Correction of Monthly SST Forecasts in CFSv2 Using the Local Dynamical Analog Method

ZHAOLU HOU,[a] JIANPING LI,[a,b] AND BIN ZUO[a]

[a] *Frontiers Science Center for Deep Ocean Multispheres and Earth System (FDOMES)/Key Laboratory of Physical Oceanography/Institute for Advanced Ocean Studies/College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao, China*

[b] *Laboratory for Ocean Dynamics and Climate, Pilot Qingdao National Laboratory for Marine Science and Technology, Qingdao, China*

(Manuscript received 19 July 2020, in final form 18 December 2020)

ABSTRACT: Numerical seasonal forecasts in Earth science always contain forecast errors that cannot be eliminated by improving the ability of the numerical model. Therefore, correction of model forecast results is required. Analog correction is an effective way to reduce model forecast errors, but the key question is how to locate analogs. In this paper, we updated the local dynamical analog (LDA) algorithm to find analogs and depicted the process of model error correction as the LDA correction scheme. The LDA correction scheme was first applied to correct the operational seasonal forecasts of sea surface temperature (SST) over the period 1982–2018 from the state-of-the-art coupled climate model named NCEP Climate Forecast System, version 2. The results demonstrated that the LDA correction scheme improves forecast skill in many regions as measured by the correlation coefficient and root-mean-square error, especially over the extratropical eastern Pacific and tropical Pacific, where the model has high simulation ability. El Niño–Southern Oscillation (ENSO) as the focused physics process is also improved. The seasonal predictability barrier of ENSO is in remission, and the forecast skill of central Pacific ENSO also increases due to the LDA correction method. The intensity of the ENSO mature phases is improved. Meanwhile, the ensemble forecast results are corrected, which proves the positive influence from this LDA correction scheme on the probability forecast of cold and warm events. Overall, the LDA correction scheme, combining statistical and model dynamical information, is demonstrated to be readily integrable with other advanced operational models and has the capability to improve forecast results.

KEYWORDS: Ocean; Statistical techniques; Seasonal forecasting; Model output statistics

1. Introduction

Models in Earth science are an important tool for analyzing physical processes and forecasting natural conditions. However, a model always has shortcomings related to its numerical and physical parameterizations, initial conditions, and external forcing. Many studies have attempted to optimize model parameterization and improve the models' vertical and horizontal resolutions (Dai et al. 2003; Feng et al. 2013; Hourdin et al. 2017; Medvigy et al. 2010; Zhu et al. 2017). However, the development and improvement of a model can be a slow and computationally demanding process. Correction methods to improve numerical models are an effective way to improve the performance of numerical models. The correction schemes for model forecast error correction can be divided into state-independent and state-dependent methods (Danforth and Kalnay 2008). Generally, statistical methods to correct model forecast errors are state-independent corrections, such as model output statistics (MOS) (Carter et al. 1989; Glahn and Lowry 1972; Xu and Yang 2012). However, state-independent corrections can reduce only the systematic component of model forecast errors (Delsole and Hou 1999; DelSole et al. 2008). Considering that state-dependent components constitute the majority of total model forecast errors (Dalcher and Kalnay 1987), state-dependent corrections are therefore needed to reduce residual components and further improve forecasts.

Earth science has benefited from the proliferation of satellite data, in situ monitoring, and numerical simulations in recent years. Large databases of valuable information have been collected for oceanic and atmospheric science. Lorenz (1969) exploited the availability of such datasets facilitates the identification of states similar to a dynamic system of interest. Analogs, phenomena of atmospheric and oceanic nonlinear evolution, have been widely used in predictability and forecast studies. Based on analogs of the data, Hamill and Whitaker (2006) provided an underlying theoretical basis for reforecast analogs and applied them in probabilistic quantitative precipitation forecasts. Lguensat et al. (2017) focused on the field of data assimilation and developed an analog data assimilation technique. The analog data also have been used to develop a model-analog method to forecast tropical oceanic and atmospheric variables (Ding et al. 2018, 2019, 2020), which provided a novel view and method of operability to forecast climate variation on seasonal scale. This method selected the nearest model states from a long control run of a coupled general circulation model as analogs to the observed initial state; then their evolution within the control run provides the model-analog forecasts. The model-analog ENSO hindcast skill matches or even exceeds traditional assimilation-initialized forecast skill when being applied to the same model (Ding et al. 2018, 2019, 2020). These researches have demonstrated the usefulness of analogs in inversions of the

Corresponding author: Jianping Li, ljp@ouc.edu.cn

DOI: 10.1175/WAF-D-20-0123.1

evolving trajectories of a dynamic system in statistics. Numerical models represent the evolution of the system in dynamics. So, can analogs be used to correct numerical model forecast results and improve the forecast skill? Once the forecast state can be regarded as a small disturbance superimposed on historical analogs, statistical techniques from analogs can be used in combination with model dynamical forecasts (Gao et al. 2006; Huang et al. 1993; Ren and Chou 2006, 2007). Following this idea, Ren and Chou (2007) used the estimated forecast error to predict a final analog correction of errors and showed that this method can to some extent reduce prediction errors and improve prediction ability in forecast results of summer mean circulation and total precipitation. The analog correction method to model forecast results does not require a new model to be built; instead, it can be applied to existing numerical models and observational data. However, the key question of the analog correction method is how to locate an analog for the interest of state effectively (Badr et al. 2016; Liu and Ren 2017; Ren et al. 2009). The analog correction scheme is a promising approach to improve model forecast skill. However, in recent years, there are few studies related to how to locate an analog.

Analogs of the observational data are located using the local dynamical analog (LDA) method in the predictability studies (Ding et al. 2016; Li and Ding 2013, 2015; Li et al. 2018). Li and Ding (2011) demonstrated the benefits of using LDAs in locating analogous states compared with other local analog-based methods. The LDA method ensures similarity between the dynamical evolution of both states instead of considering only their initial states, which contributes to locate analog with high quality. However, the traditional LDA method is used for diagnosis problems but not for forecast correction applications. Therefore, Hou et al. (2020) updated the LDA method meanwhile ensuring the physical meaning of the LDA method. In Hou et al. (2020), the LDA correction method has been applied to an intermediate coupled model of El Niño–Southern Oscillation (ENSO), which confirmed that the LDA correction method can locate high quality analogs of focused states compared to other analog methods and improve model forecast performance. The LDA method is a promising method to be applied to existing numerical models and improve forecast skill. However, the LDA correction method has not yet been applied to operational state-of-the-art air–sea coupling models.

Seasonal predictions of sea surface temperature (SST) can be applied in positive ways to benefit society and economy and are now routinely produced at many operational centers using numerical models (e.g., Barnston et al. 2012; Saha et al. 2014; Stockdale et al. 2011). However, the accuracy of SST predictions is still limited (Xue et al. 2013). As an operational seasonal forecast model, National Centers for Environmental Prediction (NCEP) Climate Forecast System version 2 (CFSv2) offers routinely forecast products up to the lead time of 9 months in recent decades and has been shown to have a certain skill in seasonal climate prediction (Saha et al. 2014). Therefore, in this paper, we will describe systematically the application steps of the LDA correction scheme and correct the SST operational forecasts of CFSv2. Besides correct the deterministic forecast, we first apply the analog correction method to model ensemble forecast and evaluate the improvement of the probability forecast skill from the LDA correction method. The rest of the paper is organized as follows: the method and data used are introduced in section 2, results are described in section 3, and section 4 presents the main conclusions and discussions.

2. Method and data

a. Data

In this work, we apply the LDA correction method to the operational forecast of SST from the CFSV2. The predictions are initialized in all calendar months from January 1982 to December 2018 (Saha et al. 2014; Xue et al. 2013). For each year, predictions were produced at $1^\circ \times 1^\circ$ horizontal resolution out to a lead time of 9 months for the period 1982–2018. In this paper, the forecast used is the deterministic (ensemble mean) forecast results. Meanwhile, the ensemble members m01–m12 of CFSv2 are also corrected by the LDA-corrected method.

For verification, the monthly SSTs from the optimum interpolation sea surface temperature, version 2 (OISSTv2), dataset are the observational data used and have a horizontal resolution of 1° in latitude and longitude (Reynolds et al. 2002).

Before applied to the analog correction experiment, the CFSv2 dataset and OISSTv2 have been processed, including removing climate state and tendency to get anomalies, and removing forecast drift. Similar to Hu et al. 2014, due to the impact of a discontinuity in the ocean component, two climate states are used to compute the anomalies both for CFSv2 predicted and OISSTv2 analyzed (Xue et al. 2013; Kumar et al. 2012). The first climatology is the average between January 1982 and December 1998, and the second one between January 1999 and December 2018. Then, SST data both from the CFSv2 forecast and OISSTv2 analysis are removed the least squares linear trend of the time dimension from all grid points. The climate drift of the forecast results in different lead time have been also removed.

b. Method

The analog correction method takes advantage of model forecast errors of analogous states in historical data, which are retrieved in advance, to correct forecast results of the state of interest whose subsequent actual conditions have not yet been observed. The key question of the analog correction method is how to locate analog. In this paper, we used the LDA method.

The LDA method was first proposed to estimate the system's predictability by Li and Ding (2011). They demonstrated the benefits of this approach in locating analogous states in comparison with other analog location methods. The LDA method ensures similarity between the dynamical evolution of both states in addition to similarities in their initial states. To apply the LDA method to the model correction, Hou et al. (2020) updated the algorithm. Given state $\mathbf{y}_a(t_i)$ (the state-of-interest, representing an observational state) and its possible analog point $\mathbf{y}_a(t_k)$ from the historical dataset pool, the initial distance d_i between these two points is given by $d_i = ||\mathbf{y}_a(t_i) - \mathbf{y}_a(t_k)||$, where $||\ ||$ represents the norm

distance. Within the initial evolutionary interval $L\Delta$, where Δ is the sampling interval of the time series (i.e., $\Delta = t_i - t_{i-1}$) and L is chosen as the number of sampling intervals over the evolutionary interval, the evolutionary distance d_e between points $\mathbf{y}(t_i)$ and $\mathbf{y}(t_k)$ is given by (Hou et al. 2020)

$$d_e = \sqrt{\frac{1}{L}\sum_{j=1}^{L} \|\mathbf{y}_a(t_{i-j}) - \mathbf{y}_a(t_{k-j})\|^2}. \quad (1)$$

Thus, the total distance d_t between state points $\mathbf{y}(t_i)$ and $\mathbf{y}(t_k)$, considering not only the initial distance, but also the evolutionary distance, is defined by adding d_i and d_e (i.e., $d_t = d_i + d_e$). The term d_e represents the dynamical information. If d_t is exceedingly small, it is highly likely that the points $\mathbf{y}_a(t_k)$ is a local dynamical analog point of $\mathbf{y}_a(t_i)$. Thus, analog states of the LDA method are identified using initial and evolutionary information at two distinct time points in the time series under a system with reduced dynamical dimension. In contrast, traditional approaches to locating analogous states only consider similarities in initial spatial structures, which ignores dynamical evolution information, meanwhile based on the spatial structure of variable field whose analogs need be found in more data.

However, the state distance between a focused state and its initial-near states in the phase space increases with the dynamical system developing, the same for the focused state and its analogs. Thus, when locating an analog, the state distance $||\mathbf{y}_a(t_{i-j}) - \mathbf{y}(t_{k-j})||$ in the lead time of j always is large compared to that $||\mathbf{y}_a(t_{i-1}) - \mathbf{y}(t_{k-1})||$ in the lead time of $j = 1$. Thus, the originated evolutional distance $d_e = \sqrt{(1/L)\sum_{j=1}^{L}\|\mathbf{y}_a(t_{i-j}) - \mathbf{y}(t_{k-j})\|^2}$ gives more weight to the state distance far from the focused time. For example, $||\mathbf{y}_a(t_{i-L}) - \mathbf{y}(t_{k-L})||$ is the developed state distance and always is larger than $||\mathbf{y}_a(t_{i-1}) - \mathbf{y}(t_{k-1})||$. Compared to the pairs t_{i-1} and t_{k-1} which are more adjacent to the focused state time t_i and t_k, $||\mathbf{y}_a(t_{i-L}) - \mathbf{y}(t_{k-L})||$ should have less weight to reduce error growth effect due to the condition of t_{i-L} and t_{k-L} far from the focused state time t_i and t_k. Thus, the formula (1) with equal weight exists some deficiency in view of error growth dynamics.

To get to unequal weight coefficient, we should acquire the averaged error growth rate of two initial-near states in the phase space. In forecast concept, the initial state to forecast is the analog state of real state. Considering that the forecast for the future state can be made simply by persistent forecast results, which can be regarded as the adjacent states in the space phase under less stringent constraints. Therefore, based on the persistent forecast results, we modified d_e as

$$d_e = \sqrt{\frac{1}{L}\sum_{j=1}^{L} \|\mathbf{y}_a(t_{i-j}) - \mathbf{y}_a(t_{k-j})\|^2 \left[\frac{\mathrm{RMSE}(1)}{\mathrm{RMSE}(j)}\right]^2}, \quad (2)$$

where RMSE(j) is root-mean-square error (RMSE) of persistent forecast compared with observation at the lead time of j. With experiment proving, this modification made error evolutional conditions more authentic and improved the correction performance.

Due to the application of the LDA method, we focus on the variable on each grid of SST field. Thus, we replace the vector variable $\mathbf{y}_a$, with the time series variable y_a. The $||\ ||$ can be regarded as the absolute operation $|\cdot|$. After getting the analog $y_a(t_k)$ of the focused states $y_a(t_i)$, model forecast error predicting from the initial time t_k to time t_{k+j} can be described as follows:

$$e_j(t_k) = y_{f,j}(t_k) - y_a(t_{k+j}), \quad (3)$$

where $y_{f,j}(t_k)$ is the model forecast value at the lead time of j from $y_a(t_k)$, f represents the forecast result. As a possible analog of $y_a(t_i)$, $y_a(t_k)$ is the model initial state to predict the state $y_{f,j}(t_k)$ at the time of t_{k+j}. Meanwhile, the subsequent observational value $y_a(t_{k+j})$ corresponding to $y_{f,j}(t_k)$ has been also observed/provided. Therefore, the model forecast error $e_j(t_k)$ can be calculated as the formula (3). For the state of interest $y_a(t_i)$, its model forecast value in the lead time of j is $y_{f,j}(t_i)$, which is known. However, $y_a(t_{i+j})$ is the future observation which is unknown. Model forecast result always deviates from its corresponding observation value. Therefore, the forecast error $e_j(t_i)$ exists and is unknown but wants to be known. The observational analog state in the historical dataset is known, and the model forecast results and forecast error are also known.

Based on the thought of the analog correction method, $e_j(t_k)$ is similar to $e_j(t_i)$, to some extent, when $y_a(t_k)$ is the analog state of $y_a(t_i)$. Therefore, $e_j(t_i)$ can be estimated or corrected by $e_j(t_k)$. Considering that the forecast errors always contain some random part, the correction of $e_j(t_i)$ is performed from averaged forecast errors from several analogs with the linear regression way. When M-nearest analog states are considered for the state of interest, the mean analog forecast error $\overline{e}_j(t_k)$ is $\sum_{m=1}^{M} e_j(t_{k_m})/M$ and the estimation of $e_j(t_i)$ can be described as follows:

$$\hat{e}_j(t_i) = \alpha_j \frac{\sum_{m=1}^{M} e_j(t_{k_m})}{M} + \beta_j, \quad (4)$$

where α_j and β_j are parameters in the lead time j which can be determined by the linear regression based on the training data. Thus, $e_j(t_i)$ can be partially eliminated by $\hat{e}_j(t_i)$ and the model forecast $y_{f,j}(t_i)$ can be corrected as

$$\hat{y}_{f,j}(t_i) = y_{f,j}(t_i) - \hat{e}_j(t_i). \quad (5)$$

The diagram of the LDA correction scheme is shown in Figs. 1 and 2 displays flowchart of the LDA correction method.

In this paper, the updated LDA correction method is applied to correct the CFSv2 SST. The model forecast SST is a field variable. Here, we locate analog using the time series in every grid as $y(t_i)$. The SST time series at each grid is considered, respectively, by the LDA-analog method. Namely, $y_a(t_i)$ represents SST time series at a grid rather than considering space domains in this paper, which is due to the time series at one grid having small dimensional information and is more likely and simple to locate analog for the state of interest. Based on the time series data at a grid, the

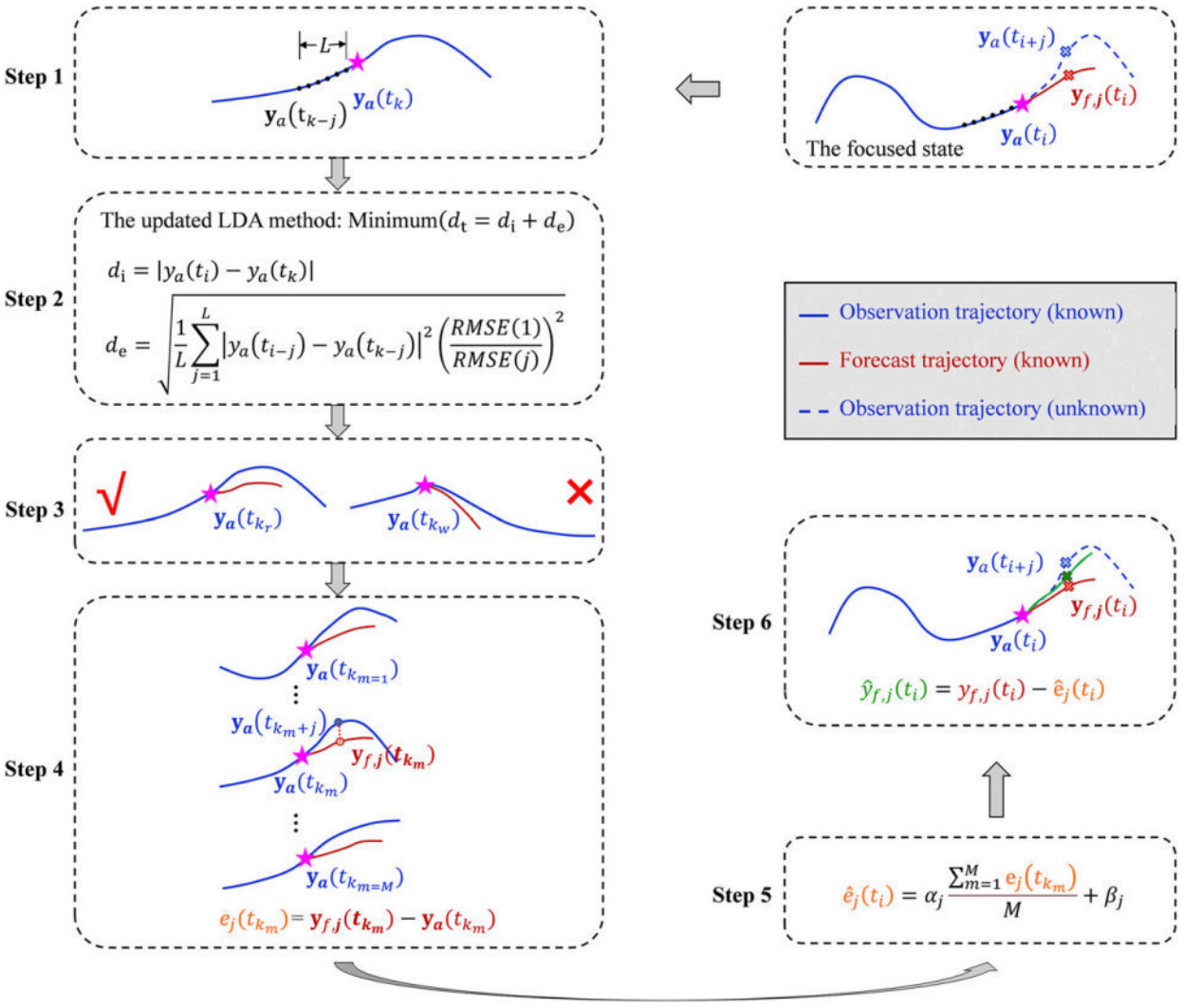

FIG. 1. The diagram of the LDA correction method to correct the model forecast. The term $\mathbf{y}_a(t_i)$ is the focused observational state at the time t_i, and $\mathbf{y}_a(t_k)$ represents its possible analog in the historical observation. The term $\mathbf{y}_{fj}(t_i)$ is the forecast at the lead time of j from the initial state $\mathbf{y}_a(t_i)$, which corresponds to $\mathbf{y}_a(t_{i+j})$.

analog is located based on the updated LDA method, including initial distance information and dynamical evolutional information. From the theoretical formula of the updated LDA method, there are some key parameters to be set.

First, the window span L in the LDA method is important for getting analog with high quality. According to Li and Ding (2013), the number of the window span L is related to the autocorrelation coefficient of the data series. Through testing, we chose the windows span $L = 6$ as lead months.

Second, for every focused state, the potential analog states are limited to historical states with the same season. In operation, the other years' calendar months of the analog are restricted to 3 months relative to that of the state of interest. For example, choosing the SST state at one grid in December 1997 as the state of interest, we regard states in October–February of 1982–95 as the possible analog states. Then, we can sort these potential states from small to large based on the distance of the LDA method. In this paper, we used $N = 6$ analogs with the smallest distance as the potential analogs.

Third, we need to choose the final analogs from the potential analogs. The forecast results of each analogs have been known, as to the forecast results of the focused state. We can calculate the correlation coefficient between the analogs' forecast results and the focused state's forecast results during the forecast period of F months ($F = 9$), that is,

$$\mathrm{corr}\langle t_{k_m}, t_i\rangle = \frac{\sum_{j=1}^{9}\left\{\left[y_{f,j}(t_i) - \mu_f(t_i)\right]\left[y_{f,j}(t_{k_m}) - \mu_f(t_{k_m})\right]\right\}}{\sqrt{\sum_{j=1}^{9}\left[y_{f,j}(t_i) - \mu_f(t_i)\right]}\sqrt{\sum_{j=1}^{9}\left[y_{f,j}(t_{k_m}) - \mu_f(t_{k_m})\right]}}, \quad (6)$$

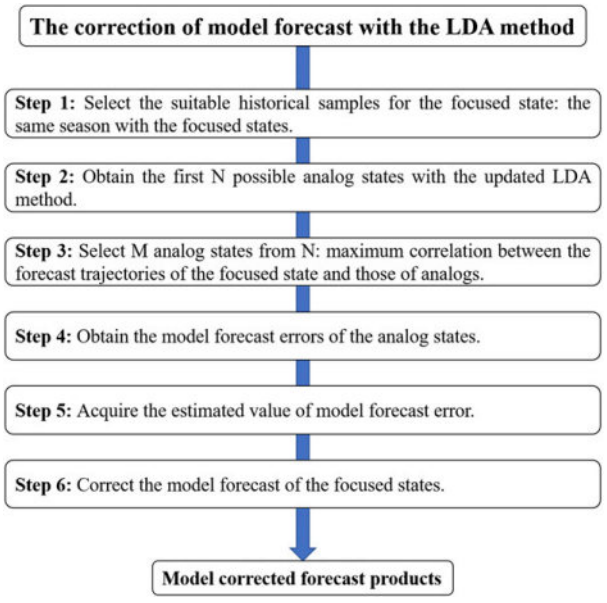

FIG. 2. Flowchart of the LDA correction method, which corresponds to Fig. 1.

where $y_{f,j}(t_i)$ is the SST forecast results initialized from t_i at the lead time of j. The term μ_f is the mean of forecast results over the forecast period of $F = 9$, and t_{k_m} is the time of the mth analog states of the focused state. Based on the correlation coefficient gotten from formula (6), We chose $M = 3$ analogs owning the highest values from six potential analogs as the final analog states and then get their average forecast error.

Fourthly, based on the averaged analog forecast errors, forecast results initialized from the focused state can be corrected with linear regression parameters α_j and β_j. The parameters α_j and β_j are determined by the linear regression process with the real forecast errors of focused states and its averaged analog forecast errors in the training period. Because of the limited amount of data, the training period is chosen by sliding selection checking off 12 months from the time of focused states. For example, when correcting the forecast results initialized from January 1997, the forecast data from CFSv2 and observational data over the period January 1982–December 1995 and January 1998–December 2018 is regarded as the training dataset. Based on the training dataset, we can get the real forecast errors and its averaged analog errors for every state. The line regression parameter can be evaluated by both forecast errors.

3. Results

In this section, to demonstrate the feasibility and effectiveness of the LDA correction method in operational seasonal prediction of SST, we corrected global monthly SST forecast from CFSv2 and used a sliding window method to divide the prediction data into two parts: training samples and testing samples. When the forecast result initiating from a focused state required correction, other forecast results whose initial times were separated from state of interest time by >12 months were used to locate analog states. The raw and corrected SST forecast from CFSv2 were assessed by calculating the temporal correlation coefficient (TCC), and the root-mean-square error (RMSE) between model predictions and OISSTv2, for all initial calendar months and at lead times from 1 to 9 months.

a. *Performance of SST field correction*

Figure 3 shows the values of TCC between observations and the CFSv2 model forecast of SST during 1982–2018 for each grid cell in the global domain at 3, 6, and 9 lead months. The first column (Figs. 3a,d,g) represents the performance of the LDA correction, the second column (Figs 3b,e,h) corresponds to the performance of the raw CFSv2 forecast, and their difference is shown in the third column (Figs. 3c,f,i). The raw CFSv2 model has a reasonable ability to predict the global SST at a lead time of 3 months, especially in most areas of the central and eastern tropical Pacific, tropical Indian ocean, and the tropical Atlantic Ocean, where the highest levels of ACC forecasting ability are found in the tropical Pacific (second column of Fig. 3). In the tropical central and eastern Pacific, especially Niño-3.4 region bounded by 5°N–5°S, from 170° to 120°W, the TCC can reach 0.7 at a lead time of 3 months, which shows that CFSv2 predicts seasonal variation of SST due to ENSO moderately well (Barnston and Tippett 2013). With increasing forecast lead time, the TCC decreases, and the areas with TCC of >0.6 are concentrated mainly in the tropical central Pacific at a lead time of 9 months. However, the raw CFSv2 forecast has weak ability to simulate SST variation of extratropical regions and specially yields a low correlation coefficient in the North Atlantic, Arctic Ocean, and the South Indian Ocean at lead times of 6 months and longer.

Corrected by the LDA correction method, the prediction skill of SST in CFSv2 is improved (first column of Fig. 3), which is most evident in the difference between the LDA-corrected forecasts and the raw forecasts (Figs. 3c,f,i). Positive differences correspond to an improvement in the TCC skill due to the LDA correction method. For lead times of 3, 6, and 9 months, TCC increases in many spatial grid cells. The improvement yielded by the LDA correction method is not evenly distributed in space and is concentrated chiefly in the tropical central and eastern Pacific, the extratropical eastern Pacific, the Southern Ocean, and the Arctic Ocean. The TCC increases by >0.12 over the western and central tropical Pacific at 3 lead months, which implies that the LDA correction method may contribute to the prediction of the central Pacific (CP) ENSO. With an increase in lead months, the area of improvement in the western and central tropical Pacific at 3 months lead gradually expands east and extratropical over the Pacific, covering key areas for the development of the CP ENSO (Kao and Yu 2009; Wang et al. 2017). These areas, where the LDA method gives a large improvement, may correspond to areas where there are large differences between the model operational forecasting ability and the predictability limit for SST. For example, Newman and Sardeshmukh (2017) suggest that the forecasting ability of SST in the equatorial

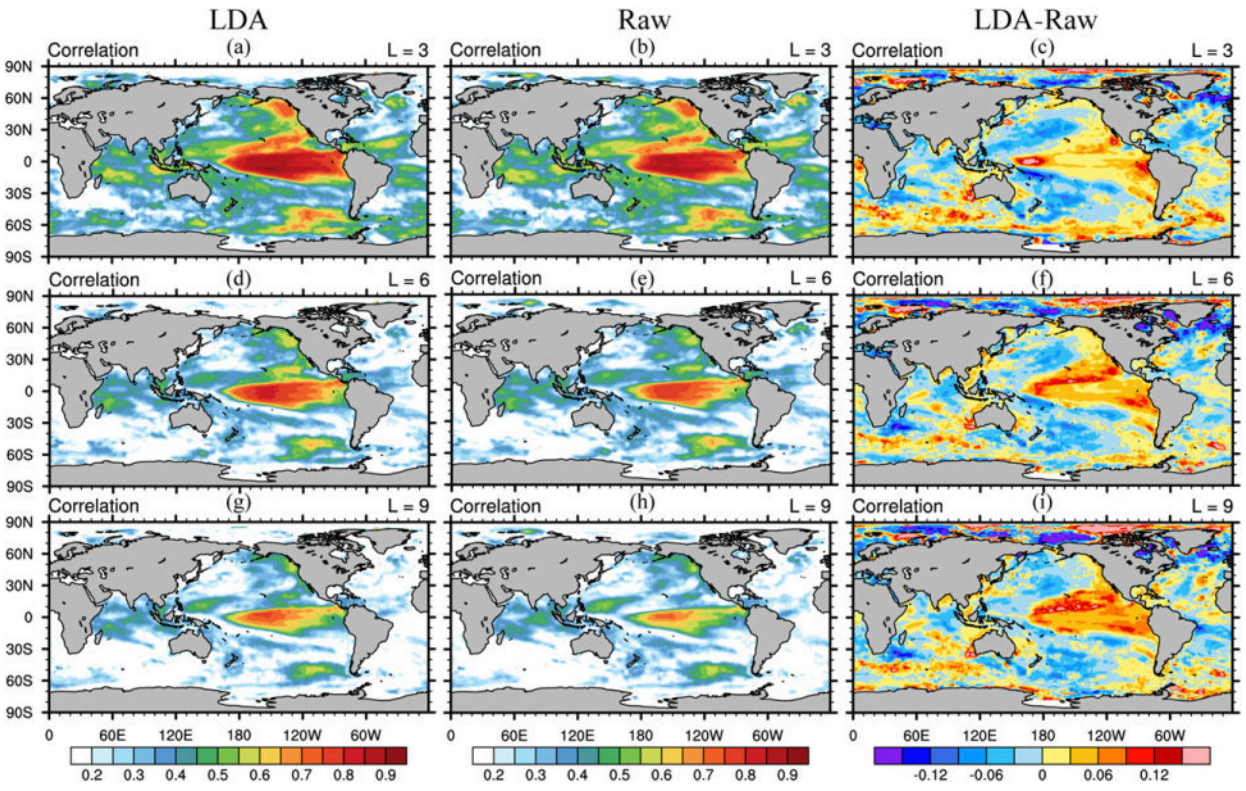

FIG. 3. The spatial distribution of temporal correlation coefficient (TCC) between observational sea surface temperature anomaly (SSTA) and (left) the CFSv2 corrected monthly prediction; (center) the CFSv2 raw prediction; and (right) the correlation coefficient difference between both results for the lead time of (a)–(c) 3, (d)–(f) 6, and (g)–(i) 9 months, during January 1982–December 2018.

Pacific is currently much lower than the predictability limit, which means that the improvement from the LDA correction method in this region is useful.

Figure 4 shows the RMSE values between observations and forecast results. RMSE values of SST increase globally with increasing lead time. Clearly, some regions, such as the Kuroshio region, Gulf Stream, and cold tongue region of the Pacific, have large RMSE values because of their large variances. After applying the LDA-correlation method, the RMSE in some regions is reduced, especially the tropical Pacific. As for the results for TCC (Fig. 3), clear improvements in RMSE are found in the eastern tropical Pacific. The RMSE value over the tropical eastern Pacific is reduced by 0.08°C at the lead time of 3, 6, and 9 months (third column of Fig. 4).

There are some regions that are worse with the correction of the LDA method, such as the western equatorial Pacific, North and South Pacific and tropical Indian Ocean. However, the western equatorial Pacific, North and South Pacific and tropical Indian Ocean are also the regions that have worse forecast performance in the raw forecast of CFSv2 in Figs. 3b, 3e, 3h, 4b, 4e, and 4h, which implies the deficiency of the CFSv2 in representing the evolutional process in these areas. In the extratropical region, the atmosphere has an important role in the evolution of SST, which brings in noise and reduce the predictability limit of SST. The CFSv2 model has low forecast skill in the extratropical region. The deficiency of the model to describe physics may lead to the increase of the randomness in the model forecast error, which is not the part of state-dependent error. The LDA correction method that mainly reduces the state-dependent error, therefore, has worse performance in these regions where the model does not have the capacity to describe its physical mechanism (shown in Fig. 13).

The CFSv2 forecast product consists of the reforecast and real-time forecast process. The CFSv2 reforecast covers January 1982–March 2011. The operation forecast has operated since May 2011. To further demonstrate the convenience and effectiveness of the LDA correction method to real-time seasonal prediction, we focused on the real-time forecast product of CFSv2. Here, we corrected the forecast results since May 2011 used the dataset from January 1982 to March 2011 as the training period to fit linear regression parameters and as the historical dataset to locate the analog. The improved performance from the LDA correction method is shown in Fig. 5 in terms of TCC and RMSE. We find the improvement is more obvious in this period than that of the results from Fig. 3 and Fig. 4. In Fig. 5, the LDA correction method can significantly enhance TCC by 0.15 in the tropical western Pacific over the whole lead times from 3 to 9 months, along with the decrease of RMSE by 0.1°C in the equatorial Pacific. It is also worth noting that the improvement of TCC skill is greater than 0.06 in the tropical Atlantic at the

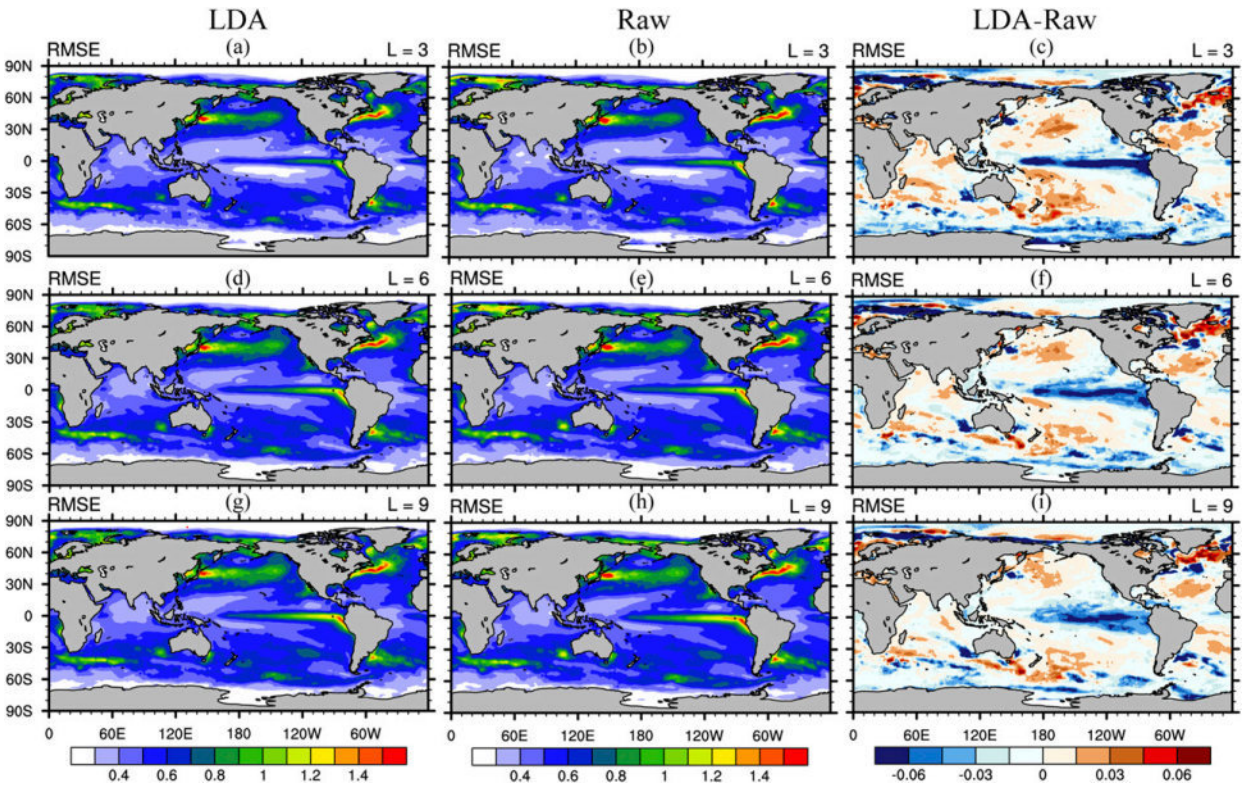

FIG. 4. As in Fig. 3, but for RMSE.

lead time of 6 months. The variation of SST in the tropical Atlantic plays an important role in South America (Cai et al. 2020), thus the improvement of the CFSv2 operational forecast skill with the LDA correction method in the tropical Atlantic contributes to early coping with seasonal climes for the people living South America.

b. *Performance of ENSO prediction correction*

The corrections yielded by the LDA method are notably well in the tropical Pacific region, especially in areas closely related to the CP ENSO. Thus, we evaluate the corrected performance of the Niño-4 index (5°N–5°S, 160°E–150°W), which is one of the most important indicators of the CP types of ENSO (Capotondi et al. 2015).

Figure 6a shows the correlation coefficient between observations and forecasts of the Niño-4 index during 1982–2018. The correlation of the corrected results is consistently higher than for the raw CFSv2 forecast, with the largest improvements being for leads of 3–5 months. The correlation coefficient averaged over 1–9 lead months is 0.84 (0.79) for the LDA-corrected (raw) forecast, and the correction difference in the lead time of 3–8 months, passes the 90% confidence level of the bootstrapping test. RMSE values are shown in Fig. 6b. The RMSE averaged over 1–9 lead months is 0.41°C for the raw results compared with 0.36°C for the LDA-corrected results, with the most significant improvement of 0.08°C occurring at leads of 3–5 months. The mean correlation and RMSE values show that the LDA correction method helps to improve the Niño-4 index forecast of the CFSV2 system and may help remedy the decrease in forecasting ability for ENSO resulting from the increased frequency of CP ENSO occurrence since 2000 (Yeh et al. 2014; Zheng et al. 2016).

The forecasting ability for ENSO varies with season, and the most well-known phenomenon is the spring forecast barrier (Duan and Wei 2013; Jin et al. 2008; Larson and Kirtman 2017). Thus, we considered the forecasting ability of the Niño-4 index at different initial calendar months during 1982–2018. Figure 7 shows the correlation coefficients and Fig. 8 the RMSE values at different calendar months. The black line represents the performance of the CFSv2 raw forecast. The correlation coefficient of the CFSv2 raw forecast is greater than 0.5 at all calendar times for different forecast lead times, demonstrating the ability of CFSv2 to forecast CP ENSO. However, there is a marked decrease in correlation coefficient when the initial times are January, February, and March, which corresponds to the ENSO forecast barrier.

The red line in Figs. 7 and 8 display the corrected results using the LDA method. The results show improvements in correlation and RMSE in the LDA-corrected forecasts for all lead months and in all initial calendar months. The improvements given by the LDA correction vary with forecast lead time. The improvement given by the LDA correction for medium (3–5 months) lead times is larger than that for short (1–2 months) and long (8–9 months) lead times.

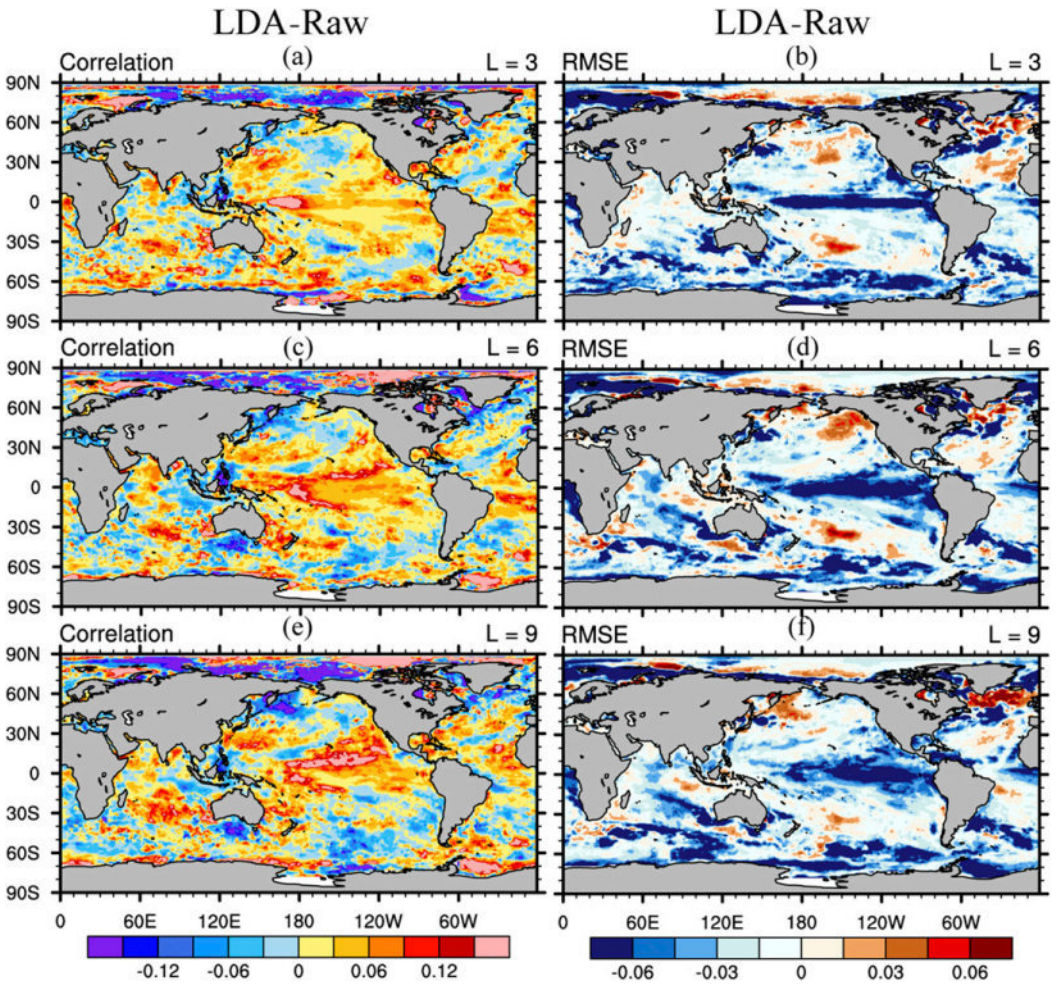

FIG. 5. The improvement of the forecast skill from the LDA correction method compared with the raw CFSv2 real-time forecast product during the period from May 2011 to December 2018. (a),(c),(e) TCC difference at the lead time of 3, 6, and 9 months, respectively, and (b),(d),(f) RMSE difference.

This feature is particularly evident in RMSE (Fig. 8). The smaller improvement given by the LDA correction method at short forecast lead times is because the forecast error of the CFSv2 raw product at these lead times is smaller. The reduction of improvement in long forecast lead times may result from the nature of the analog correction method. As noted above, the similarity of the forecast error is due to the similarity of the initial state between state of interest and analog; therefore, the similarity of the forecast errors from state of interest and analogs declines with increased forecast lead time. The correction ability of the forecast error of analogs is weak.

It should be noted that the forecast barrier of the Niño-4 index is reduced by the LDA correction method in terms of correlation coefficient (Fig. 7). Correlations of forecasts initiated from January to March are increased by 0.1 and RMSE values are reduced by 0.05°C in the lead time of 3–6 months in the LDA-corrected forecast results.

Although the LDA method yields improvements in forecasting ability over the whole period, the evolutionary trajectories of the Niño-4 index demonstrate more specific performance cases. As a real-time SST forecast case, we choose two recent strong events, El Niño event (2015/16) and La Niña event (2010/11), to assess forecast results from the raw forecast and the LDA-corrected forecast in Fig. 9. For the 2015/16 El Niño event (Fig. 9a), the raw forecast Niño-4 index initialized from February 2015 has too low a peak value than in the observations. Meanwhile, the peak of the raw forecasts occurs earlier initialized from May 2015. Encouraging results are obtained when the LDA correction method is used. The peak value of the LDA-corrected forecast becomes stronger and closer to the observations, and the peak time is improved. For the 2010/11 La Niña event, the intensity of the raw forecast results initialized from February 2011 is too small and the forecast peak value initialized from February 2011 is too larger than in the observations. As a comparison, the corrected forecast obtains a better forecast of La Niña intensity in winter.

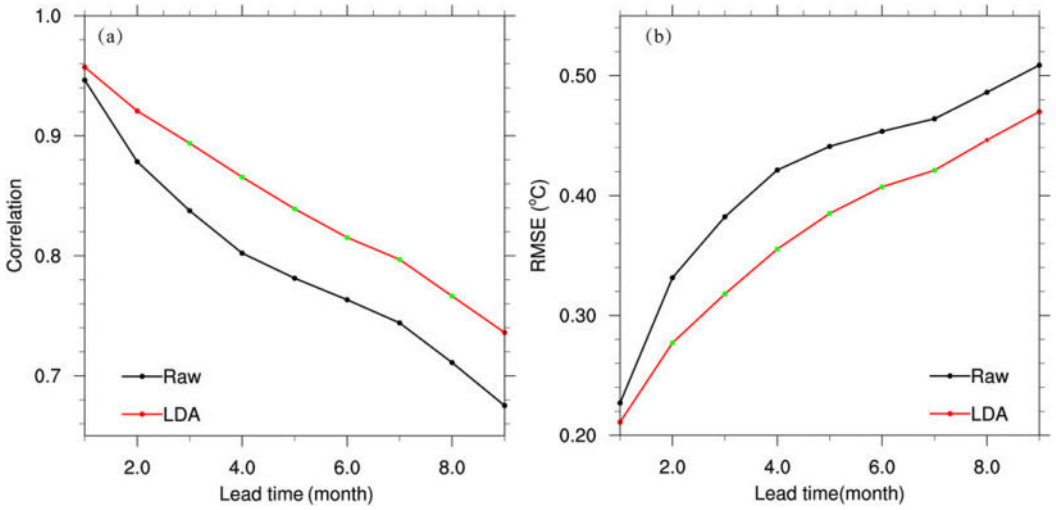

FIG. 6. (a) Temporal correlation coefficients between observations and CFSv2 raw forecasts (black), and LDA-corrected forecasts (red), for the Niño-4 index, during January 1982–December 2018. (b) Corresponding RMSE values. The green dots represent the difference between the LDA-correction skill and raw skill pass the significant level of 0.1 with the bootstrapping.

In addition, we also check the effect of historical data volume on the improved performance of the LDA correction method (no shown). For SST forecast and ENSO forecast, the results indicate the LDA correction method always improves the model forecast results regardless of the historical data amount, but the larger historical datasets to locate analog contribute the robust improvement performance.

Overall, for the deterministic (ensemble mean) forecast, the LDA correction method can further increase the correlation coefficient and reduce the RMSE of Niño-4 index forecasts at different initial calendar times by improving forecasts of the intensity of the Niño-4 index. It is particularly useful that the forecast barriers of the Niño-4 index for the initial times of January, February, and March can be reduced using the LDA correction method.

c. *Performance of ENSO ensemble prediction correction*

CFSv2 is an ensemble forecast system and can provide probability forecast products. Thus, to demonstrate the improvement effect of this LDA correction method on probability forecast skill, we corrected the model ensemble forecast product with 12 members. It needs to be noted that the forecast results averaged from the 12 members is different from the model deterministic forecast results shown in the paper. Therefore, their forecast correlation and RMSE skill are different. In this experiment, the number of the ensemble member from CFSv2 products is 12. For each member, we operate the model forecast correction using the LDA correction method like as the deterministic forecast in the paper. The corrected member forecasts are averaged to obtain the ensemble mean forecast.

Figure 10 shows the RMSE and spread of Niño-4 index forecast from the ensemble forecast results. Compared to the raw ensemble forecast results, the LDA correction method improves the forecast skill of each member and ensemble mean. The ensemble spread of the raw forecast results of the CFSv2 is lower than its corresponding RMSE, which implies that the CFSv2 ensemble system underestimate the source of error. The ensemble spread of the corrected ensemble forecast results with the LDA correction method is reduced with its corresponding RMSE skill. However, the ratio of the ensemble spread to RMSE of the corrected ensemble forecast of CFSv2 is lower than that of the raw ensemble forecast. The negative effect to the ensemble spread performance due to the LDA correction method may result from the fact that the LDA correction method reduces the state-dependent error of the model forecast, which has some differences in these ensemble members. When the state-dependent error in each member is reduced by the LDA correction method, the difference among these ensemble members decreases, therefore, the ensemble spread is lower than that of the raw model forecast.

However, besides the ensemble spread, the ensemble prediction has other evaluation indicators to describe the probability forecast skill. In ENSO prediction, we are interested in the prediction of the La Niña (cold events) and El Niño (warm events). In this part, the warm and cold events are chosen by 80% and 20% percentiles based on the observational Niño-4 index. The critical value of warm (cold) event is approximately 0.6°C (−0.6°C).

Relative operating characteristic (ROC) is used to investigate the relationship between the prediction hit rate and false alarm rate of binary events (Wilks 2011). For a time, the predicted events (cold events or warm events) can occur or not. The forecast results are evaluated by the corresponding observational conditions. For N forecast times, the corresponding

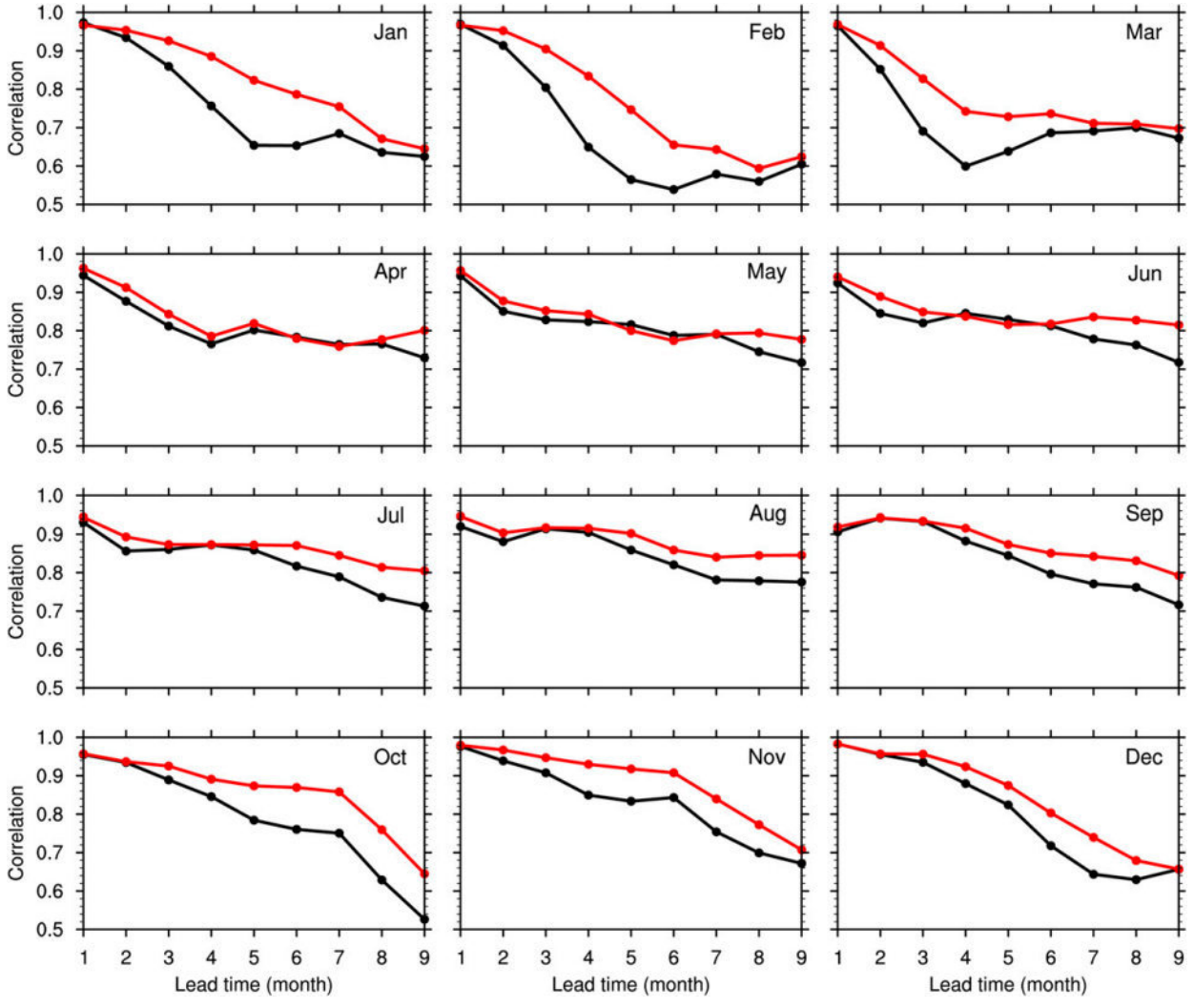

FIG. 7. Temporal correlation coefficients between observations and forecasts for the Niño-4 index, initiated in each calendar month (as marked) during January 1982–December 2018. The black line represents the CFSv2 raw prediction, and the red line is the CFSv2 forecast corrected using the LDA method.

observation shows that the number of events is H, and the number of events not occurring is NH. We define the number of forecast times as A when the observation event happens and the model forecast happens, and as B when the observation event does not happen, and the model forecast happens. Therefore, the hit rate of the kind of event can be defined as A/H; the false alarm rate is B/NH. For an event, the model members may forecast different results, thus the forecast event happening exists a ratio. Choosing different ratios as the critical probability of forecasting events, we can get different hit rates and false alarm rates, which is ROC curve as shown in Figs. 11a–c. The upper-left corner of the curve close to the coordinates indicates that the ensemble forecast has high forecast skills, while falling on the diagonal or right side of the gray indicates that the forecasting skills are poor or even no forecasting skills.

It can be seen from Figs. 11a–c that the ROC curves of cold and warm events are distributed in the upper-left corner of different forecast periods, and gradually approach the gray diagonal with the increase of forecast lead time, which indicates that the CFSv2 system has certain probability prediction skills for cold and warm events, but with the increase of forecast duration, the prediction skills gradually decrease. Compared to the raw forecast results, we find that the ROC curves of the corrected forecast results always perform better, which implies that the LDA correction method is helpful to improve ensemble probability forecast skill.

To more intuitively show the evolution of ROC prediction skill for cold and warm events with the lead time from the raw and corrected model forecast, we define an ROC area index (ROCA), which represents the area enclosed by the ROC curve and the lower-right coordinate axis (Wilks 2011):

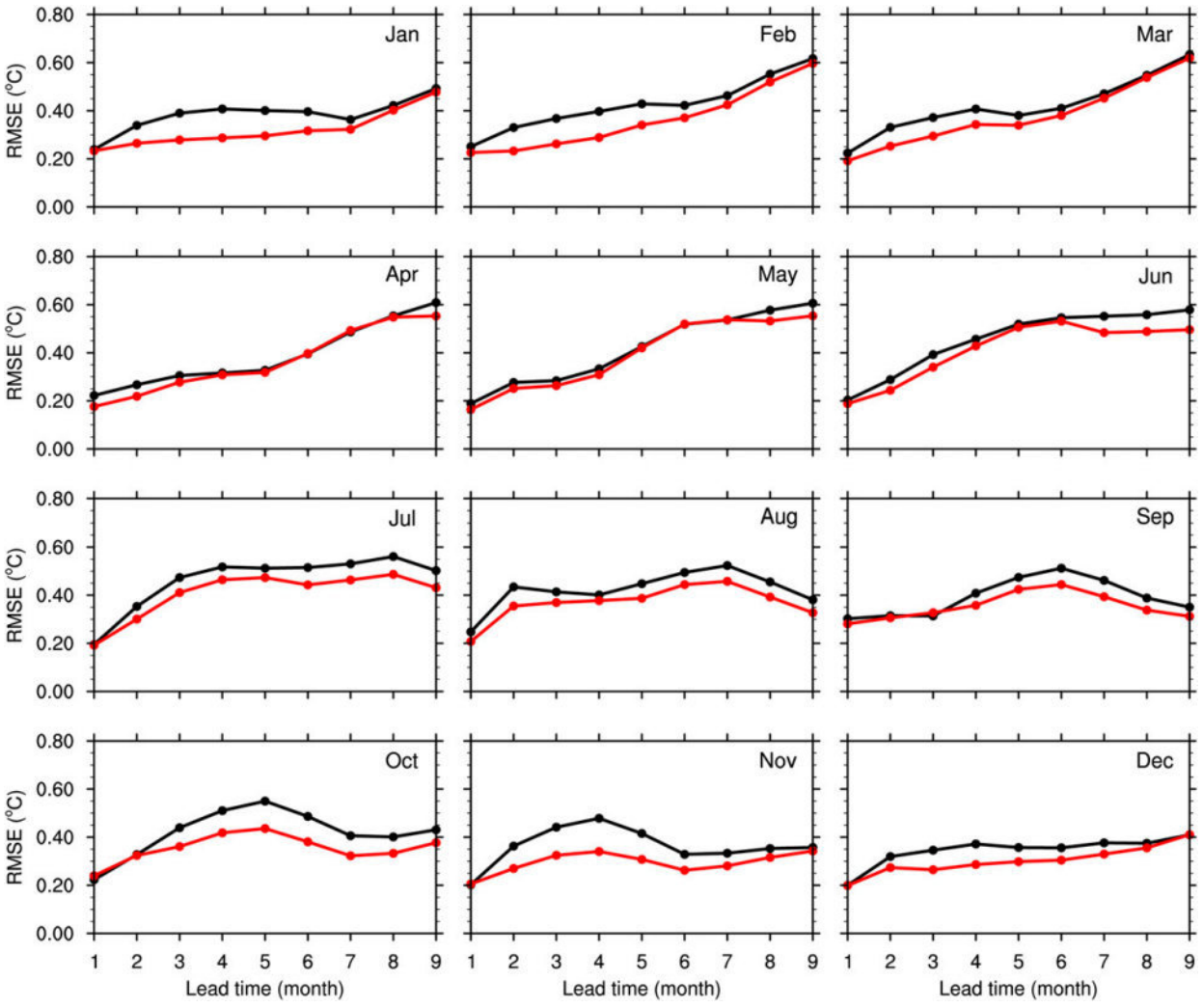

FIG. 8. As in Fig. 7, but for RMSE.

$$\mathrm{ROCA} = \sum_{i=1}^{M-1} [(H_{i+1} + H_i)(F_{i+1} - F_i)/2], \tag{7}$$

where M is the critical probability classification number. Figure 11d shows the evolution of ROCA with the forecast lead time from the raw (dotted line) and corrected (solid line) model forecast. The ROCA of cold and warm events is greater than 0.5 in all the forecast lead times, which indicates that CFSv2 has prediction skills for the two kinds of events. With the increase of the lead time, the ROCA of warm and cold events gradually decreases, which is consistent with the intuitive results in Figs. 11a–c. Comparing the solid lines with the dotted lines, we find that the ROCA of the corrected model forecast always are larger at almost all lead times, which indicates the fact that the LDA correction method contributes to the improvement of ensemble probability forecast of CFSv2.

The Brier score (BS) is most used for verification of probabilistic forecasts of dichotomous events. Thus, we use the BS to verify the prediction skill of the warm and cold event. The BS can be defined as the mean square probability error (Wilks 2011):

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N} (f_i - O_i)^2, \tag{8}$$

where the index i denotes a numbering of the N forecast–event pairs; f_i is the forecast probability of the events occurring; if the event happens, the observation $O_i = 1$, and $O_i = 0$ if the event does not occur. In this form, the BS is anticorrelated with the skill of the forecast: a higher BS represents a less accurate prediction (Murphy 1973). The BS can take on values only in the range $0 \le \mathrm{BS} \le 1$. Figure 12 shows the BS of warm and cold events evaluated for the raw model forecast (dotted line) and the corrected forecast (solid line) as a function of forecast time. At most lead times, the BS of the corrected forecast is lower than that of the raw forecast, which implies that the LDA-corrected forecast improves the performance of the model forecast to predict cold and warm events.

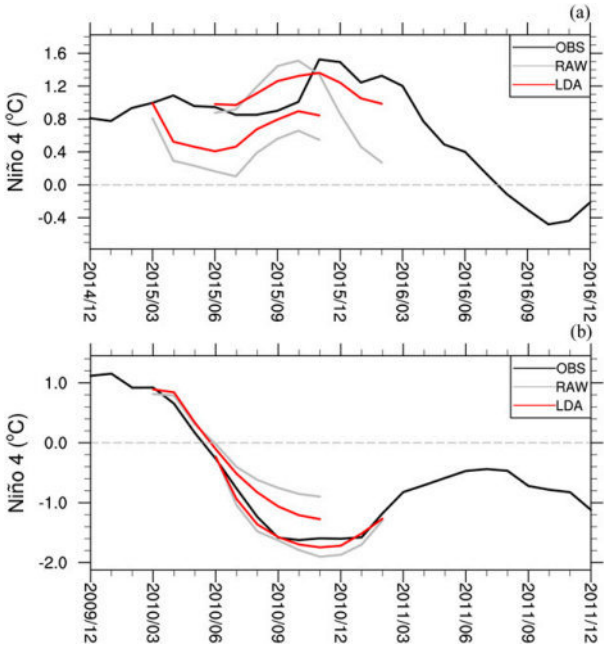

FIG. 9. Forecast Niño-4 index of the raw model forecast (gray) and the corrected model forecast by the LDA correction method (red) of (a) the 2015/16 El Niño event and (b) that of the 2009/10 La Niña event.

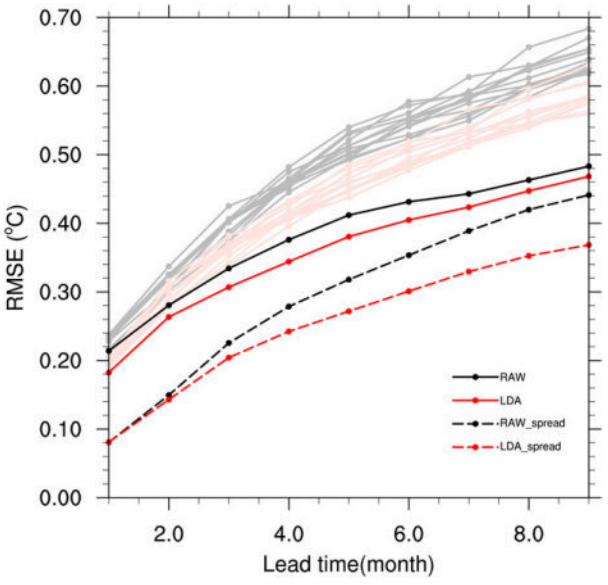

FIG. 10. The RMSE (solid line) and ensemble spread (dotted line) of the Niño-4 index from the ensemble forecast results (m01–m12) of CFSv2 over the period from January 1982 to December 2018. The gray lines are the RMSE skill of each member raw forecast results, and the light red lines represent the corrected results using the LDA correction method. The ensemble means are marked as the solid black line (raw forecast) and the solid red line (corrected forecast).

Although the corrected model forecast reduces the model forecast RMSE and spread, but has poor performance in the ratio of ensemble spread and RMSE, the LDA correction method has a positive influence on the probability forecast skill of warm and cold events.

4. Conclusions and discussion

There is forecast error in the numerical model forecast products. Model forecast correction is convenient to improve forecast skill. Considering analogs from large historical data can reproduce the evolution of a dynamical system to some extent, analogs can be applied to model forecast correction, which is still a new problem although it has been studied for a long time (e.g., Ding et al. 2020; Ren and Chou 2007). The key question of the analog correction scheme is how to locate analogs from historical dataset. This paper focused on this question and applied the updated LDA method to locate analogs. Compared with the algorithm of the LDA method used by Li and Ding (2011) and Hou et al. (2020), we adopted the unequal weight coefficients from RMSE of the persistent forecast to optimize the expression of evolutionary information of the LDA method in this paper.

In this study, we depicted the LDA correction scheme in detail and extended its application scope to the operational SST forecast products from the state-of-the-art coupled model CFSv2. The operation steps have been described in this paper. The experimental results showed that the performance of SST forecast product is improved by the LDA correction method. For the global SST forecast from CFSv2, the correction reduces RMSE values and increases correlation coefficients, especially in the tropical Pacific. The improvement of the model forecast from the LDA correction method is more obvious along with the increase of the lead time, which is useful to improve model forecast skill in the long-term lead time and increase the availability of forecast products. The ENSO phenomenon is the dominant seasonal–interannual signal in the tropical Pacific and an important forecast object. For ENSO, the LDA correction method increases the accuracy of the model forecast expressed in correlation increasing and RMSE decreasing of the central and eastern tropical SST. It is worth noting that CP ENSO represented by the Niño-4 index and its related regions own higher forecast skills by the LDA correction method. The seasonal predictability barrier of ENSO is also reduced by the LDA correction, especially for spring, which alleviates the spring forecast obstacle of the model. Besides improving the model ensemble mean forecast skill, the LDA correction method can also make a positive influence on the probability forecast results of ENSO cold and warm events, which is a new aspect of the application of the analog correction method. In this work, the LDA correction method has been demonstrated to be an effective local analog-based method to identify analogous trajectories and correct forecast product from the CFSv2 model. The LDA correction method requires low computing time and is easily applied to different operational forecast products. We consider that

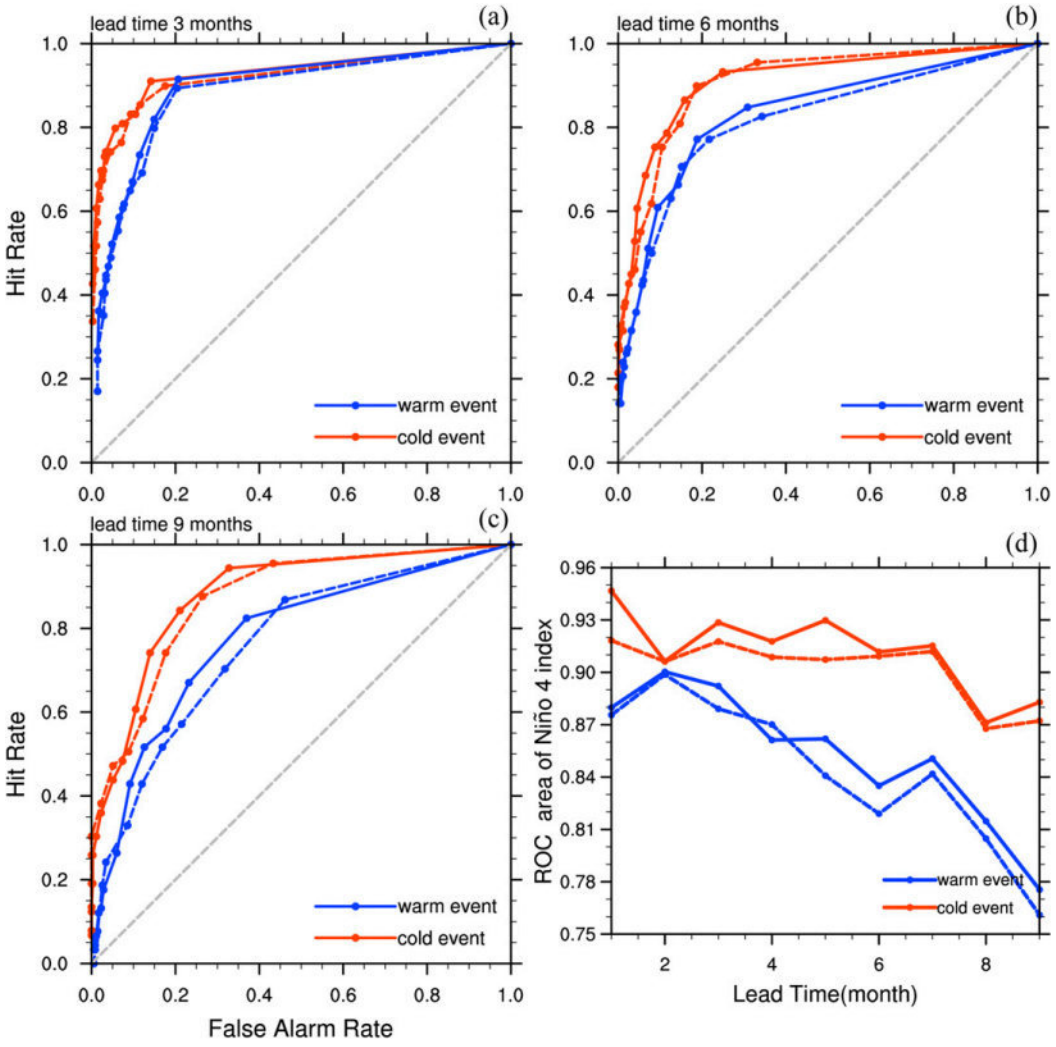

FIG. 11. Relative operating characteristic (ROC) curve for the warm (blue) and cold (red) events based on Niño-4 index in the lead time of (a) 3, (b) 6, and (c) 9 months. (d) ROC area of the Niño-4 index. The red line represents the cold event, and the blue line represents the warm event. The solid line is the corrected model forecast and the dotted line is the raw model forecast.

the LDA correction method offers a feasible approach for combining historical data and model operational forecast results.

For the analog correction method, the amount of data available is important to impact the correction performance. However, the data amount to be used to locate analogs is always limited due to the observation data. For the LDA correction method, we also evaluated its performance under a different volume of historical data (no shown). The results demonstrate the LDA correction method always improves the forecast skill of ENSO. However, as more historical data are used, the improved performance of the LDA correction method is more robust.

Besides evolutional information used by the LDA method, local analog is also an important property of the LDA method. The local feature focused by the LDA method helps to reduce the dimension of the focused system. As detailed by van den Dool (1994), the number of elements in a dataset required to guarantee the retrieval of analogs at a given precision grows exponentially with the intrinsic dimension of the state. Consequently, it is unrealistic to directly apply analog strategies to state spaces with an intrinsic dimensionality above 10, meaning that global analog forecasting operators are likely to be inappropriate for high-dimensional systems. In contrast, local analogs provide a means to decompose the analog correction of the high-dimensional state into a series of local and low-dimensional analog correction operations, which play a positive role in locating analog with quality. However, the spatial structure also contains the dynamical information of the system and needs to be used to locate the analog while the LDA method only used the time series of SST on each grid in this paper. Therefore, the LDA method will introduce the

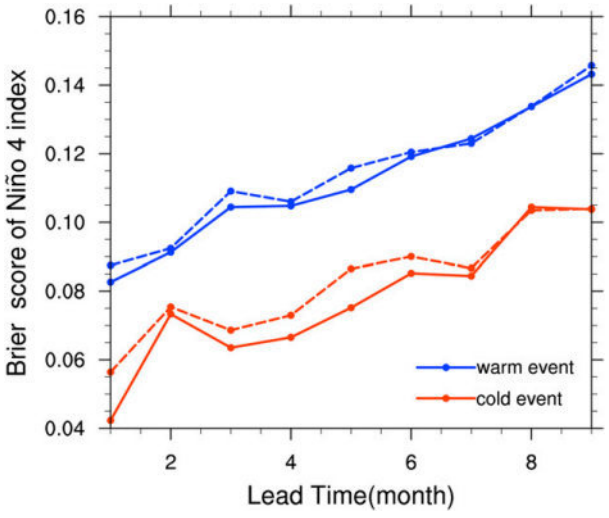

FIG. 12. Evolution of the Brier score of warm events (blue lines) and cold events (red lines) based on Niño-4 index as a function of lead time for the raw ensemble forecast (dotted lines) and the corrected forecast (solid lines) with the LDA correction method.

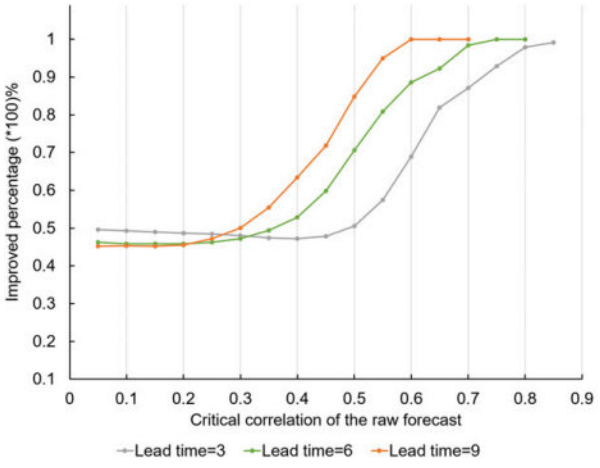

FIG. 13. The improved percentage (ordinate) of the model corrected forecast from the LDA correction method in these regions where correction skill of the raw model forecast is greater than the critical correlation (abscissa) in different lead time. The gray line is the results in the lead time of 3 months, corresponding to Fig. 3b in the paper. The green (orange) line is 6 (9) months and corresponds to Fig. 3e (3h).

spatial evolutional structure meanwhile keeping the local analog in the future.

The analog correction method improves the model forecast using statistical information from analogs by reducing model state-dependent forecast error. The analog correction method has an application premise that the model has the capacity to describe the dynamical process and its model forecast errors has similar state-dependent parts. Thus, in these regions where the raw model forecast skill is low, the improvement of forecast skill from the LDA correction method is not obvious and even negative. As a contrast, in these regions that correspond to the high forecast skill, the proportion of improved grid points from the LDA correction method is larger (as shown in Fig. 13). Therefore, the capacity of the model to depict the dynamical system and the theoretical predictability limit of the dynamical system affect the performance of the analog correction, which is valuable to be studied in the future.

Acknowledgments. This work was supported by National Natural Science Foundation of China (NSFC) Project (42005049) and the Fundamental Research Funds for the Central Universities (201962009), Shandong Natural Science Foundation Project (ZR2019ZD12, ZR2020QD056), and China Postdoctoral Science Foundation (2020M680094). We acknowledge the support of the Center for High Performance Computing and System Simulation, Qingdao Pilot National Laboratory for Marine Science and Technology.

Data availability statement. The sea surface temperature data are provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, which can be obtained from their websites (https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html). The CFSv2 data for this paper are publicly available from the NCEP Website (https://cfs.ncep.noaa.gov/).

REFERENCES

Badr, H. S., A. K. Dezfuli, B. F. Zaitchik, and C. D. Peters-Lidard, 2016: Regionalizing Africa: Patterns of precipitation variability in observations and global climate models. *J. Climate*, **29**, 9027–9043, https://doi.org/10.1175/JCLI-D-16-0182.1.

Barnston, A. G., and M. K. Tippett, 2013: Predictions of Nino3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Climate Dyn.*, **41**, 1615–1633, https://doi.org/10.1007/s00382-013-1845-2.

——, ——, M. L. L'Heureux, S. Li, and D. G. Dewitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, https://doi.org/10.1175/BAMS-D-11-00111.1.

Cai, W., and Coauthors, 2020: Climate impacts of the El Niño–Southern Oscillation on South America. *Nat. Rev. Earth Environ.*, **1**, 215–231, https://doi.org/10.1038/s43017-020-0040-3.

Capotondi, A., and Coauthors, 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Wea. Forecasting*, **4**, 401–412, https://doi.org/10.1175/1520-0434(1989)004<0401:SFBOTN>2.0.CO;2.

Dai, Y., and Coauthors, 2003: The common land model. *Bull. Amer. Meteor. Soc.*, **84**, 1013–1024, https://doi.org/10.1175/BAMS-84-8-1013.

Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus*, **39A**, 474–491, https://doi.org/10.1111/j.1600-0870.1987.tb00322.x.

Danforth, C. M., and E. Kalnay, 2008: Impact of online empirical model correction on nonlinear error growth. *Geophys. Res. Lett.*, **35**, L24805, https://doi.org/10.1029/2008GL036239.

DelSole, T., and A. Y. Hou, 1999: Empirical correction of a dynamical model. Part I: Fundamental issues. *Mon. Wea. Rev.*, **127**,

2533–2545, https://doi.org/10.1175/1520-0493(1999)127<2533:ECOADM>2.0.CO;2.

——, M. Zhao, P. A. Dirmeyer, and B. P. Kirtman, 2008: Empirical correction of a coupled land–atmosphere model. *Mon. Wea. Rev.*, **136**, 4063–4076, https://doi.org/10.1175/2008MWR2344.1.

Ding, H., M. Newman, M. A. Alexander, and A. T. Wittenberg, 2018: Skillful climate forecasts of the tropical Indo-Pacific Ocean using model analogs. *J. Climate*, **31**, 5437–5459, https://doi.org/10.1175/JCLI-D-17-0661.1.

——, ——, ——, and ——, 2019: Diagnosing secular variations in retrospective ENSO seasonal forecast skill using CMIP5 model analogs. *Geophys. Res. Lett.*, **46**, 1721–1730, https://doi.org/10.1029/2018GL080598.

——, ——, ——, and ——, 2020: Relating CMIP5 model biases to seasonal forecast skill in the tropical Pacific. *Geophys. Res. Lett.*, **47**, e2019GL086765, https://doi.org/10.1029/2019GL086765.

Ding, R., J. Li, F. Zheng, J. Feng, and D. Liu, 2016: Estimating the limit of decadal-scale climate predictability using observational data. *Climate Dyn.*, **46**, 1563–1580, https://doi.org/10.1007/s00382-015-2662-6.

Duan, W., and C. Wei, 2013: The "spring predictability barrier" for ENSO predictions and its possible mechanism: Results from a fully coupled model. *Int. J. Climatol.*, **33**, 1280–1292, https://doi.org/10.1002/joc.3513.

Feng, G. L., S. P. Sun, J. H. Zhao, and Z. H. Zheng, 2013: Analysis of stable components for extended-range (10-30 days) weather forecast: A case study of continuous overcast-rainy process in early 2009 over the mid-lower reaches of the Yangtze River. *Sci. China Earth Sci.*, **56**, 1576–1587, https://doi.org/10.1007/s11430-012-4527-8.

Gao, L., H. L. Ren, J. P. Li, and J. F. Chou, 2006: Analogue correction method of errors and its application to numerical weather prediction. *Chin. Phys.*, **15**, 882–889.

Glahn, H., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Hamill, T. M., and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.*, **134**, 3209–3229, https://doi.org/10.1175/MWR3237.1.

Hou, Z., B. Zuo, S. Zhang, F. Huang, R. Ding, W. Duan, and J. Li, 2020: Model forecast error correction based on the local dynamical analog method: An example application to the ENSO forecast by an intermediate coupled model. *Geophys. Res. Lett.*, **47**, e2020GL088986, https://doi.org/10.1029/2020GL088986.

Hourdin, F., and Coauthors, 2017: The art and science of climate model tuning. *Bull. Amer. Meteor. Soc.*, **98**, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1.

Hu, Z.-Z., A. Kumar, B. Huang, J. Zhu, and Y. Guan, 2014: Prediction skill of North Pacific variability in NCEP Climate Forecast System version 2: Impact of ENSO and beyond. *J. Climate*, **27**, 4263–4272, https://doi.org/10.1175/JCLI-D-13-00633.1.

Huang, J., Y. Yuhong, W. Shaowu, and C. Jifen, 1993: An analogue-dynamical long-range numerical weather prediction system incorporating historical evolution. *Quart. J. Roy. Meteor. Soc.*, **119**, 547–565, https://doi.org/10.1002/qj.49711951111.

Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dyn.*, **31**, 647–664, https://doi.org/10.1007/s00382-008-0397-3.

Kao, H. Y., and J. Y. Yu, 2009: Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Climate*, **22**, 615–632, https://doi.org/10.1175/2008JCLI2309.1.

Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, https://doi.org/10.1175/MWR-D-11-00335.1.

Larson, S. M., and B. P. Kirtman, 2017: Drivers of coupled model ENSO error dynamics and the spring predictability barrier. *Climate Dyn.*, **48**, 3631–3644, https://doi.org/10.1007/s00382-016-3290-5.

Lguensat, R., P. Tandeo, P. Ailliot, M. Pulido, and R. Fablet, 2017: The analog data assimilation. *Mon. Wea. Rev.*, **145**, 4093–4107, https://doi.org/10.1175/MWR-D-16-0441.1.

Li, J., and R. Ding, 2011: Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogs. *Mon. Wea. Rev.*, **139**, 3265–3283, https://doi.org/10.1175/MWR-D-10-05020.1.

——, and ——, 2013: Temporal-spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans. *Int. J. Climatol.*, **33**, 1936–1947, https://doi.org/10.1002/joc.3562.

——, and ——, 2015: Weather forecasting: Seasonal and interannual weather prediction. *Encyclopedia of Atmospheric Sciences*, 2nd ed. F. Zhang and G. North, Eds., Elsevier, 303–312.

——, J. Feng, and R. Ding, 2018: Attractor radius and global attractor radius and their application to the quantification of predictability limits. *Climate Dyn.*, **51**, 2359–2374, https://doi.org/10.1007/s00382-017-4017-y.

Liu, Y., and H. L. Ren, 2017: Improving ENSO prediction in CFSv2 with an analogue-based correction method. *Int. J. Climatol.*, **37**, 5035–5046, https://doi.org/10.1002/joc.5142.

Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2.

Medvigy, D., R. L. Walko, M. J. Otte, and R. Avissar, 2010: The ocean–land–atmosphere model: Optimization and evaluation of simulated radiative fluxes and precipitation. *Mon. Wea. Rev.*, **138**, 1923–1939, https://doi.org/10.1175/2009MWR3131.1.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600, https://doi.org/10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.

Newman, M., and P. D. Sardeshmukh, 2017: Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophys. Res. Lett.*, **44**, 8520–8529, https://doi.org/10.1002/2017GL074088.

Ren, H., and J. Chou, 2006: Analogue correction method of errors by combining statistical and dynamical methods. *Acta Meteor. Sin.*, **20**, 367–373.

——, and ——, 2007: Strategy and methodology of dynamical analogue prediction. *Sci. China. Ser. D Earth Sci.*, **50**, 1589–1599, https://doi.org/10.1007/s11430-007-0109-6.

——, ——, J. Huang, and P. Zhang, 2009: Theoretical basis and application of an analogue-dynamical model in the Lorenz system. *Adv. Atmos. Sci.*, **26**, 67–77, https://doi.org/10.1007/s00376-009-0067-3.

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

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1.

——, and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Stockdale, T. N., and Coauthors, 2011: ECMWF seasonal forecast system 3 and its prediction of sea surface temperature. *Climate Dyn.*, **37**, 455–471, https://doi.org/10.1007/s00382-010-0947-3.

van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus*, **46A**, 314–324, https://doi.org/10.3402/tellusa.v46i3.15481.

Wang, C., C. Deser, J.-Y. Yu, P. DiNezio, and A. Clement, 2017: El Niño and Southern Oscillation (ENSO): A review. *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, D. P. Manzello, I. C. Enochs, and P. W. Glynn, Eds., Vol. 8, Springer, 85–106.

Wilks, D., 2011: *Statistical Methods in the Atmospheric Sciences.* 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

Xu, Z., and Z. L. Yang, 2012: An improved dynamical downscaling method with GCM bias corrections and its validation with 30 years of climate simulations. *J. Climate*, **25**, 6271–6286, https://doi.org/10.1175/JCLI-D-12-00005.1.

Xue, Y., M. Chen, A. Kumar, Z. Z. Hu, and W. Wang, 2013: Prediction skill and bias of tropical Pacific sea surface temperatures in the NCEP Climate Forecast System version 2. *J. Climate*, **26**, 5358–5378, https://doi.org/10.1175/JCLI-D-12-00600.1.

Yeh, S. W., J. S. Kug, and S. Il An, 2014: Recent progress on two types of El Niño: Observations, dynamics, and future changes. *Asia-Pac. J. Atmos. Sci.*, **50**, 69–81, https://doi.org/10.1007/s13143-014-0028-3.

Zheng, F., X.-H. Fang, J. Zhu, J.-Y. Yu, and X.-C. Li, 2016: Modulation of Bjerknes feedback on the decadal variations in ENSO predictability. *Geophys. Res. Lett.*, **43**, 12 560–12 568, https://doi.org/10.1002/2016GL071636.

Zhu, J., A. Kumar, W. Wang, Z. Z. Hu, B. Huang, and M. A. Balmaseda, 2017: Importance of convective parameterization in ENSO predictions. *Geophys. Res. Lett.*, **44**, 6334–6342, https://doi.org/10.1002/2017GL073669.