# BRIDGING PHYSICAL HYPOTHESES AND THEIR STATISTICAL ANALYSIS THROUGH CAUSAL NETWORKS

Marlene Kretschmer, presented by Ted Shepherd

Department of Meteorology, University of Reading

This project has received funding from
the European Union’s Horizon 2020
research and innovation programme
under grant agreement 101003469

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

@Marlene_Climate

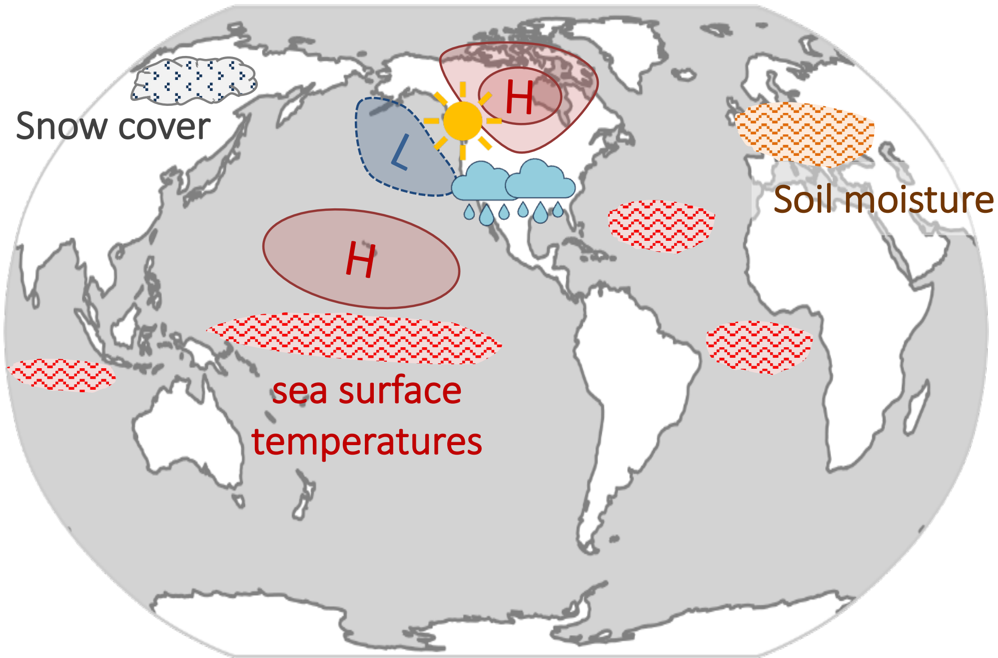
We need a ***causal understanding*** of the world, both for decision-making and for many forms of theory and research.

*What is the effect on global mean temperature if GHG emissions are increasing?*

*Will climate change lead to more intense extreme rainfall events in the UK?*

*Is El Niño increasing the chance of drought in South Africa?*

*What is the effect of melting Arctic sea ice on European climate?*



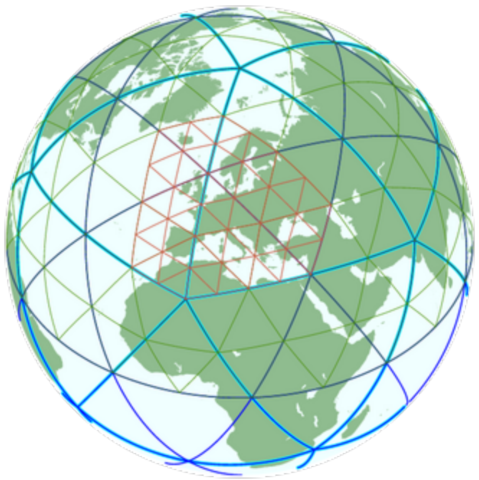

- Quantifying the causal contribution of teleconnections is key to improve our understanding of regional weather and climate variability (including attribution tasks!)
- Extracting this information from data is usually difficult!

## American Meteorological Society: *"Teleconnection"*

A significant [...] correlation in [...] widely separated points.

[...] such correlations suggest that information is propagating [...].

## Physics-based methods



As real-world experiments are usually not possible, numerical climate models are used to infer causal relationships of the climate system

Inferences about the real world depend on the realism of the climate model

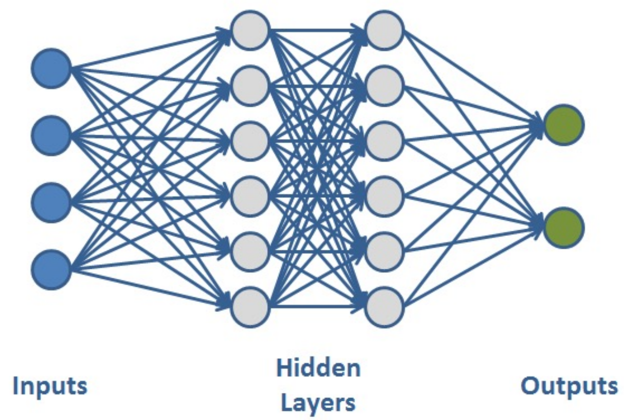
In fact, we need causality to understand how deficiencies in the model affect the outcomes

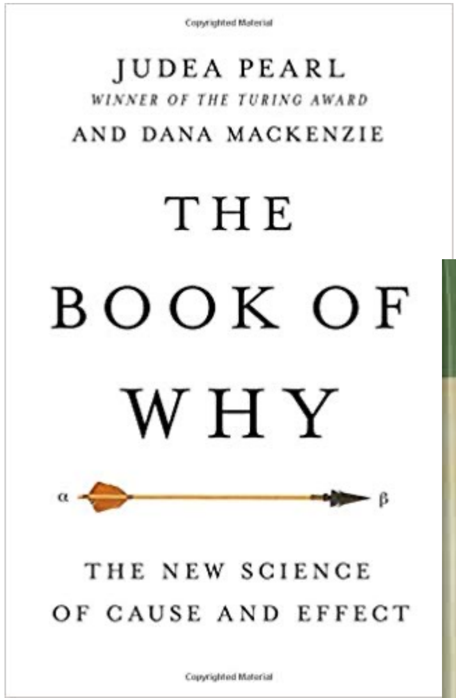
Statistics/data science are needed to study observational data

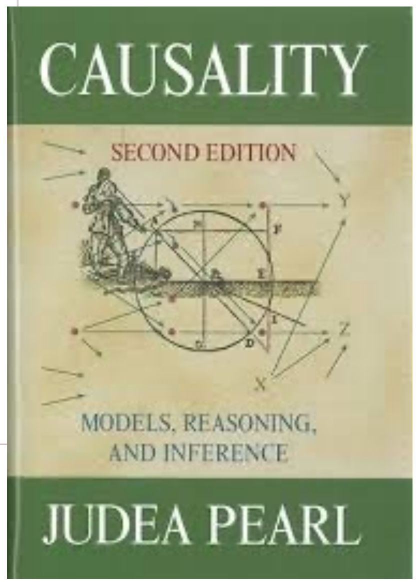
However, we are usually limited to detect statistical associations (e.g. correlations) but *correlation does not imply causation*

How can we infer ***causal*** relationships from data?

## Data-driven methods






- The concept of *causality* has long been missing in mathematics

- Causal inference: the science to extract causal information from data

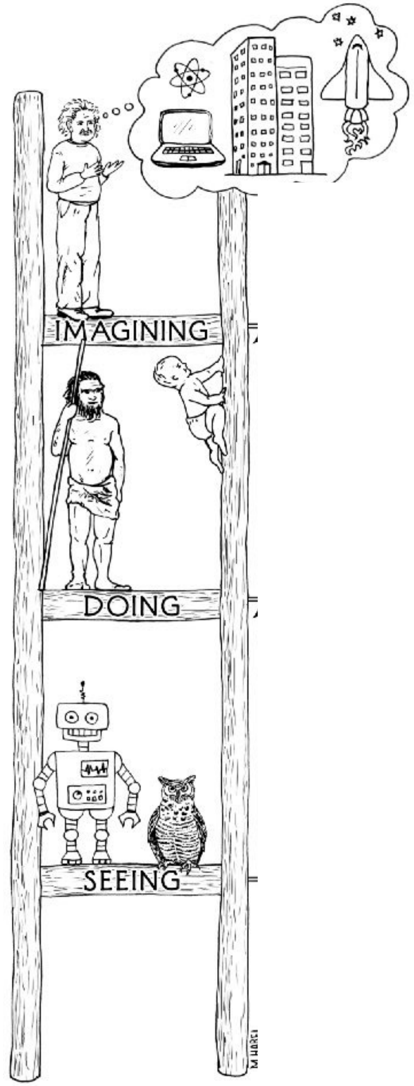
   1. learning causal relationships
   2. **quantifying causal relationships**

In this talk

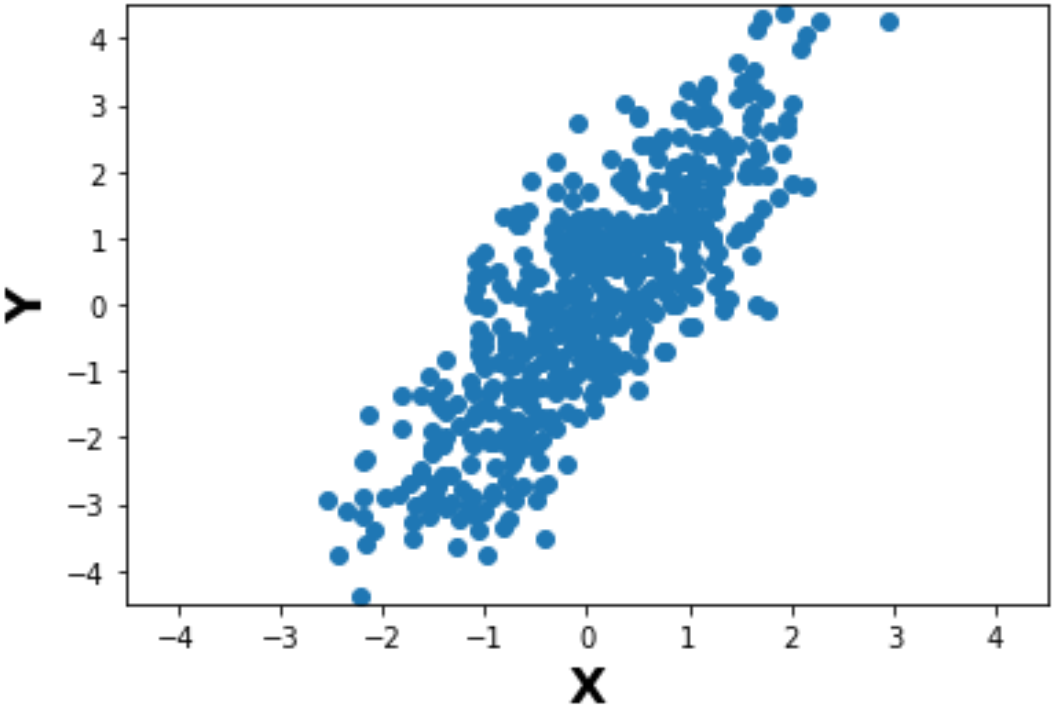
Counterfactuals: $P(y_x \mid x', y')$

Intervention: $P(y \mid do(x))$

Association: $P(y \mid x)$




Source: The book of why

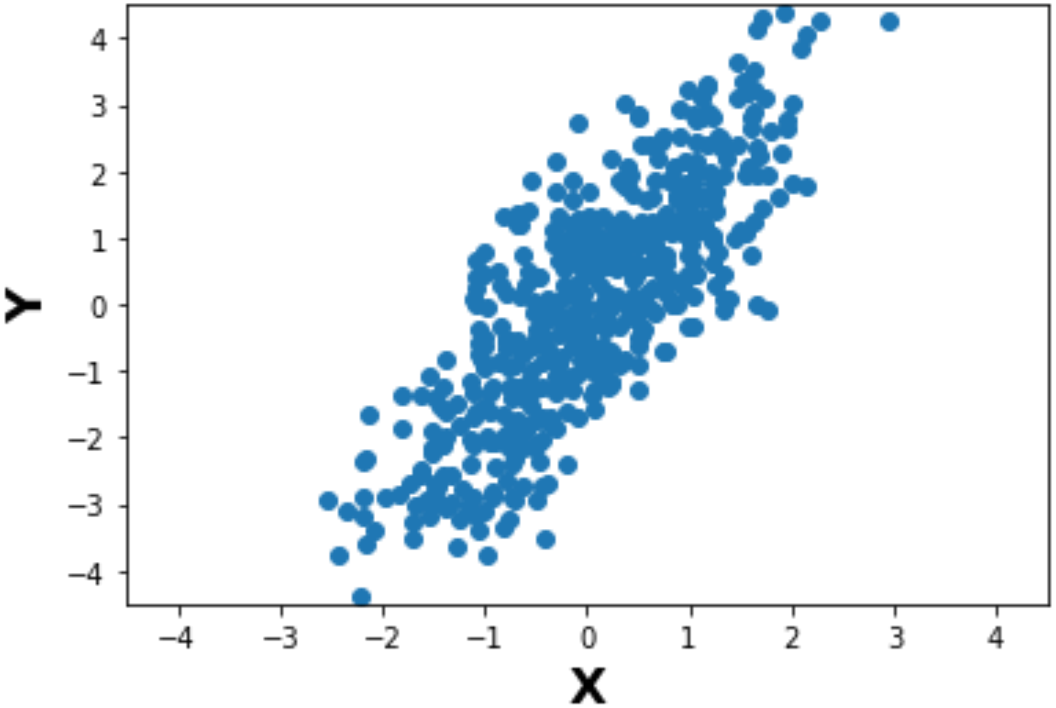



X causes Y?

Y causes X?

A common driver Z affects X and Y?

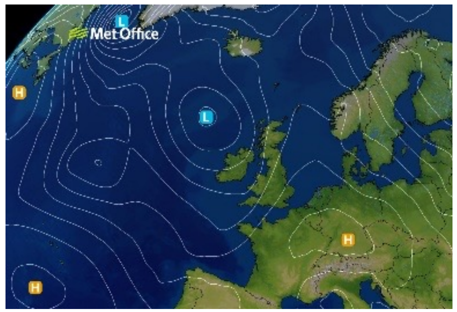
Data Doesn't Speak for Itself!

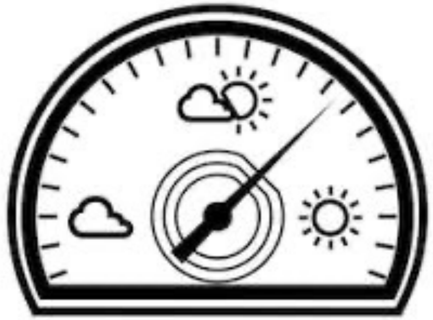



## Example

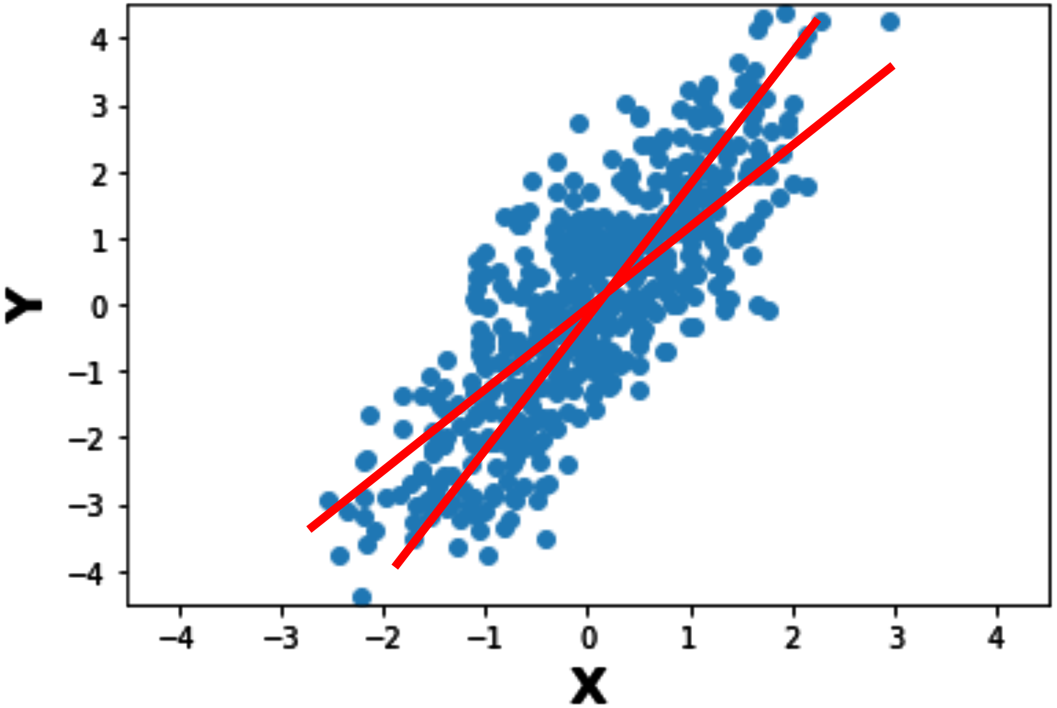
X: Pressure

Y: Barometer



Intervening in Y will *not* change X
Intervening in X *will* change Y

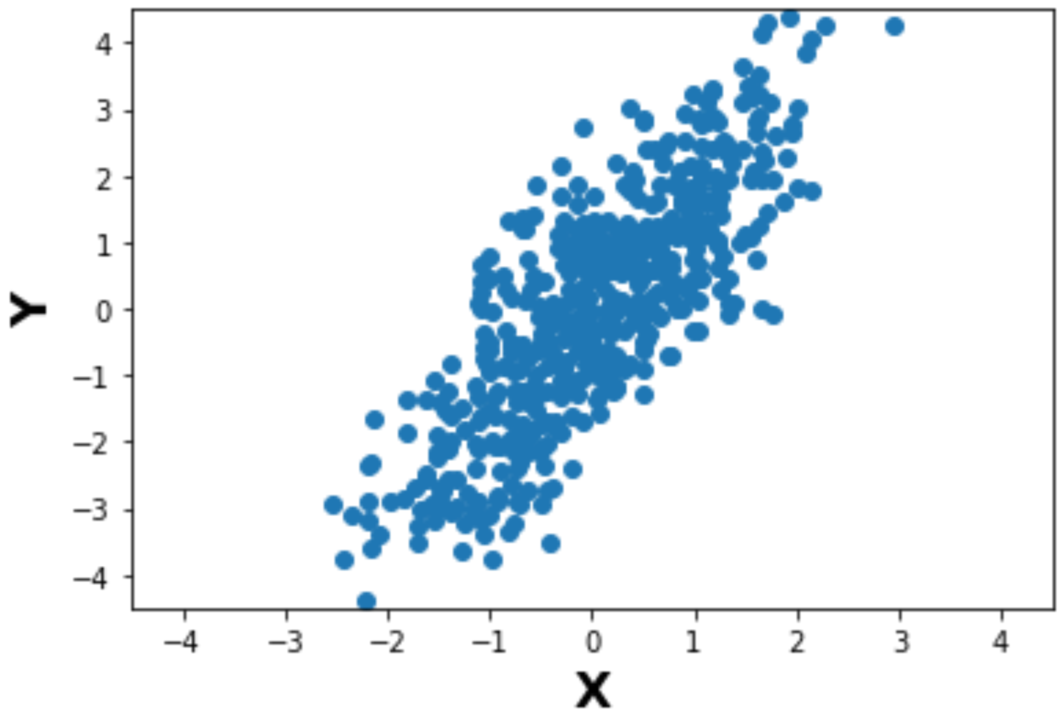



Regression involves a causal assumption, and breaks the mathematical symmetry in the data

The regression line of Y against X is shallower than the line of best fit

The regression line of X against Y is also shallower, when the axes are flipped

The two regression slopes are not reciprocals of each other

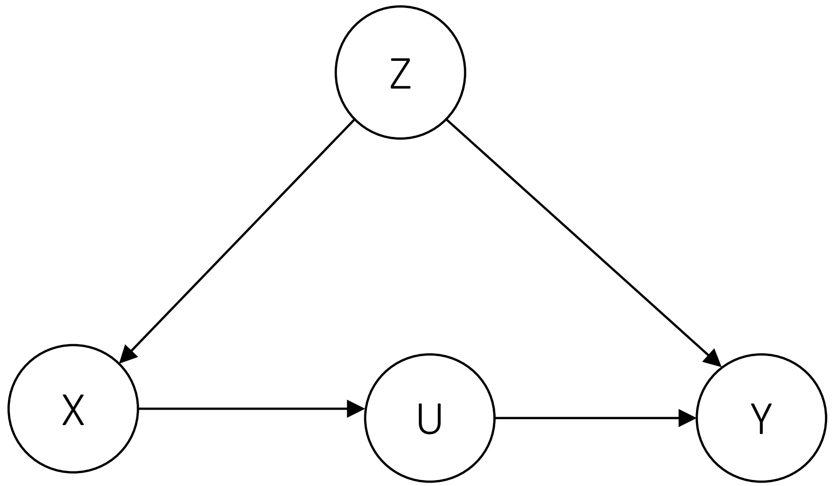



To make sense of the data, we need **<u>causal knowledge</u>** about the data-generating mechanisms

- A causal network consists of **nodes** (representing 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 the arrows!)




Paths from X to Y:

X --> U --> Y

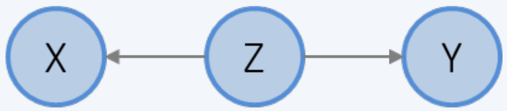
X <-- Z --> Y

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

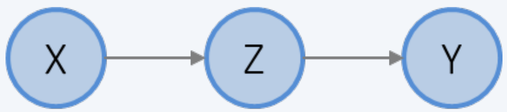
To quantify the causal effect of X on Y,

**all open** paths between them (other than the one of interest) **have to be blocked**

Z is a **common driver** of X and Y




Z is a **mediator** of X and Y



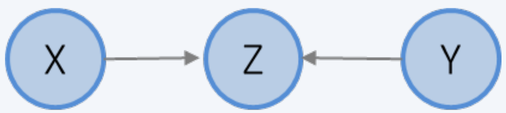

The **path** from X to Y is **open**

The **path** from X to Y is **blocked by conditioning on Z**

X and Y are **dependent**

X and Y are **independent conditional on Z**

Z is a **common effect** of X and Y



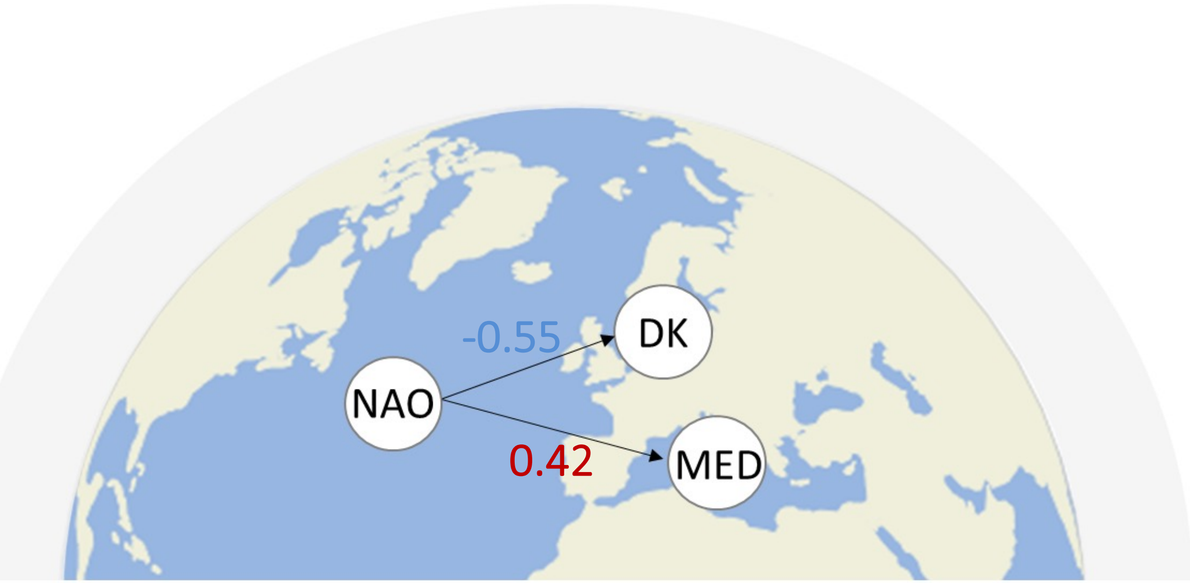

The **path** from X to Y is **blocked by Z**

The **path** from X to Y is **opened by conditioning on Z**

X and Y are **independent**

X and Y are **dependent conditional on Z**

# Common driver




(JJA mean, NCEP)

Summer precipitation in Denmark and the Mediterranean is significantly correlated

Corr(DK, MED) = -0.25

But independent conditional on NAO

→ Corr(DK, MED | NAO) = 0.001

DK = -0.55 NAO + ε

MED = 0.42 NAO + ε

The causal effects explain the correlation

-0.25 ≈ -0.55 * 0.42

# Mediator

What is the effect of ENSO on California winter precipitation?

CA = 0.05 ENSO + 0.79 Jet + ε

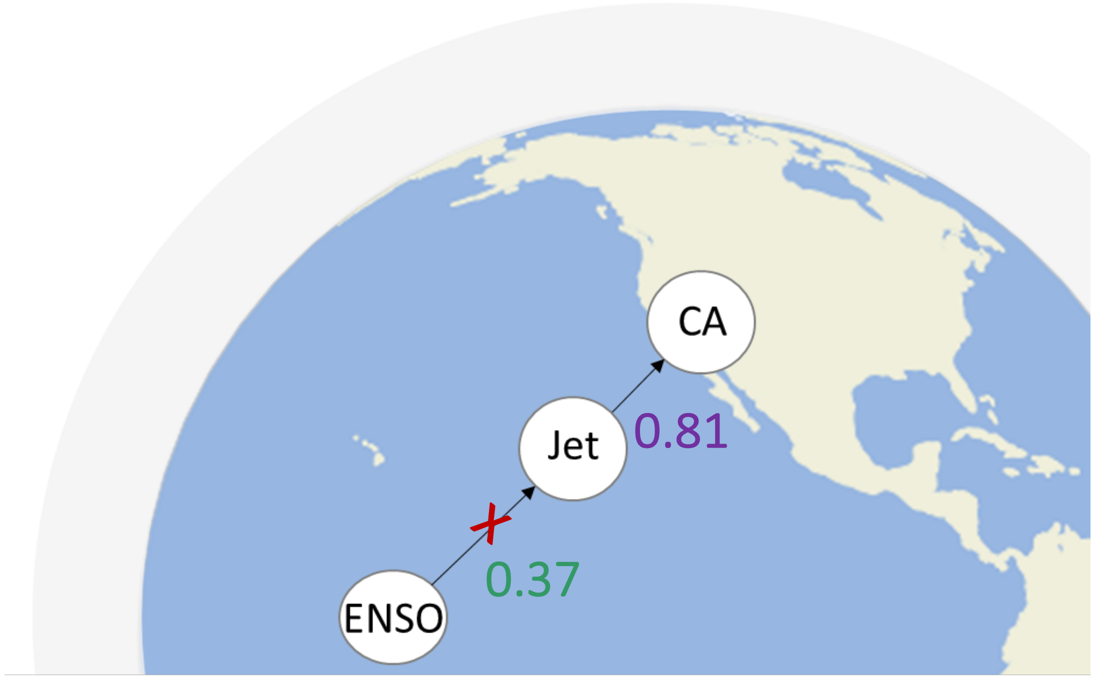
Correct way:

CA = 0.34 ENSO + ε

Or via product along pathway:

Jet = 0.37 ENSO + ε    CA = 0.81 Jet + ε

0.37 * 0.81 = 0.30



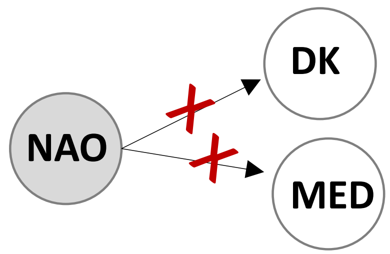

(DJF mean, NCEP)

Statistically, example 1 and 2 are indistinguishable

## X and Y are correlated

example 1: X = DK and Y = MED

example 2: X = ENSO and Y = CA

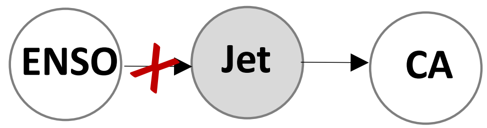



example 1

## X and Y are independent conditional on Z

example 1: Z = NAO

example 2: Z = Jet



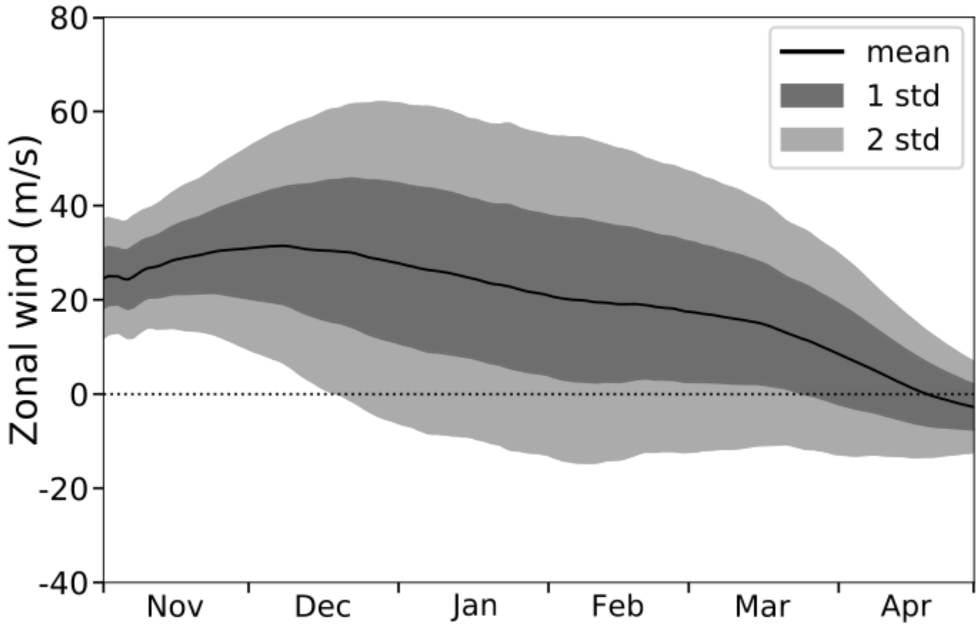

example 2

**The causal interpretation enters through our physical knowledge!**

# Common effect

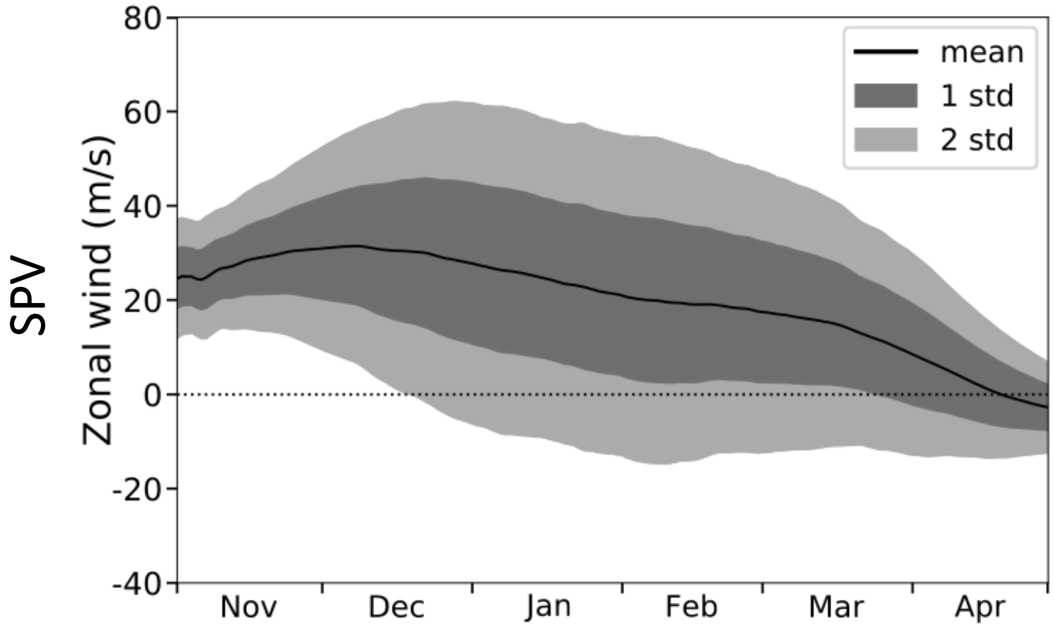
Data from the seasonal forecasting model SEAS5

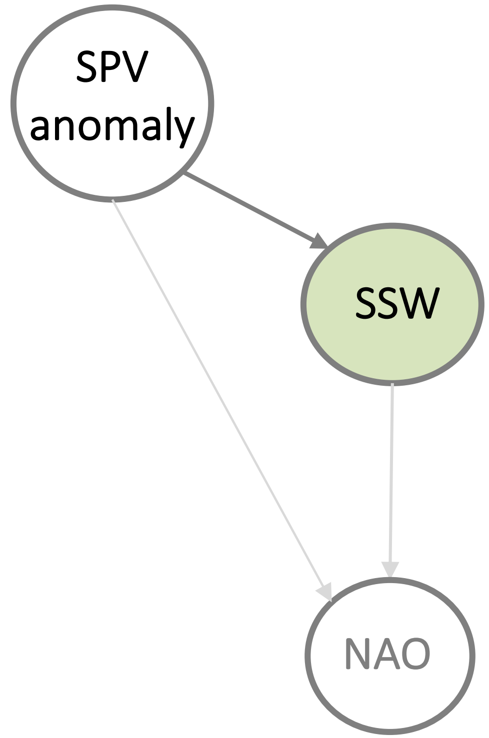
SPV



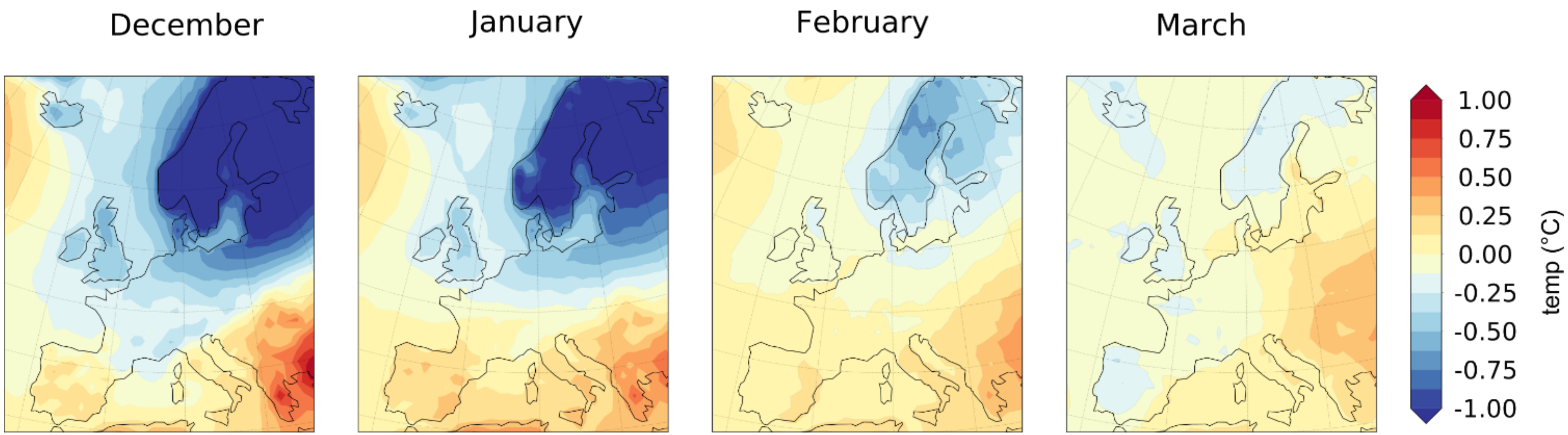

Stratospheric polar vortex (SPV)

SSWs := Days when the winds turn negative

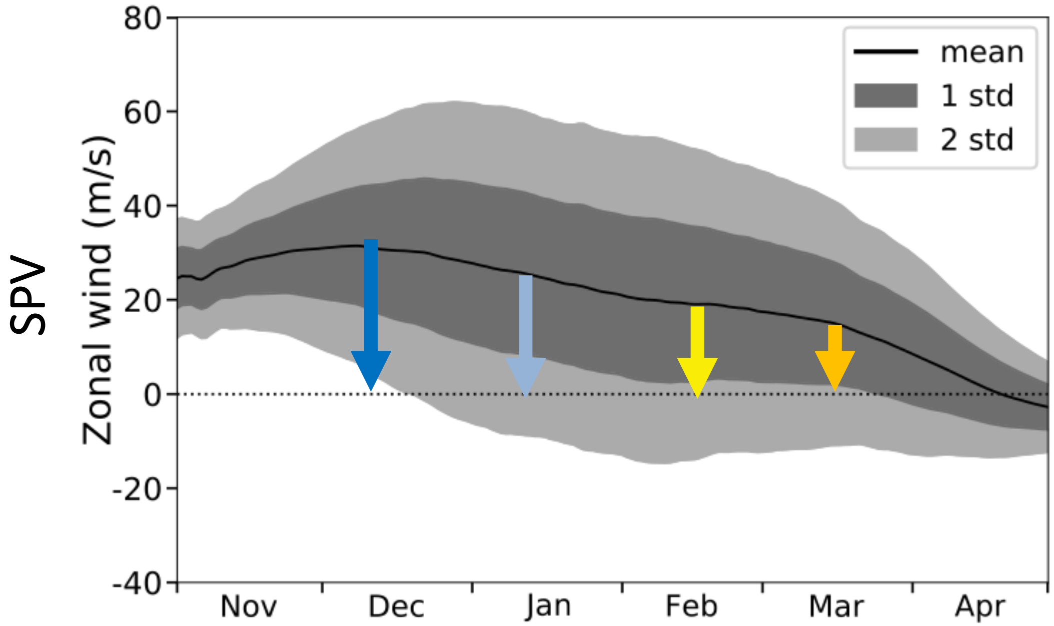


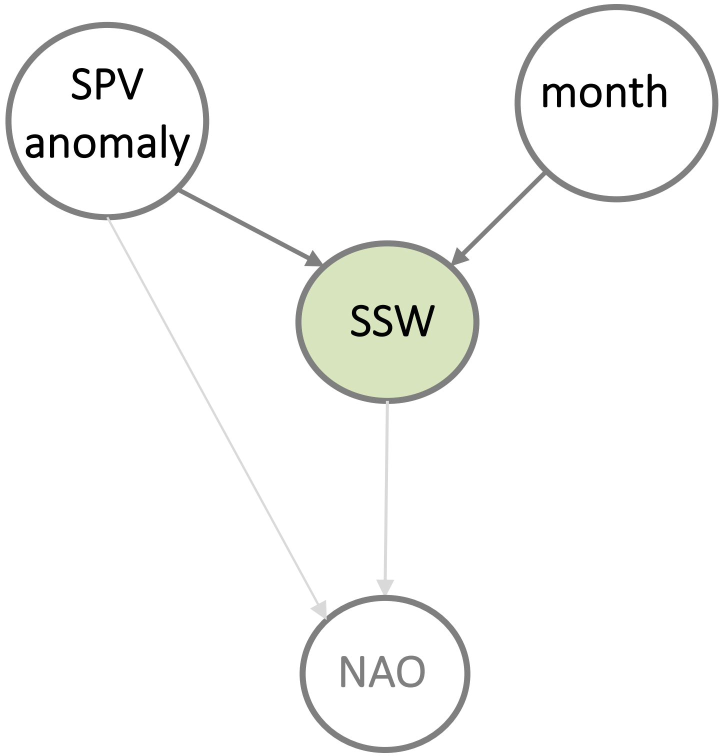



## Effect of SSWs on surface temperature



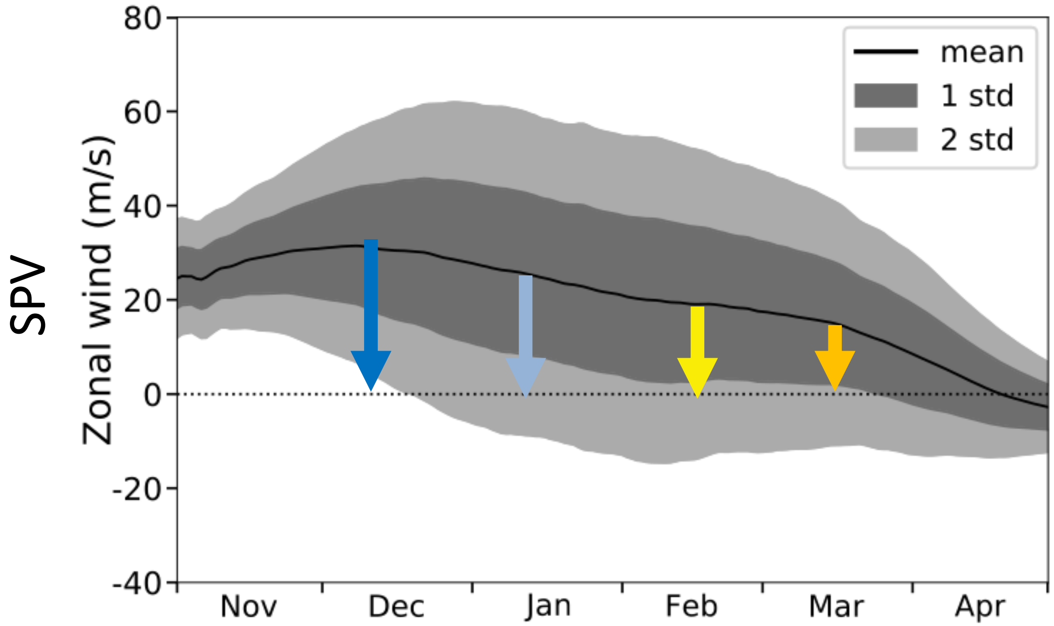
SSWs := Days when the stratospheric polar vortex (SPV) winds turn negative



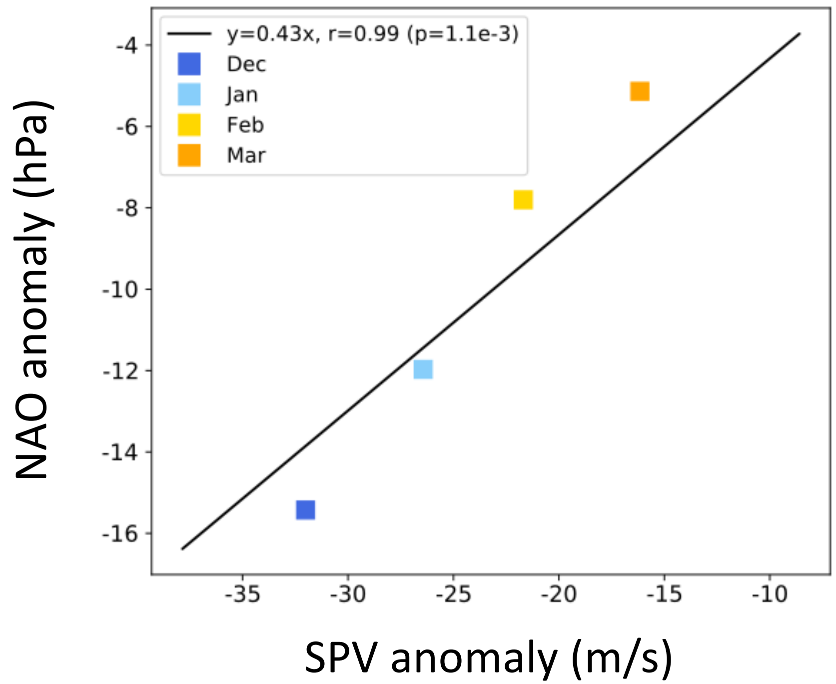




Monnin, Kretschmer, Polichtchouk, *Int. J. Clim.* (2021)

SSWs := Days when the stratospheric polar vortex (SPV) winds turn negative

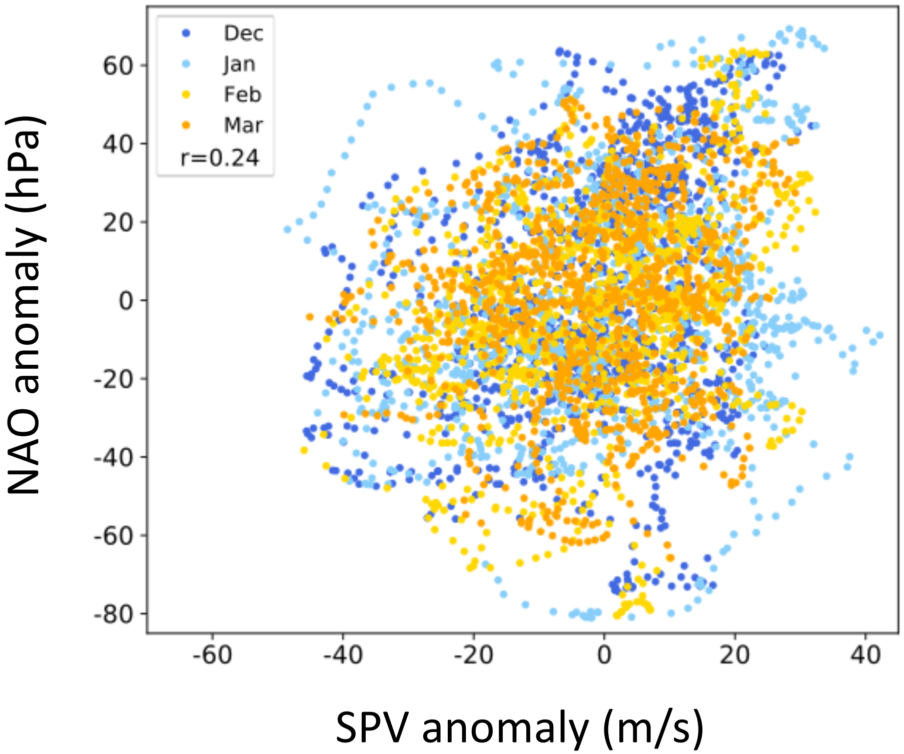



This explains the stronger impacts

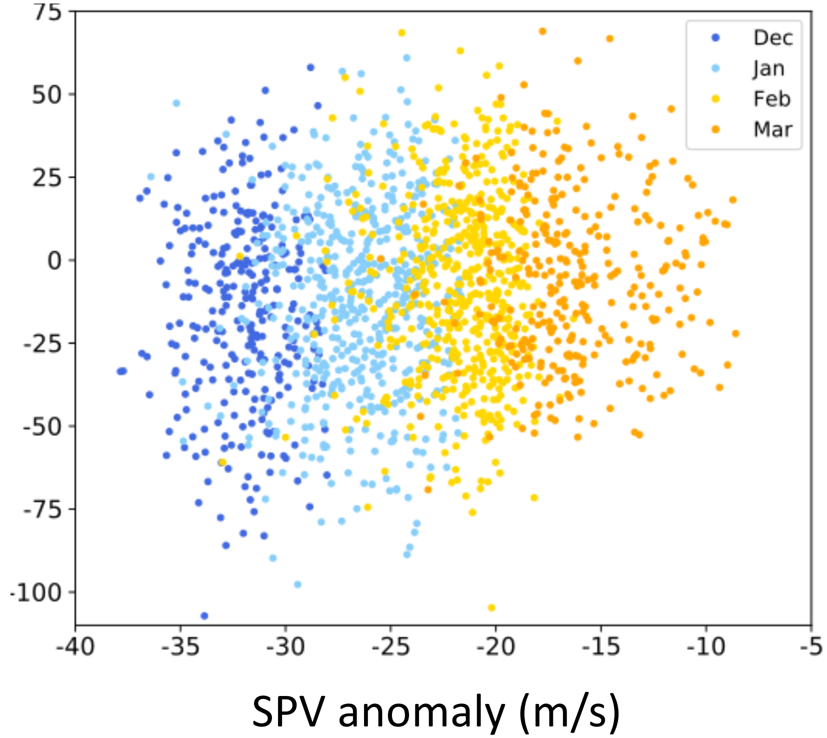


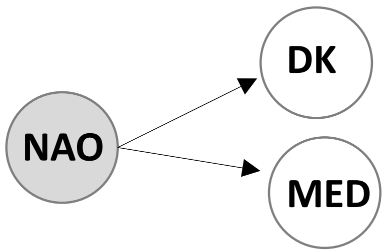

Monnin, Kretschmer, Polichtchouk, *Int. J. Clim.* (2021)

## All winter days

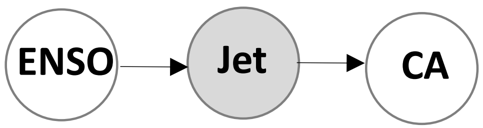



## Only SSWs

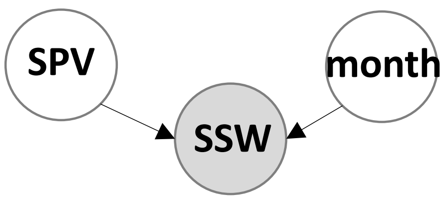





- There is an open path DK <-- NAO --> MED
- Conditioning on NAO blocks this path



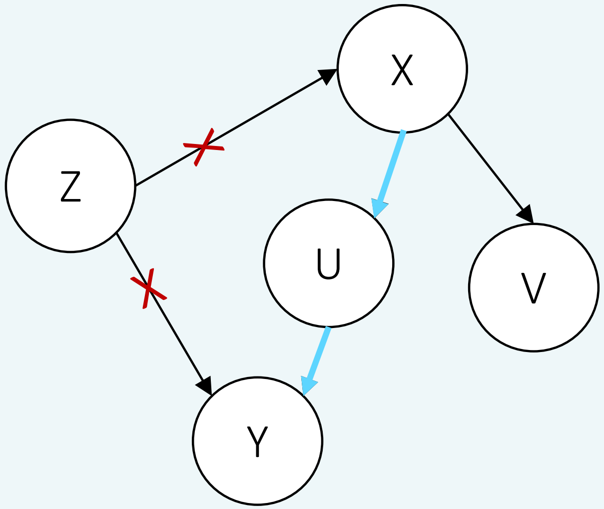

- There is an open path ENSO --> Jet --> CA
- Conditioning on Jet blocks this path



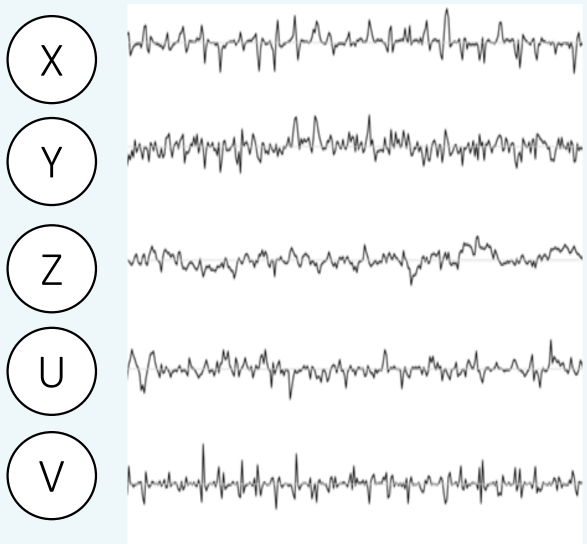

- The path SPV --> SSW <-- month is blocked
- Conditioning on SSW opens this path

## Task: 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

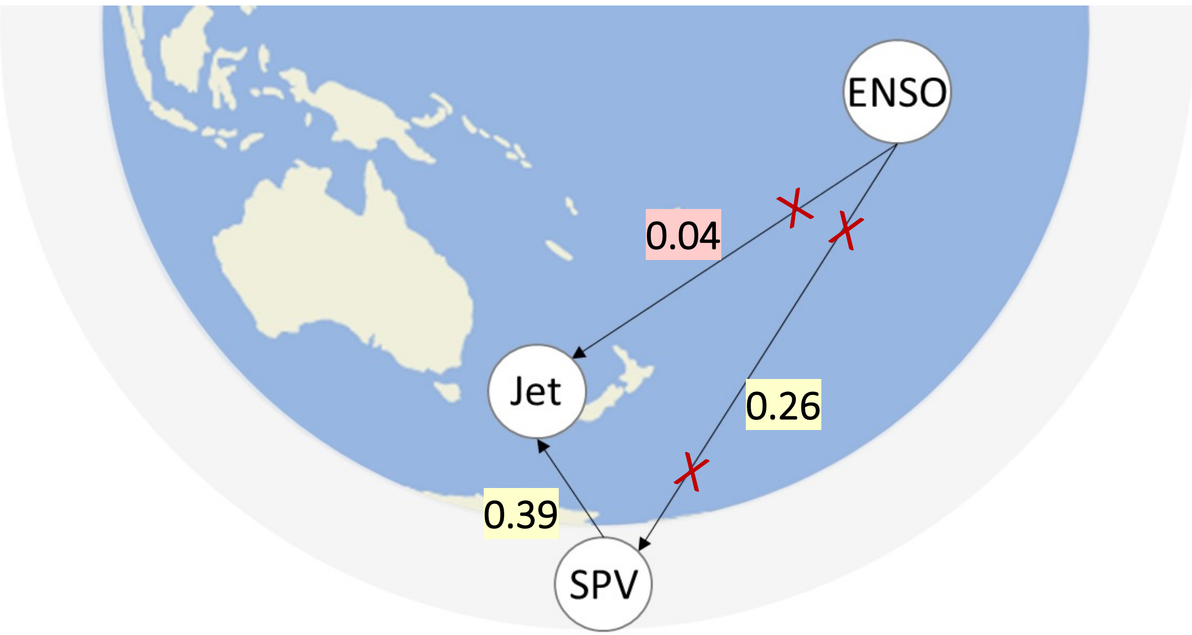
P(Y |do(X))) = P(Y | X, Z)

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

linear case:

Y = a X + b Z

# Direct and indirect pathways




(OND mean, NCEP)

Total effect of ENSO on Jet:

Jet = 0.14 ENSO + ε

Direct (tropospheric) pathway:

Jet = 0.04 ENSO + 0.39 SPV + ε

Indirect (stratospheric) pathway:

SPV = 0.26 ENSO + ε

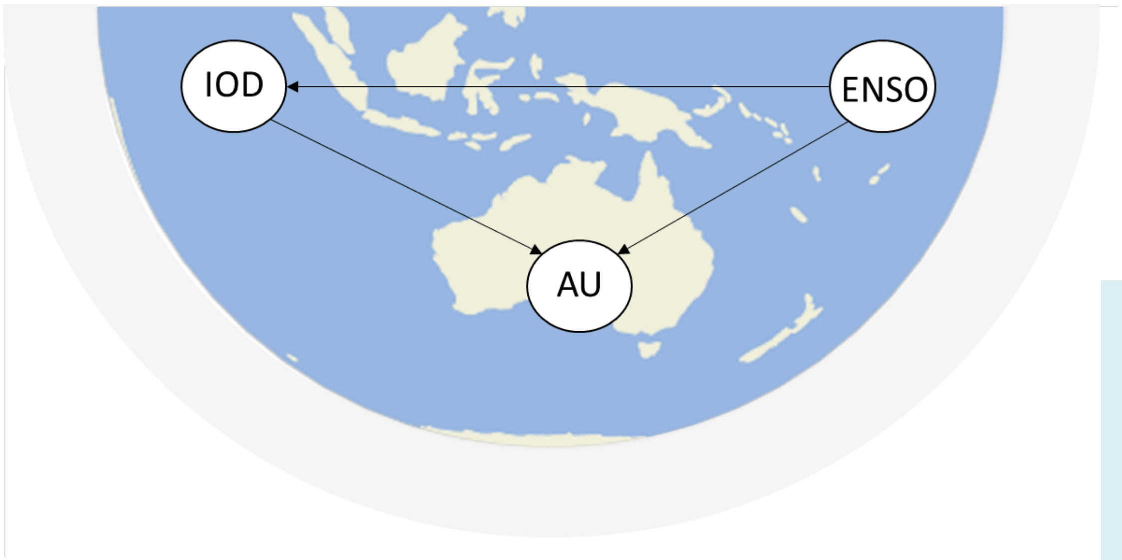
Jet = 0.39 SPV + 0.04 ENSO + ε

0.26 * 0.39 = 0.10

tropo + strato = 0.04 + 0.10

Total = 0.14

# Nonlinear case

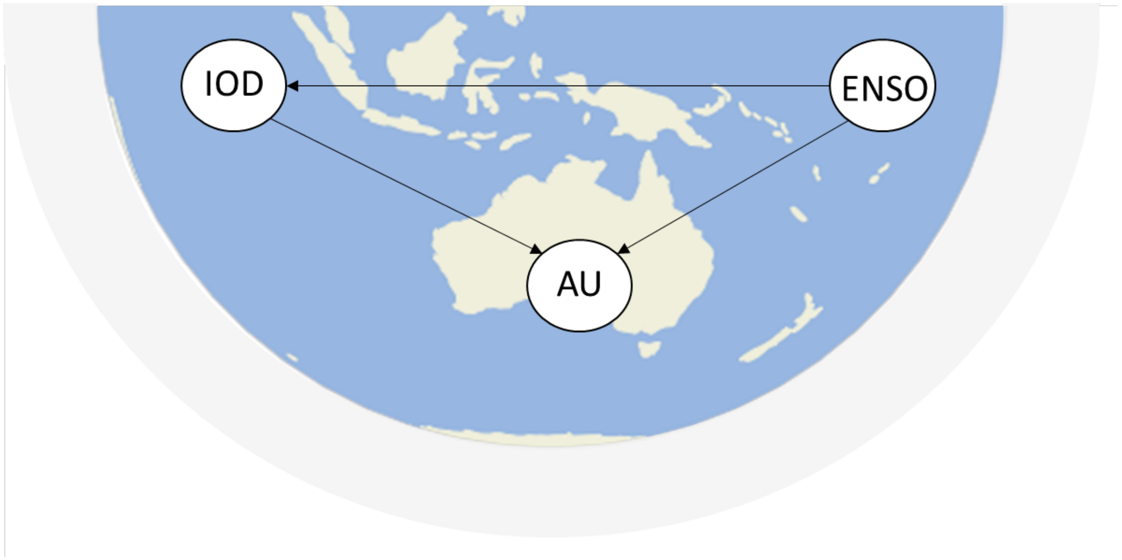



Precipitation in Australia (AU) is affected by ENSO and by the Indian Ocean Dipole (IOD)

The relationships likely involve non-linearities

(SON mean, NCEP)

Kretschmer et al., *BAMS* (2021)

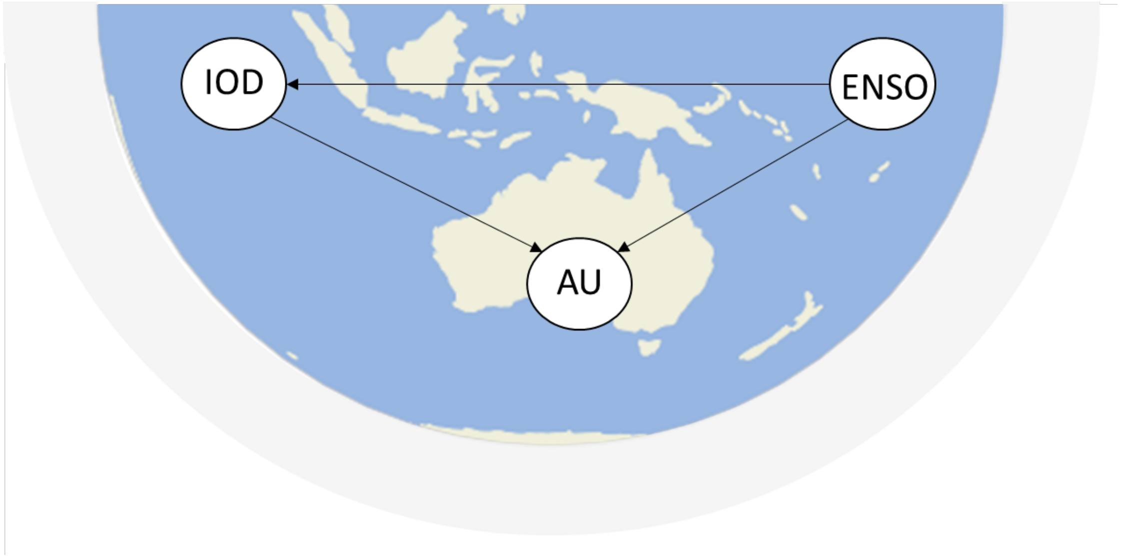



<u>We stratify the data into different categories</u>

**AU**: below/above average

**IOD**: negative/neutral/positive phase

**ENSO**: La Niña/neutral/El Niño




Conditional probabilities 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 |

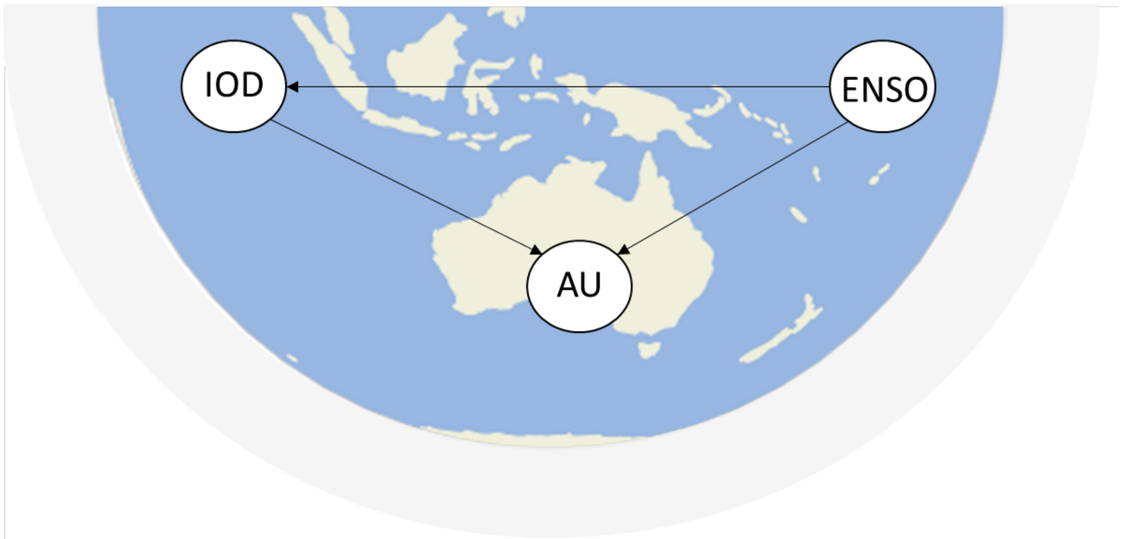
We stratify the data into different categories

**AU**: below/above average

**IOD**: negative/neutral/positive phase

**ENSO**: La Niña/neutral/El Niño

Kretschmer et al., *BAMS* (2021)




Conditional probabilities 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 |

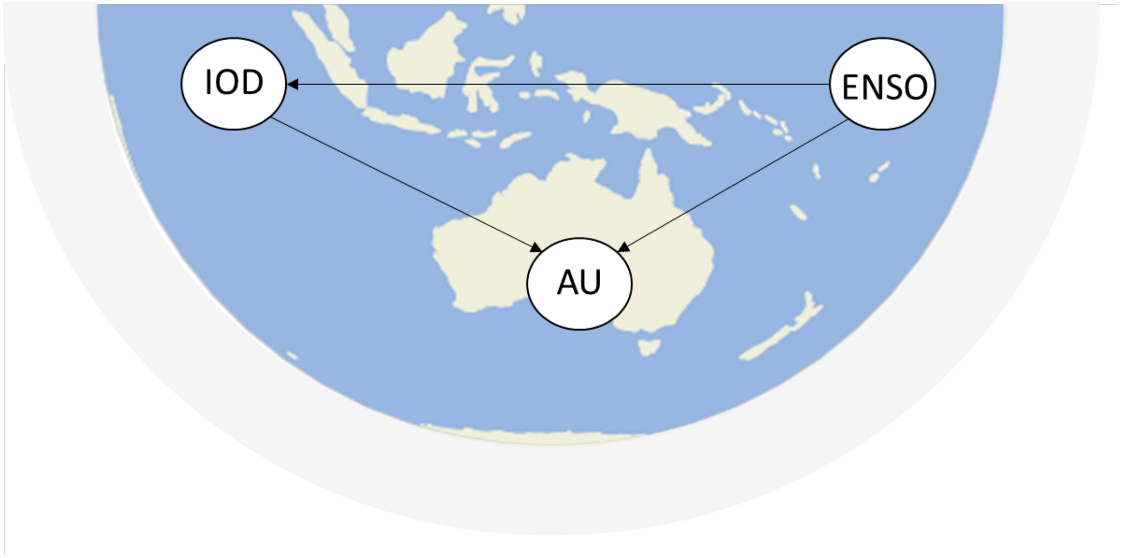
Above average precipitation is unlikely during El Niño

P(AU+ | El Niño ) = 0.22

Above average precipitation is unlikely during IOD+

P(AU+ | IOD+) = 0.30

Kretschmer et al., *BAMS* (2021)




Conditional probabilities 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+ | El Niño, IOD+) = 0.24

P(AU+ | El Niño) = 0.22

0.24/0.22 = 1.09

But what if we believed that IOD affected ENSO, rather than the other way around?

The conditional proabability tables would be unchanged, but their interpretation would be completely different.

Conditional probabilities 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 ENSO, given IOD?

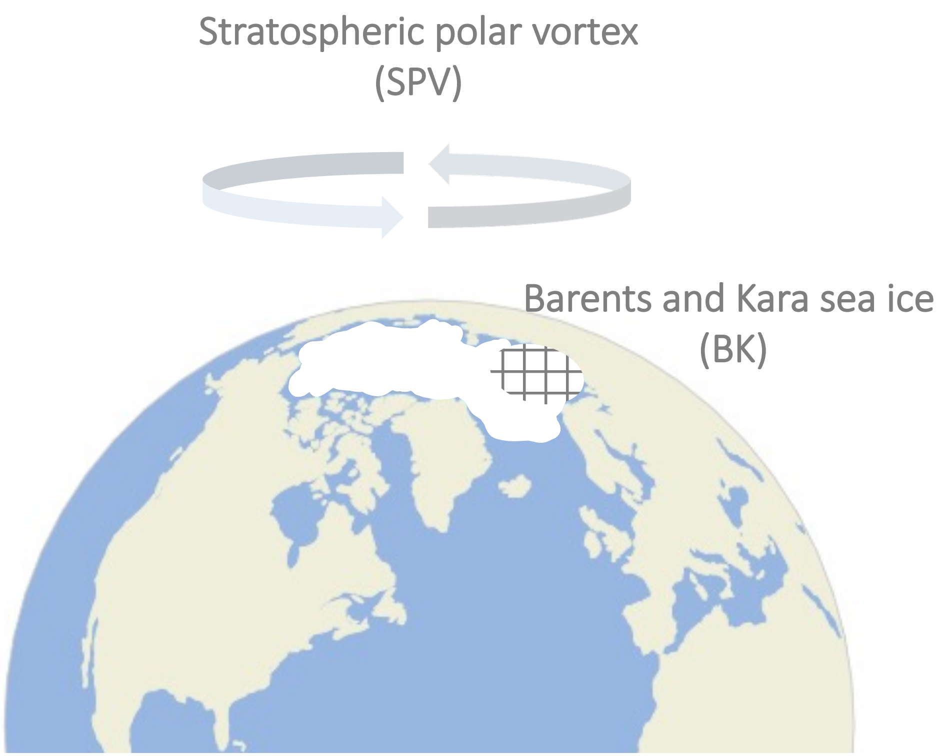
P(AU+ | El Niño, IOD+) = 0.24

P(AU+ | IOD+) = 0.30

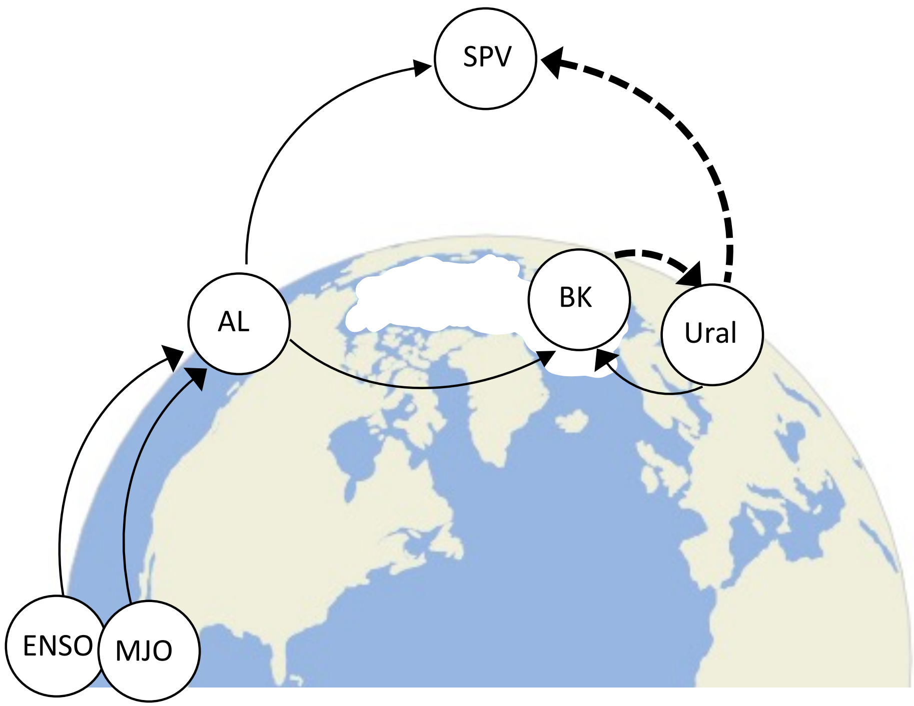
0.24/0.30 = 0.80

Interpretation of data depends on causal assumptions!

# A more complex example

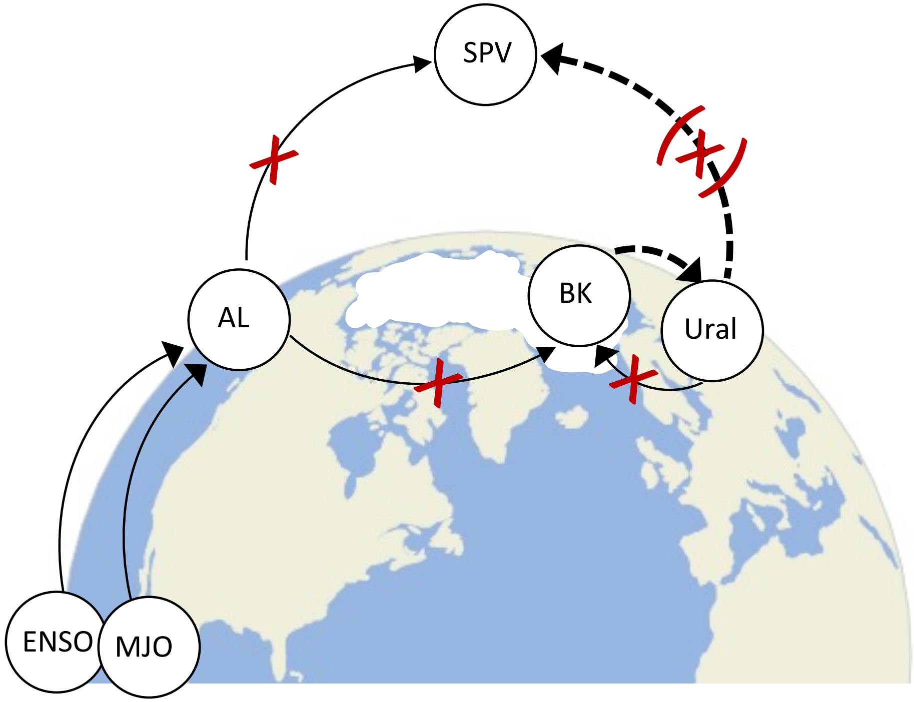



How strong is the causal effect of Barents Kara sea ice (BK) in autumn on the winter stratospheric polar vortex (SPV)?




*A reduction in Barents and Kara sea ice concentrations (BK) is assumed to enhance sea level pressure over the Ural Mountain region (Ural). This causes a weakening of the vortex (SPV). However, Ural sea level pressure also affects BK sea ice. Further, tropical Pacific variability, e.g. in the form of the El Niño–Southern Oscillation or the Madden–Julian Oscillation (ENSO/MJO), can affect the SPV via altered sea level pressure anomalies over the Aleutian Low region (AL). As the AL can also affect BK via Rossby wave propagation, it confounds the analysis of the BK to SPV pathway.*

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




Open Paths from BK to SPV:

| | |
|---|---|
| **BK --> Ural --> SPV** | URAL (after OND) |
| BK <-- AL --> SPV | |
| BK <-- Ural --> SPV | $URAL_{OND}$ |

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

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

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

nature geoscience

Explore content About the journal Publish with us Subscribe

nature > nature geoscience > articles > article

Published: 10 October 2016

**Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss**

Kelly E. McCusker, John C. Fyfe & Michael Sigmond

*Nature Geoscience* 9, 838–842 (2016) | Cite this article

4218 Accesses | 152 Citations | 81 Altmetric | Metrics

nature communications

Explore content About the journal Publish with us

nature > nature communications > articles > article

Open Access | Published: 06 March 2017

**The missing Northern European winter cooling response to Arctic sea ice loss**

James A. Screen

*Nature Communications* 8, Article number: 14603 (2017) | Cite this article

9 Altmetric | Metrics

nature geoscience PERSPECTIVE

https://doi.org/10.1038/s41561-018-0059-y

**Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across imate models**

es A. Screen[1]*, Clara Deser[2], Doug M. Smith[3], Xiangdong Zhang
l J. Kushner[5], Thomas Oudar[5], Kelly E. McCusker[6] and Lantao Sun

nature communications

Explore content About the journal Publish with us

nature > nature communications > articles > article

Published: 02 September 2014

**Weakening of the stratospheric polar vortex by Arctic sea-ice loss**

Baek-Min Kim, Seok-Woo Son, Seung-Ki Min, Jee-Hoon Jeong, Seong-Joong Kim, Xiangdong Zhang, Taehyoun Shim & Jin-Ho Yoon

*Nature Communications* 5, Article number: 4646 (2014) | Cite this article

16k Accesses | 359 Citations | 279 Altmetric | Metrics

REVIEW ARTICLE nature climate change

https://doi.org/10.1038/s41558-019-0662-y

**Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather**

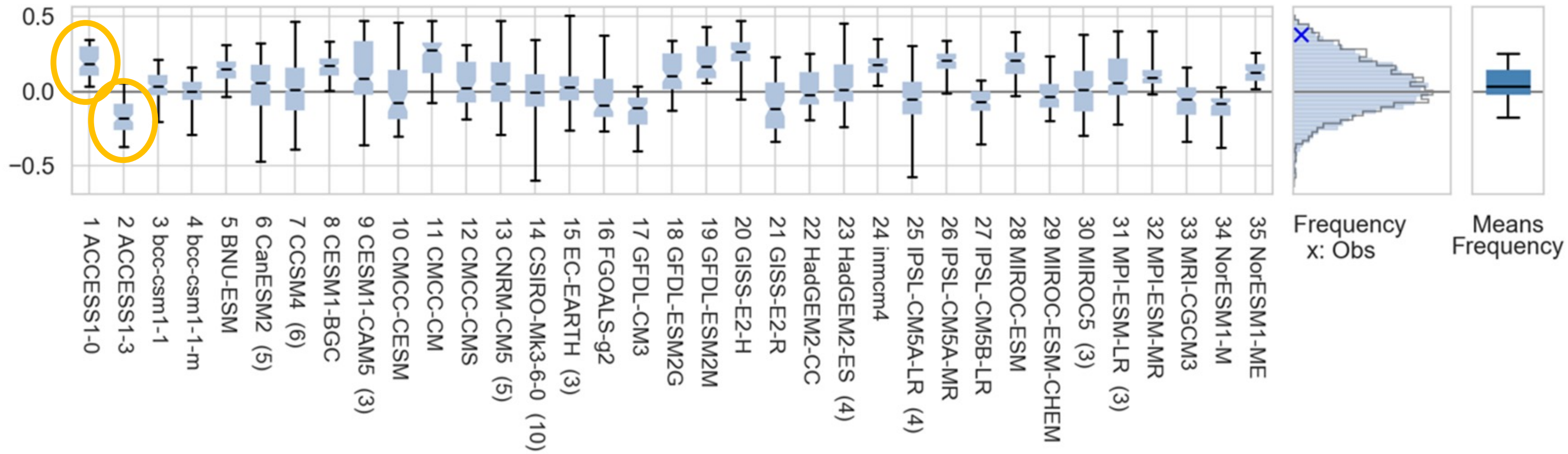
J. Cohen[1,2]*, X. Zhang[3], J. Francis[4], T. Jung[5,6], R. Kwok[7], J. Overland[8], T. J. Ballinger[9], U. S. Bhatt[3], H. W. Chen[10,11], D. Coumou[12,13], S. Feldstein[11], H. Gu[14], D. Handorf[5], G. Henderson[15], M. Ionita[5], M. Kretschmer[13], F. Laliberte[16], S. Lee[11], H. W. Linderholm[17,18], W. Maslowski[19], Y. Peings[20], K. Pfeiffer[1], I. Rigor[21], T. Semmler[5], J. Stroeve[22], P. C. Taylor[23], S. Vavrus[24], T. Vihma[25], S. Wang[14], M. Wendisch[26], Y. Wu[27] and J. Yoon[28]

$SPV_{JFM}$ = **a** $BK_{OND}$ + Confounders + ε

We estimate **a** in moving windows for different CMIP5 models in the historical runs (from 1900-2005)

Large spread of standardized regression coefficients

Effect stronger in obs. but within spread




Mean causal effect only **≈ 0.05**

SPV = **a** BK + Confounders + ε

H0: *No* influence of BK on SPV, i.e. **a** = 0

H1: Influence of BK on SPV, i.e. **a** ≠ 0

**a** ≈ 0.05 (not statistically significant)

**we cannot reject H0...**

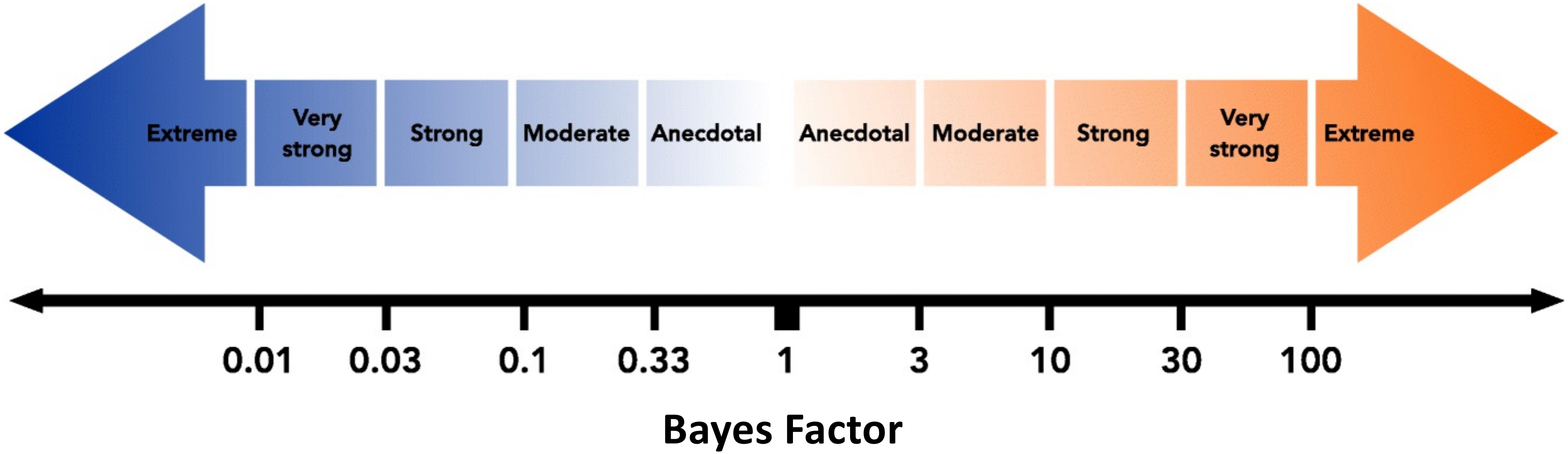
<ins>... but this does not prove H0!</ins>

$$H0: \Delta SPV = b_0\, \Delta T + \varepsilon_0$$

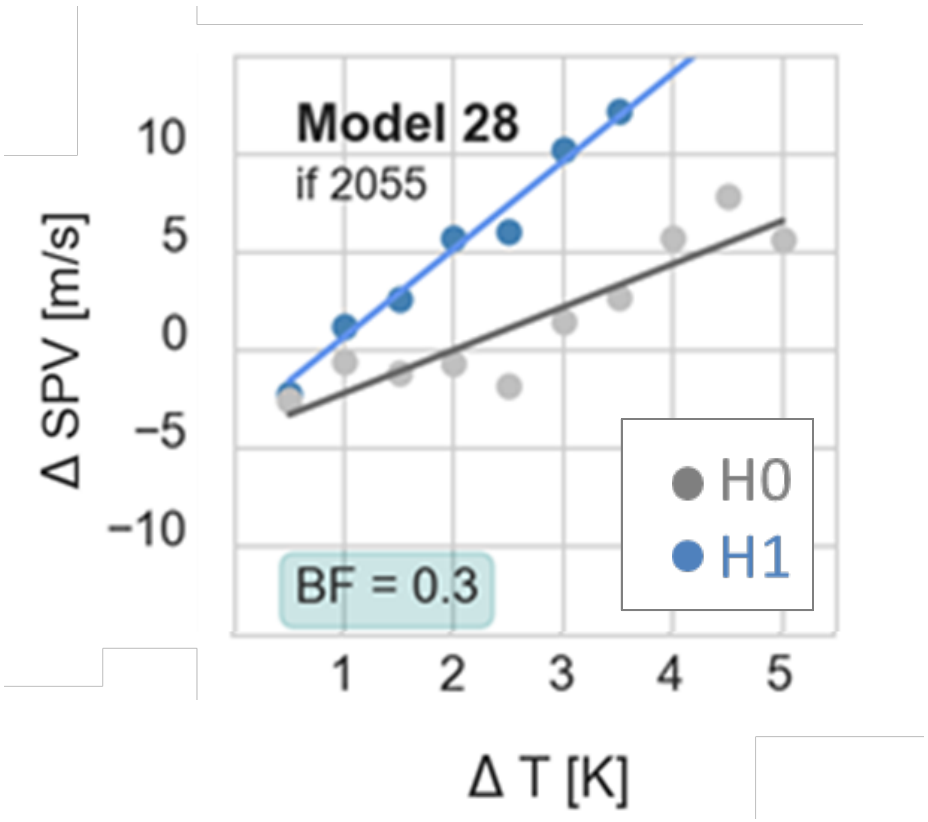
$$H1: (\Delta SPV - a\, \Delta BK) = b_1\, \Delta T + \varepsilon_1$$

with a in [0.025, 0.1]

$$BF = \frac{P(\text{data}\,|H0)}{P(\text{data}\,|H1)}$$




Quintana & Williams (2018) based on Lee & Wagenmaker (2013)




$$\mathrm{BF} = \frac{\mathrm{P(data\,|H0)}}{\mathrm{P(data\,|H1)}}$$

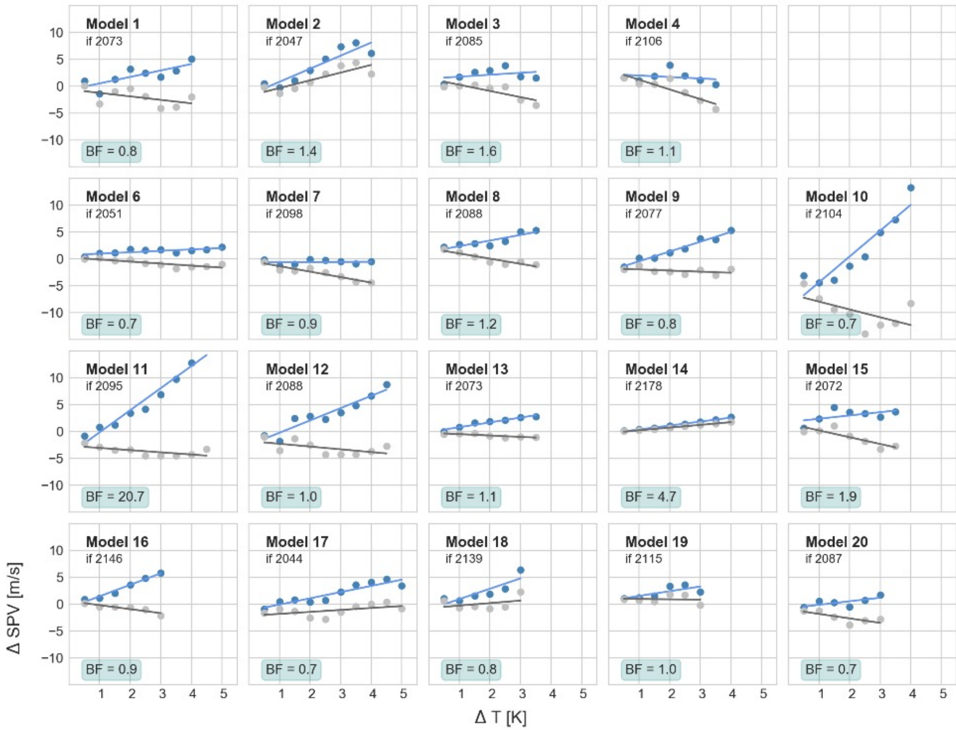
H0: $\Delta SPV = b_0\,\Delta T + \varepsilon_0$

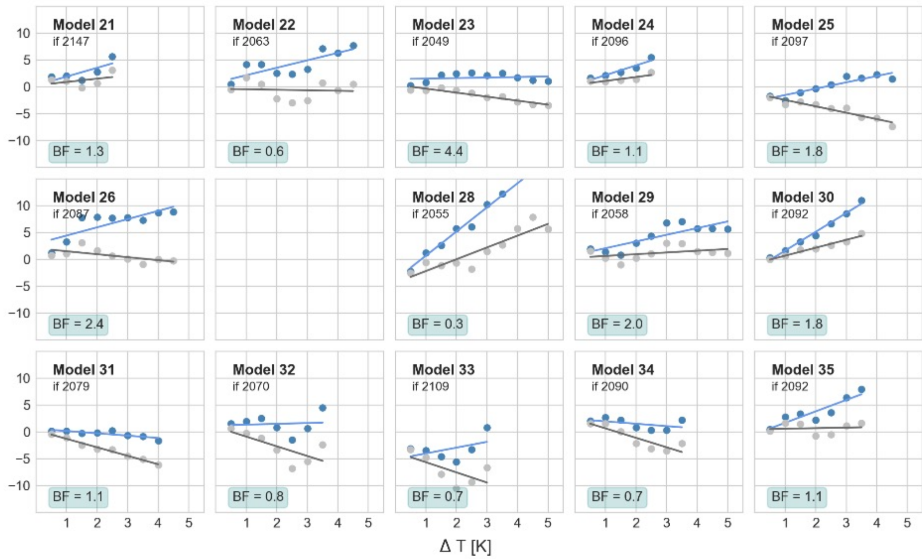
H1: $(\Delta SPV - a\,\Delta BK) = b_1\,\Delta T + \varepsilon_1$

with a in [0.025, 0.1]

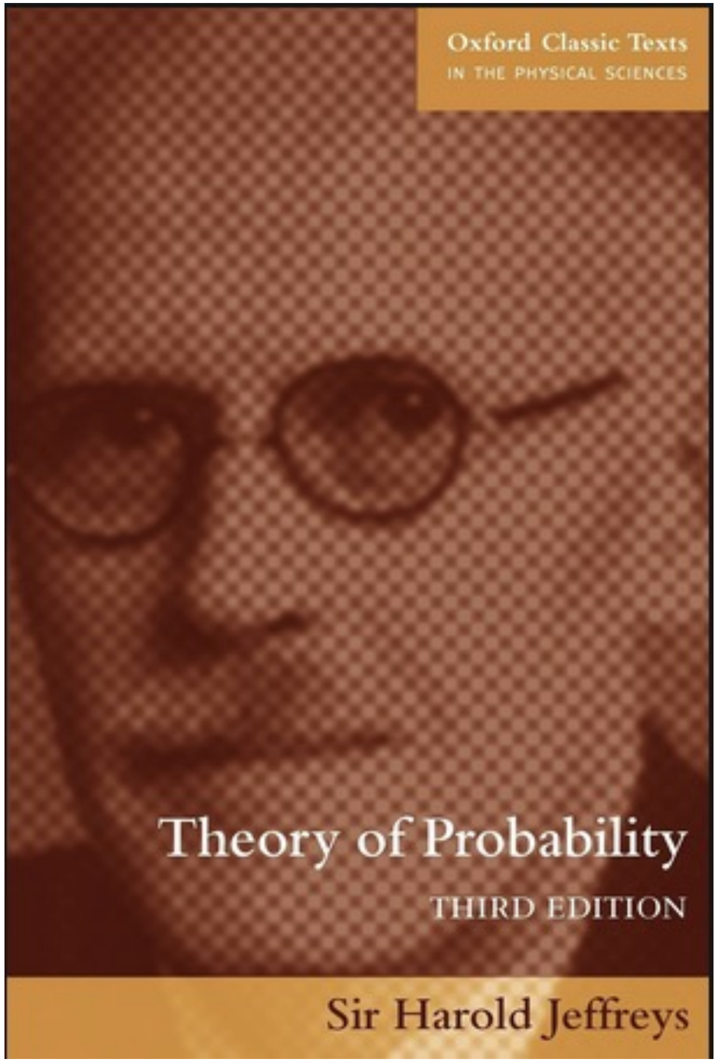
The data is slightly more likely under H1

BF is close to unity --- > no proof of H1
but neither proof against it!




Kretschmer et al. *WCD* (2020)




"There are cases where there is no positive evidence for a new parameter, but important consequences might follow if it was not zero, and we must remember that [a Bayes factor] > 1 does not prove that it is zero, but merely that it is more likely to be zero than not. Then it is worth while to examine the alternative [hypothesis] further and see what limits can be set to the new parameter, and thence to the consequences of introducing it." (Jeffreys 1961)

- Causal knowledge/hypotheses about the data-generating mechanisms are needed to interpret correlations and to extract causal effects from data

- Causal inference gives the formal rules for how to achieve this

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

- To extract causal effects from data, one needs to control for all confounding factors

- Bayes Factors can be computed to quantify under which hypothesis the data is more likely

**Scientific data analysis requires causal reasoning**

## Causal inference and teleconnections (+ jupyter notebooks)

BAMS Article

**Quantifying Causal Pathways of Teleconnections**

Marlene Kretschmer, Samantha V. Adams, Alberto Arribas, Rachel Prudden, Niall Robinson, Elena Saggioro, and Theodore G. Shepherd

## Conditioning on a common effect

Received: 25 March 2021 | Revised: 13 October 2021 | Accepted: 20 October 2021
DOI: 10.1002/joc.7426

RESEARCH ARTICLE

International Journal of Climatology RMetS

**The role of the timing of sudden stratospheric warmings for precipitation and temperature anomalies in Europe**

Erika Monnin[1] | Marlene Kretschmer[2] | Inna Polichtchouk[3]

## Quantifying the causal effect of sea ice loss

Research article 20 Nov 2020

The role of Barents–Kara sea ice loss in projected polar vortex changes

Marlene Kretschmer[1], Giuseppe Zappa[1,2], and Theodore G. Shepherd[1]
[1]Department of Meteorology, University of Reading, Reading, UK
[2]Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche, Bologna, Italy
**Correspondence:** Marlene Kretschmer (m.j.a.kretschmer@reading.ac.uk)

## Use of causal networks for regional storylines

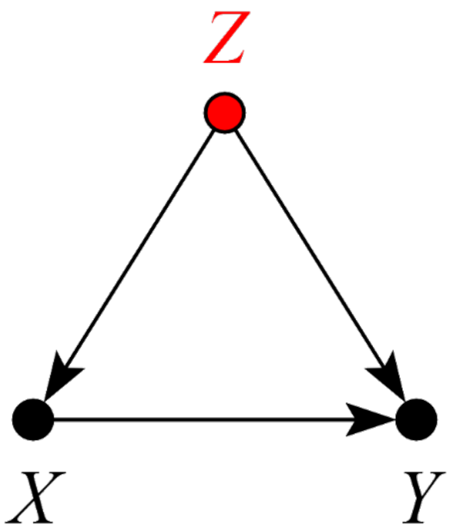
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Special Issue: *The Year in Ecology and Conservation Biology*
ORIGINAL ARTICLE

**Environmental catastrophes, climate change, and attribution**

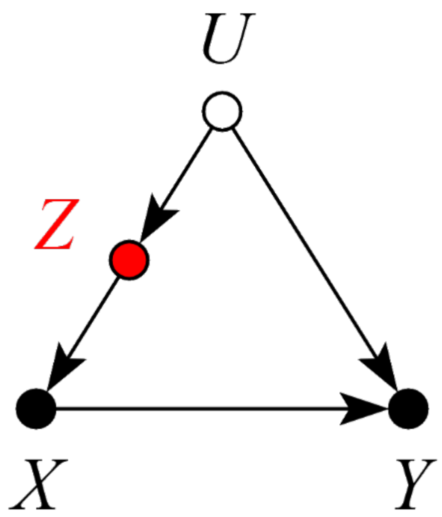
Elisabeth A. Lloyd[1] and Theodore G. Shepherd[2]
[1]Department of History and Philosophy of Science and Medicine, Indiana University, Bloomington, Indiana. [2]Department of Meteorology, University of Reading, Reading, United Kingdom


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

# Good examples of conditioning

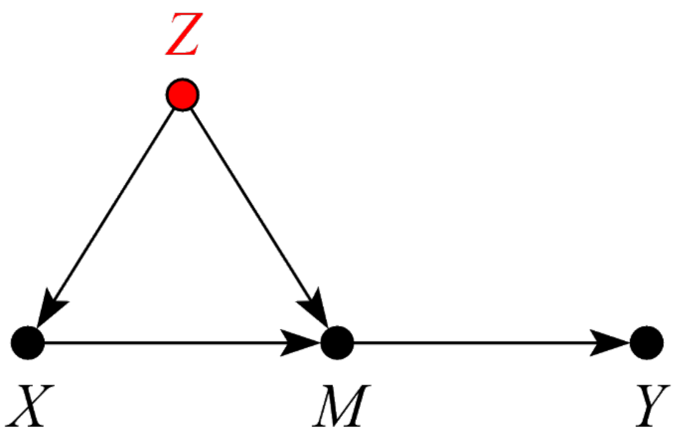



To block the path X <-- Z --> Y

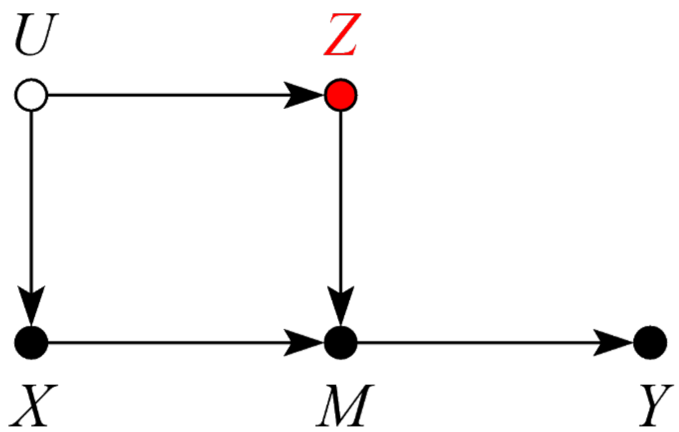



To block the path X <-- Z <-- U --> Y

# Good examples of conditioning

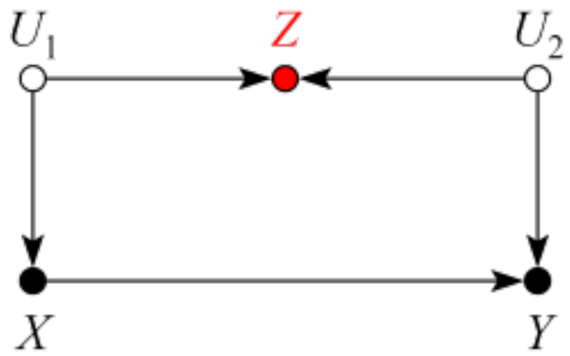



To block the path X <-- Z --> M --> Y




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

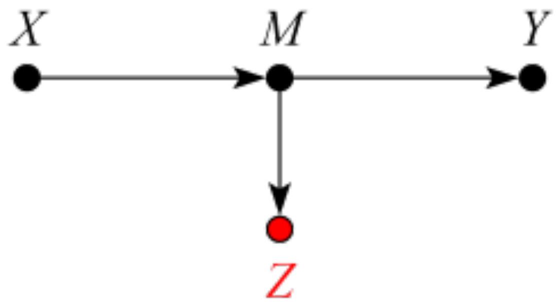
# A bad example of conditioning




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

# Bad examples of conditioning


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




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

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