# Investigating large-scale atmospheric phenomena using complex networks and nonlinear time series analysis tools

Cristina Masoller
Departamento de Fisica
Universitat Politecnica de Catalunya

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Campus d'Excel·lència Internacional

*4th Nonlinear Processes in Oceanic and Atmospheric Flows*
January 22, 2025, Institut de Ciències del Mar, Barcelona

cristina.masoller@upc.edu @cristinamasoll1

# THANKS TO

Ignacio Deza Giulio Tirabassi Dario Zappala Riccardo Silini Marcelo Barreiro

UNIVERSIDAD DE LA REPÚBLICA
URUGUAY

"LINC"
Learning about Interacting Networks in Climate

CAFE
Climate Advanced Forecasting of sub-seasonal Extremes

Agència de Gestió d'Ajuts Universitaris i de Recerca

ICREA

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

cristina.masoller@upc.edu @cristinamasoll1

# Outline

- Hilbert time series analysis to identify variations in Surface Air Temperature (SAT) reanalisis data in the last 30 years.

  $\{x_1, x_2, \ldots x_N\}$

- Bivariate analysis: mutual information of SAT anomalies and causal networks of climatic indices (ENSO, NAO, etc.)

  $\{x_1, x_2, \ldots x_N\}$ $\{y_1, y_2, \ldots y_N\}$

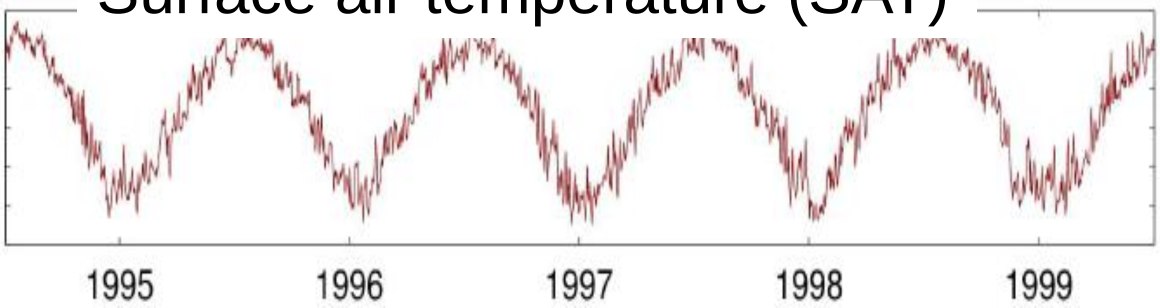
- Climate networks constructed from SAT anomalies.

cristina.masoller@upc.edu @cristinamasoll1

# The Hilbert Transform (HT)

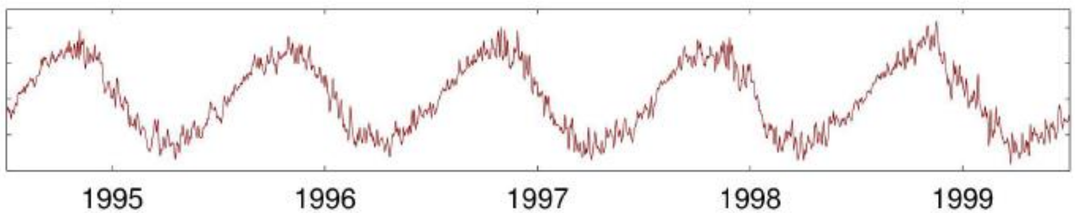
Surface air temperature (SAT)

x

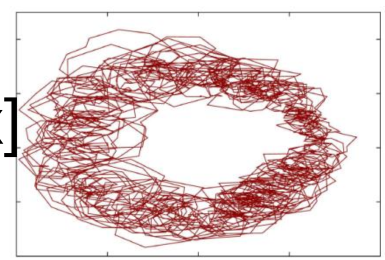



HT[sin(ωt)]=cos(ωt)

HT[x]




y=HT[x]



x

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

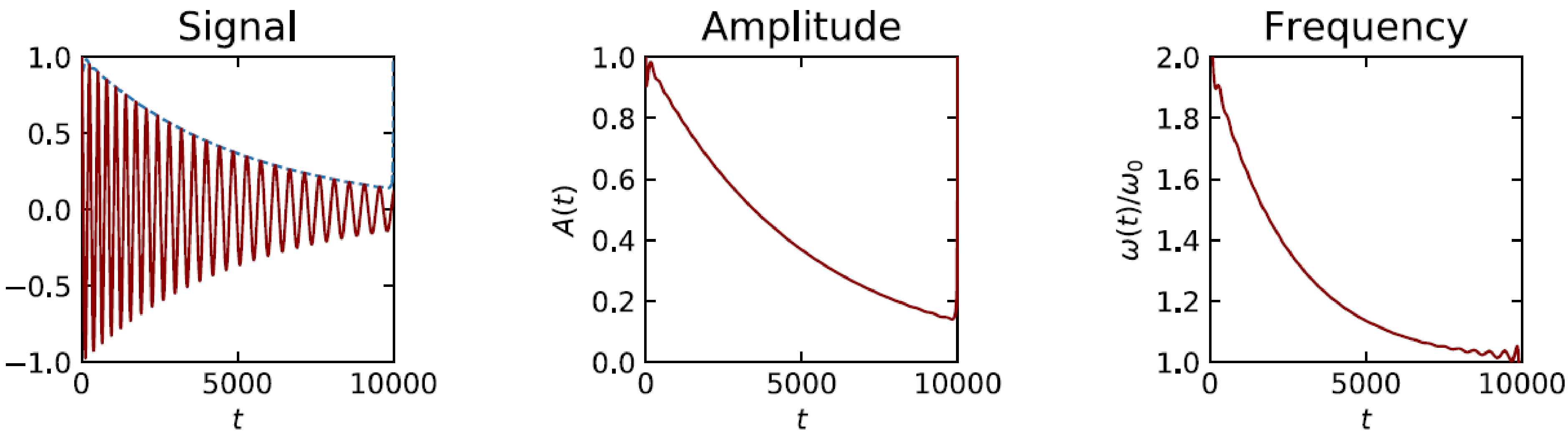
$$\varphi(t) = \arctan[y(t)/x(t)]$$

$\omega(t) = d\varphi/dt$

cristina.masoller@upc.edu @cristinamasoll1

# Example

$$x(t) = \mathrm{e}^{-\alpha t} \cos\left[\left(1 + \mathrm{e}^{-2\alpha t}\right)\omega_0 t\right].$$




<u>A word of warning</u>: only if *x(t)* is a “narrow-band” signal then *a(t)* and *ω(t)* = *dφ/dt* have clear physical meaning

– ***a(t)*** is the envelope of *x(t)*

– ***ω(t)*** is the main frequency in the Fourier spectrum

PROBLEM! Climate time series are NOT narrow-band.



cristina.masoller@upc.edu @cristinamasoll1

## Solution ?

- Isolate a narrow frequency band (usual for EEG analysis).
- However, I will show that HT directly applied to raw surface air temperature (SAT) returns meaningful results.

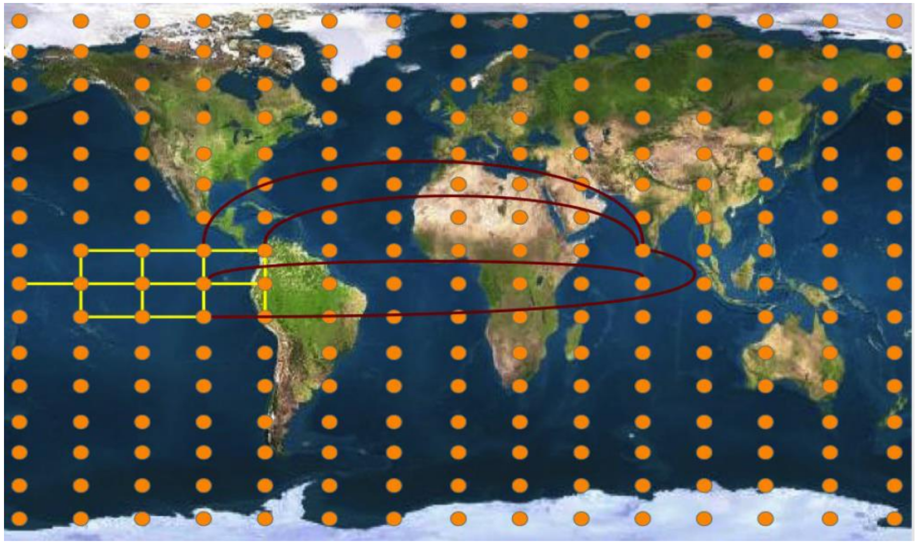
## Data

Surface Air Temperature (SAT) ERA-Interim data

- Spatial grid $2.5^0 \times 2.5^0 \Rightarrow$

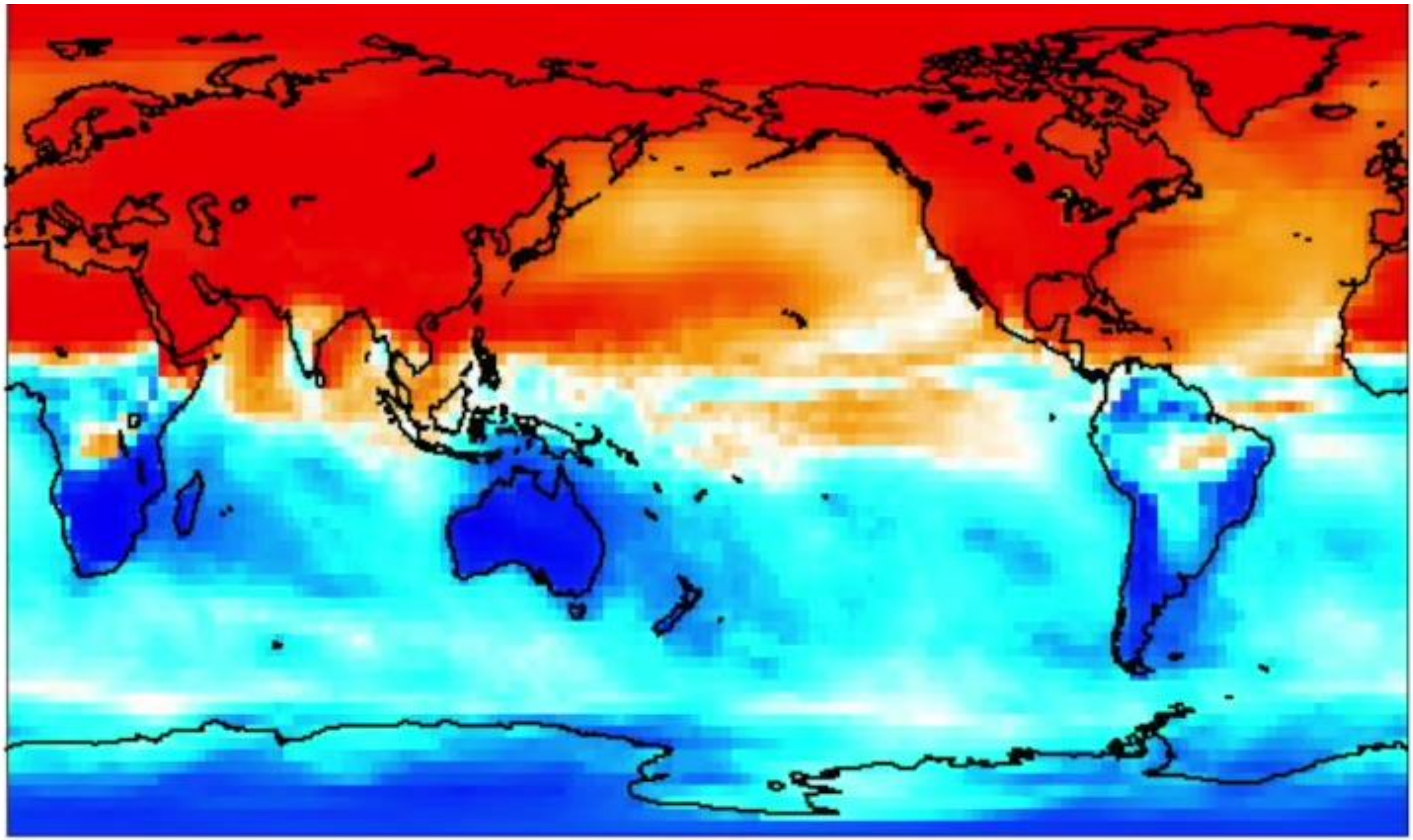
  10226 time series

- Daily resolution **1979 – 2016** $\Rightarrow$ each time series has 13700 data points



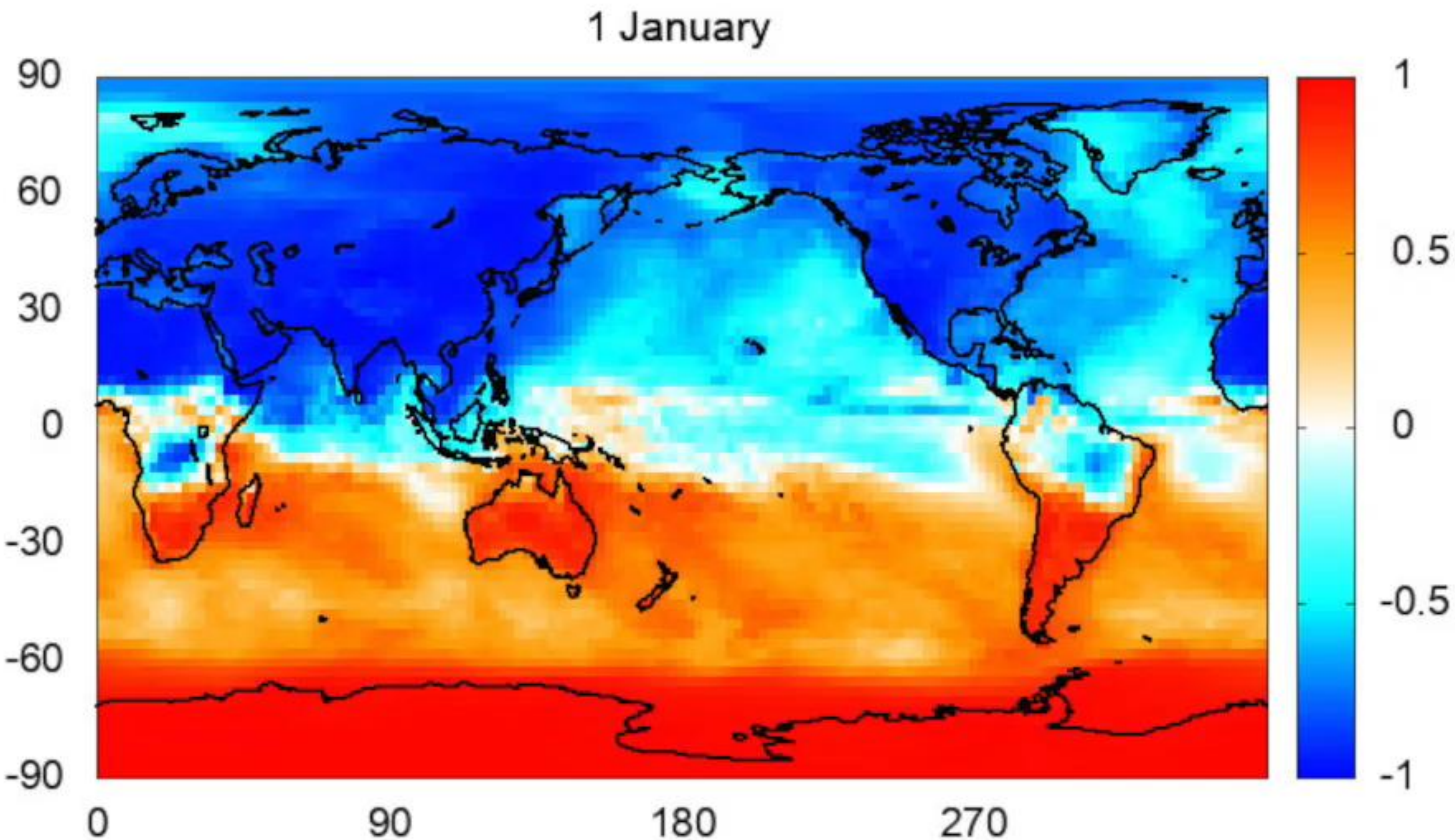
cristina.masoller@upc.edu @cristinamasoll1

**Which information carries the <u>Hilbert phase</u>?**
In color code the cos(φ) averaged over all **July 1** in the period 1979 – 2016.



cristina.masoller@upc.edu @cristinamasoll1

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




cristina.masoller@upc.edu @cristinamasoll1

## Can we use Hilbert analysis to identify and quantify regional "climate change"?

# Relative decadal variations in each region ("node")

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

$$\text{Relative variation} = \frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

Relative variation is considered **significant** if:

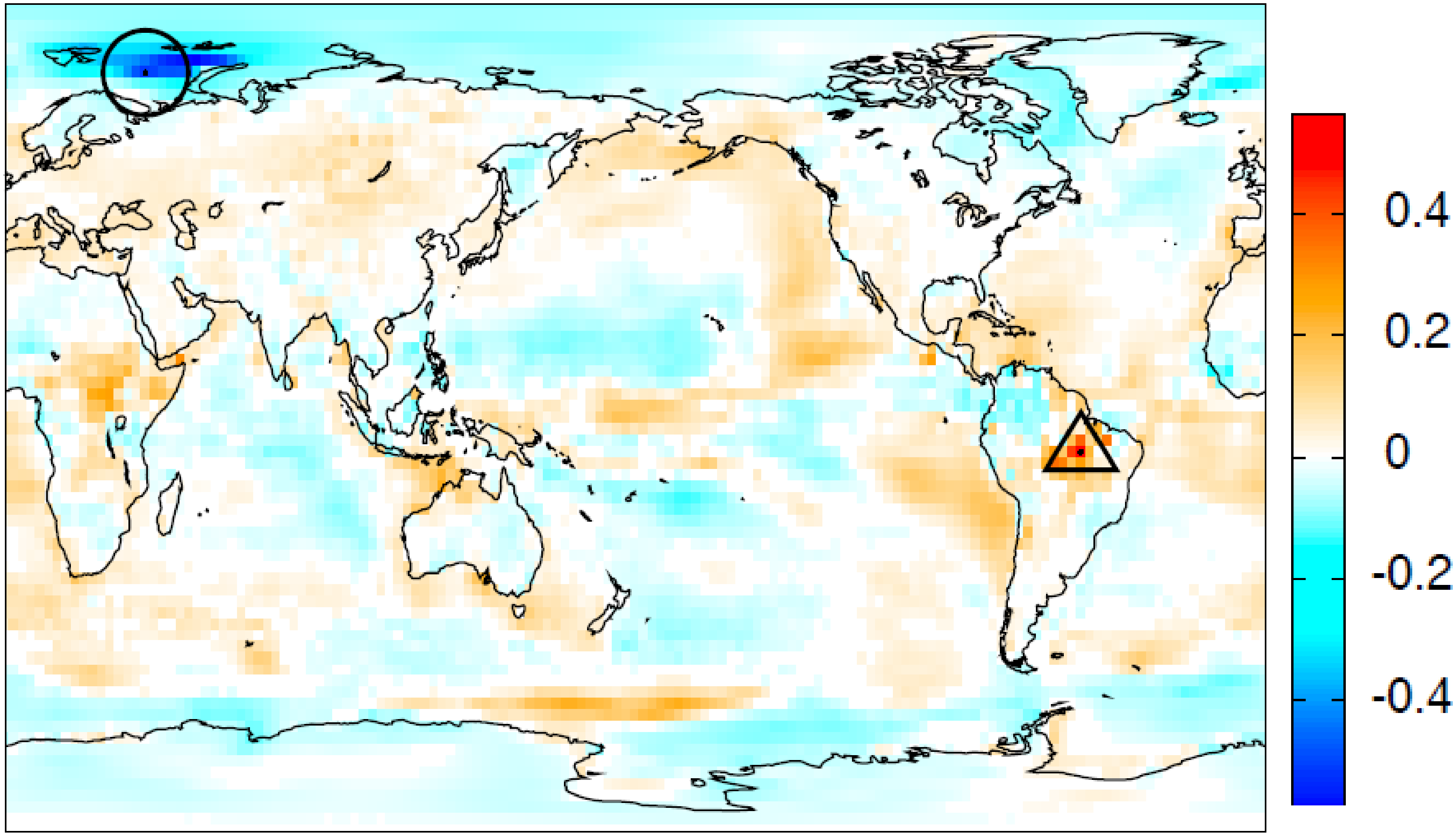
$$\frac{\Delta a}{\langle a \rangle} \geq \langle . \rangle_s + 2\sigma_s \quad \text{or} \quad \frac{\Delta a}{\langle a \rangle} \leq \langle . \rangle_s - 2\sigma_s$$

100 "surrogates"

*G. Lancaster et al, "Surrogate data for hypothesis testing of physical systems", Physics Reports 748, 1 (2018).*

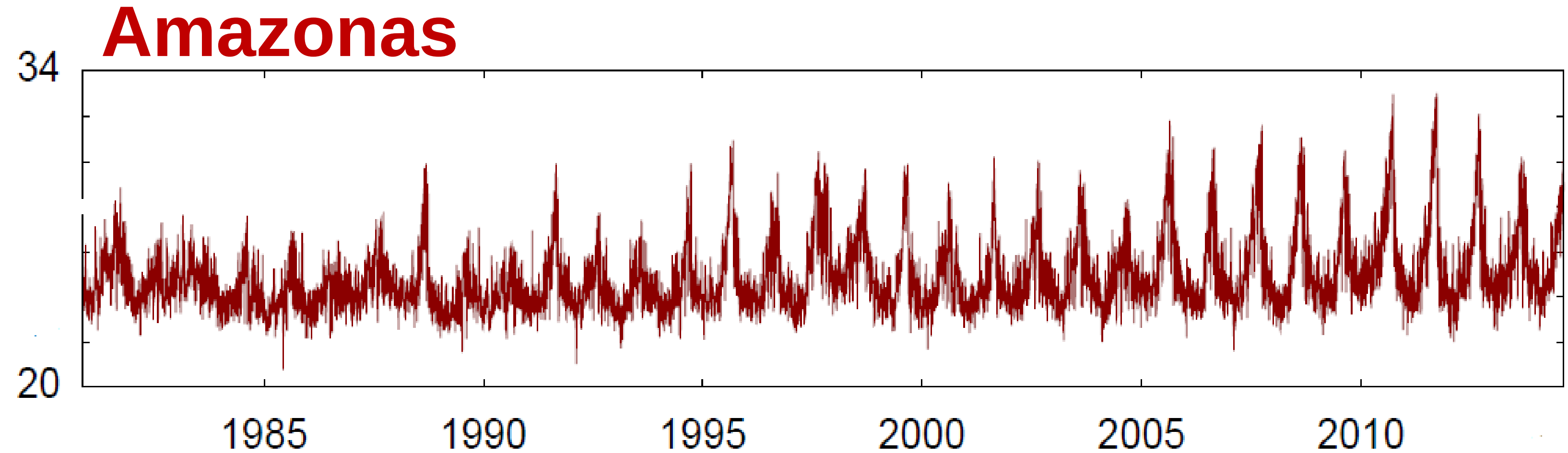
cristina.masoller@upc.edu @cristinamasoll1

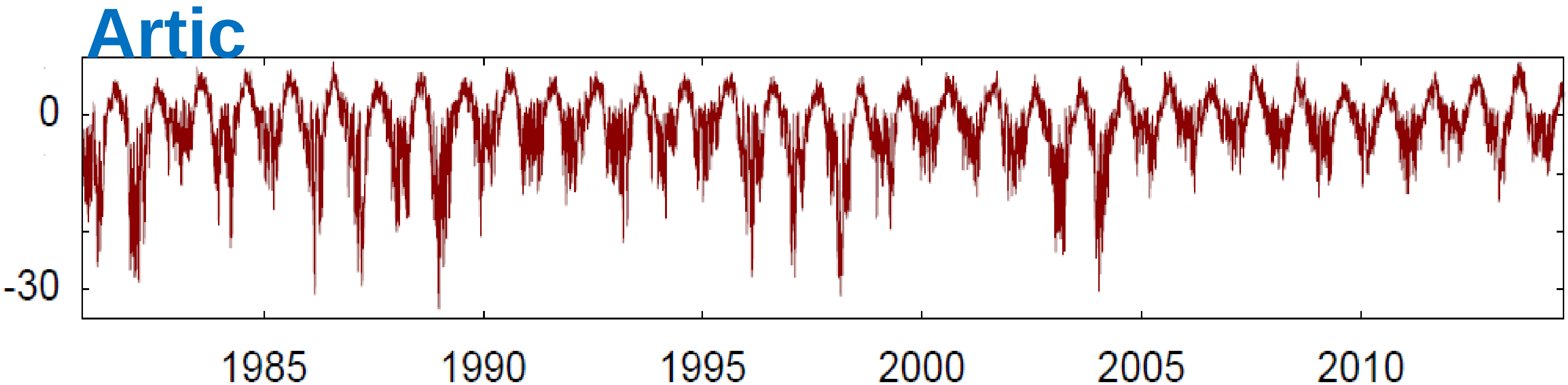
# Relative decadal variations




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

cristina.masoller@upc.edu @cristinamasoll1




- **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.

cristina.masoller@upc.edu @cristinamasoll1

# Outline

- Hilbert analysis to identify significant variations in Surface Air Temperature (SAT) reanalisis data.

  $\{x_1, x_2, \ldots x_N\}$

- Bivariate analysis: mutual information of SAT anomalies and causal networks of climatic indices (ENSO, NAO, etc.)

  $\{x_1, x_2, \ldots x_N\}$ $\{y_1, y_2, \ldots y_N\}$

- Climate networks constructed by applying ordinal analysis to SAT anomalies.

cristina.masoller@upc.edu @cristinamasoll1

## Mutual Information (MI) $\{x_1, x_2, \dots x_N\}$ $\{y_1, y_2, \dots y_N\}$

- *MI is calculated from probability distributions*

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

- It quantifies the reduction in uncertainty of one variable by knowing the other variable.

- *If X, Y are independent, MI = 0*, else ***MI >0***

- For Gaussian processes: MI = $-1/2 \log(1-\rho^2)$ where $\rho$ is the cross-correlation coefficient.



cristina.masoller@upc.edu @cristinamasoll1

# Transfer Entropy (TE) and Directionality Index (DI)

- TE: is the Conditional Mutual Information, given the "past" of one of the variables.

$$TE\ (x,y) = MI\ (x,\ y|x_\tau)$$

$$TE\ (y,x) = MI\ (y,\ x|y_\tau)$$

- MI (x,y) = MI (y,x) but TE (x,y) $\neq$ TE(y,x)

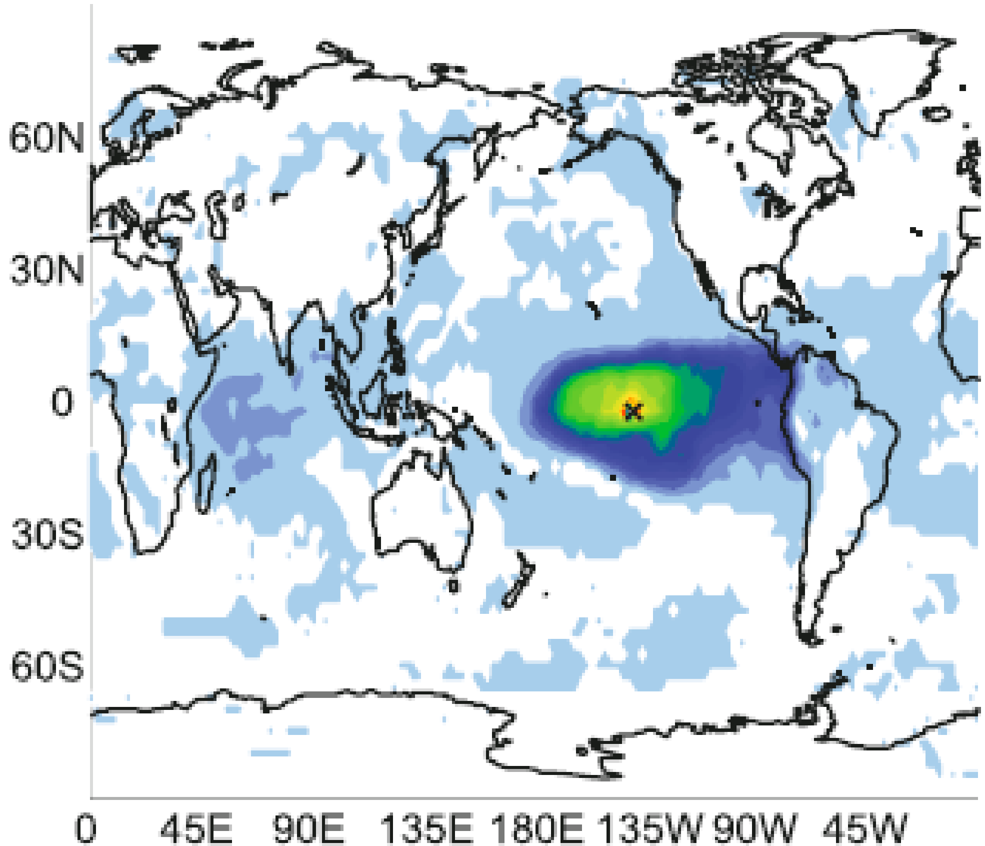
- Directionality Index: TE(x,y)-TE(y,x)

*K. Hlaváčková-Schindler et al. / Physics Reports 441 (2007) 1–46*



cristina.masoller@upc.edu @cristinamasoll1

# Mutual Information of SAT anomalies in El Niño region and other regions (white: MI not significant)




Monthly data, NCEP/NCAR, January 1949 to December 2006

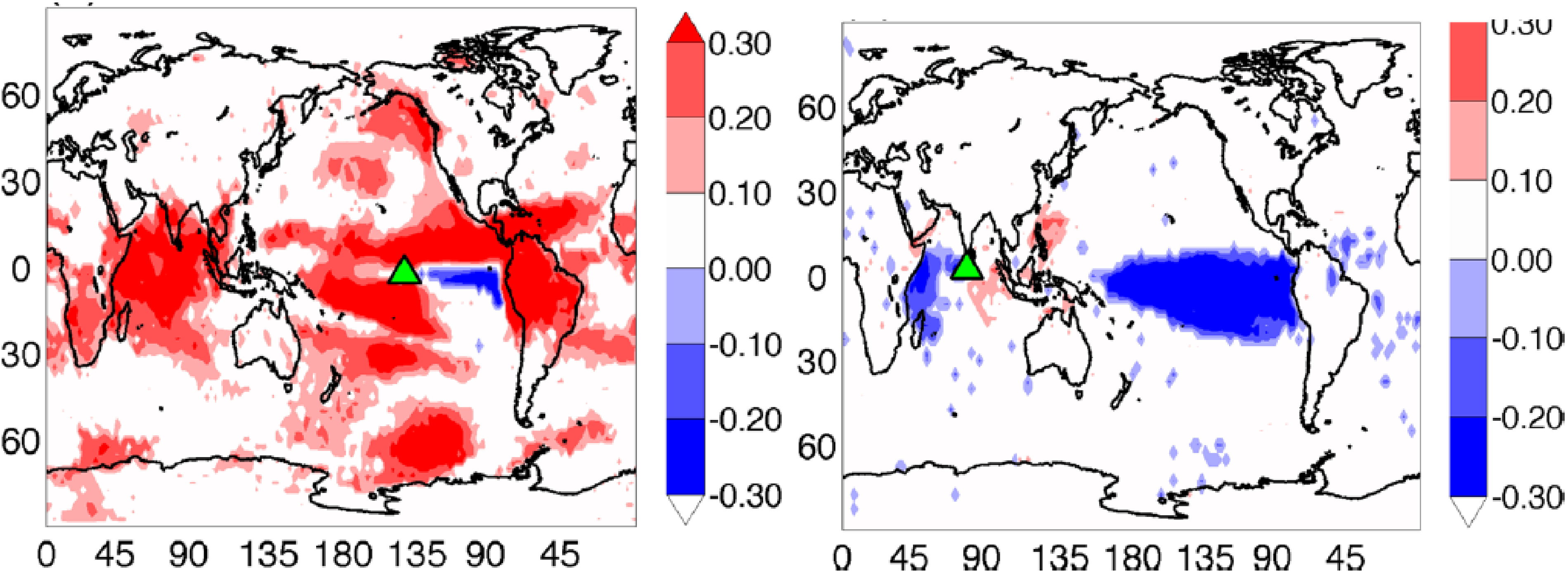
*J. I. Deza, M. Barreiro, C. Masoller, "Inferring interdependencies in climate networks constructed at inter-annual, intra-season and longer time scales", Eur. Phys. J. ST 222, 511 (2013).*



cristina.masoller@upc.edu @cristinamasoll1

# Directionality Index

Daily data, NCEP/NCAR, January 1949 to December 2013




J. I. Deza, M. Barreiro, and C. Masoller, "*Assessing the direction of climate interactions by means of complex networks and information theoretic tools*", Chaos 25, 033105 (2015).

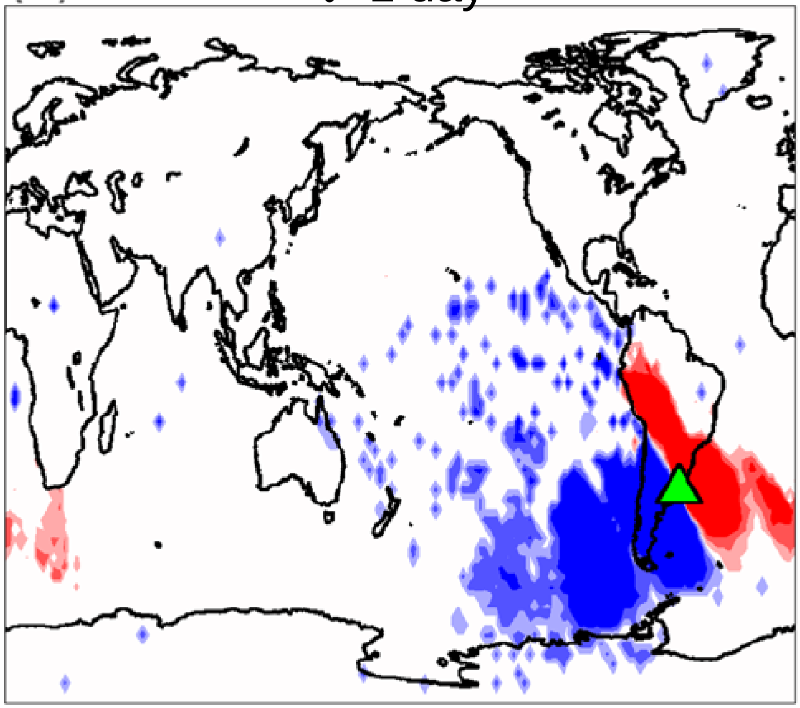


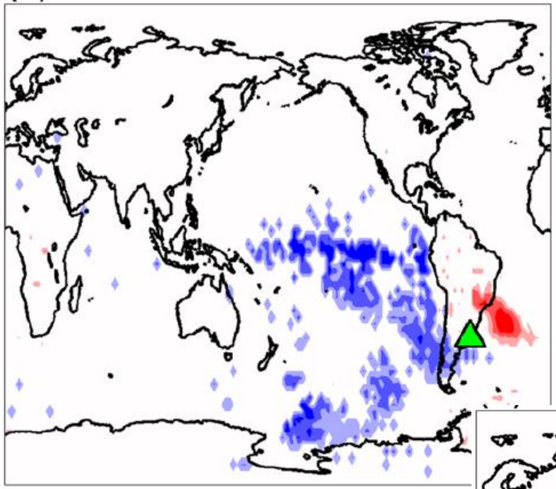
cristina.masoller@upc.edu @cristinamasoll1

# Influence of the lag time

$$TE(x,y) = MI(x, y|x_{\tau})$$
$$TE(y,x) = MI(y, x|y_{\tau})$$
$$DI = TE(x,y) - TE(y,x)$$

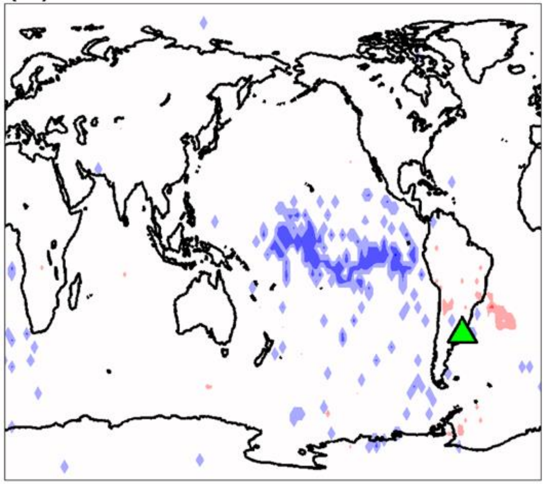
$\tau$=1 day





$\tau$=7 days

$\tau$=30 days



J. I. Deza, M. Barreiro, and C. Masoller, Chaos 25, 033105 (2015).



cristina.masoller@upc.edu @cristinamasoll1

## Problem: Transfer Entropy is computationally demanding

"simple" solution: use the expression that is valid for Gaussian distributions [ $MI = -1/2 \log(1-\rho^2)$ ]

Can this work?

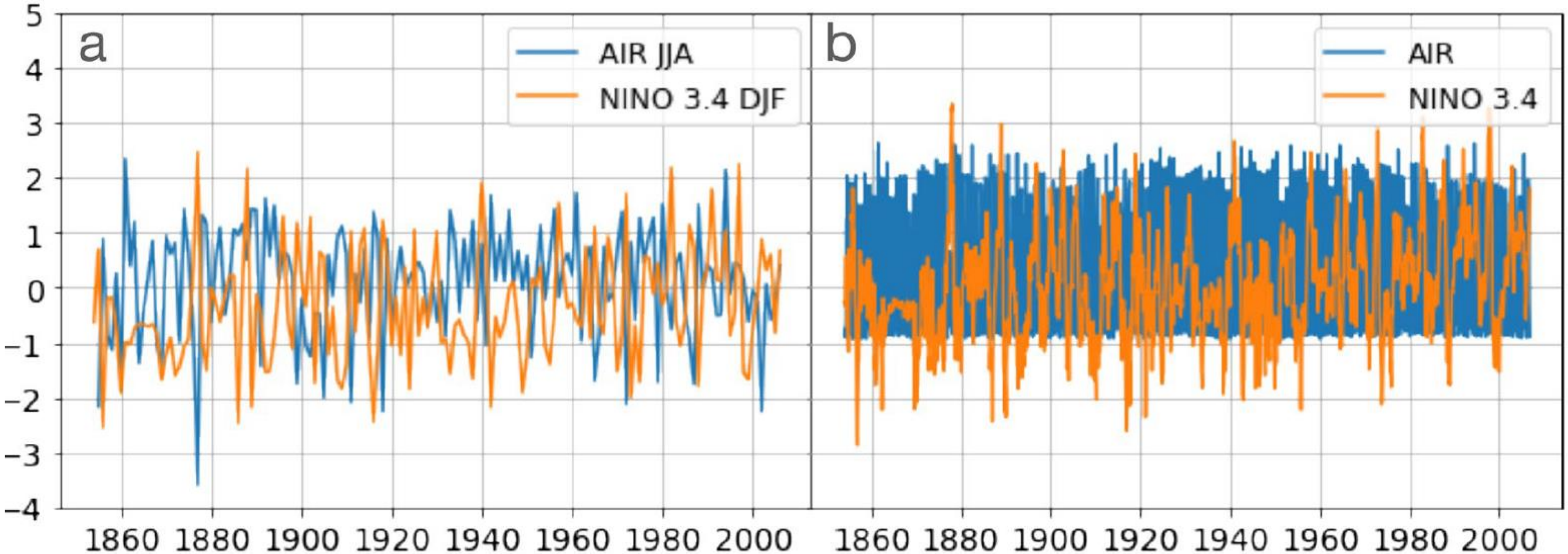
R. Silini, C. Masoller "*Fast and effective pseudo transfer entropy for bivariate data-driven causal inference*", Sci. Rep. 11, 8423 (2021).

**Analysis of NINO3.4 ←→ All India Rainfall indices**



cristina.masoller@upc.edu @cristinamasoll1

**Yearly sampled** (152) **Monthly sampled**



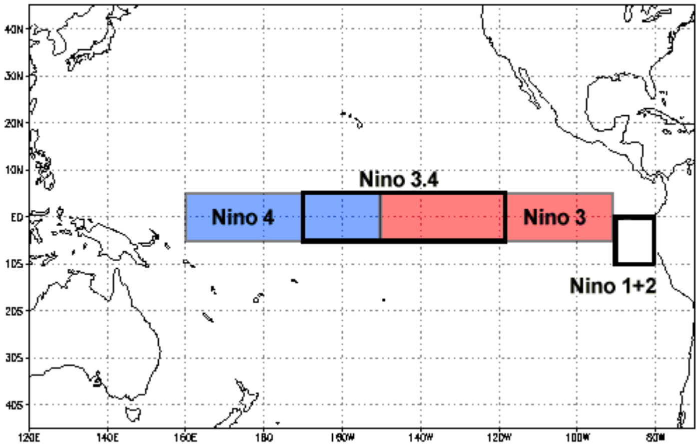

| | Yearly sampled | Monthly sampled |
|---|---|---|
| pTE | NINO3.4 ⟶ AIR (0.04 s) | NINO3.4 ⟵ AIR (0.5 s) |
| GC | NINO3.4 ⟶ AIR (0.4 s) | NINO3.4 ⟵ AIR (0.9 s) |
| TE | NINO3.4 ⇆ AIR (1 s) | NINO3.4 ⇆ AIR (68 s) |

cristina.masoller@upc.edu @cristinamasoll1

## How much time can we save?

For two time-series of 500 data points (1 data point per month, 40 years):

TE:**112 ms** but pTE: **4 ms**




8000 grid points (high resolution)
$\Rightarrow$ 64 x $10^6$ pairs

$\Rightarrow$ **829 days** (TE) vs. **29 days** (pTE).

(without “surrogate” analysis)

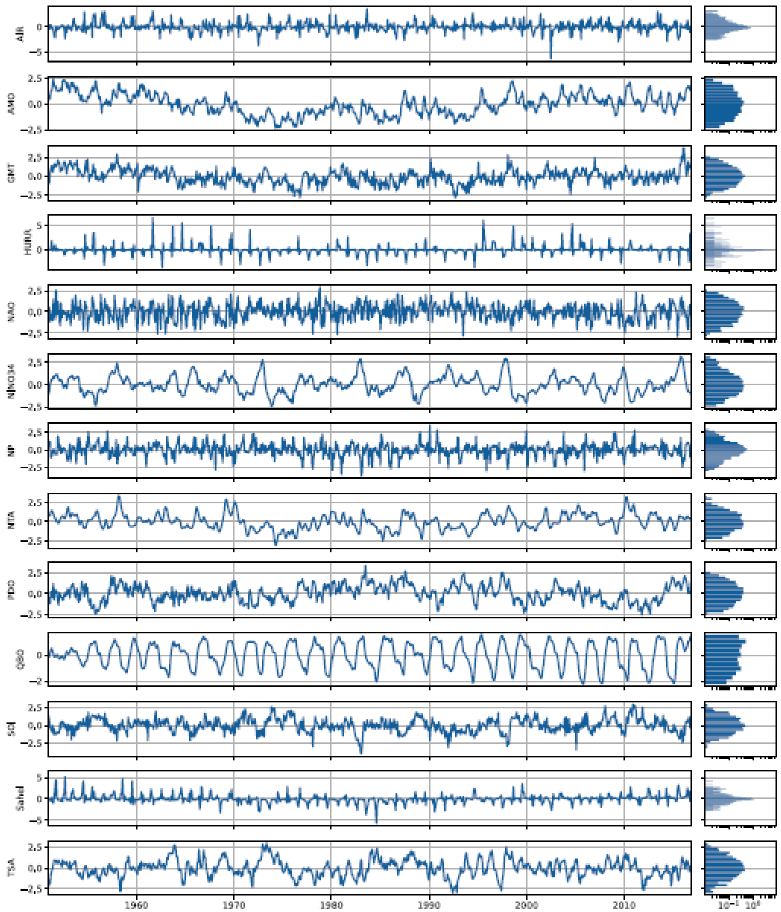
https://github.com/riccardosilini/pTE

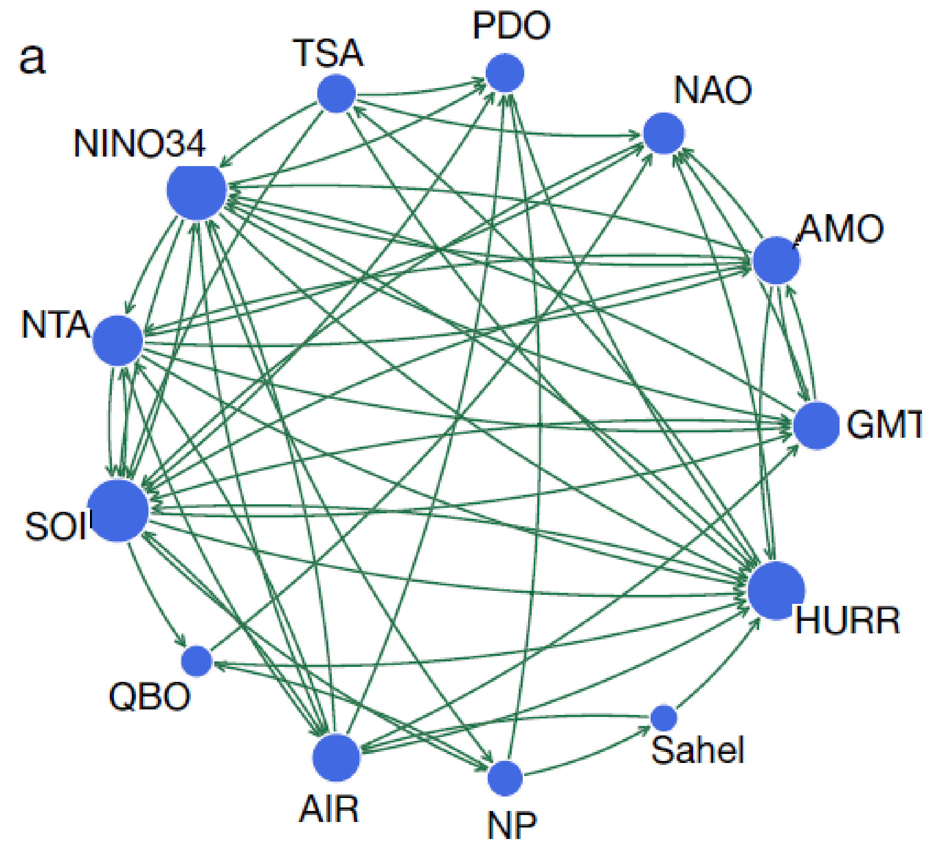


cristina.masoller@upc.edu @cristinamasoll1

# Directed network of 13 climatic indices

Constructed from pTE analysis with different lags




R. Silini, G. Tirabassi, M Barreiro, L. Ferranti, C. Masoller, "*Assessing causal dependencies in climatic indices*", Climate Dynamics 61, 79–89 (2023).



cristina.masoller@upc.edu @cristinamasoll1

# Outline

- Hilbert analysis to identify significant variations in Surface Air Temperature (SAT) reanalisis data.

  $\{x_1, x_2, \ldots x_N\}$

- Bivariate analysis: mutual information of SAT anomalies and causal networks of climatic indices (ENSO, NAO, etc.)

  $\{x_1, x_2, \ldots x_N\}$ $\{y_1, y_2, \ldots y_N\}$

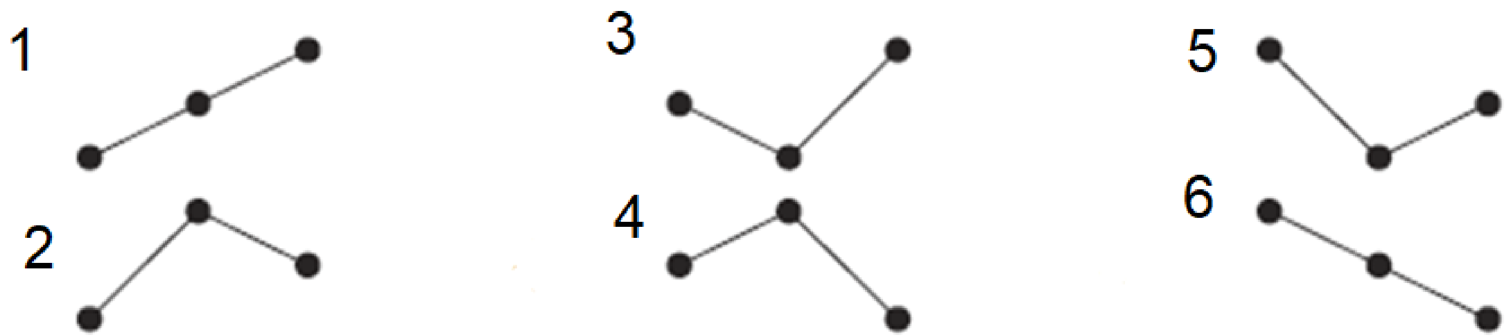
- Climate networks constructed from SAT anomalies by using *ordinal analysis*.

cristina.masoller@upc.edu @cristinamasoll1

# Ordinal analysis $\{\ldots x_i, x_{i+1}, x_{i+2}, \ldots\}$

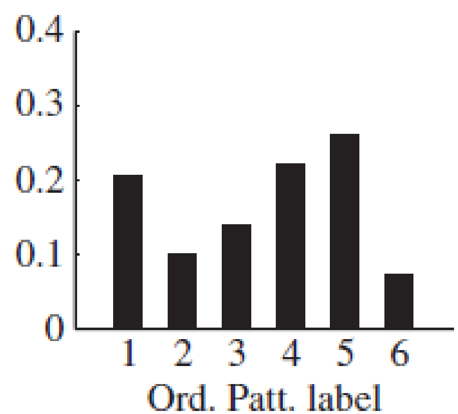
How can three data points (let's say 2, 5, 7) be ordered?

{…2, 5, 7…}  {…5, 2, 7…}  {…7, 2, 5 …}




{…2, 7, 5…}  {…5, 7, 2…}  {…7, 5, 2…}

Ordinal probabilities

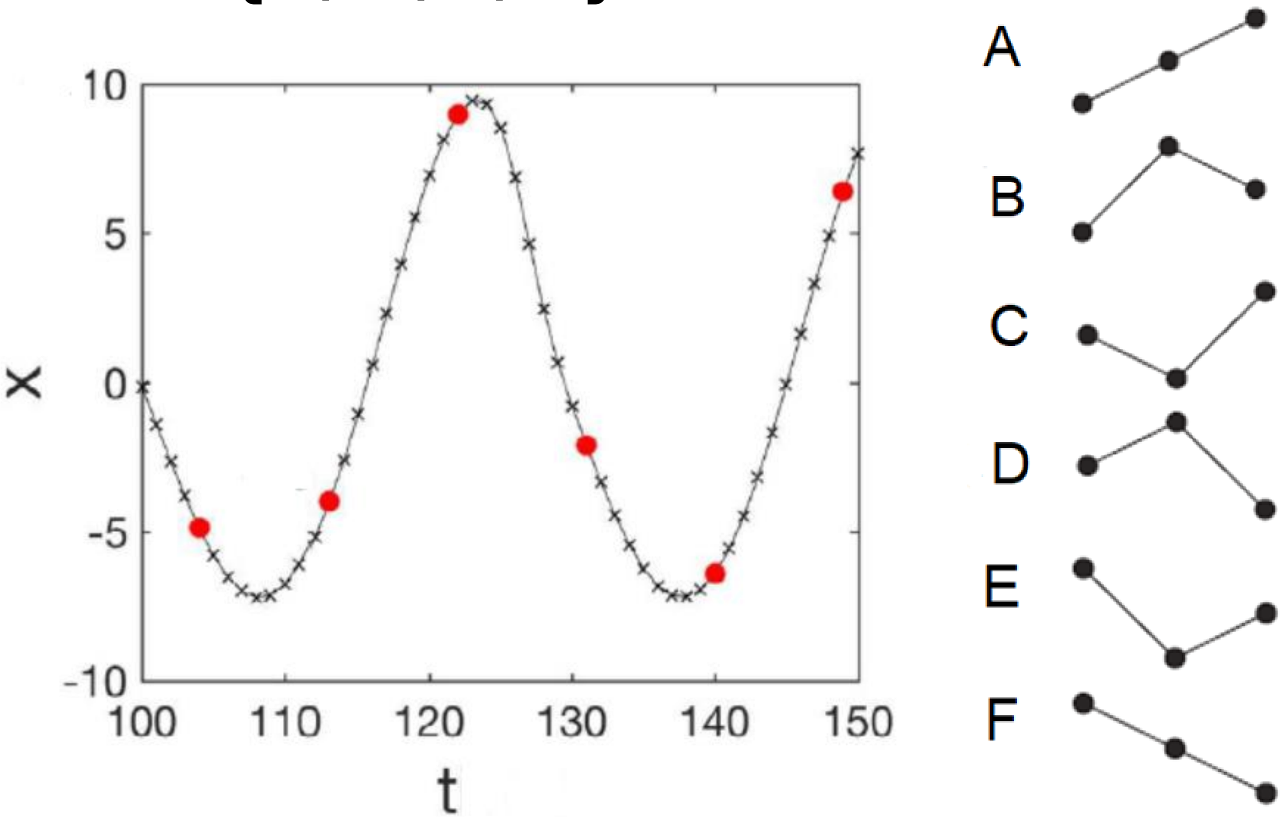



Bandt and Pompe: Phys. Rev. Lett. 2002

cristina.masoller@upc.edu @cristinamasoll1

# Which is the “message” “encoded” in the red dots?

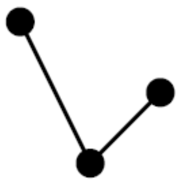
## {A, B, F, C}



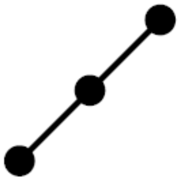

cristina.masoller@upc.edu @cristinamasoll1

# Using lagged points to define the patterns allows to select the time scale of the analysis, useful for climatological data

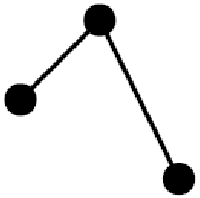
**Intra-season**

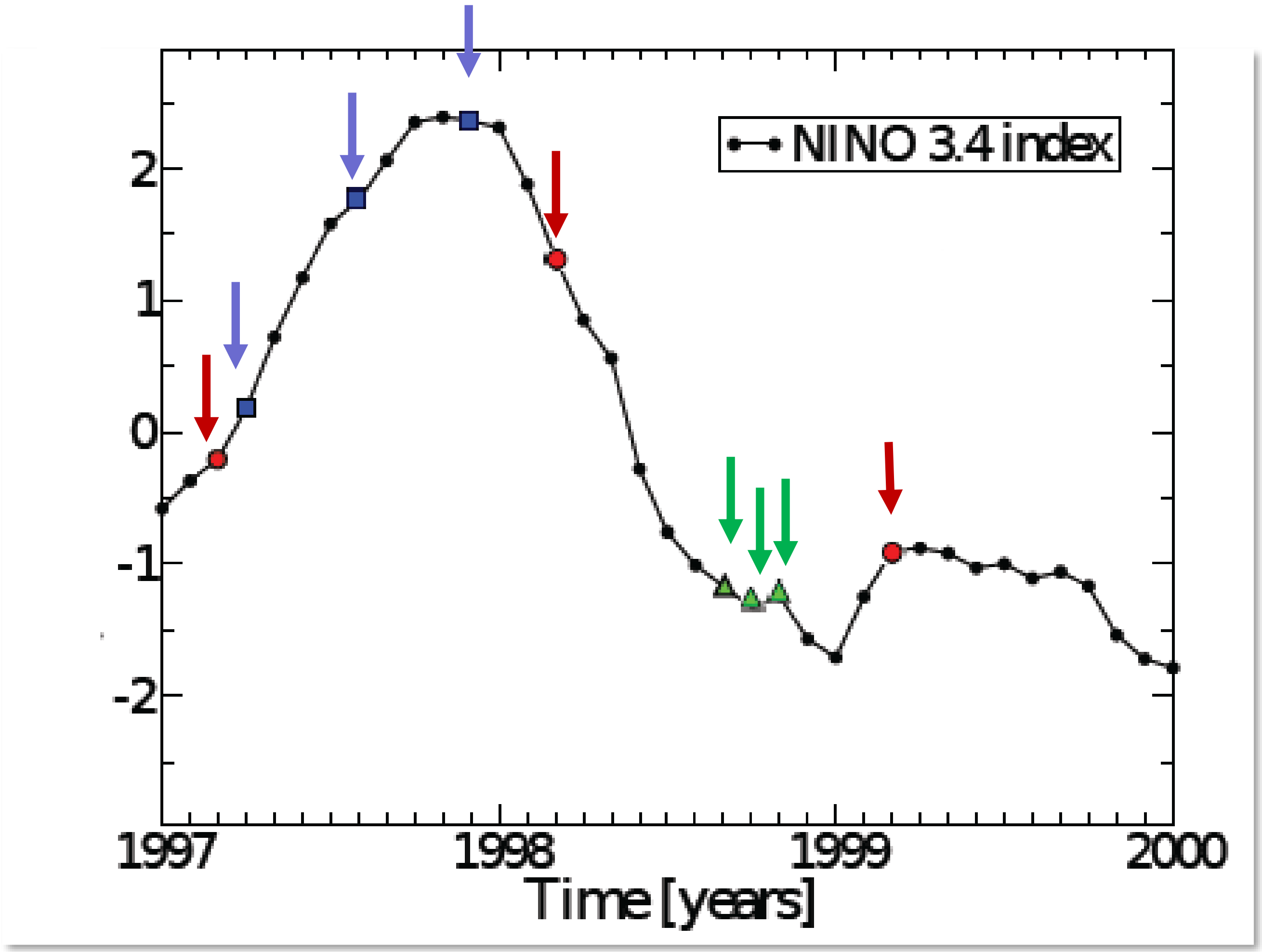


**Intra-annual**



**Inter-annual**





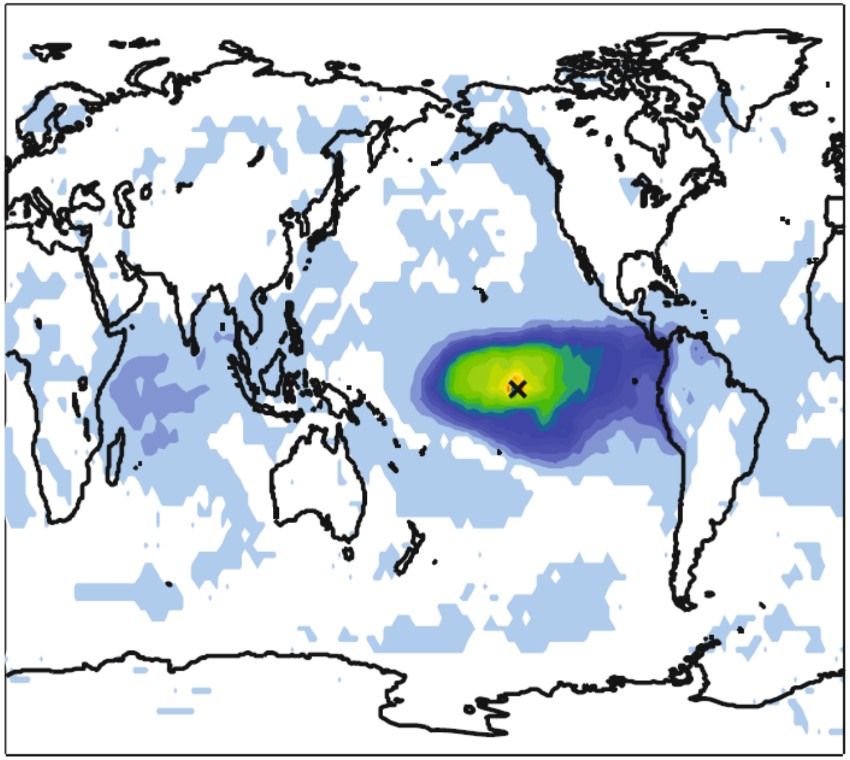



cristina.masoller@upc.edu @cristinamasoll1

# Mutual Information of SAT anomaly in El Niño region and other regions (shown before, white: MI not significant)

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

MI from (usual) probabilities of values of SAT anomalies



*J. I. Deza, M. Barreiro, C. Masoller, Eur. Phys. J. ST 222, 511 (2013)*

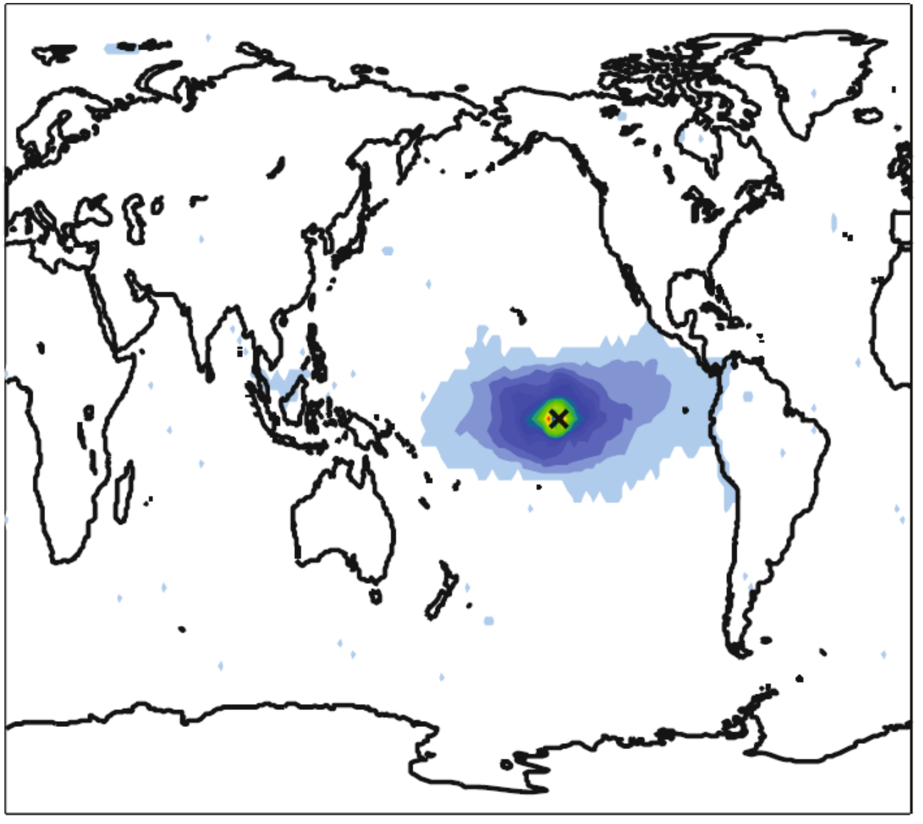


cristina.masoller@upc.edu @cristinamasoll1

# Mutual Information (color code) from probabilities of ordinal patterns (white: MI not significant)

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

MI from probabilities of ordinal patterns defined by values in 3 consecutive months.



*J. I. Deza, M. Barreiro, C. Masoller, Eur. Phys. J. ST 222, 511 (2013).*

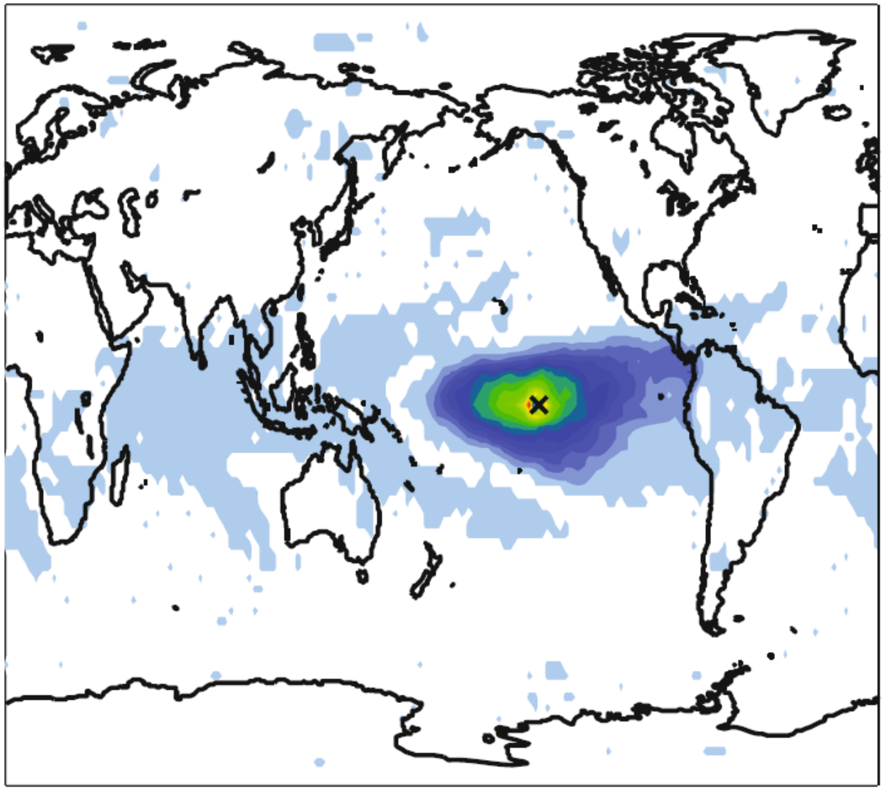


cristina.masoller@upc.edu @cristinamasoll1

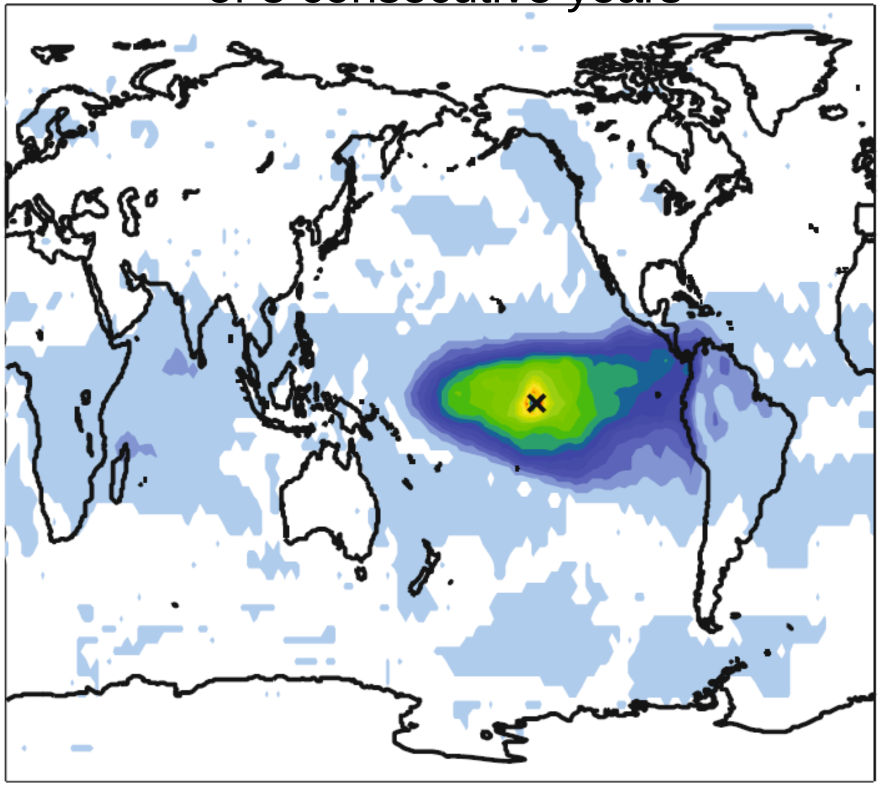
# Comparison of two ways to calculate the ordinal probabilities, used to calculate the mutual information

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

Patterns defined by 3 values within a year



Patterns defined by data values of 3 consecutive years



*J. I. Deza, M. Barreiro, C. Masoller, Eur. Phys. J. ST 222, 511 (2013).*

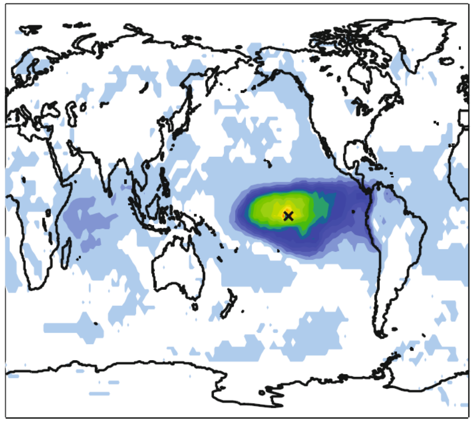


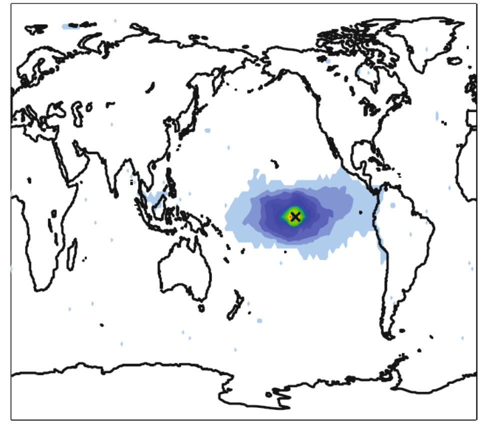
cristina.masoller@upc.edu @cristinamasoll1

# Comparison

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

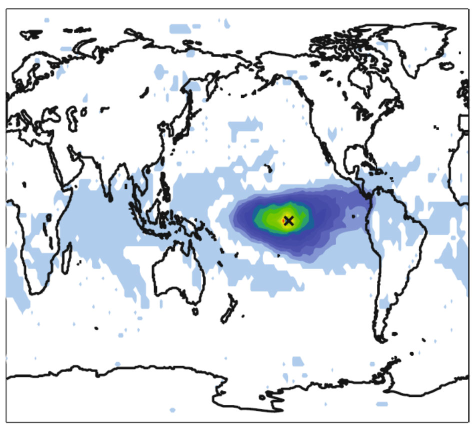
Probabilities of SAT values

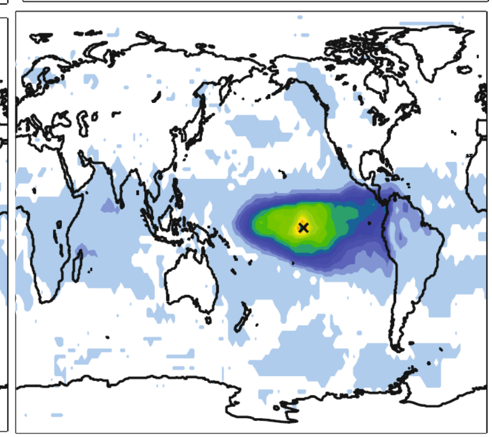




Probabilities of ordinal patterns defined by values in 3 consecutive months.

Probabilities of patterns defined by 3 values within a year.





Probabilities of patterns defined by values in 3 consecutive years.

*J. I. Deza, M. Barreiro, C. Masoller, Eur. Phys. J. ST 222, 511 (2013).*



cristina.masoller@upc.edu @cristinamasoll1

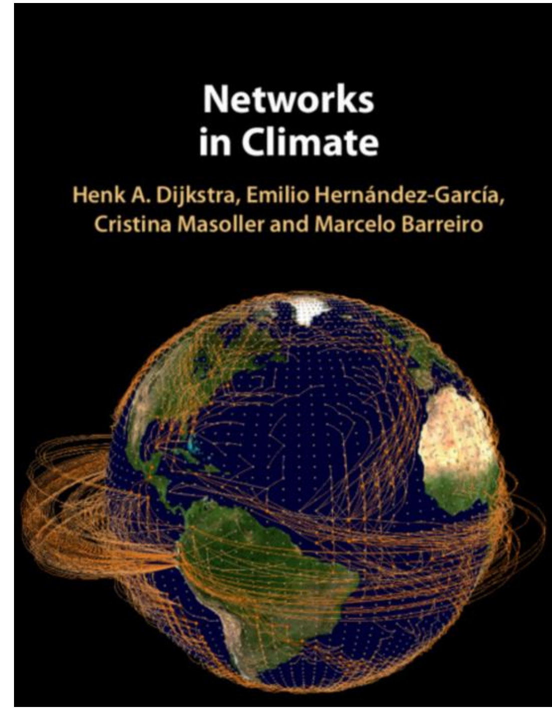
## Take home messages

- Hilbert analysis and ordinal analysis are versatile tools that can provide new insights into climate phenomena.
- Mutual information can be calculated in terms of the probabilities of ordinal patterns, allowing to select the time-scale of the analysis.
- Different large-scale spatial structures are uncovered when using different lags between the data points that define the ordinal patterns.

cristina.masoller@upc.edu @cristinamasoll1

# Thank you for your attention


- J. I. Deza, M. Barreiro, C. Masoller, Eur. Phys. J. ST 222, 511 (2013)
- J. I. Deza, M. Barreiro, C. Masoller, Chaos 25, 033105 (2015)
- D. A. Zappala, M. Barreiro, C. Masoller, Earth Syst. Dynamics 9, 383 (2018)
- R. Silini and C. Masoller, Scientific Reports 11, 8423 (2021)
- R. Silini, G. Tirabassi, M Barreiro, L. Ferranti, C. Masoller, Climate Dynamics 61, 79 (2023).
- H. Dijkstra, M. Barreiro, E. Hernandez-Garcia, C. Masoller, *Networks in Climate*, Cambridge University Press (2019)


Agència de Gestió d'Ajuts Universitaris i de Recerca

ICREA

GOBIERNO DE ESPAÑA | MINISTERIO DE CIENCIA E INNOVACIÓN


cristina.masoller@upc.edu @cristinamasoll1