

Supplementary Information for “Understanding ENSO asymmetry with a reduced stochastic model: deterministic nonlinearity and multiplicative noise”

Cristian Martinez-Villalobos, Boris Dewitte, Soong-Ki Kim, Malte F. Stuecker

Climate Dynamics

Corresponding author: Cristian Martinez-Villalobos

Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago, Chile

Data Observatory Foundation, Santiago, Chile

`cristian.martinez.v@uai.cl`

Supplementary Text 1: Adiabatic elimination of the wind variable

Here we derive the effective one-dimensional stochastic equation for the Niño 3 SST index T used in the main text. The starting point is the two-variable stochastic system for Niño 3 SST anomalies T and Niño 4 zonal wind anomalies U ,

$$\frac{dT}{dt} = A(T) + \gamma_U U + \gamma_T U T + B_{11}\eta_1(t) + B_{12}\eta_2(t), \quad (1)$$

$$\frac{dU}{dt} = \beta_1 T + \beta_2 U + B_{21}\eta_1(t) + B_{22}\eta_2(t). \quad (2)$$

The processes $\eta_1(t)$ and $\eta_2(t)$ are independent Gaussian white-noise processes satisfying

$$\langle \eta_i(t) \rangle = 0, \quad (3)$$

$$\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t'). \quad (4)$$

Equivalently, $\eta_i(t) = dW_i/dt$, where W_i are independent Wiener processes. The function $A(T)$ denotes the local SST tendency in the T equation. In the linear model,

$$A(T) = \alpha_1^L T, \quad (5)$$

whereas in the nonlinear model,

$$A(T) = \alpha_0^N + \alpha_1^N T + \alpha_2^N T^2 + \alpha_3^N T^3. \quad (6)$$

The empirical system exhibits a separation of timescales between the slowly evolving SST index T and the more rapidly adjusting zonal wind index U . In the observations, T decorrelates on a timescale of approximately one year, whereas U decorrelates on a timescale of order two weeks. This motivates treating U as a fast variable relative to T . The approximation used here is a leading-order quasi-steady, or adiabatic, elimination of U : on the slow timescale over which T evolves, the fast wind variable is assumed to remain close to its instantaneous stochastic balance[1, 2, 3, 4].

To express this idea, consider first the deterministic part of the U equation with T held fixed. The fast equation relaxes toward the conditional equilibrium

$$U_{\text{eq}}(T) = -\frac{\beta_1}{\beta_2}T, \quad (7)$$

assuming $\beta_2 < 0$, as in the fitted system. Thus, the deterministic component of the wind response is approximately proportional to the concurrent SST anomaly. The stochastic residual tendency in the U equation then perturbs this balance. Retaining the leading-order stochastic balance in Eq. (2) gives

$$0 \simeq \beta_1 T + \beta_2 U + B_{21}\eta_1(t) + B_{22}\eta_2(t). \quad (8)$$

Solving Eq. (8) for U yields

$$U \simeq -\frac{\beta_1}{\beta_2}T - \frac{B_{21}}{\beta_2}\eta_1(t) - \frac{B_{22}}{\beta_2}\eta_2(t). \quad (9)$$

This expression should be interpreted as the leading-order quasi-steady diagnostic response of the fast wind variable to the slowly varying SST anomaly and to the fast residual forcing. It is not meant as a full homogenization result for the finite-variance fast Ornstein–Uhlenbeck process. The reduction is performed on the daily-resolved equations; temporal coarse-graining to monthly resolution would define a different reduced model. Rather, it is the algebraic fast-variable balance used to project the wind tendency and its residual stochastic forcing onto the slower SST equation[1, 2, 3].

Substituting Eq. (9) into Eq. (1) gives

$$\begin{aligned} \frac{dT}{dt} &= A(T) + \gamma_U U + \gamma_T U T + B_{11}\eta_1(t) + B_{12}\eta_2(t) \\ &= A(T) - \frac{\gamma_U \beta_1}{\beta_2}T - \frac{\gamma_T \beta_1}{\beta_2}T^2 \\ &\quad + \left(B_{11} - \frac{\gamma_U B_{21}}{\beta_2} - \frac{\gamma_T B_{21}}{\beta_2}T \right) \eta_1(t) \\ &\quad + \left(B_{12} - \frac{\gamma_U B_{22}}{\beta_2} - \frac{\gamma_T B_{22}}{\beta_2}T \right) \eta_2(t). \end{aligned} \quad (10)$$

42 Defining

$$43 \quad G_1 = B_{11} - \frac{\gamma_U B_{21}}{\beta_2}, \quad E_1 = -\frac{\gamma_T B_{21}}{\beta_2}, \quad (11)$$

$$44 \quad G_2 = B_{12} - \frac{\gamma_U B_{22}}{\beta_2}, \quad E_2 = -\frac{\gamma_T B_{22}}{\beta_2}, \quad (12)$$

45 the reduced equation can be written as

$$46 \quad \frac{dT}{dt} = A_{\text{qs}}(T) + (G_1 + E_1 T) \eta_1(t) + (G_2 + E_2 T) \eta_2(t), \quad (13)$$

47 where the quasi-steady deterministic drift is

$$48 \quad A_{\text{qs}}(T) = A(T) - \frac{\gamma_U \beta_1}{\beta_2} T - \frac{\gamma_T \beta_1}{\beta_2} T^2. \quad (14)$$

49 Equation (13) shows why state-dependent stochastic forcing appears in the reduced SST equation.
 50 If $\gamma_T = 0$, then $E_1 = E_2 = 0$, and the stochastic forcing remains purely additive. This is the case for
 51 the linear two-variable model. If $\gamma_T \neq 0$, however, the wind-SST interaction term $\gamma_T UT$ maps the
 52 fast stochastic wind residuals into noise amplitudes proportional to T . The reduced one-dimensional
 53 equation contains correlated additive-multiplicative noise[5, 6, 7].

54 The final step is to express the reduced equation in the Itô convention used for the Fokker-
 55 Planck analysis in the main text. The quasi-steady substitution above treats the fast forcing as the
 56 white-noise limit of a rapidly varying process. When the resulting state-dependent noise amplitudes
 57 are written in equivalent Itô form, the drift contains the standard noise-induced correction[4, 8, 5, 9]

$$58 \quad \frac{1}{2} \sum_{i=1}^2 \frac{d}{dT} (G_i + E_i T) (G_i + E_i T). \quad (15)$$

59 Thus,

$$60 \quad \frac{1}{2} \sum_{i=1}^2 \frac{d}{dT} (G_i + E_i T) (G_i + E_i T) = \frac{1}{2} [E_1 (G_1 + E_1 T) + E_2 (G_2 + E_2 T)] \\ 61 \quad = \frac{1}{2} (E_1 G_1 + E_2 G_2) + \frac{1}{2} (E_1^2 + E_2^2) T. \quad (16)$$

62 Including this term, the effective Itô drift is

$$63 \quad A_{\text{eff}}(T) = A_{\text{qs}}(T) + \frac{1}{2} (E_1 G_1 + E_2 G_2) + \frac{1}{2} (E_1^2 + E_2^2) T. \quad (17)$$

64 The final reduced stochastic equation is

$$65 \quad \frac{dT}{dt} = A_{\text{eff}}(T) + (G_1 + E_1 T) \eta_1(t) + (G_2 + E_2 T) \eta_2(t). \quad (18)$$

For the linear model, $\gamma_T = 0$, so $E_1 = E_2 = 0$, and Eq. (18) becomes

$$\frac{dT}{dt} = \tilde{\alpha}_1^L T + G_{1L} \eta_1(t) + G_{2L} \eta_2(t), \quad (19)$$

with

$$\tilde{\alpha}_1^L = \alpha_1^L - \frac{\beta_1}{\beta_2} \gamma_U^L. \quad (20)$$

Thus, the linear reduction produces an Ornstein–Uhlenbeck process with additive noise.

For the nonlinear model, the effective drift remains cubic,

$$A_{\text{eff}}^N(T) = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3, \quad (21)$$

where

$$\tilde{\alpha}_0^N = \alpha_0^N + \frac{1}{2} (E_1 G_1 + E_2 G_2), \quad (22)$$

$$\tilde{\alpha}_1^N = \alpha_1^N - \frac{\beta_1}{\beta_2} \gamma_U^N + \frac{1}{2} (E_1^2 + E_2^2), \quad (23)$$

$$\tilde{\alpha}_2^N = \alpha_2^N - \frac{\beta_1}{\beta_2} \gamma_T^N, \quad (24)$$

$$\tilde{\alpha}_3^N = \alpha_3^N. \quad (25)$$

The reduced nonlinear equation is

$$\frac{dT}{dt} = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3 + (G_{1N} + E_1 T) \eta_1(t) + (G_{2N} + E_2 T) \eta_2(t). \quad (26)$$

This expression is the one-dimensional stochastic SST equation used in the main text. It shows that eliminating the fast wind variable modifies the deterministic SST damping and, when $\gamma_T \neq 0$, generates state-dependent stochastic forcing. Thus, the same wind-SST interaction term that contributes to nonlinear deterministic damping also produces correlated additive-multiplicative noise in the reduced SST equation.

Supplementary Text 2: Sensitivity to including an explicit SSH variable

The main analysis uses a deliberately minimal two-variable stochastic model based on the Niño 3 SST index (T) and the Niño 4 zonal wind index (U). This choice is intended to isolate the coupled SST–wind mechanism relevant for the daily evolution of eastern equatorial Pacific SST anomalies. ENSO dynamics also involve upper-ocean adjustment and thermocline variations, which are often represented in conceptual models through a thermocline-depth, warm-water-volume, or sea-surface-height-like variable. Here we test whether adding an explicit upper-ocean proxy materially changes the statistical behavior relevant to our analysis.

To do this, we repeated the empirical modeling using an extended three-variable system that includes Niño 3 sea surface height anomalies, denoted here by h , as a proxy for thermocline and upper-ocean heat-content variability. The three-variable model is fitted over the 1993–2017 period, for which the SSH product used here is available. The state vector is $\mathbf{x} = (T, U, h)$, where T is the Niño 3 SST anomaly, U is the Niño 4 zonal wind anomaly, and h is the Niño 3 SSH anomaly. As in the main analysis, we compare a linear model with a nonlinear model. The nonlinear formulation retains nonlinear terms in the T equation and also allows nonlinear terms in the SSH tendency, while the U equation is kept linear. This extended formulation is not used to derive an analytical one-dimensional stationary distribution; rather, it is used to assess whether the explicit inclusion of SSH changes the qualitative interpretation obtained from the simpler SST–wind system.

Figure S1 compares the observed probability density functions of (T), (U), and (h) with those generated by the linear and nonlinear three-variable models. The linear model reproduces the bulk variability of the variables but remains close to Gaussian and does not capture the observed positive skewness of either Niño 3 SST or Niño 3 SSH. In contrast, the nonlinear model produces a clearly asymmetric Niño 3 SST distribution and substantially improves the representation of the positively skewed SSH distribution. The improvement for Niño 4 zonal wind is more modest, consistent with the fact that the wind equation is kept linear in this extended formulation.

Including Niño 3 SSH as an additional state variable does not qualitatively change the main statistical interpretation. For the purpose of diagnosing the origin of Niño 3 non-Gaussianity in this reduced daily framework, the dominant contrast remains the transition from a linear to a nonlinear deterministic formulation. For the statistical diagnostics examined here, explicitly retaining SSH does not materially change the dominant linear-versus-nonlinear comparison pursued in the main text, while a dynamically more complete representation of ENSO would retain explicit upper-ocean adjustment.

Supplementary Text 3: Stationary first-moment consistency of the reduced nonlinear drift

Here we provide a simple consistency check on the effective nonlinear drift obtained after adiabatic elimination of the wind variable. For a stationary Itô stochastic differential equation with drift $A_{\text{eff}}(T)$, the first-moment equation is

$$0 = \langle A_{\text{eff}}(T) \rangle. \quad (27)$$

For the reduced nonlinear model derived in Supplementary Text 1, the effective drift is

$$A_{\text{eff}}^N(T) = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3. \quad (28)$$

Substituting Eq. (28) into Eq. (27) gives

$$0 = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N \langle T \rangle + \tilde{\alpha}_2^N \langle T^2 \rangle + \tilde{\alpha}_3^N \langle T^3 \rangle. \quad (29)$$

Because the Niño 3 SST anomaly is approximately centered, $\langle T \rangle \simeq 0$, this reduces to

$$0 \simeq \tilde{\alpha}_0^N + \tilde{\alpha}_2^N \langle T^2 \rangle + \tilde{\alpha}_3^N \langle T^3 \rangle. \quad (30)$$

Equivalently,

$$\tilde{\alpha}_0^N \simeq -\tilde{\alpha}_2^N \langle T^2 \rangle - \tilde{\alpha}_3^N \langle T^3 \rangle. \quad (31)$$

Using the fitted coefficients and observed moments, this balance is satisfied to a good approximation:

$$\tilde{\alpha}_0^N \approx -9.6 \times 10^{-4}, \quad \tilde{\alpha}_2^N \langle T^2 \rangle + \tilde{\alpha}_3^N \langle T^3 \rangle \approx 9.5 \times 10^{-4}. \quad (32)$$

This agreement provides a useful consistency check that the fitted effective drift is compatible with an approximately stationary, zero-mean SST anomaly distribution.

Supplementary Text 4: Derivation of stationary distributions

Here we derive the stationary distributions $p_s(T)$ associated with the effective one-dimensional stochastic equations for the Niño 3 SST index T . These equations are obtained after eliminating the fast wind variable U . We write the general equation as

$$\frac{dT}{dt} = A(T) + S\eta_1(t) + (G + ET)\eta_2(t), \quad (33)$$

where $\eta_1(t)$ and $\eta_2(t)$ are independent Gaussian white-noise processes. The total noise variance is

$$B^2(T) = S^2 + (G + ET)^2. \quad (34)$$

The corresponding Fokker–Planck equation is [4, 5, 8, 6]

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial T} [A(T)p] + \frac{1}{2} \frac{\partial^2}{\partial T^2} [B^2(T)p]. \quad (35)$$

For a stationary solution with zero probability flux, this gives

$$p_s(T) = \frac{N}{B^2(T)} \exp \left[\int^T \frac{2A(T')}{B^2(T')} dT' \right], \quad (36)$$

where N is a normalization constant.

Change of variables for CAM noise

For the CAM-noise cases, the diffusion coefficient contains a linear dependence on T . We use the change of variables

$$u = \frac{G + ET}{S}, \quad (37)$$

which implies

$$T = \frac{Su - G}{E}, \quad dT = \frac{S}{E} du. \quad (38)$$

This substitution reduces integrals of the form

$$\int \frac{T^n}{S^2 + (G + ET)^2} dT \quad (39)$$

to rational functions of u . The resulting primitive functions involve polynomial terms, $\tan^{-1}(u)$, and $\log(1 + u^2)$. The prefactor $1/B^2(T)$ in Eq. (36) contributes an additional power of $1 + u^2$, which is included in the exponents below.

Nonlinear+CAM noise model

For the Nonlinear+CAM model, the drift is

$$A_{\text{eff}}^N(T) = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3, \quad (40)$$

and the stochastic forcing is

$$S_1 \eta_1(t) + (G_{2N} + E_2 T) \eta_2(t). \quad (41)$$

Thus, in Eq. (33), $S = S_1$, $G = G_{2N}$, and $E = E_2$. Substituting Eq. (40) into Eq. (36) gives

$$p_s(T) = N_1 \exp(\delta_0^N T) \exp(\delta_1^N T^2) \exp\left[\delta_2^N \tan^{-1}\left(\frac{G_{2N} + E_2 T}{S_1}\right)\right] \\ \times \left[1 + \left(\frac{G_{2N} + E_2 T}{S_1}\right)^2\right]^{\delta_3^N}, \quad (42)$$

where N_1 is chosen so that the distribution integrates to unity, and

$$\delta_0^N = \frac{2\tilde{\alpha}_2^N}{E_2^2} - \frac{4G_{2N}\tilde{\alpha}_3^N}{E_2^3}, \quad (43)$$

$$\delta_1^N = \frac{\tilde{\alpha}_3^N}{E_2^2}, \quad (44)$$

$$\delta_2^N = \frac{2\tilde{\alpha}_0^N}{S_1 E_2} - \frac{2\tilde{\alpha}_1^N G_{2N}}{S_1 E_2^2} + \frac{2\tilde{\alpha}_2^N (G_{2N}^2 - S_1^2)}{S_1 E_2^3} - \frac{2\tilde{\alpha}_3^N G_{2N} (G_{2N}^2 - 3S_1^2)}{S_1 E_2^4}, \quad (45)$$

$$\delta_3^N = \frac{\tilde{\alpha}_1^N}{E_2^2} - \frac{2\tilde{\alpha}_2^N G_{2N}}{E_2^3} + \frac{\tilde{\alpha}_3^N (3G_{2N}^2 - S_1^2)}{E_2^4} - 1. \quad (46)$$

The final -1 in Eq. (46) comes from the prefactor $1/B^2(T)$ in Eq. (36).

Nonlinear+additive noise model

For the Nonlinear+additive model, the deterministic drift is again given by Eq. (40), but the stochastic forcing is additive:

$$S_1 \eta_1(t) + S_2 \eta_2(t). \quad (47)$$

The total noise variance is constant,

$$D_N = S_1^2 + S_2^2. \quad (48)$$

Equation (36) then gives

$$p_s(T) = \frac{2N_2}{D_N} \exp(\kappa_0^N T) \exp(\kappa_1^N T^2) \exp(\kappa_2^N T^3) \exp(\kappa_3^N T^4), \quad (49)$$

181 where N_2 is a normalization constant and

$$182 \quad \kappa_0^N = \frac{2\tilde{\alpha}_0^N}{D_N}, \quad (50)$$

$$183 \quad \kappa_1^N = \frac{\tilde{\alpha}_1^N}{D_N}, \quad (51)$$

$$184 \quad \kappa_2^N = \frac{2\tilde{\alpha}_2^N}{3D_N}, \quad (52)$$

$$185 \quad \kappa_3^N = \frac{\tilde{\alpha}_3^N}{2D_N}. \quad (53)$$

186 The factor $2N_2/D_N$ is a normalization prefactor; any constant factors can be absorbed into N_2 .

187 **Linear+additive noise model**

188 For the Linear+additive model,

$$189 \quad \frac{dT}{dt} = \tilde{\alpha}_1^L T + G_{1L}\eta_1(t) + G_{2L}\eta_2(t). \quad (54)$$

190 The total noise variance is constant,

$$191 \quad D_L = G_{1L}^2 + G_{2L}^2. \quad (55)$$

192 The stationary solution is Gaussian:

$$193 \quad p_s(T) = \frac{2N_3}{D_L} \exp\left(\kappa_1^L T^2\right), \quad (56)$$

194 with

$$195 \quad \kappa_1^L = \frac{\tilde{\alpha}_1^L}{D_L}. \quad (57)$$

196 Since $\tilde{\alpha}_1^L < 0$, this is a normalizable Gaussian distribution.

197 **Linear+CAM noise model**

198 For the Linear+CAM model, the deterministic drift is linear but the stochastic forcing includes the
199 same state-dependent component used in the Nonlinear+CAM model:

$$200 \quad \frac{dT}{dt} = \tilde{\alpha}_1^L T + G_{1L}\eta_1(t) + (G_{2N} + E_2 T)\eta_2(t). \quad (58)$$

201 Thus, in Eq. (33), $S = G_{1L}$, $G = G_{2N}$, and $E = E_2$. Using Eq. (36), the stationary solution is

$$202 \quad p_s(T) = N_4 \exp\left[\delta_2^L \tan^{-1}\left(\frac{G_{2N} + E_2 T}{G_{1L}}\right)\right] \left[1 + \left(\frac{G_{2N} + E_2 T}{G_{1L}}\right)^2\right]^{\delta_3^L}, \quad (59)$$

203 where N_4 is a normalization constant and

$$204 \quad \delta_2^L = -\frac{2\tilde{\alpha}_1^L G_{2N}}{G_{1L} E_2^2}, \quad (60)$$

$$205 \quad \delta_3^L = \frac{\tilde{\alpha}_1^L}{E_2^2} - 1. \quad (61)$$

206 Again, the final -1 in Eq. (61) comes from the prefactor $1/B^2(T)$ in the zero-flux stationary solution.

207 The Linear+additive and Nonlinear+additive stationary distributions are recovered as the
 208 regular $E_2 \rightarrow 0$ limits of the corresponding CAM solutions. These limits are derived explicitly in
 209 Supplementary Text 5.

Supplementary Text 5: Additive limits of the CAM stationary distributions

Here we show that the stationary probability density functions of the CAM-noise models reduce to additive-noise stationary distributions in the limit $E_2 \rightarrow 0$. The derivation is perturbative in E_2 and focuses on the cancellation of the singular terms in the stationary exponent.

Linear+CAM limit

We begin from the Linear+CAM stationary distribution,

$$p_s(T) = N_4 \exp \left[-\frac{2\tilde{\alpha}_1^L G_{2N}}{G_{1L} E_2^2} \tan^{-1} \left(\frac{G_{2N} + E_2 T}{G_{1L}} \right) \right] \left[1 + \left(\frac{G_{2N} + E_2 T}{G_{1L}} \right)^2 \right]^{\tilde{\alpha}_1^L / E_2^2 - 1}. \quad (62)$$

Define

$$\alpha = \tilde{\alpha}_1^L, \quad B = G_{1L}, \quad G = G_{2N}, \quad E = E_2, \quad (63)$$

with $\alpha < 0$. Then Eq. (62) becomes

$$p_s(T) = N_4 \exp \left[-\frac{2\alpha G}{B E^2} \tan^{-1} \left(\frac{G + E T}{B} \right) \right] \left[1 + \left(\frac{G + E T}{B} \right)^2 \right]^{\alpha / E^2 - 1}. \quad (64)$$

Taking logarithms and collecting all T -independent terms into a constant $C(E)$ gives

$$\begin{aligned} \log p_s(T) = C(E) &- \frac{2\alpha G}{B E^2} \tan^{-1} \left(\frac{G + E T}{B} \right) \\ &+ \left(\frac{\alpha}{E^2} - 1 \right) \log \left[1 + \left(\frac{G + E T}{B} \right)^2 \right]. \end{aligned} \quad (65)$$

Let $\epsilon = E T$ and $D_0 = B^2 + G^2$. For fixed T as $E \rightarrow 0$,

$$\log \left[1 + \left(\frac{G + \epsilon}{B} \right)^2 \right] = \log \left(\frac{D_0}{B^2} \right) + \frac{2G}{D_0} \epsilon + \frac{B^2 - G^2}{D_0^2} \epsilon^2 + O(\epsilon^3), \quad (66)$$

$$\tan^{-1} \left(\frac{G + \epsilon}{B} \right) = \tan^{-1} \left(\frac{G}{B} \right) + \frac{B}{D_0} \epsilon - \frac{B G}{D_0^2} \epsilon^2 + O(\epsilon^3). \quad (67)$$

Substituting Eqs. (66)–(67) into Eq. (65), the terms proportional to ϵ/E^2 cancel:

$$-\frac{2\alpha G}{D_0 E^2} \epsilon + \frac{2\alpha G}{D_0 E^2} \epsilon = 0. \quad (68)$$

The remaining linear term is $O(E)$, while the leading quadratic contribution is

$$\frac{2\alpha G^2}{D_0^2 E^2} \epsilon^2 + \frac{\alpha(B^2 - G^2)}{D_0^2 E^2} \epsilon^2 = \frac{\alpha}{D_0} T^2. \quad (69)$$

Combining these terms,

$$\log p_s(T) = C(E) + \frac{\alpha}{D_0} T^2 + O(E), \quad (70)$$

where $C(E)$ is independent of T . After normalization, the limiting stationary distribution is

$$p_s(T) \longrightarrow \sqrt{\frac{-\alpha}{\pi D_0}} \exp\left(\frac{\alpha}{D_0} T^2\right). \quad (71)$$

Returning to the original notation,

$$p_s(T) \longrightarrow \sqrt{\frac{-\tilde{\alpha}_1^L}{\pi (G_{1L}^2 + G_{2N}^2)}} \exp\left[\frac{\tilde{\alpha}_1^L}{G_{1L}^2 + G_{2N}^2} T^2\right]. \quad (72)$$

This is the additive-noise limit of the Linear+CAM benchmark. If the additive component in the CAM benchmark is identified with G_{2L} rather than G_{2N} , Eq. (72) becomes the Linear+additive stationary distribution used in the main text.

Nonlinear+CAM limit

We now consider the Nonlinear+CAM model with drift

$$A(T) = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3, \quad \tilde{\alpha}_3^N < 0, \quad (73)$$

and diffusion coefficient

$$D_E(T) = S_1^2 + (G + ET)^2, \quad (74)$$

where $G = G_{2N}$ and $E = E_2$. For compactness, define

$$a_0 = \tilde{\alpha}_0^N, \quad a_1 = \tilde{\alpha}_1^N, \quad a_2 = \tilde{\alpha}_2^N, \quad a_3 = \tilde{\alpha}_3^N, \quad (75)$$

and

$$D_0 = S_1^2 + G^2. \quad (76)$$

The Nonlinear+CAM stationary distribution can be written as

$$p_s(T) = N \exp(\gamma_0 T) \exp(\gamma_1 T^2) \exp\left[\gamma_2 \tan^{-1}\left(\frac{G + ET}{S_1}\right)\right] \\ \times \left[1 + \left(\frac{G + ET}{S_1}\right)^2\right]^{\gamma_3}, \quad (77)$$

with

$$\gamma_0 = \frac{2a_2}{E^2} - \frac{4Ga_3}{E^3}, \quad (78)$$

$$\gamma_1 = \frac{a_3}{E^2}, \quad (79)$$

$$\gamma_2 = \frac{2a_0}{S_1 E} - \frac{2a_1 G}{S_1 E^2} + \frac{2a_2 (G^2 - S_1^2)}{S_1 E^3} - \frac{2a_3 G (G^2 - 3S_1^2)}{S_1 E^4}, \quad (80)$$

$$\gamma_3 = \frac{a_1}{E^2} - \frac{2a_2 G}{E^3} + \frac{a_3 (3G^2 - S_1^2)}{E^4} - 1. \quad (81)$$

Let $\epsilon = ET$. For fixed T as $E \rightarrow 0$, we use

$$\begin{aligned} \log \left[1 + \left(\frac{G + \epsilon}{S_1} \right)^2 \right] &= \log \left(\frac{D_0}{S_1^2} \right) + \frac{2G}{D_0} \epsilon + \frac{S_1^2 - G^2}{D_0^2} \epsilon^2 \\ &\quad - \frac{2G (3S_1^2 - G^2)}{3D_0^3} \epsilon^3 - \frac{S_1^4 - 6S_1^2 G^2 + G^4}{2D_0^4} \epsilon^4 + O(\epsilon^5), \end{aligned} \quad (82)$$

$$\begin{aligned} \tan^{-1} \left(\frac{G + \epsilon}{S_1} \right) &= \tan^{-1} \left(\frac{G}{S_1} \right) + \frac{S_1}{D_0} \epsilon - \frac{S_1 G}{D_0^2} \epsilon^2 \\ &\quad + \frac{S_1 (3G^2 - S_1^2)}{3D_0^3} \epsilon^3 + \frac{S_1 G (S_1^2 - G^2)}{D_0^4} \epsilon^4 + O(\epsilon^5). \end{aligned} \quad (83)$$

Substituting Eqs. (82)–(83) into $\log p_s(T)$ and collecting powers of T gives

$$\log p_s(T) = C(E) + c_1(E)T + c_2(E)T^2 + c_3(E)T^3 + c_4(E)T^4 + O(E|T|^5), \quad (84)$$

where $C(E)$ is independent of T . The singular powers of E^{-1} cancel, leaving

$$c_1(E) = \frac{2a_0}{D_0} - \frac{2GE}{D_0}, \quad (85)$$

$$c_2(E) = \frac{a_1}{D_0} - \frac{2a_0 GE}{D_0^2} - \frac{(S_1^2 - G^2) E^2}{D_0^2}, \quad (86)$$

$$c_3(E) = \frac{2a_2}{3D_0} + O(E), \quad (87)$$

$$c_4(E) = \frac{a_3}{2D_0} + O(E). \quad (88)$$

Thus, as $E \rightarrow 0$,

$$c_1(E) \rightarrow \frac{2a_0}{D_0}, \quad c_2(E) \rightarrow \frac{a_1}{D_0}, \quad c_3(E) \rightarrow \frac{2a_2}{3D_0}, \quad c_4(E) \rightarrow \frac{a_3}{2D_0}. \quad (89)$$

In the variance-matched additive formulation, the additive amplitude S_2 is defined so that

$$S_2^2 = G^2 + E^2 \langle T^2 \rangle. \quad (90)$$

Therefore, in the limit $E \rightarrow 0$,

$$D_0 = S_1^2 + G^2 \longrightarrow S_1^2 + S_2^2. \quad (91)$$

The limiting exponent is

$$P(T) = \kappa_0^N T + \kappa_1^N T^2 + \kappa_2^N T^3 + \kappa_3^N T^4, \quad (92)$$

with

$$\kappa_0^N = \frac{2\tilde{\alpha}_0^N}{S_1^2 + S_2^2}, \quad \kappa_1^N = \frac{\tilde{\alpha}_1^N}{S_1^2 + S_2^2}, \quad (93)$$

and

$$\kappa_2^N = \frac{2\tilde{\alpha}_2^N}{3(S_1^2 + S_2^2)}, \quad \kappa_3^N = \frac{\tilde{\alpha}_3^N}{2(S_1^2 + S_2^2)}. \quad (94)$$

Since $\tilde{\alpha}_3^N < 0$, the limiting additive-noise density is normalizable. The convergence of the normalized density then follows from the convergence of the exponent above and the fact that the negative quartic term provides an integrable tail bound for sufficiently small E . Hence

$$p_s(T) \longrightarrow N_{\text{add}} \exp \left[\kappa_0^N T + \kappa_1^N T^2 + \kappa_2^N T^3 + \kappa_3^N T^4 \right], \quad (95)$$

where N_{add} is chosen so that the density integrates to one. This is the Nonlinear+additive stationary distribution.

Summary

The calculations above show that, in the limit $E_2 \rightarrow 0$, the CAM stationary distributions reduce to their corresponding additive-noise stationary distributions:

$$\text{Linear} + \text{CAM} \longrightarrow \text{Linear} + \text{additive}, \quad \text{Nonlinear} + \text{CAM} \longrightarrow \text{Nonlinear} + \text{additive}. \quad (96)$$

More precisely, the Linear+CAM limit is the additive linear model with the same additive amplitudes used in the Linear+CAM benchmark. The Nonlinear+CAM limit is the variance-matched Nonlinear+additive model with total additive noise variance $S_1^2 + S_2^2$.

Supplementary Text 6: Moment-closure approximation for Niño 3 skewness

Here we derive the leading-order analytical approximation for the skewness of the Niño 3 SST index used in the main text. The approximation is based on the Nonlinear+additive reduced model, which retains the effective nonlinear SST drift but replaces the state-dependent stochastic forcing by variance-matched additive noise. This is the reduced model used to interpret intermodel differences in Niño 3 skewness.

The Nonlinear+additive model is

$$\frac{dT}{dt} = A_{\text{eff}}^N(T) + S_1\eta_1(t) + S_2\eta_2(t), \quad (97)$$

where

$$A_{\text{eff}}^N(T) = \tilde{\alpha}_0^N + \tilde{\alpha}_1^N T + \tilde{\alpha}_2^N T^2 + \tilde{\alpha}_3^N T^3, \quad (98)$$

and $\eta_1(t)$ and $\eta_2(t)$ are independent Gaussian white-noise processes. The total additive noise variance is

$$D_N = S_1^2 + S_2^2. \quad (99)$$

We assume that the stationary distribution is approximately centered, so that

$$\langle T \rangle \simeq 0, \quad \sigma^2 = \langle T^2 \rangle, \quad (100)$$

where σ is the Niño 3 SST standard deviation. To avoid confusion with the stochastic amplitude S_1 , we denote the skewness by $\mathcal{S}$:

$$\mathcal{S} = \frac{\langle T^3 \rangle}{\sigma^3}. \quad (101)$$

For a one-dimensional Itô stochastic differential equation with additive noise, the stationary moment equations can be obtained by multiplying the Fokker–Planck equation by powers of T and integrating by parts [4]. Equivalently, for the n th moment,

$$0 = n \langle T^{n-1} A_{\text{eff}}^N(T) \rangle + \frac{n(n-1)}{2} D_N \langle T^{n-2} \rangle. \quad (102)$$

For $n = 3$, this gives

$$0 = 3 \langle T^2 A_{\text{eff}}^N(T) \rangle + 3 D_N \langle T \rangle. \quad (103)$$

Since $\langle T \rangle \simeq 0$, the additive-noise term does not contribute directly to the third-moment equation. Therefore,

$$0 \simeq \langle T^2 A_{\text{eff}}^N(T) \rangle. \quad (104)$$

Substituting Eq. (98) into Eq. (104) yields

$$0 \simeq \tilde{\alpha}_0^N \langle T^2 \rangle + \tilde{\alpha}_1^N \langle T^3 \rangle + \tilde{\alpha}_2^N \langle T^4 \rangle + \tilde{\alpha}_3^N \langle T^5 \rangle. \quad (105)$$

To close this expression at low order, we neglect fourth- and higher-order cumulants while retaining the leading contribution from skewness. With $\langle T \rangle \simeq 0$, this gives

$$\langle T^4 \rangle \simeq 3\sigma^4, \quad \langle T^5 \rangle \simeq 10\langle T^3 \rangle \sigma^2 = 10\mathcal{S}\sigma^5. \quad (106)$$

Using

$$\langle T^2 \rangle = \sigma^2, \quad \langle T^3 \rangle = \mathcal{S}\sigma^3, \quad (107)$$

Eq. (105) becomes

$$0 \simeq \tilde{\alpha}_0^N \sigma^2 + \tilde{\alpha}_1^N \mathcal{S}\sigma^3 + 3\tilde{\alpha}_2^N \sigma^4 + 10\tilde{\alpha}_3^N \mathcal{S}\sigma^5. \quad (108)$$

Solving for $\mathcal{S}$ gives the leading-order skewness approximation

$$S_0 \simeq -\frac{\tilde{\alpha}_0^N + 3\tilde{\alpha}_2^N \sigma^2}{\sigma (\tilde{\alpha}_1^N + 10\tilde{\alpha}_3^N \sigma^2)}. \quad (109)$$

This is the expression used in the main text to diagnose Niño 3 skewness from the reduced stochastic model.

Equation (109) should be interpreted as a diagnostic leading-order approximation, not as an exact expression for the skewness. It makes explicit how skewness arises from the asymmetric components of the effective drift, $\tilde{\alpha}_0^N$ and $\tilde{\alpha}_2^N$, while the variance σ^2 controls how strongly the nonlinear drift terms are sampled by the stationary distribution. The terms $\tilde{\alpha}_1^N$ and $\tilde{\alpha}_3^N$ enter through the effective restoring strength in the denominator. In a purely linear additive system, the asymmetric drift coefficients vanish, $\tilde{\alpha}_0^N = \tilde{\alpha}_2^N = 0$, and Eq. (109) gives zero skewness, as expected.

Supplementary Tables

Table S1: Additional fitted and derived parameters used in the direct univariate SST fits and in the analytical stationary probability density functions. The direct-fit parameters correspond to univariate regressions of the Niño 3 SST tendency on T alone, where T is expressed in $^{\circ}\text{C}$. The PDF parameters are those appearing in the analytical stationary distributions in Eqs. (23)–(26) of the main text. Normalization constants are not listed because they are determined by imposing unit integral on each stationary density.

Parameter	Symbol	Linear	Nonlinear	Units
<i>Direct univariate SST fits</i>				
Direct linear SST damping	$\hat{\alpha}_1^L$	-2.899×10^{-3}	–	day^{-1}
Direct nonlinear constant term	$\hat{\alpha}_0^N$	–	-1.081×10^{-3}	$^{\circ}\text{C day}^{-1}$
Direct nonlinear linear term	$\hat{\alpha}_1^N$	–	-3.229×10^{-3}	day^{-1}
Direct nonlinear quadratic term	$\hat{\alpha}_2^N$	–	1.328×10^{-3}	$(^{\circ}\text{C})^{-1} \text{ day}^{-1}$
Direct nonlinear cubic term	$\hat{\alpha}_3^N$	–	-1.72×10^{-4}	$(^{\circ}\text{C})^{-2} \text{ day}^{-1}$
Direct univariate residual noise amplitude	B_1^{dir}	7.026×10^{-2}	–	$^{\circ}\text{C day}^{-1/2}$
<i>Analytical stationary PDF parameters</i>				
Linear+additive quadratic exponent coefficient	κ_1^L	-5.747×10^{-1}	–	$(^{\circ}\text{C})^{-2}$
Linear+CAM arctangent exponent coefficient	δ_2^L	1.283×10^1	–	dimensionless
Linear+CAM power-law exponent coefficient	δ_3^L	-4.724×10^1	–	dimensionless
Nonlinear+additive linear exponent coefficient	κ_0^N	–	-3.795×10^{-1}	$(^{\circ}\text{C})^{-1}$
Nonlinear+additive quadratic exponent coefficient	κ_1^N	–	-6.678×10^{-1}	$(^{\circ}\text{C})^{-2}$
Nonlinear+additive cubic exponent coefficient	κ_2^N	–	1.576×10^{-1}	$(^{\circ}\text{C})^{-3}$
Nonlinear+additive quartic exponent coefficient	κ_3^N	–	-1.270×10^{-2}	$(^{\circ}\text{C})^{-4}$
Nonlinear+CAM linear exponential coefficient	δ_0^N	–	4.815×10^1	$(^{\circ}\text{C})^{-1}$
Nonlinear+CAM quadratic exponential coefficient	δ_1^N	–	-2.047	$(^{\circ}\text{C})^{-2}$
Nonlinear+CAM arctangent exponent coefficient	δ_2^N	–	-4.514×10^2	dimensionless
Nonlinear+CAM power-law exponent coefficient	δ_3^N	–	4.931×10^1	dimensionless

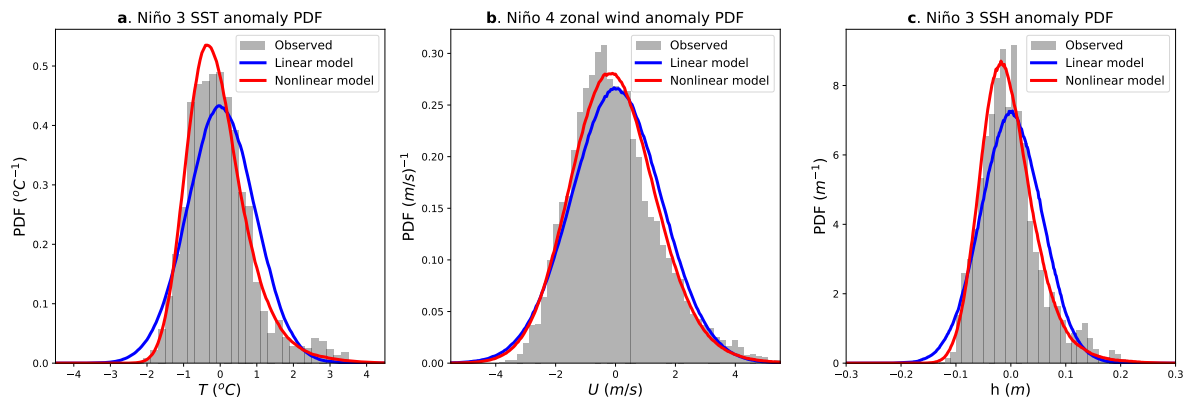


Figure S1: Probability density functions of **a** Niño 3 SST anomalies (T), **b** Niño 4 zonal wind anomalies (U), and **c** Niño 3 sea surface height anomalies (h), comparing observations with the linear and nonlinear three-variable models. The three-variable model includes Niño 3 SSH as an explicit upper-ocean proxy and is fitted over the 1993–2017 period. The linear model remains close to Gaussian and does not reproduce the observed positive skewness of Niño 3 SST or Niño 3 SSH. The nonlinear model substantially improves the representation of the asymmetric SST and SSH distributions, while changes in the Niño 4 wind distribution are more modest. This indicates that adding an explicit SSH variable does not qualitatively change the main linear-versus-nonlinear contrast emphasized in the two-variable SST–wind framework.

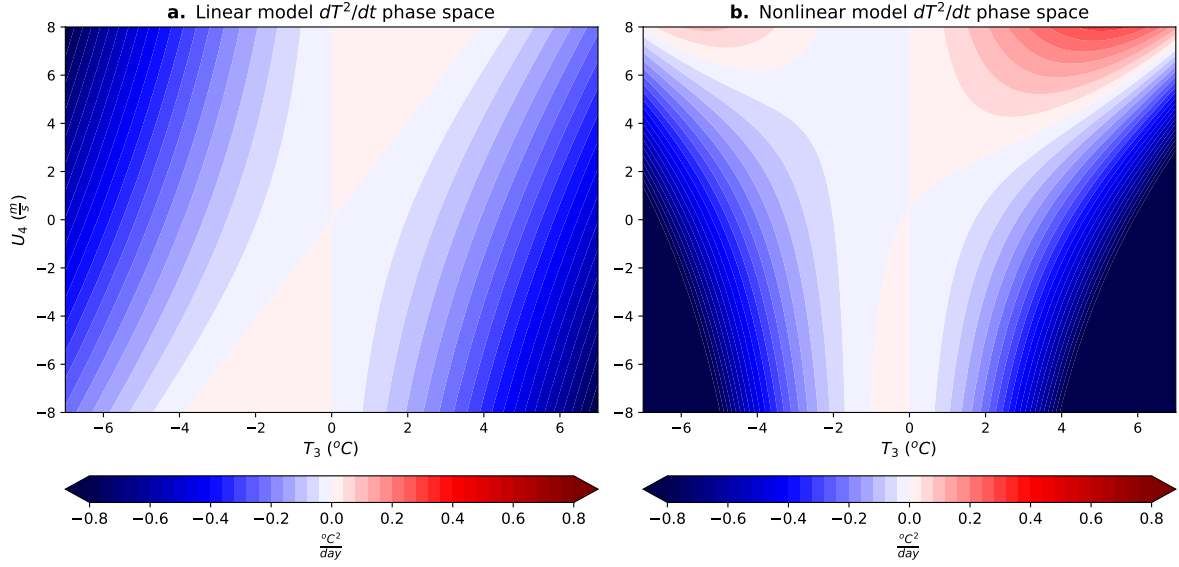


Figure S2: Phase space of the growth tendency of the Niño 3 SST index, dT^2/dt , as a function of concurrent Niño 3 SST T and Niño 4 zonal wind U indices in the **a** linear and **b** nonlinear models. A positive (negative) value of dT^2/dt implies that a given Niño 3 SST anomaly grows (decays) in magnitude deterministically.

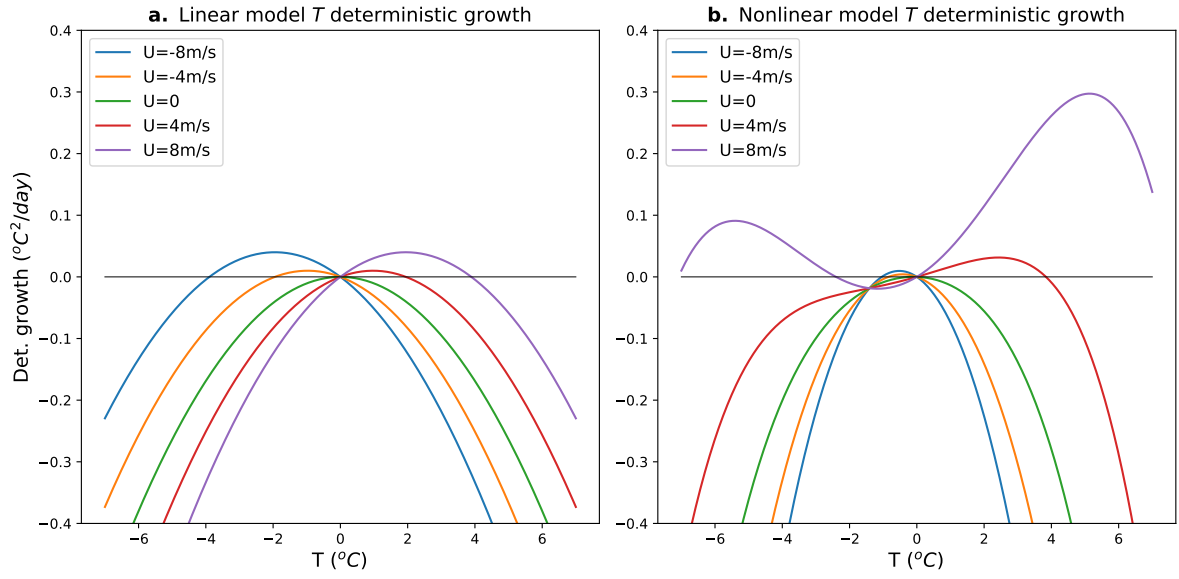


Figure S3: Growth tendency of the Niño 3 SST index, dT^2/dt , in the **a** linear and **b** nonlinear models as a function of T , for selected values of the Niño 4 zonal wind index U .

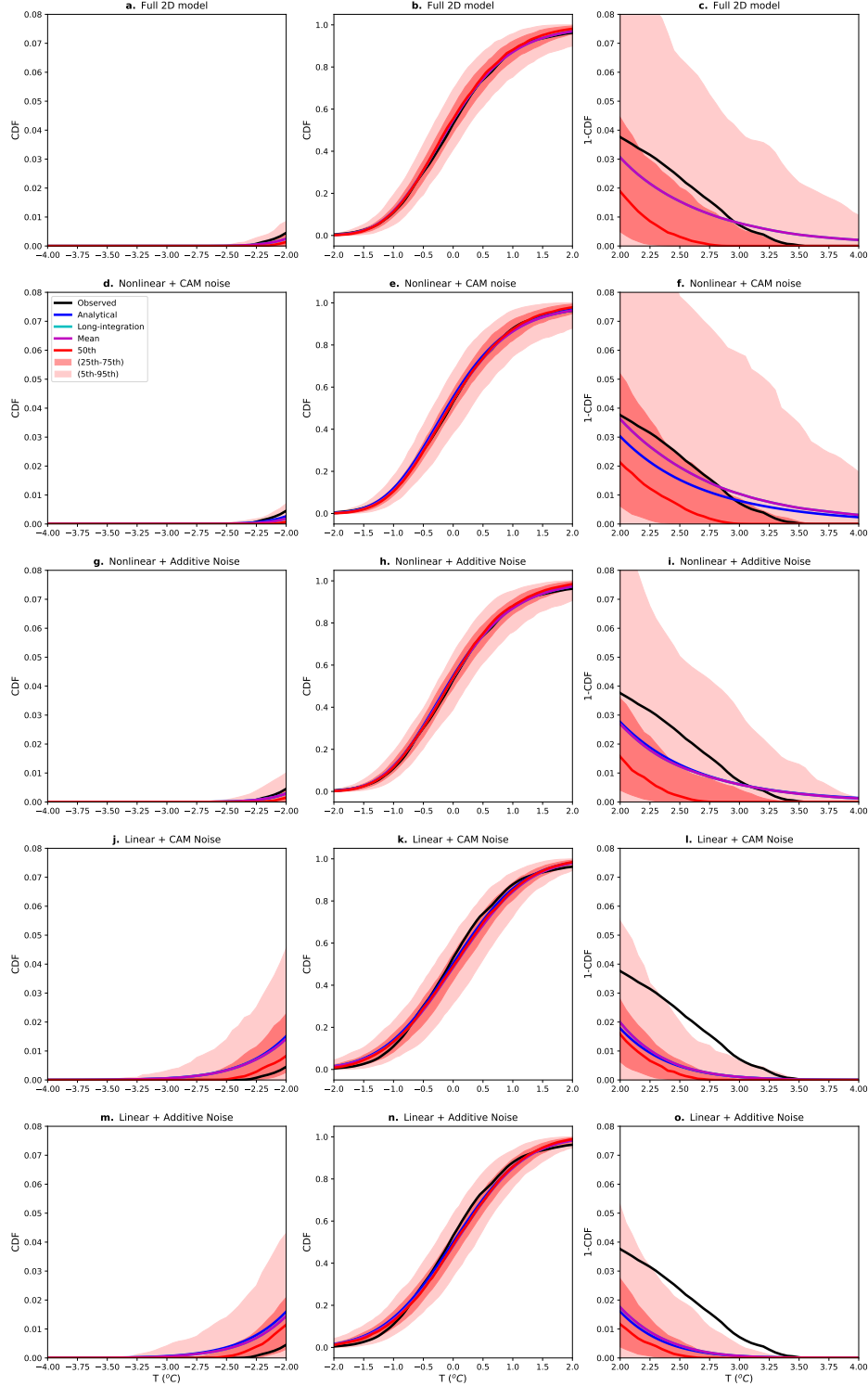


Figure S4: Cumulative density functions of the Niño 3 index in observations and in the different models considered: the full two-dimensional model, the one-dimensional Nonlinear+CAM noise model, the one-dimensional Nonlinear+additive noise model, the one-dimensional Linear+CAM noise model, and the one-dimensional Linear+additive noise model. Curves show the observed CDF (black), the CDF derived from the analytical PDFs (blue; except for the full two-dimensional model, for which an analytical solution is not derived), the CDF numerically calculated from a long integration of 200 epochs of 38 years of daily values (cyan), the mean of the 200 CDFs calculated from 38-year epochs (magenta; in most cases indistinguishable from the long integration), and the 50th percentile of the 200 CDFs calculated from 38-year epochs (red). Shading denotes the 25th–75th percentile range (dark red) and 5th–95th percentile range (light) of the 200 CDFs calculated from 38-year epochs. The center panels show the central part of the CDFs, while the panels on either side zoom in on the tails. For the positive tail, the exceedance probability, $1 - \text{CDF}$, is shown instead.

References

- [1] Majda, A. J., Timofeyev, I. & Vanden Eijnden, E. Models for stochastic climate prediction. *Proceedings of the National Academy of Sciences* **96**, 14687–14691 (1999).
- [2] Franzke, C., Majda, A. J. & Vanden-Eijnden, E. Low-Order Stochastic Mode Reduction for a Realistic Barotropic Model Climate. *Journal of the Atmospheric Sciences* **62**, 1722–1745 (2005).
- [3] Pavliotis, G. A. & Stuart, A. M. *Multiscale Methods*, vol. 53 of *Texts Applied in Mathematics* (Springer, New York, NY, 2008).
- [4] Gardiner, C. *Stochastic Methods*.
- [5] Sardeshmukh, P. D. & Sura, P. Reconciling Non-Gaussian Climate Statistics with Linear Dynamics. *Journal of Climate* **22**, 1193–1207 (2009).
- [6] Sardeshmukh, P. D. & Penland, C. Understanding the distinctively skewed and heavy tailed character of atmospheric and oceanic probability distributions. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**, 036410 (2015).
- [7] Martinez-Villalobos, C., Vimont, D. J., Penland, C., Newman, M. & Neelin, J. D. Calculating State-Dependent Noise in a Linear Inverse Model Framework. *Journal of the Atmospheric Sciences* **75**, 479–496 (2018).
- [8] Majda, A. J., Franzke, C. & Crommelin, D. Normal forms for reduced stochastic climate models. *Proceedings of the National Academy of Sciences* **106**, 3649–3653 (2009).
- [9] Ewald, B. & Penland, C. Numerical Generation of Stochastic Differential Equations in Climate Models. In *Handbook of Numerical Analysis*, vol. 14, 279–306 (Elsevier, 2009).