International Journal of Climatology

International Journal of Climatology | RMetS

RESEARCH ARTICLE

Enhancing ENSO Ensemble Forecast Skill by a Coupled Conditional Nonlinear Optimal Perturbation Method

Lei Hu[1] | Wansuo Duan[2] | Rong Feng[2]

[1]Institute of Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing, China | [2]State Key Laboratory of Earth System Numerical Modeling and Application, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Correspondence: Rong Feng (fengrong@lasg.iap.ac.cn)

Received: 19 November 2025 | **Revised:** 9 February 2026 | **Accepted:** 16 March 2026

Keywords: ensemble forecast | ENSO | initial perturbations | nonlinearity

ABSTRACT

This study conducts ensemble forecasting experiments for El Niño–Southern Oscillation (ENSO) events spanning 1982–2015, comparing two perturbation generation methods: the coupled condition nonlinear optimal perturbation (C-CNOP) and singular vector (SV). We specifically focus on the sea temperature component of the C-CNOP, referred to as CP-T. The results demonstrate that the CP-T ensemble mean forecast outperforms the SV ensemble mean forecast in capturing both the temporal evolution of Niño3.4 sea surface temperature anomalies (SSTAs) and spatial patterns of SSTAs across the tropical Pacific. This is especially pronounced during El Niño events with strong nonlinearity and at longer lead times, effectively extending the lead times for skillful forecasts. Furthermore, it is revealed that the CP-T ensemble-mean perturbations, which incorporate nonlinear effects, can better capture the nonlinear development of analysis errors and appropriately adjust the feedback between sea temperature and wind field, resulting in higher forecast skill than the SV ensemble mean forecast. Therefore, the C-CNOP method is a valuable approach that not only appropriately considers the effect of initial coupling uncertainties but also incorporates the effect of nonlinearity, significantly improving ENSO forecasting skill.

1 | Introduction

The El Niño–Southern Oscillation (ENSO) is a dominant ocean–atmosphere coupled phenomenon in the tropical Pacific. This phenomenon often triggers extreme weather and climate anomalies across China and worldwide, resulting in devastating natural disasters. Therefore, accurate prediction of ENSO events is highly important. Over the past several decades, substantial progress has been achieved in enhancing ENSO prediction skill (Luo et al. 2016; Tang et al. 2018; Barnston et al. 2019; Yeager et al. 2022). Despite significant advancements, the forecast skill remains constrained by uncertainties in initial conditions, imperfect model formulations and inherent unpredictability. A highly promising approach to mitigating these uncertainties involves the implementation of ensemble forecasting.

Ensemble forecasting has been acknowledged by the World Meteorological Organization (WMO) as a critical strategy for advancing numerical weather prediction in the future. In traditional ensemble forecasting, the focus lies on initial uncertainties and the forecast skill is largely determined by how the initial perturbations are sampled (Du et al. 2018). In operational forecasting, the bred vector (BV) (Toth and Kalnay 1993) and the singular vector (SV) methods (Mureau et al. 1993; Molteni et al. 1996) are widely popular methods for yielding growing-type initial perturbations. Notably, the SV method, adopted in the European Centre for Medium-Range Weather Forecasts (ECMWF), has significantly advanced the numerical weather prediction. This method identifies the fast-growing perturbations within the forecast window with clear dynamic significance. The SV method has been widely applied

© 2026 Royal Meteorological Society.

in studies of ENSO forecasting and predictability using simple or intermediate models (Xue et al. 1997; Fan et al. 2000). However, as these SVs were calculated using tangent linear and adjoint models, the enormous computational workload makes it challenging to apply directly to complex coupled climate models. Kleeman et al. (2003) proposed an approach for calculating SVs, referred to as climatically relevant singular vectors (CSVs). This approach utilizes the ensemble method to compute SVs when the tangent linear and adjoint models are unavailable, and has been utilized in ENSO ensemble forecasting using a fully coupled global climate model (Tang et al. 2006).

As is well known, SVs are derived from a linear model framework, which fails to adequately capture the effects of nonlinear processes (Kleeman et al. 2003; Mu et al. 2003). To address this limitation, Duan and Huo (2016) proposed the orthogonal conditional nonlinear optimal perturbation (O-CNOP) ensemble forecasting method, which fully considers the impacts of nonlinear physical processes and has been utilized in studies of tropical cyclone ensemble forecast (Duan and Huo 2016; Huo and Duan 2018; Zhang et al. 2023). Nonetheless, similar to SVs, the computation of O-CNOP requires the availability of tangent linear and adjoint models, which may be absent in many coupled global climate model. To overcome these challenges, Duan et al. (2024) further developed the coupled CNOP (C-CNOP) method based on the O-CNOP method, which can not only consider nonlinear fast-growing initial perturbations, but also be implemented in numerical coupled models even in the absence of tangent linear and adjoint models. In particular, this approach can capture initial coupling uncertainties across different components of the Earth system. Duan et al. (2024) applied the C-CNOP approach to ENSO ensemble forecasting and significantly improved the forecast skill by successfully

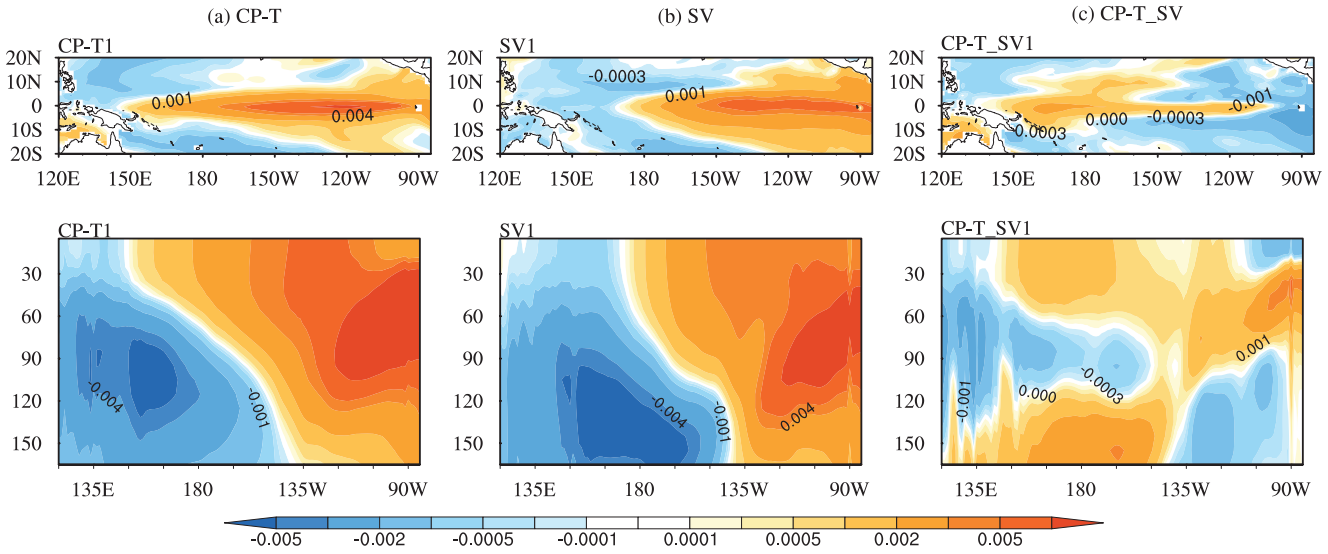

FIGURE 1 | Spatial structures of (a) the leading CNOP, (b) the leading SV and (c) their difference (CNOP minus SV) in the tropical Pacific Ocean for ENSO events. The top panels show SST anomalies, and the bottom panels display equatorial subsurface temperature anomalies. Note that the 'leading SV' refers to the first singular vector corresponding to the largest singular value obtained from the singular value decomposition, while the 'leading CNOP' refers to the first mode obtained from the EOF decomposition of the growing initial error ensemble generated by the C-CNOP algorithm. [Colour figure can be viewed at wileyonlinelibrary.com]

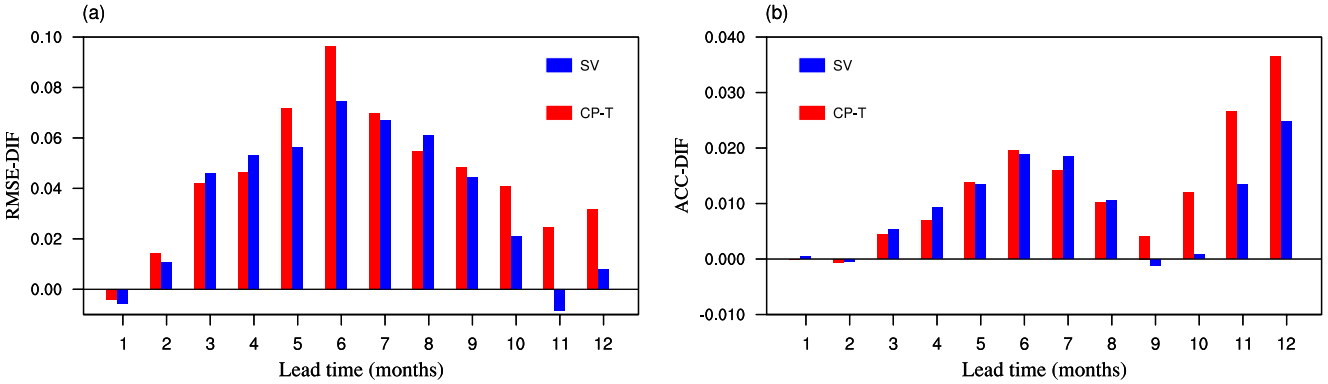

FIGURE 2 | (a) The RMSE-DIF (units: °C) and (b) ACC-DIF of the predicted Niño3.4 SSTA as a function of lead times, comparing the CP-T and SV ensemble mean forecasts. Differences are statistically significant at the 95% confidence level (Student's *t*-test). [Colour figure can be viewed at wileyonlinelibrary.com]

capturing initial coupling uncertainties, even the weakest ones during boreal spring and summer. In this study, we will further explore the important role of incorporating nonlinear effects in improving ENSO ensemble forecast skill by comparing the C-CNOP and SV methods. Since the SV method cannot adequately describe the initial coupling uncertainties between various variables (Kleeman et al. 2003), the SV method optimizes only the SST field. Following the principle of control variables, we focus solely on the effect of initial sea temperature uncertainties on ENSO predictability. Accordingly, we utilize the sea temperature component of the C-CNOP to generate ensemble initial perturbations (hereafter referred to as CP-T), thereby ensuring variable consistency. As previously stated, the SVs are derived from a linear model framework. We will investigate whether utilizing the C-CNOP method, which incorporates nonlinear effects, leads to a significant improvement in ENSO forecasting skill. In particular, since El Niño events are characterized by strong nonlinearity, we will assess whether this improvement is particularly evident during these events?

To address the aforementioned question, the paper is structured as follows. Section 2 provides a brief introduction to the model and data, as well as the SV and C-CNOP methods. All the ensemble forecast results are compared in Section 3. The analysis of physical mechanisms is presented in Section 4. Finally, a summary and discussion are provided in Section 5.

2 | Model, Data and Method

2.1 | Model and Data

The Community Earth System Model (CESM) is applied in this study, which is an atmosphere–ocean–sea ice–land coupled model. The atmospheric component is the Community Atmosphere Model 5 (CAM5) (0.9°×1.25° horizontal resolution, 30 vertical layers). The Parallel Ocean Program (POP; Smith et al. 2010) as the oceanic component has an equatorial resolution of 1°×0.27° and 60 vertical levels. The CESM suitably simulates the climatological state and interannual variability in the tropical Pacific, and reproduces the basic features of ENSO (Yao et al. 2019).

The sea temperature data of Global Ocean Data Assimilation System (GODAS) dataset is assimilated into the initial conditions of the oceanic component using the nudging method, and the initial analysis fields of the predictions are obtained. The assimilation is applied globally across the ocean, with vertical coverage extending from 15 m to 400 m depth. The assimilation spans 36 years (1980–2015). Using these initial analysis fields, we conduct tropical Pacific sea temperature forecasts through CESM integration. These forecast results serve as control forecasts (labelled Ctrl in figures) for subsequent analysis.

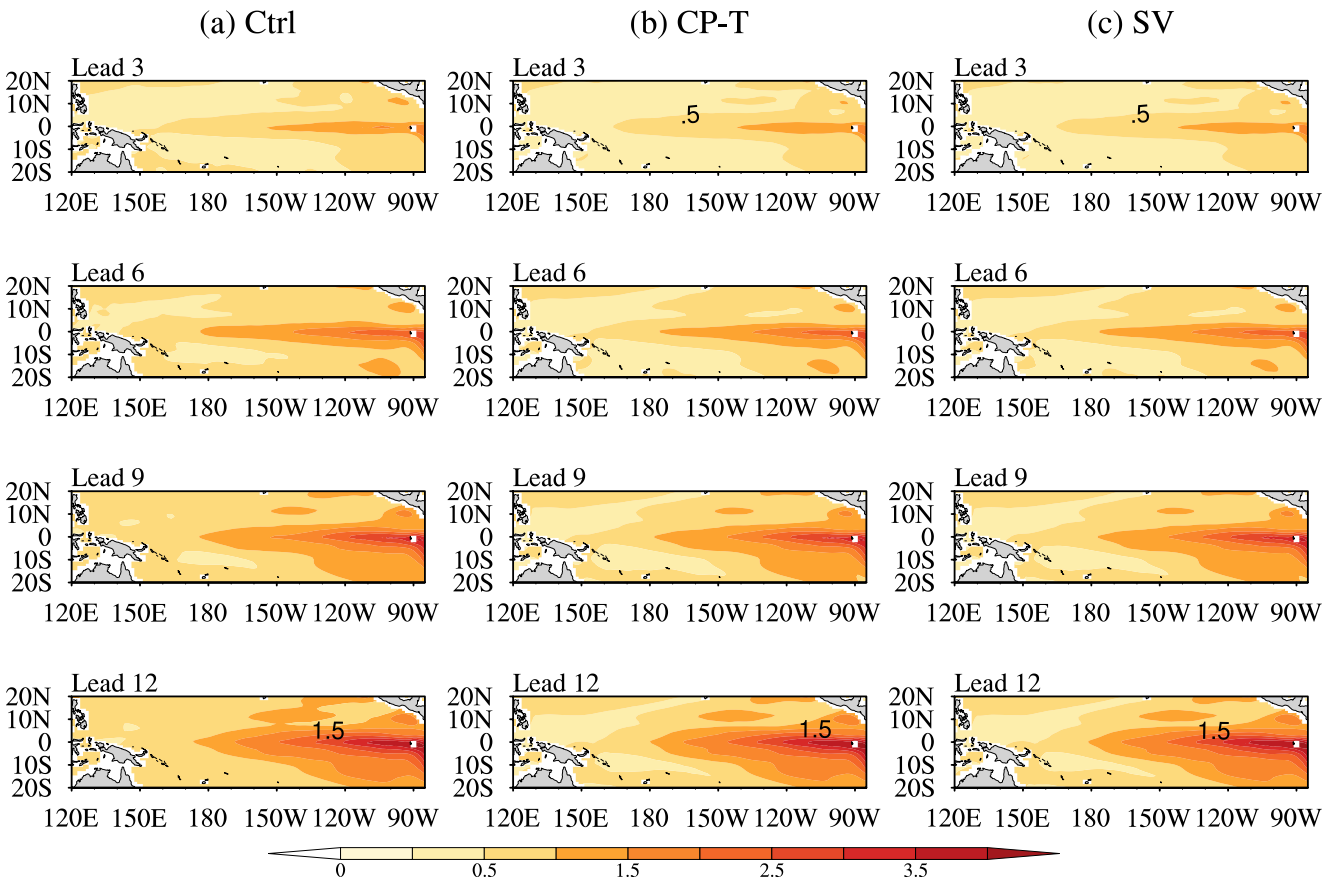

FIGURE 3 | The horizontal distributions of the RMSE between the observed and predicted SSTA at 3-, 6-, 9- and 12-month lead times during 1982–2015. (a) The control forecast; (b) CP-T and (c) SV ensemble mean forecast (units: °C). [Colour figure can be viewed at wileyonlinelibrary.com]

2.2 | The Methods Used to Yield Ensemble Initial Perturbations

2.2.1 | The C-CNOP Method

The C-CNOP method (Duan et al. 2024) is utilized to generate initial perturbations for ensemble forecasting, which takes into account the interaction uncertainties between fast-varying variables (e.g., atmospheric weather variability) and slow-varying variables (e.g., oceanic sea temperature evolution). Under constant external forcing conditions, the prediction error at the terminal time results from combined effects of the initial errors in both the slow-and fast-varying variables. The optimal initial error that maximizes the terminal prediction error can be identified by solving the following optimization problem:

$$J(u_{i,n}) = \mathrm{MAX}(1 \le n < N; i < N)\left\|X_{T_{i,n}} - X_{T_i}\right\| / \left\|X_{0_{i,n}} - X_{0_i}\right\|, \tag{1}$$

where $u_{i,n} = (x_{i,n}, y_{i,n})$ represents the initial errors in both the slow- (x) and fast-varying (y) variables that maximize the prediction error at forecast time T. What requires particular emphasis is that it not only seeks to maximize the absolute prediction error at the forecast time T but, more fundamentally, demands the most significant error growth throughout the forecast period. X_T and X_0 denote the concerned variable at the forecast time and initial time, respectively. i and n represent the numbers of observation and prediction series, and N represents the total number of series. They can be broadly selected from the model dataset to guarantee the diversity of time series. Consequently, the initial errors $u_{i,n}$ (i.e., the C-CNOP) derived from Equation (1) are computed over a finite yet statistically robust set of prediction series, and represents the statistically optimal initial error, which incorporates the dynamic interactions of uncertainties between the fast- and slow-varying variables.

Following the C-CNOP method, we selected 10 El Niño events as 'observations' from a 150-year CESM integration. For each one-year 'observation', the remaining 149 years of model data serve as the 149 'predictions' at 12-month lead time. The prediction error $J(t)$ is defined as

$$J(t) = \sqrt{\frac{1}{N}\sum_{(i,j)}\left[T^{p}_{(i,j)}(t) - T^{o}_{(i,j)}(t)\right]^2}, \tag{2}$$

where T^o and T^p denote observed and predicted sea surface temperature (SST), respectively. And grid points (i, j) span the tropical Pacific (20° S–20° N, 120° E–80° W), with N being the total

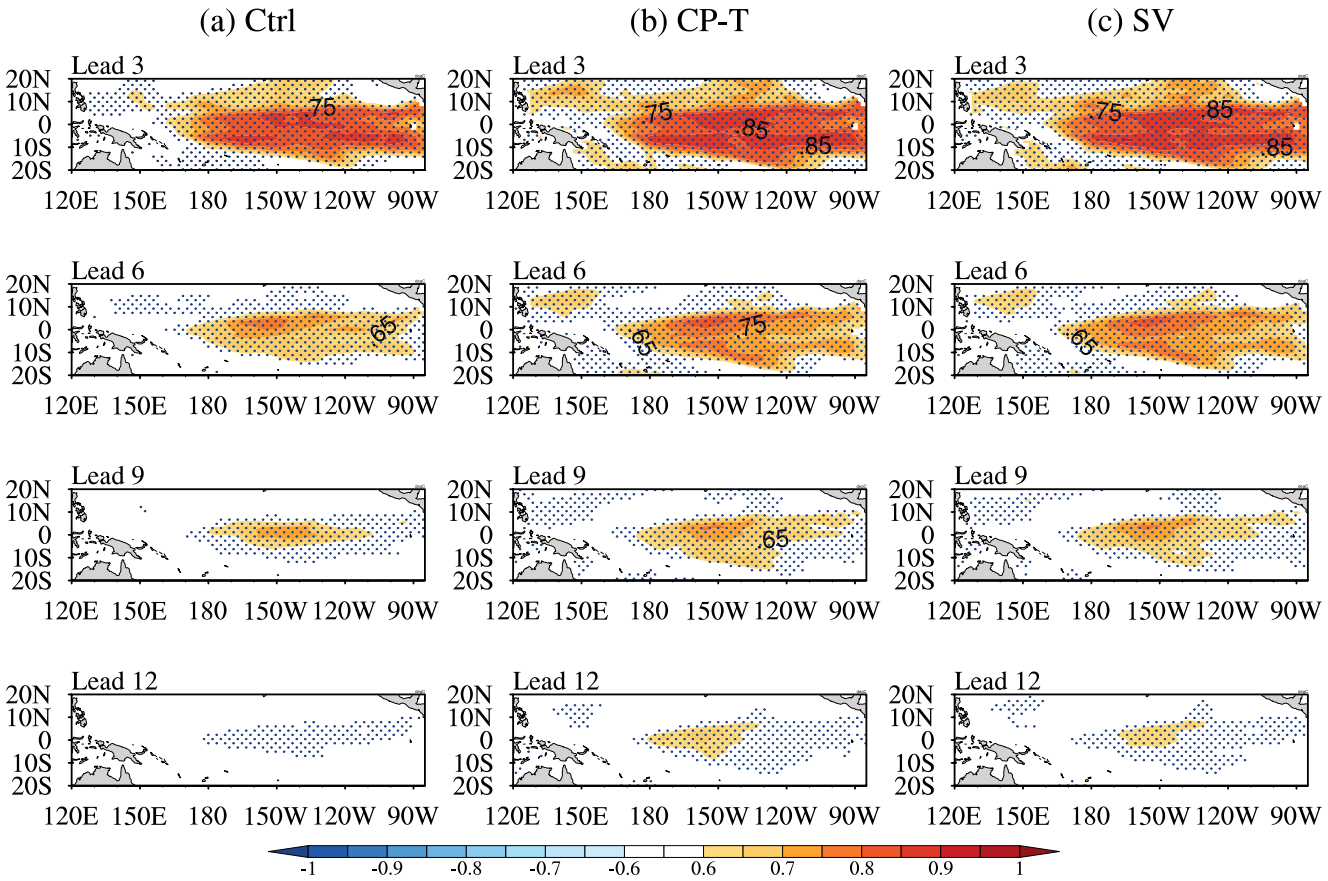

FIGURE 4 | The horizontal distributions of the ACC between the observed and predicted SSTA at 3-, 6-, 9- and 12-month lead times. (a) The control forecast, (b) CP-T and (c) SV. The dotted area denotes regions where the ACCs are statistically significant at the 99% confidence level using a r test. [Colour figure can be viewed at wileyonlinelibrary.com]

number of grid points. These prediction errors arise solely from initial errors.

The C-CNOPs are calculated for forecasts initialized in January, April, July and October. For each starting month, we derived 10 C-CNOPs corresponding to 10 'observations'. Additionally, we also calculated the second to fifth fast-growing initial errors for each 'observation'. Consequently, this yields totally 50 fast-growing initial errors for each starting month, collectively referred to as C-CNOPs. These C-CNOPs encompass global components such as the atmosphere, ocean and land surface. As introduced earlier, we only focus on slow variability, specifically selecting the sea temperature component of C-CNOP (T_0) to yield initial perturbations for ENSO ensemble forecasting. By performing empirical orthogonal function (EOF) decomposition on T_0 for a given starting month, we extract the basis of the phase space for these fast-growing initial perturbations. The dimension of the leading EOF modes (denoted as E_i) is determined by the SVD explained variance, ensuring that the EOF explained variance exceeds 90% for each starting month. Then, using a linear combination of these modes, we construct the phase space of the fast-growing initial perturbations at each starting month. The perturbation samples, denoted as CP-T $= a_1E_1 + a_2E_2 + \ \ldots \ + a_iE_i$, can be obtained from this space, where a_i represents constant coefficients. By varying the coefficients a_i, a diverse set of initial perturbations can be generated. These CP-T perturbations encompass global horizontal coverage and 17 vertical layers, extending to approximately 165 m—a depth that corresponds to the thermocline base in the tropical Pacific.

2.2.2 | The SV Method

Despite the adoption of the SV method for generating initial perturbations in ensemble forecasting, the requirement of adjoint calculations limits its widespread application in climate forecasting. As introduced by Kleeman et al. (2003), CSVs provide an alternative approach to calculate SVs by utilizing an ensemble-based method, which eliminates the dependency on tangent linear and adjoint models. This method effectively filters out weather-scale noise through ensemble averaging while retaining climatically relevant rapid growth modes. In this study, we implement the CSV calculation following the methodology described in Liu et al. (2022) to derive the initial perturbations.

We obtain several modes (denoted as F_i) and linearly combine them to yield the SVs, denoted as SV $= b_1F_1 + b_2F_2 + \ \ldots \ + b_iF_i$, where b_i denotes constant coefficients. The dimension of a_i is determined by the SVD explained variance, ensuring that the EOF explained variance exceeds 90% for each starting month. By varying the coefficients b_i, we can produce an ensemble of initial perturbations for ENSO ensemble forecasting. It is noted that, the dimensions of a_i and b_i may vary across different starting months, based on a truncation criterion requiring the cumulative

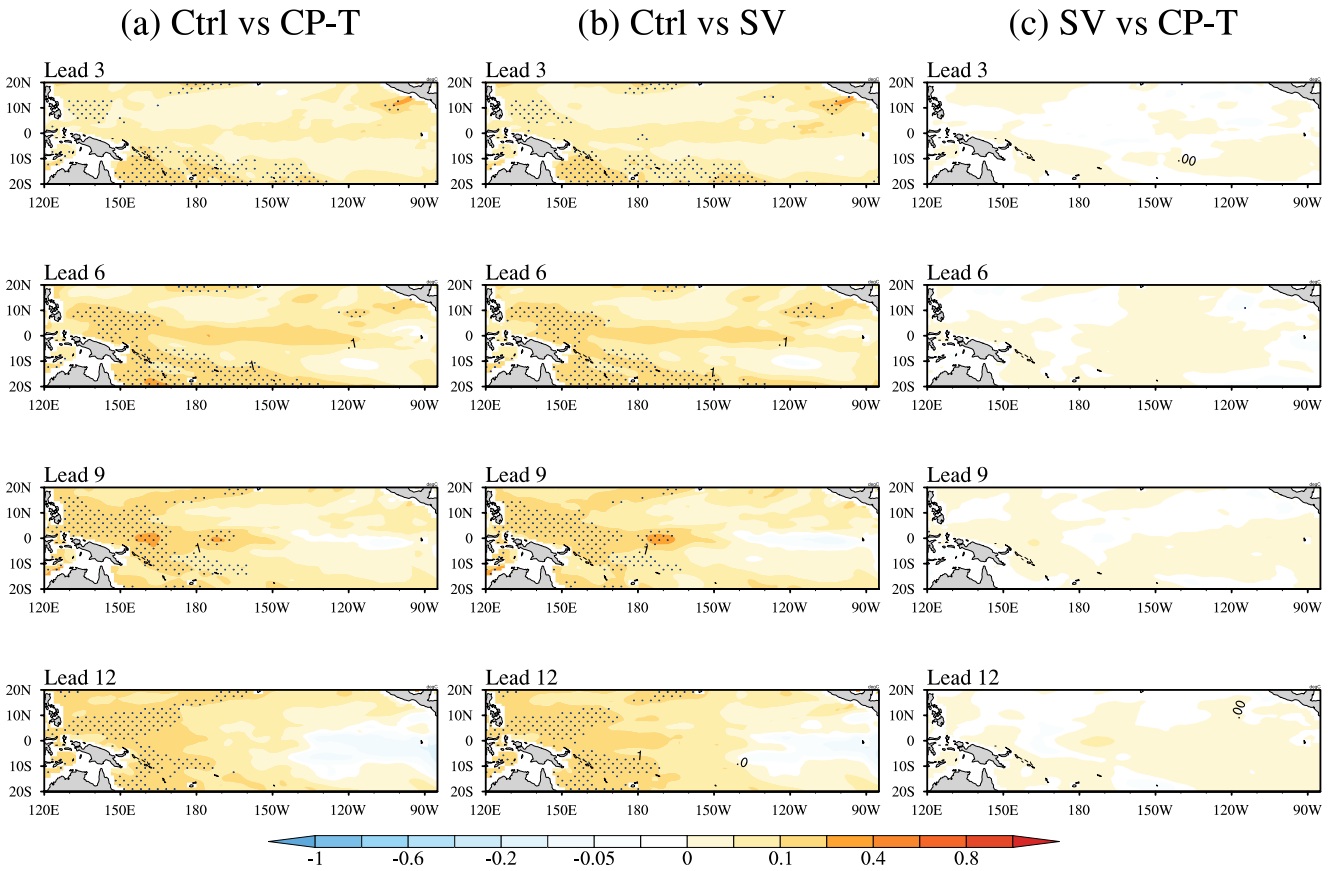

FIGURE 5 | The horizontal distributions of the difference of the SSTA RMSE between different experiments at 3-, 6-, 9- and 12-month lead times during 1982–2015. The difference of the RMSE between (a) the control forecast and CP-T, (b) the control forecast and SV and (c) SV and CP-T (units: °C). The dotted area denotes regions where differences are statistically significant at the 90% confidence level using a *t*-test. [Colour figure can be viewed at wileyonlinelibrary.com]

explained variance to exceed 90%. The horizontal and vertical extents of the SVs are configured to be identical to those of the CP-Ts. To facilitate comparison, both the CP-Ts and SVs are scaled to the identical amplitudes according to the characteristic amplitude of initial analysis errors for each starting month, as quantified by the L2 norm. Specifically, let x_{true} denotes the true initial state and x_{analysis} denotes the initial state of the analysis field. The initial analysis error is defined as $x' = x_{\text{analysis}} - x_{\text{true}}$. The L2 norm of this error, serving as the characteristic amplitude of initial analysis errors, is then calculated as

$$\|x'\| = \sqrt{\sum_{i=1}^{N} (x_i')^2}, \tag{3}$$

where N represents the total number of grid points, and x_i' denotes the error value at the ith grid point. Subsequently, the amplitudes of CP-T and CSV are normalized and scaled based on the magnitude of the initial analysis error to determine the final perturbation amplitudes.

2.2.3 | The Initial Perturbations Generated by the C-CNOP and SV Methods

Figure 1 presents the spatial patterns of the leading SV and CNOP-type initial perturbations. They both exhibit a prominent west–east dipole pattern in the tropical Pacific, which is also apparent in the subsurface layer. Nevertheless, significant differences exist between these modes. For the CNOP-type perturbations, the large-value areas are concentrated in the central-eastern Pacific, whereas in the SV-type perturbations, these regions are shifted eastward. Along the vertical section, the CNOP-type perturbations exhibit significantly stronger positive anomalies in the eastern Pacific compared to the SV-type perturbations. These structural differences emerge because the CNOP and SV represent the fastest-growing perturbations in nonlinear and linear frameworks, respectively. Given their distinct spatial patterns, we hypothesize that these differences may lead to varying performances in ENSO ensemble forecasting. In the following sections, we will evaluate this hypothesis by conducting ENSO ensemble forecasts using both CNOP and SV-type initial perturbations and compare their forecast skills.

3 | The Performance of C-CNOP and SV for ENSO Ensemble Forecasting

In this study, ensemble forecasting experiments were conducted from January 1982 to December 2015, initiated in January, April, July and October of each year, with a lead time of 12 months. To create the ensemble forecast members, we add and subtract five CP-T perturbations to and from the

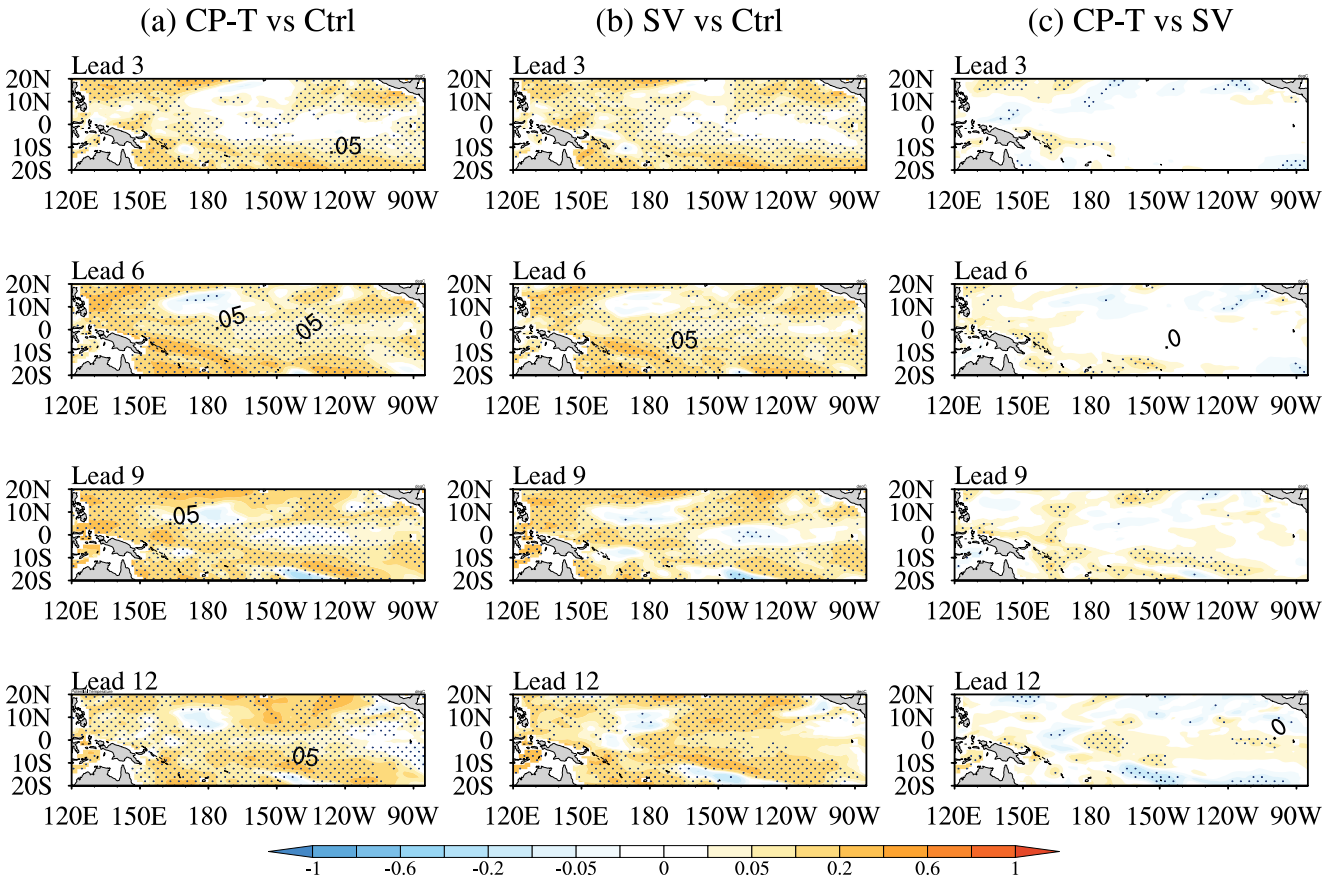

FIGURE 6 | Horizontal distributions of the difference in ACC between the observed and predicted SSTA at 3-, 6-, 9- and 12-month lead times. Difference of the ACC between (a) CP-T and the control forecast, (b) SV and the control forecast and (c) CP-T and SV. The dotted area indicates regions where differences are statistically significant at the 90% confidence level based on Fisher's Z-transformation test. [Colour figure can be viewed at wileyonlinelibrary.com]

control forecast's initial analysis field, followed by a 12-month model integration. These perturbations are synthesized as linear combinations of the leading EOF modes described in Section 2.2.1, with five distinct samples generated by randomly varying the expansion coefficients. Subsequently, a 12-month model integration is performed for each member. Combining these perturbed forecasts with the control forecast results in 11-member CP-T ensemble forecast. Using the same approach, we generate the SV ensemble forecast by adding and subtracting five SV perturbations to and from the initial analysis field, yielding an ensemble of same size.

3.1 | Ensemble Forecasting for Tropical Pacific SSTAs

3.1.1 | Temporal Variability of Niño3.4 SSTA

To investigate the ensemble forecasting skill, we employed two commonly used deterministic forecast evaluation metrics for Niño3.4 SSTA: root mean square error (RMSE) and anomaly correlation coefficient (ACC). The RMSE measures the deviation between observations and predictions, while the ACC assesses the forecast skill in capturing the temporal evolution of observations. To facilitate

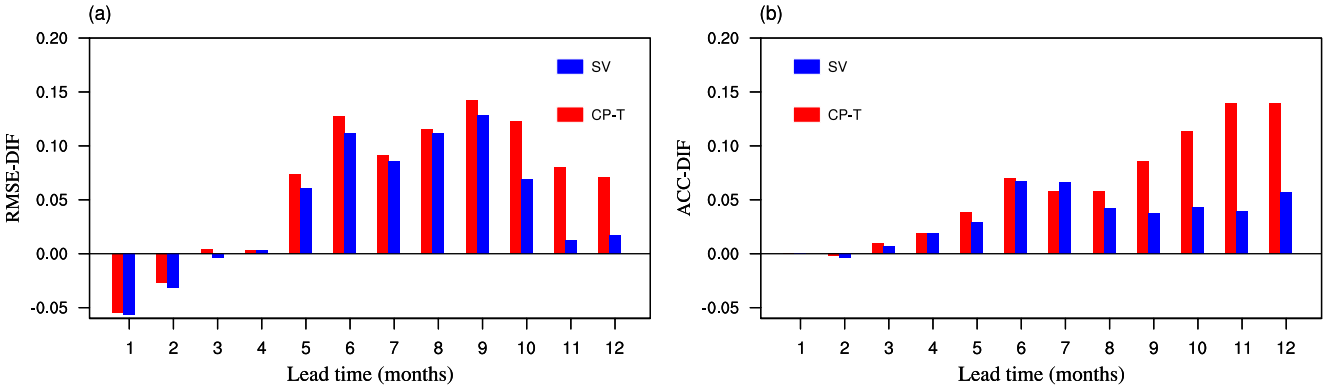

FIGURE 7 | (a) RMSE-DIF (units: °C) and (b) ACC-DIF of the predicted Niño3.4 SSTA for El Niño events as a function of lead times. Differences are statistically significant at the 95% confidence level (Student's *t*-test). [Colour figure can be viewed at wileyonlinelibrary.com]

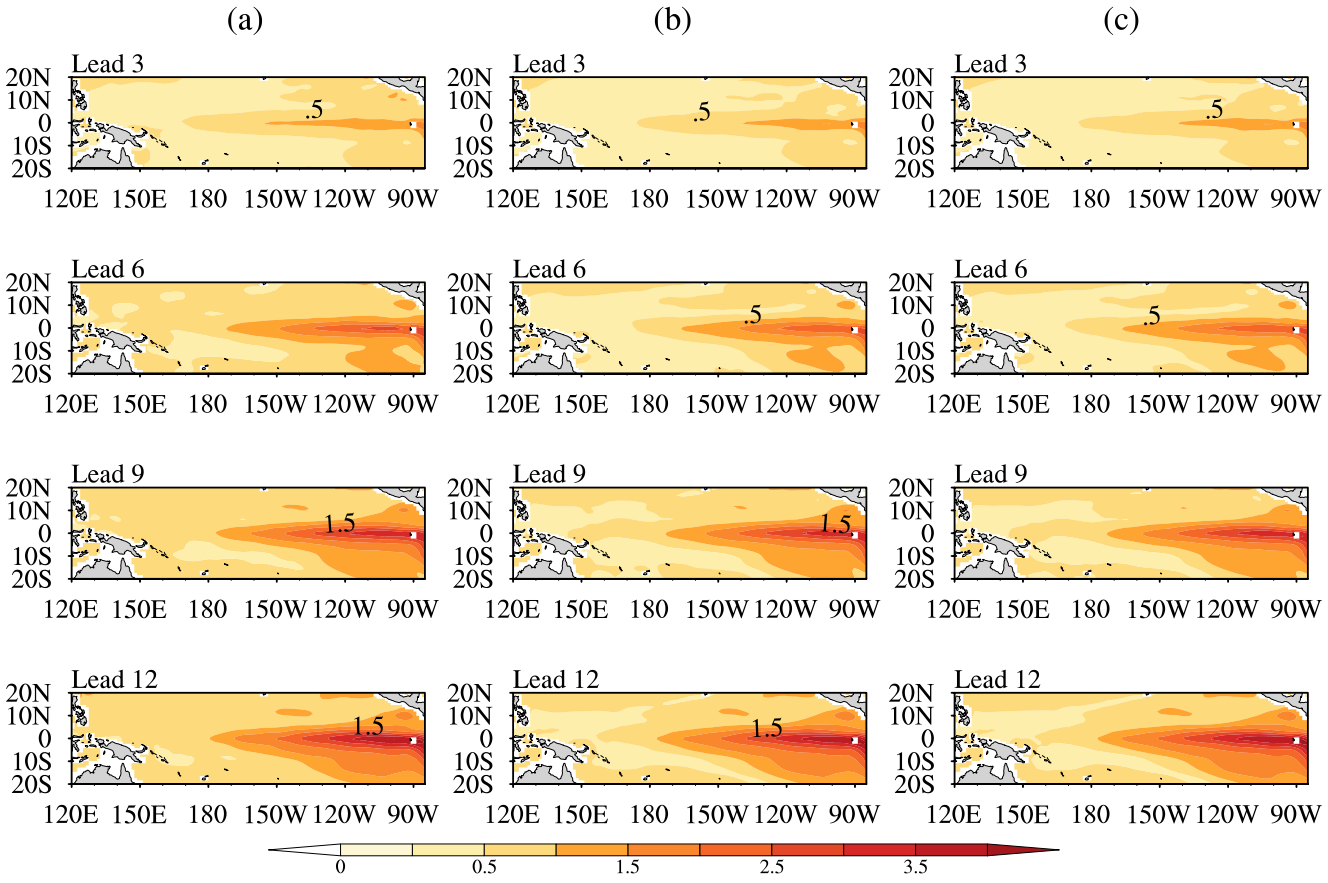

FIGURE 8 | As in Figure 3, but for the El Niño events from 1982 to 2015. (a) Ctrl. (b) CP-T. (c) SV. [Colour figure can be viewed at wileyonlinelibrary.com]

a more intuitive comparison of forecast skills between the CP-T and SV ensemble mean forecasts, we computed the differences in RMSE between the control forecast and the ensemble mean forecast (RMSE-DIF), as well as the difference in ACC between the ensemble mean forecast and the control forecast (ACC-DIF). As shown in Figure 2, the evolution of RMSE-DIF and ACC-DIF for CP-T and SV with forecast lead time is presented. Both RMSE-DIF and ACC-DIF are positive, which indicates that the ensemble forecast outperforms the control forecast, with larger values corresponding to greater improvements over the control forecast. The improvement of the CP-T ensemble mean forecast over the control forecast is superior to that of the SV ensemble mean forecast at most lead times. Moreover, with increasing lead times, the performance gap between the CP-T and SV ensemble mean forecasts widens, highlighting the growing superiority of the CP-T method with longer lead times. This may be attributed to the fact that, generally, errors at shorter lead times can be considered to grow linearly; however, as lead times increase, nonlinear effects become more prominent. This explains why the CP-T ensemble mean forecast that exhibits more nonlinear characteristics has a greater advantage over the SV ensemble mean forecast at longer lead times.

3.1.2 | The SSTA Spatial Variability in the Tropical Pacific

Figures 3 and 4 depict the spatial distributions of RMSE and ACC for (a) control forecasts, (b) CP-T ensemble mean forecasts and (c) SV ensemble mean forecasts in predicting winter SSTA variability (1982–2015). Figures 5 and 6 present the spatial distributions of RMSE and ACC differences between different forecasts. In term of RMSE, the large-value areas are mainly located in the tropical central-eastern Pacific. As lead times increase, these regions gradually expand. Both ensemble forecasts significantly reduce RMSE within the tropical Pacific, demonstrating improved forecast skill compared to the control forecast (Figure 5a,b). Furthermore, the CP-T method shows superior improvement to the SV method in the central-eastern Pacific (Figure 5c), with this improvement becoming more evident at longer lead times, although only a limited number of grid points meet the threshold for statistical significance. Regarding the ACC, the results show that high ACC values are mainly concentrated in the central-eastern tropical Pacific across all forecasts, although these values decrease with increasing lead times (see Figure 4). Both the CP-T and SV ensemble mean forecasts demonstrate higher ACC values than the control forecast across the entire tropical Pacific. Notably, they maintain significant skill (ACC > 0.6) in parts of the central Pacific even at 12-month lead time, whereas the control forecast shows negligible forecast skill. Particularly, Figure 6c reveals the ACC values in CP-T ensemble mean forecast are higher than those in the SV ensemble mean forecast. This suggests superior forecast skill in the CP-T ensemble mean forecast in predicting the spatial variability of winter SSTAs, especially at longer lead times.

3.2 | Ensemble Forecasting Skill for El Niño Events

The previous section focuses on the ensemble forecast skills regarding both the temporal and spatial variability of SSTA in the

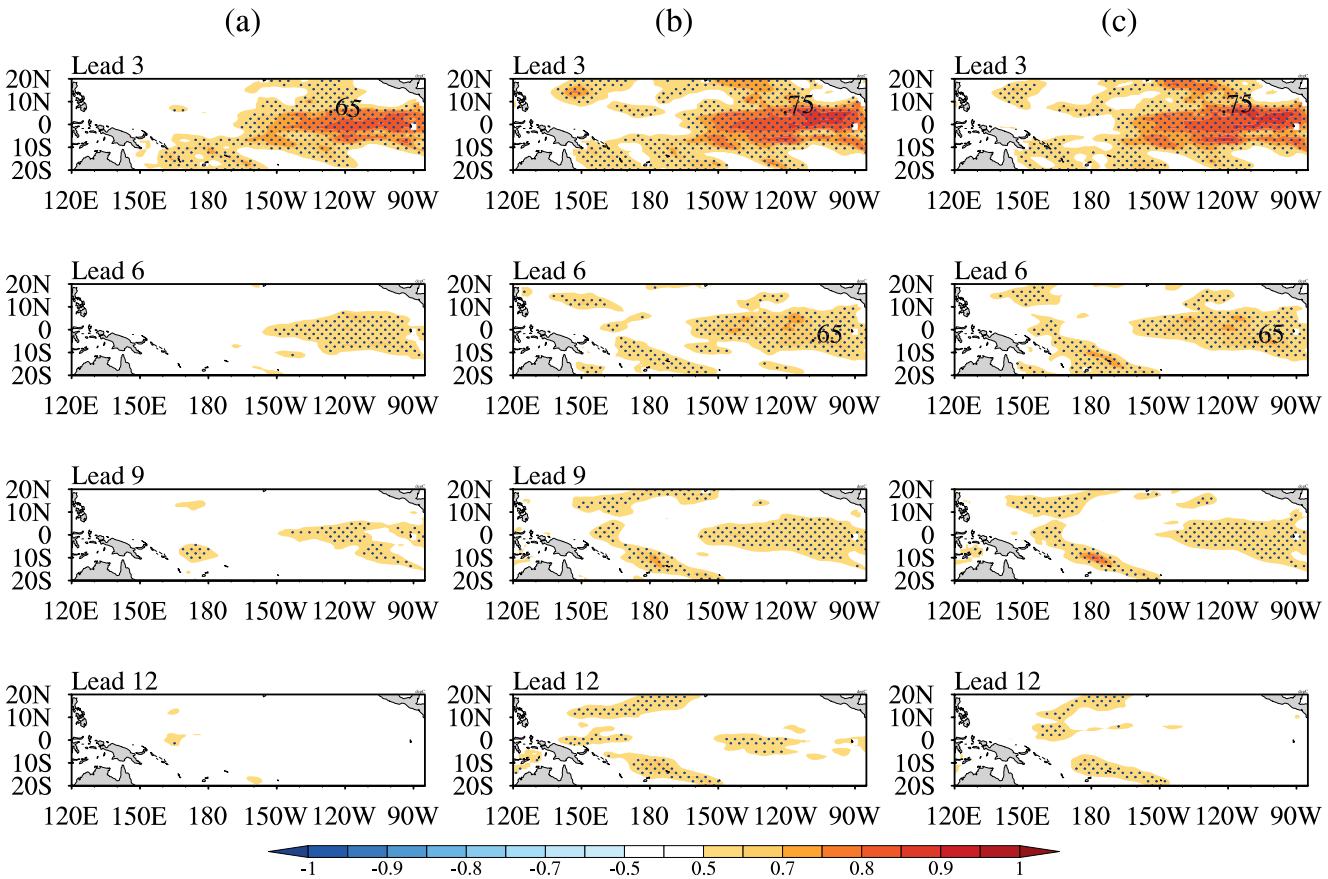

FIGURE 9 | As in Figure 4, but for the El Niño events from 1982 to 2015. (a) Ctrl. (b) CP-T. (c) SV. [Colour figure can be viewed at wileyonlinelibrary.com]

tropical Pacific from 1982 to 2015. The results demonstrate that the CP-T ensemble mean forecast exhibits higher forecast skill compared to the SV ensemble mean forecast, with this advantage becoming more pronounced as lead times increase. Given the highly nonlinear characteristics of El Niño events, this section will further compare the forecast skills of the CP-T and SV ensemble mean forecasts, specifically focusing on El Niño events, to investigate the advantage of the CP-T with nonlinear effects considered.

3.2.1 | Temporal Variability of Niño3.4 SSTA

Figure 7 depicts the evolution of RMSE-DIF and ACC-DIF for the CP-T and SV ensemble mean forecasts during the El Niño events from 1982 to 2015. The results indicate that, regardless of the ACC or RMSE, the CP-T ensemble mean forecast skill is superior to the SV ensemble mean forecast at most lead times. Consistent with the previous section, the advantage of the CP-T ensemble mean forecast over the SV ensemble mean forecast becoming increasingly prominent with increasing lead times during El Niño events. Furthermore, compared to the overall forecasting skill from 1982 to 2015 in Figure 2, the advantage of the CP-T ensemble mean forecast over the SV ensemble mean forecast is more pronounced for El Niño events, emphasizing the significant role of capturing the nonlinear effects in forecasting El Niño events.

3.2.2 | The SSTA Spatial Variability in the Tropical Pacific

Figures 8 and 9 illustrate the spatial patterns of RMSE and ACC for the SSTA prediction during the mature phase of El Niño from 1982 to 2015. Figures 10 and 11 present the differences in RMSE and ACC between different forecasts. It can be observed that the large values of RMSE are primarily concentrated in the central-eastern Pacific region, and as lead times increase, RMSE gradually increase, exceeding 1.5 at a 12-month lead time. Figure 10 indicates that ensemble forecasts significantly reduce prediction errors relative to the control forecast in the tropical Pacific at all lead times. Specifically, the CP-T ensemble mean forecast exhibits higher forecast skill than the CSV ensemble mean forecast, further reducing prediction errors in the central-eastern Pacific. Regarding the ACC of SSTA, the results also indicate that ensemble forecasts have higher forecast skills, particularly at longer lead times, thereby effectively extending the lead time of skillful forecast. Particularly, the ACC of the CP-T ensemble mean forecast is higher than that of the SV ensemble mean forecast, especially at a 12-month lead time. This further confirms that the superiority of the CP-T ensemble mean forecast over the SV ensemble mean forecast becomes more prominent with increasing lead times, consistent with earlier analytical conclusions. Moreover, when comparing the CP-T and SV ensemble mean forecasts, the advantage of CP-T is particularly pronounced

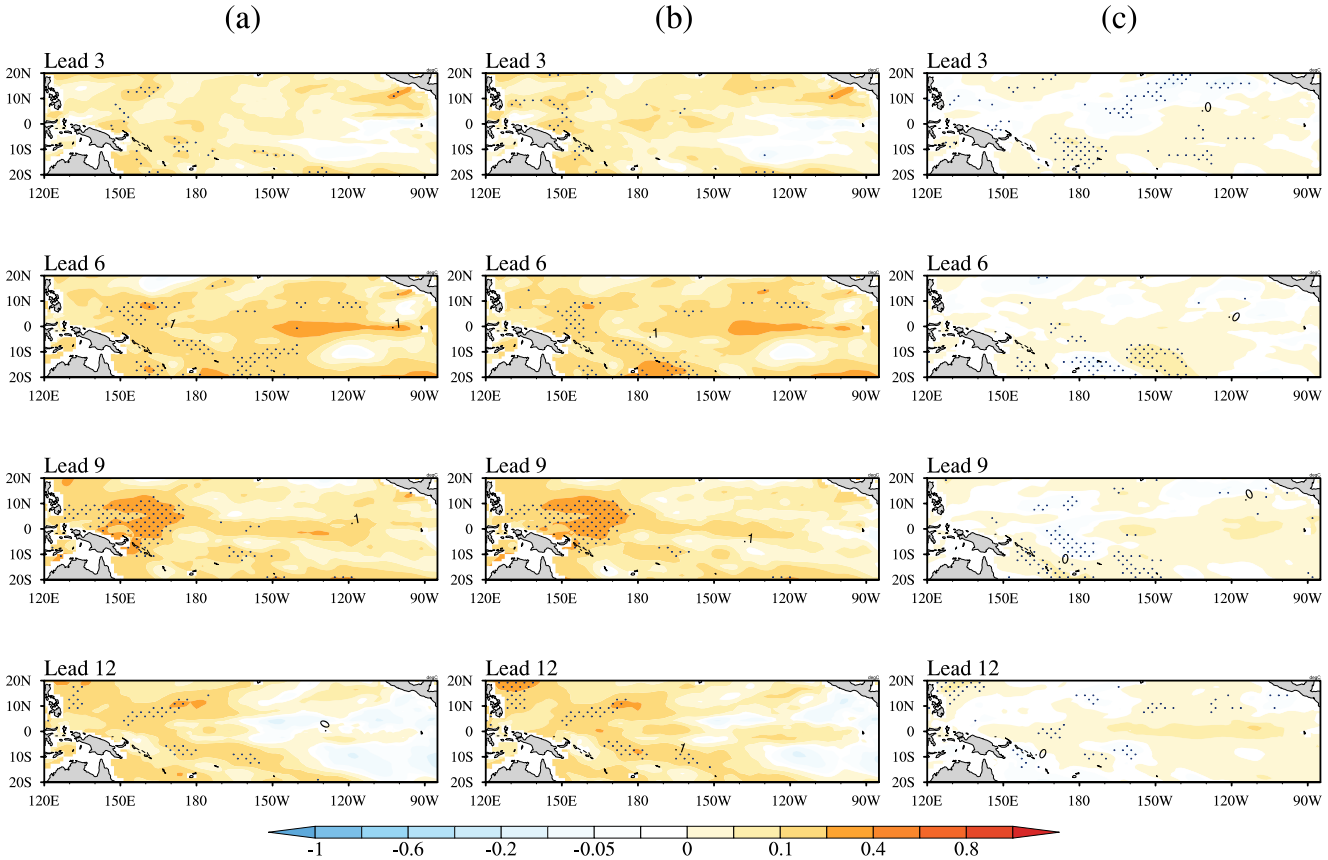

FIGURE 10 | As in Figure 5, but for the El Niño events from 1982 to 2015. (a) Ctrl versus CP-T. (b) Ctrl versus SV. (c) SV versus CP-T. [Colour figure can be viewed at wileyonlinelibrary.com]

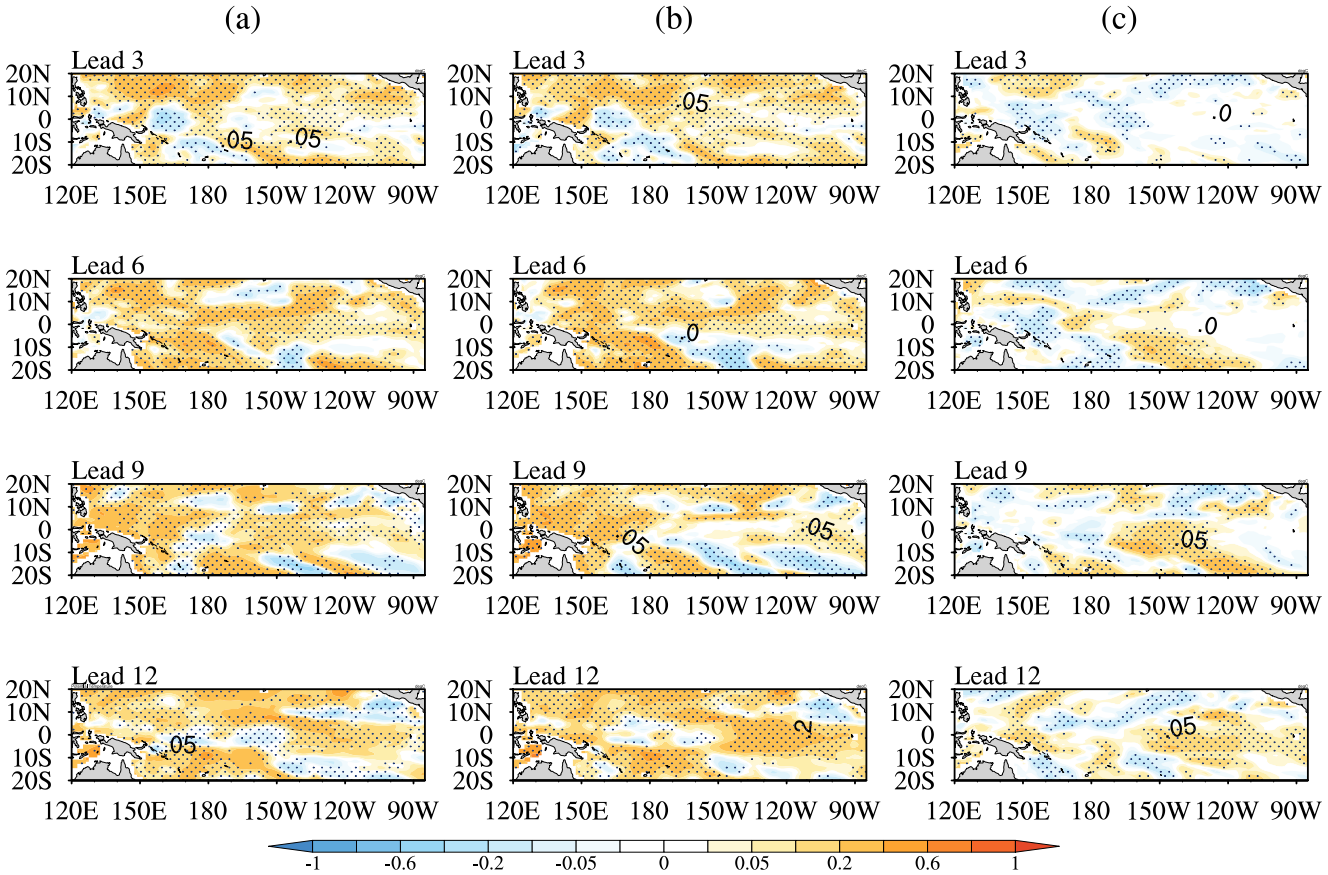

FIGURE 11 | As in Figure 6, but for the El Niño events from 1982 to 2015. (a) CP-T versus Ctrl. (b) SV versus Ctrl. (c) CP-T versus SV. [Colour figure can be viewed at wileyonlinelibrary.com]

for El Niño events relative to the overall forecast skill during 1982–2015. This enhanced performance is evident from the RMSE comparisons (Figures 5c vs. 10c) and ACC analyses (Figures 6c vs. 11c), underscoring the critical importance of capturing strong nonlinear effects in El Niño prediction.

This study systematically examines forecast skills from 1982 to 2015, with particular focus on the forecast skill for El Niño events through comprehensive analysis of both temporal and spatial variability of SSTA in the tropical Pacific. The results show that the CP-T ensemble mean forecast exhibits higher forecast skill than the SV ensemble mean forecast, particularly during El Niño events characterized by strong nonlinearity and at longer lead times. Consequently, the CP-T, which incorporates more nonlinear features, is better able to capture the nonlinear development of ENSO events compared to the SV method.

3.2.3 | The Pattern of the Mature-Phase SSTA for El Niño Events

The previous section primarily examined forecast errors and correlation coefficients regarding the CP-T and SV ensemble mean forecast skills. However, for El Niño events, accurately predicting the spatial patterns during the mature phase is also crucial. In this section, we will compare mature-phase El Niño SSTAs between the CP-T and SV ensemble mean forecasts, using January as an example.

Figure 12 shows the spatial distribution of SSTA during the mature phase of El Niño events for observations, the control forecast and the CP-T and SV ensemble mean forecasts at different lead times. The results indicate that both the control forecast and ensemble forecasts effectively capture the spatial distribution and intensity of the positive SSTAs during the peak El Niño; however, the intensity varies among different forecasts. To better highlight the differences between various ensemble forecasts, we conduct difference processing (see Figure 13). It can be revealed that the SSTA in the CP-T and SV ensemble mean forecasts is generally warmer than the observed SSTA in the central-eastern tropical Pacific. Additionally, the difference between the CP-T and SV forecasts shows a cold bias, which becomes more pronounced with increasing lead times. This suggests that the CP-T ensemble mean forecast predict the mature phase with a more close intensity to the observations and effectively improves the forecasting skill for SSTA intensity compared to the SV ensemble mean forecast, with the improvement becoming more obvious at longer lead times. This further indicates that incorporating nonlinear effects in the CP-T is more advantageous for longer lead times.

3.3 | Reliability of the Ensemble Forecasting

As emphasized by Buizza et al. (2005), the ensemble spread and RMSE are nearly equal in magnitude in a reliable ensemble forecasting system. The closer the ratio of the RMSE to ensemble

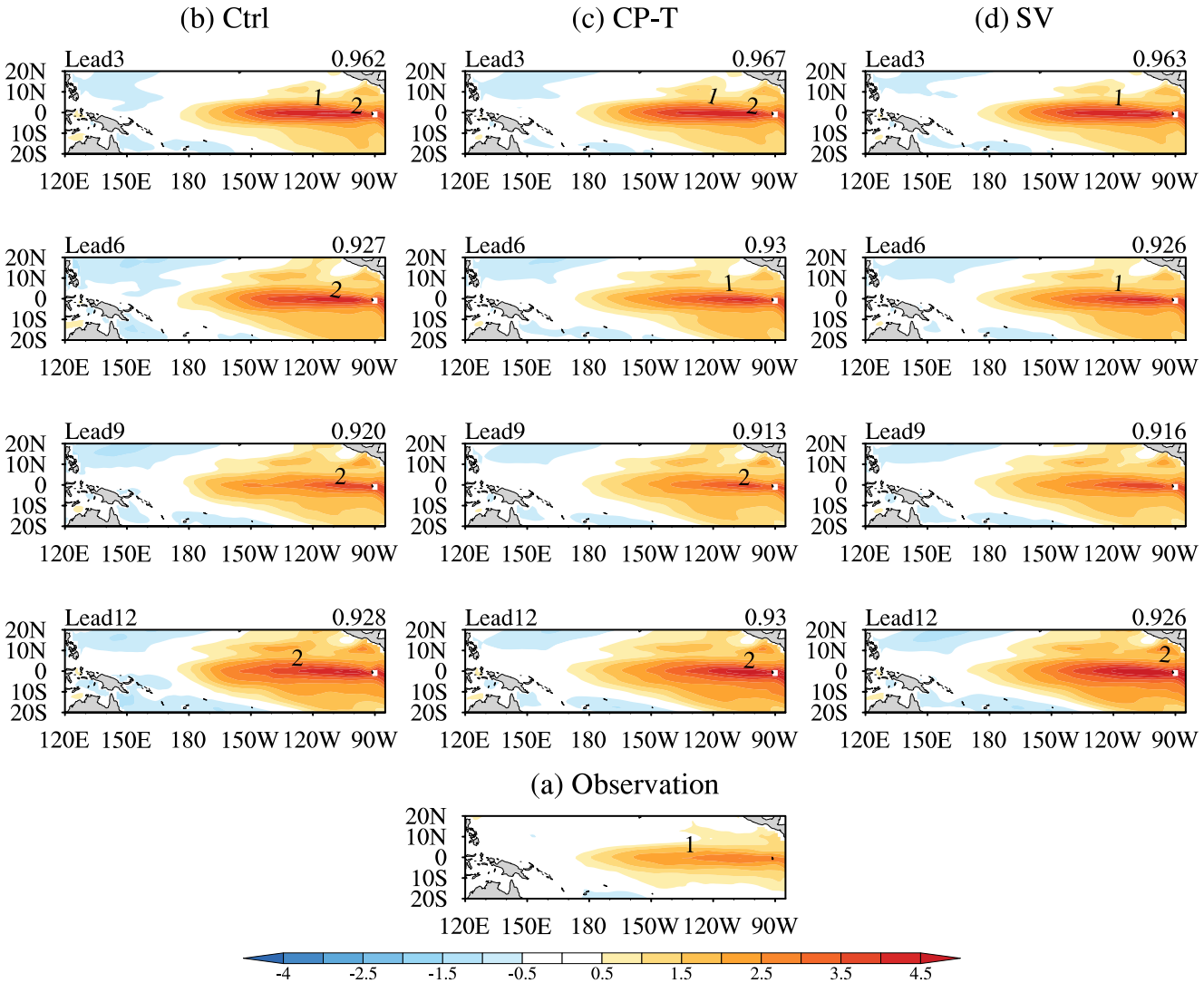

FIGURE 12 | Composite of the SSTA for mature phase of the El Niño events derived from the (a) observation, predictions by (b) the control forecast, (c) those by the CP-T and (d) SV ensemble mean forecasts. From top to bottom, the predictions are generated at 3-, 6-, 9- and 12-month lead times (units: °C). [Colour figure can be viewed at wileyonlinelibrary.com]

spread is to 1, the more reliable the ensemble forecasting system's estimation of the prediction uncertainties.

Figure 14 depicts the evolution of the 'reliability index' (RI; defined as RMSE/SPREAD) for CP-T and SV ensemble forecasts with respect to the Niño3.4 index. Although the CP-T ensemble forecast does not achieve an RI of exactly 1, its values are closer to 1 than those of the SV ensemble forecast at most lead times. This suggests that the CP-T system exhibits superior reliability, with its ensemble spread more accurately quantifying the prediction uncertainties of the ensemble mean Niño3.4 SSTA predictions. Figure 15 illustrates the spatial patterns of RI for both ensemble forecasts in the tropical Pacific. The RI for both ensemble forecasts are all approximately 1 in the tropical central-western Pacific; however, the RI significantly exceeds 1, and rises sharply to around 2 in the tropical southeast Pacific. Even so, the CP-T ensemble forecast maintains RI values closer to 1 compared to the SV ensemble forecast, indicating that it has higher reliability and provides a more reliable estimate of prediction uncertainties, particularly in the tropical central-western Pacific. In this study, the results suggest that initial perturbations incorporating nonlinear effects can enhance the reliability of ensemble forecasting. Duan et al. (2024) proposed the C-CNOP approach, which not only incorporates nonlinear effects but also takes into account the initial coupling uncertainties of ocean–atmosphere variability, demonstrating a more reliable consistency relationship compared to the CP-T. This indicates that considering the effects of initial coupling uncertainties is helpful for ensemble forecasts to obtain a more reliable estimate of forecast uncertainties on the basis of the CP-T method.

4 | The Physical Mechanisms Analysis

Based on the preceding analysis, the CP-T ensemble mean forecast exhibits higher skill than the SV ensemble mean forecast. This advantage is particularly pronounced during El Niño events and for forecasts at longer lead times. Additionally, the CP-T ensemble forecast shows higher reliability, providing more dependable estimates of forecast uncertainties. In the following

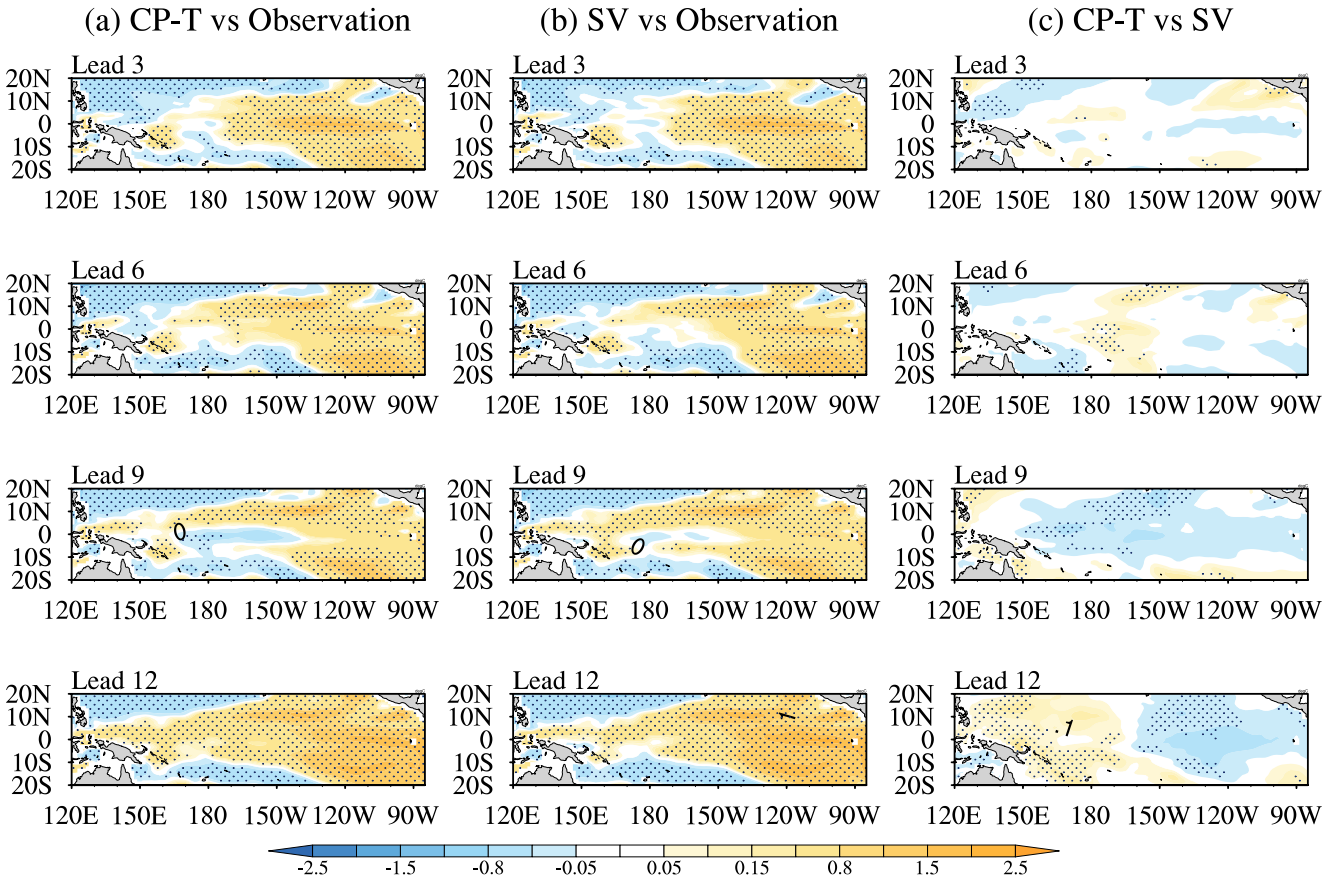

FIGURE 13 | The difference of SSTA for mature phase of the El Niño events derived from the between the (a) CP-T and observations, (b) SV and observations and (c) CP-T and SV. From top to bottom, the predictions are generated at 3-, 6-, 9- and 12-month lead times (units: °C). The dotted area denotes regions where differences are statistically significant at the 99% confidence level using a *t*-test. [Colour figure can be viewed at wileyonlinelibrary.com]

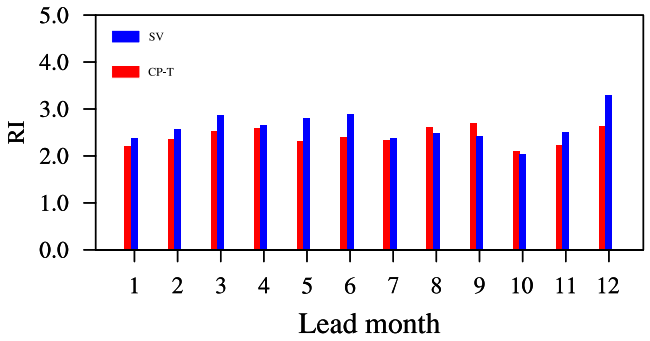

FIGURE 14 | The temporal variability of the RI for CP-T and SV. [Colour figure can be viewed at wileyonlinelibrary.com]

sections, we explore the underlying reasons why the CP-T has higher forecasting skills than the SV from two perspectives.

4.1 | The Relationship Between Ensemble-Mean Perturbation Evolution and Analysis Error Growth

Previous studies have demonstrated that the quality of ensemble forecasts is largely affected by the methods used to generate initial perturbation samples. By superimposing initial perturbation samples, which incorporate uncertainty information from the initial analysis field, onto control forecasts, the resulting perturbation samples can more effectively encompass the true development path, and thus obtain a high forecast skill. In other words, it is particularly important whether this set of perturbation samples can capture the evolution characteristics of the analysis errors.

In this section, by investigating the relationship between the ensemble-mean perturbation evolution and analysis error growth, we identify which set of initial perturbations generated by the CP-T and SV methods is more capable of capturing the evolution characteristics of analysis errors.

Assuming the initial analysis field is $x(0)$, the control forecast at time t, $t \in [0, T]$, is obtained by integrating the nonlinear operator M with the initial analysis field:

$$x^{\text{cont}}(t) = M(x(0)). \tag{4}$$

Assuming there are N perturbation members at the initial time, denoted as $x_i(0)$, $i = 1, 2, 3, \ldots, N$, by integrating the nonlinear operator M, we obtain N forecast values at time t:

$$x_i^f(t) = M\left(x_i(0)\right)(i = 1, 2, \ldots, N). \tag{5}$$

The ensemble mean forecast field can be derived as follows:

$$\bar{x}(t) = \frac{1}{N}\sum_{i=1}^{N} x_i^f(t). \tag{6}$$

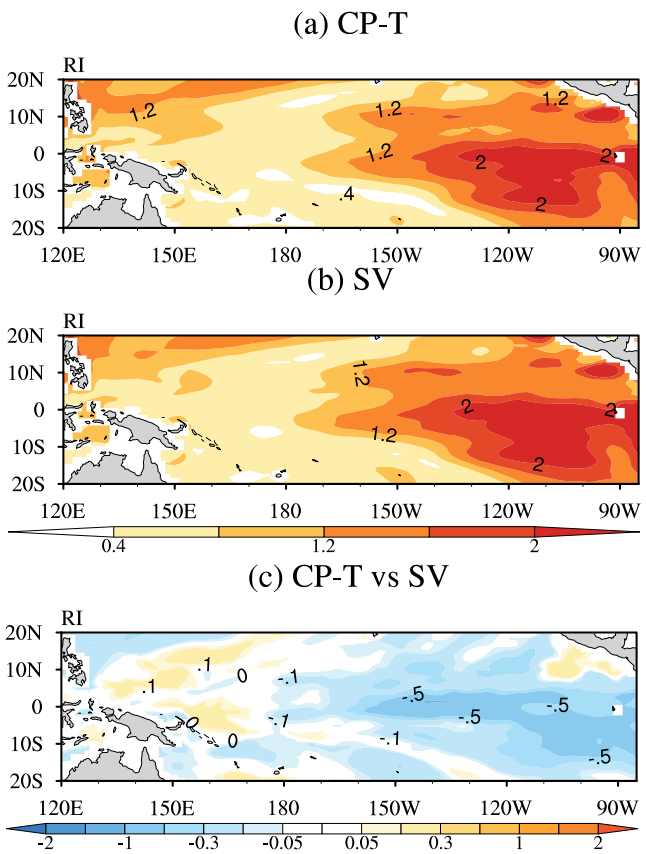

FIGURE 15 | The spatial distribution of RI for CP-T and SV ensemble forecast, and the RI differences between CP-T and SV. (a) CP-T. (b) SV. (c) CP-T versus SV. [Colour figure can be viewed at wileyonlinelibrary.com]

Assuming the true value of the state variable at time t is $x^{\text{true}}(t)$ then the nonlinear development of the analysis error $u^{\text{cont}}(t)$ can be described as

$$u^{\text{cont}}(t) = x^{\text{cont}}(t) - x^{\text{true}}(t). \tag{7}$$

Similarly, the nonlinear development of ensemble-mean perturbations $\bar{u}(t)$ can be written as follows:

$$\bar{u}(t) = x^{\text{cont}}(t) - \bar{x}(t). \tag{8}$$

Then, Equations (7) and (8) are expressed as

$$u^{\text{cont}}(t) - \bar{u}(t) = \bar{x}(t) - x^{\text{true}}(t). \tag{9}$$

Equation (9) implies that the difference between the analysis error growth and the evolution of the ensemble-mean perturbations is exactly equal to the difference between the ensemble mean forecast field and the true field.

Therefore, for $u^{\text{cont}}(t)$ and $\bar{u}(t)$, when the correlation coefficient between $u^{\text{cont}}(t)$ and $\bar{u}(t)$ is higher and the difference is smaller, it indicates that the ensemble mean perturbation can better capture the nonlinear development of initial analysis errors, making the ensemble mean forecast field closer to the true field.

Figure 16 shows the temporal evolution of the correlation coefficient between ensemble-mean perturbations (CP-T vs. SV) and analysis errors with lead times, along with cumulative absolute SSTA differences in the tropical Pacific Ocean. The results indicate that there is little difference in the correlation coefficients between the CP-T and SV methods during the first 4 months. However, the CP-T method exhibits higher correlation coefficients in subsequent lead months, which is more pronounced at longer lead times. Meanwhile, CP-T presents smaller cumulative absolute SSTA differences compared to SV. Thus, the CP-T ensemble-mean perturbations better capture the nonlinear development of analysis errors compared to the SV ensemble-mean

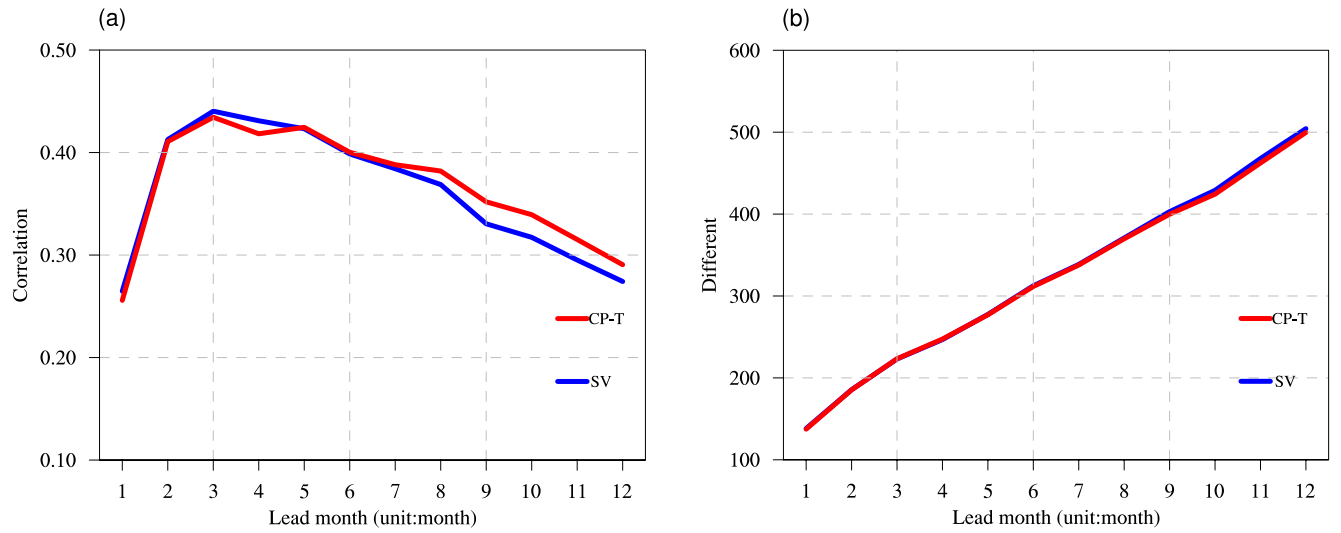

FIGURE 16 | Temporal evolution of (a) correlation between analysis errors and ensemble-mean perturbations for CP-T and SV methods, and (b) cumulative absolute SSTA differences between analysis errors and ensemble-mean perturbations in the tropical Pacific Ocean (units: °C). [Colour figure can be viewed at wileyonlinelibrary.com]

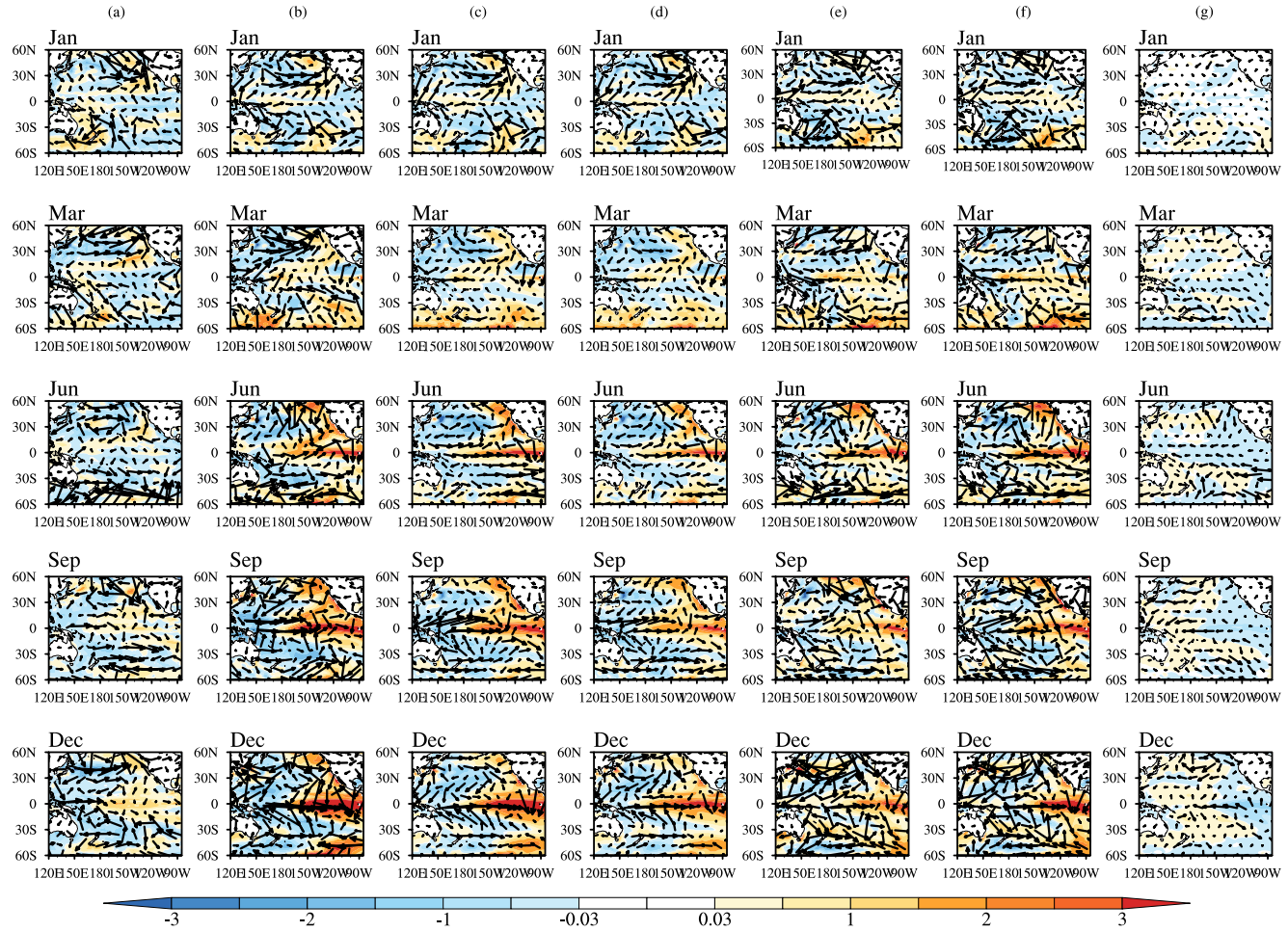

FIGURE 17 | The temporal evolution of the wind anomalies (units: $m\,s^{-1}$) and SSTAs (units: °C) for 1986 El Niño event derived from the (a) observations, (b) the control forecast, (c) SV, (d) CP-T, the different between (e) CP-T and observations, (f) SV and observations and (g) CP-T and SV. [Colour figure can be viewed at wileyonlinelibrary.com]

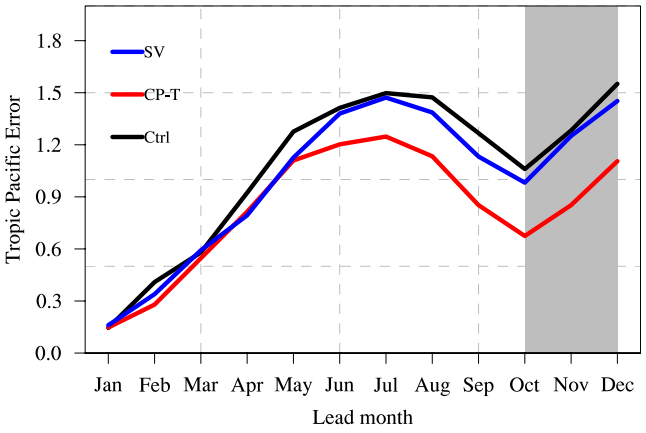

FIGURE 18 | The temporal evolution of the forecast errors in tropical Pacific SSTA (units: °C). [Colour figure can be viewed at wileyonlinelibrary.com]

perturbations. In other words, the ensemble mean forecast field produced by the CP-T method is closer to the true field than that generated by the SV method. In fact, the SV initial perturbations are derived within a linear framework, whereas the CP-T method incorporates nonlinear effects. Consequently, the CP-T ensemble-mean perturbations can better capture the nonlinear development of analysis errors, leading to higher forecast skill compared to the SV ensemble forecast.

4.2 | Dynamic Physical Process Analysis

In the following section, we will explore which physical processes in the CP-T ensemble mean forecast contribute to its superior forecast skill compared to the SV ensemble mean forecast. We focus on forecasts initiated in January, using the 1986 El Niño event as a representative case, and other events yield similar conclusions.

Figure 17 shows the evolution of SSTAs and wind anomalies initiated in January for the 1986 El Niño event. Both the control forecast and the ensemble mean forecasts successfully predict the occurrence of the El Niño event. However, compared to the observed SSTAs, all these forecasts exhibit a warm bias in the central-eastern tropical Pacific, especially the control forecast. Notably, the difference between the CP-T and SV ensemble mean forecasts in these regions is negative (as shown in Figure 17g), indicating that the CP-T ensemble mean forecast effectively improves the intensity prediction of El Niño events. Quantitatively,

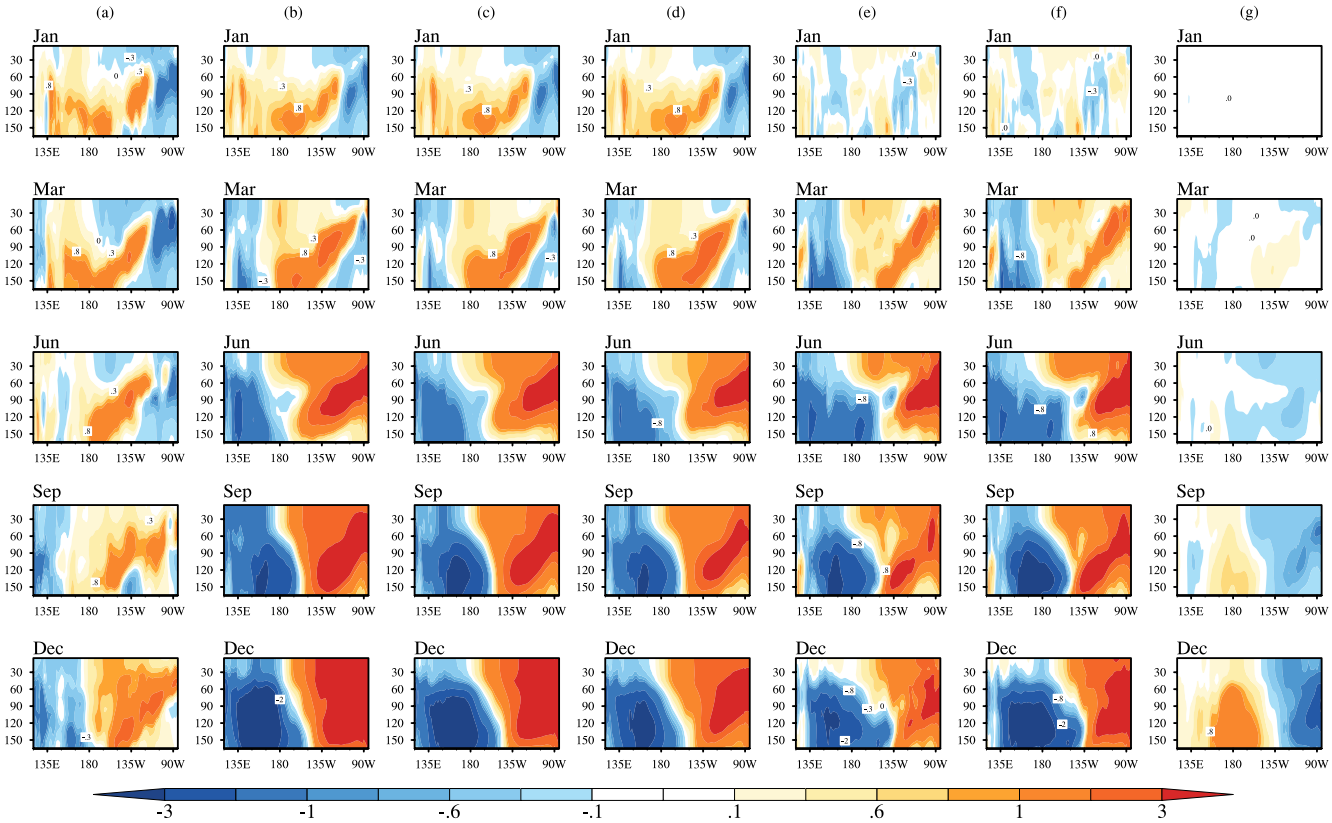

FIGURE 19 | The temporal evolution of the equatorial (5° S–5° N) subsurface temperature anomalies averaged by the meridian 5° S–5° N for 1986 El Niño event derived from the (a) observations, (b) the control forecast, (c) SV, (d) CP-T, the different between (e) CP-T and Observations, (f) SV and observations and (g) CP-T and SV (units: °C). [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 18 depicts the temporal evolution of forecast errors for tropical Pacific SSTAs. The CP-T ensemble mean forecast errors are significantly lower than the SV ensemble mean forecast errors. Furthermore, spatial correlation coefficients of SSTAs between forecasts and observations in the tropical Pacific reveal higher values for the CP-T method (0.80) compared to both the control forecast (0.72) and SV method (0.73). These results demonstrate that the CP-T ensemble mean forecast outperforms the control forecast and the SV ensemble mean forecast in terms of both prediction errors and spatial correlation coefficients. Similarly, as illustrated in Figure 19, the CP-T ensemble mean forecast significantly reduces the subsurface sea temperature bias present in the SV ensemble mean forecast when compared to observations, which aligns with the SSTA results.

So, what physical processes contribute to the higher skill of the CP-T ensemble mean forecast? As shown in Figure 17, the SV ensemble mean forecast displays warmer SSTAs relative to observations in the tropical central-eastern Pacific in the initial stage. At the same time, westerly wind anomalies in the western Pacific (relative to observations) force downwelling Kelvin waves that propagate eastward, thereby strengthening the transport and accumulation of warm water in the eastern Pacific and deepening the thermocline in that region. Subsequently, the Bjerknes positive feedback promotes the development of warm anomalies in the equatorial central-eastern Pacific. However, the CP-T method induces easterly wind anomalies in the tropical Pacific relative to the SV method (see Figure 20). This results in a reduction in the transport of warm water to the eastern Pacific to some extent. Consequently, the warm anomalies in the CP-T ensemble mean forecast are weaker than those in the SV ensemble mean forecast, which are closer to observations and result in higher forecast skill. Furthermore, a Victoria mode (VM)-like appears in the North Pacific in observations, where positive SSTAs near the Gulf of Alaska can trigger local convective anomalies. This promotes the atmosphere to generate a baroclinic response, leading to changes in the sea surface wind field. The resulting wind anomalies will guide the positive SST anomalies to gradually propagate towards low latitudes through the wind-evaporation-SST (WES) feedback mechanism (Xie and Philander 1994), thereby promoting the development of warm SSTAs at the equator. The SV ensemble mean forecast exhibits a stronger VM-like mode than the observations, resulting in stronger warm anomalies at the equator. In contrast, the VM-like mode in the CP-T ensemble mean forecast is weaker compared to that in the CSV ensemble mean forecast. As a result, the CP-T method effectively inhibits the excessive development of equatorial warm SSTAs, making it closer to the observations. In summary, the CP-T ensemble mean forecast successfully weakens the overestimation of warm SSTAs in the SV ensemble mean forecast by adjusting the feedback between sea temperature and the wind field, ultimately achieving higher forecast skill.

5 | Summary and Discussion

In this paper, we conduct ensemble forecasting experiments using the initial perturbation approaches of CP-T and SV based

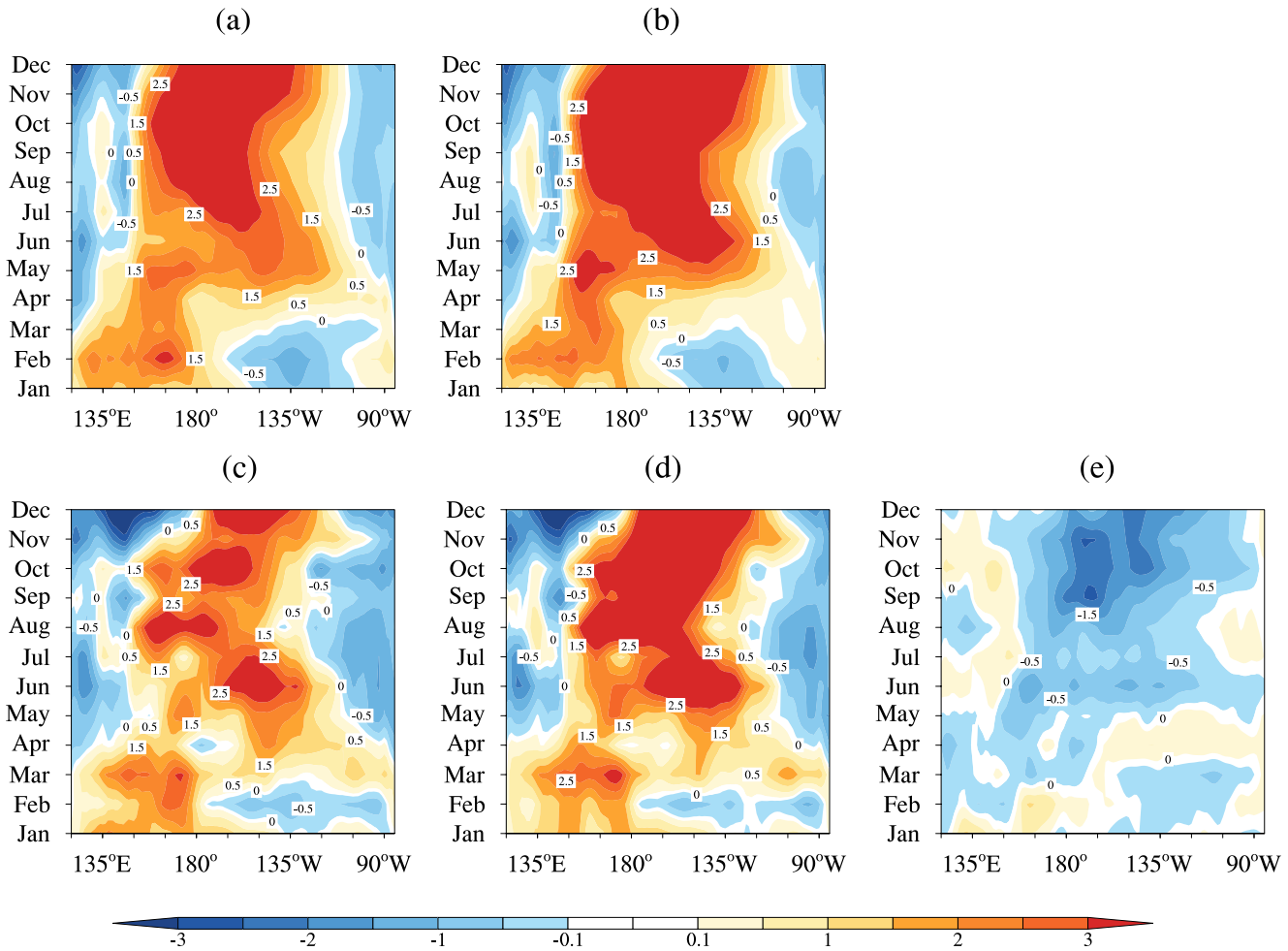

FIGURE 20 | The temporal evolution of the equatorial zonal wind anomalies averaged by the meridian (5° S–5° N) for El Niño event derived from the (a) CP-T, (b) SV, the different between (c) CP-T and observations, (d) SV and observations and (e) CP-T and SV (units: m s^{-1}). [Colour figure can be viewed at wileyonlinelibrary.com]

on the CESM model to investigate how incorporating nonlinear effects can improve the skill of ENSO ensemble forecasting.

The results indicate that, from 1982 to 2015, the CP-T ensemble mean forecast improves the forecast skill for Niño3.4 SSTAs compared to the SV ensemble mean forecast. Moreover, as lead times increase, the advantages of the CP-T ensemble mean forecast become more pronounced. This may be attributed to the fact that generally, errors tend to grow linearly at shorter lead times, while nonlinear effects become more significant at longer lead times. Consequently, the forecast skill of the CP-T ensemble mean forecast, which incorporates these nonlinear effects, becomes increasingly advantageous with longer lead times. Additionally, by comparing the forecast skill of SSTA spatial variability at different lead times, it is observed that the CP-T ensemble mean forecast effectively extends the skillful forecast lead times compared to the control forecast and the SV ensemble mean forecast.

On the other hand, recognizing that the El Niño events are characterized by strong nonlinearity, we conduct a further investigation into the forecast skills during El Niño events to explore the advantages of the CP-T incorporating the nonlinear effects. The results indicate that the CP-T ensemble mean forecast significantly improves the forecast skill of Niño3.4 SSTAs compared to the SV ensemble mean forecast. Moreover, the ACC of the CP-T ensemble mean forecast is higher than that of the SV ensemble mean forecast, especially at a lead time of 12 months, effectively extending the skillful forecast lead time. The advantages of the CP-T ensemble mean forecast over the SV ensemble mean forecast become more significant during El Niño events with strong nonlinearity and at longer lead times. Furthermore, regarding the prediction of the spatial structure of El Niño events, the CP-T ensemble mean forecast significantly improves the intensity forecast skill during the mature phase, making the forecast results closer to observations.

The study also conducted a reliability assessment of various forecasts. The results indicate that the CP-T ensemble forecast better captures the consistent relationship between ensemble mean forecast errors and ensemble spread compared to the SV ensemble forecast, from both temporal and spatial variability. This is especially evident in the tropical southeastern Pacific, where the CP-T ensemble forecast exhibits stronger reliability. Therefore, for ENSO prediction, the CP-T ensemble forecast that incorporates nonlinear effects provides better estimates of forecast uncertainties compared to the SV ensemble forecast.

This suggests that initial perturbations considering nonlinear features can further enhance the reliability of ensemble forecasting.

Based on the above analysis, we further examine the reasons why the CP-T ensemble mean forecast has higher forecast skill than the SV ensemble mean forecast. Our examination focuses on two primary aspects. First, we examine the relationship between the evolution of ensemble-mean perturbations and analysis error growth. The results indicate that the CP-T ensemble-mean perturbations can better capture the nonlinear development of analysis errors, resulting in higher forecast skill compared to the SV ensemble forecast. Second, through case analysis, the advantage of the CP-T method is examined from the perspective of physical mechanisms, highlighting its ability to properly describe the feedback processes between sea temperature and wind field, making the CP-T ensemble mean forecast closer to observations.

However, it is important to note that the conclusions drawn above are mainly based on ensemble mean forecast results, focusing on the role of the CP-T approach in improving the deterministic forecast skills. These promising results encourage us to further investigate the role of the CP-T approach in improving the probabilistic forecast skills by utilizing more ENSO events or extending the time series. Additionally, to quickly demonstrate the dynamical rationality of the CP-T approach and to save the computational burden of ensemble forecasts, we initially attempted a smaller number of perturbation samples in ensemble forecast experiments. Therefore, further investigation is needed to determine the optimal number of ensemble members required to achieve the highest forecasting skill using the CP-T approach in the coupled model.

Author Contributions

Lei Hu: investigation, formal analysis, writing – original draft, validation, writing – review and editing. **Wansuo Duan:** methodology, investigation, formal analysis. **Rong Feng:** methodology, formal analysis, writing – review and editing, supervision.

Acknowledgements

We thank the three anonymous reviewers for their insightful comments and suggestions. This work was supported by the National Natural Science Foundation of China (Grant Nos. 42330111, 42475060, 42505066 and 42450178), National Key Scientific and Technological Infrastructure project 'Earth System Numerical Simulation Facility' (EarthLab) and Fundamental Research Funds of the Institute of Atmospheric Physics, Chinese Academy of Sciences.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 42330111, 42475060, 42505066 and 42450178), National Key Scientific and Technological Infrastructure project 'Earth System Numerical Simulation Facility' (EarthLab) and Fundamental Research Funds of the Institute of Atmospheric Physics, Chinese Academy of Sciences.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The GODAS dataset is available at https://www.cpc.ncep.noaa.gov/products/GODAS/. The CESM model outputs are available from the corresponding author on reasonable request.

References

Barnston, A. G., M. K. Tippett, M. Ranganathan, and M. L. L'Heureux. 2019. "Deterministic Skill of ENSO Predictions From the North American Multimodel Ensemble." *Climate Dynamics* 53: 7215–7234. https://doi.org/10.1007/s00382-017-3603-3.

Buizza, R., P. L. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei. 2005. "A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems." *Monthly Weather Review* 133: 1076–1097.

Du, J., J. Berner, R. Buizza, et al. 2018. "Ensemble Methods for Meteorological Predictions." In *Handbook of Hydrometeorological Ensemble Forecasting*, edited by Q. Duan, F. Pappenberger, J. Thielen, et al., 1–52. Springer.

Duan, W. S., L. Hu, and R. Feng. 2024. "Coupled Conditional Nonlinear Optimal Perturbations and Their Application to ENSO Ensemble Forecasts." *Science China Earth Sciences* 67, no. 3: 826–842.

Duan, W. S., and Z. H. Huo. 2016. "An Approach to Generating Mutually Independentinitial Perturbations for Ensemble Forecasts: Orthogonal Conditional Nonlinear Optimal Perturbations." *Journal of the Atmospheric Sciences* 73: 997–1014.

Fan, Y., M. R. Allen, D. L. T. Anderson, and M. A. Balmaseda. 2000. "How Predictability Depends on the Nature of Uncertainty in Initial Conditions in a Coupled Model of ENSO." *Journal of Climate* 13: 3298–3313.

Huo, Z. H., and W. S. Duan. 2018. "The Application of the Orthogonal Conditional Nonlinear Optimal Perturbations Method to Typhoon Track Ensemble Forecasts." *Science China Earth Sciences* 62: 376–388.

Kleeman, R., Y. M. Tang, and A. M. Moore. 2003. "The Calculation of Climatically Relevant Singular Vectors in the Presence of Weather Noise as Applied to the ENSO Problem." *Journal of the Atmospheric Sciences* 60: 2856–2868.

Luo, J. J., C. X. Yuan, W. Sasaki, S. K. Behera, and S. Masson. 2016. "Current Status of Intraseasonal–Seasonal-to-Interannual Prediction of the Indo-Pacific Climate." In *Indo-Pacific Climate Variability and Predictability*, edited by S. K. Behera and T. Yamagata, 63–107. World Scientific.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis. 1996. "The ECMWF Ensemble Prediction System: Methodology and Validation." *Quarterly Journal of the Royal Meteorological Society* 122: 73–119.

Mu, M., W. S. Duan, and B. Wang. 2003. "Conditional Nonlinear Optimal Perturbation and Its Applications." *Nonlinear Processes in Geophysics* 10: 493–501.

Mureau, R., F. Molteni, and T. N. Palmer. 1993. "Ensemble Prediction Using Dynamically Conditioned Perturbations." *Quarterly Journal of the Royal Meteorological Society* 119: 299–323.

Smith, R., P. Jones, B. Briegleb, et al. 2010. "The Parallel Ocean Program (POP) Reference Manual Ocean Component of the Community Climate System Model (CCSM) and Community Earth System Model (CESM)." LAUR-10-01853, 141, 1–140.

Tang, Y. M., R. Kleeman, and S. Miller. 2006. "ENSO Predictability of a Fully Coupled GCM Model Using Singular Vector Analysis." *Journal of Climate* 19: 3361–3377.

Tang, Y. M., R. H. Zhang, T. Liu, et al. 2018. "Progress in ENSO Prediction and Predictability Study." *National Science Review* 5: 826–839.

Toth, Z., and E. Kalnay. 1993. "Ensemble Forecasting at NMC: The Generation of Perturbations." *Bulletin of the American Meteorological Society* 74: 2317–2330.

Xie, S. P., and S. G. H. Philander. 1994. "A Coupled Ocean-Atmosphere Model of Relevance to the ITCZ in the Eastern Pacific." *Tellus. Series A, Dynamic Meteorology and Oceanography* 46: 340–350.

Xue, Y., M. A. Cane, and S. E. Zebiak. 1997. "Predictability of a Coupled Model of ENSO Using Singular Vector Analysis. 1. Optimal Growth in Seasonal Background and ENSO Cycles." *Monthly Weather Review* 125: 2043–2056.

Yao, Z. X., Y. Tang, T. Lian, et al. 2019. "Roles of Atmospheric Physics and Model Resolution in the Simulation of Two Types of El Nino." *Ocean Modelling* 143: 101486. https://doi.org/10.1016/j.ocemod.2019.101468.

Yeager, S. G., N. Rosenbloom, A. A. Glanville, et al. 2022. "The Seasonal-To-Multiyear Large Ensemble (SMYLE) Prediction System Using the Community Earth System Model Version 2." *Geoscientific Model Development* 15, no. 16: 6451–6493. https://doi.org/10.5194/gmd-15-6451-2022.

Zhang, H., W. Duan, and Y. Zhang. 2023. "Using the Orthogonal Conditional Nonlinear Optimal Perturbations Approach to Address the Uncertainties of Tropical Cyclone Track Forecasts Generated by the WRF Model." *Weather and Forecasting* 38: 1907–1933.