Functional connectivity: can we find a common ground?

Anna Alińska, Juan Camilo Avendaño Diaz, Agnieszka Tymorek, Paulina Anna Dąbrowska, Agnieszka Dębska, Marcin Koculak, Emilia Kolada, Aleksander Molak, Grzegorz Link, Çağdaş Topçu, Małgorzata Wierzba, Natalia Z. Bielczyk

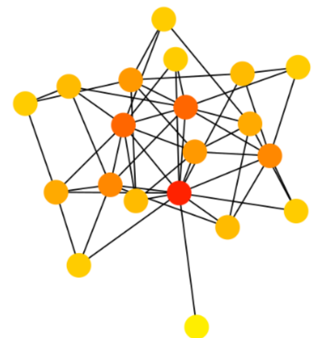
Paris, June 11th 2018

What is functional connectivity

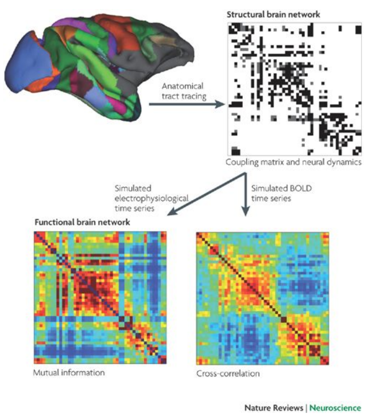
- Complex systems characterized as graphs (Leonhard Euler, 1736)

  Nodes – processes/agents

  Links – relations



- *Functional connectivity* – a tool to define and further study graphs in brain networks



Sporns & Bullmore, 2009

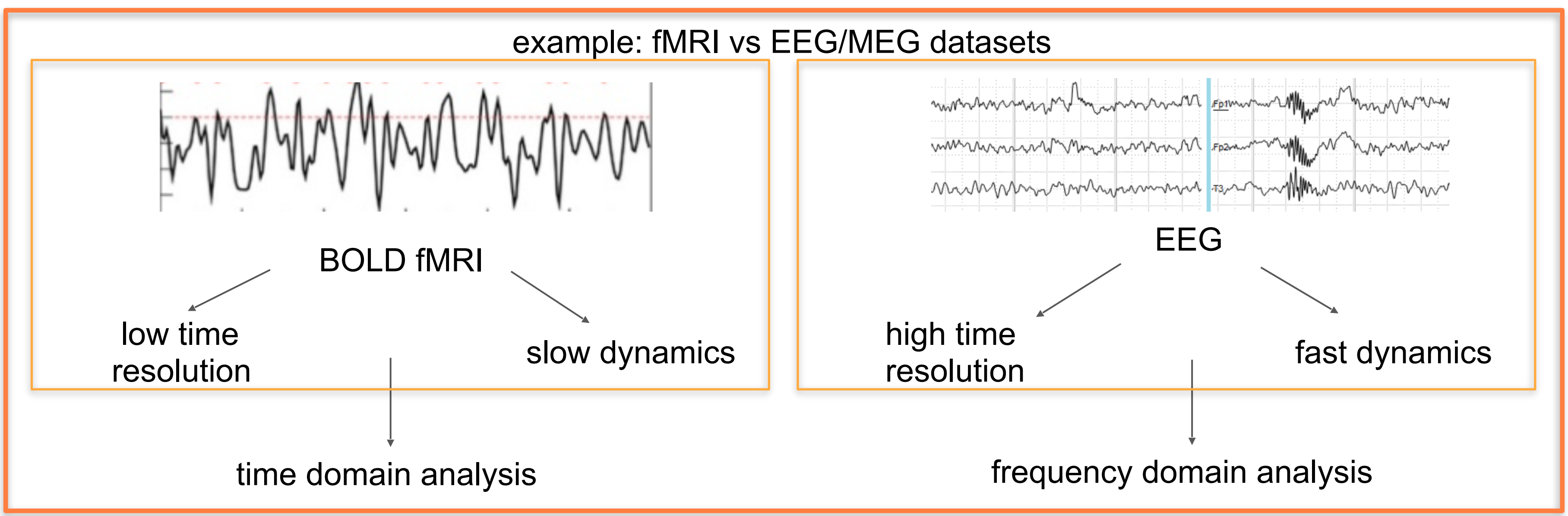
- Typically defined as a *statistical association* between nodes in the network

Research questions

- Can we compare the performance of the methods measuring functional connectivity across different domains?
- How much does the method is limited by the data properties?
- How to choose the right method to operationalize functional connectivity for a particular problem?

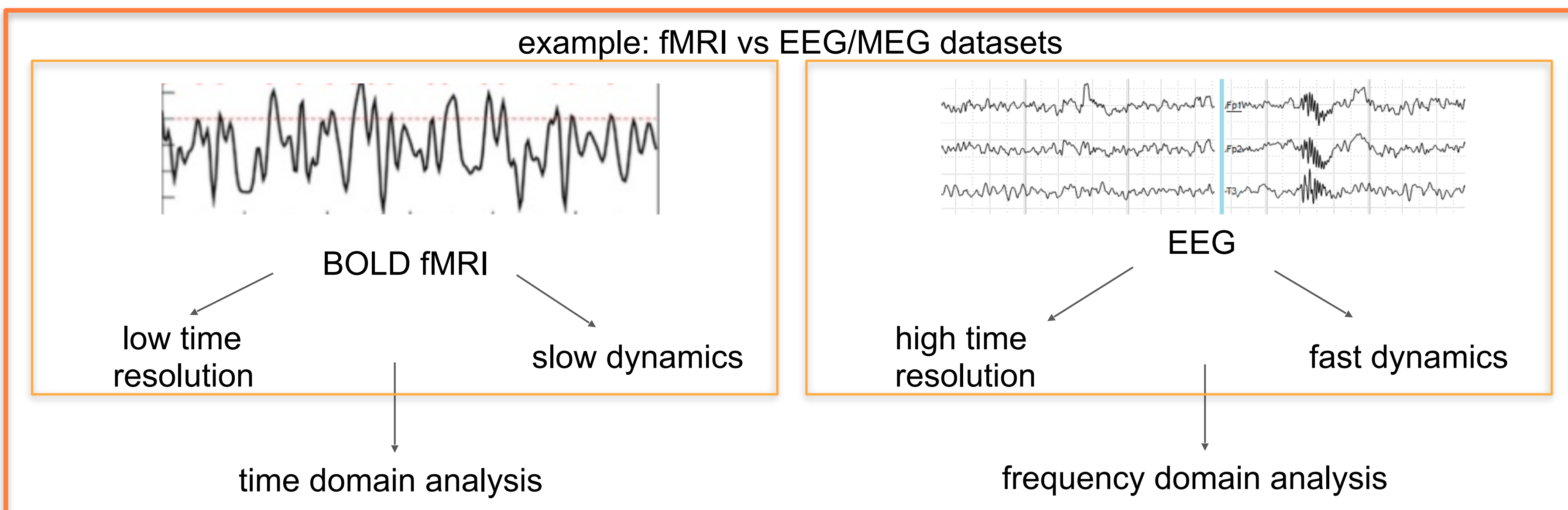
Research questions

- Can we compare the performance of the methods measuring functional connectivity across different domains?
- How much does the method is limited by the data properties?
- How to choose the right method to operationalize functional connectivity for a particular problem?



Research questions

- Can we compare the performance of the methods measuring functional connectivity across different domains?
- How much does the method is limited by the data properties?
- How to choose the right method to operationalize functional connectivity for a particular problem?
- How do other methods perform in relation to Pearson correlation in the neuroscience research?

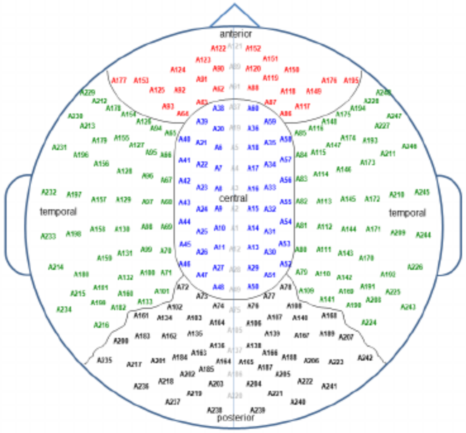


Datasets

[1] Classic resting state neuroimaging datasets coming from the Human Connectome Project (van Essen et al., 2013):

- **functional Magnetic Resonance Imaging** datasets:
  - > preprocessed using FSL with ICA-AROMA correction for movement artifacts
  - > parcellated using new cortical parcellation by Glasser et al. (2016) using a trade-off between machine learning and neuroanatomy:
    - (a) visual cortex (42 ROIs)
    - (b) somatosensory cortex (21 ROIs)

- **magnetoencephalography** datasets:
  - > channel-based ROIs:
    - (a) anterior (23 ROIs)
    - (b) posterior (51 ROIs)
    - (c) central (58 ROIs)
    - (d) left lobe (41 ROIs)
    - (e) right lobe (42 ROIs)



Datasets

[2] Datasets beyond neuroscience:

• **Google trends:**

> Google trends weekly summary statistics for the past 5 years, normalized to the interval 0 - 100, 100 denoting the peak over the 5-year period of time

> 34 search terms, 4 categories: regular rhythm (e.g.,“tax”), peaking at a particular event (e.g., “UFO”), irregular (“sport”), and drifting (“cyber-security”)

• **Stock Exchange:**

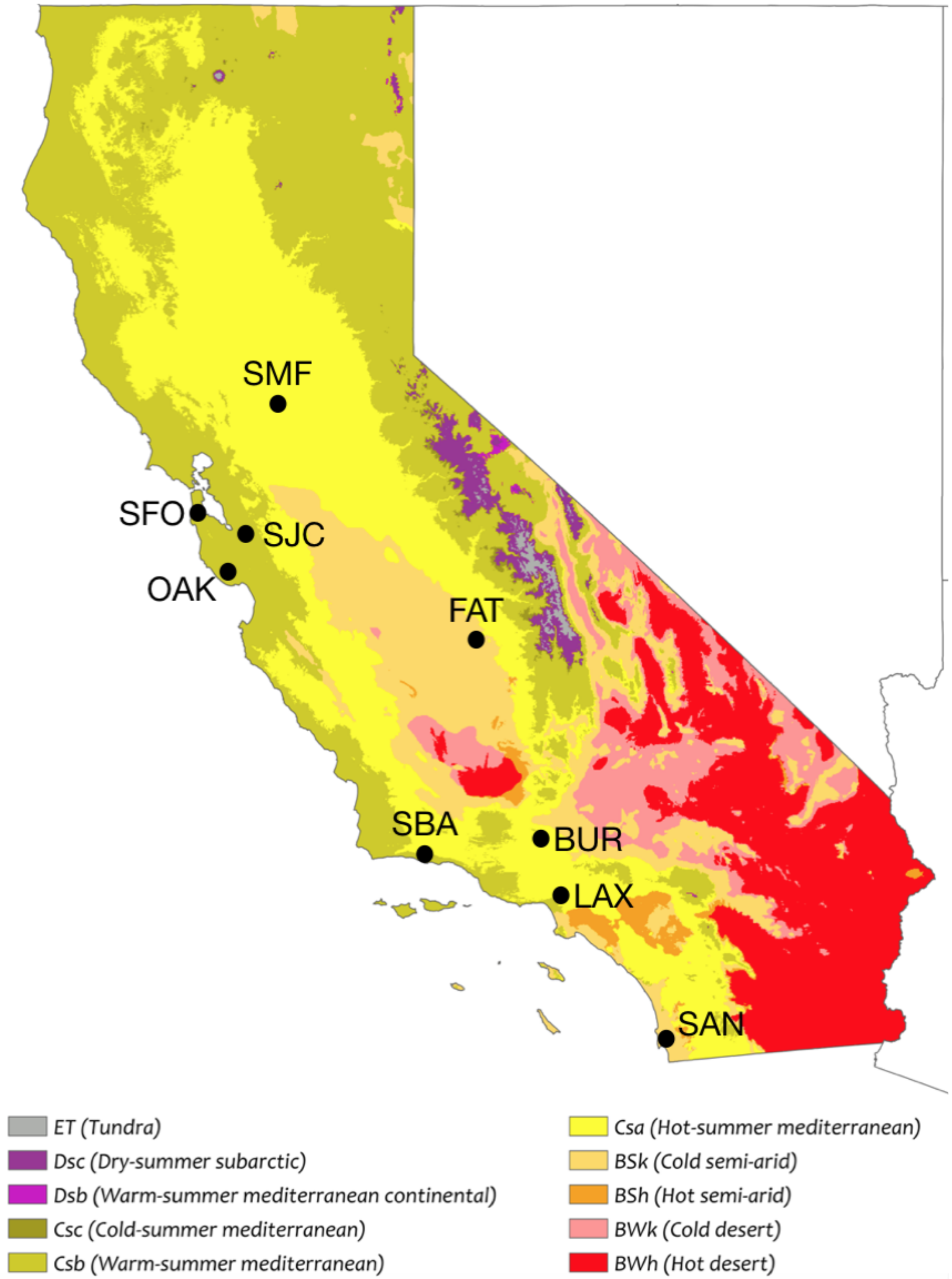
> selection from US companies (S&P500 index)

> stock exchange normalized closing daily prices

• **Weather:**

> Airport weather stations (California, US), Global Historical Climatology Network-Daily database

> Weather variables (e.g., precipitation, temperature, wind speed)


SMF
SFO
SJC
OAK
FAT
SBA
BUR
LAX
SAN
ET (Tundra)
Dsc (Dry-summer subarctic)
Dsb (Warm-summer mediterranean continental)
Csc (Cold-summer mediterranean)
Csb (Warm-summer mediterranean)
Csa (Hot-summer mediterranean)
BSk (Cold semi-arid)
BSh (Hot semi-arid)
BWk (Cold desert)
BWh (Hot desert)

Methods

Functionnal Connectivity Methods

Time domain

- Pearson correlation
- Partial correlation
- Partial partial correlation
- Mutual information
- Distance measures

Frequency domain

- Coherence

Functional Connectivity Methods

Pearson Correlation

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Possible issues: pair of variables = isolated system

Functional Connectivity Methods

Partial correlation

Regression coefficient vectors

$$\mathbf{w}_X^* = \arg\min_w \left\{ \sum_{i=1}^{N} (x_i - \langle \mathbf{w}, \mathbf{z}_i \rangle)^2 \right\},$$

$$\mathbf{w}_Y^* = \arg\min_w \left\{ \sum_{i=1}^{N} (y_i - \langle \mathbf{w}, \mathbf{z}_i \rangle)^2 \right\}$$

$$e_{X,i} = x_i - \langle \mathbf{w}_X^*, \mathbf{z}_i \rangle,$$
$$e_{Y,i} = y_i - \langle \mathbf{w}_Y^*, \mathbf{z}_i \rangle$$

Residuals

Possible issues: can be overconservative

Possible solution? **Partial partial correlation**

$$e_{X,i} = x_i - \alpha \langle \mathbf{w}_X^*, \mathbf{z}_i \rangle,$$
$$e_{Y,i} = y_i - \alpha \langle \mathbf{w}_Y^*, \mathbf{z}_i \rangle$$

$\alpha \in [0,1]$

$\alpha = 0.5$

Functional Connectivity Methods

Mutual Information

$$H(X) = -\sum_{i=1}^{n} p_i \log(p_i).$$

$$\begin{aligned} H(X,Y) &= H(X) - H(X \mid Y) \\ &= H(Y) - H(Y \mid X) \end{aligned}$$

Possible issues: the type of the relationship is unknown

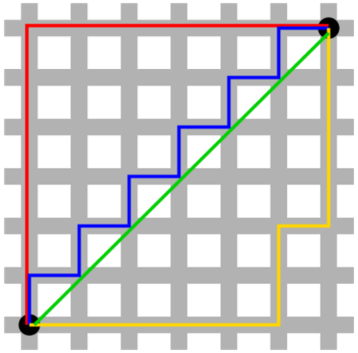
Functional Connectivity Methods

Euclidian Distance

$$d_E(x_1, x_2) = \sqrt{\sum_{i=1}^{N} (x_{1,i} - x_{2,i})^2}$$

Manhattan Distance

$$d_M(x_1, x_2) = \sum_{i=1}^{N} |x_{1,i} - x_{2,i}|$$



Possible issues: only absolute values between the variables are taken into account

Inverse distance – the closer the variables, the stronger the connection

Functional Connectivity Methods

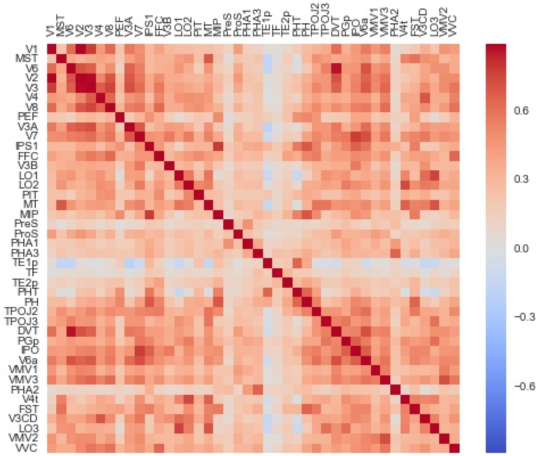
Coherence

$$C_{xy}(f) = \frac{G_{xy}(f)}{\sqrt{G_{xx}(f)G_{yy}(f)}}$$

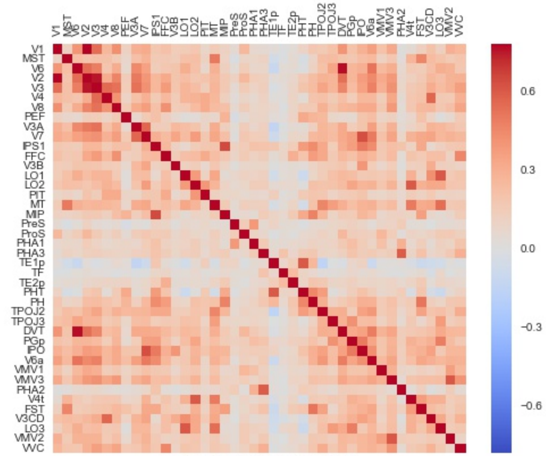
Results

Results: methods in time domain

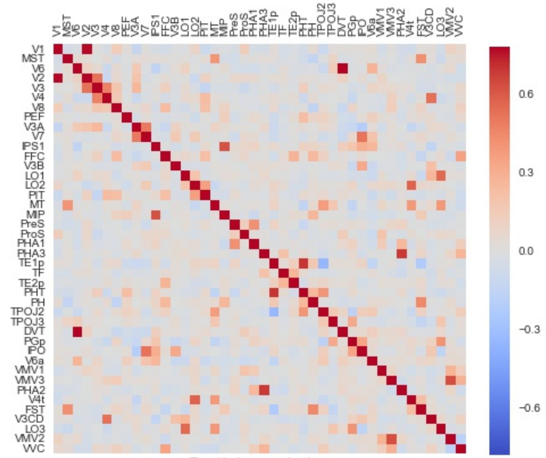
Example: fMRI data, visual system:



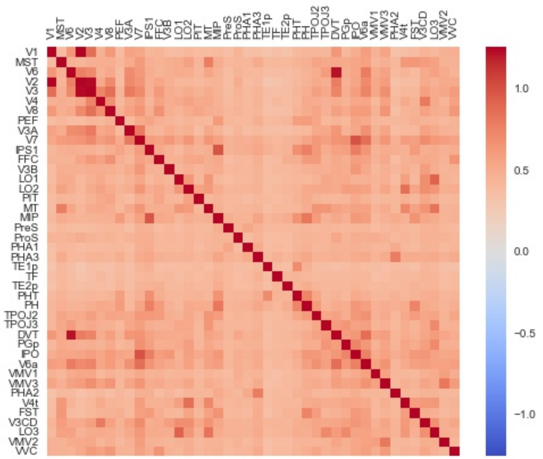
Pearson correlation



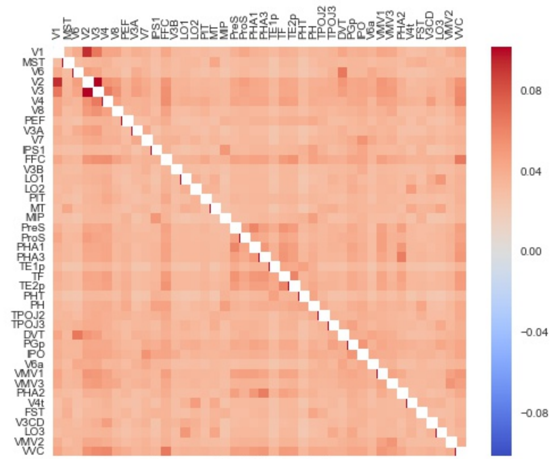
Partial partial correlation



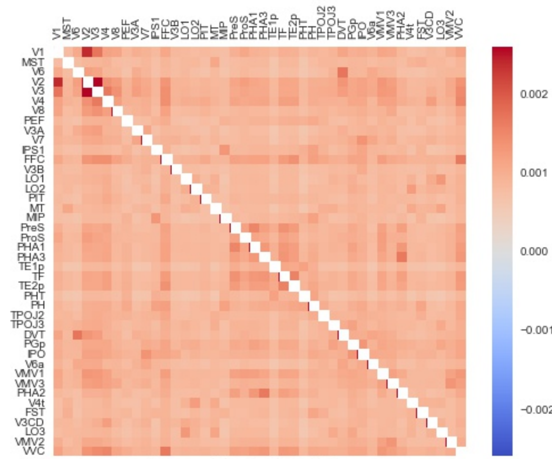
Partial correlation



Mutual information



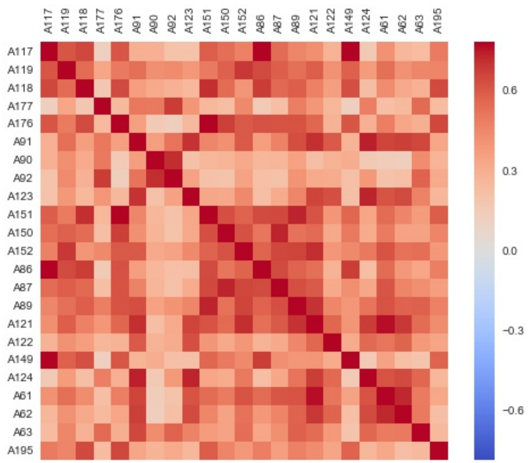
Inverse Euclidean distance



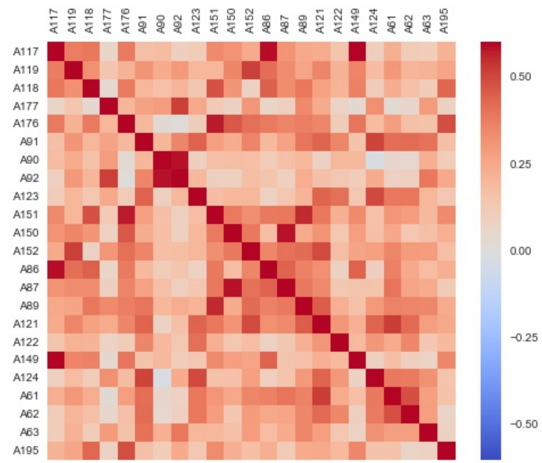
Inverse Manhattan distance

Results: methods in time domain

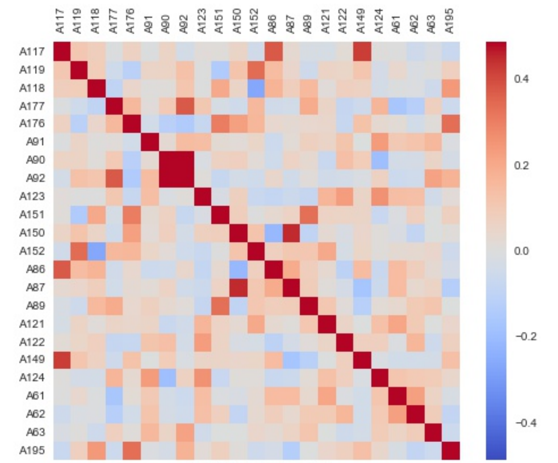
Example: MEG data, Anterior



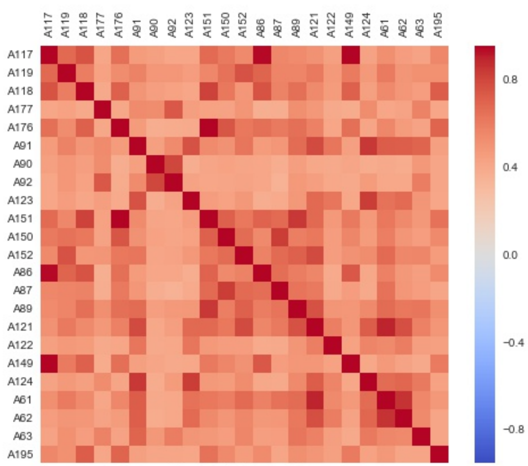
Pearson correlation



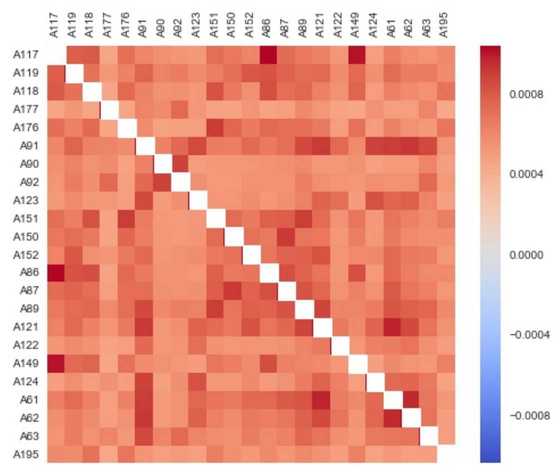
Partial partial correlation



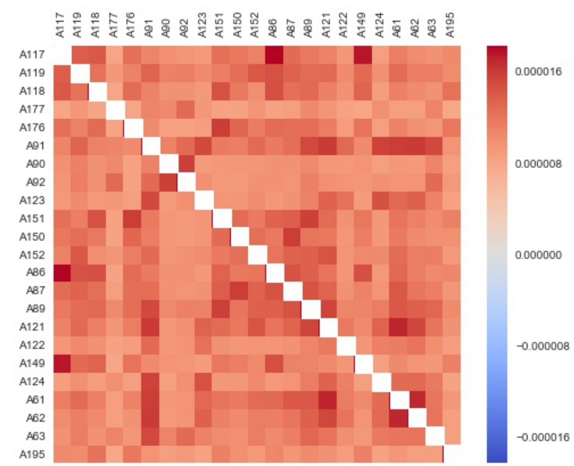
Partial correlation



Mutual information



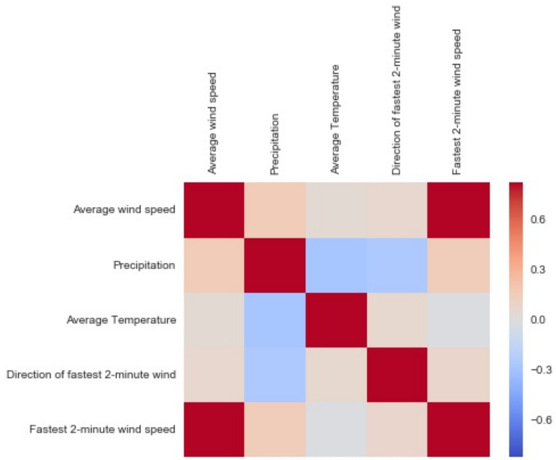
Inverse Euclidean distance



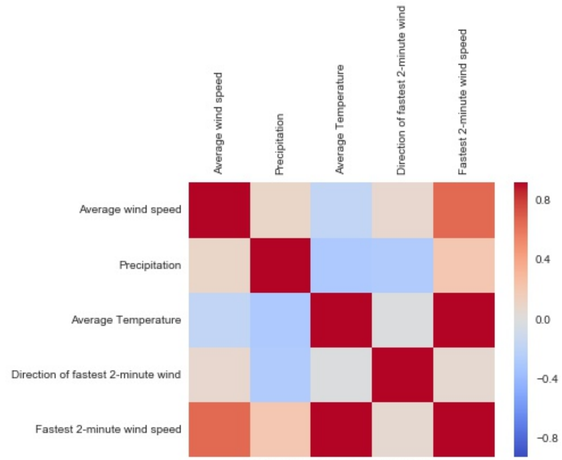
Inverse Manhattan distance

Results: methods in time domain

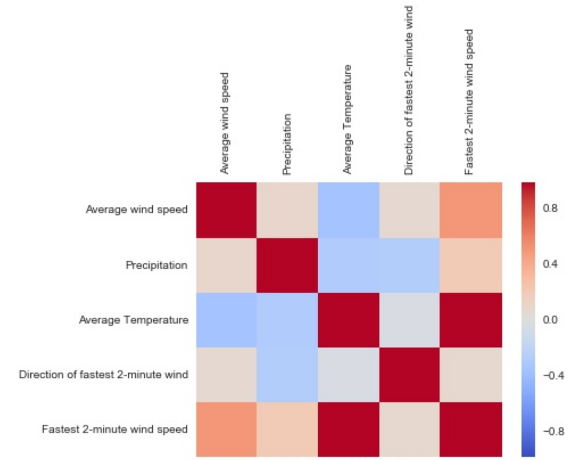
Example: Weather



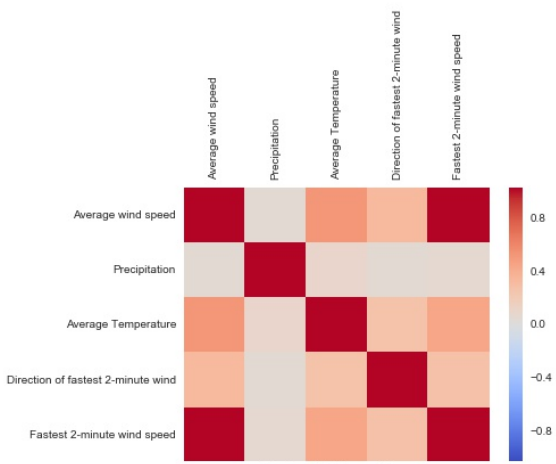
Pearson correlation



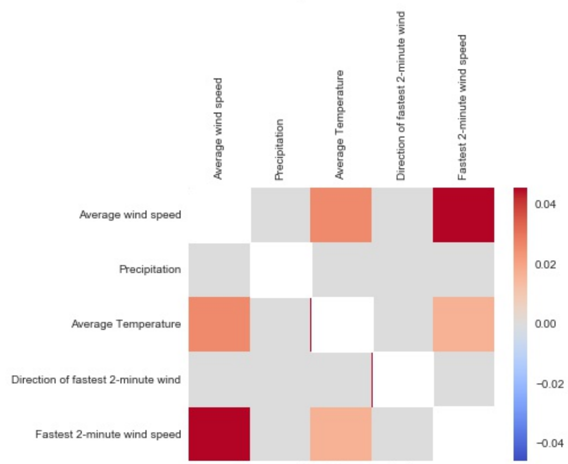
Partial partial correlation



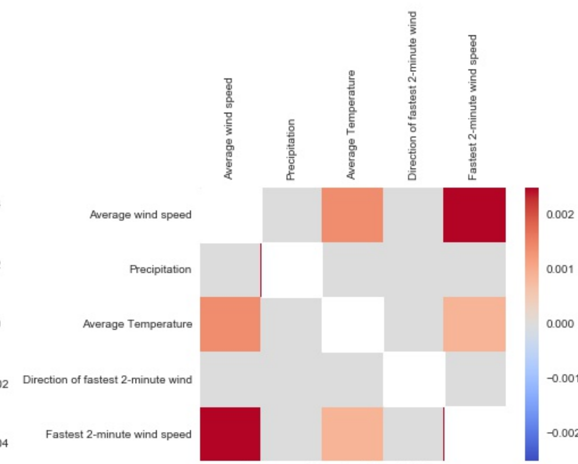
Partial correlation



Mutual information



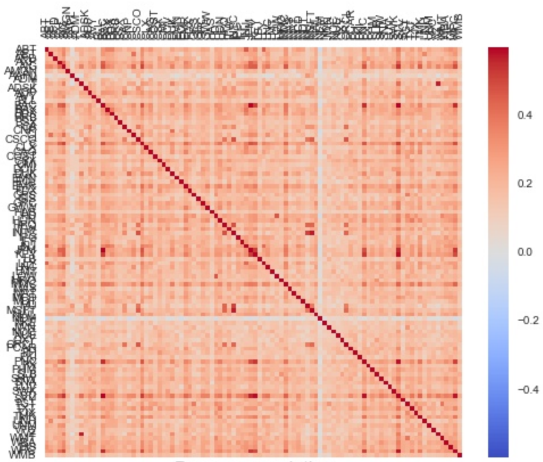
Inverse Euclidean distance



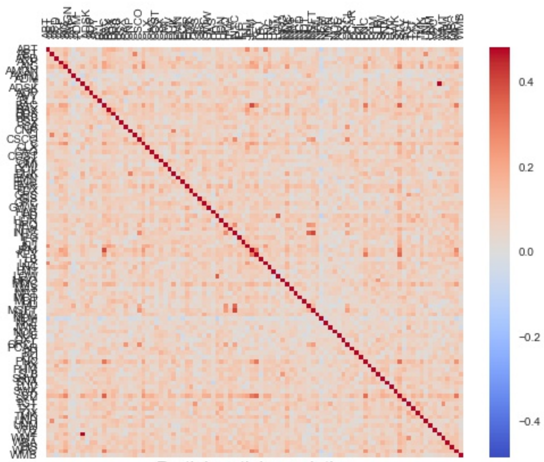
Inverse Manhattan distance

Results: methods in time domain

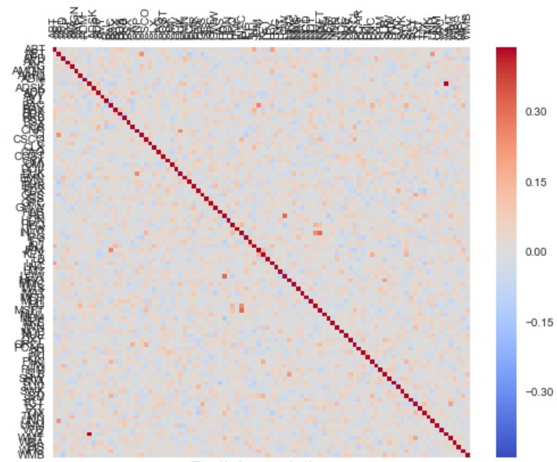
Example: Stock Exchange



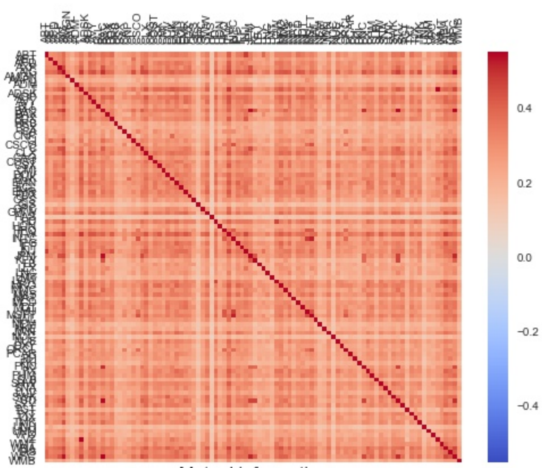
Pearson correlation



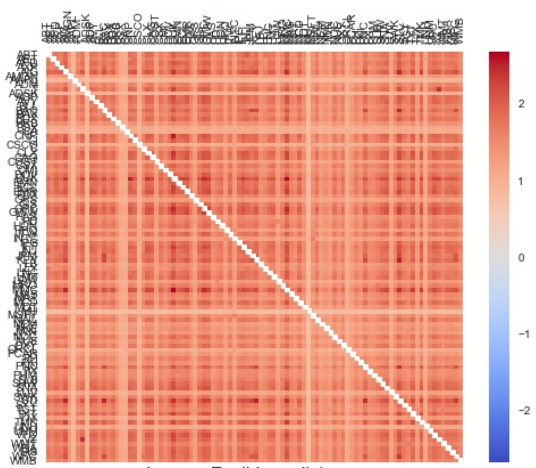
Partial partial correlation



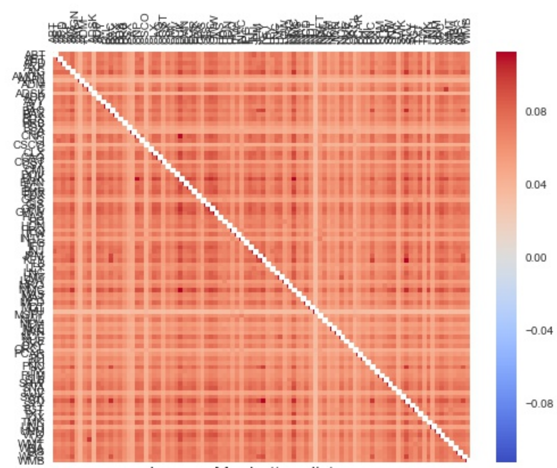
Partial correlation



Mutual information



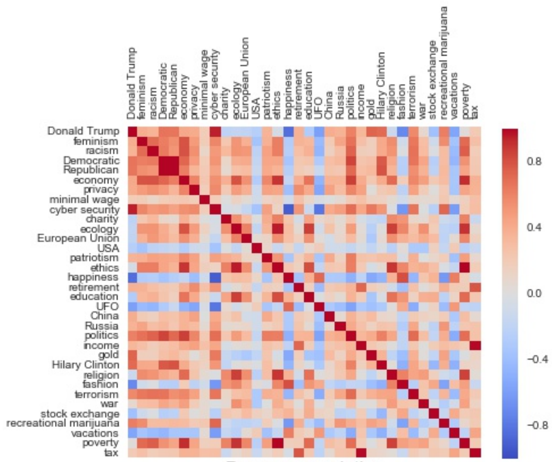
Inverse Euclidean distance



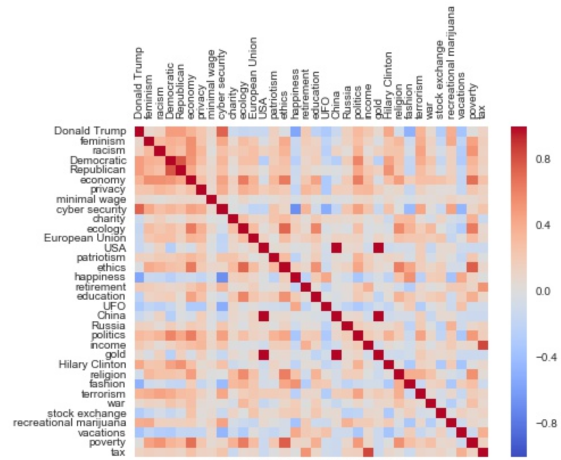
Inverse Manhattan distance

Results: methods in time domain

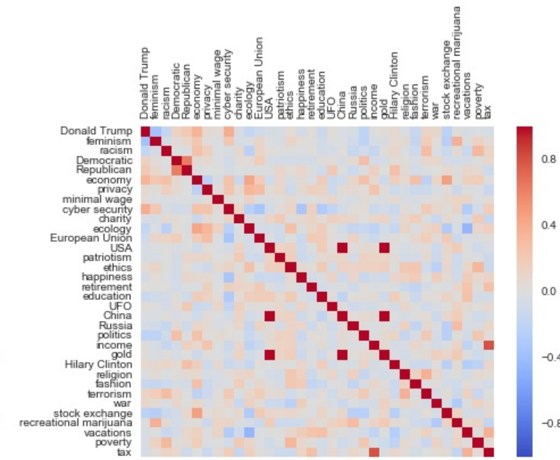
Example: Google trends



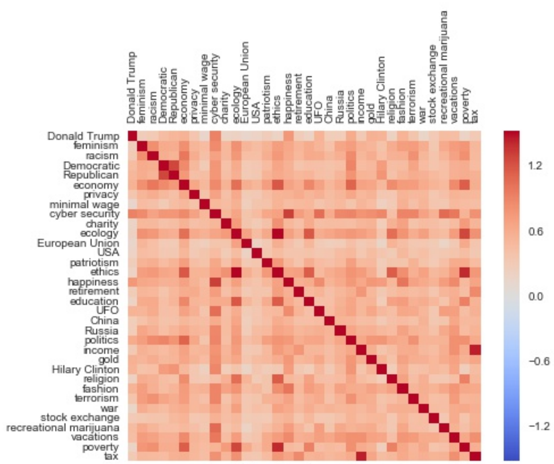
Pearson correlation



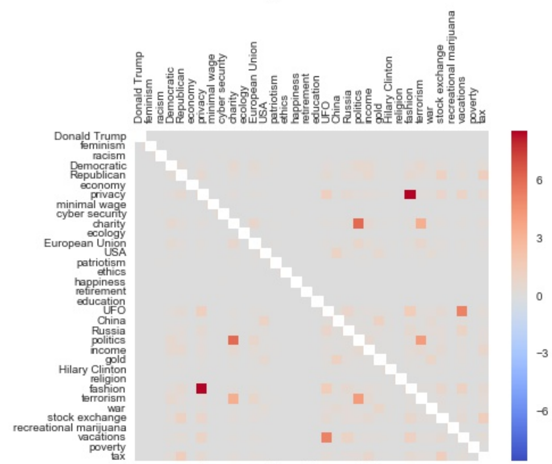
Partial partial correlation



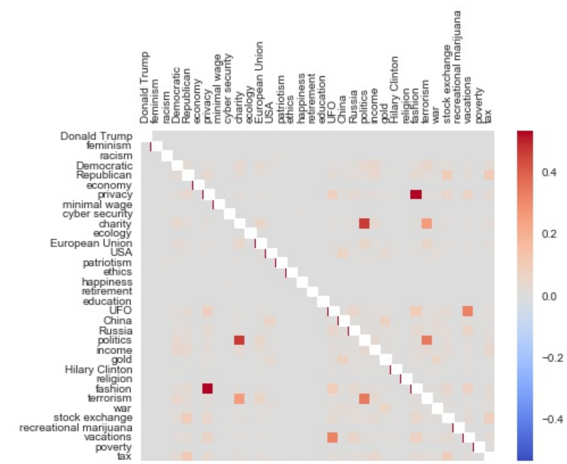
Partial correlation



Mutual information



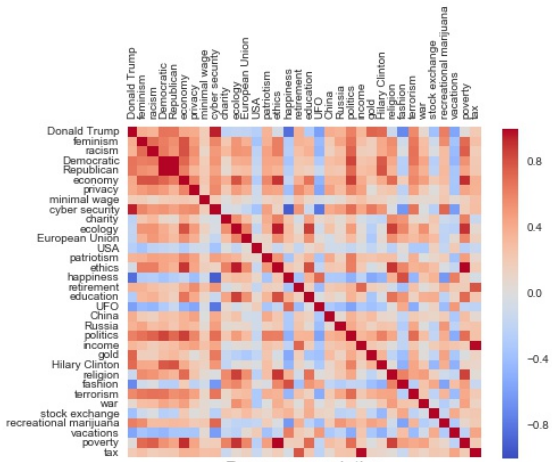
Inverse Euclidean distance



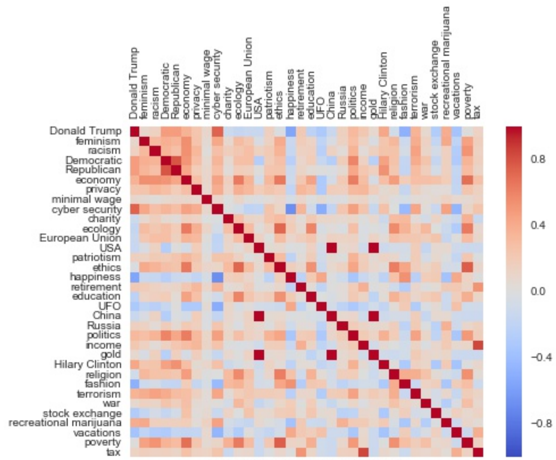
Inverse Manhattan distance

Results: methods in time domain

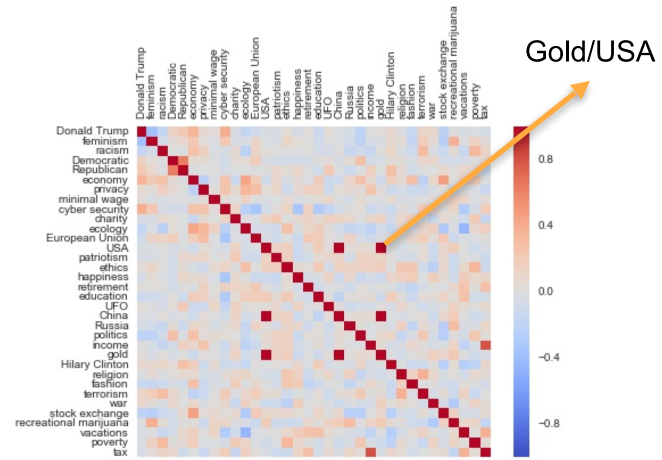
Example: Google trends



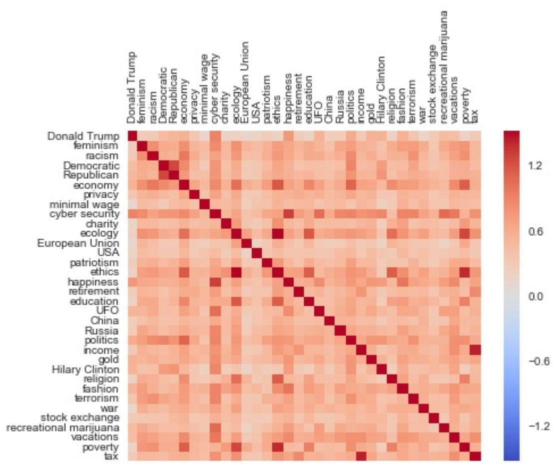
Pearson correlation



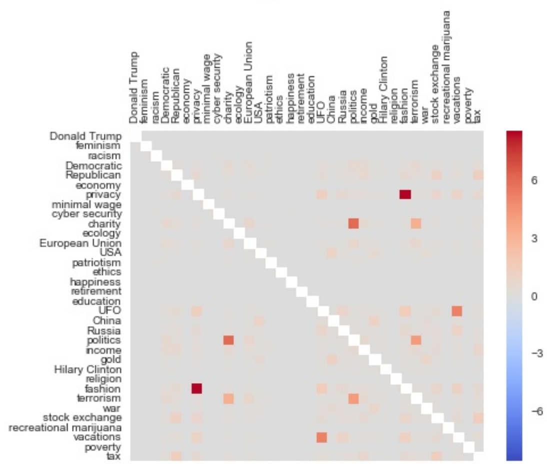
Partial partial correlation



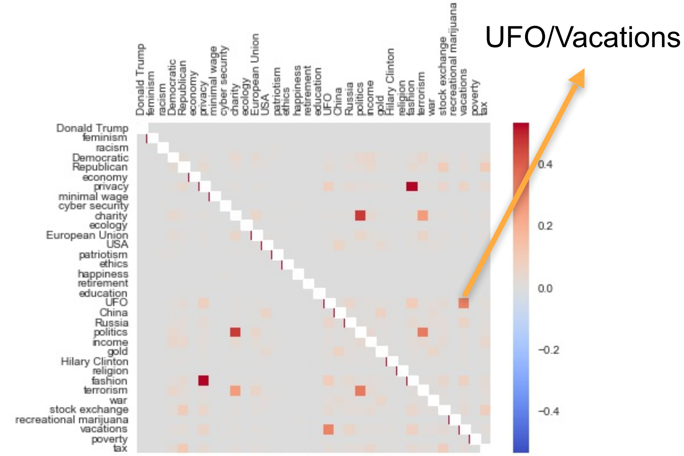
Partial correlation



Mutual information



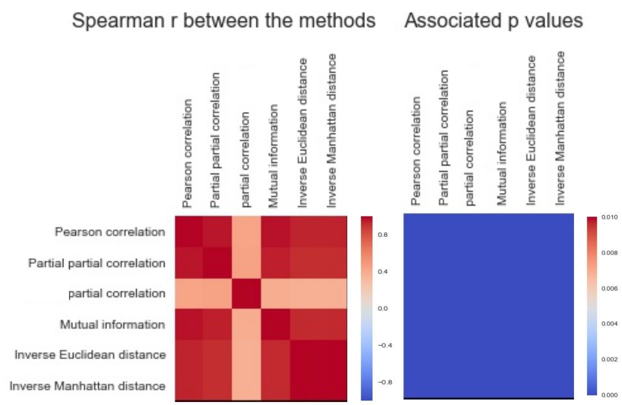
Inverse Euclidean distance



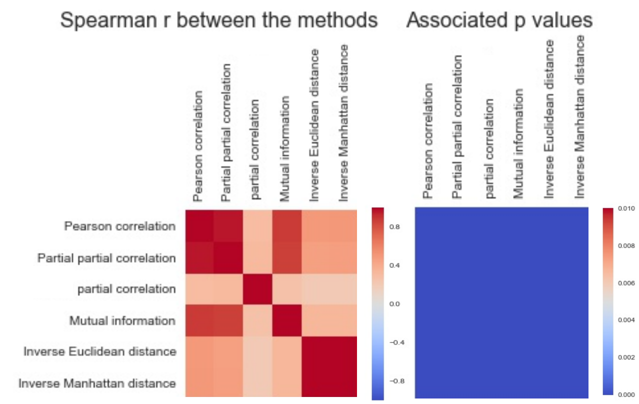
Inverse Manhattan distance

Results: FC methods across datasets

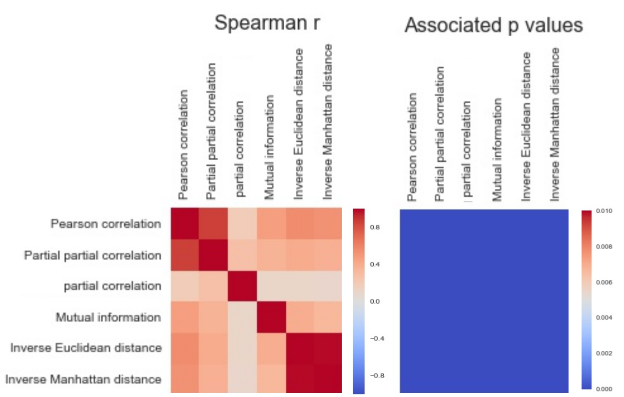
MEG, Anterior



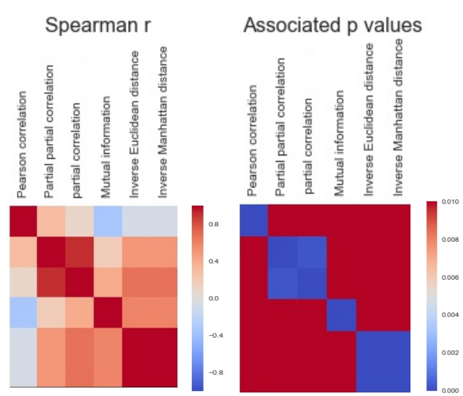
fMRI, Visual



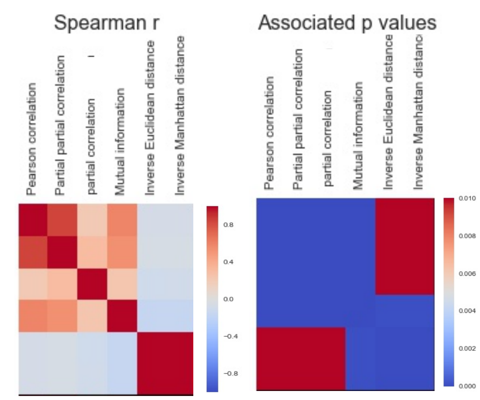
Stock Exchange



Weather

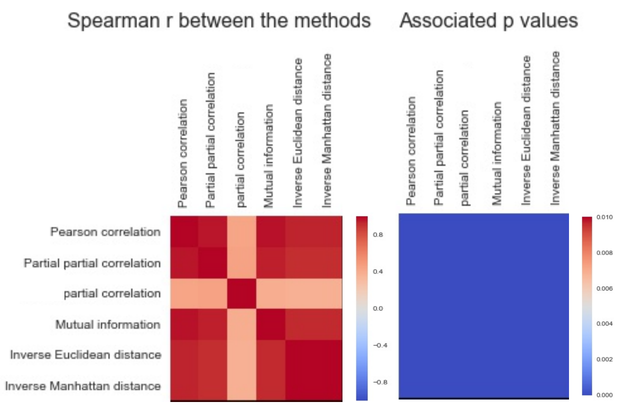


Google trends

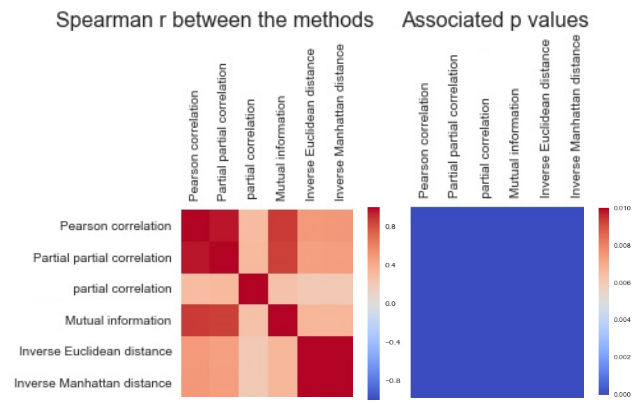


Results: FC methods across datasets

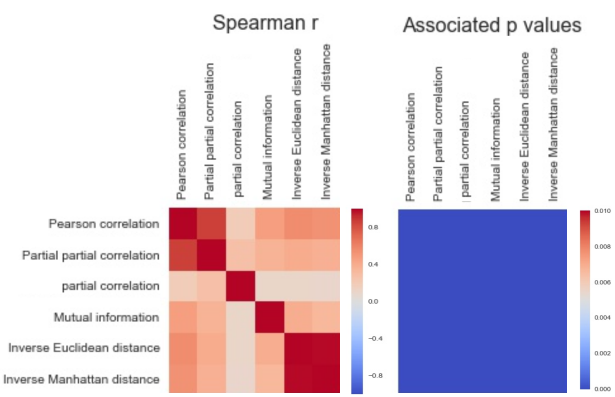
MEG, Anterior



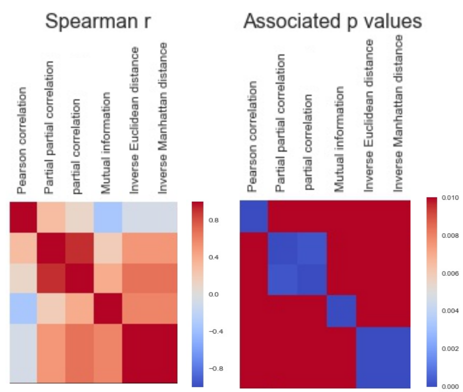
fMRI, Visual



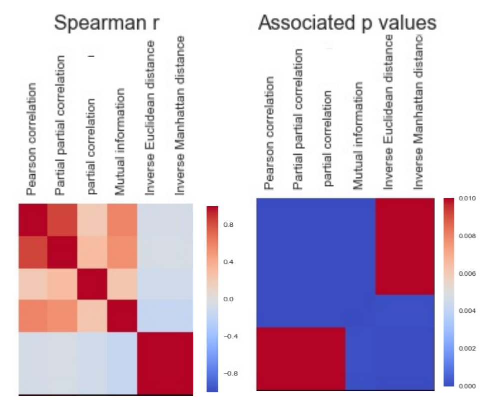
Stock Exchange



Weather

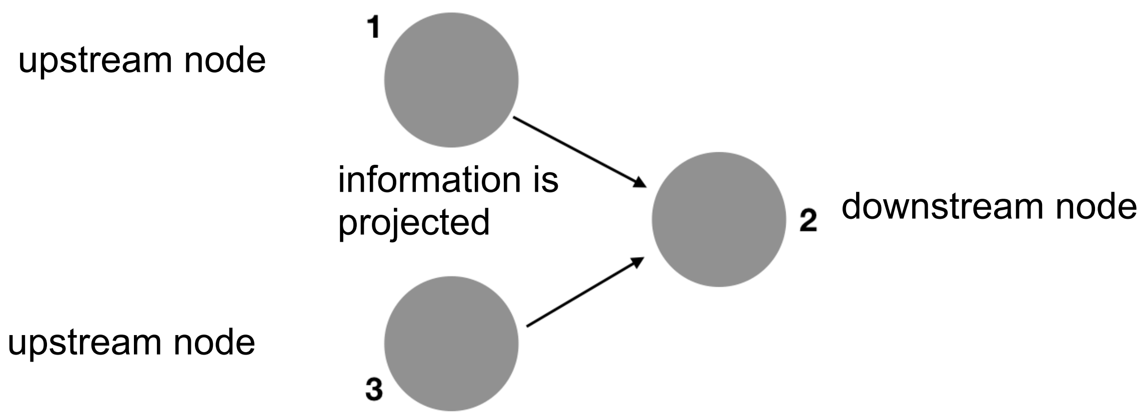


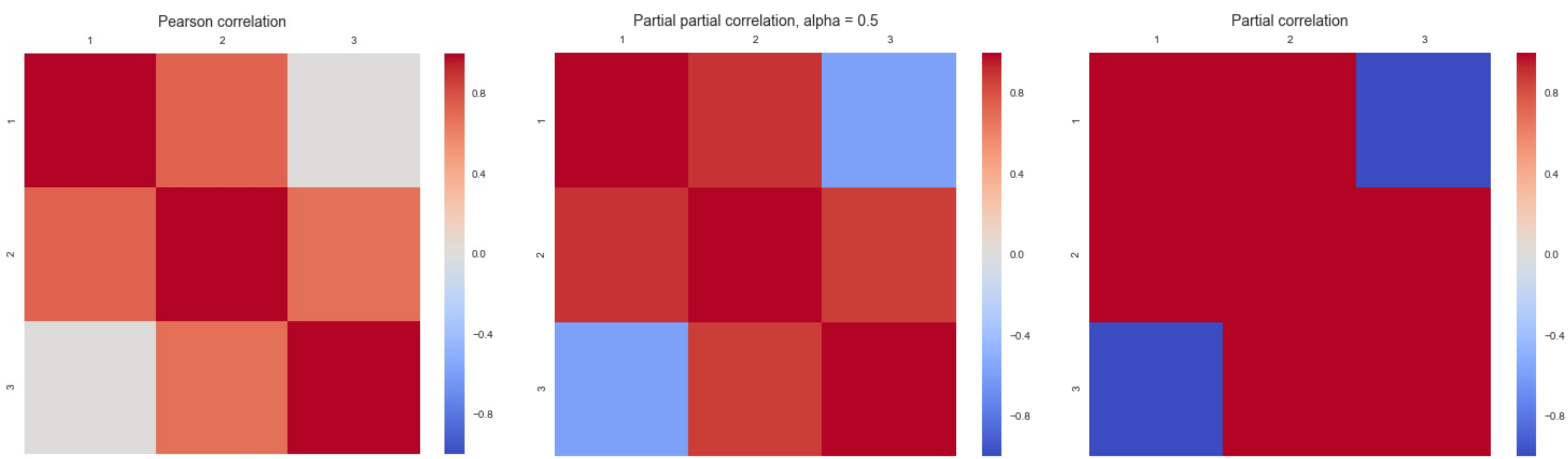
Google trends



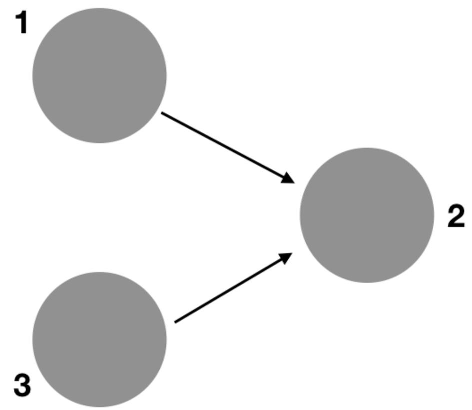
For datasets beyond neuroimaging, each method shows a very specific performance, which may suggest that they represent complementary information.

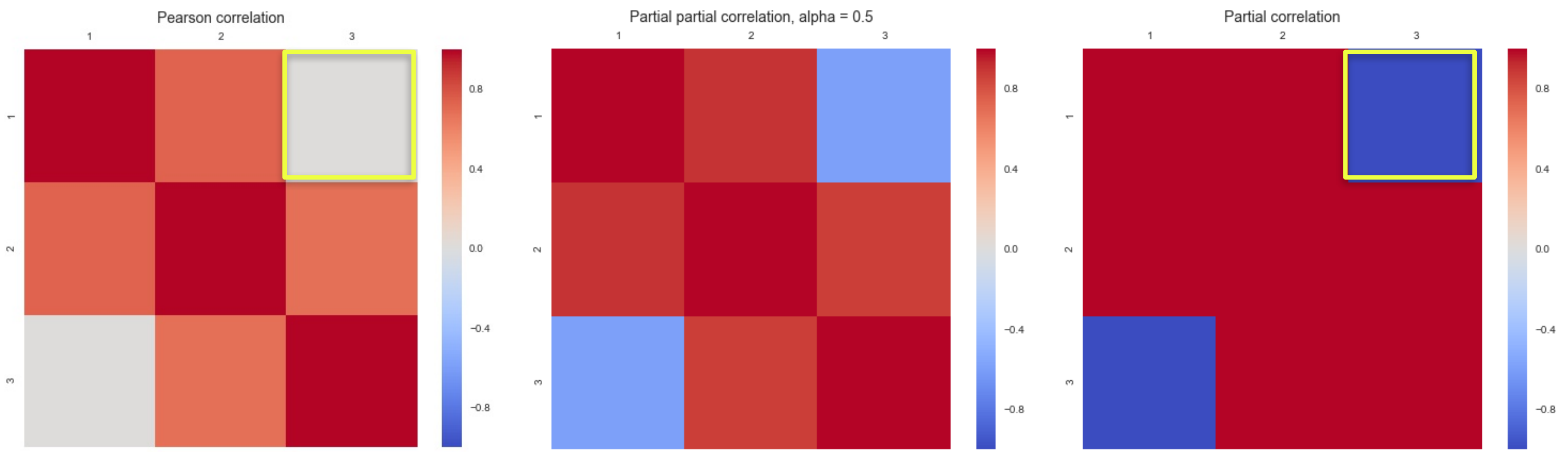
Berkson’s Paradox





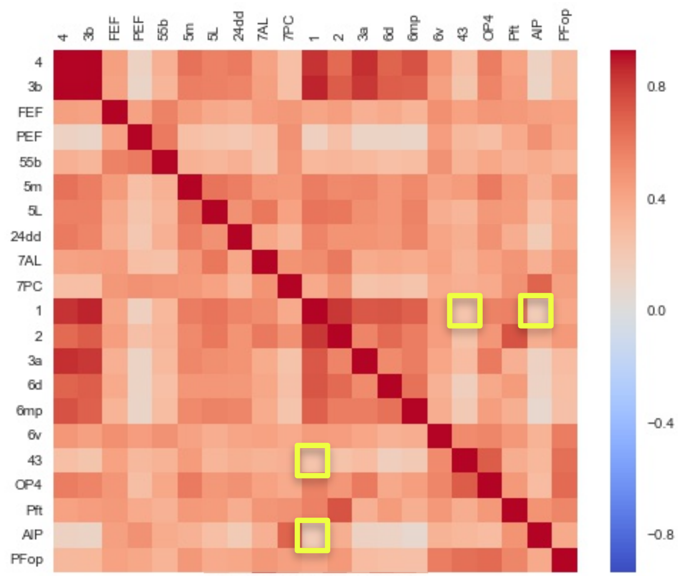
Berkson’s Paradox



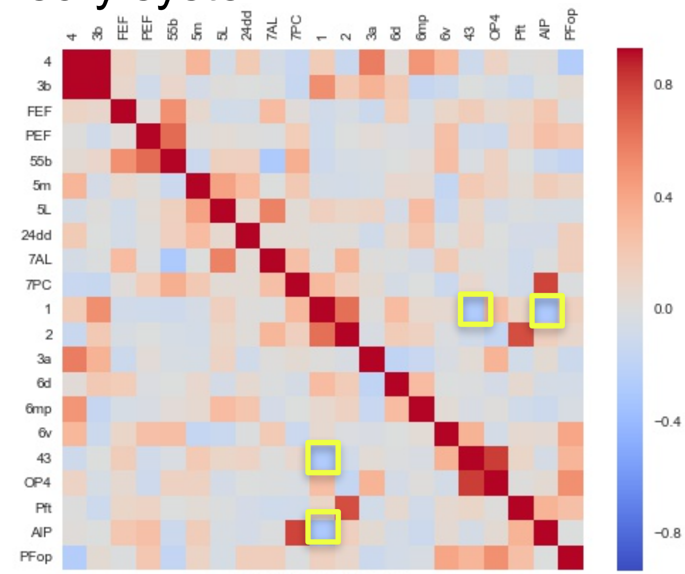


Results: Berkson's Paradox

fMRI data, somatosensory system:

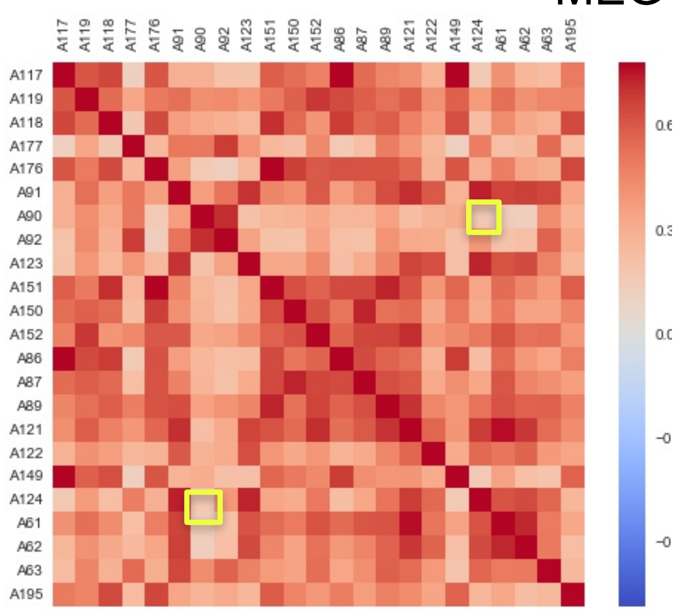


Pearson correlation

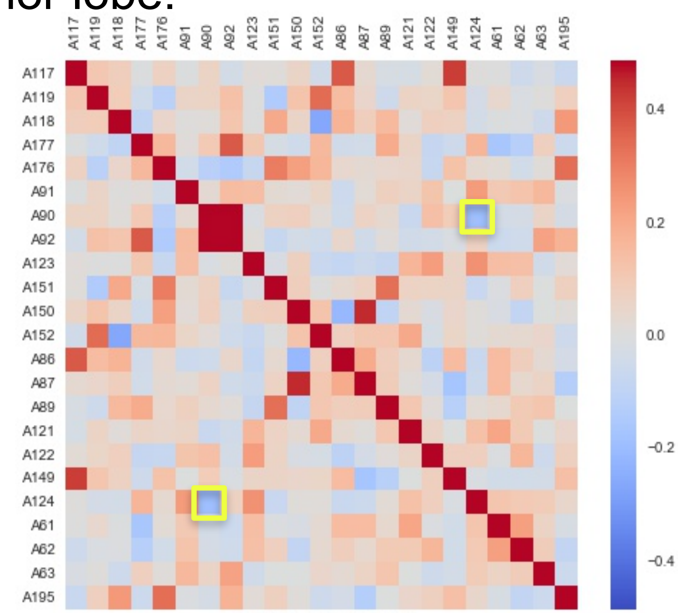


Partial correlation

MEG data, anterior lobe:



Pearson correlation

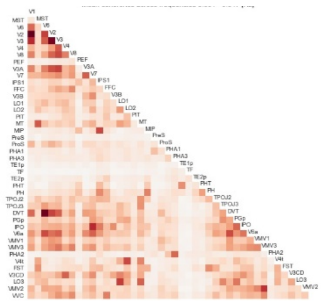


Partial correlation

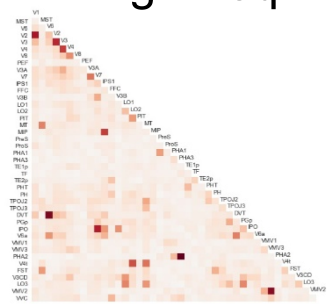
Results: Coherence

fMRI data, visual system (TR: 0.7-2sec)

Low freq

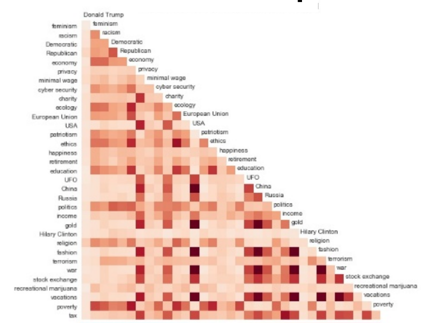


High freq

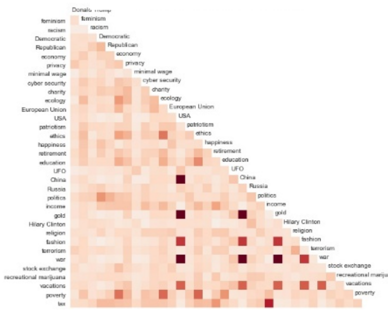


Google trends (weekly)

Low freq

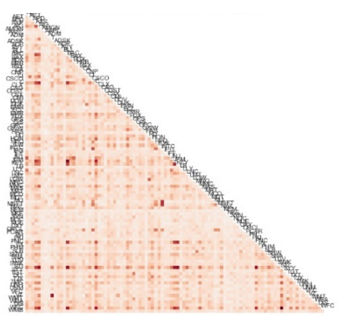


High freq

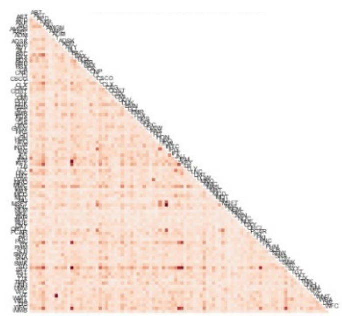


Stock Exchange (daily)

Low freq

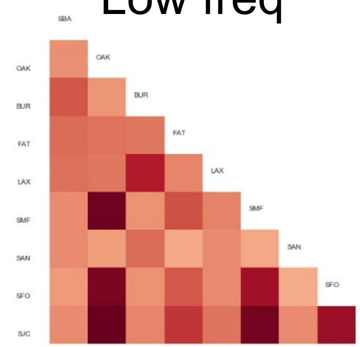


High freq

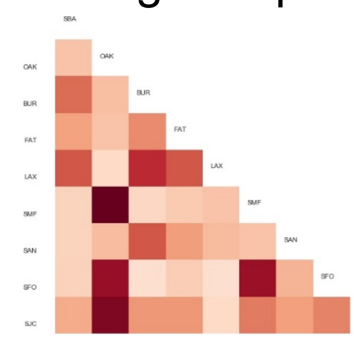


Weather (daily)

Low freq



High freq



Low -> 0Hz to half of the Nyquist frequency

High -> From half of the Nyquist frequency to the maximum frequency in the data.

Discussion

1. Distance based measures don't work (influence of detrending will be examined). They don't seem to reveal the underlying structure of the networks.
2. Partial correlation showed little correlation with other FC methods. It seems to be too conservative, specially for large networks. Partialized correlations may help.
3. Berkson's paradox is a real problem.
4. For datasets beyond neuroimaging, each method shows a very specific performance. Maybe they are complementary. It could be worth trying them separately.
5. We are in the process of extending this work by comparing Hierarchical functional connectomes generated by different FC methods, and by comparing dynamic FC across datasets.
6. We will add analysis on more datasets (e.g., mRNA, social interactions)

Thank you!

Thank you!

Acknowledgements

Daniele Marinazzo, PhD

Andrew Reid, PhD

organizers of the AoN Brainhack Warsaw 2017

Thank you!

Acknowledgements

Daniele Marinazzo, PhD

Andrew Reid, PhD

organizers of the AoN Brainhack Warsaw 2017

All the codes and results available at

https://github.com/cryptofan/FunctionalConnectivity_AoNBrainhackWarsaw/

Thank you!

Acknowledgements

Daniele Marinazzo, PhD

Andrew Reid, PhD

organizers of the AoN Brainhack Warsaw 2017

All the codes and results available at

https://github.com/cryptofan/FunctionalConnectivity_AoNBrainhackWarsaw/

Please contact us:

Anna Alińska: alinska.anna@gmail.com

Juan Camilo Avendaño: juancamiload@gmail.com

Agnieszka Tymorek: agnieszkatymorek@gmail.com

Natalia Bielczyk: natalia.bielczyk@gmail.com