Como extraer información de señales complejas. Ejemplos de análisis de datos climáticos, ópticos y biomédicos

Cristina Masoller
Departamento de Fisica
Universitat Politecnica de Catalunya

Slides:
https://www.fisica.edu.uy/~cris/trefemac.pdf

Outline

- Time series analysis tools
  - Ordinal analysis
  - Hilbert analysis
- Applications:
  - Lasers
  - Brain
  - Climate

UPC

About me

- Originally from Montevideo, Uruguay
- Bachelor and Master degrees from Facultad de Ciencias, Universidad de la Republica
- PhD degree from Bryn Mawr College, Pensylvania, USA
- Since 2004 @ *Departamento de Fisica*, *Universitat Politecnica de Catalunya*.

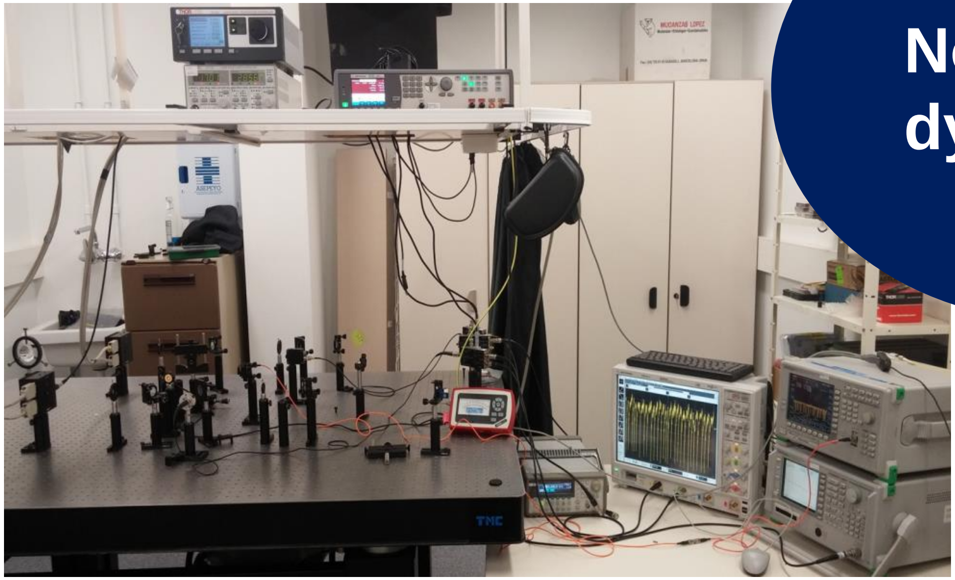
Where are we?

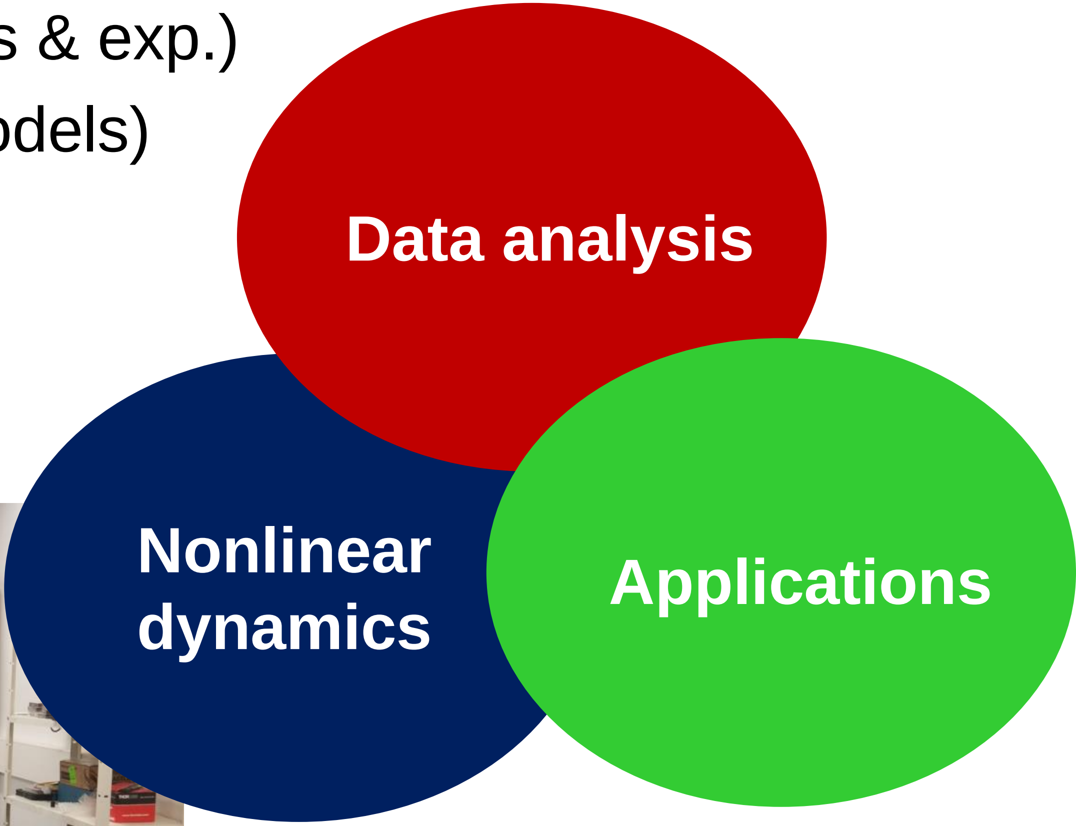

Viernes, 25 de septiembre de 2009 Diari de Terrassa

El edificio Gaia centraliza grupos científicos consolidados y emergentes.

What do we study?

- laser dynamics (models & exp.)
- neuronal dynamics (models)
- complex networks
- climate data analysis
- biomedical signals





Thanks to my co-authors

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Maria Masoliver, Carlos Quintero, Jordi Tiana, Carme Torrent (*ordinal analysis: neuronal, optical and brain signals*)
- Dario Zappala and Marcelo Barreiro (*climate data*)

How to extract information from complex signals?

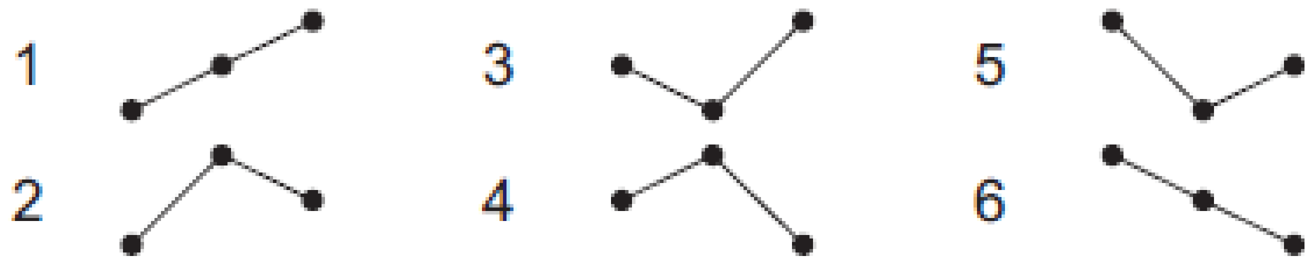
First analysis tools: ordinal symbolic analysis

It allows to identify “patterns” in data

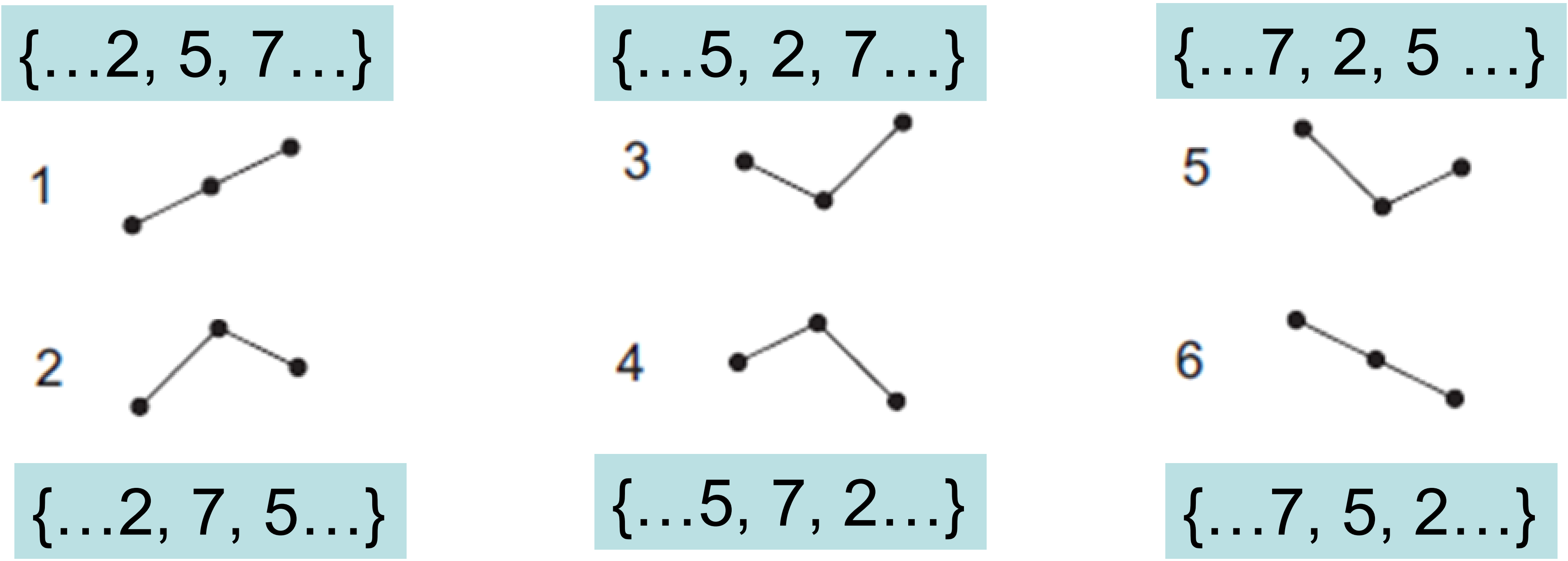
Ordinal analysis: brief introduction

- Consider a time series $x(t)=\{\ldots x_i,\ x_{i+1},\ x_{i+2},\ \ldots\}$

- Which are the possible order relations among three data points?



Bandt and Pompe: Phys. Rev. Lett. 2002

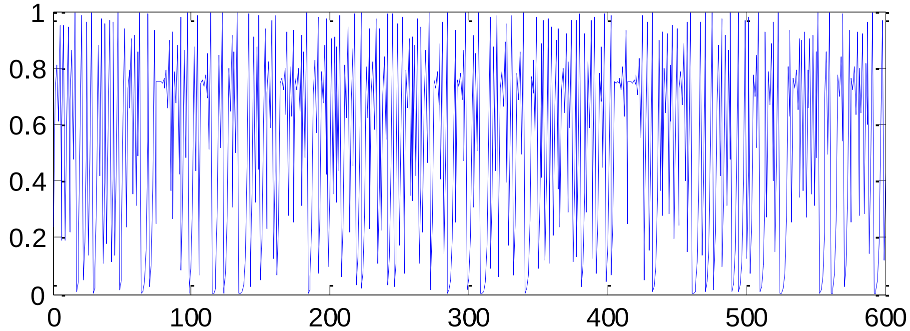


- Analyze the frequency of occurrence of “ordinal patterns” (“ordinal probabilities”)
- They allow to identify temporal structures
- Robust to noise.
- Drawback: information about actual data values is lost.

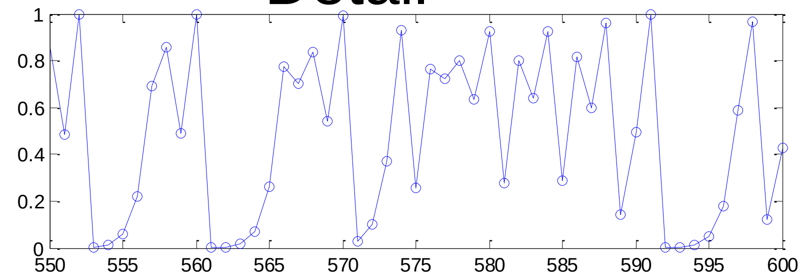
Example: Logistic map

$$x(i+1) = r\, x(i)[1 - x(i)]$$

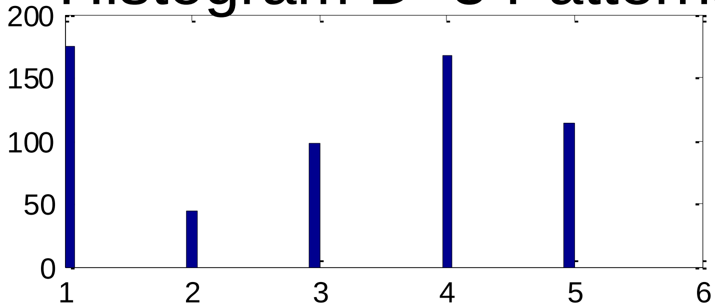
Time series *r*=4



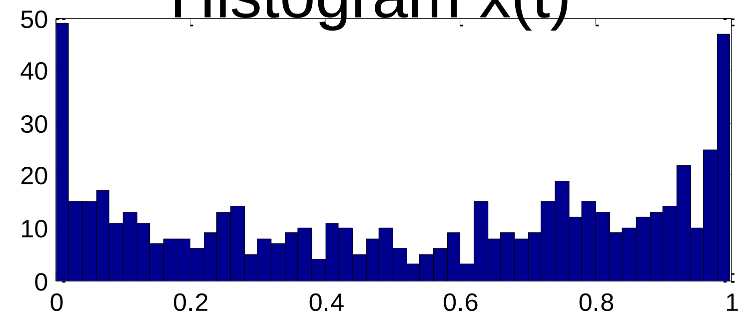
Detail

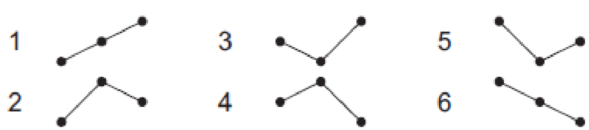


Histogram D=3 Patterns



Histogram x(t)



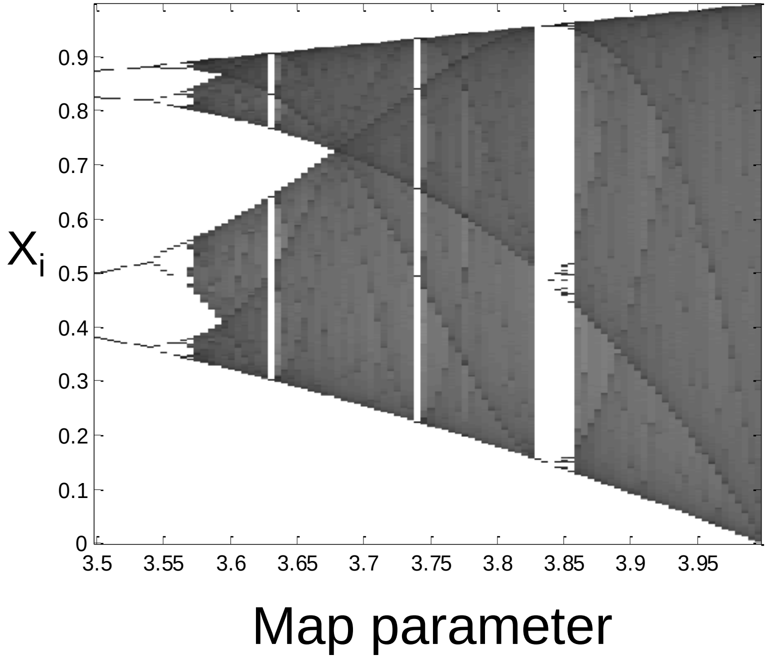


↑ **forbidden**

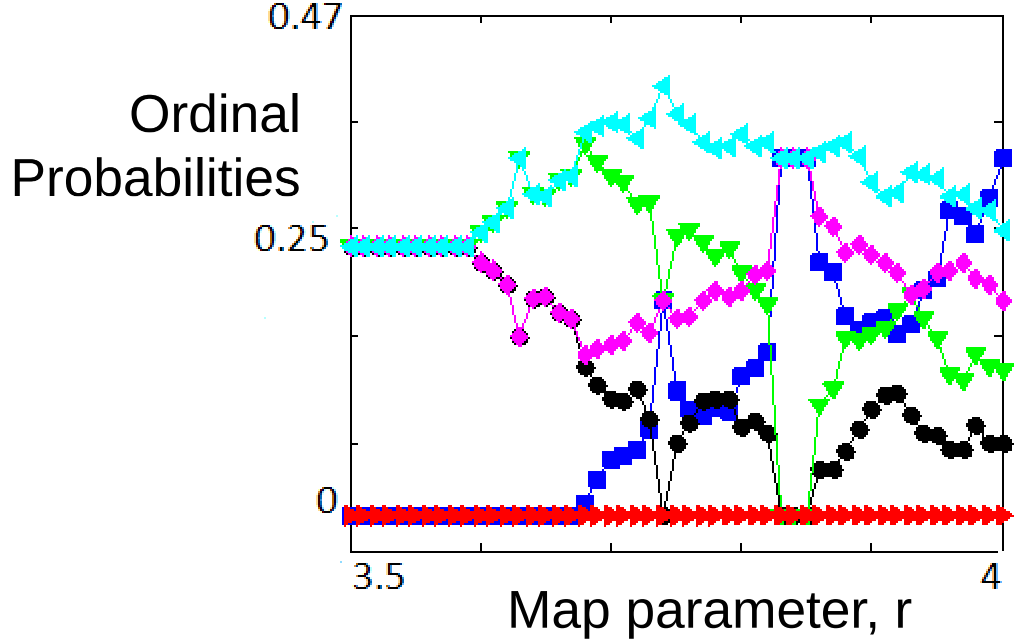
Ordinal analysis yields information about more expressed and less expressed oscillation patterns in the data.

Ordinal analysis yields complementary information

Normal bifurcation diagram

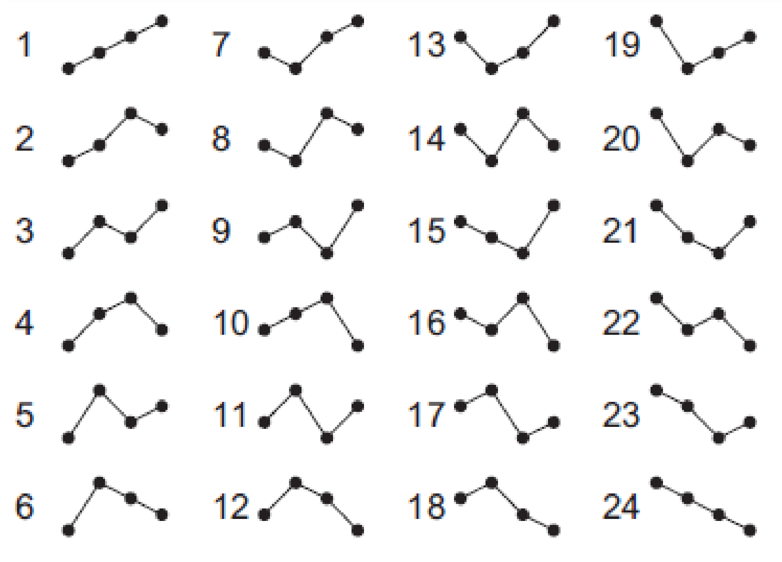


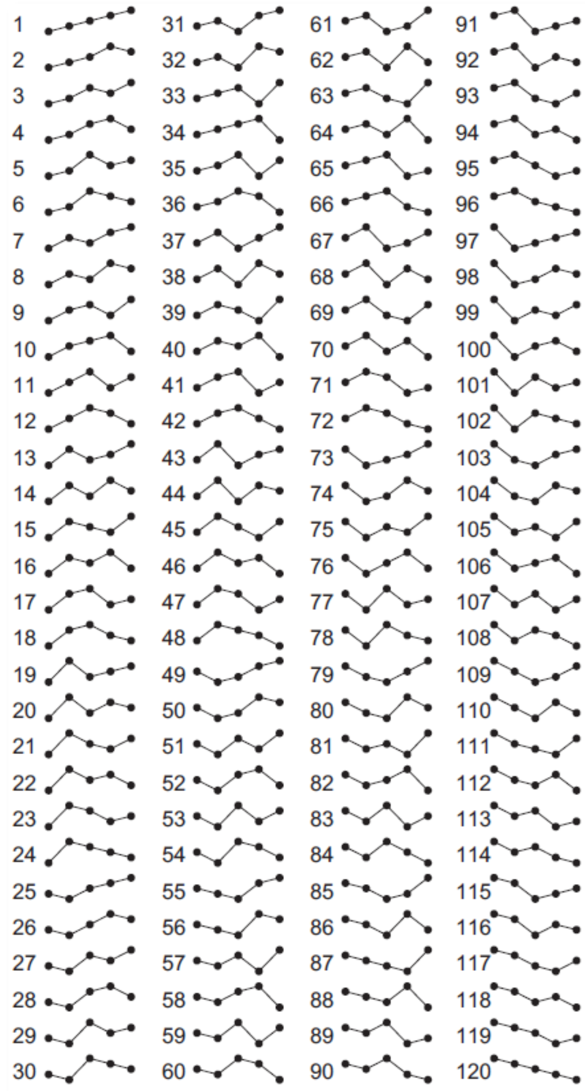
Ordinal bifurcation diagram



Pattern **6 (210)** is always forbidden; pattern **1 (012)** is more frequently expressed as r increases

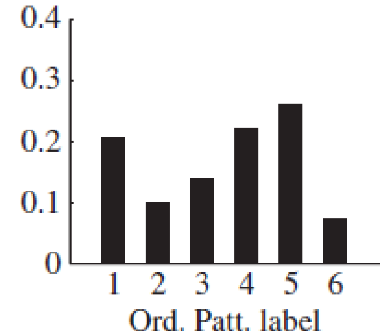
The number of patterns increases as D!





U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

Are the *D*! ordinal patterns equally probable?

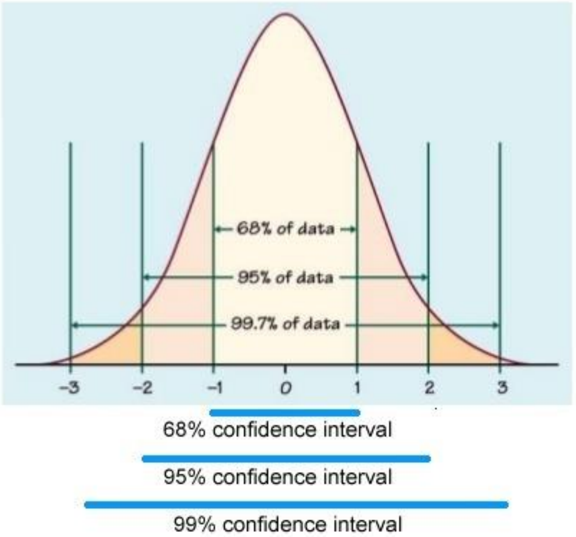


- **Null hypothesis**:

  $p_i = p = 1/D!$ for all $i = 1 \ldots D!$

- If at least one probability **is not** in the interval $p \pm 3\sigma$ with $\sigma = \sqrt{p(1-p)/N}$ and N the number of ordinal patterns:

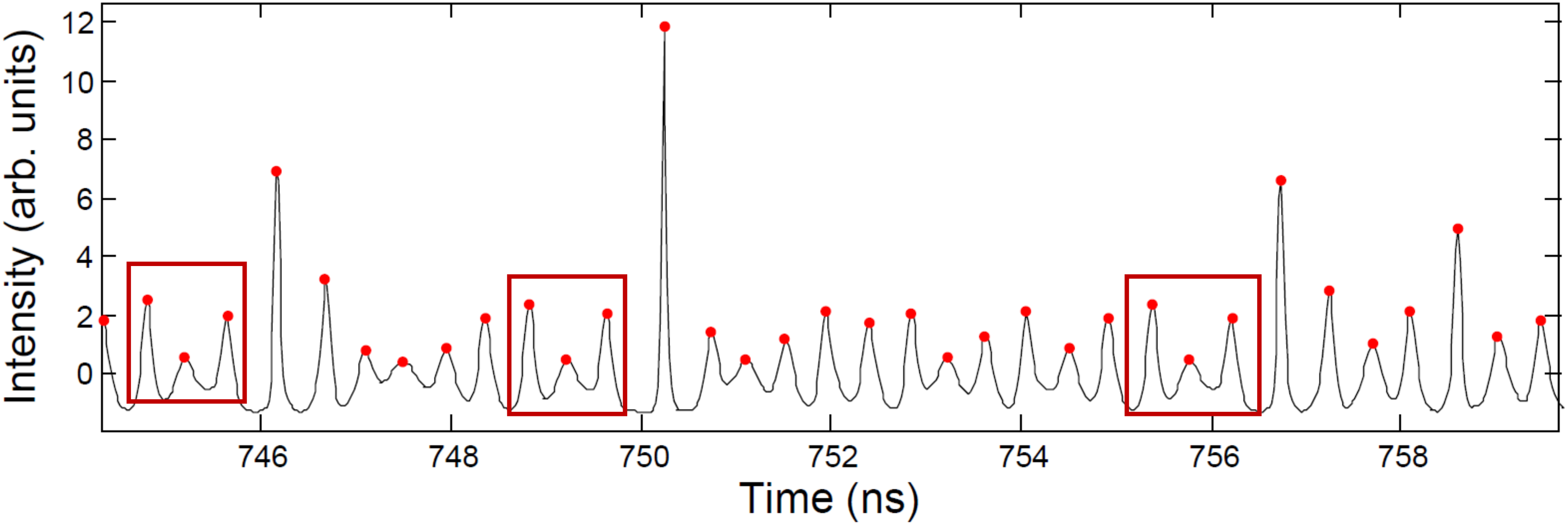
  We **reject** the NH with 99.74% confidence level.



- Else

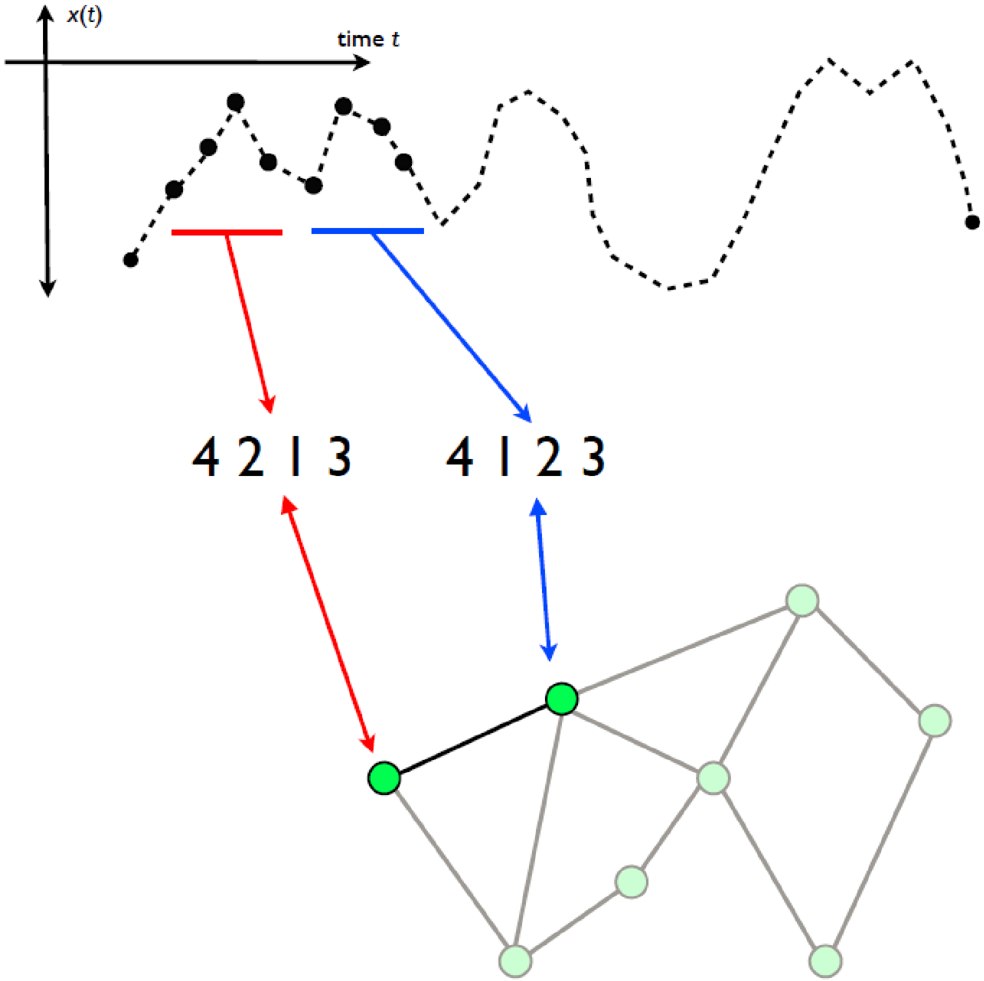
  We **fail to reject** the NH with 99.74% confidence level.

Example: intensity pulses emitted by a chaotic laser



N. Martinez Alvarez et al, Eur. Phys. J. Spec. Top. 226, 1971 (2017).

Ordinal analysis can be used to transform a time series into a graph, or network, that is weighted and directed



- D! nodes
- Weigh of node i: the probability of pattern i ($\sum_i p_i=1$)
- Weight of the link i→j: probability of transition i→j (for each *i*: $\sum_j w_{ij}=1$)

Adapted from M. Small
(The University of Western Australia)

Measures to characterize the graph

- Entropy computed from node weights (**permutation entropy**)

$$s_p = -\sum p_i \log p_i$$

- Average node entropy (entropy of the link weights)

$$s_n = \frac{1}{M}\sum_{i=1}^{M} s_i \qquad s_i = -\sum_{j=1}^{M} w_{ij} \log w_{ij}$$

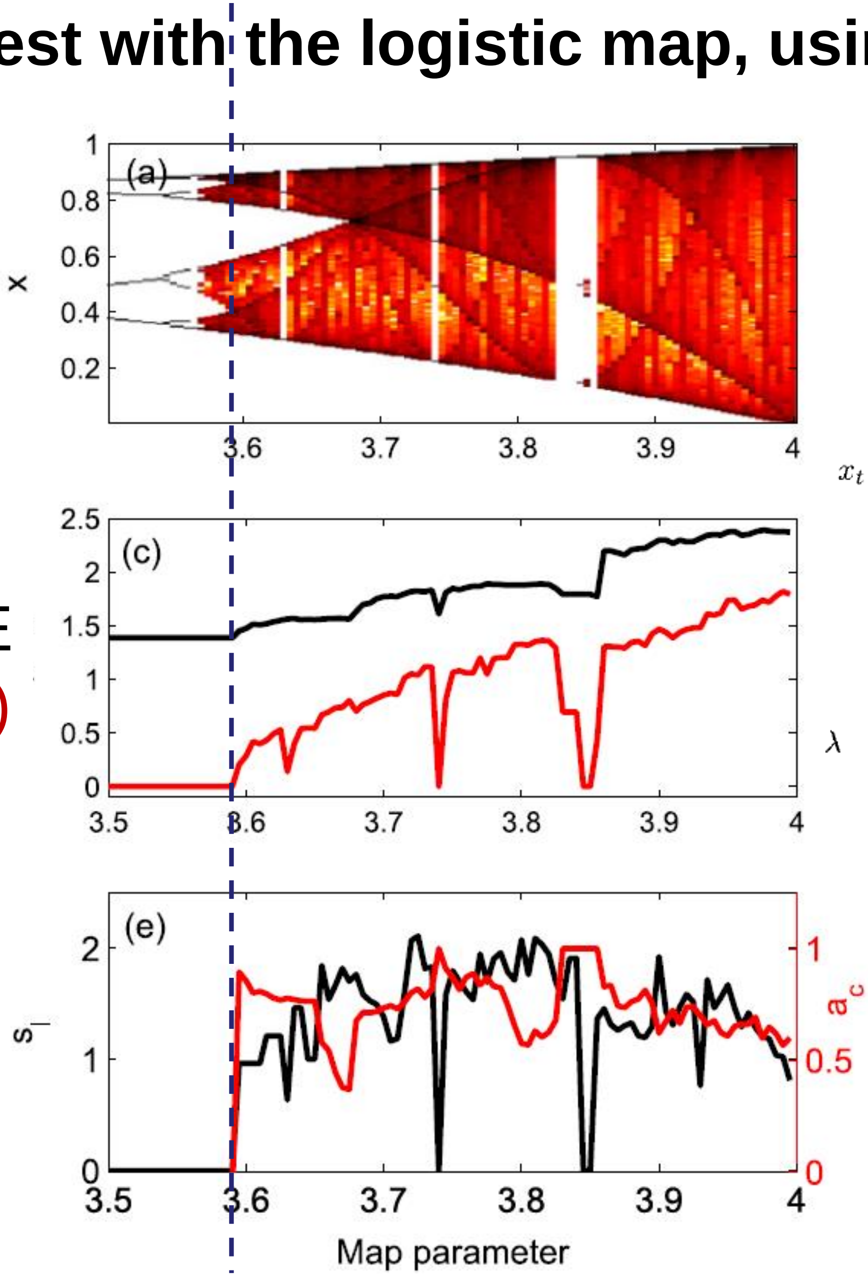
- Asymmetry coefficient: normalized difference of transition probabilities, P('01'→ '10') - P('10'→ '01'), etc.

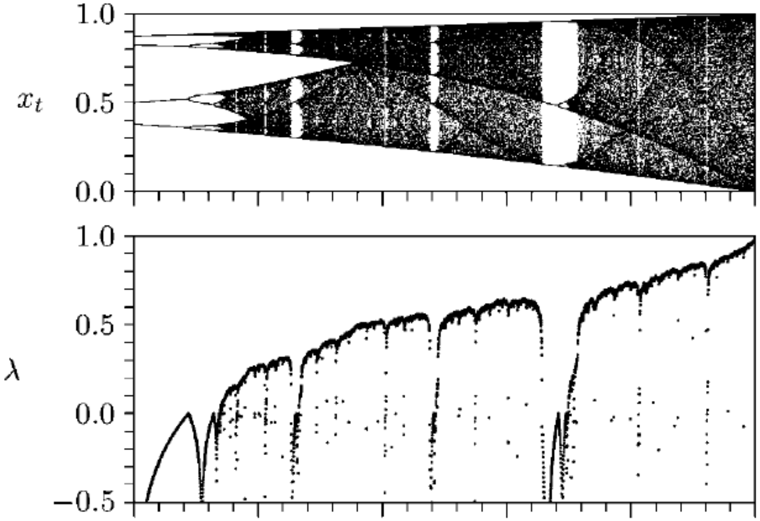
$$a_c = \frac{\sum_i \sum_{j \neq i} \left| w_{ij} - w_{ji} \right|}{\sum_i \sum_{j \neq i} \left(w_{ij} + w_{ji} \right)}$$

(0 in a fully symmetric network;
1 in a fully directed network)

A first test with the logistic map, using D=4 ordinal patterns

PE
S(TPs)



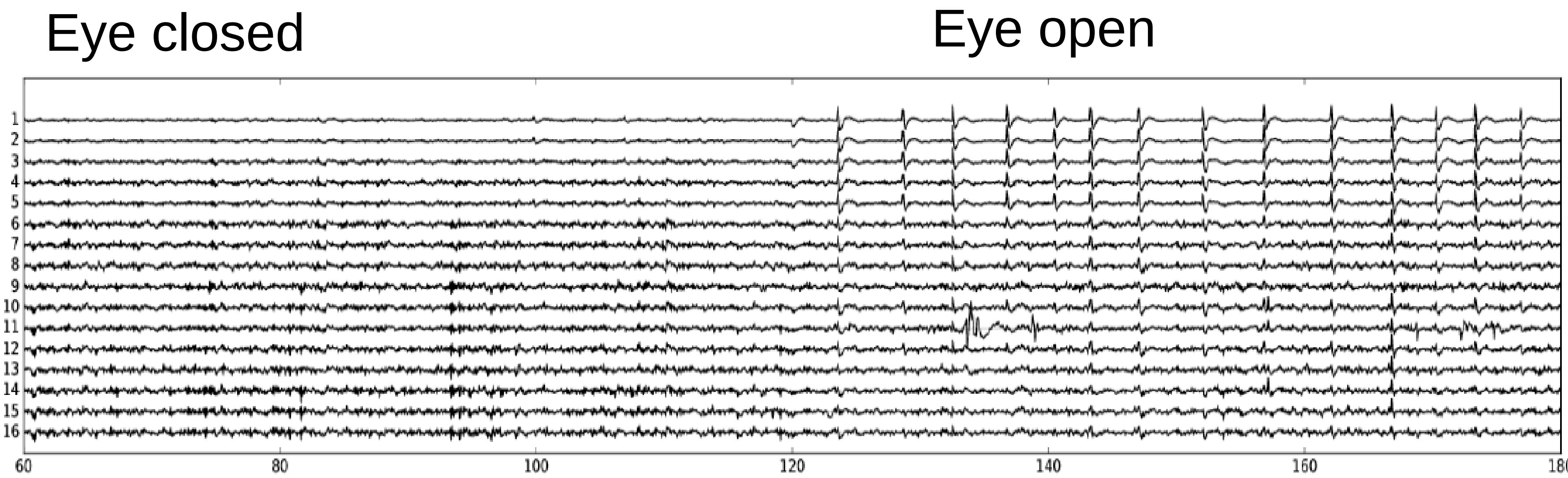


C. Masoller et al., New J. of Phys. 17, 023068 (2015)

First example of application on empirical data: distinguishing eyes closed (EC) and eyes open (EO) brain states from the analysis of EEG signals

BitBrain (company in Zaragoza) — PhysioNet (free)

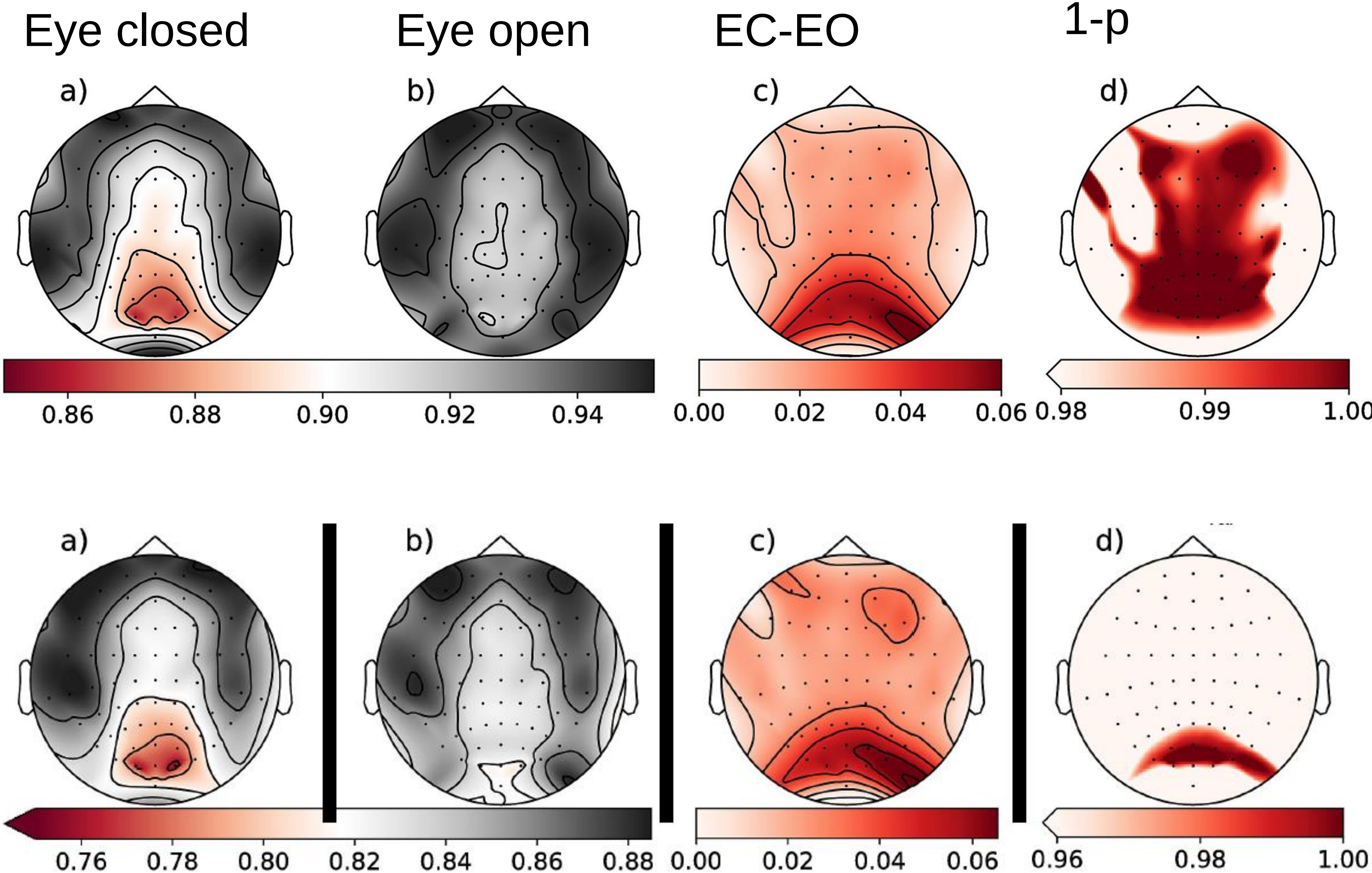
	DTS1	DTS2
Sampling rate(Hz)	256	160
Time task(seg)	120	60
Total points	30720	9600
Number of electrodes	16	64
Number of subjects	70	109



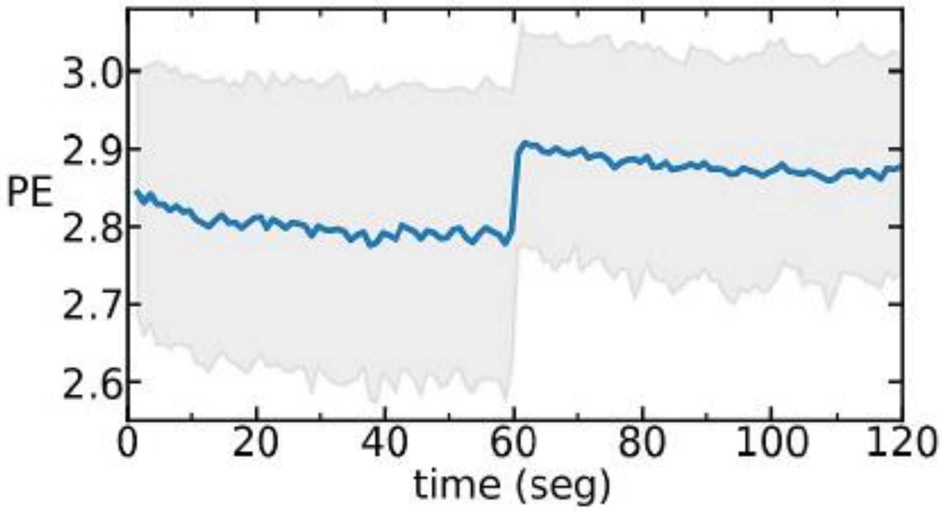
Ordinal analysis was applied to the raw data; similar results were found with filtered data using independent component analysis.

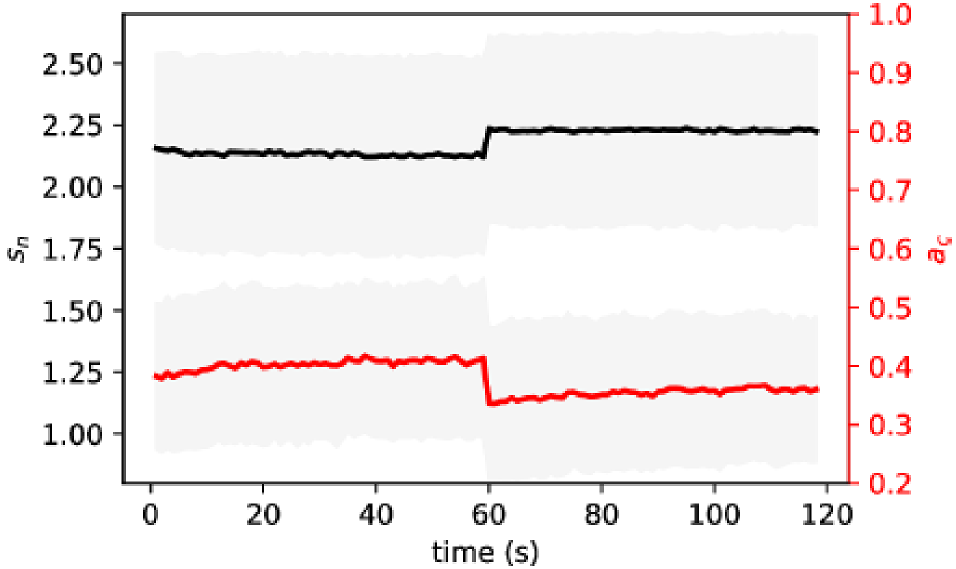
Permutation entropy (top) and node entropy (bottom)

PhysioNet dataset



"Randomization": the entropy increase and the asymmetry coefficient decreases when the person opens the eyes





Time window = 1 s
(160 data points)

C. Quintero-Quiroz et al. Chaos 28, 106307 (2018).

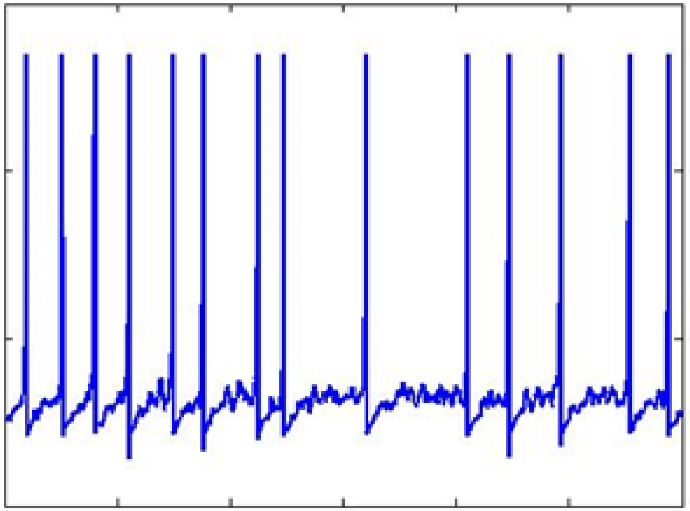
Second example of application: how similar these time series are?

Optical spikes

Time

μs or shorter

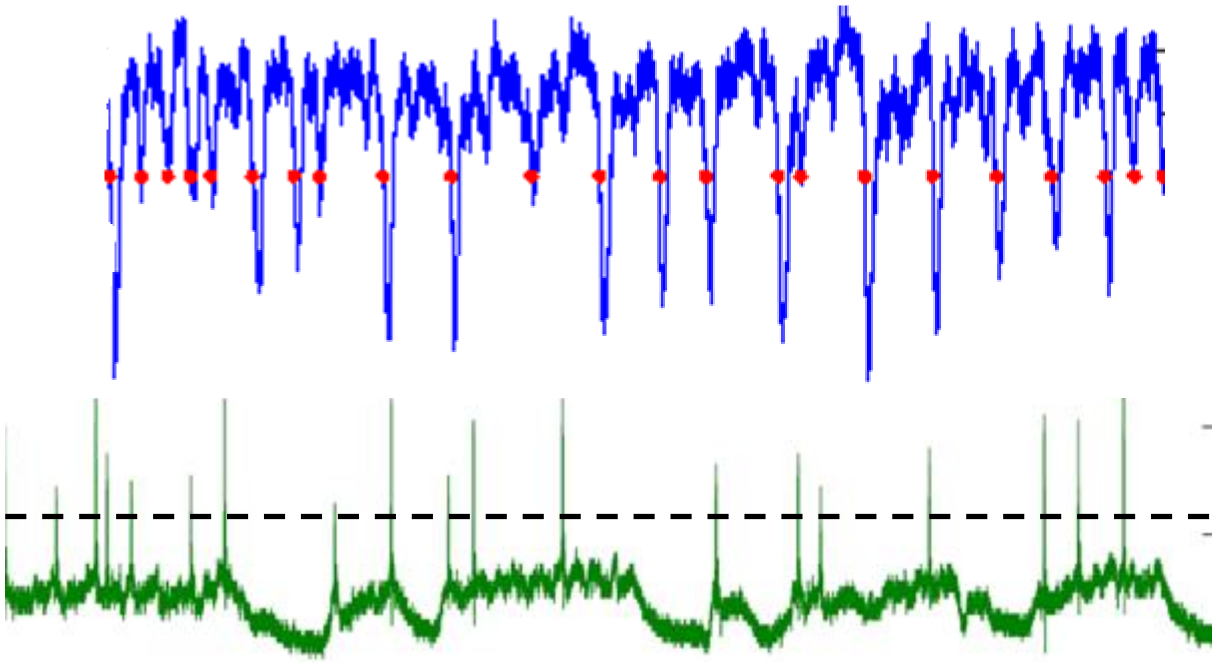
Neuronal spikes



Time

ms

Threshold crossings define ``events'' in a time series

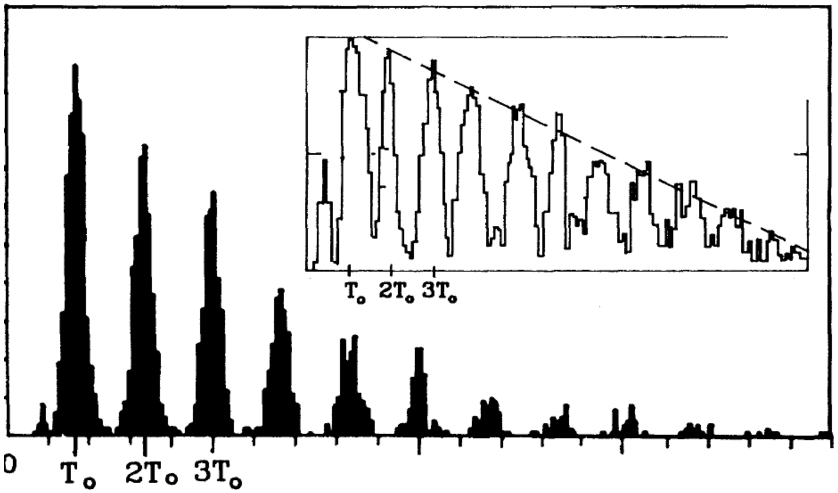


Analysis of sequence of inter-spike-intervals (ISIs):

$$\Delta T_i = t_{i+1} - t_i$$

ISI distribution indicates that neurons and lasers have a similar response to external periodic forcing

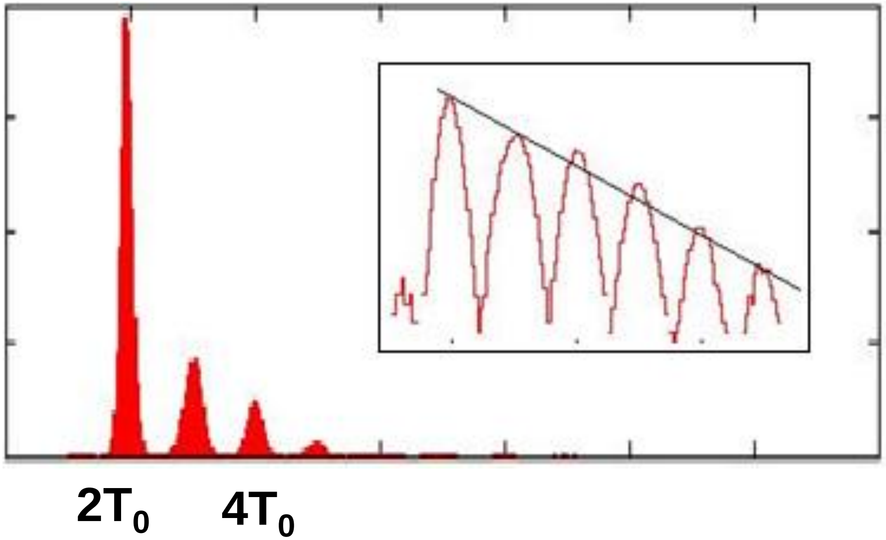
Neuron data



Single auditory nerve fiber of a squirrel monkey with a sinusoidal sound stimulus applied at the ear.

A. Longtin et al., PRL 67, 656 (1991)

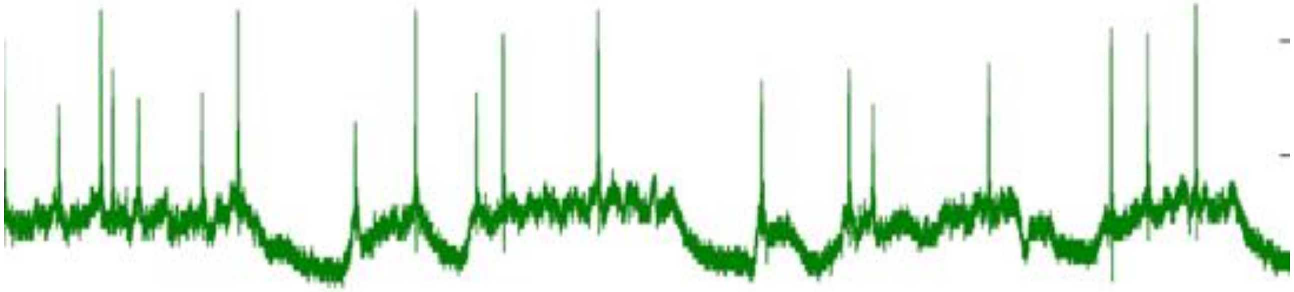
Laser data



Data recorded in our lab when a sinusoidal signal is applied to the laser current.

A. Aragoneses et al.,
Optics Express 22, 4705 (2014)

How neurons encode information?



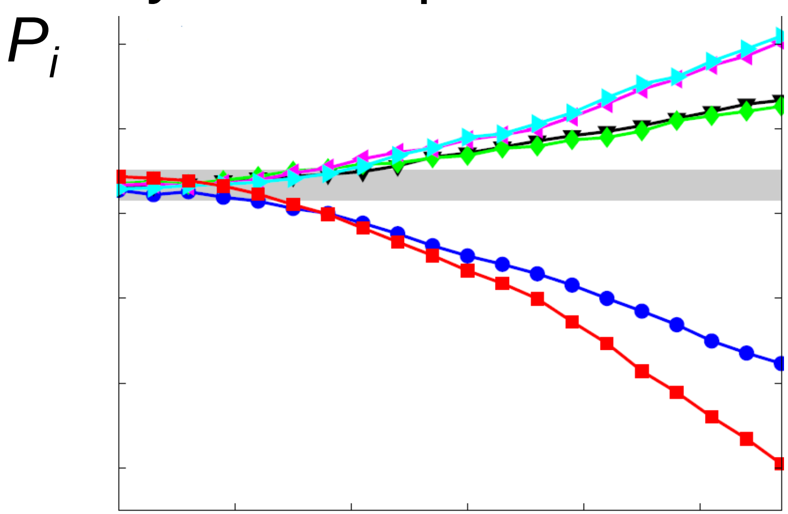
- In the spike rate?
- Is the **timing** of the spikes relevant?

Comparison of laser spikes and simulated neuronal spikes

FHN model with Gaussian white noise and **weak** sinusoidal input: spikes are noise-induced

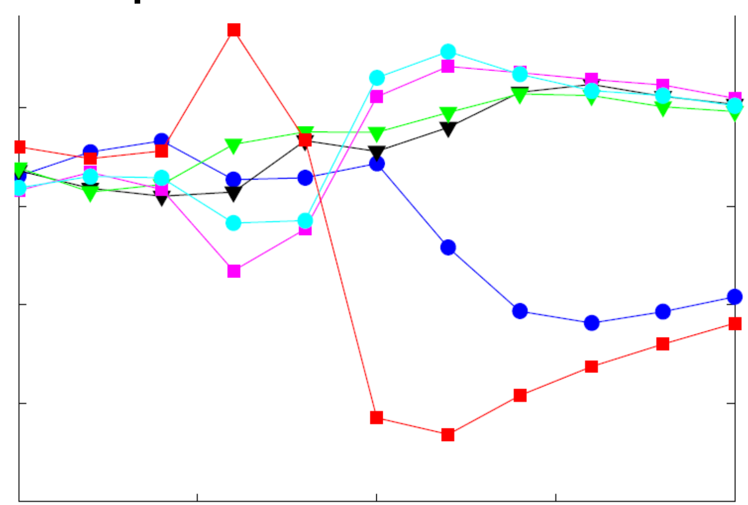
$$\epsilon \frac{dx}{dt} = x - \frac{x^3}{3} - y,$$
$$\frac{dy}{dt} = x + a + \boxed{a_o \cos(2\pi t/T)} + D\xi(t),$$

Synthetic spikes



Modulation amplitude

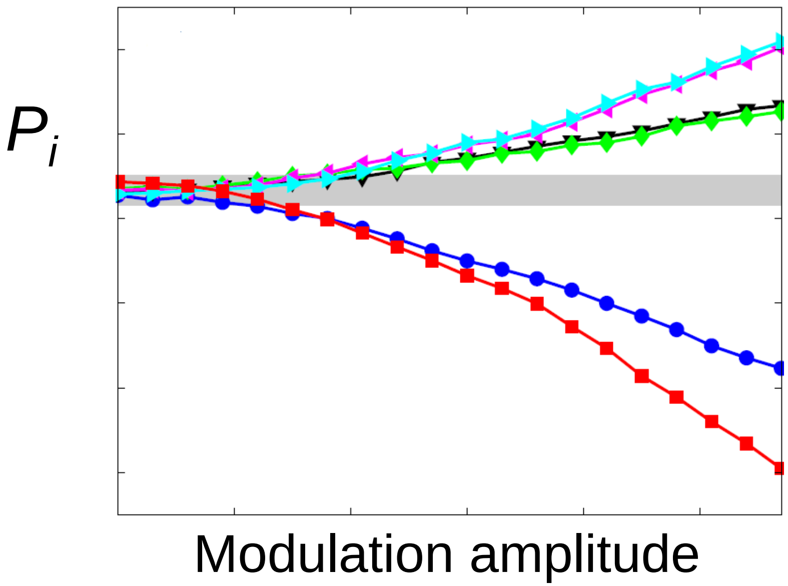
Empirical laser data

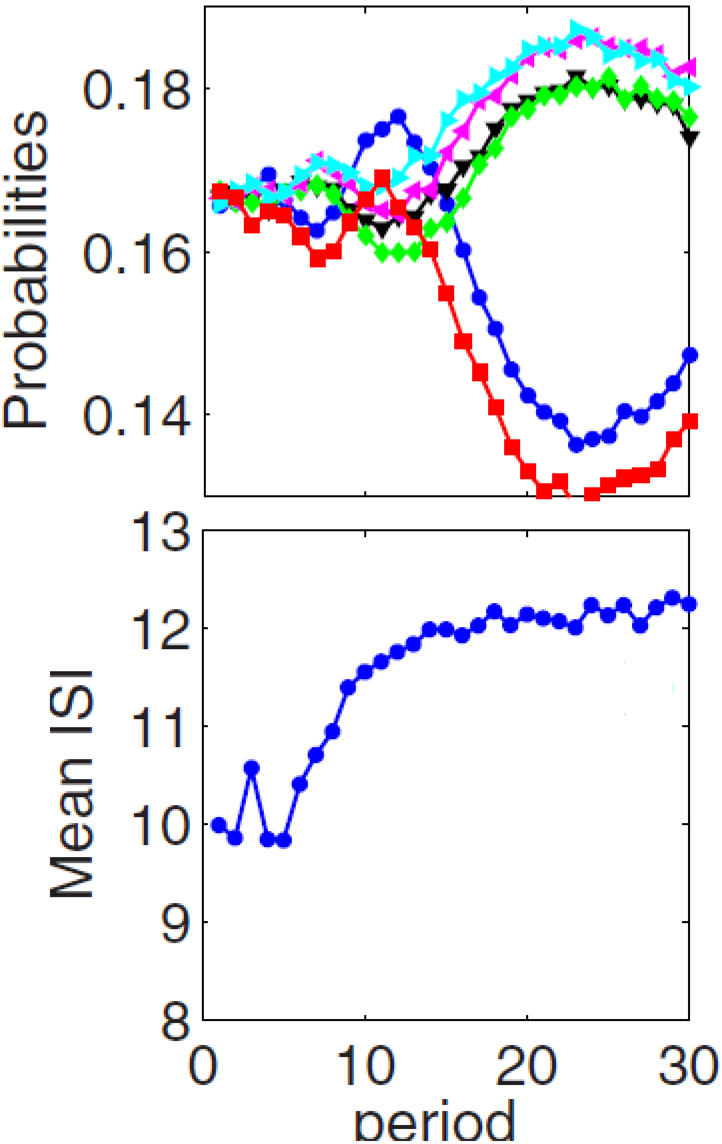


Modulation amplitude

J. A. Reinoso, M. C. Torrent, and C. Masoller, Phys. Rev. E. 94, 032218 (2016)
J. Tiana-Alsina, C. Quintero-Quiroz, C. Masoller, New J. of Phys 21, 103039 (2019)

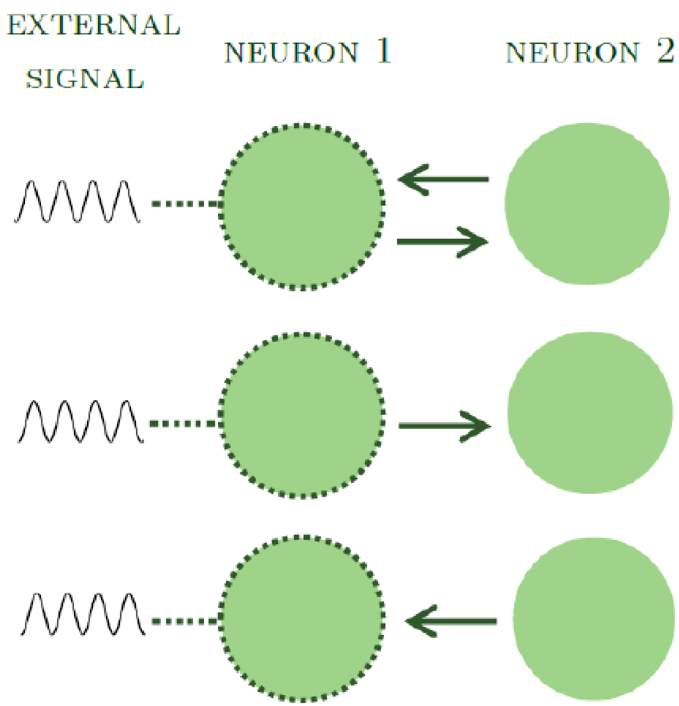
Understanding the neural code: how sensory neurons encode a weak signal in the presence of noise?

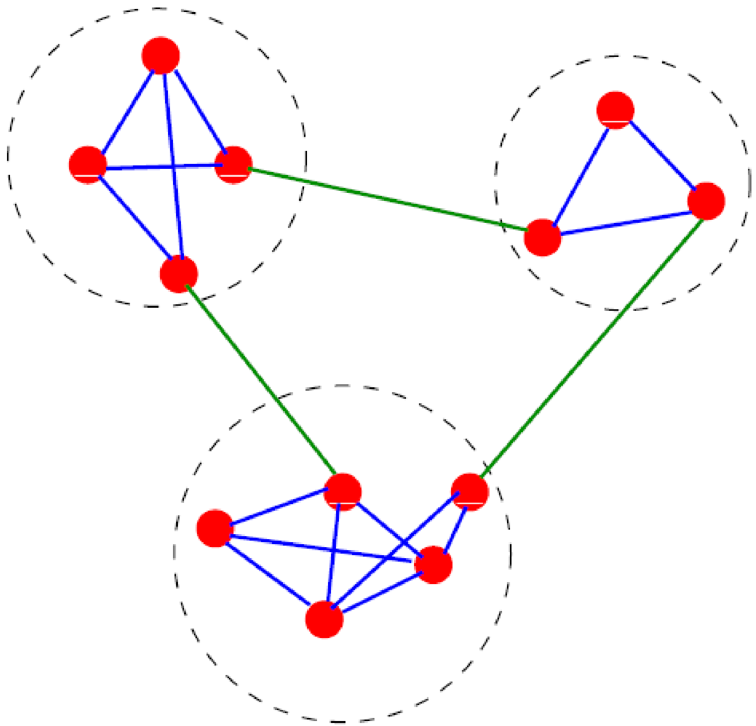




- The probabilities depend on the amplitude and on the period T of the signal.
- For large T, the mean ISI does not depend on T.

Is this encoding mechanism robust to neuronal coupling?

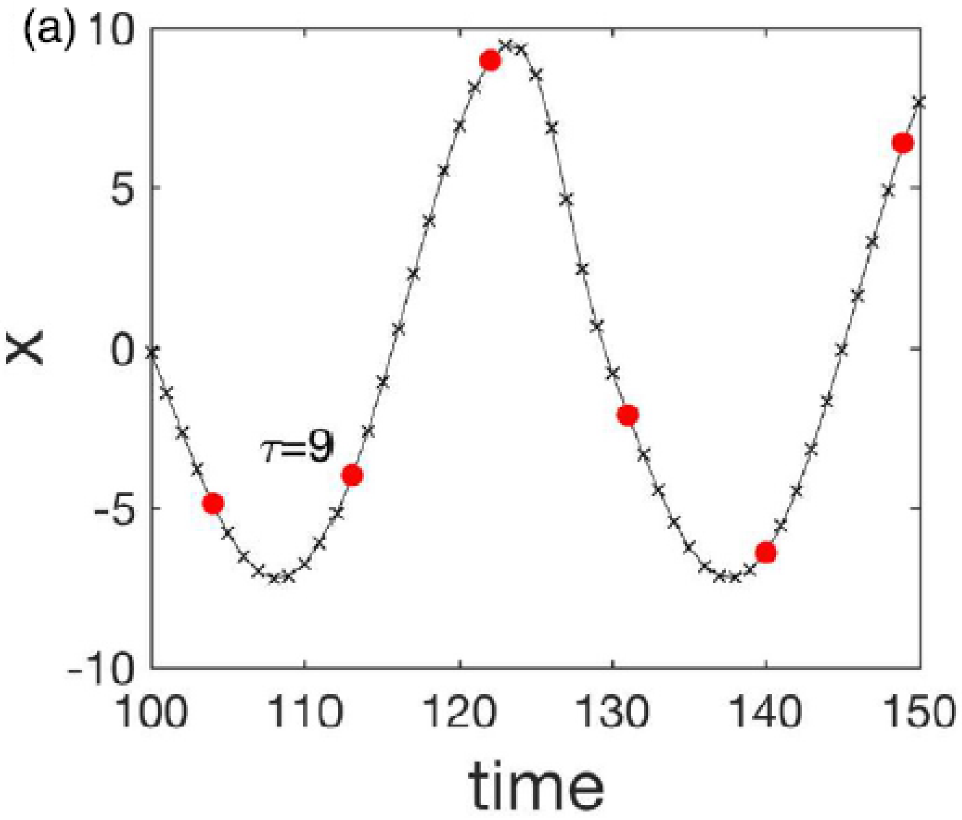




Yes!

M. Masoliver, C. Masoller, Sci. Rep. 8, 8276 (2018)
M. Masoliver, C. Masoller, Commun. Nonlinear Sci. Numer. Simulat. 88, 105023 (2020)
C. Estarellas, M. Masoliver, C. Masoller, C.R. Mirasso, Chaos 30, 013123 (2020)

Ordinal patterns can be defined using a lag between data points (varying the effective “sampling time”)



Y. Zou, R.V. Donner, N. Marwan et al. / Physics Reports 787 (2019) 1–97

By using lags, ordinal analysis allows to separate times scales of climatic interactions

Example. We calculate the **mutual information,** a nonlinear correlation between two time series, $x_i(t)$ and $x_j(t)$, computed from probability distributions, p_i, p_j, p_{ij}, extracted from $x_i(t)$ and $x_j(t)$

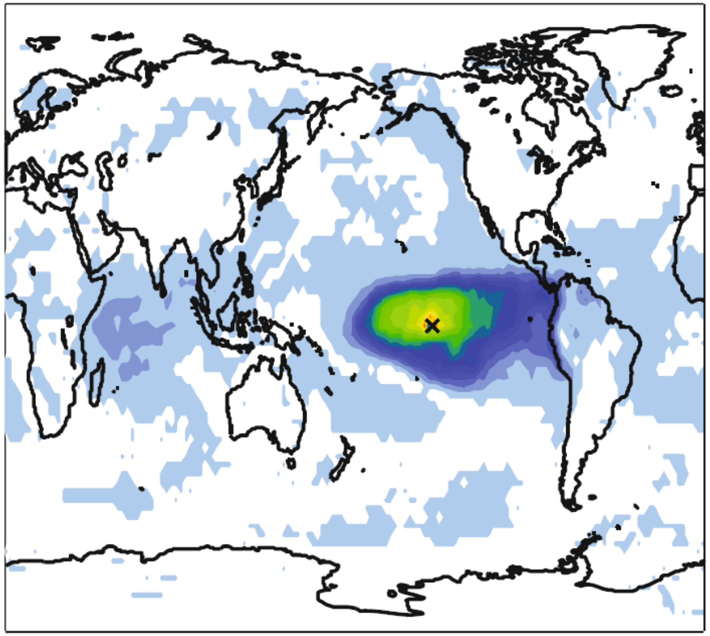
$$M_{ij} = \sum_{m,n} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)p_j(n)}$$

What are $x_i(t)$ and $x_j(t)$?

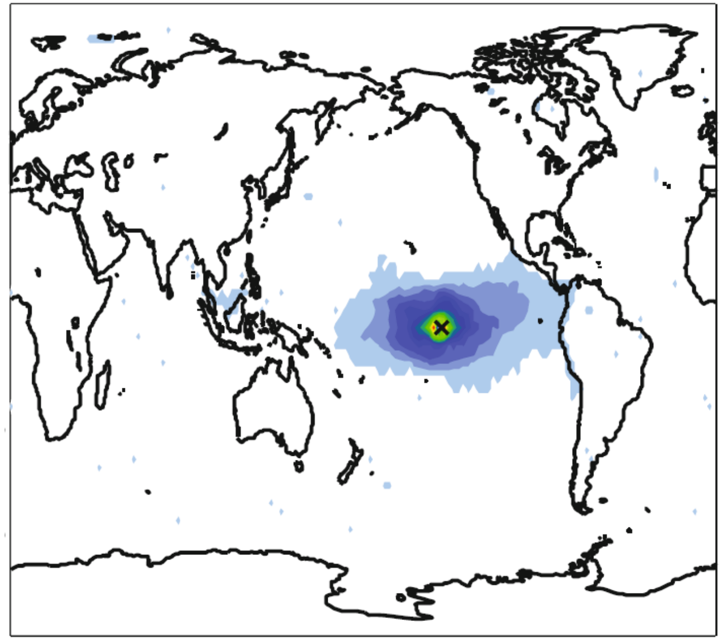
Climatic time series (surface air temperature anomalies) recorded at a reference point (in El Niño region), and at another geographical region.

What do we obtain?   Depends on how we “extract” p_i, p_j, p_{ij}.

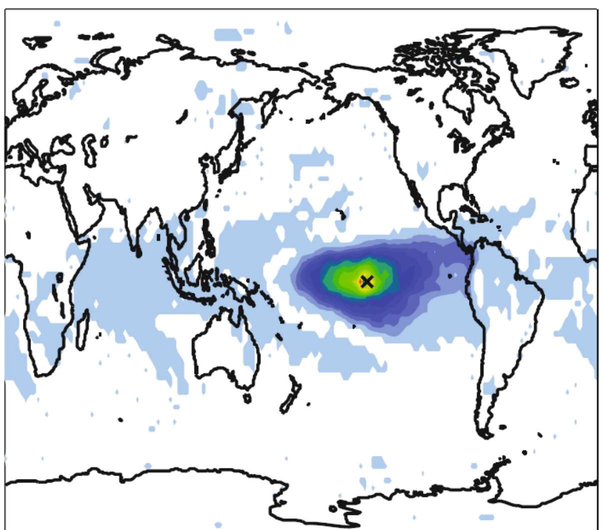
Histograms of values

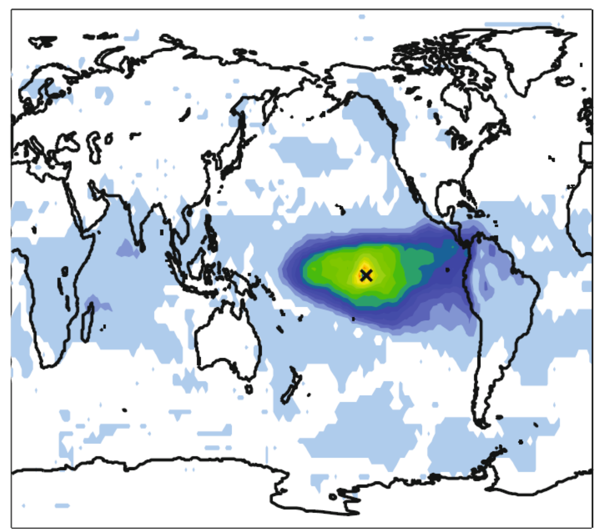


Lag=1 month ordinal patterns



Inter-annual ordinal patterns





Lag=12 months ordinal patterns

Deza, Barreiro and Masoller, Eur. Phys. J. ST 222, 511 (2013)

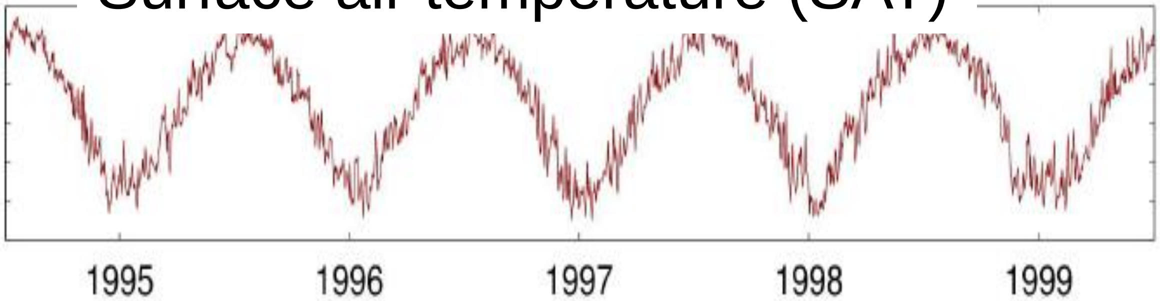
Second analysis tool: Hilbert analysis

It provides an instantaneous phase, amplitude and frequency for each data point of a scalar oscillatory time series

The Hilbert transform

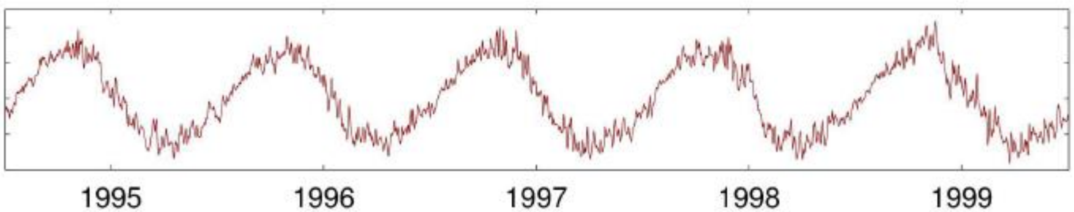
Surface air temperature (SAT)

x

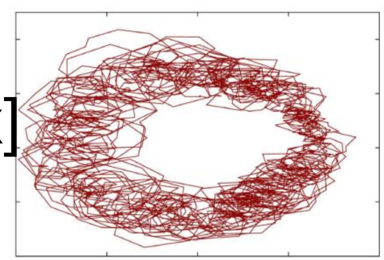


- HT[sin(ωt)]=cos(ωt)

HT[x]



y=HT[x]



x

$$a(t) = \sqrt{[x(t)]^2 + [y(t)]^2}$$

$$\varphi(t) = \arctan[y(t)/x(t)]$$

Can we use the Hilbert amplitude, phase, frequency, to identify and quantify regional climate change?

- A word of warning: only if $x(t)$ is a narrow-band signal $a(t)$ and $\omega(t) = d\varphi/dt$ have clear physical meaning
  - **$a(t)$** is the envelope of $x(t)$
  - **$\omega(t)$** is the main frequency in the Fourier spectrum
- Problem: climate time series are not narrow-band
- Usual solution (e.g. brain signals): isolate a narrow frequency band
- However, HT directly applied to surface air temperature uncovers the “hot spots” where changes in atmospheric dynamics are more pronounced.

The data: surface air temperature (SAT)

- Spatial resolution $2.5^0 \times 2.5^0 \Rightarrow$ 10226 time series
- Daily resolution 1979 – 2016 $\Rightarrow$ 13700 data points

Where does the data come from?

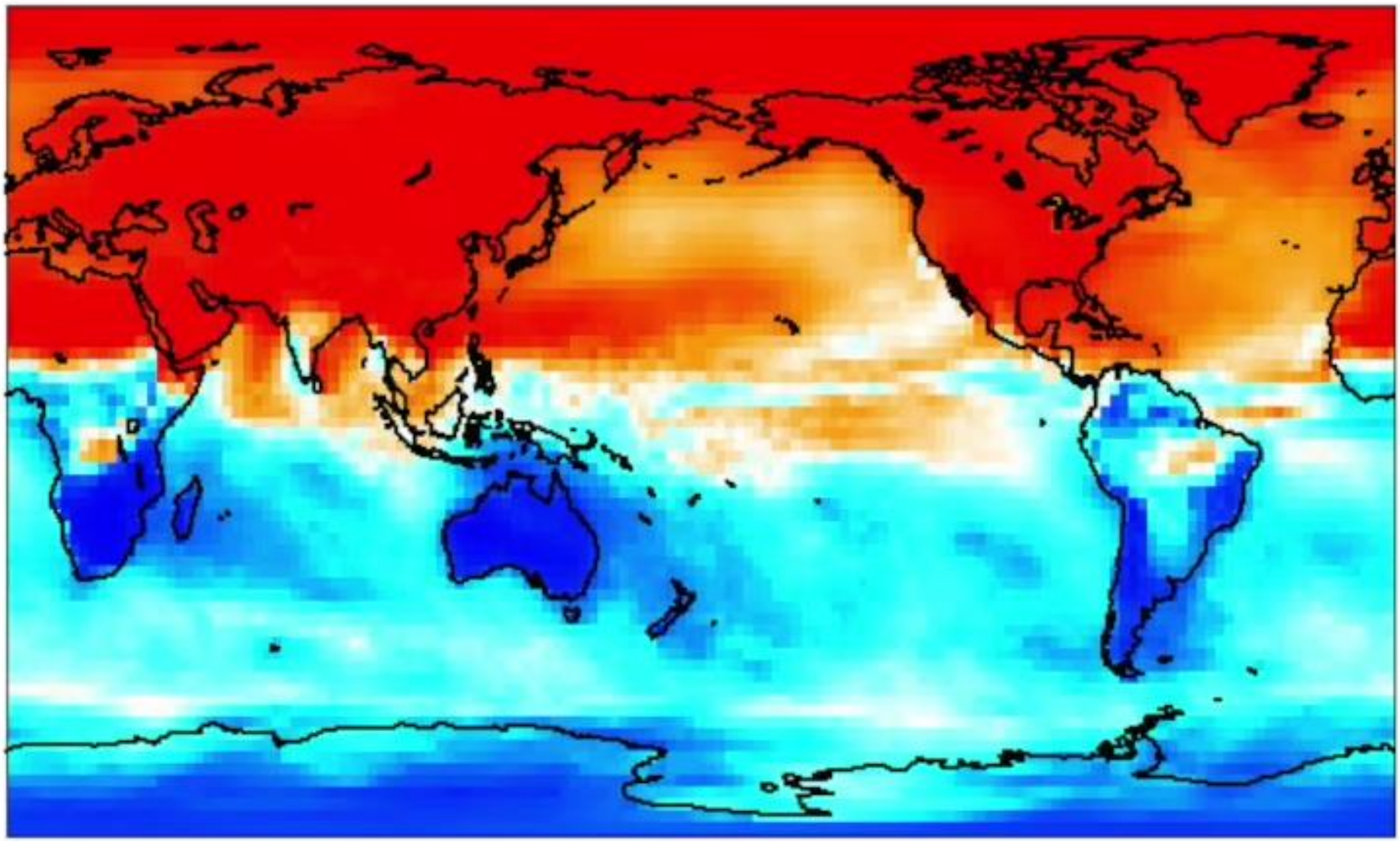
- European Centre for Medium-Range Weather Forecasts (ECMWF, ERA-Interim).
- Freely available.
- Reanalysis = general atmospheric circulation model feed with empirical data, where and when available (data assimilation).

Features extracted from each SAT time series

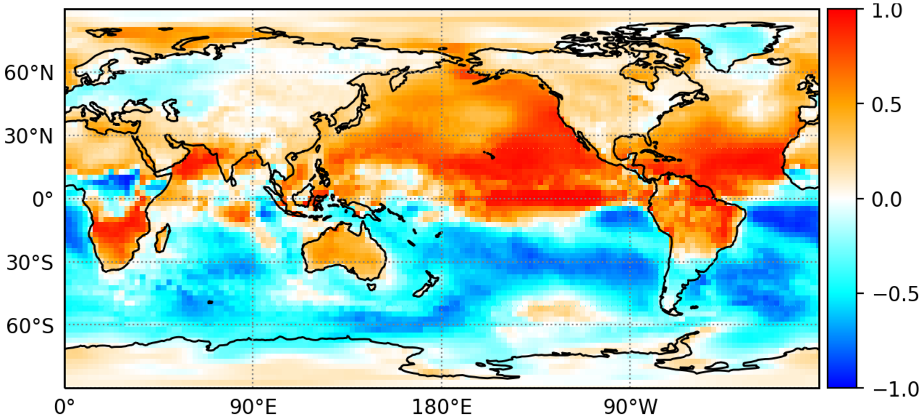
- Time averaged amplitude, $\langle a \rangle$
- Time averaged frequency, $\langle \omega \rangle$
- Standard deviations, σ_a, σ_ω

Which information carries the Hilbert phase? In color code the cosine of the Hilbert phase on an average typical year

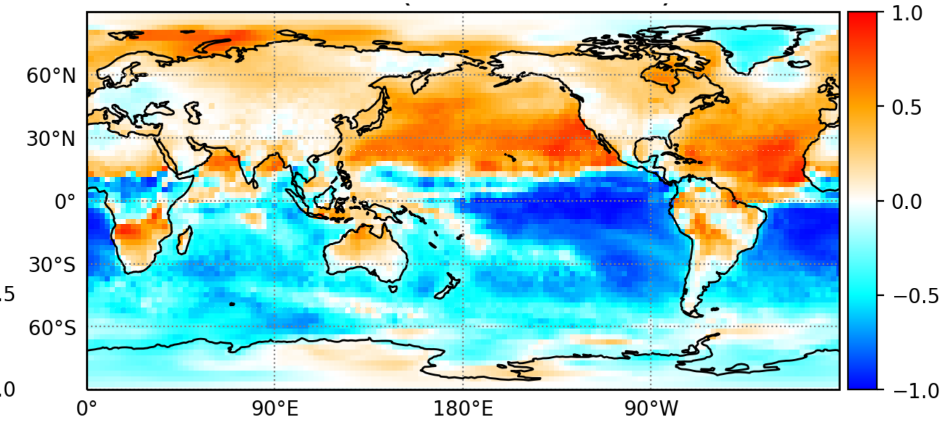
1 July



Cosine of Hilbert phase during a El Niño period (October 2015)

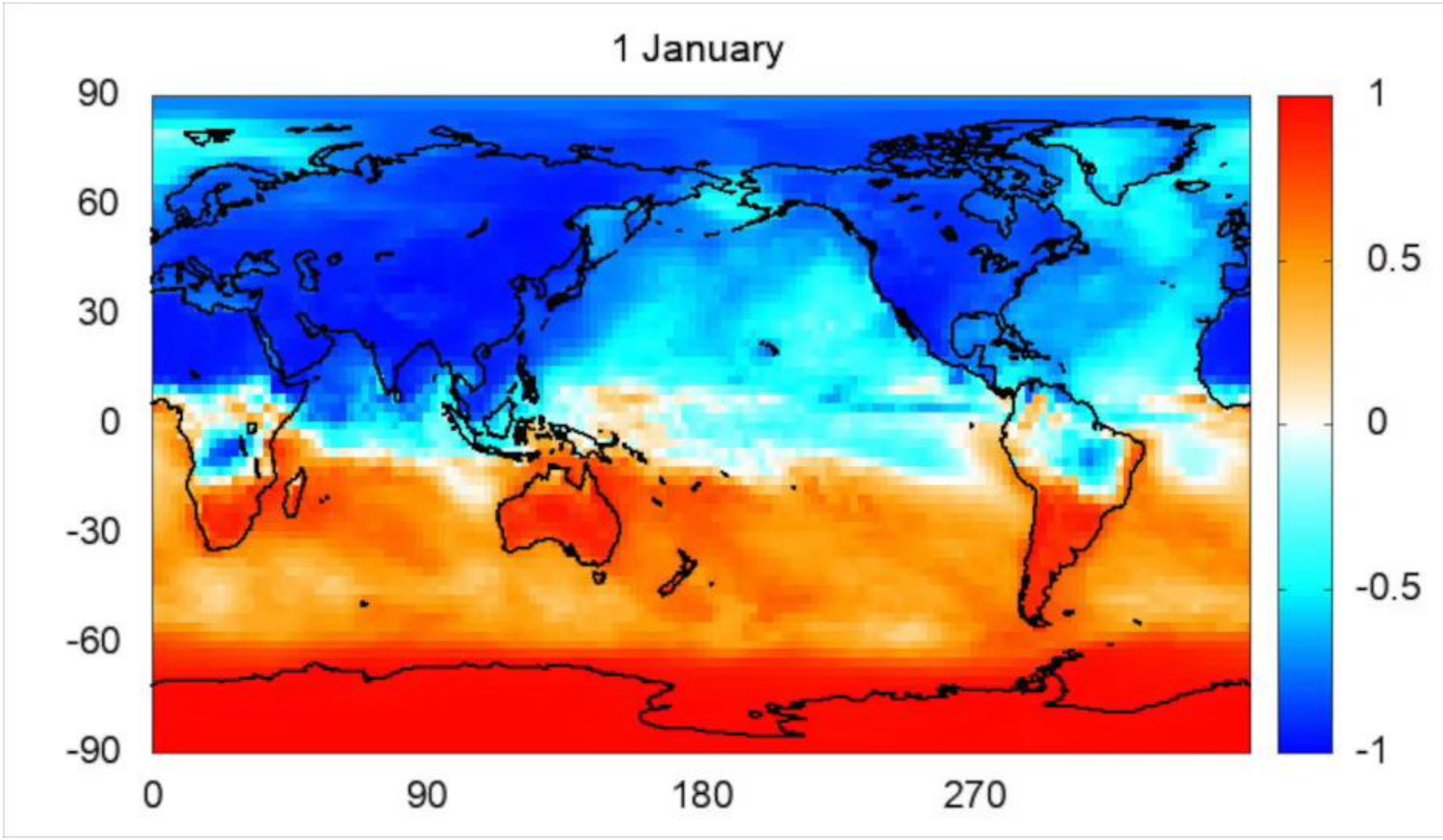


Cosine of Hilbert phase during a La Niña period (October 2011)

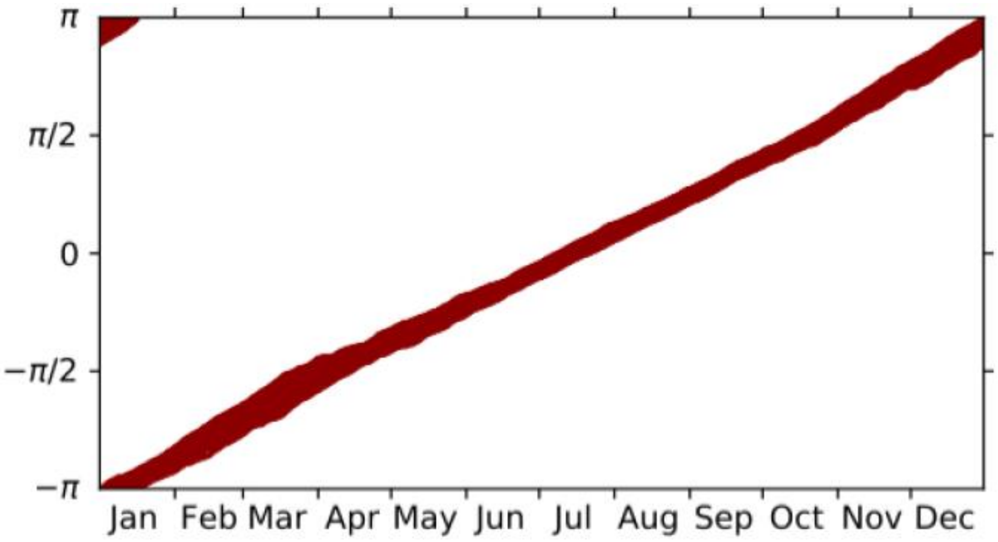


D. A. Zappala, M. Barreiro, C. Masoller, Earth Syst. Dynam. 9, 383–391 (2018).

How seasons evolve?
Temporal evolution of the cosine of the Hilbert phase

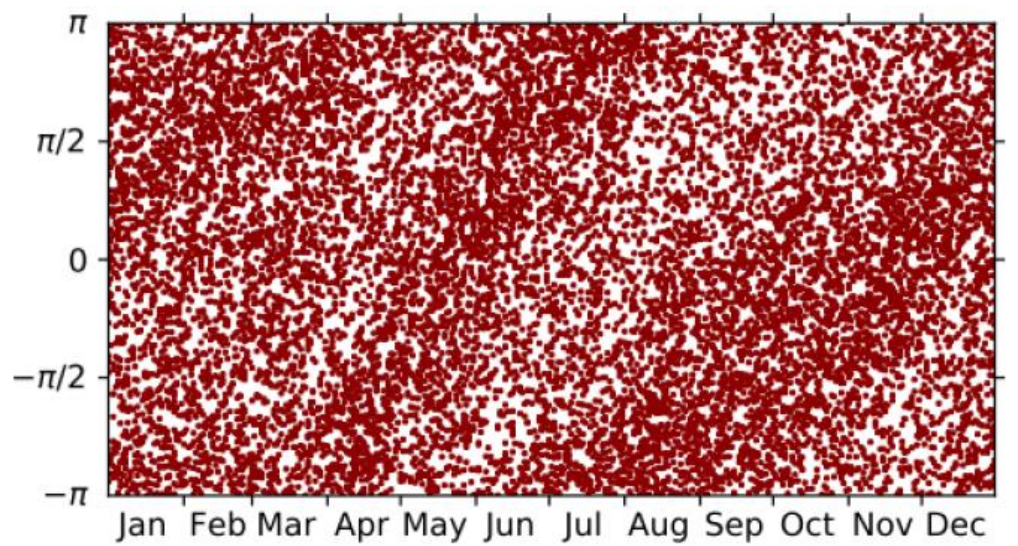


Hilbert phase vs day of the year relation



in a NH continental region

but in a tropical region



D. A. Zappala, M. Barreiro, C. Masoller, Chaos 29, 051101 (2019).

Relative decadal variations

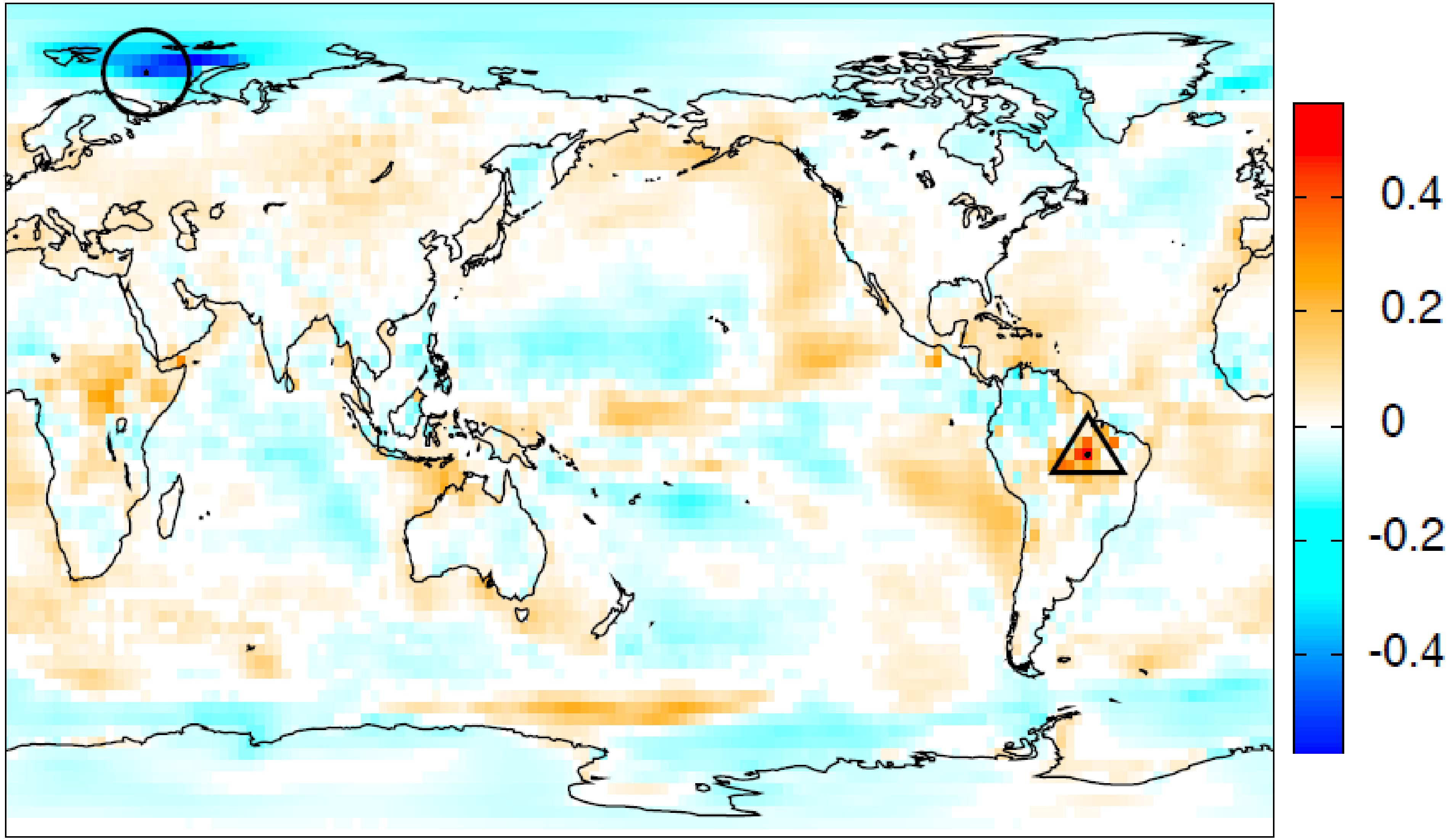
$$\Delta a = \langle a \rangle_{2016-2007} - \langle a \rangle_{1988-1979}$$

$$\frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

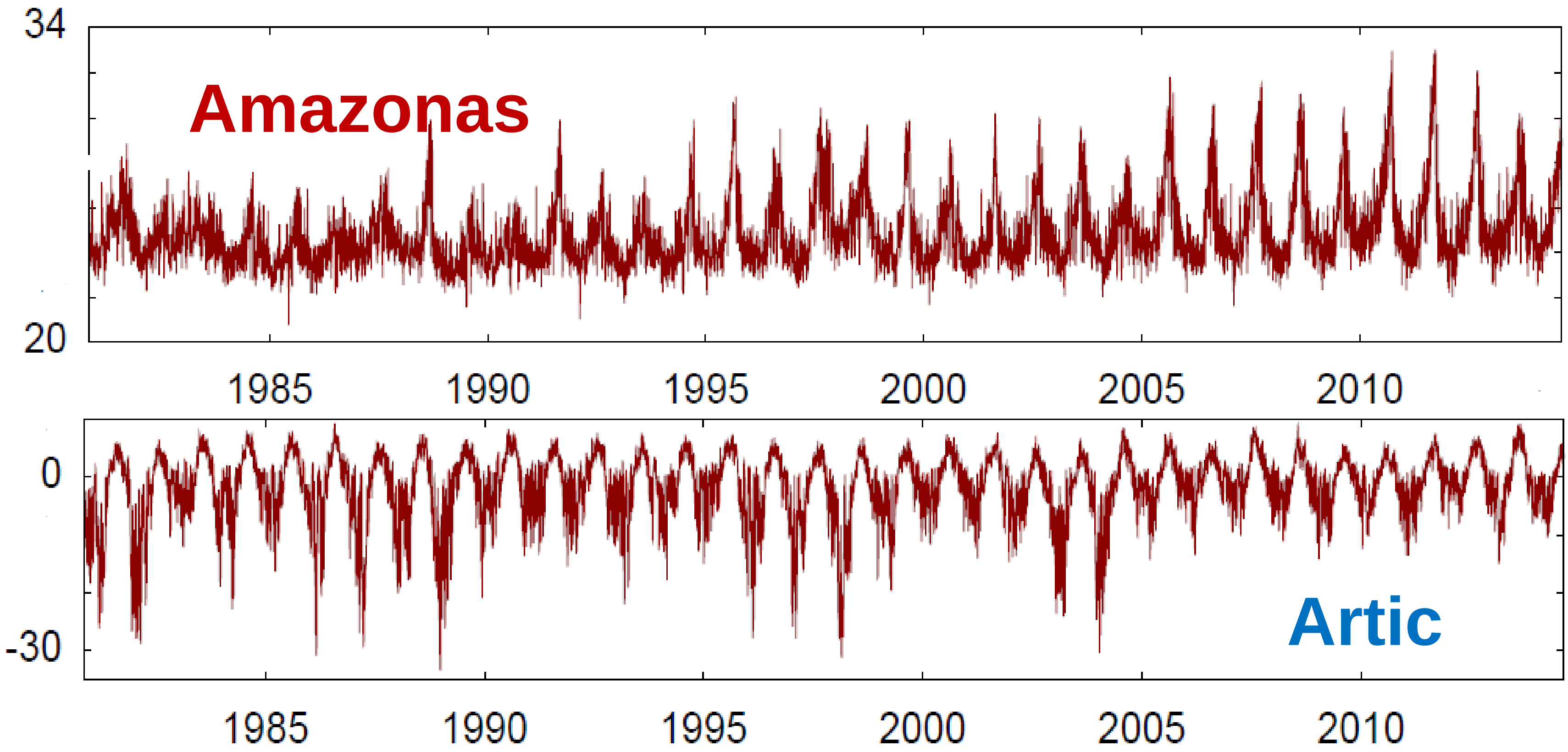
Relative variation is considered significant if:

$$\frac{\Delta a}{\langle a \rangle} \geq \langle . \rangle_s + 2\sigma_s \quad \text{or} \quad \frac{\Delta a}{\langle a \rangle} \leq \langle . \rangle_s - 2\sigma_s$$

100 surrogates



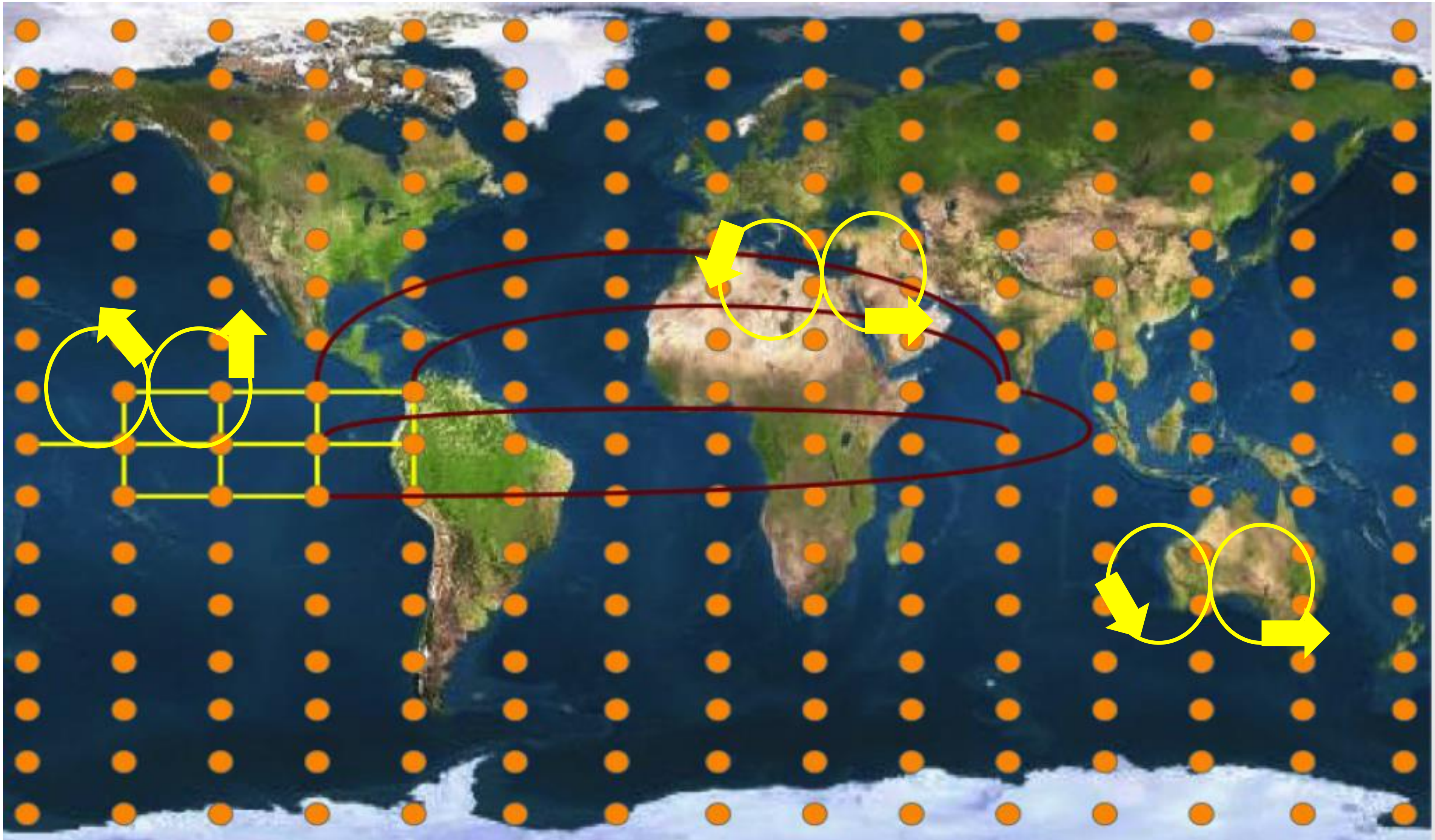
0.4
0.2
0
-0.2
-0.4



- **Decrease of precipitation**: the solar radiation that is not used for evaporation is used to heat the ground.

- **Melting of sea ice**: during winter the air temperature is mitigated by the sea and tends to be more moderated.

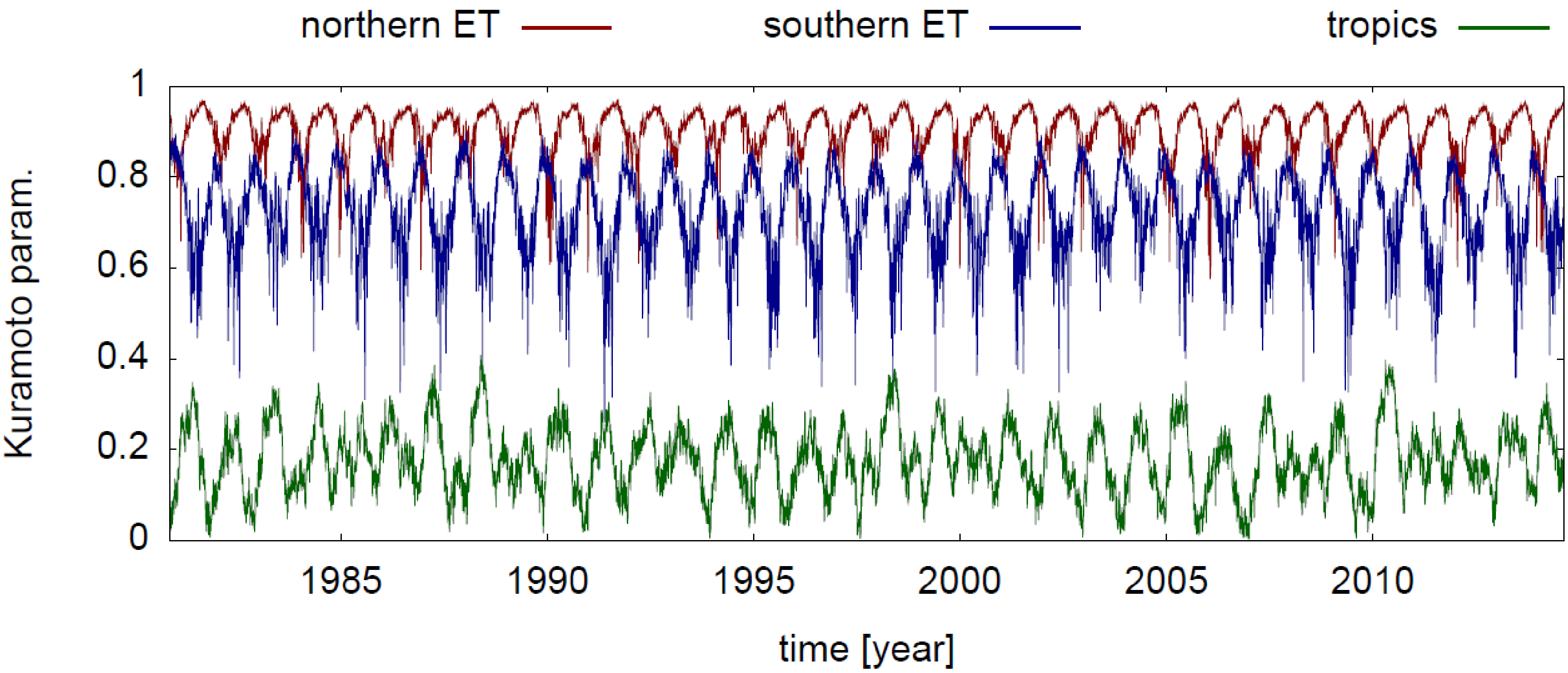
How to quantify the synchronization of climatic oscillations?



Quantifying synchronization in atmospheric data

Kuramoto order parameter

$$r(t) = \left| \frac{1}{N} \sum_{j=1}^{N} e^{i\theta_j(t)} \right|$$



Summarizing

Take home messages

- Ordinal analysis and Hilbert analysis are useful tools to extract information of complex signals.

- They provide *complementary* information to that provided by other linear or nonlinear methods.

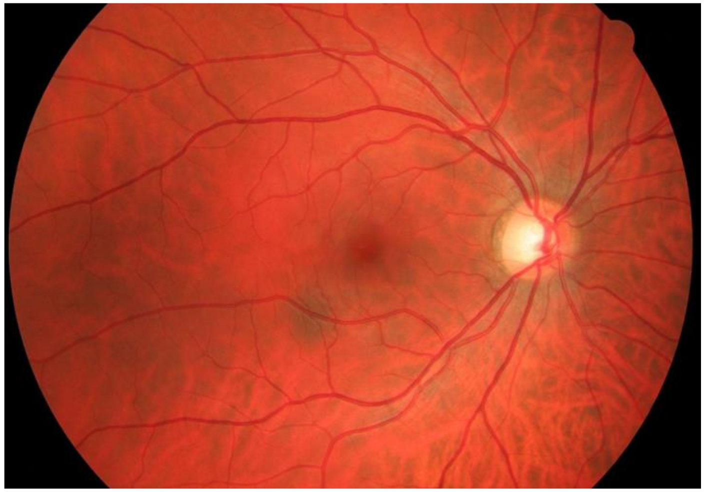
- Ordinal analysis was used to
  - Identify transitions in EEG data (eyes closed – eyes open)
  - Identify similarities in laser and neuronal dynamics
  - Uncover a plausible "neural coding" mechanism

- Hilbert analysis was used to
  - To identify climate changes in the last three decades
  - To quantify the synchronization level

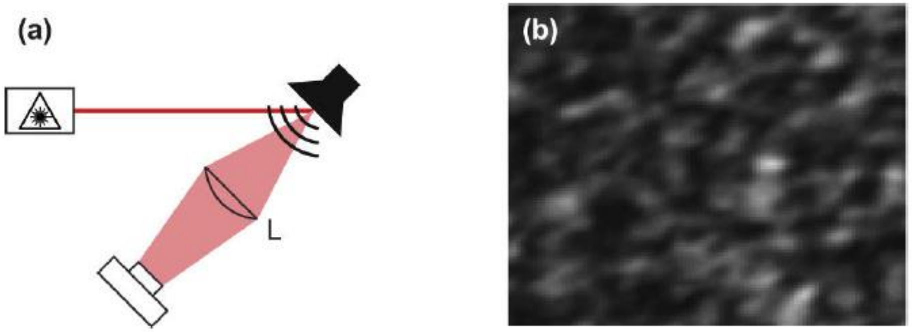
- Both analysis tools were applied directly to the raw data.

Other research lines, ongoing research

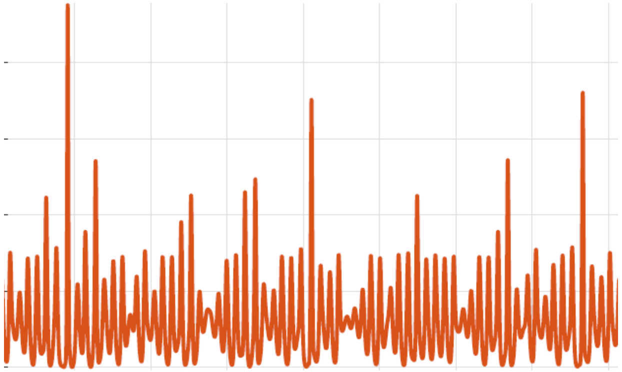
- Machine learning analysis of complex images (retina, speckle)
- Machine learning time series prediction (optical pulses, climatic time series)



K. Garside et al., PLoS ONE 14, e0217413 (2019).
P. Amil, et al., PLoS ONE 14, e0220132 (2019).



C. Barcellona et al., Opt. Express 28, 8716 (2020)



P. Amil et al., Chaos 29, 113111 (2019)

Thank you for your attention

<cristina.masoller@upc.edu>

- C. Masoller et al., New J. Phys. 17, 023068 (2015)
- J. A. Reinoso, M. C. Torrent, C. Masoller, Phys. Rev. E. 94, 032218 (2016)
- D. A. Zappala, M. Barreiro, C. Masoller, Earth Syst. Dynamics 9, 383 (2018)
- C. Quintero et al., Chaos 28, 106307 (2018)
- D. A. Zappala, M. Barreiro, C. Masoller, Chaos 29, 051101 (2019)
- J. Tiana-Alsina, C. Quintero, C. Masoller, New J. of Phys 21, 103039 (2019)
- M. Masoliver, C. Masoller, Commun. Nonlinear Sci. Numer. Simulat. 88, 105023 (2020)

Agència de Gestió d'Ajuts Universitaris i de Recerca

ICREA

GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN