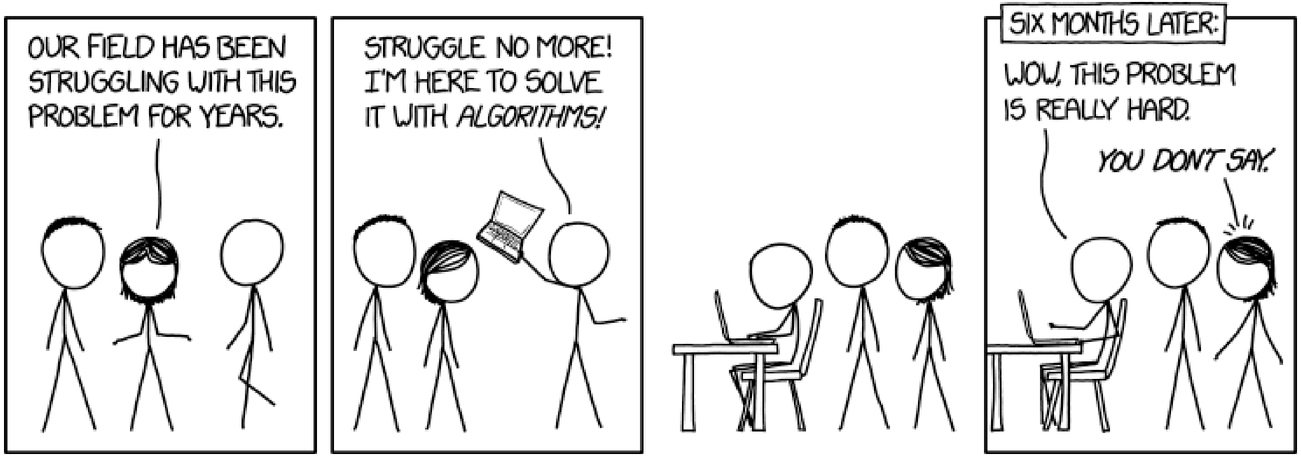
Causal Inference and Causal Discovery in Climate Science

Marlene Kretschmer[1]

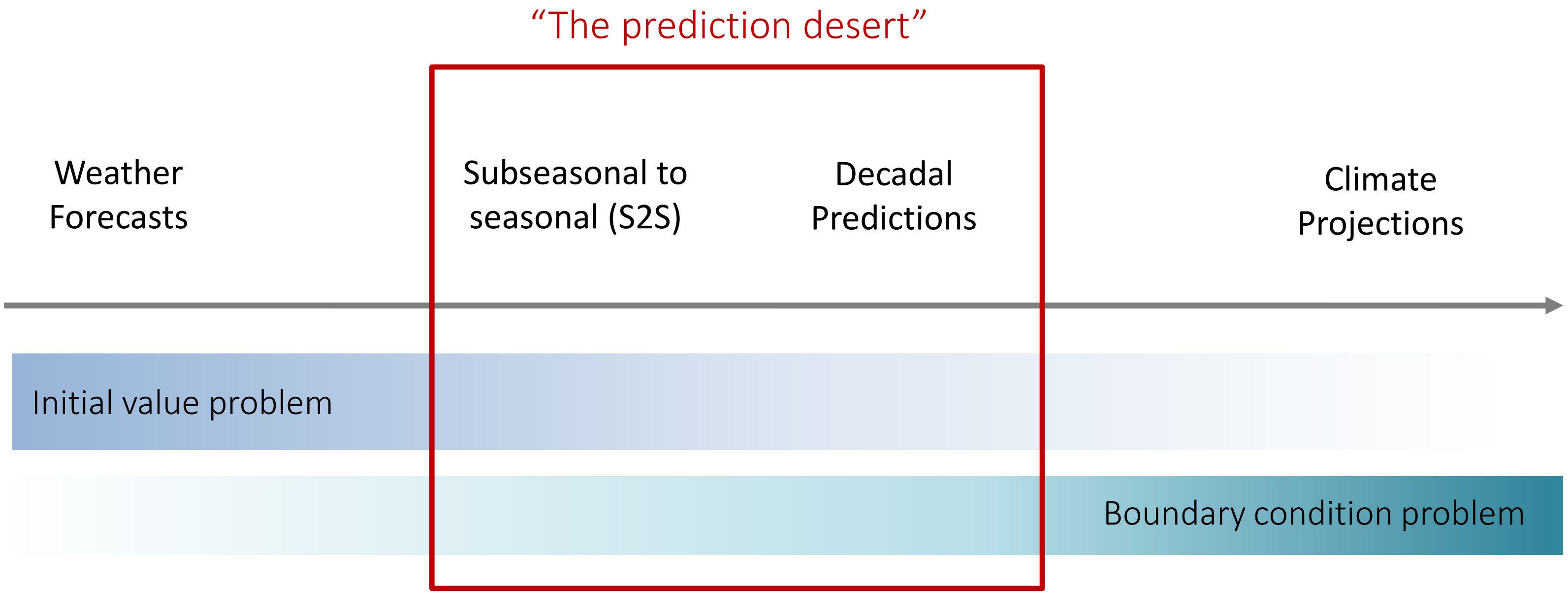
Ted Shepherd[1], Elena Saggioro[1], Alberto Arribas[2],
Rachel Prudden[2], Niall Robinson[2], Sam Adams[2]

[1]University of Reading, [2]Met Office Informatics Lab

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


OUR FIELD HAS BEEN STRUGGLING WITH THIS PROBLEM FOR YEARS.
STRUGGLE NO MORE! I'M HERE TO SOLVE IT WITH *ALGORITHMS!*
SIX MONTHS LATER:
WOW, THIS PROBLEM IS REALLY HARD.
YOU DON'T SAY.

MOTIVATION



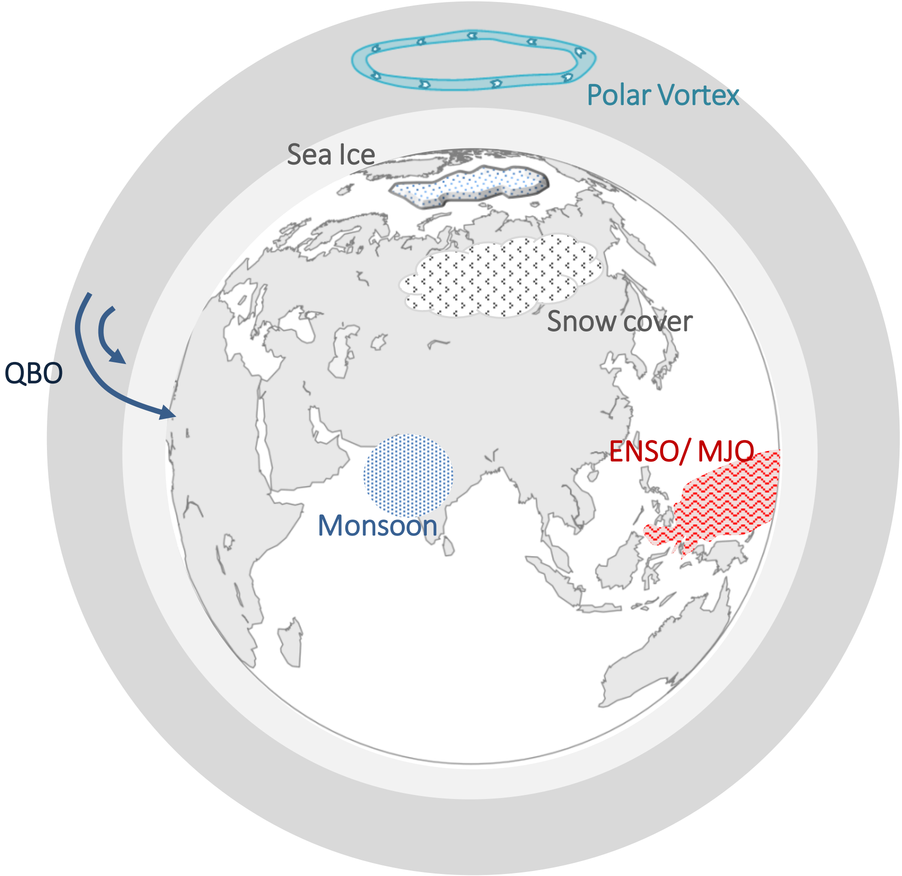
Based on Fig. 2 (Box 11.1) in AR5-WG1.

INTRODUCTION

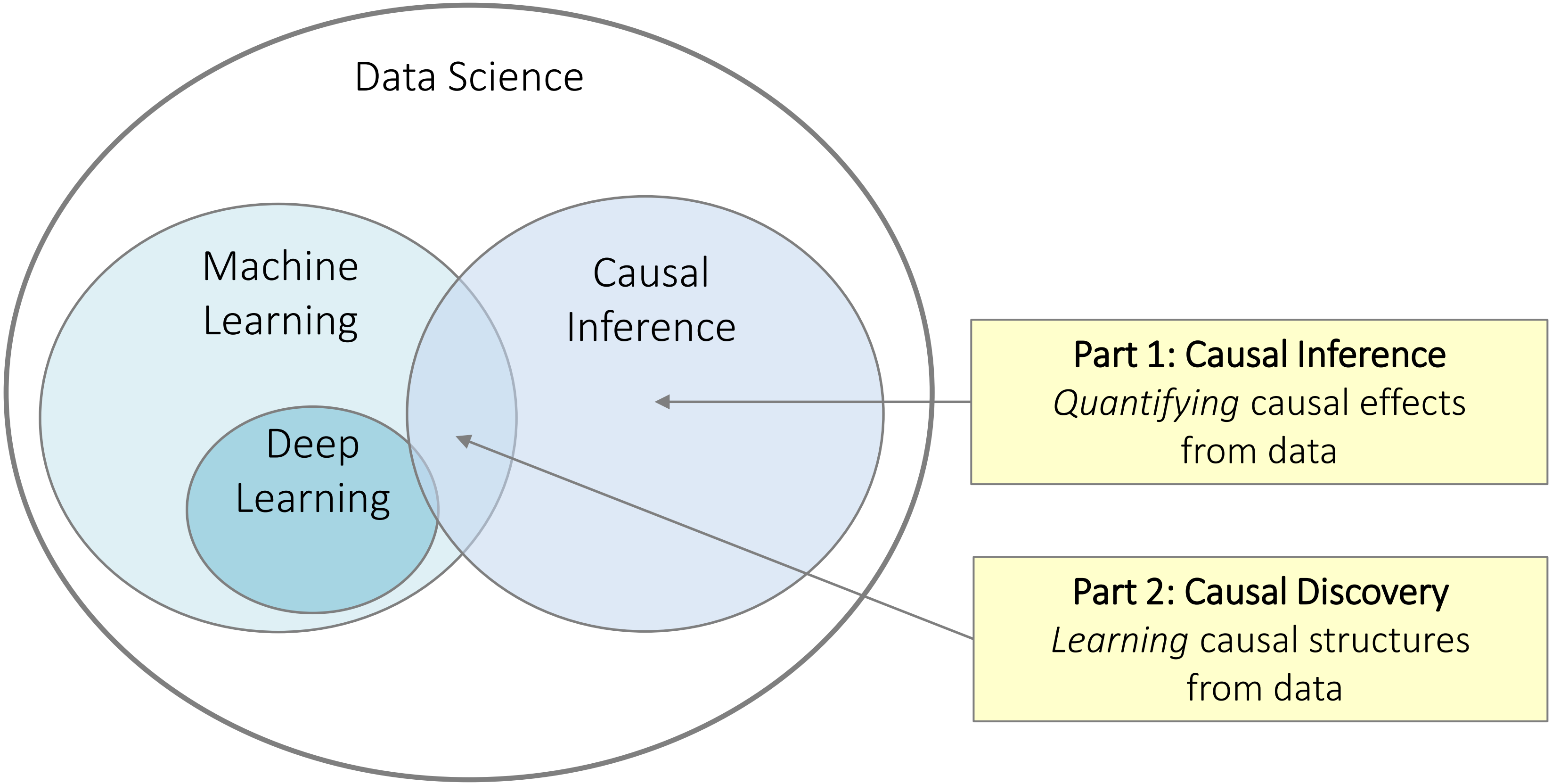
- **Motivation:**
  Improved *understanding* of teleconnections is key to reduce and communicate uncertainties about regional weather and climate predictions

- **Typical questions:**
  How much does ENSO contribute to temperature variability in region A?
  Which processes drive precipitation in region B?

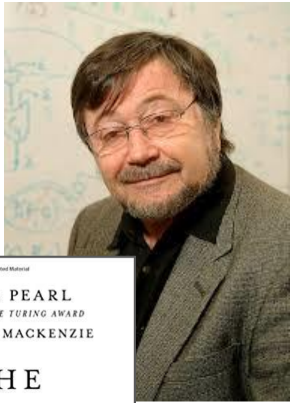
- **Challenge:**
  Extracting the (causal) information from data



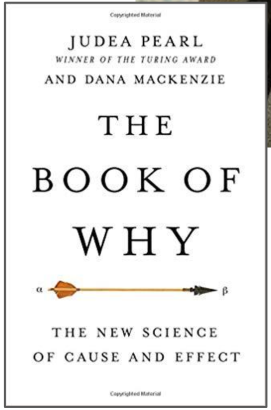
OVERVIEW


Data Science
Machine Learning
Causal Inference
Deep Learning
Part 1: Causal Inference
Quantifying causal effects from data
Part 2: Causal Discovery
Learning causal structures from data

Judea Pearl

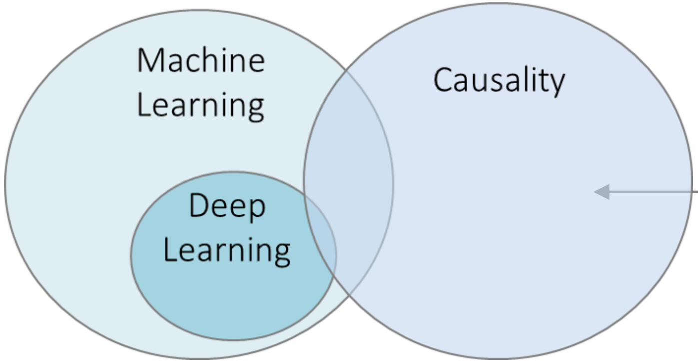


[…] deep learning, I see they're all stuck there on the level of associations. Curve fitting. […] no matter how skillfully you manipulate the data [...], it's still a curve-fitting exercise, albeit complex and nontrivial.



Yann LeCun

[…] sometimes people have accurate models of a phenomenon without any intuitive explanation or causation […] sometimes, requiring explainability is counterproductive.



Part 1: Causal Inference

Quantifying causal effects from data

TELECONNECTIONS

AMS Definition

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

[...] the name refers to the fact that such correlations suggest that information is propagating [...].

Problem

Large gap between

our physical understanding (= causal) and

our statistical description (= correlational)

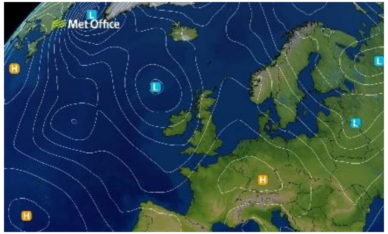
of teleconnections.

How can we reconcile correlations with causation?

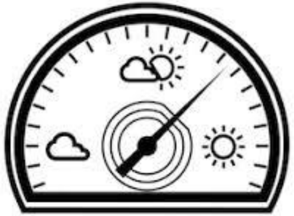
CAUSALITY IN STATISTICS

Example

X: Pressure



Y: Barometer



Intervening in Y will *not* change X:

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

Intervening in X will cause changes in Y:

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

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

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

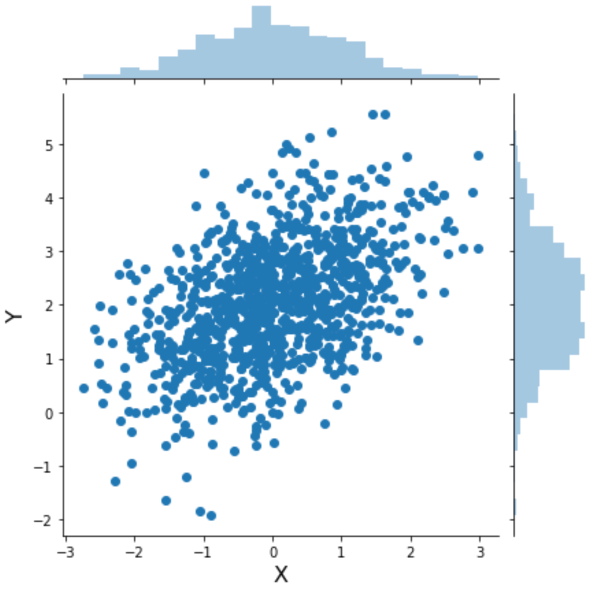
Causal Inference:
Predict the effect of an intervention from data
(without actually doing the intervention)

<u>Causal knowledge (e.g. X --> Y) is required!</u>

J. Pearl, "Causality" (2000)

Toy Example

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



We want the *interventional* cond. distribution

P(Y | **do(X)** = 1)

We can only compute the *observational* cond. distribution

P (Y | X = 1)

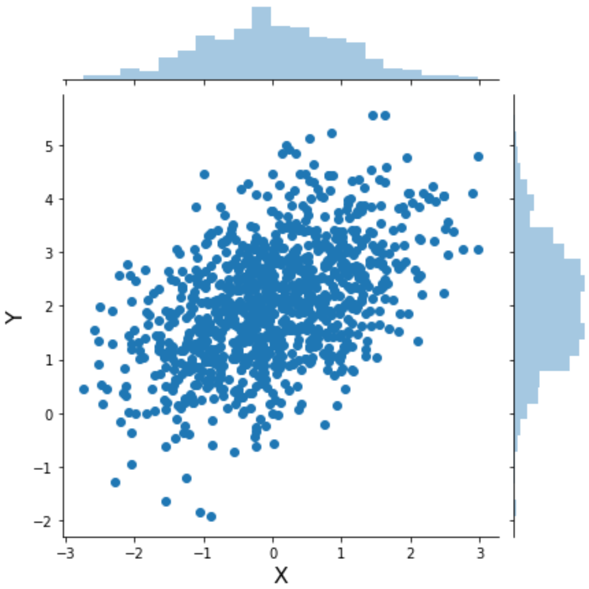
But these are usually not the same

P(pressure| barom. = 1) ≠ P(pressure | do(barom.) = 1)

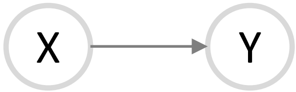
Answer is only possible if we have <u>causal knowledge</u> about the mechanisms that generated the data

Toy Example

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



Expert knowledge



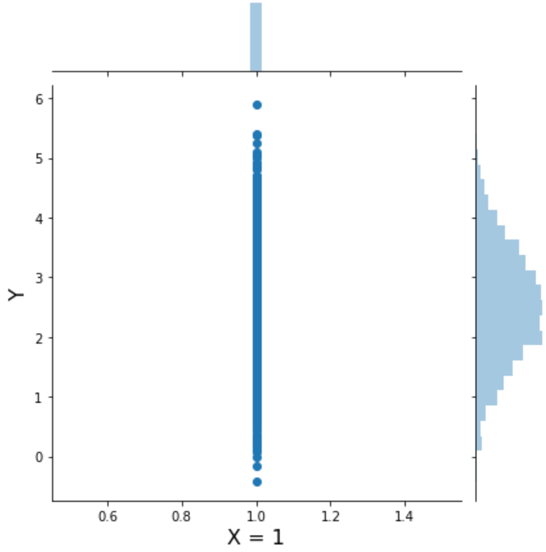
Quantify effect from data

$X = \varepsilon_x$

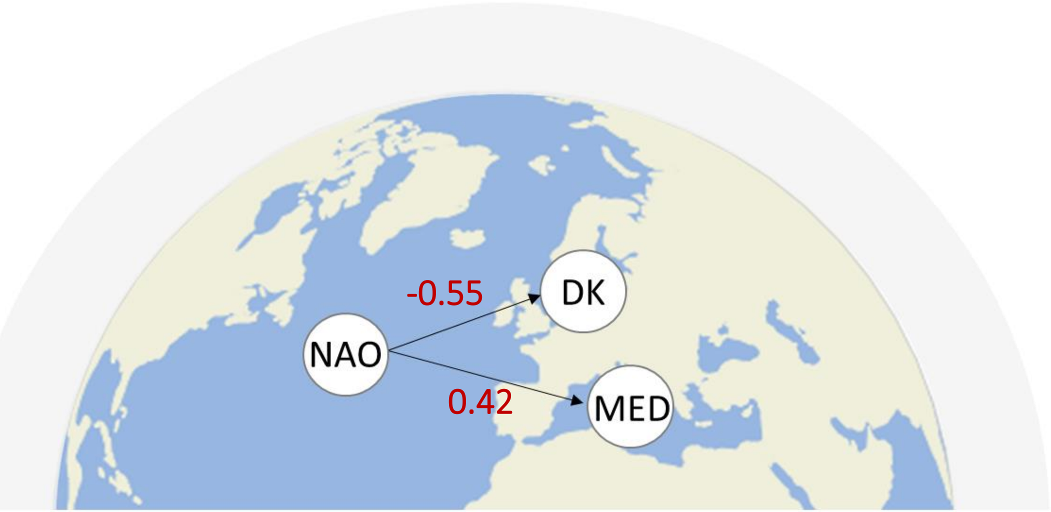
$Y = \mathbf{0.5} * X + 2 + \varepsilon_y$

Answer causal questions

$P(Y \mid do(X) = 1) = P(Y \mid X = 1)$



Example 1: Common Driver



(JJA mean, NCEP)

Precipitation in Denmark and the Mediterranean are significantly correlated

Corr(DK, MED) = -0.25

→ Corr(DK, MED | NAO) = 0.001

DK and MED are independent conditional on NAO

DK = - 0.55 NAO + ε

MED = +0.42 NAO + ε

The causal effect explains the association

-0.25 ≈ -0.55 * 0.42

Kretschmer et al. (in revision at *BAMS*)

EXAMPLE 2: MEDIATOR

What is the effect of ENSO on California precipitation?

CA = 0.0 ENSO + 0.81 Jet + ε

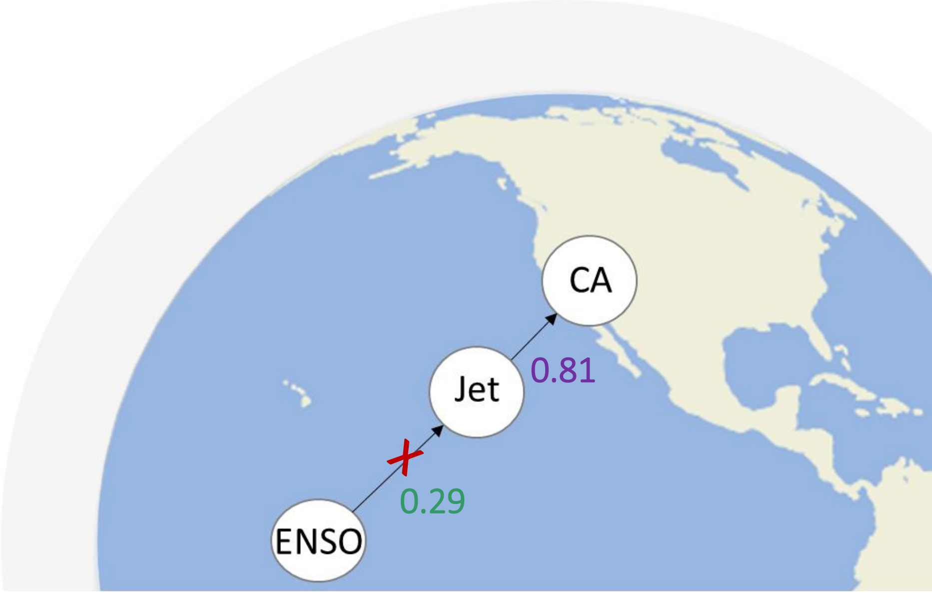
Correct way:

CA = 0.24 ENSO + ε

Or via product along pathway:

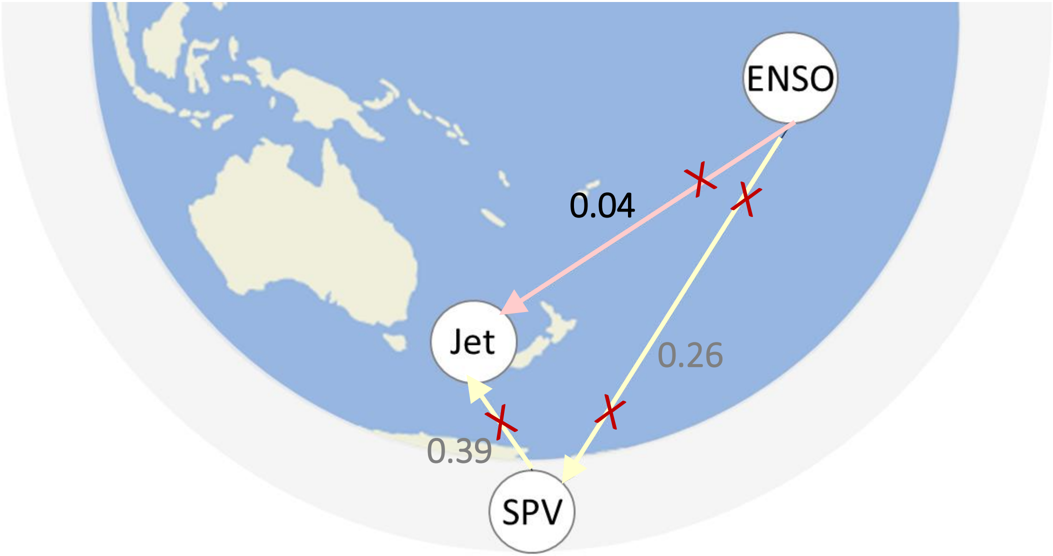
Jet = 0.29 ENSO + ε   CA = 0.81 Jet + ε

0.29 * 0.81 = 0.24



(DJF mean, NCEP)

EXAMPLE 3: INDIRECT AND DIRECT EFFECTS



(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

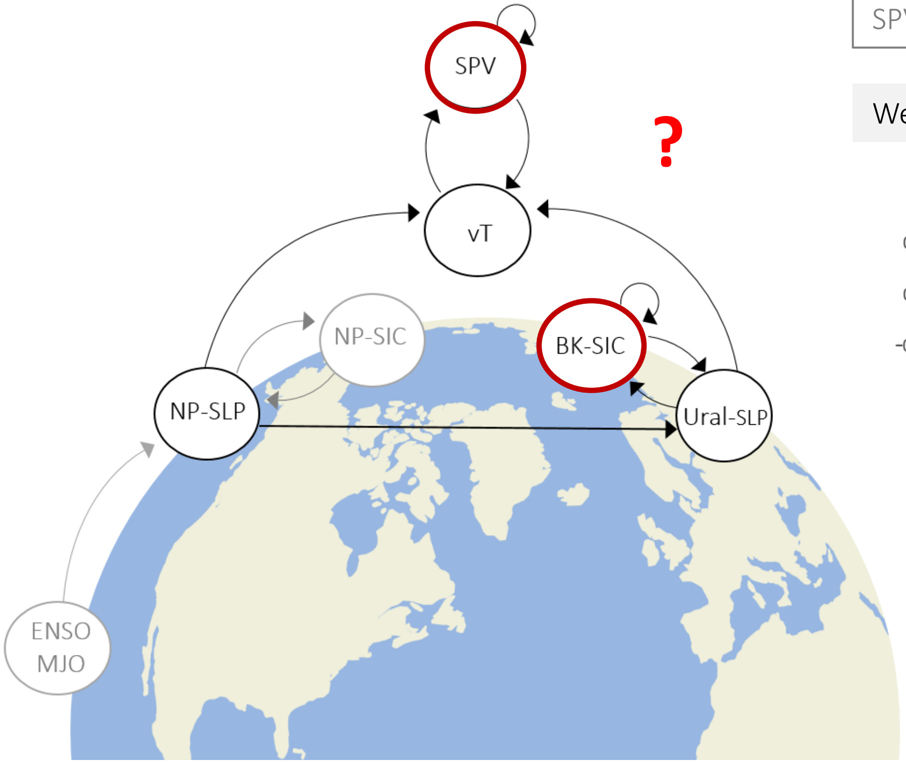
Jet = 0.39 SPV + 0.04 ENSO + ε

0.10 = 0.26 * 0.39

tropo + strato = 0.04 + 0.10

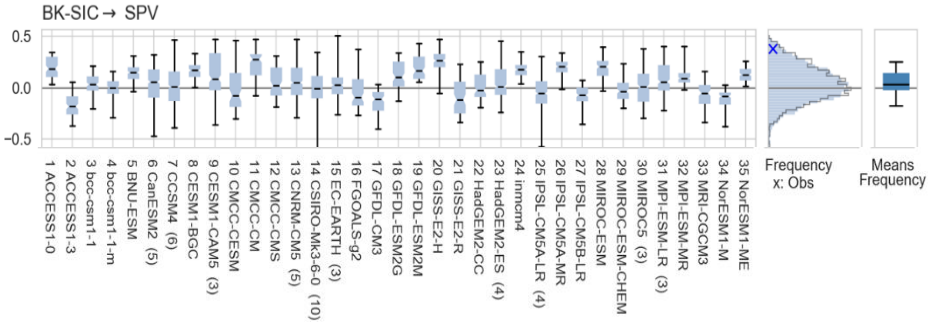
Total = 0.14

EXAMPLE 4: INFLUENCE OF SEA ICE ON THE POLAR VORTEX



SPV_{JFM} = **a** $BK\text{-}SIC_{OND}$ + $Ural\text{-}SLP_{SON}$ + ε

We estimate **a** for each model in the historical CIMP5 runs (1900-2005)



Mean causal effect is only very small, **a ≈ 0.05** but has large implications for the future SPV

EXAMPLE FROM NWP

Received: 23 October 2019 | Revised: 7 March 2020 | Accepted: 9 March 2020 | Published on: 5 May 2020

DOI: 10.1002/qj.3788

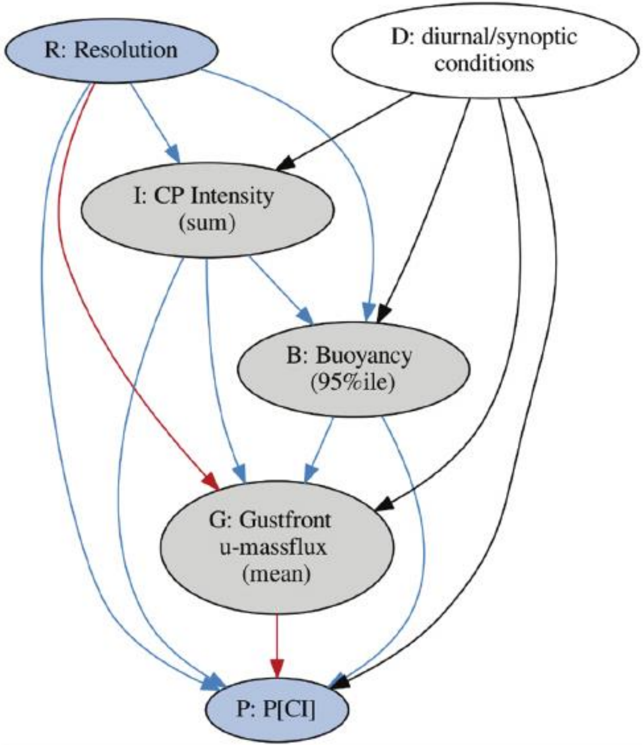
Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Cold-pool-driven convective initiation: using causal graph analysis to determine what convection-permitting models are missing

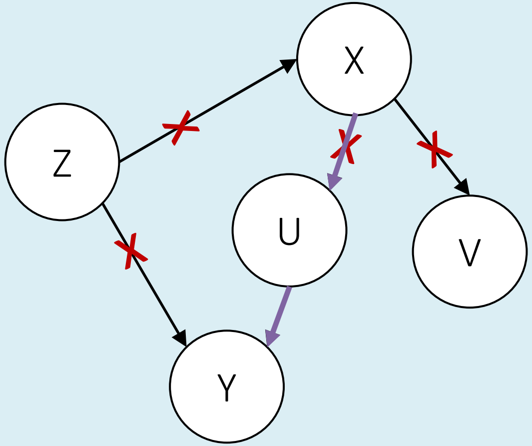
Mirjam Hirt[1] | George C. Craig[1] | Sophia A. K. Schäfer[2] | Julien Savre[1] | Rieke Heinze[3]

Causal network of how resolution can impact convective initiation



Steps of Causal Inference (= "Knowledge Guided Statistics")

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



2. Draw logical implications and test if data support them

U and V must be independent conditioned on X

$U \perp V \mid X$

linear:

Corr (U, V| X)= 0 ?

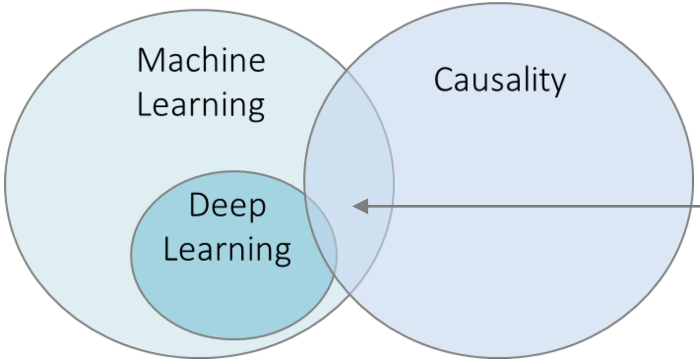
3. Use CI rules to estimate causal effects

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

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

linear:

Y = a X + b Z

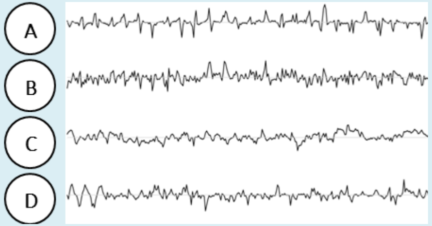


Part 2: Causal Discovery

Learning causal structures from data

Causal Discovery

Input: Time-series data



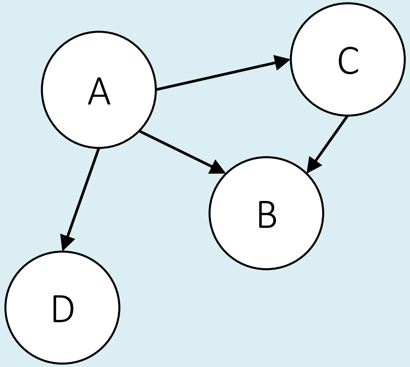
Causal Discovery

PCMCI Algorithm

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

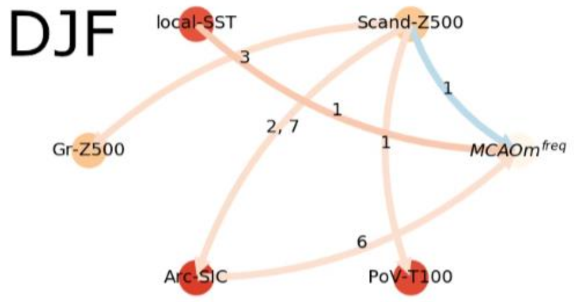
Can deal with auto-correlation, regime-dependence, instantaneous links, ...

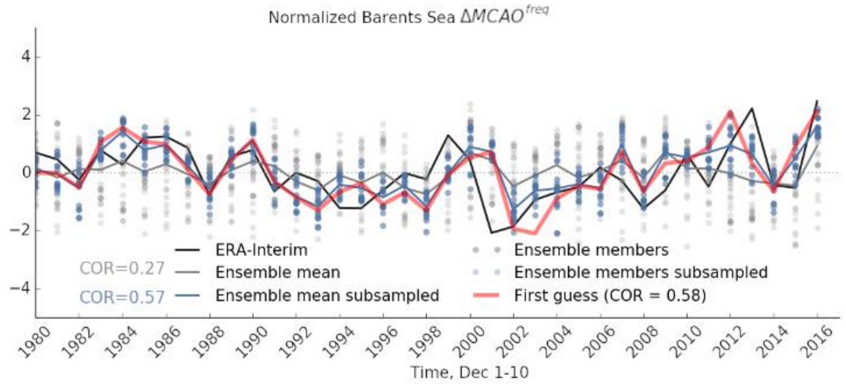
Output: Causal Structure



EXAMPLE: ENSEMBLE SUBSAMPLING

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

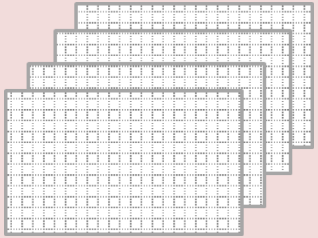




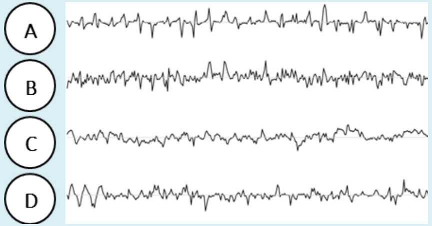
Polkova et al. (in revision, *QJRMS*)

Links to Machine Learning

Input: Gridded Climate data



Machine Learning



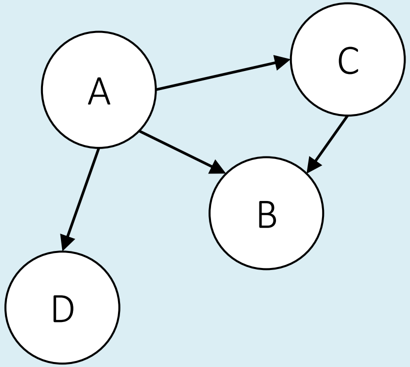
Causal Discovery

PCMCI Algorithm

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

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

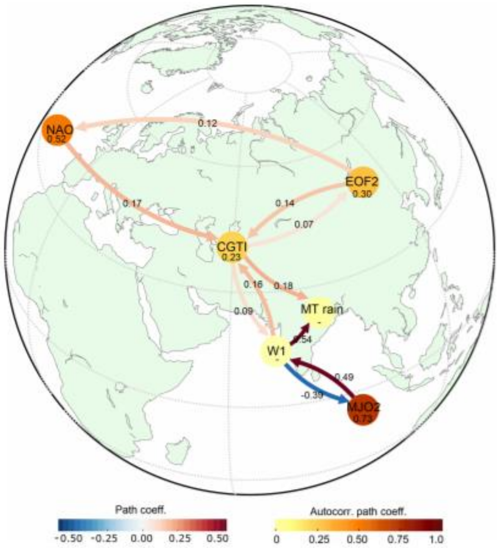
Output: Causal Structure



Applications (Domain Knowledge Required)

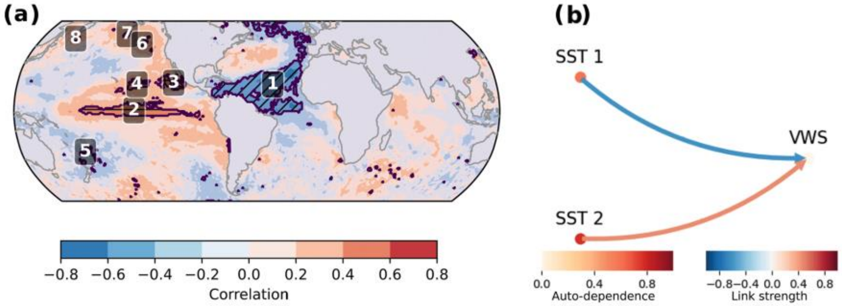
Indian Summer Monsoon

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



Hurricane Activity

Pfleiderer et al. (2020), *WCD*



Morocco Crop yield

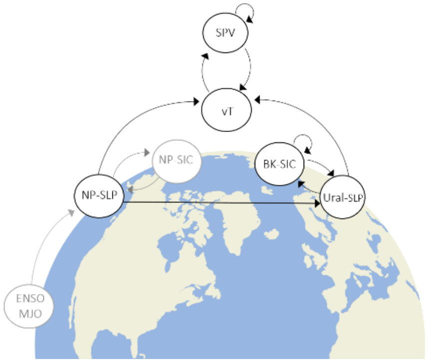
Lehmann et al. (2020), *GRL*

SUMMARY

Part 1: Causal Inference

Quantifying causal effects from data

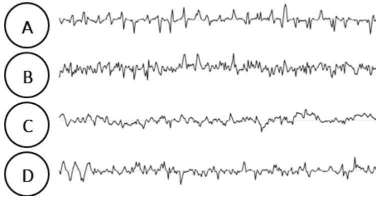
Domain expertise is explicitly built in and guides the data analysis



Easy and traceable approach to quantifyy teleconnection pathways (e.g. Using MLR)

Part 2: Causal Discovery

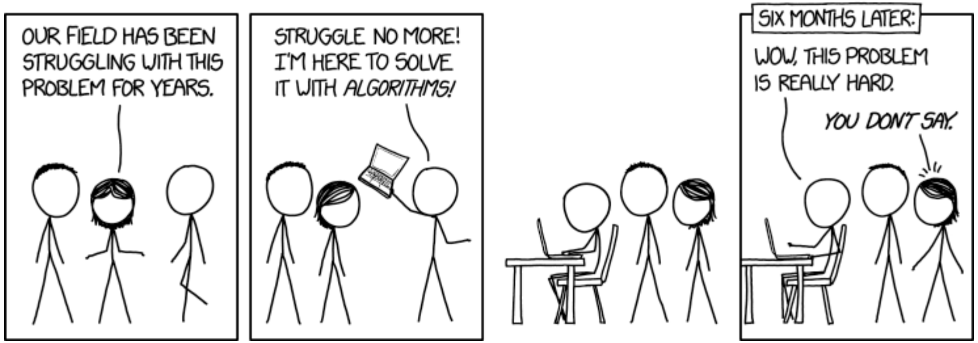
Learning causal structures from data



Causal discovery algorithms can help to close knowledge gaps and to test hypotheses

Outlook

Thank You!



Causal Networks as a framework to reconcile physics with data science.

Key to build fully interpretable ML models for high-stake decision making?

There is a rich literature on causal inference but so far only few applications to climate science problems

Combining causality with ML/DL, (e.g. to quantify causal effects) is a cutting edge research topic