Causality and information transfer in systems with extreme events

Milan Paluš

Department of Complex Systems
Institute of Computer Science, Czech Acad. Sci.
Prague, Czech Republic
E-Mail: mp@cs.cas.cz

Support:
Praemium Academiae, Czech Academy of Sciences
Czech Science Foundation, project GA19-16066S
Causality analysis:
Can we identify causes of observed phenomena?
Granger causality

C. Granger, 2003 Nobel prize in economy

- causal variable can help to forecast the effect variable after other data has been first used
- generalization using information theory
  - conditional mutual information (transfer entropy)
  - \( I(y(t); x(t + \tau)|x(t), x(t - \eta), \ldots, x(t - (n - 1)\eta)) \)
    - \( y(t) \) – the cause (predictor)
    - \( x(t + \tau) \) – the effect (the future of the influenced variable)
    - \( x(t), x(t - \eta), \ldots, x(t - (n - 1)\eta) \) – condition – removes the influence of history of the influenced variable
- for Gaussian systems CMI \( \equiv \) TE \( \equiv \) Granger causality
- causality interpreted as information transfer
Mutual information

- mutual information
  
  \[ I(X; Y) = H(X) + H(Y) - H(X, Y) \]

- average amount of common information, contained in the variables \( X \) and \( Y \)

- measure of general statistical dependence

- \( I(X; Y) \geq 0 \)

- \( I(X; Y) = 0 \) iff \( X \) and \( Y \) are independent
Conditional mutual information

- Conditional mutual information $I(X; Y|Z)$ of variables $X$, $Y$ given the variable $Z$

\[ I(X; Y|Z) = H(X|Z) + H(Y|Z) - H(X, Y|Z) \]

- $Z$ independent of $X$ and $Y$

\[ I(X; Y|Z) = I(X; Y) \]

- $I(X; Y|Z) = I(X; Y; Z) - I(X; Z) - I(Y; Z)$

- “net” dependence between $X$ and $Y$ without possible influence of $Z$
random variable $X$ with sets of values $\Xi$ and PDF’s $p(x)$

- Shannon entropy

$$H(X) = - \sum_{x \in \Xi} p(x) \log p(x)$$

- Rényi entropy

$$H_\alpha(X) = \frac{1}{1 - \alpha} \log \sum_{x \in \Xi} p(x)^\alpha,$$

where $\alpha > 0$, $\alpha \neq 1$.

- As $\alpha \to 1$, $H_\alpha(X)$ converges to $H(X)$ which is Shannon entropy.

- Rényi’s measure satisfies $H_\alpha(x) \leq H_{\alpha'}(x)$ for $\alpha > \alpha'$. 

M. Paluš mp@cs.cas.cz
Rényi order parameter $\alpha > 0$, $\alpha \neq 1$

$$p(x)^\alpha$$

- $\alpha > 1$ “amplifies” center of PDF
- $0 < \alpha < 1$ “amplifies” tails of PDF
- tails – extreme events

Can conditional mutual information (transfer entropy) in Rényi concept uncover causes of extreme events?
DATA

- near surface air temperature (SAT) anomalies
- daily station data, example Frankfurt (50° 02’ 47”N, 8° 35’ 54”E, 112 m above sea level)
- daily NAO index
- daily blocking index for 9°E
Causality in Rényi concept

CAUSALITY btw NAO and WINTER (DJF) TEMPERATURE

CAUSALITY

NAO -> TEMPERATURE

MEASURED DATA
SURROGATE DATA

CAUSALITY

TEMPERATURE -> NAO

CMI

α - Z-score

CMI_α / CMI_1

Renyi entropy order α

M. Paluš
mp@cs.cas.cz
Causality in Rényi concept

CAUSALITY btw BLOCKING and MARCH-AUGUST TEMPERATURE

CAUSALITY
BLOCKING -> TEMPERATURE

CAUSALITY
TEMPERATURE -> BLOCKING

Renyi entropy order $\alpha$

CMI$_\alpha$/CMI$_1$

MEASURED DATA
SURROGATE DATA

Renyi entropy order $\alpha$

Z SIG LEVEL 2$\sigma$

M. Paluš
mp@cs.cas.cz
Causality in Rényi concept: NAO vs blocking

**CAUSALITY**

**NAO -> TEMPERATURE**

CMI $\alpha$ - Z-score

**CAUSALITY**

**BLOCKING -> TEMPERATURE**

CMI $\alpha$ / CMI$_1$

significant in tail area

M. Paluš
mp@cs.cas.cz
Experiment with Rényi CMI

RESULTS

- NAO significant causality mainly for $\alpha > 1$
- blocking significant causality mainly for $\alpha < 1$
- Rényi CMI normalized by Shannon CMI for $\alpha < 1$ much larger for blocking

CAN WE INFER

- NAO is causal to temperature, but not causing extremes
- blocking is causal mainly with respect to extremes

COMPARE

- distribution of SATA for different NAO/blocking conditions
- SATA distribution with Gaussian with the same mean and variance
- tails of data vs Gaussian distribution

M. Paluš mp@cs.cas.cz
Conditional distributions

NAO - WINTER (DJF) TEMPERATURE ANOMALY

TEMPERATURE ANOMALY [$^\circ$C]

DATA HISTOGRAM

GAUSSIAN

NAO neutral

NAO positive

NAO negative

M. Paluš
mp@cs.cas.cz
Conditional distributions

Blocking - March-August temperature anomaly

Data histogram vs. Gaussian distribution

No blocking

Blocking

Temperature anomaly [°C]
RESULTS

- NAO shifts means/whole histogram
- NAO does not specifically cause tails higher than normal
- blocking only slightly shifts means/whole histogram
- blocking causes tails higher than normal

- Information-theoretic approach to causality in Rényi concept seems promising in identification of causes of extreme events
- many technical problems, however, should be solved in order to avoid false results
- research in progress, publication in preparation
CONCLUSION

Thank you for your attention

Interested in postdoc, PhD position?

Interested in part-time long distance job?
(Applicable in Covid-19 border closure conditions.)

mp@cs.cas.cz
http://www.cs.cas.cz/mp/

Milan Paluš mp@cs.cas.cz