{M}\sum_{m=1}^{M=100} T^p(i,t,m) \quad (3)$$

$$\sigma_p(i,t) = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}\left(T_p(i,t,m) - \mu_p(i,t)\right)^2} \quad (4)$$

$$\lambda_p(i,t) = \left|\frac{\mu_p(i,t)}{\sigma_p(i,t)}\right| \quad (5)$$

where μ_p, σ_p, and λ_p are functions of initial time i and lead time t. T is again the index of Niño3.4 SSTA, and the subscripts p denotes predictions (forecasts). Note that instead of using EM, the square of ensemble mean, denoted by EM2 or μ_p^2, is used as a potential skill measure in this study since it is a better indicator of the magnitude of ENSO signal as suggested by *Tang et al.* [2008a].

3.3. Information-Based Measures of Potential Predictability

[13] We now give a review of information-based measures of potential predictability. Further details can be found in relevant literature [e.g., *DelSole*, 2004; *DelSole and Tippett*, 2007; *Kleeman*, 2002; *Tang et al.*, 2008a]. In general, the information-based potential predictability is a measure of the difference between two probability distributions: the forecast distribution $p(v|\Theta)$ and climatological distribution $q(v)$, with

$$p(v \mid \Theta) = \int p(v \mid i)p(i \mid \Theta)di \quad (6)$$

where conditional probability $p(A|B)$ denotes the probability of A event when B event has happened, and i, Θ, and v denotes the initial state, the corresponding observation, and the forecast respectively. Equation (6) means that the forecast distribution $p(v|\Theta)$ can be theoretically obtained by the initial analysis probability $p(i|\Theta)$ and the transition probability $p(v|i)$ of a perfect model system [*DelSole and Tippett*, 2007]. If the variable v has no additional predictive information from climatology, the forecast and climatological distributions will be identical, i.e., $p(v|\Theta) = q(v)$.

[14] The variance σ_q^2 of the model climatological distribution $q(v)$ was obtained from 100 ensemble members over the 148 years in this study, and σ_q^2 for 12 calendar months was estimated respectively.

[15] Entropy is a measure of dispersion level (e.g., uncertainty). The entropy of a continuous distribution $p(x)$ is defined as

$$H(x) = -\int p(x)\ln p(x)dx \quad (7)$$

where the integral is understood to be a multiple integral over the domain of x. Larger entropy is associated with smaller probability and larger uncertainty.

[16] A natural measure of predictability is the predictive information (PI), defined as the difference between the entropy of the climatological and forecast distributions

$$PI = H(v) - H(v \mid \Theta) \tag{8}$$

Considering (7), then

$$PI = -\int q(v)\ln[q(v)]dv + \int p(v \mid \Theta)\ln[p(v \mid \Theta)]dv. \tag{9}$$

[17] The first term on the right hand side of equation (9) denotes the entropy of the prior distribution $q(v)$ (climatological distribution), measuring the uncertainty of a prior time when no extra information is provided from observation or model, whereas the second term represents the entropy of the posterior distribution $p(v|\Theta)$ (forecast distribution), measuring the uncertainty after the observation and associated prediction becomes available. Thus a large PI indicates that the posterior uncertainty will decrease because of useful information being provided by a prediction (e.g., the larger $p(v|\Theta)$ the smaller uncertainty); that is, the prediction is to be more reliable in a "perfect model" context.

[18] Another potential predictability measure, predictive power (PP), is defined by *Schneider and Griffies* [1999] based on PI, which has the same interpretation as PI, i.e., smaller PP corresponds to a more uncertain prediction, and vice versa

$$PP = 1 - \exp(-PI). \tag{10}$$

[19] An alternative measure of the difference between two distributions is relative entropy (RE)

$$RE = \int p(v \mid \Theta)\ln\left(\frac{p(v \mid \Theta)}{q(v)}\right)dv \tag{11}$$

where q denotes the climatological distribution and p is that for the prediction.

[20] In the case where the pdfs are Gaussian distributions, which is a good approximation in many practical cases (including ENSO prediction), the relative entropy may be calculated exactly in terms of the predictive and climatological variances, and the difference between their means. The resulting analytical expression for the relative entropy RE is given by *Kleeman* [2002] and *Tang et al.* [2008a]

$$RE = \frac{1}{2}\left\{\ln\left[\frac{\det\left(\sigma_q^2\right)}{\det\left(\sigma_p^2\right)}\right] + trace\left(\sigma_p^2\sigma_q^{-2}\right) + \left(\mu_p - \mu_q\right)^T\sigma_q^{-2}\left(\mu_p - \mu_q\right) - n\right\} \tag{12}$$

where q and p are the climatological and predictive covariance matrices respectively; *det* is the determinant operator and *tr* is the trace operator; μ_q and μ_p are the climatological and predictive mean state vectors of the system, and n is the number of degree of freedom; RE is composed of two components: (1) a reduction in climatological uncertainty by the prediction (the first two terms plus the last term on the right-hand side of (12)) and (2) a difference in the predictive and climatological means (the third term on the rhs of (12)). These components can be interpreted respectively as the dispersion and signal components of the utility of a prediction [*Kleeman*, 2002]. A large value of RE indicates that more information that is different from the climatological distribution is being supplied by the prediction, which could be interpreted as making it more reliable [*Tang et al.*, 2008a].

[21] For a Gaussian distribution, a univariate state vector with a climatological mean of zero, the covariance matrices are scalar variances in equation (12). RE, PI, and PP can be simplified as [*DelSole*, 2004]

$$PI = \frac{1}{2}\ln\left(\frac{\sigma_q^2}{\sigma_p^2}\right) \tag{13}$$

$$RE = \frac{1}{2}\left[\underbrace{\ln\left(\frac{\sigma_q^2}{\sigma_p^2}\right) + \frac{\sigma_p^2}{\sigma_q^2} - 1}_{Dispersion} + \underbrace{\frac{\mu_p^2}{\sigma_q^2}}_{Signal}\right] = PI + \frac{1}{2}\left[\frac{\sigma_p^2}{\sigma_q^2} - 1 + \frac{\mu_p^2}{\sigma_q^2}\right] \tag{14}$$

$$PP = 1 - \left(\frac{\sigma_p^2}{\sigma_q^2}\right)^{1/2}. \tag{15}$$

[22] A key difference between relative entropy (RE) and predictive information (PI) is that RE vanishes if and only if the forecast and climatological distributions are identical (i.e., same mean and spread), while PI is zero as long as the two distributions have the same spread. Remarkably, predictive information and relative entropy are invariant with respect to linear transformations of the state.

[23] The averaged RE and PI ($\overline{RE}$ and $\overline{PI}$) over all predictions (initial conditions) are identical to MI, as mentioned before. For seasonal climate prediction, the total variance (i.e., climate variance) can be decomposed into signal (S) variance and noise (N) variance, if the signal and noise are assumed to be independent with each other [e.g., *Kumar and Hoerling*, 1998; *Shukla et al.*, 2000], namely,

$$Var(T) = Var(S) + Var(N) \tag{16}$$

where

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

$$Var(S) = \frac{1}{M}\sum_{i=1}^{M}\left(\overline{X}_i - \overline{\overline{X}}\right)^2$$

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

with $X_{i,j}$ being the jth member of the ensemble prediction starting from the ith initial condition, K the ensemble size and M the total number of initial conditions (predictions); and $\overline{X}_i = \frac{1}{K}\sum_{j=1}^{K}X_{i,j}, \overline{\overline{X}} = \frac{1}{M}\sum_{i=1}^{M}\overline{X}_i$.

Table 1. Summary of Prediction Skill Measures Used in This Study[a]

Potential Skill Measure	Actual Skill Measure			
	R(t)	RMSE(t)	CIP(i)	MSEIP(i)
Ensemble-based measure	EM2(i, t)	ES(i, t)	ER(i, t)	–
Information-based measure	RE(i, t)	PI(i, t)	PP(i, t)	MI(t)

[a]Prediction skill measures are functions of either lead time (t), initial time (i), or both.

[24] Without the loss of generality, the climatological mean is assumed to be zero, and thus (16) can be expressed by

$$\overline{\sigma_q^2} = \overline{\mu_p^2} + \overline{\sigma_p^2} \tag{17}$$

where the overbar denotes the expectation over all predictions (initial conditions). Equations (14) and (17) can easily verify the property of *MI*, for example,

$$MI = \overline{RE} = \overline{PI} + \frac{1}{2}\left[\frac{\overline{\mu_p^2} + \overline{\sigma_p^2} - \overline{\sigma_q^2}}{\sigma_q^2}\right] = \overline{PI}. \tag{18}$$

[25] Using (17), the information-based potential predictability measures *MI*, ($\overline{RE}$ or $\overline{PI}$) and $\overline{PP}$ can be rewritten as a function of the mean signal and noise, or their ratio $\overline{\lambda_p}$

$$MI = \overline{RE} = \overline{PI} = \frac{1}{2}\ln\left(\frac{\overline{\sigma_q^2}}{\overline{\sigma_p^2}}\right) = \frac{1}{2}\ln\left(1 + \frac{\overline{\mu_p^2}}{\overline{\sigma_p^2}}\right)$$
$$= \frac{1}{2}\ln\left(1 + \overline{\lambda_p^2}\right) \tag{19}$$

$$PP = 1 - \left(\frac{\overline{\sigma_p^2}}{\overline{\sigma_p^2} + \overline{\mu_p^2}}\right)^{1/2} = 1 - \left(\frac{1}{1 + \overline{\lambda_p^2}}\right)^{1/2}. \tag{20}$$

[26] For Gaussian variables, *MI* follows a theoretical relationship with the actual correlation skill *R* as below [*DelSole*, 2004],

$$MI_{th} = -0.5\ln\left(1 - R^2\right) \tag{21}$$

where the correlation skill *R* is also related to the square of signal-to-noise ratio or ensemble ratio λ_p^2 as proved by *Kleeman and Moore* [1997]

$$R^2 = \frac{1}{1 + \dfrac{\overline{\sigma_p^2}}{\overline{\mu_p^2}}} = 1 - \frac{1}{1 + \overline{\lambda_p^2}}. \tag{22}$$

[27] Equations (16)–(22) build the connections between information-based potential predictability and ensemble-based potential predictability: all the averaged information-based potential predictability measures are equivalent to the ensemble-based potential measure λ_p^2. The theoretical correlation skill *R* also is related to ensemble ratio as indicated in equation (22). On the other hand, the RMSE skill should have a good relationship with the ensemble spread in a perfect ensemble prediction system. Because λ_p^2 includes the signal variance (EM2) and the noise variance (ES), both fundamental to the potential predictability. Thus, in the following sections we mainly explore the features and variations of EM2 and ES, in particular their relationship to the actual forecast skill.

[28] The actual forecast skills and the potential predictability measures used in this study are summarized in Table 1. R, RMSE, CIP, and MSEIP are actual forecast skill measures because they depend on observations whereas EM2, ES, ER, RE, PI, PP, and MI are potential predictability skill measures that do not involve observational data.

4. Characteristics of Potential Predictability Measures

4.1. The Characteristics of RE, PI, and PP

[29] We first explore the properties of the information-based potential predictability measures through the long-term ENSO ensemble forecast with the LDEO5 model. Figure 1 shows RE, PI, and PP in the Nino3.4 region, as a function of lead time and initial condition, for the time period from 1970 to 2003 (Note: the analysis for the entire period of 1856–2003 has been performed, but only plotted the period of 1970–2003 in Figure 1 for a clearer presentation). Several features here are similar to those presented by *Tang et al.* [2008a], where hindcasts from the period of 1981–1998 were obtained with two hybrid ENSO models. These common features are as follows: (1) Large RE peaks are related to strong ENSOs. For example, strong El Niños that occurred in 1972/73, 1982/83, and 1997/98, and La Niñas in 1974/75, 1988/89, and 1999/2000 all have corresponding peaks in the RE plot (Figure 1a); on the other hand, it is difficult to connect ENSO variability with PP since large PP occurred frequently in Figure 1c. (2) RE declines significantly as the lead time of prediction increases, whereas PP displays relatively smooth variations with the lead time and initial conditions. However, one feature is different from that of *Tang et al.* [2008a]. In Figure 1b, PI shows a similar low frequency variability as RE, but similar high frequency variability as PP, as compared to *Tang et al.* [2008a]. The difference is due to the climatological variance σ_q^2. In this study σ_q^2 has an annual cycle, which is more reasonable than the fixed value in previous work, but similar results can be obtained if the annual cycle is not considered in σ_q^2. Mathematically, if we consider σ_q^2 as a fixed value, PI and PP in equations (13) and (15) depend only on the noise variance $\overline{\sigma_p^2}$, and RE depends on both $\overline{\sigma_p^2}$ and signal variance $\overline{\mu_p^2}$. In that case, the high frequency features of PI and PP are obviously determined only by the noise variance, and the low frequency variability features of RE thus is dominated by signal variance. In this study, although both signal and noise variances are involved in all the formulas of the information-based predictability measures, the high frequency features of PI (at short lead times) and PP suggest that PI and PP still depend more on the noise variance than the signal variance; and the low frequency variations of RE depends more on the signal variance. Thus, the finding of *Tang et al.* [2008a] is still valid: RE is a better indicator of ENSO variability than PI and PP.

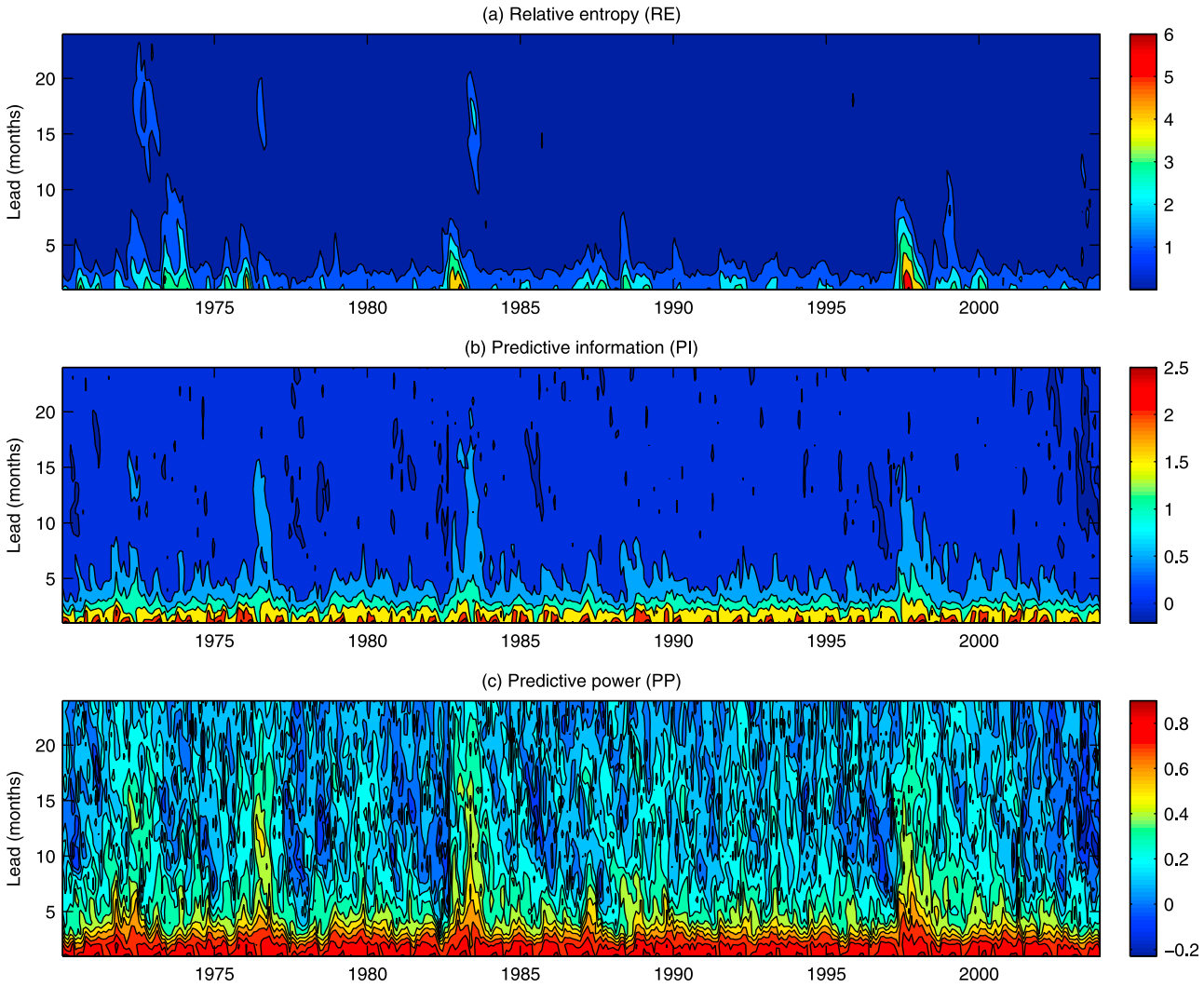

Figure 1. (a) Relative entropy (RE), (b) predictive information (PI), and predictive power (PP) of Niño-3.4 SSTA index as a function of initial time and lead time (months), from January 1970 to December 2003 for the LDEO5 model.

4.2. The Characteristics of EM2, ES, and ER

[30] Shown in Figure 2 are the temporal variations of EM, ES, and ER, as a function of initial condition and lead time. As can be seen in this figure, EM2 presents some features similar to RE but fails to characterize the predictability of some ENSO events such as La Niñas in 1974/1975 and 1999/2000. ER has a low resolution, especially for leads greater than 5 months, making it less informative in characterizing potential predictability associated with ENSO events. On the other hand, the ensemble spread (ES) displays strong high frequency noise that enshrouds signal components significantly. Another feature in Figure 2b is that large and small ES slopes occur annually, indicating a strong seasonal variability of ES. Along each slope, ESs start from different initial time (i) and vary with different lead time (t). However, ESs actually correspond to the same verification time [the time at the end of the forecast or the target time (v), i.e., $v = i + t$]. The target-dependence of ES has been reported by *Karspeck et al.* [2006] when they studied variations of ES with the ZC model (LDEO4). They suggested that the seasonal variability of ES is consistent with the seasonal cycle in ENSO signals; ES can be better presented as a function of the calendar month at the target month of the forecast (regardless of the lead time) than as a function of the initialization month.

[31] An inverse relationship between ES and ENSO signals has been reported in some previous studies, which was explained by the delayed oscillator mechanism [e.g., *Cai et al.*, 2003; *Zhou et al.*, 2008; *Cheng et al.*, 2010a], i.e., a large (small) error growth occurs with weak (strong) signals at the initial time. To further demonstrate this, we binned ES according to the value of Nino3.4 index at initial times in Figure 3a. An inverse ES – signal relationship can be observed for lead times up to 18 months, which is consistent with previous studies. However, if ES is grouped by SSTA index at the target time, the opposite appears as displayed in Figure 3b, i.e., ES increases with the magnitude of Nino3.4 index. Figure 3 suggests that not only the SSTA at the initial

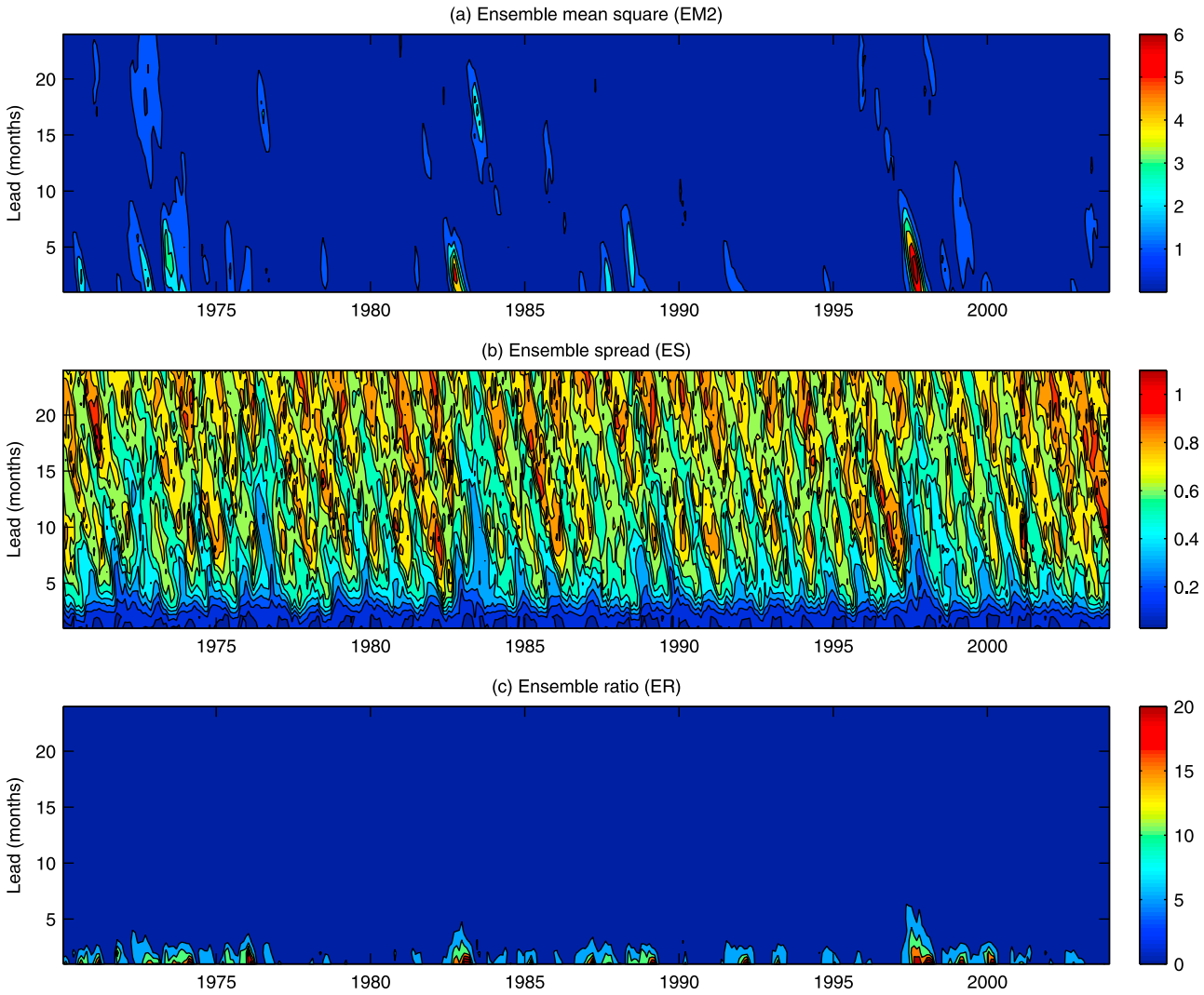

Figure 2. Same as Figure 1 but for the ensemble-based potential predictability measures: (a) ensemble mean square (EM2), (b) ensemble spread (ES), and (c) ensemble ratio (ER).

time, but also that at the target time, have significant impacts on the potential predictability measure ES.

4.3. Characteristics of Mutual Information

[32] According to information theory, equation (18), the averaged value of RE and PI over a long time period should be close to each other, approaching an overall potential predictability measure called the mutual information (MI). MI as a function of lead time is an indicator of the overall predictability of the target variable in a forecast model. This property is well represented in the present ensemble predictions of the LDEO5 model. In Figure 4a, two sets of MIs that were computed from the averaged RE and PI respectively indeed have same values at any given lead time, which is in agreement with equation (18). In addition, MIs decrease smoothly as lead time increases, holding the monotonical property perfectly as they asymptotically approach the minimum value. At lead times of 12 months and longer, MIs vary very smoothly, suggesting that ENSO predictability quickly approaches the minimum at lead times around 12 months and it stays at that level for longer leads in the LDEO5 model.

[33] Figures 4c and 4d show strong relationships between MI and the actual prediction skill (anomaly correlation skill R and RMSE skill). Large MIs are related to small RMSEs but high correlation skill, and vice versa. The monotonical property of MI and its good relations with actual prediction skills indicate that MI is a good indicator of overall predictability.

[34] To examine equation (21) for the LDEO5 model, the MI obtained from the averaged RE and PI are compared with the estimated MI from actual R. As seen in Figure 4b, except for the three largest MIs (corresponding to lead times of 1–3 months), the estimated MIs are very close to the MIs from the averaged RE and PI. The scattered points are distributed along the "perfect" diagonal line for lead times of 4–24 months. A comparison of the LDEO5 model with the hybrid models by *Tang et al.* [2008a] reveals that the LDEO5 model provides a much better validation for the theoretical relationship in equation (21). The largest three MI points are related to the high correlation skills of short

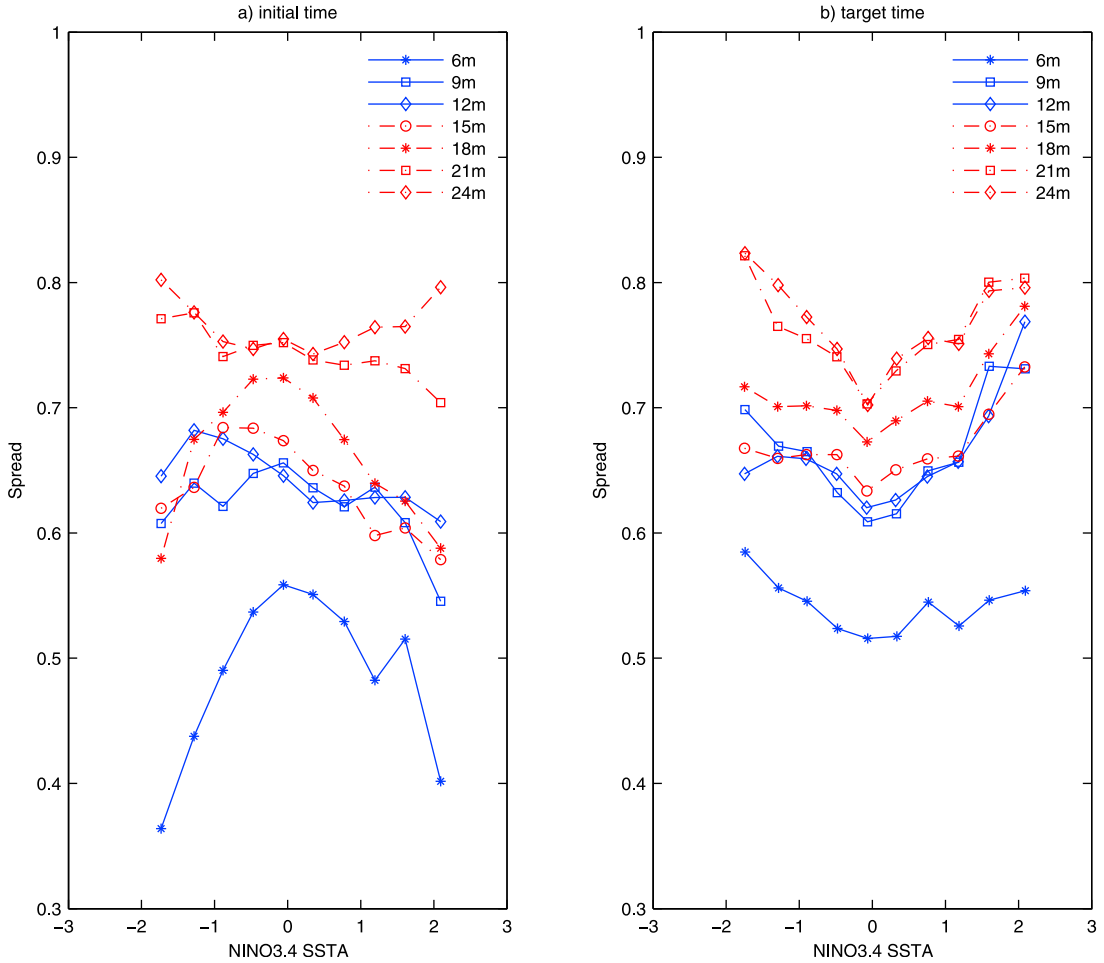

Figure 3. (a) Ensemble spread (°C) as a function of NINO3.4 SSTA index at lead times of 6, 9…, 24 month, respectively. ES is binned by SSTA at the initial time. (b) Same as Figure 3a but at the target time.

lead times (1–3 months), at which the ensemble spread may have not yet fully developed, leading to a large RE/PI. This can be further demonstrated in a "perfect" model scenario, in which the observation is replaced by a randomly selected ensemble member that removes the impact of model bias. Figure 5b presents the MI in this scenario, showing a perfect relationship between MI and correlation skill as indicated by the diagonal line. Comparison between Figures 4 and 5 indicates that there are still some imperfections in using the LDEO5 model to study potential predictability measures, but such imperfections are not very significant. In the following sections, we will use the perfect model scenario to study the potential predictability. Thus the "actual" prediction skill noted below always means the skill using perfect model scenario unless otherwise indicated.

5. Relationship of Potential Predictability Measure and Actual Forecast Skill

[35] In the proceeding sections, we have examined some important features and properties of information-based measures as well as their capability in characterizing the "overall" actual forecast skill. Here we examine the "individual" relationship of potential predictability measures and actual forecast skill (CIP and MSEIP). The identified relationships may not only contribute to the theory of ENSO predictability, but also provide a practical utility in estimating the confidence that we can place in future predictions using the same ENSO forecast model.

[36] Theoretically, a close relation between CIP and RE or EM2 can be expected according to a qualitative analysis as follows. If the observation O is the sum of a predictable signal μ_p and a small unpredictable part ε, $O = \mu_p + \varepsilon$), then the best forecast is for μ_p that is close to ensemble mean EM for a reliable forecast system, and the expected squared correlation CIP^2 between O and μ_p is

$$CIP^2 = \frac{< O\mu_p >^2}{< O^2 >< \mu^2 >} = \frac{< (\mu_p + \varepsilon)\mu_p >^2}{< O^2 >< \mu_p^2 >} = \frac{\mu_p^2}{< O^2 >} \quad (23)$$

RE is decomposed into the signal component (SC) and the dispersion component (DC) according to equation (14)

$$RE = SC + DC \approx \frac{1}{2}\frac{\mu_p^2}{< O^2 >} + DC. \quad (24)$$

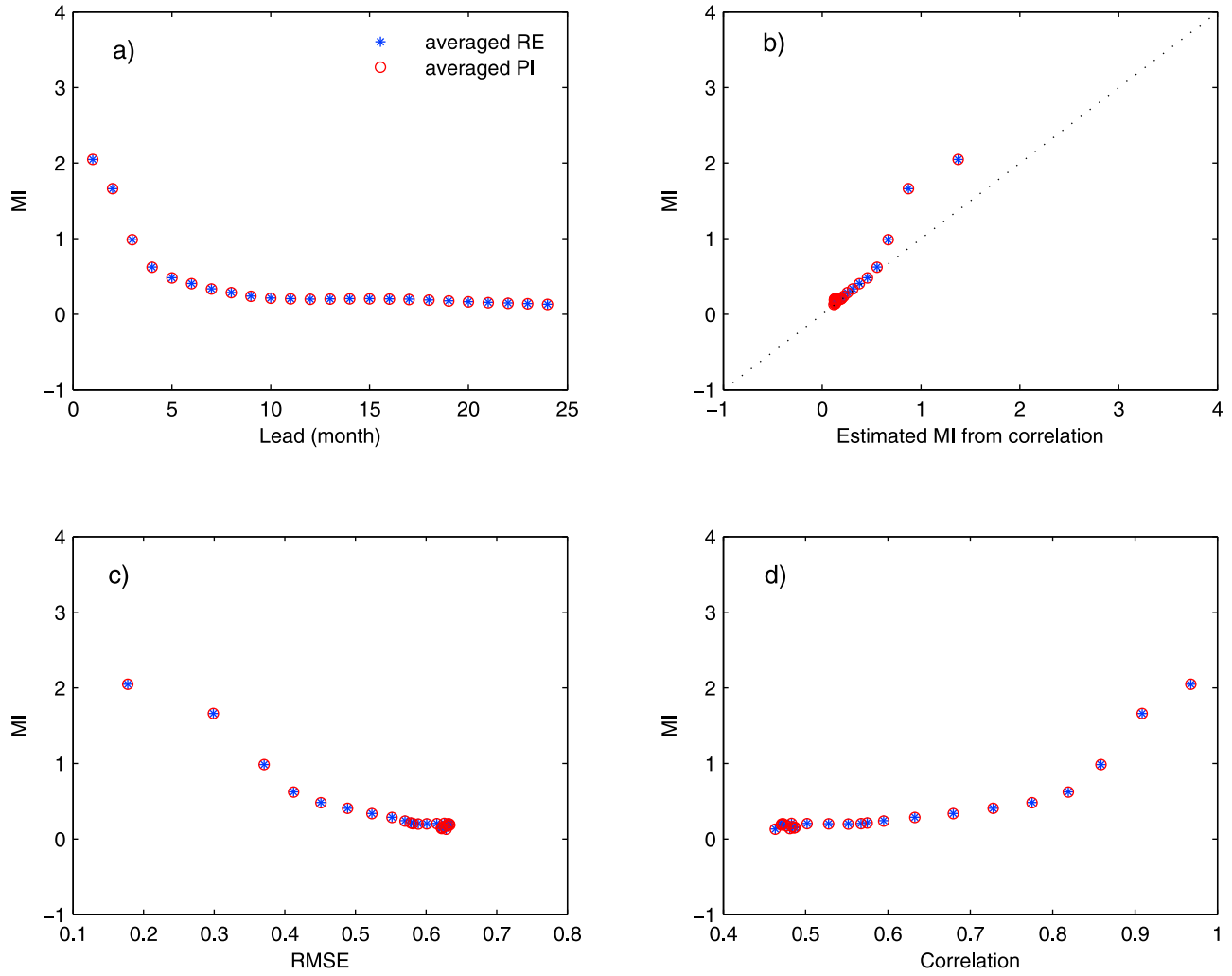

Figure 4. (a) Variation of MI calculated by averaged RE and PI over all initial conditions as functions of lead times. (b) MI from averaged RE and PI against the estimated MI from correlation skill using equation (21). (c) Averaged RE and PI versus RMSE skill. (d) Averaged RE and PP versus correlation skill.

[37] From equations (23) and (24), CIP and RE is expected to show a close relationship when the dispersion component DC is smaller than the signal component SC (If DC is a constant, the conclusion still holds). Note that σ_q^2 in DC is also partly related to μ_p^2. To compare the contributions of SC and DC in RE, Figure 6 shows the temporal variations of RE, SC, and DC over the 148 years. A 2 year running mean was applied to the data to filter out high-frequency variations. Decadal/interdecadal variability can be seen in time series of RE and SC while DC shows very weak variability after applying the 2 year running mean. During the higher SC periods (i.e., the end of 19th century and the end of the 20th century), the signal component SC is much larger than the dispersion term DC; whereas during the weaker SC period (i.e., 1910–1940), DC and SC have comparable contributions to RE. Therefore, the variations of the phase and the amplitude of RE are mainly determined by the signal component SC, but DC is important when ENSO signals are weak. This may explain why in Figures 1 and 2 RE works for weaker ENSO events but EM2 does not, due to ignored contribution of DC in the latter.

[38] The significant contribution of ENSO signal to ENSO predictability is in agreement with many recent ENSO predictability studies [e.g., *Kirtman and Schopf*, 1998; *Tang et al.*, 2008a; *Cheng et al.*, 2010a], that is, stronger ENSOs have a higher predictability. The fundamental reason for decadal/interdecadal ENSO predictability are still not very clear, but the strength of ENSO signal is a key factor [e.g., *Tang et al.*, 2008a; *Cheng et al.*, 2010a].

5.1. Relationship on the Decadal/Interdecadal Time Scale

[39] To explore the decadal/interdecadal relationships, we extract low-frequency components (>10 years) using a fast Fourier transform (FFT) filter. Shown in Figures 7a–7c are the decadal/interdecadal variations of the low-frequency components of averaged RE (or PI, PP) over all leads of 24 months, along with the correlation skill CIP. In-phase relationships can be seen between RE (or PI, PP) and CIP, with positive correlation coefficient of 0.77/0.56/0.54, respectively, which are statistically significant at the 99% confidence level. When REs were relatively large during the late 19th and the late 20th centuries, the correlation skill CIP

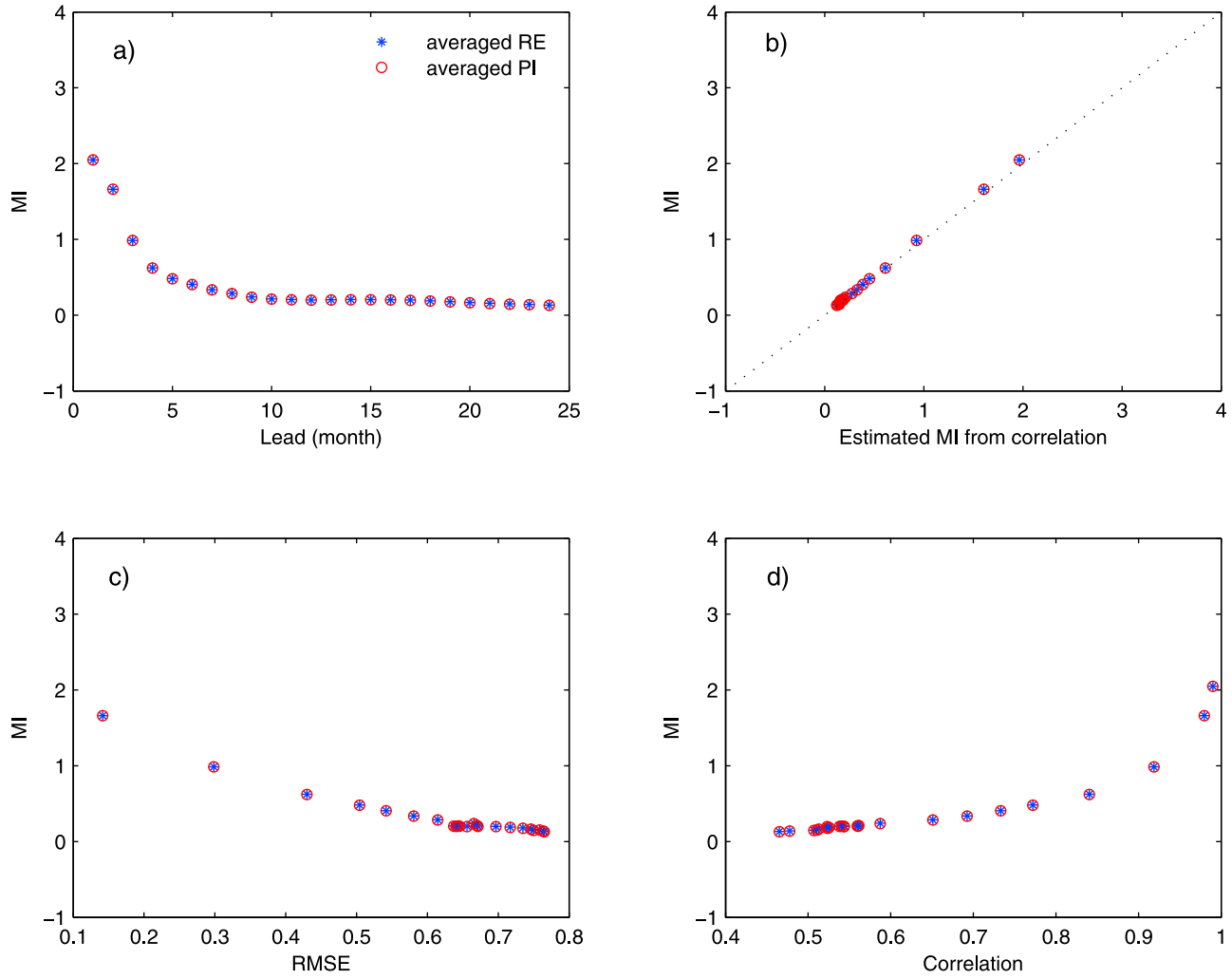

Figure 5. Same as Figure 4 but using the "perfect" model assumption, i.e., observation is replaced by an ensemble member, which is selected randomly from the 100 ensemble member.

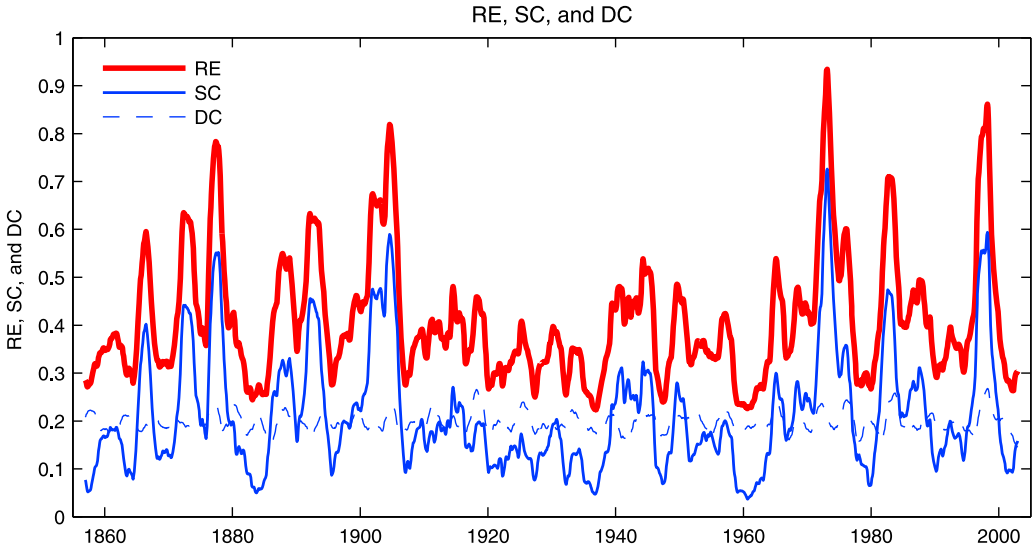

Figure 6. Temporal variations of relative entropy (red solid), the signal component (SC; blue solid) and the dispersion component (DC; blue dashed) of RE (a 2 year running mean is applied).

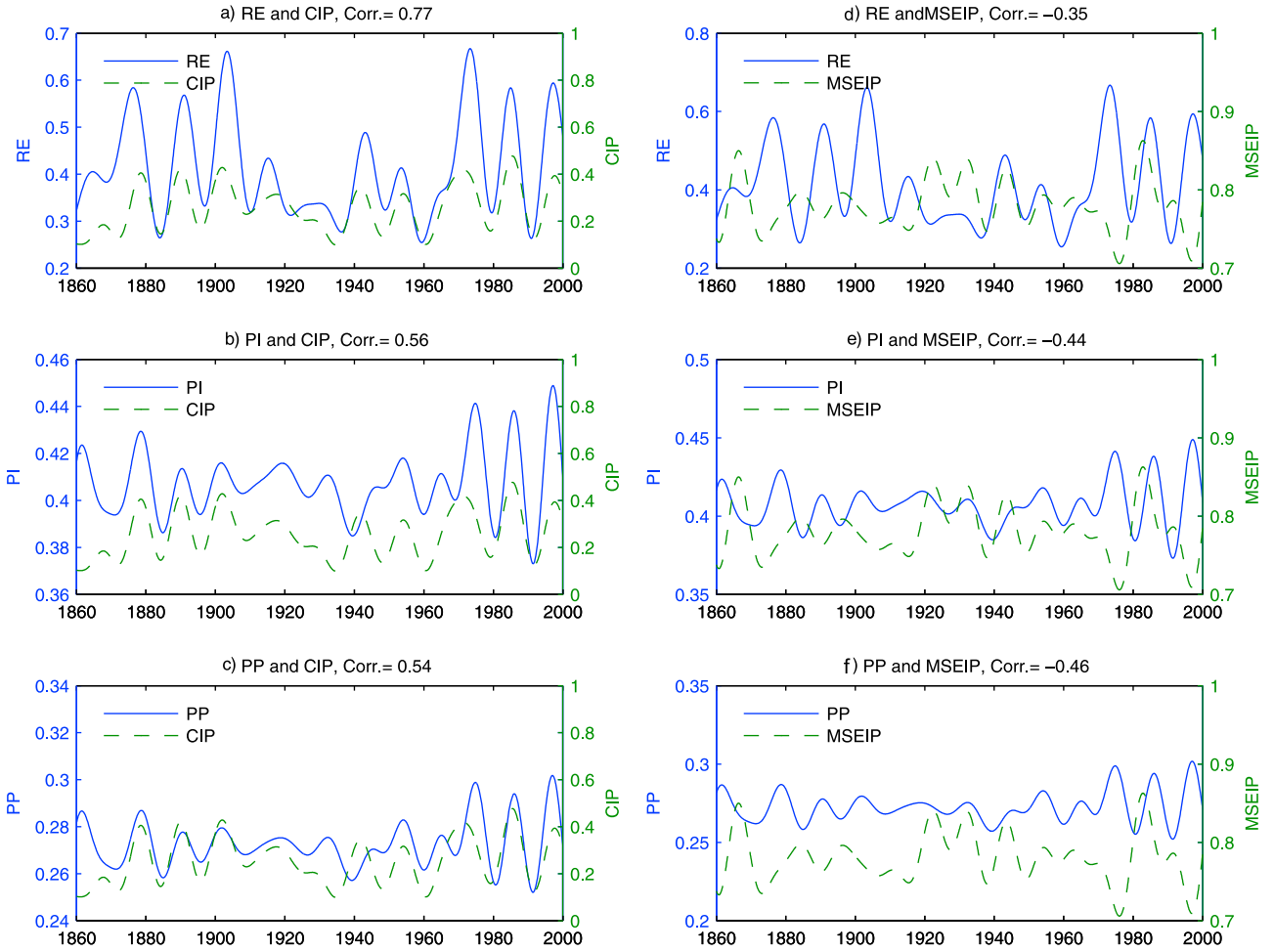

Figure 7. Decadal/interdecadal variations of (a) RE and CIP, (b) PI and CIP, and (c) PP and CIP. Decadal/interdecadal variations of (d) RE and MSEIP, (e) PI and MSEIP, and (fc) PP and MSEIP (a 10 year low-pass FFT filter is applied).

was also high; when REs were small during the time period of 1910–1955, CIP was also low. Compared with PI and PP, RE has the closest in-phase relationship with the correlation skill CIP. On the other hand, Figures 7d–7f depicts an inverse relationship of RE (or PI, PP) with MSEIP on the decadal/interdecadal time scales. PI and PP have a closer relationship with MSEIP skill than RE, because they are more related to noise variance than signal variance as discussed in Section 4.1. Therefore, Figure 7 indicates that on the decadal/interdecadal time scales, RE has a better in-phase relationship with the correlation skill CIP than PI and PP, while PI and PP have better relationships with the MSEIP skill measure than RE. Similarly, EM2 and ER have closer relations to CIP than to MSEIP as compared to ES, as shown in Figure 8, because they both are dominated by the signal component inherent to correlation-based skills. Also, there are similar relationships between ensemble-based potential predictability measures and the actual forecast skills at the interannual time scale (not shown).

5.2. Relationship on All Time Scales

[40] On all time scales from seasonal to decadal, scatterplots of the information-based measures (RE, PI, and PP) and the actual deterministic measures (CIP and MSEIP) are given in Figure 9 using all original samples without filtering. Again, among three information-based measures, RE has the best relationship with CIP with a highly significant correlation coefficient of 0.55. A comparison of these results with that of *Tang et al.* [2008a] reveals that the LDEO5 model offers a higher correlation coefficient between CIP and RE. In addition, there are still some uncertain relationships in Figure 9a, which are consistent with the "triangular relationship" found in previous studies [e.g., *Tang et al.*, 2005, 2008a]. In other words, when RE is large, the correlation prediction skill is certainly high; but when RE is small, the RE-CIP relationship exhibits more uncertainties. Figure 9d shows a similar triangular relationship between MSEIP and RE: MSEIP is small when RE is large, whereas MSEIP is uncertain when RE is small. Generally, RE has a weak relationship with MSEIP because the signal in RE can bias the close relation between noise and MSE. On the other hand, there is an inverse relationship between MSEIP and PI or PP in Figure 9e (or PP in Figure 9f) with a correlation coefficient of −0.33 (−0.38), which is much better than the RE-MSEIP relationship over all time scales.

[41] As expected, the ensemble-based predictability measures EM2 and ER, as shown in Figure 10, also have better

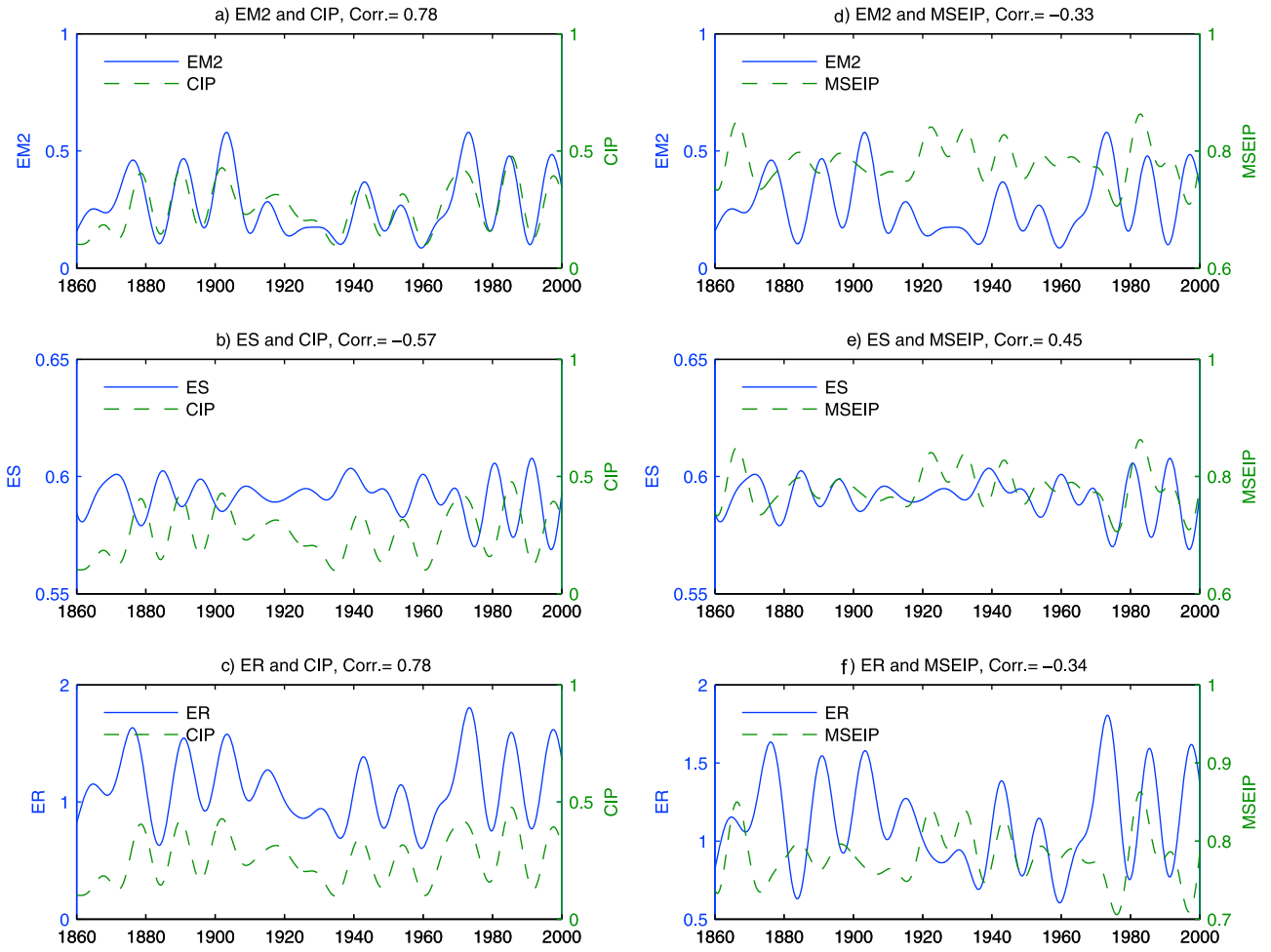

Figure 8. Same as Figure 7 but for ensemble-based predictability measures.

relations with CIP than ES, while ES has a better relation with MSEIP than others.

6. Discussion and Summary

[42] One important task of ENSO predictability studies is to seek good potential predictability measures, by which the uncertainty of individual prediction skill can be estimated without involving observations. In this study, newly developed information-based potential measures (RE, PI, and PP) and classic ensemble-based potential measures (EM2, ES) were explored based on their capability in quantifying the actual prediction skill (CIP and MSEIP) using long-term ensemble predictions of the LDEO5 model. Emphasis was placed on establishing stable and robust relationships between potential predictability and actual prediction skill at various time scales, which offer a practical means of estimating the confidence level of individual predictions. The decadal/interdecadal relationship between potential predictability and actual predictability has not been addressed in previous studies of ENSO predictability due to the lack of sufficiently long retrospective predictions.

[43] The ensemble in this study was produced by using the optimal initial perturbation of SV1_SST and the perturbation of the stochastic optimal wind field during the forecast period. It was found that other ensemble construction methods such as using either the SV1_SSTA or the realistic stochastic wind did not offer reliable ensemble predictions with the LDEO5 model [*Cheng et al.*, 2010b]. From the analysis of the 148 year ensemble predictions, a good in-phase relationship was found between the relative entropy RE and the correlation skill CIP across the time scales from seasonal to decadal/interdecadal.

[44] The mutual information MI is a good measure for the overall forecast skill (correlation R and RMSE). The signal component in ENSO predictions is much stronger than the dispersion component (noise), and thus the predictability is often dominated by ENSO signal strength. The good relationship between RE and CIP is due to the fact that they both emphasize the signal component. It is particularly true for ENSO where RE is determined mainly by SC. The dispersion component DC, though insignificant when signals are strong, can play a comparable role to SC in the predictability of weak ENSO episodes. On the other hand, the SC in RE biases its relationship to MSEIP, which may be the reason why the relationships of these measures are often weaker than PI and PP, and why PP/PI better than RE in measuring MSE-based skill for all time scales and for individual forecasts. For a practical predictability evaluation, all of the information measures (RE and PI/PP) should be explored because they characterize actual prediction skill from different aspects.

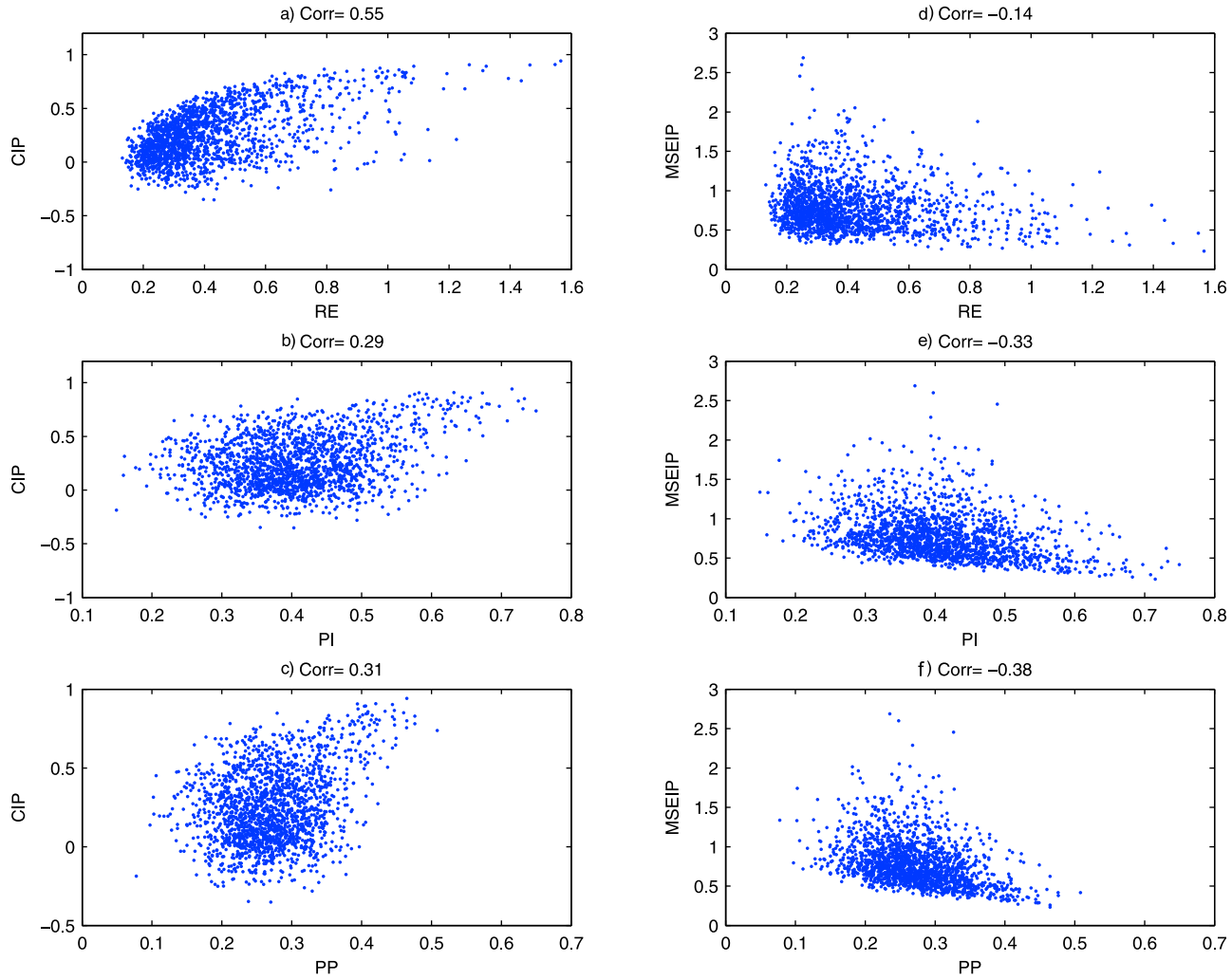

Figure 9. The correlation skill CIP against (a) RE, (b) PI, and (c) PP. MSEIP against (d) RE, (e) PI, and (f) PP.

[45] Because SC and DC are two important factors that determine predictability measures, it is reasonable to divide all those measures in Table 1 into two groups: (1) signal factor (EM2, ER, RE, CIP) and (2) ensemble spread (ES, PI, PP, MSEIP). High correlation coefficients can be found between any two measures within each group. The better potential predictability measures associated with the individual correlation skill are RE, EM2, and ER which are dominated by signal variance or SC, while the best potential predictability measure related to MSEIP is ES due to the contribution of noise variance or DC.

[46] One interesting result of this study is that ES highly depends on initial time and target time, which is consistent with the previous finding using the LDEO4 model. ES has a very strong annual cycle, most likely associated with the background SSTA. ES is often large when SSTA has large variance during the target time of boreal winter and autumn, while it is small during the target time of boreal spring and summer. This is consistent with our previous results from SV analysis by *Cheng et al.* [2010a], where the error growth rate has similar target-season-dependent feature, i.e., large error growth rate occurs during target months in boreal autumn and winter, and small error growth rate occurs for target times in spring and summer. ES and error growth rate are closely related to each other. In a perfect model, ES should be close to RMSE skill. Thus, ES has relatively poor relation to the correlation skill but has the best relation to RMSE skill.

[47] Cautions should be borne in mind. The results and findings presented here are based on the LEDO5 model and the chosen metrics of skill, and thus they might be model and/or metric dependent. For example, we have explored ENSO predictability using CIP and MSEIP to measure prediction skill in this study. Even though the chosen metrics have been widely used in the field of predictability study, they may not be able to completely characterize all properties of predictability. These concerns need to be further addressed through more comprehensive analyses. Nevertheless, this work is the first exploration of information-based potential predictability of ENSO over the past one and half centuries, which not only provides insights into ENSO predictability, but also offers a practical means to estimate the confidence level for individual forecasts.

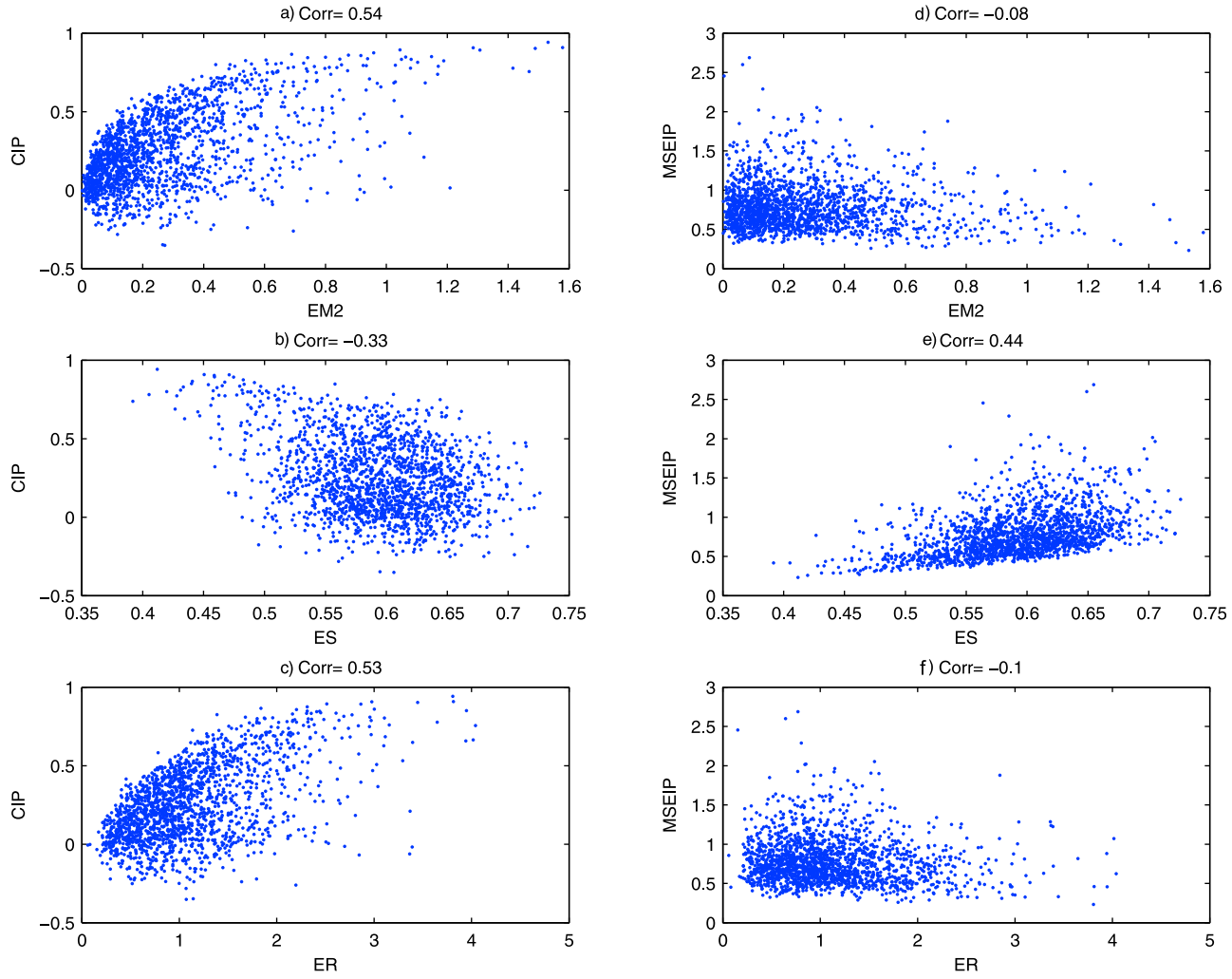

Figure 10. The correlation skill CIP against (a) EM2, (b) ES, and (c) ER. MSEIP against (d) EM2, (e) ES, and (f) ER.

[48] **Acknowledgments.** This work is supported by a NSERC (Natural Sciences and Engineering Research Council of Canada) Discovery Grant and an open grant of the State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou, China. YC is also supported by National Basic Research Program of China (973 Programs: 2010CB951902; 2010CB950501), National Scientific Research Program of China (2010GXS5B148), and National Science Foundation of China (41175074). DC is supported by research grants from National Basic Research Program (2007CB816005) and National Science Foundation of China (40730843).

References

An, S.-I. (2004), Interdecadal changes in the El Nino-La Nina asymmetry, *Geophys. Res. Lett.*, *31*, L23210, doi:10.1029/2004GL021699.

An, S.-I. (2009), A review of interdecadal changes in the nonlinearity of the El Niño-Southern Oscillation, *Theor. Appl. Climatol.*, *97*, 29–40, doi:10.1007/s00704-008-0071-z.

An, S.-I., and B. Wang (2000), Interdecadal change of the structure of the ENSO mode and it impact on the ENSO frequency, *J. Clim.*, *13*, 2044–2055, doi:10.1175/1520-0442(2000)013<2044:ICOTSO>2.0.CO;2.

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.

Cai, M., E. Kalnay, and Z. Toth (2003), Bred vectors of the Zebiak–Cane Model and their potential application to ENSO predictions, *J. Clim.*, *16*, 40–56, doi:10.1175/1520-0442(2003)016<0040:BVOTZC>2.0.CO;2.

Chen, D., M. A. Cane, S. E. Zebiak, R. Cañizares, and A. Kaplan (2000), Bias correction of an ocean-atmosphere coupled model, *Geophys. Res. Lett.*, *27*, 2585–2588, doi:10.1029/1999GL011078.

Chen, D., M. A. Cane, A. Kaplan, S. E. Zebiak, and D. Huang (2004), Predictability of El Niño over the past 148 years, *Nature*, *428*, 733–736, doi:10.1038/nature02439.

Chen, Y.-Q., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. Sarachik (1997), A study of the predictability of tropical Pacific SST in a coupled atmosphere/ocean model using singular vector analysis, *Mon. Weather Rev.*, *125*, 831–845, doi:10.1175/1520-0493(1997)125<0831:ASOTPO>2.0.CO;2.

Cheng, Y., Y. Tang, X. Zhou, P. Jackson, and D. Chen (2010a), Further analysis of singular vector and ENSO predictability from 1876–2003—Part I: Singular vector and the control factors, *Clim. Dyn.*, *35*, 807–826, doi:10.1007/s00382-009-0595-7.

Cheng, Y., Y. Tang, P. Jackson, D. Chen, and Z. Deng (2010b), Ensemble construction and verification of the probabilistic ENSO prediction in the LDEO5 model, *J. Clim.*, *23*, 5476–5497, doi:10.1175/2010JCLI3453.1.

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.

DelSole, T., and M. K. Tippett (2008), Predictable components and singular vectors, *J. Atmos. Sci.*, *65*, 1666–1678, doi:10.1175/2007JAS2401.1.

Fan, Y., M. R. Allen, D. L. T. Anderson, and M. A. Balmaseda (2000), How predictability depends on the nature of uncertainty in initial conditions in a coupled model of ENSO, *J. Clim.*, *13*, 3298–3313, doi:10.1175/1520-0442(2000)013<3298:HPDOTN>2.0.CO;2.

Gill, A. E. (1980), Some simple solutions for heat-induced tropical circulation, *Q. J. R. Meteorol. Soc.*, *106*, 447–462, doi:10.1002/qj.49710644905.

Goswami, B. N., and J. Shukla (1991), Predictability of the coupled ocean–atmosphere model, *J. Clim.*, *4*, 3–22, doi:10.1175/1520-0442(1991)004<0003:POACOA>2.0.CO;2.

Kaplan, A., M. A. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan (1998), Analyses of global sea surface temperature 1856–1991, *J. Geophys. Res.*, *103*(C9), 18,567–18,589, doi:10.1029/97JC01736.

Karspeck, A. R., A. Kaplan, and M. A. Cane (2006), Predictability loss in an intermediate ENSO model due to initial error and atmospheric noise, *J. Clim.*, *19*(15), 3572–3588, doi:10.1175/JCLI3818.1.

Kirtman, B. P., and P. S. Schopf (1998), Decadal variability in ENSO predictability and prediction, *J. Clim.*, *11*, 2804–2822, doi:10.1175/1520-0442(1998)011<2804:DVIEPA>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 midlatitude atmosphere, *J. Atmos. Sci.*, *65*, 263–275, doi:10.1175/2007JAS2234.1.

Kleeman, R., and A. M. Moore (1997), A theory for the limitation of ENSO predictability due to stochastic atmospheric transients, *J. Atmos. Sci.*, *54*, 753–767, doi:10.1175/1520-0469(1997)054<0753:ATFTLO>2.0.CO;2.

Kleeman, R., J. P. McCreary, and B. A. Klinger (1999), A mechanism for generating ENSO decadal variability, *Geophys. Res. Lett.*, *26*, 1743–1746, doi:10.1029/1999GL900352.

Kumar, A., and M. P. Hoerling (1998), Annual cycle of Pacific–North American seasonal predictability associated with different phases of ENSO, *J. Clim.*, *11*, 3295–3308, doi:10.1175/1520-0442(1998)011<3295:ACOPNA>2.0.CO;2.

Kumar, A., A. B. 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.

Moore, A. M., and R. Kleeman (1998), Skill assessment for ENSO using ensemble prediction, *Q. J. R. Meteorol. Soc.*, *124*, 557–584, doi:10.1002/qj.49712454609.

Scherrer, S. C., C. Appenzeller, P. Eckert, and D. Cattani (2004), Analysis of the spread–skill relations using the ECMWF ensemble prediction system over Europe, *Weather Forecast.*, *19*, 552–565, doi:10.1175/1520-0434(2004)019<0552:AOTSRU>2.0.CO;2.

Schneider, T., and S. M. 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. S., et al. (2000), Dynamical seasonal prediction, *Bull. Am. Meteorol. Soc.*, *81*, 2593–2606, doi:10.1175/1520-0477(2000)081<2593:DSP>2.3.CO;2.

Tang, Y., and Z. Deng (2010a), Tropical Pacific upper ocean heat content variations and ENSO predictability during the period from 1881–2000, *Adv. Geosci.*, *18*, 87–108.

Tang, Y., and Z. Deng (2010b), Low-dimensional nonlinearity of ENSO and its impact on predictability, *Phys. D*, *239*, 258–268, doi:10.1016/j.physd.2009.11.006.

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

Tang, Y., R. Kleeman, and S. Miller (2006), ENSO predictability of a fully coupled GCM model using singular vector analysis, *J. Clim.*, *19*, 3361–3377, doi:10.1175/JCLI3771.1.

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. 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. Moore (2008b), Measuring the potential predictability of ensemble climate predictions, *J. Geophys. Res.*, *113*, D04108, doi:10.1029/2007JD008804.

Tang, Y., Z. Deng, X. Zhou, Y. Cheng, and D. Chen (2008c), Interdecadal variation of ENSO predictability in multiple models, *J. Clim.*, *21*, 4811–4833, doi:10.1175/2008JCLI2193.1.

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.

Toth, Z., and E. Kalnay (1993), Ensemble forecasting at NMC: The generation of perturbations, *Bull. Am. Meteorol. Soc.*, *74*, 2317–2330, doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

Xue, Y., M. A. Cane, and S. E. Zebiak (1997a), Predictability of a coupled model of ENSO using singular vector analysis, part I: Optimal growth in seasonal background and ENSO cycles, *Mon. Weather Rev.*, *125*, 2043–2056, doi:10.1175/1520-0493(1997)125<2043:POACMO>2.0.CO;2.

Xue, Y., M. A. Cane, and S. E. Zebiak (1997b), Predictability of a coupled model of ENSO using singular vector analysis, part II: Optimal growth and forecast skill, *Mon. Weather Rev.*, *125*, 2057–2073, doi:10.1175/1520-0493(1997)125<2057:POACMO>2.0.CO;2.

Zebiak, S. E., and M. A. Cane (1987), A model El Niño-Southern Oscillation, *Mon. Weather Rev.*, *115*, 2262–2278, doi:10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2.

Zhang, Y., J. M. Wallace, and D. S. Battisti (1997), ENSO-like interdecadal variability: 1900–93, *J. Clim.*, *10*, 1004–1020, doi:10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2.

Zhou, X., Y. Tang, and Z. Deng (2008), The impact of nonlinear atmosphere on the fastest error growth of ENSO prediction, *Clim. Dyn.*, *30*(5), 519–531, doi:10.1007/s00382-007-0302-5.

D. Chen, Lamont-Doherty Earth Observatory, Columbia University, 106B Oceanography, 61 RT 9W, Palisades, NY 10964, USA.

Y. Cheng and Y. Tang, Environmental Science and Engineering, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada. (ytang@unbc.ca)