Supplementary Material for "Climatology and natural and forced changes of ENSO variability"

SI1 OLS ill-suited in the presence of CAM noise

Ordinary least-squares regression (OLS) is possibly the most common technique to fit a model/curve to data. However, here we demonstrate that caution should be exercised as to the suitability of this method. Sometimes it just does not work. For the purpose of this demonstration, we consider the linear stochastic differential equation (SDE) featuring the so-called "correlated additive and multiplicative" (CAM) noise [1] written in the Itô form:

$$dx = a(x-b)dt + (c(x-b)+d)dW. \tag{S1}$$

This is a reparametrized version of the SDE in [2] such that b appears as a "location" parameter. The probability density distribution of x evolving according to this SDE can then be written by substituting the formulae of the parameter transformation into eq. (2) of [2], yielding:

$$p(x) = N_0 \frac{2e^{2a(\log(f)+d/f)/c^2}}{f^2}, \quad f(x;b,c,d) = -bc + d + cx, \tag{S2}$$

where N_0 is a normalization constant. Because of the term $\log(f)$, it is required that $f > 0$, implying a lower boundary of the distribution: $b - d/c$. Further constraints can be derived as: $a/c^2 < 1/2$ and $ad/c^3 < 0$ (which we obtained by using the software suit "Mathemtaica"). See $p(x)$ in Fig. S2 (a) for the choices of $a = -1$, $b = 1$, $c = 0.2$, $d = 1$, the only scenario that we are considering for this demonstration. The SDE can be discretized by applying the Euler-Maruyama stochastic integrator scheme [3], the simplest of its kind, yielding:

$$x_n = x_{n-1} + a\Delta t(x_{n-1}-b) + (d + c(x_{n-1}-b))\sqrt{\Delta t}\xi_{n-1}. \tag{S3}$$

For the parameter choices given above, $\Delta t = 0.01$ and a simulation span $N = 10^6$ are suitable to reproduce $p(x)$ numerically (Fig. S2 (a)). Equation (S3) can be rearranged expressing the normal random variable ξ_{n-1} explicitly (or rather that multiplied by d):

$$\xi_{n-1} \sim \frac{x_n - x_{n-1} - a\Delta t(x_{n-1}-b)}{(1+\frac{c}{d}(x_{n-1}-b))\sqrt{\Delta t}}. \tag{S4}$$

Given an observational data series, x_n, $n = 1, \ldots, N$, that is thought to be governed by (S1), one could attempt to infer the model parameters a, b, c, d – as in OLS – by minimizing the "sum of squares" (of the right hand side of (S4)):

$$SS \sim \sum_{n=2}^{N} \xi_{n-1}^2. \tag{S5}$$

This is at least what An et al. [4] did, despite that OLS – linear or nonlinear – assumes the random increments to be *additive* [5]. This is clearly not the case here. And in this instance, it seems to cause some peculiar phenomenon, as showcased in Fig. S1 (a), looking at the dependences $SS(a,b,c',d)$ and $SS(a,b,c,d')$ (in which the primed symbols denote independent variables). These functions of one variable are slices of the function $SS(a',b',c',d')$ of four variables (or an approximation of that considering that $N = 10^6$) traversing the point in parameter/variable space corresponding to the true value of the parameters. For some respective ranges of the variables, the graph of the functions are very jagged, presumably they are *sporadically fratal* curves (as described in Sec. 5.2.3 of [6]). This most likely implies uncountably many minima of the function $SS(a',b',c',d')$. Around the true values, the 1D dependences are smooth. However, one might truly have no good idea of the actual values, and, thus, cannot start the minimum search from a "safe basin of attraction". We also notice by the $SS(a,b',c,d)$ dependence that there is no minimum near the true parameter value b. Therefore, if this is not a sign of the proposed minimum search being an underdefined problem, then that of very considerable biases.

Next, we look at similar slices of the negative log-likelihood function, $-\log(L)$, defined by eq. (9) in Methods 4.2, shown in Fig. S1 (b). This looks rather promising: the slices have minima just about at the

true values, c showing the largest bias. It is also c for which the slice has multiple minima, owing to the couple of wiggles of $-\log(L)$ for small values of c. This is luckily not such a fundamental problem as the fractality of SS discussed above; the wiggles are due to difficulties with determining the normalization constant N_0 accurately as an integral. For vanishing values of c, N_0 is increasing beyond any limit and finite number representation by a digital computer introduces inaccuracies. Still, a minimum search of $-\log(L)$ to the end of inferring the parameter values from data is *not* successful. This is in fact due to the underdetermination of the problem for all four parameters of (S1). Equation (3) and (5) of [7] providing $p(x)$, unlike our formula (S2), reveals that only *two* nondimensional parameters do fully determine the density function. (After all, the fact alone that all four slices have a minimum does not mean that the function of four variables does have a unique minimum right about there because its minimum might be a 1D object.) Therefore, MLE that is based on $p(x)$, cannot determine more than two (dimensional) parameters. This is not a problem in the case of the simple SDE (S1), because we can easily estimate two parameters as follows:

$$\tilde{b} = S_1, \quad \tilde{a} = \frac{S_3 - S_2}{\Delta t(S_2 - S_1^2)},$$

$$S_1 = \overline{x_n}, \quad S_2 = \overline{x_n^2}, \quad S_3 = \overline{x_n x_{n-1}},$$

where the overbar denotes arithmetic average. Furthermore, a connection between the remaining two parameters can be established:

$$S_2 = C_1^2 S_2 - 2 * C_1 C_2 S_1 + C_2^2 + C_3^2 S_4 + 2C_3 C_4 S5 + C_4^2 S_6,$$

where

$$S_4 = \overline{\xi_n^2} = 1, \quad S_5 = \overline{x_n \xi_{n-1}^2}, \quad S_6 = \overline{x_n^2 x_{n-1}^2},$$

$$C_1 = \Delta t\tilde{a} + 1, \quad C_2 = \Delta t\tilde{a} + \tilde{b}, \quad C_3 = \sqrt{\Delta t}(d - c\tilde{b}), \quad C_4 = \sqrt{\Delta t}c.$$

The connection is a quadratic equation for d depending on c. The +ve root is to be taken for $\tilde{d}$, with which we end up with a function of a single variable for the likelihood, or $-\log(L(\tilde{a}, \tilde{b}, c', \tilde{d}(c')))$. It is shown in Fig. S2 (b), where we can visually verify that the minimum is indeed just about the true value $c = 0.2$. The estimate is in fact $\tilde{c} = 0.215$. This is the estimate to such a precision for any realisation of $N = 10^6$, which advises us about the bias. We deem it rather small. At the same time, $\tilde{a} = -0.9876$, $\tilde{b} = 0.9971$, $\tilde{c} = 1.0021$, as yielded by the above estimator formulae and connection. This suggests that the estimator of c is the most (or, perhaps, the only) biased one. Furthermore, the problem remains of course that small c's cannot be estimated reliably in this way. This problem, however, should not be present for the other MLE estimation technique outlined in Sec. 4.2.

In fact, the classical MLE estimation technique presented in this section is not applicable to the ROM1 model (1a)-(1b), because the PDF $p(T)$ is most likely cannot be established analytically. Furthermore, we would need further "tricks" to estimate as many parameters as many are not determined by the classical MLE method.

SI2 Forced changes of skewness in view of the ROM1

Here we provide an alternative analysis to that presented in Sec. 2.3 of the main article on the ROM2. This one is most likely wrong, because the ROM1 is missing the hetero-quadratic nonlinearity ($a_{2,Th} \neq 0$) arising due to vertical nonlinear dynamical heating, which is believed to be present in the CESMv2 and MPI-ESM. This situation, however, provides an opportunity of demonstrating what model misspecification could cause. On the one hand, we will observe that the coefficient $a_{2,TT}$ of the homo-quadratic term would be misestimated, thereby recovering some of the skewness lost with missing $a_{2,Th}$. On the other hand, out of the three conceptual models of ENSO considered (Sec. 1), it is only the ROM1 for which we are able to obtain an approximate formula (S6) of the skewness that can predict well the parametric estimate of the forced change of ENSO skewness in the CESMv2.

Only in the CESMv2 do we have a strong signal of decreasing b **of the ROM1** under *stronger* greenhouse forcing (Fig. S3). This could be thought to be responsible for the declining positive skewness in that model (Fig. 4). But perhaps also a forced change of deterministic nonlinearity ($a_{2,TT}$) plays a role; this is at least what we see in view of the monthly data – but not the pentad data! We seem to be able to rule out the importance of the change of noise correlations, on the other hand, as the auto-correlation

(AC) of e.g. ξ_T is *increasing* (Fig. S3) which should enhance skewness whereas the opposite is true. In the historical or relatively weakly forced period, on the other hand, b seems stationary, but $a_{2,TT}$ decreases slightly and D_T increases, both of which should reduce skewness γ, the latter, D_T, as per [8], yet γ is stationary. Q1: Are the changes of $a_{2,TT}$ and D_T too small to impact on the skewness, or is there a problem with the simple justification regarding D_T?

We must point out that estimates of $a_{2,TT}$ in the different models vastly differ: being about a fifth in the CESMv2 as compared with the MPI-ESM. Yet, the skewness in the two models, seen in Fig. 4 (b) and (e) are not that different. This makes sense naively considering that estimates of $a_{1,TT}$ are also vastly different. It is easy to check numerically that the skewness depends both on $a_{2,TT}$ and $a_{1,TT}$. Therefore, we are inclined to conclude that we are seeing a *compensation effect* here.

In any case, the compensation, agreeing with the numerical experience, implies that a decreasing $|a_{1,TT}|$ would *increase* γ, which falsifies a naive reading of [8]. This also means that under strong forcing, the increase of $|a_{1,TT}|$ should decrease γ – competing with the decreasing b. Q2: Which one is the dominant factor, then, if there is one?

The situation is different in the MPI-ESM. There, $a_{2,TT}$ is steadily increasing and b is practically zero throughout the simulation period. Consistently, the skewness increases, as discussed above. Perhaps the increasing auto-correlation of ξ_T also contributes beside $a_{2,TT}$, although only after year 2000, not in the rather weakly forced historical period. It could be thought that the change of $a_{2,TT}$ is not an artefact in relation with the said compensation effect because $a_{1,TT}$ does not change when $a_{2,TT}$ already changes.

Considering the CESM2-LE, answering the two questions, Q1 and Q2, Q1 pertaining to the weakly forced historical period and Q2 to the strongly forced future period, is attempted here by performing *dimensional analysis* of the ROM1. **See the derivation of the dimensionless ROM2 in Sec. 2.3 as well as the definitions of dimensionless parameters all of which have $|a_{1,TT}|$ in the denominator of a fraction raised to various powers.**

Concerning Q2, this inverse proportionality with $|a_{1,TT}|$ explains why the increase of $|a_{1,TT}|$ reduces the skewness of T, γ_T. However, we would also like to know if the increase of $|a_{1,TT}|$ or possibly the decrease of b in CESMv2 is the more dominant effect. This would be a straightforward analysis if γ_T depended linearly on $|a_{1,TT}|$ and b. However, while it might be the case wrt. b, because of the said *inverse* proportionality and the various powers $|a_{1,TT}|$ is raised to in defining the π's, it looks unlikely wrt. $|a_{1,TT}|$. In the case of nonlinear dependence, though, the analysis is not much more complicated, only that beside the contributions (wrt. some reference, say, the first value in a time series, pertaining to the background state) of the respective variables, there is also a *nonlinear interaction* term. Let us pretend for now as if noise correlations did not matter. Considering Fig. S5 (d), first, it seems that while the individual contributions are –ve, the interaction term is +ve, opposing. Second, it is also clear that the changes of $|a_{1,TT}|$ and b cause most of the changes of γ_T under strong late 21st c. forcing. Third, none of them is dominant, but, intriguingly, their contributions are approximately equal. Then, the discrepancy between the grey curve of total change and the purple one accounting for the changes of only $|a_{1,TT}|$ and b appears to be largely due to a precipitous drop of $a_{2,TT}$ somewhat after 2050 (Fig. 4). That is, considerable changes of D_T, A, Φ, $a_{1,hT}$, $a_{1,hh}$ do not make much difference. In fact, γ_T does not show measurable dependence on A, Φ, $a_{1,hh}$, or α, $\alpha_{1,hh}$, at all. We have come up with an approximate formula for the skewness:

$$\gamma_T = \gamma_n + \gamma_s = \varphi_n f_n(\alpha_{2,TT}) + \varphi_s f_s(\beta), \tag{S6a}$$

$$f_n(\alpha_{2,TT}) \approx \alpha_{2,TT} \tag{S6b}$$

$$f_s(\beta) \approx \beta \tag{S6c}$$

$$\varphi_n \approx c_1 + c_2\alpha_{1,hT} + c_3\alpha_{1,hh}^2, \tag{S6d}$$

$$\varphi_s \approx c_4, \tag{S6e}$$

with constant c's, fairly accurate (Figs. 5 (a), S4 (b)) in the regime of the changes of the parameters of the ROM1 fitted to the monthly CESM2-LE data. Notably, because of the additive terms, changes of γ_T due to changes of $\alpha_{2,TT}$ alone and β alone (or $a_{2,TT}$ alone and b alone), clearly, γ_n and γ_s ('n' ('s') for 'nonlinearity' ('state dependence')), are *superimposable* (Fig. S4 (a)). (They would be superimposable, of course, even if the dependence of γ_T on either of $\alpha_{2,TT}$ and β, i.e., $f_n(\alpha_{2,TT})$ or $f_s(\beta)$, is nonlinear, which is actually true but noticeable for rather larger changes of those.) Then, the "factor" φ_n for $\alpha_{2,TT}$ depends on a subset of other π's, while b features a constant factor, $\varphi_s = c_4$. Although, both factors (but also $f_n(\alpha_{2,TT})$ and $f_s(\beta)$ with some restrictions of the functional forms) most probably depend on

(possibly all, except for $\alpha_{2,TT}$ and β) π's in the full range of permissible parameter values. (In fact, we started out with quadratic polynomials for all of $f_n(\cdot)$, $f_s(\cdot)$, $\varphi_n(\cdots)$, $\varphi_s(\cdots)$, not on a theoretical basis but just using a truncated Taylor series for *approximation*, counting 22 terms (without intercepts/offsets in $f_n(\cdot)$, $f_s(\cdot)$), and eliminated terms successively when we saw (in diagrams like Fig. 5 (b)) approximately identical pairs of time series or pairs symmetrical to the t-axis – one of the pair(s) randomly selected were eliminated. The process involved several iterations of term elimination. Only (S6a) do we believe to be exact. Otherwise, because of the certain dependence of φ_n on $\alpha_{1,hT}$ and $\alpha_{1,hh}$, parameters of the h-equation, we can point out that in modelling ENSO (well, its skewness, at least), it is essential to account for the dynamics of the mixed layer depth (h) and its coupling to the SST (T) (a counterexample being [7]). Fig. 5 (b) is testimony to this claim.

Aided by eq. (S6), finally we turn to our other question, Q1, pertaining to changes of the skewness, or the lack of change, under the weaker historical forcing. Considering the definition of the π's on which the factor φ_n depends, we can say that γ_T *increases* with D_T. (Again, a naive reading of [8] would thus be a mistake.) Therefore, chances are that effects of changes of D_T and $a_{2,TT}$ largely cancel one another. Fig. 5 (c) shows that this is indeed the case, although the cancelling effects themselves, individually, are rather small.

Yet, a gross problem with the analysis here is that the ROM1 fitted to the CESM2-LE would not feature the level of skewness. It is about half of that of the CESM2-LE. It is another matter that the CESM2-LE itself features an unrealistically small skewness. See further discussions in Sec. 3 regarding the problem of model misspecification.

SI3 Forced changes of ENSO variability in view of the DROM+

References

[1] Philip Sura, Matthew Newman, Cécile Penland, and Prashant Sardeshmukh. Multiplicative Noise and Non-Gaussianity: A Paradigm for Atmospheric Regimes? *Journal of the Atmospheric Sciences*, 62(5):1391–1409, 2005.

[2] Tamás Bódai and Christian Franzke. Predictability of fat-tailed extremes. *Phys. Rev. E*, 96:032120, Sep 2017.

[3] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1995.

[4] Soon-Il An, Soong-Ki Kim, and Axel Timmermann. Fokker–Planck dynamics of the El Niño-Southern Oscillation. *Scientific Reports*, 10(1):16282, Oct 2020.

[5] Wikipedia contributors. Regression analysis — Wikipedia, the free encyclopedia, 2024. [Online; accessed 16-February-2024].

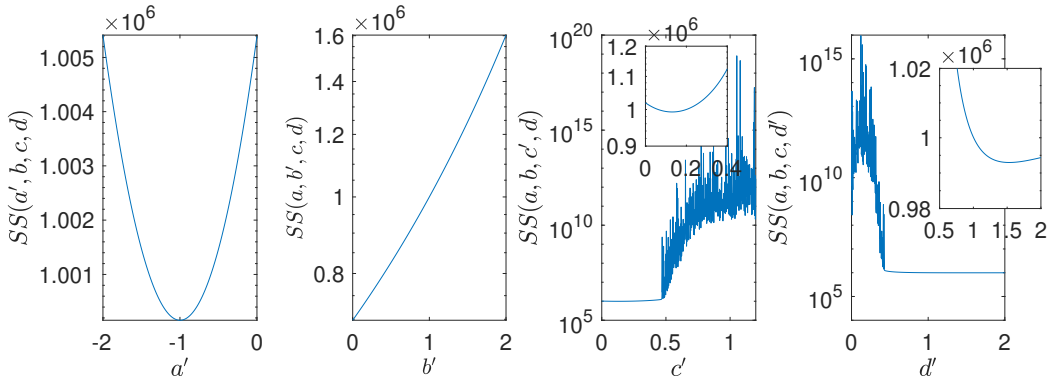
[6] Y.-C. Lai and T. Tél. *Transient Chaos*. Springer, New York, 2011.

[7] Marco Bianucci and Riccardo Mannella. On the determination of the optimal parameters in the CAM model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(3):033113, 03 2021.

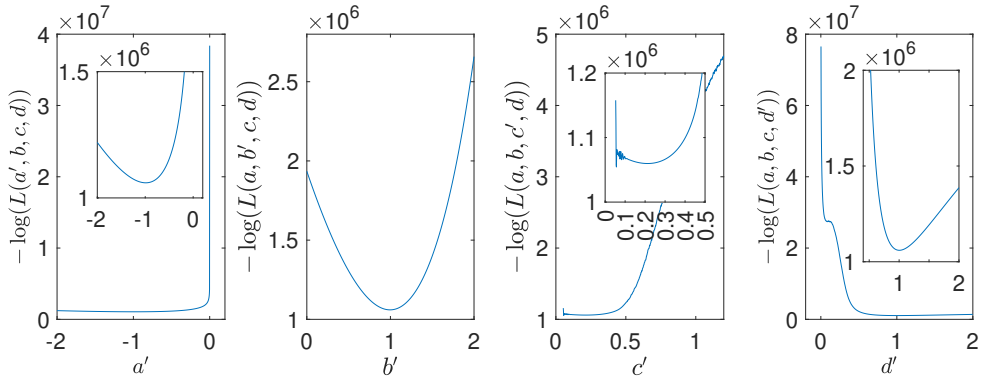
[8] Chao Liu, Soon-Il An, Fei-Fei Jin, Malte F. Stuecker, Wenjun Zhang, Jong-Seong Kug, Xinyi Yuan, Jongsoo Shin, Aoyun Xue, Xin Geng, and Soong-Ki Kim. ENSO skewness hysteresis and associated changes in strong El Niño under a CO2 removal scenario. *npj Climate and Atmospheric Science*, 6(1):117, 2023.

[9] Clara Deser, Adam S. Phillips, Isla R. Simpson, Nan Rosenbloom, Dani Coleman, Flavio Lehner, Angeline G. Pendergrass, Pedro DiNezio, and Samantha Stevenson. Isolating the Evolving Contributions of Anthropogenic Aerosols and Greenhouse Gases: A New CESM1 Large Ensemble Community Resource. *Journal of Climate*, 33(18):7835 – 7858, 2020.

[10] David Ruelle. A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22(4):855, 2009.

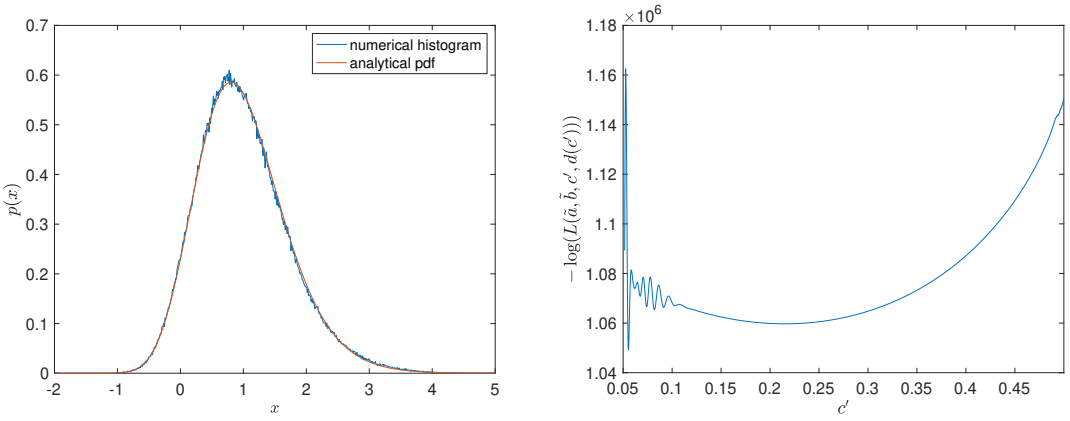


(a)



(b)

Figure S 1: Slices of the "sum of squares" SS (S5) (top) and the negative log-likelihood (9), functions of a single variable, quantities to minimise in the methods of regression and Maximum Likelihood Estimation. They pertain to our demo example (S1).



(a) (b)

Figure S 2: Analytic and simulated probability density function $p(x)$ (a) and the negative log-likelihood depending on just one parameter value (b) for our demo example (S1).

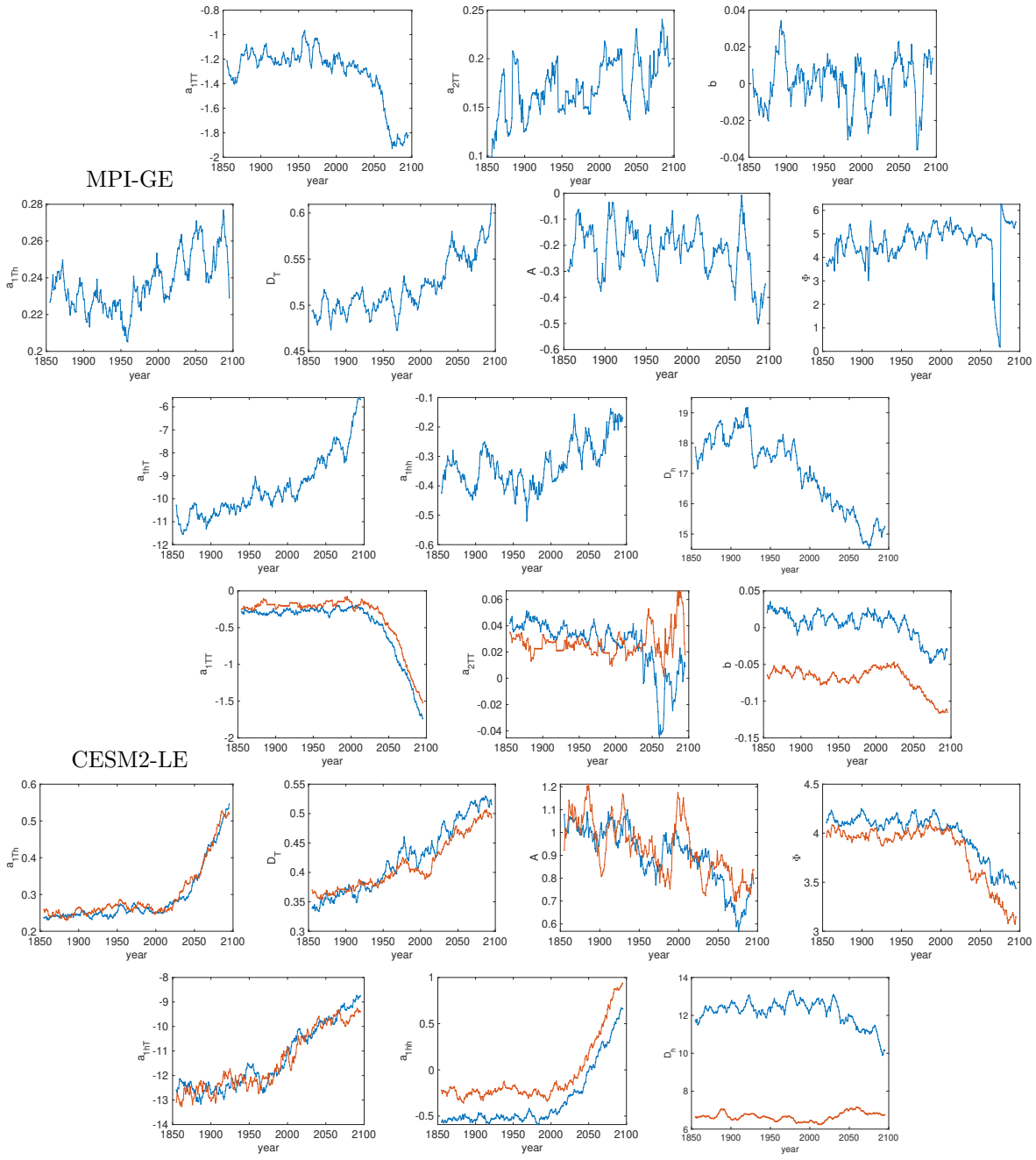


Figure S 3: Like Fig. 4, but for the *parameters* of the ROM1 (1).

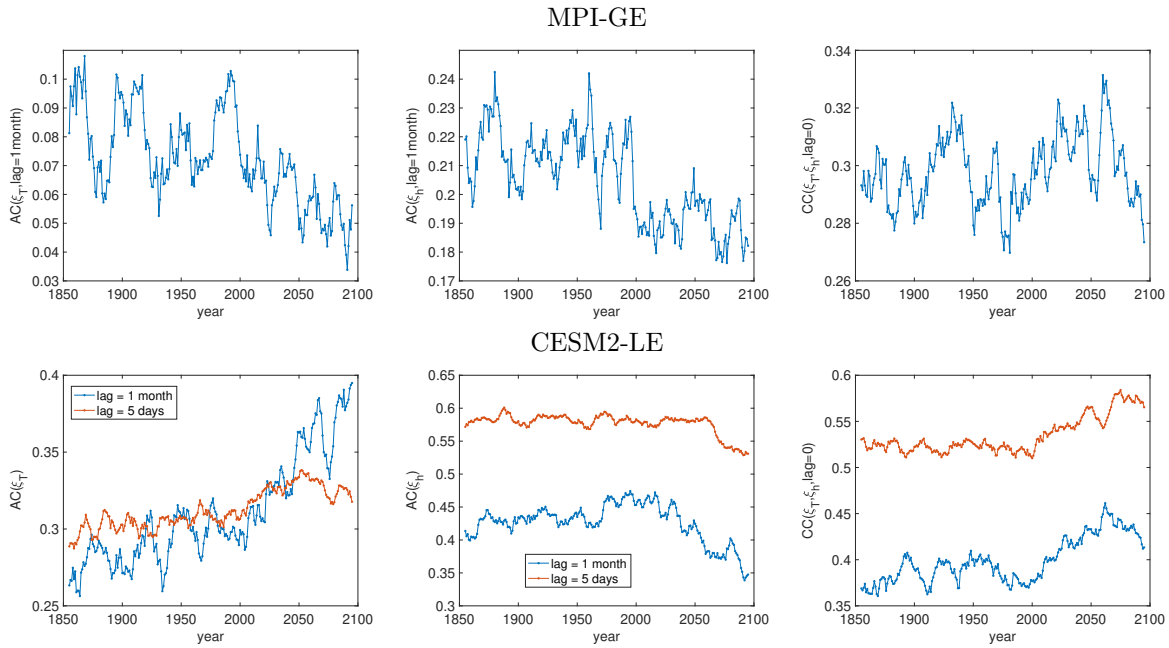
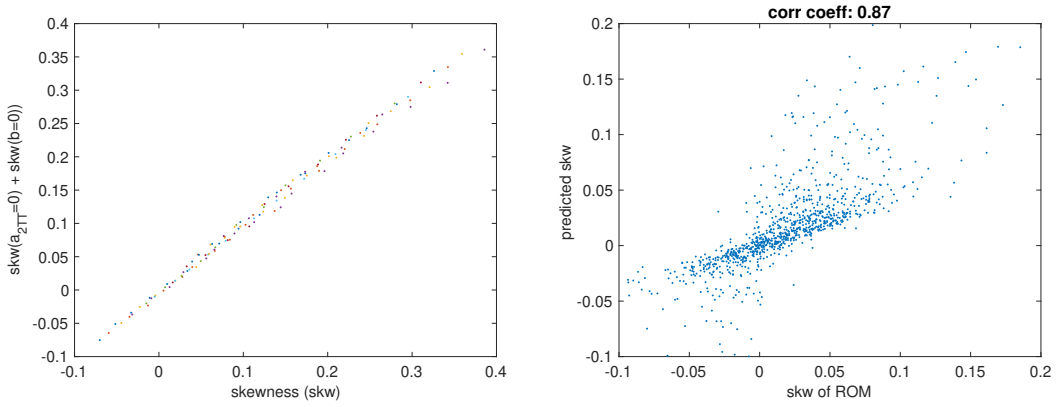


Figure S 3: Like Fig. 4, but for the *correlations* of the random increments ξ_T, ξ_h of the ROM1 (1). 'AC' is for 'auto-correlation'; 'CC' is for 'cross-correlation'.



(a) (b)

Figure S 4: Monte Carlo (MC) experiments in relation with our new skewness formula (S6). (a) Demonstration of the superposition of contributions to the skewness made by nonzero values if $a_{2,TT}$ and b. (b) Scatter plot to validate eq. (S6). For the two different diagrams, the output of two different MC experiments are used. (a) 11 equally spaced sample values of $a_{2,TT}$ (b) are defined between 0 and 3 (–3 and 3); the other parameters are fixed at values of: $a_{1,TT} = -1$, $a_{1,Th} = 0.3$, $a_{1,hT} = -13$, $a_{1,hh} = 0.2$, $A = 1.2$, $\Phi = 4.4$, $D_T = 0.3$, $D_h = 15$, loosely corresponding to figures seen in Table 1. The ranges for the variations of $a_{2,TT}$ and b are large enough so that the dependence of γ on them is already clear to be nonlinear. The simulations of the ROM1 are run over a span of 50,000 years with a small $\Delta t = 1/10$ [month] for accuracy. On the vertical axis we measure a *prediction* of the skewness from a subset of the simulation output assuming superposition holds. Markers are lined up tightly around the line of unit slope going through the origin, indicating superposition. (b) This is rather a true MC experiments with random realisations of the parameter values within respective ranges using a uniform probability density. The ranges are corresponding to the ranges roughly realised by the CESM2-LE data as seen in Fig. S3: $a_{2,TT} : [-0.02, 0.05]$, $b : [-0.03, 0.03]$, $a_{1,TT} : [-1.8, -0.2]$, $a_{1,Th} : [0.1, 0.6]$, $a_{1,hT} : [-13, -9]$, $a_{1,hh} : [-0.6, 0.2]$, $A : [0.7, 1.1]$, $\Phi = 4.4$ (no need to vary due to symmetry), $D_T : [0.5, 1.5]$, $D_h : [10, 13]$. Parameter value combinations not permitted for stable dynamics do come up (notice how the forced change of $a_{1,hh}$ goes from –ve to +ve values, resulting a change from –ve to +ve feedback.), which are simply ignored. Simulation time span and step size are the same as for the other experiment. On the horizontal (vertical) axis we measure the skewness directly estimated from the simulation time series of T (predicted by eq. (S6)). The markers are not tightly around a straight line, however, the linear correlation coefficient is 0.87; plus see Fig. 5 (a) as for the validation.

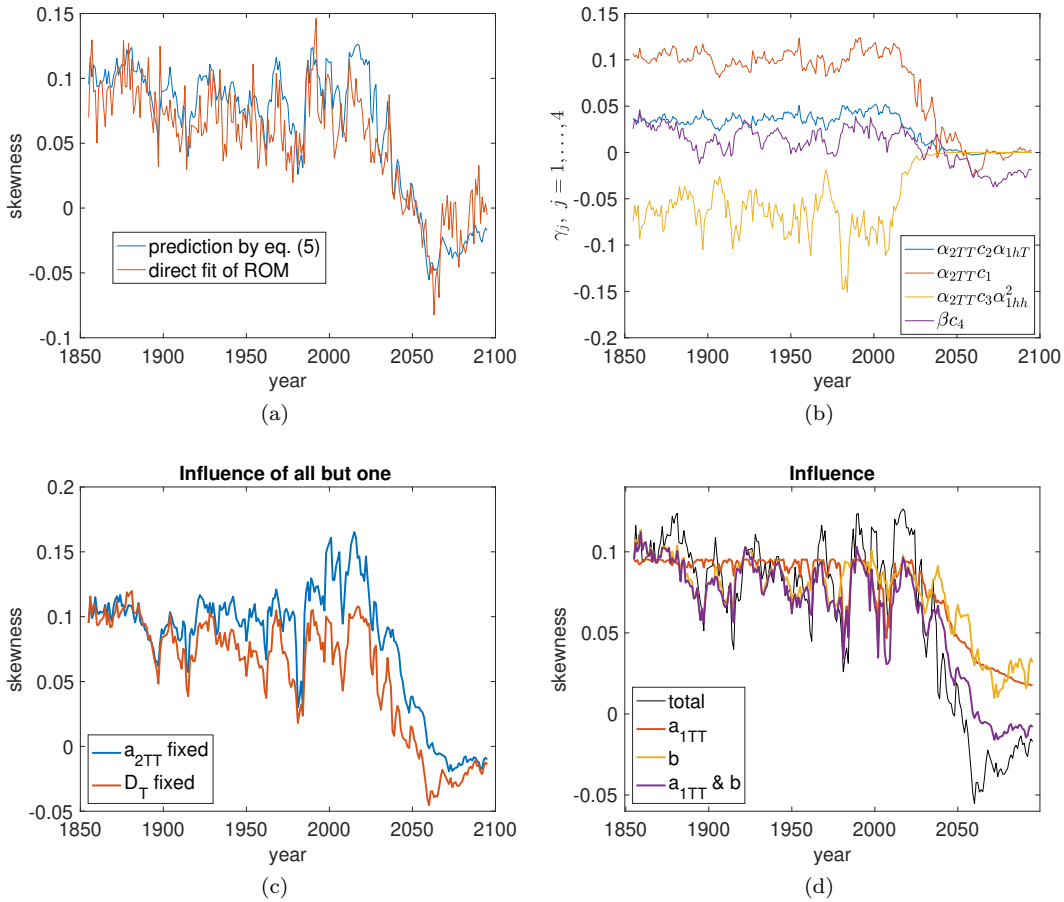
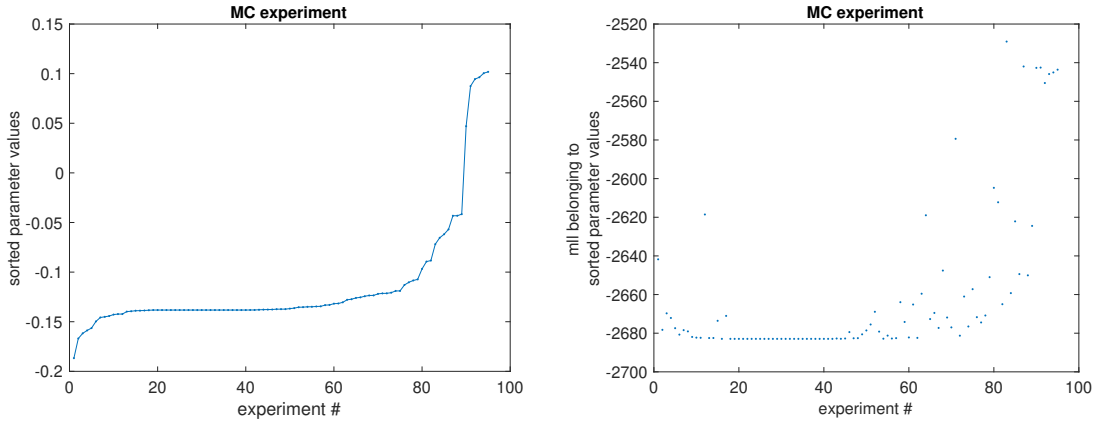


Figure S 5: Predictions of ENSO skewness using our new formula (S6), including the influence of various ROM1 parameters/physical factors. (a) Validation of (S6) by comparing the direct measurement of the skewness of the ROM1 fitted to the monthly CESM2-LE data (time series copied from Fig. S13 (a)) to the prediction by the formula using the time series of the ROM1 parameters (as seen in Fig. S3) inferred by the same fit. (b) The four terms of (S6), j corresponding to the constants c_j. (c) So called "all but one" scenarios [9], when the influence of a forcing factor is assessed through fixing it to a constant value while all other forcing factors are as originally realised. (d) A "reciprocal" approach to (c) in that we fix all forcing factors except retaining the original values of, say, one. The latter would show the "contribution" of that factor to the grand total in the specific sense that this contribution is a response to that factor in a system where the responses to *all* forcing factors are *linear* [10]. When the response characteristics are nonlinear, then a "single forcing" experiment can be misleading because of nonlinear interactions in the case of multiple forcings acting at the same time, something that the "single forcing" experiment cannot indicate. Of course, if the sum total of all possible "single forcing" experiments do not add up to the outcome when all forcing factors are active, it is an indication of nonlinearity.



(a) (b)

Figure S 6: Visualisation of the outcome of a typical Monte Carlo experiment in which the parameter inference method (MLE) is initialised randomly. There is a "dominant solution" that belongs to a quasi-uniform minimum of the negative log-likelihood (nll). Other outcomes might or might not belong to a different local minimum of the nll; perhaps their initial condition might just not be attracted strongly enough to the solution.

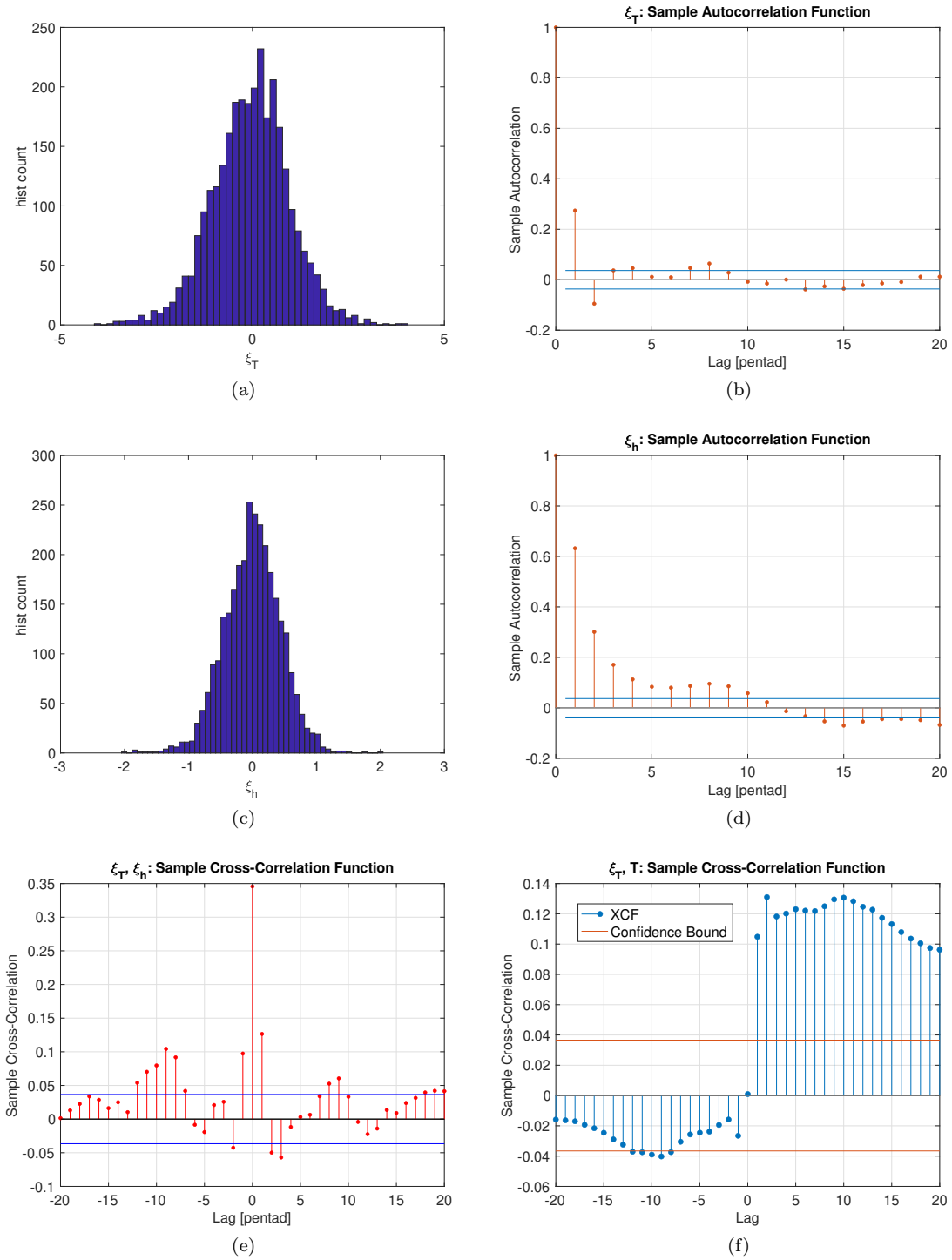
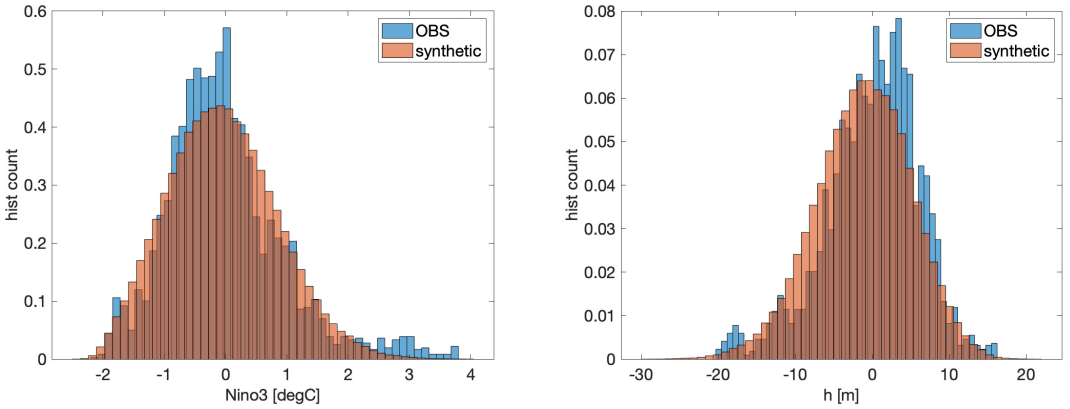
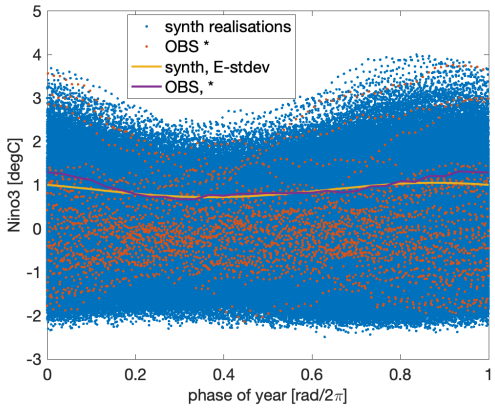


Figure S 7: Diagnostics of the fitted ROM1 model (1a)-(1b) to the SODAv3.3.1 data, I: normality and auto- and cross-correlation of residuals. Pairs of horizontal blue lines bracket correlation data points that are not significantly non-zero. We used Matlab's `autocorr` and `crosscorr`. Unfortunately, not only the residuals ξ_T, ξ_h correlate, but also the state variable T seems to causally affect ξ_T, in violation with the assumption of the MLE fitting (Methods 4.2).

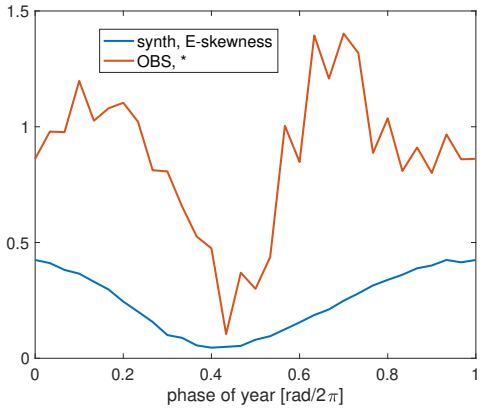


(g)

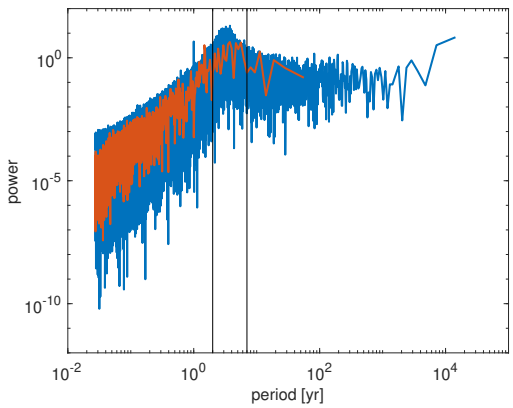
(h)



(i)



(j)



(k)

Figure S 7: Diagnostics of the fitted ROM1 model (1a)-(1b) to the SODAv3.3.1 data, II: histograms of the state variables, seasonal dependence of the standard deviation and skewness of Niño3, power spectrum/periodogram of Niño3. The pair of vertical black lines in the periodogram indicate the 2-7 years band.

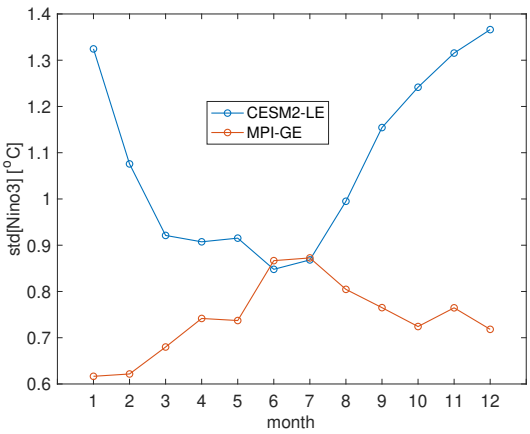


Figure S 8: Historical seasonal dependence of the standard deviation of Niño3 in two Earth system models.

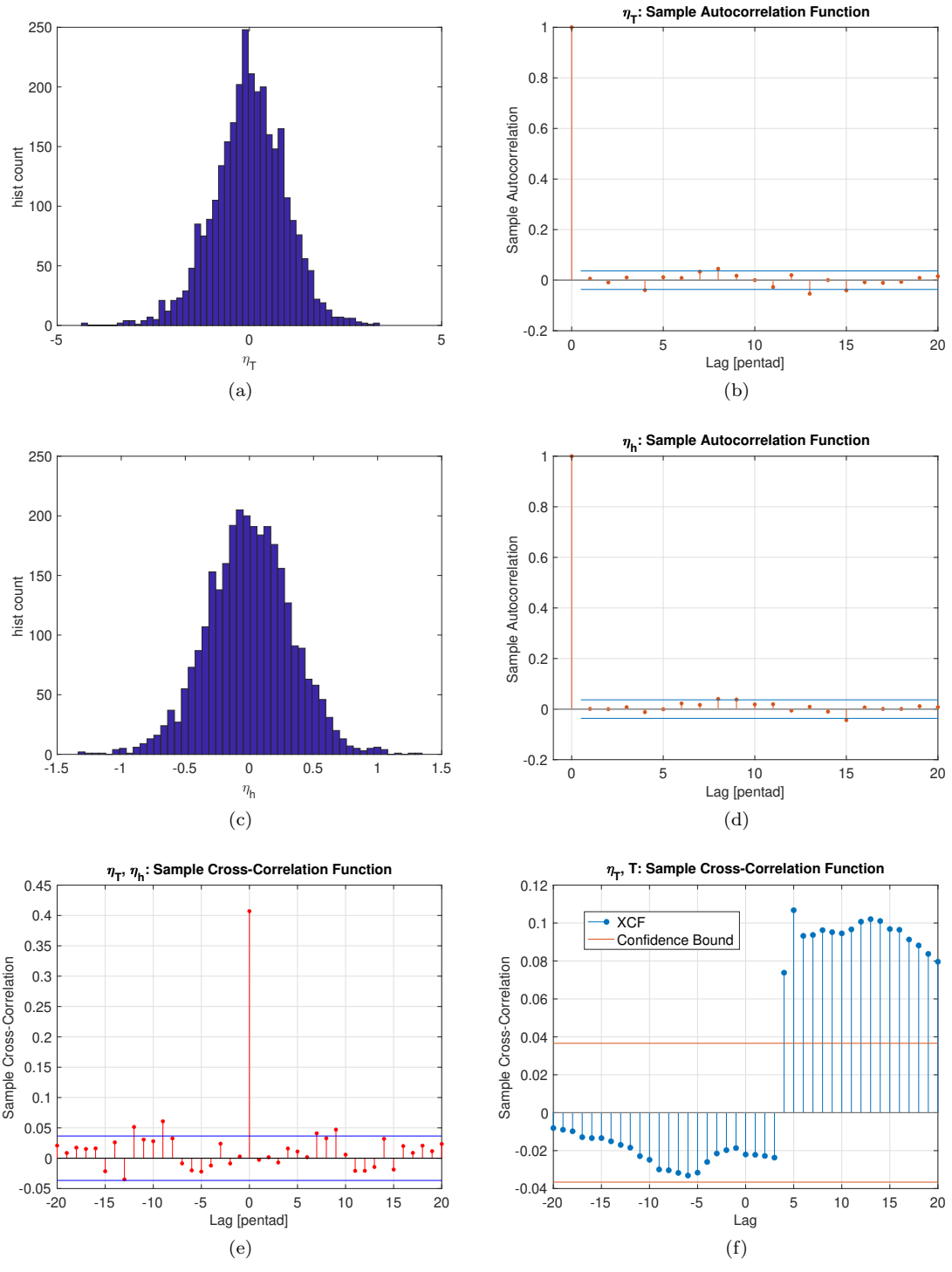


Figure S 9: Diagnostics of the fitted VAR(3) model to the residuals ξ_T, ξ_h corresponding to Fig. S7: normality and auto- and cross-correlation of residuals η_T, η_h. In show of improvement: there are (typically) no significant auto- and cross-correlations.

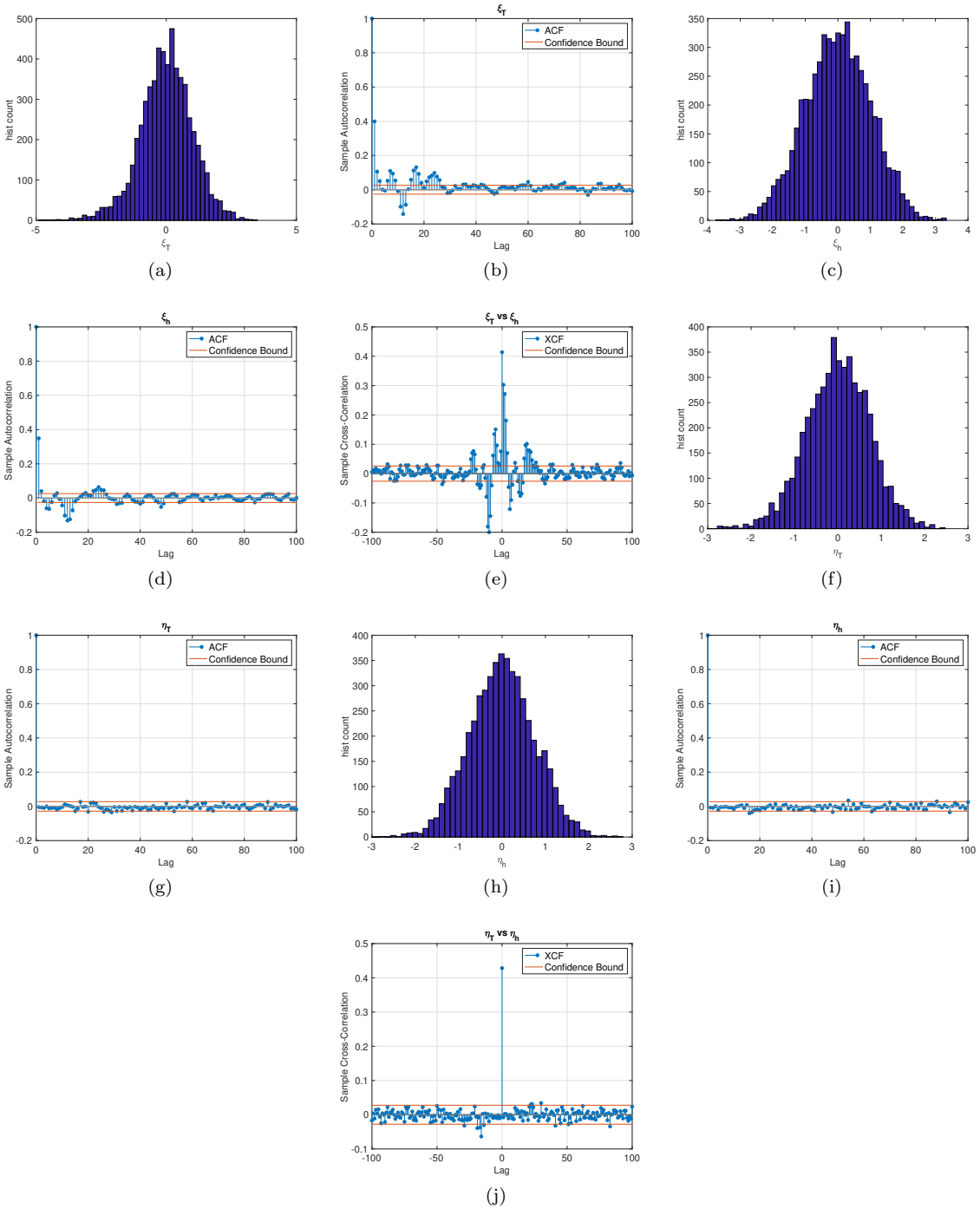


Figure S 10: Diagnostics of the fitted ROM1 model (1a)-(1b) to the monthly CESM2-LE data, I.

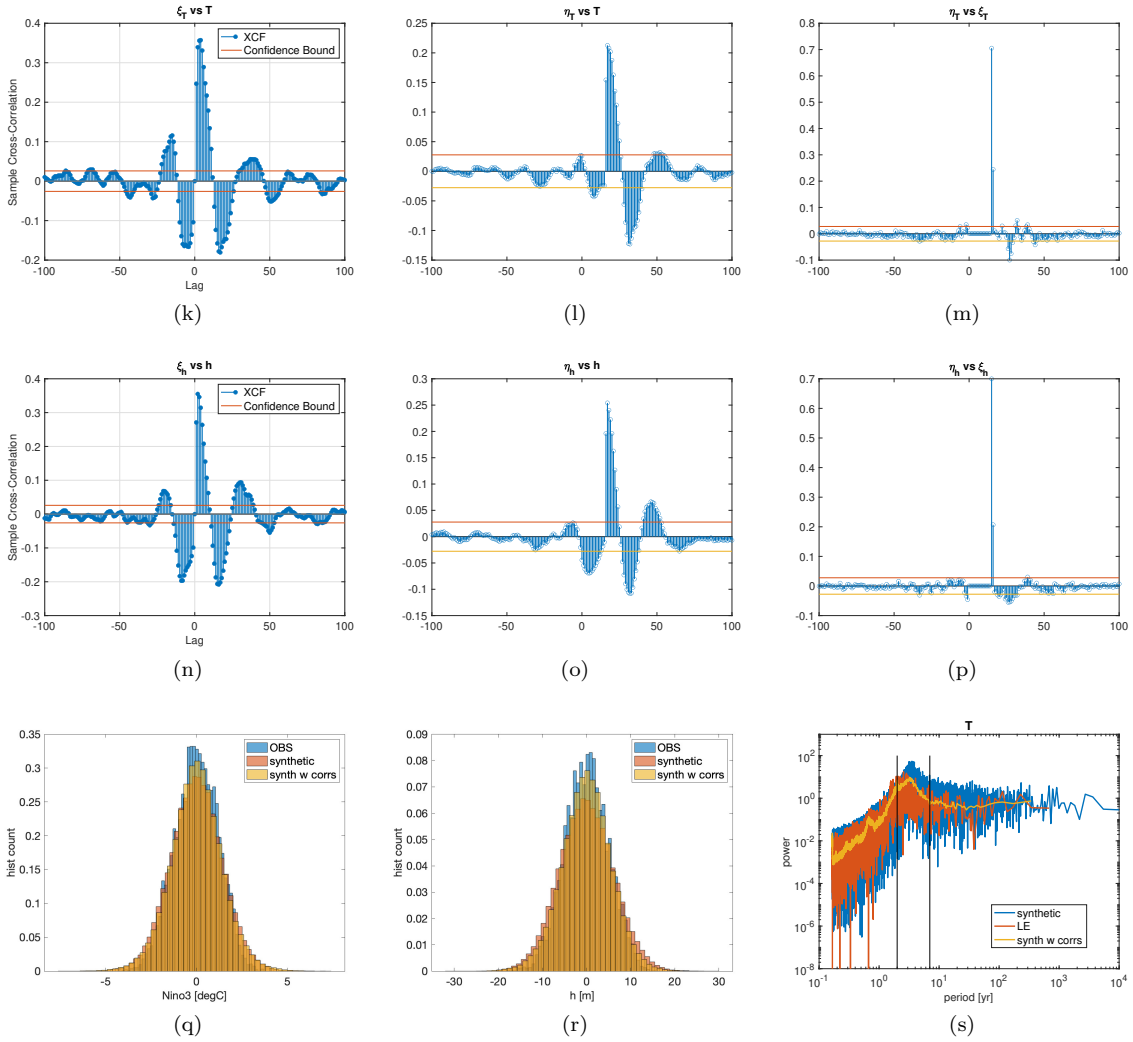


Figure S 10: Diagnostics of the fitted ROM1 model (1a)-(1b) to the monthly CESM2-LE data, II. Notice how a VAR($p \gg 3$) was used as opposed to the case of observational data (Fig. S9). This was needed to have not much significant causal effect of T on η_T as seen in panel (l).

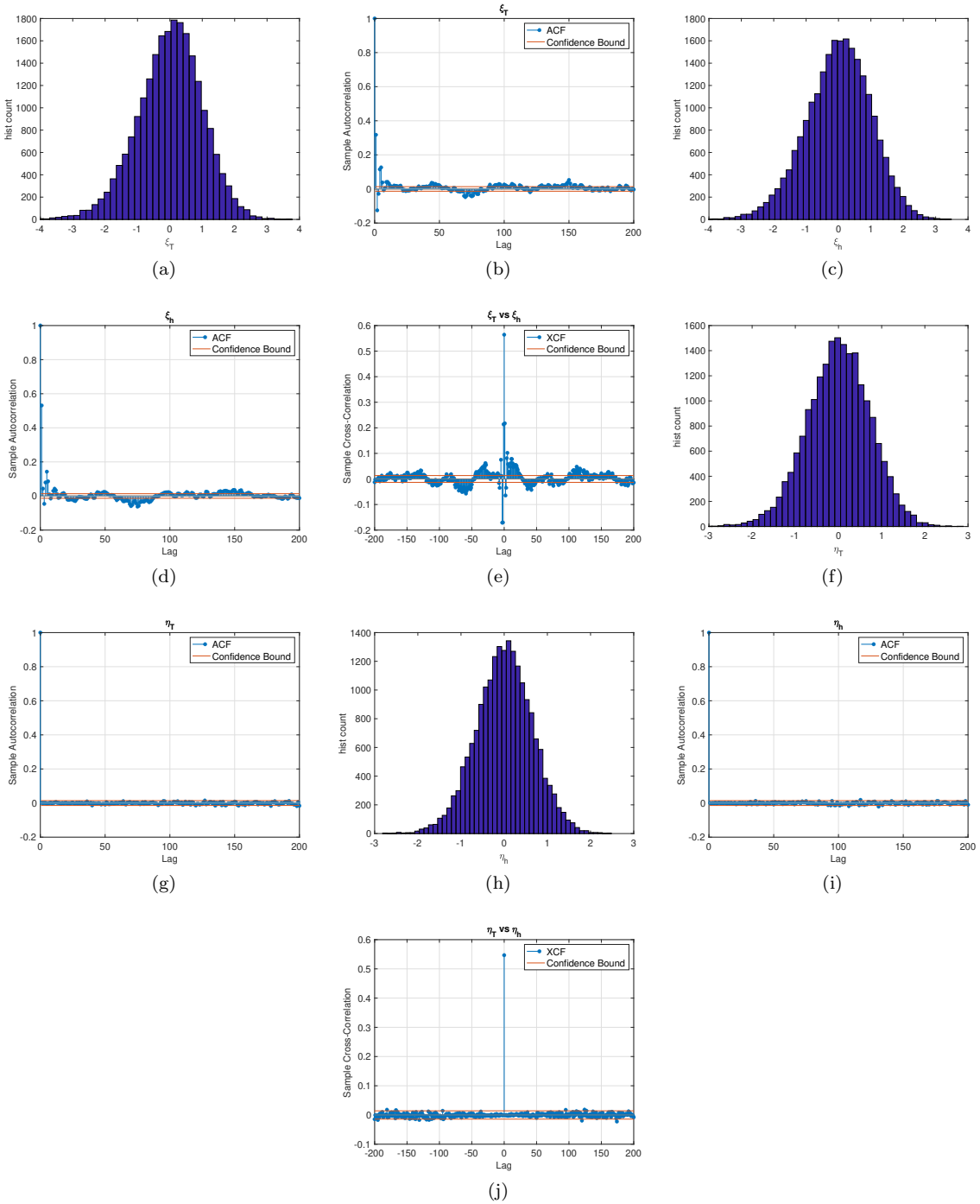


Figure S 11: Diagnostics of the fitted ROM1 model (1a)-(1b) to the pentad CESM2-LE data, I.

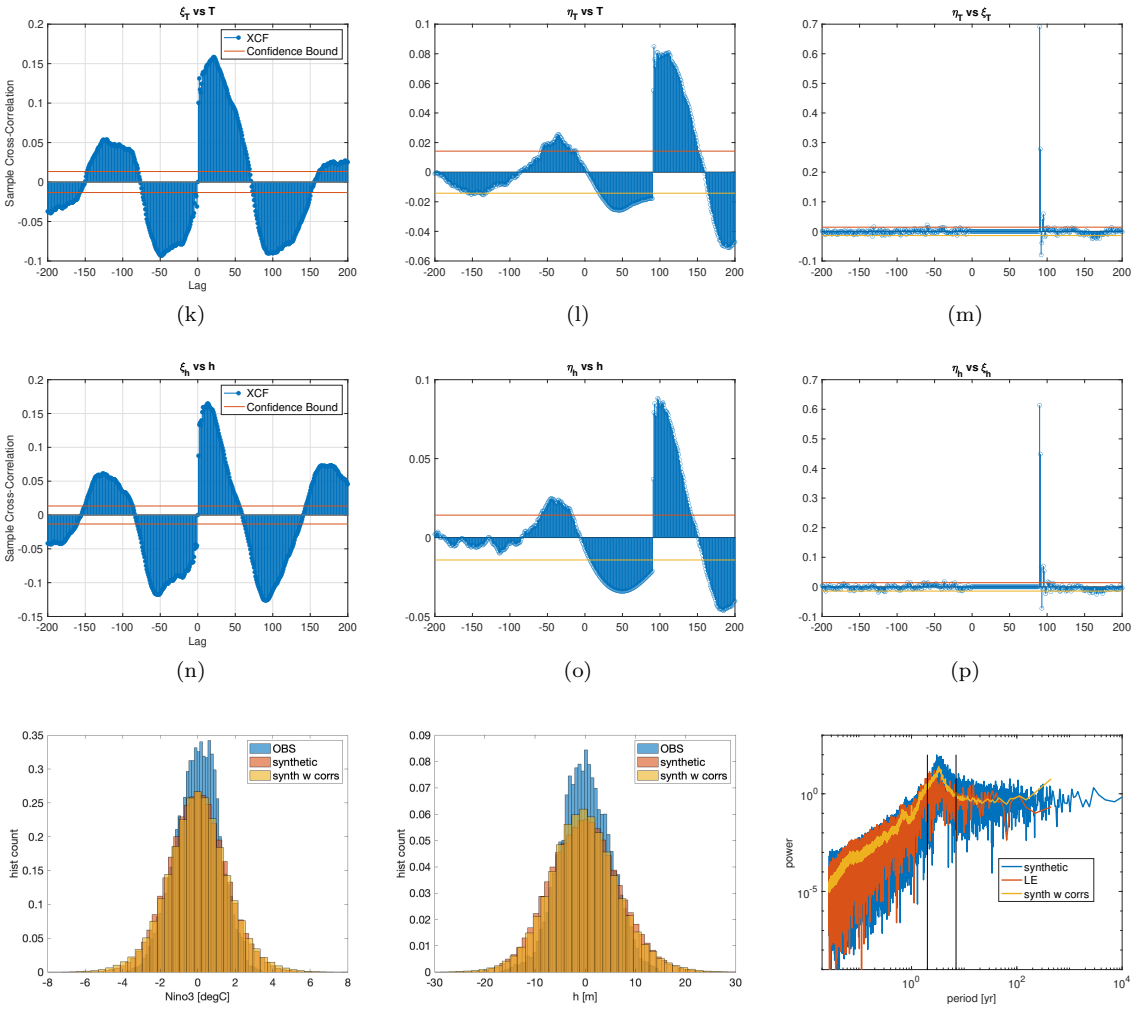


Figure S 11: Diagnostics of the fitted ROM1 model (1a)-(1b) to the pentad CESM2-LE data, II.

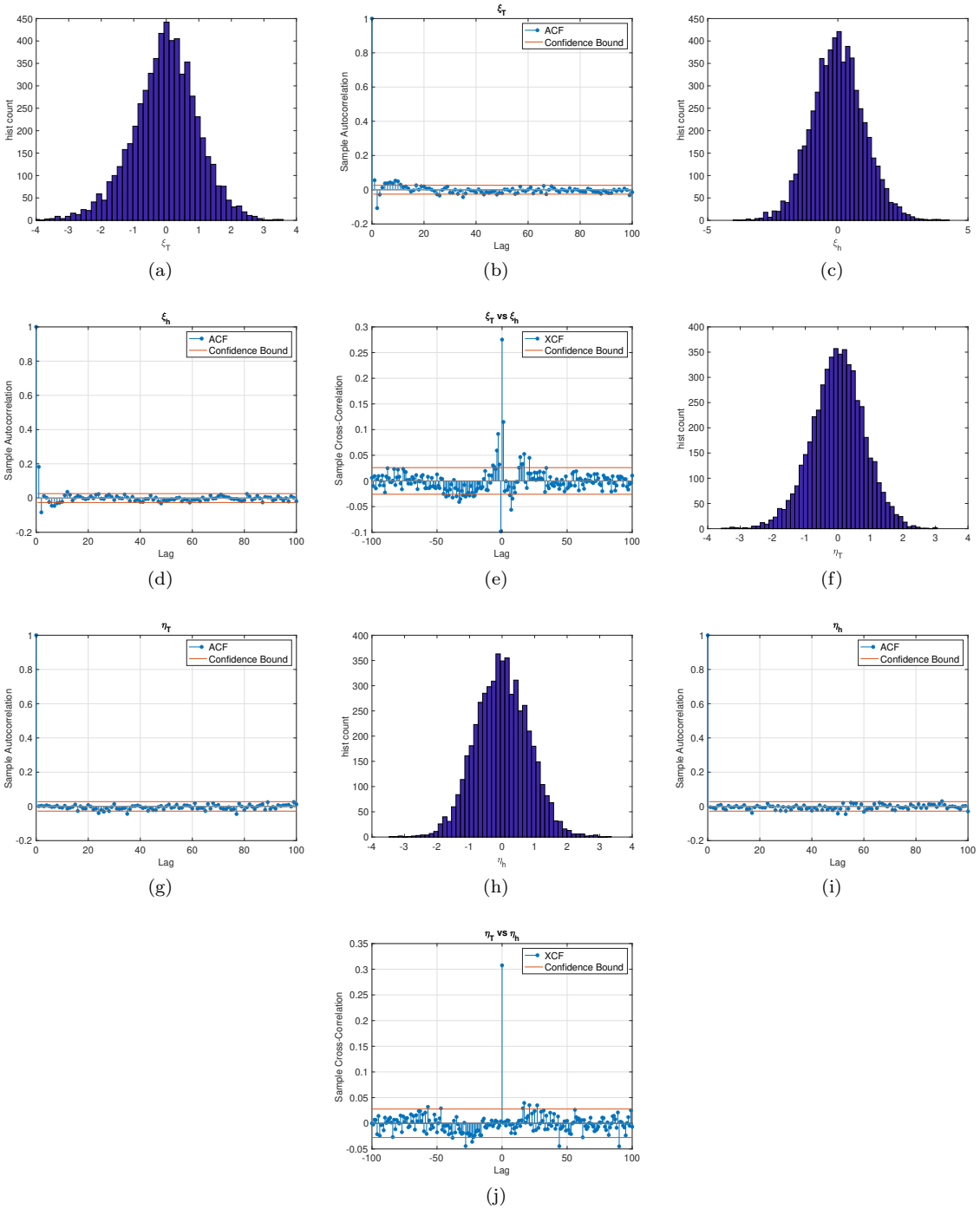


Figure S 12: Diagnostics of the fitted ROM1 model (1a)-(1b) to the monthly MPI-GE data, I.

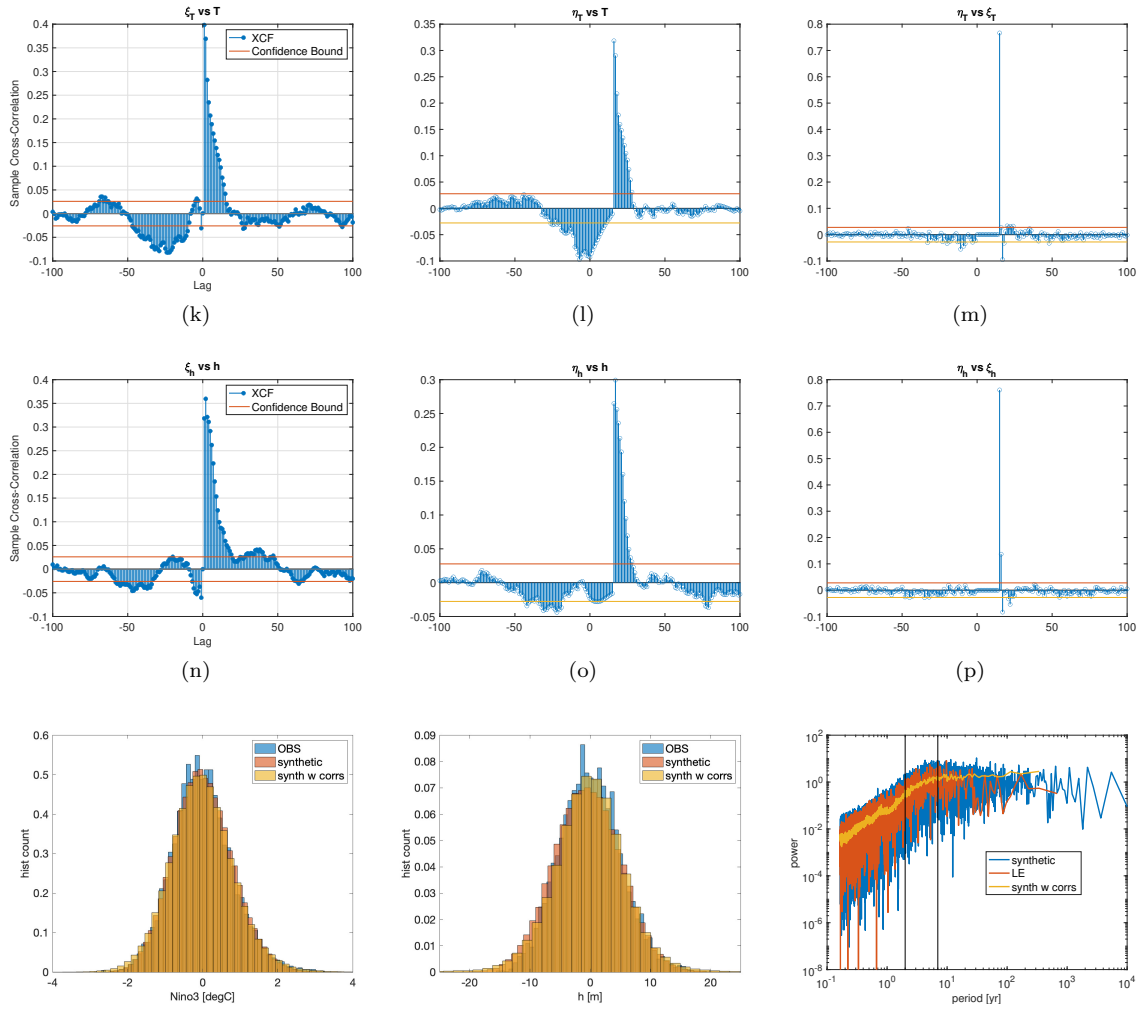


Figure S 12: Diagnostics of the fitted ROM1 model (1a)-(1b) to the monthly MPI-GE data, II.

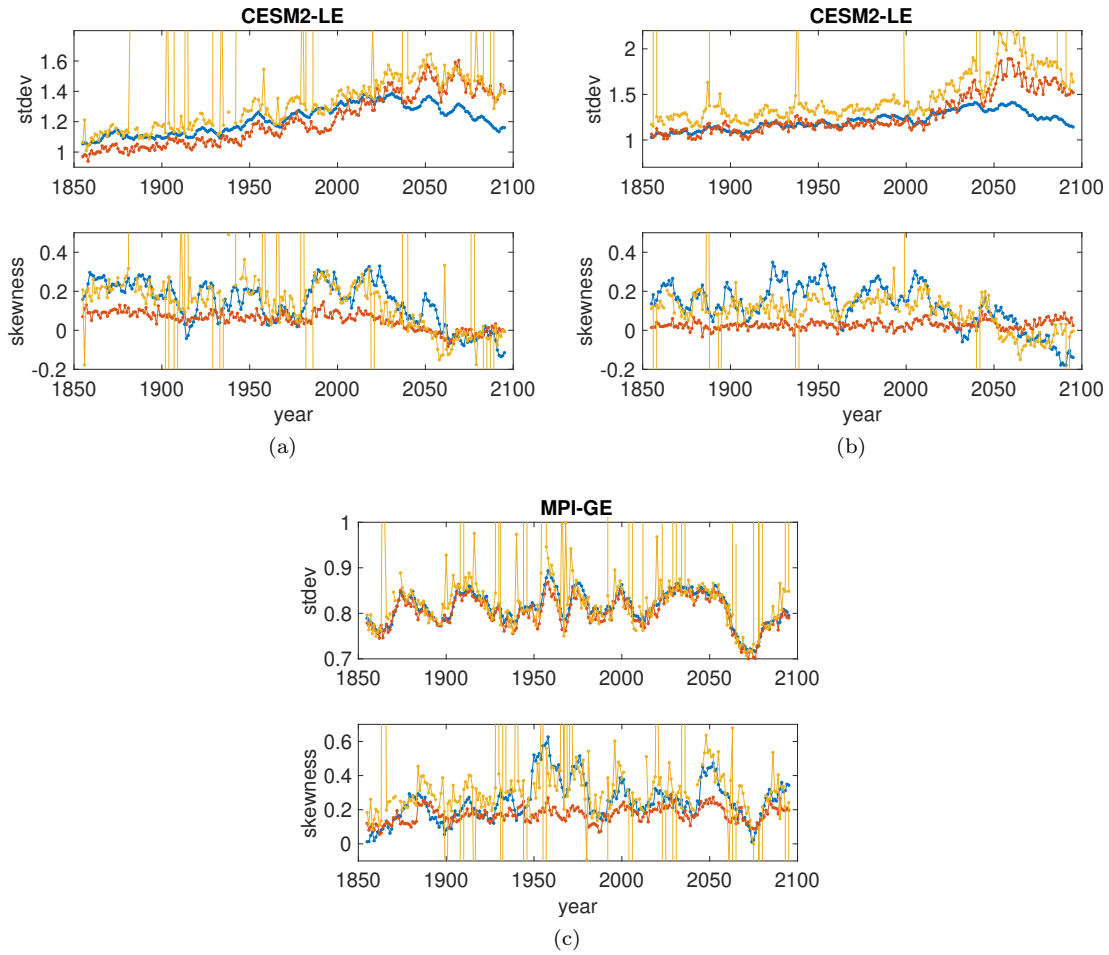


Figure S 13: Diagnostics of the fitted ROM1 model (1a)-(1b) to the CESM2-LE and MPI-GE data sets corresponding to Fig. 4, III. Panel(s) (b) ((a) and (c)) is (are) based on pentad (monthly) data. Curves in blue, vermillion and gold belong in that order to the direct estimate from the LE data (copying curves from Fig. 4), skewness of the ROM1 without- and with taking noise correlations into account. For the ROM1 results, the skewness was estimated from simulations of spans of 10,000 years using a time step Δt the same as the data resolution from which the ROM1 parameters were estimated. We do not know why the ROM1 fit to the CESM2-LE data is not yielding accurate standard deviation for the strongest of forcing magnitudes examined. Vertical lines of the gold curve are due to outliers as a result of a divergent simulation. The latter is probably due to a too large Δt (read the captions of Figs. 2, S4) and/or parameter value combinations not permitted for stable dynamics, possibly because of the unreliability of the parameter estimates. The latter might very well be down to "bifurcations" suffered by the (nonautonomous) system – the Earth system model is indeed a much more intricate system than the ROM1.

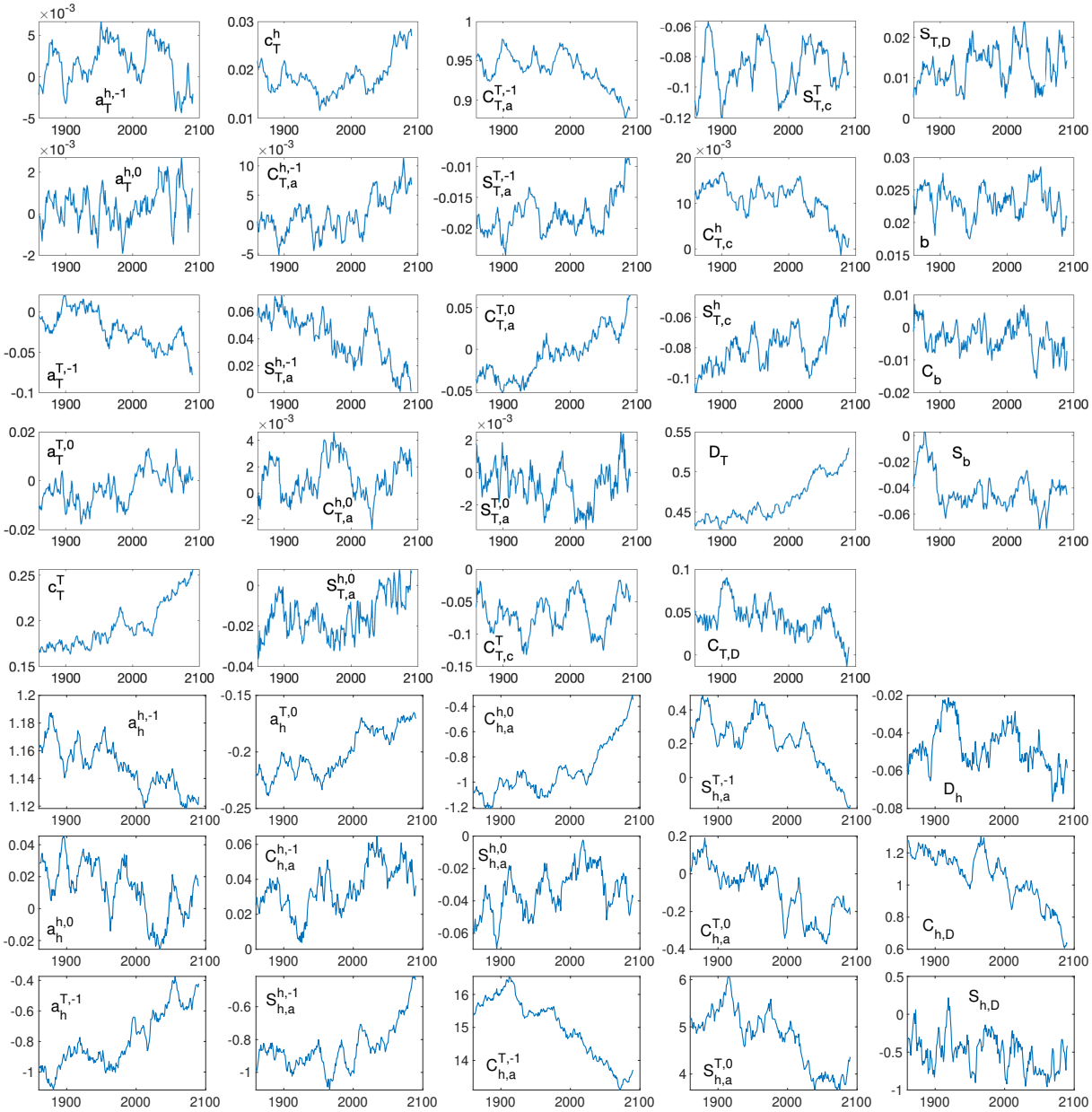


Figure S 14: Like Fig. 4, but for the *parameters* of the DROM+ (3) fitted to the MPI-GE data.

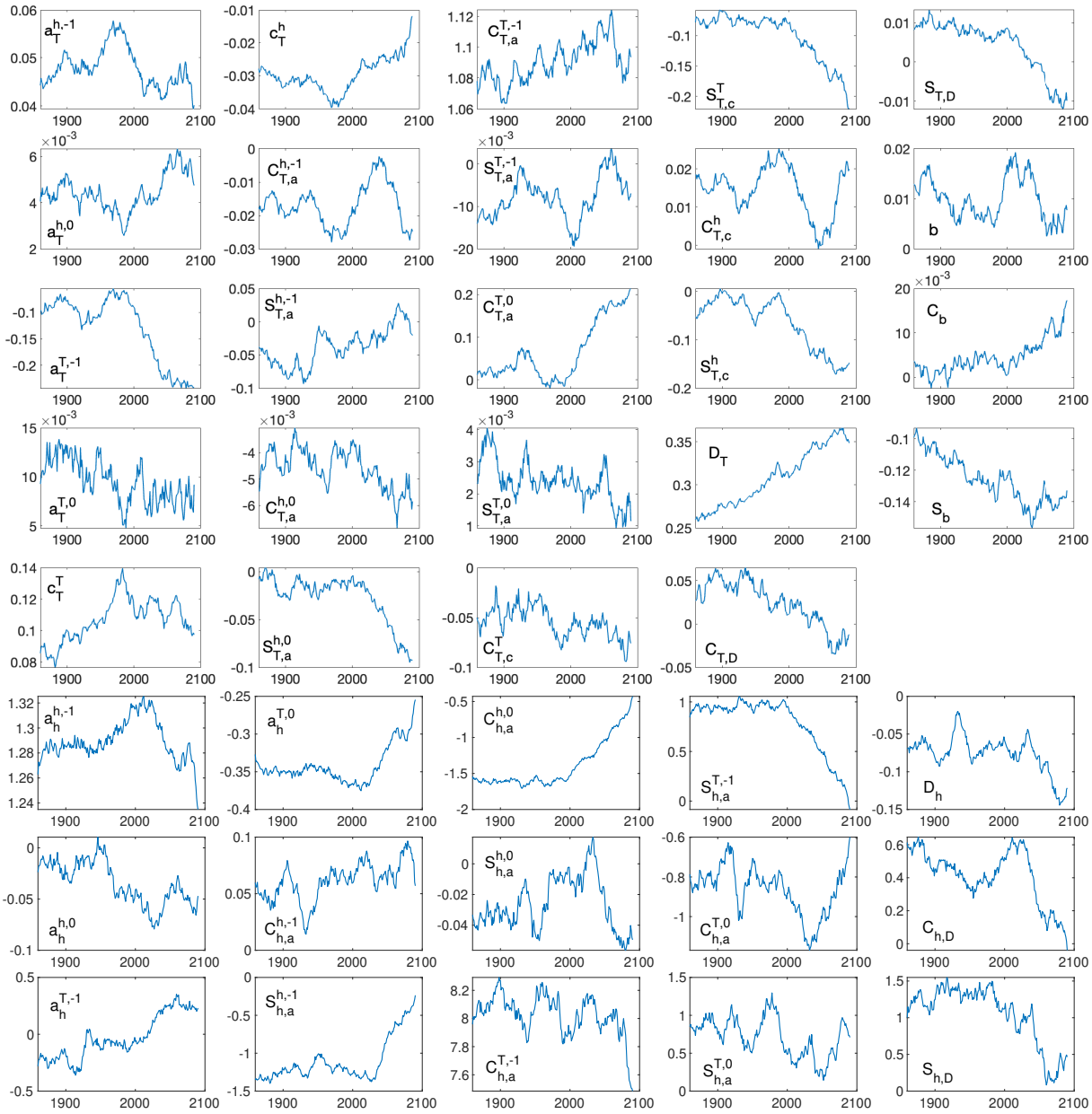


Figure S 15: Like Fig. 4, but for the *parameters* of the DROM+ (3) fitted to the CESM2-LE data.

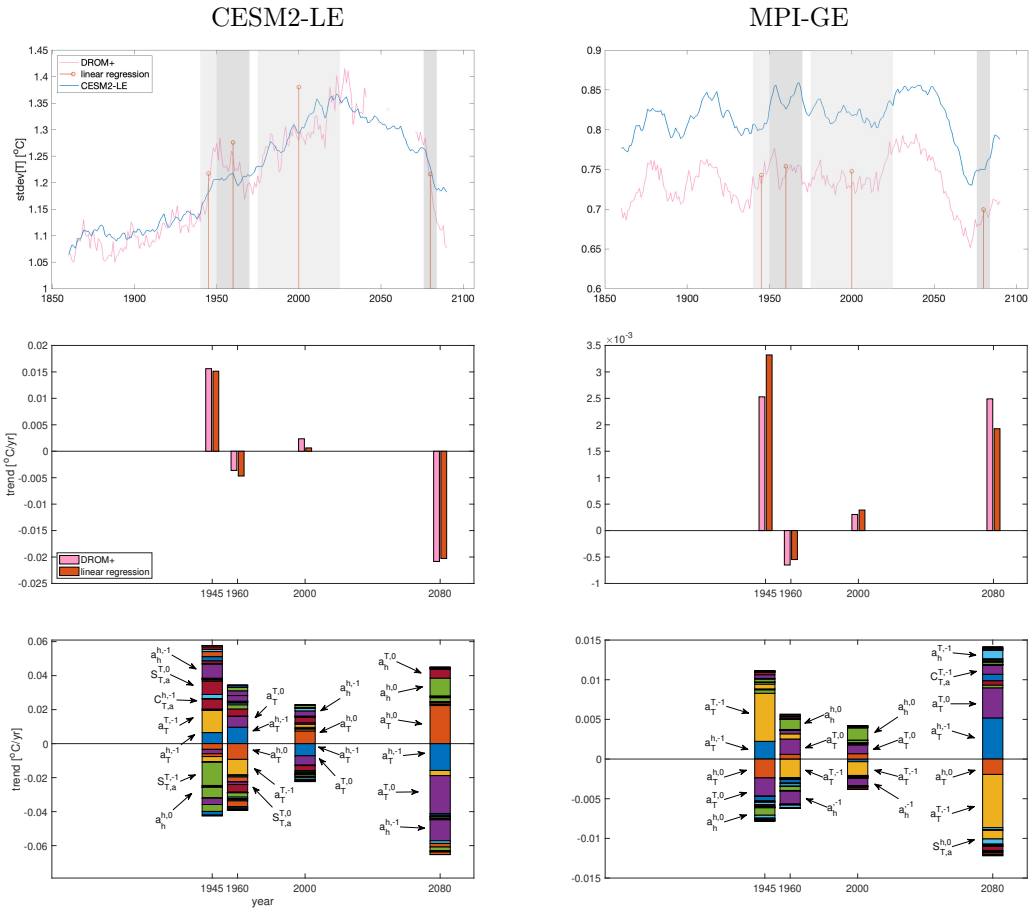


Figure S 16: Forced changes of ENSO standard deviation, std$[T]$, in view of the DROM+, $D = 1$. Top: Time series of std$[T](t)$, directly evaluated for the CESMv2 and MPI-ESM, and that of the DROM+ fitted to the respective LE data. Middle: Linear trends of std$[T](t)$ in time windows indicated by grey shaded strips in the diagrams of the time series above. It is that of the DROM+ fitted to the respective LE data as well as trends approximated by linear regression (Methods 4.4). Bottom: contributions to the linear trend of std$[T](t)$ by trends of parameters of the DROM+ determined using the regression coefficients (Methods 4.4).

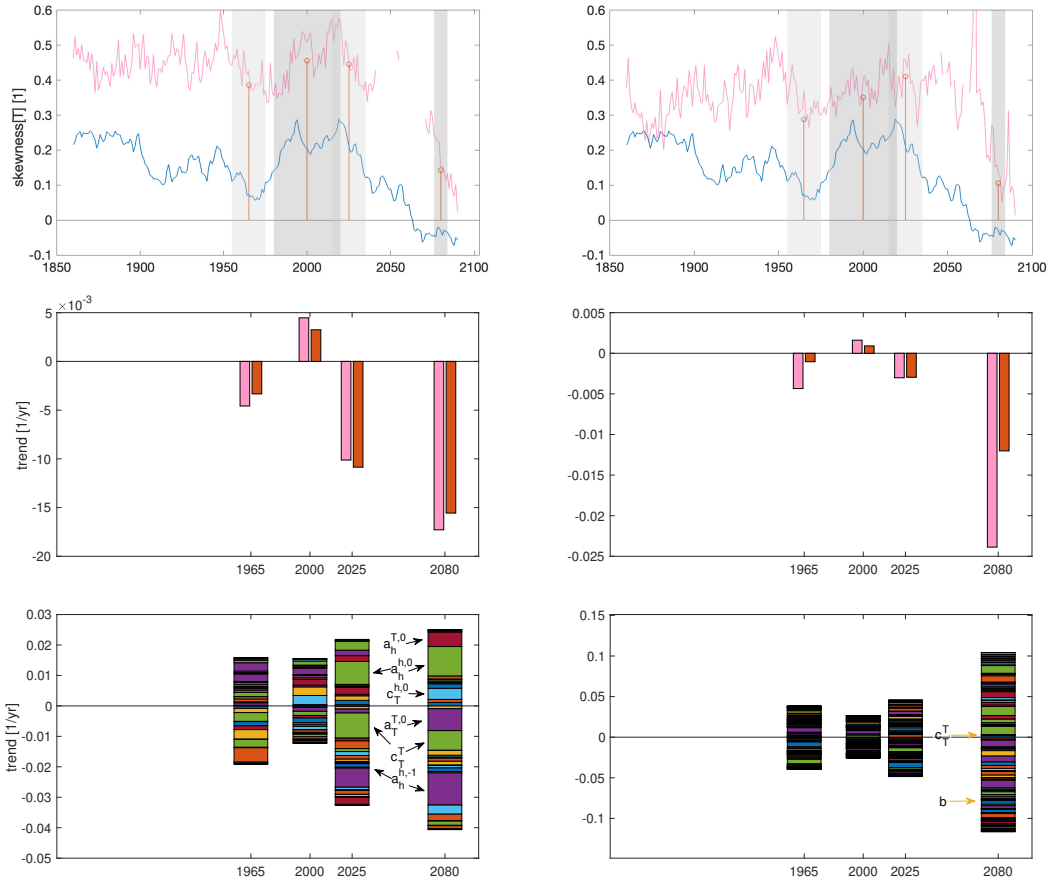


Figure S 17: Same as Fig. 16 but for the skewness, and considering the CESMv2 only. Left: $D = 1$; right: $D = 2$ [month].

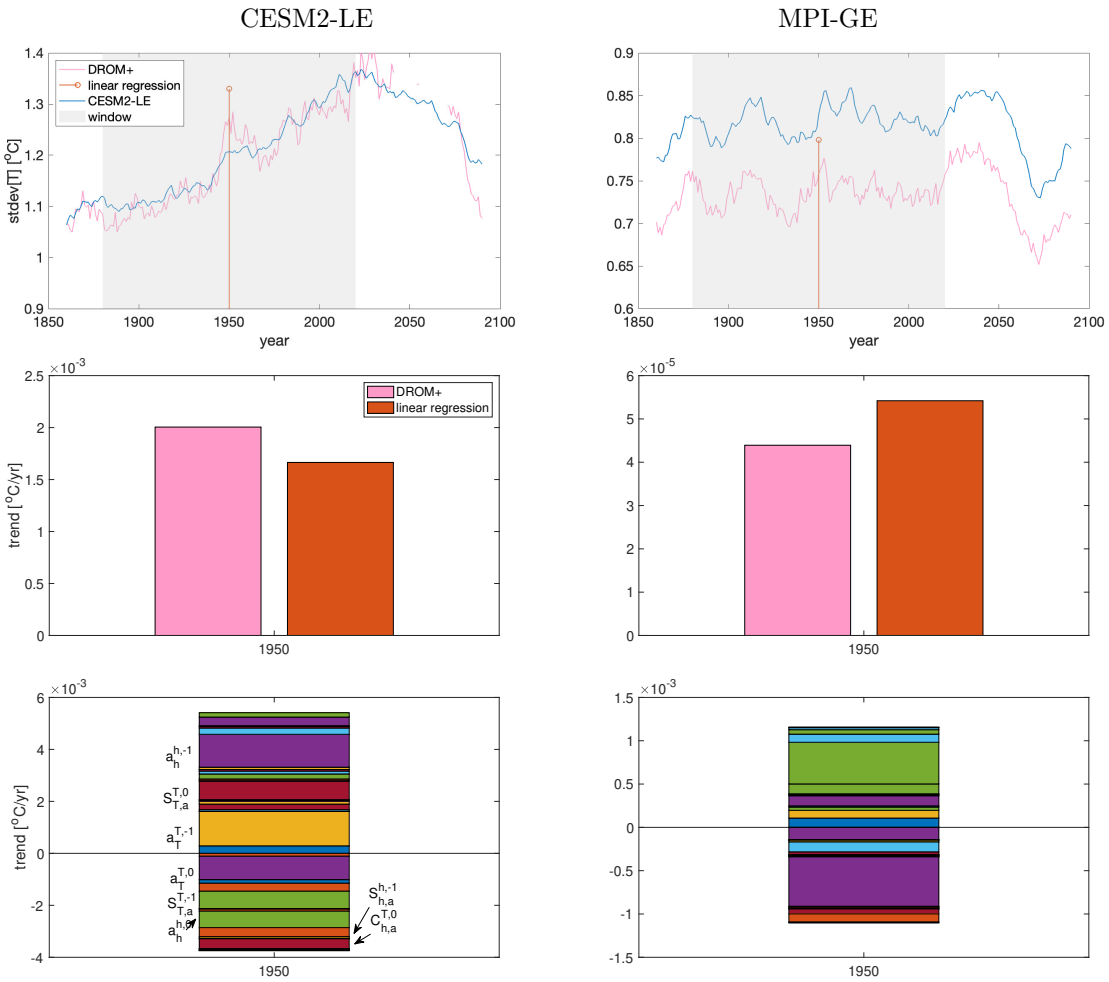


Figure S 18: Same as Fig. 16 but trends (middle) and contributions to them (bottom) are evaluated for a rather wide time window (top).

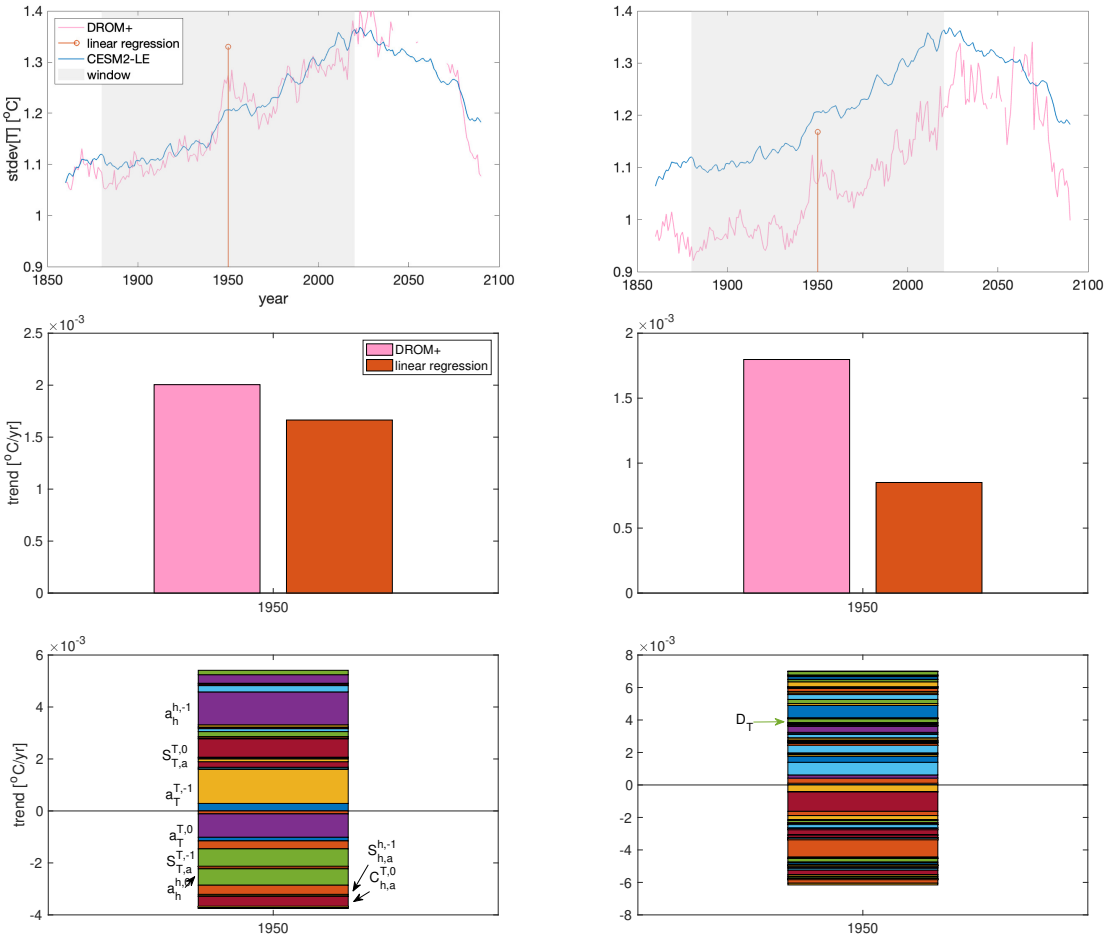


Figure S 19: Same as Fig. S18 but only for the CESMv2. This is to compare the cases of $D = 1$ (left) and 4 [month] (right).

Sl. N0.	Model	No. of Members	Period	Scenario
1	CESM2	100	1850-2100	SSP 370
2	MPI-ESM	63	1850-2100	RCP 8.5, RCP4.5
3	MIROC6	50	1850-2100	SSP5-8.5
4	CanESM5	50	1850-2100	SSP585, SSP370, SSP245
5	CanESM2	50	1950-2100	RCP 8.5
6	CESM1	40	1920-2100	RCP 8.5
7	CSIRO-MK3.6	30	1850-2100	RCP 8.5
8	GFDL ESM2M	30	1950-2100	RCP 8.5
9	GFDL-CM3	20	1920-2100	RCP8.5
10	MIROC-ES2L	10	1850-2100	SSP-585, SSP370, SSP245
11	MPI-ESM1-2-LR	10	1850-2100	SSP-585, SSP370, SSP245
12	ACCESS-ESM1-5	10	1850-2300	SSP-585, SSP370, SSP245

Table 1: Data sets of initial condition large ensembles used to analyse the modelling robustness of ENSO variance (amplitude) in Figs. S20 and S21. Annual mean data is used.

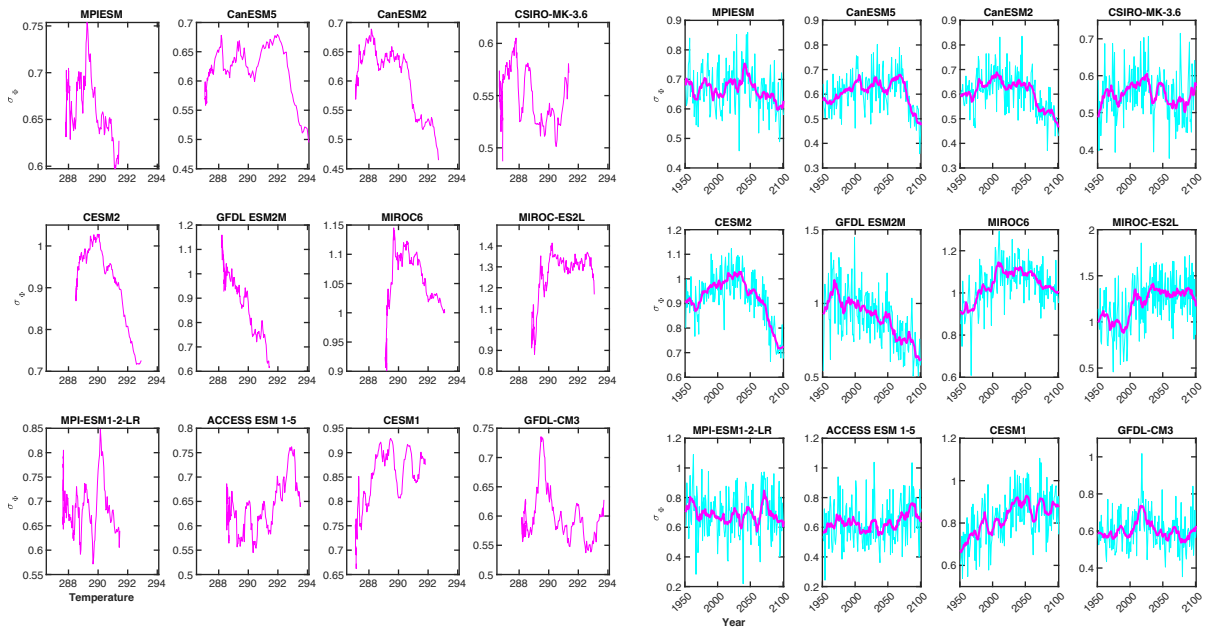


Figure S 20: ENSO standard deviation (σ_Φ) as a function of time being the forced change (left) or the forced change of the global mean surface temperature (right). Only data belonging to the strongest forcing scenario used to create these diagrams.

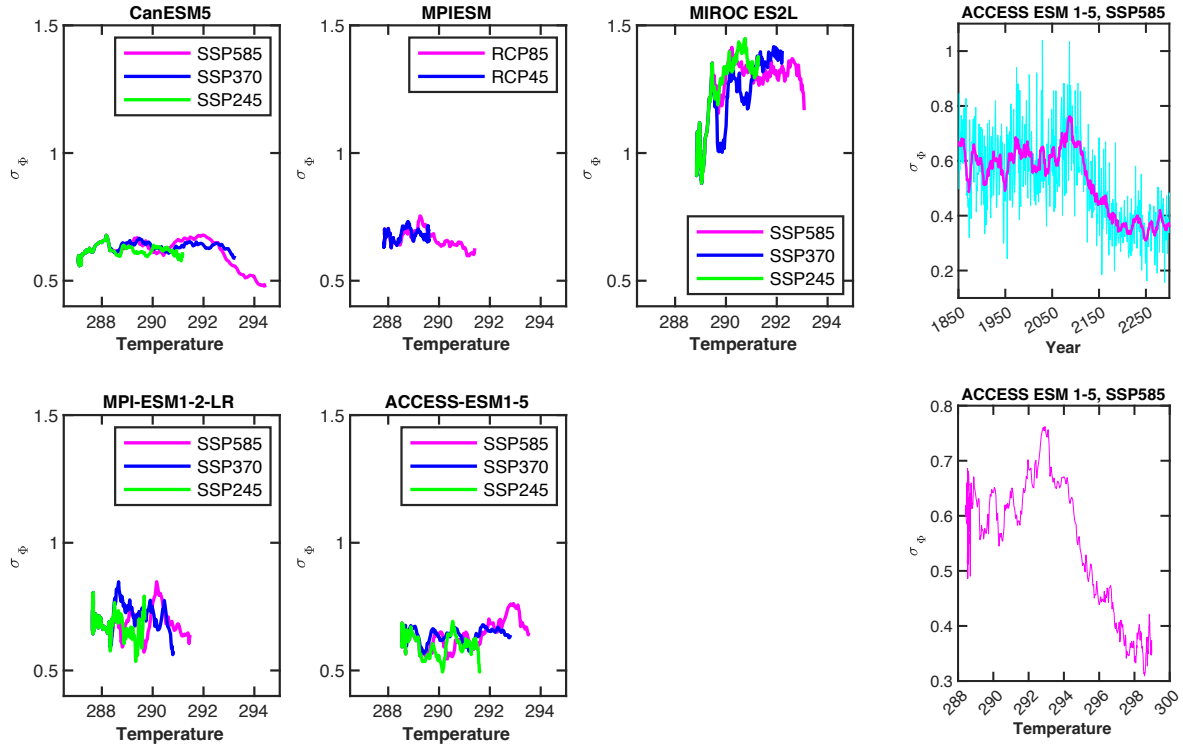


Figure S 21: Left: Similar to Fig. S20 but all available data is used as indicated in Table S1. Only those models are featured for which more than one forcing scenarios are applied. Extreme right: For one particular model, we have a run extended out to the year 2300. Here we can see an eventual decline of ENSO amplitude also for this model.