

😊 Add icon 🖼️ Add cover

# 文献調査

	☰ Constraint-based	☰ Score-based	☰ FCM-based	☰ Gradient-based	☰ Miscellaneous
1	PC	GES	LiNGAM	NOTEARS	FRITL
2	FCI	FGS	ANM	GraN-DAG	HCM
3	FCI with TBK	A-star search	SAM	GOLEM	Kg2Causal

## 一般的な因果探索

### 1.Constraint-based

制約ベースの因果探索手法において、条件付き独立性（Conditional Independence, CI）の検定は中心的な目的である。

観測データの確率分布が基礎となる因果グラフに忠実（faithful）である場合、条件付き独立性の検定によって因果スケルトンを復元することができる（Marx & Vreeken, 2019）。

したがって、制約ベースの手法では、変数間の条件付き独立性を検定し、エッジの有無を確認する。

### 2.Score-based

スコアベースの因果探索アルゴリズムは、全ての可能な有向非循環グラフ（DAG）の空間を探索し、データを最もよく説明するグラフを見つけることを目的とする。

一般的に、スコアベース手法は二つの主要な構成要素を持つ。(i) 候補グラフ  $G'$  の探索状態または探索空間を調べるための探索戦略、(ii) 候補因果グラフを評価するためのスコア関数である。探索戦略とスコア関数の組み合わせによって、全ての可能なDAG空間での探索を最適化する。代表的なスコア関数はバイズ情報量規準であり、他にも、BDeu、BGe、MDLなどが一般的なスコア関数として利用される。

### 3.FCM-based

関数的因果モデル（Functional Causal Model, FCM）に基づく手法は、変数間の因果関係を特定の関数形式で表現する。

FCMでは、各変数をその親の関数と独立したノイズ項の和として表す。

データ分布や関数クラスに追加の仮定を課すことで、FCMベースの手法は同じ同値類に属する異なるDAGを識別することができる。

$$X = f(PA_X) + E$$

### 4.Gradient-based

組合せ最適化として扱われていた構造学習問題を連続的な問題へと変換し、最小二乗目的関数および有向非循環グラフ (DAG) の代数的特徴を利用し、勾配ベースの最適化手法を用いて解く。

## 時系列因果探索

### 概要

時系列データとは、一定期間にわたって収集された観測値の集合である。

これまでに議論した手法は、独立同分布 (I.I.D.) または時間に依存しないデータからの因果探索に特化したものである。

しかし、現実世界の多くの分野では、データはしばしば時系列 (非I.I.D.) の形で得られる。

このようなデータに対しては、条件付き独立性検定 (CI testing)、構造方程式モデル/関数的因果モデル (SEM/FCM)、グレンジャー因果 (Granger, 1969)、およびディープニューラルネットワークなどに基づく多様な因果探索手法が存在する。

### 時系列データ

時系列データとは、一定の時間間隔で測定された一連の観測値の集合である。

時刻  $t$  における時系列変数  $X_j$  の観測値は  $X_j^t$  と表される。

時系列データの例として、小売売上高、株価、気候データ、患者の心拍数、脳活動記録、気温測定値などが挙げられる。

一般に、時系列データは次のような特徴を持つ。

i. **トレンド (Trend)** : データが長期的に上昇または下降する傾向を示す場合、トレンドが存在するという。このような変化は必ずしも線形ではなく、上昇傾向から下降傾向に変わることもある。

ii. **季節性 (Seasonality)** : 時系列データの周期的な特性を指す。データが日・週・月・四半期・年などの時間スパンに基づいて規則的に変動する場合、季節性が存在する。

例えば気温データでは、夏に高く冬に低い傾向が一般的である。時系列分析では、このような季節性を活用してより堅牢なモデルを構築する。

iii. **自己相関 (Autocorrelation)** : 自己相関 (または自己相似) は、ある時系列とその遅延版との間の類似度を示す。

時系列データは通常自己相関的であり、過去が現在および未来に影響を与える。

iv. **定常性と非定常性 (Stationarity & Non-stationarity)** : 定常性とは、確率過程の結合分布が時間平行移動によって変化しないことを指す。

もし時系列が因果関係  $X_i \rightarrow X_j$  を持ち、任意の時刻  $t$  においてこれが成立するならば、すべての  $t' \neq t$  においても成立する場合、その時系列は定常である。

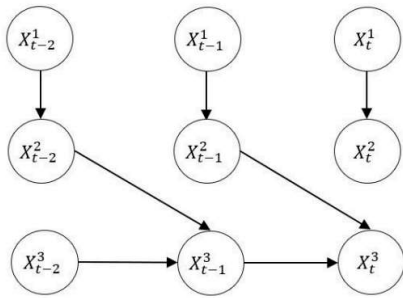
一方、非定常時系列ではこの条件が成り立たず、ある時点での  $X_i \rightarrow X_j$  が他の時点では成立しない可能性がある。

### 出力グラフの分類

#### 完全時系列因果グラフ, Full-time Causal Graph

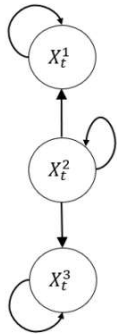
完全時系列因果グラフは、即時因果エッジおよび遅延因果エッジの両方を表し、因果間の遅延がグラフ上に明示される。

場合によっては、モジュール変化 (CM) も併せて示されることがある。



## 要約因果グラフ, Summary Causal Graph

要約因果グラフは、変数間の因果関係のみを示し、時間的依存性を明示的には表さない。



## アルゴリズム

### Full-time Causal Graph

#### 1. PCMCI / PCMCI+

**理論的基盤：**条件付き独立性検定（MCI）に基づく制約ベース（Constraint-based）の時系列因果探索手法。

**長所：**高次元データにも適用可能で、線形・非線形系の両方で滞後および同時因果関係を検出できる。

**短所：**一部の因果方向が定向できず、同時因果関係の識別には限界がある。

#### 2. VarLiNGAM

**理論的基盤：**構造方程式モデル（SEM）と関数的因果モデル（FCM）に基づく非ガウス線形 SVAR モデル。

**長所：**滞後および即時因果効果を同時に推定でき、因果強度の定量化が可能。

**短所：**線形・非ガウスデータにのみ適用可能で、潜在交絡を扱えない。

#### 3. DYNOTERS

**理論的基盤：**NOTEARS に由来する連続最適化型のスコアベース手法（微分可能な非巡回制約を使用）。

**長所：**高次元時系列データにもスケラブルで、滞後および即時エッジを同時に学習可能。

**短所：**アンダーサンプリング（欠測間隔）に弱く、非線形データには不向き。

#### 4. NAVAR

**理論的基盤：**グレンジャー因果（Granger causality）に基づく加法構造モデル。

**長所：**実装が容易で、非線形の滞後関係も検出可能。

**短所：**NNが使われ、アンダーサンプリング（欠測間隔）に弱く、解釈性が弱い。

## Summary Causal Graph

1. oCSE

**理論的基盤：** 最適因果エントロピー原理（Optimal Causation Entropy）に基づく情報理論的アプローチ。

**長所：** 比較的少ないサンプル数でも安定して因果構造を推定できる。

**短所：** グラフが密な場合、計算時間が非常に長くなる。

2. NBCB

**理論的基盤：** 制約ベースとノイズベースを組み合わせたハイブリッド手法（因果エントロピー利用）。

**長所：** マルコフ同値類（MEC）に制約されず、より柔軟な構造推定が可能。

**短所：** 完全な全時系列因果グラフ（Full-time Causal Graph）を推定できない。

3. PCTMI

**理論的基盤：** 条件付き独立性に基づく制約ベース手法で、PCとER原理によるラグ方向付けを組み合わせたもの。

**長所：** アルゴリズムの計算複雑度が低く、効率的に因果構造を得られる。

4. GrangerPW

**理論的基盤：** グレンジャー因果に基づく統計的時系列分析手法。

**長所：** 実装が容易で、滞後関係の存在を直接的に検定できる。

	≡ Method	≡ Outcome Graph	≡ Assumptions	≡ Technique Used	≡ Advantages	≡ Disadvantages
1	PCMCI	Full-time graph with lagged edges	Data stationarity, causal sufficiency	Constraint-based, momentary CI (MCI) test	Handles high-dimensional; applicable on linear/nonlinear/continuous/discrete data	Contemporaneous edges not identified; some causal links remain unoriented
2	PCMCI+	Full-time graph with lagged and contemporaneous edges	Causal sufficiency	Constraint-based, MCI test	Detects both contemporaneous and lagged causal links	Some causal links remain unoriented
3	LPCMCI	Time-series Directed Maximal Ancestral Graphs (DMAGs)	Causal faithfulness, absence of selection bias	Constraint-based	Handles large-scale time-series data, latent confounders, and discrete/continuous data; incorporates prior knowledge	May lead to wrong conclusions if faithfulness is violated; can't distinguish all members of MEC

4	<b>VarLiNGAM</b>	Full-time causal DAG (with causal strengths)	Linear DGP, no unobserved confounders, non-Gaussian noise with non-zero variances	FCM/SEM-based	Detects both lagged and instantaneous temporal edges	Hidden confounders not handled; inapplicable to nonlinear data
5	<b>DyNOTEARS</b>	Full-time causal graph	Acyclicity, fixed and identical network structure through time	Score-based approach	Scalable to high-dimensional datasets; produces both lagged and instantaneous edges	Undersampling not handled well; inapplicable to nonlinear data
6	<b>NTS-NOTEARS</b>	Full-time causal graph	No latent confounders; fixed and stationary DGP	Gradient-based, uses 1-D CNNs	Incorporates prior knowledge	Inapplicable when latent confounders exist
7	<b>NAVAR</b>	Full-time causal graph (lagged edges only)	Additive structure	Granger-causality-based	Easy to implement; can detect nonlinear relations	Produced score-matrix analyzable only via CausaMe platform
8	<b>TCDF</b>	Full-time causal graph	Temporal precedence (cause precedes effect)	Attention-based CNNs	Easy to implement; can detect hidden confounders	Performs poorly on short time series; does not ensure acyclicity
9	<b>tsFCI</b>	Partial ancestral graph	Causal Markov condition, faithfulness	Constraint-based, FCI adaptation for time series data	Handles hidden confounders	Non-stationarity property is not handled
10	<b>CD-NOD</b>	Summary or full-time graph incorporating changing modules	Pseudo causal sufficiency, faithfulness, no selection bias	Constraint-based	Detects changing modules	Some causal directions may not be identifiable if identifiability conditions not satisfied

11	TiMINo	Causal summary time graph	Independent noise, identifiability	FCM/SEM-based	Applicable to multivariate, linear, nonlinear, and instantaneous cases	Rigid model assumptions; sometimes remains undecided
12	oCSE	Summary causal graph	Faithfulness, acyclicity	Optimal Causation Entropy principle	Requires relatively few samples	Very long runtime for dense graphs
13	NBCB	Summary causal graph	Adjacency faithfulness	Hybrid (constraint and noise-based), causation entropy	Not restricted to MEC	Does not estimate a full-time causal graph
14	PCTMI	Summary causal graph	Causal Markov condition and faithfulness	Constraint-based; skeleton via PC; ER principle for lagged edges	Lower complexity	Does not estimate a full-time causal graph

## 評価指標と精度

### 1.SHD (Structural Hamming Distance)

推定されたグラフを真の因果グラフに変換するために必要な「辺の追加・削除・反転」の総数を表す指標である。

値が小さいほど推定グラフが真の構造に近く、SHD=0で完全一致を意味する。

### 2.SID (Structural Intervention Distance)

因果効果の観点から2つのDAGの近さを測る指標で、誤った介入分布の数を基準として評価する。

生成グラフ中の誤りが介入効果推定に与える影響の大きさを反映する。

### 3.FDR (False Discovery Rate)

発見された因果関係のうち、誤検出（偽陽性）の割合を示す指標である。

値が低いほど、余分なエッジ（誤りの因果関係）が少ないことを意味する。

### 4.TPR (True Positive Rate)

真の因果関係のうち、正しく検出された割合を表す。

値が高いほど、モデルが多くの正しい因果エッジを見つけられていることを示す。

### 5.FPR (False Positive Rate)

存在しない因果関係を誤って「ある」と判定した割合を表す。

値が低いほど、偽の因果関係を生成しにくい高精度なモデルであることを意味する。

### 6.F1スコア (F1 Score)

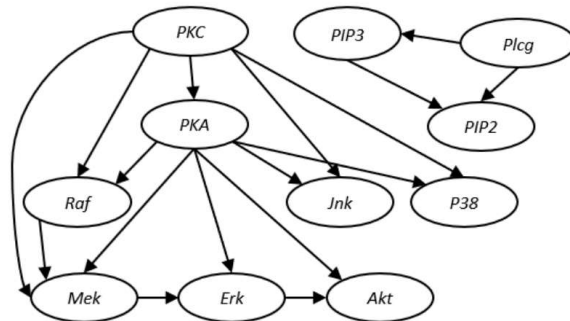
精度 (Precision) と再現率 (Recall) の調和平均で、両者のバランスを評価する指標である。

特にデータが不均衡な場合に有効で、1に近いほど性能が良いことを示す。

## Datasets

遺伝子

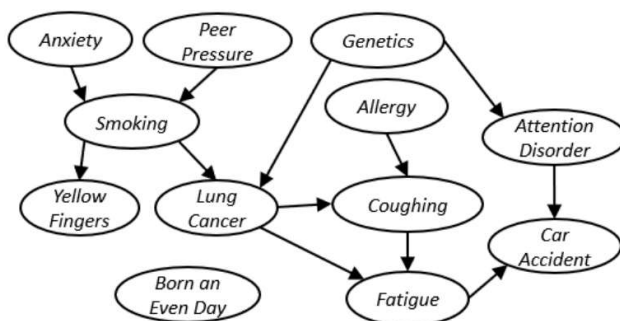
SACHS : d=11, e=17, n=853



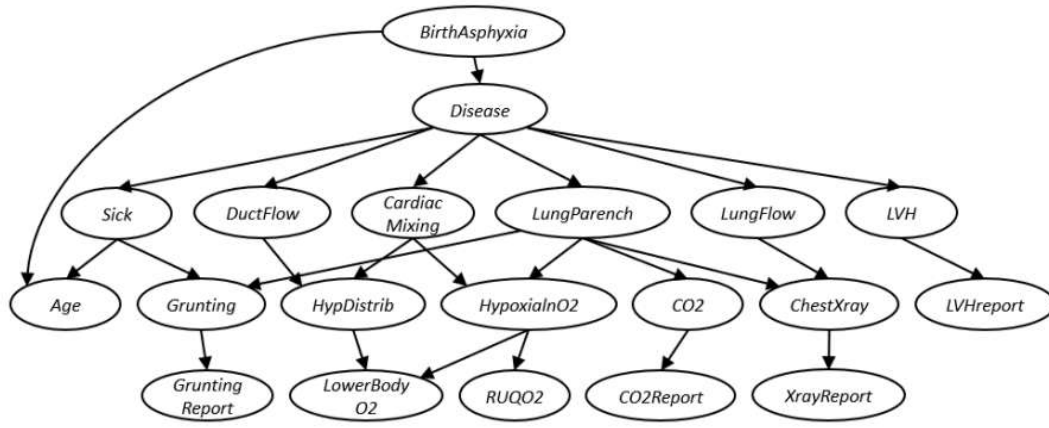
肺がん

LUCAS: d=12, e=12, n=10000

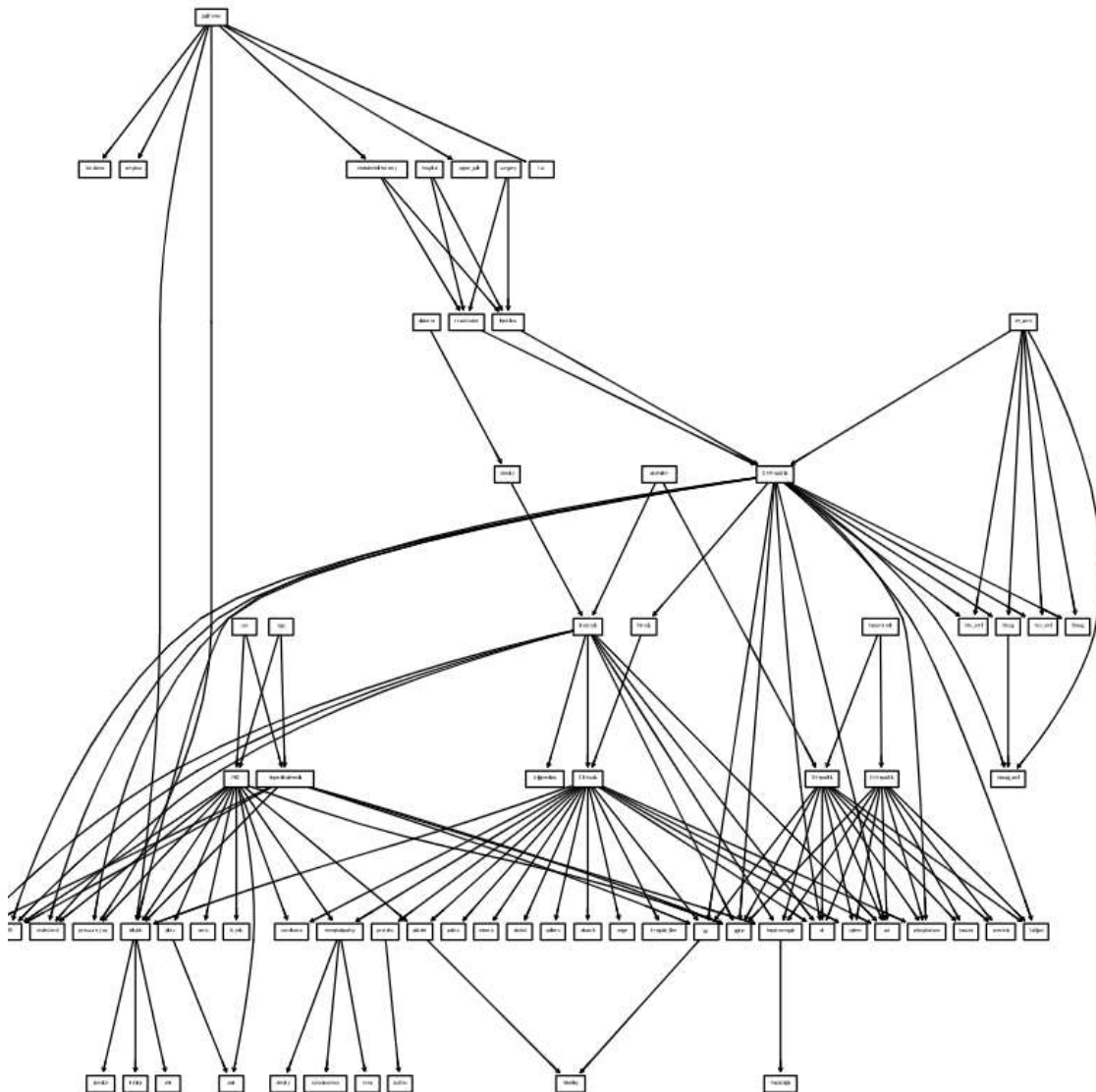
ASIA: d=8, e=8, n=10000



CHILD: d=20, e=25



HEPAR2: d=70, e=123



気候

CauseMe - Earth Sciences data

N=10~12, T=1000~5000

DREAM3

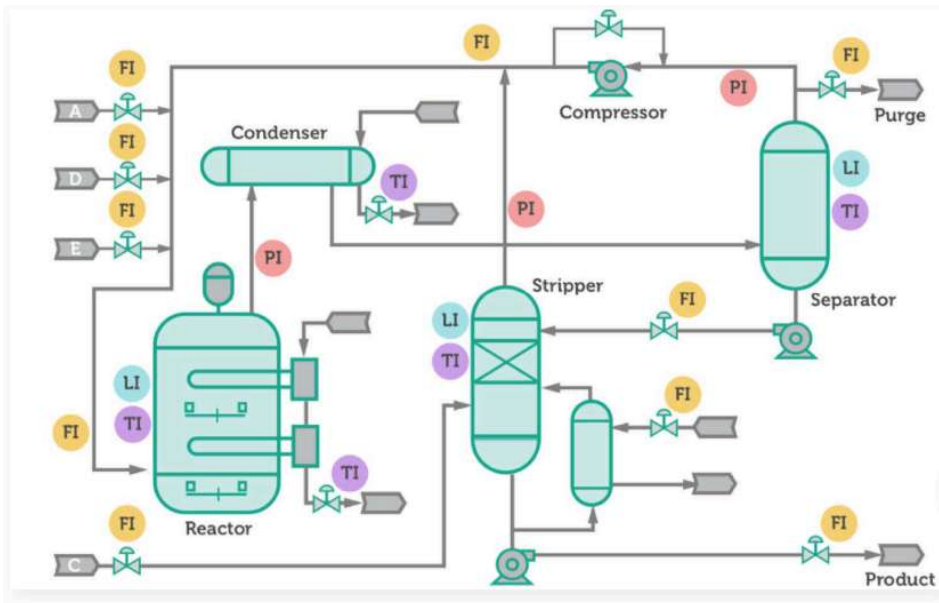
DREAM3 challenge datasets: DREAM3 (Prill et al. (2010)) is a simulated gene expression dataset often used for evaluating time-series causal discovery algorithms. It has five different datasets of E. coli and yeast gene networks (Ecoli1, Ecoli2, Yeast1, Yeast2, and Yeast3), each consisting of a maximum of 100 variables. Bussmann et al. (2021) used this dataset to evaluate their approach. Every dataset has 46 time series and every time series consists of only 21 timesteps.

<https://github.com/bartbusmann/NAVAR>.

製造業

d=50

variables are sampled every 3 minutes for 25 hours→約500



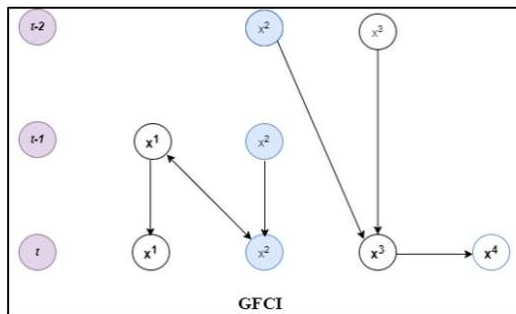
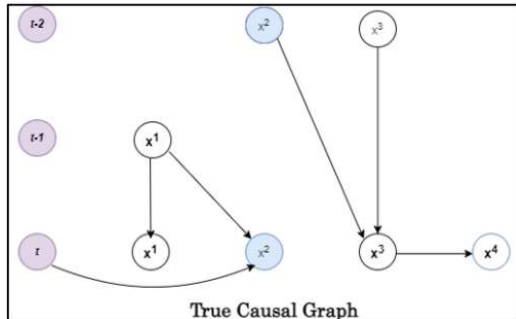
Real Yahoo Services Network Traffic Dataset

This is a univariate time-series dataset containing the traffic to Yahoo services. The anomalies are labeled by humans. This dataset consists of 67 different timeseries each containing about 1400 timestamps.

実験結果

Methods	ASIA			CHILD			ALARM			HEPAR2		
	SHD	TPR	FDR	SHD	TPR	FDR	SHD	TPR	FDR	SHD	TPR	FDR
PC	5	0.6	0.3	43	0.24	0.86	55	0.67	0.6	172	0.35	0.75
GES	<b>4</b>	<b>0.63</b>	0.38	34	0.38	0.89	56	<b>0.74</b>	0.61	<b>70</b>	0.5	0.23
LiNGAM	7	0.25	0.6	23	0.28	0.63	43	0.43	0.55	111	0.1	0.32
Direct-LiNGAM	<b>4</b>	0.5	<b>0</b>	28	0.12	0.82	40	0.39	0.5	110	0.1	0.07
NOTEARS	12	0.13	0.83	<b>22</b>	0.16	0.64	41	0.17	0.38	123	-	-
DAG-GNN	7	0.25	0.5	24	0.24	0.7	<b>39</b>	0.196	<b>0.31</b>	123	0	1
GraN-DAG	7	0.13	<b>0</b>	24	0.04	<b>0</b>	44	0.044	0.75	122	0.008	<b>0</b>
GOLEM	11	0.25	0.75	49	0.2	0.88	60	0.26	0.71	157	0.05	0.89
MCSL	19	0.5	0.82	140	<b>0.56</b>	0.91	464	0.72	0.93	1743	<b>0.45</b>	0.97

Methods	Syn-6 data			fMRI data		
	SHD	TPR	FDR	SHD	TPR	FDR
PCMRI	14	0.63	0.69	61	0.52	0.82
PCMRI+	<b>11</b>	0.50	<b>0.63</b>	<b>26</b>	0.43	<b>0.61</b>
VarLiNGAM	15	0.50	0.73	30	0.48	0.66
DyNOTEARS	28	<b>0.83</b>	0.83	95	<b>0.81</b>	0.84
TCDF	10	0	1	27	0.38	0.64



<https://www.bnlearn.com/bnrepository/discrete-small.html#asia>.

<https://www.causality.inf.ethz.ch/data/LUCAS.html>.

<https://github.com/M-Nauta/TCDF/tree/master/data/fMRI>.


## 参考文献

 A Survey on Causal Discovery Methods for I.I.D. and Time S...

The ability to understand causality from data is one of the major milestones of human-level intelligence. Causal Discovery (CD) algorithms can identify the...

<https://arxiv.org/abs/2303.15027>



 <https://proceedings.mlr.press/v213/howard23a/howard23a.pdf>

<https://proceedings.mlr.press/v213/howard23a/howard23a.pdf>