

# CAUSAL INFERENCE AND CAUSAL DISCOVERY FOR STUDYING TELECONNECTION PATHWAYS

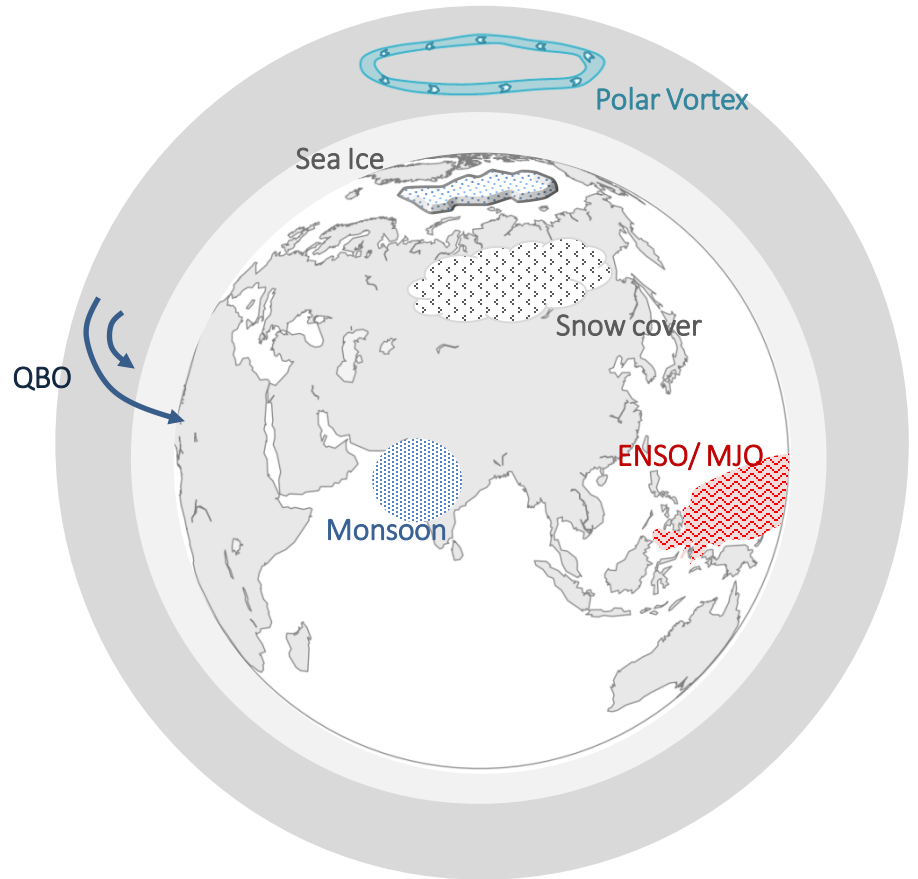
Marlene Kretschmer<sup>1</sup>

Ted Shepherd<sup>1</sup>, Elena Saggioro<sup>1</sup>, Alberto Arribas<sup>2</sup>,  
Rachel Prudden<sup>2</sup>, Niall Robinson<sup>2</sup>, Sam Adams<sup>2</sup>

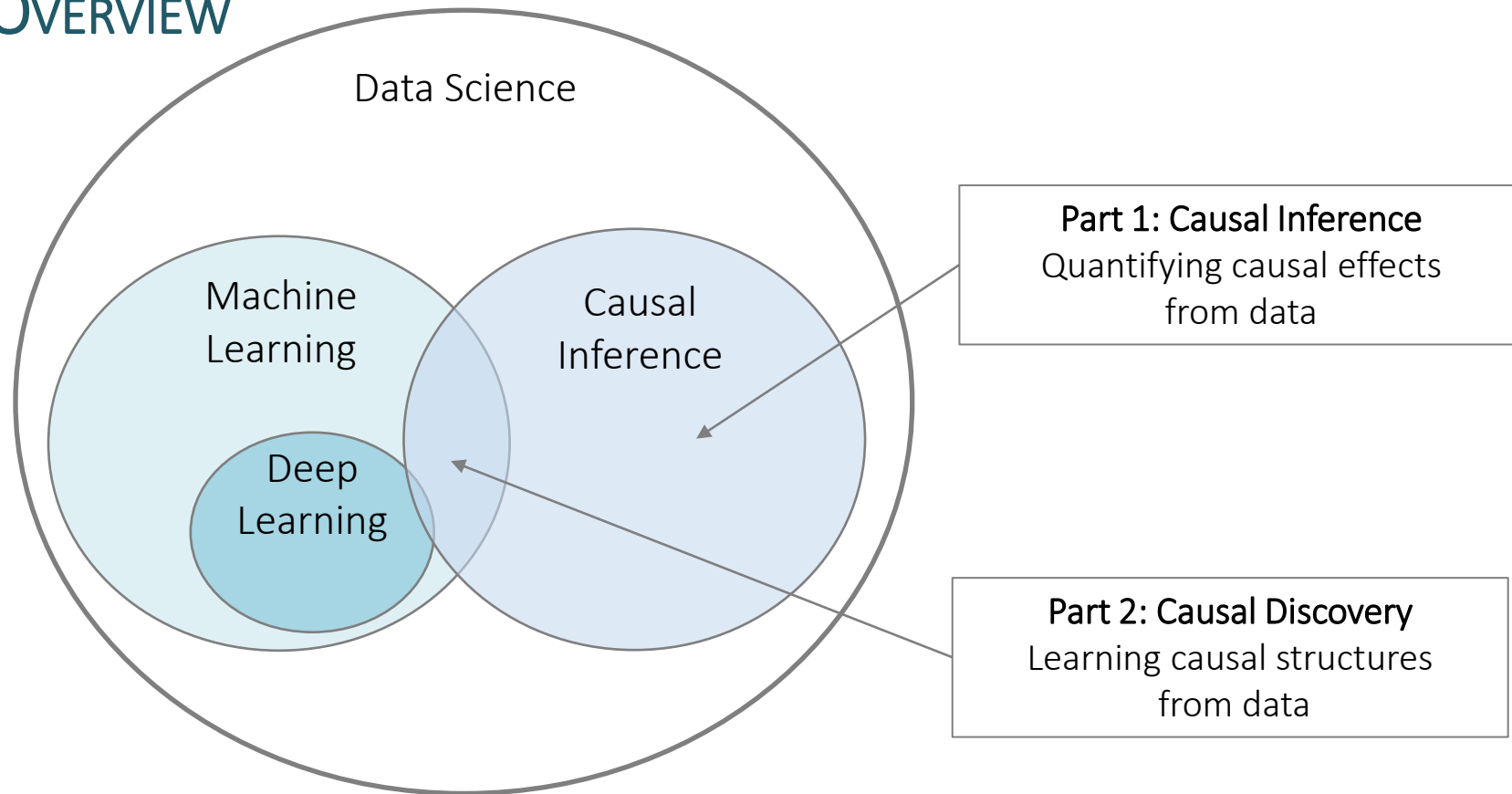
<sup>1</sup>University of Reading, <sup>2</sup>Met Office Informatics Lab

# INTRODUCTION

- Motivation:  
Improved understanding of teleconnections is key to reduce uncertainties about regional weather and climate predictions
- Typical questions:  
How much does ENSO contribute to temperature variability in region A?  
Which processes drive precipitation in region B?
- Challenge:  
Extracting the (causal) information from data



# OVERVIEW



# TELECONNECTIONS

## AMS Definition

A significant [...] correlation in the fluctuations of a field at widely separated points.

[...] the name refers to the fact that such correlations suggest that information is propagating between the distant points through the atmosphere.

Large gap between our physical understanding (= causal) and our statistical description (= correlational) of teleconnections.

# CAUSALITY IN STATISTICS

X causes Y, if intervening in X  
(while keeping everything else fixed)  
changes Y

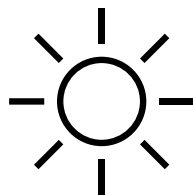
$$P(Y \mid do(X) = x) \neq P(Y)$$

## Causal Inference:

Predict the effect of an intervention from  
observed data  
(without actually doing the intervention)

## Example

X: Temperature



Y: Thermometer



Changes in X will cause changes in Y:

$$P(Y) \neq P(Y \mid do(X) = x)$$

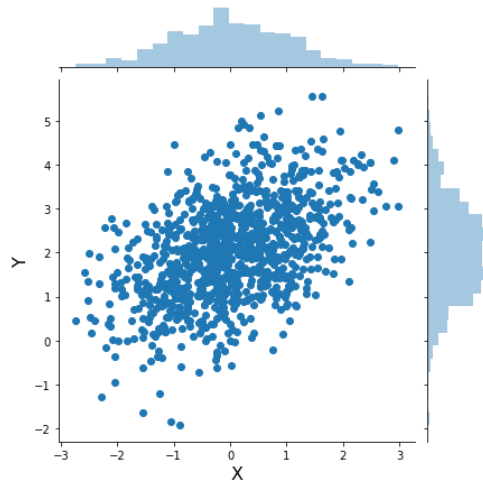
Intervening in Y will not change X:

$$P(X) = P(X \mid do(Y) = y)$$

# TOY EXAMPLE

Question: What happens to Y if X is changed to  $x=1$ ?

Answer only possible if we have knowledge about the underlying causal mechanisms that generated the data



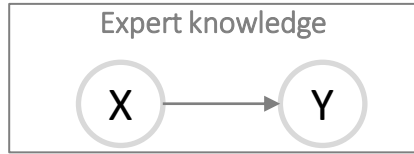
We can compute the observational cond. distribution  
 $P(Y | X = 1)$

We want the interventional cond. distribution  
 $P(Y | \mathbf{do}(X) = 1)$

But these are usually not the same  
 $P(\text{Temp} | \text{Thermo} = 1) \neq P(\text{Temp} | \mathbf{do}(\text{Thermo}) = 1)$

# TOY EXAMPLE

Question: What happens to Y if X is changed to  $x=1$ ?

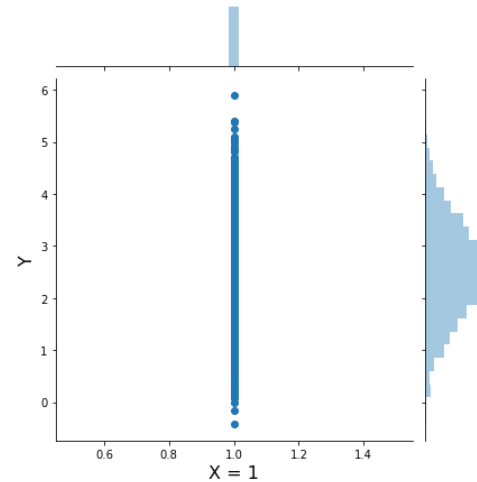
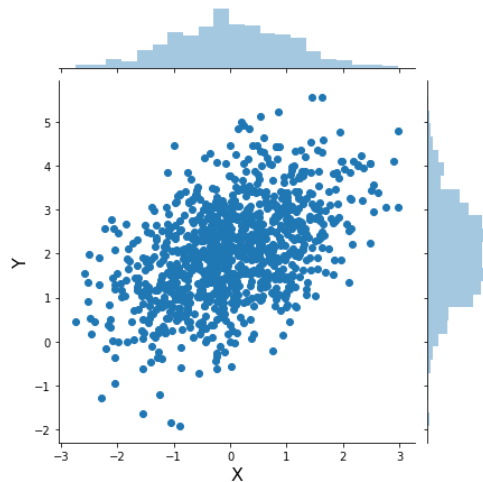


Estimate from data

$$X = \epsilon_X$$
$$Y = 0.5 * X + 2 + \epsilon_Y$$

Answer causal questions

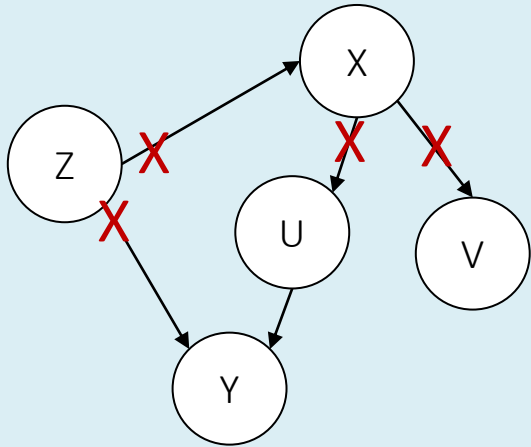
$$P(Y \mid \text{do}(X) = 1) = P(Y \mid X = 1)$$



To answer causal questions, we need causal knowledge (i.e. hypotheses)

# CAUSAL INFERENCE (= “KNOWLEDGE GUIDED STATISTICS”)

Use knowledge to set a (plausible) causal model



Draw logical implications and test if data support them

U and V must be independent conditioned on X

$$\text{Corr}(U, V \mid X) = 0$$

Use CI rules to estimate causal effects

To estimate the causal effect from X to Y, need to control for Z

$$P(Y \mid \text{do}(X)) = P(Y \mid X, Z)$$



# EXAMPLE 1: COMMON DRIVER



Precipitation in Denmark and the Mediterranean are significantly correlated

$$\text{Corr}(\text{DK}, \text{MED}) = -0.25$$

DK and MED are independent conditional on NAO

$$\text{DK} = 0.01 \text{ MED} - 0.55 \text{ NAO} + \varepsilon$$

$$\text{MED} = 0.0 \text{ DK} + 0.42 \text{ NAO} + \varepsilon$$

The causal effect explains the association

$$-0.55 * 0.42 \approx -0.25$$

(JJA mean, NCEP)

## EXAMPLE 2: MEDIATOR

What is the effect of ENSO on California precipitation?

$$CA = \mathbf{0.0} \text{ ENSO} + 0.81 \text{ Jet} + \varepsilon$$

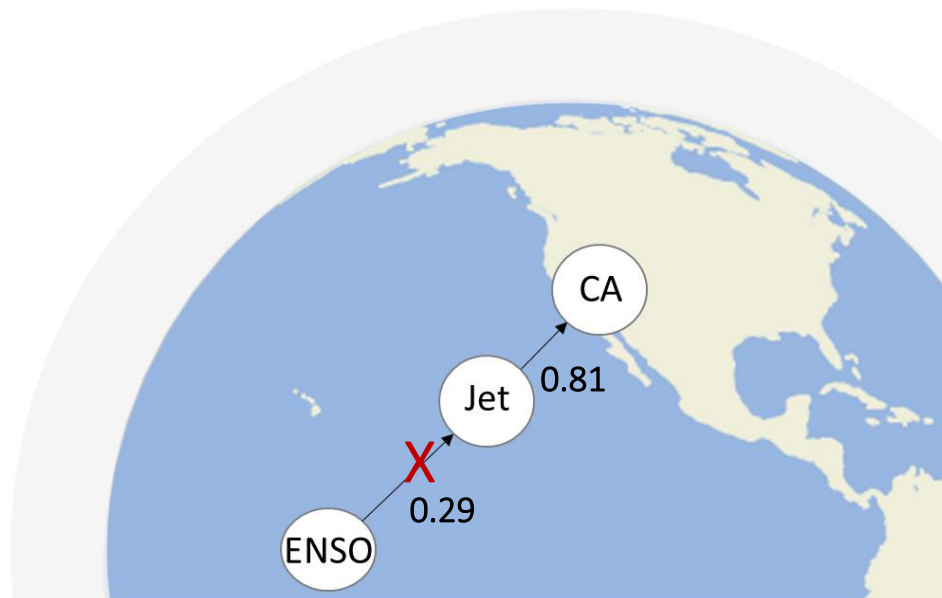
Correct way:

$$CA = \mathbf{0.24} \text{ ENSO} + \varepsilon$$

Or via product along pathway:

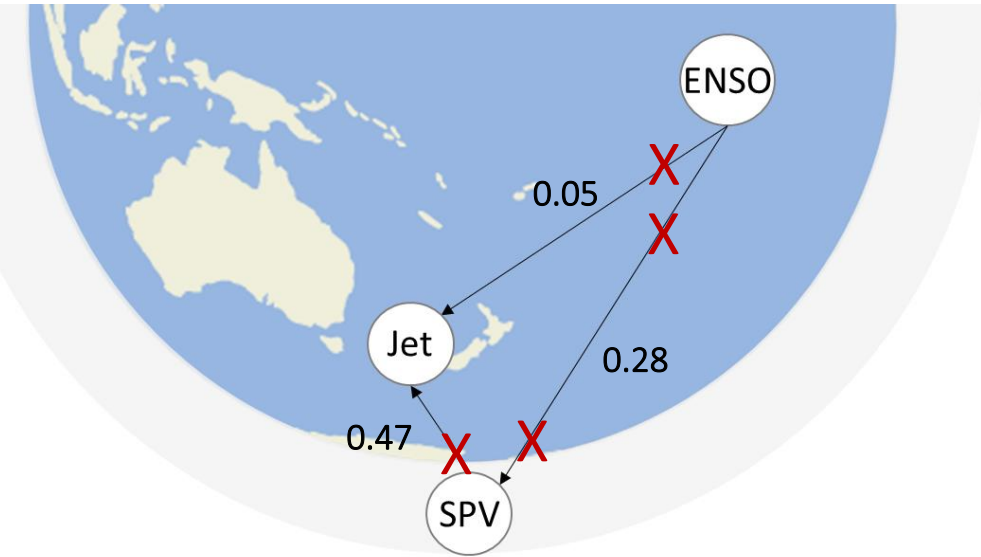
$$\text{Jet} = \mathbf{0.29} \text{ ENSO} + \varepsilon \quad CA = \mathbf{0.81} \text{ Jet} + \varepsilon$$

$$0.29 * 0.81 = 0.24$$



(DJF mean, NCEP)

# EXAMPLE 3: INDIRECT + DIRECT EFFECTS



(OND mean, NCEP)

Total effect

$$\text{Jet} = \mathbf{0.18} \text{ ENSO} + \epsilon$$

Direct (tropospheric) pathway:

$$\text{Jet} = \mathbf{0.05} \text{ ENSO} + a \text{ SPV} + \epsilon$$

Indirect (stratospheric) pathway :

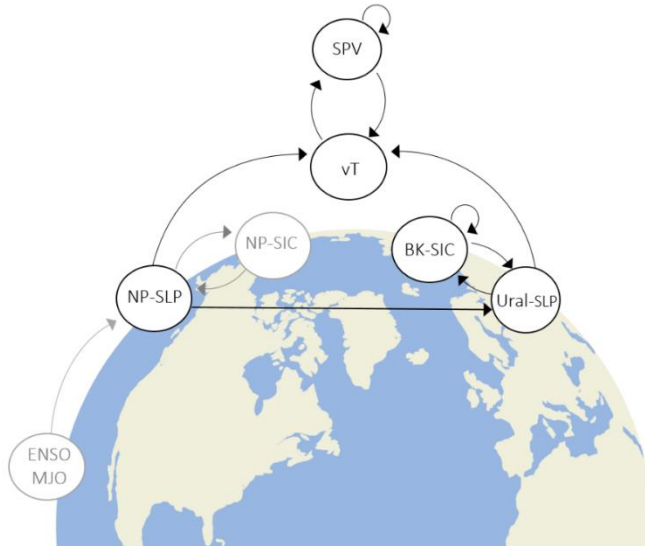
$$\text{SPV} = \mathbf{0.28} \text{ ENSO} + \epsilon$$

$$\text{Jet} = \mathbf{0.47} \text{ SPV} + a \text{ ENSO} + \epsilon$$

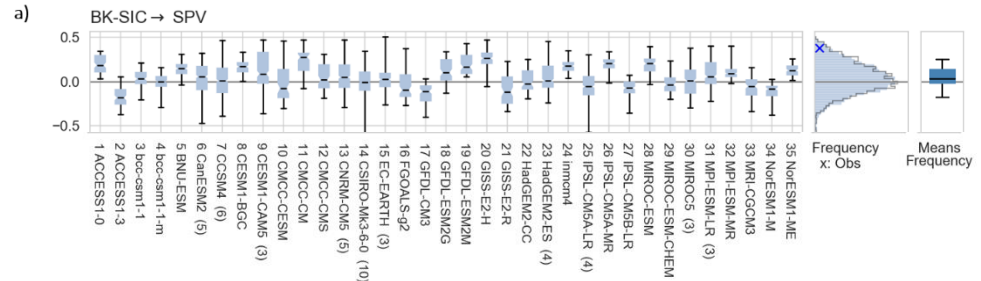
$$\mathbf{0.13} = 0.28 * 0.47$$

# OPPORTUNITIES & APPLICATIONS

- Easy, transparent and traceable approach to quantify teleconnection pathways
- Process-based model evaluation and bias adjustments



What is the role of Barents and Kara sea ice loss in future polar vortex changes in CMIP5 models?

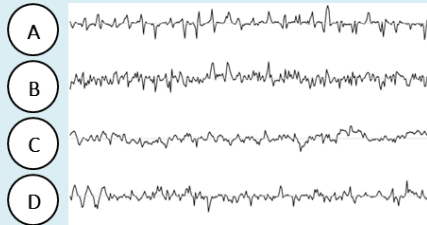


# OPPORTUNITIES & APPLICATIONS

- Easy, transparent and traceable approach to quantification teleconnection pathways
- Process-based model evaluation and bias adjustments (e.g. Kretschmer et al. (2020))
- Understanding signal to noise issues
- Robust emergent constraints
- Physically self-consistent storylines of regional climate change (e.g. Shepherd.(2018))
- Climate modelling (e.g. Hirt et al.(2020))
- Framework to connect the AI and the climate dynamics communities

# PART 2: CAUSAL DISCOVERY

Input: Time-series data



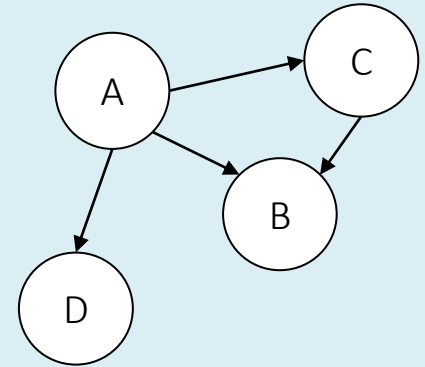
Causal Discovery

PCMC1 Algorithm

$\text{corr}(A_{t-\tau}, B_t \mid \text{Iterate through combinations of conditions})$

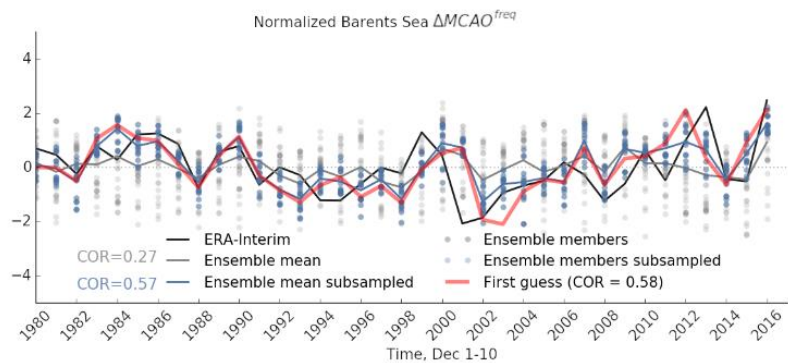
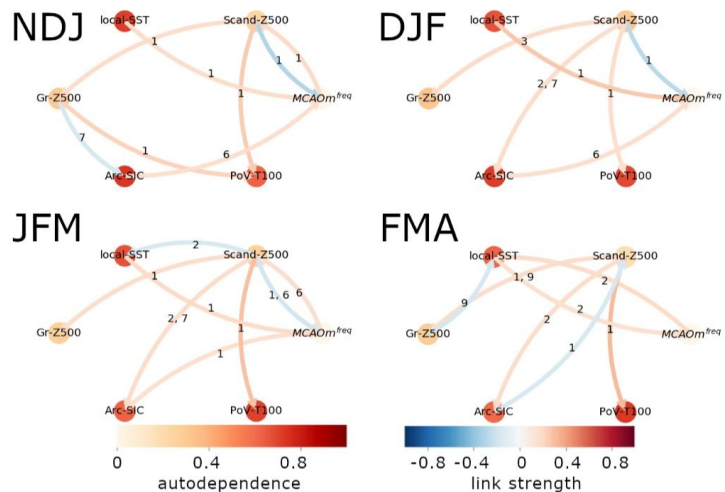
Can deal with auto-correlation,  
regime-dependence,  
instantaneous links, ...

Output: Causal Structure



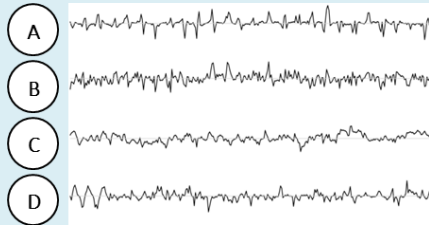
# EXAMPLE: ENSEMBLE SUBSAMPLING

“Predictors and prediction skill for marine cold air outbreaks over the Barents Sea”



# CAUSAL DISCOVERY

Input: Time-series data



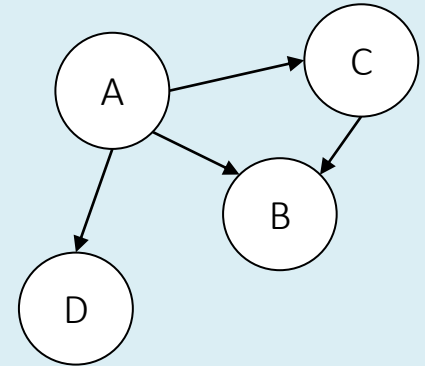
Causal Discovery

PCMCI Algorithm

$\text{corr}(A_{t-\tau}, B_t \mid \text{Iterate through combinations of conditions})$

Can deal with auto-correlation,  
regime-dependence,  
instantaneous links, ...

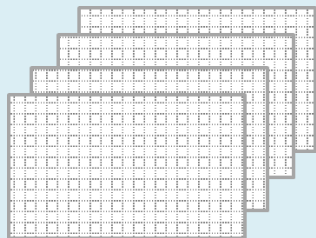
Output: Causal Structure



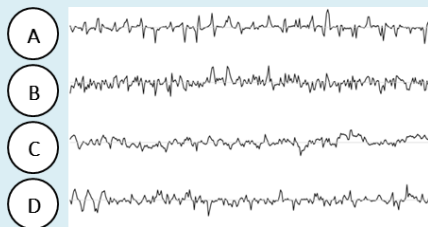


# CAUSAL DISCOVERY

Input: Gridded Climate data



Machine Learning



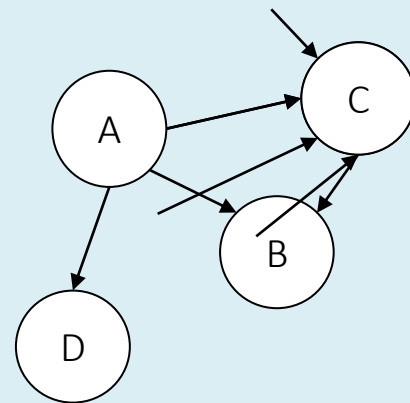
Causal Discovery

PCMCI Algorithm

$\text{corr}(A_{t-\tau}, B_t \mid \text{Iterate through combinations of conditions})$

Can deal with auto-correlation, regime-dependence, instantaneous links, ...

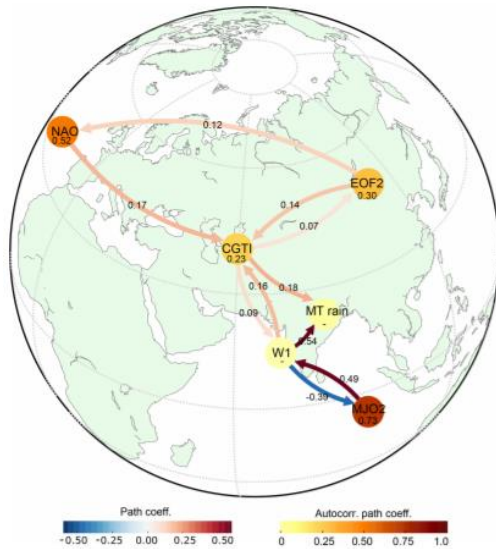
Output: Causal Structure



# APPLICATIONS (DOMAIN KNOWLEDGE REQUIRED)

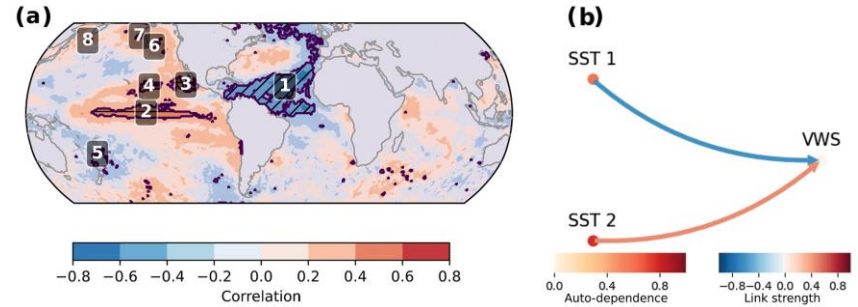
## Indian Summer Monsoon

Di Capua et al. (2019), *ESD*



## Hurricane Activity

Pfleiderer et al. (2020), *WCD*



## Morocco Crop yield

Lehmann et al. (2020), *GRL*



# CONCLUSIONS

THANK YOU!

