The El Niño Southern Oscillation (ENSO) has significant impact on global climate and seasonal prediction. A simple modeling framework is developed here that automatically captures the statistical diversity of ENSO. First, a stochastic parameterization of the wind bursts including both westerly and easterly winds is coupled to a simple ocean–atmosphere model that is otherwise deterministic, linear, and stable. Second, a simple nonlinear zonal advection with no ad hoc parameterization of the background sea-surface temperature (SST) gradient and a mean easterly trade wind anomaly representing the multidecadal acceleration of the trade wind are both incorporated into the coupled model that enabled anomalous warm SST in the central Pacific. Then a three-state stochastic Markov jump process is used to drive the wind burst activity that depends on the strength of the western Pacific warm pool in a simple and effective fashion. It allows the coupled model to simulate the quasi-regular moderate traditional El Niño, the super El Niño, and the central Pacific (CP) El Niño as well as the La Niña with realistic features. In addition to the anomalous SST, the Walker circulation anomalies at different ENSO phases all resemble those in nature. In particular, the coupled model succeeds in reproducing the observed episode during the 1990s, where a series of 5–y CP El Niños is followed by a super El Niño and then a La Niña. Importantly, both the variance and the non-Gaussian statistical features in different Niño regions spanning from the western to the eastern Pacific are captured by the coupled model.

Significance

The El Niño Southern Oscillation (ENSO) has significant impact on global climate and seasonal prediction. A simple modeling framework is developed here that automatically captures the statistical diversity of ENSO. In addition to simulating different types of El Niño and La Niña with realistic features, the model succeeds in capturing both the variance and the non-Gaussian statistical properties in different Niño regions spanning the Pacific. Particularly, the observed episode during the 1990s, where a 5–y central Pacific El Niño is followed by a super El Niño and then a La Niña, is reproduced by the model. Key features of the model are state-dependent stochastic wind bursts and nonlinear advection of sea-surface temperature that allow effective transitions between different ENSO states.
of the wind bursts including both westerly and easterly winds is coupled to the simple ocean–atmosphere system, where according to the observational evidence (27–29) the amplitude of the wind bursts depends on the strength of SST in the western Pacific warm pool. Such a coupled model is fundamentally different from the Cane–Zebiak model (30) and other nonlinear models (31, 32), where the internal instability rather than the external wind bursts triggers the ENSO cycles. It is shown that (33) in addition to recovering the traditional moderate El Niño and super El Niño as well as the La Niña, the coupled model is able to capture key features of the observational record in the eastern Pacific. Second, a simple nonlinear zonal advection with no ad hoc parameterization of the background SST gradient is introduced that creates a coupled nonlinear advective mode of SST. In addition, due to the recent multidecadal strengthening of the easterly trade wind, a mean easterly trade wind anomaly is incorporated into the stochastic parameterization of the wind activity. The combined effect of the nonlinear zonal advection, the enhanced mean easterly trade wind anomaly, and the effective stochastic noise facilitates the intermittent occurrence of the CP El Niño (34). Finally, a three-state Markov jump process is developed to drive the stochastic wind bursts. It emphasizes the distinct properties of the wind activity at different ENSO phases and effective fashion, which is in essence different from previous works (29, 35–37) where the wind burst parameterization includes many detailed structures. This simple stochastic switching process allows the coupled model to simulate different types of ENSO events with realistic features.

Subsequent sections present the coupled ocean–atmosphere system, the stochastic wind burst model, and the Markov jump process as well as the statistical diversity of ENSO simulated by the coupled model. Details of model derivations, choice of parameters, and sensitivity tests are included in SI Appendix.

**Coupled ENSO Model**

**ENSO Model.** The ENSO model consists of a nondissipative atmosphere coupled to a simple shallow-water ocean and SST budget (33, 34).

**Interannual atmosphere model.**

\[
-yv - \partial_x \theta = 0
\]
\[
yu - \partial_y \theta = 0
\]
\[-(\partial_x u + \partial_y v) = E_v/(1 - Q).\]

**Interannual ocean model.**

\[
\partial_t U - c_1 \text{YV} + c_1 \partial_x H = c_1 \tau_x
\]
\[
YU + \partial_y H = 0
\]
\[
\partial_t H + c_1 (\partial_x U + \partial_y V) = 0.
\]

**Interannual SST model.**

\[
\partial_t T + \mu \partial_x (UT) = -c_1 \xi E_v + c_1 \eta H,
\]
with
\[
E_v = \alpha_v T
\]
\[
\tau_x = \tau(u + u_v).
\]

In Eqs. 1–4, \(x\) is zonal direction and \(\tau\) is interannual time, whereas \(y\) and \(Y\) are meridional direction in the atmosphere and ocean, respectively. The \(u, v\) are zonal and meridional winds, \(\theta\) is potential temperature, \(U\) and \(V\) are zonal and meridional currents, \(H\) is thermocline depth, \(T\) is SST, \(E_v\) is latent heating, and \(\tau_x\) is zonal wind stress. All variables are anomalies from an equilibrium state and are nondimensional. The coefficient \(c_1\) is a nondimensional ratio of timescales, which is of order \(O(1)\). The term \(u_v\) in Eq. 4 describes stochastic wind burst activity. The atmosphere extends over the entire equatorial belt \(0 \leq x \leq L_A\) with periodic boundary conditions, whereas the Pacific ocean extends over \(0 \leq x \leq L_O\) with reflection boundary conditions for the ocean model and zero normal derivative at the boundaries for the SST model.

The above model retains a few essential processes that model the ENSO dynamics in a simple fashion. Latent heating \(E_v\), proportional to SST, is depleted from the ocean and forces an atmospheric circulation. The resulting zonal wind stress \(\tau_x\) in return forces an ocean circulation that imposes feedback on the SST through thermocline depth anomalies \(H\). This thermocline feedback \(\eta\) is more significant in the eastern Pacific, as shown in Fig. 1.

The coupled model introduces unique theoretical elements such as a nondissipative atmosphere consistent with the skeleton model for the Madden–Julian oscillation (MJO) (38, 39), valid here on the interannual timescale and suitable to describe the dynamics of the Walker circulation (40–42). In addition, the meridional axes \(y\) and \(Y\) are different in the atmosphere and ocean as they each scale to a suitable Rossby radius. This allows for a systematic meridional decomposition of the system into the well-known parabolic cylinder functions (43), which keeps the system low dimensional (44). Here, models 1–3 are projected and truncated to the first parabolic cylinder function of the atmosphere (38) and the ocean (1), respectively. Details are included in SI Appendix.

The coupled system in Eqs. 1–4 without the nonlinear zonal advection in Eq. 3 was systematically studied in ref. 33. It succeeds in recovering the traditional El Niño with occasionally super El Niño and capturing the ENSO statistics in the eastern Pacific as in nature. If the stochastic wind burst \(u_v\) is further removed, the resulting coupled system is linear, deterministic, and stable. Therefore, such a coupled model is fundamentally different from the Cane–Zebiak model (30) and other nonlinear models (31, 32) where the internal instability rather than the external wind bursts plays the role of triggering the ENSO cycles.

The observational significance of the zonal advection has been shown for the CP El Niño (5, 45). Different from the previous works (46, 47) where the advection is mostly linear and requires ad hoc parameterization of the background SST gradient, a simple nonlinear advection is adopted in Eq. 3 that contributes significantly to the SST tendency. Such nonlinear advection provides the mechanism of transporting anomalous warm water to the central Pacific region by the westward ocean zonal current. Importantly, when stochasticity is included in the wind activity...
$u_p$, this nonlinear zonal advection involves the contribution from both mean and fluctuation, the latter of which is usually ignored in the previous works. The combined effect of this nonlinear advection, a mean easterly trade wind anomaly, and effective stochastic noise was shown to facilitate the intermittent occurrence of the CP El Niño with realistic features (34).

**Stochastic Wind Burst Process.** Stochastic parameterization of the wind activity is added to the model that represents several important ENSO triggers such as westerly wind bursts (WWBs) (28, 29, 48) and easterly wind bursts (EWBs) (49) as well as the convective envelope of the MJO (27, 50). It also includes the recent multidecadal strengthening of the easterly trade wind anomaly. The wind bursts $u_p$ read as

$$u_p = a_p(\tau) s_p(x) \phi_0(y),$$  \hspace{1cm} [5]

with amplitude $a_p(\tau)$ and fixed zonal spatial structure $s_p(x)$ shown in Fig. 1. Here, $\phi_0(y)$ equals the first parabolic cylinder function of the atmosphere (SI Appendix). Both the wind burst perturbations (29) and the strengthening of the trade wind anomaly (12, 13) are localized over the western equatorial Pacific according to the observations and for simplicity they share the same zonal extent.

The evolution of wind burst amplitude $a_p$ reads as

$$\frac{da_p}{d\tau} = -d_p(a_p - \bar{a}_p(T_W)) + \sigma_p(T_W) W(\tau),$$  \hspace{1cm} [6]

where $d_p$ is dissipation and $W(\tau)$ is a white-noise source, representing the intermittent nature of the wind bursts at an interannual timescale. The amplitude of the wind burst noise source $\sigma_p$ depends on $T_W$ (Eq. 7), which is the average of SST anomalies in the western half of the equatorial Pacific ($0 \leq x \leq L_O/2$). Note that this state-dependent wind amplitude has a different viewpoint from those in previous works (37, 51, 52) that rely on the eastern Pacific SST. The term $\bar{a}_p < 0$ represents the mean strengthening of the easterly trade wind anomaly. Corresponding to $\bar{a}_p < 0$, the direct response of the surface wind associated with the Walker circulation at the equatorial Pacific band is shown in Fig. 1C, which is similar to the observed intensification of the Walker circulation in recent decades (12, 13).

---

**Fig. 2.** PDFs of T−3, T−3.4, and T−4 from the coupled model (A–C) and those of Niño 3, 3.4, and 4 from the National Oceanic and Atmospheric Administration (NOAA) observational record (1982/01–2016/02) (E–G). D shows the power spectrum of $T_e$. The model statistics are computed based on a 5,000-y-long simulation.

**Fig. 3.** A sample period-series T−3, T−4, T−3.4; the stochastic wind activity amplitude $a_p$; and the states in the Markov process. The red curve on top of $a_p$ shows the 120-d running average indicating the low-frequency part of the wind activity. The two dashed boxes correspond to the periods shown in Figs. 4 and 5.
A Three-State Markov Jump Process. Due to the fact that the ENSO diversity is associated with the wind activity with distinct features (33, 34), a three-state Markov jump process (53–55) is adopted to model the wind activity. Here, state 2 primarily corresponds to the traditional El Niño and state 1 to the CP El Niño whereas state 0 represents discharge and quiescent phases. The following criteria are used in determining the parameters in Eq. 6 in each state. First, strong wind bursts play an important role in triggering the traditional El Niño (27–29), which suggests a large noise amplitude $\sigma$ in state 2. Second, the observational fact that an enhanced easterly trade wind has accompanied the CP El Niño since the 1990s indicates a negative (easterly) mean $\bar{u}$ in state 1. To obtain the CP El Niño, the amplitude of $\bar{u}$ and the stochastic noise must be balanced (34). This implies a moderate noise amplitude in state 1, which also agrees with observations (56). Finally, only weak wind activity is allowed in the quiescent state and the discharge phase with La Niña (state 0). Thus, the three states are given by

\[
\begin{align*}
\text{State 2:} & \quad \sigma_{p2} = 2.6, \quad d_{p2} = 3.4, \quad \hat{a}_{p2} = -0.25, \\
\text{State 1:} & \quad \sigma_{p1} = 0.8, \quad d_{p1} = 3.4, \quad \hat{a}_{p1} = -0.25, \\
\text{State 0:} & \quad \sigma_{p0} = 0.3, \quad d_{p0} = 3.4, \quad \hat{a}_{p0} = 0,
\end{align*}
\]

respectively, where $d_p = 3.4$ represents a relaxation time around 10 d. Note that the same mean easterly trade wind anomaly as in state 1 is adopted in state 2 due to the fact that both the traditional and the CP El Niño occurred during the last 25 y. Because the amplitude of the stochastic noise dominates the mean easterly wind in state 2, this mean state actually has little impact on simulating the traditional El Niño events. On the other hand, to guarantee no El Niño event occurring in the quiescent phase, no mean trade wind anomaly is imposed in state 0. With such a choice of the parameters, both the amplitude and the timescale of the wind burst activity are similar to those in nature.

The local transition probability from state $i$ to State $j$ with $i \neq j$ for small $\Delta \tau$ is defined as

\[
P(\sigma_p(\tau + \Delta \tau) = \sigma_{pj} | \sigma_p(\tau) = \sigma_{pi}) = \nu_{ij} \Delta \tau + o(\Delta \tau),
\]

and the probability of staying in state $i$ is given by

\[
P(\sigma_p(\tau + \Delta \tau) = \sigma_{pi} | \sigma_p(\tau) = \sigma_{pi}) = 1 - \sum_{j \neq i} \nu_{ij} \Delta \tau + o(\Delta \tau).
\]

Importantly, the transition rates $\nu_{ij}$ (with $i \neq j$) depend on $T_W$, implying the state dependence of the wind bursts (57–59). A transition $\nu_{ij}$ (with $i < j$) from a less active to a more active state is more likely when $T_W \geq 0$ and vice versa. This allows, for example, a rapid shutdown of wind burst activity followed by extreme El Niño events, as in nature.

The transition rates are chosen in accordance with the observational record. A higher transition probability from state 2 to state 0 is adopted compared with that to state 1, representing the situation that the traditional El Niño is usually followed by the La Niña rather than the CP El Niño (e.g., years 1963, 1965, 1972, 1982, 1987, and 1998). Likewise, starting from the quiescent phase, the model has a preference toward the occurrence of the CP El Niño rather than the eastern Pacific super El Niño, as observed in years 1968, 1990, and 2001. See SI Appendix for details. It is also shown in SI Appendix that the model statistics are robust with respect to the perturbation of the parameters, which indicates that a crude estimation of the transition rates is sufficient for obtaining the ENSO diversity with realistic features.
The parameterization of the wind activity in Eqs. 5–9 emphasizes the distinct properties of the wind activity at different ENSO phases and describes effective state-dependent transitions in a simple and effective fashion. It is different from those adopted in previous works that involve many detailed structures, such as the central location and peak time of each wind burst event (29, 35, 36) and the separation of the linear and nonlinear parts of the wind activity (37).

**Results**

Now we present the statistical diversity of ENSO produced by the coupled model. To compare the model simulations with the observational record, we define three indexes from the model simulations: T-3, T-3.4, and T-4, which are the averaged SSTs over the regions of Niño 3 (150W–90W), Niño 3.4 (170W–120W), and Niño 4 (160E–150W), respectively.

**Fig. 2** shows the probability density functions (PDFs) of T-4, T-3.4, and T-3 as well as the power spectrum of T_0, which is the averaged SST over the eastern Pacific. These statistics are formed based on a 5000-y-long model simulation. The power spectrum of T_0 peaks at the interannual band (3–7 y), which is consistent with that in nature (60). The variance of all three T indexes almost perfectly matches those of the three Niño indexes (from the NOAA). Particularly, the fact that the variance of SST in the Niño 4 region is roughly half that in the other two regions is captured by the coupled model. In addition, consistent with observations, the PDFs of T-4 and T-3.4 show negative and positive skewness, respectively. The presence of a fat tail together with the positive skewness in T-3 indicates the extreme El Niño events in the eastern Pacific (61). Note that, despite the correct skewed direction, the skewness of T-3 seems to be underestimated compared with that in Niño 3. However, the Niño indexes are calculated based only on a 34-y-long monthly time series (1982/01–2016/02), which may not be sufficient to form unbiased statistics. In fact, the single super El Niño event during 1997–1998 accounts for a large portion of the skewness in Niño 3. Therefore, the statistics of the model are qualitatively consistent with those in nature. SI Appendix contains sensitivity tests, which show that the model statistics are robust to parameter variations.

These non-Gaussian statistics indicate the significant role of both the nonlinear zonal advection and the state-dependent noise in the wind burst process, because otherwise the PDFs of T indexes will all become Gaussian. In addition, in the absence of the CP El Niño, positive skewness is found in T-4 from the coupled model (33).

**Fig. 3** shows the time series of the three T indexes, the stochastic wind amplitude $u_p$, and the state transitions. Examples of the ENSO diversity are demonstrated in Figs. 4 and 5. First, the coupled model succeeds in simulating traditional El Niño events in the eastern Pacific including both the quasi-regular moderate El Niño (e.g., $t = 383, 392, 397, 406$) and the super El Niño (e.g., $t = 211, 228, 402$). These traditional El Niño events are typically followed by a reversal of conditions toward La Niña with weaker strengths but longer durations. Particularly, the super El Niño appears roughly every 15–25 y (T-3 index), as in nature. These traditional El Niño events start with a realistic buildup of SST and thermocline depth anomalies in the western Pacific in the preceding year, which switches the system to the active state (state 2) and increases wind burst activity over the warm pool region. At the onset phase of these El Niño events, a strong series of WWBs (Fig. 4F) triggers enhanced thermocline depth and SST anomalies in the western Pacific, which then propagate eastward and intensify in the eastern Pacific at the peak of the El Niño events. Meanwhile, the western edge of the warm pool is cooled (62). However, there are many phases (e.g., $t = 400, 405$) where wind burst activity builds up without triggering an El Niño event, implying that wind burst activity in the model is a necessary but nonsufficient condition for the El Niño development (52, 63–65). In Fig. 4, there is also a special event around $t = 231$ that resembles the observational record in year 2014. That is, a strong WWB tends to trigger a traditional El Niño. However, an intense EWB occurs right after the WWB, which serves as the key dynamical factor that stalled the El Niño development (49). On the other hand, in the presence of the mean easterly wind anomaly and moderate wind activity, CP El Niño events are simulated by the coupled model (state 1), where the duration of these CP El Niño events varies from 1–2 y (e.g., $t = 208, 382, 390$) to 4–5 y (e.g., $t = 222–227$). Particularly, the period from $t = 222$ to $t = 230$ resembles the observed ENSO episode during the 1990s, where a series of 5-y CP El Niño events is followed by a super El Niño and then a La Niña.

Budget analysis indicates the distinct mechanisms of the two types of El Niño in the coupled model. The flux divergence $-\mu_i\partial_i (UT)$ appears to be the main contributor to the occurrence of the CP El Niño, where the westward anomalous zonal ocean current tends to transport the anomalous warm water to the central Pacific region and leads to the eastern Pacific cooling with a shallow thermocline depth (Fig. 4H and I from $t = 222$ to $t = 227$). On the other hand, the dominant factor for the traditional El Niño is the thermocline feedback $\gamma_1 \partial H$, which is nearly proportional to the SST in the eastern Pacific and dominates the budget of the SST tendency $dT/dt$ (Fig. 4 C, D, and I around $t = 201$ and $t = 228$). All these findings in the model recover the analysis from the observational data (3, 8–11).

**Fig. 6** shows the anomalous Walker circulation at three different ENSO phases, which all agree with the observational record (5). Particularly, the surface wind $u + u_p$ converges and forms the rising branch of the circulation in the central and eastern Pacific at the CP and the traditional El Niño phases ($t_1$ and $t_2$), respectively. On the other hand, at the La Niña phase $t_3$, a descending branch of the anomalous Walker cell is found in the eastern Pacific, along with the divergence of the surface wind.

**Conclusions**

A simple dynamical model is developed that automatically captures key features and the statistical diversity of ENSO. Systematic strategies are designed for incorporating several major causes of the ENSO diversity into a simple coupled ocean–atmosphere model that is otherwise deterministic, linear, and stable. First, a stochastic parameterization of the wind bursts including both westerly and easterly wind is coupled to the ocean–atmosphere system, where the amplitude of the wind bursts depends on the SST in the western Pacific warm pool. Second, a simple nonlinear zonal advection and a mean easterly trade wind anomaly are incorporated into the coupled system that enables anomalous warm SST in the central Pacific. Finally, an effective three-state stochastic Markov jump process is developed to drive the stochastic wind activity, which allows the occurrence of different types of ENSO.
The statistics produced by the coupled model resemble those in nature. The power spectrum of $T_2$ peaks at the interannual band. The variance of the three T indexes almost perfectly matches those in the observational record of Niño 3, 3.4, and 4 indexes and the directions of the skewness of the T and the Niño indexes are the same in all three regions. Importantly, the fat tail together with the positive skewness in the PDF of the T indexes implies the intermittent occurrence of the extreme El Niño events. The coupled model succeeds in simulating both the CP and the traditional El Niño events, including quasi-regular moderate El Niño and super El Niño, as well as the La Niña with realistic features. Particularly, the model is able to reproduce the observed ENSO episode with diversity in the 1990s. In addition, the anomalous Walker circulations at different ENSO phases all resemble those in nature.

ACKNOWLEDGMENTS. The research of A.J.M. is partially supported by the Office of Naval Research Grant Office of Naval Research, Multidisciplinary University Research Initiative (ONR MURI) N00014-16-1-2161 and the New York University Abu Dhabi Research Institute. N.C. is supported by a postdoctoral fellowship through A.J.M.’s ONR MURI Grant. The authors thank the National Oceanic and Atmospheric Administration for providing Niño indexes.
Simple Stochastic Dynamical Models Capturing the Statistical Diversity of El Niño Southern Oscillation (Supporting Information Appendix)

Nan Chen and Andrew J. Majda

Department of Mathematics, and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012 USA

This SI Appendix consists of three sections. Section 1 involves a brief model derivations and the low-order meridional truncation. The parameter values in the coupled model and the mathematical formulae of the anomalous Walker circulation are also included in this section. Section 2 contains the details of determining the transition rates in the three-state Markov jump process according to the observations. Section 3 includes sensitivity test.

1 Model derivations, meridional truncation and parameter choices

The coupled model considered in this article is derived from a more complicated model that consists of the skeleton model in the atmosphere [1, 2] coupled to a shallow water ocean in the long-wave approximation and a sea surface temperature (SST) budget [3]. Then an asymptotic expansion with respect to a small factor $\epsilon$ that is the ratio of intraseasonal time scale over the interannual one is applied and the result is Eq. (1)-(4) in the article. The details of model derivation are contained in the SI Appendix of [4]. For the convenience of statement, we summarize the coupled model below.

1. Atmosphere model:

$$
- y v - \partial_x \theta = 0, \\
- y u - \partial_y \theta = 0, \\
-(\partial_x u + \partial_y v) = E_q/(1 - \bar{Q}).
$$  (1.1)

2. Ocean model:

$$
\partial_t U - c_1 Y V + c_1 \partial_x H = c_1 \tau_x, \\
Y U + \partial_Y H = 0, \\
\partial_Y H + c_1 (\partial_x U + \partial_Y V) = 0.
$$  (1.2)

3. SST model:

$$
\partial_t T + \mu \partial_x (UT) = -c_1 \zeta E_q + c_1 \eta H.
$$  (1.3)
1.1 Meridional truncation

In order to compute the solutions of the coupled model, we consider the model in its simplest form, which is truncated meridionally to the first parabolic cylinder functions [6].

Different parabolic cylinder functions are utilized in the ocean and atmosphere due to the difference in their deformation radii. The first atmospheric parabolic cylinder function reads \( \phi_0(y) = (\pi)^{-1/4} \exp(-y^2/2) \), and the third one that will be utilized as the reconstruction of solutions reads \( \phi_2 = (4\pi)^{-1/4}(2y^2 - 1) \exp(-y^2/2) \). The oceanic parabolic cylinder functions read \( \psi_m(Y) \), which are identical to the expressions of the atmospheric ones except depending on the \( Y \) axis.

In the atmosphere we assume a truncation of moisture, wave activity and external sources to the first parabolic cylinder function \( \phi_0 \). This is known to excite only the Kelvin and first Rossby atmospheric equatorial waves, of amplitude \( K_A \) and \( R_A \) [1, 2]. In the ocean, we assume a truncation of zonal wind stress forcing to \( \psi_0 \), \( \tau_x = \tau_x \phi_0 \). This is known to excite only the the Kelvin and first Rossby atmospheric oceanic waves, of amplitude \( K_O \) and \( R_O \). Similarly, for the SST model we assume a truncation \( \psi_0 \), \( T = T \phi_0 \). The ENSO model truncated meridionally reads:

1. **Atmosphere model:**

   \[
   \begin{align*}
   \partial_t K_A &= \chi_A (E_q - \langle E_q \rangle)(2 - 2\bar{Q})^{-1}, \\
   -\partial_x R_A / 3 &= \chi_A (E_q - \langle E_q \rangle)(3 - 3\bar{Q})^{-1},
   \end{align*}
   \] (1.4)

2. **Ocean model:**

   \[
   \begin{align*}
   \partial_t K_O + c_1 \partial_x K_O &= \chi_O c_1 \tau_x / 2, \\
   \partial_t R_O - (c_1 / 3) \partial_x R_O &= -\chi_O c_1 \tau_x / 3,
   \end{align*}
   \] (1.5)

3. **SST model:**

   \[
   \partial_t T + \mu \partial_x ((K_O - R_O)T) = -c_1 \zeta E_q + c_1 \eta H,
   \] (1.6)

where \( \chi_A \) and \( \chi_O \) are the projection coefficients from ocean to atmosphere and from atmosphere to ocean, respectively, due to the different extents in their meridional bases. The latent heating is linearized with \( E_q = \alpha_q T \) in the Pacific band and zero outside. Due to the absence of dissipation in the atmosphere, the solvability condition requires a zero equatorial zonal mean of latent heating forcing \( \langle E_q \rangle \) [7, 8]. Note that when meridional truncation is implemented, a projection coefficient \( \chi \approx 0.65 \) appears in front of the nonlinear term [2], which here is absorbed into the nonlinear advection coefficient \( \mu \) for the notation simplicity and the parameter \( \mu \) in the Table below has already taken into account this projection coefficient.

Now instead of solving the coupled system (1.1)–(1.3), we solve the system (1.4)–(1.6). Periodic boundary conditions are adopted for the atmosphere model (1.4). Reflection boundary conditions are adopted for the ocean model (1.5),

\[
K_O(0, t) = r_W R_O(0, t), \quad R_O(L_O, t) = r_E K_O(L_O, t),
\] (1.7)

where \( r_W = 0.5 \) representing partial loss of energy in the west Pacific boundary across Indonesian and Philippine and \( r_E = 0.5 \) representing partial loss of energy due to the north-south propagation of the coast Kelvin waves along the eastern Pacific boundary. Note that \( r_E \) here is different from the one taken in [4] (\( r_E = 1 \)), where a perfect reflection is assumed. For the SST model, no normal derivative at the boundary of \( T \) is adopted, i.e. \( dT/dx = 0 \).
To prevent nonphysical boundary layers in the finite difference method, the coupled model is solved through an upwind scheme, where some details of discretization is included in the SI Appendix of [4, 5]. The total grid points in the ocean and in the atmosphere are \( N_O = 56 \) and \( N_A = 128 \), respectively, which are doubled compared with that in [4] for the purpose of resolving some small scale interactions due to the nonlinearity. The time step is \( \Delta t = 4.25 \) hours. The ratio \( \Delta t/\Delta x \) is approximately 0.115 under the nondimensional values.

The physical variables can be easily reconstructed in the following way.

\[
\begin{align*}
    u &= (K_A - R_A)\phi_0 + (R_A/\sqrt{2})\phi_2, \\
    \theta &= -(K_A + R_A)\phi_0 - (R_A/\sqrt{2})\phi_2, \\
    v &= (4\partial_x R_A - HA - S^\theta)(3\sqrt{2})^{-1}\phi_1, \\
    U &= (K_O - R_O)\psi_0 + (R_O/\sqrt{2})\psi_2, \\
    H &= (K_O + R_O)\psi_0 + (R_O/\sqrt{2})\psi_2.
\end{align*}
\]

(1.8)

See [2, 4, 5] for more details. The variables in (1.8) are utilized in showing the Hovmoller diagrams in Figure 4 and 5 of the main article.

### 1.2 Choices of parameters values

Two tables are included below. Table 1 summarizes the variables in the coupled model and lists the associated units and the typical unit values. Table 2 shows the nondimensional values of the parameters that are utilized in the meridional truncated model (1.4)–(1.6).

<table>
<thead>
<tr>
<th>Variable</th>
<th>unit</th>
<th>unit value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x ) zonal axis</td>
<td>([y]/\delta)</td>
<td>15000km</td>
</tr>
<tr>
<td>( y ) meridional axis atmosphere</td>
<td>( \sqrt{c_A/\beta} )</td>
<td>1500km</td>
</tr>
<tr>
<td>( Y ) meridional axis ocean</td>
<td>( \sqrt{c_O/\beta} )</td>
<td>330km</td>
</tr>
<tr>
<td>( t ) time axis intraseasonal</td>
<td>( 1/\delta\sqrt{c_A\beta} )</td>
<td>3.3 days</td>
</tr>
<tr>
<td>( \tau ) time axis interannual</td>
<td>([t]/\epsilon)</td>
<td>33 days</td>
</tr>
<tr>
<td>( u ) meridional wind speed anomalies</td>
<td>( \delta c_A )</td>
<td>5 m/s⁻¹</td>
</tr>
<tr>
<td>( v ) meridional wind speed anomalies</td>
<td>( \delta [u] )</td>
<td>0.5 m/s⁻¹</td>
</tr>
<tr>
<td>( \theta ) potential temperature anomalies</td>
<td>( 15\delta )</td>
<td>1.5K</td>
</tr>
<tr>
<td>( q ) low-level moisture anomalies</td>
<td>([\theta])</td>
<td>1.5K</td>
</tr>
<tr>
<td>( a ) envelope of synoptic convective activity</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>( \bar{H} a ) convective heating/drying</td>
<td>([\theta]/[t])</td>
<td>0.45 K.day⁻¹</td>
</tr>
<tr>
<td>( E_q ) latent heating anomalies</td>
<td>([\theta]/[t])</td>
<td>0.45 K.day⁻¹</td>
</tr>
<tr>
<td>( T ) sea surface temperature anomalies</td>
<td>([\theta])</td>
<td>1.5K</td>
</tr>
<tr>
<td>( U ) zonal current speed anomalies</td>
<td>( c_O\delta_O )</td>
<td>0.25 m/s⁻¹</td>
</tr>
<tr>
<td>( V ) zonal current speed anomalies</td>
<td>( \delta\sqrt{c[U]} )</td>
<td>0.56 cm/s⁻¹</td>
</tr>
<tr>
<td>( H ) thermocline depth anomalies</td>
<td>( H_O\delta_O )</td>
<td>20.8 m</td>
</tr>
<tr>
<td>( \tau_x ) zonal wind stress anomalies</td>
<td>( \delta\sqrt{\beta/c_A H_O\rho_O c_O^2\delta_O} )</td>
<td>0.00879 N.m⁻²</td>
</tr>
<tr>
<td>( \tau_y ) meridional wind stress anomalies</td>
<td>([\tau_y])</td>
<td>0.00879 N.m⁻²</td>
</tr>
</tbody>
</table>

Table 1: Definitions of model variables and units in the meridional truncated model.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>description</th>
<th>Nondimensional values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>ratio of ocean and atmosphere phase speed</td>
<td>0.05</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Froude number</td>
<td>0.1</td>
</tr>
<tr>
<td>$c_1$</td>
<td>ratio of $c/\epsilon$</td>
<td>0.5</td>
</tr>
<tr>
<td>$\chi_A$</td>
<td>Meridional projection coefficient from ocean to atmosphere</td>
<td>0.32</td>
</tr>
<tr>
<td>$\chi_O$</td>
<td>Meridional projection coefficient from atmosphere to ocean</td>
<td>1.30</td>
</tr>
<tr>
<td>$L_A$</td>
<td>Equatorial belt length</td>
<td>8/3</td>
</tr>
<tr>
<td>$L_O$</td>
<td>Equatorial Pacific length</td>
<td>1.16</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>wind stress coefficient</td>
<td>6.529</td>
</tr>
<tr>
<td>$r_W$</td>
<td>Western boundary reflection coefficient in ocean</td>
<td>0.5</td>
</tr>
<tr>
<td>$r_E$</td>
<td>Eastern boundary reflection coefficient in ocean</td>
<td>0.5</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Latent heating exchange coefficient</td>
<td>8.5</td>
</tr>
<tr>
<td>$\alpha_Q$</td>
<td>mean vertical moisture gradient</td>
<td>0.9</td>
</tr>
<tr>
<td>$\mu$</td>
<td>nonlinear zonal advection coefficient</td>
<td>0.08</td>
</tr>
<tr>
<td>$d_p$</td>
<td>dissipation coefficient in the wind burst model</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Table 2: Nondimensional values of the parameters.

1.3 Anomalous Walker circulation

In the atmospheric model (see [1] for the original version of the skeleton model), only the first baroclinic mode is included in the vertical direction, which has a profile of $\cos(z)$ function. Also recall that the coupled model is projected to the leading parabolic cylinder function in the meridional direction, which has a Gaussian profile that centers at the equator. Thus, the meridional derivative at the equator is $\partial_y \phi_0(y) = 0$ and the mass conservation equation reduces to

$$ \tilde{u}_x(x, z) + \tilde{w}_z(x, z) = 0, \quad (1.9) $$

where $\tilde{u}(x, z)$ and $\tilde{w}(x, z)$ are the zonal and vertical velocities, respectively, which are functions of both $x$ and $z$. Recall that the zonal velocity can be written as [1]

$$ \tilde{u}(x, z) = u(x) \cos(z). \quad (1.10) $$

To satisfy the mass conservation condition (1.9), the vertical velocity is given by

$$ \tilde{w}(x, z) = w(x) \cos(z) = -u_x(x) \sin(z), \quad (1.11) $$

where $w(x) = -u_x(x)$. In the dimensional form (variables with notation $\cdot^D$),

$$ w^D(x) = -\frac{[H_v]}{[L]} u_x^D(x), \quad (1.12) $$

where $[H_v] = 16/\pi \text{km}$ is the vertical length scale and $[L] = 15000 \text{km}$ is the horizontal length scale with nondimensional range $x \in [0, 1.17]$, $z \in [0, \pi]$. The pair $\left( \tilde{u}(x, z), \tilde{w}(x, z) \right)$ forms the anomalous Walker circulation above the equatorial Pacific ocean as shown in Figure 4 of the main article.
2 Details of the transition rates in the three-state Markov jump process

Recall the governing equation of the stochastic wind burst amplitude $a_p$,

$$\frac{da_p}{d\tau} = -d_p(a_p - \hat{a}_p(T_W)) + \sigma_p(T_W)\dot{W}(\tau), \quad (2.1)$$

As discussed in the main article, a three-state Markov jump process is adopted for the parameters in (2.1),

State 2: \quad \sigma_{p2} = 2.6, \quad d_{p2} = 3.4, \quad \hat{a}_{p2} = -0.25, \quad \hat{a}_{p0} = 0, \quad (2.2)

State 1: \quad \sigma_{p1} = 0.8, \quad d_{p1} = 3.4, \quad \hat{a}_{p1} = -0.25, \quad \hat{a}_{p0} = 0, \quad (2.3)

State 0: \quad \sigma_{p0} = 0.3, \quad d_{p0} = 3.4, \quad \hat{a}_{p0} = 0, \quad (2.4)

where State 2 corresponds to the traditional El Niño and State 1 to the CP El Niño while State 0 stands for quiescent phases. We assume all the three states can switch between each other. The detailed forms of the transition rates are shown below, which are functions of $T_W$, the averaged SST over the western Pacific. These transition rates are determined in accordance with the observational facts [13] as will be discussed below.

- The transition rates from State 2 to State 1 and from State 2 to State 0 are given by respectively

\[
\nu_{21} = \frac{1}{10} \cdot \frac{1 - \tanh(2T_W)}{4}, \quad (2.5)
\]
\[
\nu_{20} = \frac{9}{10} \cdot \frac{1 - \tanh(2T_W)}{4}. \quad (2.6)
\]

Starting from State 2, the probability of switching to State 0 is much higher than that to State 1. This comes from the fact that a traditional El Niño is usually followed by a La Niña rather than a CP El Niño (e.g., year 1963, 1965, 1972, 1982 and 1998). Typically, the La Niña event has a weaker amplitude and a longer duration compared with the preceding El Niño. This actually corresponds to a discharge phase of the ENSO cycle with no external wind bursts (State 0).

- The transition rates from State 1 to State 0 and from State 1 to State 2 are given by respectively

\[
\nu_{10} = \frac{1 - \tanh(2T_W)}{12}, \quad (2.7)
\]
\[
\nu_{12} = \frac{1 + \tanh(2T_W)}{40}. \quad (2.8)
\]

Although the denominator of $\nu_{10}$ is smaller than that of $\nu_{12}$, quite a few CP El Niño events are associated with a slight positive $T_W$ in the model, which means the transition rate $\nu_{12}$ is not necessarily smaller than $\nu_{10}$. In fact, with the transition rates given by (2.7)–(2.8), the results show in the main article that more events are transited from state 1 to 2 than from state 1 to 0. This is consistent with the observations (e.g., year 1981 and 1995), implying that the CP El Niño is more likely to be followed by the classical El Niño than the quiescent phase.
The transition rates from State 0 to State 1 and State 2 are given by

\[ \nu_{01} = \frac{2}{3} \cdot \frac{1 + \tanh(2T_W)}{7}, \quad (2.9) \]

\[ \nu_{02} = \frac{1}{3} \cdot \frac{1 + \tanh(2T_W)}{7}. \quad (2.10) \]

Again, the transition rates to State 1 and 2 are different. This is due to the fact that after a quiescent period or discharge La Niña phase, more events are prone to become CP El Niño as an intermediate transition instead of directly forming another traditional El Niño (e.g., year 1969, 1977, 1990 and 2002).

Note that in (2.5)–(2.10), the transition rate \( \nu_{ij} \) from a more active state to a less active state (with \( i > j \)) is always proportional to \( 1 - \tanh(2T_W) \) while that from a less active state to a more active state (with \( i < j \)) is always proportional to \( 1 + \tanh(2T_W) \). These are consistent with the fact stated in the main article that a transition from a less active to a more active state is more likely when \( T_W \geq 0 \) and vice versa.

### 3 Sensitivity test

With the optimal parameters of the transition rates shown in (2.5)–(2.10), the variance of the three T indices almost perfectly match those of the Nino indices and the non-Gaussian statistical characteristics in different Nino regions are recovered. Since most of the general circulation models tend to be sensitive to parameter perturbations [14, 15, 16], it is important to test the robustness of the coupled model studied in the main article. To this end, some perturbations are added to the transition rates and the statistics with the suboptimal rates are shown in the following. In each of the panel below, the variable with an asterisk stands for the optimal value given by (2.5)–(2.10). The maximum perturbation of each transition rate is \( \pm 25\% \).

In Figure S1–S3, the variance, skewness and kurtosis of T-3, T-3.4 and T-4 are shown as functions of perturbed transition rates. It is clear that all these statistics are fairly robust with respect to the parameter perturbations, where the variance of different T indices remain nearly the same as those in nature and the non-Gaussian features are still significant. The only parameter that is slightly sensitive is the transition rate from State 1 to State 2, i.e., \( \nu_{12} \). In fact, an underestimated \( \nu_{12} \) modifies not only the frequency of both the CP and traditional El Niño but also the development of the nonlinear advective mode [5], which affect the PDFs in different Nino regions.

Other sensitivity tests with respect to the parameter perturbations in the coupled ocean-atmosphere were shown in the previous study [4]. The results there also indicate the robustness of the coupled model.
Figure S1: Sensitivity test. The variance of the three T indices as functions of suboptimal transition rates, where the variable with asterisk stands for the optimal value in (2.5)–(2.10).
Figure S2: Sensitivity test. The skewness of the three T indices as functions of suboptimal transition rates, where the variable with asterisk stands for the optimal value in (2.5)–(2.10).
Figure S3: Sensitivity test. The kurtosis of the three T indices as functions of suboptimal transition rates, where the variable with asterisk stands for the optimal value in (2.5)–(2.10).
References


