LPCMCI: Causal Discovery in Time Series with Latent Confounders

Andreas Gerhardus[1] and Jakob Runge[2]

[1]German Aerospace Center (DLR), Institute of Data Science, Jena, Germany (andreas.gerhardus@dlr.de)
[2]German Aerospace Center (DLR), Institute of Data Science, Jena, Germany (jakob.runge@dlr.de)

The quest to understand cause and effect relationships is at the basis of the scientific enterprise. In cases where the classical approach of controlled experimentation is not feasible, methods from the modern framework of causal discovery provide an alternative way to learn about cause and effect from observational, i.e., non-experimental data. Recent years have seen an increasing interest in these methods from various scientific fields, for example in the climate and Earth system sciences (where large scale experimentation is often infeasible) as well as machine learning and artificial intelligence (where models based on an understanding of cause and effect promise to be more robust under changing conditions.)

In this contribution we present the novel LPCMCI algorithm for learning the cause and effect relationships in multivariate time series. The algorithm is specifically adapted to several challenges that are prevalent in time series considered in the climate and Earth system sciences, for example strong autocorrelations, combinations of time lagged and contemporaneous causal relationships, as well as nonlinearities. It moreover allows for the existence of latent confounders, i.e., it allows for unobserved common causes. While this complication is faced in most realistic scenarios, especially when investigating a system as complex as Earth's climate system, it is nevertheless assumed away in many existing algorithms. We demonstrate applications of LPCMCI to examples from a climate context and compare its performance to competing methods.

Related reference:
Gerhardus, Andreas and Runge, Jakob (2020). High-recall causal discovery for autocorrelated time series with latent confounders. In Advances in Neural Information Processing Systems 33 pre-proceedings (NeurIPS 2020).

LPCMCI: Causal Discovery in Time Series with Latent Confounders

Andreas Gerhardus & Jakob Runge
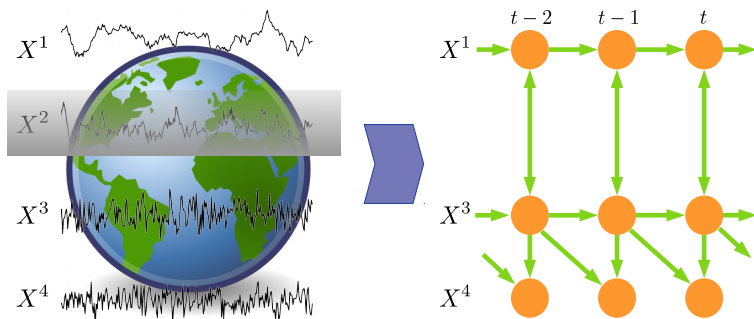**Institute of Data Science, German Aerospace Center (DLR)**

DLR Deutsches Zentrum für Luft- und Raumfahrt — German Aerospace Center

EGU General Assembly 2021

(1) Problem setting

Learn causal structure of multivariate time series in form of a causal graph from observational data …

… while allowing for arbitrary non-linear functional relationships and **latent (hidden) confounders**.
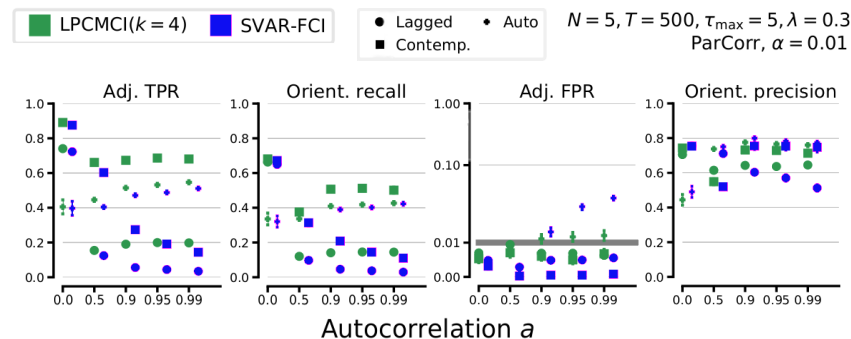
(2) Our contribution

Our novel **LPCMCI algorithm**(*) strongly **outperforms competing methods in terms of detection power** [1].

(*) Available at GitHub:
https://github.com/jakobrunge/tigramite/tree/master/neurips2020

(3) Numerical experiments

(4) Real data applications

An application to a river discharge dataset is demonstrated in our paper [1], more applications are subject of future work.

References:

[1] Gerhardus, A. and Runge, J. (2020). High-recall causal discovery for autocorrelated time series with latent confounders. In *Advances in Neural Information Processing Systems 33*.

[2] Baseline method SVAR-FCI: Malinsky, D. and Spirtes, P. (2018). Causal Structure Learning from Multivariate Time Series in Settings with Unmeasured Confounding. In *Proceedings of 2018 ACM SIGKDD Workshop on Causal Discovery*.

LPCMCI: Causal Discovery in Time Series with Latent Confounders

Display material for contribution EGU21-8259 at EGU General Assembly 2021

Andreas Gerhardus and Jakob Runge

andreas.gerhardus@dlr.de, jakob.runge@dlr.de

April 29, 2021

Climate Informatics Group
Institute of Data Science
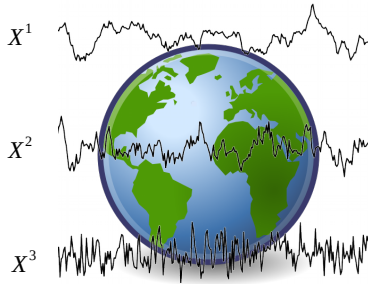German Aerospace Center

Knowledge for Tomorrow

DLR

Introduction

Motivation: Complex dynamics of the climate system

System of interest:

X^1

X^2

X^3

Goal:
Contribute to a better understanding of Earth's complex weather and climate system.

Climate Informatics in general:

Use modern tools of machine learning, statistics, and data science to aid climate and Earth system sciences.

Focus of the Climate Informatics Group @DLR Jena[*]:

- Development of methods
- Provisioning of open-source software implementations[†] for application by domain scientists
- Methods based on the modern **causal inference** framework

[*]www.climateinformaticslab.com
[†]https://github.com/jakobrunge/tigramite

Causal discovery

Causal inference and causal discovery

Causal inference:

- Defines notions of *cause* and *effect* in a mathematical framework.

- Casts causal questions within this framework.

- Specifies assumptions and develops methods for answering these questions.

Important sub-field: Causal discovery

- Specifies assumptions and develops methods for learning cause and effect relationships from observational data.

Correlation is not causation:
Statistical dependencies in observational data do not by themselves imply causal relationships.

> ⇒ Need assumptions to connect stat. dependencies and causation

Working definition of causality:
Variable X causes variable Y if an experimental manipulation that changes X (and only X) leads to a change of Y.

> ⇒ experimental mode of inferring causation

A theory of causality:
Framework of causal inference, largely developed and popularized by Judea Pearl, Peter Spirtes, Clark Glymour, Richard Scheines.

Textbooks: [Pearl, 2000, Spirtes et al., 2000, Peters et al., 2017].

Why is causal knowledge important?

Scientific understanding:
Knowledge of cause and effect relationships is an essential part of the
physical understanding of natural processes.

Robust prediction & forecasting:
Predictive systems consistent with the underlying causal structures are
thought to be more robust under changing environmental conditions.

Evaluation the effect of actions:
Questions of the type *What will happen if do ...?* are of causal nature.

Attribution:
Questions of the type *Why did this event happen?* are of causal nature.

Learning causal relationships from statistical independencies

Independence-based causal discovery:
Learn causal graph from statistical tests of (conditional) independencies*
in observational data

$$\Rightarrow \textit{CI-based} \text{ causal discovery}$$

Enabling assumptions:

1. Data is generated by a (unknown) structural causal model[1]

2. No *accidental* independencies[2] $\Rightarrow$ more on this later

3. Optional: No unobserved confounders[3] $\Rightarrow$ more on this later

---

[1] This implies the *causal Markov assumption*.
[2] Known as *faithfulness*.
[3] Known as *causal sufficiency*.

See [Pearl, 2000, Spirtes et al., 2000, Peters et al., 2017] for more details.

Causal graphs and (conditional) independencies

Fact:

The structure of the causal graph often has observable implications in terms of (conditional) independencies in the observed data.

Intuition:

- Statistical dependencies derive from causal relationships
- Conditioning can block and open the *flow of information*

Example:

- X influences Y: $X \not\!\perp\!\!\!\perp Y$
- Y influences Z: $Y \not\!\!\perp\!\!\!\perp Z$
- X influences Z through Y: $X \not\!\!\perp\!\!\!\perp Z$
- Knowing Y, X does not say more about Z: $X \perp\!\!\!\perp Z \mid Y$

Causal graphs and (conditional) independencies

Example:

- X influences Y: $\qquad\qquad\qquad\qquad\qquad\qquad$ $X \not\perp\!\!\!\perp Y$

- Y influences Z: $\qquad\qquad\qquad\qquad\qquad\qquad$ $Y \not\perp\!\!\!\perp Z$

- X influences Z through Y: $\qquad\qquad\qquad\quad$ $X \not\perp\!\!\!\perp Z$

- Knowing Y, X does not say more about Z: $\quad$ $X \perp\!\!\!\perp Z \mid Y$

General rule: d-separation
Graphical criterion to read off all (conditional) independencies implied by the structure of a given causal graph [Pearl, 1985, Pearl, 1988].

Assumption of no accidental independencies:
There are no independencies beyond those implied by the causal graph.

CI-based causal discovery without unobserved confounders

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

CI-based causal discovery without unobserved confounders

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 1:

Test decisions:

- $X \not\perp\!\!\!\perp Y$
- $Y \not\perp\!\!\!\perp Z$
- $X \perp\!\!\!\perp Z$

Possible causal graphs:

$X \longrightarrow Y \longleftarrow Z$

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 2:

Test decisions:

- $X \not\perp\!\!\!\perp Y$
- $Y \not\perp\!\!\!\perp Z$
- $X \not\perp\!\!\!\perp Z$
- $X \perp\!\!\!\perp Z \mid Y$

Possible causal graphs:
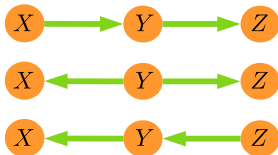
observationally equivalent graphs

Unobserved confounders make causal discovery more difficult

Without unobserved confounders:

$X \not\perp Y \qquad \Rightarrow$

$X \longrightarrow Y$ or

$X \longleftarrow Y$

With unobserved confounders:

$X \not\perp Y \qquad \Rightarrow$

$X \longrightarrow Y$ or

$X \longleftarrow Y$ or

$X \longleftarrow L \longrightarrow Y$

Unobserved confounders make causal discovery more difficult

Without unobserved confounders:

$X \not\perp\!\!\!\perp Y \qquad \Rightarrow$

$X \longrightarrow Y$ or

$X \longleftarrow Y$

With unobserved confounders:

$X \not\perp\!\!\!\perp Y \qquad \Rightarrow$

$X \longrightarrow Y$ or

$X \longleftarrow Y$ or

$X \longleftrightarrow Y$
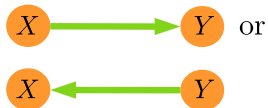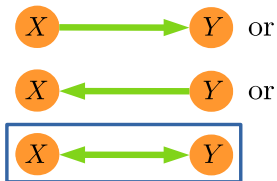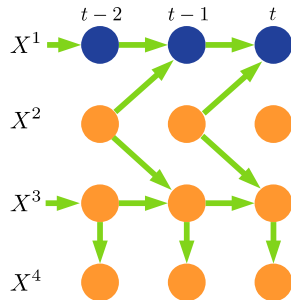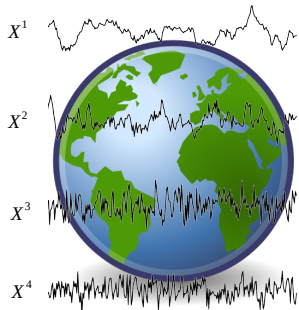
Causal discovery for time series
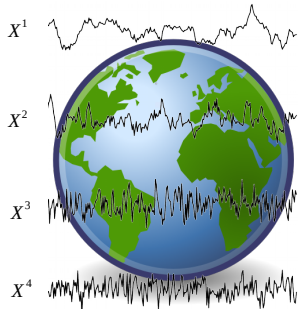
Particularities:

- Variables are resolved in time
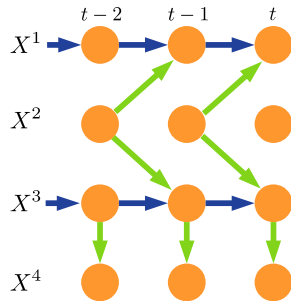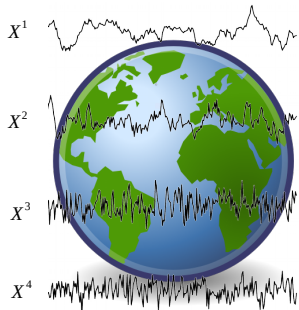
Particularities:

- Variables are resolved in time
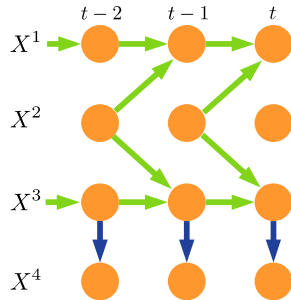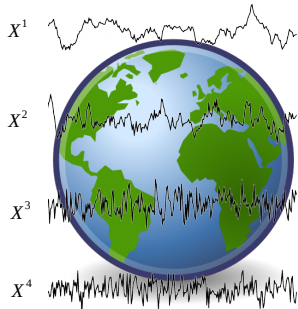- Autocorrelation

CI-based causal discovery for time series
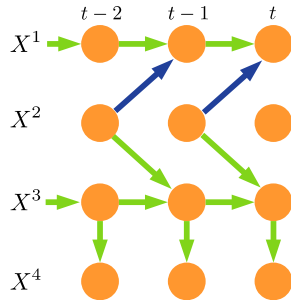
Particularities:

- Variables are resolved in time
- Autocorrelation

Additional assumption:

- Stationary causal structure

Particularities:

- Variables are resolved in time
- Autocorrelation

Additional assumption:

- Stationary causal structure

CI-based causal discovery for time series

Additional statistical challenges:

- High dimensionality   (resolving in time)

- Ill-calibrated statistical tests of independence   (autocorrelation)

- Low detection power   (autocorrelation)

 ⇒ standard algorithms often yield bad statistical performance

CI-based causal discovery for time series

Additional statistical challenges:

- High dimensionality   (resolving in time)

- Ill-calibrated statistical tests of independence   (autocorrelation)

- Low detection power   (autocorrelation)

    $\Rightarrow$ standard algorithms often yield bad statistical performance

Our contribution:

Statistical problems alleviated by specialized algorithms[†] developed by the Climate Informatics Group @DLR Jena:

- PCMCI   time-lagged links only & no unobserved confounders   [Runge et al., 2019]

- PCMCI^{+} no unobserved confounders   [Runge, 2020]

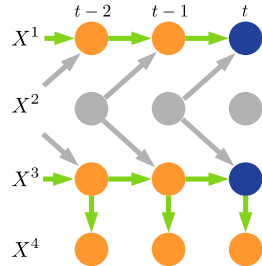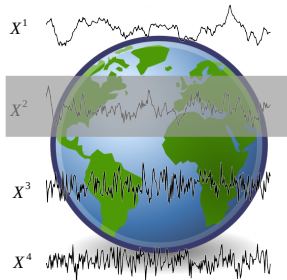- LPCMCI (Latent-PCMCI)   [Gerhardus and Runge, 2020]

[†] available at: https://github.com/jakobrunge/tigramite
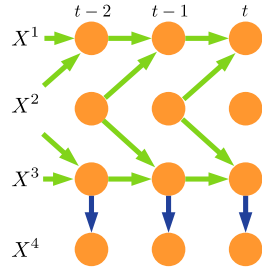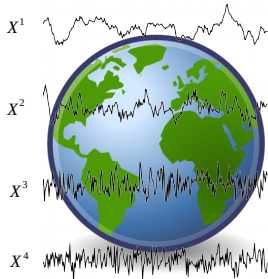
LPCMCI allows for:

- Contemporaneous links   (also PCMCI$^+$ does)

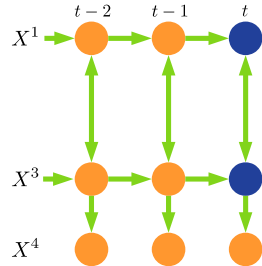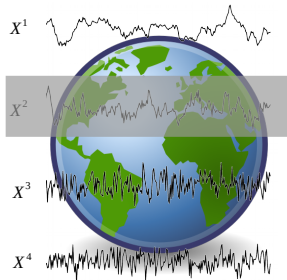**LPCMCI allows for:**

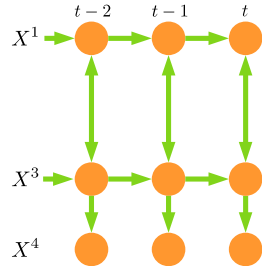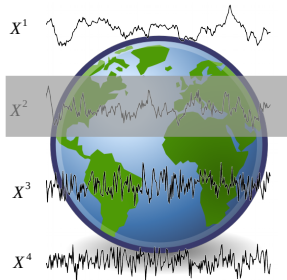- Contemporaneous links
- Unobserved confounders

(also PCMCI$^+$ does)

LPCMCI allows for:

- Contemporaneous links
- Unobserved confounders

(also PCMCI$^+$ does)

LPCMCI allows for:

- Contemporaneous links
- Unobserved confounders

(also PCMCI$^+$ does)

Basic idea:

More powerful CI tests by iterative learning of and subsequent conditioning on direct causes.

Results of numerical experiments:

For autocorrelated continuous data LPCMCI shows strong gains in recall as compared to the current state of the art algorithm*

*the SVAR-FCI algorithm by [Malinsky and Spirtes, 2018]
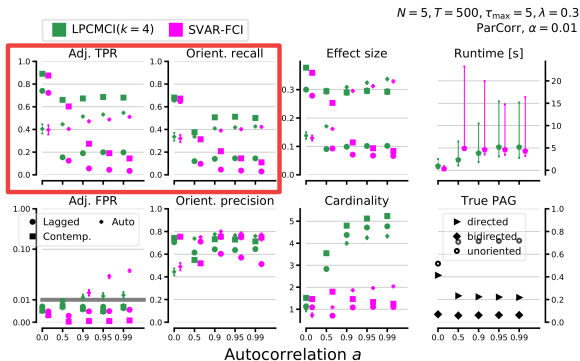
References i

📄 Gerhardus, A. and Runge, J. (2020).
High-recall causal discovery for autocorrelated time series with latent confounders.
In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. F., and Lin, H., editors, Advances in Neural Information Processing Systems, volume 33, pages 12615–12625. Curran Associates, Inc.

📄 Malinsky, D. and Spirtes, P. (2018).
Causal Structure Learning from Multivariate Time Series in Settings with Unmeasured Confounding.
In Le, T. D., Zhang, K., Kıcıman, E., Hyvärinen, A., and Liu, L., editors, Proceedings of 2018 ACM SIGKDD Workshop on Causal Disocvery, volume 92 of Proceedings of Machine Learning Research, pages 23–47, London, UK. PMLR.

Pearl, J. (1985).
A Constraint - Propagation Approach to Probabilistic Reasoning.
In Kanal, L. N. and Lemmer, J. F., editors, Proceedings of the First Conference on Uncertainty in Artificial Intelligence, UAI'85, page 31–42, Arlington, Virginia, USA. AUAI Press.

Pearl, J. (1988).
Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.
Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Pearl, J. (2000).
Causality: Models, Reasoning, and Inference.
Cambridge University Press, New York, NY, USA.

Peters, J., Janzing, D., and Schölkopf, B. (2017).
Elements of Causal Inference - Foundations and Learning Algorithms.
Adaptive Computation and Machine Learning Series. The MIT Press, Cambridge, MA, USA.

Runge, J. (2020).
Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets.
In Adams, R. P. and Gogate, V., editors, Proceedings of the Thirty-Sixth Conference on Uncertainty in Artificial Intelligence, UAI 2020, virtual online, August 3-6, 2020, page 579. AUAI Press.

📄 Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019).
Detecting and quantifying causal associations in large nonlinear time series datasets.
Science Advances, 5(11).

📄 Spirtes, P., Glymour, C., and Scheines, R. (2000).
Causation, Prediction, and Search.
MIT Press, Cambridge, MA, USA.