Spatiotemporal complexity and time-dependent networks in sea surface temperature from mid- to late Holocene

4 authors, including:

Fabrizio Falasca
New York University
25 PUBLICATIONS   **171** CITATIONS

SEE PROFILE

Julien Crétat
University of Burgundy
65 PUBLICATIONS   **1,120** CITATIONS

SEE PROFILE

Annalisa Bracco
Georgia Institute of Technology
163 PUBLICATIONS   **5,107** CITATIONS

SEE PROFILE

Spatiotemporal complexity and time-dependent networks in mid- to late Holocene simulations

Fabrizio Falasca*[1], Julien Crétat[2,3], Pascale Braconnot[2], and Annalisa Bracco[1]

[1]Georgia Institute of Technology, Atlanta, USA
[2]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ,
Université Paris-Saclay, 91191 Gif-sur-Yvette, France
[3]Biogéosciences/CRC, CNRS-UB, Université de Bourgogne, Dijon, France

Abstract

In climate science regime transitions include abrupt changes in modes of climate variability and shifts in the connectivity of the whole system. While important, their identification remains challenging. This paper proposes a new framework to investigate regime transitions and connectivity patterns in spatiotemporal climate fields. Firstly, local regime shifts are quantified by means of information entropy; secondly, their spatial heterogeneity is examined by identifying the underlying spatial domains of the entropy field; finally, a weighted, direct and time-dependent network is inferred to capture the linkages between these domains. The spatiotemporal variability in sea surface temperature (SST) in two simulations of the last 6000 years is investigated with the proposed approach. The largest regional regime shifts emerge as abrupt transitions from low to high-frequency SST oscillations, or vice versa, in both simulations. Furthermore, the variability in time of the two climate networks is studied in terms of their network density. Generally, rapid and sudden transitions in the degree of connectivity of the system are observed in both simulations but, in most cases, at different times, with few exceptions. This suggests that our ability to predict the climate system may be hampered by its inherent complexity resulting from internal variability.

---

*Correspondence: fabrifalasca@gatech.edu

1 Introduction

Our ability to predict Earth's climate is limited by the nonlinear and intrinsically complex interactions between its components. Quantifiable changes in the system dynamics, i.e. dynamical transitions or *climate regime shifts*, are manifestations of such complexity. A global-scale example of major climate shifts are the Snowball Earth glaciations that may have occurred 650-700 million years ago [Hoffman and Schrag, 2002, Pierrehumbert et al., 2011, Lucarini and Bódai, 2017, 2019]. In addition, the climate system has experienced a large variety of regional regime shifts, such as the Dansgaard-Oeschger events [Dansgaard et al., 1982, Johnsen et al., 1992] or the abrupt hydrological fluctuations in the African and Asian monsoon regions [Gasse, 2000, Morrill et al., 2003], the sudden cooling of the North Pacific sea surface temperature in the mid-Holocene [Mix et al., 1999] or the 4.2-kiloyear BP aridification event during the Holocene epoch [deMenocal, 2001] to cite a few [1].

In recent years, methods stemming from non-equilibrium statistical mechanics, information theory, and network analysis have proven valuable for understanding and investigate complex climate transitions [Balasis et al., 2013, Donner et al., 2011a]. Examples include the work by Donges et al. [Donges et al., 2011], who investigated climate fluctuations in Northern and Eastern Africa and linked them to rapid changes in hominid evolution, or by Donges et al. [Donges et al., 2015], who showed that nonlinear regime shifts of the Asian monsoon were associated to the migration and collapse of ancient societies. In the study of dynamical systems, the permutation entropy metric [Bandt and Pomp, 2002] is often used in nonlinear time series analysis to quantify the complexity of a dynamical system, with changes in complexity being linked to regime shifts. This metric has been applied recently to climate fields to studying the complexity of a proxy record of the El Niño Southern Oscillation (ENSO) Saco et al. [2010] and to identifying anomalies and post-processing issues of deep polar ice-cores [Garland et al., 2018]. An alternative, more traditional measure of complexity in dynamical systems is provided by Lyapunov exponents [Eckmann and Ruelle, 1985] that quantify the rate of separation of infinitesimally close trajectories. However, recently developed entropy quantifiers are easier to estimate and preferable when dealing with arbitrary real-world observations and large datasets [Bandt and Pomp, 2002]. Independently of the methodology used, the majority of the studies employing nonlinear data analysis techniques for studying climate dynamics have focused so far on single time series (i.e., on paleoclimate proxies or their modeled equivalent). Here, we propose a novel approach that allows for extending nonlinear time series analysis to the investigation of dynamical transitions in spatiotemporal fields. The methodology couples an entropy quantifier recently proposed by Corso et al. [Corso et al., 2018] with δ-MAPS a dimensionality reduction scheme developed by Fountalis et al. [Fountalis et al., 2018]. Given a spatiotemporal field, our framework allows to quantify the local (i.e., for each grid cell) time evolution in information entropy and identify spatially contiguous sets of grid cells (domains) with highly correlated entropy variability. Changes in the information entropy of a time series reflect changes in its recurrent structure and hence in its system's dynamics. It follows that the identified domains consist of grid cells sharing the same dynamical transitions. Studying the average entropy signal inside domains enables to assess the timing, strength and location of climate regime shifts. In this paper we generally focus on abrupt regime shifts, broadly defined as fast and sudden transitions in the system's dynamics. A strong abrupt regime shift in the sea surface temperature (SST) field, and reported in this work, would consist in a sudden transition in regional dynamics from high- to

---

[1]For a review of abrupt regional shifts and tipping points in the climate system, the reader is referred to the work by Alley et al. [Alley et al., 2003] and Lenton et al. [Lenton et al., 2008].

low- frequency variability, or vice versa, that occurs over a period much shorter than the length of the integrations, for example over 200 to 300 years.

Additionally, the dynamics of the climate system establish linkages between domains, and complex network science provides a well-established methodology to quantify these connections [Donner et al., 2009, Donges et al., 2009]. Here we adopt a functional network inference that allows to quantify the magnitude of connections between domains, their directionality, and their evolution in time [Fountalis et al., 2018]. In our framework, domains are considered as static nodes of the climate network and links quantify their connections [Barabási, 2016, Newman, 2010]. The strength of our approach relies in exploring climate variability simultaneously in space and time.

To showcase this novel methodology, we analyze two global ocean-atmosphere coupled simulations of the last 6000 years performed with different versions of the IPSL Earth system model [Dufresne and et al., 2013, Braconnot et al., 2019a]. These simulations test the response of the climate system to the slow variations in Earth's orbit and to changes in trace gases. They produce a wide range of variability resulting from the combination of the response to external forcing (Earth's orbit and gases) and intrinsic climate variability [Braconnot et al., 2019a,b]. These two long, global climate simulations provide an ideal testbed for assessing the efficiency of the new method and its ability in identifying major shifts in the simulations.

The paper is organized as follows: Section 2 introduces the proposed approach; the climate simulations analyzed are described in Section 3; results are presented in Section 4 and a comparative application to a reanalysis data set follows in Section 5. A discussion of the method and its preliminary applications concludes the work.

2   Methodology and strategy

The proposed strategy builds upon concepts of recurrence plots, information entropy and δ-MAPS. Here we briefly introduce these tools and present the rationale behind and strategy followed by our comprehensive approach. More details on each component can be found in Marwan et al. [Marwan et al., 2007], Corso et al. [Corso et al., 2018] and Fountalis et al. [Fountalis et al., 2018].

2.1   Recurrence plots and information entropy

A recurrence plot (RP) is a nonlinear time series analysis method introduced by Eckmann et al. [Eckmann et al., 1987] that allows for visualizing a d-dimensional dynamical system in terms of its recurrences, a fundamental property of these systems [Poincaré, 1890, Cvitanović et al., 2016]. Given a trajectory $\mathbf{x}_i$ of a dynamical system in a d-dimensional state space at time i, a RP is an $N \times N$ matrix of 1 and 0 given by:

$$\mathrm{RP}_{i,j}(\epsilon) = \Theta(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \ \mathbf{x}_i \in \mathbb{R}^d, \ i,j \in [1, N] \ . \tag{1}$$

Here ϵ is the threshold distance and defines the neighborhood of a state $\mathbf{x}_i$, Θ is the Heaviside function, $\|\cdot\|$ is a norm and N is the number of states considered.

The analysis of structures (such as diagonal, vertical or horizontal lines) in a RP is known as

recurrence quantification analysis (RQA) and has found numerous applications [Marwan et al., 2007]. RQA allows for the definition of different measures of complexity of time series. Here we focus on one such measure, developed by Corso et al. [Corso et al., 2018] and based on the probability of occurrence of microstates in a RP. A microstate of size M is defined as an $M \times M$ matrix inside the RP. The total number of configurations of 1 and 0 in a microstate of size M is $M^* = 2^{M^2}$. It is then possible to define a probability of occurrence P_k of a microstate k as $P_k = \frac{n_k}{M^*}$, with n_k being the number of occurrences of the k-th microstate in the RP. The information entropy of the time series is then given by:

$$S(M^*) = -\sum_{k=1}^{M^*} P_k \log P_k \ . \tag{2}$$

The maximum entropy corresponds to the case in which all microstates appear with the same probability $P_k = \frac{1}{M^*}$. It can be analytically calculated as $S_{max}(M^*) = M \log 2$ and used to normalize the entropy, so that $S = 0(1)$ implies total predictability (unpredictability) of the system dynamics. The number of admissible microstates M^* grows exponentially as a function of M, but Corso et al. [Corso et al., 2018], showed that in practice only few of them actually populate the RP. It is therefore reasonable to sample $\bar{M}$ of the possible microstates in the RP and to expect a rapid convergence to $S(M^*)$. In this case, the maximum entropy will be equal to $S_{max}(\bar{M}) = \log \bar{M}$. The computed information entropy is a measure of the complexity of the time series examined and from now on the two terms will be used interchangeably.

There are other methods to estimate complexity of time series, for example related to RQA [Marwan et al., 2007], recurrence network analysis [Donner et al., 2011b] and information theory [Balasis et al., 2013, Bandt and Pomp, 2002]. Some of these quantifiers have been applied to climate science to analyze regime shifts and tipping points [Donges et al., 2011]. These methods usually require embedding the time series in a higher dimensional space, a procedure known as "state space reconstruction" [Takens, 2011]. Among these methods the entropy quantifier proposed by Corso et al. [Corso et al., 2018] has clear advantages when dealing with large spatiotemporal fields ($\sim 10^4$ time series in this work given the focus on two-dimensional spatial fields that evolve in time): (a) it avoids the state space reconstruction that may lead to false results [Marwan, 2011, M. Riedl, 2013] and (b) it relies on just one parameter, the vicinity threshold ϵ. Additionally, Corso et al. [Corso et al., 2018] showed that given a generic time series, results are (a) well correlated with its maximum Lyapunov exponent and (b) have a weak dependence on the distance threshold ϵ.

Here, for a given climate field $\mathbf{X}(t)$, we compute the information entropy and its variability for all the time series embedded in a spatial grid to derive its spatiotemporal entropy field $\mathbf{S_X}(t)$. In order to reduce its dimensionality, we then identify sets of grid cells (domains) that behave homogeneously in terms of their complexity using δ-MAPS [Fountalis et al., 2018], briefly described below.

2.2 δ-MAPS

2.2.1 Dimensionality reduction

Given a generic spatiotemporal field, δ-MAPS aims at identifying spatiotemporal patterns, defined as spatially contiguous sets of grid cells with highly correlated activity and referred to as *domains*.

The underlying hypothesis of the domain identification algorithm is that domains have epicenters or cores where their local homogeneity is maximum. The first step is then to identify these cores and expand them. Formally, each grid cell i of a spatiotemporal field $\mathbf{S}(t)$, is associated with a time series $s_i(t)$. Linear trends are removed from the time series. The similarity between two grid cells i and j is quantified by their Pearson correlation $r_{i,j}$ at lag $\tau = 0$.

Cores are identified by defining a $K-$neighborhood $\Gamma_K(i)$ including a grid cell i and its K closest (in terms of geodesic distance) cells. We then compute the local homogeneity of a cell i, defined as the average pairwise correlation in $\Gamma_K(i)$. A grid cell i is marked as a core if its local homogeneity is a local maximum and greater than a threshold δ. Finally, cores are then iteratively expanded and merged to identify domains [Fountalis et al., 2018]. The total number of domains D will depend on the threshold δ and its minimum sized on the K parameter. Details on the methodology are discussed in Appendix A and more in depth in [Fountalis et al., 2018] and [Falasca et al., 2019].

2.2.2 Network Inference

The temporal variability of domains is quantified by computing their signals. Given a spatiotemporal field $\mathbf{X}(t)$, the signal $X_A(t)$ of a domain A is the weighted cumulative anomaly of all time-series within that domain:

$$X_A(t) = \sum_{i=1}^{|A|} x_i(t) \cos \phi_i \ , \tag{3}$$

where $x_i(t)$ is a time series of length T associated to grid cell i with latitude ϕ_i and $|A|$ is the cardinality of the set.

Given D domains, the network is inferred by considering each possible pair of domains A and B and computing their Pearson correlation $r_{A,B}(\tau)$ for a lag range $\tau \in [-\tau_{max}, \tau_{max}]$. The statistical significance is then tested for the $\frac{1}{2}D(D-1)(2\tau_{max}+1)$ correlations adopting the False Discovery Rate (FDR) proposed by Benjamini and Hochberg [Benjamini and Hochberg, 1995] against multiple testing, and excluding autocorrelations by using the Bartlett's formula [Box et al., 2011].

z Two domains A and B are connected if there exists at least one significant correlation between the two at any lag in the range $\tau \in [-\tau_{max}, \tau_{max}]$, denoted as $R_{A,B}(\tau)$. If $R_{A,B}(\tau)$ includes the lag $\tau = 0$, the link is left undirected, and treated here as a two-way link $A \leftrightarrow B$. If $R_{A,B}(\tau)$ is strictly positive (negative) the link will be directed from A to B (B to A). Finally a weight $w_{A,B}$ is assigned to each link and based on the covariance between the two signals $X_A(t)$ and $X_B(t)$ at the lag τ^* at which their significant correlation $r_{A,B}(\tau)$ is maximized.

In this work, we quantify the degree of connectivity of the climate network, along with its time variability, with the network density metric. Given a weighted, direct network N between D domains, with weights $w_{A,B}$, we define its density ρ_N as:

$$\rho_N = \frac{\sum_{A,B,A\neq B}^{D} |w_{A,B}|}{\sum_{A,B,A\neq B}^{D} \sigma_A \sigma_B} \ , \tag{4}$$

where σ_A denotes the standard deviation of the signal $X_A(t)$ of domain A.

The denominator is never zero and accounts for the case in which all possible connections exist with

correlation $|r_{A,B}(\tau^*)| = 1$. Therefore, if $\rho_N = 1$ the network is a complete direct weighted graph, with each node connected to every other node and with their weights maximized. If $\rho_N = 0$ the network has no links.

3   Strategy to study climate variability in long transient simulations

3.1   Model simulations

We analyze two 6000-years long transient simulations of the global climate system realized with the IPSL model, Vlr01 and Sr02 hereafter. The Vlr01 simulation was run with the standard IPSL-CM5A-LR version [Dufresne and et al., 2013] utilized also in the Coupled Model Intercomparison Project Phase 5 (CMIP5). The atmospheric model resolution is $3.75°$ in longitude, $1.875°$ in latitude and 39 vertical levels. The ocean model resolution is $2°$, with spatial refinement at the equator and in the Arctic, and 31 vertical levels. The Sr02 simulation has a higher horizontal resolution in the atmosphere ($2.5°$ in longitude, $1.125°$ in latitude) and includes the new 11-layer ORCHIDEE land surface model and a dynamical vegetation module [Braconnot et al., 2019a]. The latter allows both the Leaf Area Index and the vegetation type to vary in each grid cell as a function of climate and CO_2.
1000-years long simulations were performed to produce initial conditions in equilibrium with the external forcing [Braconnot et al., 2019b]. Both simulations are forced by yearly-updated Earth's orbital parameters [Berger, 1978], and atmospheric composition derived from ice-core reconstructions [Braconnot et al., 2019a, et al., 2017]. When showing time-series we always report the simulation time on the x-axis (i.e., time going from 1 to 6000 with 6000 = year 1950).
The focus of the analysis is on sea surface temperature (SST) fields, saved as monthly averages.

3.2   Strategy

Given a spatiotemporal climate field, the objectives of our approach are (a) to identify spatially contiguous regions undergoing regime shifts, and (b) to investigate the time dependent connectivity patterns between those regions.

This can be achieved as follows:

- We consider a climate field $\mathbf{X}(t)$ and assign to each grid cell i a (constant) value of the vicinity threshold ϵ_i. The selection of ϵ_i for a time series $x_i(t)$ is based on the rationale that its entropy $s_i(\epsilon_i)$ should be invariant under small perturbations around its value such that $s_i(\epsilon_i) \sim s_i(\epsilon_i \pm \delta_i)$, $\delta_i << \epsilon_i$. Details on the heuristics can be found in Appendix B. For every grid cell i we have a (linearly) detrended, anomaly time series $x_i(t)$ with T data points. We consider W years of data, and compute its recurrence entropy s_i (see section 2.1). We do so every ΔW years and for each i. This step results in the computation of a spatiotemporal complexity field $\mathbf{S_X}(t)$.

- Given $\mathbf{S_X}(t)$, its dimensionality is reduced by identifying its domains (see section 2.2.1) and for each domain A the average entropy (complexity) signal is computed as $S_A(t) = \frac{\sum_{i=1}^{|A|} s_i(t)}{|A|}$,

where $s_i(t)$ is the entropy time series at grid cell i and $|A|$ the number of grid cells inside the domain. The information on climate regime shifts of each domain A is then contained in $S_A(t)$. Changes in information entropy are directly linked to structural changes in the RP and therefore to transitions in the system's dynamics. The domains identified can then be regarded as spatially contiguous sets of grid cells undergoing regime transitions or *shifts* in a homogeneous way.

- Complexity is a metric that quantifies the variability of a given time series and the D domains can be thought as the true, semi-autonomous components of the system in the 6000 years analyzed (all the grid points within a domain are characterized by highly correlated variability). This allows to consider them as "static" nodes of the climate network and investigate their connections in time, given that the D domains identified can be connected through oceanic and/or atmospheric teleconnections. Their connections are captured by the network topology in the $\mathbf{X}(t)$ field. We consider now Y years of data and infer a weighted and direct network between domains. We do so every ΔY years. This step results in the computation of a time dependent network between all D domains.

This strategy is applied to the two SST fields presented in section 3.1. The time resolution is monthly, and the length of the simulations is 6000 years.

It should be noted that the domain identification and network inference act on different fields: first, the dimensionality of the entropy field $\mathbf{S_X}(t)$ is reduced by identifying its underlying spatial components, then the connectivity patterns between these domains identified in the entropy field are inferred in the original climate field $\mathbf{X}(t)$.

3.3   Parameters

We compute the complexity of (linearly) detrended SST anomalies using time windows of $W = 100$ years every $\Delta W = 20$ years.This choice of W allows for the investigation of periods with decadal and multidecadal variability and shifts between them. We choose $\Delta W = 20$ years as a compromise between having a reasonable time resolution in the entropy field and restricting the large amount of computations. The recurrence entropy of each time series is computed by sampling $\bar{M} = 10000$ microstates of size $M = 4$ (as in Corso et al. [Corso et al., 2018]). The dimensionality reduction in the spatiotemporal entropy field $\mathbf{S_X}(t)$, allows for the identification of regions undergoing the same dynamical transitions. For this step we use a value of $\alpha = 0.01$ to select the threshold δ (as done in Fountalis et al. [Fountalis et al., 2018] and Falasca et al. [Falasca et al., 2019]) and $K = 2$ and $K = 6$ for the low- and high-resolution simulations, respectively. Results are qualitative similar for other values of K close to the selected ones. Networks between the D domains are inferred using $Y = 100$ years of (linearly) detrended anomalies every $\Delta Y = 3$ years. This results in a total of 1967 networks computed in each simulation. We choose $Y = 100$ years for consistency with the entropy computation. A choice of $\Delta Y = 3$ years is feasible given the reduced dimensionality of the dataset (i.e., D domains instead of all grid cells) and has the advantage of quantifying temporal shifts in connectivity patterns with high precision. For the network inference we set $\tau_{max} = 12$ months and and the FDR parameter $q = 0.04$. Therefore, for every network at most 4% of the inferred links may be false positives.
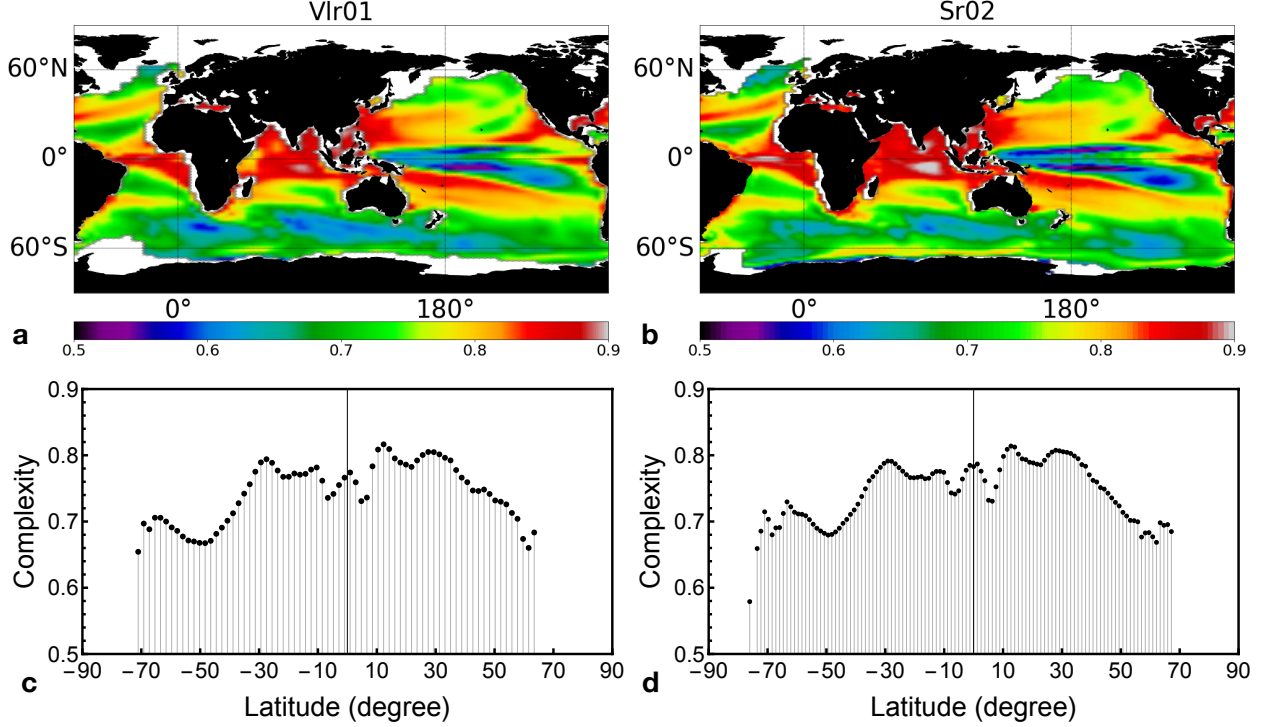
7

Figure 1: (a,b) Mean complexity field for Vlr01 and Sr02 simulations. (c,d) Average latitudinal dependence of complexity in both simulations.

4 Results

4.1 Mean complexity pattern

We first focus on the mean state of the complexity field in the two simulations. Figure 1a,b shows the mean spatiotemporal entropy field $\langle \mathbf{S}_{SST}(t) \rangle$.

The spatial distribution and magnitude of the SST mean complexity is nearly identical across simulations (Figure 1a,b). There is a marked latitudinal gradient between the tropics and mid- to high latitudes (Figure 1c,d) resulting mostly from the large complexity in the equatorial Atlantic and Indian Ocean basins. In the inter-tropical belt, the SST mean complexity depicts significant basin-dependency in both magnitude and spatial gradients. The tropical Indian Ocean is characterized by very large SST mean complexity, with values ranging from ~ 0.9 ($1 =$ maximum complexity) in its eastern equatorial part to ~ 0.8 in its western tropical part. Similar values are found in the equatorial Atlantic, also with a small east-to-west gradient. The tropical Pacific mean complexity, on the other hand, exhibits non-trivial spatial features. There is a steep west-to-east gradient in complexity, with the highest values in the upwelling cold tongue and significantly lower mean complexity in its western part. Additionally, the lowest global mean complexity is located in the central Pacific on two hemispheric-symmetric stripes.

The equatorial Pacific hosts the strongest mode of interannual climate variability, ENSO, and the potential predictability of tropical SST anomalies is, on average, greater in the Pacific than in the
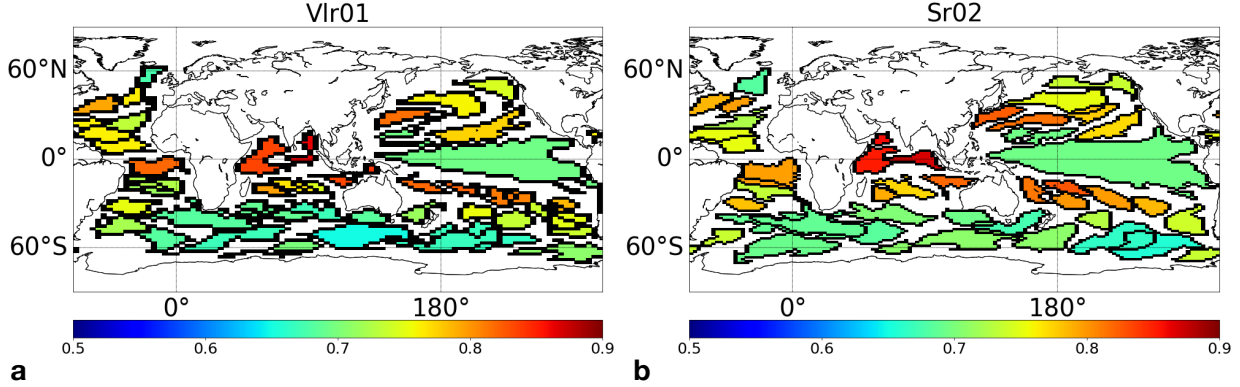
8

Figure 2: (a, b) Domains in the SST entropy field for the lower (Vlr01) and higher (Sr02) resolution simulations, respectively. Colors quantify the averaged complexity values inside each domain.

two remaining basins due to the ENSO quasiperiodicity. The analysis suggests that in the 6000 years considered, ENSO-related predictability in the IPSL model is on average larger in the central Pacific, east of 180°E, in two symmetric bands north and south of the Equator.

In the extra-tropics, the SST variability displays relatively low mean complexity outside the western boundary current pathways and the eastern boundary upwelling systems. The strongest shifts in complexity in both simulations, however, are found at high latitudes, as we will show next.

4.2   Exploring climate regime shifts in SST

To investigate the variability in the information entropy field $\mathbf{S}_{SST}(t)$, we reduce its dimensionality using $\delta-$MAPS. The identified domains are shown in Figure 2 for both simulations and are similar in terms of their spatial characteristics and mean complexity. Following from Figure 1, domains with large mean complexity are found in the Indian Ocean, Equatorial Atlantic basins, the Southern branch of the Horse Shoe pattern extending from the South Pacific Convergence Zone through Indonesia, and along the path of major boundary currents.
We report here on two domains that undergo major shifts in local dynamics in both simulations, to exemplify how the proposed strategy allows to quantify regime transitions in terms of timing and spatial characteristic. The two domains have been chosen because they exhibit considerable jumps in entropy ($\Delta S \sim 0.1$) in periods as short as 200 to 300 years. For each domain we show the time series of SST anomalies in centuries of highest and lowest complexity. A more comprehensive investigation of all regime shifts, their interrelationships and causes is beyond the scope of this work, that focuses on the methodology, and is left for future studies.

First, we focus on the domain in the Vlr01 simulation with the largest regime shift (Figure 3). It connects the Indian and South Atlantic Oceans (Figure 3a) and embeds the Agulhas retroflection current. Its mean complexity is 0.68 (Figure 3b), a value similar to that in the higher resolution

9

simulation in which, however, shifts occur at different times (Figure 3 and Appendix C). Therefore, for this domain, we can conclude that regime shifts are likely due to the internal variability of the system rather than variations in the orbital parameters or gas concentrations. Every point in the complexity signals is computed based on the detrended SST anomalies over 100 years. Panels c and d in Figure 3 show the monthly SST anomalies relative to the century with the lowest and largest complexity, respectively. In roughly 400 years, from year 5120 to 5520, the domain undergoes a sudden, rapid transition from high to low and back to high-frequency SST variability. Specifically, from year 5281 to 5380, we observe the least complex SST variability with a clear multidecadal oscillation (Figure 3c). This multidecadal oscillation has a periodicity of 33.3 years, which is significant at the 95% level according to a red noise test [Imkeller and Storch, 2001, Dijkstra et al., 2019]. The system then jumps to its highest complexity state in the 6000-year record between year 5421 and 5520 (Figure 3d).

Figure 4 presents the complexity signal for a domain crossing the Drake Passage in the Sr02 simulation. The mean complexity of this region is 0.73, higher than the majority of domains identified south of 30°S (see Figure 2b). Again, the SST anomalies relative to the centuries with lowest and largest complexity are also shown (Figure 4c,d). Low-frequency oscillations in SST anomalies are observed from year 3361 to 3460 when the lowest complexity state and significant, energetic spectral peaks of 20 years are found. The century between simulation year 1981 and 2080, on the other hand, displays high-frequency variability and corresponds to the highest complexity state. For this domain, the largest jump in complexity appears over a period of about 300 years, from 3200 to 3500 (Figure 4b).

We note that results are constrained by the choices of time windows for the entropy computation (i.e., $W = 100$ years windows every $\Delta W = 20$ years) and that new insights could be obtained using shorter sliding windows (i.e., $\Delta W << 20$) or longer periods (i.e. $W > 100$ years). However, regime transitions are also reflected in changes in the topology of the networks between domains, and here, climate networks are inferred every 3 years, allowing for the investigation of regime shifts at very high temporal resolution (see next section).

4.3 Exploring variability in SST climate networks

In our approach, the domains in Figure 2 emerge as the true semi-autonomous components of monthly SST variability. They are considered as *true* components since they are spatiotemporal patterns of the system under study and largely independent of the spatial resolution of the grid and as *semi-autonomous* since they can interact via coupled atmospheric-ocean dynamics.
Here we investigate their time-dependent connectivity patterns using the network density metric ρ_N proposed in Eq. 4. Changes in the network density allow for the quantification of transitions in the degree of connectivity on a global scale. We focus on abrupt regime transitions.

We draw attention to two points: (a) a system with no regime transitions would imply constant values of the network density over time and (b) high correlations between the two densities would be obtained if the external forcings (i.e., orbital parameters and trace gases) were the dominant mechanisms for global regime transitions In the 6000 years explored here, the climate system undergoes relatively large dynamical changes in its network density from minima of 0.10 (0.07) to maxima of 0.19 (0.18) for the low (high) resolution simulations, respectively, and the correlation between the two densities is not significant (c.c. = 0.05) (Figure 5). Therefore, both simulations
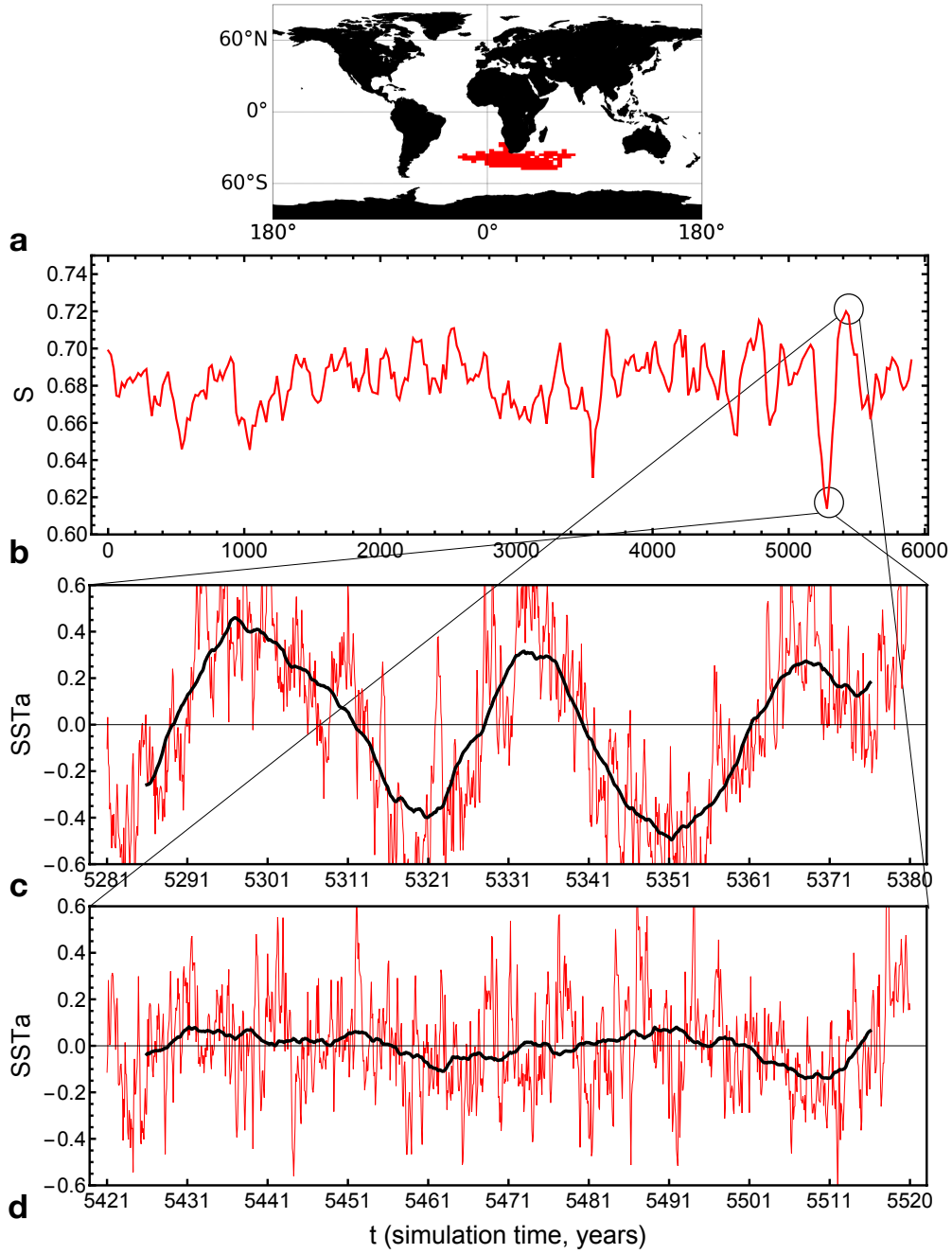
Figure 3: (a) Domain with the largest regime shift in the low resolution (Vlr01) simulation. (b) Mean complexity signal for the domain in (a) over time (x-axis: time in years; 6000 = year 1950; the year reported indicates the first year of the period considered i.e., year = 1 for the period going from year 1 to 100). (c, d) Associated monthly SST anomalies corresponding to centuries of minimum and maximum complexity, respectively; x-axis: time in years. The black lines indicate a 10-years moving average.
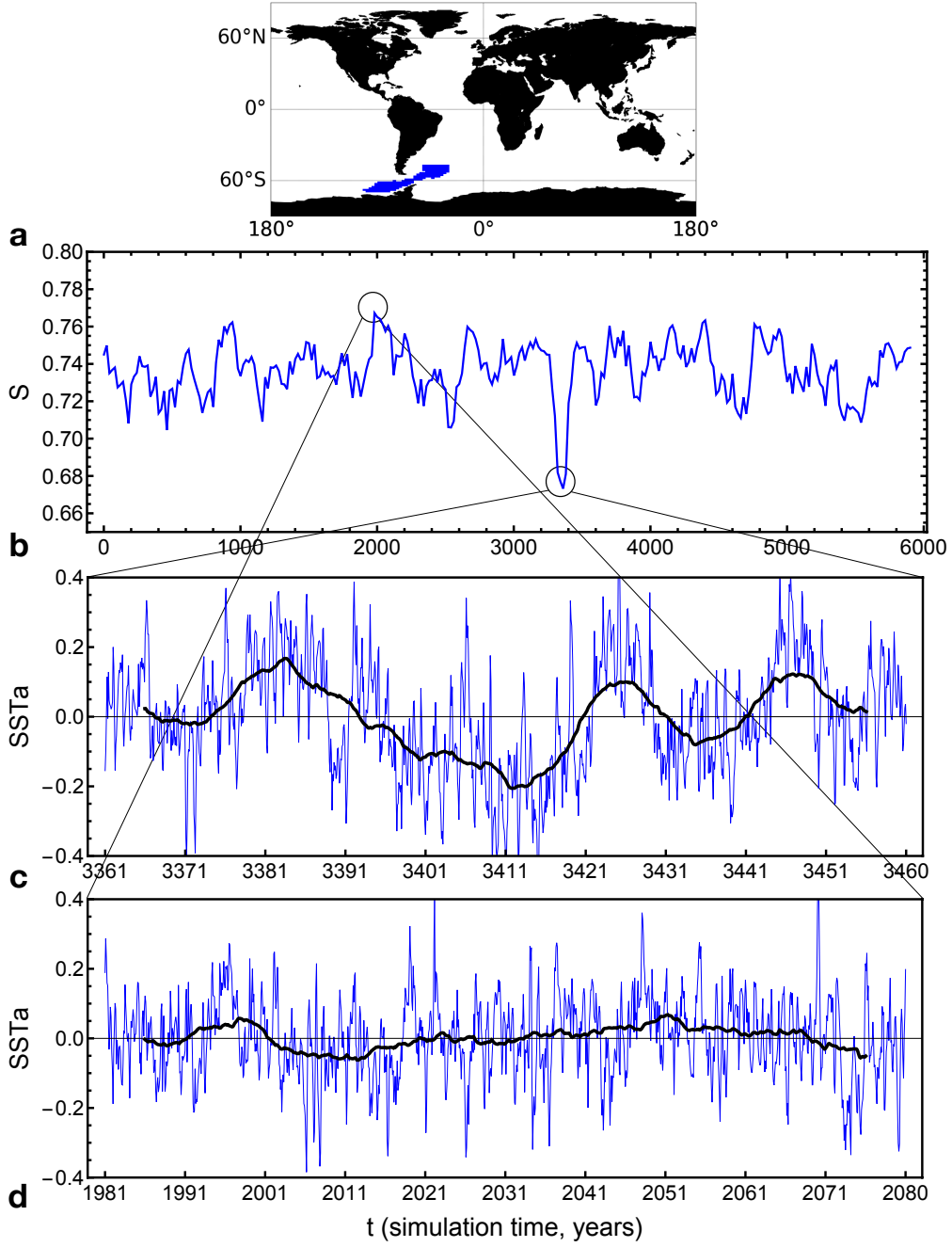
11

Figure 4: (a) a domain in the high resolution (Sr02) simulation with a large regime shift. (b) Mean complexity signal over time for the domain shown in (a) (x-axis: time in years; 6000 = year 1950; the year reported indicates the first year of the period considered i.e., year = 1 for the period going from year 1 to 100). (c, d): Monthly SST anomalies corresponding to centuries of minimum and maximum complexity, respectively; x-axis: time in years. The black lines indicate a 10-years moving average.
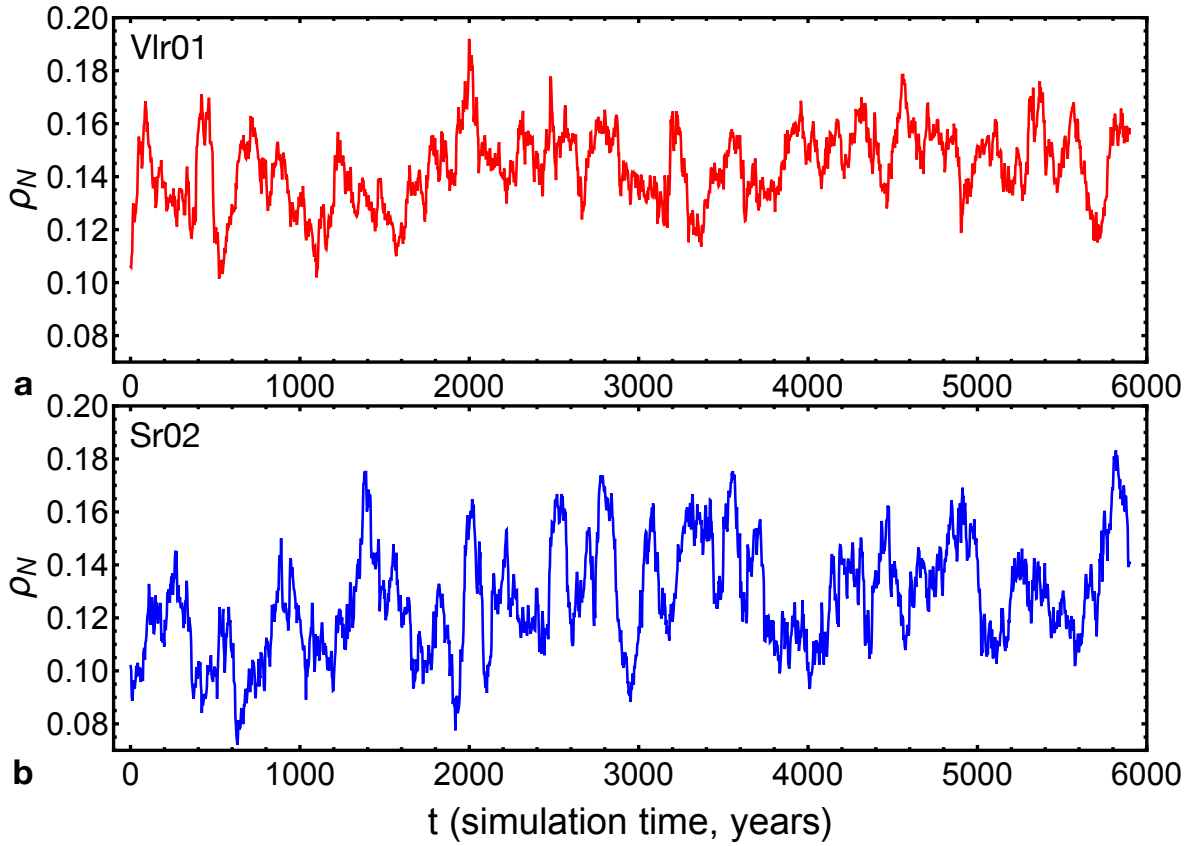
Figure 5: Evolution of the network density ($\rho_N = 1$ implies a complete weighted graph, $\rho_N = 0$ implies that the network has no links) in the (a) Vlr01 and (b) Sr02 simulations. x-axis: time in years; 6000 = year 1950; the year reported indicates the first year of the period considered i.e., year = 1 for the period going from year 1 to 100).
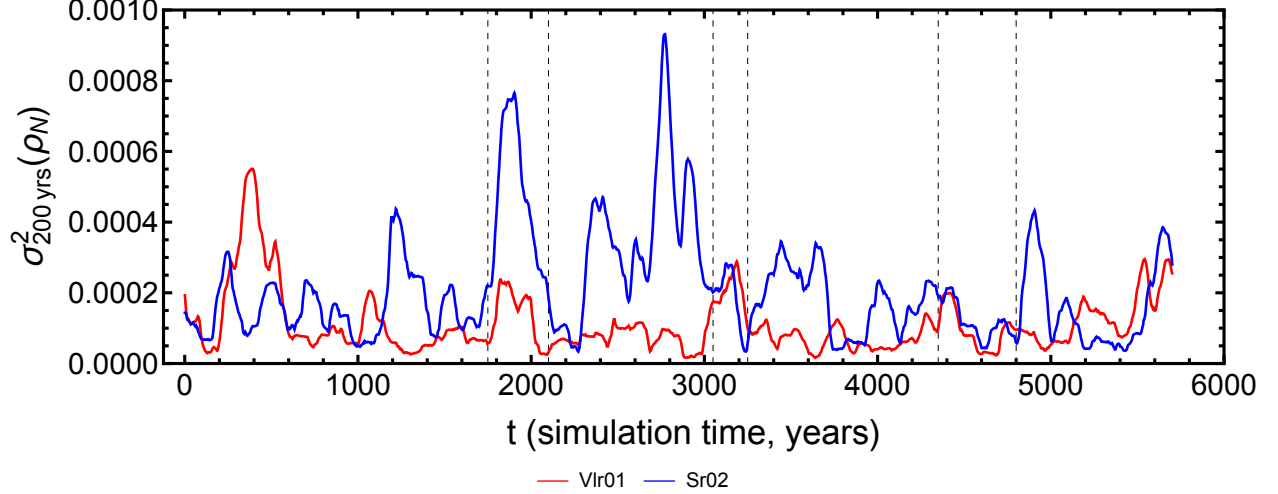
Figure 6: Network density variance in moving windows of 200 years every 3 years (minimum time step) respectively in the Vlr01 (red) and Sr02 (blue). x-axis: time in years; 6000 = year 1950; the year reported indicates the first year of the period considered i.e., year = 1 for the period going from year 1 to 100). Dashed lines indicate periods with qualitative similar shifts at the same time in both simulations.

display regime transitions driven mostly by internal variability. The mean network densities are very similar across simulations (i.e., 0.14 and 0.13 respectively for Vlr01 and Sr02). The variability in density, quantified by its standard deviation σ, on the other hand, is greater in the higher resolution run ($\sigma = 0.014$ and $\sigma = 0.021$ respectively for Vlr01 and Sr02).

Positive outliers in network densities (i.e., values greater than $\langle \rho_N \rangle + 2.5\sigma$) are indicative of large, sudden transitions to periods of high connectivity and are found in both simulations at different timings.

To quantitatively detect large abrupt shifts in connectivity, we consider values of the network density inside moving windows of 200 years every 3 years (3 years being time step used to compute the networks) and compute their variance for each window. Jumps in variance in periods as short as 200 years indicate robust, abrupt transitions in the climate connectivity of the system. Results are shown in Figure 6.

Abrupt changes in the variability of the network density are, on average, smaller in the Vlr01 than in the Sr02 simulation, with the maximum from year 1171 to 1270.

The density in the Sr02 undergoes frequent, abrupt changes approximatively from year ~ 2000 to ~ 4000 and exhibit fewer transitions in the first and last 2000 years of the simulation. The curves are not correlated (correlation $r = 0.04$), and major transitions in the two climate networks can be considered as an emergent [Dijkstra, 2019], unpredictable phenomena, arising from the internal variability of the system rather than from the (common) external forcing.

There are some exceptions to this view, such as the shifts identified approximatively around the years ~ 1800, ~ 3200 and ~ 5800. Around these periods, both simulations display large transitions in global connectivity. Nonetheless, the shifts around year ~ 3200 is of a different type in the two simulations: around this year, Vlr01 exhibits a large, abrupt increase in connectivity, while Sr02 the

14

opposite (Figure 5).

Interestingly, around the year ∼ 1800 (∼ 4200 BP) both simulations undergo a rapid, strong increase in connectivity of the network. The timing of this event agrees with the transition between the so-called Northgrippian and Meghalayan ages in the Holocene [et al., 2018], that was characterized by severe aridification and cooling that extended from North America to the Mediterranean basin and North Africa, all the way to India and China [deMenocal, 2001, Staubwasser et al., 2003, Rousseau et al., 2018-2019].

This suggests a possible dominant role of the external forcing in causing this specific transition. However, more ensemble members are needed to assess the significance of this result. Furthermore, the investigation of multiple fields including air-temperature and precipitation may shed light on the evolution, connectivity and wide-spread impacts of this regime shift. This is left for future work.

5   SST complexity in the "real" world

The model simulations analyzed in this study allow to explore SST variability at spatial and temporal resolution unavailable in the observational records. However, climate models are far from perfect, and results are compromised by structural model errors [Frigg et al., 2014]. Conclusions may be more informative when presented together with model biases [Thompson and Smith, 2019] and, therefore, we present a brief, qualitative comparison of the global SST complexity in the IPSL model and in an observational data set. For a first assessment of SST climate networks in the IPSL model we refer the reader to Fountalis et al. [Fountalis et al., 2015].

We consider the last 50 years of simulation, from 1900 to 1950, and compare the SST complexity reproduced by the IPSL model in the higher resolution configuration (Sr02) with the COBEv2 reanalysis [Hirahara et al., 2014] saved at 1° degree spatial resolution and as monthly averages (Figure 7). COBEv2 is also a model product and data availability in the first 50 years of the 20th century is rather poor [Hirahara et al., 2014]. Nonetheless, we expect the SST variability in COBEv2 to be more realistic than IPSL.

The complexity field shown in Figure 7 for the Sr02 simulation is qualitatively similar to the mean complexity fields shown in Figure 1. This highlights that SST complexity is quasi-stationary and trends are small in the mid- to late Holocene transient simulations.

The average latitudinal dependence of complexity in the observational data (Figure 7b,d) suggests larger predictability in the tropics (from ∼ 25°S to ∼ 25°N) and lower at higher latitudes (from ∼ 60°S to ∼ 25°S and ∼ 25°N to ∼ 60°N).This is due to the low entropy values in the equatorial Atlantic and Pacific basins. Therefore, the complexity of the simulated SST in the tropical Atlantic is largely overestimated by the model. Large complexity characterizes the Indian Ocean in both data sets, but is less uniform and weaker in COBEv2 than in the model.

The complexity of the ENSO region emerges in the reanalysis as confined to the central and eastern Pacific and is more uniform than in the model. Interestingly, two hemispheric-symmetric stripes marked with low complexity can be identified also in COBEv2 for the ENSO area, suggesting larger predictability potential for the El Niño events just South of the equator and East of 180°E.
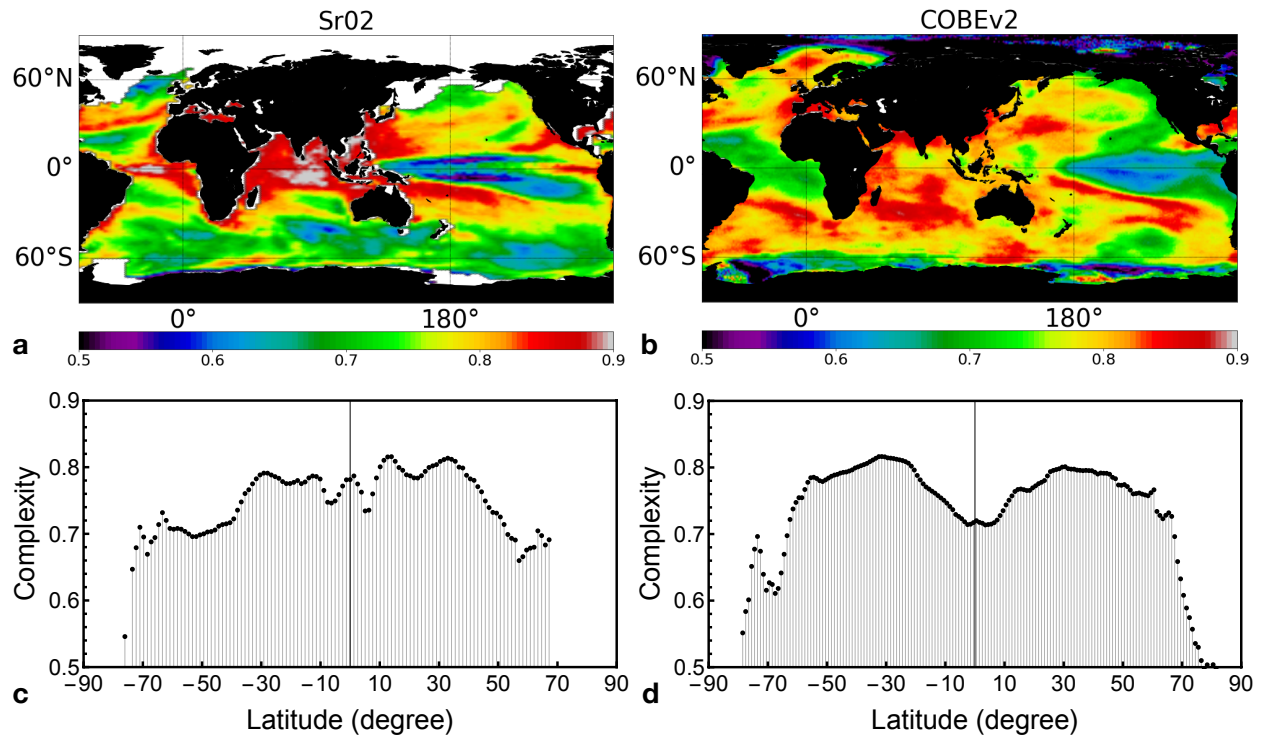
Figure 7: (a,b) Complexity field for the Sr02 simulation and the COBEv2 reanalysis for the 1900-1950 period. (c,d) Average latitudinal dependence of complexity in both fields.

6 Discussion

This work presents a new framework stemming from dynamical system theory and complex network analysis to investigate regime transitions in spatiotemporal climate fields in a comprehensive way. The approach couples a complexity quantifier [Corso et al., 2018] to a dimensionality reduction methodology [Fountalis et al., 2018] in order to (a) identify spatial regions (i.e., domains) of grid cells with highly correlated evolution of their complexity and (b) study the time-dependent, weighted, direct network between these regions. The approach is applied to two 6000-years long simulations of the IPSL Earth system model. Both simulations develop highly non-trivial spatiotemporal variability, which may stem at time from common forcing (changes in Earth's orbit and trace gases) but more often from intrinsic, internal variability. We tested the proposed strategy focusing on decadal to multidecadal SST variability in the last 6000 years, providing an analysis of regime transitions and time-dependent networks that simultaneously embraces spatial and temporal scales.

The results of our analysis can be summarized as follows:

- SST variability in the IPSL model is characterized by relatively low mean complexity (i.e., high predictability) in high latitudes and high complexity (i.e., low predictability) in the equatorial Indian and Atlantic Oceans, and along major eastern and western boundary currents. The ENSO region displays a non-trivial spatial configuration of mean complexity, with high values in the cold tongue and low values and higher predictability in two hemispheric-symmetric bands in the central Pacific.

- The dimensionality reduction of the entropy field allows for the identification of its underlying domains. Each domain is a set of points with highly correlated variability in information entropy. A domain's complexity signal, in turn, allows for exploring regime shifts in time. We provided evidence for transitions in local system dynamics in the two simulations that are identified as abrupt jumps of large amplitude from low to high-frequency SST variability and vice versa in periods as short as 200 to 300 years. Our results suggest that, at least in the modeled climate, multidecadal oscillations can rapidly emerge and fade due to the system's internal dynamics (i.e., independent of the external forcings). More theoretical and data driven work is needed to better understand the physical underlying mechanisms behind these shifts and their expression and impacts on the overall system and not limited to SSTs.

- The identified domains can be considered as the true semi-autonomous components of the system in the 6000-years long simulation. Domains can interact via coupled atmospheric-ocean dynamics, giving rise to complex connectivity patterns. These connections, and their variability, can be studied as a time-dependent, weighted and direct network between domains. The time evolution of the network density reveals that the whole system underwent large shifts in the degree of connectivity during the last 6000 years. These shifts occur, in most cases, at different times in the two simulations, implying that most transitions in connectivity may be unpredictable, emergent phenomena [Dijkstra, 2019]. This suggests that changes in connectivity patterns in timescales as long as 1000 years are dominated by the internal variability of the system rather than by external forcings, at least in the last 6000 years. Finally, we identified a common (among the two simulations) large shift around year 1800 (i.e., 4200 years BP) consistent with previous findings in paleoclimate proxies [Rousseau et al., 2018-2019]. A larger ensemble is required to confirm if this shift in connectivity is a response

17

to external forcing. Exploring large regime shifts in different variables and quantifying their underlying causal associations using state-of-the-art methodologies [Runge et al., 2019] is left for future work.

- In comparing the SST complexity of the IPSL model with that of a reanalysis product over the first half of the 20th century, differences were mainly found in the tropical Atlantic and Pacific basins. The modeled complexity in the Atlantic around the Equator and in the cold tongue is largely overestimated, while the ENSO domain extends too far west covering most of the equatorial Pacific. We note that evaluating a climate model's ability in reproducing the observed complexity is a very stringent (but useful) test: high similarity could only be obtained if models and observations shared the same underlying dynamics. Such evaluation allows for testing directly the dynamics of the model rather than comparing signal amplitudes or mean states and opens interesting research avenues.
Additionally, while dynamical transitions in SST emerge at relatively long-time scales (i.e., hundreds of years), future work could focus on variables expected to show regime shifts at shorter timespans. This would include precipitation or cloud fraction fields in high temporal resolution datasets and/or model integrations.

Transitions or shifts in local and, more importantly, global system dynamics are expected in complex, nonlinear spatiotemporal climate fields. While climate models are more often used for future projections and projected climate shifts cannot be validated, long simulations covering past millennia allow for exploring profound changes in the system dynamics that may have already taken place and for which we have sparse records. In this work we investigated the relationships between these transitions, external forcing and internal climate variability. Our results suggest that shifts in local dynamics and global connectivity patterns in sea surface temperature may often appear as emergent and abrupt phenomena, independently of the external forcing fields because of the internal chaotic dynamics of the system.

Acknowledgements

Data availability

The climate simulations used in this manuscript are freely available (`https://doi.org/10.14768/20191028001.1`). Python codes for recurrence infomation entropy are available at `https://github.com/FabriFalasca/NonLinear_TimeSeries_Analysis`. δ-MAPS is available at `https://github.com/FabriFalasca/py-dMaps`. Additionally, some Mathematica codes for network inference between domains are also available at `https://github.com/FabriFalasca/network_Inference`.

A δ-MAPS

Core identification.   The domain identification algorithm starts from epicenters or cores. Given a K-neighborhood $\Gamma_K(i)$ of a cell i, we compute its local homogeneity as:

$$r_K(i) = \frac{\sum_{m \neq n \in \Gamma_K(i)} r_{m,n}}{K(K+1)} \ . \tag{A.1}$$

A similar notation is used for the homogeneity of a set of grid cells:

$$r(A) = \frac{\sum_{m \neq n \in A} r_{m,n}}{|A|(|A|-1)} \ , \tag{A.2}$$

$|A|$ being the cardinality of the set and and $r_{m,n}$ the correlation (at lag $\tau = 0$) between the time series embedded in grid cells m and n. A grid cell is marked as a core if its local homogeneity is a local maximum and greater than a threshold δ. More formally, a grid cell i is a core if $r_K(i) > \delta$ and $r_K(i) > r_K(j)$, $i \neq j$, $\forall j \in \Gamma_K(i)$. Cores are then iteratively expanded and merged to identify domains [22].

Domain identification.   Given one or more identified cores c, a spatial grid G and a threshold δ, a domain A is the maximal set of grid cells satisfying three constraint: (a) $c \in A$, (b) $I_G(A) = 1$ and (c) $r(A) > \delta$. $I_G(A) = 1$ denotes that the set A is spatially contiguous. Fountalis et. al [Fountalis et al., 2018] proved that this problem is NP-Hard and relied on a heuristic to solve it. The heuristic algorithm iteratively expands and merges cores to find the full extent of domains. The process starts from the domain with largest homogeneity, let it be A. The expansion algorithm considers all adjacent grid cells of A, and then adds the grid cell i which maximizes the homogeneity $r(A \cup i)$ and for which $r(A \cup i) > \delta$. If two domains A and B are adjacent, the merging algorithm determines that the two domains should be merged whenever $I_G(A \cup B) = 1$ and $r(A \cup B) > \delta$. This process stops when no more merging and expansions are possible. Two heuristics to infer the δ threshold and the K parameter has been proposed in [Fountalis et al., 2018] and [Falasca et al., 2019]. Given a significance level α, the threshold δ is computed as a sample average of the statistically significant cross-correlations between randomly chosen grid cells in the dataset considered.

B   Heuristic for the vicinity threshold ϵ

A crucial step in the computation of a recurrence plot (RP) is the definition of the vicinity threshold ϵ. The ϵ parameter defines the neighborhood of each state $\mathbf{x}_i$ in a trajectory of a dynamical system. It must be carefully chosen since a value that is too small may include noisy fluctuations in the RP, and one that is too large could hide the recurrence structure of the time series [Marwan et al., 2007]. Several "rules of thumb" for the choice of the ϵ threshold have been proposed: $\sim 5\%$ of the maximal state space diameter [Mindlin and Gilmore, 1992], no more than $\sim 10\%$ of the mean (or maximum) state space diameter [Koebbe and Mayer-Kress, 1992, Zbilut and Jr., 1992] or a value ensuring a recurrence point density of $\sim 1\%$ [Zbilut et al., 2002]. Moreover, choosing ϵ strongly depends on the system under study [30].

Here we present a new simple heuristic tailored for the selection of a different ϵ_i for each (not-embedded) time series $x_i(t)$ of a spatiotemporal field $\mathbf{X}(t)$.

The proposed strategy relies on the following steps:

1. we consider a monthly spatiotemporal climate field $\mathbf{X}(t)$ embedded in a two-dimensional grid at each time step t. For every grid cell i we have a linearly detrended, anomaly time series $x_i(t)$ with T data points. We define an ϵ_i as a percentage ρ of the standard deviation σ_i of $x_i(t)$. This step defines a new spatial field $\mathbf{E}(\rho)$, referred to as ϵ-map.

2. For a given ϵ-map $\mathbf{E}(\rho)$, we compute the RP and its information entropy for every $x_i(t)$ in $\mathbf{X}(t)$ to get a spatial entropy field $\mathbf{S} = \mathbf{S}(\rho)$. Since the entropy field depends on ρ, we assess its sensitivity to ρ as follows: first, we compute $\mathbf{S}(\rho)$ for ρ ranging from $\rho_{min} = 0.01$ to $\rho_{max} = 1.50$ every $\Delta\rho = 0.01$; second, we define a pairwise distance matrix between all $\mathbf{S}(\rho)$. Given two fields, $\mathbf{S}(\rho_1)$ and $\mathbf{S}(\rho_2)$, we compute their distance as $D(\mathbf{S}(\rho_1), \mathbf{S}(\rho_2)) = \sum_i |s_i(\rho_1) - s_i(\rho_2)|$, where $s_i(\rho)$ denotes the entropy at grid cell i.

3. Finally, we select an optimal $\rho = \rho^*$ such that the computed entropy field does not depend on perturbations around ρ^*. This assures that, for every $s_i(\epsilon_i)$ at grid cell i, $s_i(\epsilon_i) \sim s_i(\epsilon_i \pm \delta_i)$, $\delta_i << \epsilon_i$.

In case of results presented in section 5, we obtain a value of $\rho^* = 0.75$ both for COBEv2 and for the Sr02 simulation. We then assign to each grid cell i a value $\epsilon_i = \rho^* \cdot \sigma_i$, with $\rho^* = 0.75$.

For results in section 4, we face an additional problem: we have to compute a spatiotemporal entropy field and not just one time-instant. Therefore, for this step we compute $\mathbf{E}(\rho)$ using the complete simulation length. We consider a time series $\hat{x}_i(t)$ of monthly SST defined at grid cell i (with length $T = 72000$ months) and compute a new time series $x_i(t)$ representing the evolution of detrended anomalies at all times t. SST anomalies $x_i(t)$ are computed by removing both the seasonal cycle and a linear trend. This is done for each 100-yr non-overlapping window instead of the 6000-yr timeseries to account, at first order, for changes in SST seasonality and (small) trends induced by orbital changes from mid- to late Holocene [29]. The parameter ϵ_i is thus constant in time. Given a $x_i(t)$, this implies a notion of neighborhood as a function of all possible states explored in the 6000-yr long simulation (i.e., the closeness between two states is conditioned on distances between all states in the 6000-yr simulation).

The heuristic is then tested for the first $W = 100$ years of the simulation. This analysis is found to
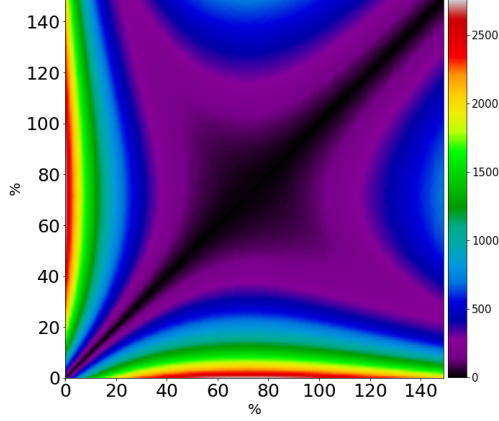
Figure B.1: Pairwise distance matrix between the spatial complexity fields computed for the first 100 years as a function of the percentage of standard deviation of the 6000 years detrended anomalies time series defined at each grid cell. Similar results are obtained for the COBEv2 dataset and for other centuries in the simulations.

be independent of resolution. A global minimum in the pairwise distance matrix is identified around the values of $\rho^* = 75\%$ (see Figure B.1). We then assign to each grid cell i a value $\epsilon_i = \rho^* \cdot \sigma_i$, with $\rho^* = 0.75$.

It should be noted that in the analyzed simulations the standard deviations of time series in each grid cell i do not undergo profound changes: the ϵ-map $\mathbf{E}$ obtained using the 6000 years of the simulations is qualitative similar to the one obtained using a randomly chosen period of 100 years. Therefore, the optimization problem should not depend on the century analyzed. We verified this assumption by re-computing the heuristic for three, randomly picked centuries. Results were indistinguishable from those shown in Figure B.1.

C Shifts for a domain in the South Atlantic ocean in the Sr02 simulation

Here we present a domain in the Sr02 simulation similar to the one shown in Figure 3 for the Vlr01 simulation. The two complexity signals shown in Figure C.1 and Figure 3 have same mean (0.68) but shifts are found at different times.
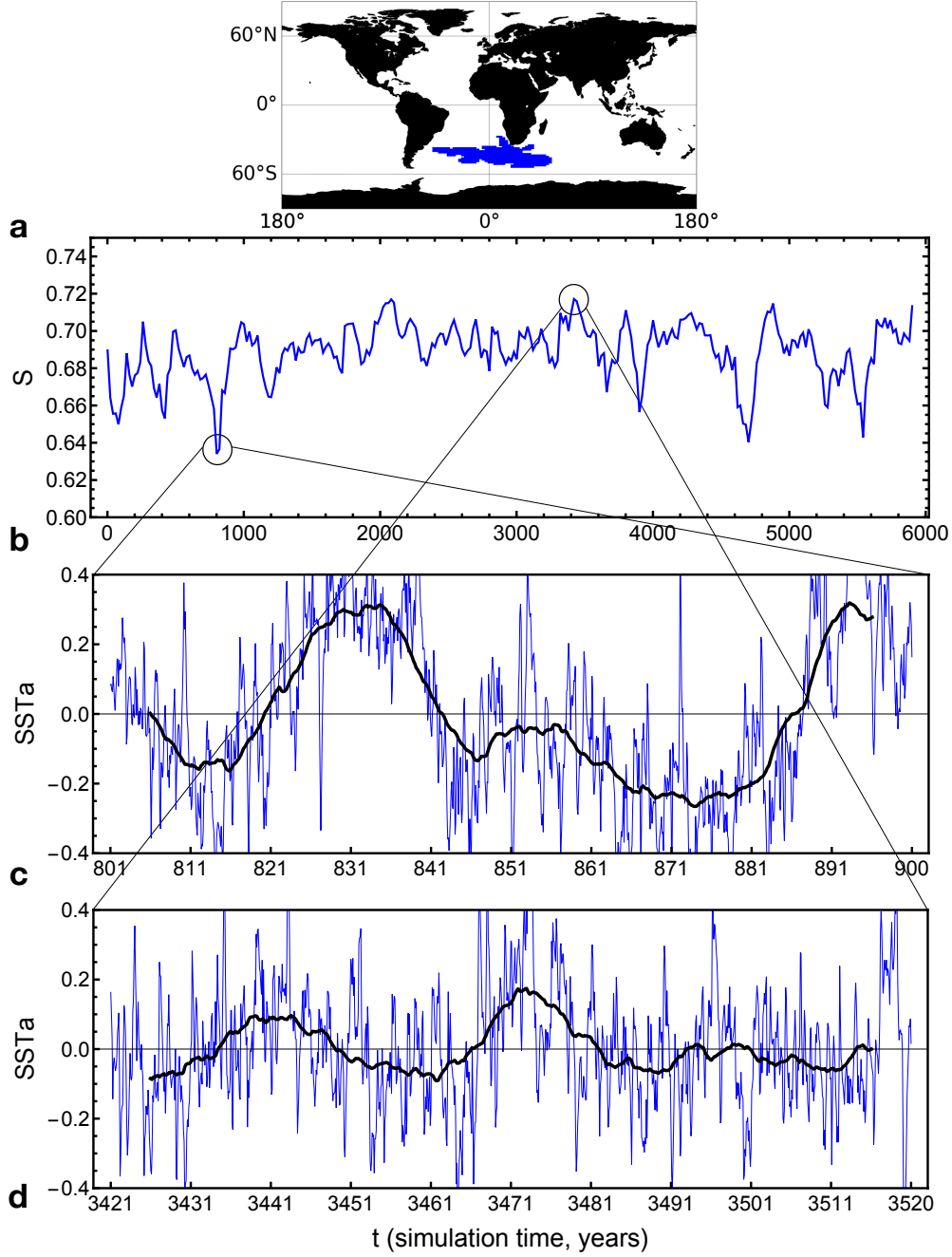
Figure C.1: (a) Domain in the high resolution (Sr02) simulation. (b) Mean complexity signals for the domain in (a) over time (x-axis: time in years; 6000 = year 1950; the year reported indicates the first year of the period considered i.e., year = 1 for the period going from year 1 to 100). (c, d) Associated monthly SST anomalies corresponding to centuries of minimum and maximum complexity, respectively; x-axis: time in months. The black lines indicate a 10-years moving average.

References

P. F. Hoffman and D. P. Schrag. The snowball earth hypothesis: testing the limits of global change. *Terra nova*, 14:129, 2002.

R. T. Pierrehumbert, D. Abbot, A. Voigt, and D. Koll. Climate of the neoproterozoic. *Earth Plan. Sci.*, 39:417, 2011.

V. Lucarini and T. Bódai. Edge states in the climate system: exploring global instabilities and critical transitions. *Earth Plan. Sci.*, 30:R32, 2017.

V. Lucarini and T. Bódai. Transitions across melancholia states in a climate model: reconciling the deterministic and stochastic points of view. *Phys. Rev. Lett.*, 122:15, 2019.

W. Dansgaard, H.B. Clausen, N. Gundestrup, C.U. Hammer, S.F. Johnsen, P.M. Kristinsdottir, and N. Reeh. A new greenland deep ice core. *Science*, 218:1273–1277, 1982.

S. J. Johnsen, H.B. Clausen, W. Dansgaard, K. Fuhrer, N. Gundestrup, C.U. Hammer, P. Iversen, J. Jouzel, B. Stauffer, and J.P. Seffensen. Irregular glacial interstadials recorded in a new greenland ice core. *Nature*, 359:311–313, 1992.

F. Gasse. Hydrological changes in the african tropics since the last glacial maximum. *Quaternary Science Reviews*, 19:189–211, 2000.

C. Morrill, J. Overpeck, and J.E. Cole. A synthesis of abrupt changes in the asian summer monsoon since the last deglaciation. *Holocene*, 13:4, 465–476, 2003.

A.C. Mix, D.C. Lund, N.G. Pisias, P. Bóden, L. Bornmalm, M. Lyle, and J. Pike. Rapid climate oscillations in the northeast pacific during the last deglaciation reflect northern and southern hemisphere sources. *Mechanisms of Global Climate Change at Millennial Time Scales, AGU Monograph, American Geophysical Union, Washington, DC*, page 27–148, 1999.

P.B. deMenocal. Cultural responses to climate change during the late holocene. *Science*, 292: 5517,667–673, 2001.

R.B. Alley, J. Marotzke, W.D. Nordhaus, J.T. Overpeck, D.M. Peteet, R.A. Pielke Jr., R.T. Pierrehumbert, P.B. Rhines, T.F. Stocker, L.D. Talley, and J.M. Wallace. Abrupt climate change. *Science*, 299:2005–2010, 2003.

T.M. Lenton, H. Held, E. Kriegler, J.W. Hall, W. Lucht, S. Rahmstorf, and H.J. Schellnhuber. Tipping elements in the earth's climate system. *Proc. Natl Acad. Sci.*, 105:1786–1793, 2008.

G. Balasis, R.V. Donner, S.M. Potirakis, J. Runge, C. Papadimitriou, I.A. Daglis, K. Eftaxias, and J. Kurths. Statistical mechanics and information-theoretic perspectives on complexity in the earth system. *Entropy*, 15:4844–4888, 2013.

R.V. Donner, M. Small, J.F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths. Recurrence-based time series analysis by means of complex network methods. *Int. J. Bifurcation Chaos*, 14: 1019–1046, 2011a.

J.F. Donges, R.V. Donner, M.H. Trauth, N. Marwan, and H.J. Schellnhuber. Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. *Proc. Natl. Acad. Sci.*, 108:20422–20427, 2011.

J.F. Donges, R.V. Donner, N. Marwan, S.F.M. Breitenbach, K. Rehfeld, and J. Kurths. Non-linear regime shifts in holocene asian monsoon variability: potential impacts on cultural change and migratory patterns. *Clim. Past*, 11:709–741, 2015.

C. Bandt and B. Pomp. Permutation entropy: a natural complexity measure for time serie. *Phys. Rev. Lett.*, 88:174102, 2002.

P. M. Saco, L.C. Carpi, A. Figliola, E. Serrano, and O.A. Rosso. Entropy analysis of the dynamics of el niño/southern oscillation during the holocene. *Physica A*, 389:5022, 2010.

J. Garland, T.R. Jones, M. Neuder, V. Morris, J.W.C. White, and E. Bradley. Anomaly detection in paleoclimate records using permutation entropy. *Entropy*, 20:931, 2018.

J.P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57:617, 1985.

G. Corso, T. de L. Prado, G. Z. dos Santos, J. Kurths, and S. R. Lopes. Quantifying entropy using recurrence matrix microstates. *Chaos*, 28:083108, 2018.

I. Fountalis, C. Dovrolis, A. Bracco, B. Dilkina, and S. Keilholz. δ-MAPS from spatio-temporal data to a weighted and lagged network between functional domain. *Appl. Netw. Sci.*, 3:21, 2018.

R.V. Donner, S.M. Barbosa, J. Kurths, and N. Marwan. Understanding the earth as a complex system - recent advances in data analysis and modelling in earth sciences. *Eur. Phys. J. Special Topics*, 174:1–9, 2009.

J.F. Donges, Y. Zou, N. Marwan, and J. Kurths. Complex networks in climate dynamics. *Eur. Phys. J. Special Topics*, 174:157–179, 2009.

A. L. Barabási. Network science. *Cambridge, UK: Cambridge University Press*, 2016.

M. Newman. Networks: An introduction. *Oxford, UK: Oxford University Press*, 2010.

J-L. Dufresne and et al. Climate change projections using the IPSL-CM5 earth system model: From CMIP3 to CMIP5. *Clim. Dyn.*, 40:2123–2165, 2013.

P. Braconnot, D. Zhu, O. Marti, and J. Servonnat. Strengths and challenges for transient mid- to late holocene simulations with dynamical vegetation. *Clim. Past*, 15:997–1024, 2019a.

P. Braconnot, J. Crétat, O. Marti, Y. Balkanski, A. Caubel, A. Cozic, M.A. Foujols, and S. Sanogo. Impact of multiscale variability on last 6000 years indian and west african monsoon rain. *Geophys. Res. Lett.*, 46(23):14021–14029, 2019b.

N. Marwan, M. C. Romano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438:237–329, 2007.

J.P. Eckmann, S.O. Kamphorst, and D. Ruelle. Recurrence plots for the analysis of complex systems. *Europhys. Lett.*, 5:973–977, 1987.

H. Poincaré. Sur le probleme des trois corps et les équations de la dynamique. *Acta Mathematica*, 13:1–271, 1890.

P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay. Chaos: Classical and quantum. *ChaosBook.org, Niels Bohr Institute, Copenhagen*, 2016.

R.V. Donner, J. Heitzig, J.F. Donges, Y. Zou, and N. Marwan. The geometry of chaotic dynamics – a complex network perspective. *Eur. Phys. J N*, 2011b.

F. Takens. Detecting strange attractors in turbulence. *D. Rand, L.S. Young (Eds.), Dynamical Systems and Turbulence,Lecture Notes in Mathematics*, 898:366–381, 2011.

N. Marwan. How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos*, 21(4):1003–1017, 2011.

N. Wessel M. Riedl, A. Muller. Practical considerations of permutation entropy. *Eur. Phys. J Special Topics*, 222:249–262, 2013.

F. Falasca, A. Bracco, A. Nenes, and I. Fountalis. Dimensionality reduction and network inference using δ-MAPS: Application to the cesm large ensemble sea surface temperature. *JAMES*, 11(6), 2019.

Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B*, 257:289–300, 1995.

G.E. Box, G.M. Jenkins, and G.C. Reinsel. Timeseries analysis: forecasting and control. *John Wiley Sons*, pages 289–300, 2011.

A. Berger. Long-term variations of caloric solar radiation resulting from the earth's orbital elements. *Quaternary Res.*, page 139–167, 1978.

B. L. Otto-Bliesner et al. The PMIP4 contribution to CMIP6 - part 2: Two interglacials, scientific objective and experimental design for holocene and last interglacial simulations. *Geosci. Model Dev.*, 10:3979–4003, 2017.

P. Imkeller and J.S. Von Storch. Stochastic climate models. *Springer Science Business Media*, 49, 2001.

H. A. Dijkstra, E. Hernández-Garcia, C. Masoller, and M. Barreiro. Networks in climate. *Cambridge Universtity Press*, 2019.

H. A. Dijkstra. Numerical bifurcation methods applied to climate models: analysis beyond simulation. *Nonlin. Processes Geophys.*, 26:359–369, 2019.

M. Walker et al. Formal ratification of the subdivision of the holocene series/epoch (quaternary system/period): two new global boundary stratotype sections and points (gssps) and three new stages/subseries. *Subcommission on Quaternary Stratigraphy (SQS)*, 41(4):213–223, 2018.

M. Staubwasser, F. Sirocko, P.M. Grootes, and M. Segl. Climate change at the 4.2 ka BP termination of the indus valley civilization and holocene south asian monsoon variability. *Geophys. Res. Lett.*, 30(8), 2003.

D.D. Rousseau, G. Zanchetta, H. Weiss, and M. Bini. The 4.2 ka BP climatic event [special issue]. *Clim. Past*, 30(8), 2018-2019.

R. Frigg, S. Bradley, H. Du, and L.A. Smith. Laplace's demon and the adventures of his apprentices. *Philosophy of Science*, 81:31–59, 2014.

E. Thompson and L.A. Smith. Escape from model-land. *Economics: The Open-Access, Open-Assessment E-Journal*, 13:1–15, 2019.

I. Fountalis, A. Bracco, and C. Dovrolis. ENSO in CMIP5 simulations: network connectivity from the recent past to the twenty-third century. *Clim. Dyn.*, 45:511–538, 2015.

S. Hirahara, M. Ishii, and Y. Fukuda. Centennial-scale sea surface temperature analysis and its uncertainty. *Journal of Climate*, 27:57–75, 2014.

J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic. Detecting and quantifying causal association in large nonlinear time series datasets. *Science Advances*, 5 (11), 2019.

G. B. Mindlin and R. Gilmore. Topological analysis and synthesis of chaotic time series. *Physica D*, 58:229–242, 1992.

M. Koebbe and G. Mayer-Kress. se of recurrence plots in the analysis of time-series dats. *Proceedings of SFI Studies in the Science of Complexity, Redwood City, Addison-Wesley, Reading, MA, 1992, edited by M. Casdagli and S. Eubank*, XXI, 1992.

J. P. Zbilut and C.L. Webber Jr. Embedding and delays as derived from quantification of recurrence plots. *Physics Letter A*, 171:199–203, 1992.

J. P. Zbilut, J.M. Zaldívar-Comenges, and F. Strozzi. Recurrence quantification based lyapunov exponents for monitoring divergence in experimental data. *Physics Letter A*, 297:173–181, 2002.