EGU General Assembly 2023

Global wildfire synchronicity patterns as revealed by complex network analysis

Catharina Elisabeth Graafland[1], Andrina Gincheva[2], Marco Turco[2] and Joaquín Bedia[3,4]*

[1]Instituto de Física de Cantabria (IFCA), CSIC–Universidad de Cantabria, Santander, Spain
[2]Regional Atmospheric Modelling Group, University of Murcia, Murcia, Spain
[3]Departamento de Matemática Aplicada y Ciencias de la Computación, Universidad de Cantabria, Santander, Spain
[4]Grupo de Meteorología y Computación,Universidad de Cantabria, Unidad Asociada al CSIC, Santander, Spain
(*Presenter, bediaj@unican.es)

Abstract EGU23-9608

OVERVIEW

1
Introduction
Global wildfires as a complex system

2
Data Acquisition and processing
Burned area data

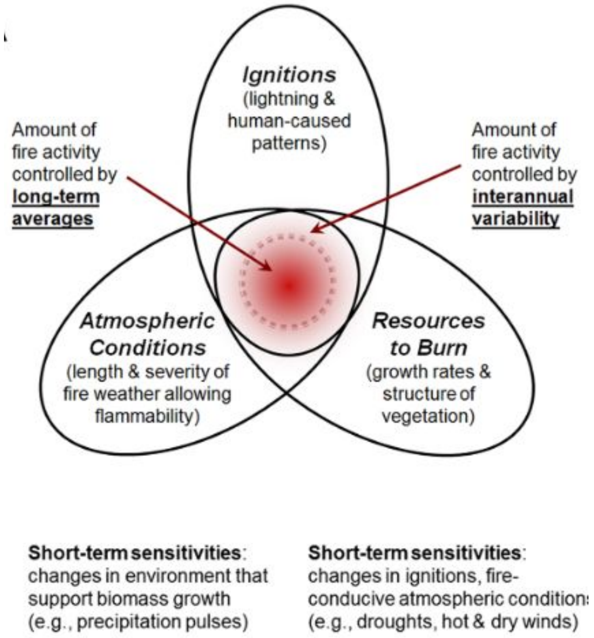
3
Methods
Overview of correlation and bayesian networks

4
Main results
Method fitting and results

5
Conclusions
Assessment of the methodology and main findings

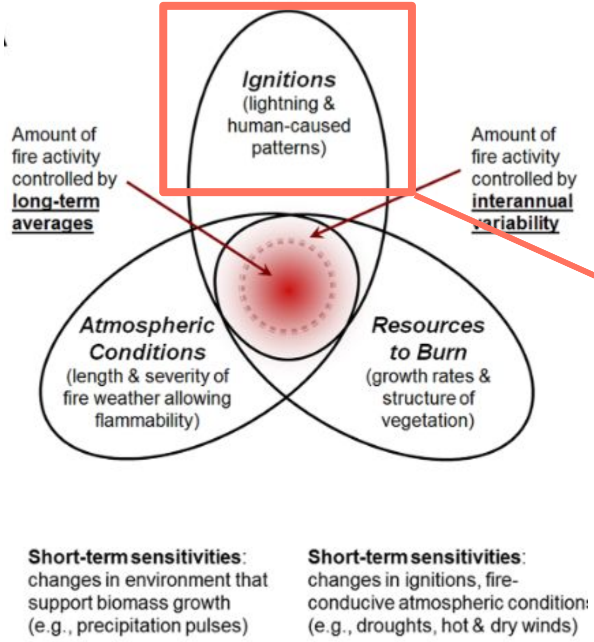
1

INTRODUCTION



Fire activity on a global scale is a complex system, determined by multiple factors operating simultaneously at different spatial and temporal scales:

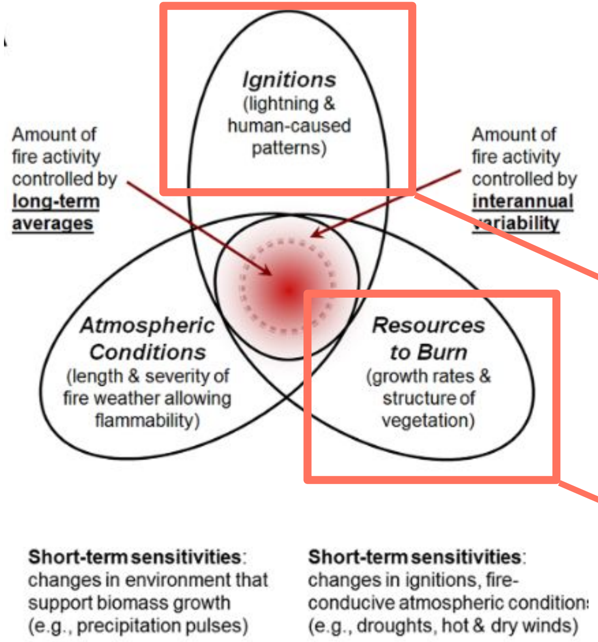
FROM: Krawchuk, M.A., Moritz, M.A., 2010. Constraints on global fire activity vary across a resource gradient. Ecology 92, 121–132.



Fire activity on a global scale is a complex system, determined by multiple factors operating simultaneously at different spatial and temporal scales:

- IGNITION PATTERNS
  - Human influence
  - Thunder storms etc.

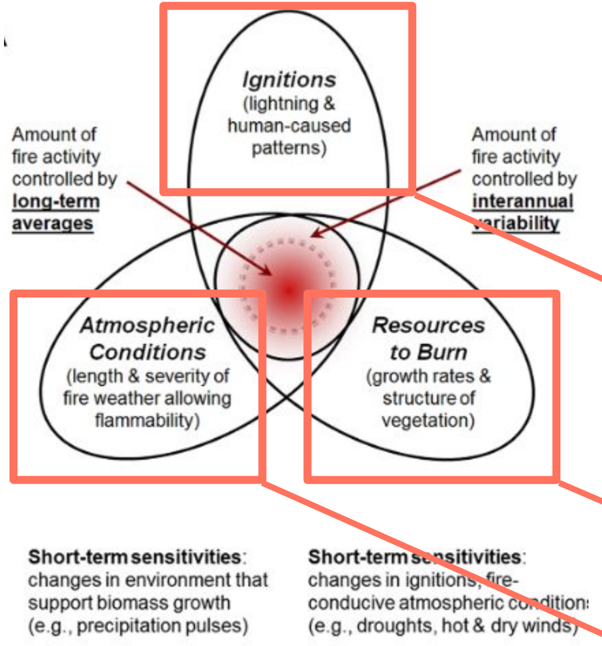
FROM: Krawchuk, M.A., Moritz, M.A., 2010. Constraints on global fire activity vary across a resource gradient. Ecology 92, 121–132.



Fire activity on a global scale is a complex system, determined by multiple factors operating simultaneously at different spatial and temporal scales:

- IGNITION PATTERNS
  - Human influence
  - Thunder storms etc.
- FUEL AVAILABILITY
  - Net Primary Productivity
  - Vegetation distribution etc.

FROM: Krawchuk, M.A., Moritz, M.A., 2010. Constraints on global fire activity vary across a resource gradient. Ecology 92, 121–132.

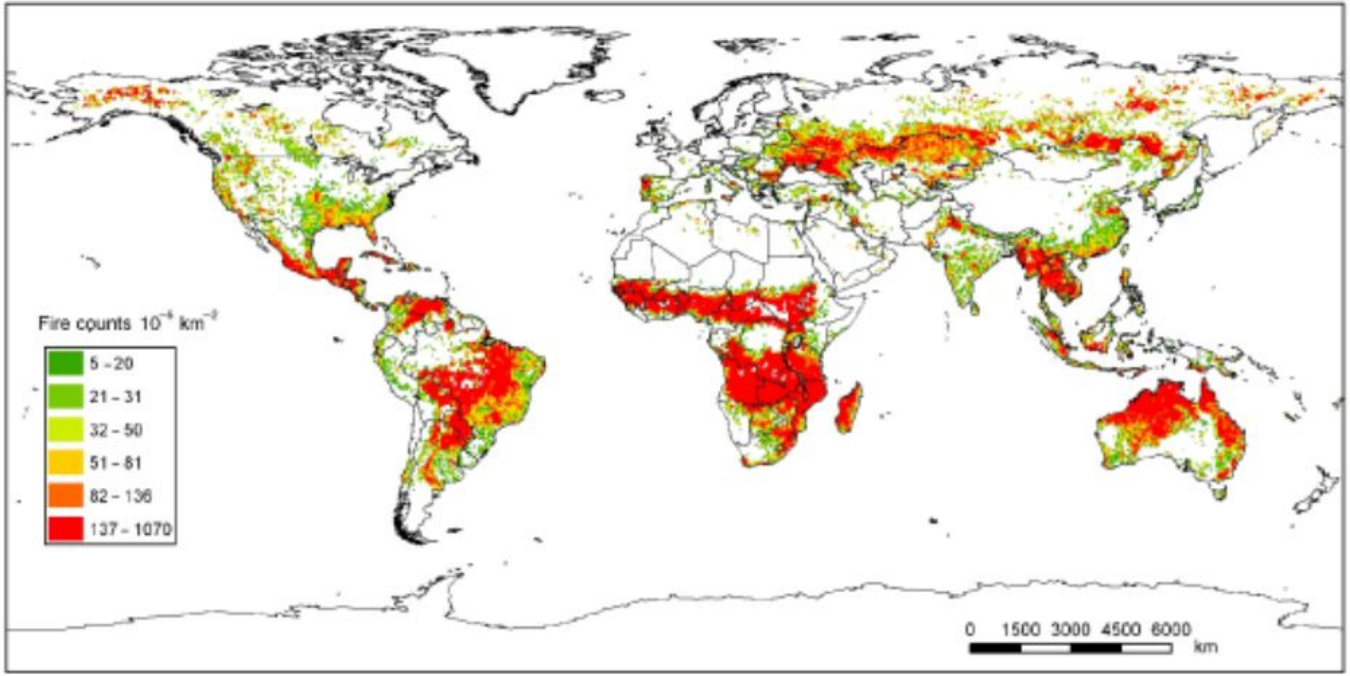


Fire activity on a global scale is a complex system, determined by multiple factors operating simultaneously at different spatial and temporal scales:

- IGNITION PATTERNS
  - Human influence
  - Thunder storms etc.
- FUEL AVAILABILITY
  - Net Primary Productivity
  - Vegetation distribution etc.
- CLIMATE
  - Drought, heatwaves
  - Fire danger
  - Climate change

FROM: Krawchuk, M.A., Moritz, M.A., 2010. Constraints on global fire activity vary across a resource gradient. Ecology 92, 121–132.

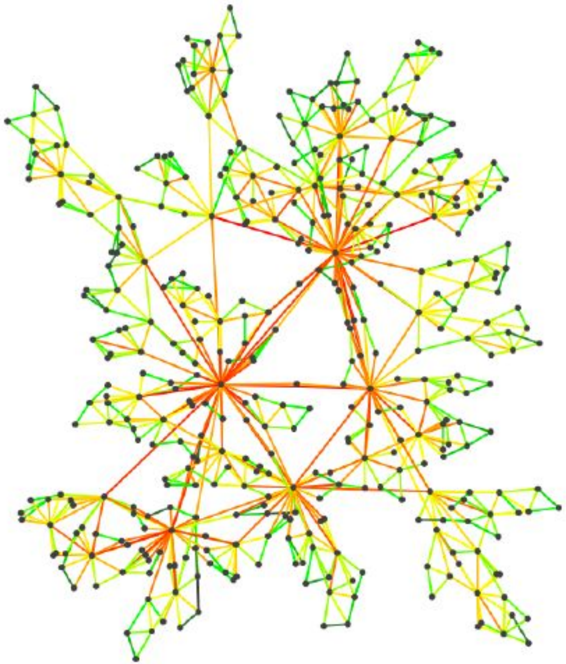
As a result, there is great spatial variability in the overall distribution of global fire activity.



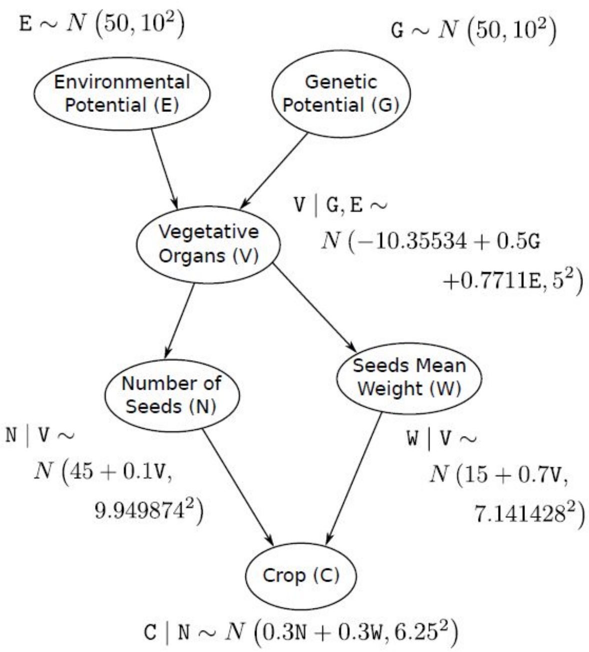
FROM: Chuvieco, E., Giglio, L., Justice, C., 2008. Global characterization of fire activity: toward defining fire regimes from Earth observation data. Global Change Biology 14, 1488–1502.

We propose two approaches based on **complex networks** for studying this phenomenon and the possible teleconnections

Correlation networks (CNs)



Bayesian networks (BNs)



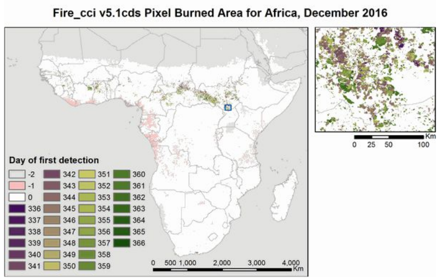
2

Data acquisition and processing

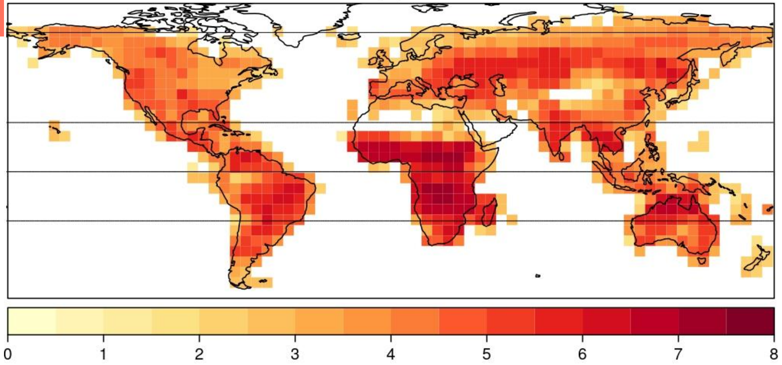
Copernicus
Europe's eyes on Earth

Climate Change Service

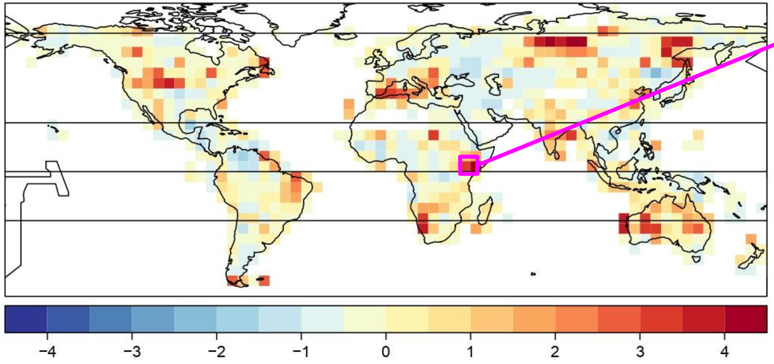
- *Fire burned area from 2001 to present derived from satellite observations.*
- Native resolution
  - Spatial: 0.25º
  - Temporal: monthly
- Data used:
  - Burned area (BA) → **Target variable**
  - Fraction of burnable area (BAF) → **Used for masking/filtering**



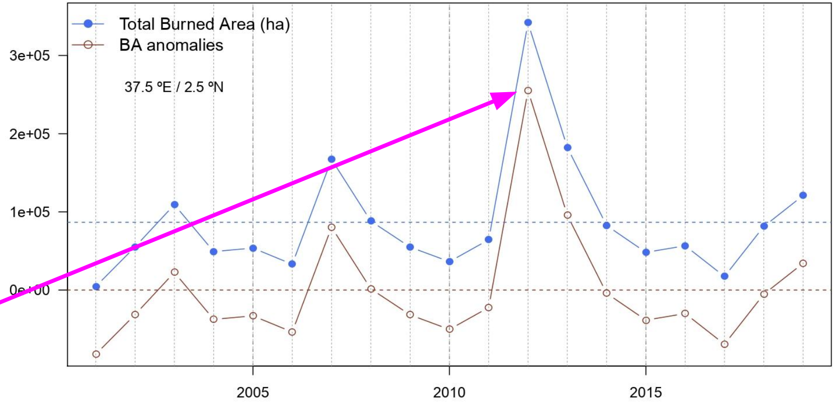
Mean burned area (log10) 2001-2019



Standardized anomalies (2012)



Annual Burned Area time series 2001-2019



Total dataset size:
19 years x 645 grid cells

3

Methods

Correlation networks

Construction

$$\rho_s = \frac{cov(R(X), R(Y))}{\sigma_{R(X)}\sigma_{R(Y)}}$$

Weighted adjacency matrix

$$(A_{ij})^w = \begin{cases} 0 & \text{si} \quad \rho_{ij} \leq \tau_c \\ w_c & \text{si} \quad \rho_{ij} > \tau_c \end{cases}$$

Unweighted adjacency matrix

$$(A_{ij})^u = \begin{cases} 0 & \text{si} \quad \rho_{ij} \leq \tau_c \\ 1 & \text{si} \quad \rho_{ij} > \tau_c \end{cases}$$

Correlation networks

igraph

Construction

$$\rho_s = \frac{cov(R(X), R(Y))}{\sigma_{R(X)}\sigma_{R(Y)}}$$

Weighted adjacency matrix

$$(A_{ij})^w = \begin{cases} 0 & \text{si} \quad \rho_{ij} \leq \tau_c \\ w_c & \text{si} \quad \rho_{ij} > \tau_c \end{cases}$$

Unweighted adjacency matrix

$$(A_{ij})^u = \begin{cases} 0 & \text{si} \quad \rho_{ij} \leq \tau_c \\ 1 & \text{si} \quad \rho_{ij} > \tau_c \end{cases}$$

Selection of threshold T_c is critical to determine network properties

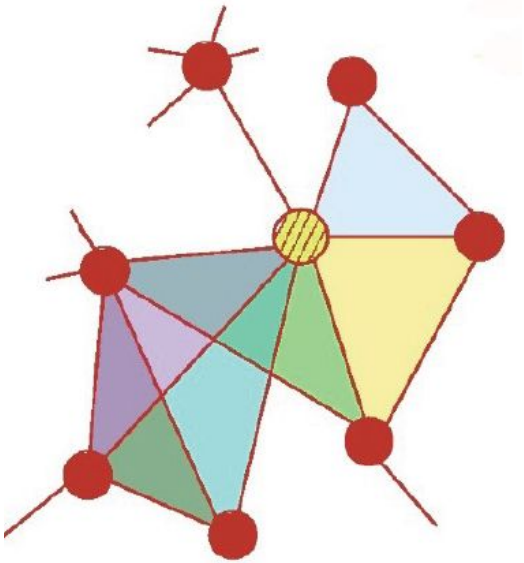
Correlation networks

Global connectivity measures

Global clustering coefficient

Measures the proportion of “clustered” nodes

$$C = \frac{\sum_{i,j,k} A_{ij} A_{jk} A_{ik}}{\sum_i k_i (k_i - 1)}$$



Correlation networks

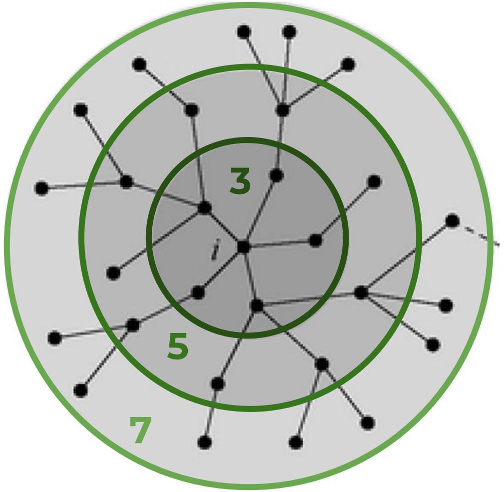
Global connectivity measures

Diameter

Longitude of the longest graph “geodesic*”

$$D = \max\left(\{g_{ij}\}_{i,j=1}^{m}\right)$$

*The geodesic g_{ij} is the shortest path between two nodes (i,j) of the network



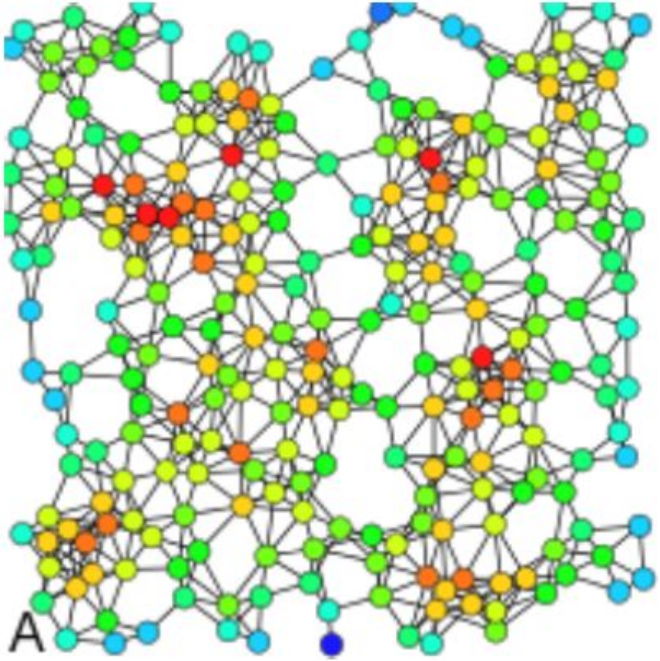
Correlation networks

Centrality measures

Degree

Number of links connecting node *i* with the rest of the network

$$k_i = \sum_{j}^{m} A_{ij}$$



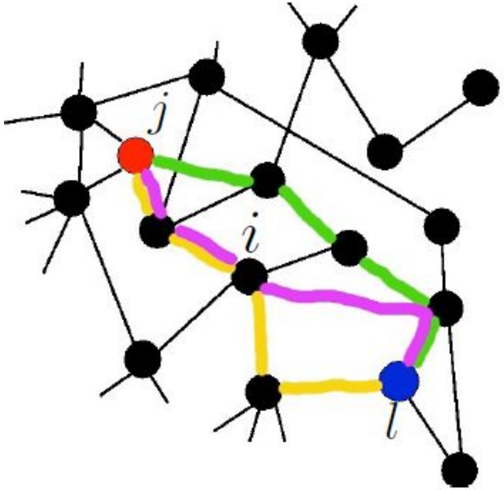
Correlation networks

Centrality measures

Betweenness

Proportion of geodesics* passing through node *i*

$$B_i = \sum_{j,k \neq i}^{m} \frac{g_{jk}^{i}}{g_{jk}}$$



*The geodesic g_{ij} is the shortest path between two nodes (*i*,*j*) of the network

Correlation networks

Centrality measures

- Strength is similar to ‘degree’, but summing the *weights* of all links connecting node *i* with the rest of the network
- Weights can be based on correlation or (geographical) distance:

Correlation-based strength

$$(S_i)^c = \sum_j^m (A_{ij})^w = \sum_j^m (w_c)_{ij}$$

Distance-based strength

$$(S_i)^d = \sum_j^m (w_d)_{ij}$$

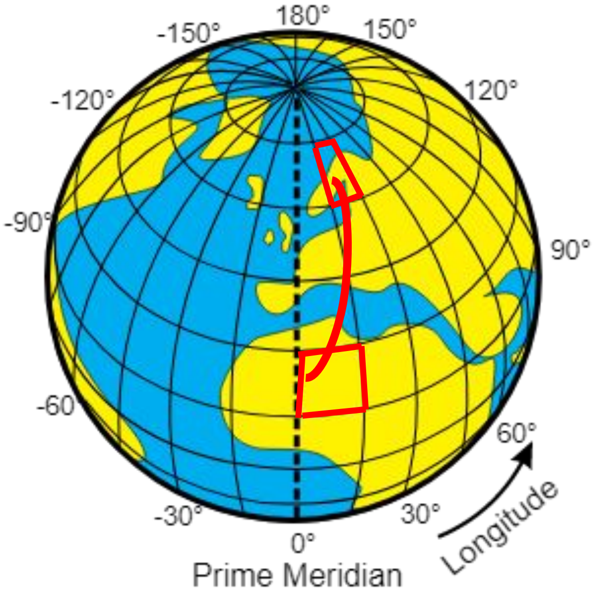
Correlation networks

Centrality measures

Area weighted connectivity

Fraction of the earth's surface to which the node is connected

$$AWC_i = \frac{\sum_j^m A_{ij} \cos(\lambda_j)}{\sum_j^m \cos(\lambda_j)}$$



Correlation networks

Centrality measures

Ratio between the distance-based strength of node *i* and its degree

Mean geographical link distance

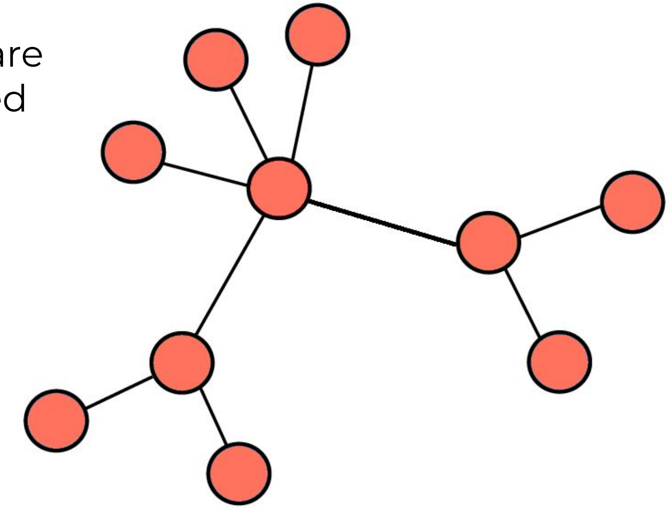
$$MD_i = \frac{(S_i)^d}{k_i} = \frac{1}{k_i}\sum_{j}^{k_i}(w_d)_{ij}$$

Correlation networks

Communities

- Communities are clusters of nodes that are highly connected to each other compared to the rest of the network
- Community Detection Algorithm (clustering) based on betweenness between links (“greedy search”)

$$B_l = \sum_{j,k \neq l}^{m} \frac{g_{jk}^{l}}{g_{jk}}$$



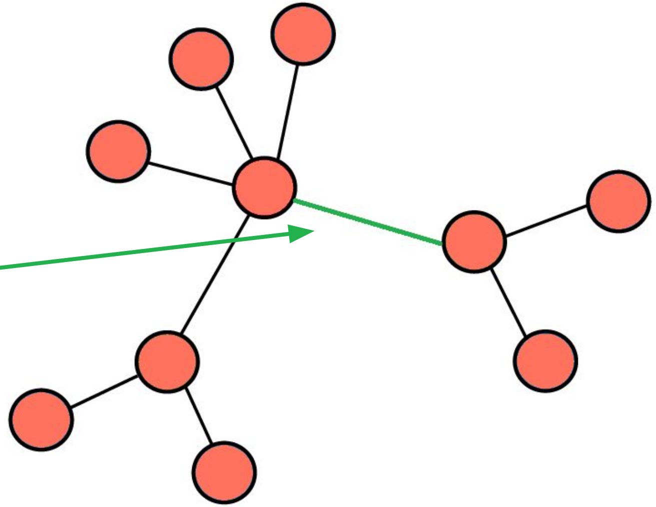
Correlation networks

Communities

Community Detection Algorithm (clustering)
based on betweenness between links

$$\boxed{B_l = \sum_{j,k \neq l}^{m} \frac{g_{jk}^{l}}{g_{jk}}}$$

max

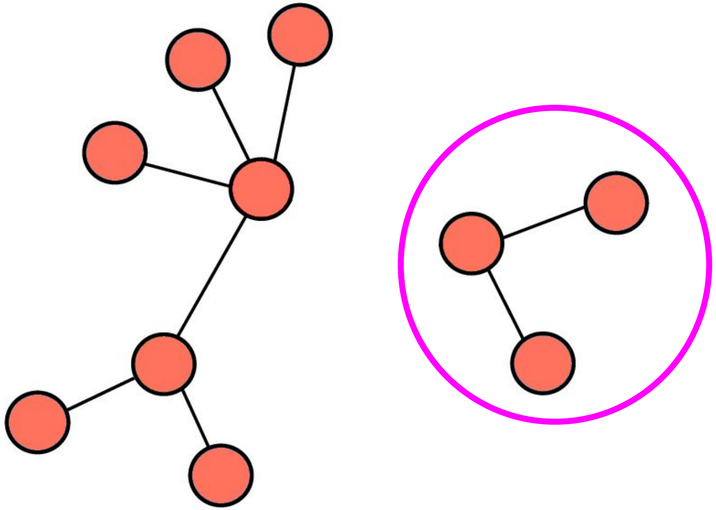


Correlation networks

Communities

Community Detection Algorithm (clustering) based on betweenness between links

$$B_l = \sum_{j,k \neq l}^{m} \frac{g_{jk}^{l}}{g_{jk}}$$

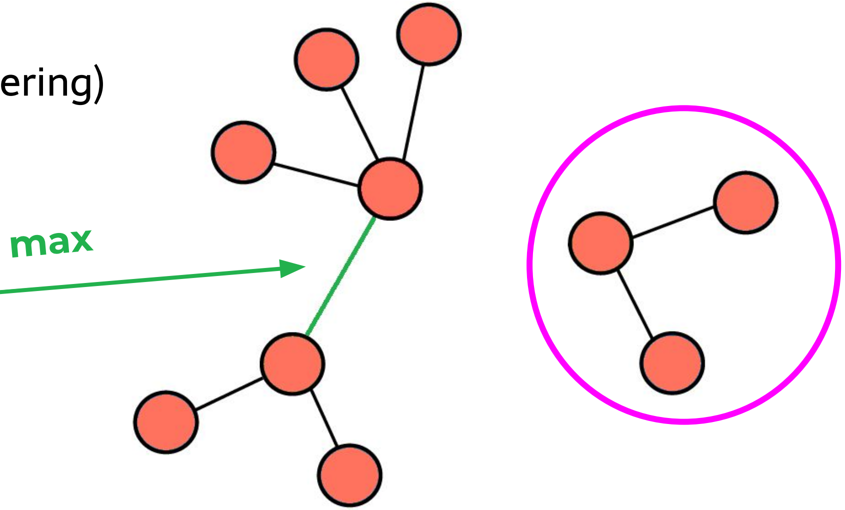


Correlation networks

Communities

Community Detection Algorithm (clustering) based on betweenness between links

$$B_l = \sum_{j,k \neq l}^{m} \frac{g_{jk}^{l}}{g_{jk}}$$

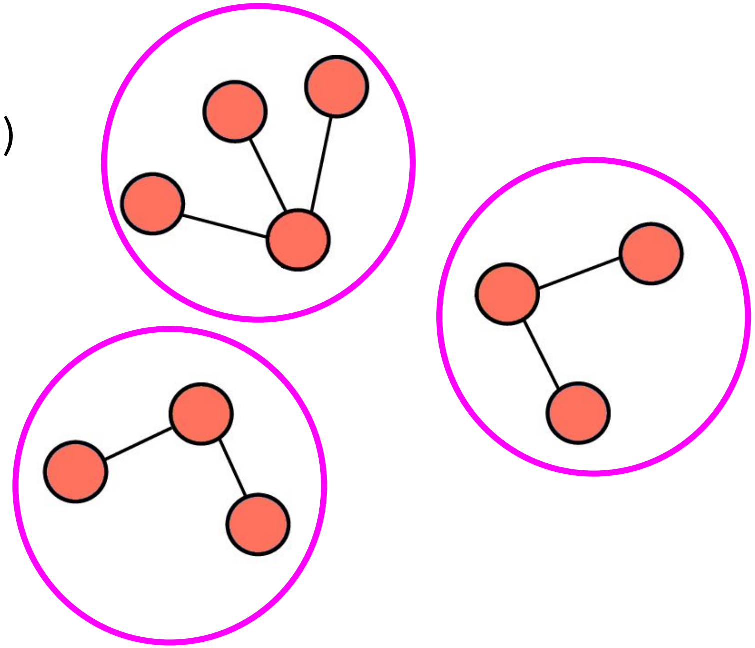


Correlation networks

Communities

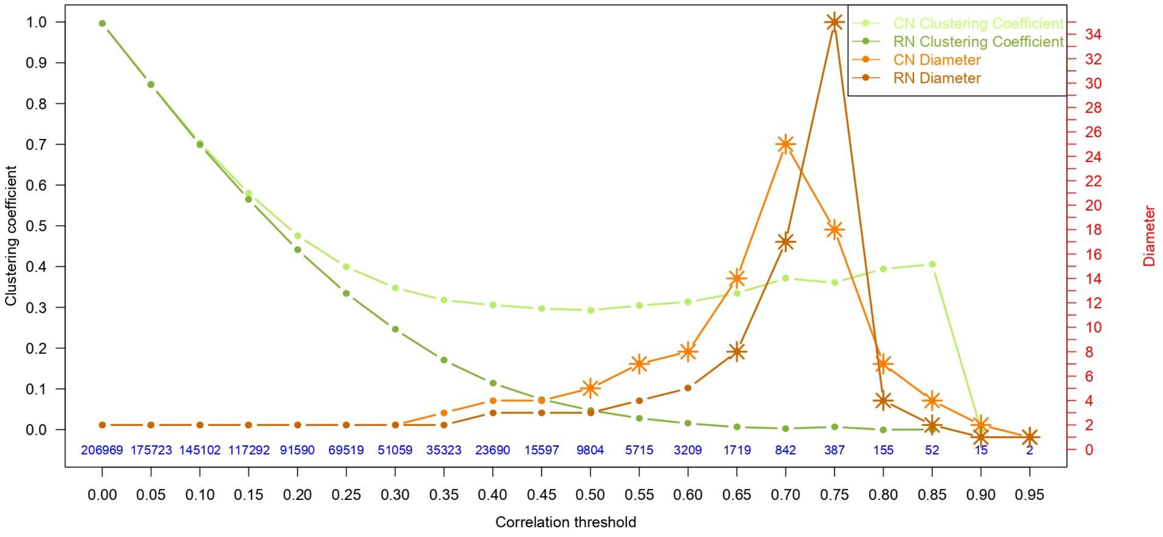
Community Detection Algorithm (clustering) based on betweenness between links

$$B_l = \sum_{j,k \neq l}^{m} \frac{g_{jk}^{l}}{g_{jk}}$$



Correlation networks

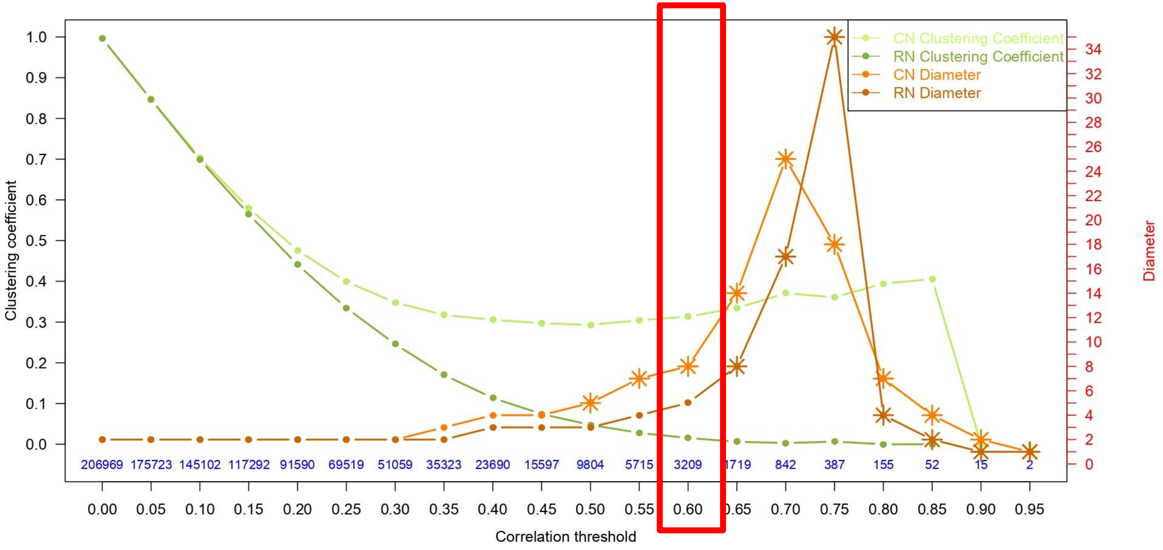
Optimal Correlation threshold choice τ_c



- Different thresholds are intercompared against a “random” network considering Diameter and Clustering coefficient

Correlation networks

Optimal Correlation threshold choice τ_c



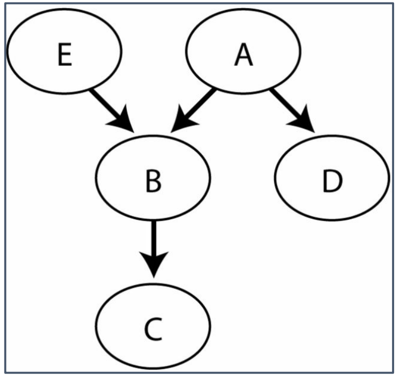
- Different thresholds are intercompared against a “random” network considering Diameter and Clustering coefficient
- Tau=0.6 provides best overall results: i) “stable” clustering and ii) mean connectivity between nodes, thus allowing for a better pattern search

Bayesian networks

Bayesian network

A probabilistic graphical model that represents a set of variables (that is, each BA pixel) and their conditional dependencies via a directed acyclic graph (DAG)

bnlearn package
bnlearn.org



Joint Probability Factorization (each letter is a burned area cell):

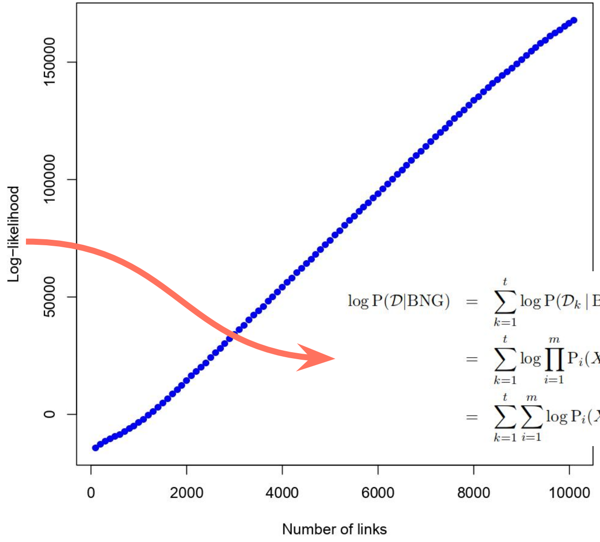
$$\boxed{P(A,B,C,D,E) = P(A)P(B|A,E)P(C|B)P(D|A)P(E)}$$

The joint probability density function can be written as a product of the individual density functions, conditional on their parent variables

$$\mathrm{P}(\Theta \mid \mathcal{G}, \mathcal{D}) = \prod_{i=1}^{d} \mathrm{P}(\Theta_i \mid \Pi_i, \mathcal{D})$$

Bayesian networks

1. Structure (DAG) learning: hill climbing algorithm (automatic)
2. Parameter learning: Gaussian Bayesian network, considering the gaussian response of log10 BA anomalies
3. Optimization: log-likelihood estimation for different network sizes (links)



$$\begin{aligned}\log P(\mathcal{D}|\mathrm{BNG}) &= \sum_{k=1}^{t} \log P(\mathcal{D}_k\,|\,\mathrm{BNG}) \\ &= \sum_{k=1}^{t} \log \prod_{i=1}^{m} P_i(X_i = d_i^k\,|\,\Pi_{X_i} = d^k_{\Pi_{X_i}}) \\ &= \sum_{k=1}^{t}\sum_{i=1}^{m} \log P_i(X_i = d_i^k\,|\,\Pi_{X_i} = d^k_{\Pi_{X_i}})\end{aligned}$$

The results are not optimal -> low number of samples (~20 years)

Optimal network search by comparing log-likelihood scores (considering also the computational cost)

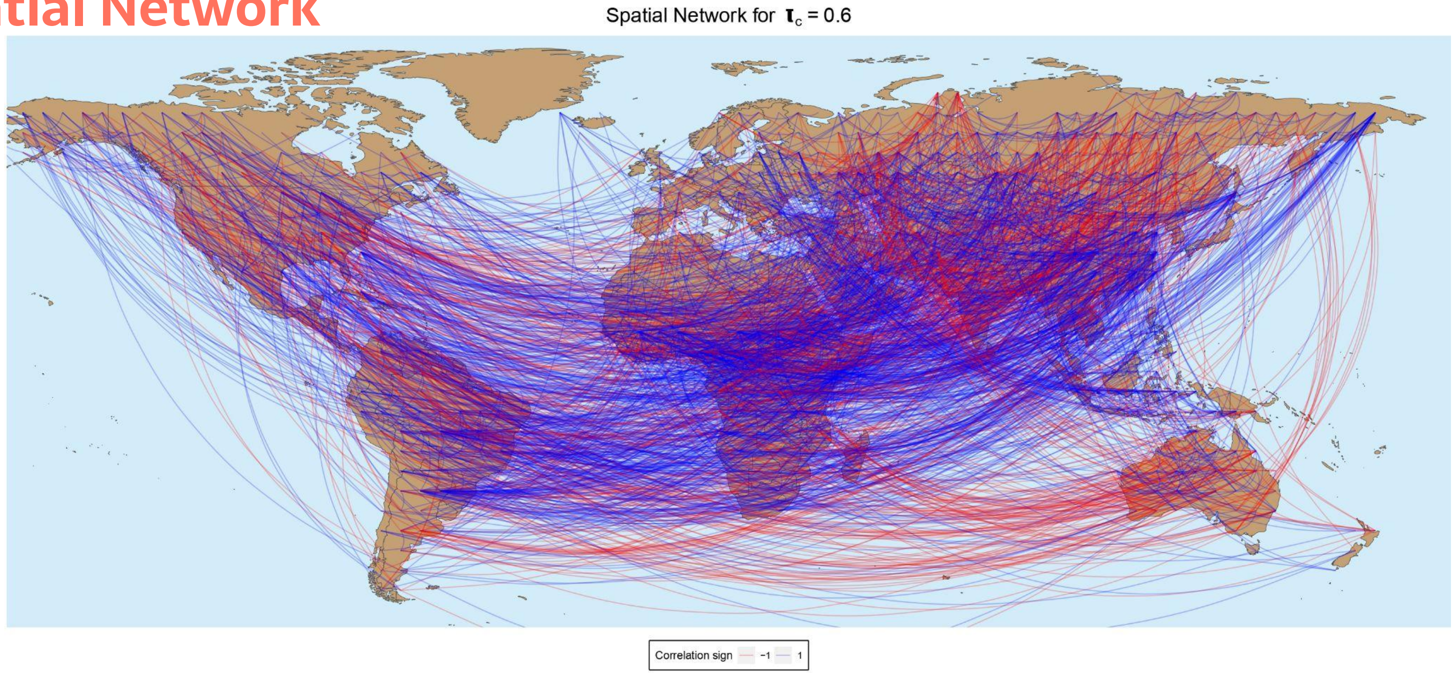
A **compromise** solution is a network with **2000 links**.

4

Results

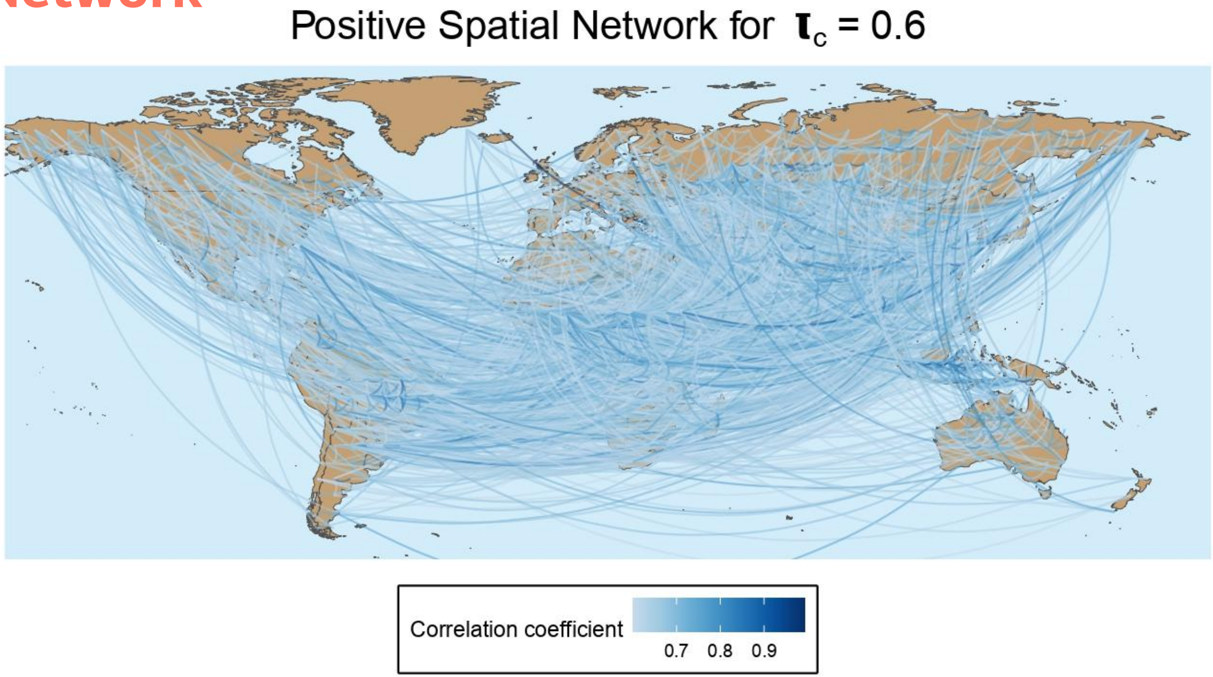
Correlation networks

Spatial Network



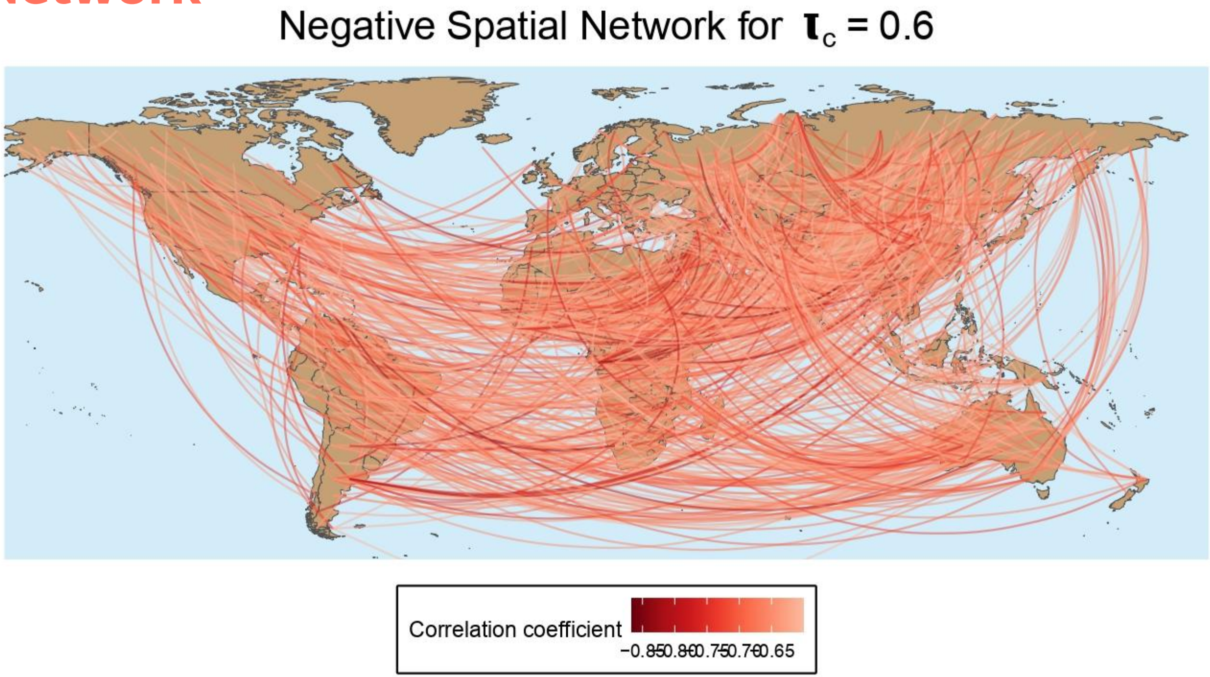
Correlation networks

Spatial Network



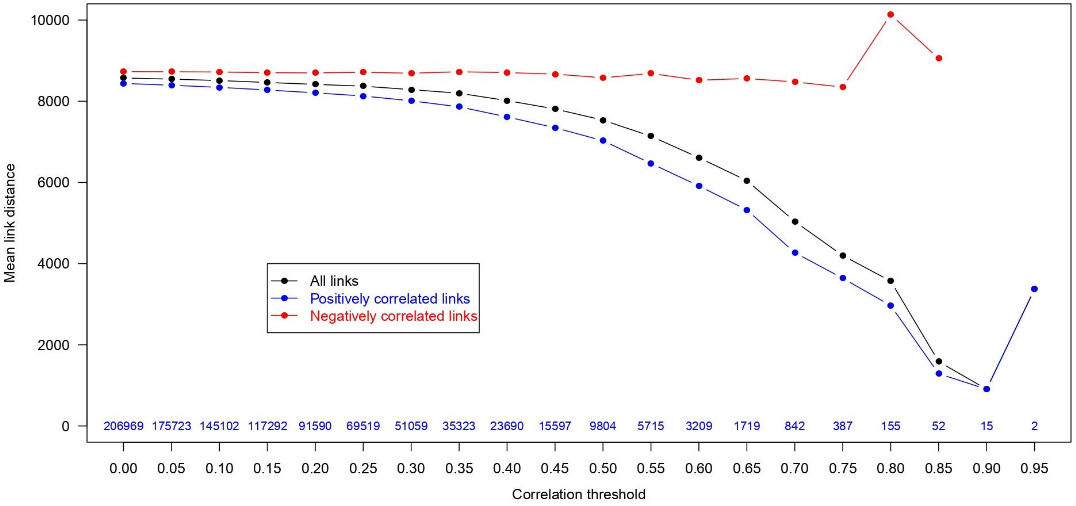
Correlation networks

Spatial Network



Correlation networks

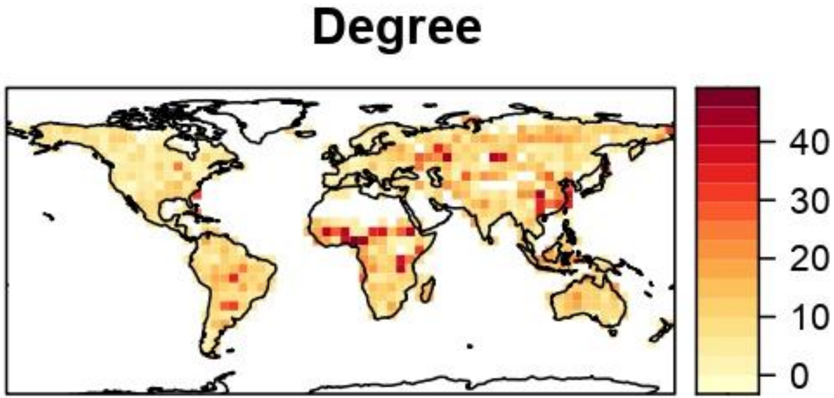
Optimal Correlation threshold choice τ_c

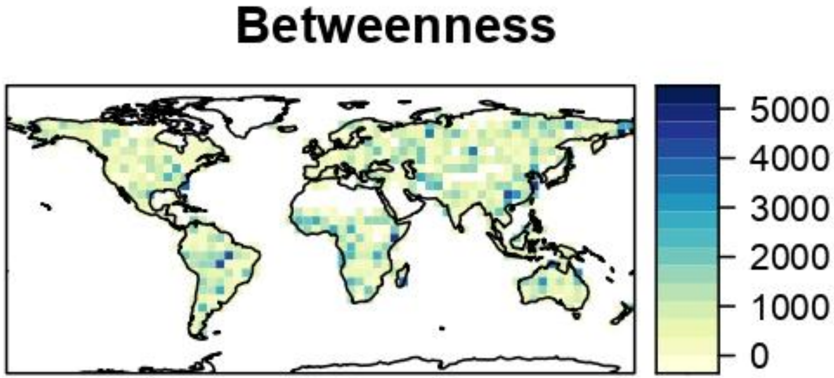


- Mean link distance by sign indicates the prevalence of strong positive local correlations and more stable stable negative long-distance relationships.

Correlation networks

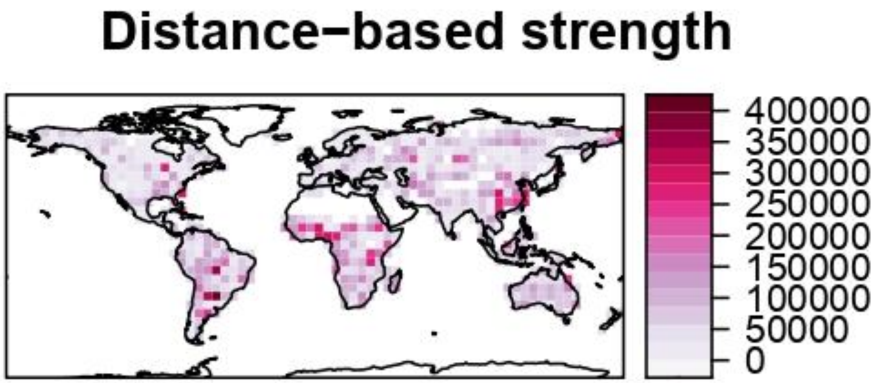
Centrality measures

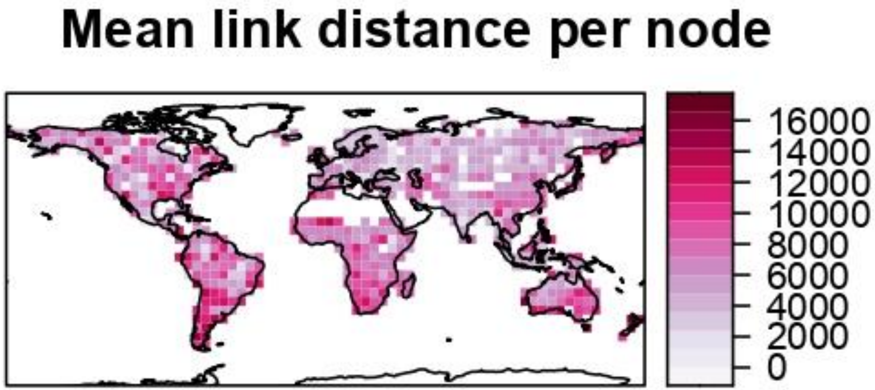




Correlation networks

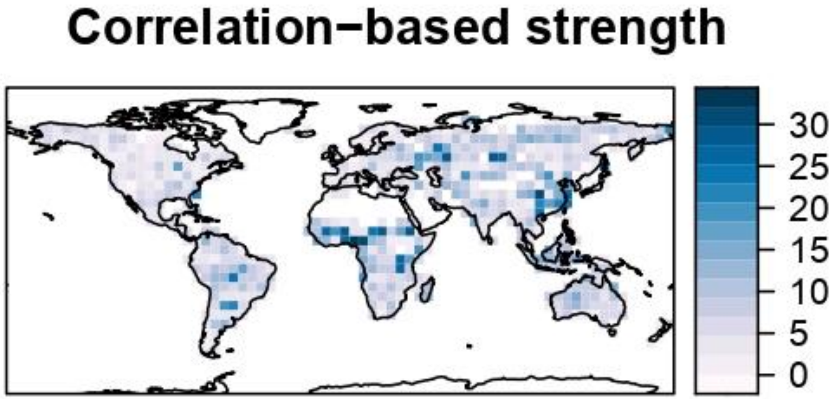
Centrality measures

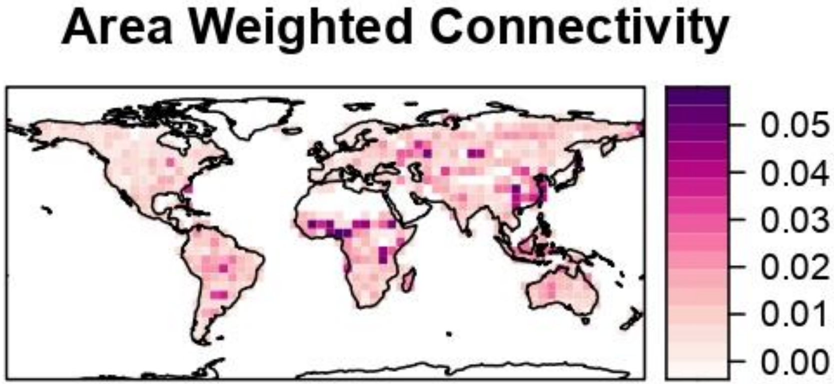




Correlation networks

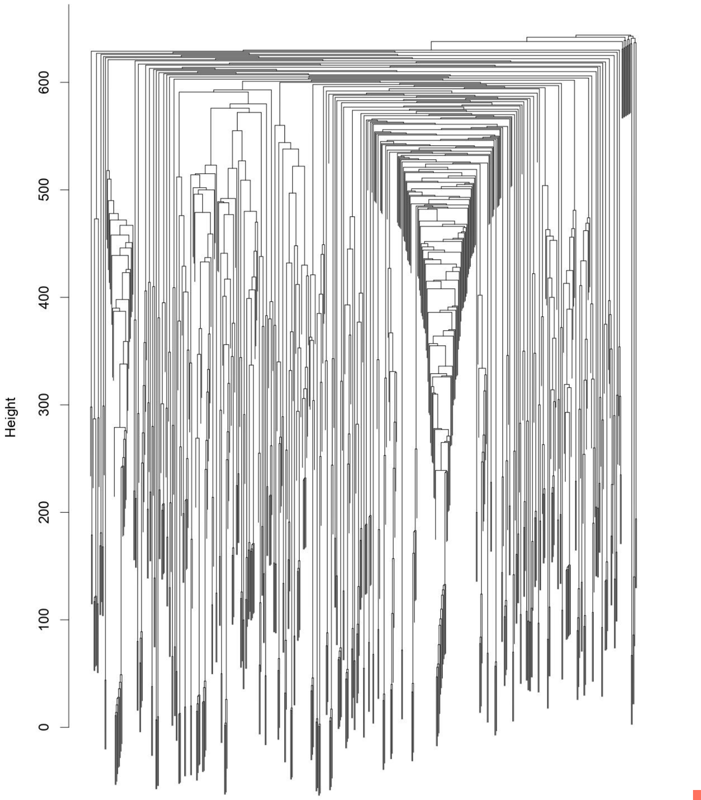
Centrality measures



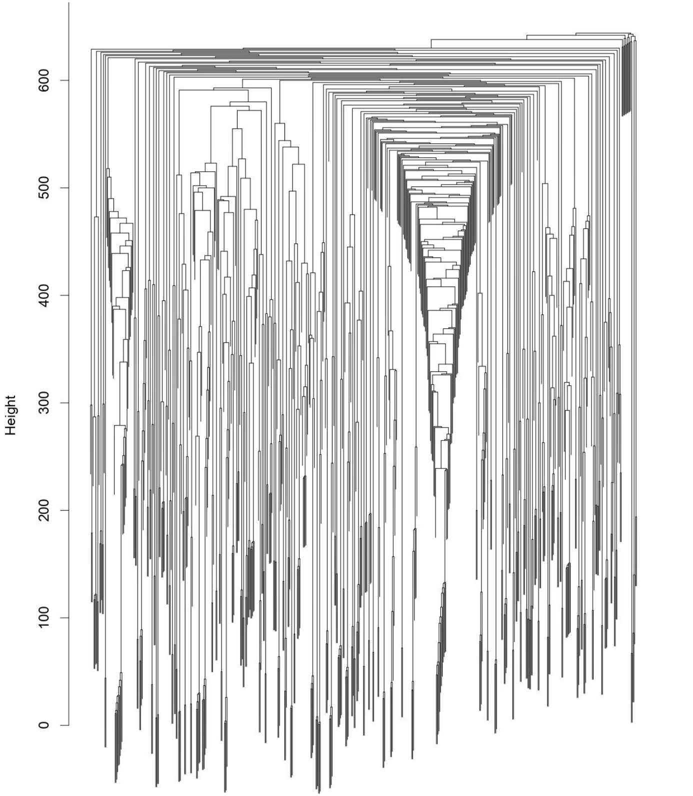


Correlation networks

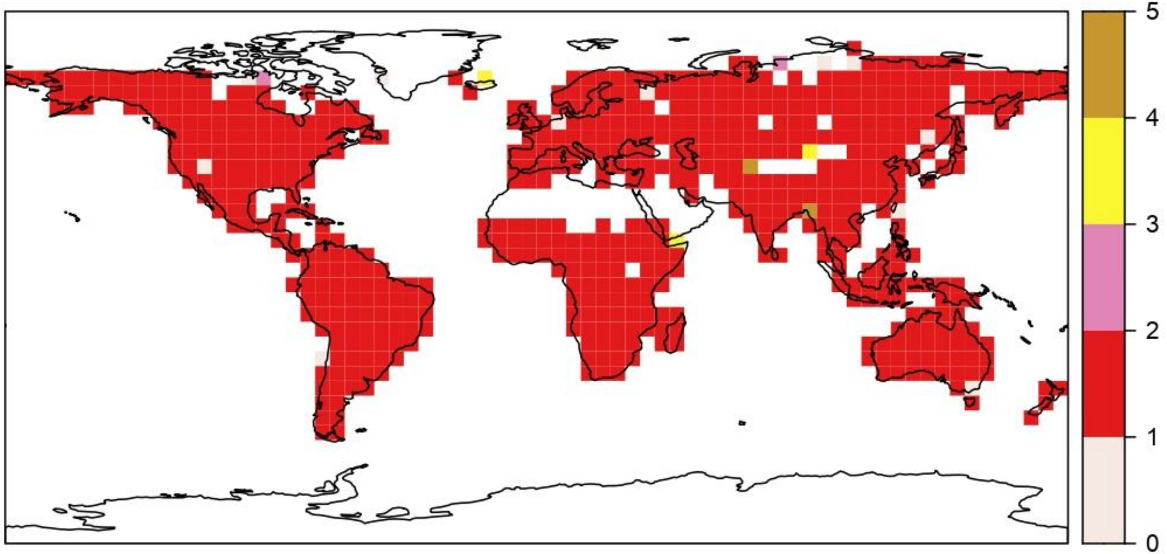
Community detection : $\tau_c = 0.6$



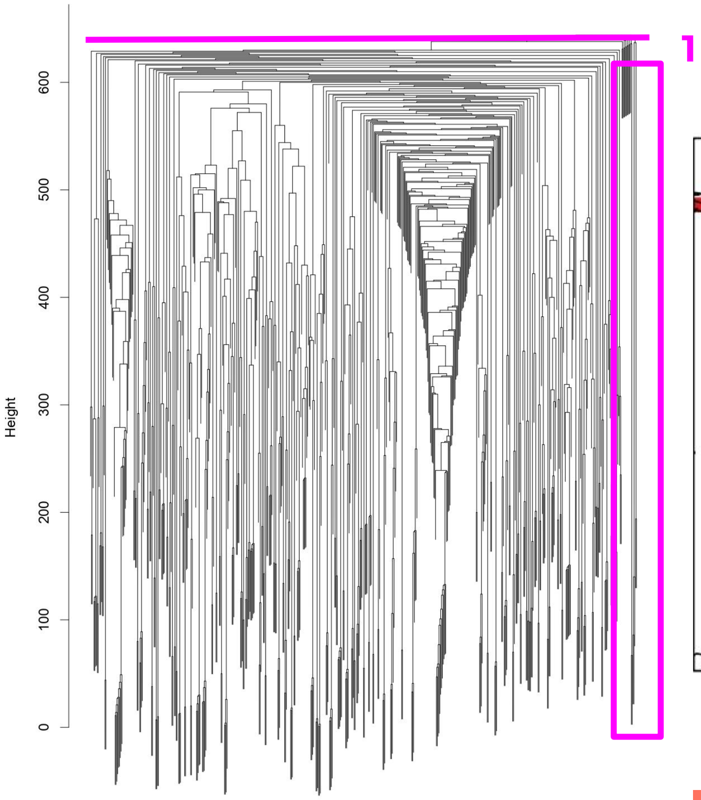
Correlation networks



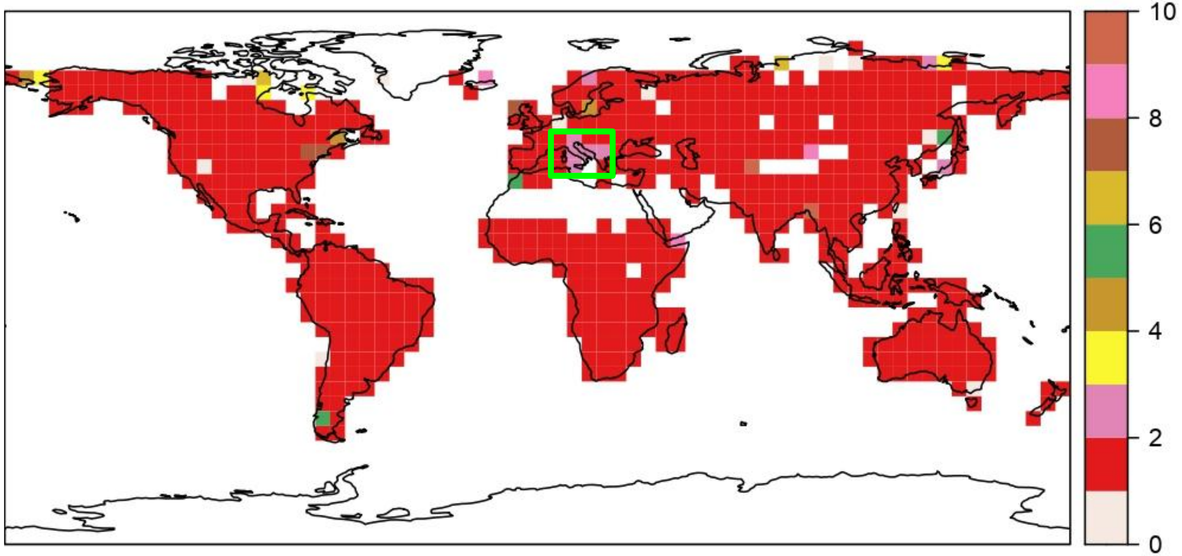
Community detection : $\tau_c = 0.6$



Correlation networks

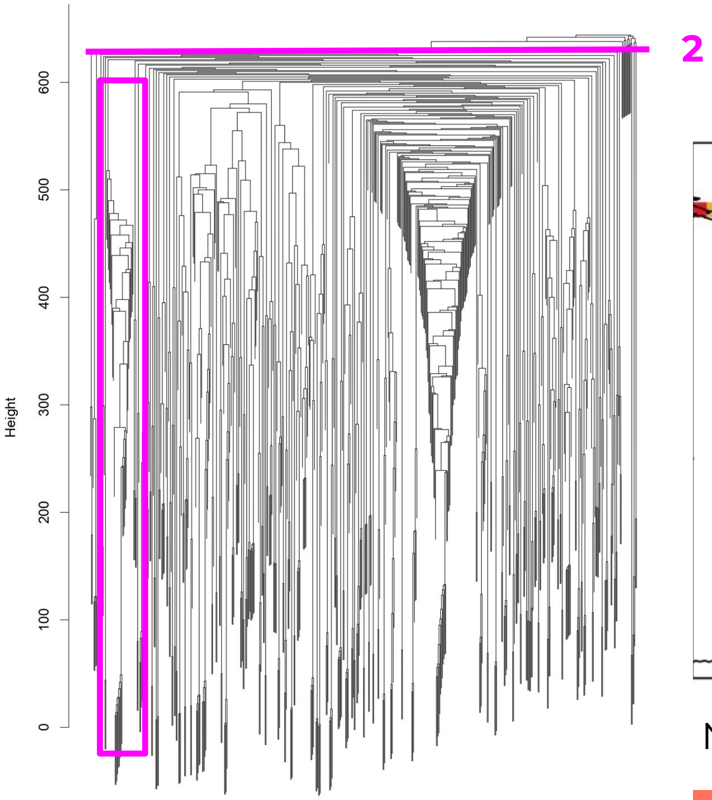


Community detection : $\tau_c = 0.6$

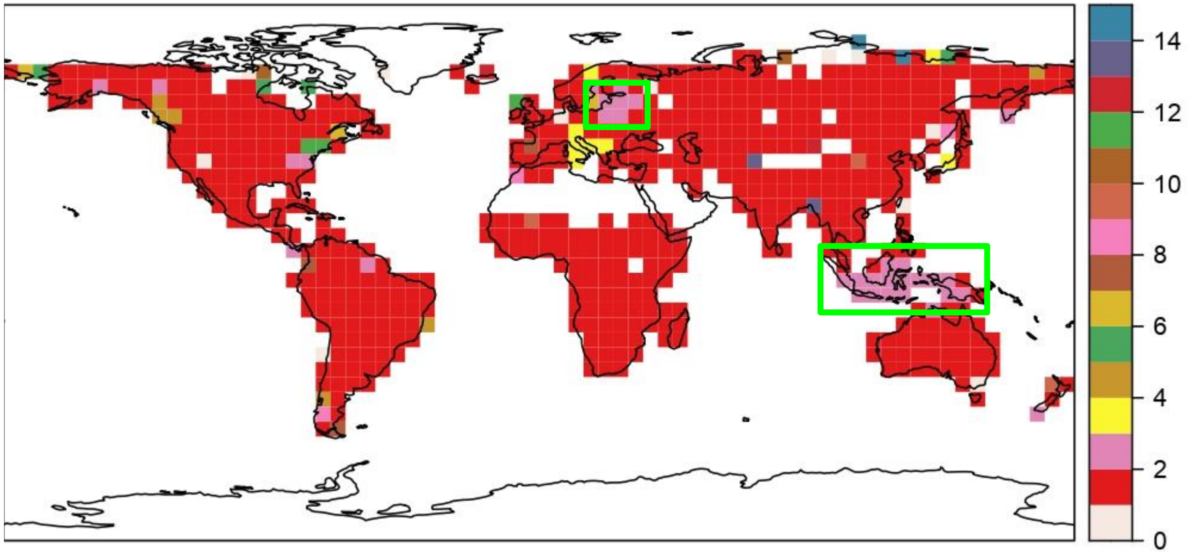


Mediterranean emerges soon as a distinct community

Correlation networks

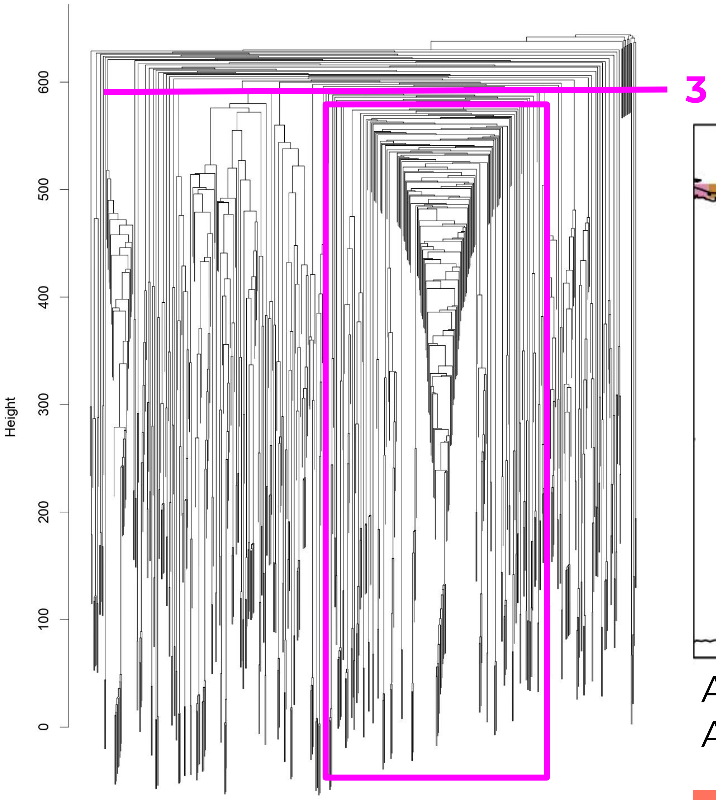


Community detection : τ_c = 0.6

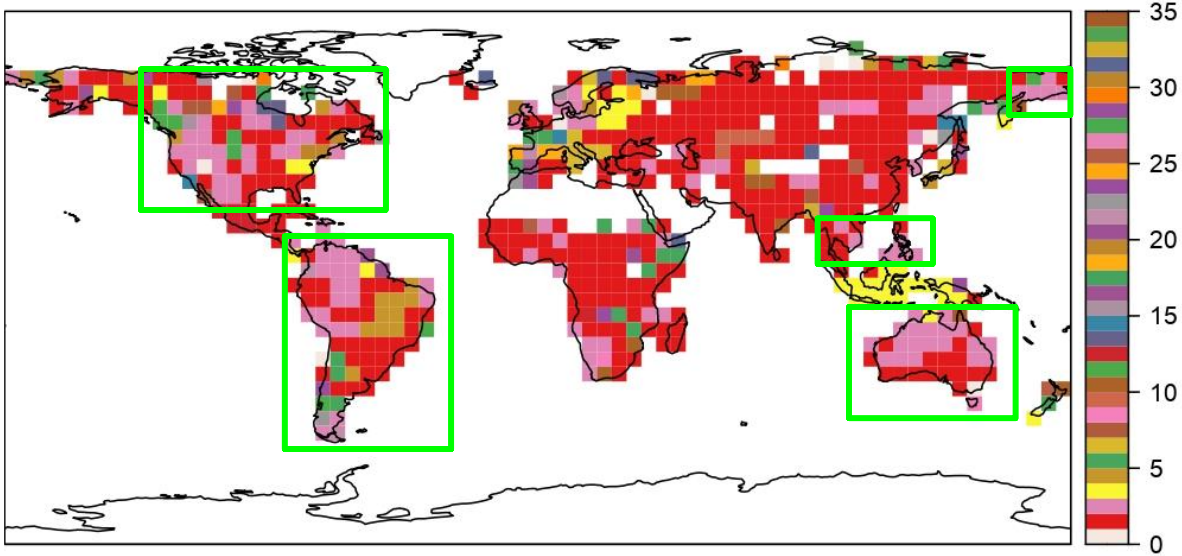


Northern Europe and Indonesia form a robust community

Correlation networks



Community detection : $\tau_c = 0.6$

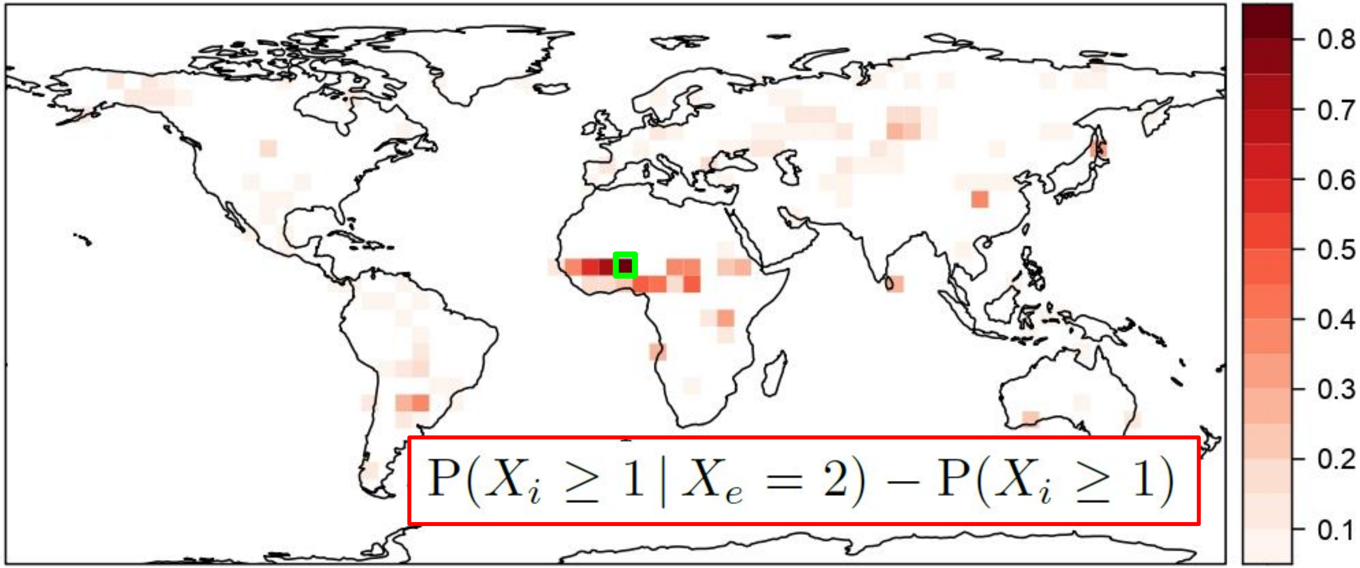


At lower cutoff thresholds Australia Western USA and South America belong to the same cluster

Bayesian networks

Bayesian inference

Evidence in high-degree pixel over Africa

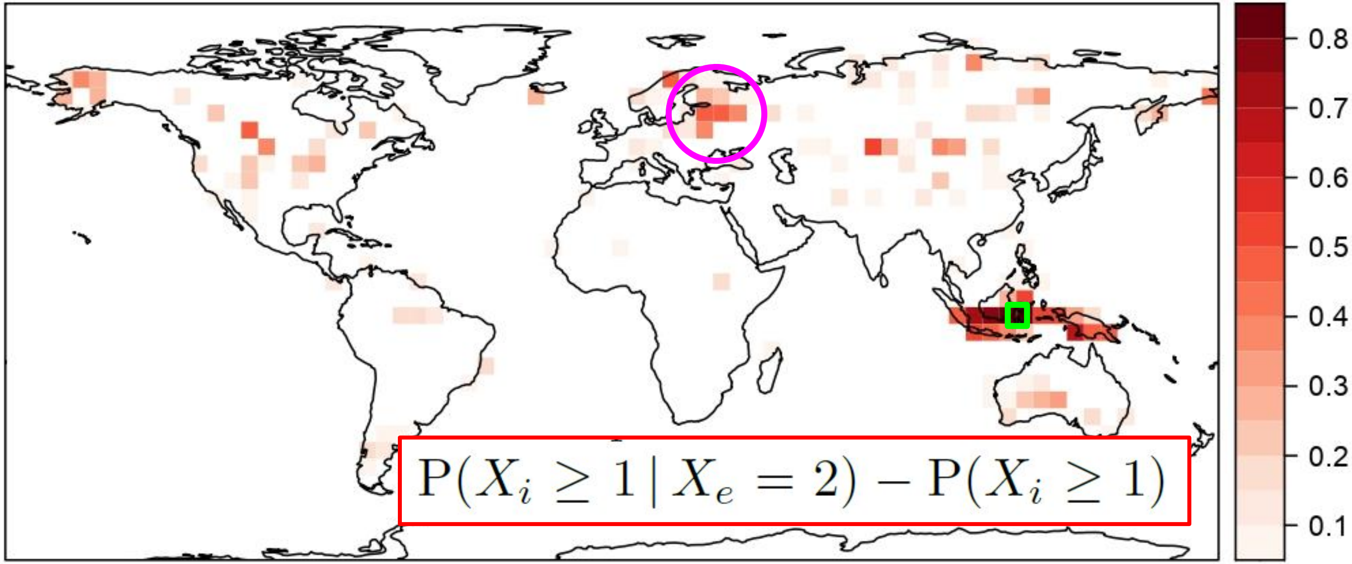


This result does not provide any significant teleconnection pattern beyond local influence

Bayesian networks

Bayesian inference

Evidence in pixel over Indonesia

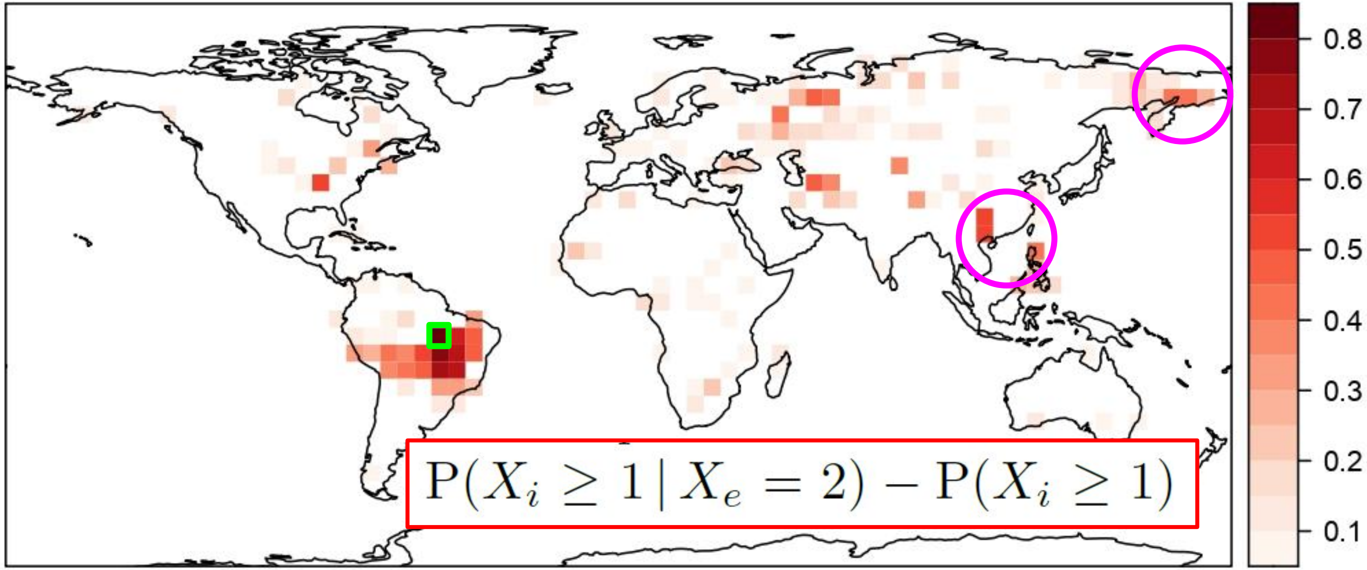


- This result is consistent with correlation network community detection
- It unveils a positive link between Indonesia and Northern Europe

Bayesian networks

Bayesian inference

Evidence in pixel in Amazon Basin (green box)



- This result is consistent with correlation network community detection
- It unveils a positive link between South America, NE Siberia and SE Asia

5

Conclusions

1. The fire database contains an underlying spatial structure.
2. Both approaches, although different in construction, provide **consistent results**. The robustness of the synchronicities found is confirmed.
3. Bayesian networks seem a preferable option, being able to eliminate redundancies inherent to correlation networks and to encode conditional dependencies.
4. Synchronicity in annual fire activity is observed between distant areas, such as equatorial Africa and South America, Indonesia and Northern Europe, or the Amazon Basin and the Philippines.
5. Complex networks offer a suitable approach for investigating wildfire synchronicities, and have the potential for investigating lagged teleconnections too

On-going work...

1. Replace Burned Area by historical Fire danger records (e.g. mean fire season FWI) allowing for a larger sample size -> More robust networks expected
2. Underlying mechanisms for teleconnections found are being currently investigated -> climate teleconnection patterns

Code for reproducibility can be found at:

https://github.com/CatharinaG/Complex_wildfire

GitHub

Thanks for your interest!