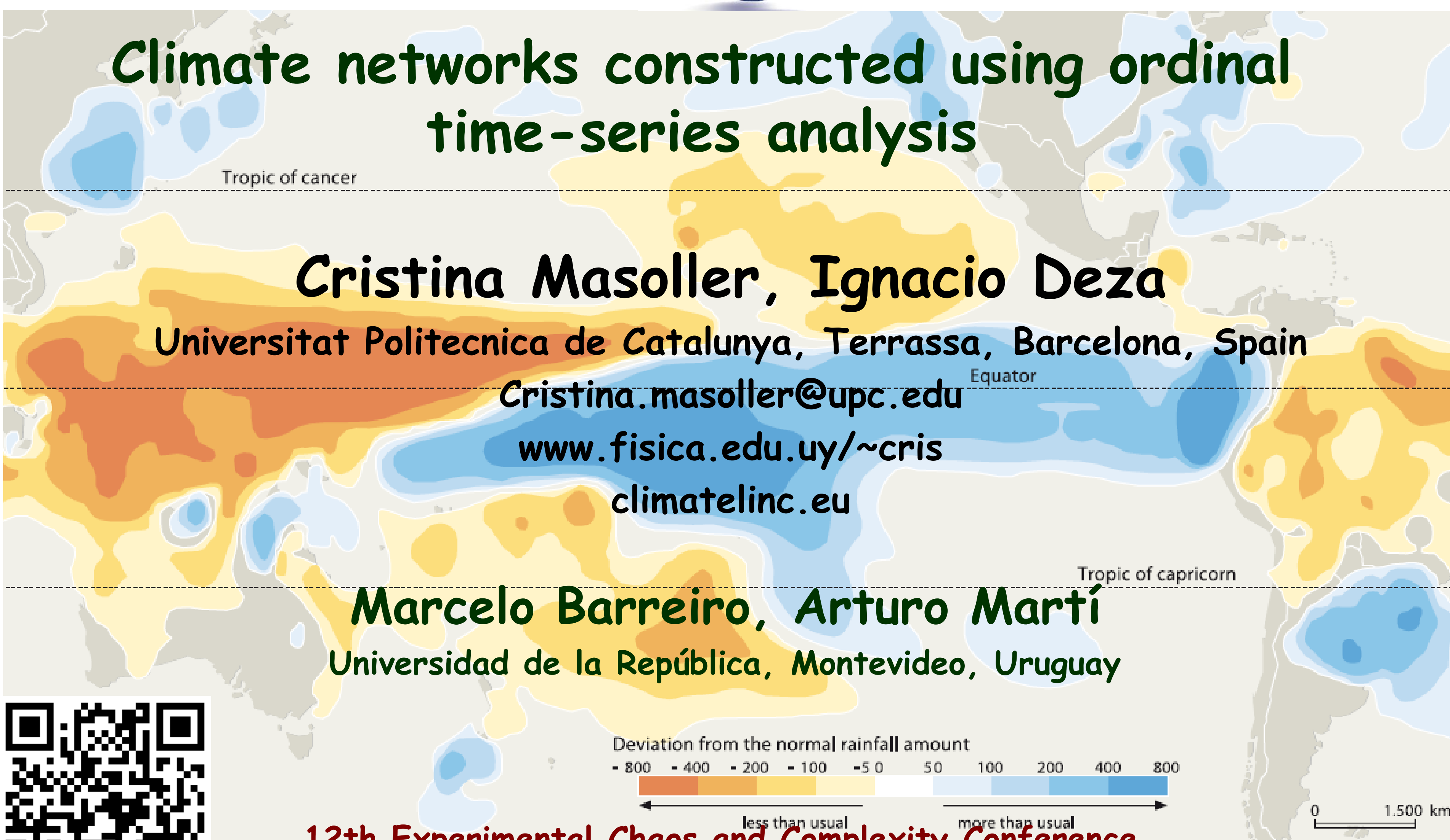
"LINC"
Learning about Interacting Networks in Climate

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

# Climate networks constructed using ordinal time-series analysis


**Cristina Masoller, Ignacio Deza**
Universitat Politecnica de Catalunya, Terrassa, Barcelona, Spain
Cristina.masoller@upc.edu
www.fisica.edu.uy/~cris
climatelinc.eu

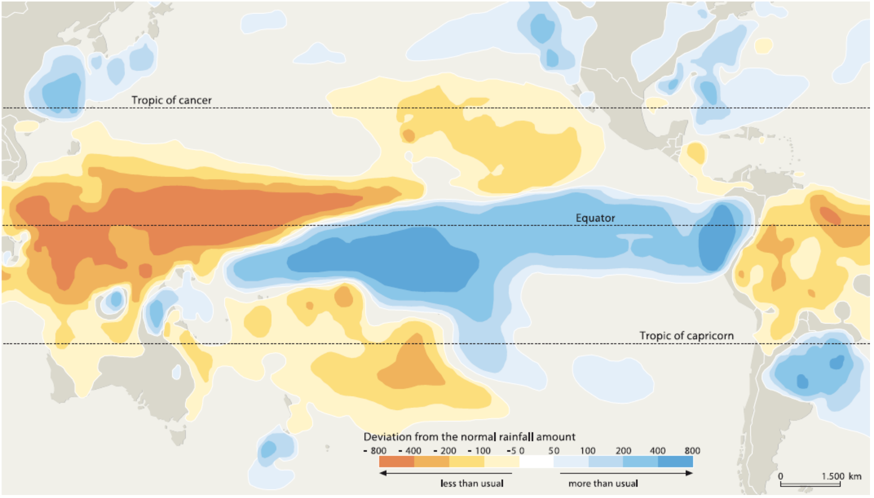
**Marcelo Barreiro, Arturo Martí**
Universidad de la República, Montevideo, Uruguay



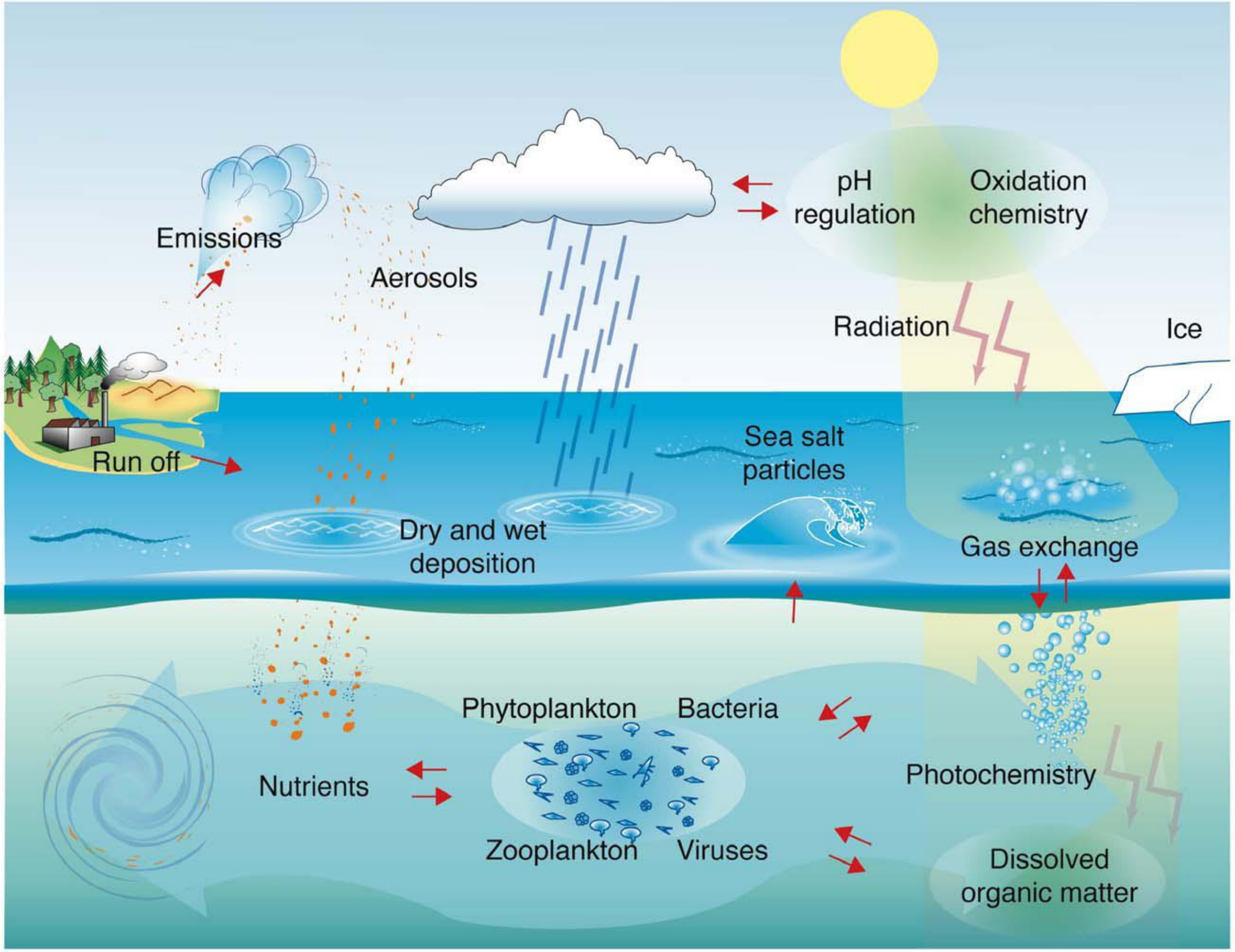
12th Experimental Chaos and Complexity Conference
University of Michigan at Ann Arbor, May 2012

# Outline




- Climate Networks
- Ordinal time-series analysis
- Results
- Conclusions

# The climate: a highly nonlinear and complex system




Adapted from Elliott and Maltrud, Los Alamos Nat. Lab.

# With a wide range of time scales

- hours to days,
- months to seasons,
- decades to centuries,
- and even longer time-scales...

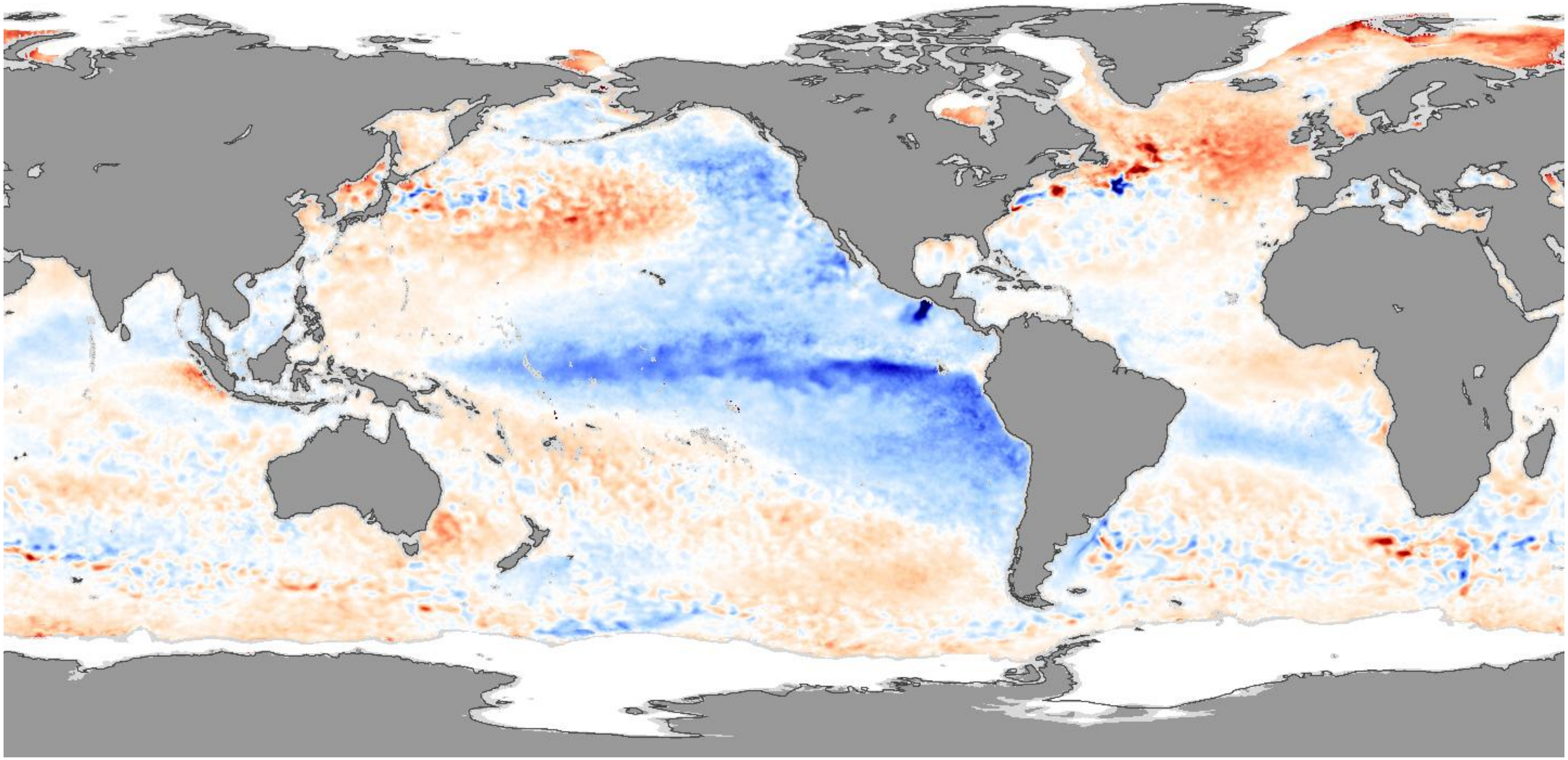
Example: El Niño/La Niña-Southern Oscillation (**ENSO**)

- Occurs across the tropical Pacific Ocean with ≈ 5 years periodicity.
- variations in the surface temperature of the tropical eastern Pacific Ocean (warming: El Niño, cooling: La Niña)
- variations in the air surface pressure in the tropical western Pacific (the Southern Oscillation).
- The two variations are coupled:

  El Niño (ocean warming) -- high air surface pressure,

  La Niña (ocean cooling) -- low air surface pressure.

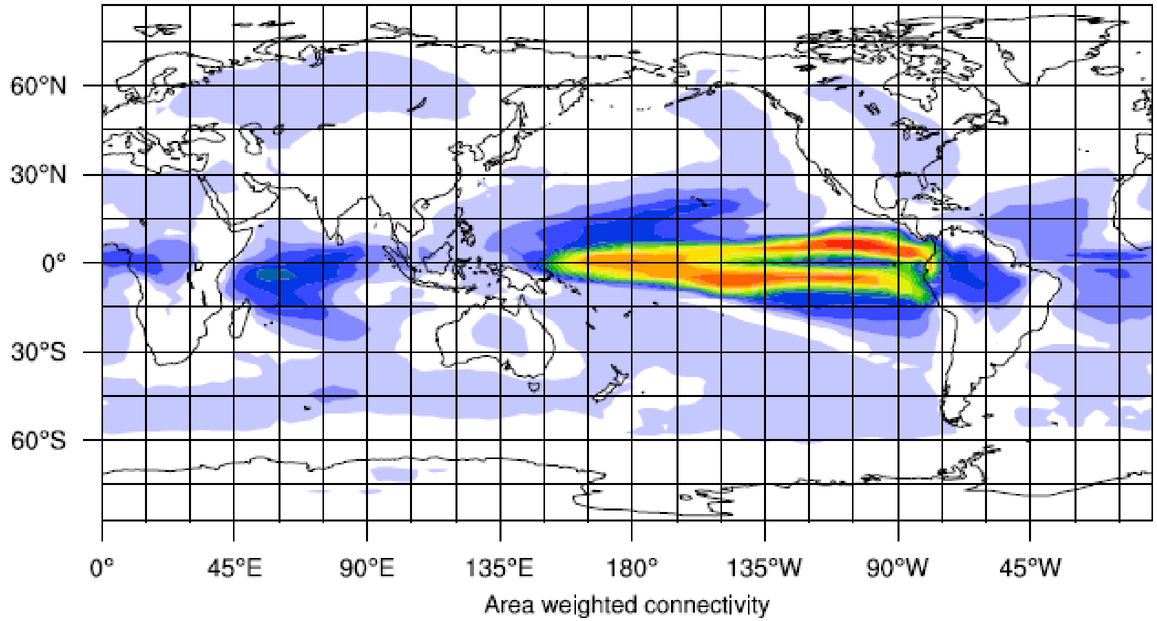
# Surface Sea Temperature anomalies during La Niña
(November 2007)



Source: Wikipedia

# Constructing climate networks

- over a regular grid (**nodes**) covering the Earth's surface.
- interdependencies are reflected as **links**
- Each grid point represents an area of 2.5 degree latitude by 2.5 degrees longitude (about 250 kms by 250 kms at the equator).




Donges et al, Eur. Phys. J. Spe. Top. 2009: ***Understanding the Earth as a Complex System***

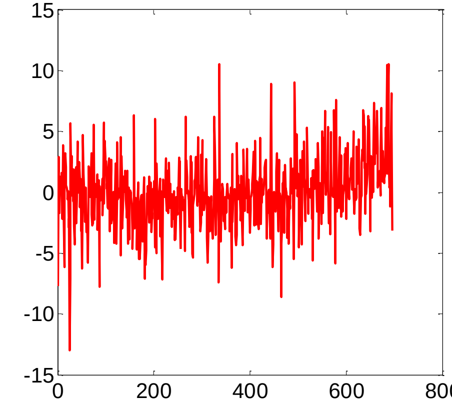
- Total number of grip points: 10,226 (93 x 144 -2)
- graphical representation of the **area-weighted connectivity**
  **AWC = fraction of the Earth to which a node is connected.**

# The data: Surface Air Temperature (SAT) Anomalies

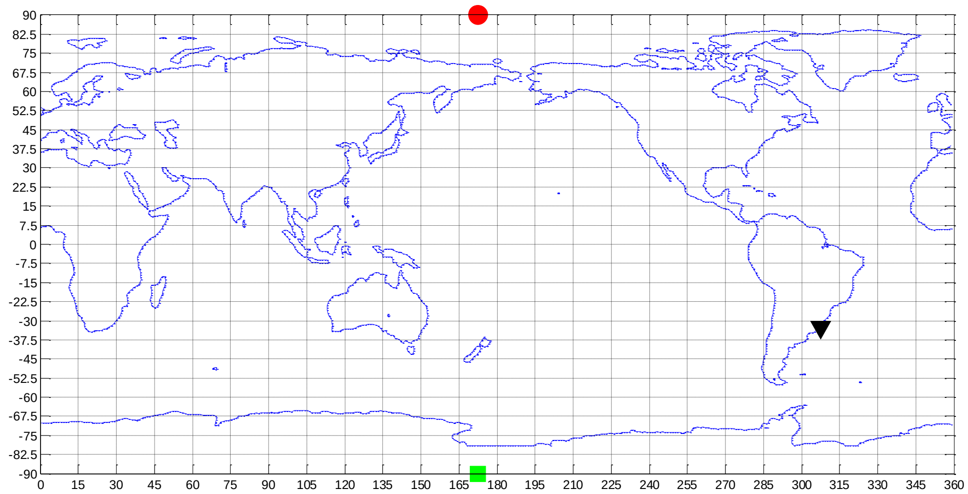
January 1949 -- December 2006

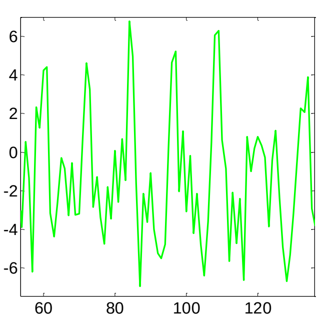
**In each 'node' 696 data points (58 years x 12 months)**

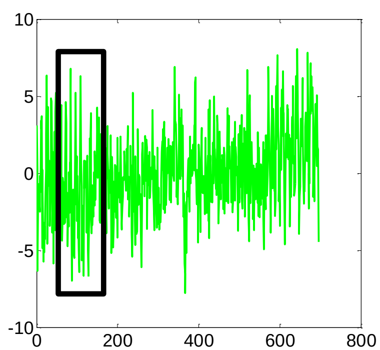
Anomalies = <u>annual cycle removed</u>



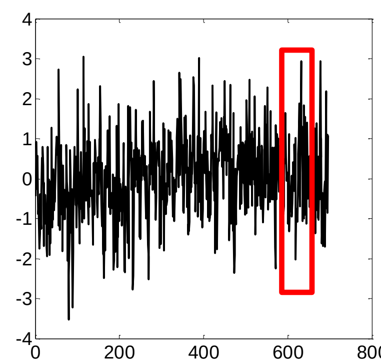

Time (months)

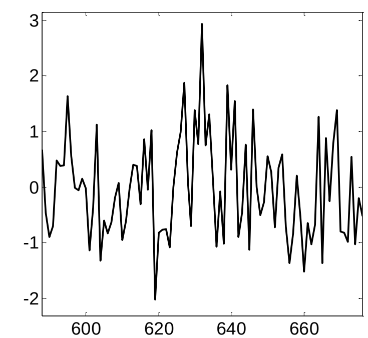



Time (months)




C. Masoller Time (months)

# Where does the data comes from?

- Reanalysis of National Center for Environmental Prediction, National Center for Atmospheric Research (NCEP-NCAR, USA).
- Reanalysis = run a sophisticated model of general atmospheric circulation and feed it with the available experimental data, in the different points of the earth, at their corresponding times.
- This process restricts the solution of the model to one as close to reality as possible in regions where there are data available, and to a solution physically "plausible" in regions where no data is available.

# Three steps for constructing climate networks

- Nonlinear approach

1. Transform the time-series of SAT anomalies into a sequence of patterns (or words) by using ordinal time-series analysis

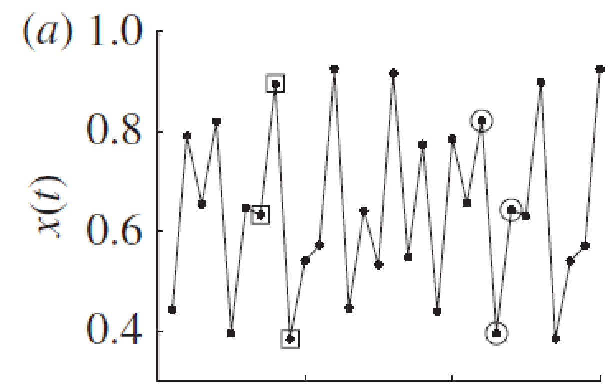
2. Quantify the degree of 'statistical interdependency' between the sequences in two nodes by using a nonlinear measure

3. Define if there is a link between two nodes by using threshold

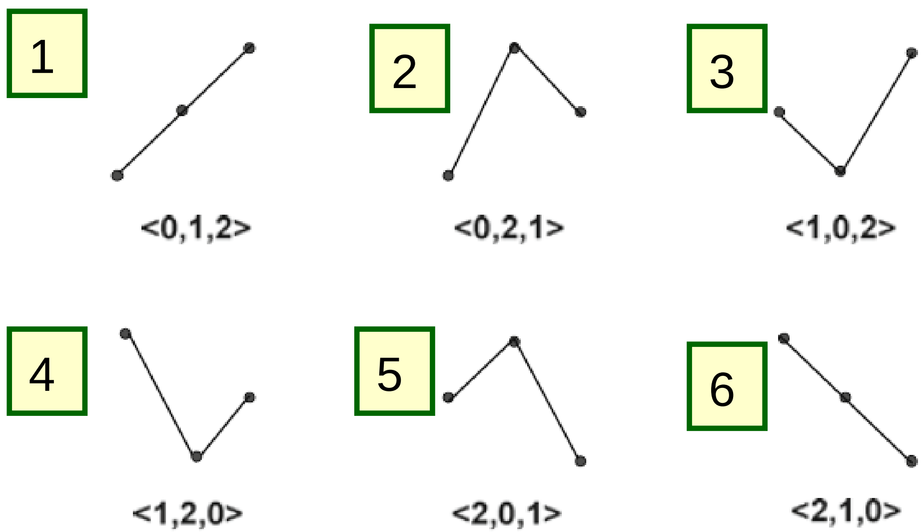
# Step 1: Ordinal Time-Series Analysis

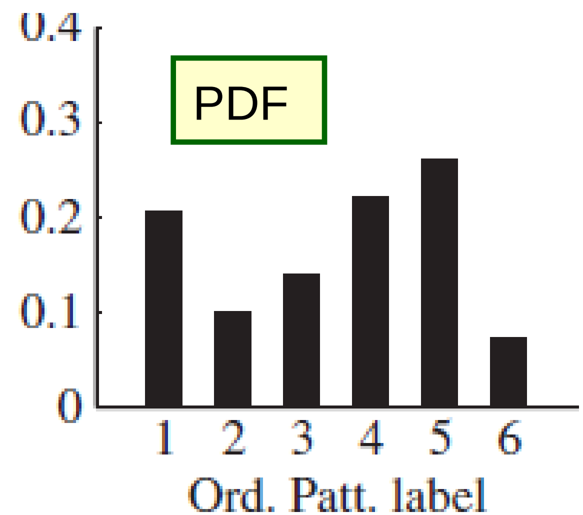
Ordinal Patterns (OPs) of length **D** take into account the **order relations** of D values in a sequence of values (Brandt & Pompe, PRL 2002):

$$\{\ x(1), x(2), \ldots, \underbrace{x(t+0), x(t+1), x(t+2)}, \ldots, x(N-1), x(N)\ \}$$




Geometrical representation of 6 OPs of length 3:

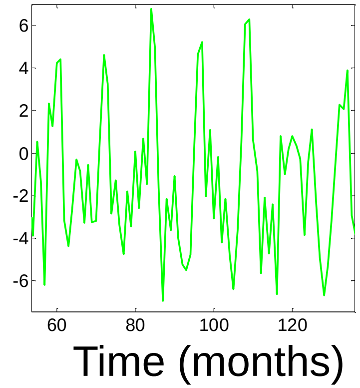



- Number of possible OPs: D!
- Good statistics if: N >> D!

# Ordinal pattern analysis of *climatological* data

- The central paradigm is that in climatological data there are repeated patterns of oscillations.




One can construct the OPs comparing monthly-averaged SAT anomalies on:

consecutive years or consecutive months

$$[x_i(t), x_i(t+12), x_i(t+24)]$$

(inter-annual time-scale)

$$[x_i(t), x_i(t+1), x_i(t+2)]$$

(intra-season time-scale)

# Step 2: quantify the degree of 'statistical interdependency'

- Nonlinear measure: the Mutual Information

$$M_{i\,j} = \sum_{m,n=1}^{N_{bin}} p_{i\,j}(m,n) \log \frac{p_{i\,j}(m,n)}{p_i(m)\, p_j(n)}$$

- $M_{ij} = 0 \Leftrightarrow \{x_i\}$ and $\{x_j\}$ are independent: $p_{i\,j}(m,n) = p_i(m)\, p_j(n)$
- $M_{ij} = M_{ji} \Leftrightarrow$ the links are symmetric.
- Number of bins to estimate the pdfs = number of possible OPs
- Length of OP D=4 $\Rightarrow$ # of OPs 4! =24 << 696 = 12 months x 58 years.

# Step 3: define the links between the nodes

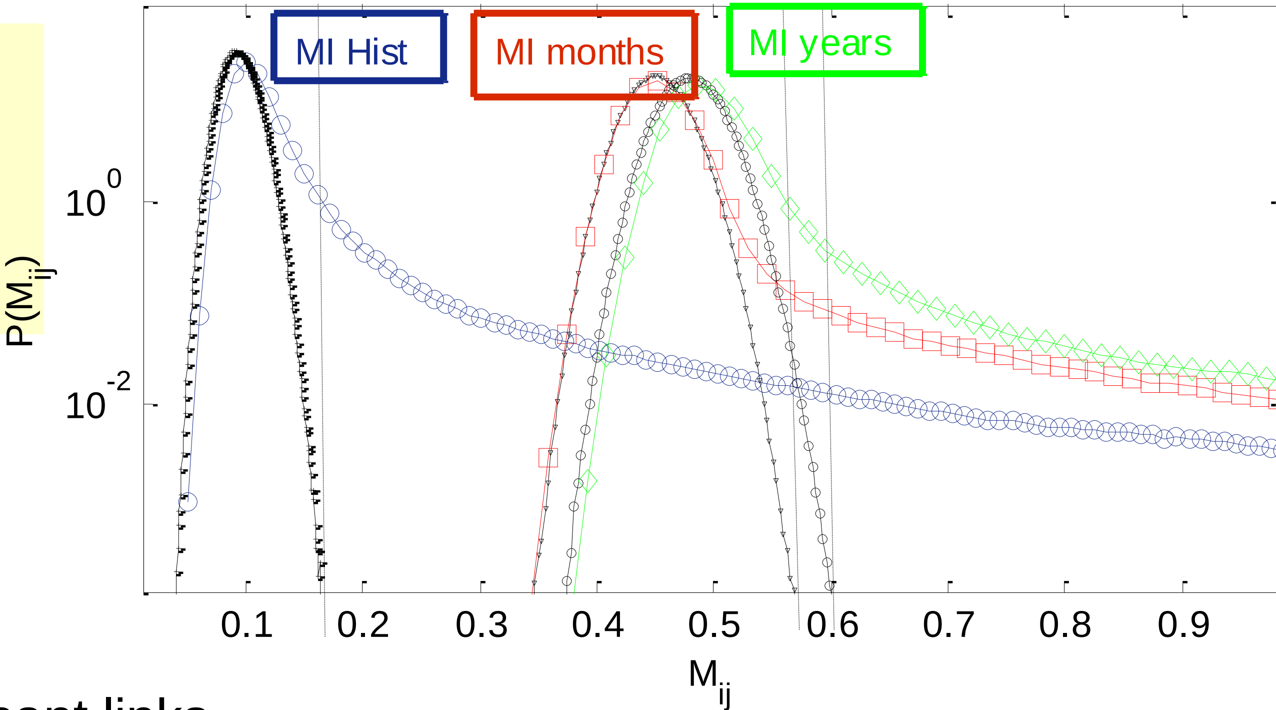
- Significant test:

To check that we only take into account significant links, we first compute the mutual information of surrogate data.

If $M_{ij} \leq \max\left(M_{ij}^{\text{surrogate}}\right) \Rightarrow M_{ij} = 0$

Then, we use a global threshold $\tau$ to define the links:

"i" $\leftrightarrow$ "j" $\Leftrightarrow M_{ij} > \tau$




We adjust $\tau$ such that the network has a certain link density $\rho$ ($\rho$ = 0.01, 0.001).

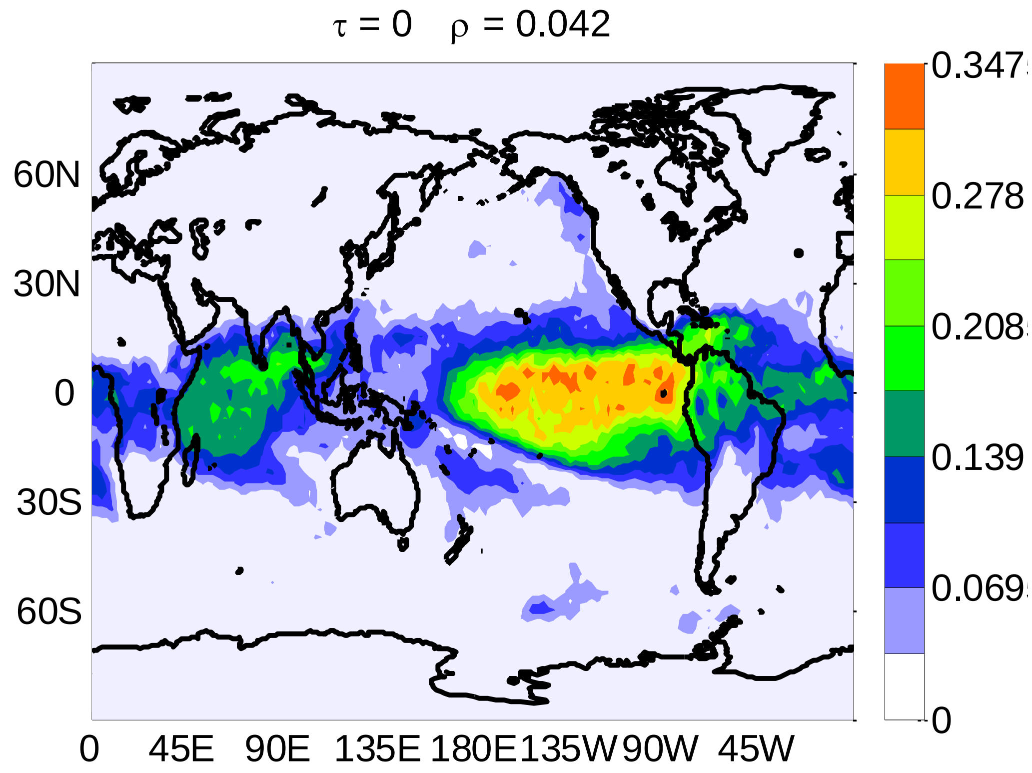
$\tau$ = 0: all significant links

# Results

## OPs formed by concatenating four consecutive years

**D=4** $[x_i(t), x_i(t+12), x_i(t+24), x_i(t+36)]$

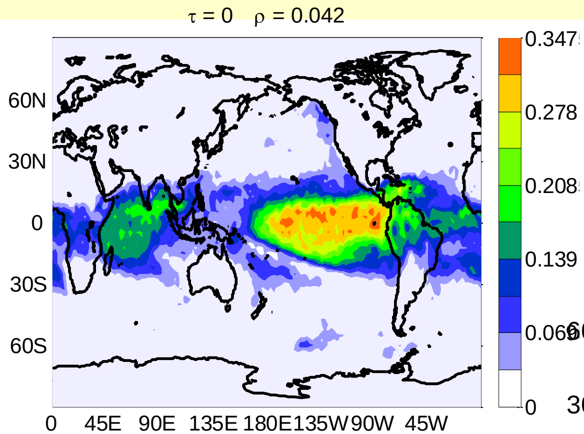
$\tau$=0 (all the ***significant* links**)



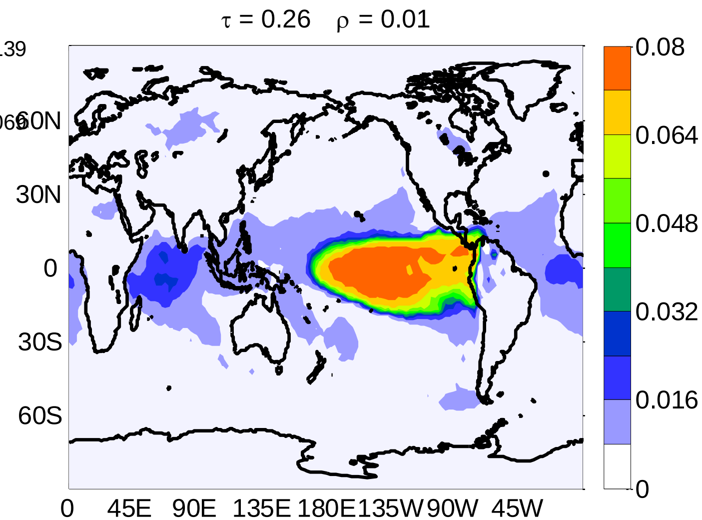

Colors code the Area Weighted Connectivity

Barreiro, Martí and Masoller, Chaos 21, 013101 (2011)

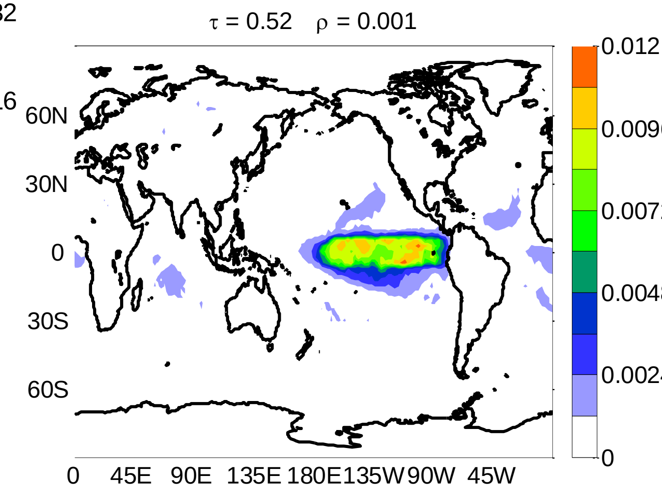
# When increasing the threshold




All the significant links

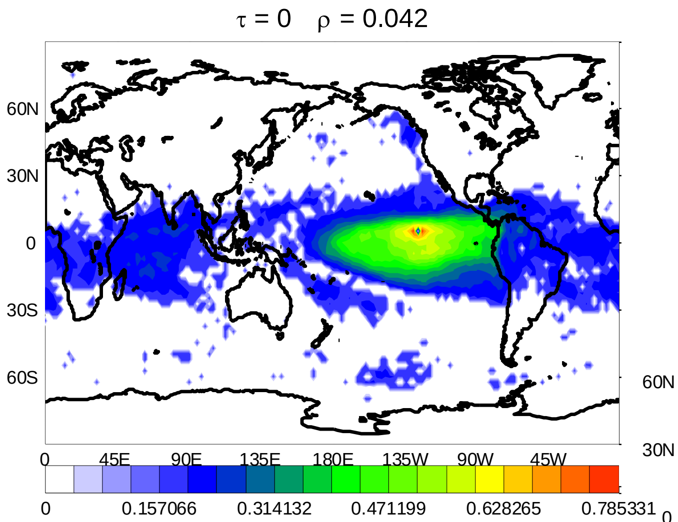


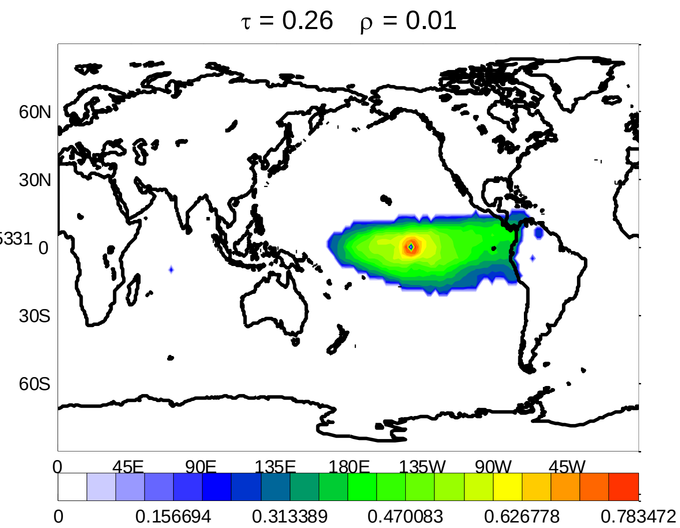

1% of the total number links

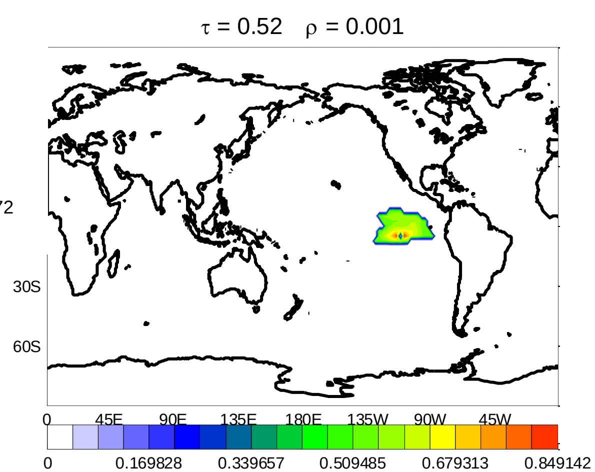



0.1 % (the strongest links)

# Where is the "hub" and where is connected to?



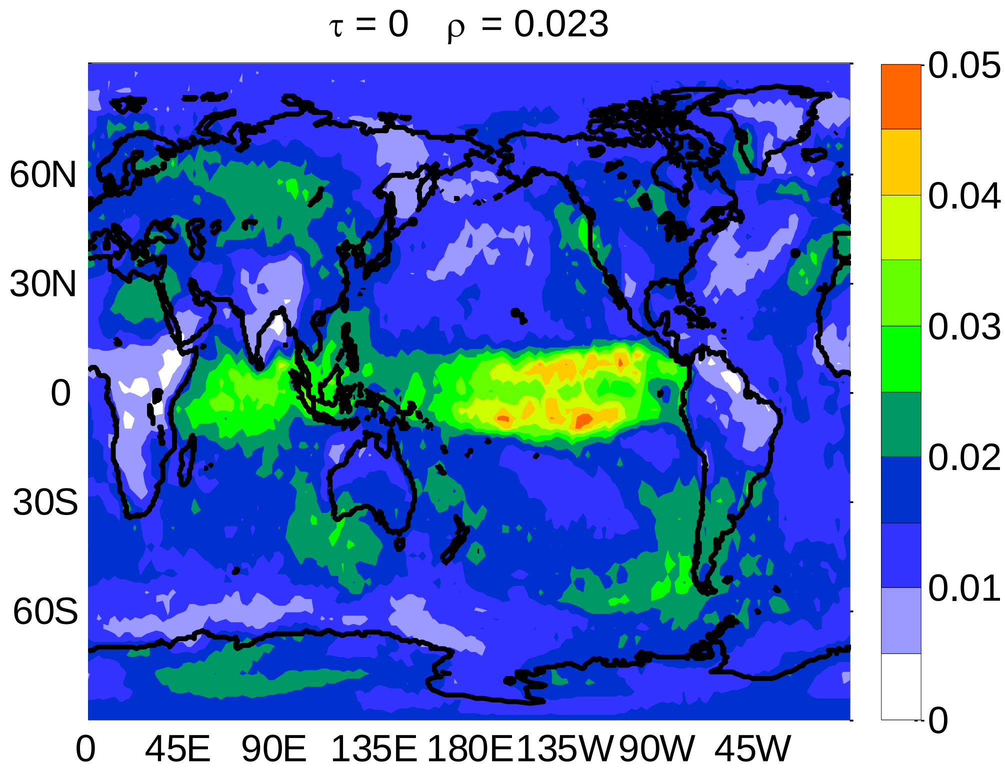




Colors code the mutual information

# OPs formed by concatenating four consecutive months

**D=4** $[x_i(t), x_i(t+1), x_i(t+2), x_i(t+3)]$

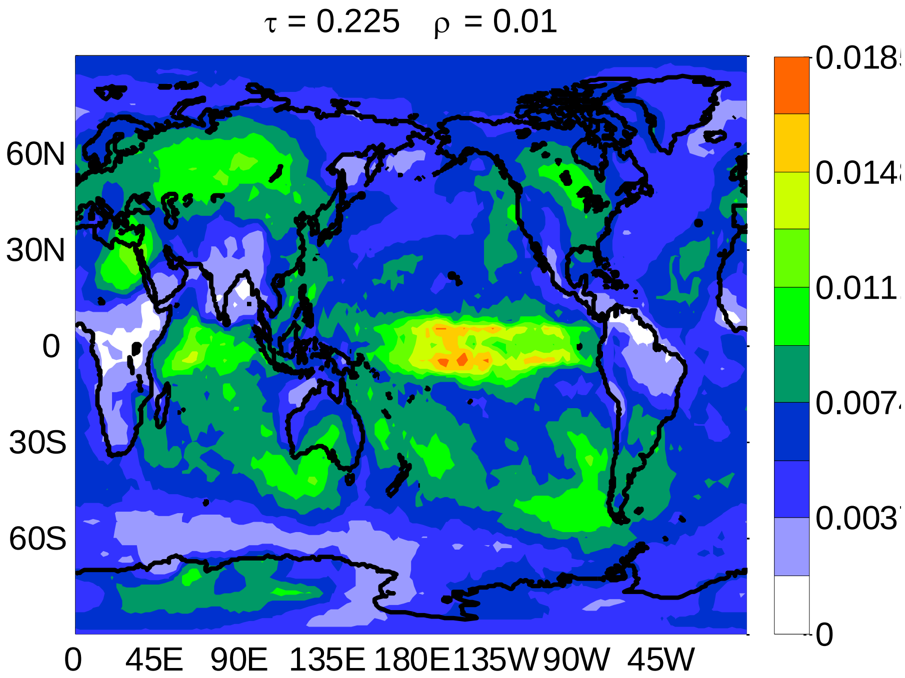
$\tau$=0 (all the ***significant* links**)

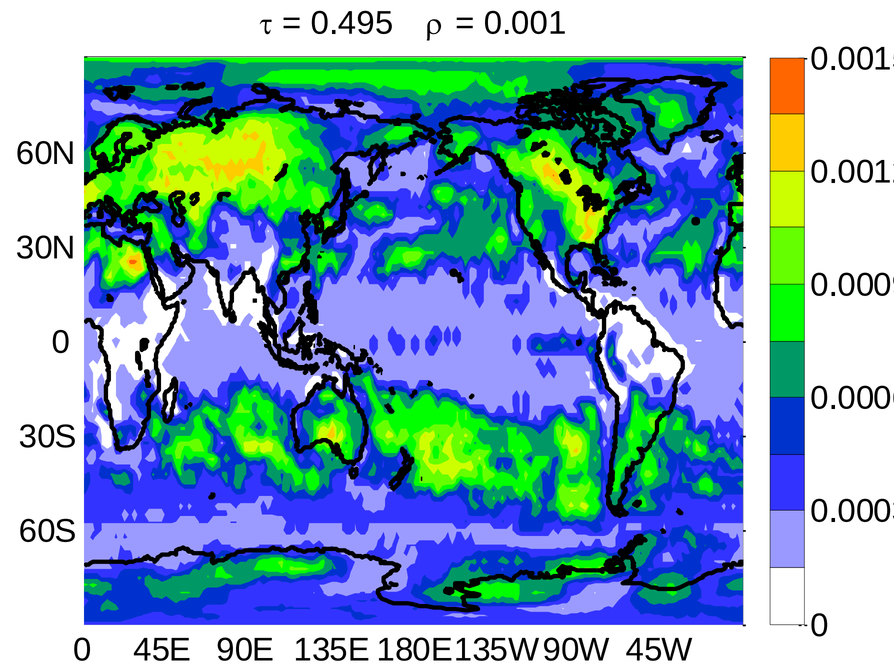



Colors code the Area Weighted Connectivity

# When increasing the threshold:

**D=4** $[x_i(t), x_i(t+1), x_i(t+2), x_i(t+3)]$

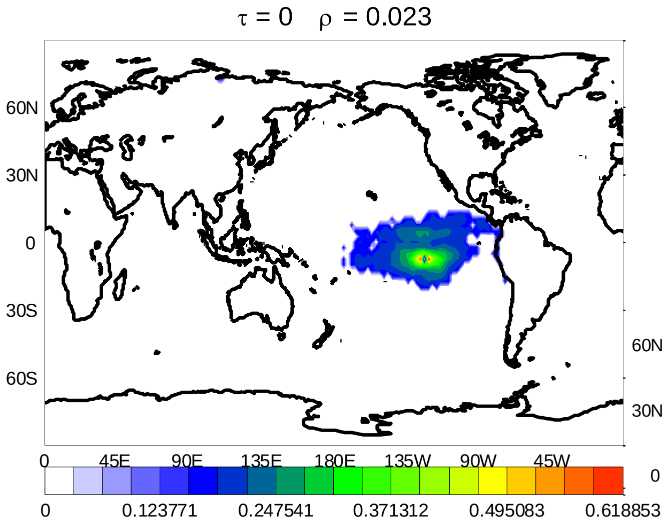


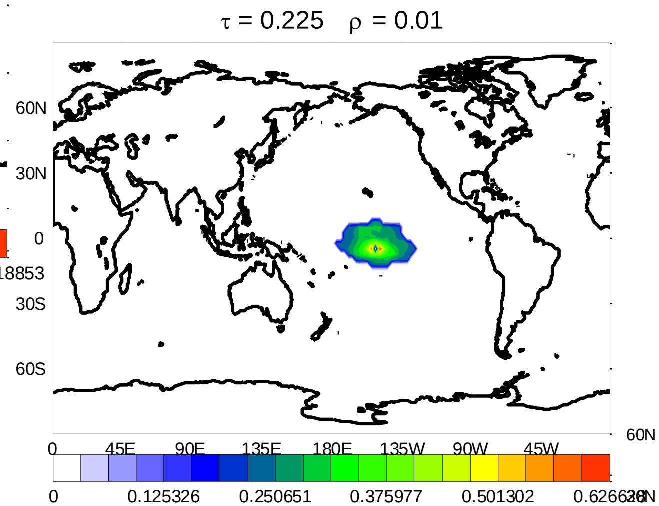



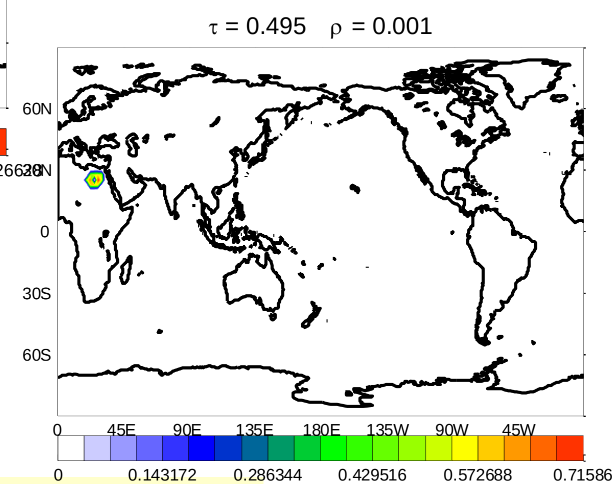

- 1% and 0.1% connectivity: different network structure

Colors code the area weighted connectivity

# Where is the "hub" and where is connected to?



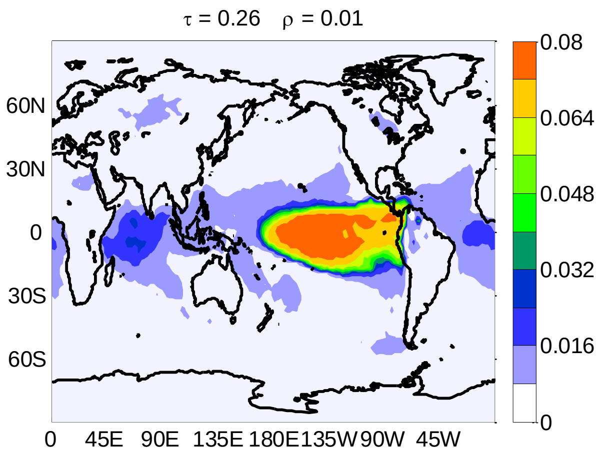



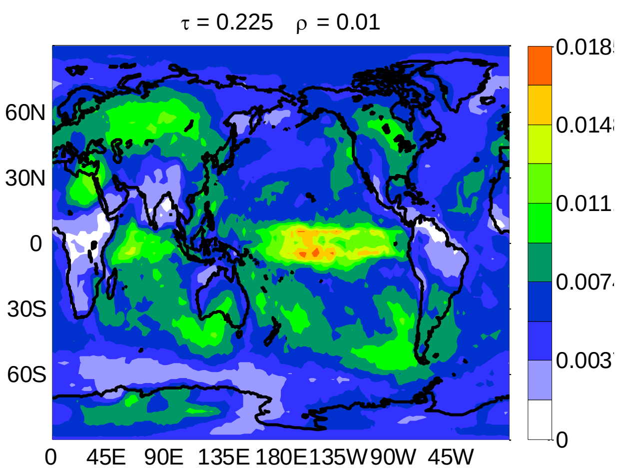




Colors code the mutual information

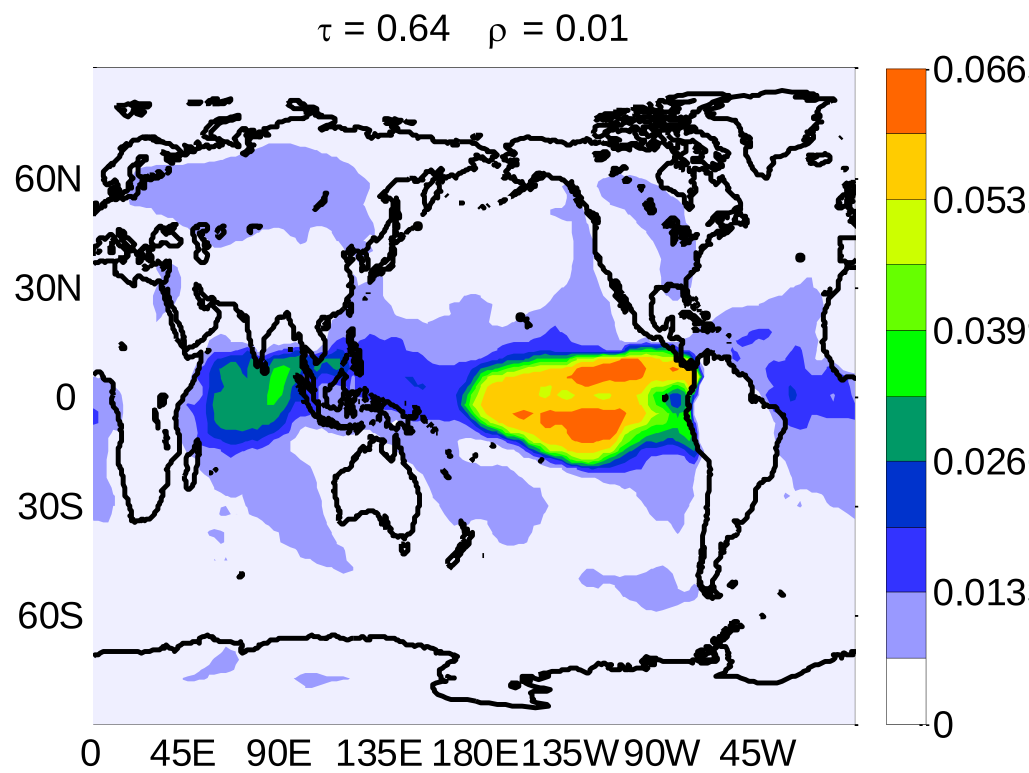
# Comparison (I/II)

## using the |cross-correlation| as a measure of statistical interdependency




$$C_{i\,j} = \frac{\sum_{t=1}^{N}(x_i(t)-\bar{x}_i)(x_j(t)-\bar{x}_j)}{\sigma(x_i)\sigma(x_j)}$$



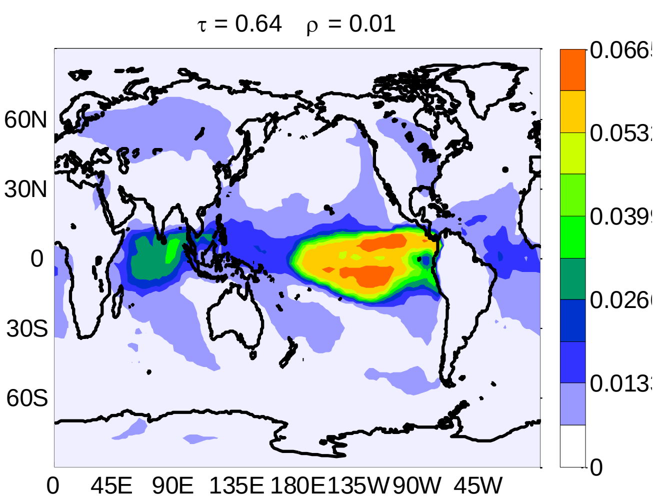

A. A. Tsonis and K. L: Swanson PRL 2008

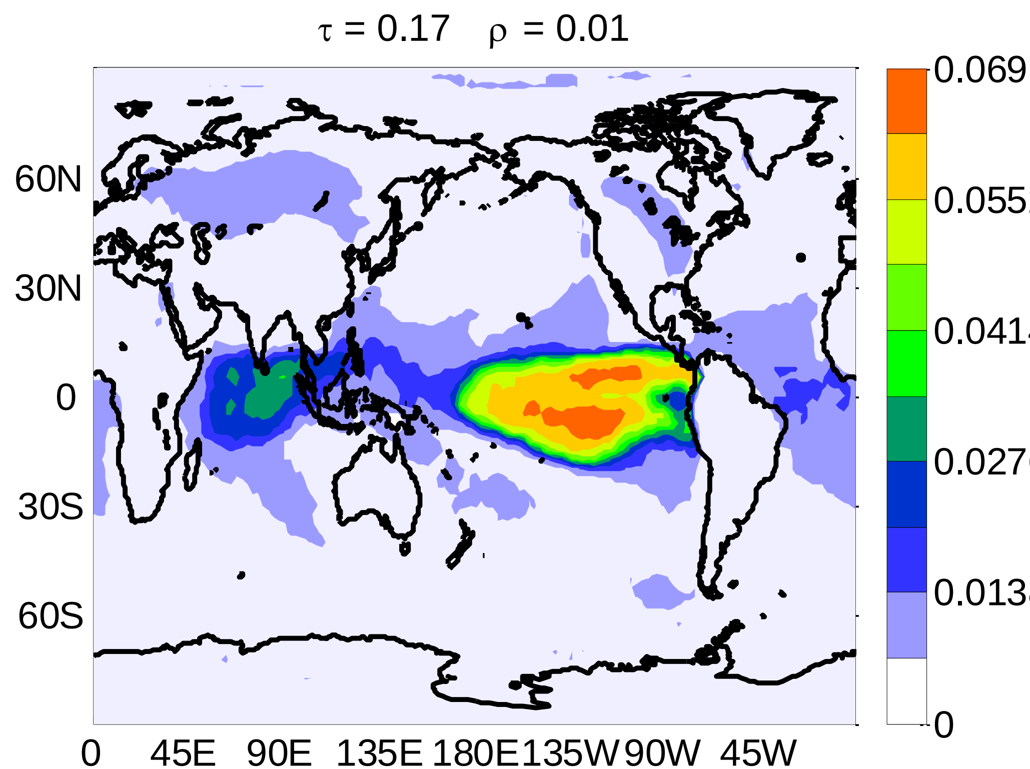
# Comparison (II/II)

## using the MI with PDFs calculated from histograms of SAT anomalies

$$C_{ij} = \frac{\sum_{t=1}^{N} (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)}{\sigma(x_i)\sigma(x_j)}$$

$$M_{ij} = \sum_{m,n=1}^{N_{bin}} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)\, p_j(n)}$$




J. F. Donges et al, Eur. Phys. J. Spe. Top. 2009, EPL 2009

# Summary and future work

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- We applied the method of **ordinal time-series analysis** to the study of monthly-averaged Surface Air Temperature anomalies.
- We found different network structures at different time scales (intra-season and inter-annual).
- The success of the methodology is based on **probability distribution functions (PDFs) that take into account the repeated patterns of oscillations** present in the Earth climate.
- Work in progress: detection of link directionality and inclusion of lag-times.

THANK YOU FOR YOUR ATTENTION