Quantitative Causality Analysis, Causality-Aided Discovery, and Causal AI-Based Ocean/Atmosphere Prediction

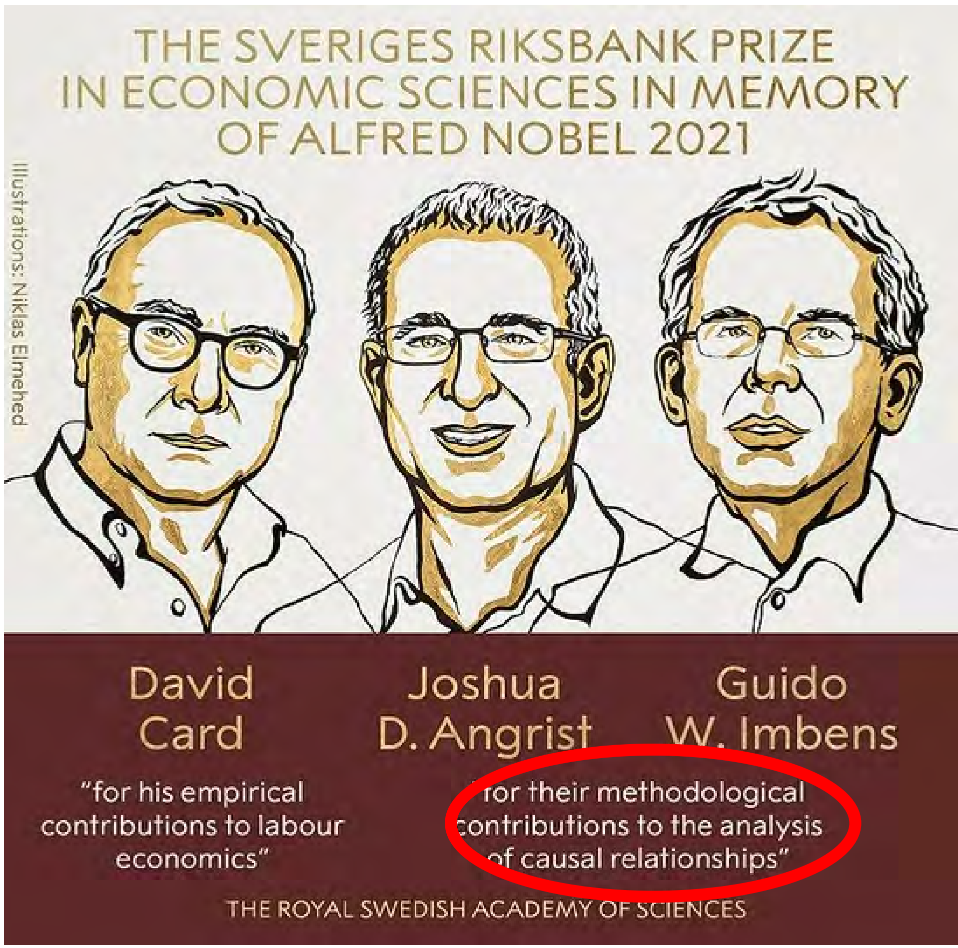
X. San Liang (梁湘三)

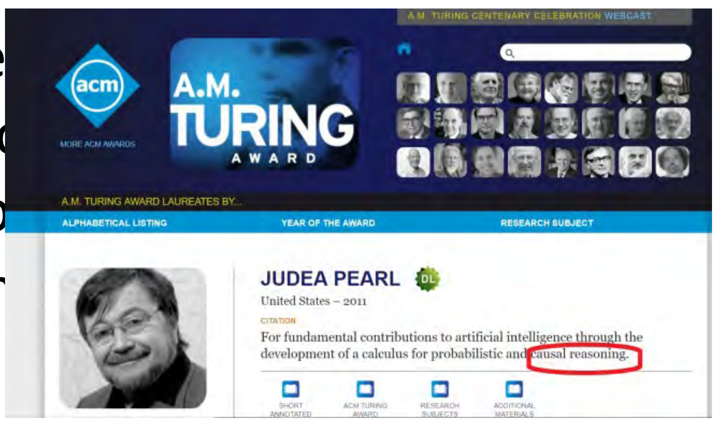
Dept. Atmos. & Oceanic Sci.
Fudan University (复旦大气海洋系)
The Artificial Intelligence Group
Southern Marine Laboratory
(南方实验室人工智能组)

6th NPOCE, CAS Institute of Oceanology
March 8, 2023

Introduction

Causality analysis is a problem lying at the heart of science
因果分析是科学研究的一个核心问题





Introduction

Outstanding problem in AI prediction　No causality!

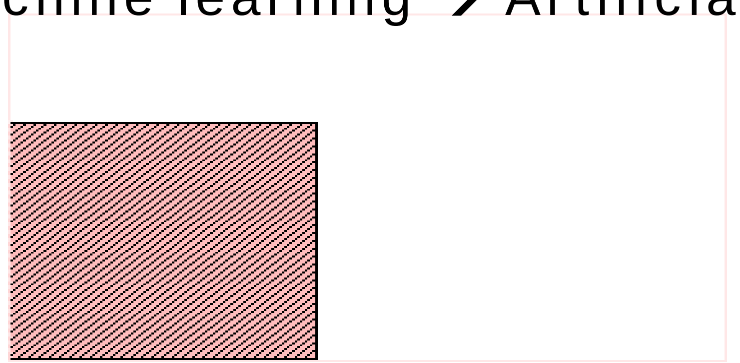
Pearl (2018), Bengio (2019), Schölkopf (2019)：因果分析对人工智能的重要作用

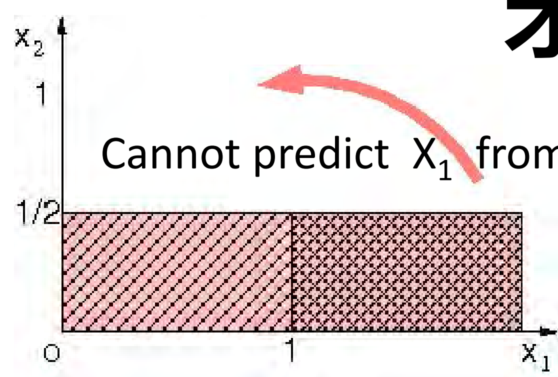
- “To achieve human level intelligence, learning machines need the guidance of a model of reality, similar to the ones used in **causal inference** tasks.” (Pearl, 2018)
- ”He believes it won’t realize its full potential, and won’t deliver a true AI revolution, until it can go beyond pattern recognition and learn more about **cause and effect**.” (Bengio, 2019)
- ”Finally, machine learning is also bad at thinking in the sense of Konrad Lorenz, i.e., acting in an imagined space. I will argue that **causality**, with its focus on modeling and reasoning about interventions, can make a substantial contribution towards understanding and resolving these issues and thus take the field to the next level.” (Schölkopf, 2019)

Causality analysis → Machine learning

Machine learning → Artificial intelligence

有因果
才能预报



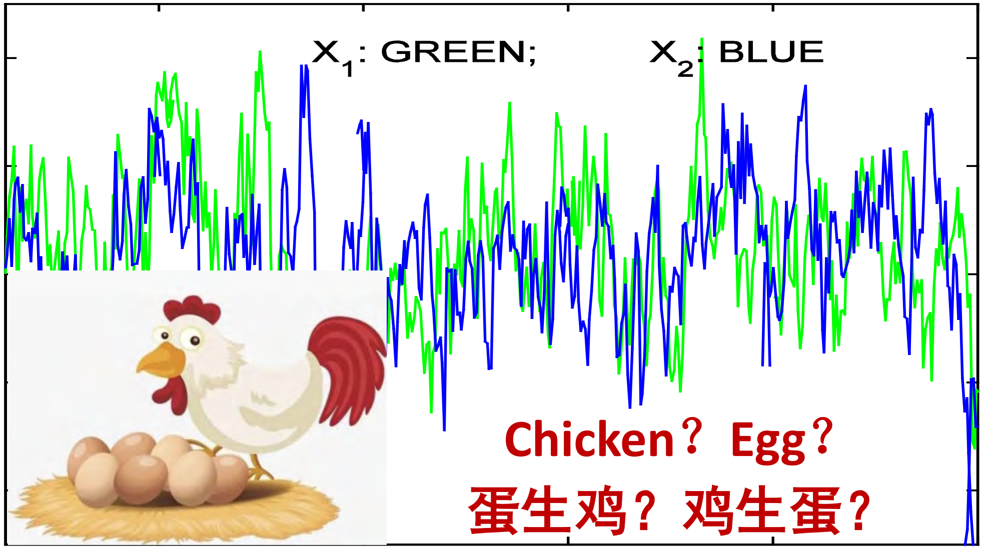


Outline

1. Causality Analysis *ab initio*
2. Causality-Aided Discovery
3. Causal AI
4. Causal AI-Based Ocean/Atmosphere Prediction

Applications

1. Causality Analysis *ab initio*



"The mathematization of causality is a relatively recent development, and has become increasingly important in data science and machine learning."
----Peters, Janzing, Schökopf (2017)

A continuing challenge since Granger ('69)
Keywords：
Math, Statitistcs, New field

"A real physical notion that can be derived *ab initio*"
----Liang (2016)

Real physical notion
Derivable from 1st principles (rather than axiomatic/empirical)
Born from atmosphere-ocean science

1. Causality Analysis *ab initio*

physicstoday

Home | Print Edition | Daily Edition | About | Jobs | Subscribe

Quantitative causality

A repurposed tool from information theory reveals whether two correlated behaviors share a causal link.

Steven K. Blau 十一月 2014

‹ PREVIOUS POST | PHYSICS UPDATE | NEXT POST ›

Cause and correlation are two different notions that are often confused. When phenomena *A* and *B* are causally related, their time evolutions are correlated. But correlation does not imply causality—an external agent such as an alarm clock, for example, can cause the correlated waking of two sleepers; the two isolated wakings, however, are not causally related. When the dynamics governing *A* and *B* are known, an information-theoretic notion called information flow rigorously determines the causal relations between *A* and *B*: If the information flow from *A* to *B* is zero, *A* has no effect on *B*; otherwise, *A* does affect *B*. Now X. San Liang of the Nanjing University of Information Science and Technology in China has shown how to obtain the information flow, not from *a priori* known dynamics but from correlations in the time-series graphs that detail the evolutions of *A* and *B*. He applied his result to a problem of practical interest for climate scientists—the relation between El Niño and the Indian Ocean Dipole (IOD), an aperiodic oscillation in sea surface temperature. Liang used correlations between time series—of sea surface temperatures in the Indian Ocean and of an index, called Niño4, that measures the overall strength of El Niño—to calculate the information flow from the IOD to El Niño shown in the figure. For a large swath of the northern Indian Ocean, the sign of the information flow is positive, which, according to information theory, means the IOD causes El Niño to be less predictable. That unusual causal link, suggests Liang, may be the reason climate scientists only recently recognized an influence of the IOD on El Niño. (X. S. Liang, *Phys. Rev. E.*, in press.)

Now X. San Liang ... in China...

Physics Update:

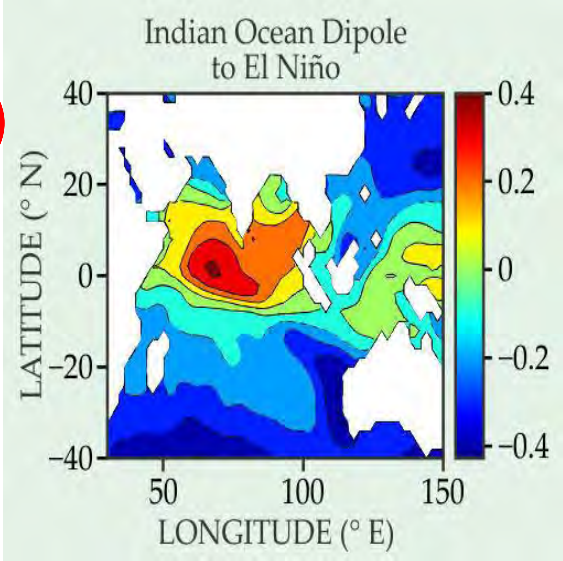
Learning the rules of kirigami

Precision plasma wakefield acceleration

Quantitative causality

Detecting shielded uranium in the field

Sound design for electric vehicles



http://scitation.aip.org/content/aip/magazine/physicstoday/news/10.1063/PT.5.7124

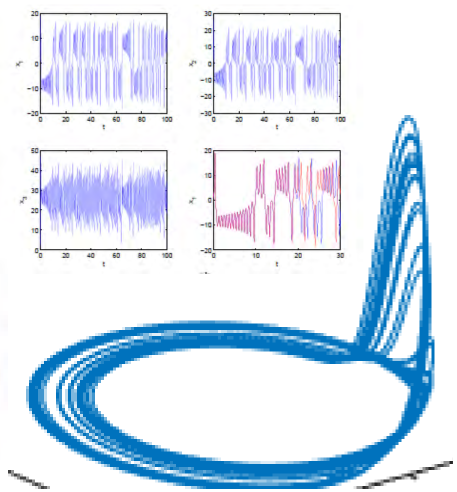
Liang, 2014: Unraveling the cause-effect relation between time series. PRE 90, 052150, 1-11

Liang, 2016: Information flow & causality as rigorous notions *ab initio*. PRE 94, 052201, 1-28

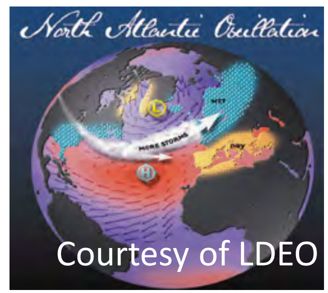
Liang, 2021: Normalized multivariate time series causality analysis and causal graph reconstruction. Entropy, 23, 679.

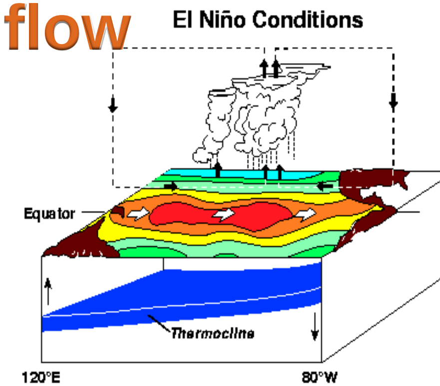
1. Causality Analysis *ab initio*

Atmosphere-Ocean system tends to be **CHAOTIC**



Information flow

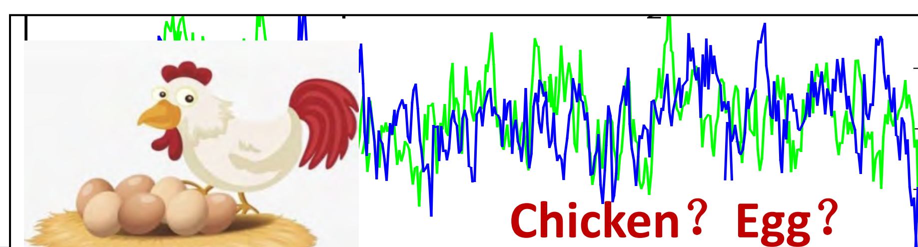




Uncertainty transference

➔ **Information flow**

➔ **Causality quantification**



1. Causality Analysis *ab initio*

Entropy **2013**, *15*, 327-360; doi:10.3390/e15010327

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Review

The Liang-Kleeman Information Flow: Theory and Applications

X. San Liang [1,2]

[1] School of Marine Sciences and School of Mathematics and Statistics, Nanjing University of Information Science and Technology (Nanjing Institute of Meteorology), 219 Ningliu Blvd, Nanjing 210044, China; E-Mail: sanliang@courant.nyu.edu; Tel.: +86-25-5869-5695; Fax: +86-25-5869-5698

[2] China Institute for Advanced Study, Central University of Finance and Economics, 39 South College Ave, Beijing 100081, China

Theorem 1 (Liang, 2016)

For an *n*-dimensional system

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}, t) + \mathbf{B}(\mathbf{X}, t)\dot{\mathbf{W}}$$

The information flow rate from X_2 to X_1 is :

$$T_{2\to1} = -E\left[\frac{1}{\rho_1}\int_{\Re^{n-2}} \frac{\partial F_1 \rho_{\backslash 2}}{\partial x_1} dx_3 \ldots dx_n\right] + \frac{1}{2} E\left[\int_{\Re^{n-2}} \frac{\partial^2 g_{11}\rho_{\backslash 2}}{\partial x_1^2} dx_3 \ldots dx_n\right]$$

where E is mathematical expectation.

$|T_{2\to1}| > 0$: X_2 is causal to X_1, otherwise not.

(Statistical significance test is needed in real applications.)

1. Causality Analysis *ab initio*

Theorem 2 (Principle of nil causality)

If the evolution of X_1 is independent of X_2, then $T_{2\to1}$=0.

前人总想在实际问题中尽量让他们的经验公式往这一准则上靠，但在我们的框架下，这是一条**定理!**

Theorem 3 （Liang 2018）

The so-obtained causality measure/information flow is **invariant** upon arbitrary nonlinear coordinate transformation.（对任意非线性坐标变换保持不变）

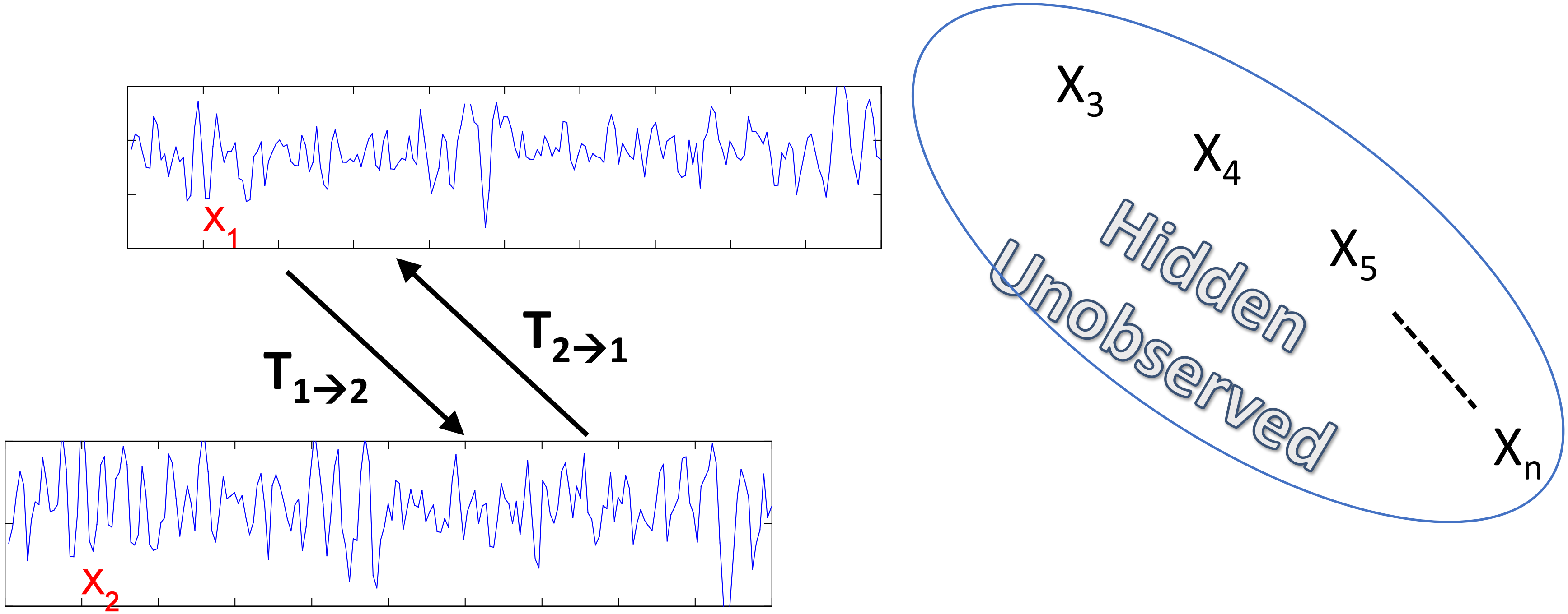
Intrinsic Physical Property!

内在的物理性质

1. Causality Analysis *ab initio*

Theorem 3 (Liang, 2018)

The above formula for information flow is **invariant** upon *any nonlinear* coordinate transformation of $(x_3, x_4, \dots, x_n)$.



Theorem 4 (Bivariate case, Liang, 2014)

For time series X_1 & X_2, the MLE of the causality is

$$\boxed{T_{2\to1}=\frac{C_{11}C_{12}C_{2,d1}-C_{12}^2C_{1,d1}}{C_{11}^2C_{22}-C_{11}C_{12}^2}}$$

Multivariate case（Liang 2021）

$$T_{2\to1}=\frac{1}{\det\mathbf{C}}\cdot\sum_{j=1}^{n}\Delta_{2j}\mathbf{C}_{j,d1}\cdot\frac{C_{12}}{C_{11}}$$

where

C_{ij} : sample covariance between X_i X_j

$C_{i,dj}$: sample covariance between X_i, $\dot{X}_j$ $\quad \dot{X}_i(n)=\frac{X_i(n+1)-X_i(n)}{\Delta t}$

$$\boxed{T_{2\to1}=\frac{r}{1-r^2}\left(r'_{2,d1}-rr'_{1,d1}\right)}$$

$$r=\frac{C_{12}}{\sqrt{C_{11}C_{22}}},\qquad r'_{i,dj}=\frac{C_{i,dj}}{\sqrt{C_{ii}C_{jj}}}$$

- **Causation ➔ Correlation**
- **Correlation $\nRightarrow$ Causation**

“有因果必有相关，有相关不一定有因果”

expressing the long-standing debate in philosophy since Berkeley (1710) using a concise math formula

1. Causality Analysis *ab initio*

Validations –

Validated with benchmark dynamical systems such as

baker transformation, Hénon map, Kaplan-Yorke map, stochastic potential flow…

Baker transformation

For $\Phi : \Omega \mapsto \Omega, \qquad \Omega = [0,1]$

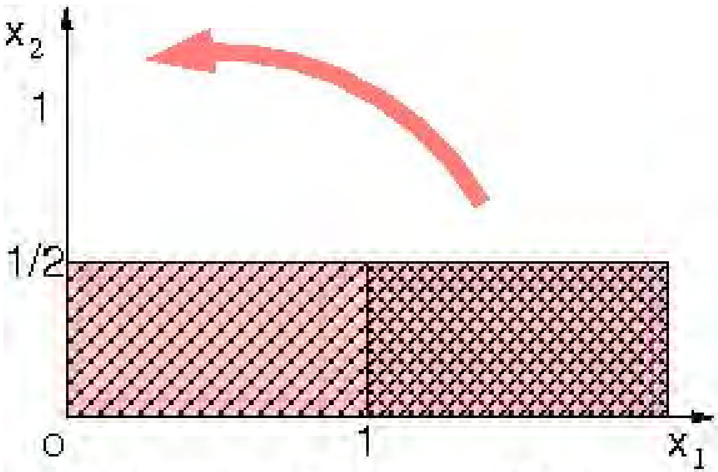
$$\Phi(x_1, x_2) = \begin{cases} (2x_1, \frac{x_2}{2}), & 0 \le x_1 \le \frac{1}{2}, 0 \le x_2 \le 1 \\ (2x_1 - 1, \frac{1}{2}x_2 + \frac{1}{2}, & \frac{1}{2} \le x_1 \le 1, 0 \le x_2 \le 1 \end{cases}$$

$$T_{2\to1} = 0$$

$$T_{1\to2} = R\left(0, \tfrac{1}{2}\right) + R\left(\tfrac{1}{2}, 1\right) > 0$$

where

$$R(a,b) = \int_a^b \rho_2 \cdot \log \frac{\int_0^1 \rho(\lambda, x_2) d\lambda}{\int_a^b \rho(\lambda, x_2) d\lambda} dx_2$$



(Lasota and Mackey, 1994)

1. Causality Analysis *ab initio*

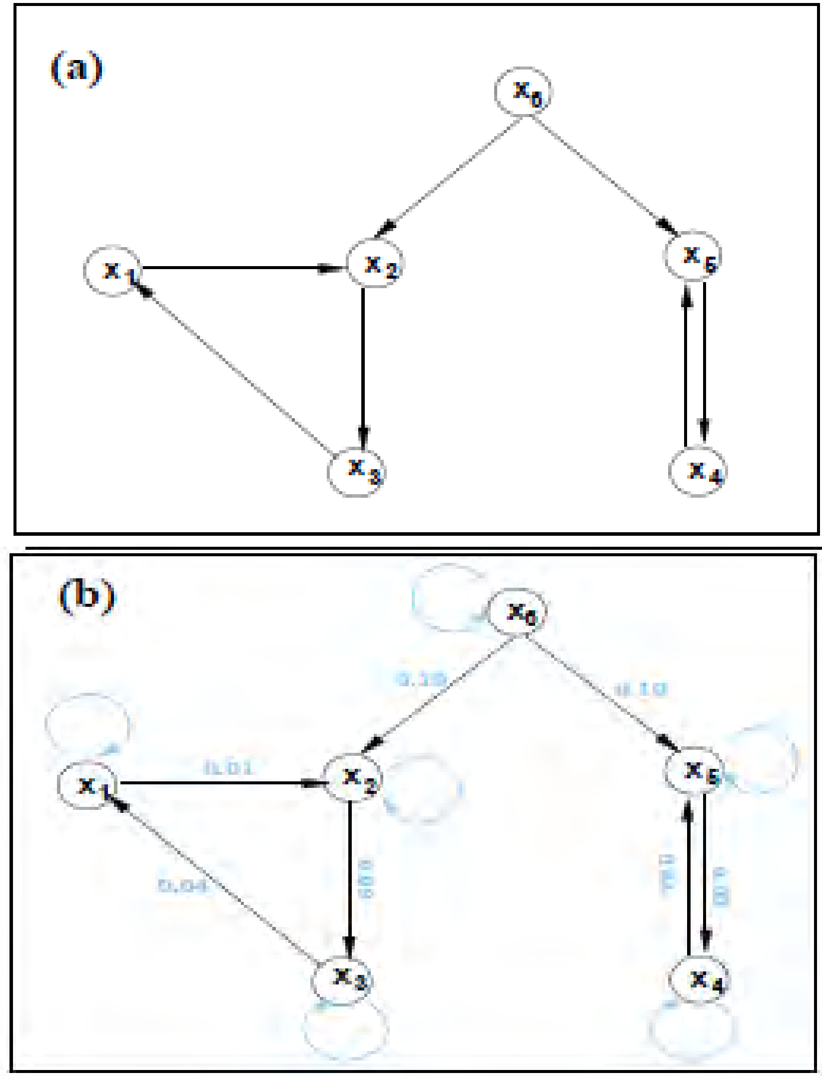
Validations – Series buried in noises (singal-to-noise ratio < 0.01)

$$\mathbf{X}(n+1) = \boldsymbol{\alpha} + \mathbf{A}\mathbf{X}(n) + \mathbf{B}\mathbf{e}(n+1)$$

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -0.6 & 0 & 0 & 0 \\ -0.5 & 0 & 0 & 0 & 0 & 0.8 \\ 0 & 0.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.7 & 0.4 & 0 \\ 0 & 0 & 0 & 0.2 & 0 & 0.7 \\ 0 & 0 & 0 & 0 & 0 & -0.5 \end{pmatrix},$$

$$\boldsymbol{\alpha} = (0.1, 0.7, 0.5, 0.2, 0.8, 0.3)^T,$$

$$\begin{pmatrix} \backslash & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & \backslash & 0.09 & 0.00 & 0.00 & 0.00 \\ 0.05 & 0.00 & \backslash & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & \backslash & 0.04 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.05 & \backslash & 0.00 \\ 0.00 & 0.19 & 0.00 & 0.00 & 0.18 & \backslash \end{pmatrix},$$



1. Causality Analysis *ab initio*

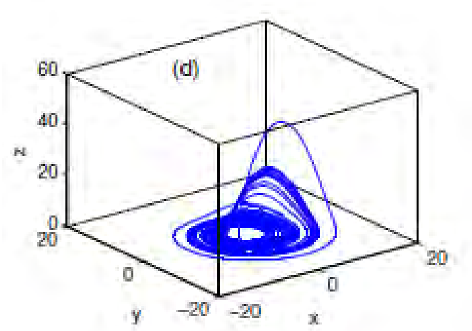
Validations – Series nearly synchronized

$$\begin{cases} dx_1/dt = -\omega_1 x_2 - x_3 \\ dx_2/dt = -\omega_1 x_1 + 0.15 x_2 \\ dx_3/dt = 0.2 + x_3[x_1 - 10] \end{cases}$$

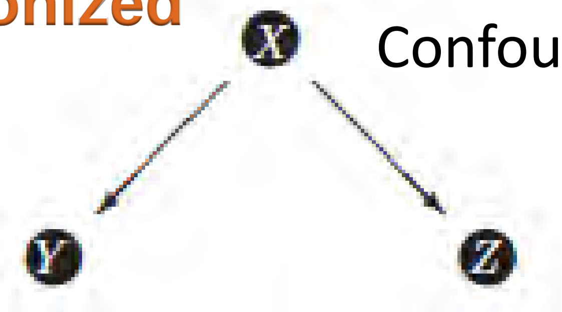
$$\begin{cases} dy_1/dt = -\omega_2 y_2 - y_3 + \varepsilon[x_1 - y_1] \\ dy_2/dt = -\omega_1 y_1 + 0.15 y_2 \\ dy_3/dt = 0.2 + y_3[y_1 - 10] \end{cases}$$

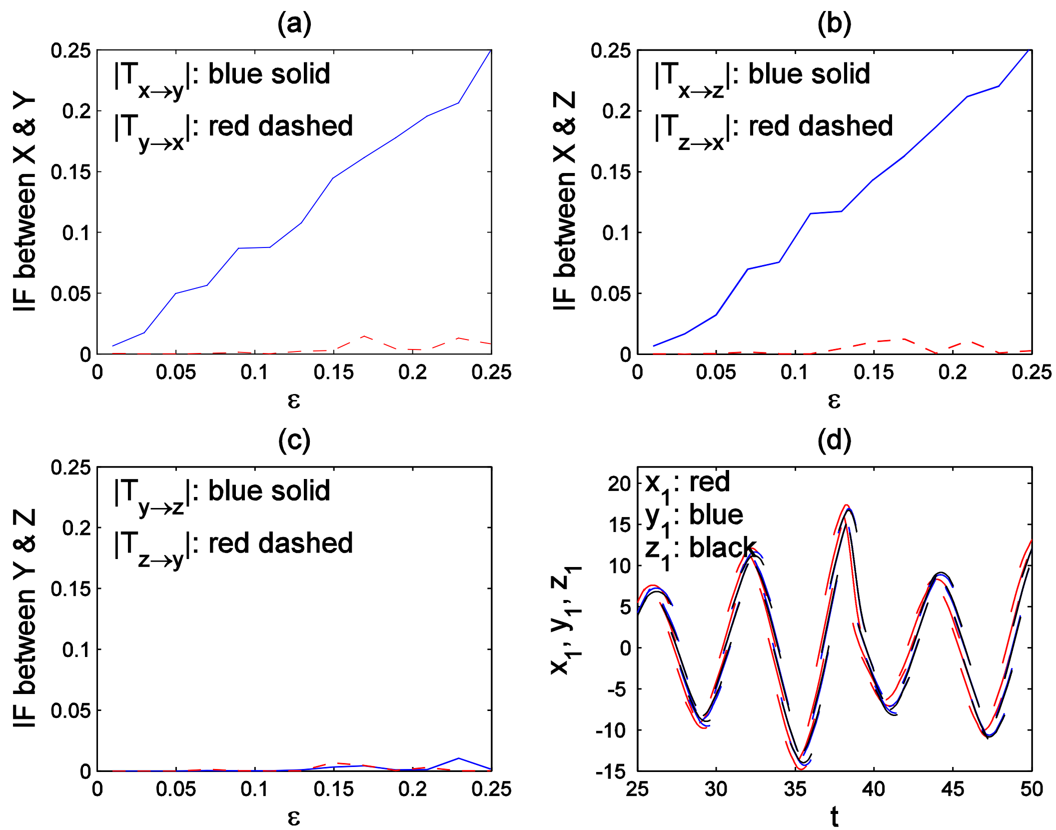
$$\begin{cases} dz_1/dt = -\omega_2 z_2 - z_3 + \varepsilon[x_1 - z_1] \\ dz_2/dt = -\omega_1 z_1 + 0.15 z_2 \\ dz_3/dt = 0.2 + z_3[z_1 - 10] \end{cases}$$

Paluš et al. (2018)



Confounding





Performance in computation

30 series, each with 100,000 steps

IF: <0.3s for one relation; < 1 second for all relations

gctest: 1200s for one relation; ~17 days for all relations

1. Causality Analysis *ab initio*

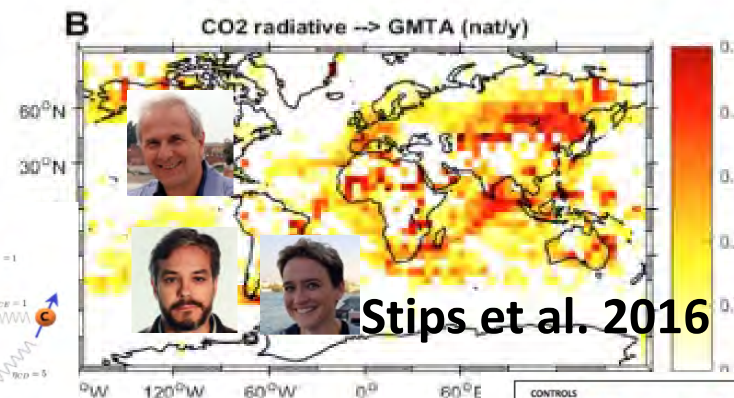
◆ Applications

- CO2 vs. global warming
- Hurricane prediction
- El Niño prediction
- Pacific-N. Amer. Pattern
- Quantum mechanics
- Wall turbulence
- Soil moisture-precipitation
- Pollutant source tracing
- Neuroscience
- Financial economics

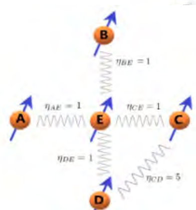
Bernanke Paulson Geithner

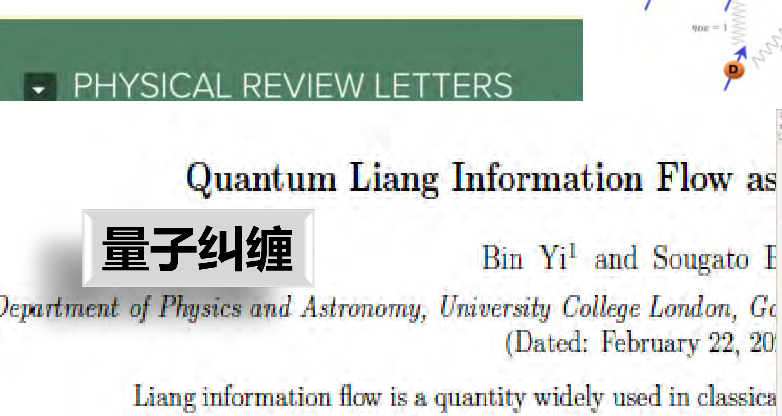
Emergency loan?

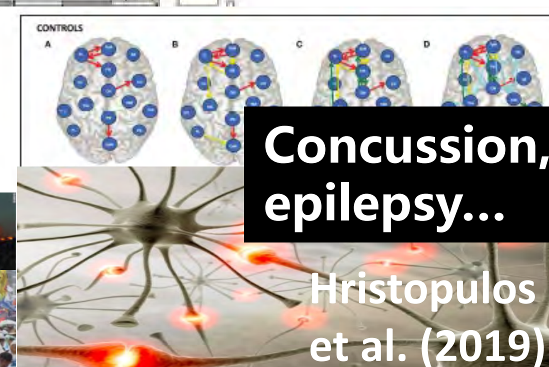
No!



CO_2 vs. Global Warming: Causal cycle!







1. Causality Analysis *ab initio*

最近代表性应用

（1）量子信息

Yi & Bose, 2022: Quantum Liang information flow as causation quantifier. ***Phys. Rev. Lett.***, 129, 020501.

（2）神经科学

Cong et al., 2023: Altered default mode network causal connectivity patterns in autism spectrum disorder revealed by Liang information flow analysis. ***Human Brain Mapping*** DOI: 10.1002/hbm.26209

获重点基金项目支持

资助类别：重点项目
项目名称：因果分析、大气可预报性传递及其在台风突然转折路径预报中的应用
负责人：梁湘三

Science Europe—国家自然科学基金委员会 Joint Policy Workshop

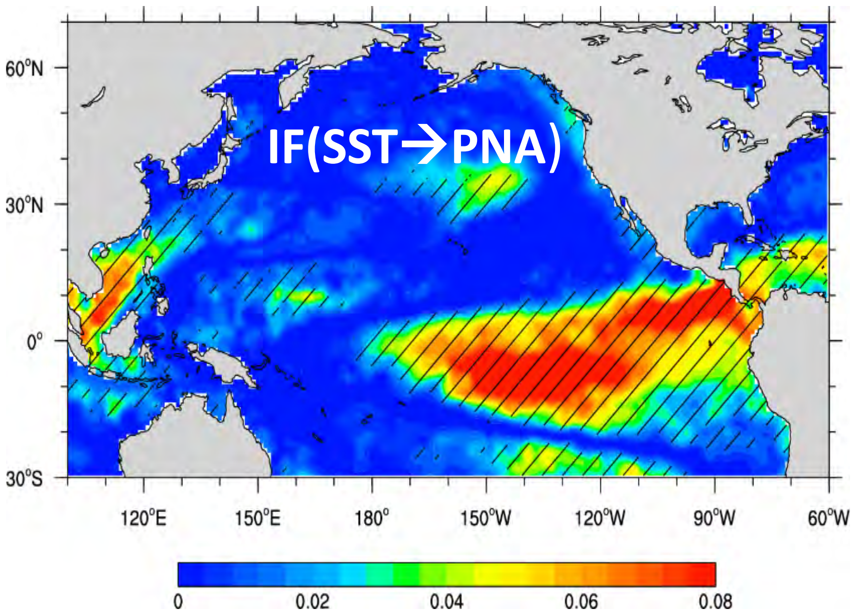
作为亮点工作之一出现在"New Theory"中

Outline

1. Causality Analysis
2. Causality-Aided Discovery
3. Causal AI
4. Causal AI-Based Ocean/Atmosphere Prediction

2. Causal Discovery – PNA vs. SCS

因果分析 ➔ 南海科学新发现



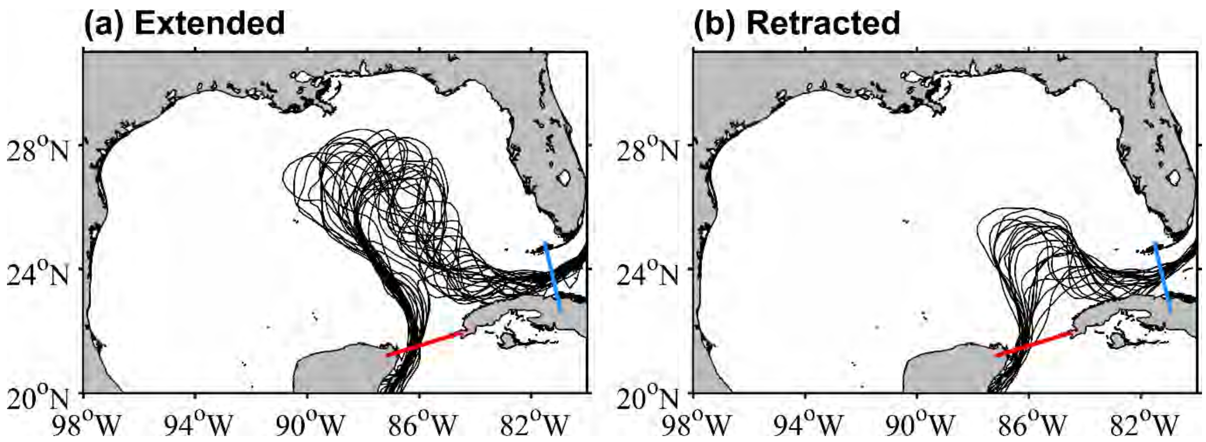
South China Sea affects the weather in N. America!
(Zhang & Liang 2022; 2023)

About PNA

- A dominant low-fequency modes of the boreal winter midlatitude atmosphere (e.g., Wallace & Gutzler 1981)
- Extensively investigated (e.g., Hoskins & Karoly 1981). Major mechanism: ENSO

1. Zhang, Yinchen, **X. San Liang*,** 2022: The causal role of South China Sea on the Pacific-North American teleconnection pattern. *Climate Dynamics*. 59, 1815-1832.
2. Zhang, Yinchen, **X. San Liang***, 2023: The distinct PNA pattern induced by the South China Sea. *Climate Dynamics*. DOI: 10.1007/s00382-022-06607-4

2. Causal Discovery – GoM Loop Current



Causality analysis indicates that

- Both of the transports thru YC & FS are causal to the LC path variation.
- YC affects LC directly.
- YC may also affect LC indirectly via FS.

Yang et al., Causal Relations Between the Loop Current Penetration and the Inflow/Outflow Conditions Inferred with a Rigorous Quantitative Causality. *Deep-Sea Res.* (Revised)

2. Causal Discovery – PM2.5 Sourcing

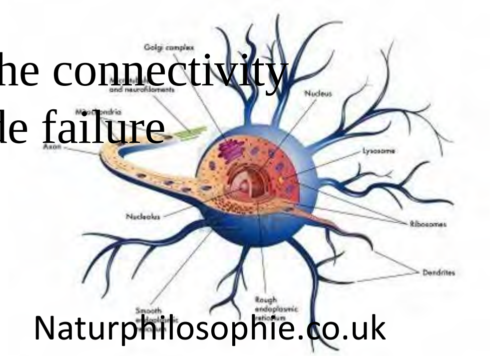
董晟宏等，2019

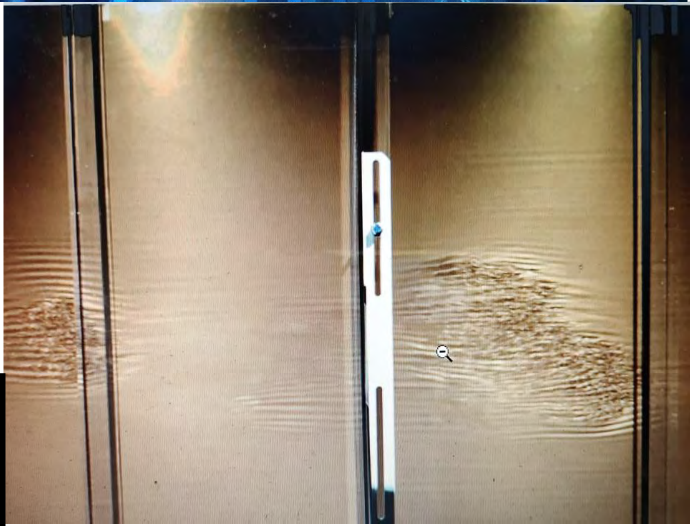
赵远冰、梁湘三，2019

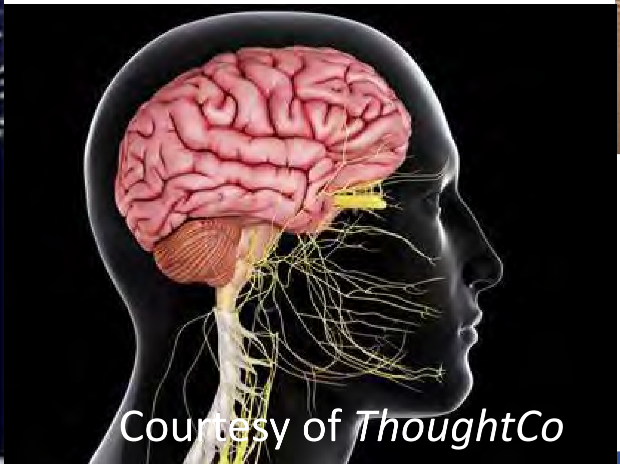
2018年12月以来，韩国首都遭受重度污染。“中国雾霾来袭，微颗粒物浓度上涨”、“又是中国雾霾，出门记得戴口罩”……

2. Causal Discovery – Network Dynamics

- Complex networks provide a framework for studying many social, biological and engineering systems
- A goal is to understand how individuals collaborate to produce the collective behavior
- One question to ask is whether the connectivity of a network is robust to local node failure







Where is most sensitive?

2. Causal Discovery – Network Dynamics

Theorem 6

Under a linear assumption, the maximum likelihood estimator of the IF from node 1 to the network is $\hat{T}_{1\to 2..n} = Tr_{\natural}\left[\tilde{\mathbf{C}}^{-1}(\hat{\mathbf{A}}\mathbf{C})^T\right] - Tr_{\natural}\left[\hat{\mathbf{A}}\right]$

$\mathbf{C}$ is the covariance matrix

$$\tilde{\mathbf{C}} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & c_{22} & c_{23} & \cdots & c_{2n} \\ 0 & c_{23} & c_{33} & \ldots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & c_{2n} & c_{3n} & \ldots & c_{nn} \end{pmatrix}. \qquad \begin{pmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{in} \end{pmatrix} = \mathbf{C}^{-1} \begin{pmatrix} c_{1,di} \\ c_{2,di} \\ \vdots \\ c_{n,di} \end{pmatrix}$$

$c_{j,di}$ is the covariance between the series $\{x_j(k)\}$ and $\{(x_i(k+1) - x_i(k))/\Delta t\}$

$Tr_{\natural}$ means the trace of a matrix with the first term removed

Corollary $T_{1\to 2\ldots n} \neq T_{1\to 2} + T_{1\to 3} + \ldots + T_{1\to d}$

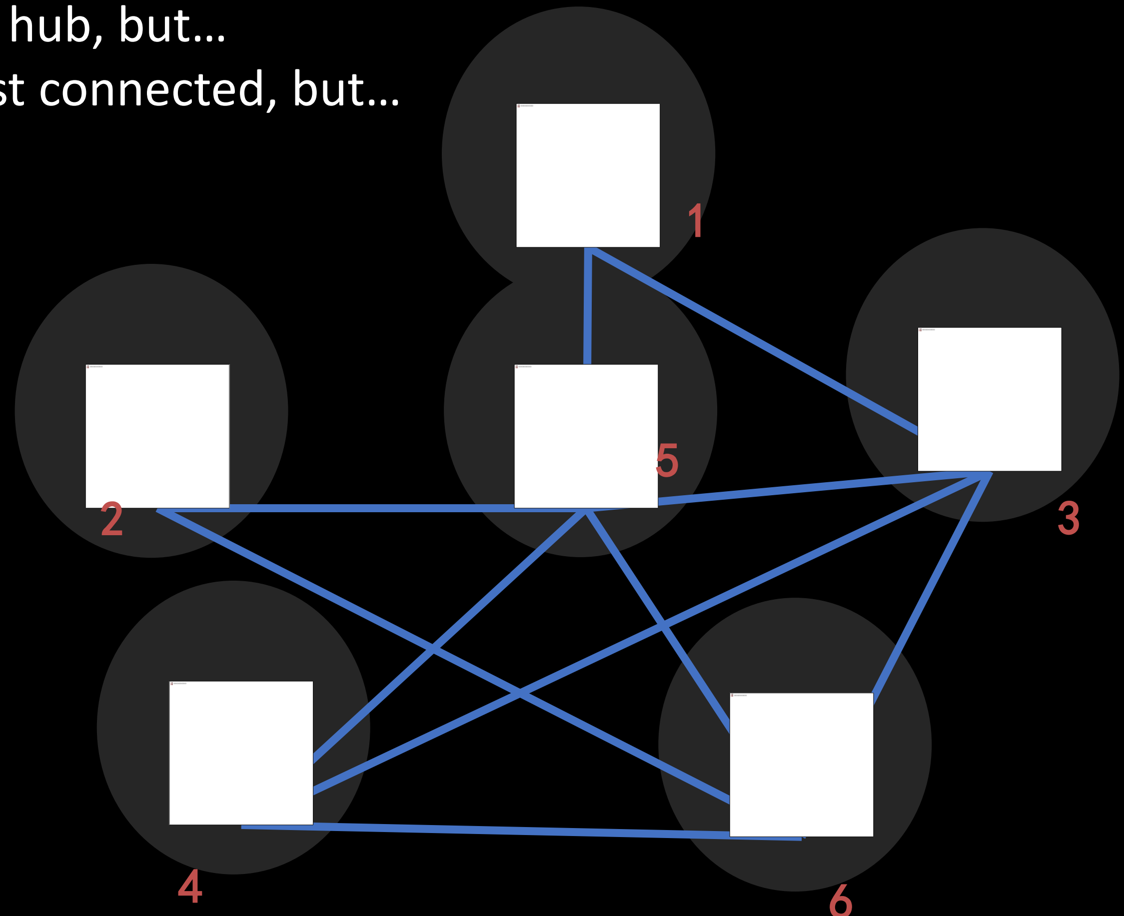
Collective phenomenon:
A group is not the addition of the individual members

"1+1 > 2"

$\hat{T}_{1\to\text{network}}$	$\hat{T}_{2\to\text{network}}$	$\hat{T}_{3\to\text{network}}$	$\hat{T}_{4\to\text{network}}$	$\hat{T}_{5\to\text{network}}$	$\hat{T}_{6\to\text{network}}$
0.35	4.11	2.06	1.95	2.48	0.74

Node #5 is the hub, but…

Node #2 is least connected, but…



restart

Node 1

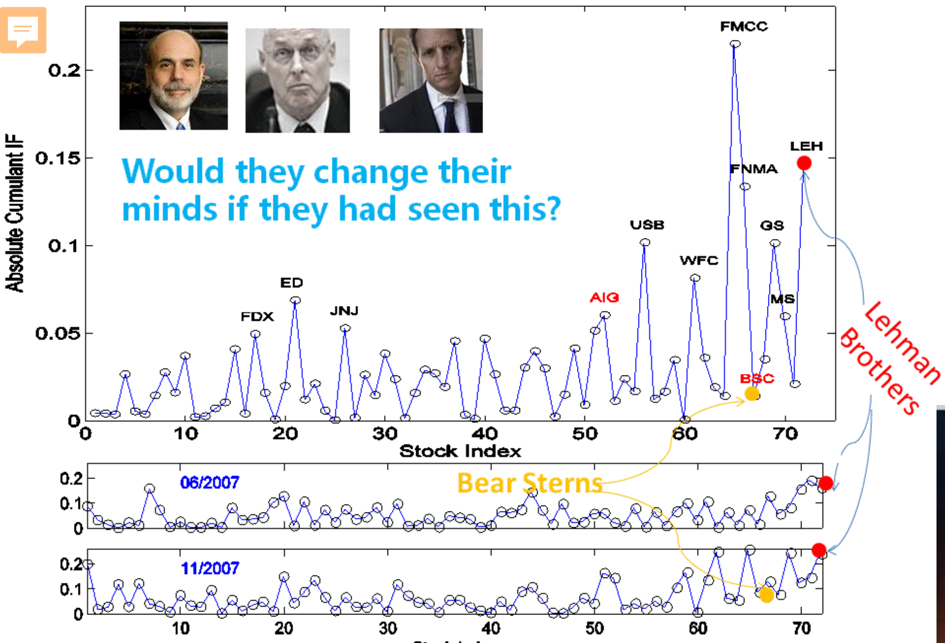
Node 2

Node 3

Node 4

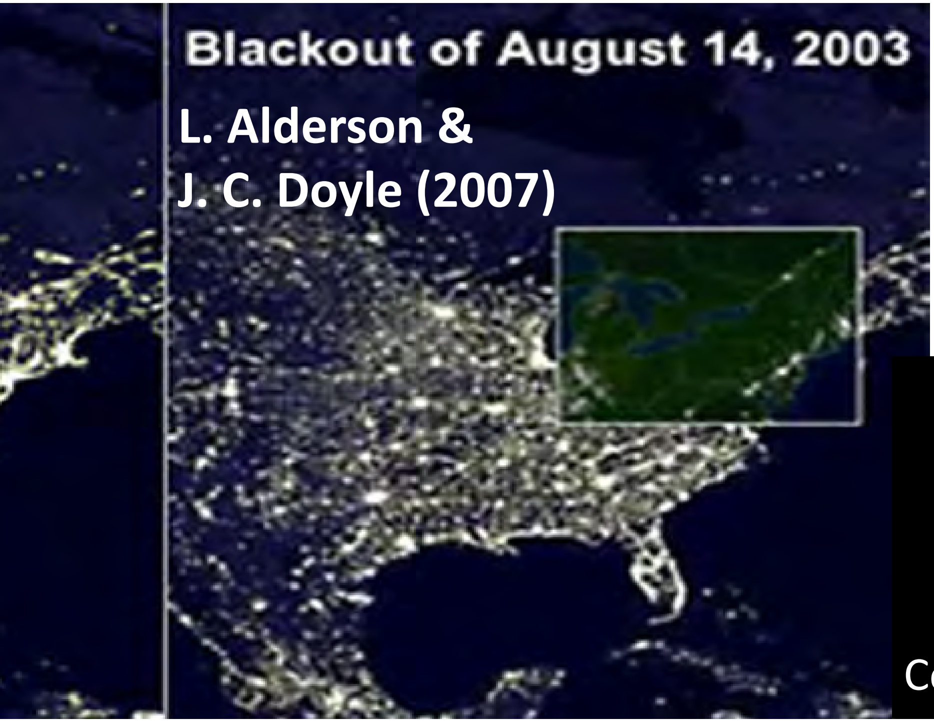
Node 5

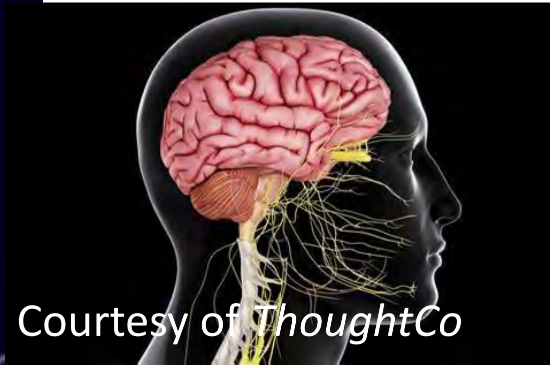
Node 6



Liang (2021)

X.S. Liang: Causality and information flow ab initio and their applications





Where is most sensitive?

Outline

1. Causality Analysis
2. Causality-Aided Discovery
3. Causal AI
4. Causal AI-Based Ocean/Atmosphere Prediction

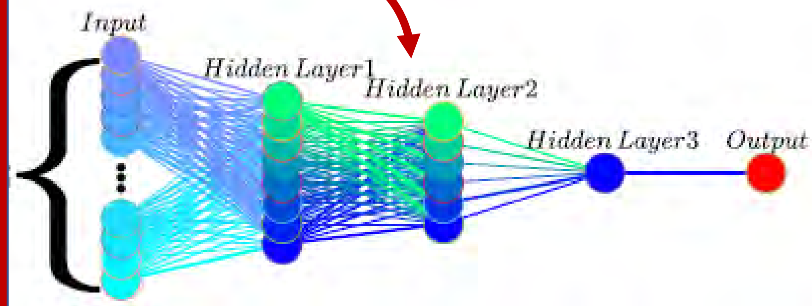
2. Causal AI

- **Established the framework for causal deep learning**
- **Fulfilled a preliminary interpretable causal AI algorithm**

Quantitative causality analysis *ab initio*

Modify the nodal weights to remove spurious association

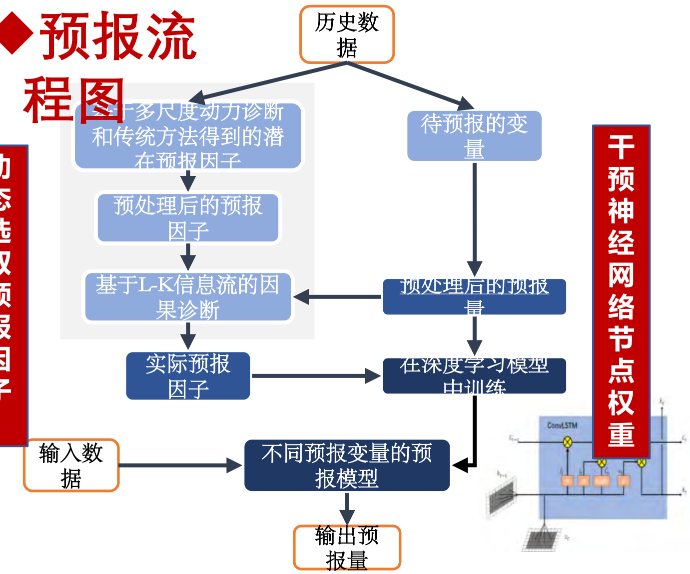
Dynamically select predictors



$y + ce_1 + de_2$

ious!

◆预报流程图



动态选取预报因子

干预神经网络节点权重

2. Causal AI

Data-driven Causal Discovery for Constructing Fuzzy Cognitive Maps: An Industrial Case Study

Marios Tyrovolas, *Student Member, IEEE*, X. San Liang,
and Chrysostomos Stylios, *Senior Member, IEEE*

Abstract—Artificial intelligence (AI) is one of the most disruptive digital enablers of the Industry 4.0 era, enabling the development of novel techniques for various industrial tasks such as anomaly detection. Despite the AI models' effectiveness, their black-box nature makes their decisions interpretation difficult, hindering their deployment in essential applications where humans make the final judgment. As a result, it is considered necessary to develop more interpretable AI models. In this context, there has recently been a surge in research efforts focusing on developing post-hoc methods for explaining AI models, which are problematic because the given explanations are prone to errors. For this reason, the scientific community has turned to developing intrinsic interpretable models such as Fuzzy Cognitive Maps (FCMs) whose decisions can be explained without additional techniques. Nonetheless, current FCM-based models are prone to capture spurious correlations presented in data, resulting in suboptimal predictive and explanatory performance. This article introduces a novel approach for constructing FCMs based on a causal inference tool, the Liang-Kleeman Information Flow (L-K IF) analysis, that identifies the actual causal relationships and rules out spurious correlations. Numerical simulations were conducted to compare the proposed model against state-of-

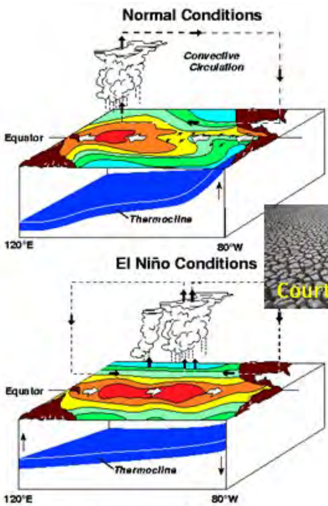
TABLE VI: ACCURACY-INTERPRETABILITY TRADE OFF AND AGGREGATE POWER FOR EACH CLASSIFIER

Model	Trade-off	Average Success
EBM	0.020614	1.959736
IF-FCM	0.052[illegible]3	1.69723
LR	0.046599	1.695915
LTCN	0.208946	1.694406
GNB	0.122065	1.538987
FCMMC	0.353438	1.515836
FCMB	0.554197	1.328987
FCM-A	-	-
FCN-FW		

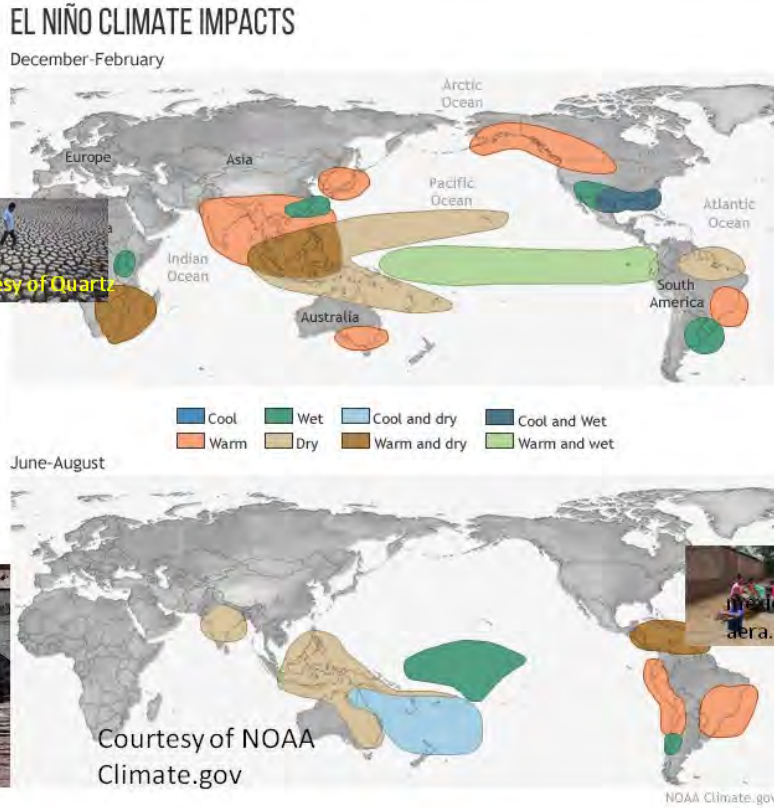
Outline

1. Causality Analysis
2. Causality-Aided Discovery
3. Causal AI
4. Causal AI-Based Ocean/Atmosphere Prediction

Forecasting – El Niño Modoki



Courtesy of NOAA



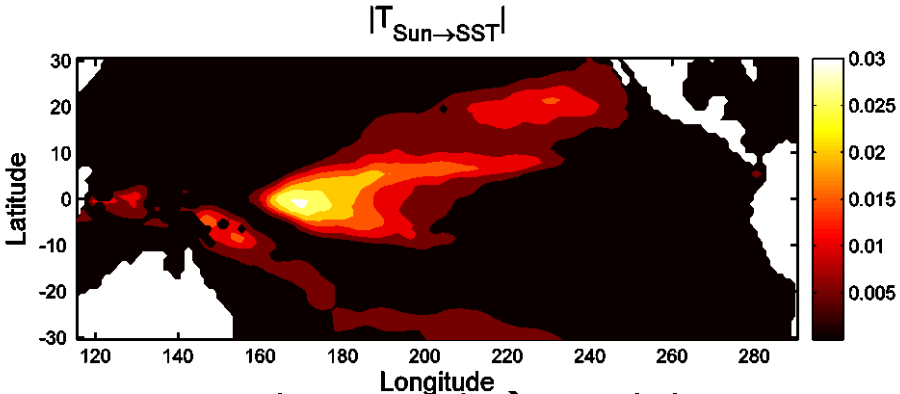
Courtesy of NOAA
Climate.gov

X.S. Liang: ENSO Modoki thus far can be mostly predicted 10 years ahead of time

El Niño Modoki
(Ashok et al., 2007)
Central-Pacific (CP)-type El Niño
(Yu & Kao, 2007)
Date Line El Niño
(Larkin & Harrison, 2005)
Warm Pool El Niño
(Kug, Jin, & An, 2009)
...

Present coupled models have **lower forecast skill** for El Niño Modoki (Ashok & Yamagata, 2009)

4. Forecasting – El Niño Modoki



SSN(n-45 yrs) → SST(n)

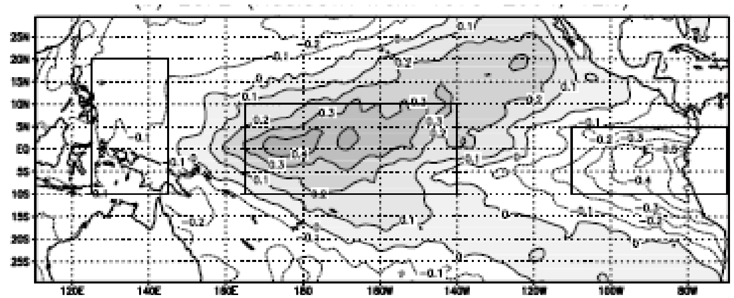
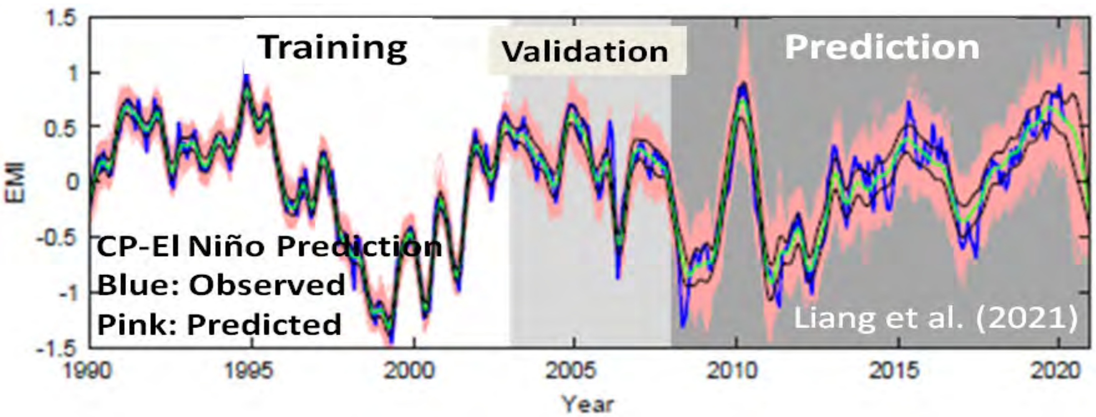


Figure 1. Left: The El Niño Modoki as shown in the Fig. 2b of Ashok et al. (2007, J. Geophys. Res. 112, C11007). Right The causal pattern from solar activity to the sea surface temperature as revealed by our causal analysis/attribution technique. Notice the high similarity between the two.

Decadal Prediction of the El Nino Modoki

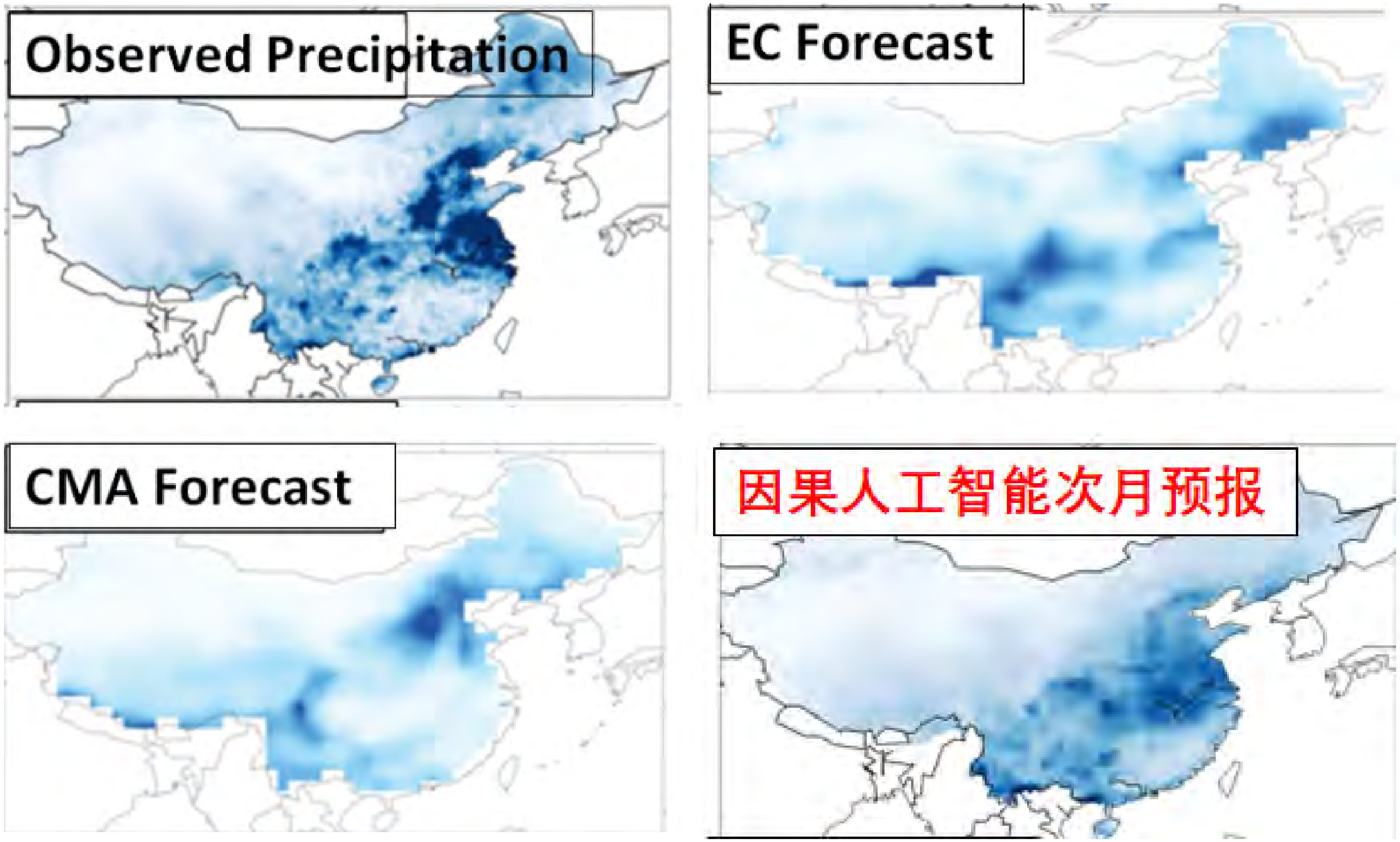


"absolutely astounding"
"breathtaking and counterintuitive"
"fairly striking"

Liang, X. San*, Fen Xu, Yineng Rong, Renhe Zhang, Xu Tang, Feng Zhang, 2021: El Niño Modoki can be mostly predicted more than 10 years ahead of time. *Sci. Rep.* 11:17860

4. Forecasting – Precipitation

Forecasts of monthly precipitation in China
e.g., precipitation July 2021



4. Forecasting – Precipitation

2022年全国次月降水预测竞赛（6-10月），战胜了传统的资深团队，获得了优秀奖

预测技术方法名称:

基于梁氏信息流的因果人工智能对次月降水的预报

团队人员: **容逸能、马继望、李渊、杨洋、徐芬、马伟翔、赵宇慧、张殷宸、张蓝心、郭子彧、付冠琦、何蔚邦**

申报单位: **南方海洋科学与工程广东省实验室（珠海）**

4. Forecasting – Typhoon

◆ 台风：全球十大自然灾害之首

台风“山竹”

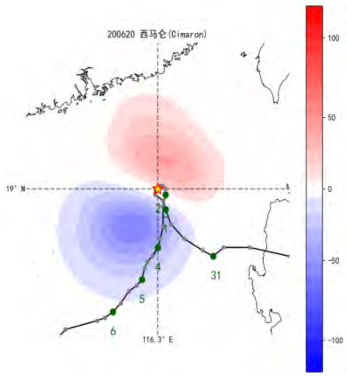
台风“山竹”造成广东、海南等5省近300万人受灾，5人死亡，1人失踪，160.1万人紧急避险转移和安置；1200余间房屋倒塌，800余间严重损坏，近3500间一般损坏；农作物受灾面积174.4千公顷，其中绝收3.3千公顷；直接经济损失52亿元。

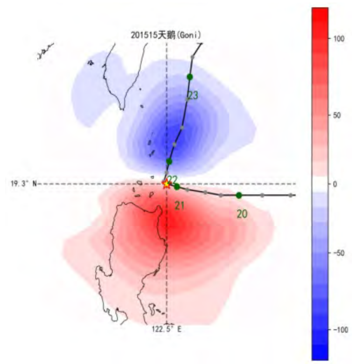
台风“威马逊”

2014 年超强台风威马逊登陆我国就导致88 人死亡失踪，1189.9 万人受灾，直接经济损失446.5亿元。

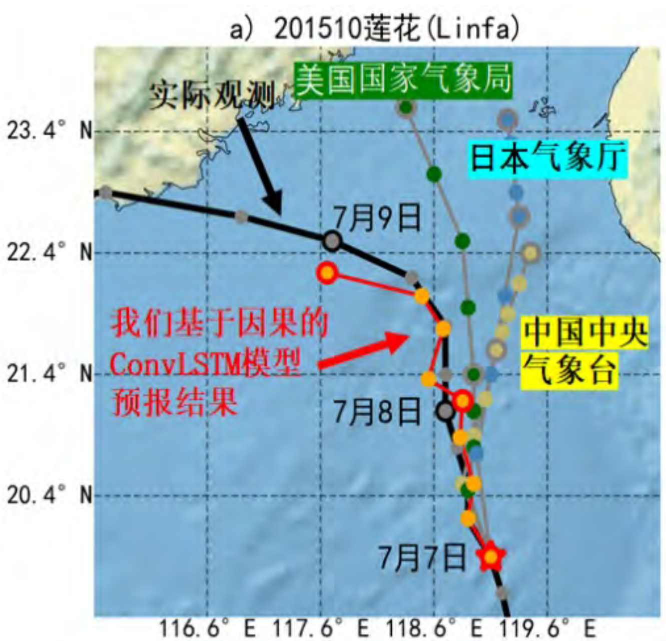
频繁的台风灾害对我国，特别是对南部沿海各省的社会经济造成巨大损失。对台风活动的准确预测对我国的减灾防灾有着重大意义

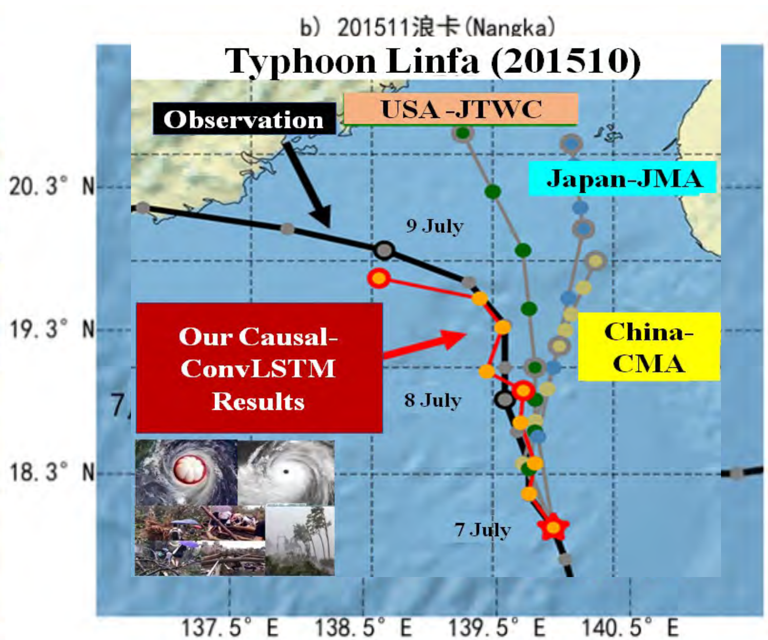
台风突然转折路径预报





- 作为十大自然灾害之首，台风每年带来巨大的经济和人员损失。
- 我国投入巨大。如建立了上海台风所、亚太台风研究中心。
- 至今未被攻克的难题：台风突然转折的路径的预报。





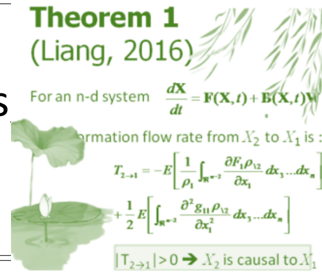
容逸能博士论文：《基于梁氏-克里曼信息流的人工智能方法在台风路径预报中的应用》

Summary

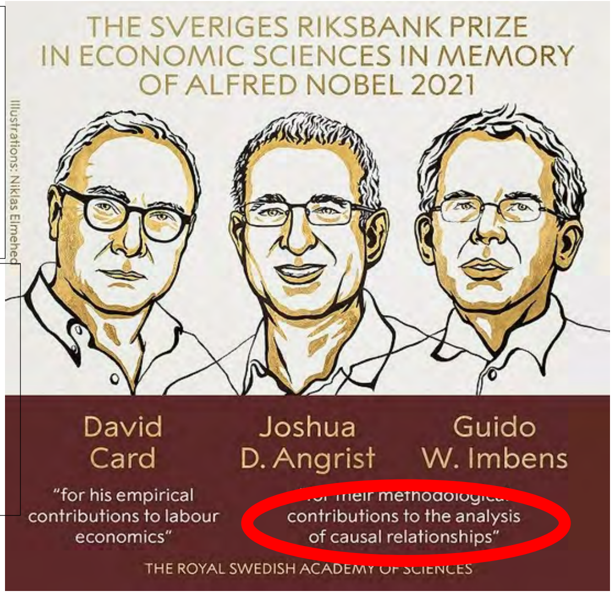
Causality analysis is a fundamental problem lying at the heart of science

Causality is "**a real physical notion that can be derived *ab initio***" ----Liang (2016, PRE 94, 052201)

- A real notion in physics
- Can be rigorously derived from first principles rather than axiomatically proposed as ansatz
- Born from atmosphere-ocean science



- Quantified causality from X_2 to X_1, $T_{2\to1}$, is explicitly obtained
- $T_{2\to1}$ invariant upon arbitrary coordinate transformation
- In the linear limit, reduces to a very concise formula. A corollary is: *causation implies correlation, but not vice versa*.



For two time series X_1 & X_2, the maximum likelihood estimator of the causality from X_2 *to* X_1 is (Liang 2014):

$$\boxed{T_{2\to1} = \frac{C_{11}C_{12}C_{2,d1} - C_{12}^2 C_{1,d1}}{C_{11}^2 C_{22} - C_{11}C_{12}^2}}$$

In multivariate case (Liang, 2022),

$$T_{2\to1} = \frac{1}{\det \mathbf{C}} \cdot \sum_{j=1}^{n} \Delta_{2j} \mathbf{C}_{j,d1} \cdot \frac{C_{12}}{C_{11}}$$

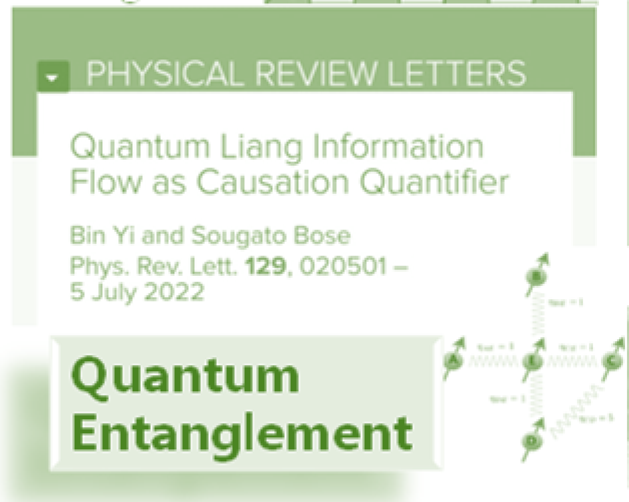
where

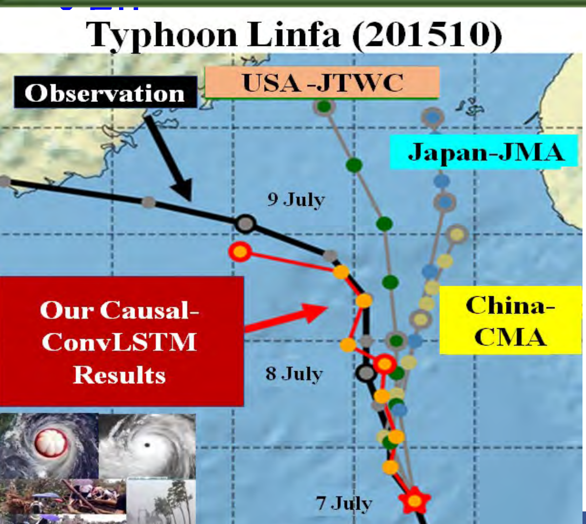
C_{ij} : sample covariance between X_i and X_j

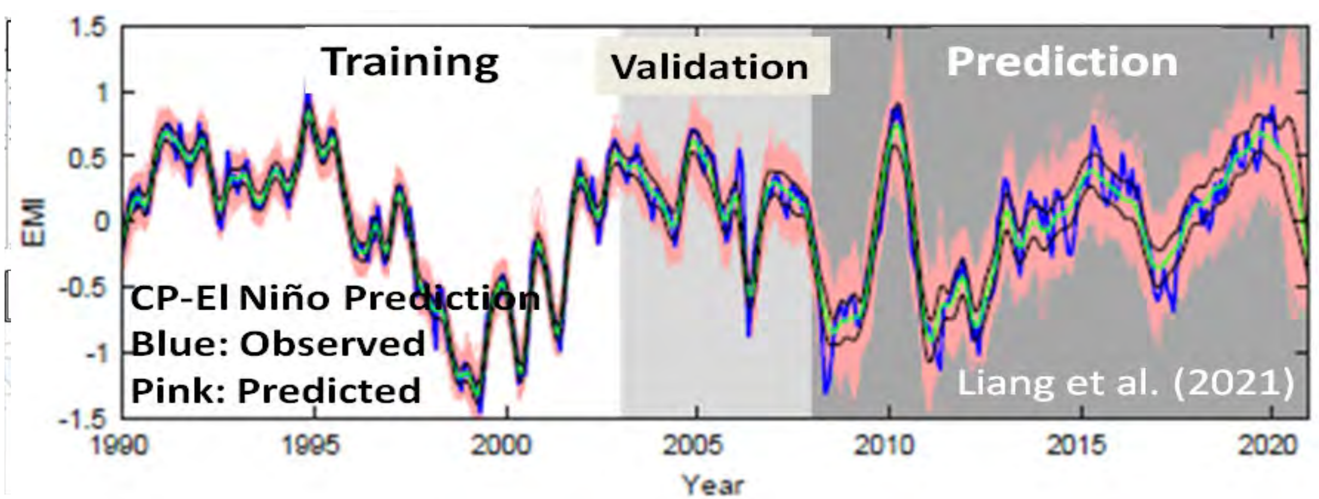
$C_{i,dj}$: sample covariance between X_i & $\dot{X}_j$, where $\dot{X}_i(n) = \frac{X_i(n+1) - X_i(n)}{\Delta t}$

Summary

- CO2 vs. global warming
- Typhoon prediction
- El Niño prediction
- Pacific-N. Amer. Pattern
- Quantum mechanics
- Wall turbulence
- Soil moisture-precipitation
- Pollutant source tracing
- Neuroscience
- Financial economics
- Deep learning (causal AI)







Take-Home Points

- 因果性是真实的物理概念，能从第一性原理推导出来

- 因果分析已定量化，并得到显示解

- 已广泛地应用于全球变化、气象、金融、智能预报、神经科学、量子物理等研究

physicstoday

Home | Print Edition | Daily Edition | About | Jobs

Quantitative causality

A repurposed tool from information theory reveals whether two correlated behaviors share a caus

Steven K. Blau 十一月 2014

PREVIOUS POST | PHYSICS UPDATE | NEXT POST

Cause and correlation are two different notions that are often confused. When phenomena *A* and *B* are causally related, their time evolutions are correlated. But correlation does not imply causality—an external agent such as an alarm clock, for example, can cause the correlated waking of two sleepers; the two isolated wakings, however, are not causally related. When the dynamics governing *A* and *B* are known, an information-theoretic notion called information flow rigorously determines the causal relations between *A* and *B*: If the information flow from *A* to *B* is zero, *A* has no effect on *B*; otherwise, *A* does affect *B*. Now X. San Liang of the Nanjing University of Information Science and Technology in China has shown how to obtain the information flow, not from *a priori* known dynamics but from correlations in the time-series graphs that detail the evolutions of *A* and *B*. He applied his result to a problem of practical interest for climate scientists—the relation between El Niño and the Indian Ocean Dipole (IOD), an aperiodic oscillation in sea surface temperature. Liang used correlations between time series—of sea surface temperatures in the Indian Ocean and of an index, called Niño4, that measures the overall strength of El Niño—to calculate the information flow from the IOD to El Niño shown in the figure. For a large swath of the northern Indian Ocean, the sign of the information flow is positive, which, according to information theory, means the IOD causes El Niño to be less predictable. That unusual causal link, suggests Liang, may be the reason climate scientists only recently recognized an influence of the IOD on El Niño. (X. S. Liang, *Phys. Rev. E.*, in press.)

$$T_{2\to1} = -E\left[\frac{1}{\rho_1}\int_{\Re^{n-2}} \frac{\partial F_1\rho_{\backslash 2}}{\partial x_1} dx_3 \ldots dx_n\right] + \frac{1}{2}E\left[\int_{\Re^{n-2}} \frac{\partial^2 g_{11}\rho_{\backslash 2}}{\partial x_1^2} dx_3 \ldots dx_n\right.$$

$$\hat{T}_{2\to1} = \frac{1}{\det \mathbf{C}} \cdot \sum_{j=1}^{n} \Delta_{2j} \mathbf{C}_{j,d1} \cdot \frac{C_{12}}{C_{11}}$$

Thanks for Your Attention

X. San Liang (梁湘三)

Fudan University, Dept. Atmos. Ocean Sci.
Southern Marine Laboratory, The AI Group
复旦大学大气海洋系
南方海洋实验室前沿研究中心人工智能组

E-mail: xsliang@fudan.edu.cn

URL: http://ncoads.cn/

南方海洋实验室**人工智能**组根据国家海洋发展的战略需求，针对现有海洋大气运动的不确定性诸问题，开展Liang-Kleeman信息流、因果分析等新理论研究，发展因果人工智能算法，进行人工智能预报等国际前沿科学研究，从而突破目前数值模式的可预报性局限，逐步建立新一代的智能海洋大气预报系统。

拟聘岗位：博士后、副研究员及以上**（具有全球竞争力的待遇）**

招聘条件（副高及以上）：1.具有博士学位，物理海洋、动力气象、计算数学、物理、力学（控制论、优化、神经网络、动力系统）、计算机科学或其它相关专业； 2.具有高水平大学/研究机构副高级及以上职称或相当经历、相当层次的优秀人才

联系地址：majiwang@sml-zhuhai.cn

南方海洋科学与工程广东省实验室（珠海）

Southern Marine Science and Engineering Guangdong Laboratory(Zhuhai)