

RESEARCH ARTICLE | JULY 19 2023

Causal interaction in high frequency turbulence at the biosphere–atmosphere interface: Structure–function coupling

Leila Constanza Hernandez Rodriguez  ; Praveen Kumar  *Chaos* 33, 073144 (2023)<https://doi.org/10.1063/5.0131469>View
OnlineExport
Citation

CrossMark

AIP Advances

Why Publish With Us?

**25 DAYS**
average time
to 1st decision**740+ DOWNLOADS**
average per article**INCLUSIVE**
scope[Learn More](#)

Causal interaction in high frequency turbulence at the biosphere–atmosphere interface: Structure–function coupling

Cite as: Chaos 33, 073144 (2023); doi: 10.1063/5.0131469

Submitted: 20 October 2022 · Accepted: 6 June 2023 ·

Published Online: 19 July 2023



View Online



Export Citation



CrossMark

Leila Constanza Hernandez Rodriguez^{1,a)}  and Praveen Kumar^{1,2,b)} 

AFFILIATIONS

¹Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Champaign, Illinois 61801, USA

²Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Champaign, Illinois 61801, USA

Note: This article is part of the Focus Issue, Theory-informed and Data-driven Approaches to Advance Climate Sciences.

a)Now at: Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

b)Author to whom correspondence should be addressed: kumar1@illinois.edu

ABSTRACT

At the biosphere–atmosphere interface, nonlinear interdependencies among components of an ecohydrological complex system can be inferred using multivariate high frequency time series observations. Information flow among these interacting variables allows us to represent the causal dependencies in the form of a directed acyclic graph (DAG). We use high frequency multivariate data at 10 Hz from an eddy covariance instrument located at 25 m above agricultural land in the Midwestern US to quantify the evolutionary dynamics of this complex system using a sequence of DAGs by examining the structural dependency of information flow and the associated functional response. We investigate whether functional differences correspond to structural differences or if there are no functional variations despite the structural differences. We base our analysis on the hypothesis that causal dependencies are instigated through information flow, and the resulting interactions sustain the dynamics and its functionality. To test our hypothesis, we build upon causal structure analysis in the companion paper to characterize the information flow in similarly clustered DAGs from 3-min non-overlapping contiguous windows in the observational data. We characterize functionality as the nature of interactions as discerned through redundant, unique, and synergistic components of information flow. Through this analysis, we find that in turbulence at the biosphere–atmosphere interface, the variables that control the dynamic character of the atmosphere as well as the thermodynamics are driven by non-local conditions, while the scalar transport associated with CO₂ and H₂O is mainly driven by short-term local conditions.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0131469>

The biosphere–atmosphere interface is a dynamic system where the propagation of fluctuations among hydrometeorological variables, such as air temperature, H₂O, and CO₂, create different types of causal structures over time. The causal structure is represented by a directed acyclic graph (DAG), where non-linear perturbations propagate among variables through dynamic channels for communication between a set of lagged sources of information. In the companion paper,¹ we found that different types of DAGs arise in the dynamics of daytime turbulence using high frequency (i.e., 10 Hz) hydrometeorological data and that there are patterns of similar behavior at the biosphere–atmosphere interface. In this study, we argue that the daytime dynamics are supported by the information flow among its multiple components and we aim to explore if the information flow changes over time

for different causal structures. We implemented a causal history analysis based on a SUR framework, where the total information is decomposed into complementary synergistic information, S, information only provided in the presence of all variables; unique, U, contributions from the self- and cross-dependencies; and redundant information, R, overlapping information from the sources. We explore self-dependence, the influence of a variable's own history on its present state, and cross-dependence, the influence of all other variables arising through interactions in the DAG. Our results show that functional differences are a reflection of causal structural differences to maintain the dynamics of the system, and we discuss how interdependencies in the system are the cause of long-range dependencies, which is a property of stationary processes.

I. INTRODUCTION

High frequency data at the biosphere–atmosphere interface encode the dynamics of a complex and open dissipative system where turbulent processes occur. In this complex system, multiple variables, such as horizontal (WS) and vertical wind speed (U_z), air temperature (T), water vapor (H_2O), and carbon dioxide (CO_2), exhibit nonlinear dynamical dependence. In the companion paper,¹ we use high frequency (10 Hz) data to identify the causal structure of dynamic interdependencies among these components. Such dependencies arise as fluctuations in a variable that instigate variations in other linked variables at a future time,² thereby creating a causal structure of interdependencies in a system that can be represented by a directed acyclic graph (DAG).³ In a DAG, the edges linking states of variables, which are denoted as nodes, are used to capture paths for information flow in a forward direction in time. These edges can be seen as dynamic channels for communication between a set of lagged sources of information, named “parents,” influencing the current state of a target variable.^{4,5} We combined information-theoretic measures to estimate causal influences and explored the evolution of the causal structure in the land–atmosphere exchange during a clear sky day and during a solar eclipse that reflects a transient intervention on the whole system dynamics. We used high frequency (10 Hz) data from an eddy covariance flux tower located in an intensively managed agricultural landscape in the Midwestern US. We estimated a series of consecutive DAGs over non-overlapping windows using the Tigramite algorithm,^{6–9} which are based on information theory metrics.¹⁰ We then used a distance-based classification in conjunction with a k-means clustering approach to cluster the DAGs with similar structures.

Our results show that the system is likely to experience a temporary reduction in its connectivity due to a sudden reduction in solar radiation, which is eventually recovered. We observed a temporal “decoupling” of variables, meaning a significant reduction of self-dependence and cross-connectivity, such as for T and CO_2 .¹ We proposed the concept of “system resilience to information flow,”¹ which refers to the ability of a system to remain within a stable range of the organizational state of its causal structure, sustaining the flow of information even under the effect of an intervention.

From the study of the evolution of the causal structure, we found how information is encoded in the dynamics that arise from interdependencies among components of the system.¹ Does the evolution of the causal structure reflect different underlying behavioral dynamics? To answer this question we analyze the representative dynamics of the clusters using the DAG nearest to the cluster centroid. Given that each of those representative DAGs is formed by edges linking the lagged variables through which information flows, we ask what does change in information flow over time tell us? Here, we aim to explore the relationship between the causal structure and the associated function of the system. We argue that the function of the system is the ability to sustain its dynamics through the different types of interactions that support the information flow among its multiple components. We aim to explore if the information flow is changing over time, is it because (1) the associated functional differences are a reflection of structural differences or (2) structural differences do not lead to functional

differences. Here, we hypothesize that functional outcomes arise due to different types of interactions among component variables and the associated structure of connectivity between them that supports the information flow, these places constraints on and support the processes.²

To test our hypothesis, we build upon the causal structure analyses¹ to estimate the information flow in each DAG cluster using a multivariate causal history approach.^{5,11} The information transfer from the causal history of a multivariate system, developed by Jiang and Kumar,^{4,5} accounts for all interactions among the variables through the entire time history of the system as they influence the outcome of any variable at the current time, t . Using this approach, we characterize the joint influence of lagged self- and cross-dependencies in determining the current state of each variable.^{4,5} While self-dependency refers to how a variable’s own history influences its present state, cross-dependency refers to the influence of all other variables arising through interactions as represented in a DAG.⁵ Using a partitioning time lag τ_C , that can be varied, the causal history can be partitioned into immediate and distant causal histories, each of which can be further partitioned into self- and cross-dependencies.¹¹ Furthermore, these interactions can be partitioned and characterized as unique, synergistic, and redundant (SUR) information flow to capture information that a variable contributes uniquely to the outcome of a target variable, or jointly, or redundantly with another variable, respectively. This partitioning of interactions is called the partial information decomposition (PID) approach^{12,13} and uses momentary partial information decomposition (MPID)¹¹ as the basis for estimation.^{4,5} Using this SUR framework, we estimate the contribution of each information type for each target variable in the system. We also determine whether the dynamics of the turbulent system exhibit long- or short-term memory, and how the influence of the self- and cross-feedback interactions from the immediate and distant causal history sustains the short- and long-term dynamics of the system.

While a few previous studies of causality in turbulence looked at information from the lens of transfer entropy,^{14–17} we propose that the use of a SUR framework provides more insights regarding the underlying dynamics of the system, in terms of the type and amount of information flowing among the system components. To answer our research question, we further compare our results between the different clusters of DAGs as identified in the companion paper¹ to determine if the different causal structures as captured by the representative DAGs reflect differences in their self-organized dependencies and associated information flow. Based on our results, we further discuss the relationship between interactions in the causal history of a variable and the observed long-range dependency as exhibited through spectral analysis.

The paper is organized as follows: In Sec. II, we summarize the concepts and metrics to estimate the information flow in the SUR framework. In Sec. III, we present the results of the causal history analysis for each of the clusters found in the exploration of the evolution of the causal structure of turbulence.¹ Section IV discusses the relationship between the evolution of the causal structure and the corresponding functionality of the system. Section V concludes the study while mentioning the potential implementation of our findings and future work.

II. CHARACTERIZING CAUSAL DEPENDENCY

A. Quantifying information

Here, we review the basic information-theoretic measures based on Shannon's entropy¹⁰ used to quantify the information flow from nonlinear dependencies among multiple variables. The uncertainty of a dynamic variable X_t is quantified by *Shannon's entropy*,

$$H(X_t) = - \sum_{x_t \in X_t} p(x_t) \log p(x_t), \quad (1)$$

where $p(x_t)$ is the probability of X_t . Now, if we consider a variable Y_t , the *conditional entropy* accounts for the uncertainty of X_t that remains given the knowledge of another variable Y_t ,

$$H(X_t | Y_t) = - \sum_{x_t \in X_t, y_t \in Y_t} p(x_t, y_t) \log \frac{p(x_t, y_t)}{p(y_t)}, \quad (2)$$

where $p(x_t, y_t)$ corresponds to the joint probability of X_t and Y_t . The *mutual information* accounts for the dependency between X_t and Y_t , symmetrically measuring the reduction of uncertainty of one variable given the knowledge of the other

$$\begin{aligned} I(X_t; Y_t) &= \sum_{x_t \in X_t, y_t \in Y_t} p(x_t, y_t) \log \frac{p(x_t, y_t)}{p(x_t)p(y_t)} \\ &= H(X_t) - H(X_t | Y_t) = H(Y_t) - H(Y_t | X_t). \end{aligned} \quad (3)$$

Given an additional variable Z_t , which is influenced by the union of X_t and Y_t , the mutual information takes the form of $I(Z_t; X_t, Y_t)$. Further, the *partial information decomposition* (PID)¹³ decomposes this total information shared with Z_t from X_t and Y_t into three components. The components of PID are *Synergistic information*, S , which is the information jointly provided by X_t and Y_t ; *Redundant information*, R , which is the overlapping information between X_t and Y_t ; and, *unique information*, U_X and U_Y , which is the information provided by X_t and Y_t individually. They satisfy

$$I(Z_t; X_t, Y_t) = S + R + U_X + U_Y. \quad (4)$$

In other words, the SUR framework decomposes the mutual information between the three variables. To estimate PID for time series data, we use the approach developed by Goodwell and Kumar.¹²

B. Estimating information flow using DAGs

The causal structure of a dynamical system is represented as a DAG,³ a structural causal model of the dynamics of a multivariate complex system. We consider that the main function of the causal structure is a cause and a consequence of information flow among interacting variables. Here, we describe the information-theoretic measures used to estimate information flow using DAGs for multivariate time series,⁵ which represent the temporal dependencies of the system. In a multivariate system with N variables, $\tilde{X}_t = \{X_t, Y_t, Z_t, \dots\}_N$, varying in time t , the current state of any variable $Z_t \in \tilde{X}_t$ encapsulates the result of all earlier interactions of its dynamics.^{4,5} Such prior dynamics are called causal history, denoted as $\tilde{X}_t^- = \{\tilde{X}_{t-1}, \tilde{X}_{t-2}, \tilde{X}_{t-3}, \dots\}$.^{4,5} We estimate information flow using a DAG representation for time series,

$G = (\tilde{X}_{t+1}^-, E)$, where each node refers to the state at time t of a variable in $\tilde{X}_t$ and E is the collection of edges. If a directed edge E links two nodes $Y_{t-\tau}$ and Z_t , it is denoted by $Y_{t-\tau} \rightarrow Z_t$, where τ is a positive time lag. The parent set of a target variable at time t , Z_t , is denoted as $P_{Z_t} \equiv \{Y_{t-\tau} : Y_t \in \tilde{X}_t, \tau > 0, Y_{t-\tau} \rightarrow Z_t\}$.^{4,5} A node $Y_{t-\tau}$ can also be linked to a target indirectly through a causal path $C_{Y_{t-\tau} \rightarrow Z_t}$, which is a set of nodes connected by a sequence of edges linking from $Y_{t-\tau}$ to Z_t , that is, $C_{Y_{t-\tau} \rightarrow Z_t} = \{V_{t-\tau_V} : V_t \in \tilde{X}_t, \tau_V > 0, Y_{t-\tau} \rightarrow \dots \rightarrow V_{t-\tau_V} \rightarrow \dots \rightarrow Z_t, \tau > \tau_V\} \cup \{Y_{t-\tau}\}$. To estimate the information flow from one or multiple sources on the current state of a variable, we use one of the following possibilities: through a direct edge [Fig. 1(a)],⁷ by one⁸ Fig. 1(b) or two causal paths [Fig. 1(c)],¹¹ or by the whole causal history [Fig. 1(d)].^{4,5}

Information flow through a directed edge between two nodes, $Y_{t-\tau}$ and Z_t , denoted by $Y_{t-\tau} \rightarrow Z_t$ [Fig. 1(a)], exist only if a perturbation in $Y_{t-\tau}$ propagates to Z_t , and it is measured as the *conditional mutual information* (CMI),⁷

$$I(Z_t; Y_{t-\tau} | \tilde{X}_t^- \setminus Y_{t-\tau}) \geq 0, \quad (5)$$

where $\setminus$ is the exclusion symbol, such as $\tilde{X}_t^- \setminus Y_{t-\tau}$ represents all the nodes in the causal history excluding $Y_{t-\tau}$. CMI estimates the information flow from a lagged variable $Y_{t-\tau}$ to the current state of a target variable Z_t , conditioned on the knowledge of the rest of the dynamics of all interacting variables in the causal history $\tilde{X}_t^-$ excluding $Y_{t-\tau}$.⁵ However, the computation of CMI is infeasible due to the potentially infinite number of nodes arising from the causal history (i.e., curse of dimensionality). To make the computation possible, the *momentary information transfer* (MIT) was developed¹⁸ using the Markov property for DAG,¹⁹ which states that a node Z_t is statistically independent of the rest of the history of its parents P_{Z_t} are given, where $P_{Z_t} = \{X_{t-\tau} : X_{t-\tau} \in \tilde{X}_{t-\tau}, \tau > 0, X_{t-\tau} \rightarrow Z_t\}$.^{4,5} MIT quantifies the direct interaction between two nodes, $Y_{t-\tau}$ and Z_t , acting as source and target, respectively, and excludes any information from other nodes that may be flowing through the source or directly to the target.⁵ MIT is defined by Eq. (7) in Runge *et al.*,¹⁸

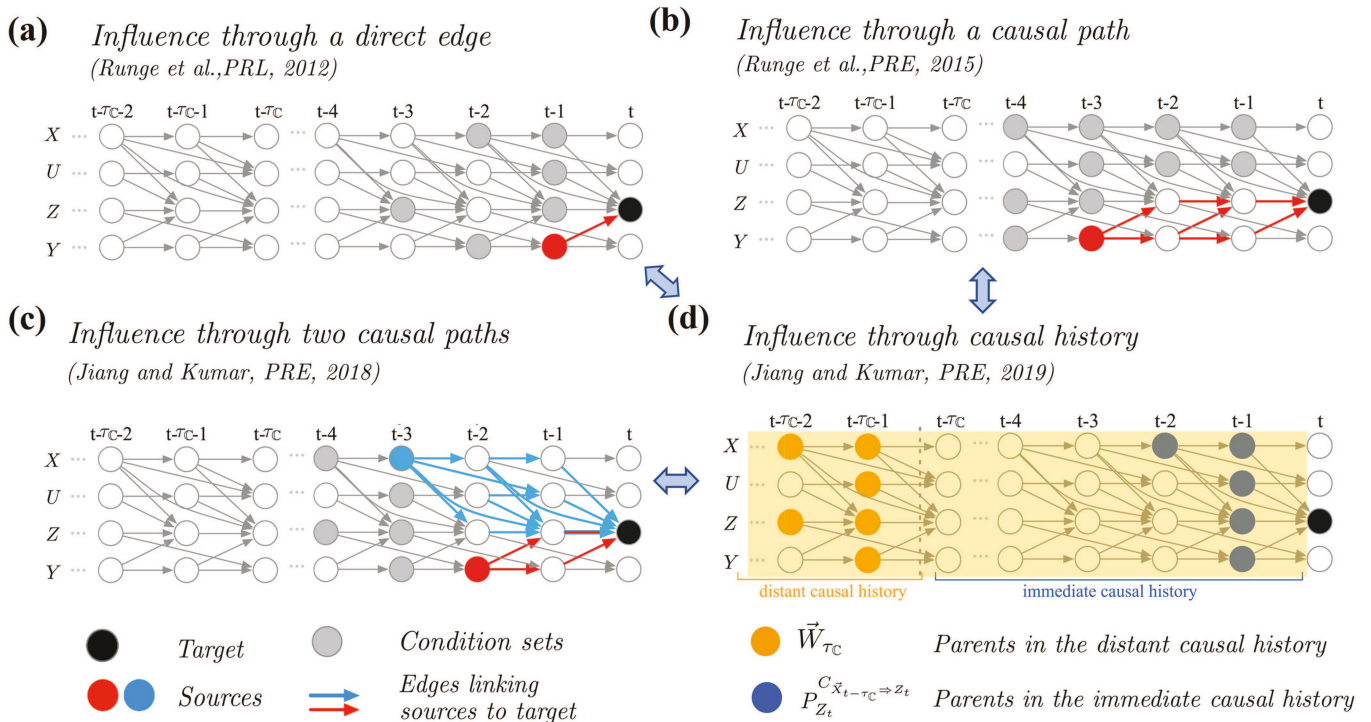
$$I_{Y_{t-\tau} \rightarrow Z_t}^{\text{MIT}}(\tau) = I(Y_{t-\tau}; Z_t | P_{Z_t} \setminus \{Y_{t-\tau}\}, P_{Y_{t-\tau}}). \quad (6)$$

In addition to a direct influence through an edge, *information flow through a causal path* [Fig. 1(b)] occurs when a lagged source node $Y_{t-\tau}$ indirectly affect a target node Z_t . The implementation of the Markov property for information flow through a causal path is called *momentary information transfer along causal path* (MITP) and is defined by Eq. (18) in Runge,⁶ as follows:

$$I_{Y_{t-\tau} \rightarrow Z_t}^{\text{MITP}}(\tau) = I(Y_{t-\tau}; Z_t | P_{Z_t} \setminus C_{Y_{t-\tau} \rightarrow Z_t}, P_{C_{Y_{t-\tau} \rightarrow Z_t}}, N_{Y_{t-\tau}}^{Z_t}, P(N_{Y_{t-\tau}}^{Z_t})), \quad (7)$$

where $P_{C_{Y_{t-\tau} \rightarrow Z_t}}$ is the parent set of the causal path $C_{Y_{t-\tau} \rightarrow Z_t}$, and $N_{Y_{t-\tau}}^{Z_t}$ is the neighbor set of the causal path.⁶

As an extension of the MITP, the *information flow through two causal paths* [Fig. 1(c)] estimates the influence of two lagged sources $X_{t-\tau_X}$ and $Y_{t-\tau_Y}$ on a target Z_t , through the corresponding causal paths $C_{X_{t-\tau_X} \rightarrow Z_t}$ and $C_{Y_{t-\tau_Y} \rightarrow Z_t}$, respectively.¹¹ The momentary interaction information for separable causal paths (MIISCPs)¹¹ quantifies



Adapted from Jiang & Kumar, 2019

FIG. 1. Illustration adapted from Jiang and Kumar⁵ that shows the information flow to a target node Z_t in a quadivariate complex system from (a) Y_{t-1} through a directed edge; (b) Y_{t-3} through the corresponding causal path $C_{Y_{t-3} \rightarrow Z_t}$; (c) X_{t-4} and Y_{t-2} through the union of the corresponding two causal paths $C_{X_{t-4}, Y_{t-2} \rightarrow Z_t}$; and (d) the entire causal history, which can be partitioned into immediate and distant causal histories based on a partition time lag τ_c .

the information transfer to a target from a preceding causal sub-graph starting with two sources with separable causal paths, and is defined by Eqs. (10) and (11) in Jiang and Kumar,¹¹

$$I_{\{X_{t-\tau_X}, Y_{t-\tau_Y}\} \rightarrow Z_t}^{\text{MIISCP}} = I(X_{t-\tau_X}; Y_{t-\tau_Y}; Z_t | \vec{W}), \quad (8)$$

where

$$\vec{W} = \vec{W}_1 \cup \vec{W}_2 \cup \vec{W}_3, \quad (9a)$$

$$\vec{W}_1 = P_{Z_t} \setminus (C_{X_{t-\tau_X} \rightarrow Z_t} \cup C_{Y_{t-\tau_Y} \rightarrow Z_t}), \quad (9b)$$

$$\vec{W}_2 = P_{C_{X_{t-\tau_X} \rightarrow Z_t}} \setminus C_{Y_{t-\tau_Y} \rightarrow Z_t}, \quad (9c)$$

$$\vec{W}_3 = P_{C_{Y_{t-\tau_Y} \rightarrow Z_t}} \setminus C_{X_{t-\tau_X} \rightarrow Z_t} \quad (9d)$$

where the condition set $\vec{W}$ represents the parents of the union set of the target and the causal paths from the two sources to the target, $C_{X_{t-\tau_X} \rightarrow Z_t}$ and $C_{Y_{t-\tau_Y} \rightarrow Z_t}$.¹¹ When the influence through two causal paths affects a target, their interaction can be characterized through PID in what is defined as *momentary partial information*

decomposition (MPID),¹¹

$$I_{\{X_{t-\tau_X}, Y_{t-\tau_Y}\} \rightarrow Z_t}^{\text{MSCP}} = S_c + R_c + U_{X,c} + U_{Y,c} \quad (10)$$

where the subscript c on the right-hand side represents that the PID is associated with causal paths.⁴ MPID accounts for the information only going through the pathways linking the sources and the target, with the influence from earlier dynamics excluded through conditioning.⁴

The state of a target variable Z_t is the result of the prior states of all interdependent variables in the system, i.e., *causal history* $\vec{X}_t^-$, with information flowing through a multitude of different pathways in the DAG [Fig. 1(d)].⁴ The *total information flow* ($\mathcal{T}$) from the causal history can be partitioned into two different complementary components. The immediate causal history arising from all the previous states from time step $t-1$ up to the time step $t-\tau_c$, called *immediate causal history* ($\mathcal{I}$), and the remaining earlier dynamics, called *distant causal history* ($\mathcal{D}$), respectively [Fig. 1(d)],⁴ and can be estimated as follows:

$$\mathcal{T} = I(Z_t; \vec{X}_t^-) = \mathcal{I}(\tau_c) + \mathcal{D}(\tau_c), \quad (11)$$

with

$$\mathcal{I}(\tau_c) = I(Z_t; C_{\vec{X}_{t-\tau_c} \rightarrow Z_t} | \vec{X}_t^- \setminus C_{\vec{X}_{t-\tau_c} \rightarrow Z_t}), \quad (12a)$$

$$\mathcal{D}(\tau_C) = I\left(Z_t; \bar{X}_t \setminus C_{\bar{X}_t - \tau_C \Rightarrow Z_t}\right), \quad (12b)$$

where immediate ($\mathcal{J}$) and distant ($\mathcal{D}$) causal histories are the partitioning of the total information $\mathcal{T}$ at time lag τ_C . Further, the Markov property of the DAG provides the feasibility of computing Eq. (12) by simplifying $\mathcal{J}(\tau_C)$ and $\mathcal{D}(\tau_C)$ as⁴

$$\mathcal{J}(\tau_C) = I\left(Z_t; P_{Z_t}^{\bar{X} \xrightarrow{C} t - \tau_C \Rightarrow Z_t} \mid \bar{W}_{\tau_C}\right), \quad (13a)$$

$$\mathcal{D}(\tau_C) = I(Z_t; \bar{W}_{\tau_C}), \quad (13b)$$

where $\bar{W}_{\tau_C}$ refers to the total information from the distant causal history, quantified as $\mathcal{D}$ in Eq. (13b).

Furthermore, the analysis of multivariate time series using DAG representation allows for the partitioning of the immediate and distant causal histories of any target node $Z_t \in \bar{X}_t$ of a system into self- and cross-dependencies.⁵ Self-dependency refers to the variable's own history influencing its present state, denoted by $\bar{Z}_{\mathcal{D}} \equiv \{Z_{t-\tau} : Z_{t-\tau} \in \bar{W}_{\tau_C}\}$. Cross-dependency refers to the influence of all other variables arising through interactions in a DAG, denoted by $\bar{Z}'_{\mathcal{D}} \equiv \bar{W}_{\tau_C} \setminus \bar{Z}_{\mathcal{D}}$. Using the PID framework, the interactions in the distant causal history $\mathcal{D}$ between a target and other variables can be quantified as a function of the partitioning time lag τ_C as follows:

$$\mathcal{D}(\tau_C) = I(Z_t; \bar{W}_{\tau_C}) = I(Z_t; \bar{Z}_{\mathcal{D}}, \bar{Z}'_{\mathcal{D}}) \quad (14a)$$

$$= S_{\mathcal{D}}(\tau_C) + R_{\mathcal{D}}(\tau_C) + U_{\text{self}, \mathcal{D}}(\tau_C) + U_{\text{cross}, \mathcal{D}}(\tau_C), \quad (14b)$$

where $S_{\mathcal{D}}$ and $R_{\mathcal{D}}$ are the synergistic and redundant information from distant causal history, respectively, and $U_{\text{self}, \mathcal{D}}$ and $U_{\text{cross}, \mathcal{D}}$ are the unique information from the self- and cross-dependencies, respectively. Likewise, for the immediate history, $\mathcal{J}$, the partitioning is given by

$$\begin{aligned} \mathcal{J}(\tau_C) &= I\left(Z_t; P_{Z_t}^{\bar{X} \xrightarrow{C} t - \tau_C \Rightarrow Z_t} \mid \bar{W}_{\tau_C}\right) = I(Z_t; \bar{Z}_{\mathcal{J}}, \bar{Z}'_{\mathcal{J}} \mid \bar{W}_{\tau_C}) \quad (15a) \\ &= S_{\mathcal{J}}(\tau_C) + R_{\mathcal{J}}(\tau_C) \\ &\quad + U_{\text{self}, \mathcal{J}}(\tau_C) \\ &\quad + U_{\text{cross}, \mathcal{J}}(\tau_C). \end{aligned} \quad (15b)$$

Therefore, from Eqs. (13) to (15), we can estimate the PID of the information from the entire causal history, $\mathcal{T}$, as follows:

$$\mathcal{T} = I(Z_t; P_{Z_t}) = \mathcal{J} + \mathcal{D} \quad (16a)$$

$$\begin{aligned} &= S_{\mathcal{J}} + R_{\mathcal{J}} + U_{\text{self}, \mathcal{J}} \\ &\quad + U_{\text{cross}, \mathcal{J}} + S_{\mathcal{D}} + R_{\mathcal{D}} + U_{\text{self}, \mathcal{D}} + U_{\text{cross}, \mathcal{D}} \end{aligned} \quad (16b)$$

$$= S_{\mathcal{T}} + R_{\mathcal{T}} + U_{\text{self}, \mathcal{T}} + U_{\text{cross}, \mathcal{T}}, \quad (16c)$$

where

$$S_{\mathcal{T}} = S_{\mathcal{J}} + S_{\mathcal{D}}, \quad (17a)$$

$$R_{\mathcal{T}} = R_{\mathcal{J}} + R_{\mathcal{D}}, \quad (17b)$$

$$U_{\text{self}, \mathcal{T}} = U_{\text{self}, \mathcal{J}} + U_{\text{self}, \mathcal{D}}, \quad (17c)$$

$$U_{\text{cross}, \mathcal{T}} = U_{\text{cross}, \mathcal{J}} + U_{\text{cross}, \mathcal{D}}. \quad (17d)$$

We use the rescaled redundancy approach of Goodwell and Kumar¹² incorporated by Jiang and Kumar^{4,5} for computing the PID for $\mathcal{D}$, $\mathcal{J}$, and $\mathcal{T}$, where the redundant information is estimated by considering the mutual dependency between two sources and ensures a non-negative information partitioning.⁵ We compute $\mathcal{D}$ and $\mathcal{J}$, along with their PIDs, using Eqs. (15) and (14), respectively. Equations (16) and (17) indicate that the information contained in the entire causal history of a variable encompasses both its immediate and distant histories, including both self- and cross-dependencies.^{4,5}

This review of information flow metrics has summarized how to quantify the dynamics that sustain the whole multivariate system.⁵ Two important aspects should be considered when implementing the previous information flow metrics. First, the dimensionality can be high for reliable estimations when computing $\mathcal{J}$, $\mathcal{D}$, and PID, even after using the Markov property for DAG.⁷ The dimensionality grows as the number of variables increases and/or as the number of lags that influence a target increases.⁵ Therefore, we use the *momentary information weighted transitive reduction* (MIWTR),⁵ which reduces the dimensionality of the DAG under analysis. Briefly, MIWTR removes edges linking $\bar{W}_{\tau_C}$ with the immediate causal history, $C_{\bar{X}_t - \tau_C \Rightarrow Z_t}$ and then exclude the nodes $\bar{W}_{\tau_C}$ not directly linked to $C_{\bar{X}_t - \tau_C \Rightarrow Z_t}$ to obtain a reduced $\bar{W}_{\tau_C}$. Second, the adequate estimation of information metrics requires a considerable data length. Jiang and Kumar⁵ found that the cardinality reduction based on MIWTR does not affect the estimation of information-theoretic measures significantly when the time series data length is sufficient. Data length greater than 1000 is considered as the minimum needed to achieve reliable estimations.⁵

III. RESULTS

We explore the relationship between causal structure and associated function, using the representative DAGs of the clusters determined in Hernandez Rodriguez and Kumar¹ (Fig. 2). As stated earlier, we define functionality as the nature of interactions as discerned through the SUR components. Our aim is to discern the structure–function duality in the causal interactions that shape the observed dynamics in the causal history. If the different causal structures give rise to different behavior in their information flow, then we may infer that the functional differences are a reflection of structural differences. Otherwise, we may infer that different structural behaviors may give rise to similar functional behavior despite the structural differences, indicating that there are no differences in terms of the ability of the system to sustain its dynamics by reorganizing information flow among its components.

A. Information flow to a target across DAG clusters

To illustrate the approach described above, we consider CO₂ and H₂O as examples of target variables. To estimate their current state (Fig. 3, left and right column, respectively), we select a varying partitioning time lag τ_C to partition the influence of the historical

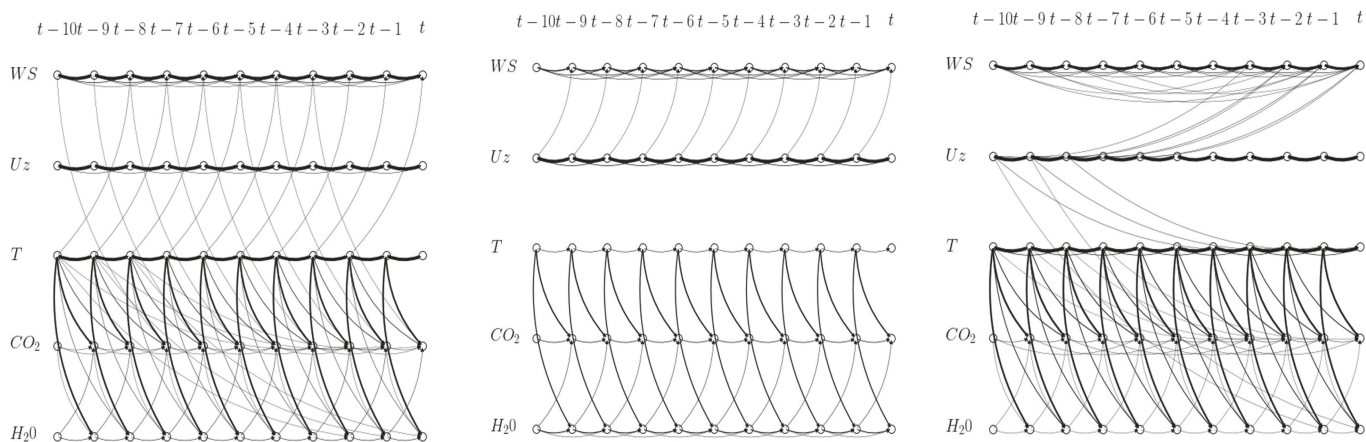


FIG. 2. DAGs from the observed time series that are closest to each cluster's centroids show the average of the dynamics of (a) cluster #1, (b) cluster #2, and (c) cluster #3. The DAGs for time series are estimated using the Tigramite package.^{6–9}

dynamics of all interacting components into immediate and distant history, respectively. By changing τ_C we can then gauge how the influence of immediate history declines or continues to persist through the time history. Furthermore, we explore the differences in self- and cross-dependencies influencing the current state of a target variable in both its immediate and distant history. As stated earlier, self-dependence refers to the variable's own history influencing its present state, and cross-dependence refers to the influence of all other variables arising through interactions as represented in the DAG [see Eq. (16)].

In Fig. 3, we compare the influences across clusters using the DAGs for when $\tau_C = 10$ on the current state of CO_2 and H_2O . That is, the immediate history accounts for the influence of the parents up to one second in the history of the variable. These results show that the distant history of the vertical wind speed, U_z , influences the current state of both CO_2 and H_2O , and it is especially noticeable for clusters #3 and #1. In general, we observe that both self- and cross-dependencies are lower in cluster #2 in comparison to the other two clusters. We also observe that the horizontal wind speed, WS , does not influence the current state of CO_2 and H_2O for clusters #1 and #2. Mechanical turbulence decoupling of the scalar transport from the dynamics is reflected as the absence of information flow from WS on the dynamics of other variables (Fig. 2, clusters #1 and #2). These results are consistent with observations during total solar eclipses regarding the slackening of horizontal wind speed due to boundary layer stabilization.^{20,21} A time-shift between the time of totality and the lagged response of ecohydrological variables is also a recurrent observation.²² It is observed that the stagnant environment with minimal wind speeds shows turbulence decoupling the canopy and atmosphere, which produced uncertainty when estimating transpiration rates.²¹ While the effect of the absence of incoming solar radiation on the upward longwave radiation is lagged only for a few minutes, the fluxes that depend on surface temperature or photosynthesis, such as latent and sensible heat, and CO_2 flux, responded to the change of the net radiation by reducing their magnitudes, and then were restored to their typical values after the increase of turbulence.^{21,22} For a clear sky day behavior or under the effect of an

intervention on the whole system dynamics, such as a solar eclipse, the analysis of changes in the structure of information flow across multiple variables helps us to characterize how the functionality of the ecohydrological system at a turbulent scale evolves. To determine the relationship between the causal structure and function of the system, we determine attributes of the interactions among components of the system.

B. Characteristics of information flow across clusters

After finding the lagged sources of information to the current state of a target variable and for a singular partitioning time lag τ_C , we now estimate the amount of information flowing from the causal history of each variable for multiple τ_C . We estimate the maximum information from the entire causal history, $\mathcal{T}$, which is invariant with respect to the partitioning time lag τ_C and the percentage of information from the distant history as a function of the partitioning time lag τ_C (Fig. 4). Typically, for small τ_C a significant fraction of causal information comes from the distant history, $\mathcal{D}$, and as τ_C increases, the information from $\mathcal{D}$ decreases while for information from the immediate history $\mathcal{I}$ increases. Furthermore, after the initial high $\mathcal{D}$, the percentage of information from distant history with respect to the total information, $\mathcal{D}/\mathcal{T}$, typically exhibits a decrease, excepting for the vertical component of the wind velocity, U_z (Fig. 4). For U_z as τ_C increases, the signal quickly dies out, which suggests that the percentage of information coming from its distant history is small. Therefore, we can say that the vertical wind velocity, U_z , highly relies on its short immediate history, which is characteristic of a short-term memory process. When the ratio $\mathcal{D}/\mathcal{T}$ increases at a higher τ_C , as observed for the horizontal wind speed, WS , in cluster #2 (Fig. 4), it indicates that the distant history provides more relevant causal information for intermediate values of τ_C , before decreasing in importance at longer time scales.

Overall, as τ_C increases, $\mathcal{D}$ accounts for less than 50% of $\mathcal{T}$, suggesting that the turbulent processes are mostly sustained by short-term causal interactions although longer-term contributions constitute an important fraction (Fig. 4). Across the three

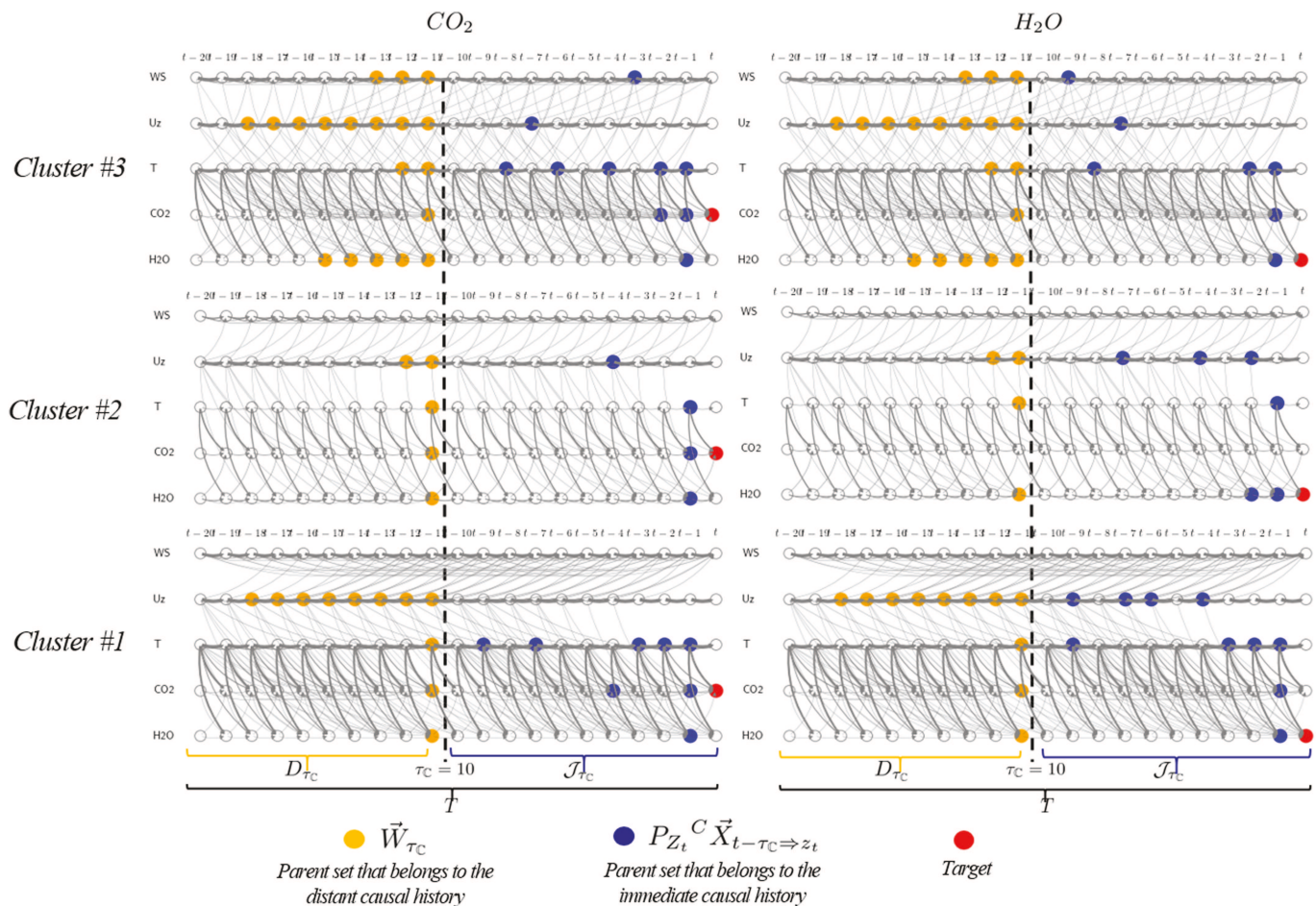


FIG. 3. Partitioned causal history for directed acyclic graphs (DAG) using the causal history approach^{4,11} with MIWTR⁵ for the present state of CO₂ (left column) and H₂O (right column). The DAGs correspond to the representative time series that is closest to the centroid of each cluster found in the causal structure analysis.¹ The nodes refer to the state of the variables at times $t - \tau$ (as marked on top), where τ is a positive time lag, here equal to 0.1s (i.e. since data is at 10Hz). We set the partitioning time $\tau_c = 10$ in this illustration to divide the influence of the causal history based on its immediate and distant history on the current state of the target variable. The nodes influencing the target (red color solid nodes) can come from the parent set in the immediate (blue color solid nodes) or distant (yellow color solid nodes) history.

clusters, we observe that the magnitude of the ratio between $\mathcal{D}$ and $\mathcal{T}$ changes depending on the variable. This suggests that the causal structure of the clusters contains information about the short- and long-term interdependencies among variables, which implies that the differences in information flow among variables reflect how the process may interact through the identified causal structure.

C. A SUR framework to infer characteristics of multivariate turbulence

Do the self-dependencies or cross-dependencies dominate the immediate and distant dynamics of each process? Knowing how the multivariate system informs the short-term dominated processes at a turbulent scale, provides insights into the oscillatory behavior observed for temperature T in cluster #2 (Fig. 4 red line in T panel at the bottom), which might be related to a periodic

component as a function of τ_c . To further investigate the self- and cross-dependencies for each variable and across clusters, here we use a partial information decomposition (PID) based causal history analysis⁵ described earlier (Sec. II). Using a partitioning of the total information, we observe how each process sustains its dynamics and its level of dependency on other components in the multivariate system (Fig. 5). For instance, for the horizontal wind speed, WS, the unique information from self-dependency, $U_{self,\mathcal{T}}$, is the main contributor sustaining its dynamics for all three clusters, followed by the synergistic, $S_{\mathcal{T}}$, and unique information from cross-dependencies, $U_{cross,\mathcal{T}}$. However, the cross-dependencies are particularly significant for WS in cluster #3, with contributions coming from both $S_{\mathcal{T}}$ and $U_{cross,\mathcal{T}}$, in comparison to the other two clusters. In the case of CO₂ and H₂O, the redundant information, $R_{\mathcal{T}}$, is the main contributor in all three clusters. Lastly, marked variables as “self-dependent” refer to variables that rely only on their own

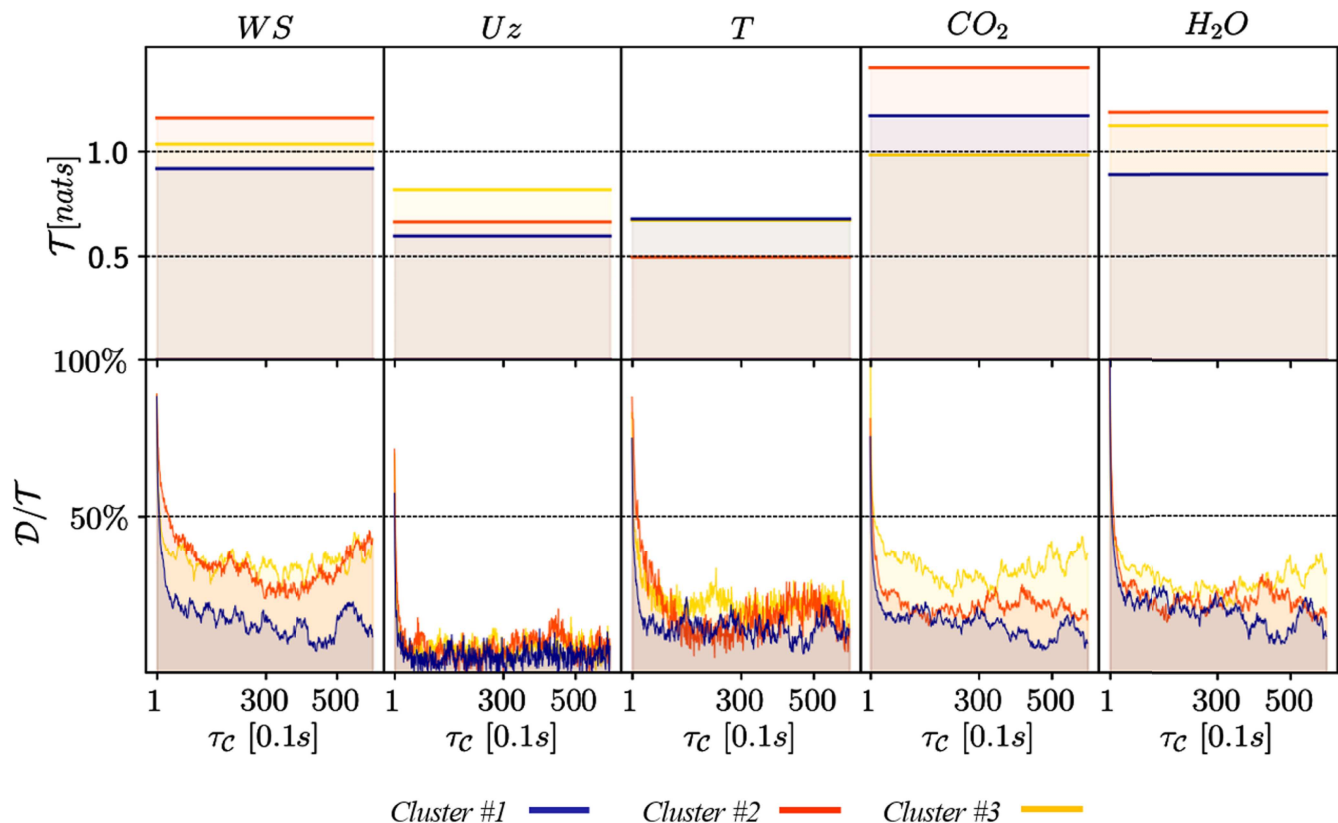


FIG. 4. Information flow from the total causal history, $\mathcal{T}$ (top), and the ratio between the distant, $\mathcal{D}$, and the total causal histories (bottom), over the partitioning time lag τ_C for the cluster #1 (blue line), cluster #2 (red line), and cluster #3 (yellow line) based on the directed acyclic graphs for time series in Fig. 2. Notice that τ_C takes values from 1 to 600, accounting for influences coming from up to one minute.

dynamics. For example, the vertical wind speed, U_z , is causally a self-dependent variable regardless of the cluster. This means that the history of the variable itself, rather than causal dependence on other variables, determines the observed present state.

Now, we explore these findings in more detail regarding their role in self- or cross-dependencies that dominate these processes, by exploring their origin in the immediate or distant histories of the variables. From our results, we can infer some unique characteristics of the system. In Fig. 6, the analysis of the immediate $\mathcal{I}$ and distant $\mathcal{D}$ causal histories provides us with three main results. First, we notice that the processes are mainly sustained by self-dependant variables in the short-term dynamics. We find that especially for the horizontal and vertical wind components, WS and U_z , which reflect the dynamic character of the atmosphere, the unique information from self-dependency in the immediate history, $U_{self,\mathcal{I}}$, is the main contributor for sustaining its dynamics. On the other hand, the redundant information in the immediate history, $R_{\mathcal{I}}$, that refers to the overlapping information from the sources is dominant for CO_2 and H_2O .

Second, from our results we can classify the variables based on their origin, whether they are driven by short-term local or non-local conditions (Fig. 6). Non-local conditions refer to variables

with strong self-dependency, such as U_z and T , whose dynamics are primarily or even solely supported by their own causal history, without the influence of other variables. This category also includes WS , which is mainly influenced by its unique information from self-dependency in the immediate and distant history, $U_{self,\mathcal{I}}$ and $U_{self,\mathcal{D}}$, respectively. Therefore, we can infer that the variables that control the dynamic character of the atmosphere, WS and U_z , as well as the thermodynamics, T , are driven by non-local conditions. On the other hand, CO_2 and H_2O , which reflect the scalar transport at the biosphere-atmosphere interface, are mainly driven by short-term local conditions given that the redundant information for the immediate history dominates, $R_{\mathcal{I}}$, and therefore their dynamics are sustained by the influence of multiple variables in the causal history.

Third, our results also show that as τ_C increases the total information is of the same order of magnitude across variables in the same cluster (Figs. 5 and 6). We notice that the variables in cluster #2 are influenced by a larger amount of information through their causal histories in comparison with clusters #1 and #3. Cluster #2 shows the causal structure of the average dynamics of the cluster under the influence of an intervention on the dynamics of the whole system. Such intervention causes a temporary decrease in

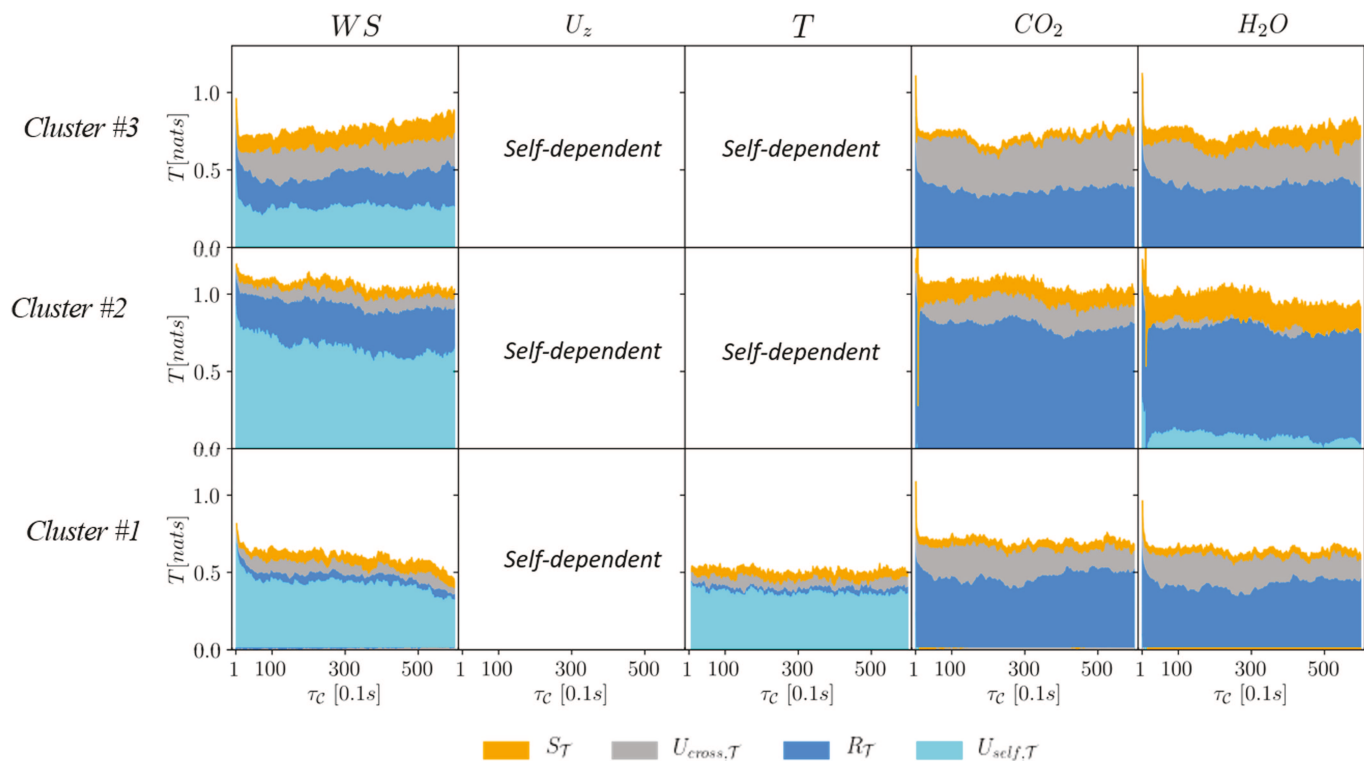


FIG. 5. Partial information decomposition (PID) for the total causal history for the DAGs for high frequency time series as shown in Fig. 2, for cluster #3 (top), cluster #2 (middle), and cluster #1 (bottom). The total information is decomposed into the complementary synergistic information, S_T , the information only provided in the presence of all variables; the redundant information, R_T , the overlapping information from the sources; and the unique contributions from the self, $U_{self,T}$, and cross-dependencies, $U_{cross,T}$.

the connectivity of the system, which refers to the reduced number of edges representing interdependencies among variables (Fig. 2).¹ Such a decrease in connectivity favors a larger amount of information flowing among components of the system (Figs. 5 and 6). Our results suggest that the system potentially reorganizes its causal structure to some sort of optimality by reducing its dimensionality to facilitate a larger amount of information flow. This finding suggests the dynamic creation of preferential channels for information flow: an attribute that links the causal structure and function of the system at turbulent scales. Therefore, we argue that such a minimal and optimal arrangement of the causal structure is shaped by the system to efficiently transfer information, i.e., diffuse fluctuations, among components of the system. Moreover, based on our results, we argue that the system shapes its causal structure as a mechanism for resilience to preserve its functionality. We previously discussed that the system has the ability to temporarily modify its causal structure to optimally meet its essential function, the transfer of information. We argue that the ecohydrological system at turbulent scales is resilient in the sense that it evolves adjusting its causal structure to protect its critical functionality from disruption. Here, resilience refers to the ability of the system to successfully adapt to interventions maintaining the functionality of the system. These support the concept of “system resilience to information flow,” proposed by

Hernandez Rodriguez and Kumar,¹ which refers to the ability of a system to remain in a stable range of the organizational state of its causal structure that sustains the flow of information, even under the effect of an intervention.

Based on the findings on Hernandez Rodriguez and Kumar¹ regarding the evolution of the causal structure followed by the causal history analysis implemented here, we argue that the system copes with an intervention using two strategies. First, the system can adjust to a minimal arrangement of causal interactions to favor a larger information flow from the causal history of each variable, as previously discussed. Second, although the system is mainly sustained by short-term dominated processes, it temporarily allows for significant support from the distant history to maintain its dynamics (Fig. 6, cluster #2). Notice that the total information among clusters is larger for WS , CO_2 , and H_2O , which are the variables that support their dynamics in both the immediate and distant causal history (Fig. 4). On the other hand, the self-dependant variables, U_z and T , highly depend on their immediate causal history regardless of the cluster.

Based on our analyses, we postulate that in the evolutionary dynamics of the ecohydrological system at turbulent scales, functional differences in the information flow of the system are the reflection of causal structural differences promoted by the system

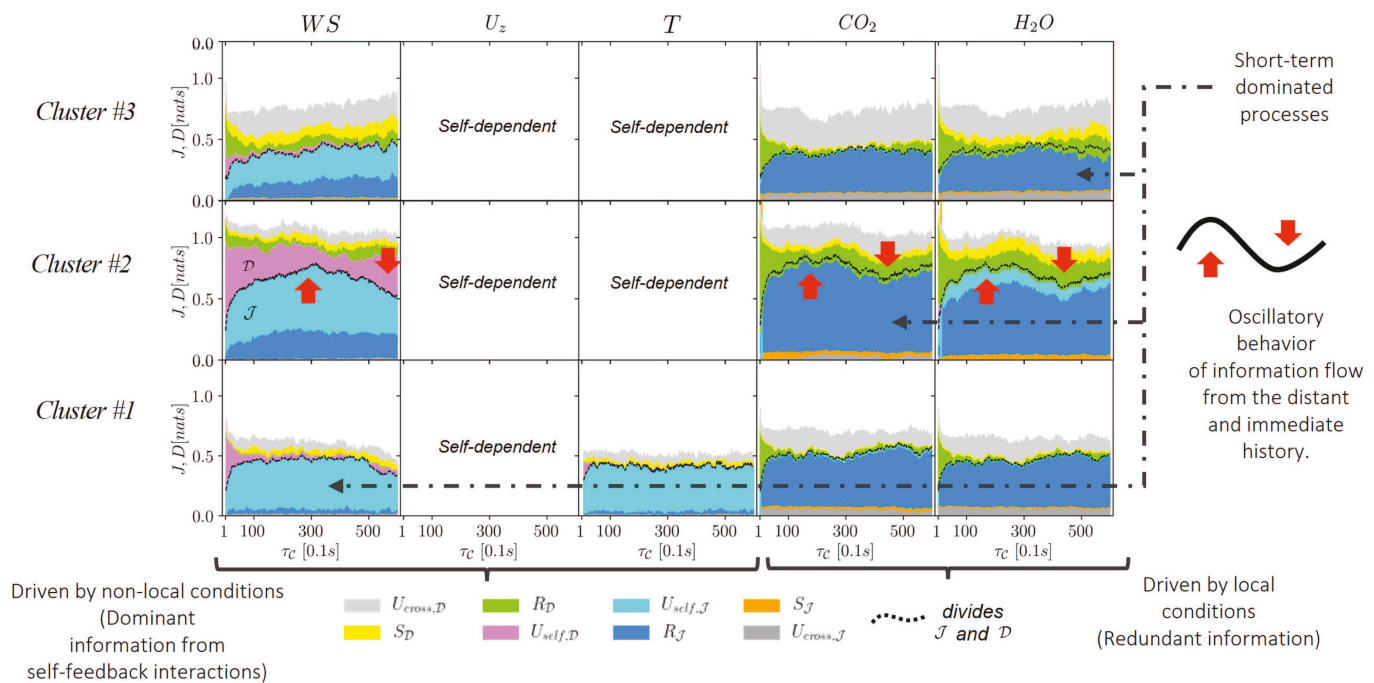


FIG. 6. Partitioning of information using PID for the self- and cross-dependencies in the immediate, $\mathcal{J}$, and distant, $\mathcal{D}$, causal histories^{5,12} of each variable (in columns) and cluster (in rows) based on the directed acyclic graphs for time series in Fig. 2. The information is decomposed into the complementary synergistic information, S_T ; redundant information, R_T ; and the unique contributions from self, U_{self} and cross-dependencies, U_{cross} . The black-dotted lines set the division between the immediate, $\mathcal{J}$ (below the black-dotted line), and distant histories, $\mathcal{D}$ (above the black-dotted line), respectively.

itself to maintain its dynamics, where functional differences are detected by the amount and type of dependencies using a SUR approach in a causal history framework.

D. Causal interdependencies as the cause for long-range dependencies

In our results for cluster #2, we noticed an oscillatory behavior in the information transferred from the immediate and distant histories as a function of τ_c for both sets of variables that have local and non-local influence. This particular behavior was first noticed in the ratio between the distant and total information for temperature, a predominantly self-dependent variable (Fig. 4). Later, we observe the same behavior for the cross-dependent variables in the PID analysis (Fig. 6). For instance, for the horizontal wind speed, WS , the contribution of the unique information from self-dependency in the immediate history, $U_{self,\mathcal{J}}$ shows an oscillatory negative correlation with the unique information from self-dependency in the distant history, $U_{self,\mathcal{D}}$ (Fig. 6). This behavior is also observed for the cross-dependent local variables, CO_2 and H_2O , which share a similar periodic signal for the redundant information, $R_{\mathcal{J}}$. Therefore, in terms of information flow from the causal history, this oscillatory behavior reveals that the system can experience transitions in the ratio of the immediate to distant history over time.

Our results suggest that the oscillatory behavior in cluster #2 is related to the frozen turbulence described by Taylor's hypothesis.²³

Eddies could be changing size and shape as they drift by the sensor. Taylor's hypothesis holds when the turbulence intensity is small relative to the mean wind speed.²³ As stated in Hernandez Rodriguez and Kumar,¹ in cluster #2, the dynamics of the horizontal WS , and vertical wind speed, U_z , are decoupled from the thermodynamics, T . In such a stagnant environment in which the system temporarily exists, the mean wind is not contributing to the vertical movement of scalars, such that eddies can change in composition as are advected past the sensor, resulting in a larger amount of information coming alternatively from the immediate and distant history of each variable carried by the eddies.

Looking at the spectral behavior of the data, we observe that the high frequency data exhibit fractal $1/f$ scaling (Fig. 7), which refers to the inverse proportionality between spectral power and frequency.²⁴ We also observe the prevalence of $1/f$ noise,²⁵ also called Flicker noise²⁶ or pink noise spectrum, which is characterized by a wide, flat plateau at high frequencies. Typically, the effect of periodic thermal or electrical fluctuations over time is used to explain the $1/f$ scaling noise in time series,²⁷ but no causal relationship has been previously established regarding the information flow in the system.

We argue that the self-dependency dominance in the causal history causes the absence of power-law dependence in the spectral plot of a variable. When there is no power-law dependence, we observe a noise plateau at high frequencies, such as in the case of temperature and the vertical wind velocity component, U_z (Fig. 7).

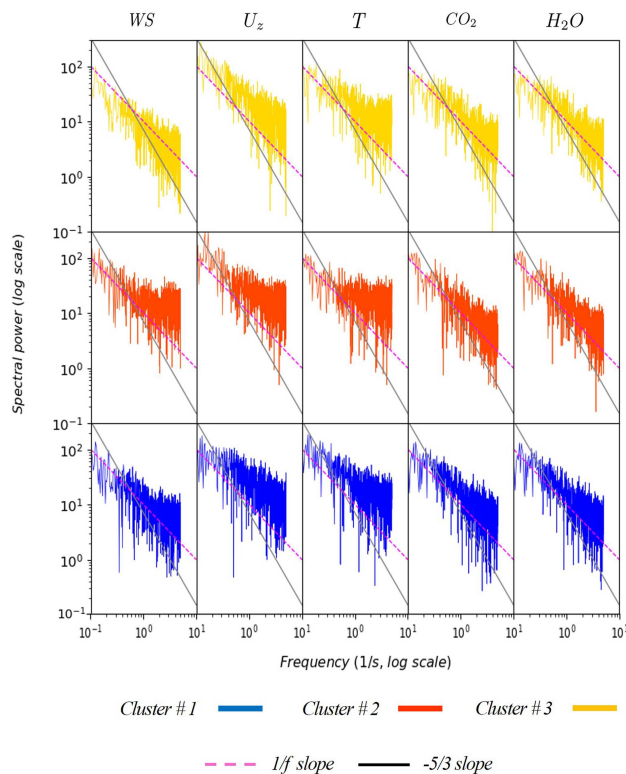


FIG. 7. Fourier power spectrum of the time series of variables at high frequency. Each row refers to each of the clusters in Fig. 2. The spectra show a noise plateau at high frequencies and $1/f$ slope at low frequencies (magenta dashed line). The Kolmogorov spectrum slope (solid black line) has the form of $E \propto k^{-5/3}$, which is supposed to fit most of the behavior at turbulence scale.^{28,29}

In other words, the temperature is flattened out at high frequencies in the spectral plot, and therefore, the temperature is not following the turbulence power law. There is a basis to infer that when cross-dependent short-term dominated processes increase their self-dependency on their distant history, some long-range dependencies appear. For instance, in cluster #2 we observe the oscillatory behavior for all variables and power-law independence at high frequencies caused by an increase in the self-dependence on its distant history. In the case of WS, it is due to the additional unique information from self-dependency in the distant history, $U_{self,\mathcal{D}}$. For CO_2 and H_2O , it is due to the larger dependence on the redundant information from the distant history, $R_{\mathcal{D}}$ (Fig. 6). Therefore, we postulate that increments of the self-dependence of a variable, especially relying on its distant causal history, are reflected in its spectral behavior as a lack of power-law dependency.

IV. DISCUSSION: DUALITY BETWEEN CAUSAL STRUCTURE AND FUNCTIONALITY

Causal structure and its functionality form a relationship that evolves to meet the information flow needs of the system. When we look at the dynamics of turbulence at the biosphere–atmosphere

interface, carbon dioxide, water vapor, and other scalars are transported away from the Earth's surface in a well-developed physical mechanism called ejection-sweep.³⁰ In turbulence, coherent structures called eddies exhibit recurrent patterns in the presence of a boundary, such as the Earth's surface.^{23,31} Momentum and energy-containing eddies, which differ on many orders of magnitude, self-sustain the cycle of inception of new eddies, and there is dissipation.^{14,23} In Lozano-Durán, Bae, and Encinar,¹⁴ causal interactions were estimated among eddies using numerical simulations to determine how the knowledge of the past states of eddies reduces the uncertainty of their future states. Their results show that energy eddies in the buffer and logarithmic layers in the atmosphere are similar and independent of the eddy size. Here, we have taken the causal look further using the dynamics of turbulence at the biosphere–atmosphere interface, beyond the univariate analysis of the exploration of the behavior of wind features, by looking at the fluctuations of a multivariate complex system that encode information about interactions in the time series at high frequency. In this multivariate system, temperature, scalars such as CO_2 and H_2O , along with wind velocity components, create a causal structure of interdependencies that sustain information flow and evolve inside a range of dynamical behavior instead of being in a fixed configuration.¹ Our analysis is based on the hypothesis that causal relationships are influenced by information flow, which in turn sustains the functionality of the ecohydrological system at high frequencies through interactions. Moreover, the causal structure shapes the network of information flow among variables in the multivariate complex system and serves as the foundation for the functional representation of the evolution of the dynamics of the system.

V. CONCLUSIONS

We explore how the causal structure of the dynamics of an ecohydrological system is related to its functionality. We use high-frequency observations and information metrics to infer causality from interactions among multiple components of an ecohydrological system. Here, we see the propagating fluctuations among interacting components in the system as information flow. Using the DAGs resulting from the clustering analysis to determine the evolution of the system causal structure,¹ our results show differences in their information flow. We conclude that functional differences in the information flow of the system are a reflection of causal structural differences promoted by the system itself to maintain its dynamics.

We demonstrate how information theory provides a framework to determine the causal structure and functionality of ecohydrological systems. The IT-based methods developed and/or used here allow the inference of the relationship between the causal structure and functionality of the evolving system. Based on the results of the causal structure analysis in Hernandez Rodriguez and Kumar,¹ and using a causal history approach^{4,5} that implements a partial information decomposition framework,^{12,13} our results suggest that at a turbulent scale, the processes that shape the dynamics of the ecohydrological system at the biosphere–atmosphere interface are mainly sustained by short-term dynamics. In addition, each process can have a predominantly local origin, such as for CO_2 and H_2O , or nonlocal origin in the case of WS, U_z , and T . Our results show

that the causal structure can be shaped by the system to meet the information flow needs of the evolving system. In addition, we find that the system can reduce its dimensionality to facilitate the transfer of large amounts of information among its components, especially under the effect of an intervention. Ultimately, we found that the cause for long-term dependencies in turbulent systems that do not follow a power law in the spectral plot is caused by self-dependencies coming from the distant causal history, even when the dynamics are mainly sustained by its short-term dynamics.

In the study of causal inference of ecohydrological systems, we recognized the importance of developing and implementing causal approaches to infer the dynamics of processes when using time series and other informative datasets. In the machine learning and big data era, we recognize the need of developing and implementing tools for causal inference in geosciences. Our analyses reveal the dynamics of the multivariate system at high frequency and provide some new avenues for causality to be integral to artificial intelligence and machine learning-driven approaches, which are also sustained and rely on observations.

ACKNOWLEDGMENTS

This research was funded by NSF Grant No. EAR-1331906 for the Critical Zone Observatory for Intensively Managed Landscapes (IMLCZO), NSF Grant Nos. OAC-1835834 and EAR-2012850 for the Critical Interface Network for Intensively Managed Landscapes (CINet), and ARPA-E Grant No. DE-AR0001225. We thank Peishi Jiang at Pacific Northwest National Laboratory (PNNL) for his support when using the Causal History open-source program <https://github.com/HydroComplexity/CausalHistory>. We also thank Steve Sargent at the Illinois State Geological Survey for the collection of the high frequency data from the flux tower at Goose Creek.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author's Contributions

All authors contributed equally to this work.

Leila Constanza Hernandez Rodriguez: Conceptualization (equal); Methodology (equal); Writing – original draft (lead); Writing – review & editing (equal). **Praveen Kumar:** Conceptualization (equal); Methodology (equal); Supervision (lead); Writing – review & editing (equal).

NOMENCLATURE

CO ₂	Carbon dioxide molar concentration (kg/m ³)
H ₂ O	Water vapor molar concentration (g/m ³)
T	Air temperature (°C)
WS	Horizontal wind speed (m/s)
U _z	Vertical wind speed (m/s)
CMI	Conditional mutual information
CMI _{ℓnn}	Conditional mutual information based on ℓnn estimation

DAG	Directed acyclic graph
DI	Dunn index
EC	Eddy covariance
FFT	Fast Fourier transform
IT	Information theory
MIT	Momentary information transfer
MITP	Momentary information transfer along causal path
MIWTR	Momentary information weighted transitive reduction
MPID	Momentary partial information decomposition
PID	Partial information decomposition
TE	Transfer entropy
TR	Transitive reduction
WTR	Weighted transitive reduction
HH:MM:SS	Hour:Minute:Second
CST	Central Standard Time
US	United States of America

DATA AVAILABILITY

The directed acyclic graph for time series are estimated using the Tigramite software package.^{7–9,32} The momentary information weighted transitive reduction (MIWTR) and the multivariate information flow are estimated by using the Causal History package^{4,5,11} <https://github.com/HydroComplexity/CausalHistory>. The codes for the evolution of the Causal Structure are available in GitHub at <https://github.com/HydroComplexity/CausalStructureEvolution>.

REFERENCES

- L. C. Hernandez Rodriguez and P. Kumar, "Causal interaction in high frequency turbulence at the biosphere-atmosphere interface: Structural behavior," *Chaos* **33**, 073108 (2023).
- A. E. Goodwell, P. Jiang, B. L. Ruddell, and P. Kumar, "Debates—Does information theory provide a new paradigm for Earth science? Causality, interaction, and feedback," *Water Resources Res.* **56**(2), e2019WR024940 (2020).
- P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman, *Causation, Prediction, and Search* (MIT Press, 2000).
- P. Jiang and P. Kumar, "Information transfer from causal history in complex system dynamics," *Phys. Rev. E* **99**, 012306 (2019).
- P. Jiang and P. Kumar, "Using information flow for whole system understanding from component dynamics," *Water Resources Res.* **55**(11), 8305–8329 (2019).
- J. Runge, "Quantifying information transfer and mediation along causal pathways in complex systems," *Phys. Rev. E* **92**, 062829 (2015).
- J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, "Escaping the curse of dimensionality in estimating multivariate transfer entropy," *Phys. Rev. Lett.* **108**, 258701 (2012b).
- J. Runge, V. Petoukhov, J. F. Donges, J. Hlinka, N. Jajcay, M. Vejmelka, D. Hartman, N. Marwan, M. Paluš, and J. Kurths, "Identifying causal gateways and mediators in complex spatio-temporal systems," *Nat. Commun.* **6**, 1–10 (2015).
- J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic, "Detecting causal associations in large nonlinear time series datasets," *arXiv:1702.07007* (2017).
- C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois, Urbana, IL, 1949). Vol. 10, p. 117.
- P. Jiang and P. Kumar, "Interactions of information transfer along separable causal paths," *Phys. Rev. E* **97**, 042310 (2018).
- A. E. Goodwell and P. Kumar, "Temporal information partitioning networks (TIPNets): A process network approach to infer ecohydrologic shifts," *Water Resources Res.* **53**, 5899–5919, <http://doi.org/10.1002/2016WR020218> (2017).

- ¹³P. L. Williams and R. D. Beer, "Nonnegative decomposition of multivariate information," [arXiv:1004.2515](#) (2010).
- ¹⁴A. Lozano-Durán, H. J. Bae, and M. P. Encinar, "Causality of energy-containing eddies in wall turbulence," *J. Fluid Mech.* **882**, A2 (2020).
- ¹⁵C. G. Belinchon, "Multiscale information transfer in turbulence," Ph.D. thesis (Université de Lyon, 2018).
- ¹⁶G. Tissot, A. Lozano-Durán, L. Cordier, J. Jiménez, and B. R. Noack, "Granger causality in wall-bounded turbulence," *J. Phys.: Conf. Ser.* **506**, 012006 (2014).
- ¹⁷R. Cerbus and W. Goldburg, "Information content of turbulence," *Phys. Rev. E* **88**, 053012 (2013).
- ¹⁸J. Runge, J. Heitzig, N. Marwan, and J. Kurths, "Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy," *Phys. Rev. E* **86**, 061121 (2012).
- ¹⁹S. L. Lauritzen, A. P. Dawid, B. N. Larsen, and H.-G. Leimer, "Independence properties of directed Markov fields," *Networks* **20**, 491–505 (1990).
- ²⁰S. L. Gray and R. G. Harrison, "Eclipse-induced wind changes over the British Isles on the 20 March 2015," *Philos. Trans. R. Soc. A* **374**, 20150224 (2016).
- ²¹D. Beverly, "Phenotypic and topographical controls of ecohydrological processes from leaf to ecosystem," Ph.D. thesis (University of Wyoming, 2021).
- ²²T. Foken, B. Wichura, O. Klemm, J. Gerchau, M. Winterhalter, and T. Weidinger, "Micrometeorological measurements during the total solar eclipse of August 11, 1999," *Meteorol. Z.* **10**, 171–178 (2001).
- ²³R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Springer Science & Business Media, 2012), Vol. 13.
- ²⁴J. W. Kirchner and C. Neal, "Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection," *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12213–12218 (2013).
- ²⁵A. Van der Ziel, *Noise in Measurements* (Wiley, New York, 1976).
- ²⁶D. Halford, "A general mechanical model for $|f|$ spectral density random noise with special reference to flicker noise $1/|f|$," *Proc. IEEE* **56**, 251–258 (1968).
- ²⁷A. Mennella, C. Baccigalupi, A. Balbi, M. Bersanelli, C. Burigana, C. Butler, B. Cappellini, G. De Gasperis, F. Hansen, D. Maino, and N. Mandolesi, "Imaging the first light: Experimental challenges and future perspectives in the observation of the cosmic microwave background anisotropy," [arXiv:astro-ph/0402528](#) (2004).
- ²⁸B. Ryden, "Basic turbulence," <http://www.astronomy.ohio-state.edu/~ryden/ast825/ch7.pdf>.
- ²⁹D. G. Ortiz-Suslow and Q. Wang, "An evaluation of Kolmogorov's $-5/3$ power law observed within the turbulent airflow above the ocean," *Geophys. Res. Lett.* **46**, 14901–14911, <http://doi.org/10.1029/2019GL085083> (2019).
- ³⁰G. Katul, G. Kuhn, J. Schieldge, and C.-I. Hsieh, "The ejection-sweep character of scalar fluxes in the unstable surface layer," *Boundary-Layer Meteorol.* **83**, 1–26 (1997).
- ³¹L. F. Richardson, *Weather Prediction by Numerical Process* (Cambridge University Press, 2007).
- ³²J. Runge, V. Petoukhov, and J. Kurths, "Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models," *J. Clim.* **27**, 720–739 (2014).