npj | Climate and Atmospheric Science

www.nature.com/npjclimatsci

ARTICLE OPEN

Check for updates

A multiscale model for El Niño complexity

Nan Chen[1], Xianghui Fang[2,3,4]✉ and Jin-Yi Yu[5]

El Niño-Southern Oscillation (ENSO) exhibits diverse characteristics in spatial pattern, peak intensity, and temporal evolution. Here we develop a three-region multiscale stochastic model to show that the observed ENSO complexity can be explained by combining intraseasonal, interannual, and decadal processes. The model starts with a deterministic three-region system for the interannual variabilities. Then two stochastic processes of the intraseasonal and decadal variation are incorporated. The model can reproduce not only the general properties of the observed ENSO events, but also the complexity in patterns (e.g., Central Pacific vs. Eastern Pacific events), intensity (e.g., 10–20 year reoccurrence of extreme El Niños), and temporal evolution (e.g., more multi-year La Niñas than multi-year El Niños). While conventional conceptual models were typically used to understand the dynamics behind the common properties of ENSO, this model offers a powerful tool to understand and predict ENSO complexity that challenges our understanding of the twenty-first century ENSO.

npj Climate and Atmospheric Science (2022)5:16 ; https://doi.org/10.1038/s41612-022-00241-x

INTRODUCTION

As one of the most striking interannual climate variations in the world, El Niño-Southern Oscillation (ENSO) manifests as a basin scale air-sea interaction phenomenon characterized by sea surface temperature (SST) anomalies in the equatorial central to eastern Pacific (EP). Although evolving in the equatorial Pacific region, ENSO can affect climate, ecosystems, and economies around the world through atmospheric pathways[1,2]. In the classical viewpoint, ENSO was often regarded as a phenomenon with cyclical attributes[3], in which the positive and negative phases are El Niño and La Niña, respectively. ENSO is known to show a significant diversity and irregularity[4,5]. The theoretical explanations of these are always grouped into two categories[5–7]. In the first category, ENSO is viewed as a self-sustained, unstable and naturally oscillatory mode of the coupled ocean-atmosphere system, in which the nonlinearity acts mainly to bound the growing eigenmode and create a finite amplitude of the ENSO cycle. In the other category, ENSO is regarded as a stable (damped) mode triggered by atmospheric random "noise" forcing. More recently, many studies have suggested that there are at least two types of ENSO[8–10]. Based on the features during their mature phase, they were named as the EP and the central Pacific (CP) types when the largest SST anomaly is located near the coast of the South America and the dateline region, respectively[9,10]. The shift of the main heating location has significant impacts on the air-sea coupling processes in the tropical Pacific, which is the way ENSO affecting the global climate, and brings serious challenges to ENSO predictions[11,12]. Thus, since the concept of CP El Niño emerged, understanding the differences in the patterns, strengths, evolution processes, physical mechanisms, and global influences between the two types of ENSO has attracted great attention.

By driving strong anomalous eastward surface currents and exciting downwelling equatorial Kelvin waves, the westerly wind bursts (WWBs), a fast atmospheric variability, play an important role in the development of El Niño events[13]. Some studies argued that ENSO is likely a result of the interplay between a self-sustaining cyclic oscillation dictated by deterministic processes and WWBs that are partially modulated by ENSO itself[14], in which the former provides a basic dynamical framework, and the latter induces the different flavors of ENSO[15]. Particularly, strong and congregated WWBs are crucial for producing extreme El Niños[14,16]. It has been shown that the stalled El Niño in the winter of 2014 and the "unexpected" extreme El Niño in 2015 could be attributed to the lack and occurrence of the interannual component of the WWBs and the related wind-SST coupling in the spring and summer of 2014 and 2015, respectively[17]. The important role of both WWBs and easterly wind bursts (EWBs) in inducing this delayed extreme El Niño was also highlighted[18,19]. Therefore, the stochastic nature of wind bursts can help to explain the irregularity of ENSO events[20–22].

The evolution of ENSO is also significantly modulated by physical processes operating on longer time scales via changing tropical Pacific background states. For example, CP El Niños are observed to occur more frequently after the twentieth century[23]. Some work[24] attributed this to the anthropogenic global warming. Others[25], however, suggested that the background-state changes observed in the tropical Pacific in the 2000s, i.e., a La Niña-like pattern with enhanced trade winds and a more tilted thermocline, are opposite from those expected to produce more frequent CP El Niño events. Based on this, it is argued that such a La Niña-like, i.e., a strengthening Walker circulation, background state in the Pacific may favor the generation of CP El Niño by suppressing convection and low-level convergence in the CP, which could shift the anomalous convection westward[26].

It should be noted that although the general circulation models (GCMs) are expected as the most ideal tool to investigate the ENSO complexity, it is still a great challenge for them to successfully simulate these ENSO characteristics. Besides, because GCMs include many factors that can influence ENSO, it is not always easy to uncover the physical processes behind model simulations. On the contrary, constructing a stochastic multiscale conceptual model that can depict the main features of all the interannual skeleton of ENSO, the intraseasonal wind bursts and

[1]Department of Mathematics, University of Wisconsin-Madison, Madison, WI, USA. [2]Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, 220 Handan Rd., Shanghai 200433, China. [3]Innovation Center of Ocean and Atmosphere System, Zhuhai Fudan Innovation Research Institute, Zhuhai 518057, China. [4]CMA-FDU Joint Laboratory of Marine Meteorology, Shanghai 200438, China. [5]Department of Earth System Science, University of California, Irvine, CA, USA.
✉email: fangxh@fudan.edu.cn

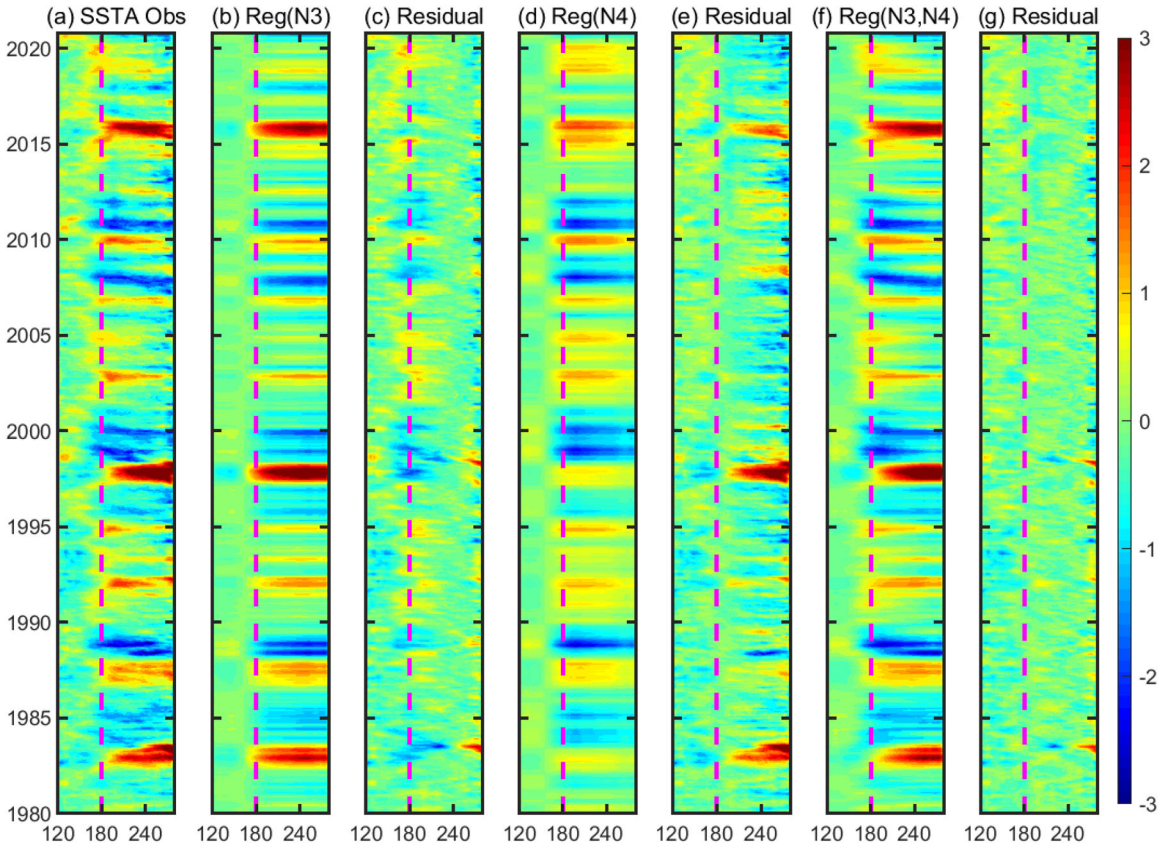

Fig. 1 Comparison between the linear regression models on reconstructing the equatorial SSTa. a The original spatiotemporal evolutions of the SST anomaly field. **b** The reconstructed SST anomaly field using the univariate linear regression based on T_E. **c** The residual between the SST fields in **a** and **b**. **d** The reconstructed SST anomaly field using the univariate linear regression based on T_C. **e** The residual between the SST fields in **a** and **d**. **f** The reconstructed SST anomaly field using the bivariate linear regression based on T_E and T_C. **g** The residual between the SST fields in **a** and **f**.

the background Walker circulation simultaneously may be a promising way to understand the causes of ENSO complexity, which is the motivation of this work.

On the subject of the interannual skeleton of ENSO, we need a model that can depict both the CP and EP SST anomalies, which are indispensable to simulate the ENSO complexity. To illustrate this point, Fig. 1a shows the observational SST anomalies in the equatorial Pacific (averaged over 5°S–5°N), along which are the regressed one with only one Niño index (Fig. 1b, d) and that with both the Niño3 (i.e., corresponding to EP in the model) and the Niño4 (i.e., corresponding to CP in the model) indices (Fig. 1f). These results clearly indicate that although the Niño3-based univariate linear regression model captures many characteristics of the ENSO variation in the EP region, it fails to realistically depict those in the CP region. In fact, the CP events in 1991, 1995, 2003–2005 and 2020 are completely missed (see Fig. 1b and the error plot in Fig. 1c). Also, although the Niño4 related part can capture most variations over the CP region, there are large residual parts in the EP, especially for the strong ENSO events, e.g., the 1982–1983, 1997–1998, 2015–2016 extreme El Ninos. In addition, due to the use of a solo SST variable in the regression, the eastward and westward propagating features (characterized by the underlying equatorial Kelvin and Rossby waves) are lost in the reconstructed ENSO spatiotemporal field. As a result, the reconstructed ENSO field contains only the standing oscillations from the CP to the EP. In contrast, the bivariate linear regression model significantly overcomes these shortcomings (Fig. 1f). It succeeds in reproducing almost the same SST variation as is observed in nature in both the CP and EP regions (see the error plot in Fig. 1g). The bivariate linear regression also facilitates the recovery of the large-scale behavior of the wave propagations across the equatorial Pacific, which allows the reconstructed SST field to highly resemble the observations.

Note that the physical mechanisms of the CP and EP El Niño are quite different[10,27]. Specifically, due to the fact that the anomalous warming center of EP type of ENSO is located in the EP, the mean thermocline is shallow and permits the perturbations on the subsurface to effectively influence the SST through upwelling processes. On the other hand, for the CP type of ENSO, the major warming center is concentrated in the CP, where the zonal mean SST gradient is strongest due to the warm pool to the west and the cold tongue to the east, the anomalous zonal current-related zonal advective feedback thus plays a very important role. Physical models of different degrees of complexity have also confirmed the important role of zonal advective feedback in causing the ENSO complexity[28–30]. Based on the above evidences, it is clear that including two degrees of freedom of the SST variation in a model, accounting for the CP and EP SST anomalies respectively, is essential to depict the large-scale features of the ENSO complexity. By adding the state-dependent noise to the classical recharge oscillator model, some studies[31,32] have suggested its efficiency on inducing the positive skewness of the SST over the EP. More recently, ref. [33] constructed a nonlinear two-box recharge oscillator model to investigate the two types of ENSO. In their work, the SST variations over the CP and EP are both dominated by the thermocline feedback, just like the original recharge oscillator model, and a zonal wind stress that is related to the T_C and T_E, respectively. By further adding a state-dependent nonlinear noise forcing, they can well capture the observed ENSO diversities. However, the zonal advective feedback, which is

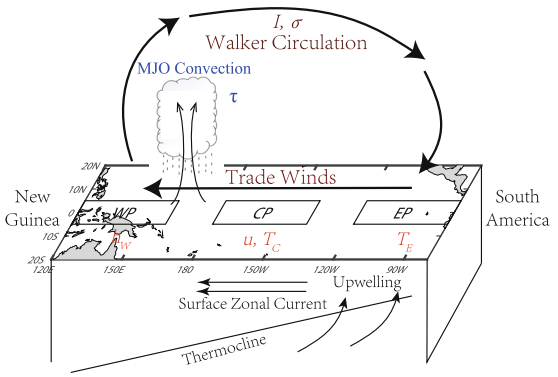

Fig. 2 Schematic diagram on the main components of the multiscale stochastic model. Specifically, it includes the interannual model that depicts the air-sea interactions over the entire western (h_W), central (u, T_C) and eastern Pacific (T_E), which is indispensable to simulate the ENSO complexity, the intraseasonal model that represents the random wind bursts and MJO (τ), and the decadal model that illustrates the variation in the background strength of the Pacific Walker circulation (I) and the related zonal advective feedback (σ).

suggested to be crucial for the SST variation over the CP, is not explicitly depicted in the model.

This article aims at developing a three-region multiscale stochastic conceptual model for the ENSO complexity to capture the general properties of the observed ENSO events as well as the complexity in patterns (e.g., CP vs. EP events), intensity (e.g., 10–20 year reoccurrence of extreme El Niños), and temporal evolution (e.g., more multi-year La Niñas than multi-year El Niños). It also aims at reproducing the observed non-Gaussian statistics in various Niño regions, e.g., the positively skewed fat-tailed probability density function (PDF) in the Niño3 region and the negatively skewed thin-tailed PDF in the Niño4 region, which facilitates the model to quantify the uncertainty and capture the extreme events in the ENSO dynamics.

RESULTS

The three-region multiscale stochastic model

In this work, a deterministic three-region conceptual model with the zonal advective feedback[34] is adopted as a starting model. It is a general extension of the classical recharge oscillator model[3] and depict the air-sea interactions over the entire western, CP and EP. Its main advantage is to efficiently describe the different SST variations in the CP and EP regions, which has been shown to be indispensable to simulate the ENSO complexity (Fig. 1). Then, a simple stochastic process describing the tropical atmospheric intraseasonal wind disturbances of the WWBs, the EWBs and the MJO, which involves a multiplicative noise that describes the modulation of the wind bursts by the interannual SST, is incorporated into the starting model. Such a stochastic parameterization of the intraseasonal variability plays a crucial role in explaining the irregularity of the ENSO events. Furthermore, a simple but effective stochastic process describing the multi-decadal variation of the background Walker circulation[35] is incorporated into the system to modulate the strength and the occurrence frequency of the EP and CP El Niños.

The three-region multiscale stochastic model is summarized as follows, where the details are described in the "Methods" section. The main components of this model and the multiscale interactions are also summarized in a schematic diagram as shown in Fig. 2. The model reads:

$$\frac{du}{dt} = -ru - \frac{a_1 b_0 \mu}{2}(T_C + T_E) + \beta_u \tau + \sigma_u \dot{W}_u, \tag{1}$$

$$\frac{dh_W}{dt} = -rh_W - \frac{a_2 b_0 \mu}{2}(T_C + T_E) + \beta_h \tau + \sigma_h \dot{W}_h, \tag{2}$$

$$\frac{dT_C}{dt} = \left(\frac{\gamma b_0 \mu}{2} - c_1(T_C)\right) T_C + \frac{\gamma b_0 \mu}{2} T_E + \gamma h_W + \sigma u + C_u + \beta_C \tau + \sigma_C \dot{W}_C, \tag{3}$$

$$\frac{dT_E}{dt} = \gamma h_W + \left(\frac{3\gamma b_0 \mu}{2} - c_2\right) T_E - \frac{\gamma b_0 \mu}{2} T_C + \beta_E \tau + \sigma_E \dot{W}_E \tag{4}$$

$$\frac{d\tau}{dt} = -d_\tau \tau + \sigma_\tau(T_C)\dot{W}_\tau, \tag{5}$$

$$\frac{dI}{dt} = -\lambda(I - m) + \sigma_I(I)\dot{W}_I. \tag{6}$$

where the interannual component (Eqs. (1)–(4)) depicts the deterministic dynamics for both the CP and EP types of ENSO, the intraseasonal component represents the random wind bursts (Eq. (5)), and the decadal component represents the variation in the background strength of the Pacific Walker circulation (Eq. (6)). In the model, T_C and T_E are the SST in the CP and EP while u is the ocean zonal current in the CP and h_W is the thermocline depth in the western Pacific (WP). The other two variables τ and I represent the intraseasonal random wind burst amplitude, including the MJO, and the background Walker circulation, respectively. The decadal variability I also stands for the zonal SST difference between the WP and CP regions that directly determines the strength of the zonal advective feedback. Besides, this is an anomaly model, which means that all the prognostic variables are the deviations from their corresponding monthly climatology during the analysis period (years 1980–2020).

Stochasticity plays a crucial role in coupling variables at different time scales and parameterizing the unresolved features in the model. First, the intraseasonal component τ is modeled by a simple stochastic differential Eq. (5) with a state-dependent (i.e., multiplicative) noise coefficient σ_τ, where $\dot{W}_\tau$ is a white noise source. The stochastic wind bursts are then coupled to the processes of the interannual variables serving as external forcings, which are the main mechanism for generating the EP events and the non-Gaussian features of T_E. In addition to the stochastic wind bursts, four Gaussian random noises $\sigma_u \dot{W}_u$, $\sigma_h \dot{W}_h$, $\sigma_C \dot{W}_C$ and $\sigma_E \dot{W}_E$ are further added to the processes describing the interannual variabilities. These random forcings effectively parameterize the additional contributions to the interannual variables that are not explicitly modeled here, such as the subtropical atmospheric forcing at the Pacific Ocean and the influences from the other Ocean basins[36]. In a more general sense, these stochastic noises can be regarded as the simplest way for stochastic parameterization, which increases the model variability such that the PDFs of the model variables can better match those of the observational data. Next, the background Walker circulation in the decadal time scale has been shown to modulate the interannual variability[14]. Since the details of the background Walker circulation consist of uncertainties and randomness, a simple but effective stochastic process is used to describe the time evolution of the decadal variability I, where $\dot{W}_I$ in (6) is a white noise source. The multiplicative noise in the process of I aims at guaranteeing the positivity of I due to the fact that the long-term average of the background Walker circulation is non-negative.

The coupled model involves a minimum nonlinearity, which nevertheless plays a crucial role in recovering the key dynamics and reproducing the non-Gaussian statistics for the CP events. The first nonlinearity is the σu term in (3), which represents that the

strength of the zonal advection is modulated by the decadal variability. Such a nonlinearity is crucial in simulating the right occurrence frequency of both the CP and EP El Niño events. Another key nonlinearity comes from the coefficient c_1 in (3), which is a quadratic function of T_C. In other words, the total damping in (3) is cubic. Such a cubic nonlinearity is justified by analyzing the observational data (see Fig. 8 and the detailed justifications in the "Methods" section). It also facilitates the recovery of the non-Gaussian statistics in the CP region, which has completely different characteristics as that in the EP region. Note that, since the coupled model is nonlinear, the long-term statistics does not necessarily have a zero mean. To guarantee the model to be an anomaly model, an extra C_u term is imposed in (3) such that all the variables have climatology with zero mean.

Finally, seasonal phase locking is a remarkable feature of the ENSO, which manifests in the tendency of ENSO events to peak during boreal winter[37,38]. Here, the effects of seasonality are added to both the wind activity and the collective damping. The former accounts for the active phase of the MJO in boreal winter[39], while the latter is due to the seasonal migration of the Intertropical Convergence Zone (ITCZ), which modulates the strength of the upwelling and horizontal advection processes to influence the evolution of the SST[40]. Thus, the three coefficients c_1, c_2 and σ_τ are all time-periodic functions.

The dimensional units and the parameters in the coupled model are summarized in Table 1. As is described in "Methods", all parameters are determined based on the observational data during 1980–2020.

Reproducing the observed ENSO statistics

Figure 3 compares several key statistics of the multiscale stochastic model with those from the observations. For a more direct comparison, the 2000-year long simulation is divided into 50 non-overlapping segments, each of which has a 40-year period as the observation. Then the average and its one standard deviation intervals are illustrated. Simulations with different random number seeds have been utilized to confirm the robustness of the statistics.

Figure 3a, b shows the power spectrums of the Niño3 and the Niño4 SST, respectively. It can be seen that the major signal of the T_C is between 2 and 3 years. The power decreases rapidly when the period is elongating. However, the signal of the T_E has a wider range, that is, the power gradually increases during 2–6 years. This means the period of the T_E is more likely to be larger than the T_C.

Table 1. Summary of the non-dimensional units and the model parameters.

$[h]$	150 m	$[T]$	7.5 °C
$[u]$	1.5 m/s	$[t]$	2 months
$[\tau]$	5 m/s	d_τ	2
γ	0.75	r	0.25
α_1	0.0625	α_2	0.125
b_0	2.5	μ	0.5
σ	$0.2I$	λ	0.1
$p(I)$	0.25 in $I \in (0, 4)$	$\sigma_I(I)$	Fig. 7e
β_E	0.15(2–0.2I)	β_u	$-0.2\beta_E$
β_h	$-0.4\beta_E$	β_C	$0.8\beta_E$
σ_u	0.04	σ_h	0.02
σ_C	0.04	σ_E	0
$\sigma_\tau(T_C, t)$	$0.9[\tan h(7.5T_C) + 1]\left[1 + 0.3\cos\left(\frac{2\pi}{6}t + \frac{2\pi}{6}\right)\right]$		
$c_1(T_C, t)$	$\left[25\left(T_C + \frac{0.75}{7.5}\right)^2 + 0.9\right]\left[1 + 0.3\sin\left(\frac{2\pi}{6}t - \frac{2\pi}{6}\right)\right]$		
$c_2(t)$	$1.4\left[1 + 0.2\sin\left(\frac{2\pi}{6}t + \frac{2\pi}{6}\right) + 0.15\sin\left(\frac{2\pi}{3}t + \frac{2\pi}{6}\right)\right]$		

This comparison is consistent with our physical understanding. Such a result indicates that the model is able to reproduce the observed irregular oscillations in both the Niño regions.

Next, Fig. 3c, d illustrates that the model perfectly recovers the remarkable non-Gaussian statistics of both the Niño3 and Niño4 SST. In particular, the observed Niño3 SST has a positive skewness and a one-sided fat tail that results from the occurrence of the extreme El Niño events. Due to the multiplicative noise in the wind burst process (5), the model is able to accurately reproduce such a highly non-Gaussian statistical feature. On the other hand, the skewness of the observed Niño4 SST is negative, and the kurtosis is 2.7, which is less than the standard Gaussian value 3, indicating the suppression of extreme El Niño events in the CP region. Thanks to the cubic and non-centered damping c_1 (Eq. (15) in the "Methods" section), the model succeeds in capturing such a skewed and light tailed distribution. Note that GCMs and the intermediate models often have great difficulties in reproducing these highly non-Gaussian PDFs, which are nevertheless one of the most important and necessary conditions for simulating the realistic ENSO complexity.

In addition to reproducing the climatology distribution functions, the model is also skillful in recovering the observed seasonal phase locking features of the EP and the CP ENSO. This can be seen in Fig. 3e, f, which show the monthly variance of the Niño3 and the Niño4 SST, respectively. Both types of ENSO onset in boreal spring, develop in summer, and peak in the following winter. The model also realistically reproduces the slight late onset (about 2 months) of the CP ENSO than the EP ENSO. The late onset reduces the growing season and is a key reason why the CP events are typically weaker (i.e., smaller variance) than the EP events[41].

Finally, Fig. 3g, h shows that the model is able to recover the observed variance of h_W and u as well. This indicates the skill of the model in quantifying the uncertainty of nature.

Reproducing the observed ENSO complexity

ENSO complexity appears in its spatial pattern, peak intensity, and temporal evolution. Table 2 compares the model results for different situations with the observation on the ENSO complexity. It also summarizes the results of the sensitivity experiments in the next subsection.

In term of the complexity in the ENSO pattern, this multiscale stochastic model produces 660 El Niños and 852 La Niñas during its 2000-year simulation based on a widely used ENSO classification method[42]. Among the El Niño events, about 60% (398 events) of them are EP events and 40% (262 events) of them are CP events. Note that during the observational period from 1950 to 2020, 14 of the 24 (i.e., 58%) major El Niños are of the EP type, and 10 of them (i.e., 42%) are of the CP type[43]. Such a comparison indicates that the model reproduces roughly the same ratio of the EP and CP events as the observations.

In term of the complexity in the ENSO intensity, there is a tendency for extreme El Niño events (e.g., the 1997–1998 and 2015–2016 ones) to occur every 10–20 years as in the observations[44]. A total of four extreme El Niño events have occurred since 1950, namely 1972–1973, 1982–1983, 1997–1998 and 2015–2016. Consistent with this reoccurrence frequency, the multiscale stochastic model produces 125 extreme El Niño events in its 2000 simulation (namely, on average every 16 years).

In term of the complexity in the ENSO evolution, it is noted that an El Niño (La Niña) event can be followed by a La Niña (El Niño) event to result in a cyclic ENSO evolution pattern or by another El Niño (La Niña) event to result in a multi-year ENSO evolution pattern[42]. Multi-year ENSO events are a major challenge for the accurate ENSO prediction[45]. During the historic period, multi-year La Niña events tend to occur more frequently than multi-year El Niño events[46]. However, the GCMs were often not able to

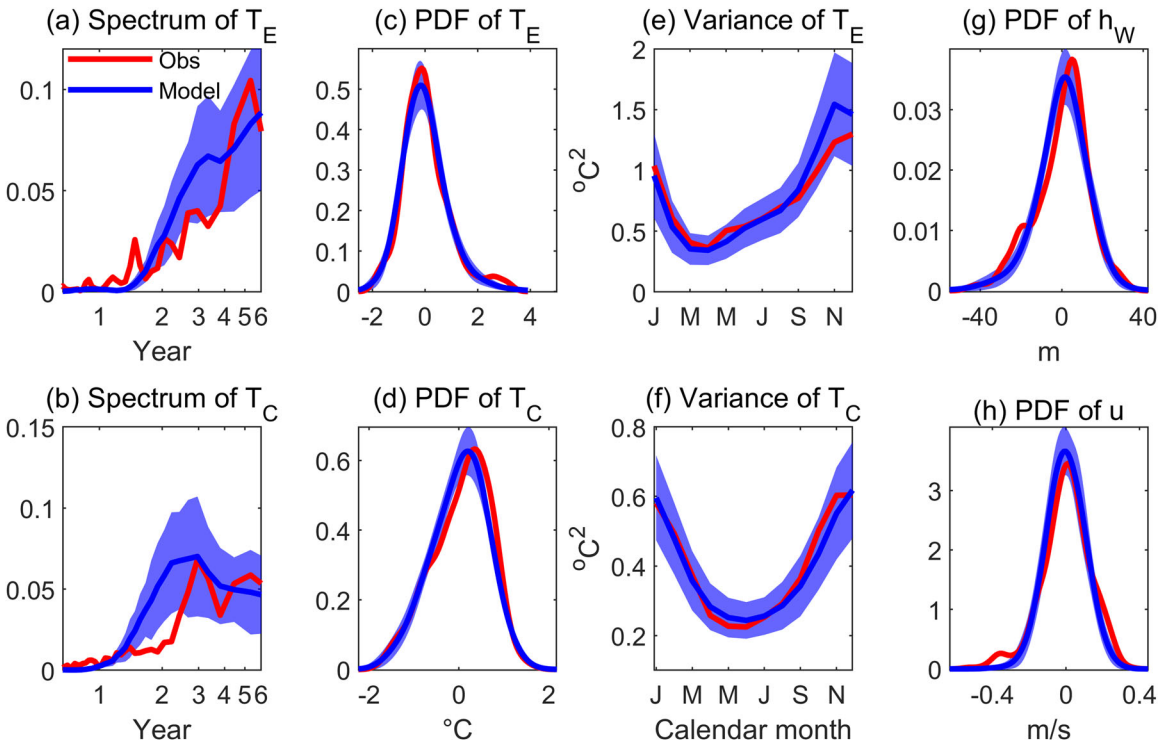

Fig. 3 Comparison of the statistics of the observations and the multiscale stochastic model. a, **b** Power spectrums of Niño3 and Niño4 SST. **c**, **d** PDFs of Niño3 and Niño4 SST. **e**, **f** The monthly variance (i.e., the seasonal cycle) of Niño3 and Niño4 SST. **g** PDF of the thermocline depth h_W in the western Pacific region. **h** PDF of the ocean zonal current u in the CP region. In each panel, red and blue curves are for the observation and model, respectively. For the model, the total 2000-year long simulation is divided into 50 non-overlapping segments, each of which has a 40-year period as the observation. Then the average (blue line) and its one standard deviation intervals (shading) are illustrated.

reproduce this asymmetric feature[46]. Such a deficiency in the operational models, together with the limited number of multi-year ENSO events in the observations, have hindered the effort to uncover the underlying dynamics of multi-year ENSO and its associate El Niño-La Niña asymmetry. In contrast, it is very encouraging to find that the multiscale stochastic model developed here is able to produce more multi-year La Niñas (209) events than multi-year El Niños (100) events in its 2000-year simulation. Thus, this model can be a useful tool to better understand the multi-year ENSO dynamics.

To use the model simulation for a better understanding of the physical processes behind the ENSO complexity, Fig. 4 compares key atmospheric and ocean variables during a particular 40-year segment of the model simulation with those observed during the past four decades (1980–2020). Both the observed and simulated T_E and T_C indices (Fig. 4a, c) clearly indicate that most of the extreme El Niño events are of the EP type. For these extreme events, the amplitude of the Niño4 SST (i.e., T_C) is significantly smaller than that of its counterpart Niño3 SST (i. e., T_E). This is consistent with the finding from the observations that the vertical thermocline process produces strong El Niño events (i.e., the EP El Niños), while the zonal advection process produces weak El Niño events (i.e., the CP El Niño). In the simulation, the time series of T_E and T_C are positively correlated with each other and the positive correlation is also found between u and h_W. The latter two, on the other hand, have negative correlations with the formers, which provide the delayed negative feedbacks according to the recharge oscillator theory. It is noticed that extreme El Niño events are preferable when the decadal variable I is close to zero (Fig. 4e). Specifically, under such circumstances, the model has a high chance to generate strong WWBs and therefore more moderate and extreme EP events are likely to be triggered. In contrast, when I becomes large, the warming center tends to occur in the CP region. Consistent with the observations[43], the slow variation of the background strength of the Pacific Walker circulation modulates the occurrence frequencies of the EP and CP ENSOs.

Another advantage of the model developed here is that it can be combined with the bivariate regression method to reconstruct the spatiotemporal evaluation of the SST field:

$$\mathrm{SST}\,(x,t) = r_C(x)T_C(t) + r_E(x)T_E(t) \tag{7}$$

which provides a clearer view of understanding the ENSO complexity dynamics from the model. Here x is the longitude and t is time. The regression coefficients $r_C(x)$ and $r_E(x)$ are determined using the observational data at each longitude grid point x. Then the Niño3 and Niño4 indices T_E and T_C from the model are plugging into the regression formula (7) to obtain the SST spatiotemporal patterns. Figure 5 shows the Hovmoller diagrams of the model simulation, including the 40-year period in Fig. 4, which clearly demonstrate the ENSO complexity.

First, the model is able to simulate realistic extreme El Niños (red color), mimicking the observations. Some examples of the extreme El Niño events are those in years 228, 251, 401 and 459, where the associated SST patterns and the profile of their precursors, i.e., the wind burst amplitudes and directions, are all similar to the observed 1982–1983 and 1997–1998 events.

Notably, the model is able to simulate the so-called delayed extreme El Niño as was observed in 2015–2016, for example, the two events during years 448–449 and 505–506 in Fig. 5. The reason for the model to generate this kind of El Niño[47] is its success in simulating the associated peculiar WWB-EWB-WWB structure[18]. Here, the first WWB tends to trigger a strong El Niño but the following EWB kills the event and postpones it until the next year when another series of the WWBs occur. Next, the model is able to simulate the traditional moderate EP El Niño (purple color; e.g., years 225, 238, 403 and 445), which are triggered by the moderate WWBs.

Table 2. Comparison of the model results for different situations with the observation on the ENSO complexity.

	El Niño	EP	CP	Extreme	Multi-year	La Niña	Multi-year
Observation (1950–2020)	24	14	10	4	5	24	8
Standard run	23.4 ± 2.9	14.1 ± 3.3	9.3 ± 2.5	4.5 ± 1.8	3.6 ± 1.5	30.3 ± 4.2	7.4 ± 2.4
$I \equiv 0$	22.9 ± 2.8	16.1 ± 3.1	6.8 ± 2.0	9.3 ± 2.1	2.8 ± 1.9	30.6 ± 3.3	7.1 ± 2.3
$I \equiv 4$	27.3 ± 4.2	15.7 ± 4.0	11.6 ± 3.4	1.2 ± 1.0	6.3 ± 2.2	27.8 ± 4.7	7.0 ± 2.4

Shown are the numbers of different ENSO events for the model and observation. For the model, the whole 2000 simulation years are divided into 28 segments (each one with the same length as the observation, i.e., 71 years) for calculation. Then the average numbers plus and minus the corresponding standard errors of the 28 segments are shown.

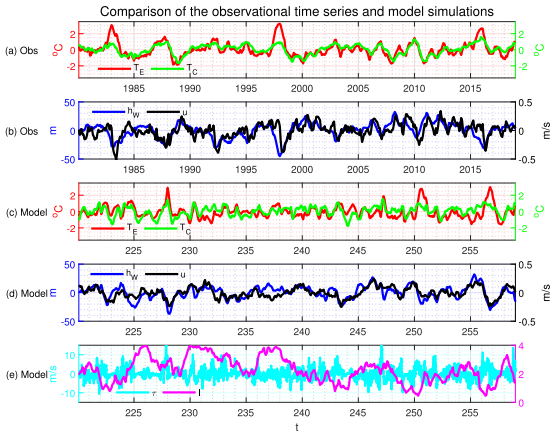

Fig. 4 Comparison of the observational time series and the model simulations. a, **b** The observational SST anomalies in the Niño3 (red) and Niño4 (green) regions, the observed thermocline depth anomaly in the western Pacific region (blue) and the observed ocean zonal current in the CP region (black). **c**, **d** Similar to **a**, **b**, but for the results from the model. **e** The time series of the intraseasonal wind bursts τ (cyan) and the decadal index I (purple) from the model.

In addition to the complexity of the EP events, one desirable feature of the model is that the simulated CP El Niño events (orange color) also highly resemble the observations. In particular, both a single CP event (e.g., years 222, 259 and 444) and a sequence of consecutive CP events (e.g., years 232–233 and 241–243) can be reproduced from the model. The latter mimics the observed CP episodes during 2003–2006.

In addition, the model can create some mixed CP-EP events (e.g., years 241, 247 and 457), which are similar to the observed ones in early 1990s.

Finally, the La Niña events from the model (blue color; e.g., years 223, 404 and 504) usually follow the El Niño ones. Some La Niña events have cold SST in the CP region while other La Niña years have cold centers locating around the EP area. The model is also able to generate multi-year La Niña events, i.e., a La Niña transitions to another La Niña, such as the one spans over year 509 and year 510.

To more clearly depict the model simulation of the El Niño complexity. Figure 6 shows the scatterplot of the peak winter (December–January–February)-mean SST anomalies along the equator, and the corresponding longitude where that peak occurs—for every El Niño event in the 2000-year standard run. Figure 6a, b exhibits two major centers, that is, one is near the dateline and the other is in the eastern region. It should be noted that this property is more consistent with the observation than the state-of-art CGCMs, which always exhibit a common bias for a farther west position[4]. Also, the events with CP SST anomaly peaks are relatively weak, while the strongest events always peak in the east Pacific, although east Pacific events can exhibit a wide range of amplitudes. These characters are physically reasonable since there is more potential for the warming amplitude in the EP due to its climatological SST being much less than the radiative-convective equilibrium temperature of about 30 °C. While the potential for the warming amplitude in the CP is relatively smaller[48].

Sensitivity analysis

The simple formula of this multiscale stochastic model and their key parameters also enable us to project the possible changes of ENSO complexity under various past and future climate regimes. Several sensitivity tests are utilized for a further understanding of the coupled multiscale stochastic model.

The first study is on the damping coefficient c_1 in (3), which reflects the collective residual part of the heat budget equation apart from the dynamical terms. Recall that a cubic damping is adopted in Table 1, i.e., c_1 is small for the small local SST and becomes larger for the large SST anomalies. Such a treatment avoids the unrealistic enlargement of the simulated SST anomalies[3]. If a linear damping is utilized, then the major change is the kurtosis of T_C, which will become larger than 3 and lead to fat tails of the Niño4 PDF. Despite that the overall model simulation remains similar, there are occasionally certain CP events that have large amplitudes, which is not the case in the observations. This indicates the necessity of adopting a cubic damping in c_1 to capture the observed non-Gaussian features in the CP region.

Next, if an additive noise σ_τ is used in the wind burst equation τ (5), i.e., σ_τ being independent with the variations on the interannual time scale, then the PDF of T_E will become nearly Gaussian. As a consequence, the occurrence of the extreme El Niño events will become much less frequent and the amplitudes of the La Niña will become stronger. This reflects the importance of the observational character, i.e., there being a deterministic part of the wind bursts modulated by the low-frequency SST variation associated with ENSO, in triggering the asymmetry of the EP El Niño.

Finally, the role of the decadal variability I is studied. Here, in addition to the standard run with a time-varying I as in (6), the other two tests both have a fixed value of I, with $I \equiv 0$ and $I \equiv 4$, respectively. For a fair comparison, the random number generators $\dot{W}_\tau$ in the wind burst Eq. (5) in the three cases are set to be the same.

Clearly, if $I \equiv 0$, i.e., the background Walker circulation and zonal thermocline slope is relatively weak, then CP events occur less frequently while EP events become dominant (Table 2 for the segments comparison with the same period of the observation). Note that, even with $I \equiv 0$ in such a case, there remains a small number of the CP El Niños in the simulation, which are triggered by the stochastic noise. Quantitatively, in this case, 646 (22.7 ± 2.8) El Niños and 861 (30.1 ± 3.1) La Niñas can be identified in the 2000 (per 71) model years. Among the El Niños, 453 (16.0 ± 3.3) events are EP type and 193 (6.7 ± 2.0) events are CP one. Also, less multi-year ENSO events are generated, i.e., 78 (2.8 ± 1.8) multi-year El

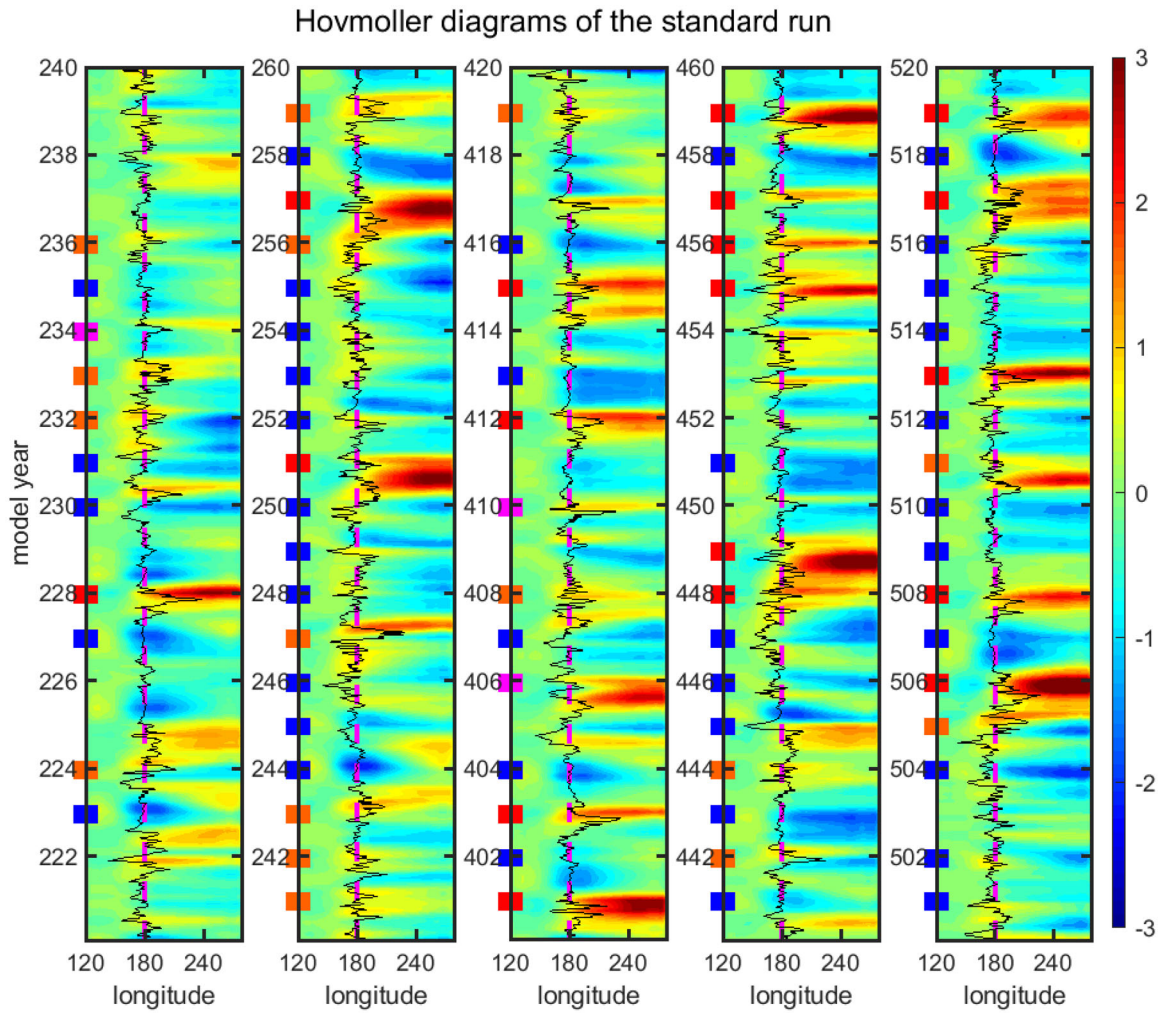

Fig. 5 Hovmoller diagrams based on the model simulations from the standard run. The equatorial SST variations are obtained by the bivariate linear regression. Here the coefficients of the regression model are obtained using the observational data. Then the Niño3 and Niño4 indices T_E and T_C from the model are plugging into the regression model to obtain the SST spatiotemporal patterns.The colored boxes on the left vertical axis (ranging from September to the next February) indicate the types of the ENSO events in boreal winter, which are based on the definitions in "Methods" section. The red, purple, orange and blue boxes are for the strong EP El Niño, the moderate EP El Niño, the CP El Niño and the La Niña events, respectively. The wind burst time series is placed on top of the Hovmoller diagram. The center of the wind burst time series is located at the dateline, where the WWBs and EWBs correspond to the time series values going toward the right and left from the dateline. The distance from the dateline represents the strength and direction of the wind bursts.

Niños and 200 (7.0 ± 2.1) La Niñas in total being found. Besides, there are 260 (9.1 ± 2.0) extreme El Niño events in total, i.e., a tendency of its occurrence for every 8 years. This frequency is twice as the standard run, consistent with the observation that 3 out of the 4 extreme El Niño events occurred before 2000. This is also consistent with the projection that an increasing frequency of extreme El Niño events will emerge due to the greenhouse warming, since a projected surface warming over the EP that occurs faster than in the surrounding ocean waters[49].

On the other hand, if $I \equiv 4$, i.e., with a relatively strong background Walker circulation and zonal thermocline slope, then the model simulations fail to capture many important features in observations (Table 2 for the segments comparison with the same period of the observation). Specifically, only 34 (1.2 ± 1.1) extreme El Niño events can be identified in this situation, which is much less than that in the standard run (125). In addition, the occurrence of the CP El Niño and La Niña will be more frequent and the resulting PDF of T_E will become nearly Gaussian, which is fundamentally different from the observed non-Gaussian PDF with a fat tail.

DISCUSSION

As was shown in the context, a three-region multiscale stochastic model is developed to show that the observed ENSO complexity can be explained by combining intraseasonal, interannual, and decadal processes. The model starts with a deterministic, linear and stable system for the interannual variabilities, which includes both the ocean heat content discharge/recharge and the ocean zonal advection. Then two stochastic processes with multiplicative noise describing the intraseasonal wind bursts and the decadal variation of the Walker circulation are incorporated. This three-region multiscale stochastic model can reproduce not only the general properties of ENSO events observed during the period of 1980–2020, but also the observed complexity in ENSO patterns (e.g., CP vs. EP El Niños), intensity (e.g., ~10–20 year reoccurrence frequency of extreme El Niño events), and evolution patterns (e.g., more multi-year La Niñas than multi-year El Niños), which are often hard to be simulated by the state-of-the-art models. The model also perfectly recovers the non-Gaussian SST statistics of nature in reproducing both the positively skewed fat-tailed PDF in the Niño3 region and the negatively skewed thin-tailed PDF in the

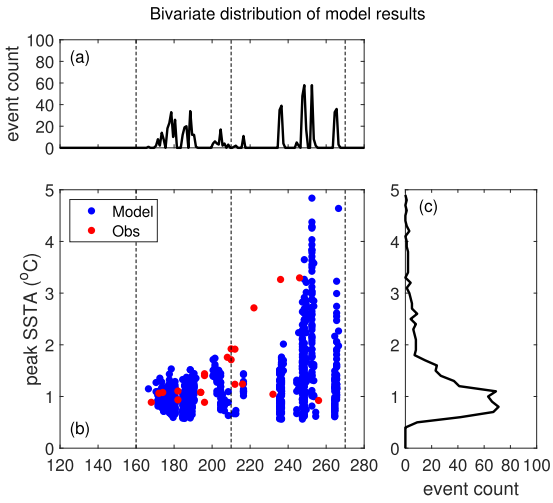

Fig. 6 Distribution of equatorial Pacific SSTa maxima for El Niño events of the standard simulations. For each of the qualified El Niño events, 660 in total, the winter-mean SST anomalies are averaged over the equatorial zone (5°S–5°N), and then the Pacific zonal maximum is located. **a** Distribution of peak SST anomaly longitudes. **b** Scatter plot of the peak SST anomaly value vs. the longitude at which it occurs. The blue (red) dots are for the model results (observations). **c** Distribution of peak SST anomaly values.

Niño4 region, which allows a systematic uncertainty quantification of the ENSO dynamics and facilitates the study of the extreme El Niño events. Actually, it is suggested that the models in the ensemble from phase 5 of the Coupled Model Intercomparison Project (CMIP5) have large deficiencies in ENSO amplitude, spatial structure, and temporal variability[50]. The multiplicative stochastically perturbed parameterization tendencies (SPPT) scheme is then included in coupled integrations of the National Center for Atmospheric Research (NCAR) Community Atmosphere Model, version 4 (CAM4). The SPPT scheme results in a significant improvement to the representation of ENSO in CAM4, improving the power spectrum and reducing the magnitude of ENSO toward that observed. Further analysis indicated that the SPPT improve the distribution of WWBs by increasing the stochastic component of WWB and reducing the overly strong dependency on SST compared to the control integration. This result is consistent with our result and verifies the importance of the state-dependent noise term on inducing the ENSO complexities in our simple model.

Except for the stochasticity of the model, the nonlinearity also plays an important role. In fact, based on a heat budget analysis of the mixed layer temperature with the observational dataset, the collective damping rate over the CP region is parameterized as a cubic polynomial function in terms of T_C. This is found to be crucial for obtaining the realistic negative-skewed PDF for the simulated T_C and therefore for simulating the realistic ENSO complexity. It should also be noted that the theoretical explanations of ENSO are always grouped into two categories[5–7]. In the first category, ENSO is viewed as a self-sustained, unstable and naturally oscillatory mode of the coupled ocean-atmosphere system, in which the nonlinearity acts mainly to bound the growing eigenmode and create a finite amplitude of the ENSO cycle. In the other category, ENSO is regarded as a stable (damped) mode triggered by atmospheric random "noise" forcing. Based on the results of this work, we conclude that the first explanation is more suitable for depicting the CP type of ENSO, since the nonlinearity plays an important role for its evolution. On the other hand, the second theory is more appropriate for explaining the development of the EP type of ENSO. In fact, the stochastic forcings, i.e., the WWBs (EWBs), are crucial for both the occurrences and the amplitudes of the EP El Niño (La Niña). This indicates that both the nonlinearity and stochastic processes are of great importance for simulating and studying the ENSO complexity.

It has been shown in this paper that the T_C and T_E time series combined with the bivariate regression technique can be utilized to simulate the spatiotemporal patterns of the SST, which are qualitatively and quantitatively similar to the observations. The stochastic forcing in the conceptual model allows the reconstructed spatiotemporal patterns to have the same level of the irregularity as nature, which outweighs most of the GCMs that are more deterministic. Thus, one direct application of this model is for prediction. The computational efficiency and the physical consistency of the model facilitates the machine learning forecast of ENSO. More specifically, the model can be easily used to create ENSO spatiotemporal patterns for several thousand years. Then the transfer learning technique[51] can be used to further improve the quality of these time series with the help of the limited but valuable observational data, which provides effective training data for the machine learning forecasts. Besides, the sensitivity analysis of the decadal variability can be utilized to effectively and quantitatively analyze the influences of the background state changes, e.g., the greenhouse warming or the historical changes at millennial time scales, on the ENSO characters. For example, if we can provide some assumed scenarios, e.g., the strength of the background Walker circulation, the corresponding ENSO complexities can be estimated by our conceptual model. Finally, the current modeling framework allows to further incorporate detailed additional physical processes for the ENSO complexity, such as the subtropical atmospheric forcing[36], which are now described by stochastic parameterizations.

METHODS

The datasets

The monthly ocean temperature and current data used here are all from the GODAS dataset[52]. The thermocline depth along the equatorial Pacific is approximated from the potential temperature as the depth of the 20 °C isotherm. GODAS dataset is available at a horizontal resolution of 1/3° × 1/3° near the tropics and has 40 vertical levels with 10 m resolution near the surface. The analysis period is from 1980 to 2020. Anomalies presented in this study are calculated by removing the monthly mean climatology of the whole period. In this work, the Niño4 (T_C) and the Niño3 (T_E) indices are the average of SST anomalies over the regions 160°E–150°W, 5°S–5°N and 150°W–90°W, 5°S–5°N, respectively. The h_W index is the average of thermocline depth anomaly over 120°E–180°, 5°S–5°N while the u index is the average mixed layer zonal current in the CP region. It should be noted that the model is constructed using observations of atmospheric and oceanic variable from 1980–2020 to determine the values of model parameters. This period is used because it is during the satellite era and its observations are more reliable than those before. However, in order to have a larger sample size of ENSO events to examine the model performance in simulating the ENSO complexity, SST observations from a longer period 1950–2020 were used in the analysis.

Next, the daily zonal wind data at 850 hPa from the NCEP-NCAR reanalysis[53] is used to describe the intraseasonal wind bursts. After removing the daily mean climatology, the anomalies are averaged over the WP region to create the wind burst index, which is shown in Fig. 7a. Note that τ lies in a faster time scale (daily) than the remaining variables (monthly), we use a stochastic process to describe its detailed behavior. Although a single daily value of τ has only a minor contribution to the SST variables, its accumulated effect over time will modulate the SST variations.

In addition to the interannual and intraseasonal data, the Walker circulation strength index is adopted to illustrate the modulation of the decadal variation on the interannual ENSO characters. It is defined as the sea level pressure difference over the CP/EP (160°W–80°W, 5°S–5°N) and over the Indian Ocean/west Pacific (80°E–160°E, 5°S–5°N)[54]. It should be stressed that the monthly zonal SST gradient between the WP and CP region is highly correlated with this Walker circulation strength index (i.e., their simultaneous correlation coefficient is around 0.85), suggesting the significant air-sea interacting characteristic over the equatorial Pacific. Since the latter is more directly related to the zonal advective feedback

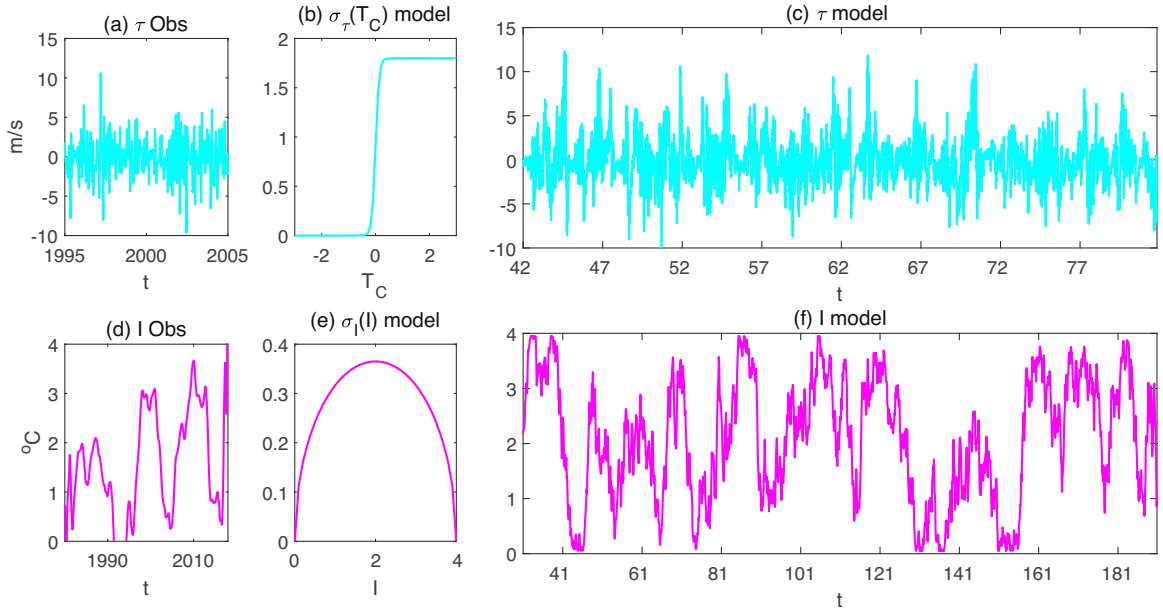

Fig. 7 Comparison between the observations and the model simulations for *τ* and *I*. a The observational wind burst time series. **b** The multiplicative noise σ_τ in the stochastic process (5) for the intraseasonal variable τ. **c** The model simulation of τ. **d**–**f** are similar to **a**–**c** but for the decadal variable *I*. In **d**, the observational *I* is smoothed by a 5-year window and then multiplied by a constant to make the final standard deviation be the same as the original one. Since *I* is a surrogate of the decadal variation of the Walker circulation, it is proportional to the zonal SST gradient between the WP and CP region. From the physical perspective, the SST over the CP is generally warmer than the EP in the decadal time scale. Thus, the *I* is always greater than 0 in our model interpretation, although there may be some extraordinary situation when the CP is colder than the WP, which is not considered in our model.

strength over the CP region, the decadal model (*I*) mainly reflects this variation.

Definitions of different types of the ENSO events

To quantify the ENSO complexity, the definitions of different El Niño and La Niña events are as follows, which are based on the average of the SST anomalies over the boreal winter (December–January–February). Following the definitions in the ref. [27], when the EP is warmer than the CP and is warmer than 0.5 °C, it is classified as the EP El Niño. Following the definitions used by the ref. [55], an extreme El Niño event corresponds to the situation that the maximum of EP SST anomaly from April to the next March is larger than 2.5 °C. When the CP is warmer than the EP and the CP SST anomaly is larger than 0.5 °C, the event is then defined as a CP El Niño. Finally, when either the CP and EP SST anomaly is less than −0.5 °C, it is defined as a La Niña event.

The starting interannual model: a deterministic three-region conceptual system

To develop the three-region multiscale stochastic model, the associated starting interannual model with four unknowns over the western, central, and EP regions is as follows:

$$\frac{du}{dt} = -ru - \frac{a_1 b_0 \mu}{2}(T_C + T_E), \tag{8}$$

$$\frac{dh_W}{dt} = -rh_W - \frac{a_2 b_0 \mu}{2}(T_C + T_E), \tag{9}$$

$$\frac{dT_C}{dt} = \left(\frac{\gamma b_0 \mu}{2} - c\right)T_C + \frac{\gamma b_0 \mu}{2}T_E + \gamma h_W + \sigma u, \tag{10}$$

$$\frac{dT_E}{dt} = \gamma h_W + \left(\frac{3\gamma b_0 \mu}{2} - c\right)T_E - \frac{\gamma b_0 \mu}{2}T_C, \tag{11}$$

where T_C and T_E are the SST in the CP and EP, respectively, while u is the ocean zonal current in the CP and h_W is the thermocline depth in the WP. All the four variables are anomalies. This starting interannual skeleton model for ENSO is constructed to depict the air-sea interactions over the entire western, CP and EP[34]. The key physics of the model are summarized as follows. First, the monthly variations of the thermocline slope and zonal wind stress over the central-western and the CP-EP regions are tightly linked through the Sverdrup balance relationships, as observed. As a result, if one obtains the variation of the thermocline depth anomalies over the WP (h_W), those over the CP (h_C) and EP (h_E) can also be diagnosed. Second, in the absence of the ocean zonal advection, the dynamic equations of the h_W and T_E degenerate to those in the recharge paradigm. Third, to introduce the zonal advective feedback, a simple equation for the mixed layer zonal current is adopted. Finally, in contrast to the recharge paradigm, which considers the thermocline feedback as the only positive feedback in the EP, the development of the SST in the CP is also influenced by the zonal advective feedback. Combining these elements yields the linear coupled system.

It can be seen that when the coefficient σ in (10) is set to be zero, i.e., ignoring the zonal advective feedback, the system will degrade into the recharge paradigm, with $T_C = T_E$, which illustrates the EP type of ENSO with no emphasis on the differences between the CP and EP regions[3]. Therefore, the three-region model can be seen as an extension of the recharge paradigm. In the model, the collective damping rate c is dominated by the time scale over which water in the equatorial band is replaced by the mean climatological upwelling, i.e., about 2 months; the parameter γ measures the strength of the thermocline feedback, which is chosen to give an SST rate of change of 1.5 °C over 2 months per 10 m of the thermocline depth anomaly over the EP. Similarly, the coefficient σ, which measures the strength of the zonal advective feedback, is chosen to give an SST rate of change of 1.5 °C over 2 months per 0.5 m/s of the zonal current anomaly over the CP, i.e., the background zonal SST difference between the WP and CP is 3 °C. The collective damping rate r in the ocean adjustment is set as 1/(8 months), which is induced by the loss of energy to the boundary currents of the west and east sides of the ocean basin. Due to the fact that, for a given steady zonal wind stress forcing, the zonal mean thermocline depth anomaly of the recharge oscillator model is about zero at the equilibrium state, i.e., $h_E + h_W = 0$, one finds that a will be about half of r. The parameter b_0, which is the high-end estimation of the thermocline tilt and is in balance with the zonal wind stress produced by the SST anomaly, is chosen to give 50 m of east-west thermocline depth difference per 1 °C of the SST anomaly. Thus, this model is nondimensionalized in a similar way as the recharge paradigm, i.e., by scales of $[h] = 150$ m, $[T] = 7.5$ °C, $[u] = 1.5$ m/s, and $[t] = 2$ months for anomalous thermocline depth, SST, mixed layer zonal current, and the time variables, respectively. Accordingly, parameters c, r, α_1 and α_2 are scaled by 1/[t], and

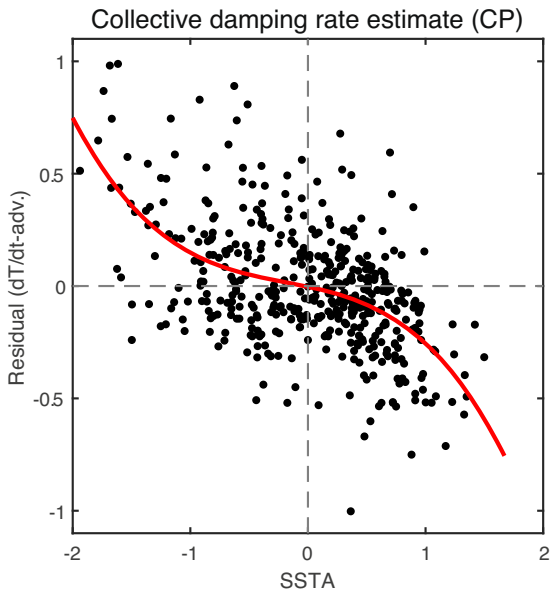

Fig. 8 The scatter plot of T_C and the residual in the temperature Eq. (14). Each black dot represents the value for 1 month. The black curve: the cubic polynomial fit of the scatter plot. The equation of the curve is $f(T_C) = -0.11{T_C}^3 - 0.06{T_C}^2 - 0.08T_C - 0.01$, which passed the 95% confidence from a Student's t test.

parameters γ and b_0 by $[h]/[t]/[T]$ and $[T]/[h]$. Their non-dimensional values are $c = 1$, $\gamma = 0.75$, $\sigma = 0.6$, $r = 0.25$, $\alpha_1 = 0.0625$, $\alpha_2 = 0.125$ and $b_0 = 2.5$, which all correspond to those used in the recharge paradigm.

In the model, the relative coupling coefficient μ is 0.5, which is smaller than the critical value (i.e., 0.7 for purely oscillating). Such a choice implies that all the eigenvalues of this four-dimensional model have negative real part, representing the negative growth rates of the solution. It also allows the model to have a pair of conjugate solutions, i.e., damped oscillating solutions, which mimic the ENSO cycles. These features facilitate the stochastic excitation nature of the ENSO events by the random wind bursts. It has been shown that this conceptual model can depict the different variations between the CP and EP well[34]. Specifically, with an increasing magnitude of the zonal advective feedback over the CP, i.e., imitating the situation for CP ENSO, the period of the system and SST magnitude over the CP and EP both decrease. Note that the decreasing amplitude is more intense over the EP, indicating an enlargement of the SST differences between the CP and EP. These results are all consistent with the observational characteristics of the CP El Niño.

The model succeeds in describing the basic two-regime dynamical behavior of the ENSO for the EP and CP events. Yet, due to the deterministic nature, it cannot reproduce the observed irregularity of ENSO in amplitude and phase as well as the regime switching behavior and the non-Gaussian PDFs. Therefore, to simulate the realistic ENSO complexity, additional processes in the intraseasonal and decadal time scales are further developed and coupled to the deterministic interannual model.

The intraseasonal model for the random wind bursts

The intraseasonal variability accounts for several important ENSO triggers, such as the WWBs, the EWBs, as well as the convective envelope of the MJO, which serve as the random input for the large-scale ENSO dynamics[14,18,20,56]. The intraseasonal component here is modeled by a simple stochastic differential Eq. (5). One important feature of (5) is that the noise coefficient σ_τ is state-dependent, i.e., a multiplicative noise. Here we assume it is positively correlated with T_C since according to observations most of the wind bursts are active in the central-west Pacific[20,57,58]. The stochastic process (5) can generate both the WWBs and the EWBs, corresponding to τ (wind burst amplitude in the model) with positive and negative values, respectively. Notably, the state-dependent noise used here is very different from the previous two-region conceptual models, where the noise dependence was on T_E[59,60] due to the lack of the state variable T_C in those models.

The intraseasonal component here is modeled by a simple stochastic differential Eq. (5), which accounts for its intermittent and unpredictable nature at interannual time scale. The variable τ is the wind burst amplitude with a unit of $[\tau] = 5$ m/s. The damping parameter $d_\tau = 2$ representing a time scale of 1 month of the wind envelope but each individual wind is random in the daily time scale. The reason to adopt 1 month is that the decorrelation time of the MJO and wind activity in the WP area is roughly around that time scale[61]. The explicit expression of the noise coefficient is:

$$\sigma_\tau(T_C) = 0.9[\tan h(7.5T_C) + 1], \tag{12}$$

which clearly indicates a positive correlation between T_C and σ_τ. The reason for adopting a hyperbolic tangent function is to prevent the unbounded growth of σ_τ when the absolute value of T_C becomes large. Note that T_C here is the non-dimensional value. The parameterization of $\sigma_\tau(T_C)$ is mainly based on the fact that an increased SST enhances the chance of the occurrence of the wind bursts. An analogous form has been used in refs. [59,61], despite the fact that the multiplicative noise dependence in those papers is on T_E since there is no T_C in those models. The profile of $\sigma_\tau(T_C)$ and a random realization of the simulated τ can be found in Fig. 7b, c, respectively. The latter bears a high resemblance with the observations (Fig. 7a).

The decadal model for the Walker circulation

Several detailed El Niño-type classification methods have been utilized to show that since 1870 the EP and CP events were alternatively prevalent every 10 or 20 years[62,63]. For example, the EP episodes were the dominant ones in the 1980s while the CP El Niño events occurred more frequently since 2000[14]. These findings indicate that the decadal variability plays an important role in driving the switching between the CP- and EP-dominant regimes. Thus, it is necessary to incorporate the decadal effect into the coupled ENSO model. The decadal model (6) proposed here is a simple stochastic process, which aims at describing the large-scale behavior, including the characteristic time scale and the amplitude. For simplicity, the decadal variable I is assumed to have no explicit dependence on the variables in the faster time scales, i.e., intraseasonal and interannual, but those ingredients are nevertheless effectively parameterized in the stochastic noise. The damping parameter $\lambda = 2/60$ is taken such that the decorrelation time $1/\lambda$ is about half of a decade. Next, despite that the observational data allows us to determine the range of I being between $I = 0$ and $I = 4$, the 40-year observational data is too short to provide unbiased information about the PDF of the decadal variability. Note that since I is a slow-varying variable and it is bounded below by $I = 0$, i.e., SST in the WP is warmer than the CP on the decadal time scale, it is unreasonable to assume a Gaussian distribution. Here, we adopt the uniform distribution function of I. This is based on the fact that the uniform distribution is the maximum entropy solution for a function in the finite interval without additional information[64]. Numerical tests have shown that replacing the uniform distribution by other empirically determined PDFs of I, such as a truncated Gaussian or a truncated bimodal distribution, only has a minor impact on the SST statistics provided that the decorrelation time of I lies in the decadal time scale and the probability of I at each point within the interval [0, 4] is non-vanishing. The resulting $\sigma_I(I)$ associated with the uniform distribution of $p(I)$ is included in Fig. 7e. Figure 7f also shows a random realization of the time series I, which clearly indicates a stochastic regime switching behavior in the decadal time scale. The parameter m is the mean of I, which can be computed from its PDF.

The stochastic decadal model is shown in (6), where I is a surrogate of the decadal variation of the Walker circulation, and also the zonal SST difference between the WP and CP regions that directly determines the strength of the zonal advective feedback. In other words, σ in (10) can be regarded to be proportional to I. Specifically, $\sigma = 0.2I$ is used here, suggesting that it could give an SST rate of change of 1.5 °C over 2 months per 0.375 m/s (when $I = 4$, i.e., the CP ENSO regime) or 1.5 m/s (when $I = 1$, i.e., the EP ENSO regime) of the zonal current anomaly over the CP.

Determining the nonlinearity in the coupled model

Note that one difference between the starting deterministic model and the coupled stochastic model is the collective damping rates. The single damping coefficient in the deterministic model is splitted into two distinct values c_1 and c_2 in the governing equations of T_C and T_E, respectively. The damping parameter:

$$c_2 = 1.4 \tag{13}$$

in the T_E equation remains as a constant since the WWBs are the main contribution to the positive skewness and the one-sided fat tail of the Niño3 PDF corresponding to a large kurtosis (Fig. 3c). On the other hand, the ocean zonal current plays a more important role in the CP. Thus, the WWBs are not the main mechanism for the non-Gaussian statistics of the SST in the CP region since otherwise the associated PDF would have a similar profile as that of T_E, which is however not the case for the PDF associated with the observational data. In fact, the PDF of the observed Niño4 SST has a different skewness direction and it has no fat tail (Fig. 3d).

To understand the contributor of the nonlinear and non-Gaussian features in the CP region, a heat budget analysis of the mixed layer temperature is performed as follows:

$$\frac{\partial T_C'}{\partial t} = -u'\frac{\partial \overline{T}_C}{\partial x} - v'\frac{\partial \overline{T}_C}{\partial y} - w'\frac{\partial \overline{T}_C}{\partial z} - u'\frac{\partial T_C'}{\partial x} - v'\frac{\partial T_C'}{\partial y} - w'\frac{\partial T_C'}{\partial z} - \overline{u}\frac{\partial T_C'}{\partial x} - \overline{v}\frac{\partial T_C'}{\partial y} - \overline{w}\frac{\partial T_C'}{\partial z} + Res, \tag{14}$$

where overbars and primes indicate monthly climatology and anomaly, respectively. The variables u, v and T_C indicate zonal current, meridional current and oceanic temperature averaged over the mixed layer (top 50 m). The vertical velocity (w) is calculated at the bottom of the mixed layer and *Res* is the residual term, which represents the collective damping[6]. T_C can be impacted by the vertical variations. One is the thermocline feedback that relates to the climatological upwelling and anomalous temperature gradient between the mixed layer mean and that below it, which is always parameterized by the thermocline depth anomalies (h). The other is the Ekman feedback that relates to the anomalous upwelling and climatology temperature gradient that can be prescribed. Figure 8 shows a scatter plot of this residual term as a function of T_C for the data at each month during the satellite observational era, which exhibits a clear nonlinear dependence. In fact, a cubic polynomial fitting curve in terms of T_C, also shown in the figure, well represents the relationship. It implies that c_1 is small for the small local SST anomalies and becomes larger for the large SST anomalies. This physically corresponds to the fact that larger damping will emerge (e.g., strong wind and precipitation) when the underlying SST is too warm[65]. Thus, in the coupled model, c_1 is depicted by a simple nonlinear equation:

$$c_1(T_C) = 25\left(T_C + \frac{0.75}{7.5}\right)^2 + 0.9. \tag{15}$$

Note that $c_1(T_C)$ is not centered at $T_C = 0$, which is consistent with the observational data in Fig. 8. Recall that the dimension of T_C is $[T] = 7.5$ °C. Therefore, c_1 is centered at -0.75 °C. This non-zero center explains the asymmetry of the SST in the CP region, which leads to a negative skewness since the damping with positive T_C is overall stronger than that with the negative one. In addition, this cubic damping prevents strong T_C in both positive and negative sides, which facilitates a kurtosis of the Niño4 SST distribution that is smaller than 3. This is consistent with the observations and it is also distinguished from the large kurtosis in the Niño3 SST.

Next, the strength of the zonal advective feedback is depicted by σ, with a modulation by the decadal variation of the Walker circulation. In fact, the zonal advective feedback is:

$$\frac{\partial T_C'}{\partial t} = -u'\frac{\partial \overline{T}_C}{\partial x}, \tag{16}$$

where $\partial \overline{T}_C$ is the background zonal SST difference between the WP and CP, which shows to directly control σ with a linear relationship and can be depicted by the decadal model (6) of variable I. Specifically, if $\partial \overline{T}_C$ is 3 °C (i.e., $I = 3$), as provided in the standard setting, it could give an SST rate of change of 1.5 °C over nearly 2 months per 0.5 m/s of the zonal current anomaly over the CP, since the distance between the WP and CP is fixed (50° longitude). Under this situation, σ will be 0.6 according to the nondimensional values. As a result, a simple relationship between σ and I can be derived, that is, $\sigma = 0.2I$, suggesting that it could give an SST rate of change of 1.5 °C over 2 months per 0.375 m/s (when $I = 4$, i.e., the CP ENSO regime) or 1.5 m/s (when $I = 1$, i.e., the EP ENSO regime) of the zonal current anomaly over the CP. This is comparable to the choice of the thermocline feedback strength γ, as discussed before.

Finally, the coefficient $C_u = 0.03$ is adopted in the coupled model. And the temporal resolution is 0.002 time unit, i.e., 2.88h, which is sufficient to resolve the intraseasonal variations.

Coupling coefficients between the interannual variables and the wind bursts

In the coupled stochastic model, the coupling coefficients β_u, β_h, β_C and β_E are determined systematically via the eigenmodes of the deterministic model. Specifically, since deterministic model is characterized by a pair of the damped oscillating modes, any external forcing that is imposed on this characteristic direction can be distributed to each of its four components u, h_W, T_C and T_E by multiplying the corresponding component of the eigenvector for the mathematical consistency. A direct calculation shows that:

$$(\beta_u, \beta_h, \beta_C, \beta_E)^T = (-0.2, -0.4, 0.8, 1.0)^T \beta_E,$$

which implies the response of T_C is positively correlated to that of T_E due to the wind burst forcing while the changes in h_W and u are anti-correlated with the SST response. These are all consistent with physics. Here the coefficient β_E (before including the seasonal phase locking effect) takes the value:

$$\beta_E = 0.15\left(2 - \frac{I}{5}\right).$$

It implies that β_E increases as the decadal variable I decreases. In other words, the wind bursts becomes stronger when I favors the EP-dominant regime.

Finally, the values of the white noise strengths are:

$$\sigma_u = 0.04 \qquad \sigma_h = 0.02 \qquad \sigma_C = 0.04 \qquad \text{and} \qquad \sigma_E = 0.$$

Here, the uncertainties in u and T_C are the largest since their actual intrinsic processes are more complicated than the simple structures used here[6,66]. On the other hand, no additional noise is imposed to the equation of T_E because T_E is largely modulated by the random wind bursts. It should be noted that the parameters are optimized within a certain range that leads to physically meaningful results. These parameters are calibrated by matching the climatological PDF with that from the observational data. The criterion of computing the difference between the two PDFs is the relative entropy, which is a widely used information measurement for assessing the statistical difference. We have tested the sensitivity of the model parameters. Adding certain perturbations to the parameters only cause small changes of the PDFs and the model simulated trajectories, provided that the perturbed parameter values remain in the possible range with physical meanings.

Seasonal phase locking

Seasonal phase locking is one of the remarkable features of ENSO, which manifests in the tendency of ENSO events to peak during boreal winter and is mainly related to the pronounced seasonal cycle of mean state[37,38]. Specifically, in the CP-EP, the climatological SST cools in boreal fall and warms in spring as a result of the seasonal motion of the ITCZ, which also modulates the strength of the upwelling and horizontal advection processes to influence the evolution of the SST anomalies[40]. Since the cool (warm) SSTs tend to coincide with decreased (increased) convective activity and upper cloud cover, a season-dependent damping term, which represents the cloud radiative feedback, can account for this seasonal variation in a simple fashion[67]. Besides, the increased wind burst activity in winter as a direct response to the increased atmospheric intraseasonal variability such as the MJO is the main constitution of the seasonal cycle in the WP[58,68]. As a result, the seasonal cycle effects can be incorporated into the parameters $\sigma_\tau(T_C)$ in (12), $c_1(T_C)$ in (15) and c_2 in (13):

$$\begin{aligned} \sigma_\tau(T_C, t) &= 0.9[\tanh(7.5T_C) + 1]\left[1 + 0.3\cos\left(\tfrac{2\pi}{6}t + \tfrac{2\pi}{6}\right)\right], \\ c_1(T_C, t) &= \left[25\left(T_C + \tfrac{0.75}{7.5}\right)^2 + 0.9\right]\left[1 + 0.3\sin\left(\tfrac{2\pi}{6}t - \tfrac{2\pi}{6}\right)\right], \\ c_2(t) &= 1.4\left[1 + 0.2\sin\left(\tfrac{2\pi}{6}t + \tfrac{2\pi}{6}\right) + 0.15\sin\left(\tfrac{2\pi}{3}t + \tfrac{2\pi}{6}\right)\right]. \end{aligned} \tag{17}$$

Recall that the time unit [t] is 2 months. Therefore, t going from 0 to 6 completes 1 year, where $t = 0$ corresponds to January. Note from (17) that the strength of the wind bursts peaks in boreal winter, which is consistent with observations for the WWBs and the MJO[39,58]. The second sinusoidal function in the collective damping c_2 represents a semiannual contribution to the seasonally modulated variance, as was suggested in a previous work[38], which does not directly link with a semiannual cycle of the SST itself.

DATA AVAILABILITY

The monthly ocean temperature and current data were downloaded from the National Centers for Environmental Prediction Global Ocean Data Assimilation System (https://www.esrl.noaa.gov/psd/data/gridded/data.godas.html). The daily zonal wind data at 850 hPa were downloaded from the NCEP-NCAR reanalysis (https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html).

CODE AVAILABILITY

Any relevant code necessary to reproduce the results presented here is available upon request.

Received: 21 December 2021; Accepted: 9 February 2022;
Published online: 11 March 2022

REFERENCES

1. Ropelewski, C. F. & Halpert, M. S. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Weather Rev.* **115**, 1606–1626 (1987).
2. Klein, S. A., Soden, B. J. & Lau, N.-C. Remote sea surface temperature variations during ENSO: evidence for a tropical atmospheric bridge. *J. Clim.* **12**, 917–932 (1999).
3. Jin, F.-F. An equatorial ocean recharge paradigm for ENSO. Part I: conceptual model. *J. Atmos. Sci.* **54**, 811–829 (1997).
4. Capotondi, A. et al. Understanding ENSO diversity. *Bull. Am. Meteorol. Soc.* **96**, 921–938 (2015).
5. Timmermann, A. et al. El Niño-Southern Oscillation complexity. *Nature* **559**, 535–545 (2018).
6. Zebiak, S. E. & Cane, M. A. A model El Niño–Southern Oscillation. *Mon. Weather Rev.* **115**, 2262–2278 (1987).
7. Wang, C. A review of ENSO theories. *Natl Sci. Rev.* **5**, 813–825 (2018).
8. Ashok, K., Behera, S. K., Rao, S. A., Weng, H. & Yamagata, T. Niño Modoki and its possible teleconnection. *J. Geophys. Res. Oceans* **112**, C11007 (2007).
9. Yu, J.-Y. & Kao, H.-Y. Decadal changes of ENSO persistence barrier in SST and ocean heat content indices: 1958–2001. *J. Geophys. Res. Atmos* **112**, D13106 (2007).
10. Kao, H.-Y. & Yu, J.-Y. Contrasting eastern-Pacific and central-Pacific types of ENSO. *J. Clim.* **22**, 615–632 (2009).
11. Barnston, A. G., Tippett, M. K., L'Heureux, M. L., Li, S. & DeWitt, D. G. Skill of real-time seasonal ENSO model predictions during 2002–11: is our capability increasing? *Bull. Am. Meteorol. Soc.* **93**, 631–651 (2012).
12. Zheng, F., Fang, X.-H., Yu, J.-Y. & Zhu, J. Asymmetry of the Bjerknes positive feedback between the two types of El Niño. *Geophys. Res. Lett.* **41**, 7651–7657 (2014).
13. Fedorov, A. V., Hu, S., Lengaigne, M. & Guilyardi, E. The impact of westerly wind bursts and ocean initial state on the development, and diversity of El Niño events. *Clim. Dyn.* **44**, 1381–1401 (2015).
14. Chen, D. et al. Strong influence of westerly wind bursts on El Niño diversity. *Nat. Geosci.* **8**, 339–345 (2015).
15. Lian, T. et al. Linkage between westerly wind bursts and tropical cyclones. *Geophys. Res. Lett.* **45**, 11–431 (2018).
16. Levine, A., Jin, F. F. & McPhaden, M. J. Extreme noise–extreme El Niño: how state-dependent noise forcing creates El Niño–La Niña asymmetry. *J. Clim.* **29**, 5483–5499 (2016).
17. Capotondi, A., Sardeshmukh, P. D. & Ricciardulli, L. The nature of the stochastic wind forcing of ENSO. *J. Clim.* **31**, 8081–8099 (2018).
18. Hu, S. & Fedorov, A. V. Exceptionally strong easterly wind burst stalling El Niño of 2014. *Proc. Natl Acad. Sci.* **113**, 2005–2010 (2016).
19. Hu, S. & Fedorov, A. V. The extreme El Niño of 2015–2016 and the end of global warming hiatus. *Geophys. Res. Lett.* **44**, 3816–3824 (2017).
20. Puy, M., Vialard, J., Lengaigne, M. & Guilyardi, E. Modulation of equatorial Pacific westerly/easterly wind events by the Madden–Julian oscillation and convectively-coupled Rossby waves. *Clim. Dyn.* **46**, 2155–2178 (2016).
21. Penland, C. & Sardeshmukh, P. D. The optimal growth of tropical sea surface temperature anomalies. *J. Clim.* **8**, 1999–2024 (1995).
22. Penland, C., Flügel, M. & Chang, P. Identification of dynamical regimes in an intermediate coupled ocean–atmosphere model. *J. Clim.* **13**, 2105–2115 (2000).
23. Lee, T. & McPhaden, M. J. Increasing intensity of el Niño in the central-equatorial pacific. *Geophys. Res. Lett* **37**, L14603 (2010).
24. Yeh, S.-W. et al. El Niño in a changing climate. *Nature* **461**, 511–514 (2009).
25. McPhaden, M., Lee, T. & McClurg, D. El Niño and its relationship to changing background conditions in the tropical Pacific ocean. *Geophys. Res. Lett.* **38**, L15709 (2011).
26. Xiang, B., Wang, B. & Li, T. A new paradigm for the predominance of standing central Pacific warming after the late 1990s. *Clim. Dyn.* **41**, 327–340 (2013).
27. Kug, J.-S., Jin, F.-F. & An, S.-I. Two types of El Niño events: cold tongue El Niño and warm pool El Niño. *J. Clim.* **22**, 1499–1515 (2009).
28. Picaut, J., Masia, F. & Du Penhoat, Y. An advective-reflective conceptual model for the oscillatory nature of the ENSO. *Science* **277**, 663–666 (1997).
29. Fang, X.-H., Zheng, F. & Zhu, J. The cloud-radiative effect when simulating strength asymmetry in two types of El Niño events using CMIP5 models. *J. Geophys. Res. Oceans* **120**, 4357–4369 (2015).
30. Fang, X. & Zheng, F. Simulating eastern-and central-pacific type ENSO using a simple coupled model. *Adv. Atmos. Sci.* **35**, 671–681 (2018).
31. Jin, F.-F., Lin, L., Timmermann, A. & Zhao, J. Ensemble-mean dynamics of the ENSO recharge oscillator under state-dependent stochastic forcing. *Geophys. Res. Lett.* **34**, L03807 (2007).
32. Levine, A. F. & Jin, F. F. A simple approach to quantifying the noise–enso interaction. part I: deducing the state-dependency of the windstress forcing using monthly mean data. *Clim. Dyn.* **48**, 1–18 (2017).
33. Geng, T., Cai, W. & Wu, L. Two types of ENSO varying in tandem facilitated by nonlinear atmospheric convection. *Geophys. Res. Lett.* **47**, e2020GL088784 (2020).
34. Fang, X.-H. & Mu, M. A three-region conceptual model for central Pacific El Niño including zonal advective feedback. *J. Clim.* **31**, 4965–4979 (2018).
35. Yang, Q., Majda, A. J. & Chen, N. ENSO diversity in a tropical stochastic skeleton model for the MJO, El Niño, and dynamic Walker circulation. *J. Clim.* **34**, 1–56 (2021).
36. Fang, S.-W. & Yu, J.-Y. A control of ENSO transition complexity by tropical pacific mean SSTs through tropical-subtropical interaction. *Geophys. Res. Lett.* **47**, e2020GL087933 (2020).
37. Tziperman, E., Zebiak, S. E. & Cane, M. A. Mechanisms of seasonal–ENSO interaction. *J. Atmos. Sci.* **54**, 61–71 (1997).
38. Stein, K., Timmermann, A., Schneider, N., Jin, F.-F. & Stuecker, M. F. ENSO seasonal synchronization theory. *J. Clim.* **27**, 5285–5310 (2014).
39. Zhang, C. Madden-Julian oscillation. *Rev. Geophys.* **43**, RG2003 (2005).
40. Mitchell, T. P. & Wallace, J. M. The annual cycle in equatorial convection and sea surface temperature. *J. Clim.* **5**, 1140–1156 (1992).
41. Yu, J.-Y., Wang, X., Yang, S., Paek, H. & Chen, M. *The Changing El Niño–Southern Oscillation and Associated Climate Extremes* (John Wiley & Sons, Inc., 2017).
42. Yu, J.-Y. & Fang, S.-W. The distinct contributions of the seasonal footprinting and charged-discharged mechanisms to enso complexity. *Geophys. Res. Lett.* **45**, 6611–6618 (2018).
43. Yu, J.-Y., Zou, Y., Kim, S. T. & Lee, T. The changing impact of el niño on us winter temperatures. *Geophys. Res. Lett.* **39**, L15702 (2012).
44. Sun, F. & Yu, J.-Y. A 10–15-yr modulation cycle of ENSO intensity. *J. Clim.* **22**, 1718–1735 (2009).
45. Xue, Y., Chen, M., Kumar, A., Hu, Z.-Z. & Wang, W. Prediction skill and bias of tropical pacific sea surface temperatures in the NCEP climate forecast system version 2. *J. Clim.* **26**, 5358–5378 (2013).
46. Fang, S.-W. & Yu, J.-Y. Contrasting transition complexity between El Niño and La Niña: observations and cmip5/6 models. *Geophys. Res. Lett.* **47**, e2020GL088926 (2020).
47. Thual, S., Majda, A. J. & Chen, N. Statistical occurrence and mechanisms of the 2014–2016 delayed super El Niño captured by a simple dynamical model. *Clim. Dyn.* **52**, 2351–2366 (2019).
48. Jin, F.-F., An, S.-I., Timmermann, A. & Zhao, J. Strong El Niño events and nonlinear dynamical heating. *Geophys. Res. Lett.* **30**, 20–1 (2003).
49. Cai, W. et al. Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Clim. Change* **4**, 111–116 (2014).
50. Christensen, H., Berner, J., Coleman, D. R. & Palmer, T. Stochastic parameterization and El Niño–southern oscillation. *J. Clim.* **30**, 17–38 (2017).
51. Torrey, L. & Shavlik, J. Transfer learning. In Soria, E. (ed.) *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, 242–264 (IGI Global, 2010).
52. Behringer, D. & Xue, Y. Evaluation of the global ocean data assimilation system at NCEP: the Pacific Ocean. In *Proc. Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, AMS 84th Annual Meeting, Washington State Convention and Trade Center, Seattle, Washington, 11–15 (2004).
53. Kalnay, E. et al. The ncep/ncar 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–472 (1996).
54. Kang, S. M. et al. Walker circulation response to extratropical radiative forcing. *Sci. Adv.* **6**, eabd3021 (2020).
55. Wang, B. et al. Historical change of El Niño properties sheds light on future changes of extreme El Niño. *Proc. Natl Acad. Sci.* **116**, 22512–22517 (2019).
56. Vecchi, G. A., Wittenberg, A. & Rosati, A. Reassessing the role of stochastic forcing in the 1997–1998 El Niño. *Geophys. Res. Lett.* **33**, L01706 (2006).

57. Tziperman, E. & Yu, L. Quantifying the dependence of westerly wind bursts on the large-scale tropical Pacific SST. *J. Clim.* **20**, 2760–2768 (2007).
58. Hendon, H. H., Wheeler, M. C. & Zhang, C. Seasonal dependence of the MJO–ENSO relationship. *J. Clim.* **20**, 531–543 (2007).
59. Levine, A. F. & Jin, F.-F. Noise-induced instability in the ENSO recharge oscillator. *J. Atmos. Sci.* **67**, 529–542 (2010).
60. Perez, C. L., Moore, A. M., Zavala-Garay, J. & Kleeman, R. A comparison of the influence of additive and multiplicative stochastic forcing on a coupled model of ENSO. *J. Clim.* **18**, 5066–5085 (2005).
61. Thual, S., Majda, A. J., Chen, N. & Stechmann, S. N. Simple stochastic model for El Niño with westerly wind bursts. *Proc. Natl Acad. Sci.* **113**, 10245–10250 (2016).
62. Yu, J.-Y. & Kim, S. T. Identifying the types of major El Niño events since 1870. *Int. J. Climatol.* **33**, 2105–2112 (2013).
63. Dieppois, B. et al. ENSO diversity shows robust decadal variations that must be captured for accurate future projections. *Commun. Earth Environ.* **2**, 1–13 (2021).
64. Kapur, J. N. & Kesavan, H. K. Entropy optimization principles and their applications. In Singh, V. P., Fiorentino, M. (eds) *Entropy and Energy Dissipation in Water Resources*, 3–20 (Springer, 1992).
65. Xie, R., Mu, M. & Fang, X. New indices for better understanding enso by incorporating convection sensitivity to sea surface temperature. *J. Clim.* **33**, 7045–7061 (2020).
66. McCreary, J. A linear stratified ocean model of the equatorial undercurrent. *Philos. Trans. Royal Soc. London. Series A, Math. Phys. Sci.* **298**, 603–635 (1981).
67. Thual, S., Majda, A. & Chen, N. Seasonal synchronization of a simple stochastic dynamical model capturing El Niño diversity. *J. Clim.* **30**, 10047–10066 (2017).
68. Seiki, A. & Takayabu, Y. N. Westerly wind bursts and their relationship with intraseasonal variations and ENSO. Part I: statistics. *Mon. Weather Rev.* **135**, 3325–3345 (2007).

ACKNOWLEDGEMENTS

The research of X.F. is supported by Guangdong Major Project of Basic and Applied Basic Research (Grant No. 2020B0301030004), the Ministry of Science and Technology of the People's Republic of China (Grant No. 2020YFA0608802) and the National Natural Science Foundation of China (Grant Nos. 42192564 and 41805045). The research of N.C. is partially funded by the Office of VCRGE at UW-Madison and ONR N00014-21-1-2904. The research of J.Y. is supported by the National Science Foundation of the United States (Grant Nos. AGS-1833075 and AGS-2109539).

AUTHOR CONTRIBUTIONS

N.C. and X.F. designed the project. N.C., X.F. and J.Y. performed the research. N.C. and X.F. wrote the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Xianghui Fang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022