Constraint-based Causal Discovery algorithms: simulation & application

IC^2*S*^{2}, Amsterdam, July 2019

nitin bhushan, laura bringmann, casper albers, linda steg

university of groningen

faculty of behavioural and social sciences

Background

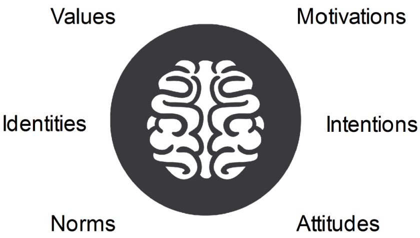
Special characteristics of psychological research

- Flag-bearer of the replication crisis.
- Technical background of psychologists more limited than of, e.g., biologists. Analyses and models should be interpretable.
- Distinction between independent and dependent variables often not clear: everything affects everything
- Establishing causal relationships is hard. True experiments are often not feasible.
- Many variables involved. Often, theories overlap each other.

Reducing household energy usage

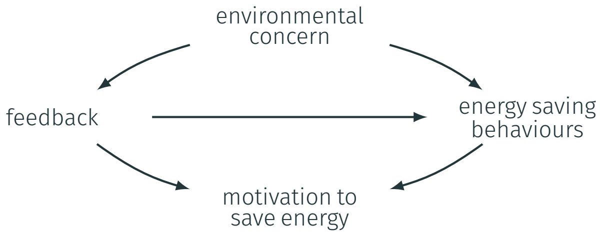
1. Technological advances
   – More efficient refrigerators, tv's, etc.

2. Psychological advances
   – This requires understanding factors determining energy behaviours

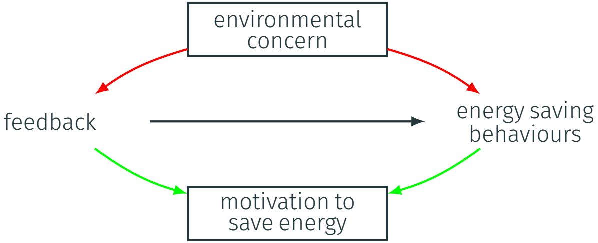


Graphical Causal Models: Directed Acyclic Graphs

Graphical Models: Directed Acyclic Graphs (DAGs)

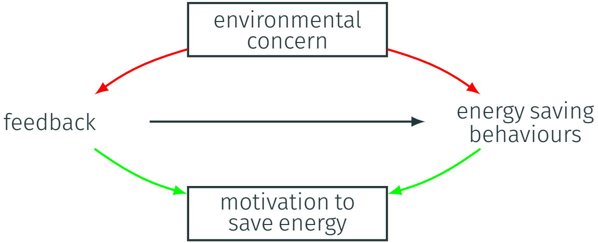


Graphical Models: Directed Acyclic Graphs (DAGs)



- Causal Identification: choosing a sufficient set of covariates such that the causal effect of interest can be estimated carefully.

Graphical Models: Directed Acyclic Graphs (DAGs)



- Causal Identification: choosing a sufficient set of covariates such that the causal effect of interest can be estimated carefully.
- Causal Discovery: in cases when when there is no clear theory, search through the entire model space to explore plausible causal structures.

Graphical Causal Models: Directed Acyclic Graphs

Causal discovery

Causal Discovery

- Competing theories can be represented with different causal graphical models, and then the task of causal search is to select the best (or better) model on the basis of the data and researchers' background knowledge.
- Constraint-based causal discovery algorithms make informed guesses of plausible causal structures by seeking to satisfy certain constraints (e.g., conditional independence)

Causal Discovery

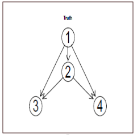
- Competing theories can be represented with different causal graphical models, and then the task of causal search is to select the best (or better) model on the basis of the data and researchers' background knowledge.
- Constraint-based causal discovery algorithms make informed guesses of plausible causal structures by seeking to satisfy certain constraints (e.g., conditional independence)
- Can be considered a promising computational in psychology, yet very little is known about their performance in psychological science.

Simulation study

We designed a simulation study to investigate:

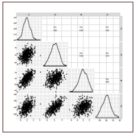
1. which method works best at retrieving the true causal structure under what circumstances?
2. how tolerant are these methods to violations of their assumptions?

Simulation design: Data generating mechanism



Truth

sample

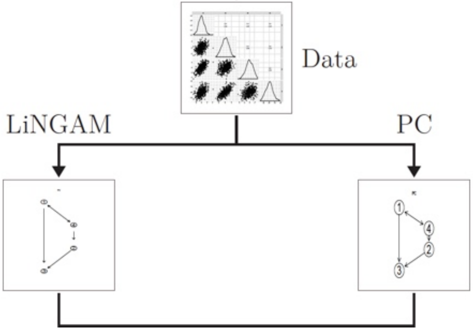


Data

$$X := BX + \epsilon$$

$$X = (I - B)^{-1}\epsilon$$

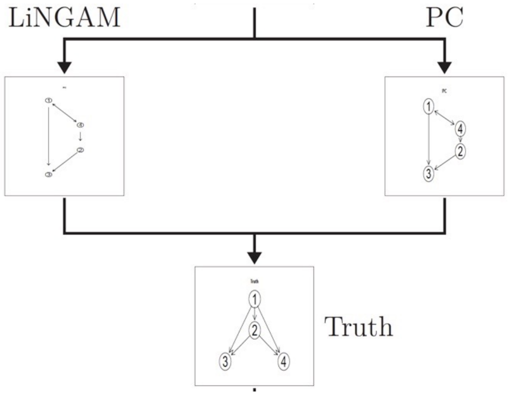
Simulation design: Methods



Non-parametric SEMs

1. PC algorithm
   - Linear model with Gaussian errors.
2. LiNGAM
   - Linear model with non-Gaussian error terms.

Simulation design: Metrics



1. Structural Hamming distance (SHD).
2. Structural Intervention distance (SID).

Scaled measures [0,1], lower values indicate better performance.

Simulation design: parameters

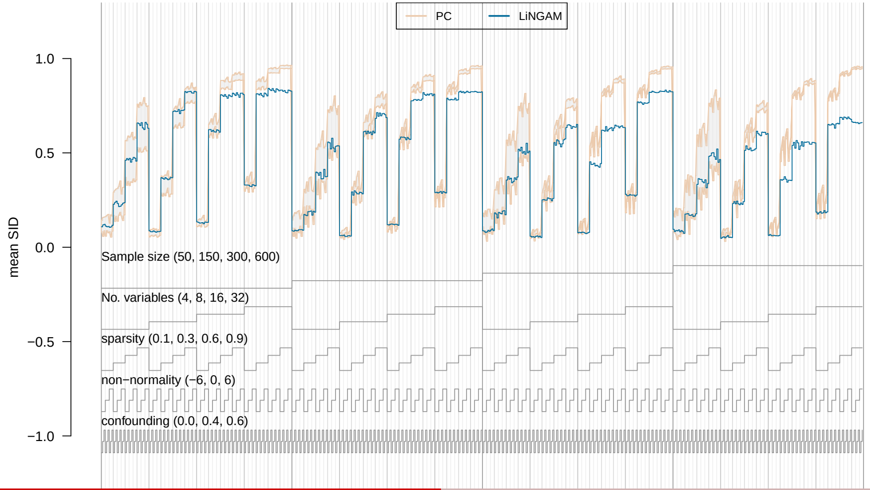
Table 1: Simulation parameters and values

Description	Expression	Value
number of replications	N_{sim}	200
sample size	n	50, 150, 300, 600
number of variables	p	4, 8, 16, 32
graph sparsity	s	0.1, 0.3, 0.6, 0.9
asymmetry of `skew-normal` distribution	α	-6, 0, 6
degree of confounding	ρ	0.0, 0.4, 0.6
causal discovery method	algo	PC, LiNGAM

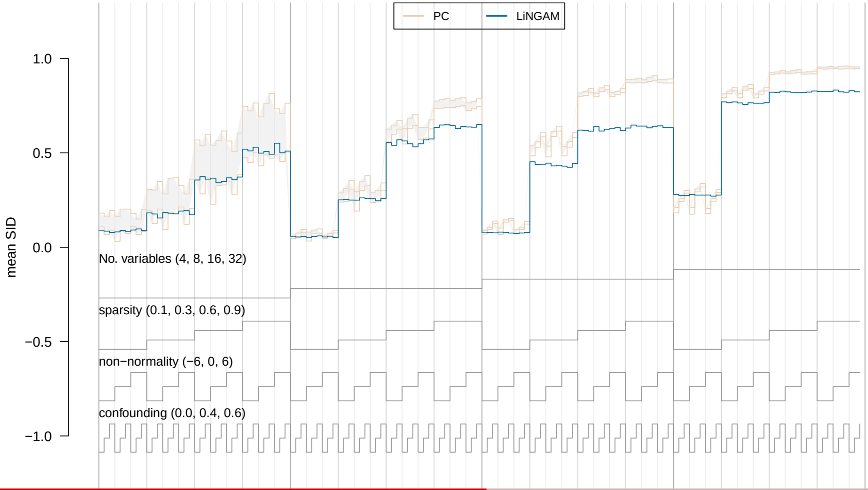
$4 \times 4 \times 4 \times 3 \times 3 \times 2$ *conditions with* 200 *replications*.

Results

Simulation results: visual exploration using nested loop graphs



Simulation results: visual exploration using nested loop graphs



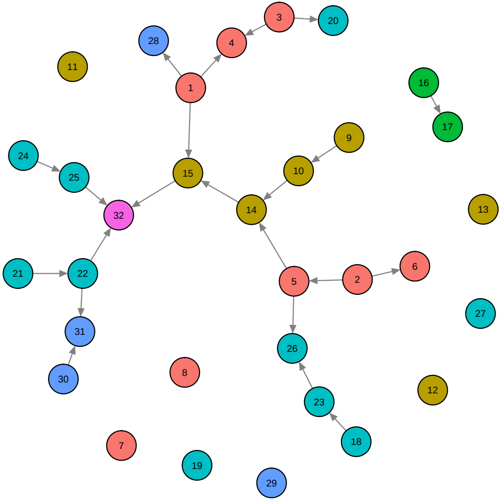
Simulation findings and limitations

- These methods work best when reality is sparse.
- Hard to separate the methods based on our settings, but LiNGAM seems to outperform PC as reality becomes less sparse.
- Both methods seem to be robust to violations of their assumptions.
- Replicability of the methods must be considered before real-world applications (use stability selection[1]).

[1] Meinshausen & Bühlmann, 2010

Application

Buurkracht [N = 313, p = 32]



Personal factors
- 1: Altruistic values
- 2: Biospheric values
- 3: Egoistic values
- 4: Hedonic values
- 5: Environmental self identity
- 6: Personal importance of sustainable energy behaviour
- 7: Need to belong
- 8: Need to be unique

Factors related to the social context
- 9: Neighbourhood entitativity
- 10: Neighbourhood homogeneity
- 11: Neighbourhood interaction
- 12: Interaction with neighbours
- 13: Neighbourhood identification
- 14: Environmental neighbourhood identity
- 15: Neighbourhood importance of sustainable energy behaviour

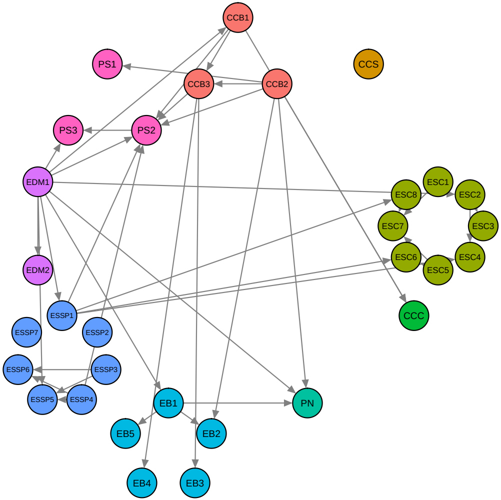
Evaluations of energy companies and the government
- 16: Group based anger
- 17: Group based distrust

Sustainable energy intentions and behaviours
- 18: Overall energy savings
- 19: Thermostat temperature
- 20: Shower time
- 21: Energy efficient appliances
- 22: Energy saving measures
- 23: Household sustainable energy intentions
- 24: Communal sustainable energy intentions
- 25: Initiative involvement intentions
- 26: Other pro environmental intentions
- 27: Other commmunal intentions

Socio−demographics
- 28: Gender
- 29: Birthyear
- 30: Education
- 31: Income

- 32: Membership

European Social Survey [N=44000, p=31]



Climate Change Beliefs
- CCB1: Climate change reality
- CCB2: Climate chance cause
- CCB3: Climate chance impact

Climate Change Salience
- CCS: Climate change salience

Energy Security Concern
- ESC1: Concern about energy reliability
- ESC2: Concern about energy affordability
- ESC3: Concern about dependency on energy imports
- ESC4: Concern about dependency on fossil fuels
- ESC5: Concern about energy security due to natural disasters
- ESC6: Concern about energy security due to insufficient power
- ESC7: Concern about energy security due to technical failures
- ESC8: Concern about energy security due to terrorist attacks

Climate Concern
- CCC: Climate change concern

Personal Norms
- PN: Personal norm to Reduce climate change

Efficacy Beliefs
- EB1: Self−efficacy
- EB2: Collective outcome expectancy
- EB3: Collective efficacy
- EB4: Institutional efficacy
- EB5: Personal outcome expectancy

Energy Supply Source Preference
- ESSP1: Preference for coal power
- ESSP2: Preference for natural gas power
- ESSP3: Preference for hydroelectric power
- ESSP4: Preference for nuclear power
- ESSP5: Preference for solar power
- ESSP6: Preference for wind power
- ESSP7: Preference for biomass power

Energy Demand Measures
- EDM1: Energy efficiency behaviour
- EDM2: Energy curtailment behaviour

Policy Support
- PS1: Support fossil fuel tax
- PS2: Support subsidy renewable energy
- PS3: Support ban least energy efficient appliances

Discussion

In the context of new lines of research, graphical causal models and causal search can be used to:

- provide a severe test of relationships between variables and improve our understanding.

Discussion

In the context of new lines of research, graphical causal models and causal search can be used to:

- provide a severe test of relationships between variables and improve our understanding.
- researchers are always advised to base the results of these methods against current scientific knowledge.

Discussion

In the context of new lines of research, graphical causal models and causal search can be used to:

- provide a severe test of relationships between variables and improve our understanding.
- researchers are always advised to base the results of these methods against current scientific knowledge.
- these methods are, at best, a helpful computational assistant which may detect expected and unexpected patterns in a dataset.
- expected patterns lead to more confidence in a theory, unexpected effects can serve as substantive hypothesis which can be put to the test using a new dataset.

5. The Step from Association to Causation

This issue is naturally of great concern to workers in observational research and has received much discussion in individual subject-matter fields. I shall confine myself to a few comments on statistical aspects of the problem.

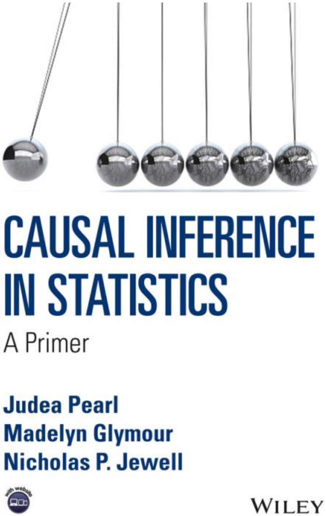
First, as regards planning. About 20 years ago, when asked in a meeting what can be done in observational studies to clarify the step from association to causation, Sir Ronald Fisher replied: "Make your theories elaborate". The reply puzzled me at first, since by Occam's razor the advice usually given is to make theories as simple as is consistent with the known data. What Sir Ronald meant, as the subsequent discussion showed, was that when constructing a causal hypothesis one should envisage as many *different* consequences of its truth as possible, and plan observational studies to discover whether each of these consequences is found to hold. If a

The Planning of Observational Studies of Human Populations
W. G. Cochran and S. Paul Chambers
Journal of the Royal Statistical Society. Series A (General)
Vol. 128, No. 2 (1965), pp. 234-266

Thank you.

n.bhushan@rug.nl

Further reading – Causality



Further reading – Graphical Models

- Lauritzen, S.L. (1996). *Graphical Models*. Oxford University Press.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT Press.
- Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge: Cambridge University Press.
- Borsboom, D., & Cramer, A. O. (2013). *Network analysis: an integrative approach to the structure of psychopathology*. Annual review of clinical psychology.

Further reading – Environmental Psychology

- Steg, L., Perlaviciute, G., & van der Werff, E. (2015). *Understanding the human dimensions of a sustainable energy transition.* Frontiers in Psychology.
- Abrahamse, W. (2007). *Energy conservation through behavioural change.* University of Groningen
- Steg, L., Bolderdijk, J.W., Keizer, K., & Perlaviciute, G.(2014). *An integrated framework for encouraging pro-environmental behaviour: the role of values, situational factors and goals.* Journal of Environmental Psychology
- Bamberg, S., Rees, J., & Seebauer, S. (2015). *Collective climate action: Determinants of participation intention in community-based pro-environmental initiatives.* Journal of Environmental Psychology.

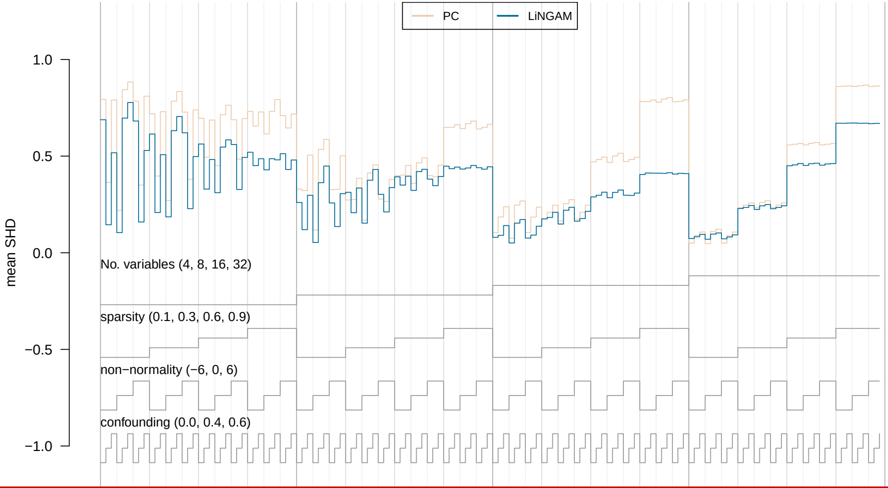
Appendix

Two nodes in a DAG are said to be d-separated if and only if the corresponding random variables are conditionally independent of their non-effects given their immediate causes.[2]

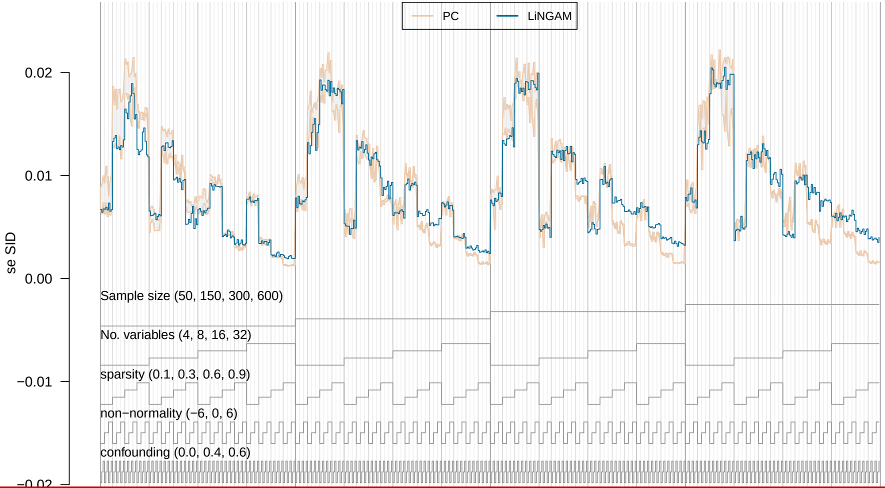
[2] for a formal definition, see S. Lauritzen, 2001; Pearl, 2009; Spirtes et al. 2001

Causal faithfulness assumes that given a graph and the associated joint probability distribution, the only independence relationships in the distribution are those that follow from d-separation.

Simulation results: visual exploration (SHD)



Simulation results: standard error (SID)



Summary : Causal discovery using graphical models

- Traditionally, causal discovery in the social sciences has proceeded through hypothesis testing.
- Standard model : researcher posits models; derives the models' testable implications; typically performs a hypothesis test.

Summary : Causal discovery using graphical models

- Traditionally, causal discovery in the social sciences has proceeded through hypothesis testing.
- Standard model : researcher posits models; derives the models' testable implications; typically performs a hypothesis test.
- Models may be straightforwardly rejected using this procedure, but if they are not rejected, we still should not categorically accept them (unconsidered models might be equally consistent with the data).

Summary : Causal discovery using graphical models

- Traditionally, causal discovery in the social sciences has proceeded through hypothesis testing.
- Standard model : researcher posits models; derives the models' testable implications; typically performs a hypothesis test.
- Models may be straightforwardly rejected using this procedure, but if they are not rejected, we still should not categorically accept them (unconsidered models might be equally consistent with the data).
- In contrast, causal discovery algorithms effectively consider the entire class of models.

(under review.) N. Bhushan, L. Bringmann, C. J. Albers. *Comparing causal discovery algorithms for psychological applications.*