

## Review

### Special Topic : Advances in El Niño Research

#### GEOSCIENCES

## Progress in ENSO prediction and predictability study

Youmin Tang<sup>2, 1\*</sup>, Rong-Hua Zhang<sup>3, 4</sup>, Ting Liu<sup>1, 2\*</sup>, Wansuo Duan<sup>5</sup>, Dejian Yang<sup>6</sup>, Fei Zheng<sup>7</sup>,  
Hongli Ren<sup>8</sup>, Tao Lian<sup>1</sup>, Chuan Gao<sup>3, 4</sup>, Dake Chen<sup>1</sup>, Mu Mu<sup>9</sup>

*1 State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Hangzhou, China*

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

*3 Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China*

*4 Qingdao National Laboratory for Marine Science and Technology, Qingdao, China*

*5 State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

*6 CMA-NJU Joint Laboratory for Climate Prediction Studies, Institute for Climate and Global Change Research, School of Atmospheric Sciences, Nanjing University, Nanjing, China*

*7 International Center for Climate and Environment Science, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

*8 Laboratory for Climate Studies & CMA–NJU Joint Laboratory for Climate Prediction Studies,  
National Climate Center, China Meteorological Administration, Beijing, China*

*9 College of Atmospheric and Oceanic Science, FuDan University, Shanghai, China*

*\*Corresponding author: Youmin Tang ([ytang@unbc.ca](mailto:ytang@unbc.ca)) and Ting Liu ([liut@sio.org.cn](mailto:liut@sio.org.cn))*

## **Abstract**

ENSO is the strongest interannual signal in the global climate system with worldwide climatic, ecological, and societal impacts. Over the past decades, the researches about ENSO prediction and predictability have attracted a broad attention. With the development of coupled models, the improvement in initialization schemes, and the advance in the theoretical studies, ENSO has become the most predictable climate mode at the time scales from months to seasons. This paper reviews in details the progress in ENSO predictions and predictability studies achieved in recent years. An emphasis is placed on two fundamental issues: the improvement in practical prediction skill and the advance in the theoretical study of the intrinsic predictability limit. The former includes the progress in the couple models, data assimilations, ensemble predictions and so on, and the latter focuses on the efforts in the study of the optimal error growth and in the estimate of the intrinsic predictability limit.

**Keywords:** ENSO Prediction and Predictability; Coupled Model; Ensemble Prediction; Optimal error growth; Probabilistic Prediction

Received:14-Mar-2018; Revised:12-Jul-2018; Accepted:02-Aug-2018

## 1 Introduction

As a critical component of global change, climate variability is potentially one of the most serious environmental issues we face today. A prominent example is the famous El Niño phenomenon, the strongest source of interannual climate variability<sup>[1]</sup>. El Niño events are characterized by abnormal warming anomalies in the eastern or central equatorial Pacific Ocean, which occurs irregularly every 2-7 years. It involves changes both in the sea surface temperature (SST) and in sea level pressure across the equatorial Pacific Ocean, so it is also referred to as the El Niño-Southern Oscillation (ENSO) phenomenon. Although ENSO originates in the tropical Pacific, its effects are not confined to regional climate but act to induce large weather and climate anomalies worldwide. It is believed that a modest predictability of global climate anomalies can be gained owing to significant impacts of the predictable ENSO<sup>[2]</sup>. Skillful ENSO prediction offers decision makers an opportunity to take into account the anticipated climate anomalies, potentially reducing the societal and economic impacts by this natural phenomenon, and assisting in the management of natural resources and the environment.

Since the 1980s, ENSO has been the focus of oceanic and atmospheric science research. With a series of international collaboration programs and initiatives such as the Tropical Oceans and Global Atmosphere (TOGA) Program and the Climate Change and Prediction (CLIVAR) Plans, significant progress of ENSO prediction has been made over the past four decades<sup>[2-5]</sup>. At present, there are more than 20 models on ENSO for real-time forecasts of half year to one year (see <http://iri.columbia.edu/climate/ENSO/currentinfo/update.html>). Overall, the current

forecast models can provide effective predictions of ENSO warm and cold events 6-12 months ahead <sup>[5]</sup>. ENSO is currently regarded as the most predictable target at the time scales of seasonal climate prediction.

Such progress should be attributed to several aspects. First, the observing systems have been greatly developed with a series of meteorological and oceanic satellites launched and international observing and research programs initiated such as TOGA, GOOS (Global Ocean Observation System), and Argo etc., offering an excellent opportunity to initialize ENSO prediction by various oceanic and atmospheric data. Theoretically, ENSO prediction is a problem of initial value. Under the perfect model assumption, ENSO prediction skill is dependent on the initialization scheme used in the prediction models. The rapid development of data assimilation methods, from simple optimal interpolation in the 1970s-1980s, to 3-dimensional variational (3Dvar) and 4-dimensional variational (4Dvar) in the 1980s-1990s, to Ensemble Kalman Filter (EnKF) and other ensemble-based methods in the 1990s, greatly advances the study of initialization, making the initial conditions of predictions as accurately as possible. Second, a hierarchy of ENSO prediction model has been developed to investigate its physical mechanisms and to improve its predictability over the last decades. These studies greatly improve model performance by improving physical parameterizations, refining spatial and temporal resolutions, and enhancing the understanding of the tropical oceanic and atmospheric processes underlying the ENSO phenomenon. Third, the theoretical predictability study, including the prediction error growth dynamics and the estimate theory of the intrinsic limit of predictability, has been well conducted and applied to

ENSO prediction. The predictability study greatly promotes the development of ensemble prediction system. In recent years, ENSO probabilistic prediction, based on the ensemble prediction, has attracted a broad attention and is routinely issued by some research and operational centers.

However, ENSO prediction still presents a great deal of uncertainty<sup>[6]</sup>. A typical case is the latest El Niño event from 2014 to 2016, which poses a severe challenge to the classic ENSO theory-based forecasting models. For example, the El Niño event did not occur in 2014 as anticipated by most models, whereas the 2015 El Niño, which is one of the strongest events in the history, was not predicted by almost all models at one year lead-time. Moreover, the 2015 El Niño is significantly different from the extreme events of 1997/98 and 1982/83 in the formation and warming pattern<sup>[7]</sup>. This challenge has stimulated a new round of ENSO research in the world.

The motivation of this paper is to review the recent progress in ENSO prediction and predictability studies. The predictability of the earth's climate can be categorized into two types: practical prediction and theoretical predictability. The former is related to prediction skill achieved by models, which can be expected to produce, to our best ability, the best prediction skill by model improvement, prediction initialization and ensemble construction etc., whereas the latter is to assess theoretical upper limit of ENSO prediction skill, which is also called the intrinsic predictability or potential predictability in the predictability study. The intrinsic predictability is an inherent characterization of a physical system rather than our ability to make skillful predictions in practice. The intrinsic predictability study can answer such challenging questions as whether the ENSO prediction skill can be

further enhanced by the improvement in the prediction system, and if so, how much room there is for improvement <sup>[8, 9]</sup>.

The remainder of this paper is organized as follows. Section 2 reviews the development of ENSO prediction models, with an emphasis on the development of operational prediction systems in China. The progress in data assimilations, initialization schemes and ensemble predictions is reviewed in section 3. Section 4 discusses the optimal prediction error growth, the Spring Predictability Barrier (SPB) and the estimate of the Intrinsic Predictability Limit (IPL). Section 5 focuses on the probabilistic Prediction of ENSO. A brief summary and discussion of the major challenges follows in Section 6.

## **2 ENSO Prediction Models**

### **2.1 Physical basis for ENSO prediction**

Since Bjerknes' seminal work <sup>[10]</sup>, the key processes responsible for interannual variability and predictability associated with ENSO have been identified. That is, the Bjerknes feedback is regarded as a major process responsible for ENSO development, involving the interactions among the SST, surface wind and thermocline. It has been recognized that subsurface thermal anomalies play important roles in ENSO-related variability and predictability in the tropical Pacific. In particular, the equatorial ocean wave dynamics and the thermal structure in the tropical Pacific enable thermal perturbations to be sustainable and to propagate around the basin, acting to have remote influences basinwide. These slowly evolving thermal anomalies in the subsurface ocean on the basin scales offer seasonal-to-interannual memory by which the coupled ocean-atmosphere

system can carry on past information with time into the future. It is the existence of these long-lasting thermal memories that provides a physical basis for ENSO prediction. Thus, oceanic information observed at the subsurface is critically important to ENSO prediction and needs to be adequately incorporated into models through data assimilation, a technique used in the prediction initialization that will be discussed in next section. One of the key factors that impacts on the ENSO prediction skill is the representation of model climatological state, which is crucial in capturing right annual cycle and characterizes ENSO asymmetry<sup>[11]</sup>. The poor performance for simulating the characteristic of ENSO asymmetry in the coupled modes is probably a main reason that fails to capture the occurrence of super El Niño event, since it can have a rectification effect on the time-mean state<sup>[12, 13]</sup>. Also, it is worth noting that ENSO can be significantly modulated by other forcing and feedback processes within and/or out of the tropical Pacific (See also Wang et al. 2018 in this issue).

## **2.2 Current status of ENSO predictions models**

Methodologically, there are two kinds of models used for ENSO predictions. The first are the statistical models, which use the historical data to construct the evolution of ENSO, typically represented by Niño3 (90-150°W, 5N-5°S) or Niño3.4 (120-170°W, 5N-5°S) SST anomaly index. The statistical models include linear statistical and nonlinear statistical models. The former is constructed using linear methods such as multiple linear regression, canonical correlation, and Markov chain and so on<sup>[14, 15]</sup>, whereas the latter is constructed mainly using neural network and other machine learning methods<sup>[16]</sup>. Both made success in ENSO prediction and some are still run in operation. However, the development and application of

statistical models have been significantly reduced due to their lacking both physical basis and the room for skill improvement.

The second kind of models are the coupled models, which have become the main tools for studying ENSO mechanisms, simulation and prediction. Since the first coupled ENSO model was developed <sup>[17, 18]</sup>, various types of coupled models have been designed and used for ENSO simulation and prediction. These coupled models include simple models <sup>[19]</sup>, intermediate coupled models <sup>[18]</sup>, hybrid coupled models <sup>[20, 21]</sup>, and fully coupled general circulation models (GCMs) <sup>[2, 4, 5]</sup>. Currently, more than 20 models with different degrees of complexity are routinely used to make real-time forecasts of ENSO. The skillful ENSO predictions can now be made six months and longer ahead <sup>[5]</sup>. There have been several excellent review papers about progress and current status of ENSO coupled models <sup>[2-5]</sup>, and readers are referred to as these for further information.

### **2.3 ENSO Prediction in China**

In China, there are two nationally operational systems for ENSO real-time prediction. One system is the ENSO Monitoring, Analysis and Prediction, called SEMAP2, developed in the National Climate Center (viz., Beijing Climate Center) of China Meteorological Administration (BCC/CMA) <sup>[22]</sup>. This system is based on the operational seasonal forecasting model (BCC\_CSM1.1m) and the physics-based statistical prediction following the dynamical mechanisms of two types of ENSO <sup>[23]</sup>. This system is composed of five sub-systems including the real-time monitoring of the tropical atmosphere-ocean, dynamical diagnosis, physics-based statistical prediction, model ensemble forecasting, and analogue-based correction of the



model prediction <sup>[24]</sup>. This system markedly improves the operational capability of ENSO monitoring and prediction in BCC/CMA. A 20-yr independent hindcast shows a good prediction skill with the correlation skill of Niño 3.4 index reaching 0.8 at 6-month lead <sup>[25]</sup>. It also has successfully predicted the 2015-16 super El Niño event at lead about six months <sup>[26]</sup>. The other operational prediction system, run in the National Marine Environmental Forecasting Center (NMEFC/SOA) of State Ocean Administration, is based on the Community Earth System Model (CESM) developed by of the US National Center for Atmospheric Research (NCAR). A nudging assimilation system for multiple oceanic data including subsurface ocean temperature is used to initialize predictions. This prediction system also has a good performance in prediction ENSO reaching 0.7 correlation skill at 6-month lead <sup>[27]</sup>.

There are also several other ENSO prediction systems run in China. One is implemented in the Institute of Oceanology, Chinese Academy of Sciences (IOCAS), called the IOCAS ICM <sup>[28]</sup>. This is an intermediate anomaly coupled model (ICM), consisting of an intermediate ocean model and an empirical wind stress model. One crucial component of the ICM is the way in which the subsurface entrainment temperature in the surface mixed layer is explicitly parameterized in terms of the thermocline variability <sup>[29]</sup>. The model is one of the coupled models that made a good prediction of the cold SST conditions in the tropical Pacific in 2010-12 <sup>[30]</sup>. Another prediction system was developed at the Institute of atmospheric physics (IAP) at the Chinese Academics of Science Sciences. It adopts the earlier version of IOCAS ICM but includes a newly-developed atmospheric and oceanic data coupled assimilation system <sup>[31]</sup> and an ensemble construction system <sup>[32]</sup>. A 20-year retrospective

hindcast experiment showed that the system can achieve good forecast skill up to one year, comparable with some of the best ENSO prediction models. In each month, the above four operational systems issue routinely ENSO predictions. Also, the predictions of the IOCAS ICM and SEMAP2.1 have been collected by the IRI/CPC ENSO plume product (<http://iri.columbia.edu/climate/ENSO/currentinfo/update.html>).

Recently, a super-ensemble of multiple-model ENSO forecast system is being developed by the Second Institute of Oceanography, State Ocean Administration of China. The system is composed of an ensemble-based coupled data assimilation system and a super-ensemble of multiple ENSO forecast models, including the LDEO5 (the Lamont-Doherty Earth Observatory version 5) model, a hybrid coupled model <sup>[21]</sup>, and two fully coupled GCMs (GFDL-CM2.1 and NCAR CESM). The system will provide both deterministic and probabilistic predictions. At present, the LDEO5 ensemble prediction system, which includes a weakly coupled assimilation system of oceanic and atmospheric observation and a stochastically optimal perturbation-based ensemble construction system <sup>[33]</sup>, has started running. A long-term hindcast experiment from 1856-2016, as shown in Fig.1, indicates that the system can capture almost all warm and cold events at 6 month leads. Its correlation skill is comparable with the current best level, and better than the latest version of LDEO5 <sup>[34]</sup>.

#### **2.4 An example taken for the 2015-16 El Niño event as a demonstration**

Here, the latest strong 2015-16 El Niño event can be taken as an example for illustrating the current status of ENSO predictions using the state-of-the-art coupled models. Figure. 2 demonstrates the models' real-time performance in predicting the

2015-16 El Niño event, collected in IRI. As observed, one striking feature associated with this 2015-16 El Niño event was the slow evolution of warm SST anomalies in the western tropical Pacific through 2014 and early 2015 <sup>[7]</sup>; subsequently, the related ocean-atmosphere anomalies were coupled and amplified in spring 2015 and developed rapidly into a warm event in late spring 2015; then this warm event eventually evolved into an extreme El Niño in the following months. Looking at the 2015 El Niño prediction by these coupled models (Fig. 2), the observed SST evolution from summer through winter 2015 were adequately depicted when predictions were initialized from August 2015 (Fig. 2a). However, there existed large uncertainties of the predictions of these coupled models. For example, the predicted intensity exhibits a wide spread across these coupled models in summer and fall 2015. In particular, almost all models failed and missed the strongest warming when the predictions were initialized at early 2015 (Fig. 2b), probably due to the SPB problem, westerly wind bursts (WWB) problem or others <sup>[7]</sup>. The uncertainty is also very apparent for ENSO prediction in 2014 (figure not shown). Although warm SST anomalies were observed to occur in the western equatorial Pacific at early 2014, the warm SST anomalies weakened in mid-2014 and did not develop into an El Niño event in late 2014. However, many coupled models predicted a strong El Niño in 2014, a false alarm that embarrassed the ENSO scientific community. The prediction cases in 2014 and 2015 clearly indicate that the real time prediction of ENSO remains challenging and problematic, even when the state-of-the-art coupled models are used. Further studies on understanding predictability and improving real-time predictions using the coupled models are clearly needed.

### 3 Initialization and Ensemble Prediction of ENSO

#### 3.1 Initialization and data assimilation

Since the memory for ENSO mainly resides in the ocean, the oceanic data assimilation plays a vital role in the history of ENSO prediction development. There was a long way in oceanic initialization study in the field of ENSO prediction, from simple to complex algorithms, surface to subsurface observations, and from single data to multiple data sources. Initially, attempts had been made to generate consistent initial conditions by using SST observations<sup>[35, 36]</sup>. Kirtman et al.<sup>[37]</sup> used an iterative ocean initialization procedure that modifies the zonal wind stress anomalies based on the simulated SST anomalies errors. With a 3DVar assimilation scheme, Tang et al.<sup>[38]</sup> assimilated SST into an OGCM with a statistically-derived correction scheme for subsurface surface temperature increment. Keenlyside et al.<sup>[39]</sup> nudged observed SST to initialize ENSO prediction from 1969-2001 and achieved good prediction skills. Recently, Merryfield et al.<sup>[40]</sup> initialized ocean states by nudging SST for the Canadian seasonal to interannual prediction system and achieved good the prediction skill comparable with the current best level.

A lot of works got to realize the importance of the subsurface processes (e.g., entrainment and mixing) in controlling interannual SST variability in the equatorial Pacific<sup>[1]</sup>. It is well known that the variation in the sea level height in the tropical Pacific is dominated by the subsurface dynamics and thermodynamics<sup>[41]</sup>. Thus, many efforts have also been contributed to assimilate sea level gauge data and altimeter data<sup>[42]</sup>, in situ profiles<sup>[43]</sup>, and even salinity observations<sup>[44]</sup>. All these studies show that the inclusion of these new observational data into the assimilation

system can significantly improve the estimation of subsurface ocean states, and therefore achieve a more reliable ENSO forecast.

More recently, the approach to assimilate atmospheric observations to initialize the ENSO prediction was initiated in a more general context of the coupled model assimilation <sup>[6, 31]</sup>. As demonstrated in Chen et al. <sup>[6]</sup>, the ocean-only initialization approach is not necessarily the optimal method for skillful forecast due to an “initial shock” at the transition from uncoupled to coupled run. The more natural method, coupled assimilation, allows the adjustment of both atmosphere and ocean states by assimilating either the atmospheric or oceanic observations. The concept of coupled assimilation is attractive due to the potential benefit of the delivery of a balanced and dynamically consistent estimate of the coupled atmosphere-ocean state. The added values of the coupled assimilation to the ocean-only assimilation have been proven in ENSO predictions <sup>[32, 45]</sup>.

### **3.2 Ensemble prediction**

Owing to the high degree of non-linearity and stochastic forcing of the atmosphere and ocean systems, the evolution of the future states of the atmosphere and the ocean contains large uncertainties. An important way to consider the prediction uncertainties is by the ensemble forecasting, namely, a group of predictions that are generated by perturbing a small amount of “errors” onto the initial conditions or model parameters <sup>[6]</sup>, and weighted (or arithmetic) average or probability distributions are taken as the forecast. Compared with a single forecast, the ensemble forecast can remove some unpredictable noises by the average and

provides a practical tool for estimating the possible uncertainties by additional information, such as the probability distribution function (PDF) of forecast.

The development of ensemble forecast is mainly the development of perturbation methods. In addition to the random perturbation methods, the optimal perturbation methods based on dynamical constraints, such as the Breeding Vector (BV), the Singular Vector (SV), Stochastic Optimal (SO), Conditional Nonlinear Optimal Perturbation (CNOP) and so on, were developed and applied. The perturbation methods derived from the data assimilation include the EnKF, the ensemble transmission Kalman filter (ETKF) and other ensemble-based filters <sup>[46]</sup>. Significant progress has been made in using these optimal perturbations to construct ensemble predictions and to study ENSO predictability.

Generally, the strategies used to produce optimal perturbations for ensemble predictions include perturbation of the initial conditions and the perturbation of model parameters that consider errors existing in physical/dynamical parameterizations. The optimal perturbation of initial condition was often constructed by SV, BV, climatologically-relevant singular vector (CSV) analysis and CNOP <sup>[33, 47-49]</sup>. The impact of model parameter uncertainties on predictions were considered in many works. For example, Zheng et al. <sup>[50]</sup> developed an ensemble perturbation method by adding some random terms to the right-hand side of the model equations of an intermediate coupled model, and these random terms or process noise sources were explicitly defined as the model errors. For the noise-free models, the stochastic noise is often considered in the framework of the stochastic optimal theory, which allows to construct a spatial-temporal coherent noise <sup>[33, 51]</sup>. In

addition, the multiple models and multi-methods are also used to construct the super-ensemble predictions which can improve ENSO prediction skill in both deterministic measures and probabilistic measures <sup>[25, 52]</sup>.

## **4 The Potential Predictability Study of ENSO**

### **4.1 The optimal prediction error growth**

An important aspect of potential predictability research is the analysis and diagnosis of the optimal growth of forecast error, which explores the dynamics of forecast error and explains the physical mechanism of forecast uncertainty. It is fundamental to the construction of ensemble prediction. The current optimal error analysis methods, as mentioned above, include SV, BV and CNOP as well as the Nonlinear Local Lyapunov Vector (NLLV) <sup>[53]</sup> etc.. These optimal error growth and optimal perturbation methods were introduced to the field of ENSO prediction at the 1990s and has since been used for ENSO predictability study <sup>[54-57]</sup>.

There are a number of studies on ENSO predictability using SV and BV as mentioned above. Although these works used different intermediate complexity dynamical models, they basically obtained the consistent results. Namely that, the first singular vector is basically dominated by a west–east dipole spanning across the equatorial Pacific, with one center located in the east and the other in the central Pacific, and the fastest error growth rate (leading singular value) is little sensitive to initial conditions and optimization time. However, the error growth is seasonal dependent with the largest growth during the spring. The ENSO states can also

influence the error growth with the largest error growth occurring on the onset of El Niño and the smallest growth during the La Niña <sup>[47, 56, 58]</sup>.

While these works all used simplified or intermediate complexity coupled models, Kleeman et al. <sup>[59]</sup> proposed a method to calculate the climatologically-relevant singular vector (CSV) for the coupled GCM, in which the atmospheric noise is filtered by running a large ensemble of integrations. Tang et al. <sup>[48]</sup> first applied the CSV to a fully coupled GCM to investigate the error growth associated with ENSO forecasts. The results show that the singular vectors share many of the properties already seen in the simpler models. However, this particular CGCM also displays some differences from the simpler models; thus subsurface temperature optimal patterns are strongly sensitive to the phase of ENSO cycle, and at times the east-west dipole in the eastern tropical Pacific basin <sup>[60]</sup>.

Besides the study of the initial errors, increasing efforts have been dedicated to explore the effect of the model errors on the forecast. In the linear regime, SO optimal <sup>[33]</sup> and forcing singular vector (FSV) <sup>[61]</sup> were employed to represent the influence of the stochastic noise on prediction error growth. Duan and Zhou <sup>[62]</sup> extended the FSV-approach to the non-linear system and proposed the concept of the Non-linear forcing singular vector (NFSV). It was found that the NFSV-related model errors have the largest negative effect on the uncertainties of ENSO prediction <sup>[63]</sup>.

## 4.2 Spring Predictability Barrier

One specifically important prediction uncertainty in ENSO prediction is SPB, which refers to a quick decrease of prediction skill during the boreal spring <sup>[64]</sup>.



From the perspective of error growth, the SPB refers to as the phenomenon that ENSO forecasting has a large prediction error; in particular a prominent error growth, during the spring when the prediction is made before spring<sup>[65]</sup>. Quite a few studies, including linear SV (LSV) and nonlinear CNOP, found that the SPB arises from the growth of initial errors<sup>[54, 55, 65]</sup>. The comparison between CNOP and LSV showed that the SPB-related initial errors determined by the CNOP, although presenting spatial patterns similar to those revealed by the LSV, cover a broader region and result in a more significant SPB<sup>[65]</sup>. It is therefore concluded that the initial errors of the CNOP structure are most likely to cause the SPB.

The concept of the two types of El Niño was proposed in early 2000s<sup>[66]</sup>. One type of El Niño, named as the “central Pacific-El Niño”, exhibits warming center in the equatorial central Pacific, while the other is named as the “eastern Pacific-El Niño” which show-warming center in the equatorial eastern Pacific. It was found in some works that the frequently occurred central Pacific-El Niño since 2000s increases the uncertainties of ENSO forecasting<sup>[67]</sup>. Tian and Duan<sup>[68]</sup> showed that central Pacific-El Niño also has the SPB and possess a SPB-related initial error pattern similar to the eastern-Pacific El Niño as illustrated in Yu et al.<sup>[65]</sup>. Both the eastern- and central-Pacific El Niño forecasting could occur SPB, but the latter has less chances to encounter a SPB<sup>[69]</sup>. Ren et al.<sup>[70]</sup> also revealed that the central-Pacific ENSO has a much weaker persistence barrier, closely related to the SPB, than the eastern-Pacific ENSO. To filter out the SPB-related initial errors, a hopeful strategy is the target observation, which aims to determine the optimal observing region for minimizing the prediction uncertainty<sup>[71]</sup>. Duan et al.<sup>[69]</sup> concluded an optimal

observational array, as shown in Fig. 3, in which the additional observations can be deployed for dealing with the challenge of ENSO prediction due to the diversity of El Niño events and the related SPB.

### 4.3 Measure of ENSO IPL

Predictability is the extent to which events can be predicted <sup>[72]</sup>, including the actual predictability and the potential predictability. The actual predictability is aiming to quantify the accuracy of model predictions against observations, measured with either deterministic or probabilistic scores. The potential predictability is often referred to as the upper limit of skill, i.e., IPL. Based on the ensemble prediction, there are a couple of metrics that can be used to quantify the IPL. Among them are variance-based measure and information-based measure, both quantifying the predictability or prediction uncertainty from different angles.

Signal-to-Noise Ratio (SNR) is widely employed in quantifying potential predictability <sup>[73]</sup>. The amplitude of signal and noise of the ENSO ensemble prediction can be approximately quantified by the variance of ensemble mean and the averaged ensemble spread over all initial conditions <sup>[74]</sup>. The prediction skill measures, such as anomaly correlation <sup>[75]</sup> or rank probability skill score <sup>[76]</sup>, are the function of SNR. The higher values of the SNR indicate less contaminations of the signal information by unpredictable random effects, and higher potential forecasting capability and higher prediction skill <sup>[77]</sup>. The amplitude of the signal is much larger than the noise in the ENSO prediction, suggesting that the predictability of ENSO is dominated by the ENSO signal <sup>[74, 78]</sup>.

Several information-based measures have been applied to qualify the IPL, including the relative entropy, predictive information, predictive power, and mutual information (MI). The essential of these information-based measures is that the difference between the prediction and climatology distribution, can quantify the extra potential information from the forecast. Specially, the relative entropy is always related with the strong ENSO events and decreases as the lead time increases. It has a better relationship with the correlation-based prediction skill than the predictive information and predictive power due to the fact that the signal component dominates the relative entropy metrics <sup>[74, 79]</sup>. The MI is an indicator of the overall potential predictability of a dynamical like the ENSO system, obtained by averaging the relative entropy or prediction information over all predictions (Fig. 4a). Theoretically, the MI has a strong relationship with actual prediction skill as shown in Fig. 4b.

Note that the MI-based potential predictability measures the statistical dependence, liner or nonlinear, between the ensemble mean prediction and hypothetical observation (an arbitrary ensemble member), whereas the SNR-based potential skill only measures their linear correlation, therefore underestimates the nonlinear statistical dependence. In other words, the information-based potential predictability measures should be better than the SNR-based measures in characterizing ‘true’ potential predictability. When the climatology and prediction distribution are both Gaussian and the prediction variances are constant, the information-based measure is equivalent to the SNR-based potential measure <sup>[80]</sup>.

Note that both of the above two kinds of measures do not involve observations, thereby measuring the potential predictability of the model forecast system.

In addition, the IPL of a chaotic system can also be quantitatively determined by the NLLV using observational data. Li and Ding<sup>[81]</sup> employed the NLLV approach to explore the temporal-spatial distribution of the tropical SST IPL. They found that the annual mean predictability limit is very large in the tropical central-eastern Pacific (>8 months), exceeding 10 months in the Niño3.4 region.

## 5 Probabilistic Prediction of ENSO

Probabilistic forecast aiming at predicting the probability distribution for the future state of the variable can express the forecast uncertainty information. It has been argued to be more informative and valuable than the deterministic forecast<sup>[82]</sup>. In practice, rather than a continuous probability distribution, only the probabilities of some discrete categorical events that are of particular interest are usually predicted. The predicted probability for an event is usually estimated as the fraction of ensemble members of forecasting of this event.

Kirtman<sup>[83]</sup> is the first to demonstrate the importance of presenting ENSO predictions in the probabilistic format and verifying ENSO ensemble predictions from the probabilistic perspective. Even when the deterministic prediction skill is insignificant, skillful probabilistic prediction is still possible for the above- and below-normal events. It was argued that the probabilistic verification is an important complement to the deterministic verification. With the same methodology of Kirtman<sup>[83]</sup>, several subsequent studies<sup>[6, 50]</sup> have also paid significant attention to

investigate the ensemble prediction skill of the ENSO from the probabilistic perspective. The receiver operating characteristic (ROC)-based diagnostics in these studies confirmed the main findings of Kirtman <sup>[83]</sup>.

However, the ROC-based analysis cannot explore the full scope of the probabilistic prediction skill since it can only reflect the resolution aspect. As argued in Yang et al. <sup>[84]</sup>, while the resolution would be strongly impacted by the intrinsic predictability of the real world, there should be, in principle, no barrier at the physical level in improving the reliability. As such, compared to the resolution, the reliability would be more sensitive to the change in the ensemble prediction system itself and would therefore be a more indicative criterion for testing the ensemble construction strategies.

More recently, along the above line of discussion, some studies have given the prominence to examine the reliability of the ENSO ensemble probabilistic predictions <sup>[33, 52]</sup>. Using the LDEO5 model, Cheng et al. <sup>[33]</sup> (2010) produced retrospective probabilistic predictions of the ENSO with using different ensemble perturbation strategies and compared their probabilistic prediction skills in terms of the reliability and the resolution. The results indicated that the reliability is very sensitive to the uncertainties sampled in initialization and in model errors (noise), whereas the resolution is not. Specifically, when either of the two uncertainties, especially the atmospheric noise uncertainty, is not sampled, the reliability suffers from a severe “overconfidence” question. When both uncertainties are sampled, this “overconfidence” is significantly reduced. The amplification of errors in ocean initial conditions can be a large source of the uncertainties <sup>[85, 86]</sup>. For example, a so-called

Multiple-ocean Analysis Ensemble (MAE)<sup>[87]</sup> initialization scheme to sample the structural uncertainty in ocean initial conditions can effectively improve the reliability of ENSO predictions<sup>[88]</sup>. The “overconfidence” was not merely associated with the immediate-complexity model used. It was also observed in comprehensive coupled GCMs (CGCMs)<sup>[52]</sup>. This cavity of the reliability seen in CGCM-based ensemble prediction systems was mainly attributed to the lack of sampling the uncertainty associated with the model errors in the forecast models. It was further found that as a very pragmatic approach of sampling model uncertainties, the multi-model ensemble (MME) can alleviate this reliability defect of the ENSO ensemble probabilistic predictions<sup>[52]</sup>.

The benefit of the MME approach was not only seen in the reliability. Kirtman and Min<sup>[52]</sup> also found the benefit of the MME in the resolution based on the analysis of the ROC score. Tippett and Barnston<sup>[89]</sup> further provided more comprehensive evidences for the MME’s effectiveness in improving the ENSO probability forecasts. The MME advantage was found to be greater than what is expected as a result of an increase in ensemble size only. Typically, the probability forecasts for ENSO are only considered for three categories (El Niño, La Niña, or neutral), which cannot reflect more detailed information regarding ENSO phase and intensity. Very recently, Tippett et al.<sup>[90]</sup> examined the ENSO probabilistic forecast skill using the North American Multi-model Ensemble Products, with more categories being tested. The results indicated that the current MME predictions have shown the capability in providing detailed probabilistic forecasts of ENSO phase and amplitude.

Compared with the probabilistic weather forecast, the study of the ENSO probabilistic prediction and verification is still at an early stage. Existing studies have demonstrated that the current ENSO probabilistic forecasting tends to be prone to an “overconfidence” and that the MME is a pragmatic yet effective approach to improve the ENSO probabilistic prediction. Further studies are required to comprehensively investigate the current status of ENSO probabilistic prediction skill in terms of the reliability and the resolution. Moreover, the ensemble generation techniques adopted in previous studies to sample the uncertainties of prediction systems are mostly too simple, usually based on random perturbations or the lagged ensemble approach, which are far from sophisticated. Therefore, it is necessary to use more refined ensemble generation techniques, for example, the SV, SO, NLLV or CNOP, to further improve the reliability and ultimately the overall skill of the ENSO ensemble probabilistic prediction. While these methods have been widely applied to the weather probabilistic forecast, and even to ENSO predictability study, they haven’t been much applied to the ENSO ensemble (and especially probabilistic) prediction study.

## **6 Summary and Discussion**

As recognized, the ENSO is the most predictable interannual signal in the climate system, providing the basis for the global short-term climate prediction. Real-time prediction information is critically important to mitigation and adaptation activities in response to ENSO events and related natural disasters. Over the past few decades, comprehensive understanding of ENSO processes and its predictability has been achieved, which provides the physical basis for ENSO prediction. So far, various

coupled ocean-atmosphere models with hierarchical complexity have been developed for ENSO simulation and prediction. The predictive skill has been improved to a level where it is possible to make successful real-time ENSO prediction at lead time of seasons and longer.

In this paper, we reviewed recent progress achieved in ENSO prediction and predictability study. We discussed several main issues responsible for the improvement in ENSO prediction skill and the advance in predictability study. First, the rapid development and huge achievement of the observing systems provide high quality oceanic and atmospheric data available. The advance in data assimilation methods and algorithms ensure the success in using observations to initialize prediction. Second, a hierarchy of ENSO prediction model has been employed with improved physical parameterizations and model resolutions. The state-of-the-art models have been widely used for operational ENSO predictions. Third, the studies of the optimal growth of forecast error provide various linear and nonlinear optimal perturbations, producing multiple approaches to construct ensemble prediction systems. The probabilistic prediction of ENSO has been run in operation in many countries, offering important information for stakeholders to make decision. Moreover, the newly developed information theory-based framework of statistical predictability promotes the theoretical study of ENSO predictability and offers effective metrics to quantify the intrinsic predictability limit of ENSO.

It has no doubt that notable progress has been made in the study of ENSO prediction and predictability. However, several specific challenges still exist in improving the ENSO prediction skill and understanding the ENSO predictability,



particularly enlightened from recent researches and predictions of 2014-16 super El Niño event <sup>[91]</sup>. Among them is the model systematic error, which is probably the most challenging issue. Focused on the birth region of ENSO, the tropical Pacific Ocean remains an area in which pronounced biases exist in model simulations compared with observations. For example, obvious discrepancies exist in ocean GCM (OGCM) simulations, including the simulated thermocline too diffuse, with a weak vertical gradient of temperature <sup>[92]</sup>. Furthermore, OGCM-based coupled simulations commonly have an unrealistic structure of interannual SST variability, with SST anomalies being underestimated over the eastern equatorial Pacific but overestimated in the central equatorial Pacific. Additionally, the real climate system in the tropical Pacific is characterized by interannual oscillations with a main period band of 4-5 years, but some coupled GCMs favor a quasi-biennial oscillation accompanied by the predominance of the westward propagation of simulated SST anomalies over the eastern and central equatorial Pacific. These discrepancies between model simulations and observations are attributed to processes that still have not well been represented in models. For example, vertical mixing and diffusion are important processes that determine how SSTs are affected by subsurface thermal conditions <sup>[93]</sup>. In ocean models, the vertical mixing/diffusion processes are generally parameterized using large-scale oceanic fields, including the K-profile parameterization (KPP)-scheme <sup>[94, 95]</sup>. It is difficult to accurately determine these effects in models and large uncertainties are inevitable in representing the related subsurface effects on SSTs, which can cause large biases in SST simulations. In addition, many processes that can affect model performance in ENSO simulations are still missing or are not adequately represented in current low-resolution climate

models, including mesoscale processes (e.g., Tropical Instable Waves, TIWs), multi-sphere processes (e.g., ocean biology), Madden-Julian Oscillation (MJO), and so on<sup>[96]</sup>. Coupled models with problems in representing physical processes cause large errors in ENSO simulations. Zhu et al.<sup>[97]</sup> demonstrated that bias in representing thermodynamical process is one of factors contributing to the 2014 El Niño false alarm. Another well-known example is the double Intertropical Convergence Zone (ITCZ) problem in the tropics, which may be solved with the better mean state simulation. To achieve this goal, the importance of the atmospheric convection scheme in ENSO prediction shouldn't be ignored<sup>[98]</sup>. These biases in simulations inevitably lead to errors in ENSO predictions, including the SPB phenomenon.

There are clear needs to improve understanding of variability and predictability associated with ENSO. The ENSO variability exhibits its diversity and asymmetry features<sup>[99]</sup>. What is not clear now is whether the predictability also varies with the different types of El Niño, which should be specifically studied. Mathematically, prediction systems themselves need to be improved not only by model parameterization processes, but also by initialization methods, ensemble-based prediction procedures and post-processing etc. For example, model components (the atmosphere, the ocean, and their interactions) need to be improved further so that relevant processes can be accurately represented in the coupled models, including various feedbacks and coupling processes that can modulate ENSO. More care needs to be taken for parameterizations of unresolved or missing processes in oceanic and atmospheric models. Also, model resolution needs to be enhanced so that processes that are missing in low resolution models can be adequately represented, including

convection, TIWs, MJO, and so on. Moreover, representation of the ENSO periodicity in climate models also needs to be quantitatively evaluated and specifically improved to make predictions of different timescales in ENSO diversity better.

In addition to model processes-related issues, ENSO prediction skills can also be enhanced by various techniques, including initializations and prediction procedures. Due to the important roles in predictions, for example, the subsurface thermal conditions in the ocean need to be coherently incorporated into prediction models through initialization procedures using data assimilation techniques. The more natural initialization method, coupled assimilation, should be developed and applied to provide initial conditions for the predictions. Also, the stochastic nondeterministic processes are prominent in the tropical Pacific, so it is necessary to treat ENSO as a stochastic system. It is still a challenge to realistically characterize some important stochastic processes, for example the WWB, in prediction models by adequate parameterization. It is expected that realistic representations of these stochastic processes can improve prediction skills. Since the ENSO prediction contain uncertainties, the adequate ensemble construction strategies that can efficiently represent the uncertainties associated with the initial and model errors should be well configured in developing ENSO prediction systems. On one hand, the prediction uncertainties can be measured and quantified by probabilistic prediction. The efforts to develop probabilistic prediction and verification should have practical value and guide meaning for the economic society. The probabilistic prediction study of ENSO should be greatly promoted in the future. On the other hand, due to model deficiencies, the post-calibration of ensemble mean is still needed to further improve

ENSO prediction by developing some empirical/statistical correction methods of model forecasts.

## Acknowledgments

This work was jointly supported by grants from the National Natural Science Foundation of China (41690124, 41690121, 41690120, 41705049, 41621064, 41530961), the National Key Research and Development Program (2017YFA0604202), the National Programme on Global Change and Air-Sea Interaction (GASI-IPOVAI-06), and the Scientific Research Fund of the Second Institute of Oceanography (JG1810).

## Reference

- [1] Philander, SGH. El Niño, La Niña, and the Southern Oscillation. Academic Press. 1990; 538-538.
- [2] Luo, J-J, Lee J-Y, Yuan CX et al. Current status of intraseasonal–seasonal-to-interannual prediction of the Indo-Pacific climate, in *Indo-Pacific climate variability and predictability*. World Scientific, 2016, 63-107.
- [3] Latif, M, Latif, Anderson, D, Barnett, T, et al. A review of the predictability and prediction of ENSO. *J Geophys Res: Oceans* 1998; **103**(C7): 14375-14393.
- [4] Jin, EK, Kinter, James, L, Wang, B, et al. Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Clim Dyn* 2008; **31**(6): 647-664.
- [5] Barnston, AG, Tippett, MK, L'Heureux, ML, et al. Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *B AM Meteorol Soc* 2012; **93**(5): 631-651.
- [6] Chen, D. and Cane, MA. El Niño prediction and predictability. *J Comput Phys* 2008; **227**(7): 3625-3640.
- [7] Lian, T, Chen, D, and Tang, YM. Genesis of the 2014–2016 El Niño events. *Sci China Earth Sci* 2017; **60**(9): 1589-1600.
- [8] Mu, M, Duan, WS, and Tang, YM. The predictability of atmospheric and oceanic motions: Retrospect and prospects. *Sci China Earth Sci* 2017; **60**(11): 2001-2012.
- [9] Kumar, A, Zhu JS. Spatial variability in seasonal prediction skill of SSTs: Inherent predictability or forecast errors. *J Climate* 2017; **31**(2): 613-621.
- [10] Bjerknes, J. Atmospheric teleconnections from the equatorial Pacific. *Mon Weather Rev* 1969; **97**(3): 163-172.
- [11] Yang, XQ, and Anderson, JL. Correction of systematic errors in coupled GCM forecasts. *J Climate* 2000; **13**(12): 2072-2085.
- [12] Liang, J, Yang, XQ, and Sun, DZ. The effect of ENSO events on the tropical Pacific mean climate: Insights from an analytical model. *J Climate* 2012; **25**(21): 7590-7606.
- [13] Liang, J, Yang, XQ, and Sun, DZ. Factors determining the asymmetry of ENSO. *J Climate* 2017; **30**(16): 6097-6106.
- [14] Penland, C. A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D*

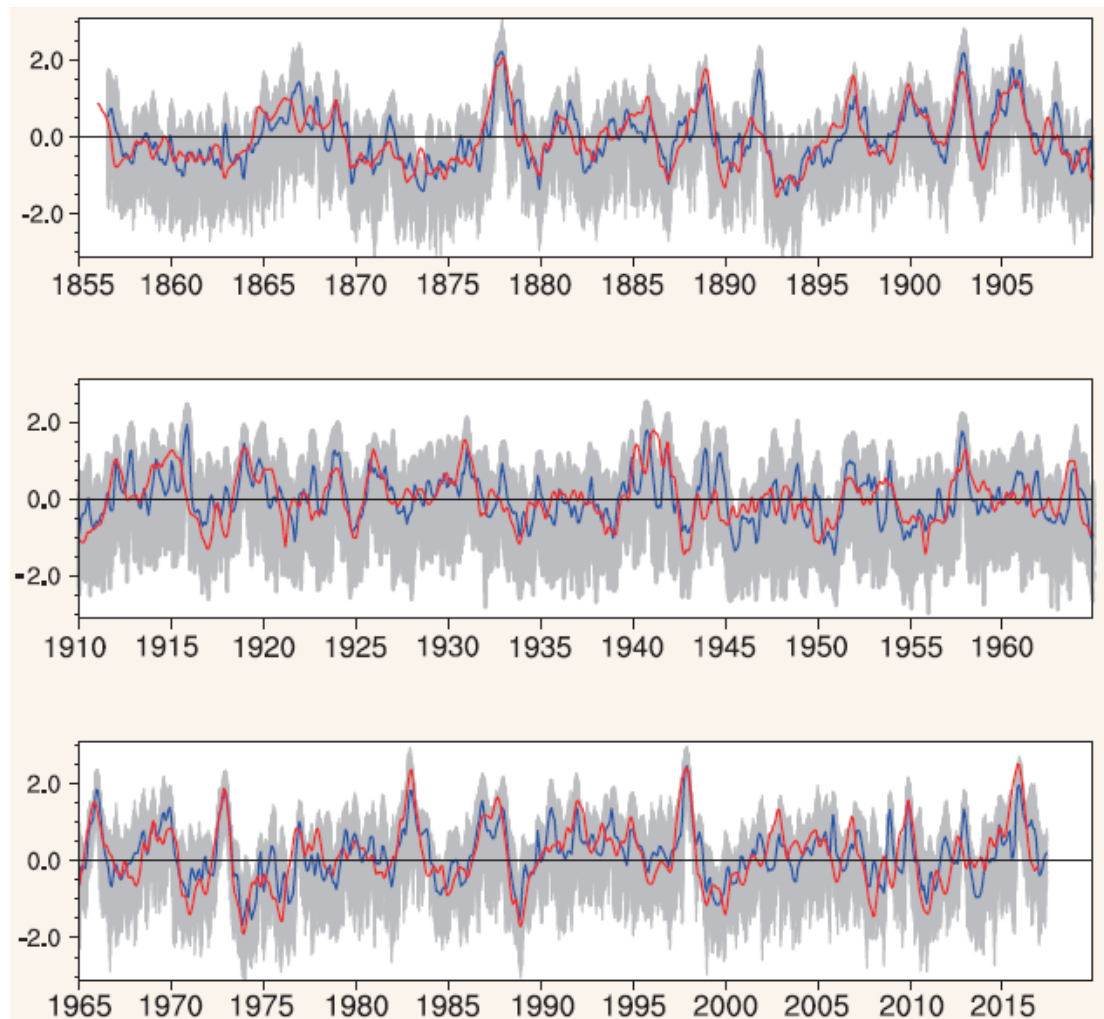
- 1996; **98**(2-4): 534-558.
- [15] Tseng, Y-H., Hu, ZH, Ding, RQ, *et al.* An ENSO prediction approach based on ocean conditions and ocean-atmosphere coupling. *Clim Dyn* 2017; **48**(5-6): 2025-2044.
  - [16] Hsieh, WW, *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*. Cambridge University Press, 2009.
  - [17] Cane, MA, Zebiak, SE, and Dolan, SC. Experimental forecasts of El Niño. *Nature* 1986; **321**(6073): 827.
  - [18] Zebiak, SE, and Cane, MA. A Model El Niño–Southern Oscillation. *Mon Weather Rev* 1987; **115**(10): 2262-2278.
  - [19] Hirst, AC. Unstable and damped equatorial modes in simple coupled ocean-atmosphere models. *J Atmos Sci* 1986; **43**(6): 606-632.
  - [20] Barnett, T, Graham, N, Pazan, S, *et al.* ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. *J Climate* 1993; **6**(8): 1545-1566.
  - [21] Tang, YM, and Hsieh, WW. Hybrid coupled models of the tropical Pacific–II ENSO prediction. *Clim Dyn* 2002; **19**(3-4): 343-353.
  - [22] Ren, HL, Liu, Y, Zuo, JQ, *et al.* The New Generation of ENSO Prediction System in Beijing Climate Centre and Its Predictions for the 2014/2016 Super El Niño Event. *Meteor Mon* 2016; **42**(5): 521-531.
  - [23] Ren, HL, and Jin, FF. Recharge oscillator mechanisms in two types of ENSO. *J Climate* 2013; **26**(17): 6506-6523.
  - [24] Liu, Y, and Ren, HL. Improving ENSO prediction in CFSv2 with an analogue-based correction method. *Int J Climatol* 2017; **37**(15): 5035-5046.
  - [25] Ren, HL, Jin, FF, Song, L, *et al.* Prediction of primary climate variability modes at the Beijing Climate Center. *J Meteorol Res* 2017; **31**(1): 204-223.
  - [26] Ren, HL, Wang, R, Zhai, P, *et al.* Upper-Ocean Dynamical Features and Prediction of the Super El Niño in 2015/16: A Comparison with the Cases in 1982/83 and 1997/98. *J Meteorol Res* 2017; **31**(2): 278-294.
  - [27] Li, Y, Chen, XR, Tan, J, *et al.* An ENSO hindcast experiment using CESM. *Acta Oceanol Sin* 2015; **37**(9): 39-50.
  - [28] Zhang, R-H, and Gao C. Processes involved in the second-year warming of the 2015 El Niño event as derived from an intermediate ocean model. *Sci China Earth Sci* 2017; **9**: 1-13.
  - [29] Zhang, R-H, and Gao C. The IOCAS intermediate coupled model (IOCAS ICM) and its real-time predictions of the 2015-2016 El Niño event. *Sci Bull* 2016; **61**(13):1061.
  - [30] Zhang, R-H, Zheng, F, Zhu, J, *et al.* A successful real-time forecast of the 2010–11 La Niña event. *Sci Rep* 2013; **3**(3): 1108.
  - [31] Zheng, F, and Zhu, J. Coupled assimilation for an intermediated coupled ENSO prediction model. *Ocean Dyn* 2010; **60**(5): 1061-1073.
  - [32] Zheng, F, and Zhu, J. Improved ensemble-mean forecasting of ENSO events by a zero-mean stochastic error model of an intermediate coupled model. *Clim Dyn* 2016; **47**(12): 3901-3915.
  - [33] Cheng, YJ, Tang, YM, Jackson, P, *et al.* Ensemble Construction and Verification of the Probabilistic ENSO Prediction in the LDEO5 Model. *J Climate* 2010; **23**(20): 5476-5497.
  - [34] Chen, D, Cane, MA, Kaplan, A, *et al.* Predictability of El Niño over the past 148 years. *Nature* 2004; **428**(6984): 733-736.
  - [35] Deber, J, and Rosati, A. A global oceanic data assimilation system. *J Phys Oceanogr* 1989; **19**(9): 1333-1347.
  - [36] Oberhuber, J, M, Roeckner, E, Christoph, M, *et al.* Predicting the '97 El Niño event with a global climate model. *Geophys Res.Lett* 1998; **25**(13): 2273-2276.
  - [37] Kirtman, BP, Shukla, J, Huang, B, *et al.* Multiseasonal predictions with a coupled tropical ocean–global atmosphere system. *Mon Weather Rev* 1997; **125**(5): 789-808.
  - [38] Tang, YM, Kleeman, R, and Moore, AM. SST assimilation experiments in a tropical Pacific Ocean model. *J Phys Oceanogr* 2004; **34**(3): 623-642.
  - [39] Keenlyside, N, Latif, M, Botzet, M, *et al.* A coupled method for initializing El Nino Southern Oscillation forecasts using sea surface temperature. *Tellus A* 2005; **57**(3): 340-356.
  - [40] Merryfield, WJ, Lee, WS, Boer, GJ, *et al.* The Canadian seasonal to interannual prediction system. Part I: Models and initialization. *Mon Wea Rev* 2013; **141**(8): 2910-2945.
  - [41] Gill AE, Atmosphere-ocean dynamics. Academic Press. 1982.

- [42] Zheng, F, Zhu, J, and Zhang, R-H. Impact of altimetry data on ENSO ensemble initializations and predictions. *Geophys Res Lett* 2007; **34**(13).
- [43] Deng, ZW, Tang, YM, and Wang, G. Assimilation of Argo temperature and salinity profiles using a bias-aware localized EnKF system for the Pacific Ocean. *Ocean Model* 2010; **35**(3): 187-205.
- [44] Zhu, J, Huang, B, Zhang, R-H. *et al.* Salinity anomaly as a trigger for ENSO events. *Sci Rep* 2014; **4**: 6821.
- [45] Lu, F, Liu, Z, Zhang, S, Liu, Y, Jacob, R. Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part II: CGCM Experiments. *Mon Wea Rev* 2015; **143**: 4645-4659.
- [46] Tang, YM, Shen ZQ, and Gao, YQ. An Introduction to Ensemble-Based Data Assimilation Method in the Earth Sciences, in *Nonlinear Systems-Design, Analysis, Estimation and Control*. 2016, InTech.
- [47] Cai, M, Kalnay, E, and Toth, Z. Bred vectors of the Zebiak–Cane model and their potential application to ENSO predictions. *J Climate* 2003; **16**(1): 40-56.
- [48] Tang, YM, Kleeman, R, and Miller, S. ENSO predictability of a fully coupled GCM model using singular vector analysis. *J Climate* 2006; **19**(14): 3361-3377.
- [49] Mu, M, and Duan, WS, A new approach to studying ENSO predictability: Conditional nonlinear optimal perturbation. *Chinese Sci Bull* 2003; **48**(10): 1045-1047.
- [50] Zheng, F, Zhu J, Wang H, *et al.* Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv Atmos Sci* 2009; **26**(2): 359-372.
- [51] Moore, AM, and Kleeman, R. Skill assessment for ENSO using ensemble prediction. *Q J Roy Meteor Soc* 1998; **124**(546): 557-584.
- [52] Kirtman, BP, and Min, D. Multimodel ensemble ENSO prediction with CCSM and CFS. *Mon Weather Rev* 2009; **137**(9): 2908-2930.
- [53] Ding, RQ and Li, JP. Nonlinear finite-time Lyapunov exponent and predictability. *Physics Lett A* 2007; **364**(5): 396-400.
- [54] Moore, AM, and Kleeman, R. The dynamics of error growth and predictability in a coupled model of ENSO. *Q J Roy Meteor Soc* 1996; **122**(534): 1405-1446.
- [55] Mu, M, Xu, H, and Duan, WS. A kind of initial errors related to “spring predictability barrier” for El Niño events in Zebiak–Cane model. *Geophys Res Lett* 2007; **34**(3).
- [56] Tang, YM, and Deng, ZW. Bred Vector and ENSO Predictability in a Hybrid Coupled Model during the Period 1881–2000. *J Climate* 2011; **24**(1): 298-314.
- [57] Hou, ZL, Li, JP, Ding, RQ, *et al.* The application of nonlinear local Lyapunov vectors to the Zebiak–Cane model and their performance in ensemble prediction. *Clim Dyn* 2017; 1-22.
- [58] Cheng, YJ, Tang, YM, Jackson, P, *et al.* Further analysis of singular vector and ENSO predictability in the Lamont model-Part II: Singular value and predictability. *Clim Dyn* 2010; **35**: 827-840.
- [59] Kleeman, R, Tang, YM, and Moore, AM. The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem. *J Atmos Sci* 2003; **60**(23): 2856-2868.
- [60] Palmer, T, and Zanna, L. Singular vectors, predictability and ensemble forecasting for weather and climate. *J Phys A: Math and Theor* 2013; **46**(25): 254018.
- [61] Barkmeijer, J, Iversen, T, and Palmer, T. Forcing singular vectors and other sensitive model structures. *Q J Roy Meteor Soc* 2003; **129**(592): 2401-2423.
- [62] Duan, WS, and Zhou, FF. Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model. *Tellus A* 2013; **65**(1): 18452.
- [63] Duan, WS, Zhao, P, Hu, JY, *et al.* The role of nonlinear forcing singular vector tendency error in causing the “spring predictability barrier” for ENSO. *J Meteorol Res* 2016; **30**(6): 853-866.
- [64] Webster, PJ, and Yang, S. Monsoon and ENSO: Selectively interactive systems. *Q J Roy Meteor Soc* 1992; **118**(507): 877-926.
- [65] Yu, YS, Duan, WS, Xu, H, *et al.* Dynamics of nonlinear error growth and season-dependent predictability of El Niño events in the Zebiak–Cane model. *Q J Roy Meteor Soc* 2009; **135**(645): 2146-2160.
- [66] Ashok, K, Behera, SK, Rao, SA, *et al.* El Niño Modoki and its possible teleconnection. *J Geophys Res: Oceans* 2007; **112**(C11).
- [67] Hendon, HH, Lim, E, Wang, GM, *et al.* Prospects for predicting two flavors of El Niño.

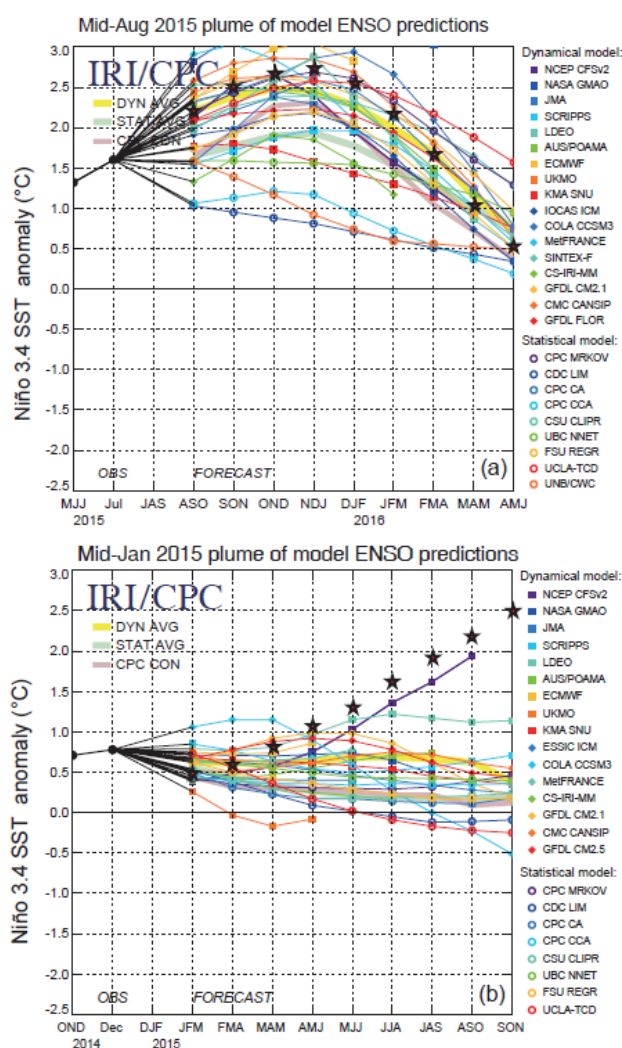
- Geophys Res Lett* 2009; **36**(19).
- [68] Tian, B. and Duan, WS. Comparison of the initial errors most likely to cause a spring predictability barrier for two types of El Niño events. *Clim Dyn* 2016; **47**(3-4): 779-792.
  - [69] Duan, WS, Li, XQ, and Tian, B. Towards optimal observational array for dealing with challenges of El Niño-Southern Oscillation predictions due to diversities of El Niño. *Clim Dyn* 2018; 1-18.
  - [70] Ren, HL, Jin, FF, Tian, B, *et al.* Distinct Persistence Barriers in Two Types of ENSO. *Geophys Res Lett* 2016; **43**(20): 10973-10979.
  - [71] Mu, M, Yu, YS, Xu, H, *et al.* Similarities between optimal precursors for ENSO events and optimally growing initial errors in El Niño predictions. *Theor Appl Climatol* 2014; **115**(3-4): 461-469.
  - [72] DelSole, T. Predictability and information theory. Part I: Measures of predictability. *J Atmos Sci* 2004; **61**(20): 2425-2440.
  - [73] Rowell D. Assessing Potential Seasonal Predictability with an Ensemble of Multidecadal GCM Simulations, *J. Climate*, 1998; **11**: 109-120.
  - [74] Tang, YM, Kleeman, R, and Moore, AM. Comparison of Information-Based Measures of Forecast Uncertainty in Ensemble ENSO Prediction. *J Climate* 2008; **21**(2): 230-247.
  - [75] Kumar, A, and Hoerling, MP. Analysis of a conceptual model of seasonal climate variability and implications for seasonal prediction. *B AM Meteorol Soc* 2000; **81**(2): 255-264.
  - [76] Kumar, A, Barnston, AG, and Hoerling, MP. Seasonal predictions, probabilistic verifications, and ensemble size. *J Climate* 2001; **14**(7): 1671-1676.
  - [77] Kumar, A, and Chen, M. Inherent predictability, requirements on the ensemble size, and complementarity. *Mon Weather Rev* 2015; **143**(8): 3192-3203.
  - [78] Kumar, A, and Hu, Z-Z. How variable is the uncertainty in ENSO sea surface temperature prediction? *J Climate* 2014; **27**(7): 2779-2788.
  - [79] Cheng, YJ, Tang, YM, and Chen, D. Relationship between predictability and forecast skill of ENSO on various time scales. *J Geophys Res* 2011; **116**(C12).
  - [80] Tang, YM, Chen, D, Yang DJ, *et al.* Methods of estimating uncertainty of climate prediction and climate change projection, in *Climate Change-Realities, Impacts Over Ice Cap, Sea Level and Risks*. 2013, InTech.
  - [81] Li, JP and Ding, RQ. Temporal-spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans. *Int J Climatol* 2013; **33**(8): 1936-1947.
  - [82] Palmer, T, The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Q J Roy Meteor Soc* 2002; **128**(581): 747-774.
  - [83] Kirtman, BP. The COLA anomaly coupled model: Ensemble ENSO prediction. *Mon Weather Rev* 2003; **131**(10): 2324-2341.
  - [84] Yang, DJ, Yang, XQ, Xie, Q, *et al.* Probabilistic versus deterministic skill in predicting the western North Pacific-East Asian summer monsoon variability with multimodel ensembles. *J Geophys Res: Atmos* 2016; **121**(3): 1079-1103.
  - [85] Xue, Y, Balmaseda, MA, Boyer, T, *et al.* A comparative analysis of upper-ocean heat content variability from an ensemble of operational ocean reanalyses. *J Climate* 2011; **25**(20): 6905-6929.
  - [86] Zhu, J, Huang, B, and Balmaseda, MA. An ensemble estimation of the variability of upper-ocean heat content over the tropical atlantic ocean with multi-ocean reanalysis products. *Clim Dyn* 2012; **39**(3-4): 1001-1020.
  - [87] Zhu, J, Huang, B, Marx, L, *et al.* Ensemble enso hindcasts initialized from multiple ocean analyses. *Geophys Res Lett* (2012); **39**(9): L09602-L09608.
  - [88] Zhu, J, Huang, B, Balmaseda, MA, *et al.* Improved reliability of enso hindcasts with multi-ocean analyses ensemble initialization. *Clim Dyn* 2013; **41**(9-10): 2785-2795.
  - [89] Tippett, MK, and Barnston, AG. Skill of multimodel ENSO probability forecasts. *Mon Weather Rev* 2008; **136**(10): 3933-3946.
  - [90] Tippett, MK, Ranganathan, M, L'Heureux, M, *et al.* Assessing probabilistic predictions of ENSO phase and intensity from the North American Multimodel Ensemble. *Clim Dyn* 2017; **68**7: 1-22.
  - [91] Mu, M, and Ren, HL. Enlightenments from researches and predictions of 2014–2016 super El Niño event. *Sci China Earth Sci* 2017; **60**(9): 1569-1571.
  - [92] Gao, C, and Zhang, R-H. The roles of atmospheric wind and entrained water temperature ( $T_e$ )

- in the second-year cooling of the 2010–12 La Niña event. *Clim Dyn* 2017; **48**(1-2): 597-617.
- [93] Zhang, R-H and Zebiak, SE. Effect of Penetrating Momentum Flux over the Surface Boundary/Mixed Layer in a z-Coordinate OGCM of the Tropical Pacific. *J Phys Oceanogr* 2001; **32**(32): 3616-3637.
- [94] Large, WG, McWilliams, JC, and Doney, SC. Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev Geophys* 1994; **32**(4): 363-403.
- [95] Zhu, Y, and Zhang, R-H. An Argo-derived background diffusivity parameterization for improved ocean simulations in the tropical Pacific. *Geophys Res Lett* 2018; **45**:1509–1517.
- [96] Zhang, R-H, Tian, F, and Wang, X. Ocean Chlorophyll-induced Heating Feedbacks on ENSO in a Coupled Ocean Physics-Biology Model Forced by Prescribed Wind Anomalies. *J Climate* 2017; **31**(5): 1811-1832.
- [97] Zhu, J., Kumar, A., Huang, B., Balmaseda, M. A., Hu, Z. Z., & Marx, L., et al. (2016). The role of off-equatorial surface temperature anomalies in the 2014 el niño prediction. *Scientific Reports*, 6.
- [98] Zhu, J., Kumar, A., Wang, W., Hu, Z., Huang, B., & Balmaseda, M. A. (1944). Importance of convective parameterization in enso predictions. *Geophysical Research Letters*, 44, 6334–6342.
- [99] Chen, D, Lian, T, Fu, CB, *et al.* Strong influence of westerly wind bursts on El Niño diversity. *Nat Geosci* 2015; **8**(5): 339-345.

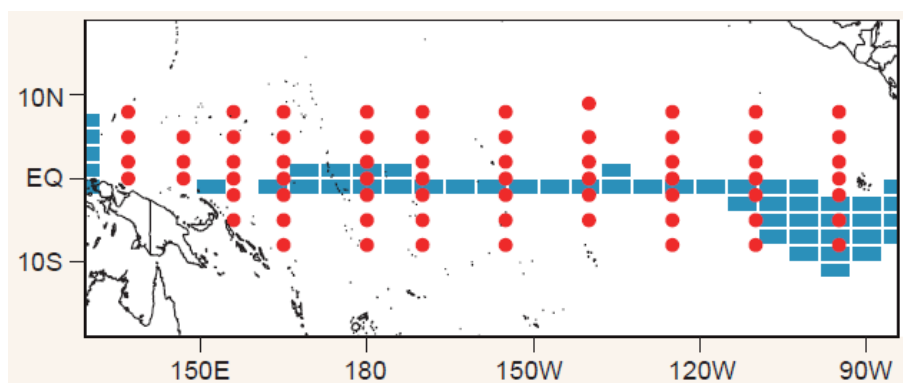




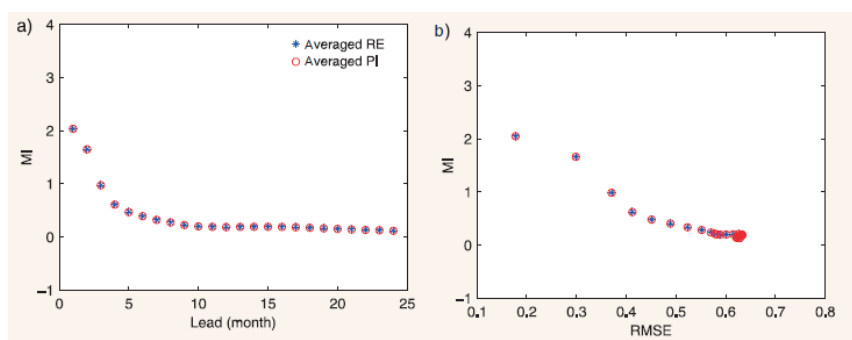
**Figure 1** Time series of observed and forecasted Niño 3.4 SST indexes at 6-month lead time. The red line represents the observations and the blue line indicates the ensemble mean, and the gray shaded area represents the prediction of 100 ensemble members.



**Figure 2** (a) The Niño 3.4 SST anomalies in 2015-16 predicted (colored lines) in mid-August 2015 using different models. Each colored line indicates a 3-month running mean of a 12-month forecast (ASO, August-September-October). (b) Same as (a), but the result in mid-Jan. The black stars indicate the observation. These figure is taken directly from the IRI website at <http://iri.columbia.edu/our-expertise/climate/forecasts/enso>.



**Figure 3** The locations of the optimal observational array (the blue rectangles) and the TAO/TRITON array (the red dots) (the TAO/TRITON array is from <http://www.pmel.noaa.gov/tao/drupal/disdsl/>). Adapted from Duan et al. (2018).



**Figure 4** (a) Variation of MI calculated by averaged relative entropy (RE) and predictive information (PI) over all initial conditions as functions of lead times. (b) Averaged RE and PI versus root mean square error (RMSE). Adapted from Cheng et al. (2011).