PERSPECTIVE

Quantitative Causality, Causality-Aided Discovery, and Causal Machine Learning

X. San Liang[1,2,4]*, Dake Chen[1,3], and Renhe Zhang[2,4]

[1]Division of Frontier Research, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China. [2]Institute of Atmospheric Sciences and CMA-FDU Joint Laboratory of Marine Meteorology, Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai, China. [3]Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China. [4]IRDR ICoE on Risk Interconnectivity and Governance on Weather/Climate Extremes Impact and Public Health, Fudan University, Shanghai, China.

*Address correspondence to: x.san.liang@gmail.com

It has been said, arguably, that causality analysis should pave a promising way to interpretable deep learning and generalization. Incorporation of causality into artificial intelligence algorithms, however, is challenged with its vagueness, nonquantitativeness, computational inefficiency, etc. During the past 18 years, these challenges have been essentially resolved, with the establishment of a rigorous formalism of causality analysis initially motivated from atmospheric predictability. This not only opens a new field in the atmosphere-ocean science, namely, information flow, but also has led to scientific discoveries in other disciplines, such as quantum mechanics, neuroscience, financial economics, etc., through various applications. This note provides a brief review of the decade-long effort, including a list of major theoretical results, a sketch of the causal deep learning framework, and some representative real-world applications pertaining to this journal, such as those on the anthropogenic cause of global warming, the decadal prediction of El Niño Modoki, the forecasting of an extreme drought in China, among others.

Introduction

Causality analysis is a fundamental problem in scientific research (e.g., [1–3]), as commented by Einstein in 1953 in response to a question on the status quo of science in China at that time (cf. the historical record in [4]). The recent rush in artificial intelligence (AI) has stimulated enormous interest in causal inference, partly due to the realization that it may take the field to the next level to approach human intelligence (see [5–7]). In the fields pertaining to this journal, assessment of the cause–effect relations between dynamic events makes a natural objective for the corresponding researches.

During the past 18 years, causality analysis in a quantitative sense has been developed independently in physics from first principles. This is quite different from the various formalisms such as the classical ones in statistics, e.g., [1]. It is the first formalism established on a rigorous footing within the framework of dynamical systems, which yields an explicit solution in closed form, allowing for quantifying and normalizing with ease the causality between dynamical events. Originally born from atmospheric predictability studies, it is hence of special use to the earth system science and, particularly, the atmosphere-ocean-climate science. In the following, we will give a brief introduction of the theory and the resulting methodology, with a focus on a concise formula for the sake of practical usage. We then show, with the aid of this methodology, how a variety of scientific discoveries can be easily made. This is followed by an introduction of the recent causal AI algorithm development and its application to a long-term prediction of El Niño Modoki and the forecasting of an extreme drought in China.

A Gentle Stroll through the Theory of Quantitative Causality Analysis

Ever since Granger's seminal work [1], data-based causal inference has been conventionally investigated as a subject in statistics or engineering science. On the other hand, Liang [8] argued that causality in the Newtonian sense is actually "a real physical notion that can be derived ab initio". This line of work can be traced back to Liang and Kleeman [9], where a discovery about the information flow with 2-dimensional deterministic systems is presented. A comprehensive study with generic systems has been fulfilled recently, with explicit formulas attained in closed forms; see [10] and [8]. These formulas have been validated with benchmark systems such as baker transformation, Hénon map, Kaplan–Yorke map, Rössler system, to name a few. They have also been applied to real world problems in the diverse disciplines such as climate science, meteorology, turbulence, neuroscience, financial economics, quantum mechanics, etc. The following is a brief introduction of the theory, originally motivated by the predictability study in atmosphere-ocean science and later on formulated within a more generic framework, i.e., the framework of stochastic dynamical systems.

To illustrate, consider a d-dimensional continuous-time stochastic system (systems with discrete-time mappings also available; see [8]):

$$\frac{dX}{dt} = F(X, t) + B(X, t)\dot{W}, \tag{1}$$

where $X = (X_1, X_2 \dots, X_d)^T$ is a d-dimensional vector of state variables, where $F = (F_1, F_2 \dots, F_d)^T$ may be arbitrary nonlinear functions of X and t, $\dot{W}$ is a vector of white noise, and $B = (b_{ij})$ is the matrix of stochastic perturbation amplitudes. Here, F and B are both assumed to be differentiable with respect to X and t. For deterministic systems such as those in meteorology and oceanography, B is zero. Liang and Kleeman [9] defined the rate of information flow/information transfer, or simply information flow, from a component X_j to another component X_i, as the contribution of entropy from X_j per unit time in increasing the marginal entropy of X_i. Hereafter by "entropy", we mean Shannon entropy, although other types of entropy have also been explored, such as Kullback–Leiber divergence [11], Von Neumann entropy [12], etc. Liang [8] proved that the rate of information flowing from X_j to X_i (in nats per unit time), denoted $T_{j \to i}$, is

$$T_{j \to i} = -E\left(\frac{1}{\rho_i}\int_{\mathbb{R}^{d-2}}\frac{\partial F_i \rho_{\backslash j}}{\partial x_i}d\mathbf{x}_{\backslash i\backslash j}\right)$$
$$+ \frac{1}{2}E\left(\frac{1}{\rho_i}\int_{\mathbb{R}^{d-2}}\frac{\partial^2\left(g_{ii}\rho_{\backslash j}\right)}{\partial x_i^2}d\mathbf{x}_{\backslash i\backslash j}\right), \tag{2}$$

where $d\mathbf{x}_{\backslash i\backslash j}$ signifies $dx_1\, dx_2 \dots dx_{i-1}\, dx_{i+1} \dots dx_{j-1}\, dx_{j+1} \dots dx_d$, E stands for mathematical expectation, $g_{ii} = \sum_{k=1}^{d} b_{ik}b_{ik}$, $\rho_i = \rho_i(x_i)$ is the marginal probability density function of X_i, and $\rho_{\backslash j} = \int_{\mathbb{R}}\rho(\mathbf{x})dx_j$.

Equation 2 has a nice property, which forms the basis of the information-flow-based causality analysis [10], and has been referred to as "principle of nil causality". It reads that if the evolution of X_i does not depend on X_j, then $T_{j \to i} = 0$. On the basis of this property, the algorithm for the information-flow-based causal inference is as follows: If $T_{j \to i} = 0$, then X_j is not causal to X_i; otherwise it is causal, and the absolute value measures the magnitude of the causality from X_j to X_i.

Another property regards the invariance upon arbitrary nonlinear coordinate transformation, indicating that the obtained information flow is an intrinsic property in nature; see [11]. As shown in Liang et al. [13] (and other publications), this is very important in causal graph reconstruction. It, together with the principle of nil causality, makes it promising toward a solution of the problem of latent confounding.

For linear systems, i.e., when $F(X) = f + AX$ and when B is constant, then

$$T_{j \to i} = a_{ij}\frac{\sigma_{ij}}{\sigma_{ii}}, \tag{3}$$

where a_{ij} is the $(i,j)^{th}$ entry of A and σ_{ij} the population covariance between X_i and X_j. Notice that if X_i and X_j are not correlated, then $\sigma_{ij} = 0$, which yields a zero causality: $T_{j \to i} = 0$. However, conversely, it is not true. We hence have the following corollary:

In the linear sense, causation implies correlation, but not vice versa.

In an explicit formula, this corollary expresses the long-standing debate on causation vs. correlation ever since Berkeley [14].

In the case with only d time series $X_1, X_2, \dots, X_d$, under the assumption of a linear model with additive and independent noises, the maximum likelihood estimator (MLE) of Eq. 2 for $T_{2 \to 1}$ is [15]:

$$\hat{T}_{2 \to 1} = \frac{1}{\det\mathbf{C}} \bullet \sum_{j=1}^{d} \Delta_{2j}C_{j,d1} \bullet \frac{C_{12}}{C_{11}} \tag{4}$$

where C_{ij} is the sample covariance between X_i and X_j, Δ_{ij} is the cofactors of the matrix $\mathbf{C} = (C_{ij})$, and $C_{i,dj}$ is the sample covariance between X_i and a series derived from X_j: $\dot{X}_{j,n} = (X_{j,n+k} - X_{j,n})/(k\Delta t)$, with $k \geq 1$ some integer.

Equation 4 is rather concise in form, involving only the common statistics, i.e., sample covariances. The transparent formula makes causality analysis, which otherwise would be complicated, very easy, and computationally efficient. Note, however, that Eq. 4 cannot replace Eq. 2; it is just an estimator (MLE) of the latter. One needs to test the statistical significance before making a causal inference based on the estimator $\hat{T}_{2 \to 1}$.

If what are given are not time series, but independent, identically distributed panel data, it has been shown that $\hat{T}_{2 \to 1}$ has the same form as Eq. 4; see [16].

Besides the information flow between 2 components, say X_1 and X_2, it is also possible to estimate the influence of one component, say X_1, on itself. Following the convention since Liang and Kleeman [9], write it as $\frac{dH_1^*}{dt}$. Then, its MLE is

$$\widehat{\frac{dH_1^*}{dt}} = \frac{1}{\det\mathbf{C}} \bullet \sum_{j=1}^{d} \Delta_{1j}C_{j,d1} \tag{5}$$

This result, first obtained by Liang [17] and later illustrated with an example in [15], provides an efficient approach to identifying self-loops in a causal graph (cf. [18]), which has been a challenging issue.

If what we want to know is the causal relation between 2 subsystems/subnetworks, rather than 2 individual components/nodes, the information flow can also be derived, in a way as above, and estimated in the maximum likelihood sense. The results are referred to [19].

Statistical significance tests can be performed for the estimators. This is done with the aid of a Fisher information matrix. See [17] and [15] for details.
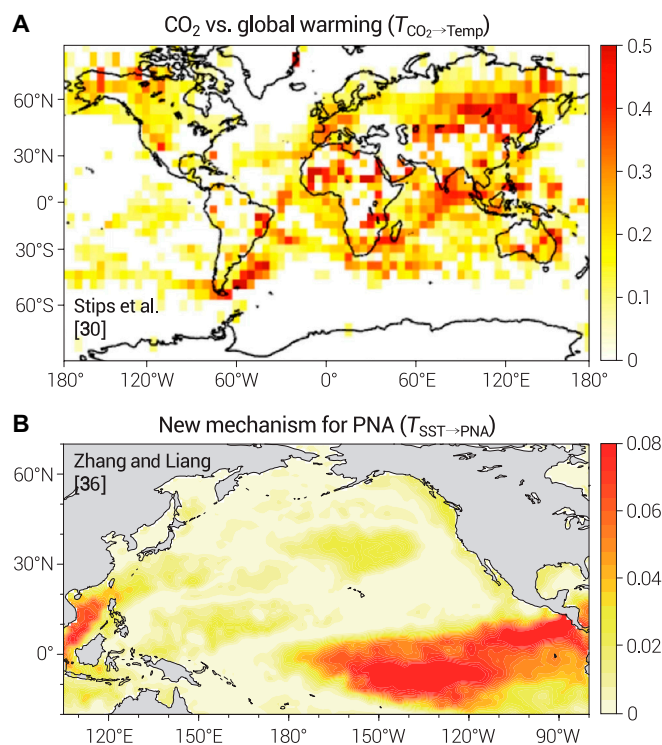
Causality in this sense can be normalized to reveal the relative importance of a causal relation. See [20] for details.

Causality-Aided Scientific Discovery

The above rigorous formalism has been successfully put to application in many real world problems and has led to important scientific discoveries that would otherwise be difficult, if not impossible, to fulfill. These include, on an incomplete list, those in the fields of global climate change [21], near-wall turbulence [22], atmosphere-ocean interaction [23], Arctic rapid warming [24], Antarctic surface mass balance [25], financial economics (e.g., [26]), soil moisture-precipitation interaction [27], climate network [28], brain disease diagnosis [29,30],

El Niño-Indian Ocean Dipole connection [17], and quantum information [12], to name a few. Among these, we want to particularly mention the study of Stips et al. [21] on greenhouse gases versus global mean surface temperature anomaly (Fig. 1A). They found that CO_2 emission does drive the recent global warming during the past century, and the causal relation is one-way; Fig. 1A shows the global causal pattern. However, on a time scale of 1,000 years or up, this one-way causality is completely reversed, becoming a causality from air temperature to carbon dioxide concentration. In other words, on the paleoclimate scale, it is global warming that drives the CO_2 emission. This interesting result is consistent with that inferred from the recent ice-core data from Antarctica.

Another discovery pertaining to the scope of this journal is the linkage of South China Sea to the Pacific–North American (PNA) teleconnection pattern. As shown in Fig. 1B, a direct application of Eq. [4] to the PNA index series and the series of the sea surface temperature reveals a clear causal pattern in the Pacific similar to the canonical El Niño, plus another center with equal causal strength but limited within South China Sea. The former verifies the well-known conclusions in previous studies that El Niño is the main drive of PNA. The latter, however, is a completely new one. This is counterintuitive, considering that South China Sea is just a marginal sea with limited size, while PNA is a Northern Hemisphere climate mode that influences the North American weather. This remarkable result, which is shown to be true later on [31], is just a straightforward application of the aforementioned Eq. 4.

Causal AI and Interpretable Machine Learning

As mentioned in the beginning, the current impetus in causality analysis is mainly driven by the AI rush. Incorporating information flow into machine learning to build interpretable AI algorithms is hence of great interest. This is not only due to the mathematical rigor and physical interpretability in the formalism but also due to the remarkable computing performance: on a laptop (DELL XPS MT530), for a network with 30 time series, it takes less than 1 s to fulfill the computation of the $30 \times 29 = 870$ causal relations, while using the MATLAB function gctest (Granger causality testing), the time needed for the computation will be more than 17 d. Efforts in this regard include predictor choosing or sparsity identification (e.g., [13,32]) and causal neural network construction in deep learning. The former can be simply put as the application of causality analysis before deep learning. The latter is more essential, requiring the building of new algorithms. The first algorithm is built by Tyrovolas et al. [33], who have successfully embedded the Liang–Kleeman information flow analysis into the fuzzy cognitive maps (FCMs). Figure 2 is a schematic of the architecture as proposed by Tyrovolas et al. [33]. In this way, the spurious correlations in the neural network are effectively removed, explanatory power enhanced, and prediction/generalization capability improved. Their numerical experiments demonstrate that the new algorithm outperforms all the state-of-the-art FCM-based models in terms of interpretability.

Causal AI-Based Atmosphere-Ocean-Climate Predictions

A recent application of the causal AI prediction is about the monthly precipitation over China. In 2022, the middle and lower reaches of the Yangtze River experienced an extreme drought in summer and autumn. Poyang Lake, the largest freshwater lake in China, turned into a prairie, as shown in the inserted subfigure in Fig. 3A. Equipped with the above causal AI technique, a team participated in a national contest on monthly precipitation prediction (at a lead time of 1 month) and rather accurately predicted the drought in the Yangtze River reaches from July to October 2022. Shown in Fig. 3A is the prediction for September (ratio of precipitation anomaly to climatology). In particular, the severe situation around Poyang Lake is clearly seen.

Another major application is the one with El Niño Modoki, or Central-Pacific-type El Niño. El Niño is a climate mode that has been linked to many hazards globewide, e.g., flooding, drought, wild fires, heat waves, etc. Its accurate forecasting is of great importance to many sectors of our society such as agriculture, energy, and hydrology, to name several. With its societal importance and the elegant setting, El Niño prediction has become a testbed for AI algorithms.

Currently, the wisdom for El Niño prediction is that it may be predictable at a lead time of 1 to 2 years. However, there still exists much uncertainty; an example is the 2014–2016 "Monster El Niño", almost all projections fell off the mark. Among the El Niño varieties, it is believed that El Niño Modoki is particularly difficult to predict. A striking breakthrough has just been made. Liang et al. [13] took advantage of the quantitative nature of the above information-flow-based causality analysis and identified a delayed causal pattern, i.e., the structure of the information flow from the solar activity to the sea surface
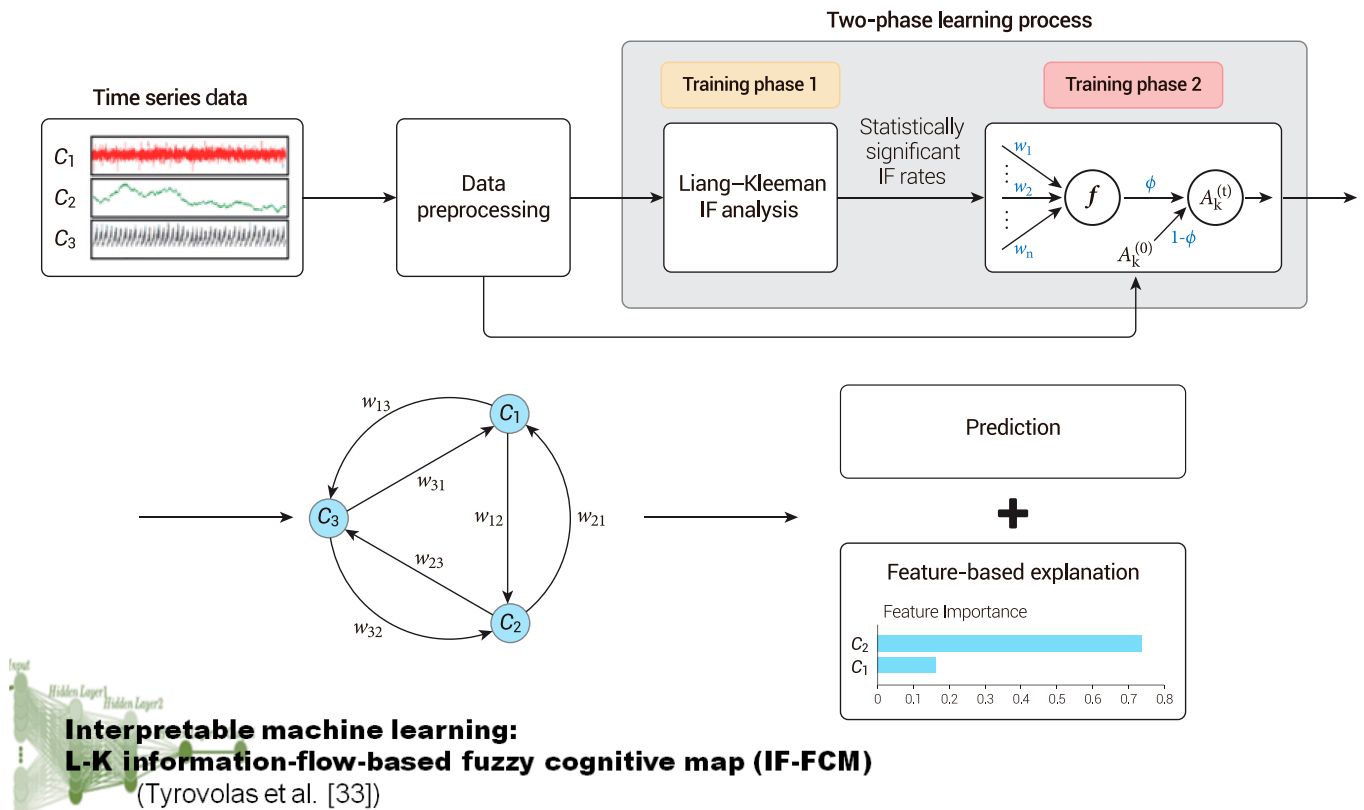

Fig. 1. (A) The causal structure between the global radiative forcing and the annual global mean surface temperature anomaly: The information flow (in nats per year) from CO_2 concentration to global mean surface temperature anomaly (from Fig. 3B of [21]). (B) The information flow (in nats per month) from sea surface temperature to the PNA teleconnection pattern index. Note the hotspot in South China Sea, which tells a new mechanism of PNA generation (redrawn from [31], Fig. 1).

Fig. 2. The architecture of a causal AI algorithm, information-flow-based FCM (IF-FCM), as proposed by Tyrovolas et al. [33] (credit: M. Tyrovolas).
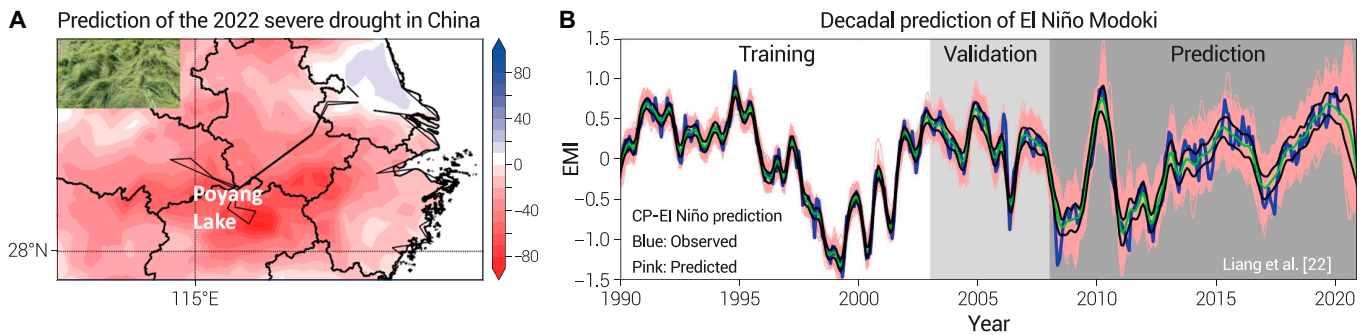

Fig. 3. (A) In the 2022 China National Contest on monthly precipitation prediction, a team equipped with the causal AI technique successfully predicted the extreme drought in the middle and lower reaches of Yangtze River. Here, the color bar indicates the ratio of precipitation anomaly (from climatology) to climatology (in percentage). Inserted is a picture of the largest freshwater lake in China in that extreme drought situation, which looks like the prairie (redrawn from [34]). (B) El Niño prediction has become a benchmark problem for the testing of machine learning algorithms. The present wisdom is that El Niño may be predicted at a lead time of 1 to 2 years. Shown here are 1,000 predictions (pink) of the El Niño Modoki index (EMI) as described in the text. Overlaid are the observed EMI (blue) and the mean of the realizations (cyan). The light shading marks the period for validation, while the darker shading marks the prediction period (from [13], Fig. 5).

temperature, very similar to the El Niño Modoki mode. They hence conjectured that, on the basis of the series of sunspot numbers, El Niño Modoki should be predictable. This is indeed the case, and, remarkably, the prediction can be at a lead time of as long as 10 years or up (Fig. 3A). This remarkable progress, among others, is a result of the rigorously formulated quantitative causality analysis.

Concluding Remarks

Assessment of the cause–effect relation between dynamic events is a major objective of scientific research. Recently, there has been a surge of interest in data-based causal inference, echoing the arrival of the era of big data. Among the efforts is the development of a rigorous formalism of causality analysis from atmosphere-ocean science. This not only opens a new field in the atmosphere-ocean science, namely, information flow, but also has led to scientific discoveries in the diverse disciplines such as quantum mechanics, neuroscience, financial economics, etc. Moreover, its quantitativeness in nature, rigor in physics, and efficiency in computation have revealed to us a promising approach to interpretable deep learning and generalization. Indeed, this has already led to the success in some notoriously difficult real-world problems. We are

expecting more realistic applications in Earth system science and, particularly, in climate projection and weather forecasting.

Data Availability

Not Applicable

References

1. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 1969;37:424–438.

2. Pearl J. *Causality: Models, reasoning, and inference*. New York, (NY): Cambridge Univ Press; 2019. ed. 2.

3. Spirtes P, Glymour C. An algorithm for fast recovery of sparse causal graphs. *Soc Sci Comput Rev*. 1991;9:62–72.

4. Hu D. *China and Albert Einstein: The reception of the physicist and his theory in China, 1917-1979*. Cambridge (MA): Harvard Univ Press; 2005.

5. Bengio Y. An AI pioneer wants his algorithms to understand the 'Why'. WIRED. 8 October 2019. https://www.wired.com/story/ai-pioneer-algorithms-understand-why.

6. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. Paper presented at: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining WSDM'18*; 2018 February 5–9; Marina Del Rey, CA, USA.

7. Schölkopf B. Causality for machine learning. In: *Probabilistic and causal inference: The works of Judea Pearl*. New York (NY): ACM; 2022. pp. 765–804.

8. Liang XS. Information flow and causality as rigorous notions *ab* initio. *Phys Rev E*. 2016;94:052201.

9. Liang XS, Kleeman R. Information transfer between dynamical system components. *Phys Rev Lett*. 2005;95(24):244101.

10. Liang XS. Information flow within stochastic dynamical systems. *Phys Rev E*. 2008;78:031113.

11. Liang XS. Information flow with respect to relative entropy. *Chaos*. 2018;28(7):075311.

12. Yi B, Bose S. Quantum Liang information flow as causation quantifier. *Phys Rev Lett*. 2022;129:020501.

13. Liang XS, Xu F, Rong YN, Zhang RH, Tang X, Zhang F. El Niño Modoki can be mostly predicted more than 10 years ahead of time. *Sci Rep*. 2021;11(1):17860.

14. Berkeley G. *A treatise concerning the principles of human knowledge*. Dublin, Ireland: Aaron Rhames; 1710.

15. Liang XS. Normalized multivariate time series causality analysis and causal graph reconstruction. *Entropy*. 2021;23:679.

16. Rong YN, Liang XS. Panel data causal inference using a rigorous information flow analysis for homogeneous, independent and identically distributed datasets. *IEEE Access*. 2021;9:47266–47274.

17. Liang XS. Unraveling the cause-effect relation between time series. *Phys Rev E*. 2014;90:052150.

18. Hyttinen A, Eberhardt F, Hoyer PO. Learning linear cyclic causal models with latent variables. *J Mach Learn Res*. 2012;13:3387–3439.

19. Liang XS. The causal interaction between complex subsystems. *Entropy*. 2022;24(1):3.

20. Liang XS. Normalizing the causality between time series. *Phys Rev E*. 2015;92:022126.

21. Stips A, Macias D, Coughlan C, Garcia-Gorriz E, Liang XS. On the causal structure between CO_2 and global temperature. *Sci Rep*. 2016;6:21691.

22. Liang XS, Lozano-Durán A. A preliminary study of the causal structure in fully developed near-wall turbulence. In: *Proceedings of the summer program*. Stanford (CA): Center for Turbulence Research, Stanford University; 2016. pp. 233–242.

23. Docquier D, Vannitsem S, Bellucci A. The rate of information transfer as a measure of ocean-atmosphere interactions. *Earth Syst Dynam*. 2023;14(3):577–591.

24. Docquier D, Vannitsem S, Ragone F, Wyser K, Liang XS. Causal links between Arctic Sea ice and its potential drivers based on the rate of information transfer. *Geophys Res Lett*. 2022;49(9):e2021GL095892.

25. Vannitsem S, Dalaiden Q, Goosse H. Testing for dynamical dependence: Application to the surface mass balance over Antarctica. *Geophys Res Lett*. 2019;46(21):12125–12135.

26. Lu X, Sun J, Wei G, Chang C-T. Causal interactions and financial contagion among the BRICS stock markets under rare events: A Liang causality analysis. *Int J Emerg Mark*. 2023; 10.1108/IJOEM-01-2023-0055.

27. Hagan DFT, Wang G, Liang XS, Dolman HAJ. A time-varying causality formalism based on the Liang-Kleeman information flow for analyzing directed interactions in nonstationary climate systems. *J Clim*. 2019;32:7521–7537.

28. Vannitsem S, Liang XS. Dynamical dependencies at monthly and interannual time scales in the climte system: Study of the North Pacific and Atlantic regions. *Tellus*. 2022;74:141–158.

29. Cong J, Zhuang W, Liu Y, Yin S, Jia H, Yi C, Chen K, Xue K, Li F, Yao D, et al. Altered default mode network causal connectivity patterns in autism spectrum disorder revealed by Liang information flow analysis. *Hum Brain Mapp*. 2023;44(6):2279–2293.

30. Hristopulos DT, Babul A, Babul S, Brucar LR, Virji-Babul N. Disrupted information flow in resting-state in adolescents with sports related concussion. *Front Hum Neurosci*. 2019;13:419.

31. Zhang Y, Liang XS. The causal role of South China Sea on the Pacific-North American teleconnection pattern. *Clim Dyn*. 2022;59:5–6.

32. Bai C, Zhang R, Bao S, Liang XS, Guo W. Forecasting the tropical cyclone genesis over the Northwest Pacific through

identifying the causal factors in cyclone-climate interactions. *J Atmos Ocean Technol*. 2018;35:247–259.

33. Tyrovolas M, Liang XS, Stylios C. Information flow-based fuzzy cognitive maps with enhanced interpretability. *Granul Comput*. 2023; 10.1007/s41066-023-00417-7.

34. Ma J, Rong Y, Zhao Y, Fu G, Guo Z, Zhang Y, Zhang L, He W. Causal AI-based prediction of the 2022 extreme drought in China. Paper presented at: Conference on Climate Projection and Climate Application Technology; 2022; November 22–23; Beijing, China (in Chinese).