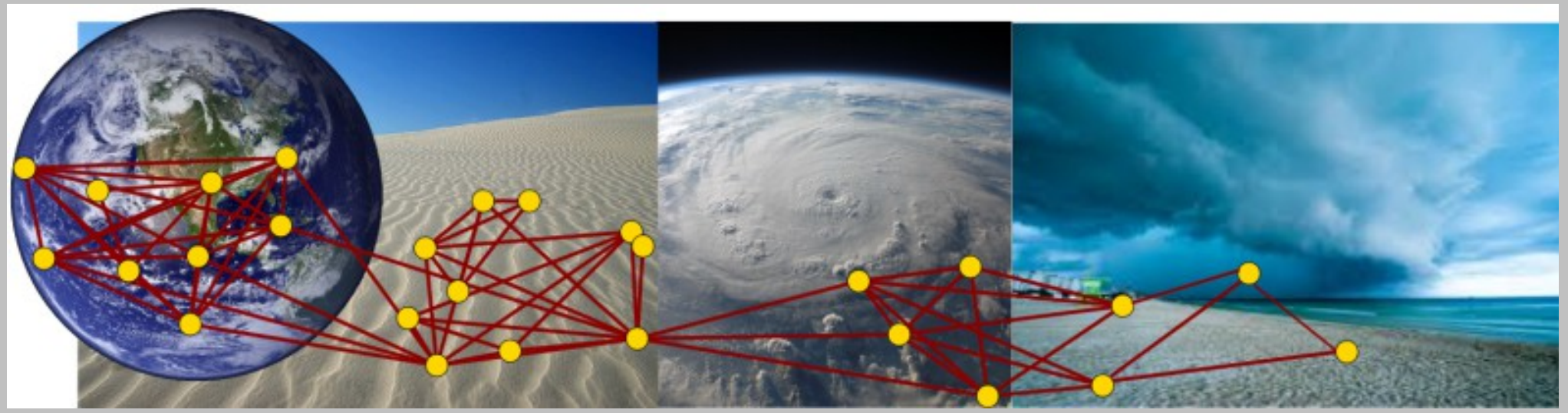
Interacting Networks in Climate



Marcelo Barreiro
Atmospheric Sciences Department
School of Sciences, Universidad de la República, Uruguay

Big Data & Environment, Buenos Aires 2015

FACULTAD DE CIENCIAS
UDELAR | fcien.edu.uy

LINC

Learning about Interacting Networks in Climate

6 academic partners
+ 3 companies

Objective:
Apply network methodologies to climate problems.

12 PhD students
3 posdocs

www.climatelinc.eu

Collaborators

Ignacio Deza

Giulio Tirabassi

Cristina Masoller

Universidad Politecnica de Cataluña, España

Universidad de la República
Uruguay

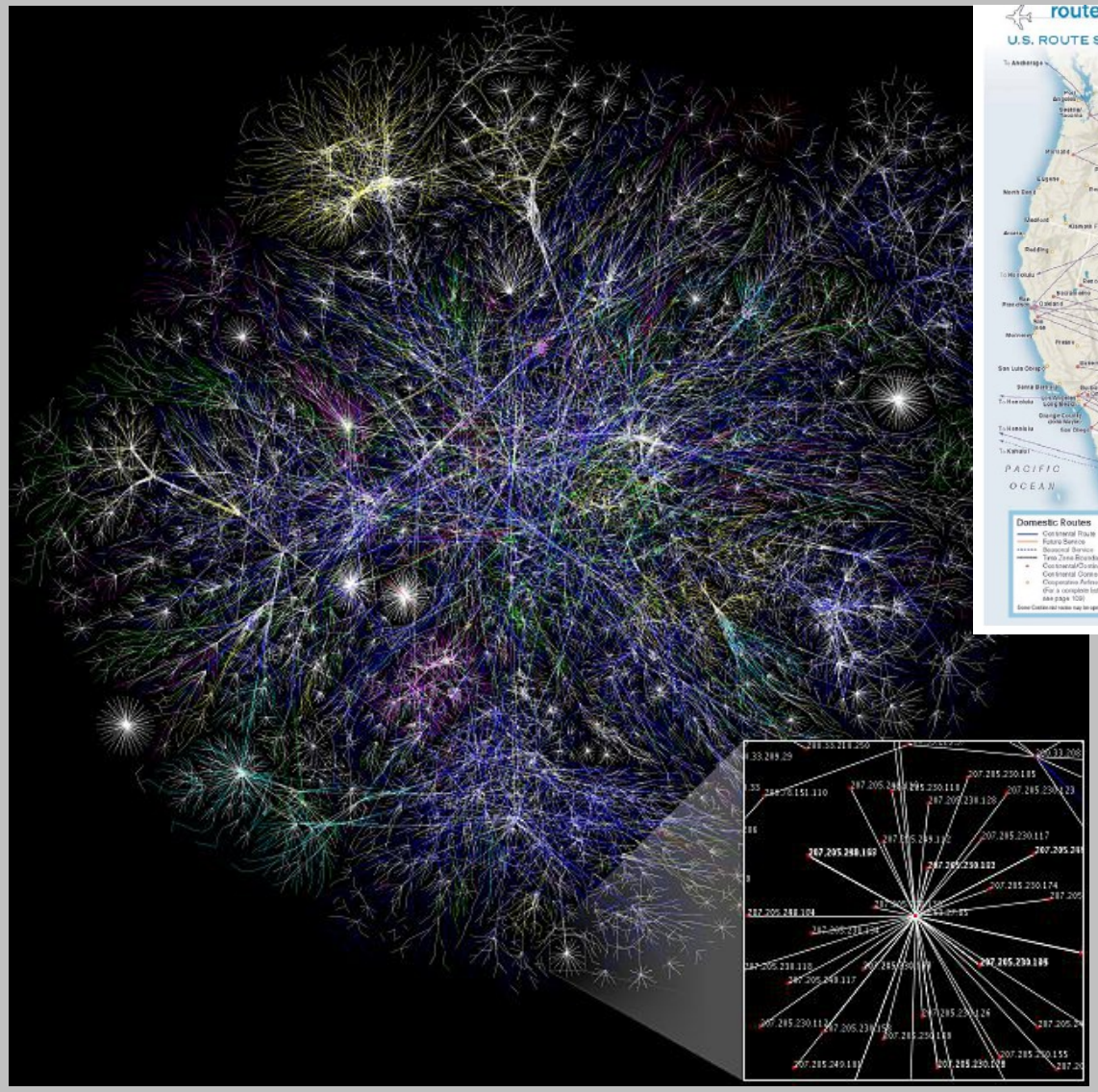
Fernando Arizmendi

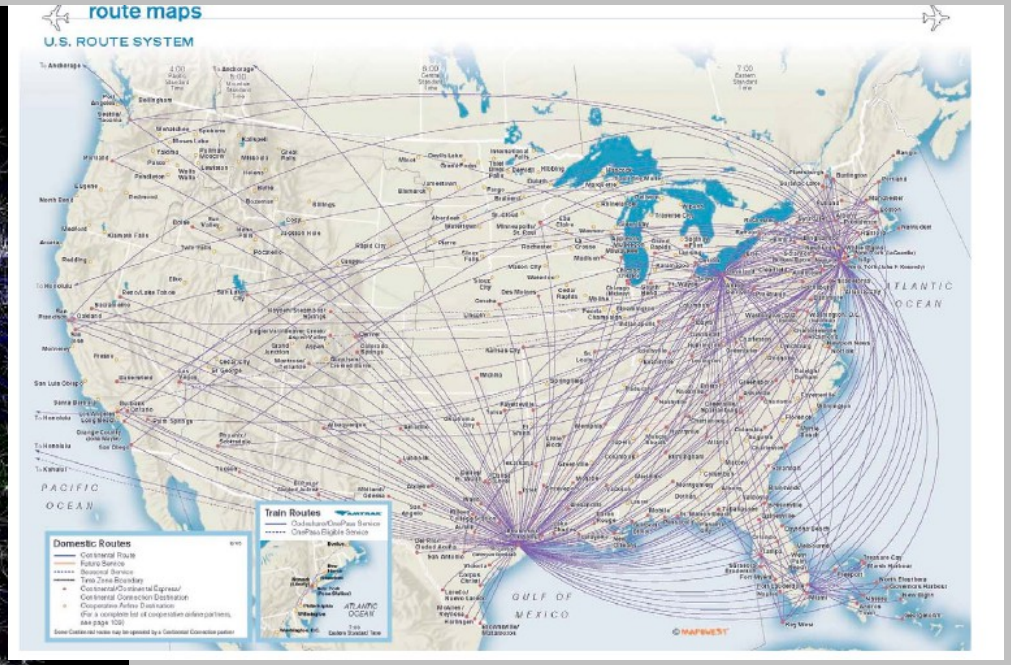
Veronica Martin

Arturo Martí

What are complex (structural) networks?

It is an interacting network of nodes (agents)





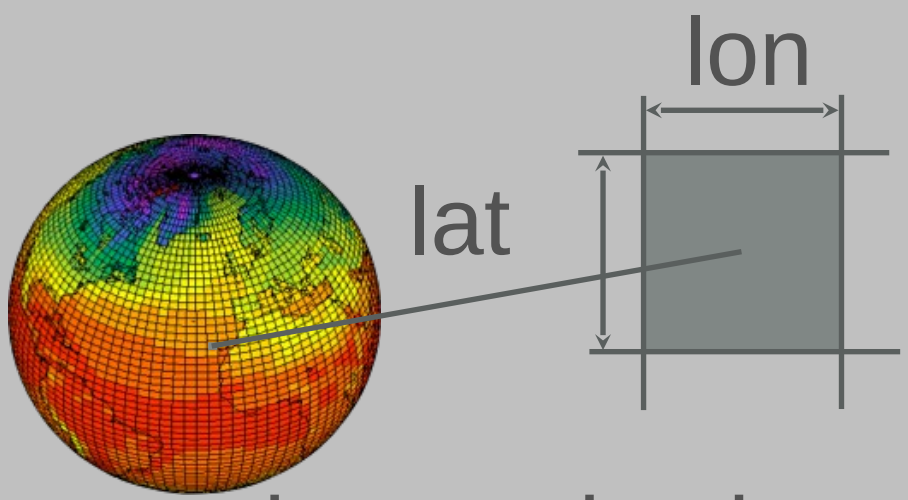
Map Continental Route
(Tsonis et al 2006)

The internet map (Opte Project)
Connection among IPs

Talk Layout

- Climate (Functional) Network Construction methodology: Similarity
  - Ordinal Patterns
  - Graphical representation
  - Connectivity Surface Air Temperature (SAT)
- Climate Network Construction methodology: Directionality
  - Directionality of SAT
- Summary

Climate networks



time series in each point

- Nodes – grid points $x_i(t)$
- Links – defined using similarity or directionality measures.

Matrices can be huge!

- Very efficient to characterize spatial patterns.
- Properties of temporal series are contained in network topology.
- Framework to validate climate models.

Previous Studies

- A. Tsonis et al 2000s → : focus on interaction among climate variability nodes (ENSO, NAO, PDO, etc) and study abrupt climate changes.
- J. Kurths et al mid-2000 → : global climate network construction from grid points and study several network measures. Focus on extremes.
- S. Havlin et al mid-2000s → : study El Niño properties and evolution using complex networks.

FACULTAD DE CIENCIAS
UDELAR | fcien.edu.uy

Methodology for climate network construction

Statistical Similarity

Directionality

Statistical similarity between time series

Person Correlation (linear)

$$C_{ij} = |\sum_{t=0}^{N} \frac{(x_i(t) - \overline{x}_i)(x_j(t) - \overline{x}_j)}{\sigma_i \sigma_j}|$$

Mutual Information (nonlinear)

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

Measures how much information about x_i we get by knowing the evolution of x_j. $P_{i,j}$ are pdfs.

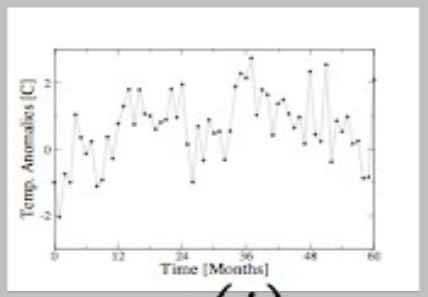
Other possible measures: transfer entropy (Runge et al 2012) or event synchronization (Malik et al 2012).

Mutual information: histograms

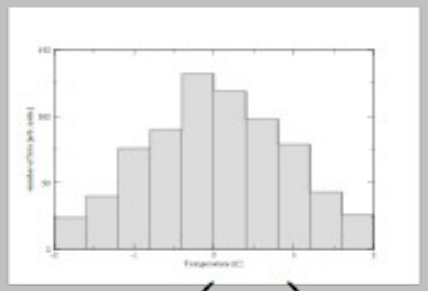
From temporal time series in i & j we can approximate their pdfs.

And their joint distribution.

Barcelona

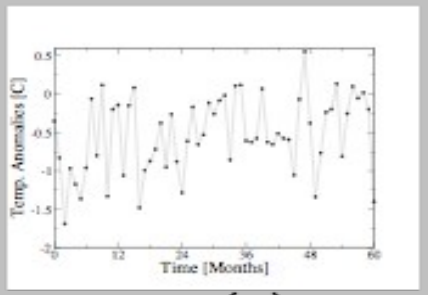


$x_i(t)$

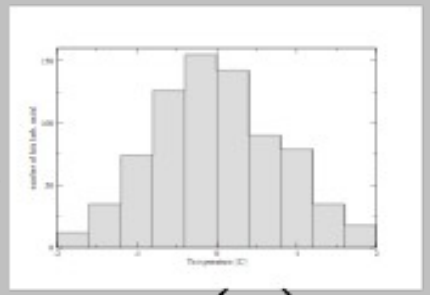


$p_i(m)$

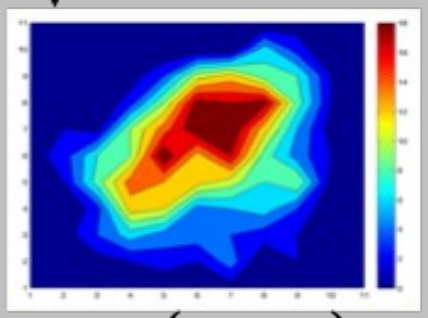
Montevideo



$x_j(t)$



$p_j(n)$



$p_{ij}(m,n)$

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

Adjacency Matrix

$$A_{ij} = \Theta(|S_{ij}| - T) - \delta_{ij}$$

Θ – heavyside function

S_{ij} – similarity measure (Mutual Information)

T – significance threshold: global or for each pair depending on significance criterion. $T \rightarrow 0$ fully connected network / $T \rightarrow 1$ network without links.

<u>The Adjacency matrix defines the network associated to a particular field.</u>

Graphical Representation

1. Local: Connections from or to a node (X).

2. Global: Area Weighted Connectivity (AWC):

$$AWC_i = \sum_j^N A_{ij} \cos(\lambda_i) / \sum_j^N \cos(\lambda_j)$$

A_{ij} = adjacency matrix

Area to which a node is connected. Maxima in AWC are called supernodes or hubs.

Other measures: e.g. closeness centrality (inverse of mean network distance of node I to all other nodes via shortest paths)

Par@graph - a parallel toolbox for the construction and analysis of large complex climate networks. Ihshaish et al 2015 (Geosci. Mod. Dev.)

Scale separation: Ordinal Patterns

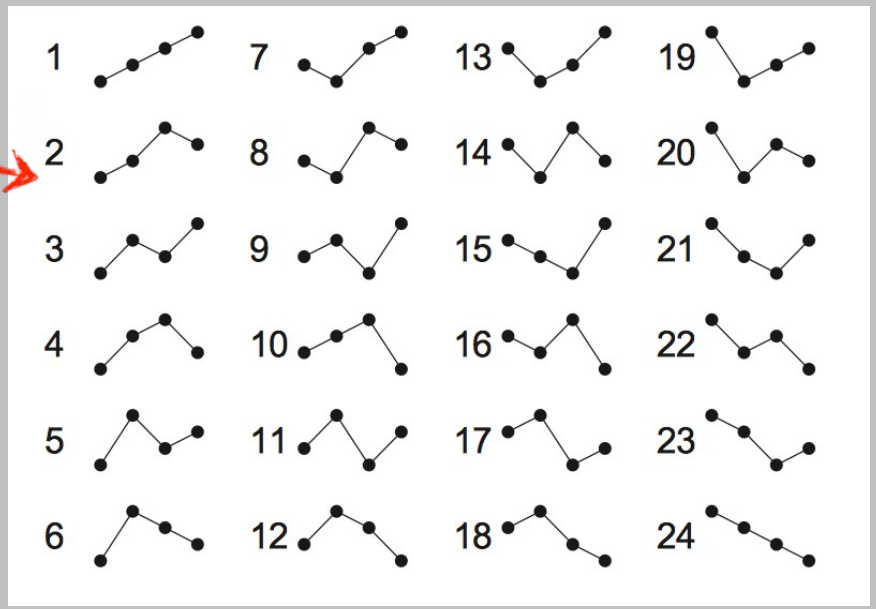
- Consider a time series

$$x_1, x_2, x_3, x_4, \ldots, x_n$$

- Ordinal Patterns (OP) are defined as a way to order the elements of the time series

OP word size 3.

OP word size 4.



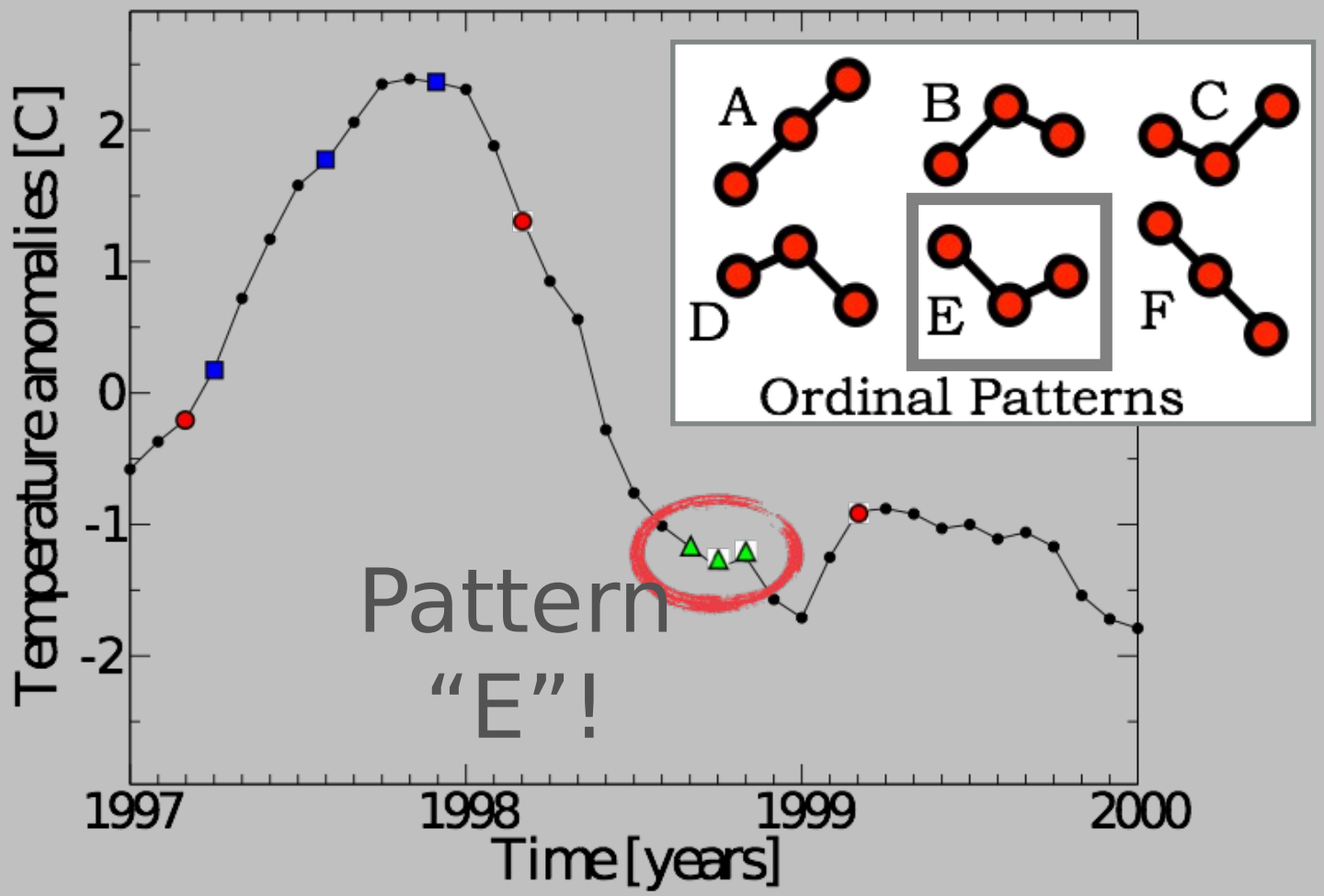
The length of the time series limits the size of the word

OP larger: more resolution

pdfs are calculated counting the number of times that each word appears in the time series.

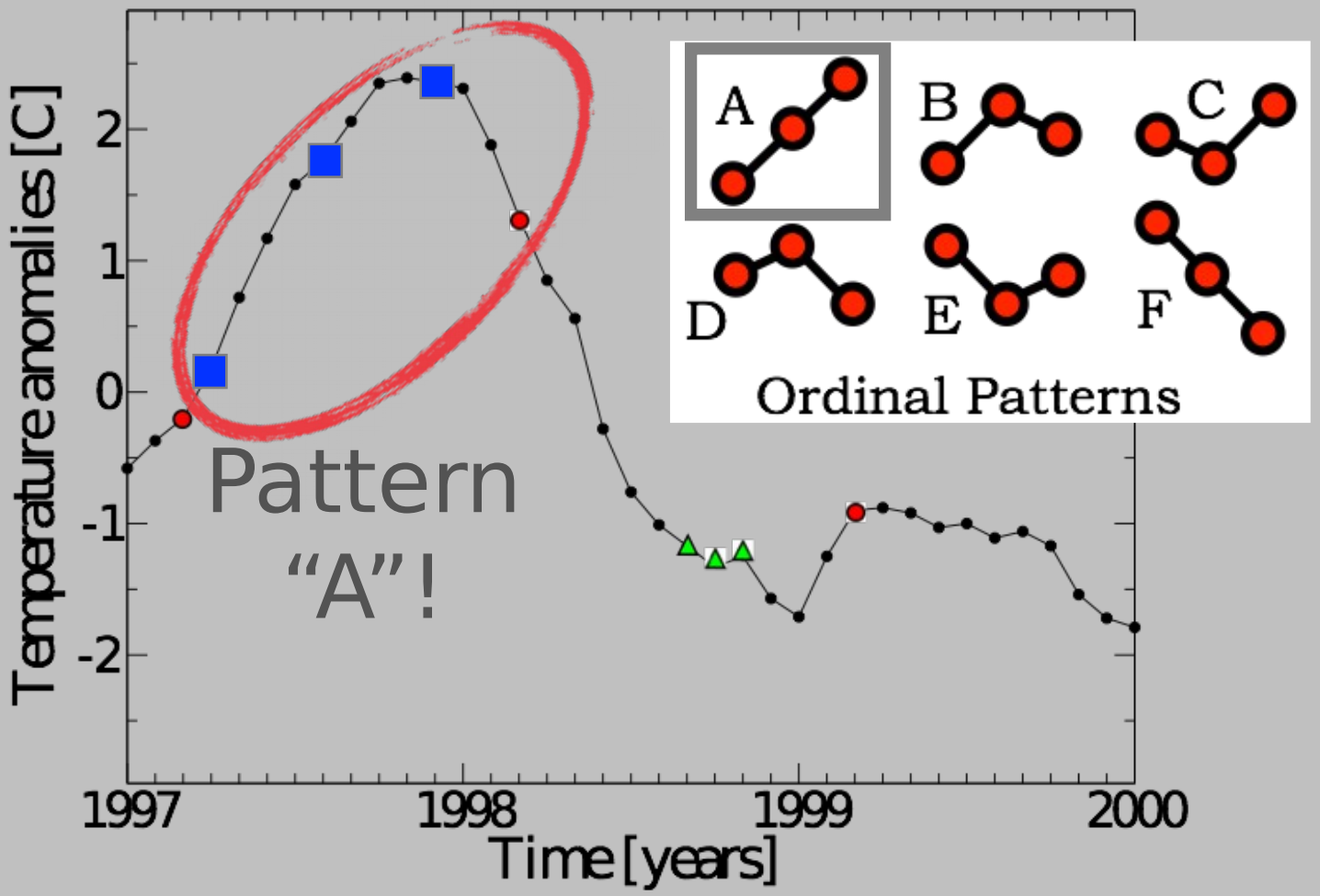
Mutual Information: Ordinal Patterns

- A new series of patterns is constructed



Example for a 3 letter word

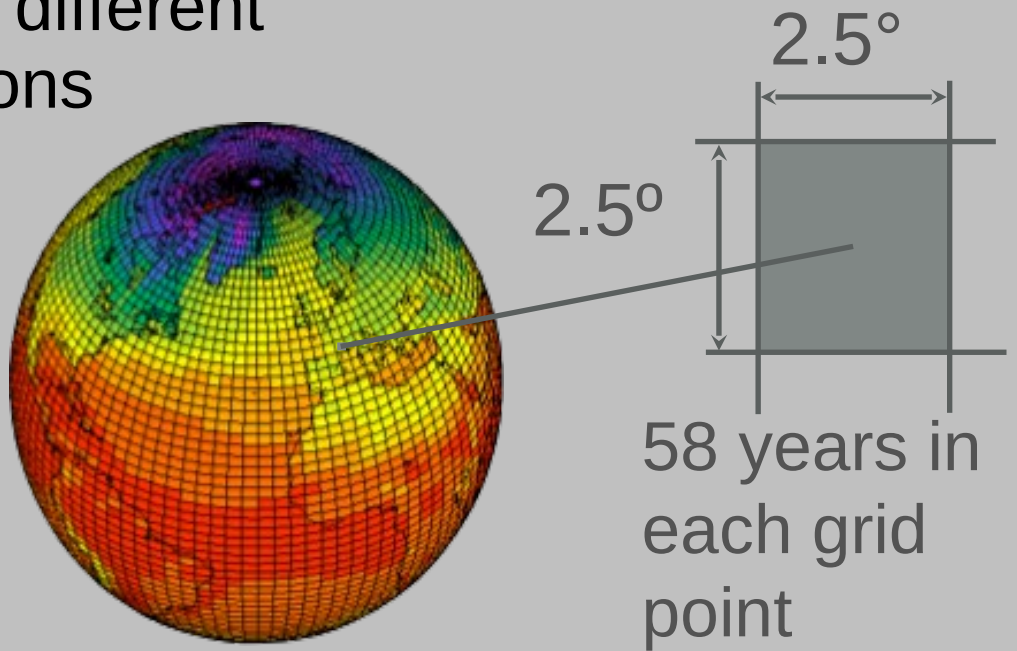
- Points not need to be contiguous. We can choose time scale of interest.



Connectivity Surface Air Temperature

Objetive

- Study connectivity of SAT on different time scales, i.e. teleconnections



Data

- SAT NCEP-CDAS 1
- ~10.000 nodes
- Monthly mean 1948-2006.

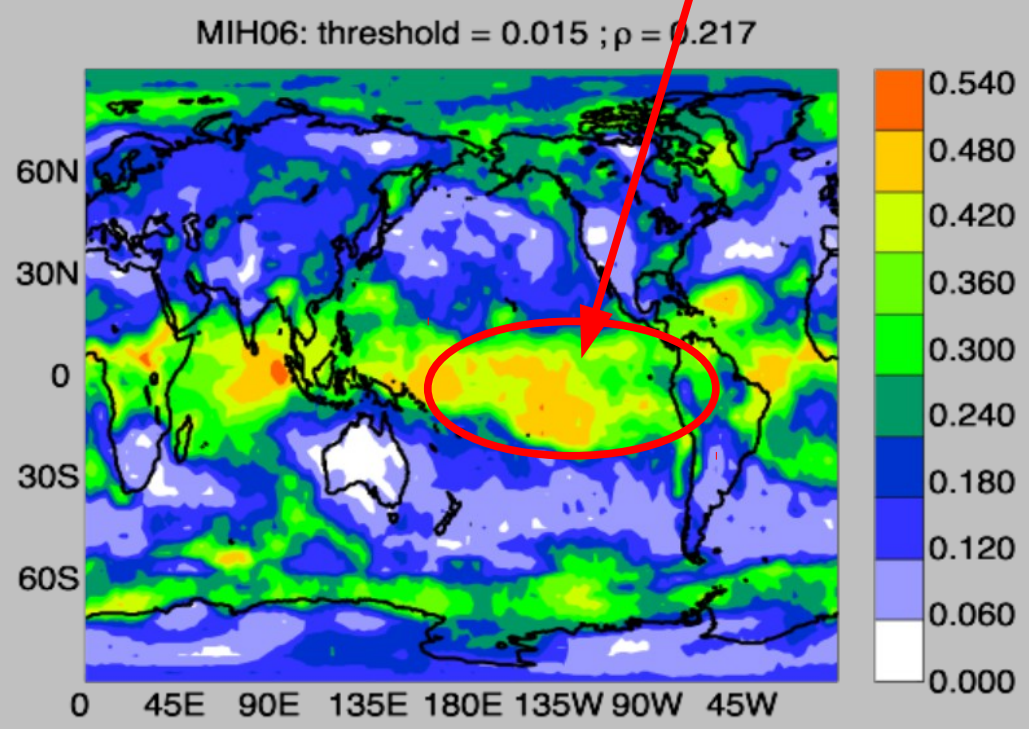
Barreiro et al 2011, Deza et al 2013, Arizmendi et al 2013

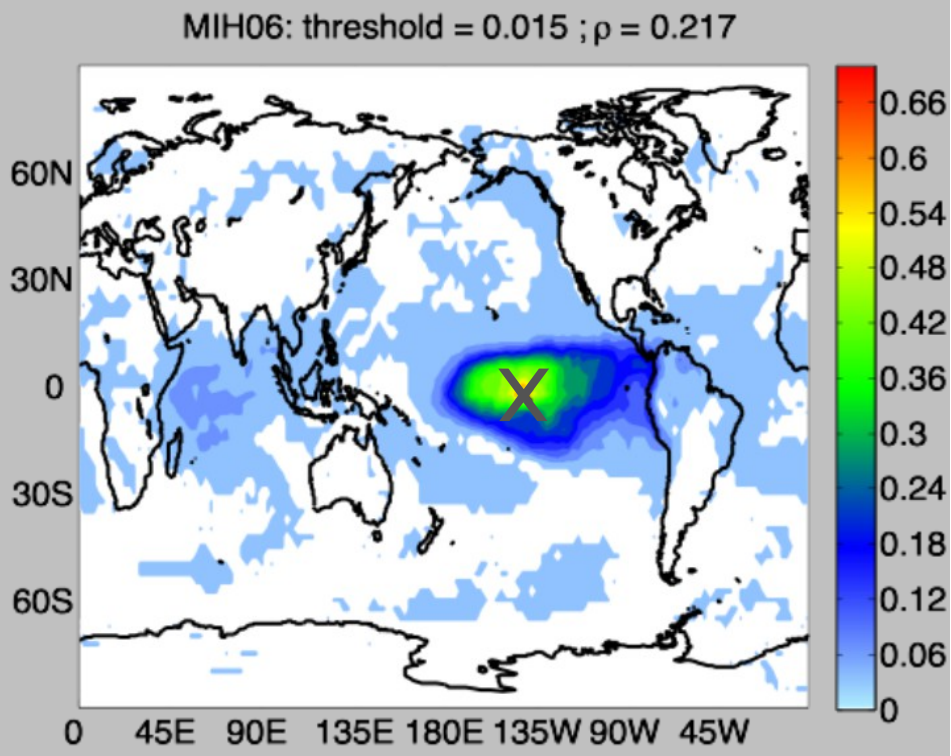
SAT – Mutual Information

AWC

Area Weighted Connectivity

Connectivity of a point





This network contains information on all time scales.

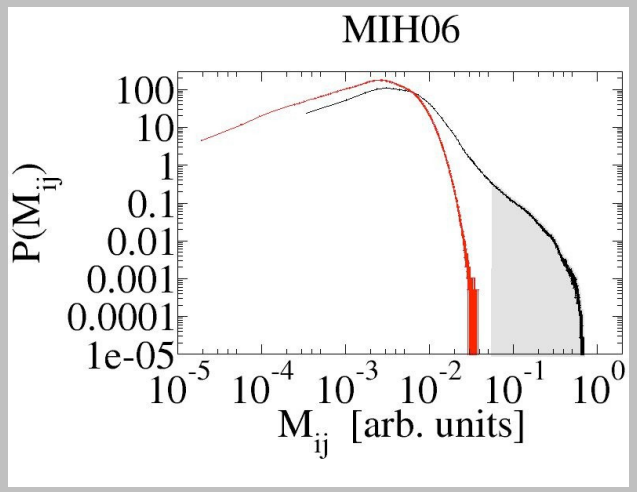
Only significant links are considered.

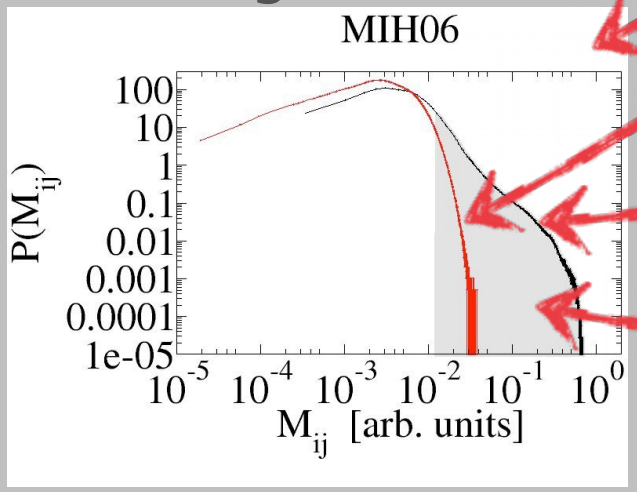
Statistical Significance of MI

Accepted area for fixed density criterion (~2.7% in this case).

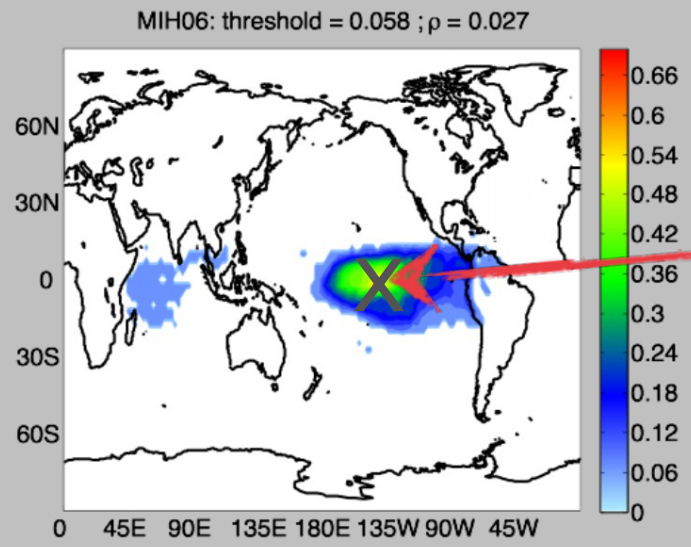
Accepted area for significance μ+3σ bootstrap surrogates (~21%)

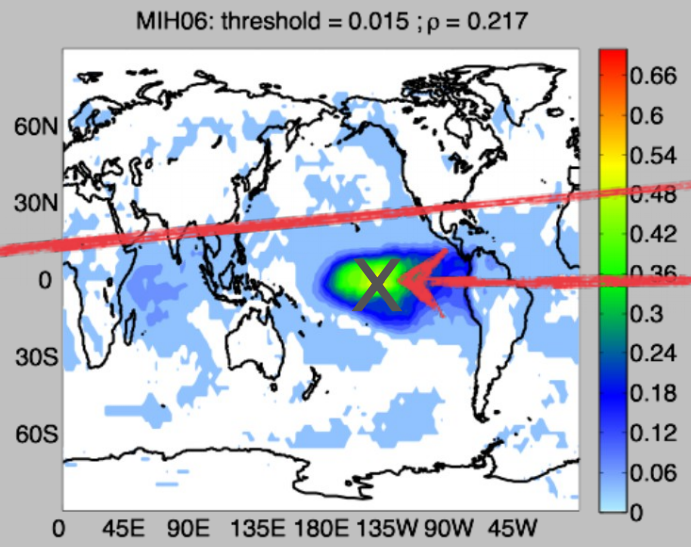
Number of bins for pdf: 6.





- Surrogate data
- Original time series
- Acepted links

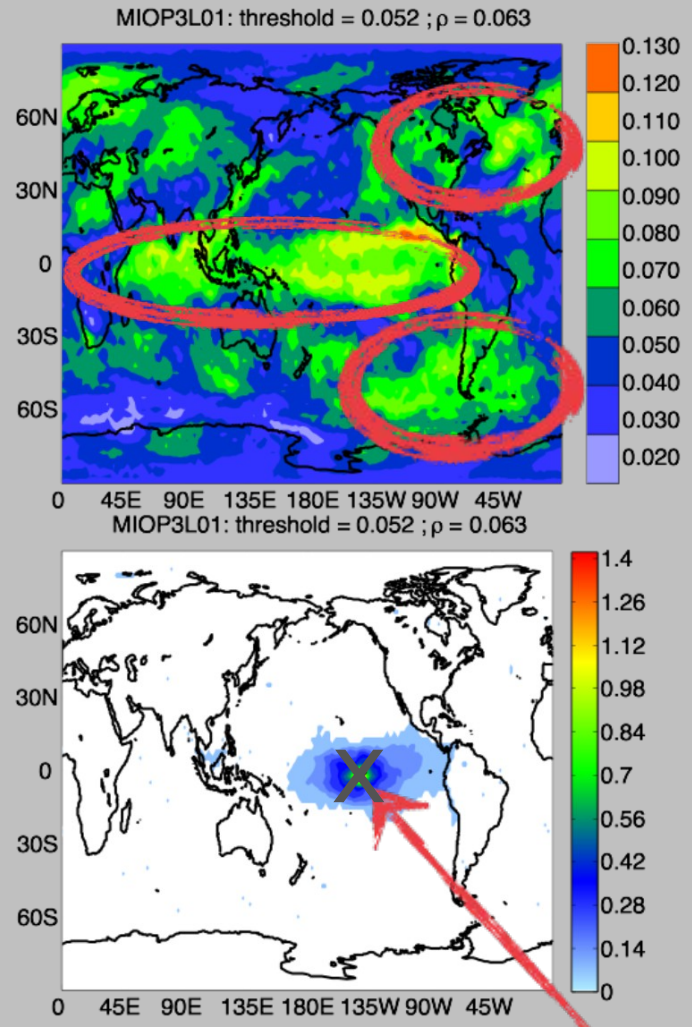




Connections to or from this point depend strongly on threshold.

Ordinal Patterns & networks on different time scales

AWC intra-seasonal



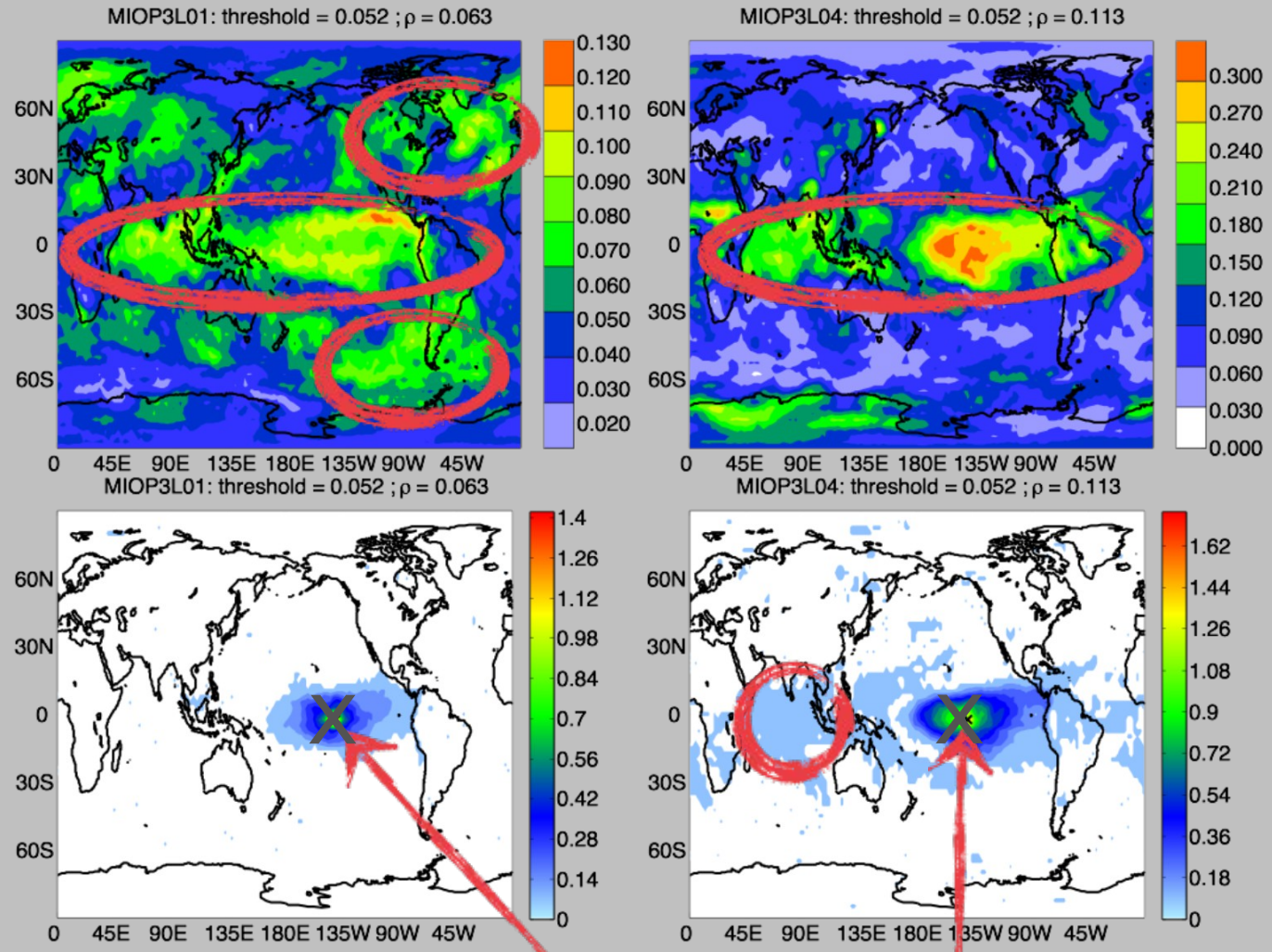
3 consecutive times are used to construct OP of 3 letters.

MI

Connections to/from this point

FACULTAD DE CIENCIAS
UDELAR | fcien.edu.uy

AWC intra-seasonal intra-annual



3 times separated by 4 months are used to construct OP of 3 letters.

MI

Connections to/from this point

FACULTAD DE CIENCIAS
UDELAR | fcien.edu.uy

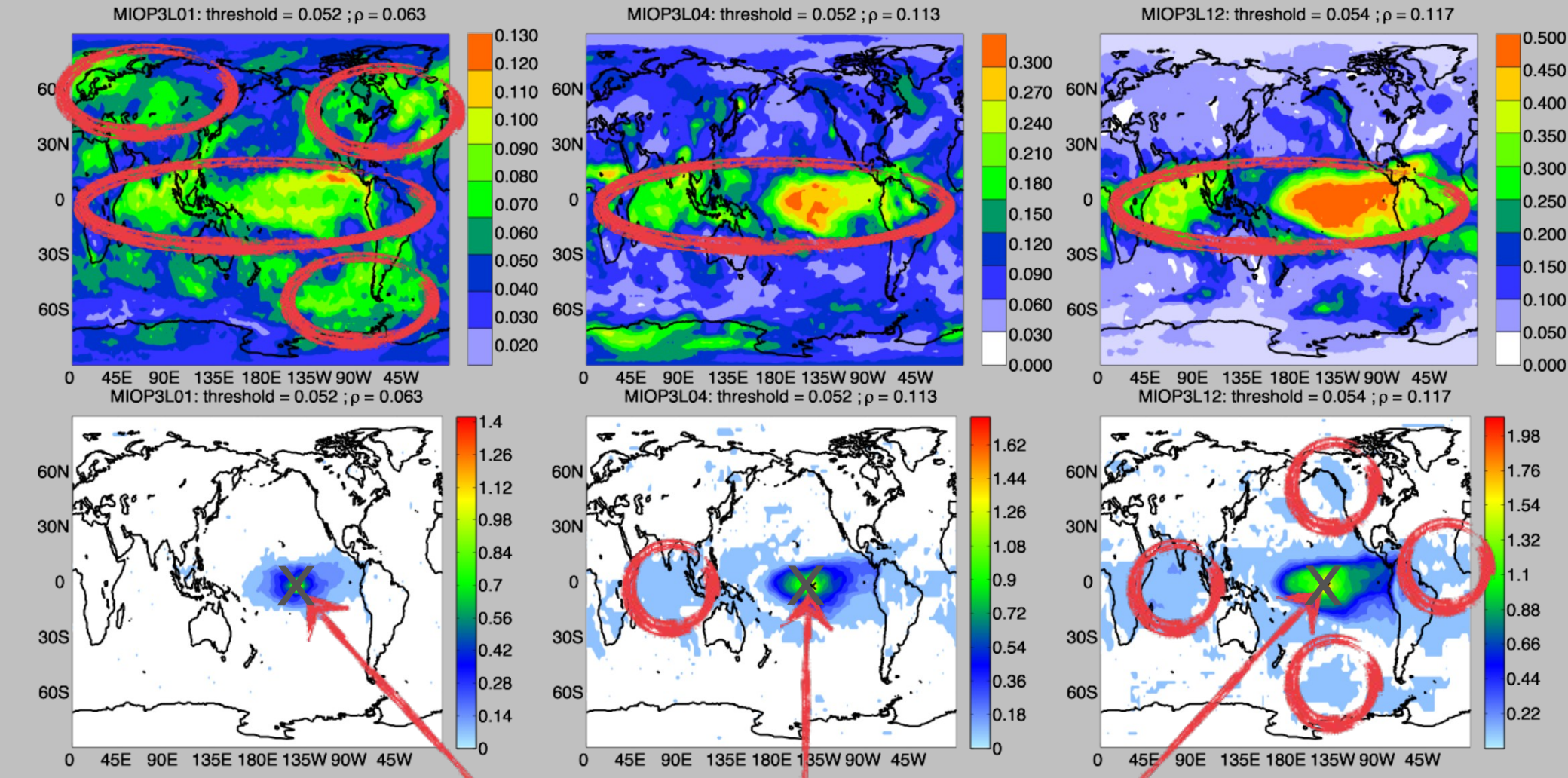
3 times separated by 12 months are used to construct OP.

AWC

intra-seasonal

intra-annual

inter-annual



MI

Connections to/from this point

FACULTAD DE CIENCIAS
UDELAR | fcien.edu.uy

Methodology for climate network construction

Statistical Similarity

Directionality

Directionality index

- The directionality index can be defined as

$$\mathrm{DI}_{XY}(\tau) = \frac{I_{XY}(\tau) - I_{YX}(\tau)}{I_{XY}(\tau) + I_{YX}(\tau)}$$

Palus (2007)

- where I_{XY} is the Conditional Mutual Information

$I_{XY}(\tau) = I(X; Y|X_\tau)$ quantifies the transfer of info from X to Y: quantity of info shared between X(t) and Y(t) given the influence of X(t-tau) on Y(t).

- $I_{YX}(\tau) = I(Y; X|Y_\tau)$ quantifies the transfer of info from Y to X

- $\tau > 0$ is a parameter chosen to determine the time scale of interest

D_{XY} determines the net direction of information flow.

Other measures like Granger Causality can also be used (Tirabassi et al 2014)

Directed Network of surface air temperature

Objetive

- Study the directionality of SAT using DI

Data

- Daily mean data SAT. NCEP-CDAS1 Reanalysis, 2.5x2.5, 1948-2013.

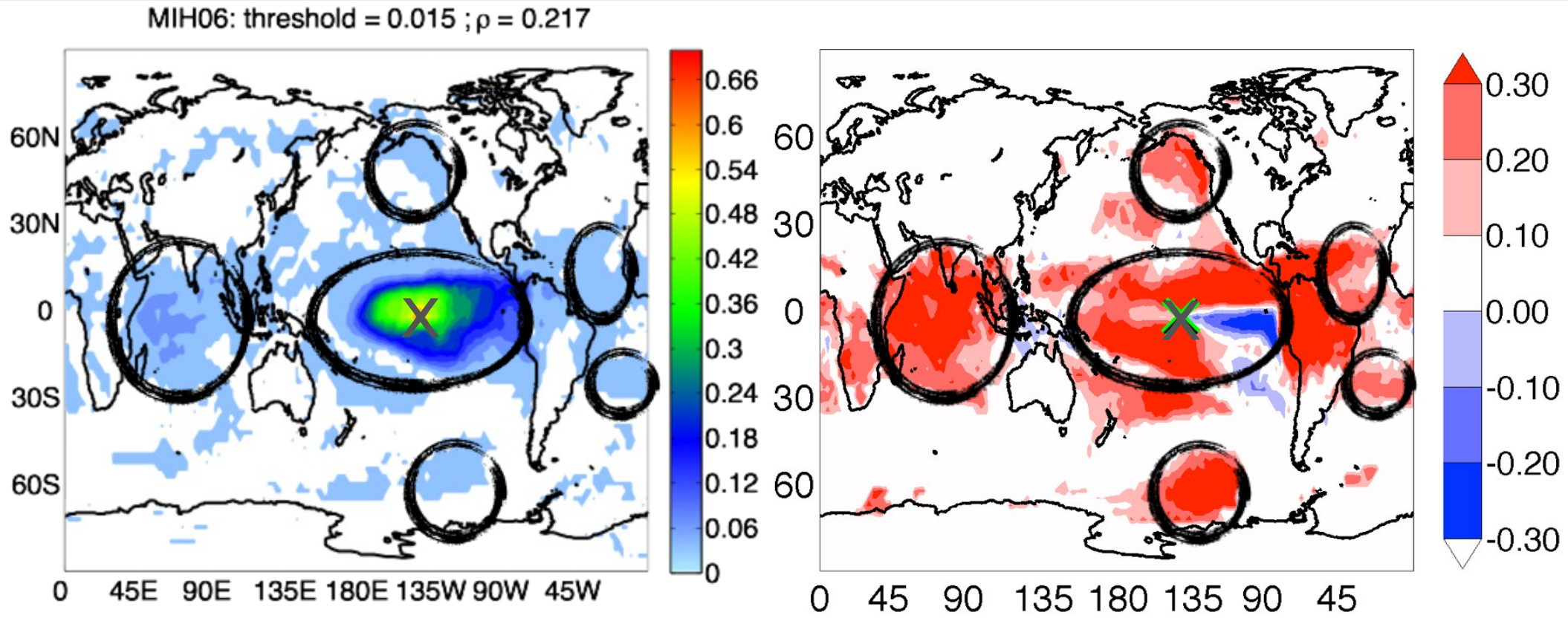
Deza et al (2015)

Tropics – Pacific ocean

Only MI

Direction

$\tau = 30$ dias



MIH

DIH

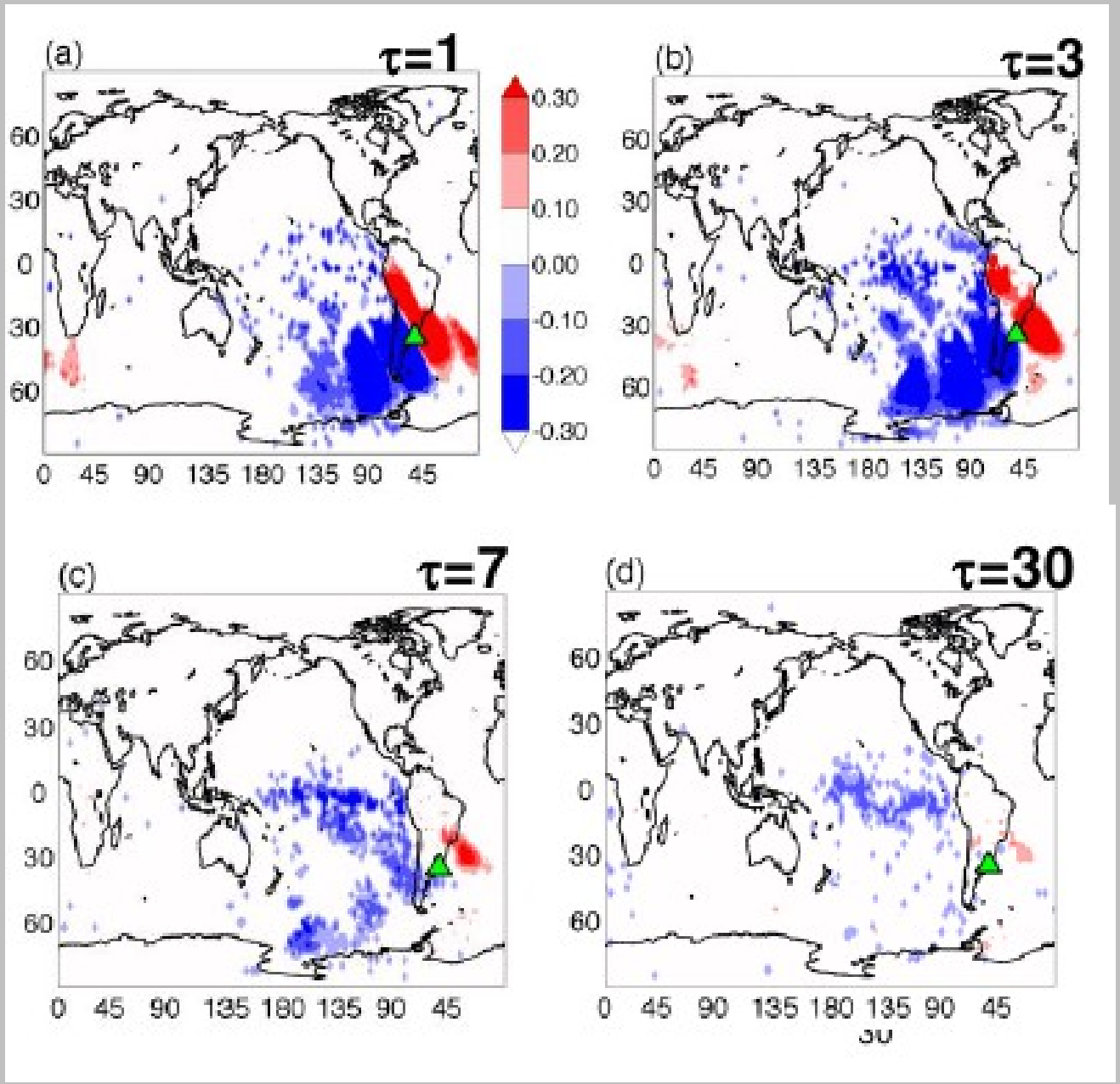
- BLUE incoming links
- RED outgoing links

Extra-tropics

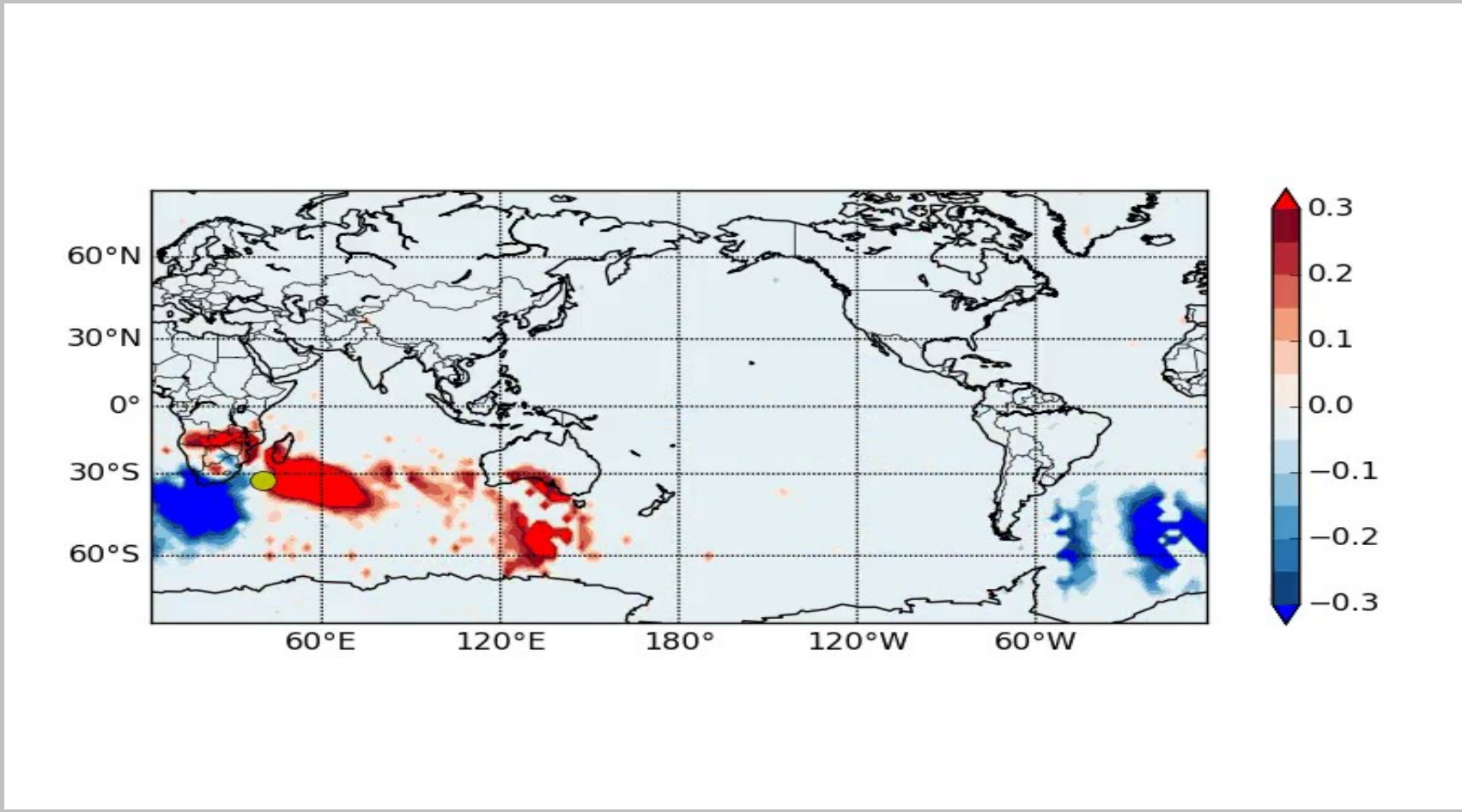
Directionality of a point over Uruguay:

Tau days: shows propagation of synoptic waves & Pacific influence.

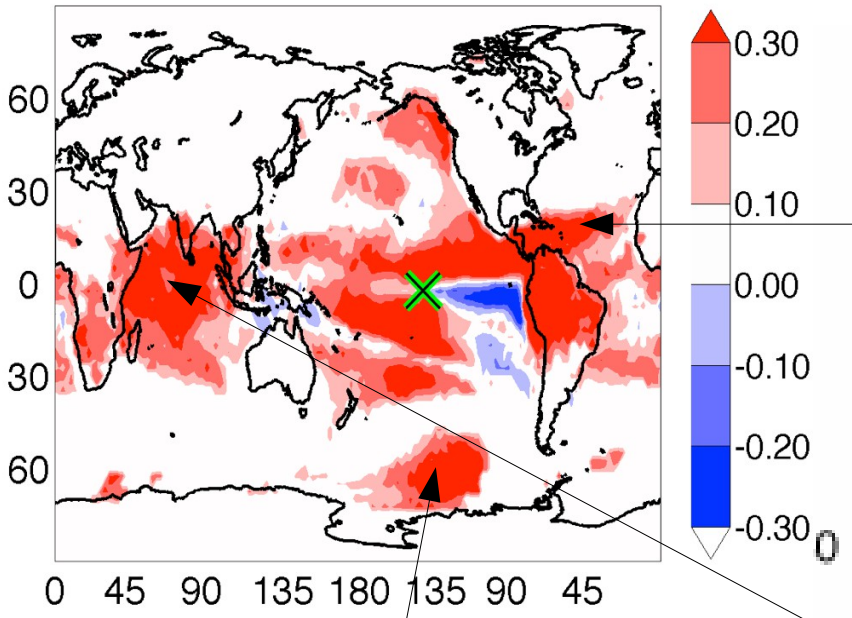
Tau monthly: only remains Pacific influence

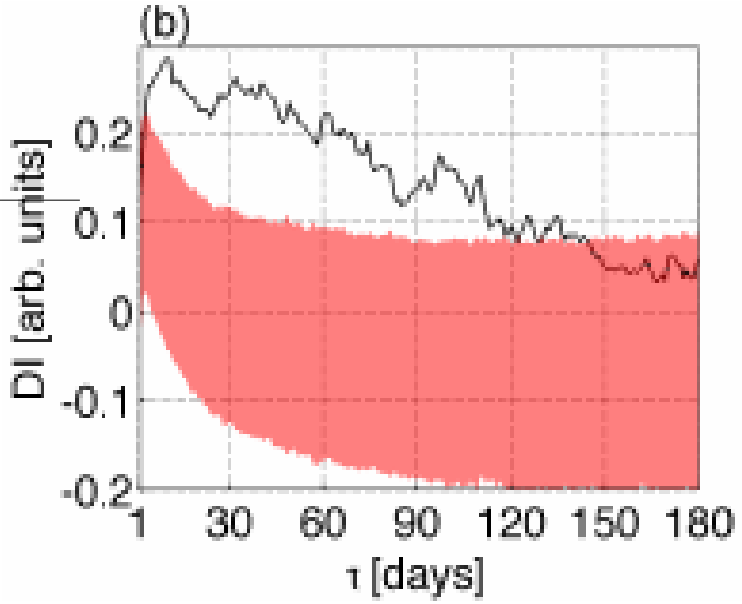


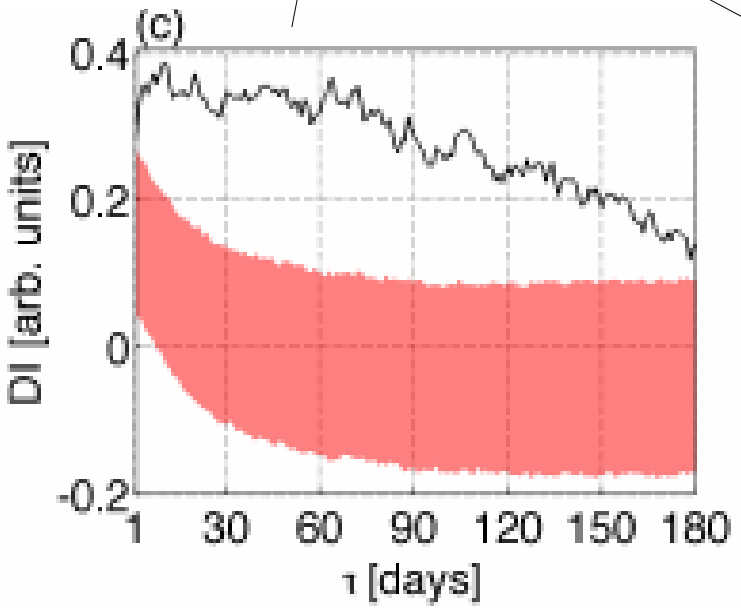
Movie – tau=3 days for different points in 30S.

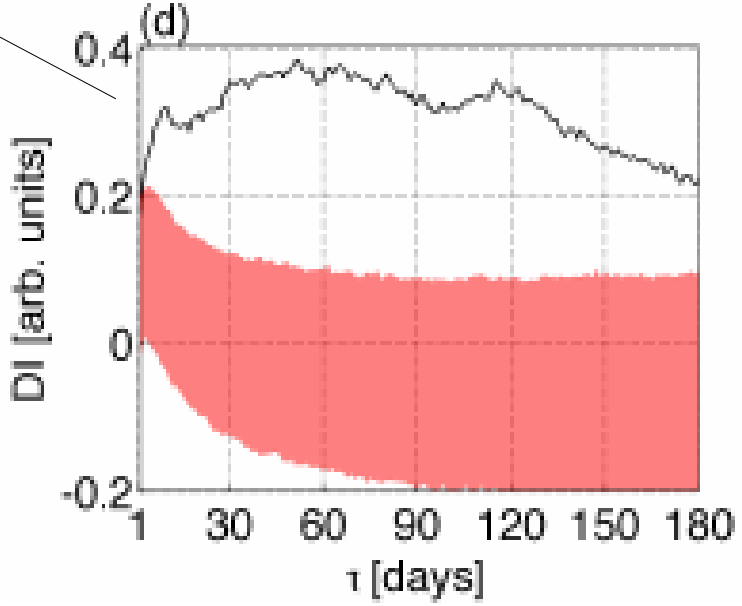


Dependence of DI on τ – shows time scale of remote connections.









Fast Time Scale ~ 10 days + a longer time scale that depends on air-sea Interaction in each basin

Summary

- Tools from complex networks and information theory provide a new methodology for analysis of (big) climate data.
- Likely adds to standard EOF analysis when there is no dominant mode of variability and the field shows spatial coherence (Dongues et al 2015)
- Multivariate (interacting) networks imply computation of very large matrices. Need to implement codes to construct such networks.
- Project LINC: www.climatelinc.eu
  - Community detection, ENSO prediction, rainfall extremes, interaction among climate modes, etc...