Comparing Constraint-based Causal Discovery algorithms in scenarios typical to psychology

IOPS, December 2018

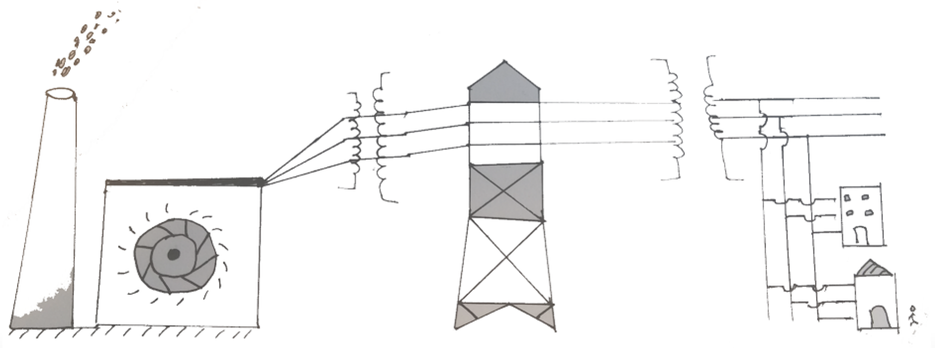
nitin bhushan, laura bringmann, casper albers

university of groningen

faculty of behavioural and social sciences

Background

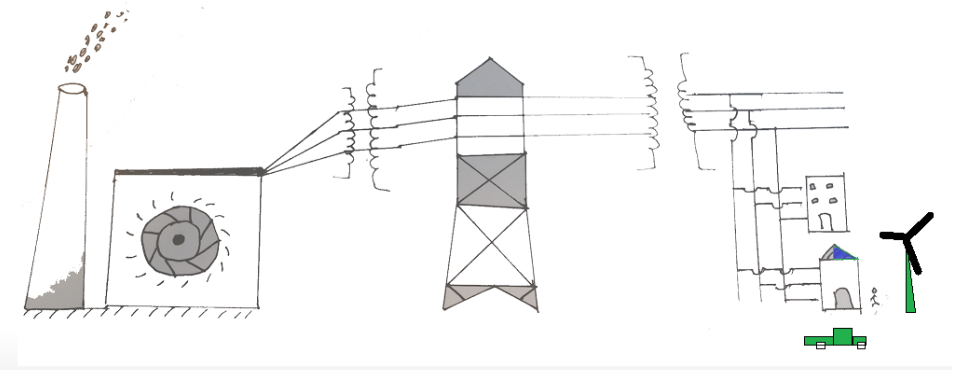
Traditional power systems



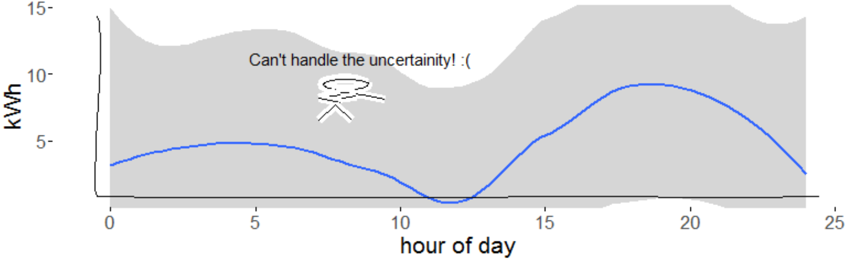
Traditional power system

Modern power systems

1

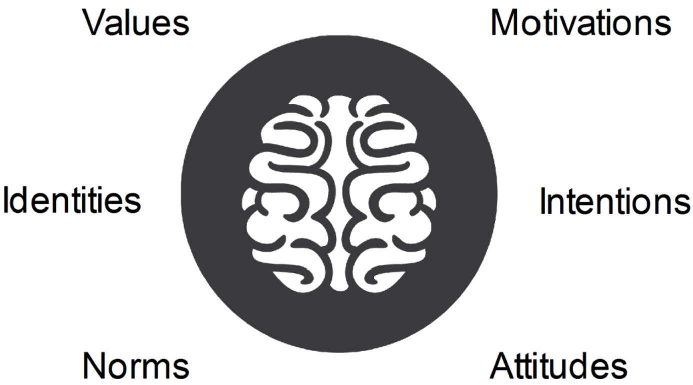


Consumption patterns are changing



(Modern?) load profiles

What are the drivers of this change?



Interventions

Design of interventions

Various behavioural interventions including block leader approaches, behavioural commitments, and different types of feedback appear to be effective to reduce household fossil energy use.[1]

[1] Abrahamse, Steg, Vlek & Rothengatter, 2005; Karlin, Zinger & Ford, 2005

Roadblocks

- Most intervention studies are based on small and non-representative samples.
- Long term effects of such interventions are poorly understood.
- Little is known about (if and) why interventions are (in)effective and how they can be improved.

[2] Allcott & Mullainathan, 2010; Frederiks et al., 2016; Vine et al., 2014

Roadblocks

- Most intervention studies are based on small and non-representative samples.
- Long term effects of such interventions are poorly understood.
- Little is known about (if and) why interventions are (in)effective and how they can be improved.

Solution: use Randomized Controlled Trials (RCTs).[2]

[2] Allcott & Mullainathan, 2010; Frederiks et al., 2016; Vine et al., 2014

RCTs: The Gold standard

- Random sampling, random assignment, manipulation.
- If done right, this is an unbiased estimator of the average causal effect (in expectation).
- Straightforward with crops, not quite so easy with people.
- In the absence of randomisation, one must consider sources of bias carefully.
- The notion of a gold standard is flawed.[3]

[3] Angus Deaton & Nancy Cartwright, Understanding and misunderstanding randomized controlled trials, 2018

RCTs: The Gold standard

- Random sampling, random assignment, manipulation.
- If done right, this is an unbiased estimator of the average causal effect (in expectation).
- Straightforward with crops, not quite so easy with people.
- In the absence of randomisation, one must consider sources of bias carefully.
- The notion of a gold standard is flawed.[3]

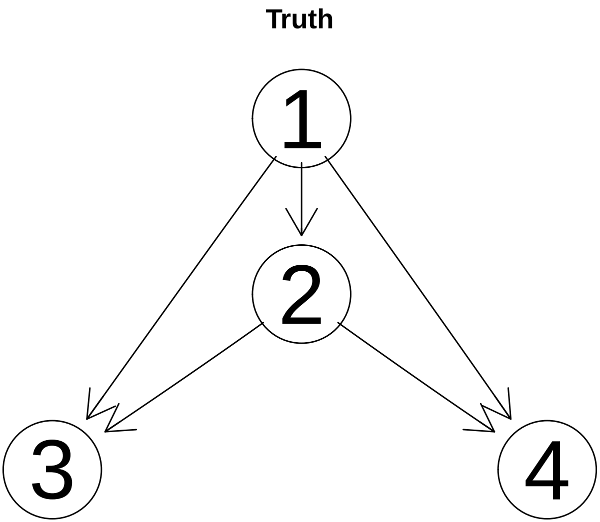
Solution: use Graphical causal models.

[3] Angus Deaton & Nancy Cartwright, Understanding and misunderstanding randomized controlled trials, 2018

Interventions

Graphical Causal Models: Directed Acyclic Graphs

Directed Acyclic Graphs (DAGs): data generating mechanisim



Graphical Models: Directed Acyclic Graphs (DAGs)

- The graphical modelling approach to causal inference, and in particular, causal directed acyclic graphs (DAGs), offers a formal tool to solve problems of causal inference when randomisation is not feasible.[4]
- These graphs encode qualitative causal assumptions and are non-parametric structural equation models.

[4] Judea Pearl, Causality, 2009

Graphical Models: Directed Acyclic Graphs (DAGs)

- The graphical modelling approach to causal inference, and in particular, causal directed acyclic graphs (DAGs), offers a formal tool to solve problems of causal inference when randomisation is not feasible.[4]
- These graphs encode qualitative causal assumptions and are non-parametric structural equation models.
- Causal Identification: choosing a sufficient set of covariates such that the causal effect of the intervention can be estimated carefully.

[4] Judea Pearl, Causality, 2009

Graphical Models: Directed Acyclic Graphs (DAGs)

- The graphical modelling approach to causal inference, and in particular, causal directed acyclic graphs (DAGs), offers a formal tool to solve problems of causal inference when randomisation is not feasible.[4]
- These graphs encode qualitative causal assumptions and are non-parametric structural equation models.
- Causal Identification: choosing a sufficient set of covariates such that the causal effect of the intervention can be estimated carefully.
- Causal Discovery: in cases when when there is no clear theory, computational algorithms can be used to explore plausible causal structures.

[4] Judea Pearl, Causality, 2009

Interventions

Directed Acyclic Graphs: causal discovery

Causal Discovery: (an informal) definition

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.

Causal Discovery: the methods

- These methods explore probabilistic causal relationships between variables of interest from observational data.[5]

[5] Eberhardt, 2016; Spirtes, Glymour, & Scheines, 2000

Causal Discovery: the methods

- These methods explore probabilistic causal relationships between variables of interest from observational data.[5]
- Constraint-based causal discovery algorithms make informed guesses of plausible causal structures by seeking to satisfy certain constraints (e.g., conditional independence)

[5] Eberhardt, 2016; Spirtes, Glymour, & Scheines, 2000

Causal Discovery: the methods

- These methods explore probabilistic causal relationships between variables of interest from observational data.[5]
- Constraint-based causal discovery algorithms make informed guesses of plausible causal structures by seeking to satisfy certain constraints (e.g., conditional independence)
- In contrast, score-based causal discovery methods aim to find the most plausible causal structure by maximising a score (e.g., a model fit criteria).
- To the best of our knowledge, there are no applications of these methods in psychology, and little is known about their performance in scenarios typical to psychology.

[5] Eberhardt, 2016; Spirtes, Glymour, & Scheines, 2000

Causal Discovery: the methods

- These methods explore probabilistic causal relationships between variables of interest from observational data.[5]
- Constraint-based causal discovery algorithms make informed guesses of plausible causal structures by seeking to satisfy certain constraints (e.g., conditional independence)
- In contrast, score-based causal discovery methods aim to find the most plausible causal structure by maximising a score (e.g., a model fit criteria).
- To the best of our knowledge, there are no applications of these methods in psychology, and little is known about their performance in scenarios typical to psychology. -> Simulation study!

[5] Eberhardt, 2016; Spirtes, Glymour, & Scheines, 2000

The aim of the study is 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?

Causal Discovery: constraint-based methods

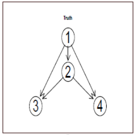
- Constraint-based causal discovery methods makes informed guesses of plausible causal structures by seeking to satisfy certain constraints. The key constraint used is conditional independence.
- Conditional independence is non-directional and is a probabilistic statement which depends on the distributional assumptions underlying the data (Dawid, 1980).
- Causal inference requires untestable assumptions.

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.[6]

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

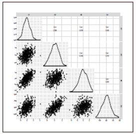
Causal faithfulness assumes that given a graph and the associated joint probability distribution, the only independence relationships in the distribution are those that follow from d-separation.

Simulation design: Data generating mechanism



Truth

sample



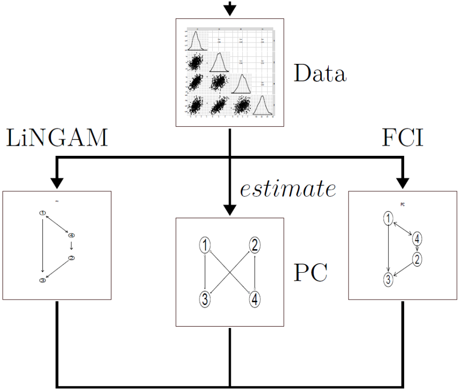
Data

$$X_i := \sum_{j=1}^{p} \beta_{i,j} X_j + \epsilon_i$$

$$X := BX + \epsilon$$

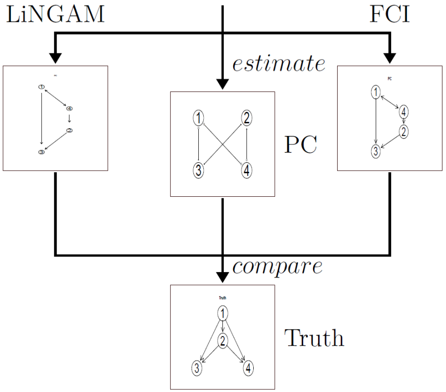
$$X = (I - B)^{-1} \epsilon$$

Simulation design: Methods



1. PC algorithm
   - Linear model with Gaussian errors.
2. LiNGAM
   - Linear model with non-Gaussian error terms.
3. FCI algorithm
   - Linear model with Gaussian errors allowing for latent confounders.

Simulation design: Metrics



1. Structural Hamming distance (SHD).
2. Structural Intervention distance (SID).
3. For both measures, lower values indicate better performance.

Simulation design: parameters

Table 1: Simulation parameters and values

Description	Expression	Value
number of replications	N_{sim}	105
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, FCI

Simulation design: parameters

Table 1: Simulation parameters and values

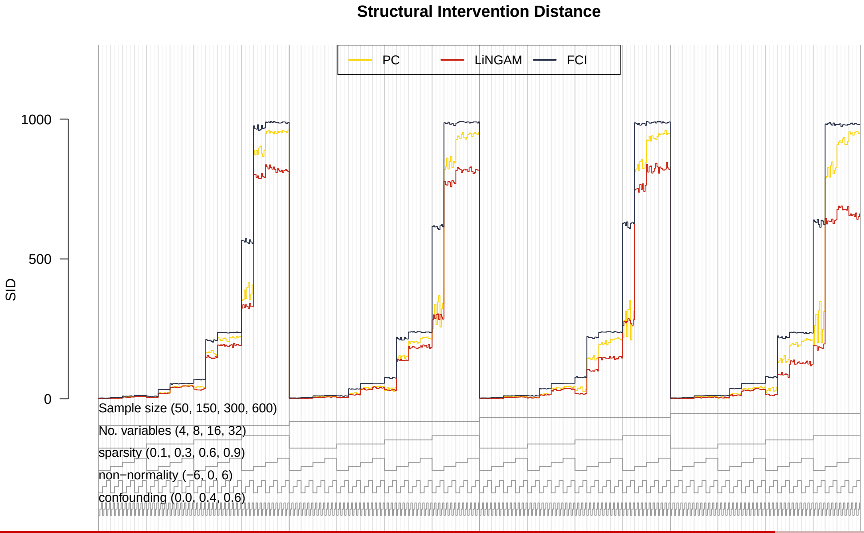
`Description`	`Expression`	`Value`
number of replications	N_{sim}	105
sample size	n	50, 150, 300, 600
number of variables	p	4, 8, 16, 32
graph sparsity	s	0.1, 0.3, 0.6, 0.9
asymmetry of `skew-normal` distribution	α	-6, 0, 6
degree of confounding	ρ	0.0, 0.4, 0.6
causal discovery method	algo	PC, LiNGAM, FCI

$$4 \times 4 \times 4 \times 3 \times 3 \textit{ conditions}$$

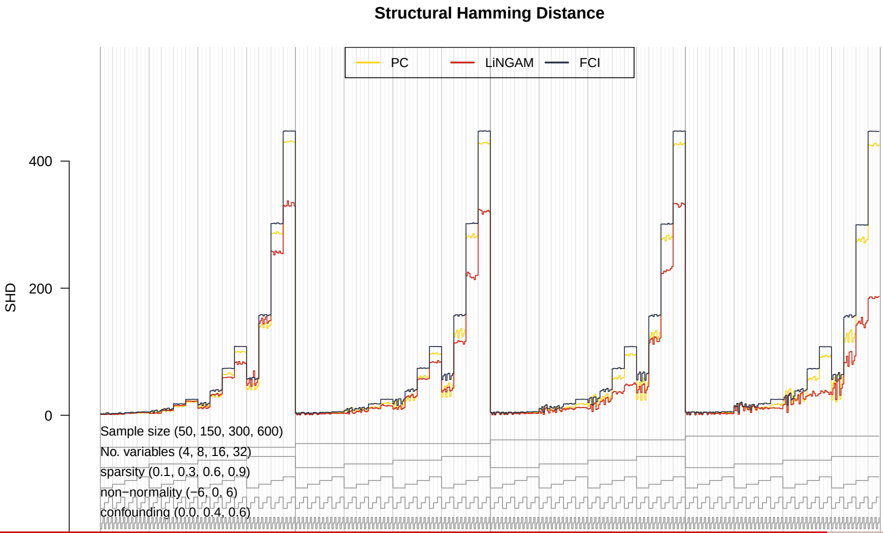
Simulation results

1. Impact of the sample size (n).
2. Impact of the number of variables (p).
3. Impact of graph sparsity (s).
4. Impact of latent confounders (ρ).
5. Impact of skewness in the error terms (α).

Simulation results: visual exploration



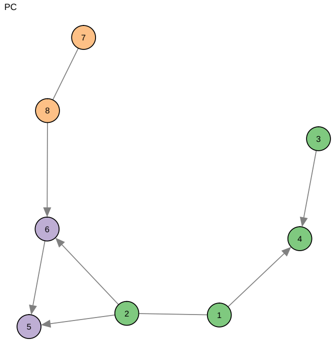
Simulation results: visual exploration



Simulation findings (in progress)

- LiNGAM seems to outperform PC and FCI.
- All methods seem to be robust to violations of their assumptions.
- The results are mainly influenced by the interaction of the number of variables p, and the graph sparsity s.
- There is no concensus on what is a suitable performance metric.
- Few successfully applications of these methods on real data-sets.

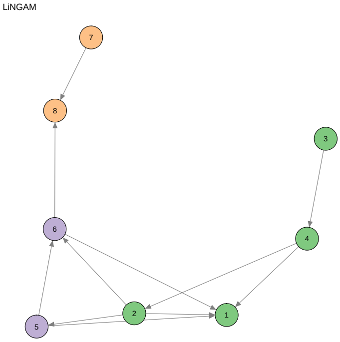
Application: Buurkracht



Values
- 1: Altruistic values
- 2: Biospheric values
- 3: Egoistic values
- 4: Hedonic values

Personal factors
- 5: Environmental self identity
- 6: Personal importance of sustainable energy behaviour

Sustainable energy intentions and behaviours
- 7: Overall energy savings
- 8: Household sustainable energy intentions



Conclusions

Summary

Graphical models are useful for:

- Exploratory analyses of (relatively) high-dimensional data.
- Identifying causal effects when experiments are not feasible.
- Exploring plausible causal structures in the absence of theory.
- Exploring temporal dynamics (e.g., the Hidden Markov Model).

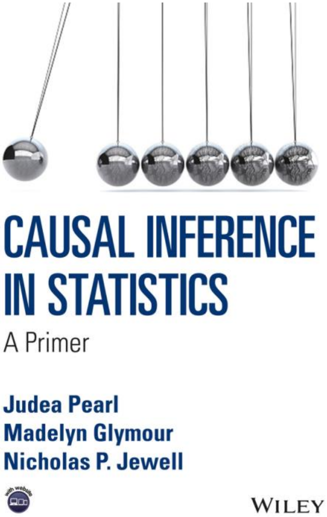
These models help (i) improve our understanding of relationships between factors influencing energy behaviours and (ii) carefully estimate effects of interventions to promote pro-environmental behaviours.

Thank you!

n.bhushan@rug.nl

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.


CAUSAL INFERENCE
IN STATISTICS
A Primer
Judea Pearl
Madelyn Glymour
Nicholas P. Jewell
WILEY

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.