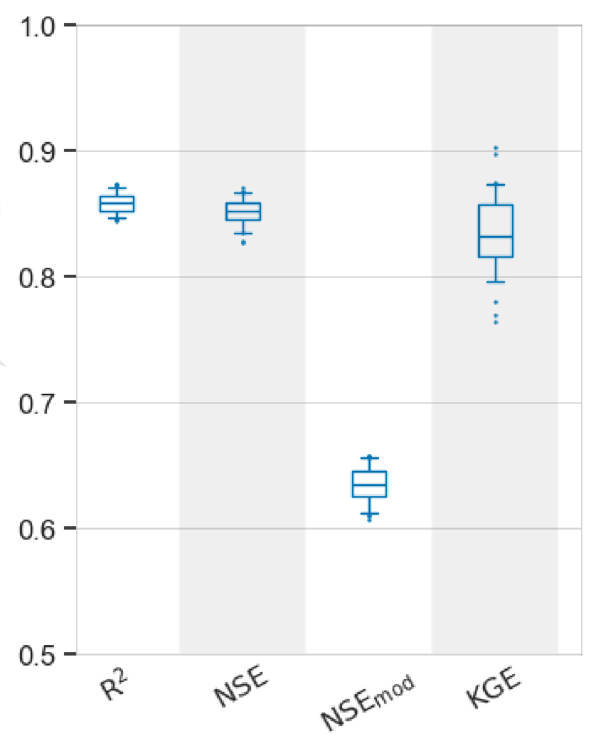
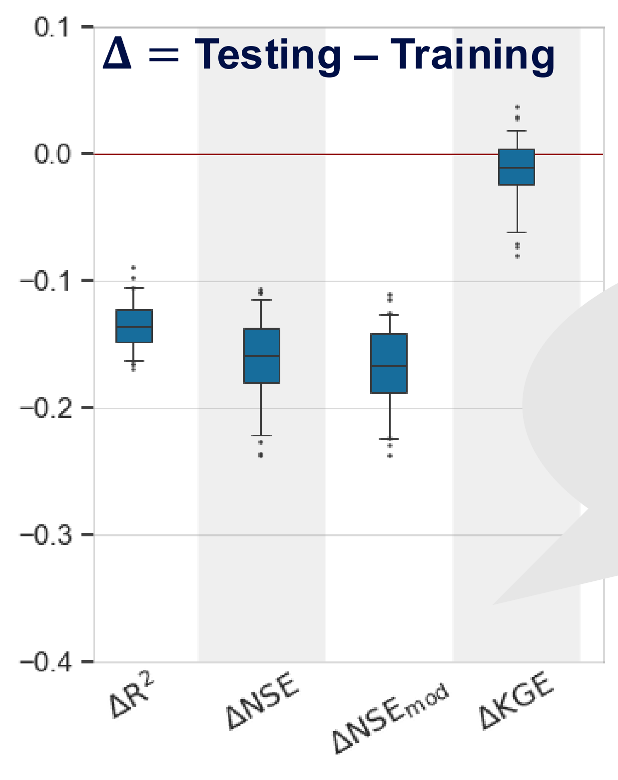
# Predictions under changing climatic conditions

- ➢ In a changing climate Robust models are needed
- ➢ Need to identify *robust* and *direct* drivers of processes

Correlation-based ML prediction of surface soil moisture

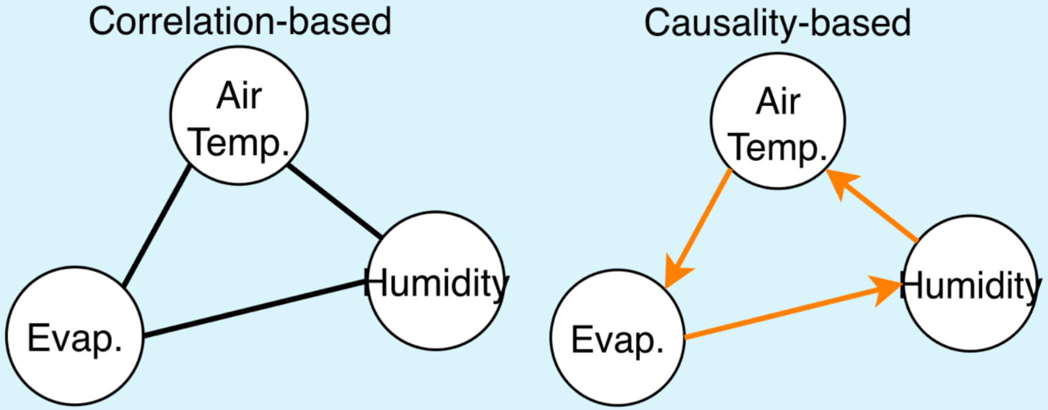
Training in normal conditions

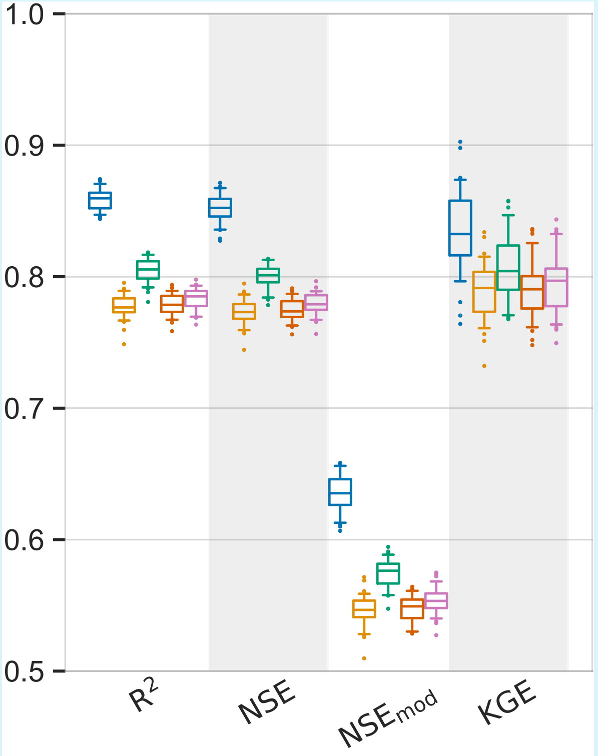


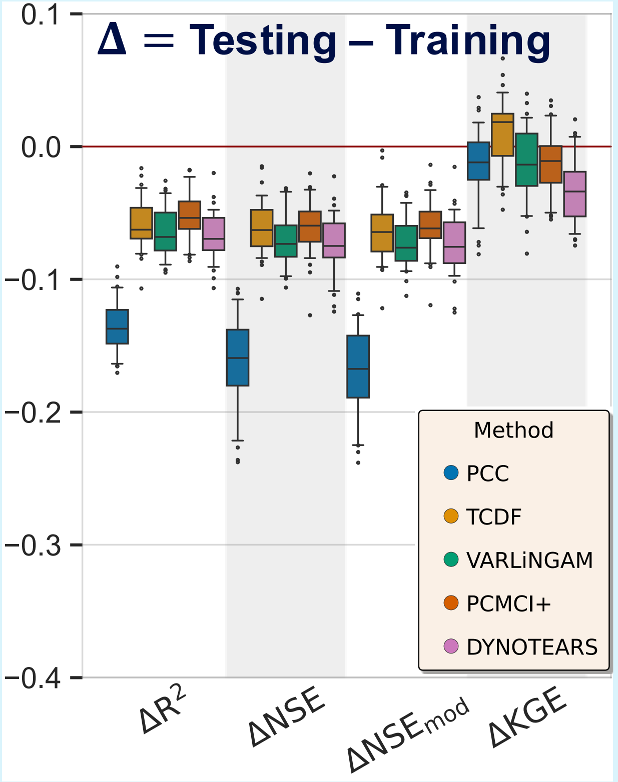

Large drop in performance when tested in drought conditions

# Causality: Science of cause-effect




**Training period**




## Diverse Causal Discovery algorithms

- TCDF
- VARLiNGAM
- PCMCI+
- DYNOTEARS

> ➢ Causality based models are more robust than correlation-based models!

- What are causal discovery methods?
- How do different causal methods work?
- Evaluation of causal methods in hydrometeorological systems
- Causality for robust time-series modelling

# PICO A.6 Cause-effect based modelling for reliable results under changing climatic conditions

**Vivek Kumar Yadav**[1,2,*], Murray Peel[2], Keirnan Fowler[2], Dongryeol Ryu[2], Bramha Dutt Vishwakarma[1]
[1]Interdisciplinary Centre for Water Research, Indian Institute of Science Bengaluru
[2]Faculty of Engineering and Information Technology, The University of Melbourne
[*]viveky@iisc.ac.in

**HS7.4: Future hydroclimatic scenarios in a changing world**

**Preprint on HESS**

Sharing is encouraged

## Objectives

i) Evaluate CD methods in a simulated hydrometeorological system to recover drivers of the system

ii) Causality for robust time-series modelling

## Methodology

GLDAS Land surface model

Timeseries data over 45 grids.
9 Climates, 5 grids

Application of 4 different CD methods and PCC

Efficacy of CD methods: Comparison of estimated drivers and reference truth

Literature review of generating equations

8 forcing variables
19 simulated variables

<u>Reference truth</u>
80 causal links,
1378 links absent

Literature review of CD methods:

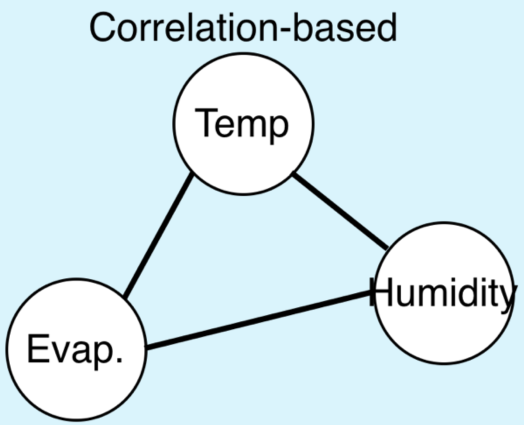
- Stochastic
- Multivariate
- Theoretically distinct

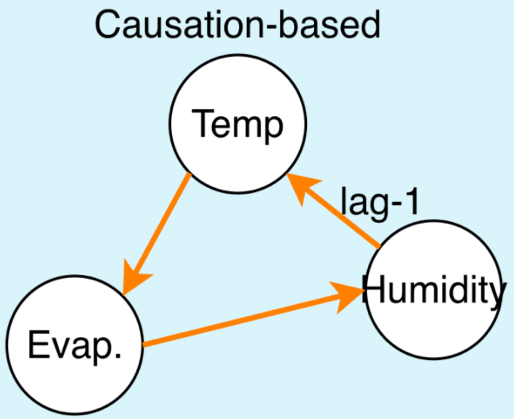
Extraction of drivers of surface SM: split sample timeseries model

# Causality: Science of cause-effect

## Causal Discovery (CD)

Discovering a <u>direct</u> cause-effect relation between entities (variables or events).



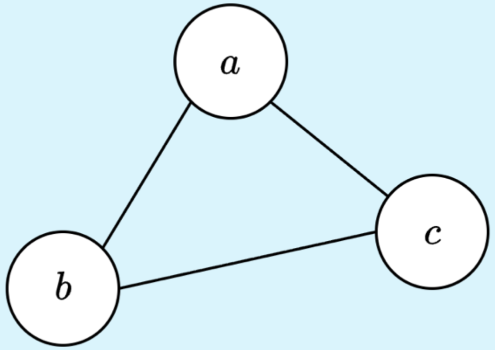

- Structure
- Directionality

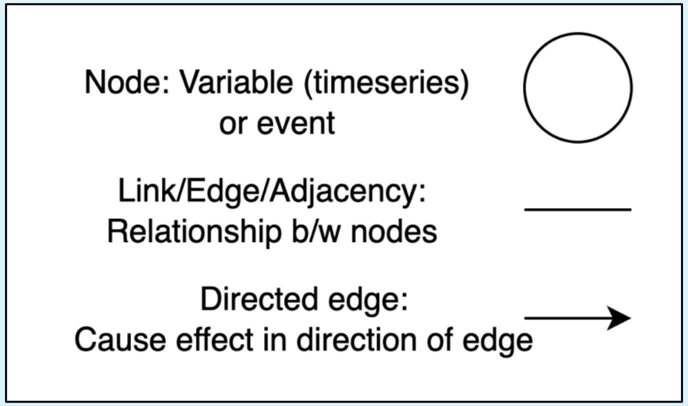
## Causal Inference

Quantifying the influence of an entity (variable or event) onto another.

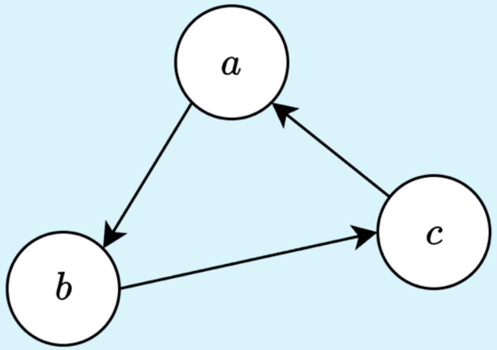
# How to represent causal relations?

Graph

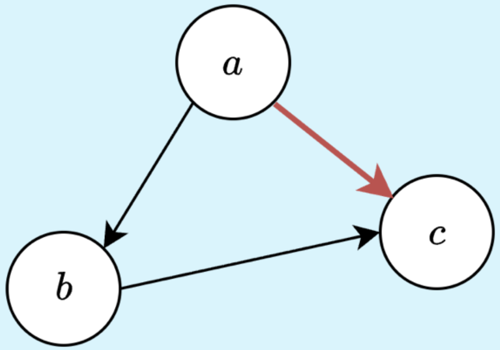



Directed Cyclic Graph




Directed-Acyclic Graph (DAG)




$\longleftrightarrow$

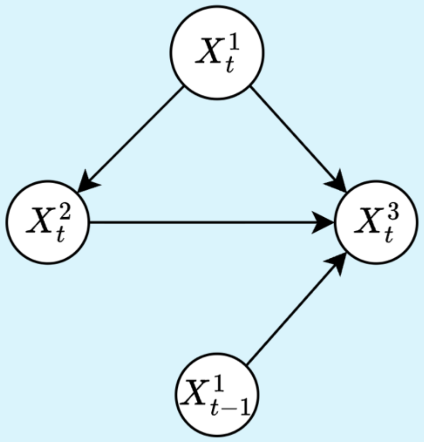
Adjacency Matrix

| | $a$ | $b$ | $c$ |
|---|---|---|---|
| $a$ | 0 | 1 | 1 |
| $b$ | 0 | 0 | 1 |
| $c$ | 0 | 0 | 0 |

# DAGs and their interesting properties

$$
\begin{gathered}
X_t^1 = \eta_t^1 \\
X_t^2 = a \cdot X_t^1 \\
X_t^3 = b \cdot X_t^1 - c \cdot X_t^2 + d \cdot X_{t-1}^1
\end{gathered}
$$

Acyclicity




Lower triangular ordering

| | $X_t^1$ | $X_t^2$ | $X_t^3$ | $X_{t-1}^1$ |
|---|---|---|---|---|
| $X_t^1$ | 0 | 1 | 1 | 0 |
| $X_t^2$ | 0 | 0 | 1 | 0 |
| $X_t^3$ | 0 | 0 | 0 | 0 |
| $X_{t-1}^1$ | 0 | 0 | 1 | 0 |

| | $X_t^3$ | $X_t^2$ | $X_t^1$ | $X_{t-1}^1$ |
|---|---|---|---|---|
| $X_t^3$ | 0 | 0 | 0 | 0 |
| $X_t^2$ | 1 | 0 | 0 | 0 |
| $X_t^1$ | 1 | 1 | 0 | 0 |
| $X_{t-1}^1$ | 1 | 0 | 0 | 0 |

Structural vector autoregressive model (SVAR)

$$
\begin{bmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ b & -c & 0 & d \end{bmatrix} \begin{bmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ X_{t-1}^1 \end{bmatrix}
$$

$$X = AB$$

# How is Casual Discovery done?

- Granger Causality
- Noise-based
- Constraint-based
- Score-based
- State-space reconstruction based: Convergent Cross Mapping
- Others

| | |
|---|---|
| **TCDF** (Nauta et al., 2020)<br><br>- Find DAG<br><br>- Method: CNN to model time-series, Attention scores to unravel causality | **VARLiNGAM** (Hyvärinen et al., 2008)<br><br>- Classic SVAR model<br><br>- Method: Exploits lower triangular ordering of adjacency matrix |
| **PCMCI+** (Runge et al., 2022)<br><br>- Find DAG<br><br>- Method: Conditional independence tests | **DYNOTEARS** (Pamfil et al., 2020)<br><br>- Classic SVAR model<br><br>- Method: Numerical optimization to minimize the model error and acyclicity score |

# TCDF

- **T**emporal **C**ausal **D**iscovery **F**ramework.
- Causal modelling: DAG.
- Main idea: CNN model with *attention* mechanism.

$X_t^j$ for all $j \in 1,2,\ldots,d$

↓

A CNN model to predict $X_t^j$ based on all remaining variables and auto-correlations: $\boldsymbol{X}_t \setminus \{X_t^j\}$

↓

Potential causes of $X_t^j$:
*attention* weights of CNN model

↓

Verification of potential causes:
Permutation Importance Validation Method

↓

Delay discovery:
Order of causal links (lag or contemp.) from verified causes to $X_t^j$ :
interpretation of CNN kernel weights

If $var\left(Y_t - f(X_{t-\tau}, Y_{t-\tau})\right)$
$<$
$var\left(Y_t - f(Y_{t-\tau})\right)$

Then $X_t$ *Granger causes* $Y_t$

# VARLiNGAM

- **V**ector **A**uto**r**egressive **Li**near **N**on-**G**aussian **A**dditive **M**odel
- Causal modelling: SVAR
- Main idea: VAR model and lower triangular ordering of adjacency matrix.

Find the coefficients of a SVAR model

Breakdown the coefficient matrix into lagged relations: $B_{t-\tau}$for $\tau > 0$ and contemporaneous relations matrix: $B_0$

$B_{t-\tau}$ estimated with least squares fit

$B_0$ estimated by exploiting lower triangular ordering of adjacency matrix

If variables are less: Matrix row operations

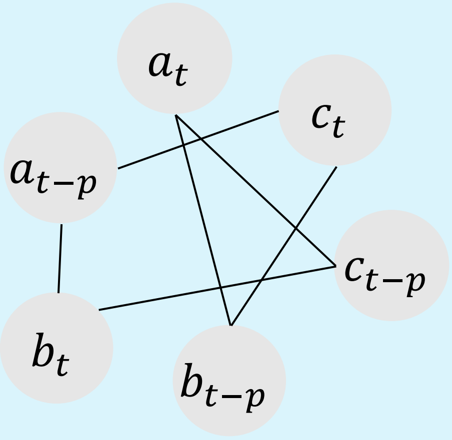
If variables are more: Independent component analysis

Combine $B_0$ and $B_{t-\tau}$ to represent causal relations

# PCMCI+

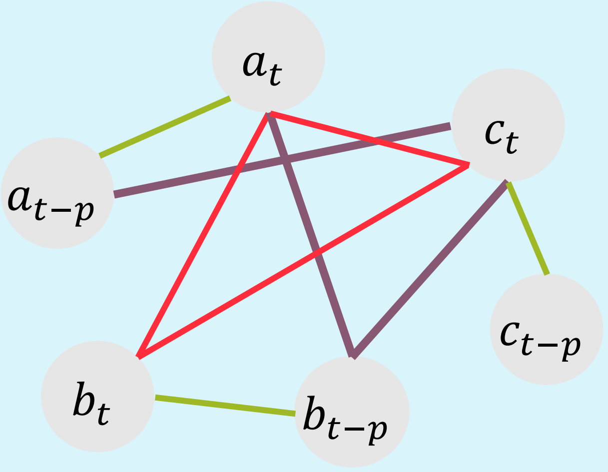
- Combination of Peter-Clark's algorithm and Momentary Conditional Independence tests.
- Causal modelling: DAG(graphs).
- Main idea: Conditional independence testing.

### I. Lagged skeleton phase




PC: $X_t^j \perp X_{t-\tau}^i \mid S$ where $S \subseteq Pa\left(X_t^j\right)\backslash\{X_{t-\tau}^j\}$ and $|S| = \tau$ for $\tau = 1,2,..,p$

### II. Contemporaneous skeleton phase




MCI: $X_t^j \perp X_{t-\tau}^i \mid S, \boxed{B_{t-\tau}^-\left(X_t^j\right)\backslash\{X_{t-\tau}^i\}B_{t-\tau}^-\left(X_{t-\tau}^i\right)}$
where $S \subseteq \hat{A}\left(X_t^j\right)\backslash\{X_{t-\tau}^j\}$ and $|S| = \tau$ for $\tau = 0,1,2,..,p$

# DYNOTEARS

- Causal modelling: SVAR.
- Numerical optimization to jointly minimize the model error and acyclicity score.

SVAR model

$$\boldsymbol{X}_t = X_t\boldsymbol{W} + X_{t-1}A^1 + \cdots + X_{t-p}A^p$$

$$\boldsymbol{X}_t = \boldsymbol{X}_t\boldsymbol{W} + \boldsymbol{X}^{-}\boldsymbol{A}$$

Acyclicity of a matrix **W**

$$h(\boldsymbol{W}) = trace(e^{W \circ W}) - d$$

$h(\boldsymbol{W}) = 0$ if and only if $\boldsymbol{W}$ is acyclic.

$F(\boldsymbol{W}, \boldsymbol{A}) =$ (Numerical Optimization)

$$\underbrace{\frac{1}{2d(T+1-p)}\|\boldsymbol{X} - \boldsymbol{X}\boldsymbol{W} - \boldsymbol{X}^{-}\boldsymbol{A}\|_F^2}_{\text{model error}} \underbrace{+ \alpha h(\boldsymbol{W}) + \frac{\rho h(\boldsymbol{W})^2}{2}}_{\text{acyclicity score}} \underbrace{+ \lambda_A\|\boldsymbol{A}\|_1}_{\text{sparsity of lagged adj matrix}} \underbrace{+ \lambda_W\|\boldsymbol{W}\|_1}_{\text{sparsity of contemp. adj matrix}}$$

# Comparing two adjacency matrices

- One-to-one comparison of adjacency matrices yields a confusion matrix.

**Confusion matrix**

| Predicted \ Actual | | |
|---|---|---|
| | True Positive | True Negative |
| | False Positive | False Negative |

Recall (*TPR*): ability to accurately identify the <u>correct links</u> in the adjacency matrix

$$= \frac{True\ Positive}{True\ Positive + False\ Negative}, \in [0, 1]$$

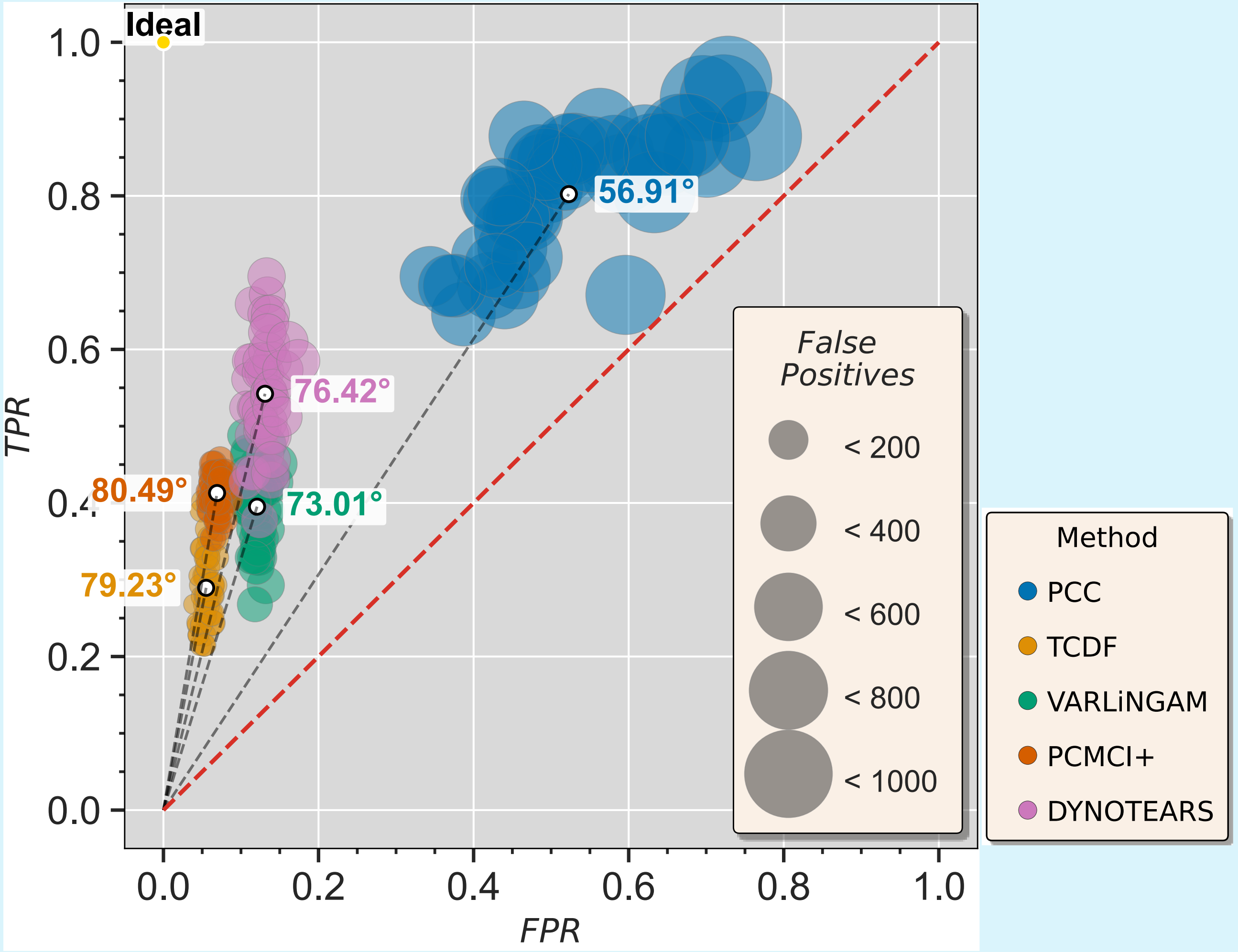
False positive ratio (*FPR*): ability to accurately identify the <u>incorrect links</u> in the adjacency matrix

$$= \frac{False\ Positive}{False\ Positive + True\ Negative}, \in [0, 1]$$

# Results

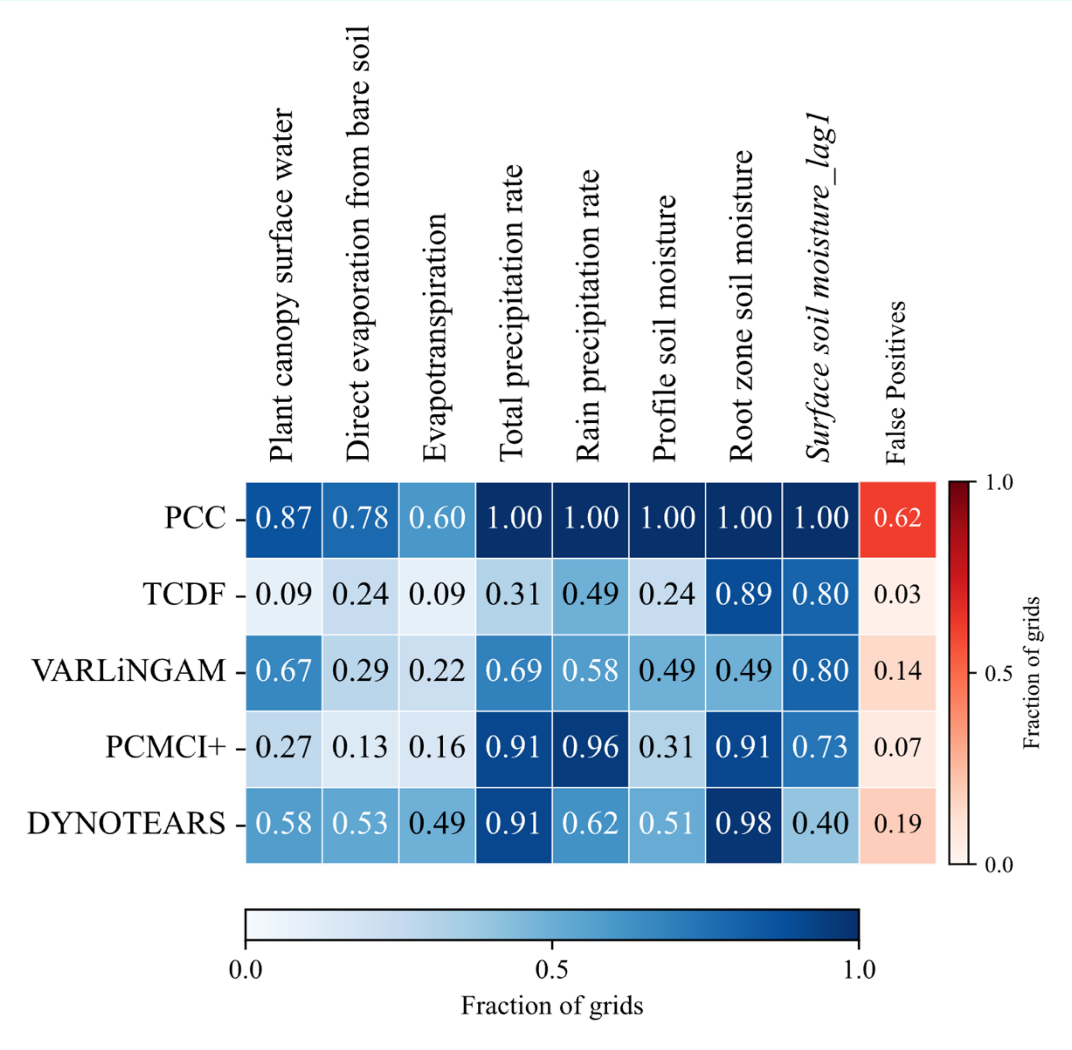
# Identifying causal drivers of the complete system

- Cost of identifying causal links is the accumulation of false positives.
- PCC shows high variance along TPR and FPR, indicating lack of robustness.
- Comparatively, CD methods show less variance along FPR, demonstrating robustness towards eliminating false positives.
- For a unit gain in FPR, CD methods have a higher TPR.



# Causal drivers of surface soil moisture

- Compared to PCC, CD methods identify fewer causal drivers.
- However, CD methods eliminate large number of correlated but *causally unrelated* variables.
- This allows a parsimonious surface soil moisture representation and model.

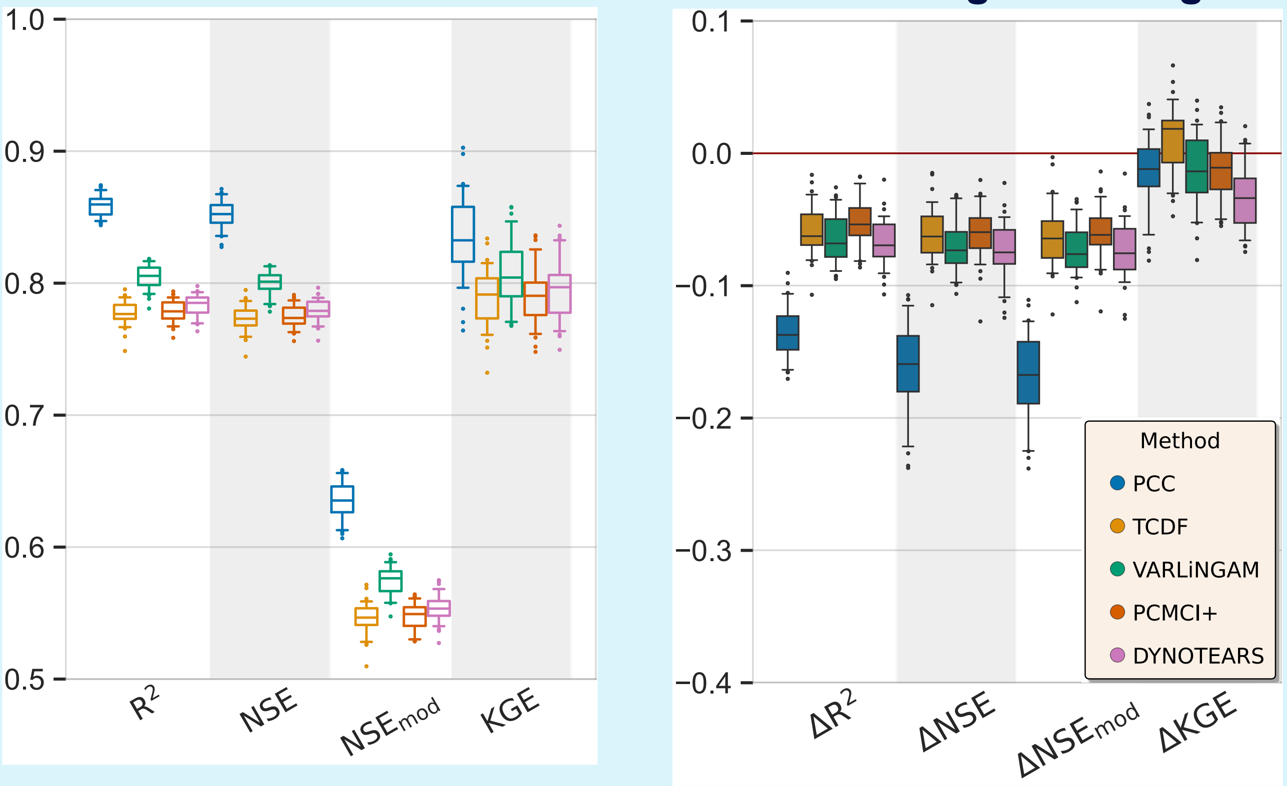


# Time-series prediction with causal drivers

| | |
|---|---|
| Location | Ganga basin, India |
| Model | Support Vector Regression |
| Training | Normal period<br>(01 Jan 2000 to 31 December 2003) |
| Testing | Drought period<br>(01 Jan 2004 to 31 December 2005) |
| Noise added | Random Gaussian (0.2 *std. dev.*)<br>(100 Monte Carlo simulations) |

**Performance metrics**

**Training period** | **$\Delta$ = Testing – Training**



# Conclusions

- Hydrometeorological systems are highly interconnected, with many cross and lagged correlations.
- Causal Discovery can help identify direct process drivers of variables in such systems.
- Robust identification of drivers is important for predictions, especially under changing conditions like drought, climate change.

# References


1. Ombadi, Mohammed, et al. "Evaluation of methods for causal discovery in hydrometeorological systems." Water Resources Research 56.7 (2020): e2020WR027251.
2. Pearl, Judea. Causality. Cambridge university press, 2009.
3. Runge, Jakob, et al. "Inferring causation from time series in Earth system sciences." Nature communications 10.1 (2019): 2553.
4. Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. The MIT press, 2017.
5. Nauta, Meike, Doina Bucur, and Christin Seifert. "Causal discovery with attention-based convolutional neural networks." Machine Learning and Knowledge Extraction 1.1 (2019): 19.
6. Pamfil, Roxana, et al. "Dynotears: Structure learning from time-series data." *International Conference on Artificial Intelligence and Statistics*. Pmlr, 2020.
7. Hyvärinen, Aapo, Shohei Shimizu, and Patrik O. Hoyer. "Causal modelling combining instantaneous and lagged effects: an identifiable model based on non-Gaussianity." *Proceedings of the 25th international conference on Machine learning*. 2008.
8. Runge, Jakob. "Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets." *Conference on uncertainty in artificial intelligence*. Pmlr, 2020.

# Preprint on HESS

Preprint

Preprints / Preprint egusphere-2025-4650

Search

https://doi.org/10.5194/egusphere-2025-4650


Abstract | Discussion | Metrics

05 Nov 2025

Cause-effect discovery in Hydrometeorological Systems: Evaluation of Causal Discovery methods

Vivek Kumar Yadav, Murray Peel, Keirnan Fowler, Dongryeol Ryu, and Bramha Dutt Vishwakarma

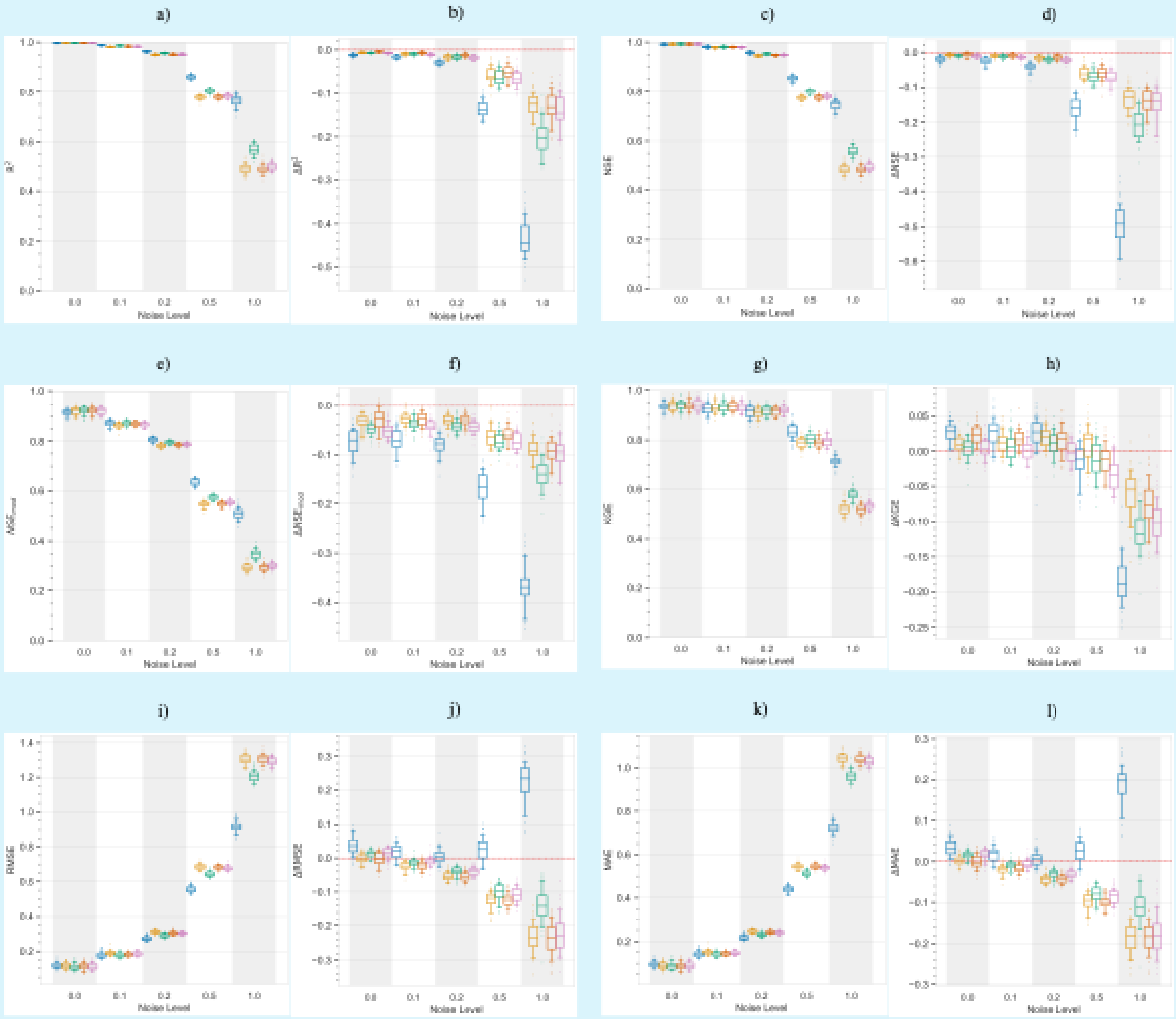
# Scan to read

# Thank you!

Sharing is encouraged

# Appendix

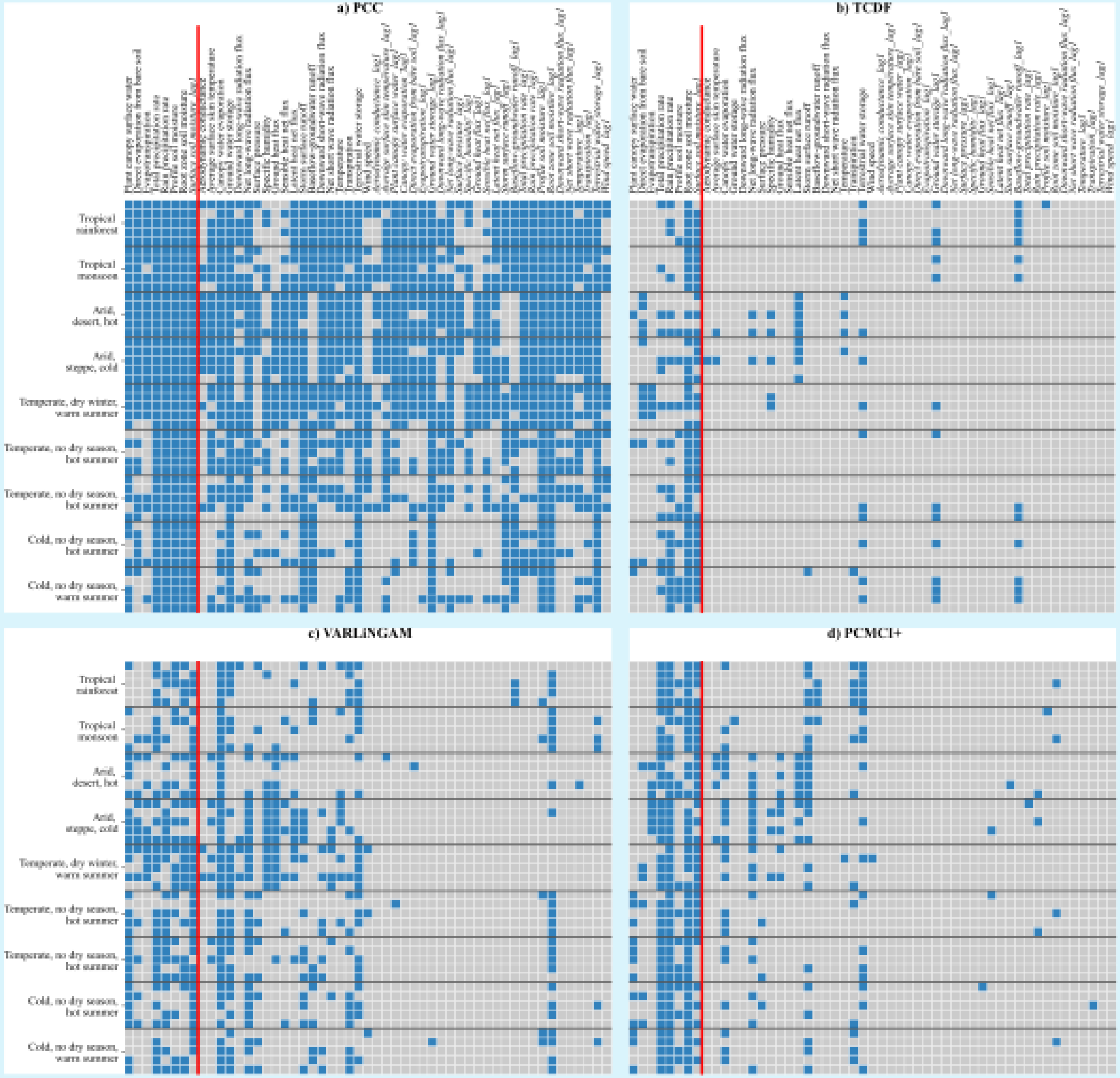
Fig. Robustness check against realization and level of added noise.



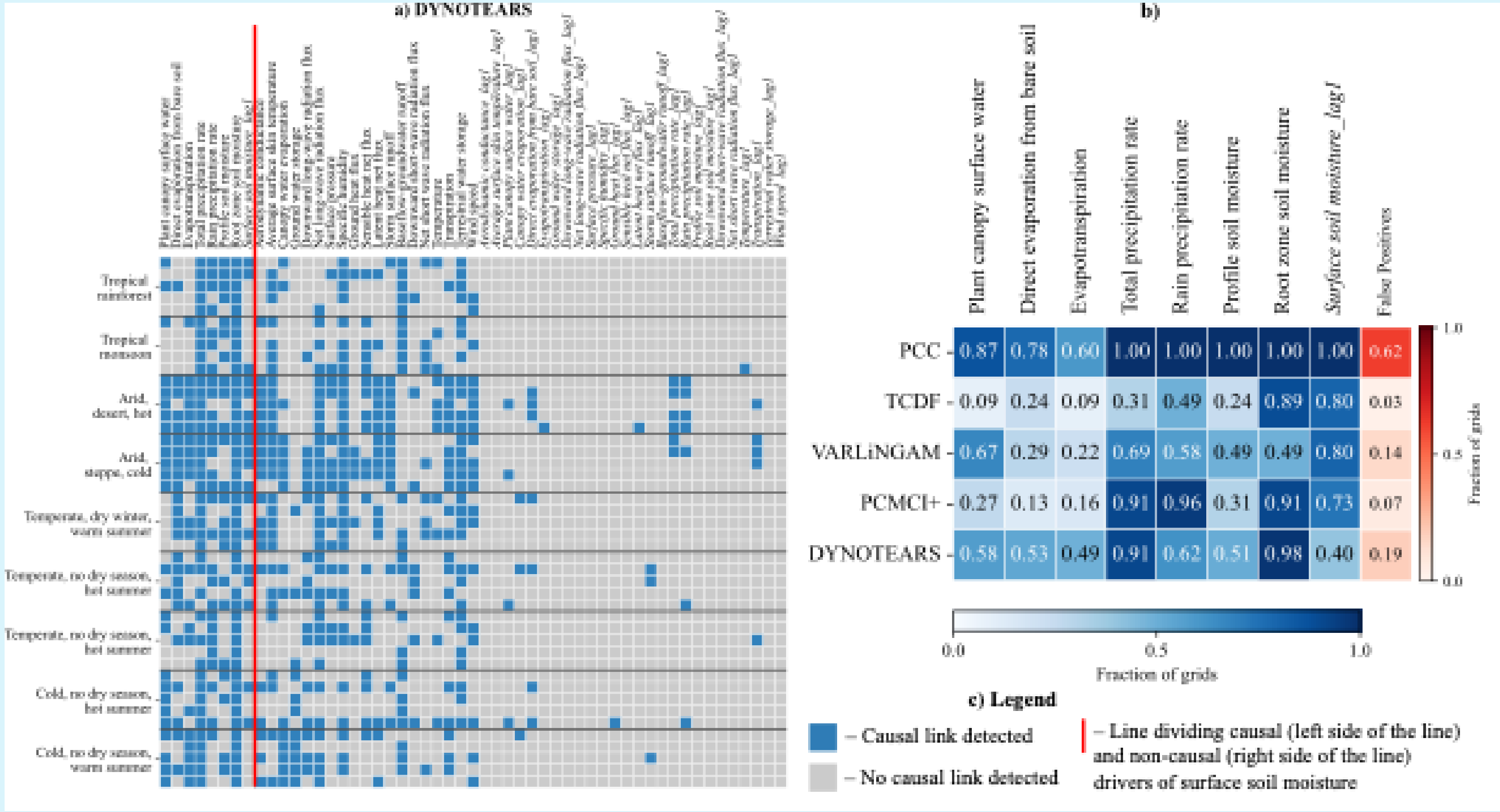

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

THE UNIVERSITY OF MELBOURNE

# Fig. Drivers of surface soil moisture across different climate regions



# Fig. Drivers of surface soil moisture across different climate regions



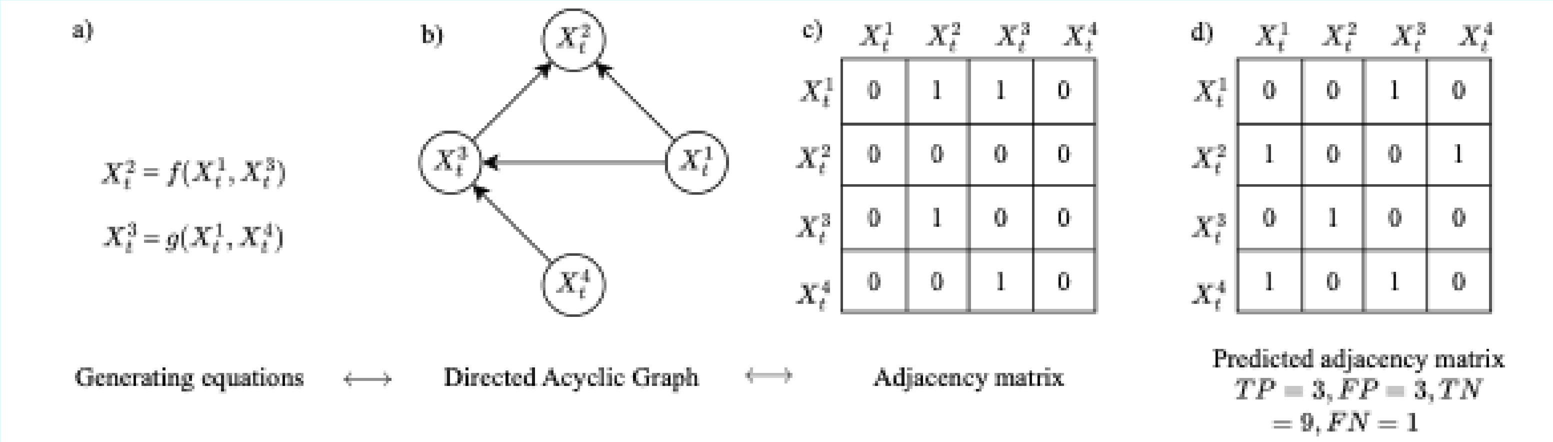



Fig. Equivalent forms of representing causal relations of a SVAR model (a) with DAG (b) and Adjacency matrix (c) along with (d) an exemplar estimated adjacency matrix.

Fig. True CLSM adjacency matrix and locations from where data was extracted for causal analysis.