CAUSALITY II

Marlene Kretschmer

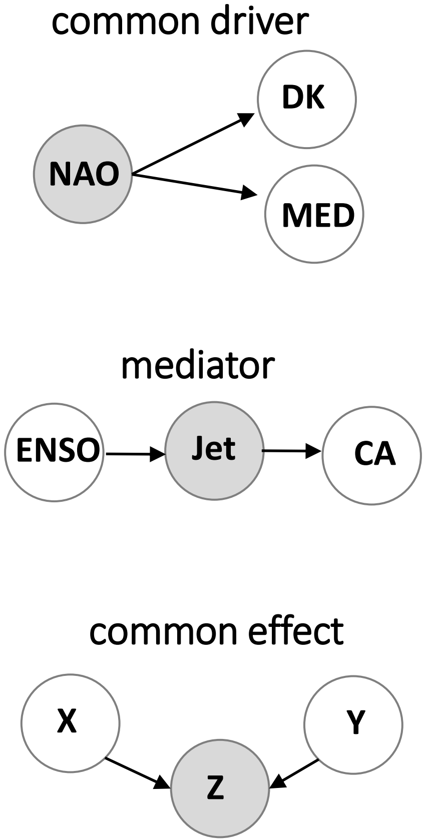
University of Reading

m.j.a.kretschmer@reading.ac.uk

@Marlene_Climate

RECAP

- Scientific data analysis requires causal reasoning
- Causal knowledge/hypotheses are best expressed using causal networks
- To extract causal effects from data, one needs to control for all confounding factors
- Causal inference gives the formal rules how to achieve this



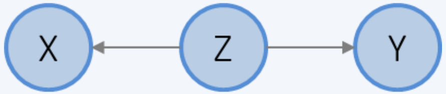
In this Lecture

1) Conditioning on a common effect
2) Controlling for confounders
3) An example from climate science
4) Non-linear dependencies
5) Conclusions

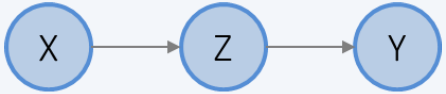
1) Conditioning on a common effect

BASIC CAUSAL STRUCTURES AND THEIR IMPLICATIONS

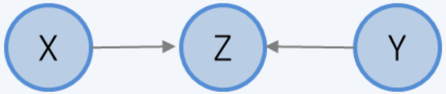
Z is a <u>common driver</u> of X and Y



Z is a <u>mediator</u> of X and Y



Z is a <u>common effect</u> of X and Y



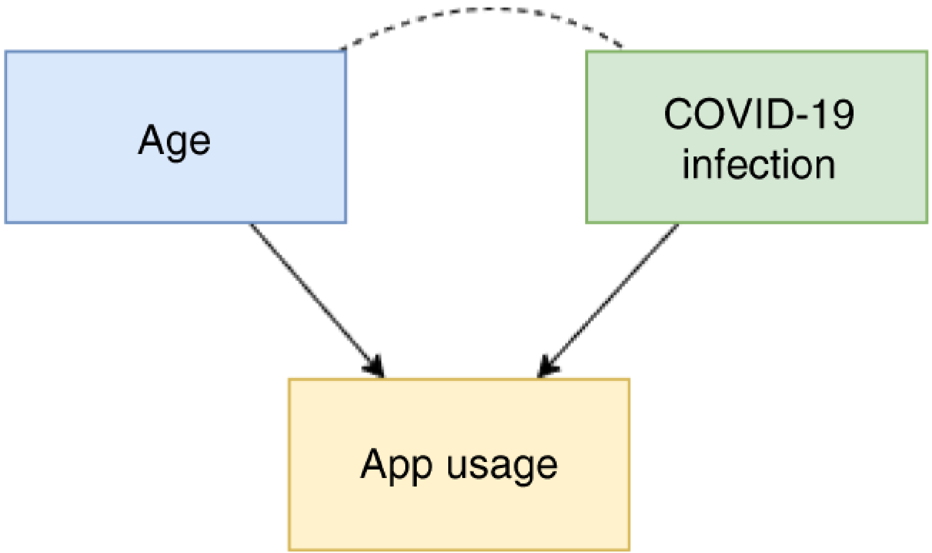
<u>Implication for data</u>

X and Y correlated
X and Y are uncorrelated conditional on Z

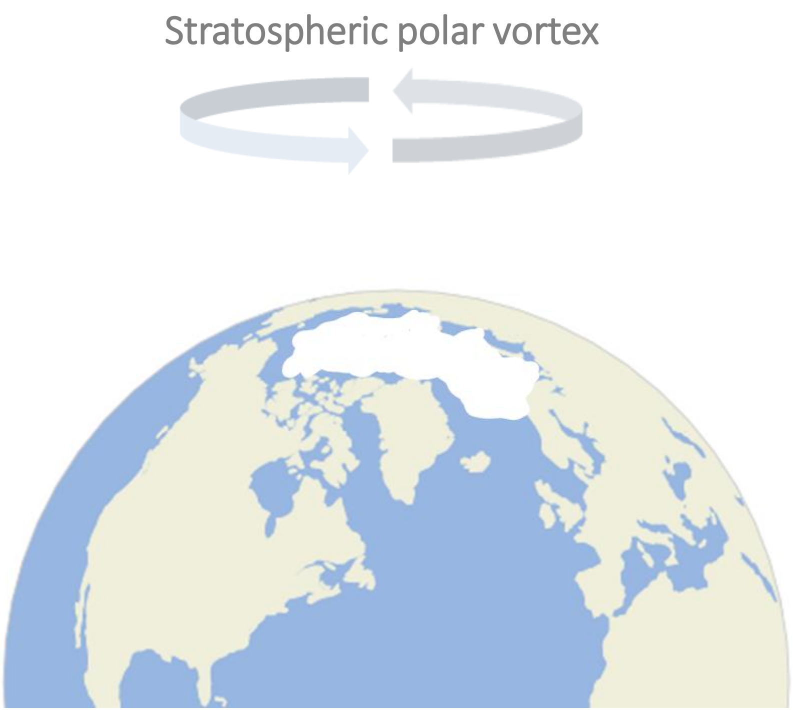
<u>Implication for data</u>

X and Y *un*correlated
X and Y are *correlated* conditional on Z

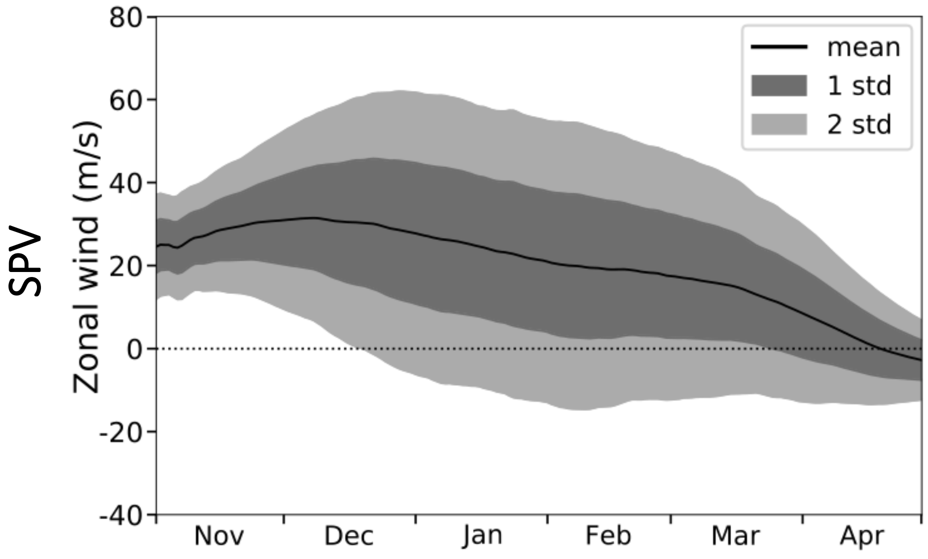
CREATING INTUITION



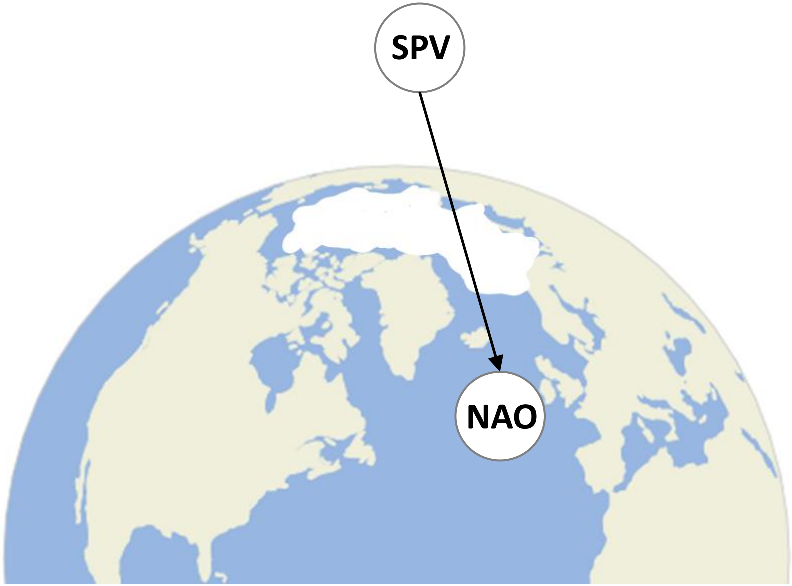
THE STRATOSPHERIC POLAR VORTEX (SPV)

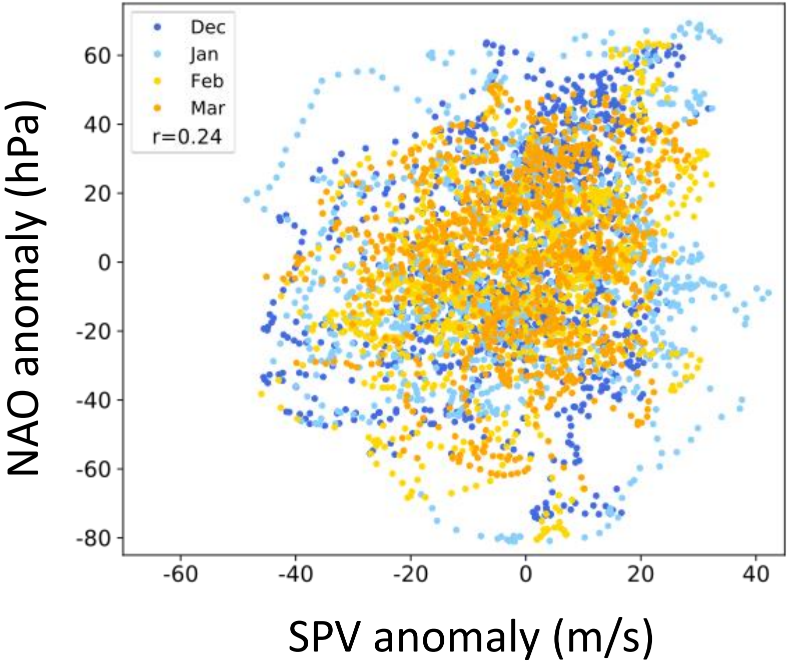


Data from the seasonal forecasting model SEAS5



The Stratospheric Polar vortex

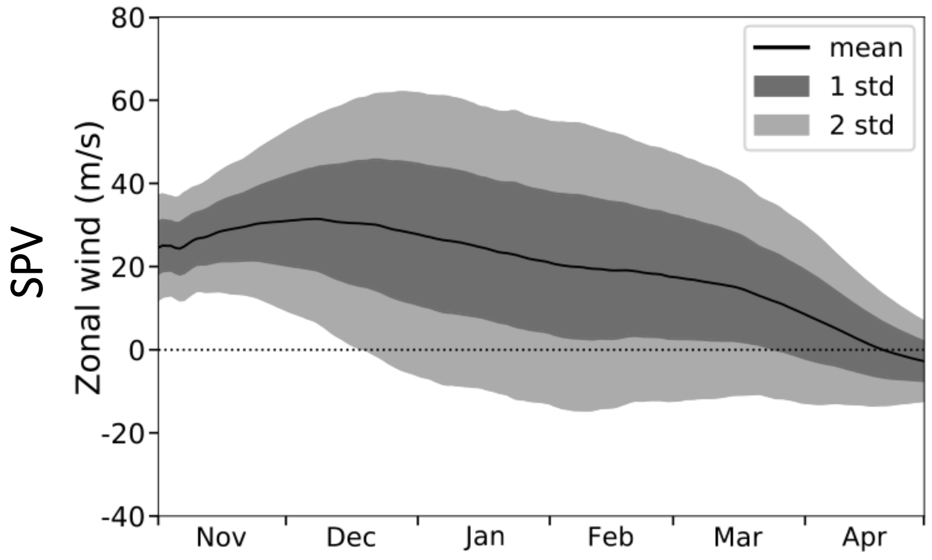


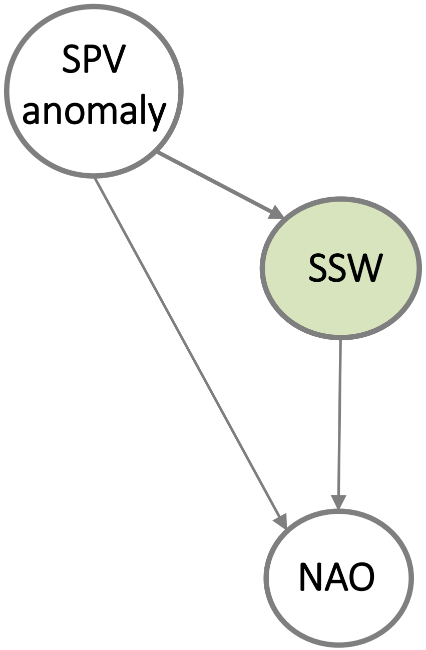


Monnin, Kretschmer, Polichtchouk , *Int. J. Clim* (2021)

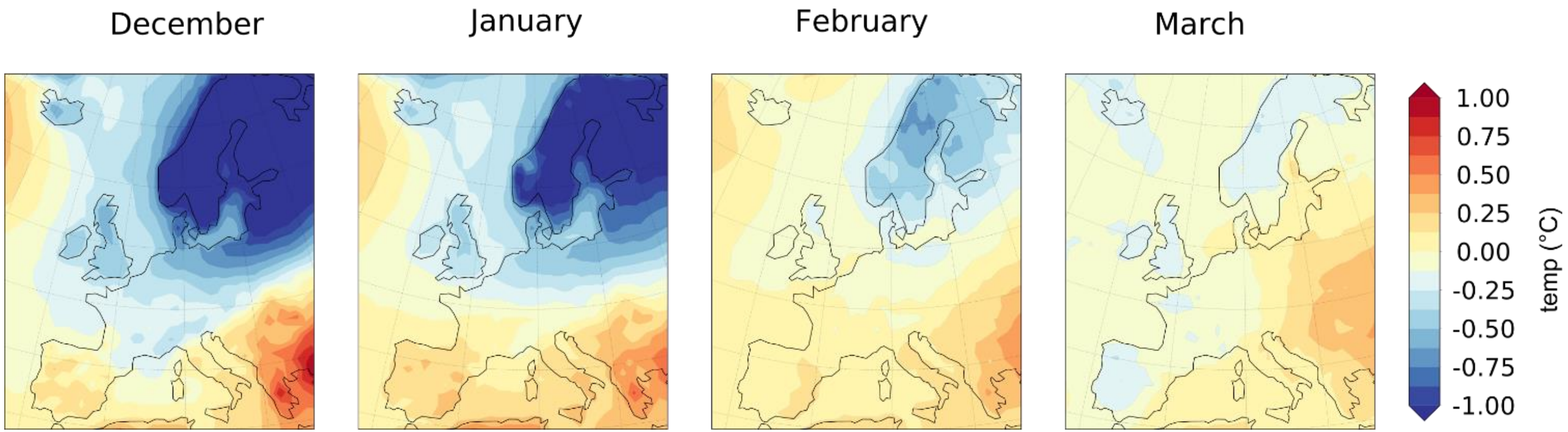
SUDDEN STRATOSPHERIC WARMINGS (SSWs)

SSWs := Days when the winds turn negative





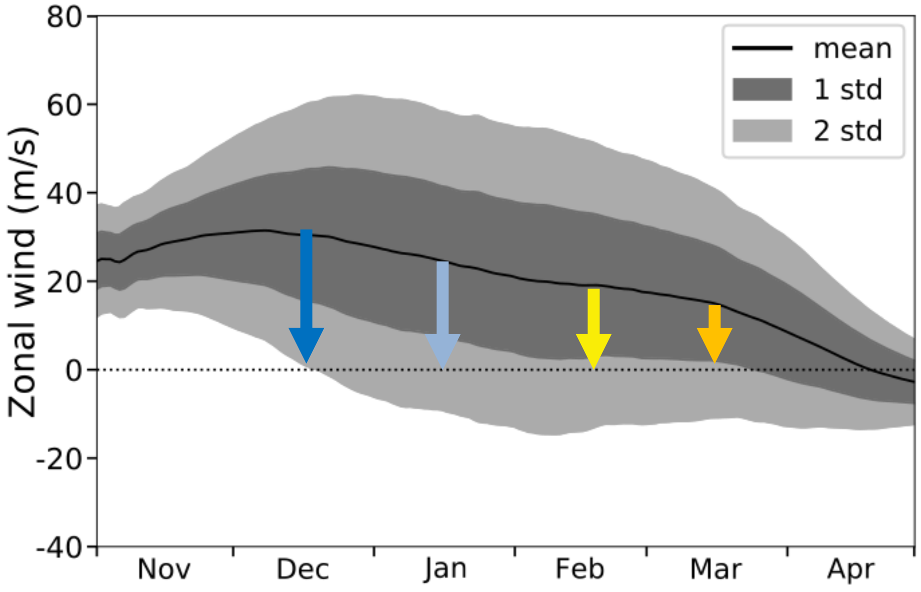
Early winter SSWs show stronger surface impacts



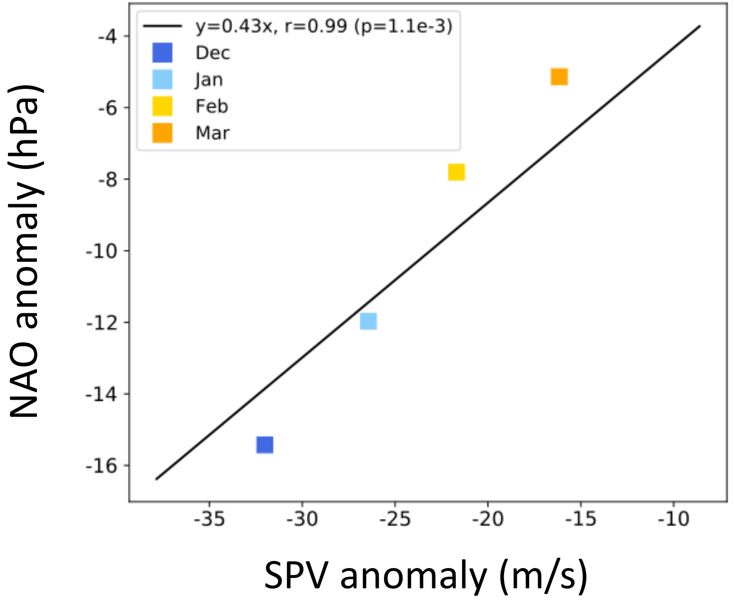
Monnin, Kretschmer, Polichtchouk , *Int. J. Clim* (2021)

EARLY WINTER SSWS HAVE STRONGER WIND ANOMALIES...

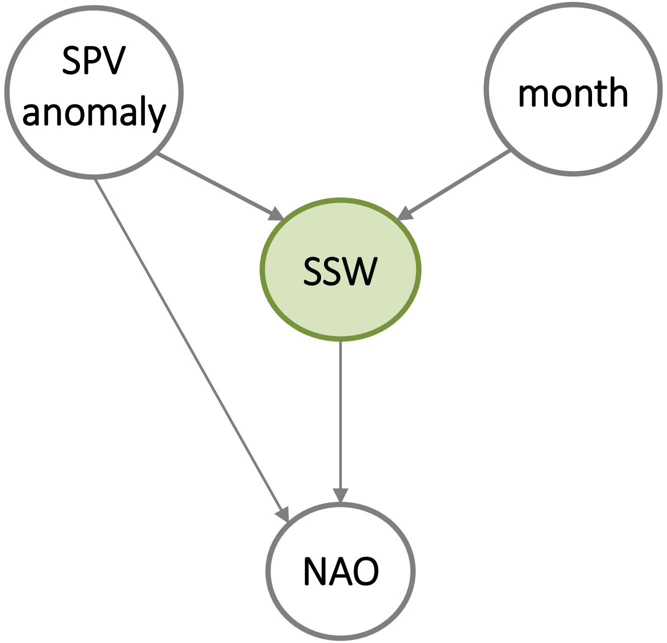
...due to the SSW definition



This explains the stronger impacts

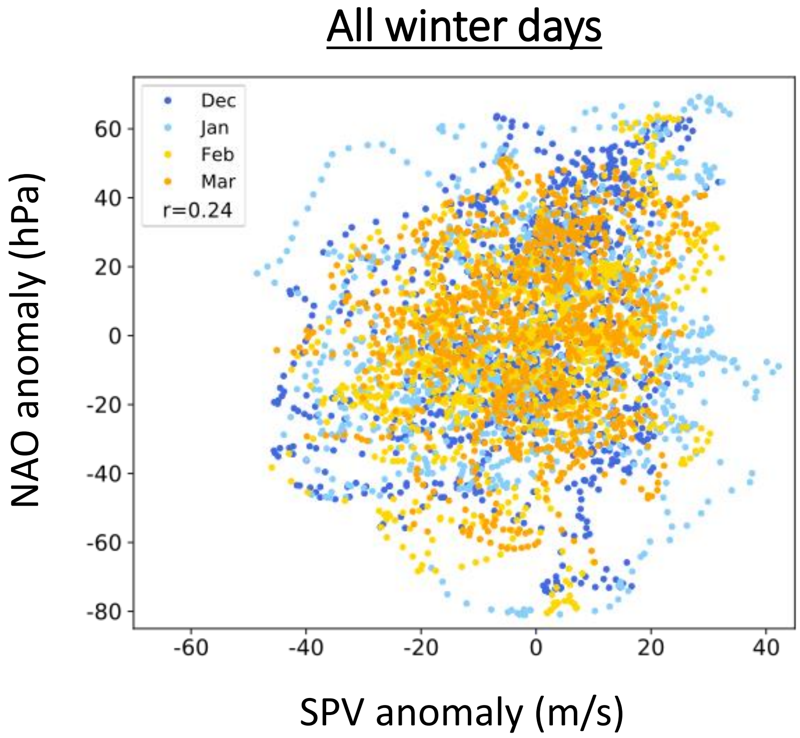


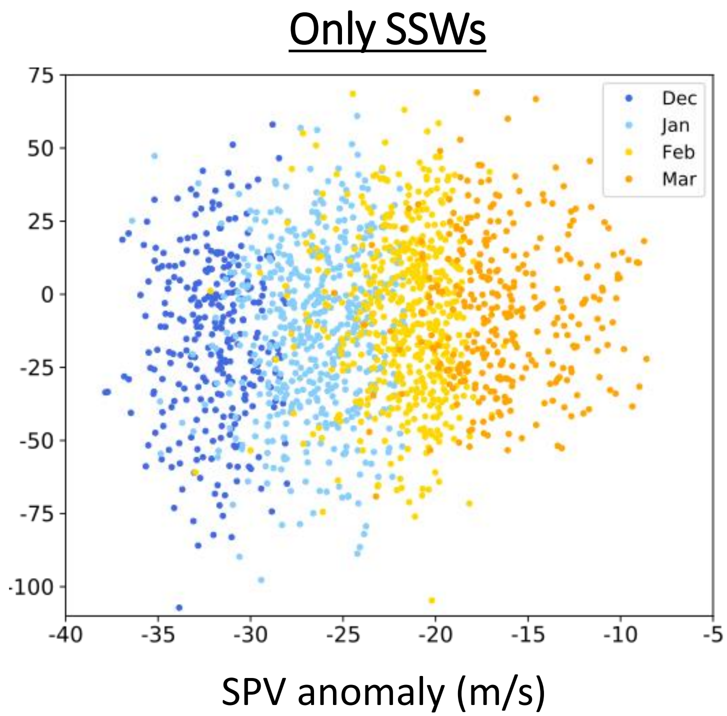
SSW IS A COMMON EFFECT



By conditioning on the common effect “SSW”, we introduce a non-causal association between “SPV anomaly” and “month”

CONDITIONING ON A COMMON EFFECT





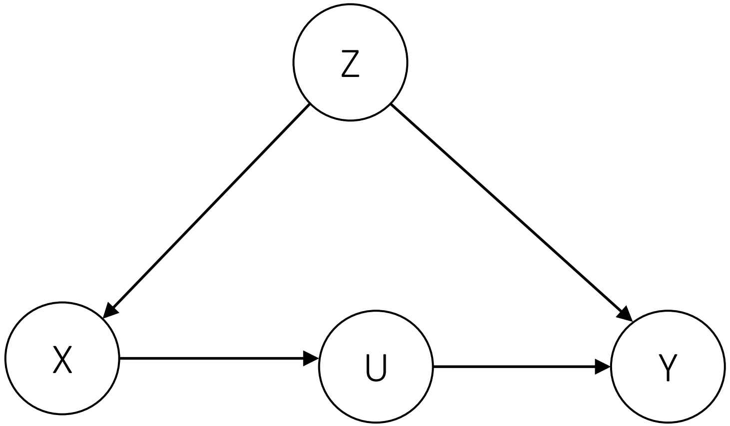
Monnin, Kretschmer, Polichtchouk , *Int. J. Clim* (2021)

2) Controlling for confounders

(in a Nutshell)

CAUSAL NETWORKS

- A causal network consists of **nodes** (representingng variables, e.g. ENSO) and **links** (indicating the direction of causality)
- causal network = directed acyclic graph (DAG)
- a sequence of links "connecting" two nodes in the network is called a **path** (regardless of the direction of arrows!)



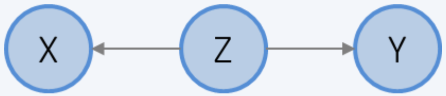
Paths from X to Y:

X --> U --> Y

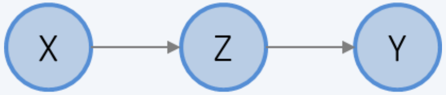
X <-- Z --> Y

BASIC CAUSAL STRUCTURES AND THEIR IMPLICATIONS

Z is a common driver of X and Y



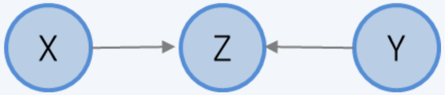
Z is a mediator of X and Y



The **path** from X to Y is **open**
X and Y are **dependent**

The **path** from X to Y is **blocked by conditioning on Z**
X and Y are **independent conditional on Z**

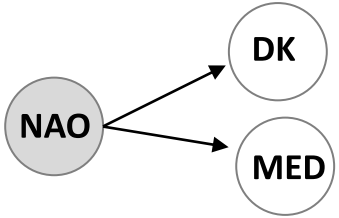
Z is a common effect of X and Y



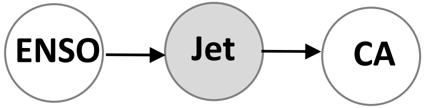
The **path** from X to Y is **blocked by Z**
X and Y are **independent**

The **path** from X to Y is **opened by conditioning on Z**
X and Y are **dependent conditional on Z**

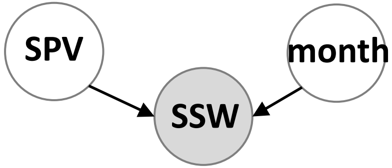
Examples



- There is an open path DK <--NAO -->MED
- conditioning on NAO blocks this path



- There is an open path ENSO --> Jet --> CA
- conditioning on Jet blocks this path



- The path SPV --> SSW <-- month is blocked
- conditioning on SSW opens this path

Rules of do-calculus

Aim: Express P(Y|do(X)) such that it does not contain any "do" expressions

1. Insertion/deletion of observations

$$P(Y \mid do(X), Z, W) = P(Y \mid do(X), Z)$$

If W is irrelevant to Y

2. Action/observation exchange

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

If Z blocks all back-door paths from X to Y

3. Insertion/deletion of actions

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

If there is no causal path from X to Y

THE BACKDOOR CRITERION

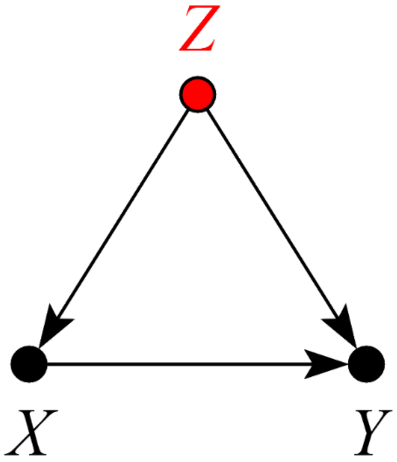
*Confounding is anything that leads to P(Y|X) being different than P(Y|do(X))

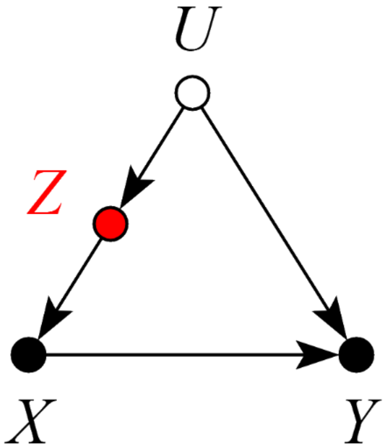
To quantify the causal effect of X on Y, one needs to control for all confounding* factors

To quantify the causal effect of X on Y,

one needs to **block all open paths** between them (other than the one of interest)

EXAMPLES (OF GOOD CONTROLS)

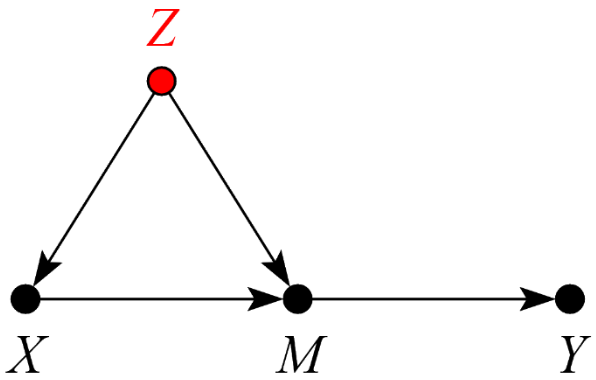


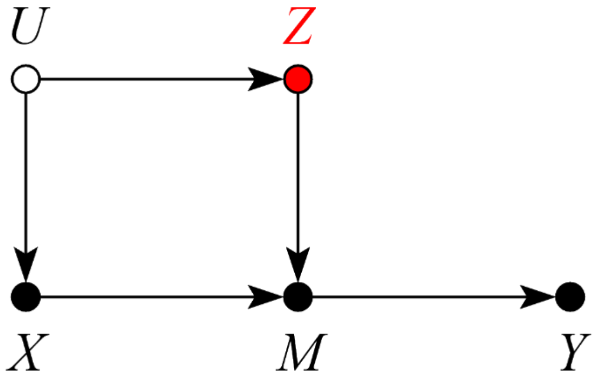


To block the path X <-- Z --> Y

To block the path X <-- Z <-- U -> Y

EXAMPLES (OF GOOD CONTROLS)

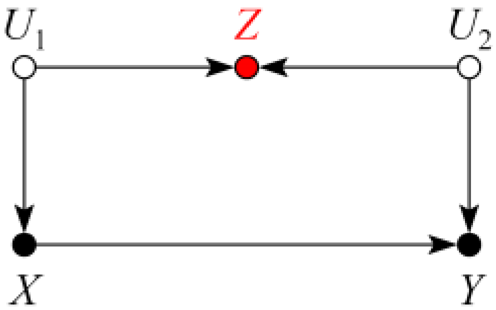




To block the path X <-- Z --> M --> Y

To block the path X <-- U -->Z -->M--> Y

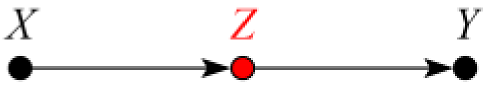
Examples (of Bad Controls)



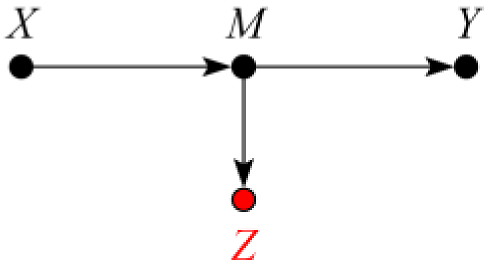
Because this opens the path
X <--U1 --> Z <--U2 --> Y

Source: http://causality.cs.ucla.edu/blog/index.php/category/back-door-criterion/

EXAMPLES (OF BAD CONTROLS)



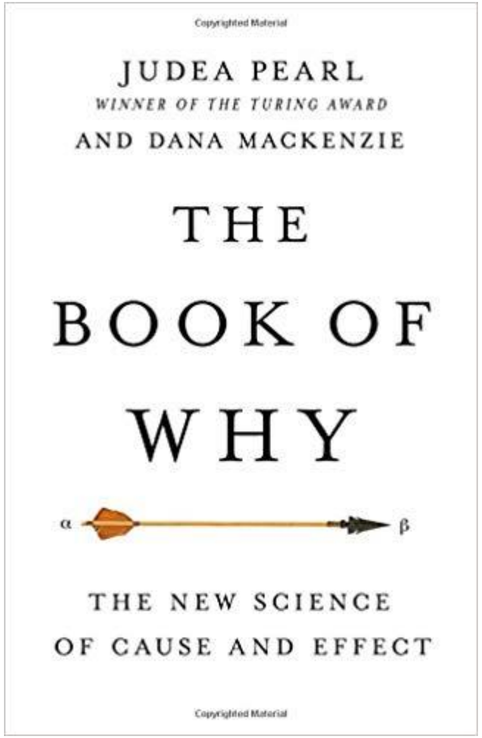
Because this blocks the path X -->Z --> Y

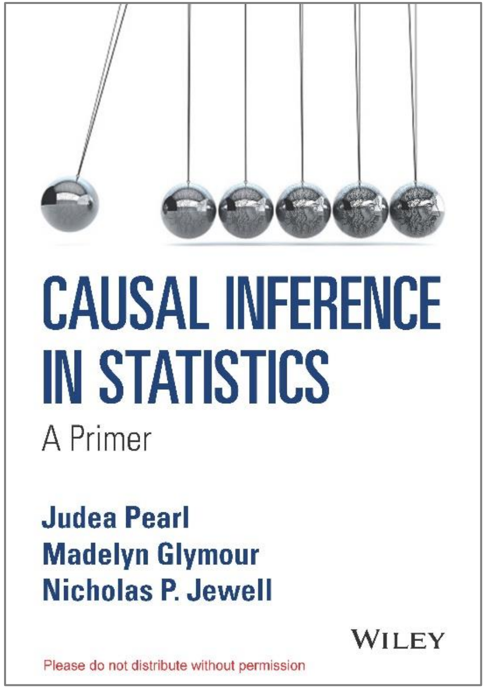


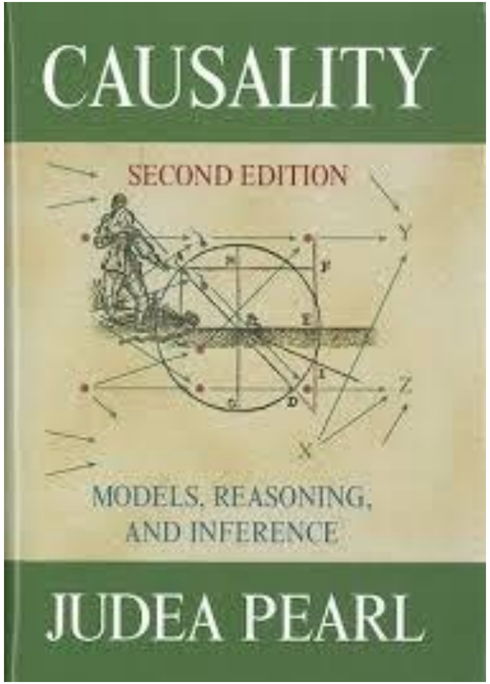
Because this (partially) blocks the path X -->M --> Y (as Z is evidence for M)

Recommendations

+ many tutorials in the internet!

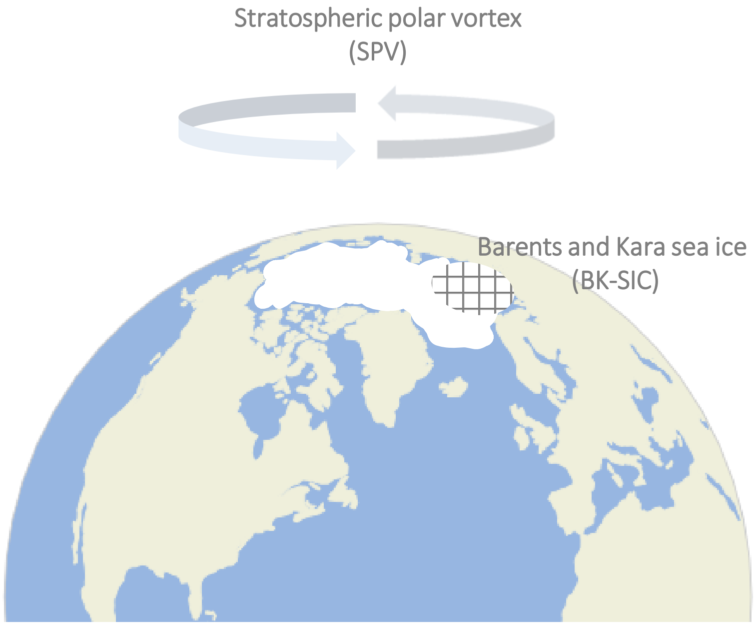






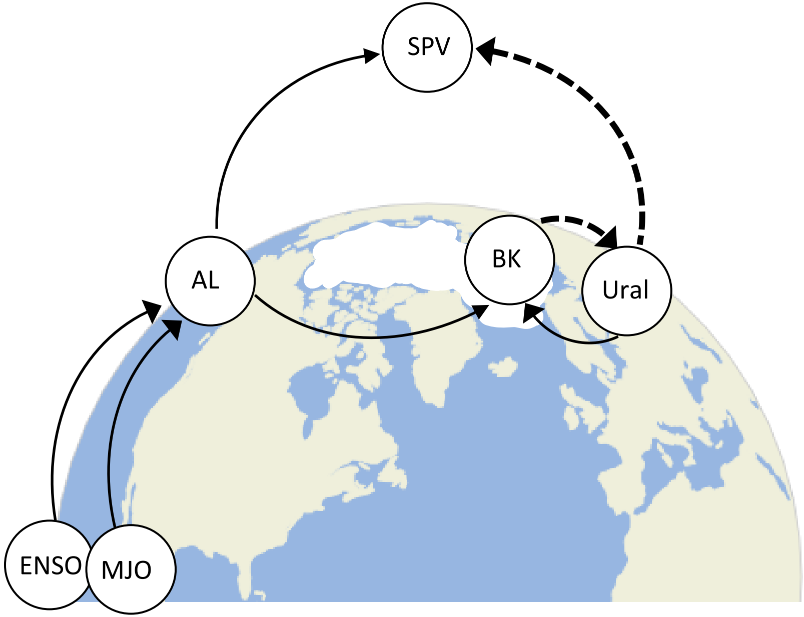
3) An example from climate science

INFLUENCE OF SEA ICE ON THE POLAR VORTEX



How strong is the causal effect of Barents Kara sea ice (BK) in autumn on the winter stratospheric polar vortex (SPV)?

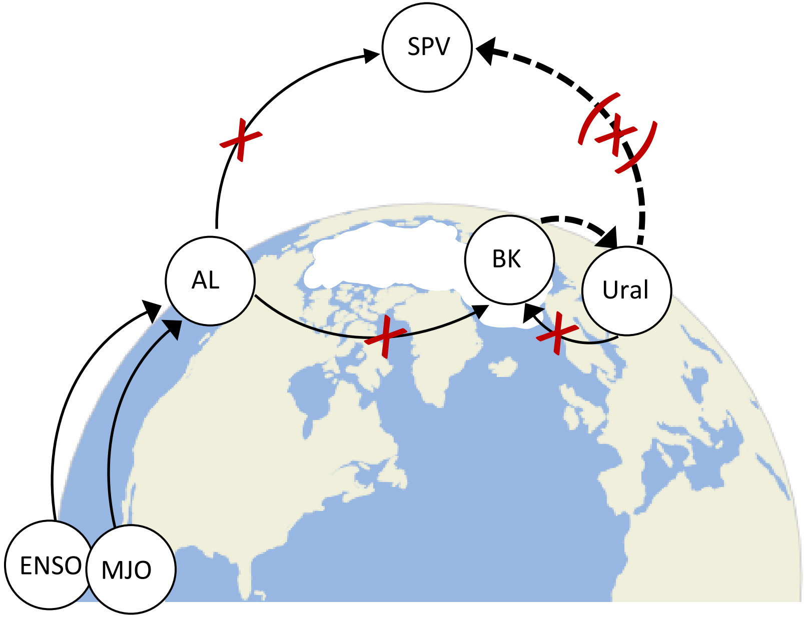
INFLUENCE OF SEA ICE ON THE POLAR VORTEX



How strong is the causal effect of Barents Kara sea ice (BK) in autumn on the winter stratospheric polar vortex (SPV)?

Kretschmer et al. *WCD* (2020), Kretschmer et al. *BAMS* (2021)

Controlling for the Back Door Paths



Open Paths from BK to SPV:

BK --> Ural --> SPV URAL (after OND)

BK <-- AL --> SPV

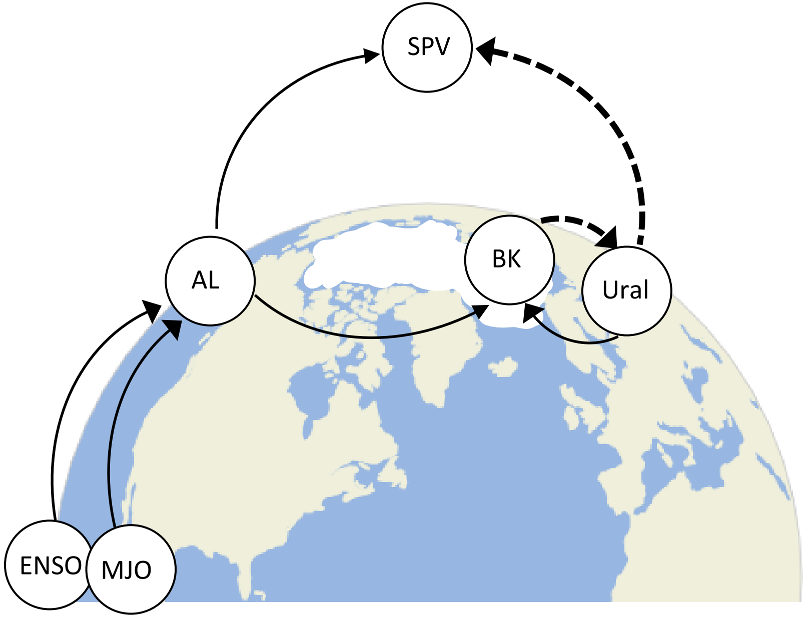
BK <-- Ural --> SPV $URAL_{OND}$

$SPV_{JFM} = a\ BK_{OND} + \text{confounders}$

$SPV_{JFM} = a\ BK_{OND} + b\ AL_{OND}$

$SPV_{JFM} = a\ BK_{OND} + b\ AL_{OND} + c\ URAL_{OND} + \varepsilon$

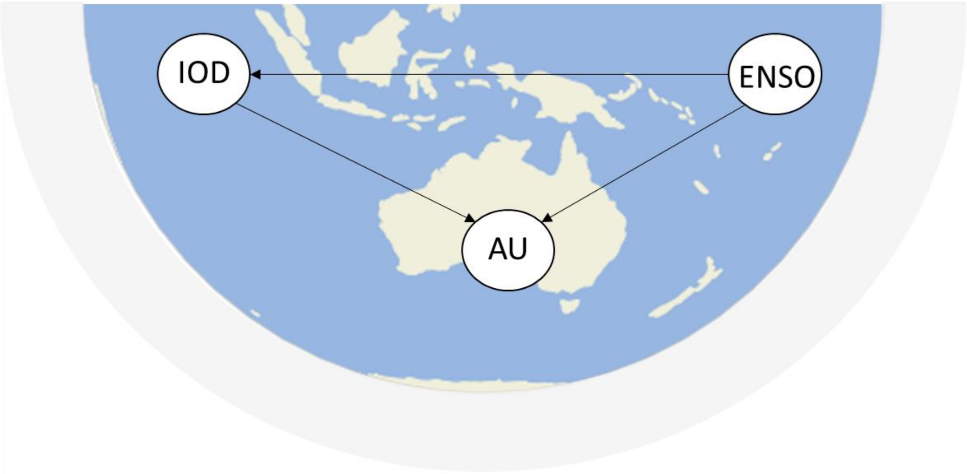
Controlling for the back door Paths



Causal networks make scientific assumptions transparent and help to identify where information is propagating

4) Non-linear dependencies

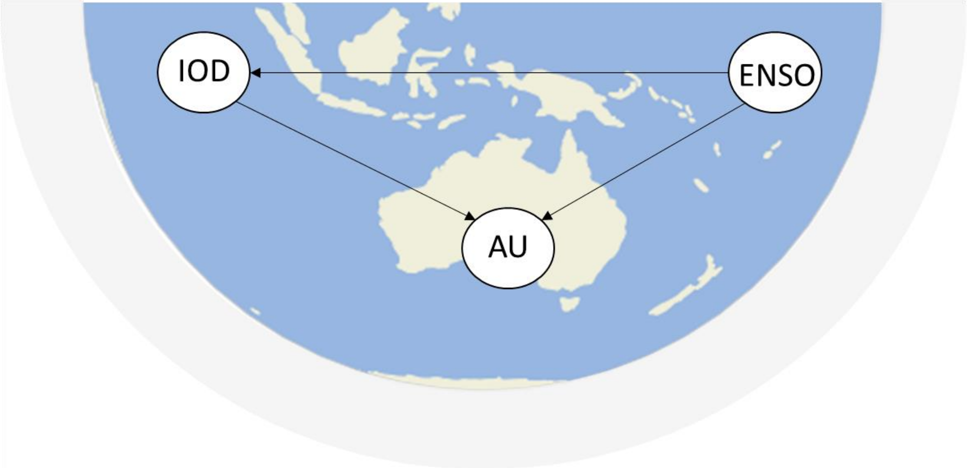
NON-LINEAR DEPENDENCIES



Precipitation in Australia (AU) is affected by ENSO and by the Indian Ocean Dipole (IOD)

The relationships likely involve non-linearities

NON-LINEAR DEPENDENCIES



We stratify the data into different categories

AU: below/above average

IOD: negative/neutral/positive phase

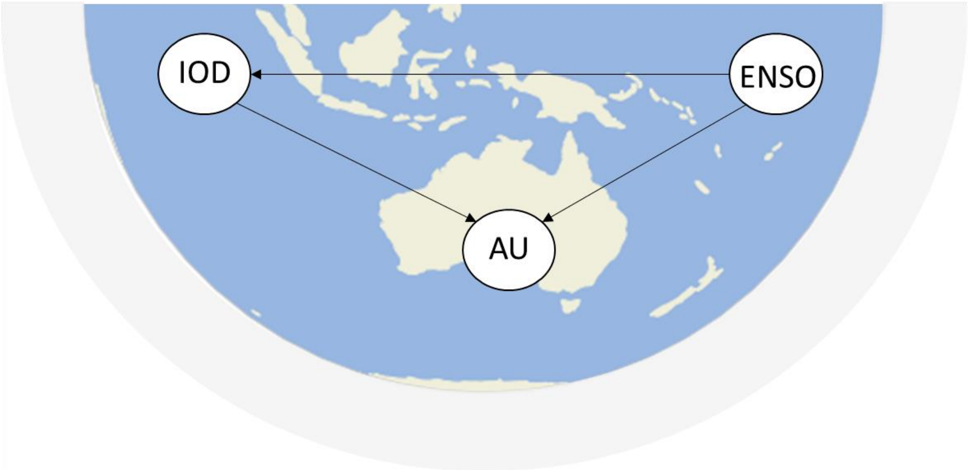
ENSO: La Niña/neutral/El Niño

Conditional probabailities for above average AU (default = ½)

	La Niña	Neutral	El Niño	Marginal
IOD -	0.83	0.50	-	0.67
Neutral	0.80	0.43	0.17	0.52
IOD +	1.0	0.25	0.24	0.30
Marginal	0.83	0.43	0.22	0.50

Kretschmer et al., *BAMS* (2021)

NON-LINEAR DEPENDENCIES



Above average precipitation is unlikely during El Niño

P(AU+ | El Niño) = 0.22

Conditional probabailities for above average AU (default = ½)

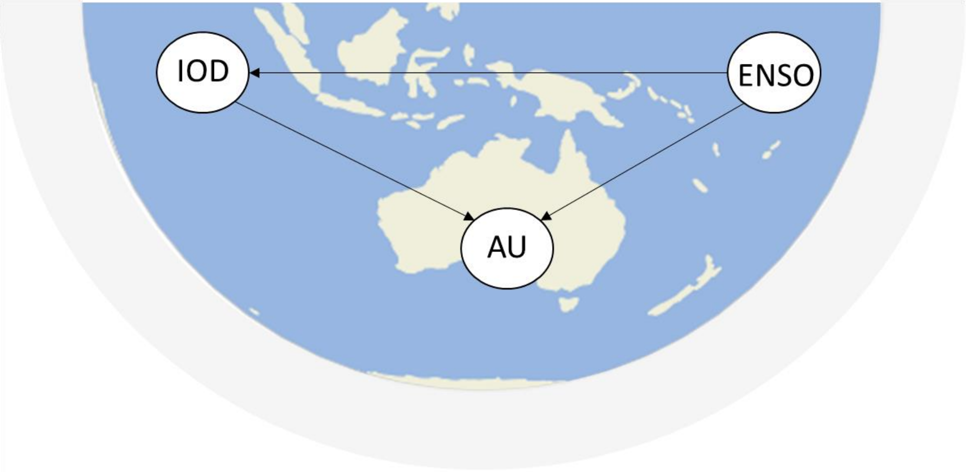
	La Niña	Neutral	El Niño	Marginal
IOD -	0.83	0.50	-	0.67
Neutral	0.80	0.43	0.17	0.52
IOD +	1.0	0.25	0.24	0.30
Marginal	0.83	0.43	0.22	0.50

Above average precipitation is likely during IOD-

P(AU+ | IOD-) = 0.67

Kretschmer et al., *BAMS* (2021)

NON-LINEAR DEPENDENCIES



Conditional probabailities for above average AU

	La Niña	Neutral	El Niño	Marginal
IOD -	0.83	0.50	-	0.67
Neutral	0.80	0.43	0.17	0.52
IOD +	1.0	0.25	0.24	0.30
Marginal	0.83	0.43	0.22	0.50

What is the added information provided by IOD, given ENSO?

P(AU+ | IOD+, El Niño) = 0.24

P(AU+ | El Niño) = 0.22

0.24/0.22 = 1.09

Interpretation of data depends on causal assumptions!

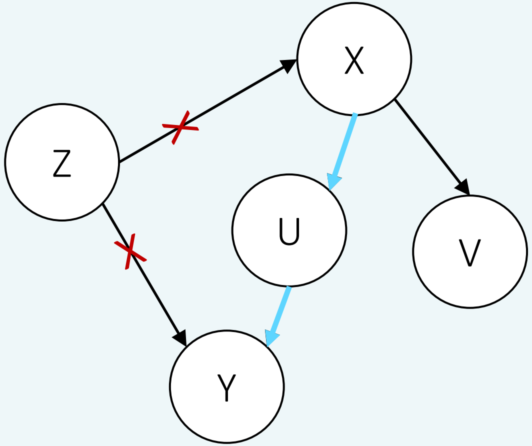
Kretschmer et al., *BAMS* (2021)

5) Conclusions

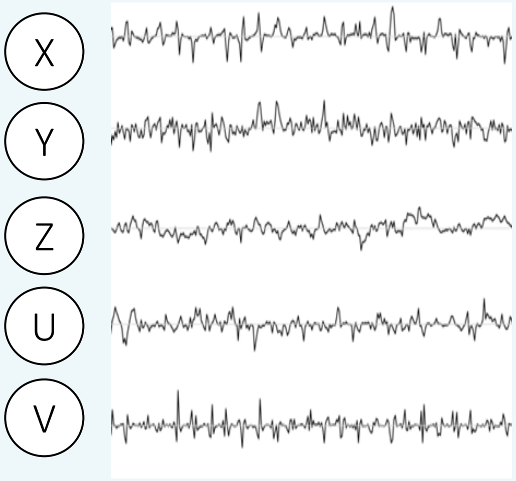
STEPS OF CAUSAL INFERENCE

Question: What is the (average) causal effect of X on Y?

1. Use expert knowledge to set a (plausible) causal model



2. Collect data



3. Control for confounders to isolate the causal effect

$$P(Y \; | do(X))) = P(Y \mid X, Z)$$

Confounding is anything that leads to P(Y|X) being different than P(Y|do(X))

linear case:

$$Y = a\, X + b\, Z$$

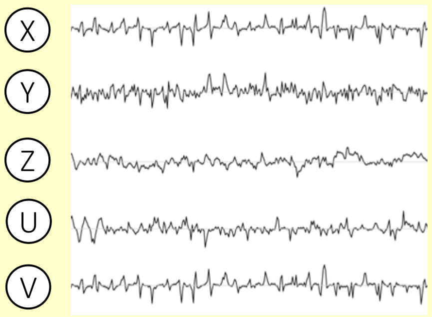
SUMMARY

- Scientific data science is never fully “objective”
- We should be transparent about our assumptions (by using causal networks)

- Causal networks help to identify where information is propagating and to extract the causal effect of interest
- Conditioning on confounders = blocking the “open” paths in the network
- Its fully non-parametric
- Implementing a causal framework only involves small changes in scientific practice but allows to draw stronger, causal statements

OUTLOOK: *LEARNING* CAUSAL STRUCTURES FROM DATA

Input: Time-series



Causal Discovery

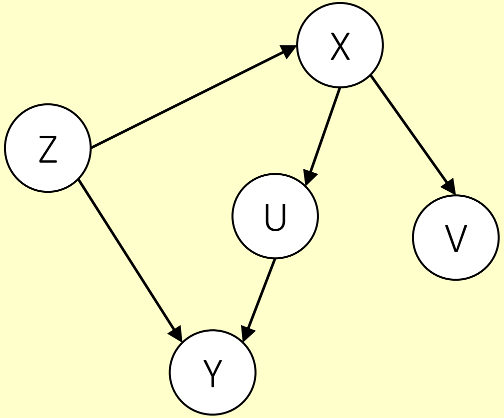
PCMCI Algorithm

corr($X_{t-\tau}$, Y_t | *Iterate through combinations of conditions*)

Identifies spurious correlations due to

- common drivers
- mediators
- auto-correlation effects

Output: causal model/network

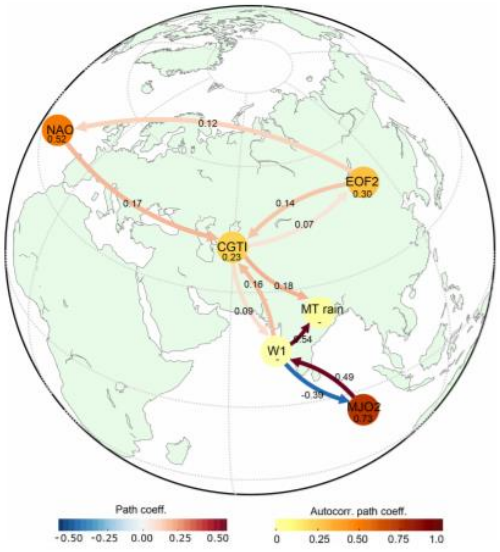


Kretschmer et *al.* (2016, 2018, 2019), Runge et *al.* (2019a, 2019b), Saggioro et *al.* (2020)

APPLICATION REQUIRES PROCESS UNDERSTANDING

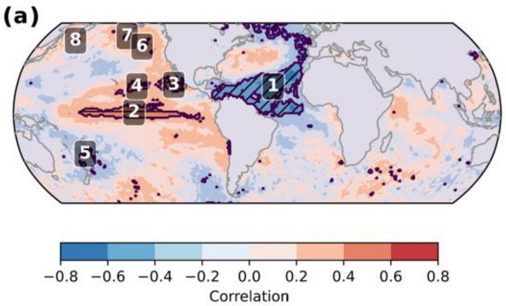
Indian Summer Monsoon

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



Hurricane Activity

Pfleiderer et al. (2020), *WCD*

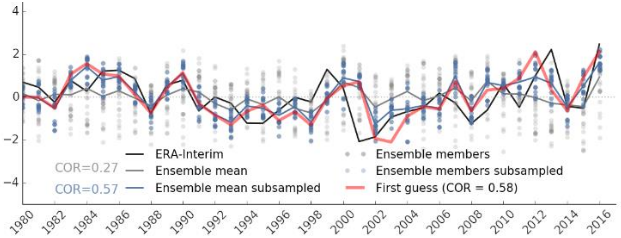


Morocco Crop yield

Lehmann et al. (2020), *GRL*

Marine cold air outbreaks

Polkova et al. (2021), *QJRMS*



SPECIAL COLLECTION AT ATMOSPHERIC SCIENCE LETTERS

<u>Guest Editors</u>
Marlene Kretschmer,
Aglaé Jezequel
Zach Labe
Danielle Touma

THANK YOU FOR YOUR ATTENTION!