

Synchronization and causality across time-scales: Complex dynamics and extremes in El Niño/Southern Oscillation

Nikola Jajcay^{1,2}, Sergey Kravtsov³, Anastasios A. Tsonis^{3,4} & Milan Paluš¹

¹*Department of Nonlinear Dynamics and Complex Systems, Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic.*

²*Department of Atmospheric Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic.*

³*Department of Mathematical Sciences, Atmospheric Science Group, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA.*

⁴*Hydrologic Research Center, San Diego, California, USA.*

Statistical inference of causal interactions and synchronization between dynamical phenomena evolving on different temporal scales is of vital importance for better understanding of underlying mechanisms and a key for modelling and prediction in natural complex systems such as the Earth's climate or human brain. This communication introduces and applies information theory diagnostics to phase and amplitude time series of different oscillatory components of observed data that characterizes El Niño/Southern Oscillation (ENSO). A suite of significant interactions between processes operating on different time scales were detected and shown to be important for emergence of extreme events. The mechanisms of these nonlinear interactions were further studied in conceptual low-order and state-of-the-art dy-

namical, as well as statistical climate models. Observed and simulated interactions exhibit substantial discrepancies, whose understanding may be the key to an improved prediction of ENSO. Moreover, the statistical framework which we apply here is suitable for inference of cross-scale interactions in human brain dynamics and other complex systems.

Better understanding of dynamics in complex systems, such as the Earth's climate or human brain, is one of the key challenges for contemporary science and society. Large amount of experimental data requires new mathematical and computational approaches. Natural complex systems vary on many temporal and spatial scales, often exhibiting recurring patterns and quasi-oscillatory phenomena. Data-driven approach to detection and recognition of relationships between subsystems in complex systems has recently become an area of active study in a range of scientific fields. Considering e.g. the climate system as a complex network of interacting subsystems¹ is a new paradigm bringing new data analysis methods helping to detect, describe and predict atmospheric phenomena². A crucial step in constructing climate networks is inference of network links between climate subsystems³. Directed links determine which subsystems influence other subsystems, i.e. uncover the drivers of atmospheric phenomena. A concise study of causal relationships and synchronization phenomena in complex systems could significantly improve understanding and predictability of such systems. Here, we investigate complex, multiple time-scale interactions in the El Niño/Southern Oscillation system in the equatorial Pacific.

El Niño/Southern Oscillation (ENSO hereafter) is a well known coupled ocean–atmosphere phenomenon, which manifests as a quasi-periodic fluctuation in sea surface temperature (El Niño)

and air pressure of the overlying atmosphere (Southern Oscillation) across the equatorial Pacific Ocean. It is comprised of two phases - the warm phase, known as *El Niño*, accompanied by high surface pressure and the cool phase, *La Niña*, accompanied by low surface pressure in the tropical western Pacific. Although the exact causes for initiating warm or cool events are not fully understood, the two components of ENSO – sea surface temperature and atmospheric pressure – are strongly related. ENSO dominates interannual climate variability⁴, strong ENSO events can cause changes in the tropical Pacific that exert a large influence on the global atmospheric circulation⁵ via associated teleconnections⁶, thus leading to significant socio-economic impacts. An important aspect of ENSO extreme events is that their positive phase, El Niño, exhibits larger magnitude than their negative phase, La Niña^{7,8}. This suggests that, at least to some extent, ENSO dynamics involve nonlinear processes^{4,9}.

ENSO events occur irregularly, with a 2–7 year span between them, but have a well defined spatial pattern and seasonal dependence, with the start of development in boreal summer and a peak in boreal winter¹⁰. The individual events generally evolve on a time scale of about two years^{5,11}; hence, the biennial variability manifesting as year-to-year alternations between “weak” and “strong” annual cycles in the Indian ocean and western Pacific sector¹², with possible global repercussions¹³, is bound to interact with ENSO. Indeed, the spectral analysis of various ENSO indices¹⁴ identifies quasi-quadrinial and quasi-biennial signals in the equatorial Pacific. In summary, ENSO is known to centrally involve processes on three distinct time scales, namely the ones associated with the annual cycle (AC), quasi-biennial (QB) processes, and low-frequency (LF) interannual processes^{15–17}.

Statistical inference of causal relationships within climate data has recently become an area of active research^{18–21}. Typically, a causal relation is sought between pairs of different variables or different modes of atmospheric variability. By contrast, Paluš^{22,23} suggested an approach — which we will follow in the present study — to examine the complexity of climatic interactions by uncovering causal relations between processes operating on different time scales within a single climatic time series. In this study, we apply this technique to discern multi-scale interactions and causality in ENSO, as represented by the NINO3.4 index (spatial average of sea surface temperatures over a box of 5°S–5°N and 170°–120°W), in observations and climate model simulations.

1 Interactions in the Observed ENSO

We examined synchronization and causal interactions in the NINO3.4 time series for the quasi-oscillatory modes with periods ranging from 5 to 96 months. We are looking for the observed interactions characterized by causality estimate exceeding the 95th percentile of the distribution of this quantity for the surrogate data samples (Fig. 1), where the surrogate time series²⁴ were generated using a Monte Carlo method, yielding synthetic time series with the same spectrum, but void of any cross-scale interactions (see Methods for more details on surrogate data).

There are three pairs of modes that exhibit phase synchronization (a process by which two or more cyclic signals tend to oscillate with a repeating sequence of relative phase angles) (Fig. 1, left). The quasi-biennial (QB) modes (periods of 1.8–2.1 yr) are synchronized with the annual cycle (AC) as well as with the combination tones (CT; periods approximately 9 and 14 months).

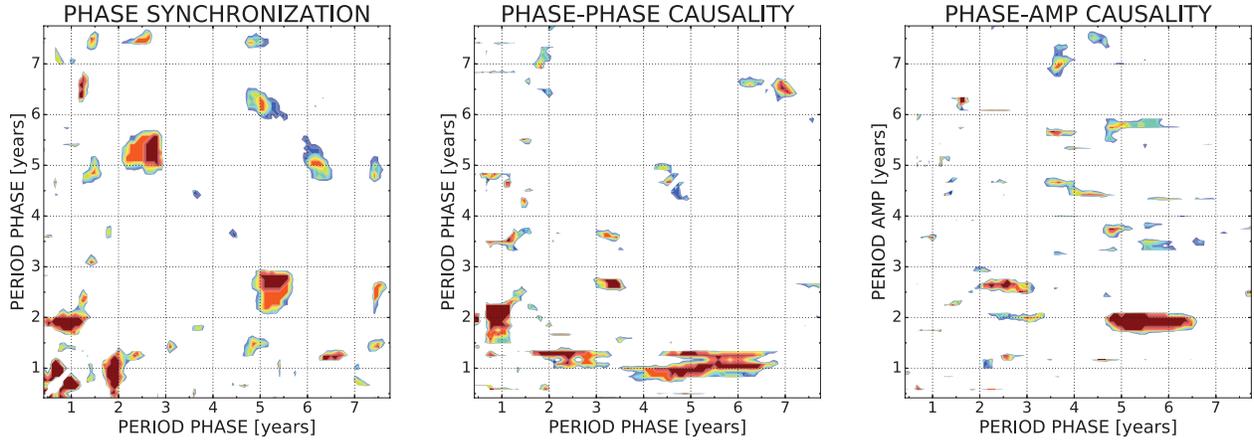


Figure 1: Cross-scale phase synchronization (left), phase–phase causality (middle) and phase–amplitude causality (right) in the observed NINO3.4 time series. The phase synchronization is a symmetrical relation, hence the plot is symmetric, while causality plots are shown with the period of the driver (master) time-series on the x -axis and driven (slave) time-series on the y -axis. Shown are (positive) significance-level deviations from the 95th percentile of the k -nearest neighbour estimates of (conditional) mutual information, tested using 500 Fourier transform surrogates (see text).

The combination tones result from the interaction of AC with low-frequency (LF) ENSO modes of periods 4–6 yr (see, e.g., Stein et al.²⁵ or Stuecker et al.²⁶). Furthermore, the AC and CT modes with periods of 8–9 months are themselves phase locked. Finally, the LF modes with periods 5–6 yr and QB modes with periods 2–3 yr also exhibit phase synchronization. These results reconfirm an important role of the annual cycle in ENSO dynamics, with strong ENSO events peaking in boreal winter^{10,27}, and point to the link between QB and LF modes which may be responsible for extreme ENSO events^{14–16} (see below).

The phase–phase causality diagnostics (Fig. 1, middle) brings in an additional information that complement the phase synchronization results, and elucidates important directed connections

between the LF, QB and AC/CT ENSO modes. In particular, the phase of LF modes affects that of the AC/CT modes, which means that the “shape” of the annual cycle depends on whether the LF mode (periods of 4–6 yr) is in its extreme warm or extreme cold phase.

Furthermore, the phase of QB modes with 2–3-yr period dictates in part the phase of the CT modes (periods of 12–16 months), while the phase of QB modes with the periods of 1.8–2.1 yr is partially slaved to the phase of the AC mode. The only pronounced phase–amplitude causality link (Fig. 1, right) is the one between the phase of the LF ENSO mode (periods of 5–6 yr) and the amplitude of QB modes (periods of 1.8–2.1 yr).

Our analysis thus identifies the three fundamental time scales in ENSO dynamics — AC, QB and LF — consistent with previous work^{15–17}, but offers further details on the interaction between these modes. Based on our results, it is natural to consider the AC and CT processes in combination to define the quasi-annual (QA) variability. The QB modes can be divided into two — ‘faster’ and ‘slower’ — sub-ranges, with the periods of 1.8–2.1 yr and 2–3 yr, respectively. Similarly, the LF processes can be divided into the ones associated with 4–5-yr and 5–6-yr periods.

We observe a pronounced connection between the (phase of) the slower LF mode and both the phase and amplitude of the faster QB mode. In particular, the slower LF mode affects the phase of the QA mode, and, therefore, — indirectly — the phase of the faster QB mode, which tends to be affected by and phase-synchronized with the QA mode; the slow LF mode also directly affects the amplitude of the faster QB mode. The connections between the phase of slow LF mode and the phase of QA mode important in the causal sequence above are both direct and indirect. In the

latter indirect case, the connection works through the phase synchronization between the slow LF mode and the slow QB mode and subsequent causal effect of the latter on the phase of the QA mode. The faster LF modes add to the picture by also affecting the phase of the QA mode, and, therefore, indirectly, the phase of the faster QB mode.

Extreme ENSO events The interactions identified above are instrumental in setting up extreme ENSO events (Fig. 2). In particular, during all of the strong El Niño events of years 72/73, 82/83 and 97/98, the QA, QB and LF modes were characterized by synchronous pronounced maxima. By contrast, during a moderate El Niño of 87/88, the three modes exhibited phase shifts, with lower-frequency modes leading the higher-frequency modes instead of being ‘stacked’ on top of one another, thus limiting the magnitude of this event. Notably, strong La Niña events do not seem to be associated with the minimum of the LF mode, but instead occur during near-neutral LF conditions when the minima of the QA modes and the minima of the whole range of QB modes synchronize. Thus, in both El Niño and La Niña cases, the behaviour of the QB modes has a vital control on the magnitude of the ENSO events.

2 Interactions in the Simulated ENSO

Interactions in CMIP5 models The Coupled Model Intercomparison Project Phase 5 (CMIP5)²⁸ is a framework for coordinated climate change experiments providing global circulation model (GCM) outputs from various modelling groups. We analysed time series of the sea surface temperature obtained from the individual runs of CMIP5 models and compared the simulated NINO3.4

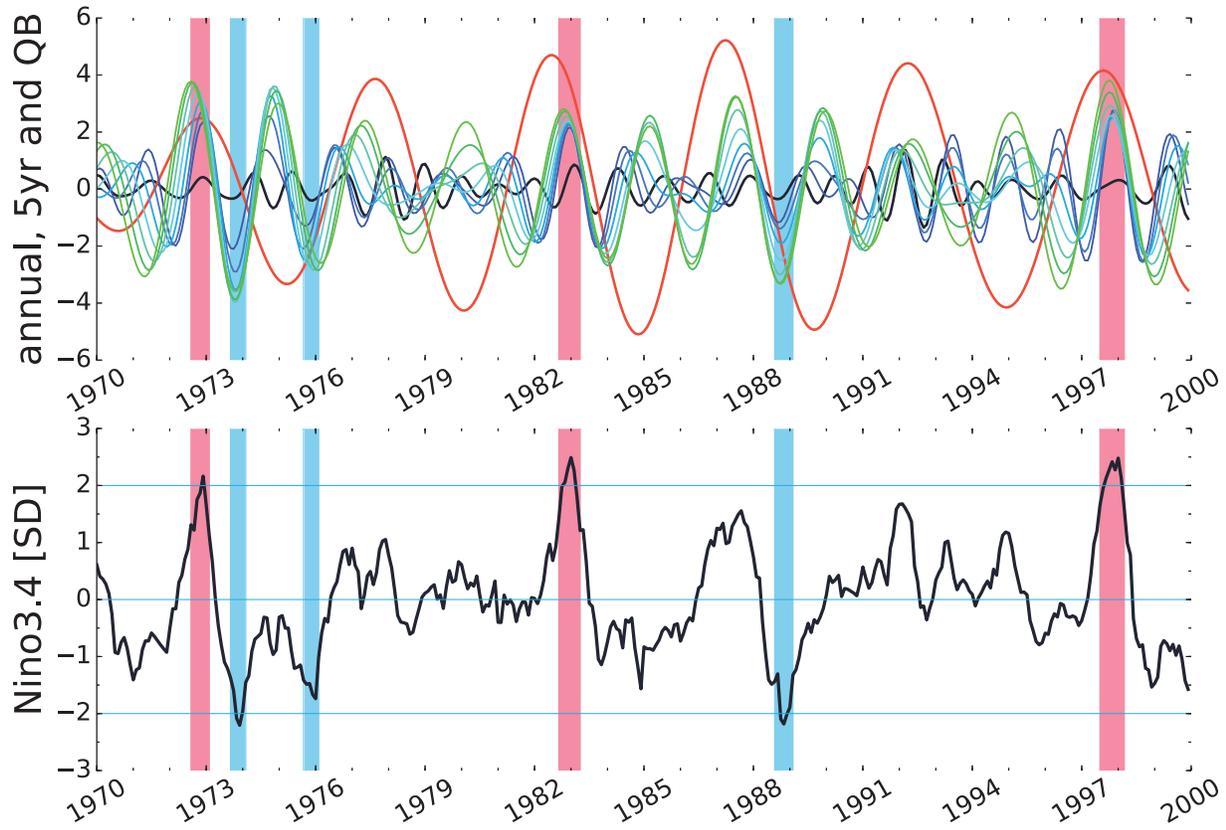


Figure 2: Top: wavelet reconstructions of the observed NINO3.4 index (1970–1999): reconstruction of the annual (black), quasi-biennial (for a range of periods from 18 to 30 months, with 2 month step; shades of blue to green) and 5-yr ENSO cycles (red). All of these reconstructions were computed via continuous complex wavelet transform (CCWT) as $A_p(t) \cos \phi_p(t)$, for the corresponding central wavelet periods p . Bottom: the full observed normalised NINO3.4 index. The years of strong El Niño and La Niña events are marked with the red and blue shading, respectively.

characteristics with the observed characteristics (Fig. 3). To start with, we measured the similarity between the observed and simulated wavelet spectra²⁹ using root-mean-square distance and Pearson correlation coefficient, with zero distance and unit correlation coefficient indicating the perfect match. The ensemble-mean values of these two measures for individual CMIP5 models are shown in the first two columns of Fig. 3 (see Supporting Information for the corresponding values from individual runs). The models exhibit great variations in their ability to match the observed NINO3.4 spectra, with correlation coefficients ranging from 0.2 to 0.8 and rms distances from 30 to 120; furthermore, there are also substantial sampling variations in the NINO3.4 spectra from multiple runs of a single model (see Supporting Information). This means that the ENSO tends to exhibit different epochs of sampling variability in models, in which its strength, spectrum and other properties may vary significantly³⁰.

Similarly to the wavelet spectra, the synchronization and causality maps (see examples in Fig. 4 analogous to the observed maps of Fig. 1), vary considerably from model to model (Fig. 3, right columns), as well as between individual runs of a single model (Supporting Information). We compared the observed and simulated maps using, once again, the standard Pearson correlation coefficient, as well as the so-called Adjusted Rand Index (ARI), which is especially well suited to measure the similarity of clustered data³¹. Both measures were computed for the pairs of interaction maps (observation vs. individual simulation) filled with ones or zeros depending on whether the significant interaction between the processes of different time scales was identified or not (so the coloured areas of maps in Figs. 1 and 4 would be filled with ones, and white areas — with zeros); hereafter, we will call the maps so constructed the thresholded binary maps.

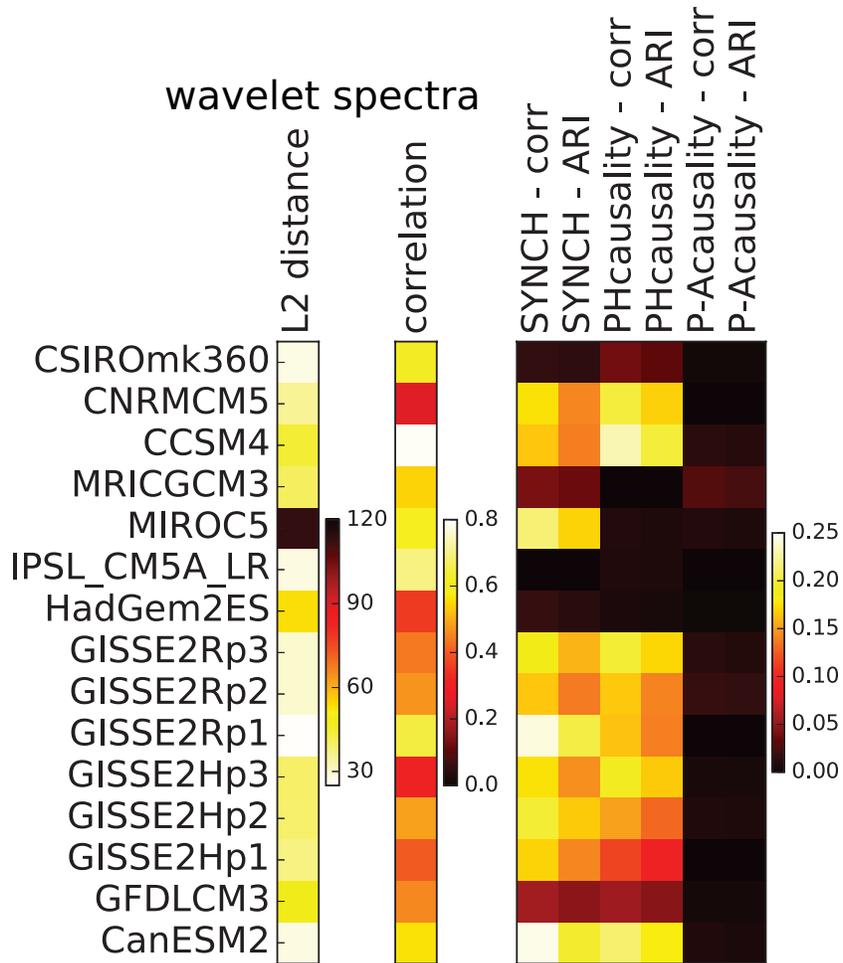


Figure 3: Measures of similarity between various characteristics of the observed and CMIP5 simulated NINO3.4 time series. Shown are ensemble averages of these characteristics over multiple runs of individual models (see the model acronyms on the left). The first two columns compare the observed and simulated wavelet spectra, using the root-mean-square (rms) distance and Pearson correlation, both computed in the frequency space. Next three pairs of columns display measures of similarity between the observed and simulated interaction maps analogous to the observed maps of Fig. 1, namely the phase synchronization, phase–phase causality and phase–amplitude causality maps. Each pair presents two distinct similarity measures: the Pearson correlation (corr) in the phase space of the corresponding map, as well as the Adjusted Rand Index (ARI; see text).

The above similarity measures averaged over the ensembles of individual model simulations (Fig. 3, right columns) are in fact fairly low. For phase synchronization, the highest similarity was detected in the CanESM2 model, at the 0.24 level. The phase synchronization map for the best run of the CanESM2 model (Fig. 4, left) indicates synchronization between the processes with the same time scales as in the observed data (Fig. 1, left), that is, between the LF, QB and AC/CT modes. This, however, is more of an exception than a rule, as the time scales of significant phase synchronization in most of the runs do not match the observed time scales.

The similarity levels between phase–phase causality maps from observations and model simulations are about the same as for phase synchronization maps, with maximum ensemble-mean correlation of 0.23 obtained for the CCSM4 model. The causality map for the best run of this model (Fig. 4, middle) is correlated with the observed map (Fig. 1, middle) at the 0.41 level, and captures correctly the observed LF–QA, QB–QA and QA–QB connections. Note that the best matches to the observed maps in terms of phase synchronization and phase–phase causality come from individual simulations of different models, meaning that neither model run was able to capture the entirety of the observed interactions. Finally, no model was able to capture the observed phase–amplitude causal connection between the LF and QB modes (Fig. 1, right). The highest ensemble-mean correlation between the observed and simulated maps is only 0.03 (MRI-CGCM3), with the highest correlation of 0.1 for one of the CSIRO-mk360 simulations (the corresponding causality map is shown in Fig. 4, right). The phase-synchronization and phase–phase causality maps for the latter run are, however, inferior to those from other models in terms of their similarity to the observed maps (not shown).

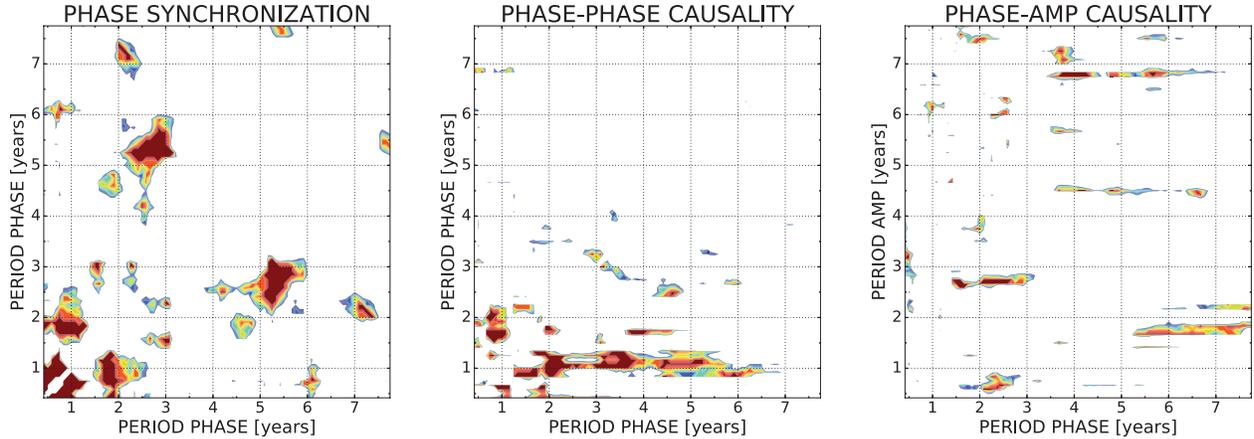


Figure 4: The same as in Fig. 1, but for the individual simulations of CMIP5 models that best match the observed structures: phase synchronization in CanESM model (left); phase–phase causality in CCSM4 model (middle); and phase–amplitude causality in CSIRO-mk360 model (right).

To summarize, the CMIP5 models exhibit great sampling variations in the simulated ENSO characteristics. Some of the simulations do exhibit certain aspects of interactions between the processes of different time scales which match the observed interactions. However, no single simulation is able to reproduce the entire sequence of causal connections inferred from the observed data.

Interactions in low-order dynamical model Next, we studied the interactions in the simulated time series representing the ENSO dynamics from low order dynamical model, namely the parametric recharge oscillator²⁵ (PRO).

Following Stein et al.²⁵, we used two traits of the parametric recharge oscillator: *neutral* model, which is purely deterministic ($\xi(t) = 0$ and $\lambda_0 = 0$) and *damped* model with the non-zero damping parameter λ_0 and Gaussian white noise $\xi(t)$ driving (see Methods for more details).

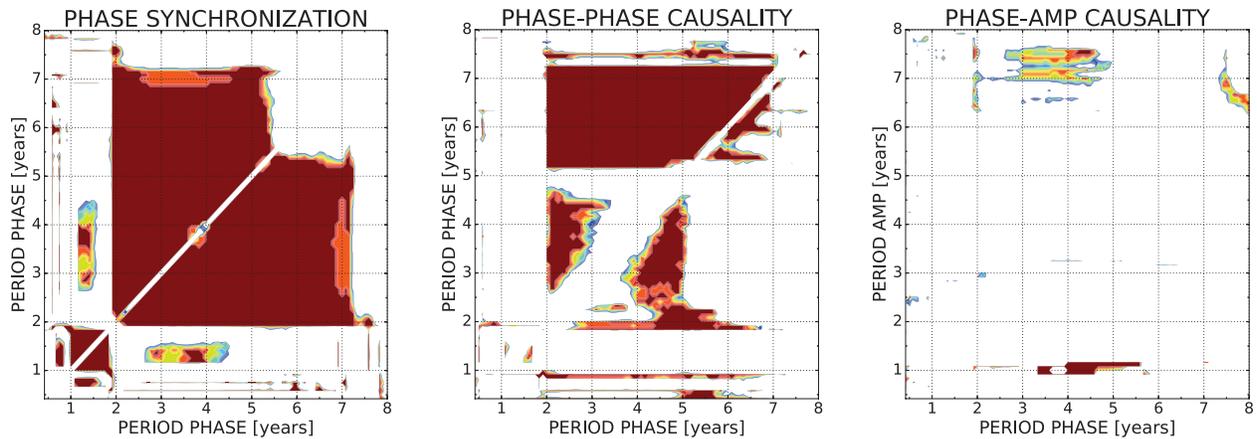


Figure 5: The same as in Fig. 1, but for modelled NINO3.4 time series using the neutral PRO model.

In the neutral case the system exhibits both the phase synchronization and phase-phase causality over a broad range of scales as shown in Fig. 5. The parametric recharge oscillator is not a multi-scale process; in fact, it contains one oscillation (ENSO) mode and a periodically changing parameter (the annual cycle). These two modes interact and create the combination tones. The main (ENSO) mode is quasi-periodic, appearing in the wavelet spectrum as a broadband peak. Since all the oscillatory modes are driven by the main PRO mode, all of them are coherent and provide a picture of mutually phase-synchronised modes over a wide range of periods, while the phase-amplitude causality is lacking completely, as depicted in Fig. 5.

For the damped PRO model, the presence of noise destroys the synchronization and causality phenomena observed in the deterministic neutral model. The corresponding results are presented in Fig. 6 in the form of aggregated thresholded binary maps. The colour shading in these plots shows the number of realizations (out of 20) in which the respective synchronization/causality relationship was significant. The interaction depicted in Fig. 6 exhibit huge variance from one model

realization to another, and no significant interactions seem to be systematically present. We thus conclude that the low-order damped PRO mode is unable to simulate synchronization and causality phenomena robustly, as significant interactions identified in each realization are driven by the noise sampling, rather than by any underlying low-order dynamics.

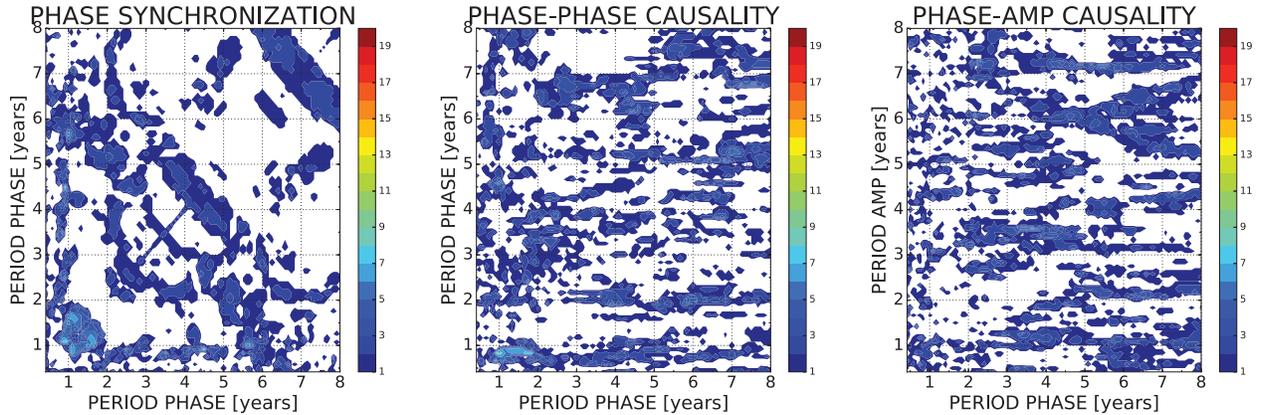


Figure 6: The same as in Fig. 1, but for modelled NINO3.4 time series using the damped PRO model. Shown in each panel are aggregates (sums) of thresholded binary maps over 20 realizations of this model.

Interaction in an empirical model. The second model of our interest was an empirical model based on the idea of linear inverse modelling. We built the inverse stochastic model according to methodology introduced by Kondrashov et al.³² (see Methods for more details). The results in the form of aggregated plots are shown in Fig. 7, which is analogous to Fig. 6 for the damped PRO model.

Unlike the low-order dynamical PRO model, the empirical stochastic model is clearly able to simulate synchronization phenomena, in which QB mode is synchronized with the AC and that the AC is phase synchronized with the combination tones. The same synchronization behaviour was observed in the observational data (see Fig. 1). The majority (though not all) of the empirical

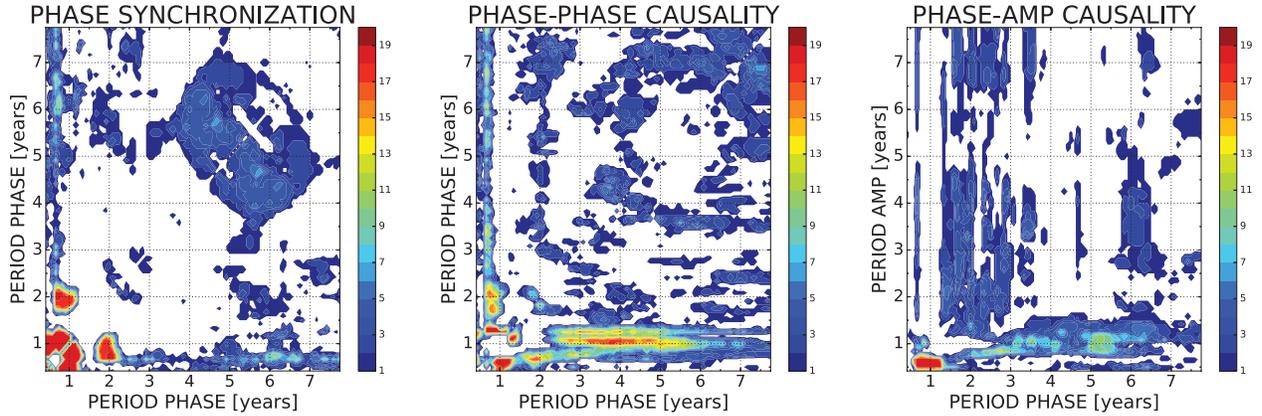


Figure 7: The same as in Fig. 6, but for modelled NINO3.4 time series from the empirical stochastic model.

model realizations were also able to correctly simulate the observed phase–phase causal relationships, with directional connection from LF mode to the AC modes, as well as from the AC modes to the modes in the QB range (compare the middle panels of Figs. 1 and 7). However, only a few realizations of the empirical stochastic model could capture the observed phase–amplitude causality from the LF mode to the QB mode; hence, we conclude that the present empirical model is not able to accurately simulate the observed phase–amplitude causal relationships in ENSO. Further improvements may perhaps be achieved by experimenting with the empirical models formulation (changing the number of variables, degree of nonlinearity etc.), but these experiments are beyond the scope of the present study.

3 Discussion and Outlook

In this communication, we examined causal interactions between different oscillatory modes comprising ENSO variability, using the tools of information theory^{22,23}. These tools enabled us to uncover an intricate network of interactions underlying the observed ENSO variability and further

elucidated, in particular, the key role of the quasi-biennial (QB) mode in ENSO dynamics. The (phase of) low-frequency (LF), interannual ENSO mode directly affects the amplitude of QB variability. It also indirectly affects the phase of QB variability, via the intermediate causal connection with the phase of the annual cycle (AC) and its combination tones (CT) associated with the LF mode³³. The AC/CT modes combined constitute the quasi-annual (QA) variability (changes in the shape of the annual cycle). The above interactions result in intermittent synchronization of QA, QB and LF modes leading to extreme ENSO events (Fig. 2).

The three important time scales — QA, QB and LF — detected by our independent analysis of ENSO interactions have been identified in previous studies as well¹⁵⁻¹⁷. Our results on synchronization of QB and QA modes are also consistent with previous work^{25,34}. The causal connection between the phase of the LF variability and that of the QA variability — that is, changes in the shape of the annual cycle depending on the state of the LF ENSO mode — is a new finding, conceptually similar to the effect of low-frequency component of North Atlantic Oscillation variability on the annual cycle of surface temperature over Europe³⁵. The novelty of our work is in applying an original methodology to obtain a new compact description of the causal and synchronization interactions instrumental in ENSO dynamics.

The observed ENSO interactions are poorly represented in the historical simulations of CMIP5 climate models, which exhibit large sampling variations in ENSO spectra and causality maps, both from model to model and among different runs of a single model. Some of the model simulations match time scales or select causality characteristics of the observed ENSO variability,

but no single simulation is able to reproduce the full picture of the observed interactions. The ENSO variability in long pre-industrial control runs of CMIP5-type models is known to exhibit multidecadal epochs characterized by different ENSO behaviour³⁰. Hence, there is still a possibility that the models possess correct ENSO dynamics, but the sample of 89 20th-century simulations considered here was simply not large enough to generate the ENSO epoch that would match the observed epoch. Analyses of long pre-industrial runs are in order to address this issue. However, the experiments with an empirical stochastic ENSO model of Kondrashov et al.³² suggest that the chances of reproducing the observed ENSO behaviour in ensemble simulations of the 20th century climate are much higher than the CMIP5 ensemble results. This implies that CMIP5 models do indeed misrepresent ENSO dynamics. The same conclusion holds for the conceptual parametric recharge oscillator mode of Stein et al.²⁵, which also fails to capture the observed cross-scale causal relationships in ENSO.

Thus, neither conceptual nor state-of-the-art dynamical climate models studied here were able to mimic the structure of the observed ENSO interactions, while the empirical models considered did quite a bit better. Understanding the discrepancies between the observed interactions and the interactions simulated by the dynamical models may be the key to an improved ENSO prediction. Other application of the framework would be in neuroscience, where cross-frequency phase–amplitude coupling has recently been observed in electrophysiological signals reflecting the brain dynamics. This cross-frequency coupling enriches the cooperative behaviour of neuronal networks and apparently plays an important functional role in neuronal computation, communication, and learning³⁶.

Methods

We analysed monthly NINO3.4 anomalies – with respect to long-term climatology – in the observed 1900–2010 sea-surface temperature data³⁷, as well as the simulated NINO3.4 time series extracted from 89 historic simulations of 15 different global climate models (see Table 1) participating in the CMIP5 intercomparison project²⁸.

Following Paluš^{22,23}, we studied interactions between the processes dominated by different time scales using the phase dynamics approach³⁸. In particular, we first applied, to the NINO3.4 time series, the complex continuous wavelet transform (CCWT) with Morlet mother wavelet²⁹ and obtained time-dependent complex wavelet coefficients $\psi_p(t)$ for a given central period p

$$\psi_p(t) = s_p(t) + i\hat{s}_p(t) = A_p(t)e^{i\phi_p(t)}. \quad (1)$$

Here $s_p(t)$ and $\hat{s}_p(t)$ are the real and imaginary parts of the wavelet coefficients, respectively. The instantaneous phase $\phi_p(t)$ and amplitude $A_p(t)$ of the variability associated with that central period p are given by

$$\phi_p(t) = \arctan \frac{\hat{s}_p(t)}{s_p(t)} \quad (2)$$

$$A_p(t) = \sqrt{s_p(t)^2 + \hat{s}_p(t)^2}. \quad (3)$$

For example, the biennial wavelet component for the observed NINO3.4 time series obtained by setting the central wavelet period to 2 years yields instantaneous phase $\phi_2(t)$ and amplitude $A_2(t)$ time series shown by blue and red curves in Fig. 8. Also shown here is the biennial reconstruction $A_2(t) \cos \phi_2(t)$ (brown). We computed analogous phase and amplitude time series for the variabil-

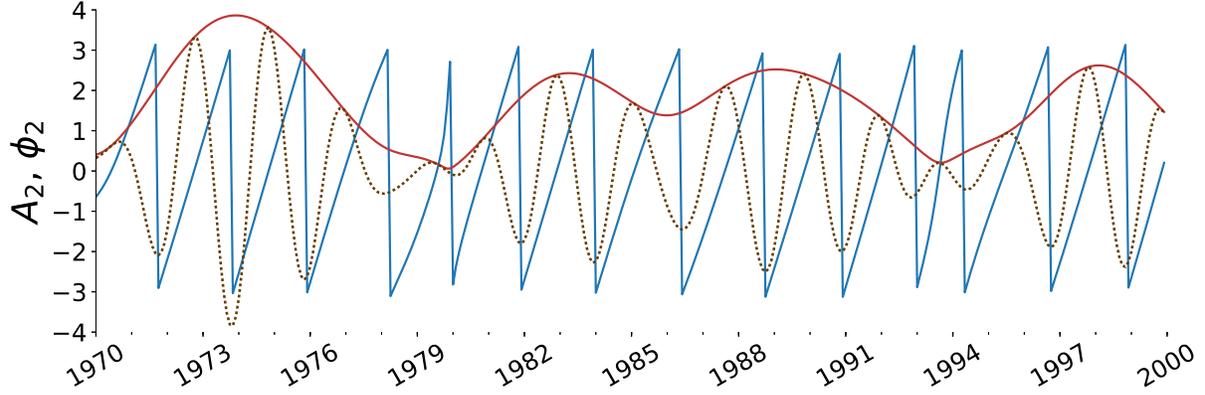


Figure 8: Biennial wavelet component of the observed NINO3.4 index (1970-1999): phase of the biennial mode $\phi_2(t)$ in blue, amplitude of the biennial mode $A_2(t)$ in red and reconstruction of the biennial mode $A_2(t) \cos \phi_2(t)$ in dotted brown.

ity associated with different central wavelet periods (over a wide range of these periods), and then studies synchronizations and causal interactions among all of these time series.

To do so, we computed the mutual information I^{39} , as well as the conditional mutual information, also known as transfer entropy⁴⁰, based on the phase and amplitude time series associated with variability at two different time scales p_1 and p_2 . In particular, mutual information $I(\phi_{p_1}(t); \phi_{p_2}(t))$ characterizes *phase synchronization* between the two time series⁴¹. Conditional mutual information $I(\phi_{p_1}(t); \phi_{p_2}(t + \tau) - \phi_{p_2}(t) | \phi_{p_2}(t))$, where τ is a forward time lag, describes the information in ϕ_{p_1} about the future of ϕ_{p_2} , which measures *phase-phase causality*. Similarly, $I(\phi_{p_1}(t); A_{p_2}(t + \tau) | A_{p_2}(t), A_{p_2}(t - \eta), A_{p_2}(t - 2\eta))$, with η being a backward time lag, is a measure of *phase-amplitude causality*. We estimated the (conditional) mutual information using the k -nearest neighbour algorithm⁴²; see Paluš^{22,23} and references therein for further details of the methodology.

We estimated statistical significance of the results using a randomization procedure²⁴ in which surrogate data time series were obtained by taking the Fourier transform of the original NINO3.4 time series, shuffling phases of its Fourier components and transforming back to the temporal domain⁴³. This procedure creates surrogate data sets with the same spectrum as that of the original data, but with no interactions between processes of different time scales.

Parametric recharge oscillator This conceptual ENSO model – due to Stein et al.²⁵ – has the following form

$$\frac{dT}{dt} = -\lambda(t)T + \omega(t)H + \xi(t) \quad (4)$$

$$\frac{dH}{dt} = -RT, \quad (5)$$

where T represents eastern equatorial Pacific SST anomalies, H represents the zonal mean equatorial Pacific thermocline depth anomalies, $\lambda(t)$ and $\omega(t)$ are the seasonally varying growth rate parameters of the oscillator, the constant R relates to the time scale of the geostrophic adjustment of the thermocline to wind stress anomalies, $\xi(t)$ is Gaussian white noise representing the forcing by the atmosphere, and finally the damping parameter, $\lambda(t) = \lambda_0 + \epsilon \cos(\omega_a t)$ with ω_a being the annual frequency. The model is integrated forward in time, using the same set of parameters as in Stein et al.²⁵, to obtain time series of NINO3.4 anomalies of the same length as the data used in this study.

Empirical model The inverse model is constructed in the space of leading empirical orthogonal functions⁴⁴ of the sea surface temperature field. The linear model takes the form

$$dx_i = \left(\mathbf{b}_i^{(0)} \mathbf{x} + c_i^{(0)} \right) dt + r_i^{(0)} dt \quad (6)$$

$$dr_i^{(l)} = \mathbf{b}_i^{(l+1)}[\mathbf{x}, \mathbf{r}^{(0)}, \dots, \mathbf{r}^{(l)}]dt + r_i^{(l+1)}dt, \text{ where } l = 1, 2, \dots, L \quad (7)$$

where \mathbf{x} is the state vector of dimension n , the matrices \mathbf{A}_i , the vectors $\mathbf{b}_i^{(0)}$, and the components $c_i^{(0)}$ of the vector $\mathbf{c}^{(0)}$, as well as the components $r_i^{(0)}$ of the residual forcing $\mathbf{r}^{(0)}$, are determined by the least squares. The set of L equations (7) is terminated at the value of L for which the L^{th} level residual r^{L+1} becomes white in time. The model explicitly accounts for the modelled process \mathbf{x} feeding back on the noise statistics. Since it is well known that extreme ENSO events tend to occur in boreal winter, the explicit seasonal dependence is included in the dynamical part of the first level of the nonlinear model, namely the matrix $\mathbf{B}^{(0)}$ and the vector $\mathbf{c}^{(0)}$ are assumed to be periodic with period $T = 12$ months:

$$\mathbf{B}^{(0)} = \mathbf{B}_0 + \mathbf{B}_s \sin(2\pi t/T) + \mathbf{B}_c \cos(2\pi t/T) \quad (8)$$

$$\mathbf{c}^{(0)} = \mathbf{c}_0 + \mathbf{c}_s \sin(2\pi t/T) + \mathbf{c}_c \cos(2\pi t/T). \quad (9)$$

The linear model (equations 6 and 7) with $L = 2$ was estimated using the ERSST v.4⁴⁵ data in the phase space of leading 20 EOFs of SST (30°S - 60°N), and was subsequently used to obtain multiple independent synthetic ENSO time series of the same length as the observational data record.

Robustness analysis In order to assess the robustness of our results, we estimated the (conditional) mutual information using equiquantal binning method (using 4 bins) and compared the results to the k -nearest neighbours estimate ($k = 64$); see Hlaváčková-Schindler et al.⁴² and references therein for a comprehensive review of (conditional) mutual information estimators and their caveats. Both estimates underwent the statistical significance testing using 500 Fourier transform

surrogates⁴³ and were computed using the NINO3.4 index data for a full available period 1870–2016³⁷. The conditional mutual information, which reflects phase–phase causality, was averaged over forward time lags 1–30 months, $\tau = 1, \dots, 30$ ³⁹.

As can be seen from the robustness analysis result in Fig. 9, both estimates provide the same picture with only small deviations, which we consider to be statistical fluctuations. Cross–scale interactions, which are of main interest here, are virtually the same – we can clearly identify the phase synchronization between the annual cycle (AC) and quasi-biennial (QB; periods of 1.8–2.1 yr) mode and between the AC and combination tones (CT; periods approximately 9 and 14 months). As for the phase–phase causality, one could distinctly detect the low-frequency (LF; periods 4–6 yr) phase driving the AC phase and that the phase of QB is partially slaved to the phase of AC.

For the second phase of robustness analysis, we divided the full NINO3.4 data period into two equally long periods, namely 1870–1943 and 1943–2016 periods, and estimated the phase synchronization and phase–phase causality for each of these periods separately. The causality estimates from the two different periods exhibit substantial differences (Fig. 10), which can either be due to the true change in ENSO dynamics between the two periods, or due to less robust estimation of the mutual information for a shorter time series.

The wavelet power spectra of the NINO3.4 time series computed for the full (1870–2016) period, as well as for the 1870–1943 and 1943–2016 sub-periods (Fig. 11) suggest that the former possibility is more likely. Indeed, the ENSO variability in the second part of the record (1943–

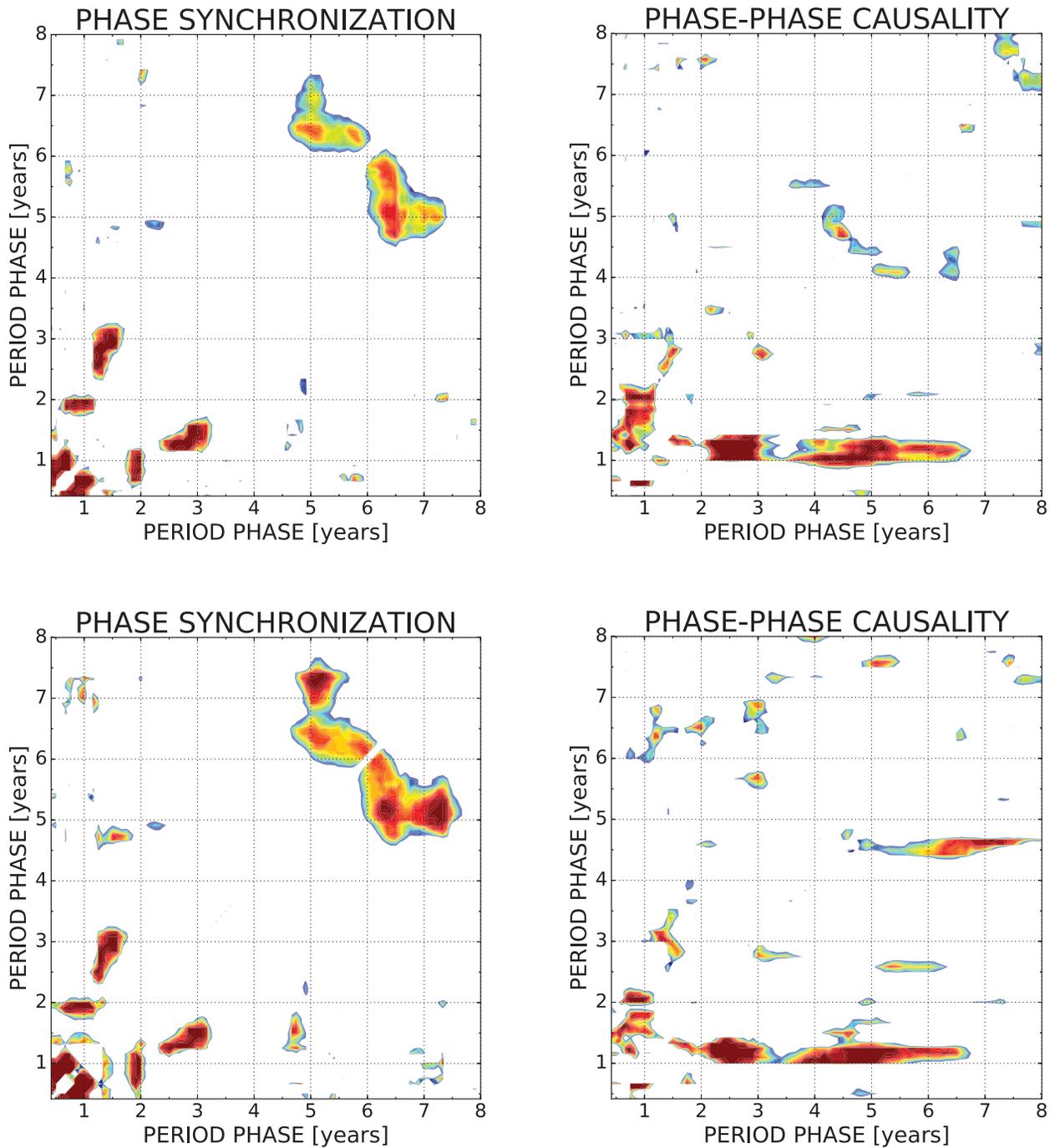


Figure 9: (left) Cross-scale phase synchronization and (right) phase–phase causality in the observed NINO3.4 time series for a full period 1870–2016. Shown are (positive) significance-level deviations from the 95th percentile of the (top row) equiquantal binning estimates and (bottom row) of the k -nearest neighbour estimates of (conditional) mutual information, tested using 500 Fourier transform surrogates (see main text for details).

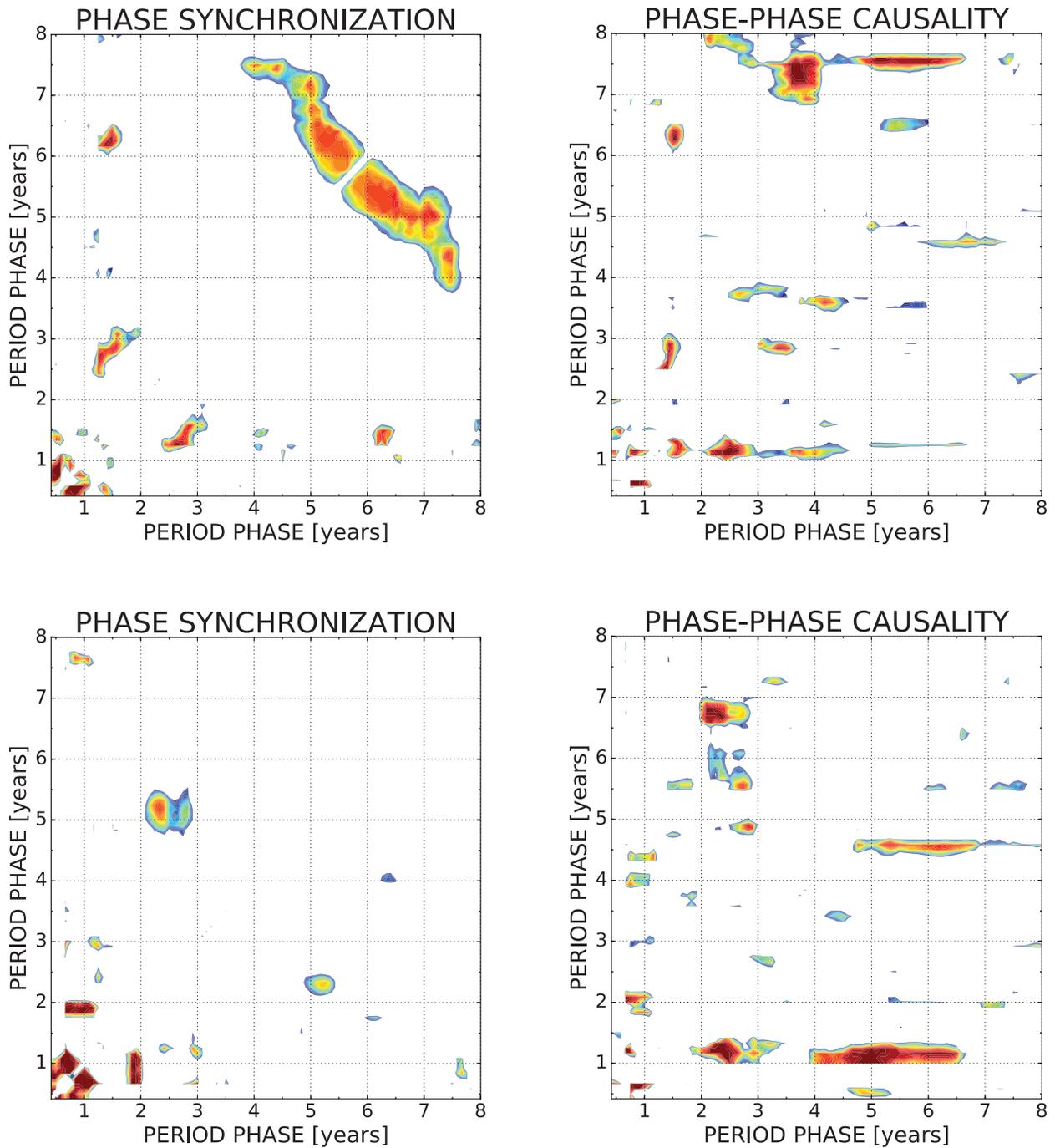


Figure 10: (left) Cross-scale phase synchronization and (right) phase–phase causality in the observed NINO3.4 time series for (top row) the first half of the full period 1870–1943 and (bottom row) the second half of the full period 1943–2016. Shown are (positive) significance-level deviations from the 95th percentile of the of the k -nearest neighbour estimates of (conditional) mutual information, tested using 500 Fourier transform surrogates.

2016) exhibited more power at QB periods compared with the first part of the record (1870–1943); thus the phase synchronization between AC and QB modes was more pronounced during 1943–2016 compared to 1870–1943, consistent with Fig. 10. Similarly, the LF ENSO variability (periods of 4–6 yr) was also more pronounced in the second part of the record, leading to a more distinct detection of LF phase AC phase causality during that period (Fig. 10). This apparent change in ENSO dynamics is in agreement with the studies which indicate that ENSO behaviour can exhibit multi-decadal epochs characterized by pronounced differences in diverse ENSO characteristics^{30,46–48}.

On the other hand, the observed phase–phase causality, where the AC partially drives the variability in the QB range as seen in the estimates of the conditional mutual information for the whole period (see Fig. 9) could not be found in neither of the half period estimates, or if present, it is considerably weaker (Fig. 10). This is most likely to be attributed to the insufficient length of the analysed time series. It is well known that the sensitivity of detection of the causal relationships heavily depends on the length of the time series and also on the dimensionality in which we are estimating the joint probability distribution^{39,49}.

1. Tsonis, A. A. & Roebber, P. J. The architecture of the climate network. *Physica A* **333**, 497–504 (2004).
2. Havlin, S. *et al.* Challenges in network science: Applications to infrastructures, climate, social systems and economics. *Eur. Phys. J.-Spec. Top.* **214**, 273–293 (2012).
3. Paluš, M., Hartman, D., Hlinka, J. & Vejmelka, M. Discerning connectivity from dynamics in climate networks. *Nonlinear Proc. Geoph.* **18**, 751–763 (2011).

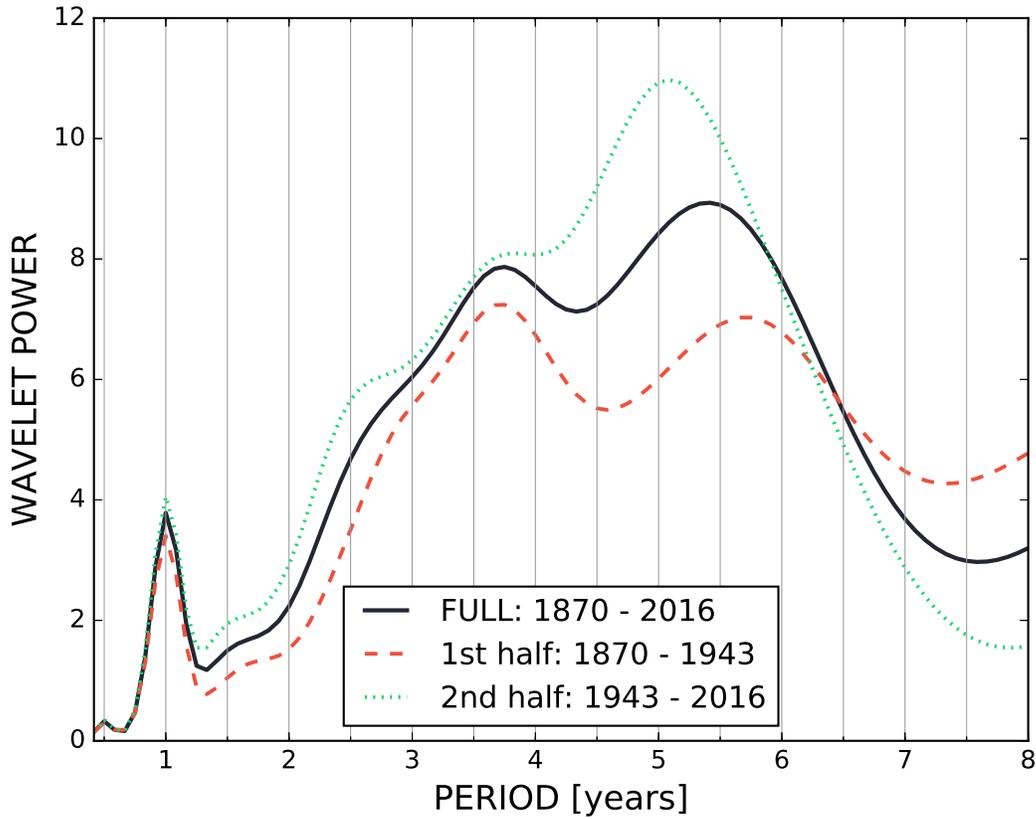


Figure 11: Power spectra of NINO3.4 SSTs, as a function of period. These three spectra are computed by time-averaging the spectral power from the wavelet analysis using the Morlet mother wavelet²⁹. Spectrum for the full period of 1870–2016 in solid black, and dashed red and dotted green for the first (1870–1943) and second (1943–2016) halves of the data, respectively.

4. Neelin, J. D. *et al.* ENSO theory. *J. Geophys. Res.-Oceans* **103**, 14261–14290 (1998).
5. Trenberth, K. E. *et al.* Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.-Oceans* **103**, 14291–14324 (1998).
6. Alexander, M. A. *et al.* The atmospheric bridge: The influence of ENSO teleconnections on air-sea interaction over the global oceans. *J. Climate* **15**, 2205–2231 (2002).
7. Burgers, G. & Stephenson, D. B. The normality of El Niño. *Geophys. Res. Lett* **26**, 1027–1039 (1999).
8. Sardeshmukh, P. D., Compo, G. P. & Penland, C. Changes of probability associated with El Niño. *J. Climate* **13**, 4268–4286 (2000).
9. Ghil, M. & Robertson, A. W. Solving problems with GCMs: General circulation models and their role in the climate modeling hierarchy. *International Geophysics* **70**, 285–325 (2001).
10. Larkin, N. K. & Harrison, D. ENSO warm (El Niño) and cold (La Niña) event life cycles: Ocean surface anomaly patterns, their symmetries, asymmetries, and implications. *J. Climate* **15**, 1118–1140 (2002).
11. Barnett, T. P. Variations in near-global sea level pressure: Another view. *J. Climate* **1**, 225–230 (1988).
12. Meehl, G. A. The annual cycle and interannual variability in the tropical Pacific and Indian Ocean regions. *Mon. Weather Rev.* **115**, 27–50 (1987).

13. Lau, K.-M. & Sheu, P. Annual cycle, quasi-biennial oscillation, and southern oscillation in global precipitation. *J. Geophys. Res.-Atmos.* **93**, 10975–10988 (1988).
14. Jiang, N., Neelin, J. D. & Ghil, M. Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Clim. Dynam.* **12**, 101–112 (1995).
15. Barnett, T. The interaction of multiple time scales in the tropical climate system. *J. Climate* **4**, 269–285 (1991).
16. Kim, K.-Y. Investigation of ENSO variability using cyclostationary EOFs of observational data. *Meteorol. Atmos. Phys.* **81**, 149–168 (2002).
17. Yeo, S.-R. & Kim, K.-Y. Global warming, low-frequency variability, and biennial oscillation: an attempt to understand the physical mechanisms driving major ENSO events. *Clim. Dynam.* **43**, 771–786 (2014).
18. Ebert-Uphoff, I. & Deng, Y. Causal discovery for climate research using graphical models. *J. Climate* **25**, 5648–5665 (2012).
19. Runge, J. *et al.* Identifying causal gateways and mediators in complex spatio-temporal systems. *Nature Comm.* **6** (2015).
20. van Nes, E. H. *et al.* Causal feedbacks in climate change. *Nature Clim. Change* **5**, 445–448 (2015).
21. Hannart, A., Pearl, J., Otto, F., Naveau, P. & Ghil, M. Causal counterfactual theory for the attribution of weather and climate-related events. *B. Am. Meteorol. Soc.* **97**, 99–110 (2016).

22. Paluš, M. Multiscale atmospheric dynamics: cross-frequency phase-amplitude coupling in the air temperature. *Phys. Rev. Lett.* **112**, 078702 (2014).
23. Paluš, M. Cross-scale interactions and information transfer. *Entropy* **16**, 5263–5289 (2014).
24. Paluš, M. From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics. *Contem. Phys.* **48**, 307–348 (2007).
25. Stein, K., Timmermann, A., Schneider, N., Jin, F.-F. & Stuecker, M. F. ENSO seasonal synchronization theory. *J. Climate* **27**, 5285–5310 (2014).
26. Stuecker, M. F., Timmermann, A., Jin, F.-F., McGregor, S. & Ren, H.-L. A combination mode of the annual cycle and the El Niño/Southern Oscillation. *Nature Geosci.* **6**, 540–544 (2013).
27. Torrence, C. & Webster, P. J. The annual cycle of persistence in the El Niño/Southern Oscillation. *Q. J. Roy. Meteor. Soc.* **124**, 1985–2004 (1998).
28. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *B. Am. Meteorol. Soc.* **93**, 485 (2012).
29. Torrence, C. & Compo, G. P. A practical guide to wavelet analysis. *B. Am. Meteorol. Soc.* **79**, 61–78 (1998).
30. Wittenberg, A. T. Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.* **36** (2009).
31. Hubert, L. & Arabie, P. Comparing partitions. *J. Classif.* **2**, 193–218 (1985).

32. Kondrashov, D., Kravtsov, S., Robertson, A. W. & Ghil, M. A hierarchy of data-based ENSO models. *J. Climate* **18**, 4425–4444 (2005).
33. Stuecker, M. F., Jin, F.-F., Timmermann, A. & McGregor, S. Combination Mode Dynamics of the Anomalous Northwest Pacific Anticyclone. *J. Climate* **28**, 1093–1111 (2015).
34. Rasmusson, E. M., Wang, X. & Ropelewski, C. F. The biennial component of ENSO variability. *J. Marine Syst.* **1**, 71–96 (1990).
35. Paluš, M., Novotná, D. & Tichavský, P. Shifts of seasons at the European mid-latitudes: Natural fluctuations correlated with the North Atlantic Oscillation. *Geophys. Res. Lett.* **32** (2005).
36. Canolty, R. T. & Knight, R. T. The functional role of cross-frequency coupling. *Trends Cogn. Sci.* **14**, 506–515 (2010).
37. Rayner, N. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.-Atmos.* **108** (2003).
38. Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: a universal concept in nonlinear sciences*, vol. 12 (Cambridge University Press, 2003).
39. Paluš, M. & Vejmelka, M. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Phys. Rev. E* **75**, 056211 (2007).
40. Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **85**, 461–464 (2000).

41. Paluš, M. Detecting phase synchronization in noisy systems. *Phys. Lett. A* **235**, 341–351 (1997).
42. Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M. & Bhattacharya, J. Causality detection based on information-theoretic approaches in time series analysis. *Phys. Rep.* **441**, 1–46 (2007).
43. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58**, 77–94 (1992).
44. Hannachi, A., Jolliffe, I. & Stephenson, D. Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.* **27**, 1119–1152 (2007).
45. Huang, B. *et al.* Extended reconstructed sea surface temperature version 4 (ERSST. v4). Part I: upgrades and intercomparisons. *J. Climate* **28**, 911–930 (2015).
46. Vecchi, G. A. & Wittenberg, A. T. El Niño and our future climate: where do we stand? *WIREs: Clim. Change* **1**, 260–270 (2010).
47. McGregor, S., Timmermann, A., England, M., Elison Timm, O. & Wittenberg, A. Inferred changes in El Niño–Southern Oscillation variance over the past six centuries. *Clim. Past* **9**, 2269–2284 (2013).
48. Li, J. *et al.* Interdecadal modulation of El Niño amplitude during the past millennium. *Nature Clim. Change* **1**, 114–118 (2011).

49. Vejmelka, M. & Paluš, M. Inferring the directionality of coupling with conditional mutual information. *Phys. Rev. E* **77**, 026214 (2008).

Acknowledgements We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. The U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support for CMIP project and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. This study was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the Program KONTAKT II, project LH14001 (N.J. and M.P.), and by NSF grants OCE-1243158 (S.K.) and AGS-1408897 (S.K. and A.A.T.). All data used in this study are publicly available at the respective sites listed in the references. All scripts used in this study are written in Python and are publicly available at github.com profile of N.J. at https://github.com/jajcayn/enso_cmi.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to M. Paluš (email: mp@cs.cas.cz).

#	model acronym	no. of realizations
1.	CCSM4	6
2.	CNRM-CM5	10
3.	CSIRO-Mk3-6-0	10
4.	CanESM2	5
5.	GFDL-CM3	5
6.	GISS-E2-Hp1	6
7.	GISS-E2-Hp2	5
8.	GISS-E2-Hp3	6
9.	GISS-E2-Rp1	6
10.	GISS-E2-Rp2	6
11.	GISS-E2-Rp3	6
12.	HadGEM2-ES	5
13.	IPSL-CM5A-LR	6
14.	MIROC5	4
15.	MRI-CGCM3	3

Table 1: CMIP5 twentieth century simulations considered in this study.