Challenges in directed network analysis documented on climate reanalysis surface air temperature data

Monster journey

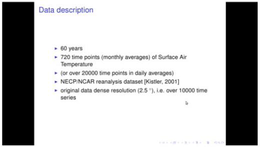
Jaroslav Hlinka

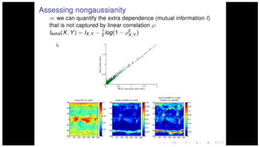
Institute of Computer Science, Academy of Sciences of the Czech Republic

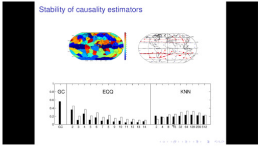
Dynamics Days
Berlin-Brandenburg 1-2. 10. 2013

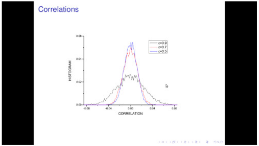
Data

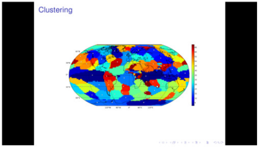
You are
HERE











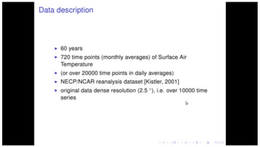
Data description

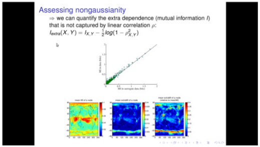
- $\sim$ 9 minutes, 213 time points of whole brain resting state brain activity
- 26 (12 males, 19-54 years) healthy volunteers
- 3T Siemens Magnetom Trio MRI scanner (GE-EPI, TR/TE=2500/30 ms, voxel size=3x3x3mm), a 3D high-resolution T1-weighted image was used for anatomical reference, slice-timing correction, motion correction, spatial normalization to MNI
- original data $\sim$ 20000 time series, dimensionality reduced to 90 time series by averaging over regions from the Automated Anatomical Labeling atlas
- orthogonalized wrt motion parameters, white matter and CSF signal

Data description

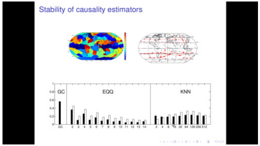
- 60 years
- 720 time points (monthly averages) of Surface Air Temperature
- (or over 20000 time points in daily averages)
- NECP/NCAR reanalysis dataset [Kistler, 2001]
- original data dense resolution (2.5 $^\circ$), i.e. over 10000 time series
- only anomalies with respect to average seasonal changes considered

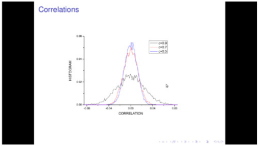
Dependence: how to measure?

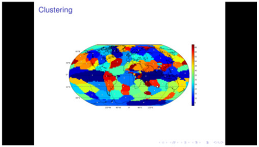




Now you are here!







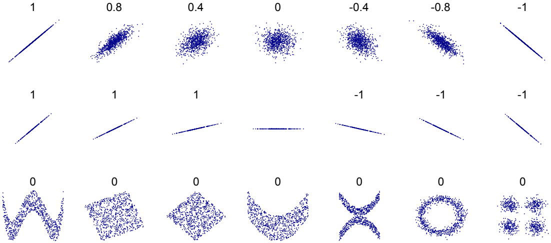
Dependence: how to measure?

Independence ($X \perp\!\!\!\perp Y$): $p(X, Y) = p(X)p(Y)$

Dependence measures:

Pearson's correlation $\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$

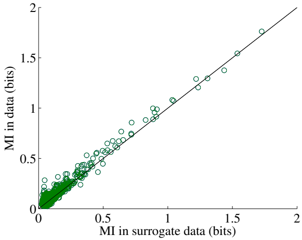
Mutual information: $I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)}\right)$

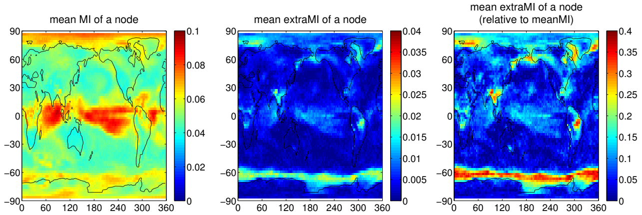


Dependence: Assessing nongaussianity

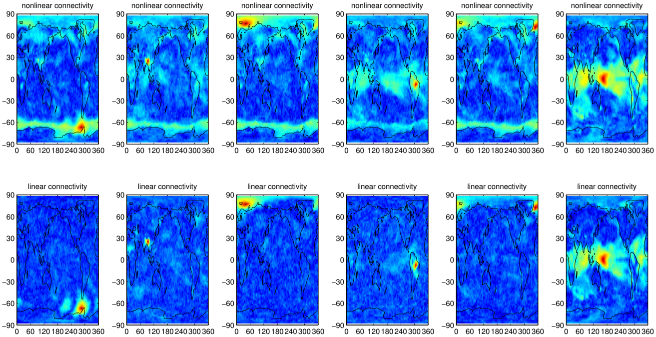
⇒ we can quantify the extra dependence (mutual information I) that is not captured by linear correlation ρ:

$$I_{extra}(X, Y) = I_{X,Y} - \frac{1}{2} log(1 - \rho^2_{X,Y})$$

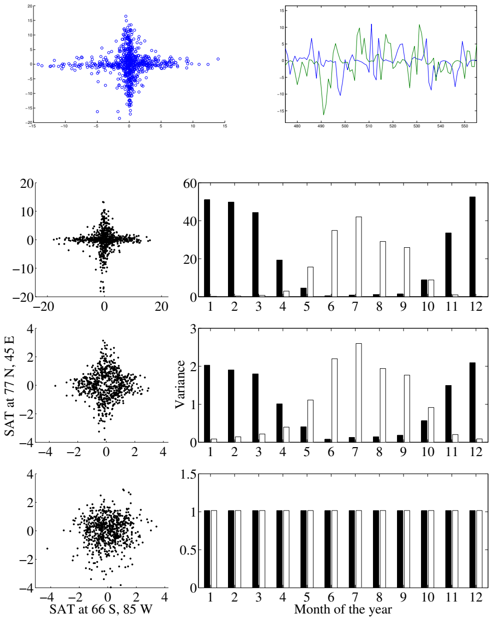




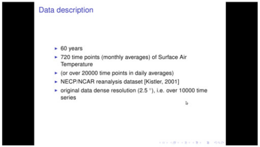
Dependence: nonlinearity examples

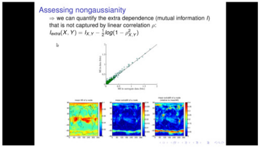


Dependence: Nongaussian ghost of seasonality

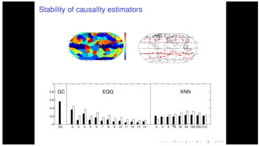


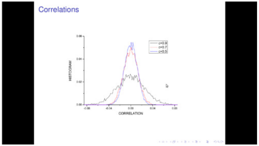
Taking into account autocorrelation

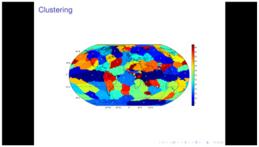




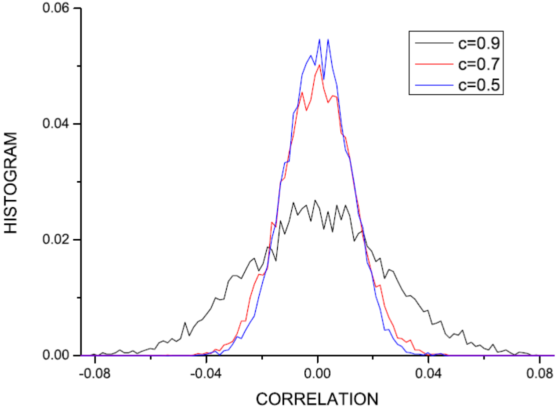
And now here



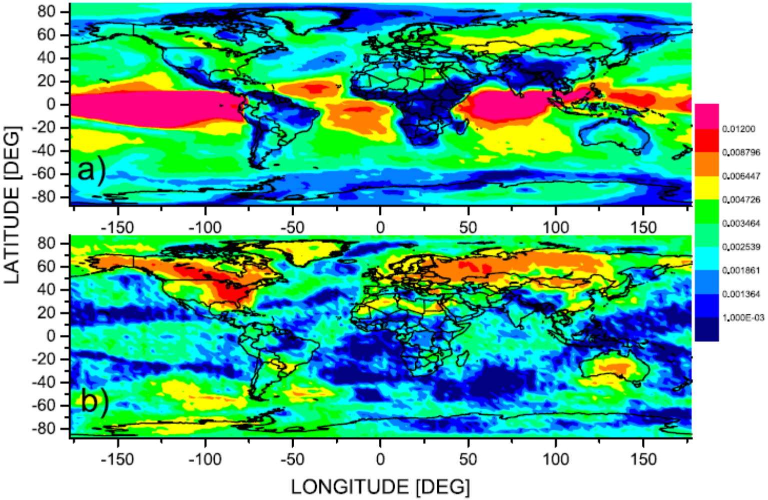




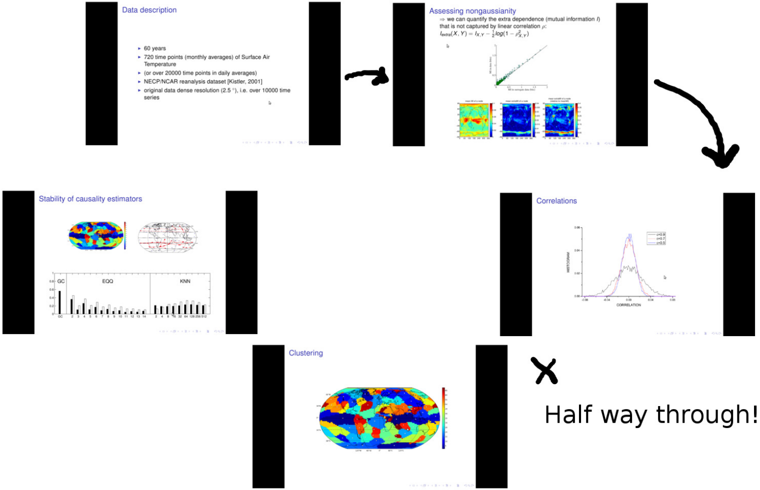
Correlation histogram for (autocorrelated) noise



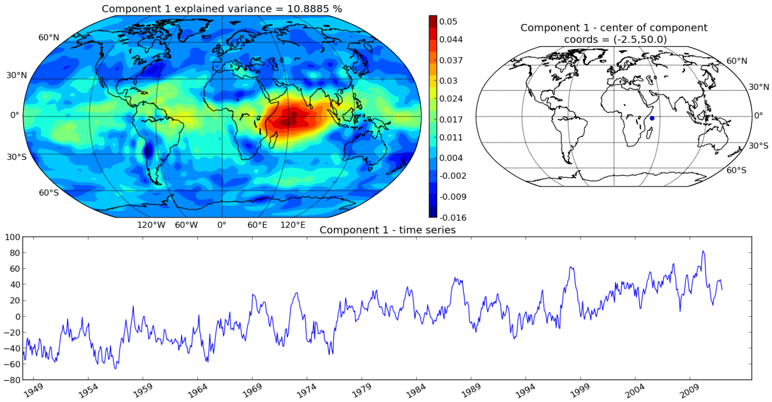
Average weighted correlation with(out) correction



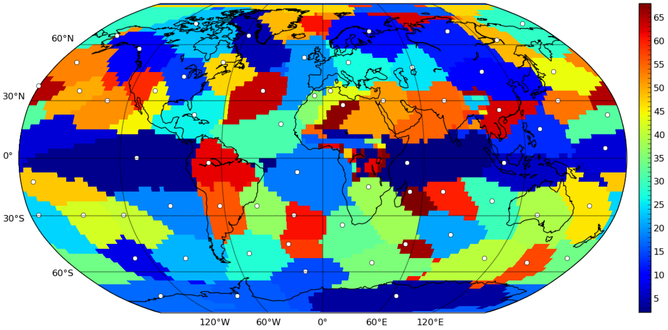
Dimensionality reduction



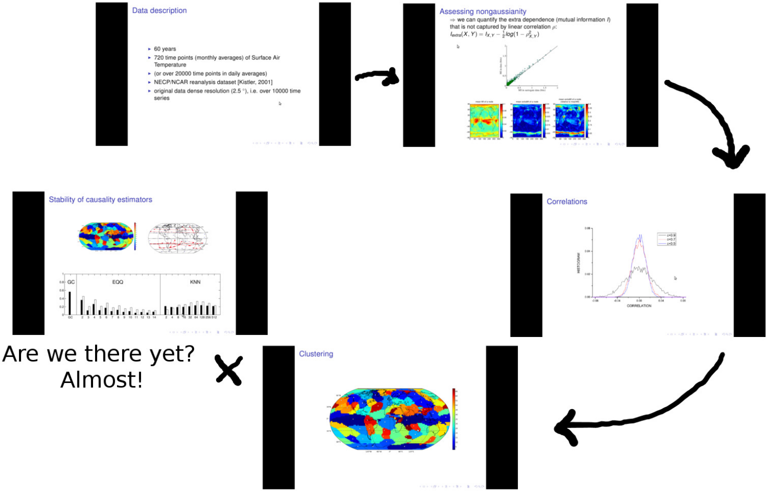
Decomposition



Clustering


60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
65
60
55
50
45
40
35
30
25
20
15
10
5

Interaction of subsystems



Causality

- Granger causality: X 'Granger causes' Y iff including the past of Y in a (linear) model of X significantly improves the model fit

$$\mathcal{F}_{Y \to X|Z} = \ln\left(\frac{|\Sigma(\varepsilon_t)|}{|\Sigma(\varepsilon'_t)|}\right)$$

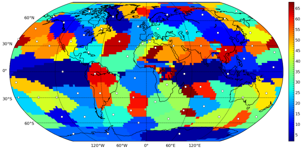
- Transfer entropy: the difference between entropies of the variable X conditioned (or not) on Y:

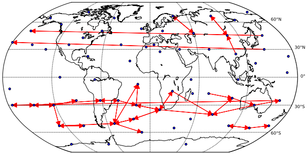
$$\mathcal{T}_{Y \to X|Z} = H(X|X^- \oplus Z^-) - H(X|X^- \oplus Y^- \oplus Z^-),$$

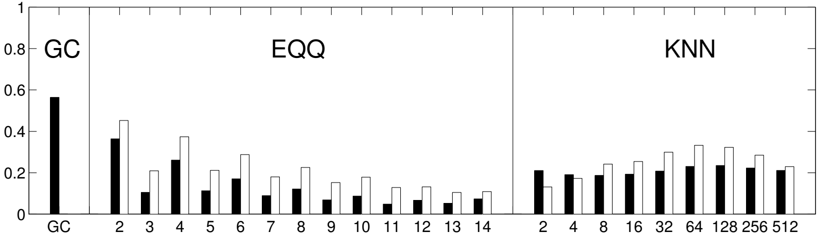
- for stationary linear Gaussian processes GC and TE **equivalent**

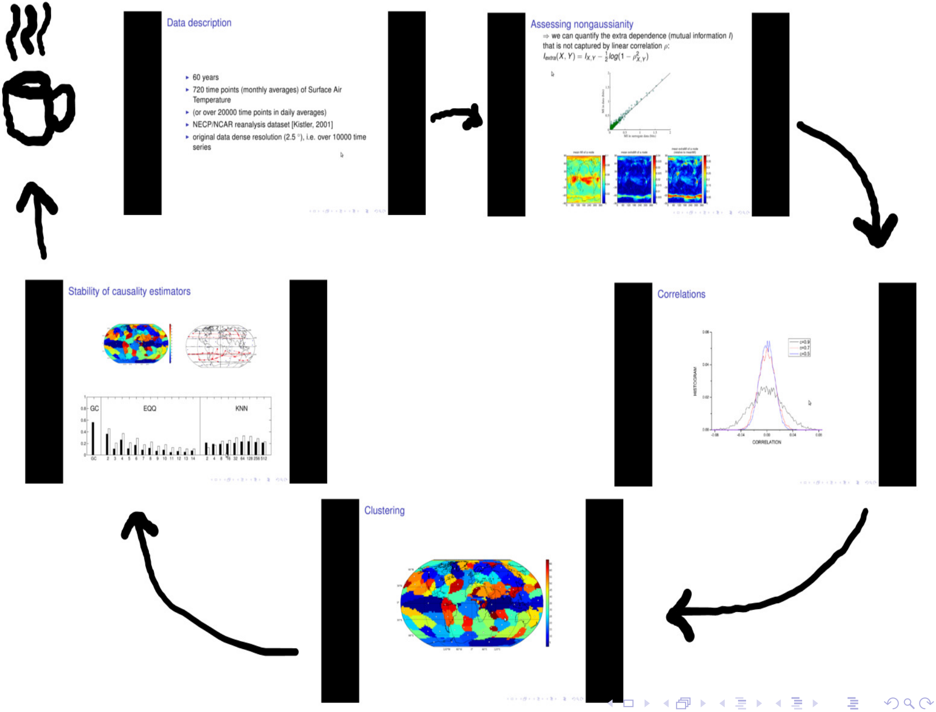
$$\mathcal{F}_{Y \to X|Z} = 2\mathcal{T}_{Y \to X|Z}$$

Stability of causality estimators








Data description
60 years
720 time points (monthly averages) of Surface Air Temperature
(or over 20000 time points in daily averages)
NECP/NCAR reanalysis dataset [Kistler, 2001]
original data dense resolution (2.5 °), i.e. over 10000 time series
Assessing nongaussianity
⇒ we can quantify the extra dependence (mutual information I) that is not captured by linear correlation ρ:
$I_{extra}(X,Y) = I_{X,Y} - \frac{1}{2}log(1-\rho^2_{X,Y})$
Correlations
HISTOGRAM
CORRELATION
c=0.9
c=0.7
c=0.5
Clustering
Stability of causality estimators
GC
EQQ
KNN

Thank you for your attention!

Thanks to my colleagues at ICS, Prague: **Milan Palus, Martin Vejmelka, David Hartman, Nikola Jajcay, Lucie Pokorna** and collaborators at IAS, Prague: **Dagmar Novotna** and PIK, Potsdam: **Jakob Runge, Juergen Kurths, Norbert Marwan**

This study was supported by the Czech Science Foundation project No. P103/11/J068. and No. 13-23940S

http://ndw.cs.cas.cz, hlinka@cs.cas.cz

References:

J. Hlinka, D. Hartman, M. Vejmelka, D. Novotna and M. Palus: Non-linear dependence and teleconnections in climate data: source. Climate Dynamics 2013

J. Hlinka, D. Hartman, M. Vejmelka, J. Runge, N. Marwan, J. Kurths and M. Palus: Reliability of inference of directed climate networks using conditional mutual information, Entropy, 2013

Hlinka, J., Palus, M., Vejmelka, M., Mantini, D., Corbetta, M. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? 2011, Neuroimage, 54, 2218-2225

Palus, M.; Hartman, D.; Hlinka, J. and Vejmelka, M. Discerning connectivity from dynamics in climate networks Nonlinear Processes in Geophysics, 2011, 18, 751-763