TEMPORAL INFORMATION PARTITIONING NETWORKS TO INFER ECOHYDROLOGIC BEHAVIORS

BY

ALLISON E. GOODWELL

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Civil and Environmental Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Praveen Kumar
Professor Ximing Cai
Professor Megan Konar
Professor Efi Foufoula-Georgiou, University of California, Irvine
Professor Hoshin Gupta, University of Arizona

Abstract

An ecohydrologic system is a complex network, in which the shifting behavior of individual components and the connectivity between them determines the dynamics. This connectivity between components can act to constrain, accentuate, or otherwise modify the variability of individuals. In an ecohydrologic system, connectivity exists in the form of many time-dependent relationships between states and fluxes related to water, energy, nutrients, soils, and vegetation. Although relationships are constrained by conservation laws, they exhibit a wide range of variability at many timescales due to non-linear interactions, threshold behavior, forcing, and feedback. Moreover, these aspects of connectivity and variability exist at a single location or over a spatial gradient. The understanding of this connectivity within the system as a whole necessitates an appropriate framework, in which evolving interactions are identified from time-series observations.

The goals of this thesis are to (i) develop a Temporal Information Partitioning Network (TIPNet) framework for understanding the joint variability of network components as characterized by time-series data, and (ii) apply this framework to understand ecohydrologic systems across climate gradients based on flux tower and weather station observations. In the TIPNet framework, nodes in the network are time-series variables, and links are information theoretic measures that quantify multivariate lagged time-dependencies from lagged "source" nodes to "target" nodes. The strength of this framework is its ability to characterize information flow between variables over short time windows, and further distinguish aspects of unique, redundant, and synergistic dependencies. Redundant information is overlapping information provided by multiple sources to a target, unique information is only provided by a single target, and synergistic information is provided only when two or more sources are known together.

Based on data from three Critical Zone Observatories, we find that network structure shifts according to conditions at sub-daily time scales and constraints imposed by seasonal energy and water availability. TIPNets constructed from 1-minute weather station data reveal shifts in time-scales and levels of uniqueness, synergy, and redundancy between wet and dry

conditions. A more complex network of synergistic interactions characterizes several-hour windows when surfaces are wet, and peaks in information flow during the growing season correspond to shifts in precipitation patterns. Networks based on half hourly flux tower data reveal seasonal shifts in the nature of forcing to carbon and heat fluxes from radiation, atmospheric, and soil subsystems. Along two study transects, we attribute variability in heat and carbon fluxes within constraints imposed by energy and moisture availability to joint interactions that are more synergistic in the spring and redundant in the fall.

Finally, we explore the nature of information flow along an elevation gradient from flux towers located along a transect to gage local versus non-local connectivity. While the strength of shared information between variables at a site reflects local connectivity, shared information between variables at different sites reflects non-local connectivity. Along two elevation transects, we find that information flow between distant sites indicates directional connectivity that is related to dominant weather patterns. At the Southern Sierra CZO in California, non-local information flow is dominantly west to east, corresponding to weather forcing from the Pacific Ocean eastward, while non-local flow has less directionality at Reynolds Creek CZO, where sites are much closer together and there is no dominant weather forcing direction along the transect. The developed framework and applications presented in this thesis reveal the common presence of multivariate process interactions at timescales from minutes to hours, many of which would not be detected using traditional approaches.

For an ecohydrologic system, the complex network of relationships dictates ecosystem resilience to perturbations such as climate change, drought, or human influences. More broadly, the methods and framework developed here contribute toward a holistic understanding of complex systems, and are applicable to a range of studies of evolving networks.

Acknowledgments

Thank you to my advisor, Dr. Praveen Kumar, for his valuable insight and support over the past several years. I am also grateful for the past and present Kumar-research group team members. Sayo Chaoka, Juan Quijano, Venkatraman Srinivasan, and Phong Le met me as a prospective student, and did not quite manage to scare me away. Debsunder Dutta, Dongkook Woo, and Qina Yan were cohorts through most of my graduate school experience and shared in many EOH project mishaps and early-CZO field efforts. Harpreet Kaur, Susana Roque, Meredith Richardson, Kunxuan Wang, Esther Lee, Peishi Jiang, Derek Wagner, and Leila Hernandez did a great job of making me feel like I knew what was going on. I also thank my past and present "hydro-family" for fun adventures, memorable conversations, and invaluable friendships. Finally, I am very thankful for my "real-family" for their unconditional love, patience, and fielding years of inquiry into "when is Allison going to be done with school and how is she paying for it".

Speaking of paying for it, I was funded for my PhD through several sources: the Support for Underrepresented Groups in Engineering (SURGE) Fellowship, the NASA Earth and Space Science Fellowship (NESSF #NNX15AN55H), a teaching assistantship, and a research assistantship as part of the NSF-WSC Resilience Under Accelerated Change (REACH) project (# CBET 1209402). This work was also made possible with support from the NSF CZO, with data and assistance from members of the Intensively Managed Landscape (IMLCZO # EAR-1331906), Reynolds Creek, and Southern Sierra CZOs. Finally, I benefited from several discussions with my PhD committee members and fellow participants of an Information Theory in the Geosciences Workshop. All of these sources are gratefully acknowledged.

TIPNet computations described in this thesis and in the appendix can be carried out using a MATLAB program available at the following link:
https://github.com/HydroComplexity/TIPNet. Weather station data is also available at that link, and CZO flux tower data is available at:
http://criticalzone.org/national/data/.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Information Theory Measures and Chaotic Logistic Networks

$H(X)$	Shannon entropy (*bits*)
I_τ	Lagged mutual information (*bits*)
$I(X,Y;Z)$	Total shared information of X and Y with Z (*bits*)
$I(X;Y;Z)$	Interaction information between 3 variables X, Y, and Z (*bits*)
$I(X,Z\|Y)$	Conditional mutual information (*bits*)
T_E	Transfer Entropy, a special case of conditional mutual information (*bits*)
T/I	Proposed measure of conditional mutual information relative to total multivariate shared information (*bits/bit*)
X_{s1}	refers to a lagged source node that provides information
X_{tar}	refers to a target node, a node at the current time step that receives information
ϵ, ϵ_z	coupling and noise strengths, respectively, for generated networks
n_{steps}	number of time steps in an interval for *pdf* computations
p	link probability for a generated network
pdf	probability density function
τ	time lag associated with information measure, units of number of time steps or measurement interval (e.g. minutes)
t	current time step, e.g. $X(t-\tau)$ indicates time-series X lagged by τ time steps
z	uniform random noise component for a generated network

Information Partitioning

I_s	Mutual information between two sources normalized by their entropies, used as measure of scaled source dependency (*bits/bit*)
R	Redundant information (*bits*)
R_{min}	Minimum bound for Redundancy (*bits*)
R_{MMI}	Maximum bound for Redundancy, equivalent to minimum mutual information (*bits*)
R_s	Rescaled Redundancy, scaled between R_{min} and R_{MMI} (*bits*)
S	Synergistic information (*bits*)
S_{MMI}	Synergistic information computed based on R_{MMI} (*bits*)
S_s	Synergistic information computed based on R_s (*bits*)

U	Unique information ($bits$)
U_{MMI}	Unique information computed based on R_{MMI} ($bits$)
U_s	Unique information computed based on R_s ($bits$)

Study Sites

CZO	Critical Zone Observatory
RC	Reynolds Creek CZO, Idaho, flux tower transect
SFP	Sangamon Forest Preserve, weather station site
SS	Southern Sierra CZO, California, flux tower transect

Time-Series Variables

G	ground heat flux (W/m^2)
H	sensible heat flux (W/m^2)
H2O	water vapor density
LE	latent heat flux (W/m^2)
LWet	leaf wetness (dielectric counts of LWS sensor)
$\mathbf{LW_{in}}$	incoming longwave radiation (W/m^2)
$\mathbf{LW_{out}}$	outgoing longwave radiation (W/m^2)
Pa	air pressure (kPa)
PPT	precipitation (mm)
RH	relative humidity
$\mathbf{SW_{in}}$	incoming shortwave radiation (W/m^2)
$\mathbf{SW_{out}}$	outgoing shortwave radiation (W/m^2)
Ta	air temperature ($^\circ C$)
$\mathbf{T_{soil}}$	soil temperature (5 cm depth) ($^\circ C$)
$\mathbf{T_{surf}}$	surface temperature ($^\circ C$)
VWC	volumetric water content
WD	wind direction (degrees from N)
WS	wind speed (m/s)

CHAPTER 1

INTRODUCTION

Ecohydrologic systems can be conceptualized as complex networks in which processes occurring in the atmospheric, soil, and vegetation systems are connected through various mechanisms. While components such as fluxes of heat, nutrients, and water are constrained by water and energy balances, they otherwise exhibit a large range of variabililty. The extent to which the variability of some components can be related to that of others indicates the level of connectivity between various processes. In a complex system of interacting parts, this connectivity can involve forcing behaviors, where a fluctuation in one variable propagates to other variables, or feedbacks that dissolve the notion of "cause" and "effect" since the system becomes partially self-regulating [Ruddell and Kumar, 2009a].

In a natural system, process connectivity [Sendrowski and Passalacqua, 2017] exists at time scales on the order of seconds, as in the case of vegetation responses to changes in temperature and moisture, to decades, as in the case of the co-evolution of landscapes and ecosystems. Moreover, connectivity can be local or non-local relative to a specific area, since forcing and feedbacks may exist over a spatial gradient due to atmospheric mixing or processes that occur from upstream to downstream locations. As limiting factors such as water, nutrient, or energy availability shift over time, fluxes and states exhibit non-linear or threshold behaviors. For example, vegetation growth in the form of photosynthetic rate may be limited by either water, nutrients, or sunlight, and will respond most drastically to fluctuations in the limiting factor. In an ecohydrologic system, the entire network of these interactions may give rise to emergent properties [Jørgensen et al., 1992, Strogatz, 2001] such as ecosystem resiliency to drought, climate change, or other environmental perturbations. As such, we seek to develop models to predict these properties under the influence of changing climate and weather patterns. While these models tend to capture the constraints imposed by water and energy availability, they do not include or reveal the full range of connectivity in terms of joint variations within the natural system.

The goal of this thesis is to characterize ecohydrologic process connectivity as a complex network in which individual variables related to energy, carbon, and water fluxes interact in

different ways. The observation and detection of these changing interactions can enhance prediction and modeling efforts, and may reveal drivers behind ecosystem responses to different types of perturbations. This work contributes to a more holistic analysis of system dynamics in a natural laboratory, in that it characterizes a network as a set of multivariate dependencies instead of pairwise links or individual components. This can benefit the understanding of many types of complex systems in which many components are jointly interacting. In this introductory section, we discuss background relevant to ecohydrologic process networks and our use of information theoretic measures for their characterization. We then outline the original contributions and the organizational structure of this thesis.

1.1 Background

A process network is a collection of linked time-series variables, where a link may indicate a flow of energy, mass, or information [Kumar and Ruddell, 2010]. We take an approach where ecohydrologic variables, such as temperatures, heat fluxes, wind, soil states, or water vapor are nodes in the network, which are linked through time-dependent information flows. In this thesis, multivariate measures based on information theory [Shannon, 1948] are used to define these time-dependent links. Various types of systems such as synapses in the human brain [Niso et al., 2013], industrial processes [Duan et al., 2013], climate-vegetation-soil relationships [Ruddell and Kumar, 2009a], and connectivity in deltaic systems [Sendrowski and Passalacqua, 2017] have all been studied as process networks.

1.1.1 Variability and dynamics in complex networks

Emergent behaviors, such as synchronization or intricate patterns in nature, are found to "emerge" from a combination of simple behaviors [Holland, 2000]. For example, fireflies exhibit synchronized blinking, and large groups of animals such as fish or birds appear to move as a single entity [Strogatz, 2001]. In these situations, knowledge of an individual component, such as a single firefly, cannot explain the large-scale behavior, such as the synchronized blinking of a multitude of fireflies. While components of an ecohydrologic system represent different types of variables and do not fully synchronize, larger-scale behaviors can involve temporal or spatial patterns or properties such as resilience, the ability of the system to return to a stable state after a perturbation. Here, the knowledge of an individual variable,

such as incoming shortwave radiation, cannot fully explain variations in air temperature or heat fluxes, much less aspects of vegetation composition.

To generate a simple example of synchronizing time-series, we link time-series variables via a coupled chaotic logistic map (Chapter 2, [Goodwell and Kumar, 2015]). As the connectivity increases via increasing node degrees or coupling strengths, nodes become completely synchronized due to induced feedbacks. In this case, although the functional form of interactions and the participating nodes are known, the resulting properties cannot be fully described based on linear measures. Between the extremes of independently acting variables and complete synchronization, we explore how nodes respond to each other or external forcing at various time scales and strengths. Identification of such dependencies from time-series can reveal the complex behavior of the system as a whole. Moreover, shifts in time dependencies or driving nodes in process networks could identify behavioral shifts due to perturbations that could indicate alterations in important components of the system.

Additional challenges in the characterization of connectivity within an ecohydrologic system include the unknown and complex nature of functional forms of relationships in addition to different interaction mechanisms. In a modeling context, model structure, limited forcing data, and rigid or time-invariant parameters do not include the many shifting dependencies that explain the full range of variability of a single component. For example, while air temperature (Ta) and wind speed (WS) both influence relative humidity (RH), their influences are through different mechanisms involving water vapor holding capacity and atmospheric mixing, respectively ([Goodwell and Kumar, 2017a], Chapter 3). In this case, each of these interactions may be represented to some degree in a model, but the joint influence of fluctuations in both driving variables is unknown. Due to these challenges of multiple components, unknown or shifting functional forms, and different types of processes, the characterization of complex ecohydrologic systems necessitates a framework that accounts for joint dependencies that represent different interaction mechanisms and time scales. We will next address the chacterization of time-dependent network links based on information theory measures.

1.1.2 Information theory and process networks

Information theory is based on Shannon Entropy $H(X)$ [Shannon, 1948], a measure of uncertainty in a random variable X:

$$H(X) = -\sum p(x) \log_2(p(x)) \tag{1.1}$$

Due to the natural logarithm, $H(X)$ has units of *bits*. Intuitively, $H(X)$ is the average number of binary ("yes" or "no") questions needed to determine the value of X. For example, if X is the result of a coin flip, $H(X) = 1$ bit, indicating the single question ("is it heads or tails?") needed to determine the outcome. While $H(X)$ represents the information contained in a variable, mutual information is the shared information between two variables, in that it is the reduction in uncertainty of one variable given knowledge of the other. Information theory enables the identification of non-linear dependencies common in natural systems that may not be detected using other statistical techniques such as correlation.

Higher-order information theory metrics such as Transfer Entropy (T_E) [Schreiber, 2000, Runge et al., 2012, Ruddell and Kumar, 2009a, Lee et al., 2012, Hlaváčková-Schindler et al., 2007], partial mutual information [Frenzel and Pompe, 2007] and interaction information [Timme et al., 2014] have been used to explore time dependent relationships in the form of forcing and feedback relationships. In Earth system studies, information theory has been used to measure complexity, interdependencies, and causality [Balasis et al., 2013]. More broadly, information theory has been applied to gain insight into complex interactions between multiple variables in fields such as neuroscience [Barrett, 2015], cardiology [Faes et al., 2015], and industrial engineering [Duan et al., 2013].

T_E, specifically, has been used in a network context and to quantify an "information flow" between lagged "source" and "target" variables [Ruddell and Kumar, 2009a, Sendrowski and Passalacqua, 2017]. T_E is a special case of conditional mutual information, in which mutual information between the source and the target is conditioned on the history of the target variable. In other words, T_E is the reduction in uncertainty of the target due to knowledge of the source that is not already reduced by the target's history. T_E has been used to reveal shifts in forcing and feedback between ecohydrologic fluxes as measured from flux tower data [Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b, Kumar and Ruddell, 2010] and process connectivity in a delta [Sendrowski and Passalacqua, 2017]. However, a limitation of T_E applications is that since it is based on a 3d *pdf*, it does not account for intermediate or multiple drivers that may account for the detected information flow [Runge, 2015, James et al., 2016]. For example, if statistically significant T_E is detected between two nodes X and Y, the knowledge of a third variable Z could reveal that the information transfer is actually due to joint forcing of X and Y by Z [Duan et al., 2013].

An important requirement of using information theory is to estimate robust probability den-

sity functions (*pdf*s) from data. In ecohydrologic applications, data may be of limited length, have gaps, or exhibit a mixed distribution with an atom-at-zero (hereafter referred to as a zero-effect). Additionally, the dominant diurnal cycle may overshadow joint fluctuations between many variables. Various filtering techniques, estimation methods, and tests of statistical significance [Lee et al., 2012, Gong et al., 2014, Goodwell and Kumar, 2015, Silverman, 1986, Sendrowski and Passalacqua, 2017, Ruddell and Kumar, 2009a] can be used to address these issues. Higher order information theory measures such as Direct T_E [Duan et al., 2013] have been proposed that condition on multiple variables to distinguish causal drivers from intermediate influences. However, these measures require 4d or higher dimensional *pdf*s, which are only robust given a very large amount of data. While the estimation of high-dimensional *pdf*s enables a more holistic network definition in which many variables are interrelated, we limit our studies to metrics based on 3d *pdf*s to enable analyses based on short sequences of data.

In addition to the robust estimation of information theory measures, the proper interpretation of joint and conditional multivariate measures is also important. Even when accurately computed, it has been shown that measures like T_E can either overestimates or underestimate of influences and be subject to "spurious" detections due to induced feedbacks or intermediate drivers as previously discussed [Smirnov, 2013]. These issues indicate the need for a thorough understanding of different types of dependencies when using information theory to infer system dynamics.

1.1.3 Temporal Information Partitioning

In a multivariate case where multiple source nodes influence a target, information provided from a source to a target node can be partitioned into redundant (R), synergistic (S) or unique (U) components [Williams and Beer, 2010, Barrett, 2015]. R is overlapping information that two or more sources share with a target, S is information obtained only when two sources are known together, and U is information that only a single source provides. An example case where two sources only provide S is the exclusive-OR (XOR) operation, for which the sources appear independent of the target unless both are known together (Figure 1.1a). Similarly, information is completely redundant when two correlated (or synchronized, as in Figure 1.1b) sources are found to provide the same information to a target. Finally, solely U exists when only a single source provides information to a target variable (Figure 1.1c). Between these extremes, a given joint relationship may be composed of a combination of U, R, and S components.

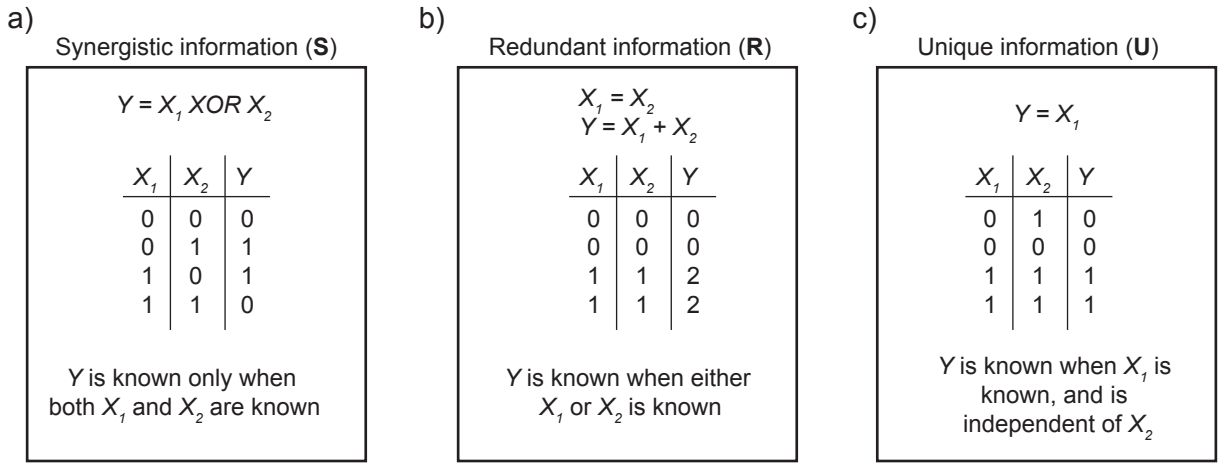

Figure 1.1: Examples of completely synergistic, redundant, and unique relationships. Only synergistic (S), redundant (R), or unique (U) information is provided to a target variable Y from sources X_1 and X_2. *(a)* For a binary XOR operation, only S exists since both X_1 and X_2 are needed to predict Y. *(b)* If sources X_1 and X_2 are redundant, or completely identical to each other, and Y is their sum, only R exists since either source could predict Y. *(c)* If X_2 is independent of Y so that Y is only a function of X_1, X_1 provides only U.

It can be shown that T_E is a combination of U and S components [Williams and Beer, 2010], such that the partitioning of information into U, R, and S provides more nuanced interpretation of information flow, specifically whether information flow as detected based on T_E arises uniquely from a given source, or together with the additional source. Other information measures, such as interaction information, can also be interpreted in terms of information partitioning.

While information partitioning allows us to define information flow in a more intuitive way than existing information theory measures, its application is less straightforward since there is no Shannon information measure to directly perform the partitioning. Several methods have been proposed to compute R, S, and U, each with different advantages and drawbacks [Williams and Beer, 2010, Harder et al., 2013, Griffith and Ho, 2015, Griffith and Koch, 2014, Barrett, 2015, Olbrich et al., 2015, Faes et al., 2015, Goodwell and Kumar, 2017a]. The most simple measure is the minimum mutual information (MMI) measure [Williams and Beer, 2010], which defines redundancy as the minimum of the mutual information values between each source and the target. However, this is actually a maximum bound for redundancy since it assumes that all information provided by the weaker of two sources is redundant, thus overestimate R in

cases where sources provide information independently. In this thesis, we give an overview of the interpretation of existing information theory measures based on information partitioning (Chapter 2) and introduce a new partitioning method that can be applied to environmental time-series data (Chapter 3). Specifically, we introduce Temporal Information Partitioning, in which information partitioning is applied to different time windows to gage how network properties shift over time.

1.1.4 Local and non-local process connectivity in ecohydrologic systems

As discussed previously, information theory measures, specifically information partitioning, provide powerful metrics to characterize time-dependent links and process networks in natural systems. Time-series data used in this thesis are from a weather station in central Illinois at the Intensively Managed Landscape CZO (Chapters 3 and 4) and flux tower transects at Reynolds Creek and Southern Sierra CZOs (Chapter 5). The 1-minute weather station dataset in Illinois enables us to construct process networks over several hour time windows, and reveal how dominant interactions and their timescales shift on very short timescales with moisture and radiation conditions. The 30-minute flux tower datasets are spaced along elevation gradients and include a wider range of time-series variables that measure radiation, heat fluxes, soil moisture and temperatures, carbon fluxes, and atmospheric states. The use of these datasets can reveal process connectivity between locations, in that some driving influences occur over a spatial gradient. In other words, variability that cannot be explained by the process network of interactions at a given site may be linked to external variability captured at other sites. Based on these data sets, we address the following questions related to process connectivity within ecohydrologic systems:

1. How do fluctuations in states such as temperature, humidity, wind, and radiation influence each other, and to what extent does the proportion of unique, synergistic, or redundant influence shift with weather or radiation conditions?

2. What are the timescales associated with detected interactions based on 1-minute data or half-hourly data, and what are the timescales of information flow between sites?

3. How do influences to heat flux partitioning and carbon fluxes shift over an elevation and precipitation gradient?

4. What are the strengths of information flow between ecohydrologic variables locally, or at a site, versus non-locally, or between sites, and how does this elucidate aspects of spatial process connectivity?

1.2 Research contributions

The original contributions of this study are as follows:

1. It is shown that information measures reflect levels of external forcing, noise, and induced feedbacks in chaotic logistic networks in various stages of synchronization that would otherwise go undetected.

2. A new measure for redundant information, Rescaled Redundancy (R_s), is developed to be appropriate for the study of environmental variables, for which time dependent links represent different types of processes.

3. This is the first application of temporal information partitioning in a network context, and specifically to study ecohydrologic systems. The developed framework, Temporal Information Partitioning Networks (TIPNets), is made available through a toolbox for educational purposes and future research.

4. The TIPNet framework is used to reveal sub-hourly interactions and feedbacks between weather variables that shift depending on radiation and moisture conditions.

5. It is shown that variability in heat and carbon fluxes within the constraints imposed by energy and moisture availability is attributed to joint interactions that change from synergy-dominated in the spring to redundancy-dominated in the fall.

6. Local and non-local information flow is established to represent spatial process connectivity between and within observation sites. It is shown that the directionality of non-local information and the proportion of local to non-local information correspond to weather patterns at two sites.

1.3 Organization

The chapters of this thesis are arranged as follows:

- In Chapter 2, chaotic-logistic networks are generated with a range of time delays (τ-distributions), coupling strengths (ϵ), topologies (Δ), and levels of noise, and information theory and variance measures are compared to the imposed network structures. Methods are evaluated in terms of correctly identified links, or statistically significant detections of information measures that correspond to imposed time dependencies. A network framework is introduced in which pairs of sources are analyzed for their joint provisions of information to target nodes. This and the interpretation of information theory measures as combinations of S, R, and U components paves the way for future applications of information partitioning to eco-hydrologic process networks.

- In Chapter 3, a new Rescaled Redundancy (R_s) measure is proposed to perform information partitioning. I analyze properties of the measure applied to increasingly complex 3-node cases involving die rolls, Gaussian sums with and without noise, a multivariate autoregressive process, and chaotic generated datasets. Temporal Information Partitioning is then applied to assess the nature of information transferred from air temperature (Ta) and wind speed (WS) to relative humidity (RH) as measured from 1-minute weather station data.

- In Chapter 4, the TIPNet framework is introduced to construct and analyze process networks based on time-series signals. TIPNet is applied to a 1-minute dataset of environmental signals from a weather station in Central Illinois. I discuss typical network behaviors over a 180-day study period in terms of time scale, interaction strength and type, and variability with weather conditions, and analyze short-term shifts in interactions that occur over several-hour periods.

- In Chapter 5, the TIPNet framework is applied to flux tower transect datasets along elevation gradients at the Southern Sierra and Reynolds Creek CZOs. Process connectivity is assessed at each site between atmospheric states, energy, water, and carbon fluxes for different points in the season. Local information flow, or information transferred between variables at a site, is compared to non-local information flow, or unique transfers between variables at different sites.

- Chapter 6 discusses conclusions of this study, broader implications, and avenues for future research.

- In the Appendix, further information is provided on *pdf* estimation, data sources, and the TIPNet Matlab interface.

CHAPTER 2

INFORMATION THEORETIC MEASURES TO INFER FEEDBACK DYNAMICS IN COUPLED LOGISTIC NETWORKS

2.1 Introduction [1]

A process network is a collection of time series variables that interact at different time scales [Kumar and Ruddell, 2010]. Each time series variable is a node in the network, and nodes are linked through time dependencies. Synapses in the human brain, industrial processes in a factory, or climate-vegetation-soil relationships can all be studied as process networks [Ruddell and Kumar, 2009a, Duan et al., 2013, Niso et al., 2013]. In each of these examples, a time dependent relationship between the history of a source node and current state of a target node defines a network link that has some strength, time scale, and directionality. A whole-network property such as average node degree, average link strength, or dominant time scale can define a "system state". It is important to correctly detect and interpret these links to reveal aspects of a network such as forcing structure, feedbacks, and shifts or breakdowns of links over time. Breakdowns or shifts in links could indicate changes in network response due to perturbations or gradual changes in the environment. With this framework, questions relating to threshold responses and overall health of a system can be readily addressed, and the system as a whole can be better understood. Process network construction requires accurate detection of time dependent links and evaluation of their importance and strength in terms of network behavior.

Studies on networks composed of oscillators and coupled chaotic logistic equations have shown that interacting nodes exhibit a wide range of dynamics depending on node coupling strengths, imposed time dependencies, and forcing [Masoller and Atay, 2011, Marti et al., 2008, Paredes et al., 2013, Rosenblum et al., 1997, Atay et al., 2004, Aguirre et al., 2014]. Time series nodes can range from being unconnected to exhibiting various types of synchronization such as complete, lagged, general, or phase synchronization. Complete and lagged synchronized nodes have coincident states

[1] This chapter is published as an article in Entropy, 2015 [Goodwell and Kumar, 2015]

either simultaneously or at a time delay, phase synchronized nodes are locked in phase but vary in amplitude, and generally synchronized nodes have some functional relationship [Rosenblum et al., 1997]. In chaotic logistic networks, the potential for these dynamics depends on delay (τ) distribution, coupling strength (ϵ) and connectivity (K_f) [Atay et al., 2004, Marti et al., 2008, Masoller and Atay, 2011]. For strongly connected networks, synchronization is largely independent of the connection topology (Δ). As a result, networks with different connection topologies, such as random, scale free, and small-world, all achieve complete synchronization at a threshold connectivity as measured by average node degree multiplied by node coupling strength ϵ [Marti et al., 2008]. When delays (τ) between nodes are uniform, the network synchronizes to a chaotic trajectory [Marti et al., 2008]. When delays are distributed over multiple τ values, the network synchronizes to a steady state, or fixed point value [Marti et al., 2008]. At lower connectivities, the network displays a range of dynamics. The complexity of network behavior increases with the incorporation of stochastic forcing or noise.

In observed or measured process networks, nodes are likely to exhibit a combination of deterministic behavior due to functional dependencies and stochastic behavior due to random influences. In this study, we aim to identify time dependencies within networks of various structures, in addition to classifying networks in terms of their forcing-feedback mechanisms. Shifts in time dependencies or driving nodes in process networks could identify behavioral shifts in response to perturbations. These shifts could indicate alterations in important structural or functional components of the system.

Identification of coupling in real networks relies on statistical measures designed to capture the diversity of time dependent interactions. Metrics used to detect synchronization and time dependencies between nodes include variance and correlation measures [Atay et al., 2004, Marti et al., 2008, Masoller and Atay, 2011], information theory measures [Ruddell and Kumar, 2009a, Duan et al., 2013], convergent cross mapping [Sugihara et al., 2012], coupling spectrums [Alizad-Rahvar and Ardakani, 2012], graphical models [Friedman et al., 2008, Eichler, 2012], and various others [Niso et al., 2013]. Variance (σ^2) measures between nodes and over time estimate relative levels of synchronization between nodes and identify the existence of complete synchronization to a single trajectory or fixed point [Masoller and Atay, 2011, Atay et al., 2004, Marti et al., 2008]. Information theory measures such as entropy $H(X)$, mutual information $I(X;Y)$, transfer entropy $TE(X \rightarrow Y)$ [Schreiber, 2000], and partial mutual information [Vlachos and Kugiumtzis, 2010, Frenzel and Pompe, 2007] quantify uncertainty of node

states and reductions in uncertainty given other node states. Information theory measures have been applied in ecohydrology [Kumar and Ruddell, 2010, Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b], neuroscience [Lee et al., 2012, Barrett and Seth, 2011], and industrial engineering [Duan et al., 2013], among others, to identify transmitters and receivers of information in addition to feedbacks within a network. When contributions from multiple "source" nodes to a "target" node are considered, shared information can be decomposed into redundant, unique and synergistic components [Williams and Beer, 2010]. Redundant information is the information shared between every source node and the target node, and unique information is that which only a single source shares with the target. In some cases, the knowledge of two source nodes may provide information to a target node that is greater than the union of the information provided by both sources individually, thus providing synergistic information.

The objective of this article is to determine the additional knowledge that information theory measures can provide over variance measures concerning process network behavior, such as distinguishing between types of drivers, locating feedbacks, and identifying redundant versus unique sources of information. How well do information theory measures capture imposed network dynamics? When a time dependent link is identified, is it critical in terms of network function, or redundant with other interactions? Time dependencies identified in real process networks could be either important aspects of system health or functioning, or redundant due to induced feedbacks. Although the existence of feedbacks within process networks can obscure what is a "cause" versus an "effect" and prevent detection of causality, the estimation of redundancy can identify groups of redundant links [Hlaváčková-Schindler et al., 2007] which can then be further evaluated in terms of their contributions to system behavior.

This study uses a method to compute information theory measures that does not assume time series variables to follow a gaussian distribution, and performs well given limited datasets with as few as 200 data points. This minimization of the data requirement is valuable because data used to form real world process networks are often sparse, fragmented, or noisy [Lee et al., 2012]. There are several proposed methods to directly evaluate redundancy, synergy, and unique information [Williams and Beer, 2010, Barrett and Seth, 2011, Bertschinger et al., 2014, Harder et al., 2013] that each have advantages and disadvantages in their interpretation [Griffith and Ho, 2015, Harder et al., 2013, Olbrich et al., 2015]. Here we instead use combinations of established information theory measures that reveal multiple aspects of information transfers.

We create chaotic-logistic networks with a range of τ-distributions, coupling strengths ϵ,

topologies Δ, and levels of noise, and compare information theory and variance measures to the imposed network structures. We evaluate our methods in terms of correctly identified links, or statistically significant detections of information measures that correspond to imposed time dependencies. In a network of observed time series data, structural properties such as driving nodes, node degrees, and coupling strengths are generally unknown and may change over time. However, process network construction can reveal some of this structure, and temporal changes in detected links indicate shifts in properties. In addition, comparisons of process networks can detect differences between inputs and outputs of a model or between measured and simulated variables. Through this analysis of generated network dynamics, we improve our ability to identify and interpret real-world process networks that range from uncoupled to completely synchronized.

2.2 Methods: definition of metrics

We evaluate network behavior with several measures that capture variability and time-dependent interactions. The standard deviation between node values, σ_{nodes}, indicates synchronization between nodes [Masoller and Atay, 2011]. In a network composed of $i = 1...N$ nodes and $t = 1....n$ time steps per node,

$$\sigma_{nodes} = \frac{1}{n}\sum_{t=1}^{n}\left[\left(\frac{\sum_{i=1}^{N}(x_i(t) - \bar{x}(t))^2}{N-1}\right)^{1/2}\right] \tag{2.1}$$

For complete synchronization, $\sigma_{nodes} = 0$. The standard deviation between time steps averaged over the N nodes, σ_{time}, indicates the temporal variation of nodes [Masoller and Atay, 2011].

$$\sigma_{time} = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{\sum_{t=1}^{n}(x_i(t) - \hat{x_i})^2}{n-1}\right)^{1/2}\right] \tag{2.2}$$

In these metrics, $\bar{x}(t)$ is the mean node value at time t, and $\hat{x_i}$ is the mean temporal value of node i. If both $\sigma_{nodes} = 0$ and $\sigma_{time} = 0$, all the nodes in the network are at the same fixed point value for all time steps. If $\sigma_{nodes} > 0$ and $\sigma_{time} = 0$, nodes are at different fixed point values. Finally, if $\sigma_{nodes} = 0$ and $\sigma_{time} > 0$, nodes are completely synchronized to each other but vary with time [Masoller and Atay, 2011]. These measures can also be applied to any pair of nodes or subsystems within a larger network, and are useful when comparing networks to a reference or baseline condition. Since σ_{nodes} and σ_{time} depend on the range of values $(x_i(t))$ taken by nodes in the network, the significance of any $\sigma > 0$ is difficult to

evaluate without further knowledge of the range of possible network behaviors.

Information theory measures involve comparing probability density functions (*pdf*s) of nodes rather than their magnitudes. Shannon entropy $H(X)$ quantifies the uncertainty or variability of a node. $H(X)$ can be computed and normalized to between 0 and 1 by dividing by the upper bound $\log(\mathcal{N})$, where $\mathcal{N}$ is the chosen number of bins into which the *pdf* $p(x)$ is discretized.

$$H(X) = \sum_{k=1}^{\mathcal{N}} p(x_k) \log \left[\frac{1}{p(x_k)} \right] \tag{2.3}$$

Mutual information $I(X_a; X_b)$ is the reduction in uncertainty of node X_a given knowledge of the state of another variable X_b, and is computed from the joint *pdf*.

$$I(X_a; X_b) = \sum p(x_a, x_b) \log \left[\frac{p(x_a, x_b)}{p(x_a)p(x_b)} \right] = H(X_a) - H(X_a|X_b) \tag{2.4}$$

Lagged mutual information $I_\tau = I(X_a(t - \tau_a); X_{tar}(t))$ quantifies the information shared between a target node X_{tar} and the τ_a-lagged history of a source node X_a. Although I is a symmetric quantity, I_τ introduces a directionality if we assume that past node states inform future states, and not vice versa. In other words, we consider past node states to be "sources" and a future state to be a "target". In a network of interacting nodes, multiple sources in the form of different nodes or a single node at different time scales can provide information to a single "target" node. The total lagged I_τ shared by two sources (X_{s1} and X_{s2}) to a target (X_{tar}) is the mutual information between one source and the target added to the conditional mutual information as follows:

$$I(X_{tar}; X_{s1}, X_{s2}) = I(X_{tar}; X_{s2}) + I(X_{tar}; X_{s1}|X_{s2}) \tag{2.5}$$

Using the partial information decomposition approach [Williams and Beer, 2010, Williams and Beer, 2011], we see that this shared information between two sources and the target can be partitioned into elements as follows:

$$I(X_{tar}; X_{s1}, X_{s2}) = U_{s1} + U_{s2} + R_{s1,s2} + S_{s1,s2} \tag{2.6}$$

$$I(X_{tar}(t); X_{s1}(t - \tau_{s1})) = U_{s1} + R_{s1,s2} \tag{2.7}$$

$$I(X_{tar}(t); X_{s2}(t-\tau_{s2})) = U_{s2} + R_{s1,s2}. \tag{2.8}$$

In Equations (2.6-2.8), U_{s1} and U_{s2} represent the unique information that only $X_{s1}(t-\tau_{s1})$ and $X_{s2}(t-\tau_{s2})$, respectively, share with $X_{tar}(t)$, $S_{s1,s2}$ is the synergistic information that arises only from the knowledge of both $X_{s1}(t-\tau_{s1})$ and $X_{s2}(t-\tau_{s2})$ together, and $R_{s1,s2}$ is the redundant information that is provided by either source node separately.

We see from substituting Equations (2.8) and (2.6) into Equation (2.5) that the conditional information term $I(X_{tar}; X_{s1}|X_{s2})$ is equivalent to $U_{s1} + S_{s1,s2}$, or the unique information component of one source node and the synergistic information due to the knowledge of both sources. The same result can be obtained by observing that conditional mutual information is equal to the interaction information or co-information [Bell, 2003, Timme et al., 2014, Williams and Beer, 2010] ($II = I(X_{tar}(t); X_{s1}(t-\tau_{s1}); X_{s2}(t-\tau_{s2})) = S_{s1,s2} - R_{s1,s2}$) added to the mutual information ($U_{s1} + R_{s1,s2}$). A positive interaction information ($II > 0$) indicates that synergy dominates over redundancy in the partitioning of shared information [Williams and Beer, 2010]. An $II < 0$ indicates dominant redundancy, or that the knowledge of any one variable "explains" correlation between the other two [Timme et al., 2014]. Conditional I_τ (i.e. $I(X_{tar}; X_{s1}|X_{s2})$) is also referred to as partial information, since it represents the part of the total mutual information that is not contained in the second source node (X_{s2}) [Frenzel and Pompe, 2007]. The conditional I_τ is computed between two sources and a target node as follows:

$$I(X_{tar}; X_{s1}|X_{s2}) = \sum_{\substack{x_{tar},\\ x_{s1},\\ x_{s2}}} p(x_{tar}(t), x_{s1}, x_{s2}) \log \left[\frac{p(x_{tar}, x_{x2}, x_{s1})}{p(x_{tar}, x_{s2}} \right] \tag{2.9}$$

where $X_{s1} = X_{s1}(t-\tau_{s1})$, $X_{s2} = X_{s2}(t-\tau_{s2})$, and $X_{tar} = X_{tar}(t)$. In Equations (2.5) and (2.9), if we consider the special case where one of the source nodes X_{s2} is the lagged history of the target node X_{tar} itself (i.e. $X_{s2}(t-\tau_{s2}) = X_{tar}(t-\tau_{tar})$), the conditional mutual information term is equivalent to the transfer entropy $TE(X_{s1}(t-\tau_{s1}) \rightarrow X_{tar})$. Transfer entropy [Schreiber, 2000] is the reduction in uncertainty of a node X_{tar} due to the knowledge of the $(t-\tau_{s1})$ history of another node X_{s1} that is not already accounted for in the $(t-\tau_{tar})$ history of X_{tar} [Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b]. TE is often interpreted as the amount of predictive information transferred between two processes [Wibral et al., 2014]. Some formulations of TE involve consideration of block lengths l and k of the histories of the transmitting node X_{s1} and receiving node X_{tar}, respectively. However, the values of l and k are

generally set equal to 1 so as to not impose additional data requirements for the computation of a higher dimensional *pdf* [Schreiber, 2000, Bollt, 2012, Paredes et al., 2013]. In this study, we relax the usual assumption in transfer entropy computations that predictive information from a source node is only conditioned on the target node's history. $T \equiv I(X_{tar}; X_{s1}|X_{s2})$ provides a generalization of TE that conditions the predictive information of the time dependency of any source, including the history of the target node itself.

In this paper, we establish network links by computing lagged mutual information I_τ using Equation (2.4) between each potential source and target node for a range of time delays. To test for statistical significance of the detected value, we randomly shuffle the target node X_{tar} to destroy time correlations while retaining other properties of the time series data [Ruddell and Kumar, 2009a, Vejmelka and Paluš, 2008, Boba et al., 2015]. We compute 100 values of $I_{\tau shuffled}$, and perform a hypothesis test at a 99% confidence level. If the detected value is less than $I_{\tau shuffled,mean} + 3 * \sigma_{shuffled}$, we dismiss the detected link as not significant.

After establishing time dependent links in the network, we compute the total and conditional I_τ provided by each pair of sources to every target node using Equations (2.9) and (2.5). We define T/I as an index to measure the non-redundant component of each link as a function of conditional and total shared information as follows:

$$\frac{T}{I}(X_{s1} \to X_{tar}) = \min_{X_{s2}} \left[\frac{I(X_{tar}; X_{s1}|X_{s2})}{I(X_{tar}; X_{s1}, X_{s2})} \right] \tag{2.10}$$

where

$$\frac{I(X_{tar}; X_{s1}|X_{s2})}{I(X_{tar}; X_{s1}, X_{s2})} = \frac{U_{s1} + S_{s1,s2}}{U_{s1} + U_{s2} + S_{s1,s2} + R_{s1,s2}} \tag{2.11}$$

Computation of T/I requires a pairwise evaluation of sources for each detected target node in the network. For a link between X_{tar} and a source X_{s1}, minimization across each each alternate source X_{s2} provides a conservative measure of the unique and synergistic components of the link. In the absence of synergistic relationships, if a source X_{s1} is completely redundant due to another source (i.e. $U_{s1} = 0$ and $S_{s1,s2} = 0$), then $T/I = 0$. If X_{s1} is the only source or is much stronger than all other sources (i.e $U_{s1} \gg U_{s2}$), then T/I approaches 1. Therefore, T/I characterizes the relative amount of unique or synergistic information provided by a link as originally determined based on statistically significant $I(X_{tar}; X_{s1})$. While $I(X_{tar}; X_{s1})$ detects a single time dependent link, conditioning on other dependencies allows for detection of unique and redundant linkages [Runge et al., 2012, Vlachos and Kugiumtzis, 2010]. High T/I values can also result from synergistic relationships, where much more information is

shared by two sources together than either shares separately. Other methods of detecting or eliminating redundant sources include direct transfer entropy [Duan et al., 2013] or causation entropy [Sun and Bollt, 2014], which involve 4d *pdf* estimation to condition on multiple source nodes.

We use the Kernel Density Estimation (KDE) [Lee et al., 2012, Silverman, 1986, Hlaváčková-Schindler et al., 2007] method to estimate the 3d *pdf* $(p(x_{tar}(t), x_{s1}(t - \tau_{s1}), x_{s2}(t - \tau_{s2}))$ required to compute conditional I_τ and the 2d and 1d *pdf*s needed for I_τ and H after testing several techniques [Ruddell and Kumar, 2009a, Paredes et al., 2013, Alizad-Rahvar and Ardakani, 2012, Lee et al., 2012] on 2-node networks of $50 \leq n \leq 2000$ data points and varying noise levels. The kernel estimator at a grid point or location y given $Y_{i=1....n}$ observations is defined as [Silverman, 1986]:

$$\hat{p}(y) = \frac{1}{nh^d} \sum_{i=1}^{n} \kappa \left[\frac{1}{h}(y - Y_i) \right] \tag{2.12}$$

The multivariate Epanechnikov kernel $\kappa = \kappa_e$ [Silverman, 1986] is as follows:

$$\kappa_e(x) = \begin{cases} \frac{1}{2} c_d^{-1}(d+2)(1 - y^T y) & \text{if } y^T y < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

in which d is the dimension of the *pdf*, c_d is the volume of a d-dimensional unit sphere, and n is the number of observations. The optimal window width h for the kernel is chosen to vary with n and d based on [Silverman, 1986] as follows:

$$h_{opt} = \begin{cases} 1.06\sigma n^{-1/5} & \text{if } d = 1 \\ 1.77\sigma n^{-1/6} & \text{if } d = 2 \\ 2.78\sigma n^{-1/7} & \text{if } d = 3 \end{cases} \tag{2.14}$$

in which σ is the standard deviation of the data. We evaluate the kernel at $\mathcal{N} = 35$ evenly spaced grid points (y) in each dimension. The KDE method performed similarly to fixed-binning [Ruddell and Kumar, 2009a, Lee et al., 2012] and partitioning methods [Lee et al., 2012] for large data sets, but the smoothing of the *pdf* due to the kernel improved performance for small data sets with $n < 200$.

2.3 Results: 2-node networks

To assess whether our estimates of I_τ and T/I accurately identify time dependencies and distinguish between types of forcing, we first consider three bivariate cases. In each case, node X_1 forces another node X_2 at a time lag of $\tau_{1,2} = 3$ via a chaotic logistic equation. However, the driving node X_1 is established in different ways.

2.3.1 Logistic Forcing:

For a chaotic logistic forcing case, X_1 is an independent chaotic logistic equation, and node X_2 is a chaotic logistic function dependent on the $\tau_{1,2} = 3$ lagged history of X_1 (illustration in Figure 2.1a, left), that is:

$$\begin{aligned} X_1(t) &= f(X_1(t-1)) \\ X_2(t) &= f(X_1(t-3)) \end{aligned} \tag{2.15}$$

where $f(X) = aX(1-X)$ with $a = 4$. This configuration results in statistically significant detected I_τ between all node pairs (Figure 2.1a). Since X_1 is self-driven at a lag of $\tau_{1,1} = 1$, and forces X_2 at a lag of $\tau_{1,2} = 3$, we detect the dominant dependency from X_1 to X_2 at a lag of $\tau_{1,2} - \tau_{1,1} = 2$ instead of the imposed $\tau_{1,2} = 3$. The self-feedback of X_1 is reflected in X_2, and we see from the time series (Figure 2.1a, left) that the nodes are shifted copies of each other. The proportion T/I is rather low between all pairs (Figure 2.1a), indicating no single source to a target is extremely strong compared to others, thus sources are likely to contain redundancies ($R_{s1,s2} > 0$). In fact, we know from the time series that any information shared between X_1 and X_2 is completely redundant given their own histories. In other words, there is no S or U component since the history of each node contains complete predictive information. This should result in $T/I = 0$, since $U_s = 0$ and $S = 0$. However, we see from Figure 2.1a that $T/I > 0$. This results from the necessarily empirical estimation of the *pdf*.

Figure 2.2a shows the data attractor (points $[x_2(t), x_1(t-2), x_1(t-3)]$) and estimated *pdf*s used to compute I_τ and T/I at the delay $\tau_{1,2} = 3$. From the 3d *pdf* (Figure 2.2a, middle), we see that there is a linear relationship (1:1 line) between $X_1(t-2)$ and $X_2(t)$ indicating that the nodes have coincident states at a time lag of 2, thus a dominant I_τ link at that lag. There is a parabolic relationship between $X_1(t-3)$ and $X_2(t)$. Due to this structure of the 3d *pdf* where $X_1(t-2)$ predicts $X_2(t)$ more directly than the imposed link $X_1(t-3)$, we detect non-zero T/I. In other words, a unique component of information is detected between

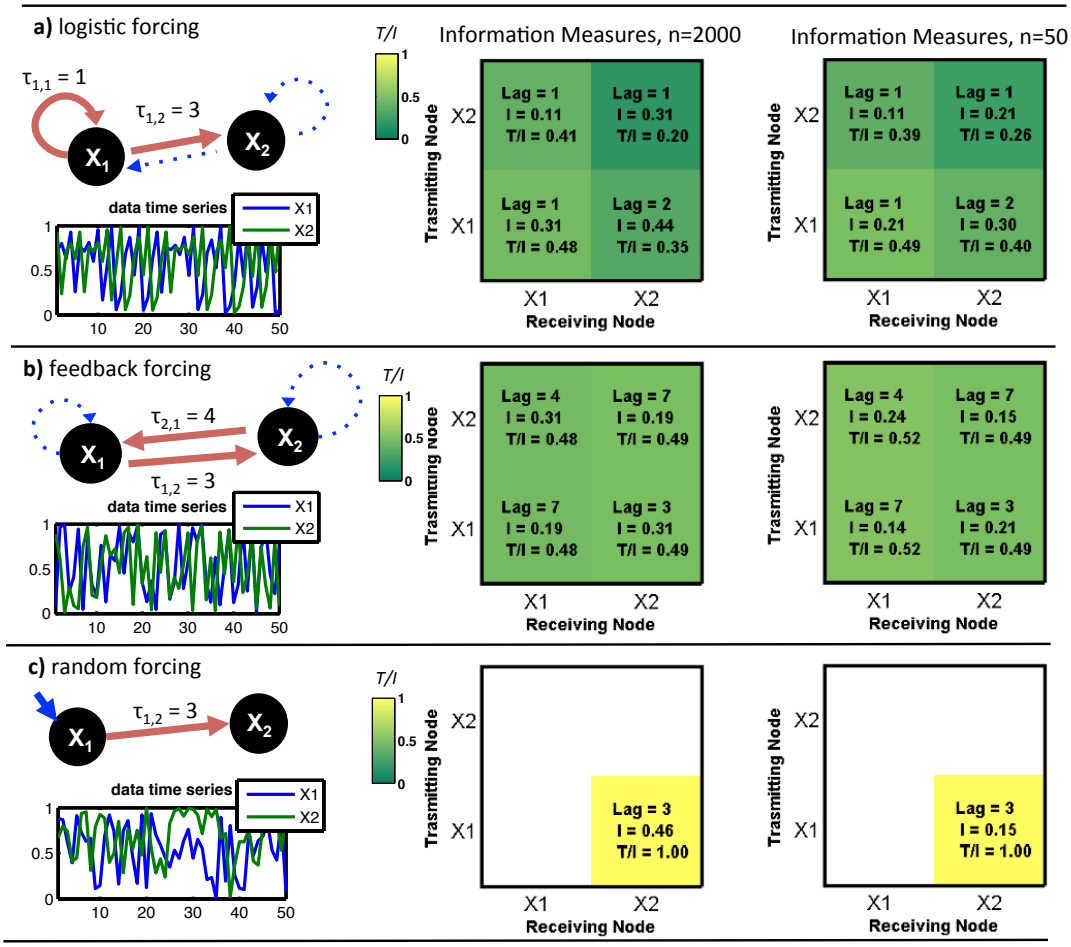

Figure 2.1: Two-node network examples with chaotic logistic, random, and feedback forcing. Node X_2 is always a function of the $\tau_{1,2}$ lagged history of X_1. Left: red arrows indicate forcing, and blue dotted arrows indicate induced feedbacks that were also detected using information measures. An example time series is shown below illustrations. Middle, Right: Information measures I_τ and T/I for each source indicate strength and uniqueness of detected dependencies (yellow indicates high T/I) for $n = 2000$ data points (middle) and $n = 50$ (right). (a) Logistic equation forcing case: dominant transfer detected from X_1 to X_2 at a lag equal to $\tau_{1,2} - \tau_{1,1}$, and self-feedback on X_2 reflects that of X_1. (b) Feedback forcing case: I and T/I detect imposed links, and self-feedbacks at lag of $\tau_{2,1} + \tau_{1,2}$. $T/I < 0.5$ for all links, indicating high level of redundancy (c) Random forcing case: imposed link from X_1 to X_2 detected. $T/I = 1$ indicates that this link constitutes unique shared information.

$x_1(t-2)$ and $x_2(t)$. However, the low T/I for all links indicates that each target node receives information from multiple sources, or a single source at multiple time lags. If we assume that sources do not provide synergistic information, $T/I < 0.5$ for all dependencies indicates that

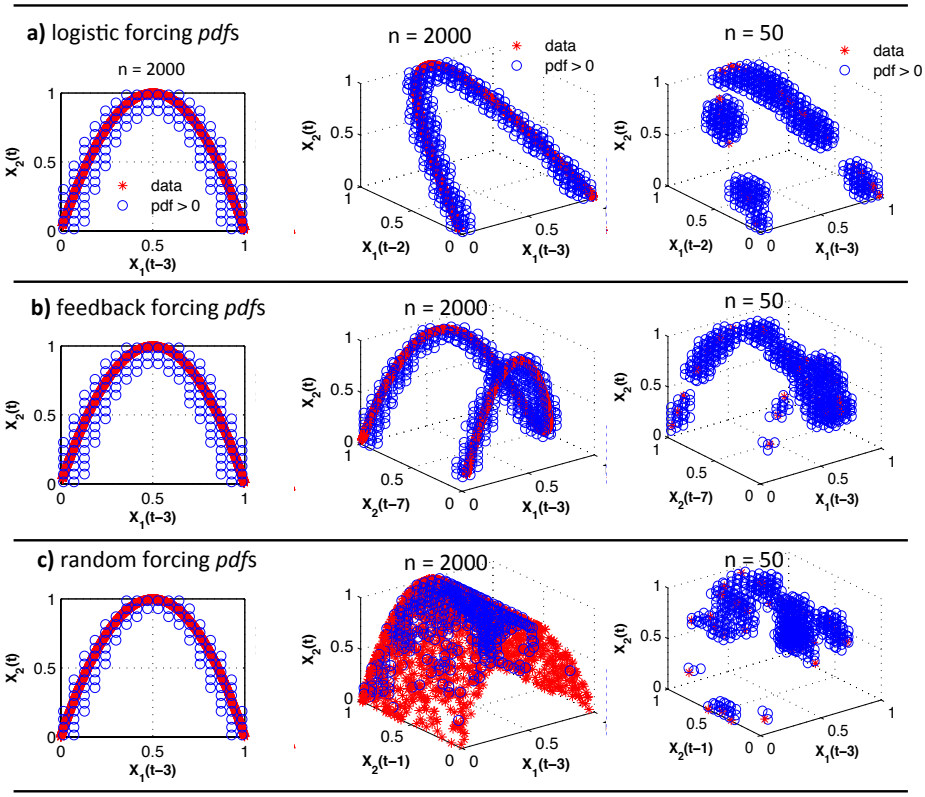

Figure 2.2: Estimated *pdf*s for each of the two-node network cases in Figure 2.1 Red stars indicate data points, blue circles are grid points (i, j, k) for which $pdf(i, j, k) > 0.0001$. *(Left)*: 2-dimensional *pdf*s used to compute I_τ between $X_1(t-3)$ and $X_2(t)$. *(Middle, Right)*: 3-dimensional *pdf* used to compute corresponding T/I of link. *(a)* Logistic forcing: $X_1(t-2)$ and $X_2(t)$ are completely synchronized (1:1 line), so there is redundancy in the transfer from $X_1(t-3)$ to $X_2(t)$. *(b)* Feedback forcing: both the histories of X_1 and X_2 inform $X_2(t)$ due to feedbacks. In *(a)* and *(b)*, significant T/I is still detected due to empirical estimation of the 3d *pdf*s. *(c)* Random forcing: knowledge of $X_2(t-1)$ (or any other lagged node history) conveys no information concerning $X_2(t)$, thus there are no other significant links and $T/I = 1$.

the stronger sources contain significant redundancies with weaker sources. This assumption is valid for our generated networks, since multiple source node histories do not inform the target beyond the union of their individual contributions.

When the *pdf* is estimated based on 50 data points instead of 2000 (Figure 2.1a, right), we detect statistically significant but lower I_τ compared to the $n = 2000$ point case, and similar values of T/I. Comparison of the 3d *pdf*s in Figure2.2a shows that fewer data points results in more spread of the empirical *pdf* due to the kernel estimator, and weaker detection of

information measures (Figure 2.1,right).

2.3.2 Feedback Forcing:

For a feedback forcing case, each node is a function of the history of the other node at lags $\tau_{2,1} = 4$ and $\tau_{1,2} = 3$ (illustration in Figure 2.1b), that is:

$$\begin{aligned} X_1(t) &= f(X_2(t-4)) \\ X_2(t) &= f(X_1(t-3)) \end{aligned} \tag{2.16}$$

Similar to the logistic forcing case, this case results in high detected I_τ and relatively low T/I between all the node pairs (Figure 2.1b). This indicates that the nodes are highly coupled, and are predictable given knowledge of either node at one of several time lags. This is also apparent given the high "self" I_τ for each node. We note that by substitution in Equation (2.16), each node can be written as a function of its own history, indicating complete source redundancy and $T/I = 0$. However, we detect $T/I > 0$ for the same reason as the previously discussed logistic forcing example. The strongest detected source to target links, in this case the second order polynomial relation between $X_1(t-3)$ and $X_2(t)$ (Figure 2.2b, left), is detected to have a unique component when compared to the weaker link (fourth order polynomial) between $X_2(t-7)$ and $X_2(t)$ (Figure 2.2b, middle). In other words, although $X_2(t-7)$ provides all predictive information regarding $X_2(t)$, $X_2(t)$ is a more simple function of $X_1(t-3)$, so we still detect statistically significant T/I from source X_1 to target X_2 (Figure2.1b, middle). When feedbacks are involved, it is not possible to distinguish between "drivers" and "receivers" within the network, but we can identify the existence of these feedbacks and their strengths and time scales. Similar to the logistic forcing case, reducing the number of data points to $n = 50$ results in more spread of the $pdf s$ and weaker detection of I_τ (Figure 2.1b, right).

2.3.3 Random Noise Forcing:

In the last 2-node network example, X_1 is a time series of randomly generated uniform noise (illustration in Figure 2.1c, left), that is:

$$\begin{aligned} X_1(t) &= z \sim U(0,1) \\ X_2(t) &= f(X_1(t-3)) \end{aligned} \tag{2.17}$$

This case results in statistically significant I_τ and $T/I = 1$ from X_1 to X_2 at the imposed delay (Figure 2.1c, middle). The high T/I indicates that the link is a dominant and unique source of information to the target X_2. Furthermore, since there are no other links detected with I_τ, we can conclude that neither synergy nor redundancy exists because all shared information is unique to the sole source X_1. In other words, because X_1 forces X_2, and X_1 is randomly generated, no information about X_2 is encoded in any other source. The *pdf*s estimated to detect I_τ and T/I in this case are shown in Figure 2.2c for the $n = 2000$ and $n = 50$ cases. From the *pdf*s, we observe that $X_2(t-1)$ and $X_2(t)$ are uncorrelated.

With these example cases, we show that information measures T/I and I_τ capture imposed time dependencies and feedbacks, and can determine the partitioning of shared information between different source histories. When a randomly generated node drives another node, we detect high T/I since there is no information provided by other sources, and all shared information between the source (random driver) and target (receiving) node is unique. When there are feedbacks involved, such as in the logistic and feedback forcing cases, high I_τ is detected between all node pairs, but T/I is low due to increased redundancies. Although T/I is often detected as significant due to *pdf* structure and its estimation, it still provides an estimate of the dominance of redundancy versus unique information, particularly in the absence of synergy. Detection of high $T/I(X_{s1} \to X_{tar})$ in the case of multiple sources indicates that the source X_{s1} is either unique or is highly synergistic with another source.

The 2-node cases with chaotic logistic driving nodes and feedbacks illustrate the detection of multiple links between nodes that were not directly imposed. However, these links are due to the imposed time dependencies and can be considered to be *induced* feedbacks. In the chaotic logistic driving case, a time-dependent source node forces a target node, causing the target node to acquire the same time dependency given its own history. In the feedback example, the imposed bi-directional feedback causes each node to have a time dependency given its own history. These *induced* feedbacks occur in highly connected networks and can propagate due to a single imposed bi-directional feedback or time dependency. Detected time dependencies that were not imposed could be characterized as "false positives" in link detection [Smirnov, 2013], but we note that these detections are expected due to the forcing-feedback structure.

In this section, each node was forced by either a chaotic logistic equation or uniform random noise. In networks of multiple interacting processes, a single node may respond to many variables. In the next section, we generate 10-node networks with $n = 200$ time series points, in which nodes may be forced by various combinations of neighbors in addition to uniform

random noise. We compute information and variance measures, and analyze synchronization and time dependencies as connectivity varies. As in the 2-node cases, I_τ and T/I measures detect dependencies between node pairs, but are also used to identify the larger connectivity structure of the network. When multiple source nodes and random noise influence a single target node, information measures between node pairs are more weakly detected. However, we show that these measures can correctly identify even weak interactions and reveal the forcing-feedback structure of a network.

2.4 Results: coupled chaotic logistic networks

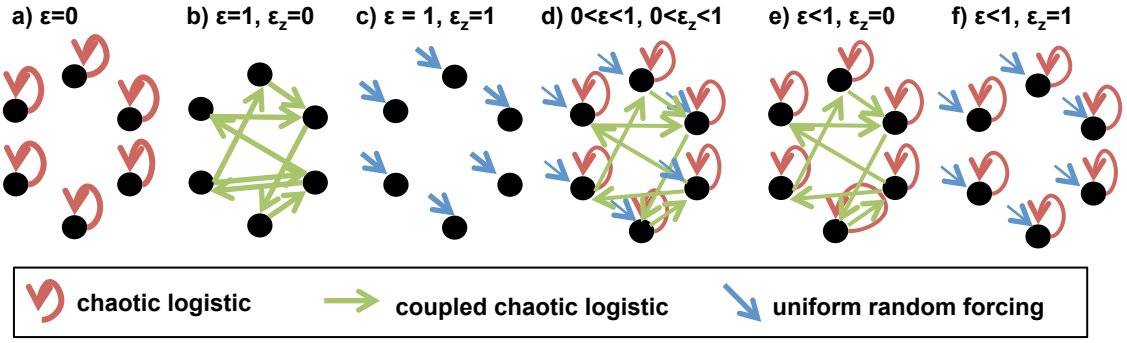

Figure 2.3: Illustration of possible network structures based on chaotic logistic mapping (Equation (2.18)). *(a)* unconnected network driven by individual chaotic logistic equations. *(b)* network driven by chaotic logistic couplings. *(c)* unconnected network driven by random forcing. *(d)* network driven by combination of forcings according to $\epsilon > 0$ and $\epsilon_z > 0$. *(e)* network driven by combination of individual and coupled chaotic logistic equations. *(f)* unconnected network driven by random forcing and chaotic logistic equation.

Previous studies have determined that chaotic logistic network synchronization capacity in terms of σ measures (σ_{nodes} and σ_{time}) depends more on the delay (τ) distribution than network topology Δ [Masoller and Atay, 2011, Atay et al., 2004, Marti et al., 2008]. For a range of Δ including small-world, scale-free, and random networks, increasing connectivity (increasing coupling strength or number of links) leads to synchronization of all connected nodes. The dynamics of the resulting synchronized trajectory depend on the τ-distribution [Marti et al., 2008]. Networks with uniform delays (e.g. $\tau = 1$ for all linked nodes) synchronize to a single chaotic logistic trajectory. In contrast, networks with heterogenous delays (e.g.

random $\tau \in \{1, 10\}$ for linked nodes) synchronize to the fixed point ($x^* = 1 - 1/a$) of the logistic equation. This type of synchronization occurs when nodes that receive from enough neighbors at different lags converge toward the fixed point x^* and all nodes approach zero amplitude [Cakan et al., 2014]. In terms of information theory measures, it has been found that information transfer can be used to predict synchronization and distinguish between origins of interaction fields, or types of forcing, in different types of generated networks [Paredes et al., 2013].

In this section, we extend the forcing mechanisms introduced in the 2-node examples to larger 10-node networks. A 10 node network is small enough for computation efficiency and to represent many systems of interest, and large enough to capture the complexity and synchronization that larger networks exhibit. We generated networks of between 5 and 50 nodes and observed that larger networks synchronize at lower connectivities, but display otherwise similar behavior. The networks are generated over a range of connectivities, with different proportions of chaotic logistic and uniform random noise forcing. We introduce randomness into the network through randomly generated "driving nodes" that act as controls, or through the addition of uniform random noise to each node in equal proportion. We compute information and variance measures for the generated networks, and use whole-network measures to summarize each individual case. In chaotic logistic networks with no random component, we observe the expected delay-dependent synchronization.

2.4.1 Network Formation

We generate networks using Equation (2.18), each with $N = 10$ nodes of $n = 200$ time series points per node, using the framework given as:

$$X_i(t) = (1 - \epsilon) f(X_i(t-1)) + \frac{(1 - \epsilon_z)\epsilon}{k_i} \sum_{j=1}^{N} w_{j,i}[f(X_j(t - \tau_{j,i}))] + \epsilon_z \epsilon z. \tag{2.18}$$

In Equation (2.18), i and j are node indices, $f(X) \equiv aX(1 - X)$, t is time step, k_i is the in-degree of node i, w is the adjacency matrix ($w_{j,i} = 1$ if X_i is a function of X_j, $w_{j,i} = 0$ otherwise), τ is the delay matrix associated with w, and z is a uniform random noise between 0 and 1. As in the 2-node example, we set $a = 4$ so that each individual $f(X_i)$ is in the chaotic regime.

Network Forcing:

The formulation of Equation (2.18) defines a node X_i to be forced by (*1*) its own lagged history, (*2*) the lagged histories of connected nodes, and (*3*) random noise. The extents to which these components influence X_i are defined by the coupling strengths ϵ and ϵ_z. For example, for $\epsilon = 1$ and $\epsilon_z = 0$, X_i is solely a function of the histories of all X_j for which $w_{j,i} = 1$ (Figure 2.3b). If $\epsilon = 0$, each node is an independent chaotic logistic time series, since X_i is only dependent on its own history (Figure 2.3a). If $\epsilon = 1$ and $\epsilon_z = 1$, the network is entirely composed of uniform random noise (Figure 2.3c). For values of $0 < \epsilon < 1$ and $0 < \epsilon_z < 1$, the network responds to all three types of forcing (Figure 2.3d).

In Equation (2.18), the imposed adjacency matrix w determines the interaction "field", or set of nodes to which each node responds. The field is homogenous over time, but is different for each node. To explore the effect of external forcing, we introduce cases in which some of the 10 nodes are randomly generated time series ($n_{drivers} > 0$). The remaining nodes are generated from Equation (2.18), and can be functions of both chaotic logistic and randomly generated nodes, depending on the adjacency matrix w. The noise component $\epsilon_z \epsilon z$ represents a different type of random forcing in that it affects each node in the network in equal proportion.

Network Topologies and Delays:

Network topologies used to generate the adjacency matrix w include random and small world. Random networks are generated based only on a link probability p, while small world topologies [Albert and Barabasi, 2002] are bi-directional cyclic networks of degree 2 (each node transmits to and receives from $k = 1$ neighbor on each side), and links are added randomly with probability p. A "theoretical weighted degree" K for each network type is the average number of incoming links per node multiplied by the coupling strength term $\epsilon(1 - \epsilon_z)$. A fractional weighted degree $K_f = \frac{K}{N-1}$ is a measure of connectivity that ranges between 0 (unconnected nodes) and 1. At $K_f = 1$, nodes are completely connected at maximum coupling strength, i.e. $\epsilon(1 - \epsilon_z) = 1$ and $p = 1$.

Four specific classes of networks were tested, combining network topologies and delay distributions (Table 1). Cases 1 and 2 are random networks, while Cases 3 and 4 have small world topologies. Cases 1 and 3 have uniform delay distributions ($\tau_{j,i} = 1$ for $w_{j,i} = 1$), and Cases 2 and 4 have random delay distributions ($\tau_{j,i} \in \{1, 10\}$ for $w_{j,i} = 1$). As expected, we find that both topologies Δ behave similarly as network connectivity K_f increases in terms of

case	structure	τ-distribution	K_f	synch type
1	random	uniform, $\tau = 1$	$\epsilon(1-\epsilon_z)p$	chaotic
2	random	random, $\tau \in \{1, 10\}$	$\epsilon(1-\epsilon_z)p$	fixed point
3	small world	uniform, $\tau = 1$	$\epsilon(1-\epsilon_z)p + (1-p)\frac{2k\epsilon(1-\epsilon_z)}{N-1}$	chaotic
4	small world	random, $\tau \in \{1, 10\}$	$\epsilon(1-\epsilon_z)p + (1-p)\frac{2k\epsilon(1-\epsilon_z)}{N-1}$	fixed point

Table 2.1: Synchronization characteristics of four network cases composed of different topologies and delay (τ) distributions.

both standard deviation and information measures. As found in previous studies, we see that network behavior is most dependent on the τ-distribution and the connectivity K_f rather than Δ. Cases 1 and 3 synchronize to a single chaotic logistic trajectory as K_f is increased (Case 1 shown in Figure 2.4a), while Cases 2 and 4 synchronize to a fixed-point value as K_f is increased (Case 2 shown in Figure 2.4b). Other network configurations tested include scale-free networks and higher degree ($k > 1$) small world networks, all of which synchronized similarly according to their τ-distribution. Due to the similarities between network topologies, we show results for only Cases 1 and 2, the random network cases. Networks of various sizes from 5-50 nodes and observed similar results in terms of synchronization and detected information measures.

We canvas a range of parameters to form the adjacency matrices w and forcing structures and generate the subsequent process networks (Table 2). For each network, ϵ and ϵ_z are constants (i.e. all nodes transmit and receive with equal coupling strengths), except for cases where $n_{drivers} > 0$. In these cases, $n_{drivers}$ of the N total nodes are randomly generated nodes that only transmit information according to $w_{j,i}$. Over 40,000 distinct networks are generated (Table 2), and we compare several categories.

2.4.2 Synchronization and information in noise free networks:

We first set $\epsilon_z = 0$ in Equation (2.18) to obtain a noise free network, and consider the range of coupling strengths $0 < \epsilon < 1$ and link probability values $0 < p < 1$ (illustrations in Figure 2.3a,b,e). We see from the generated time-series data that nodes completely synchronize for high values of K_f according to their τ-distributions (Figure 2.4). Observation of σ_{nodes} (Figure 2.5a) leads to the same conclusion that both network cases synchronize to a single trajectory as K_f increases, but at different rates. For Case 1 (uniform τ), complete

parameter	range of values	number of cases
p	[0,0.05 ... 1]	21
ϵ	[0,0.05 ... 1]	21
ϵ_z	[0, 0.01, 0.1, 0.5]	4
$n_{drivers}$	[0, 1 ... 5]	6
topology Δ	[random, small world $k = 1$]	2
τ-distributions	[random, uniform]	2
total number of networks		42,336
cases with random Δ, $n_{drivers} = 0$, $\epsilon_z = 0$		882

Table 2.2: Parameter ranges for 10-node chaotic logistic networks (42,336 total networks generated).

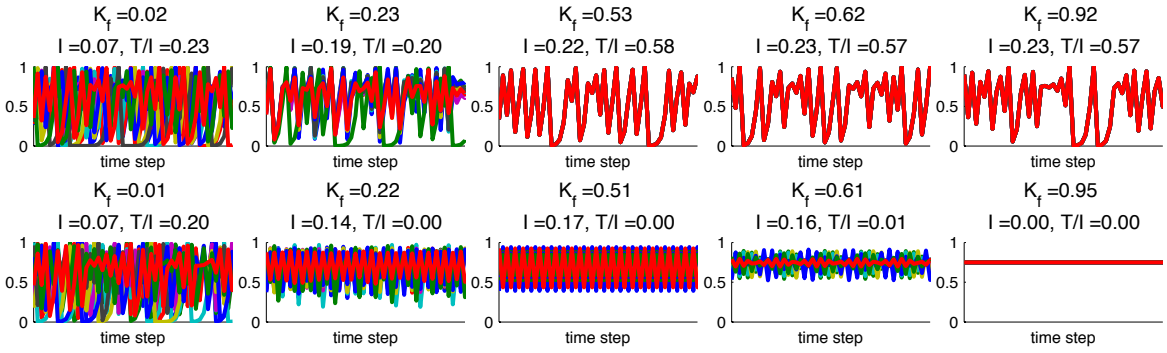

Figure 2.4: Time series for several generated networks with uniform and random delay distributions. *(a)* 50 time steps are shown for Case 1 with uniform ($\tau = 1$) delay distribution and *(b)* Case 2 with random delay distribution. Both cases approach synchronization as K_f increases.

synchronization to a time-varying trajectory is reached at $K_f = 0.4$ (Figure 2.5a,d). Case 2 (random τ-distribution) synchronizes more gradually as K_f increases, and is completely synchronized to a fixed-point trajectory for $K_f > 0.7$ (Figure 2.5a,d).

The mean values of I_τ and T/I displayed in the bar plots of Figure 2.5g,j represent the mean statistics for imposed links over all networks in each K_f interval, while the maximum I_τ and T/I values displayed in the open circles represent the maximum individual values detected within any of the networks in each K_f range. In other words, mean values represent average detections for imposed links, while the maximum represents overall maximum detected values. For Case 1, mean I_τ for imposed links approaches a constant and statistically significant

value of approximately $I_\tau = 0.25$ (Figure 2.5g) as the network synchronizes, indicating that the synchronized trajectory retains the imposed uniform $\tau = 1$ time dependency. T/I also reaches a constant non-zero value when the network synchronizes, indicating that multiple sources are detected, but the imposed lag is dominant compared to others.

For unsynchronized networks ($K_f < 0.4$), the low value of mean T/I indicates that most target nodes have multiple sources that lead to redundancies. These redundant sources may not be imposed links from the adjacency matrix, but arise due to induced feedbacks as illustrated in the 2-node example cases. However, for low $K_f < 0.2$ in Case 1, we see high maximum individual values of T/I (open red circles in Figure 2.5j). These high maximum values result from cases in which a target node receives from one source (its own history in the unconnected chaotic logistic case) very strongly, and other sources very weakly, so that T/I is close to 1. Maximum values of I_τ and T/I that are much higher than average values indicate that most imposed links become redundant, but there is at least one less connected node that receives more unique information. For Case 1, I_τ links are weaker on average for less synchronized networks, and more redundant. However, maximum individual I_τ and T/I values are highest for less synchronized networks, representing cases of high coupling strength but low link probability ($\epsilon = 1$ and $p \ll 1$) in which a single node is a dominant influence on a target.

For Case 2, slightly lower I_τ values are detected for the range of connectivities (Figure 2.5g). However, we observe very high maximum individual I_τ values over the non-synchronized high connectivity range ($0.5 < K_f < 0.8$). From the time series (Figure 2.4b), we see that this is because nodes are generally phase-locked at a time lag of 2 in this range of K_f values, so they are completely predictable based on their own histories. As expected, these high I_τ values are associated with lower T/I (Figure 2.5j) because of many redundant sources. At complete synchronization ($K_f > 0.8$) for Case 2 networks, we see that no I_τ or T/I are detected on average. Case 1 and Case 2 networks show similar behavior at low connectivities, but information measures diverge for the different types of synchronization. T/I decreases as Case 2 networks synchronize, indicating the increase in redundant links. In contrast, T/I increases as Case 1 networks synchronize due to the dominant $(t-1)$ "self"-dependency that arises on the path to synchronization.

We define a correctly detected link as a statistically significant value of I_τ detected for a node pair (X_i, X_j) at the imposed lag time $\tau_{j,i}$. For unsynchronized networks with mid-range connectivities ($0.1 < K_f < 0.6$), we correctly identify nearly all imposed links according to $\tau_{i,j}$ for both network cases (Figure 2.5m). Even for very low-connectivity networks, over half

of imposed links are detected. When Case 1 networks are synchronized, a link is detected between every node pair due to the common trajectory, regardless of the imposed w. This leads to a 100% correct link detection rate, but also a 100% "false" detection [Smirnov, 2013] rate of unimposed links. As discussed in the 2-node examples, these detections are all due to induced feedbacks, thus we do not show their increasing number as networks synchronize. For Case 2, as networks synchronize to a very low amplitude phase-locked state, we cease to detect any of the imposed links.

2.4.3 Influence and Detection of External Drivers:

When nodes in a process network contain complete predictive information, as in the previously discussed noiseless case, complete synchronization occurs at high connectivities. However, real process networks are likely to involve some proportion of nodes that are unpredictable due to influences from outside of the network. These unpredictable nodes may act only as drivers and do not respond to the network dynamics. To simulate these conditions, we generate networks in which one or more nodes ($n_{drivers}$ of N total nodes) have independent dynamics, and act only as sources. While chaotic logistic nodes approach synchronization with increasing K_f, the random driving nodes remain independent of this behavior. However, even chaotic logistic nodes do not entirely synchronize due to their varying dependencies on the drivers (Figure 2.6). From observation of σ_{nodes} and σ_{time} for the $n_{drivers} = 1$ case (Figure 2.5b,e), we see similar trends as in the $n_{drivers} = 0$ case, but no complete synchronization.

For Case 1, we detect similar mean and maximum I_τ values for imposed links over the K_f range (Figure Figure 2.5h), but maximum T/I values decrease as K_f increases (Figure 2.5k). This indicates that imposed sources become increasingly redundant with other sources as connectivity increases, and that imposed sources are weaker than induced feedbacks that arise as the non-driving nodes partially synchronize. The decreasing maximum T/I behavior for Case 1 networks with $n_{drivers} = 1$ is very similar to the Case 2 networks with $n_{drivers} = 0$. When Case 1 networks are prevented from completely synchronizing due to a random driver, no single source becomes dominant as in the case where $n_{drivers} = 0$. Case 2 networks with $n_{drivers} = 1$ have lower detected mean and maximum I_τ (Figure 2.5h) than the $n_{drivers} = 0$ case, and lower mean T/I (Figure 2.5k). However, the maximum individual T/I values are higher at higher K_f values for Case 2, reflecting the influence of the random driving node. If a random driver forces a target node, the shared information is not redundant with any other source, except in the case where an induced feedback exists to form an indirect link. For

example, if the random node forces an intermediate node that in turn forces the target node, some of the information shared between the the random and target node is also encoded in the intermediate node history, resulting in some redundancy. However, random sources do not propagate feedbacks as do time-dependent drivers, leading to more unique transfers of information.

When $n_{drivers} = 5$, we observe a continuing trend of decreased synchronization capacity (Figure 2.5c,f), decreased shared information I_τ (Figure 2.5i), decreased average T/I, and increased maximum individual T/I (Figure 2.5l). We also see that the fraction of correctly detected imposed links decreases with increased randomness in the network (Figure 2.5m,n,o). For the set of $n_{drivers} = 5$ networks, uniform and random τ-distribution cases are nearly indistinguishable, and only slight synchronization is observable from σ measures (Figure 2.5c,f). Although target nodes receive from a similar number of source nodes as in previous cases, the randomness of some of the sources results in high values of T/I, reflecting unique contributions of information. Essentially, any given source node may only share a small amount of information with a target node, but this information is more likely to be unique if the source is random. On average however, linkages are increasingly redundant (low T/I) as connectivity increases and feedbacks are created.

2.4.4 Influence of Noise in Network

In real networks, variability cannot always be attributed to the behavior of other nodes, but may be caused by noise. In this section, we set $\epsilon_z > 0$ in Equation (2.18) to represent sources of variability such as measurement noise. While randomly generated driving nodes force other network components according to connectivity as determined by adjacency matrix w, a noise component represents random variability z applied to each node. Similar to cases where $n_{drivers} > 0$, randomness due to $\epsilon_z > 0$ prevents complete synchronization. The introduction of noise as 10% of the coupling strength ($\epsilon_z = 0.1$) to the case with no random drivers (Figure 2.8,left) results in similar synchronization behavior as the initial noiseless scenario (Figure 2.5, left), except that nodes do not completely synchronize. In contrast to the $n_{drivers} > 0$ cases, all nodes contain time dependencies in addition to the noise components, and we observe that nodes tend to synchronize to an equal degree (Figure 2.7) as K_f increases. When the noise component is increased to $\epsilon_z = 0.5$, the network further loses capacity to synchronize (Figure 2.8b,e), and Cases 1 and 2 are nearly indistinguishable.

The shared information I_τ for the $\epsilon_z = 0.1$ case (Figure 2.8g) is similar to the initial

noiseless case (Figure 2.5g) for Case 1 networks, but lower for Case 2 networks. While mean detected I_τ increases with K_f for Case 1 networks up to a constant value, we observe an opposite trend in T/I, in which the maximum detected T/I decreases with increased K_f until it reaches a constant low value at $K_f = 0.6$. As the nodes partially synchronize, the noise components cause scatter in the *pdf*s that results in similar strengths of information measures between sources. The detected T/I is similar to the case with 1 external driver (Figure 2.8k). At low K_f, a node may receive a large amount of unique information from a source, but feedbacks result in redundancy at high K_f.

Increasing the noise component to $\epsilon_z = 0.5$ (Figure 2.8b,e) results in similar σ_{nodes} and σ_{time} as the $n_{drivers} = 5$ case, in which the two cases are not distinguishable. However, the mean I_τ and T/I (Figure 2.8h,k) are very small compared to the case with random driving nodes. For Case 1, there is a threshold connectivity value around $K_f = 0.7$ at which no I_τ is detected for any imposed link. This is due to the high noise in addition to many source nodes. Nodes would synchronize to a chaotic trajectory if not for the noise, and the spread of the resulting *pdf* does not allow for significant detection of any sources. For Case 2, maximum detected I_τ is statistically significant even at high $K_f(> 0.7)$ values, because nodes tend toward synchronization to a phase-locked trajectory in which $X_i(t) = X_i(t-2)$. Although T/I is very low on average for $\epsilon_z = 0.5$ networks (Figure 2.8k), the maximum detected T/I is high over the range of connectivities. Similar to the case of multiple random drivers, when a target node receives from a single source that is partially random, the information due to the random component is more likely to be unique, resulting in a high T/I.

A final case in which $n_{drivers} = 5$ and $\epsilon_z = 0.5$ combines the influences of random driving nodes and noise. In this case, little synchronization is detected based on σ measures (Figure 2.8c,f) for either Cases 1 or 2. Shared information I_τ is statistically significant over the range of K_f, but very small (Figure 2.8i). However, the maximum individual T/I values tend to be large over the entire K_f range, similar to the previous cases with high noise levels in the form of either random drivers or noise.

As noise and randomness are introduced in the networks, fewer imposed links are correctly identified (Figure 2.8m,n,o). However, for high $K_f > 0.5$, a higher fraction of imposed links is detected in the case with random drivers and noise (Figure 2.8o) than the case with only noise (Figure 2.8n). This is because the random drivers transmit information more strongly than source nodes composed of both noise and chaotic logistic components, so are more likely to be detected at higher K_f values. Detection of links improves with longer time series datasets, but we consider only networks of $n = 200$ data points to reflect realistic data availability. For

all of the network cases generated, some links are not detected at low K_f values because they are very weak. At high K_f, links other than those imposed are detected due to feedbacks induced by the high connectivity.

2.4.5 Summary of structure and synchronization of networks:

The addition of randomness or noise to a connected network prevents complete synchronization. This random component could be in the form of driving nodes that do not directly participate in feedbacks, or in the form of noise inherent to each individual node. Driving nodes remain independent of the synchronization of the rest of the network, while nodes in a noisy but feedback-connected network synchronize to an equal degree. Measures of σ_{nodes} and σ_{time} are useful to gage the relative level of synchronization, and particularly distinguish between uniform and random delay τ-distributions in noiseless network cases through their detection of synchronization and amplitude death. However, they do not convey information about time dependencies and redundancies within the network, and do not distinguish between high and low connectivities when there is a high level of randomness.

Information measures such as lagged mutual information ($I(X_{tar}; X_{s1})$), conditional information given other sources ($I(X_{tar}; X_{s1}|X_{s2})$), and total shared information between multiple sources ($I(X_{tar}; X_{s1}, X_{s2})$) detect time dependencies between node pairs in a process network, and enable detection of dominant drivers, and unique and redundant sources of information. Even for completely synchronized nodes, I_τ detects time dependencies within a single trajectory, as in the noiseless Case 1 (uniform τ-distribution) networks. For unsynchronized networks with detected time dependencies (significant I_τ), T/I further conditions on other source nodes and time scales to reveal redundancies and unique links. A $T/I > 0$ indicates that the detected link is not completely redundant given the history of another source node, which could be the target's own history, as is detected with transfer entropy. In the case of a network forced only by feedbacks, there may be high I_τ between node pairs, but low T/I due to redundancies in the synchronizing nodes. In contrast, for a network forced randomly or by a node with no time dependencies, target nodes may share information with only one source, or completely unique sources. In these cases, we detect both significant I_τ and T/I, indicating a high level of unique information transfer.

We define a correctly detected link as statistically significant value of lagged mutual information I_τ detected between two nodes (X_i, X_j) that corresponds to an imposed link according to $w_{i,j}$ and $\tau_{i,j}$. In a weakly connected network with little noise, we identify nearly

all imposed links. As connectivity increases, if nodes tend to synchronize to a time-dependent trajectory, we increasingly detect "false" links between nodes that are not defined to be connected, or are connected at a different time scale. These false links are actually feedbacks induced by the imposed time dependencies. When a network begins to synchronize, it is not possible to distinguish links due to imposed network structure from those due to induced feedbacks. In cases where nodes synchronize to a fixed point trajectory, we cease to detect any links. In networks with noise, some imposed links are correctly detected at high levels of connectivity since the random components provide unique information that prevents complete synchronization.

2.5 Discussion

The 2-node and 10-node network scenarios presented here represent a small fraction of potential network dynamics that could be observed in real-world networks. The 2-node networks help us understand induced feedbacks and types of forcing, and the resulting interpretation from analysis of the data. The 10-node networks capture general features of a larger network dynamic arising from multiple feedbacks. Process networks based on measured or simulated nodes of time series exhibit a wide range of connectivities and time-varying interactions. Coupling strengths and timescales vary between nodes and shift over time, and thresholds may exist for which certain couplings break down while others become dominant. Additionally, shared information between two or more variables could be synergistic, if the knowledge of two nodes together provides more information than their union separately. Although we present relatively simple cases in this study, the metrics used for analysis allow for the detection of a range of behaviors such as complete or partial synchronization, weak or strong time dependencies, and redundancy or uniqueness of shared information. The information theoretic measures used in this study may be compared with efficient statistical learning methods applied to graphical models, such as the graphical lasso [Friedman et al., 2008] or methods that combine graphical models with conditional dependencies [Eichler, 2012], particularly in cases with many time series nodes.

In real-world process networks in which noise and other drivers prevent complete synchronization, some detected time dependencies are likely to correspond to causality (i.e. "correctly detected links"), while others represent induced feedbacks. When nodes in the network are highly connected with feedbacks present, detected links are identified as redundant, and it is difficult to distinguish critical interactions from induced feedbacks. This feature of process

networks indicates a future challenge in terms of connecting network time dependencies to system functionality.

Figure 2.9 categorizes the range of possible whole-network or subsystem behaviors that can be extended to an observed process network. If a network is not completely synchronized, nodes could be lag-synchronized, or transferring and receiving information at different time scales and strengths. Real-world process networks consist of measured time-series data for which the underlying mechanisms are partially or completely unknown. There may be unmeasured or hidden driving and receiving nodes, and network connectivity can shift over time. The weakening of a single link may result in decreased redundancy in the form of induced feedbacks throughout an entire network. For real process network analysis, the measures presented in this study can aid in comparing observations to simulation results, evaluating system states, or assessing the influence of noise or bias on time dependencies.

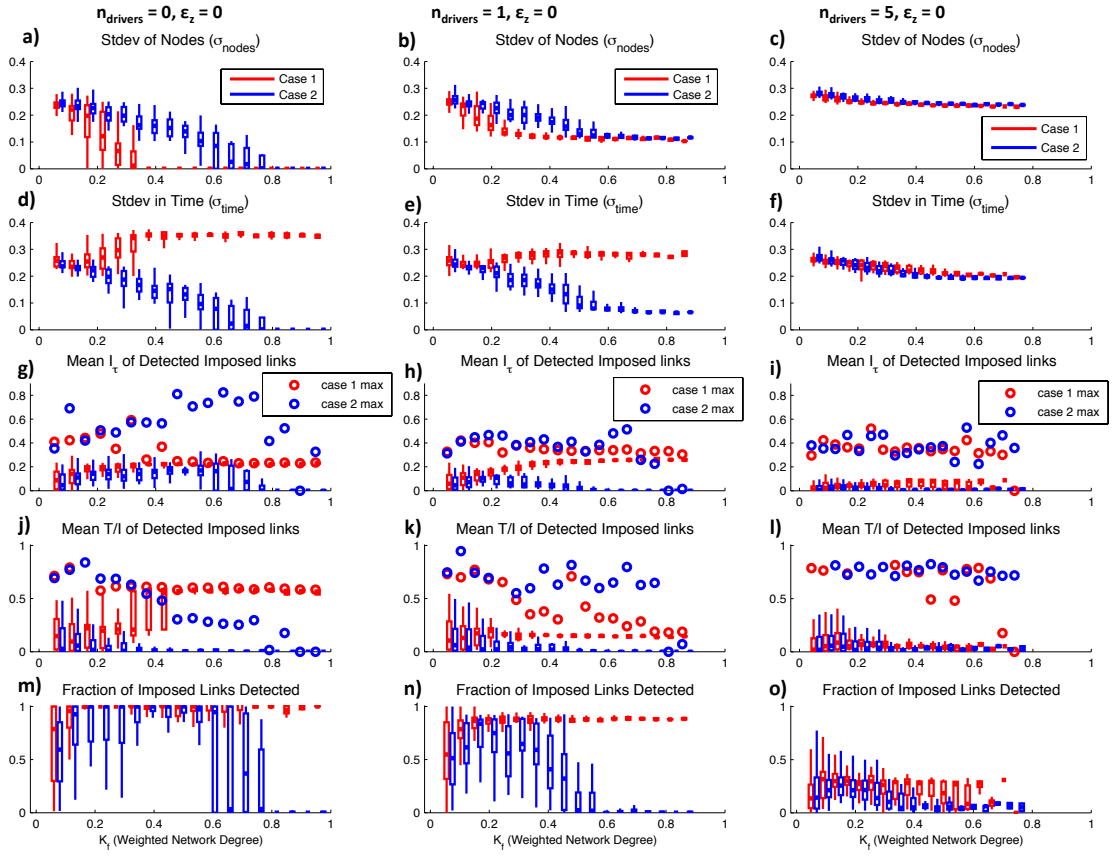

Figure 2.5: Behaviors of network configurations with a range of connectivities K_f. 882 networks over K_f range for *(left)* the case with no randomly generated driving nodes, *(middle)* $n_{drivers} = 1$, and *(right)* $n_{drivers} = 5$. *(a-c)* Standard deviation across nodes σ_{nodes}. *(d-f)* Standard deviation across time σ_{time}. *(g-i)* Box plots show mean I_τ detected for all imposed linkages for all networks in each K_f range, and open circles are maximum detected I_τ of any imposed link. *(j-l)* mean T/I over all networks and maximum detected T/I as in *(g-i)*. *(m-o)* Fraction of all imposed links that were correctly identified as time dependencies through detected I_τ.

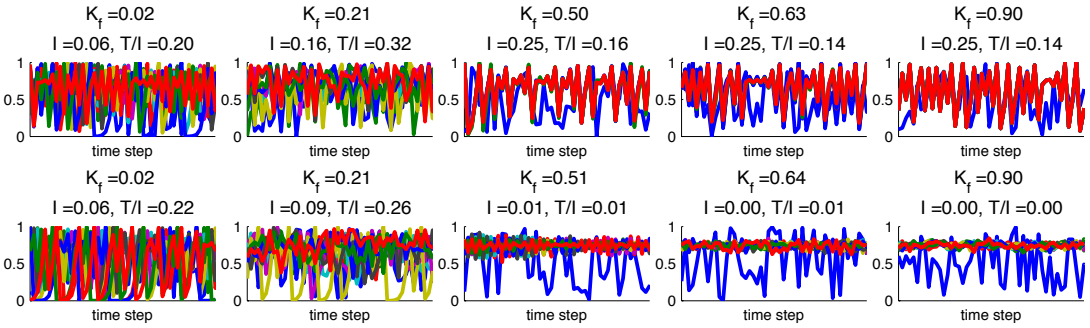

Figure 2.6: Time series for several generated networks with an external driver. 50 time steps shown for $n_{drivers} = 1$ for *(a)* Case 1 and *(b)* Case 2.

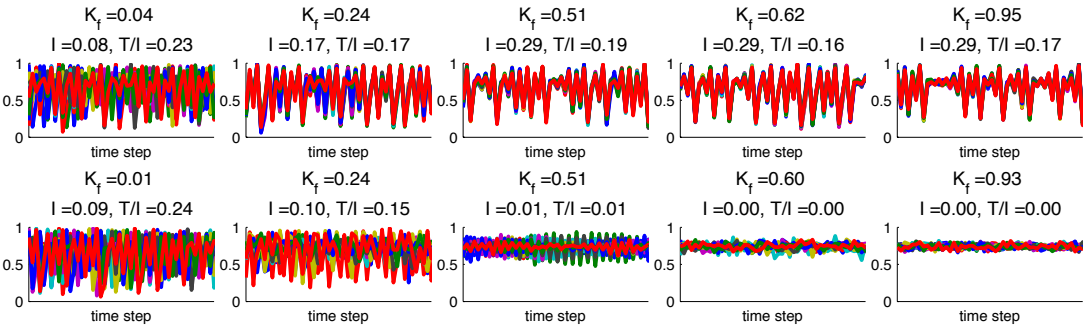

Figure 2.7: Time series for several generated networks with a noise component, $\epsilon_z = 0.1$. 50 time steps shown with $\epsilon_z = 0.1$ and no random driving nodes for *(a)* Case 1 *(b)* Case 2.

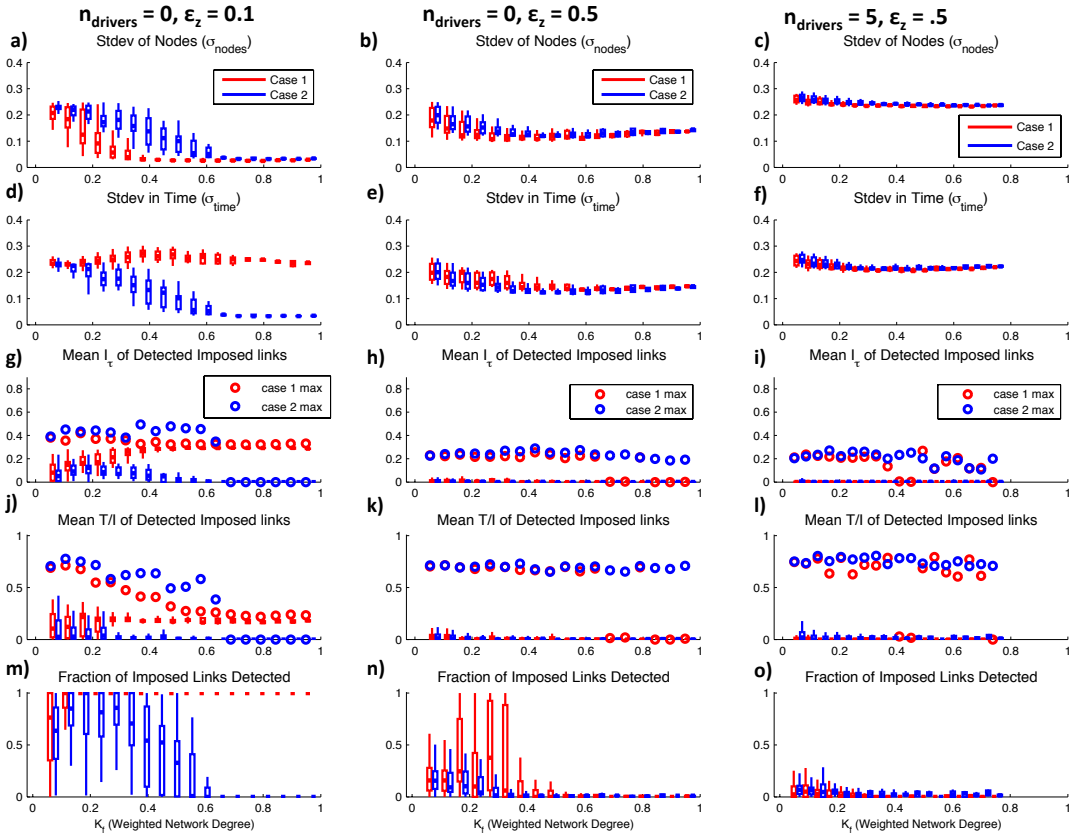

Figure 2.8: Behaviors of network configurations with different combinations of external drivers and noise. 882 networks over K_f range for *(left)* $n_{drivers} = 0$ and $\epsilon_z = 0.1$, *(middle)* $n_{drivers} = 0$ and $\epsilon_z = 0.5$, and *(right)* $n_{drivers} = 5$ and $\epsilon_z = 0.5$. *(a-c)* Standard deviation of nodes. *(d-f)* Standard deviation in time. *(g-i)* Box plots indicate mean I_τ, and open circles are maximum detected I_τ. *(j-l)* mean and maximum detected T/I values as in *(g-i)*. *(m-o)* Fraction of all imposed links that were correctly identified as peak time dependencies through detected I_τ.

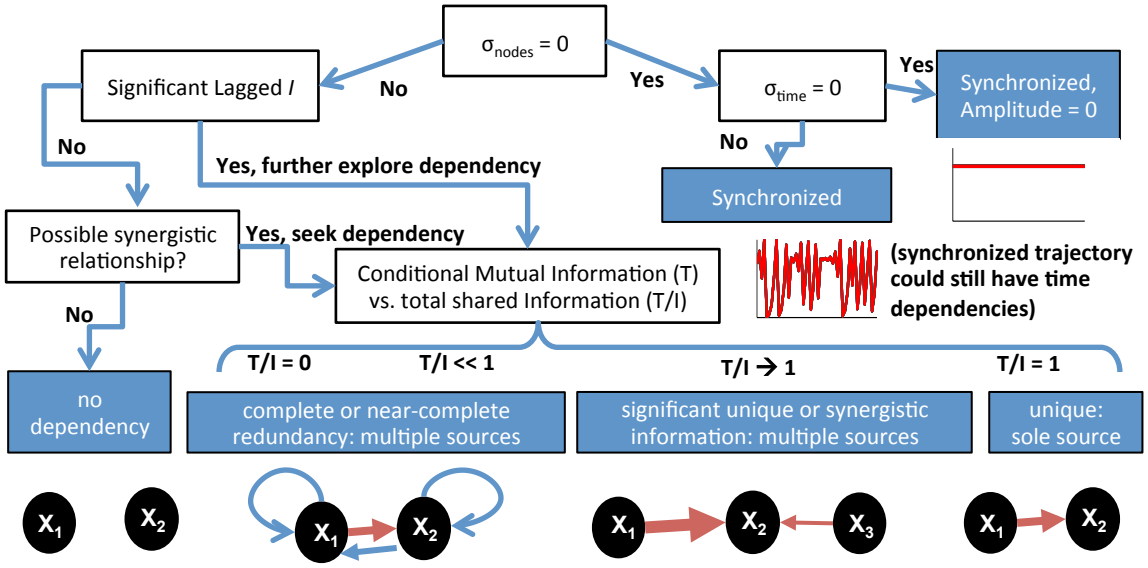

Figure 2.9: Flow chart of range of network dynamics that can be identified using information theory measures. Nodes are synchronized if $\sigma_{nodes} = 0$ and synchronized to zero-amplitude trajectories if $\sigma_{time} = 0$. In asynchronous cases, the absence of statistically significant I_τ indicates a disconnected network in the case of no synergistic shared information. Otherwise, the dependencies between nodes can be further explored with conditional and total information measures (T/I). If $T/I = 0$, multiple sources are completely redundant with each other. If $T/I = 1$, there is only 1 unique source providing information to the target node. In between, sources can be partially redundant, synergistic or unique.

CHAPTER 3

TEMPORAL INFORMATION PARTITIONING: CHARACTERIZING SYNERGY, UNIQUENESS, AND REDUNDANCY IN INTERACTING ENVIRONMENTAL VARIABLES

3.1 Introduction [1]

Multivariate observation methods are necessary to understand dynamical behaviors within systems of interacting variables. The behavior of a single variable within one such system can be attributed to external forcing, or to direct and indirect dependencies that arise from forcing and feedback interactions. The characterization of these dependencies enables a deeper understanding of system behavior. It additionally provides potential for predictive modeling, and may reveal shifts in dynamical responses to perturbations or stresses. Information theory, based on Shannon Entropy $H(X) = -\sum p(x) \log_2 p(x)$, for a random variable X with probability distribution function (*pdf*) $p(x)$, provides effective metrics with which to determine the characteristics of such dependencies. Information measures have been used in various geoscience contexts to measure complexity, dependencies, and driving or causal mechanisms [Balasis et al., 2013]. In ecohydrology, [Ruddell and Kumar, 2009a], [Ruddell and Kumar, 2009b], and [Kumar and Ruddell, 2010] used information theory in a network context to reveal shifts in forcing and feedback over drought and healthy growing seasons based on flux tower data. From a systems perspective, [Sharma and Mehrotra, 2014] used information measures to formulate prediction models of natural systems based on observed data. More broadly, information theory has been applied to gain insight into complex interactions between multiple variables in diverse fields such as neuroscience [Barrett, 2015], cardiology [Faes et al., 2015], and industrial engineering [Duan et al., 2013].

The studies mentioned above employ various information measures to either characterize the variability or complexity of a single time-series variable or to detect information "flow" between τ-lagged "source" variables to "target" variables, where directionality between sources and targets are defined based on the time lag τ between them. In a complex system, we expect that many lagged source variables $(X_{s1}, X_{s2}...X_{sn})$ may influence the behavior of

[1] This chapter is in review for Water Resources Research [Goodwell and Kumar, 2017a]

a given target (X_{tar}). This indicates that information metrics detected between individual sources and targets ($X_s \rightarrow X_{tar}$) do not reflect the presence of other sources that may be intermediate drivers or causal influences. For example, a variable may be synchronized with another variable that is actually driving the target, or many source variables may be jointly driving the target. While multivariate mutual information $I(X_{s1}, X_{s2}...X_{sn}; X_{tar})$ and conditional mutual information $I(X_{s1}; X_{tar}|X_{s2}...X_{sn})$ quantify the total and conditional amounts of information that multiple sources provide to a target, respectively, they do not separate unique influences or indicate in what ways sources may be acting jointly.

Recent research on information partitioning [Williams and Beer, 2010] has enabled more precise classification of the nature of shared information between sources and target variables. Information partitioning categorizes shared information quantities between multiple sources and a target as either unique, synergistic, or redundant. Redundant information (R) is the information that multiple sources provide to a target such that they overlap in their information content. Unique information (U) from a source X_s refers to the information it shares with a target X_{tar} that is not redundant with information provided by another source. Synergistic information (S) refers to the information that two sources provide to a target only when known jointly. A traditional example of synergistic information is the binary *XOR* operation [Williams and Beer, 2010, Griffith and Ho, 2015, Bertschinger et al., 2014], in which no uncertainty is reduced for the target when the sources are known individually, but all is reduced when both are known together. This partitioning of shared information into unique, synergistic, and redundant components opens up tremendous potential for exploring deep and nuanced dependencies in environmental signals. However, this has not yet been explored.

Our goal in this paper is to establish a methodological framework for information partitioning that is relevant to the study of interacting environmental variables. When two sources inform a target, the target may receive information uniquely from each source, redundantly from both sources, synergistically from both sources, or in some combination of these. This partitioning can reveal much about the function or process that physically links the two sources with the target, in addition to any relationship between sources. For example, both wind speed (WS) and air temperature (Ta) influence relative humidity (RH), but through the different mechanisms of atmospheric mixing and water vapor holding capacity, respectively. Although the functional form of a joint relationship such as this may be unknown, we would expect both WS and Ta to exert at least partially unique influences to RH. In contrast, variables such as Ta and soil temperature may be somewhat more synchronized

or otherwise correlated, and provide a higher proportion of redundant information to RH. This interpretation of the nature of a dependency in a multivariate context will allow us to understand detected dependencies in natural systems where functional forms of relationships between variables are not necessarily known.

In practice, U, R, and S components are not directly computable from Shannon Entropy measures, and several metrics have been proposed to perform information partitioning [Barrett, 2015, Griffith and Ho, 2015]. For example, in a cardiovascular application, [Faes et al., 2015] used existing R and S metrics to show that heart period, arterial pressure, and respiration flow are intertwined elements that control heart rate together, and that there is higher R between subsystems during fast-paced breathing. However, these measures take a limited perspective based on minimum mutual information (MMI) [Barrett, 2015], in which R is defined as the quantity of information that the weaker source shares with the target. In other words, a quantity of information "looks the same" to the target, regardless of the source from which it came, so that only the stronger of two sources can provide unique information. However, we consider that two sources may influence the same target partially or completely independently, such as in a natural system where variations in air temperature and windspeed drive relative humidity in different ways. In this vein, we propose a new measure of R called Re-scaled Redundancy (R_s) that scales redundant information based on source dependencies. This re-scaling is important for application to environmental time-series datasets because it reflects the simultaneous existence of many physical processes that are not necessarily redundant. In this study, we consider information partitioning between information shared from two sources to a target. The methods presented here could be generalized to the partitioning of information from many sources if sufficient data is available to compute higher order pdfs.

This paper is organized as follows. In Section 3.2, we present several examples with varying complexities to build the case for our information partitioning methodology as a novel approach to study interactions between environmental signals. We first introduce temporal information partitioning and discuss it in the context of traditional information measures, specifically Transfer Entropy (T_E). We use a hypothetical example to illustrate how a widely used current information partitioning method (namely, minimum mutual information) overestimates redundancy, and propose an alternative approach. We compare methods with Gaussian sum and multivariate autoregressive cases to explore how information components depend on joint source variability, differing source strengths, noise, and dynamical dependencies.

Section 3.3 focuses on the application of information partitioning to time-series datasets. While the information partitioning approach has been applied to time-series data in neural and cardiovascular applications [Erramuzpe et al., 2015, Faes et al., 2015], these have assumed either Gaussianity or simple *pdf* estimation techniques, and use minimum mutual information as a redundancy measure. We use robust *pdf* estimation methods to explore the influence of noise and source correlations on established information theory measures and S, R, and U components using generated chaotic datasets. Here we reveal several differences between analytically derived information components and those detected based on empirically estimated *pdf*s.

In Section 3.4, we apply our technique of information partitioning with R_s to assess the nature of information transferred from air temperature (Ta) and wind speed (WS) to relative humidity (RH) as measured from 1-minute environmental signals at a weather station. This is the first rigorous application of information partitioning to environmental datasets that involve noise, sparse data, an atom-at-zero effect, and other properties typical of environmental signals that constitute challenges in *pdf* estimation and subsequent information partitioning. Our analysis shows that information partitioning can reveal aspects of interactions that cannot be obtained using linear correlation or other types of information theory measures.

3.2 Information partitioning into synergistic, unique, and redundant components

3.2.1 Existing information measures

Multivariate mutual information $I(X_{s1}, X_{s2}...X_{sn}; X_{tar})$ captures the reduction in uncertainty of a target variable X_{tar} when the lagged source variables $X_{s1}...X_{sn}$ are known. In this paper, we consider the simplest case which involves two sources, X_{s1} and X_{s2}, such that the quantity $I(X_{s1}, X_{s2}; X_{tar})$ is the information "whole", since it represents the total amount of information (*bits*) that the sources share with the target. This is a function of the *pdf* of the three variables as follows:

$$I(X_{s1}, X_{s2}; X_{tar}) = \sum p(x_{s1}, x_{s2}, x_{tar}) \log_2 \left(\frac{p(x_{s1}, x_{s2}, x_{tar})}{p(x_{s1}, x_{s2}) p(x_{tar})} \right) \tag{3.1}$$

Conditional mutual information is the uncertainty of the target that is reduced due to the knowledge of a source (e.g. X_{s1}) beyond that which is reduced due to knowledge of another

source (e.g. X_{s2}), and is computed as follows:

$$I(X_{s1}, X_{tar}|X_{s2}) = I(X_{s1}, X_{s2}; X_{tar}) - I(X_{s2}; X_{tar}). \tag{3.2}$$

When the second source X_{s2} is considered to be the history of the target (i.e. $X_{s2} = X_{tar}(t-\tau)$), the conditional mutual information is equivalent to Transfer Entropy, that is, $T_E(X_{s1} \rightarrow X_{tar}|X_{tar}(t-\tau))$ [Schreiber, 2000]. T_E is equivalent to Granger causality for a Gaussian case [Stramaglia et al., 2012], and has been applied in many studies to address the problem of distinguishing between forcing and feedback-induced interactions [Ruddell and Kumar, 2009a]. Some studies define T_E conditioned on other source variables X_{s2} besides the target history $X_{tar}(t-\tau)$ to address systems with multiple drivers or indirect causality [Runge, 2015]. Due to these features, T_E has at times been interpreted as a measure of causality. However, this causal influence is only in a statistical sense, and does not necessarily indicate forcing mechanisms within the system. Particularly in systems with feedback, the notion of "causality" is obscured and T_E does not distinguish between feedbacks, induced feedbacks, and intermediate drivers [Goodwell and Kumar, 2015].

Another seemingly unintuitive feature of T_E, and conditional information in general, is the potential for $I(X_{s1}; X_{tar}|X_{s2})$ to be greater than$I(X_{s1}; X_{tar})$. This indicates that the information provided to the target by the source conditioned on the target history is greater than the information provided by the source alone. In other words, X_{s1} provides information to X_{tar} *together with* X_{s2}, rather than *overlapping with* X_{s2}. This occurs, for example, when the knowledge of two variables are both needed to predict a third, such as in the binary *XOR* operation. We next briefly introduce the concept of information partitioning and its utility in both interpreting traditional information measures such as T_E [Williams and Beer, 2011], and revealing properties of interacting variables.

3.2.2 Existing information partitioning methodology

In a system where two sources share information with a target, the total information quantity $I(X_{s1}, X_{s2}; X_{tar})$ can be partitioned into unique, synergistic, and redundant components [Barrett, 2015, Williams and Beer, 2010]. This partitioning, also called the partial information decomposition, is as follows [Williams and Beer, 2010]:

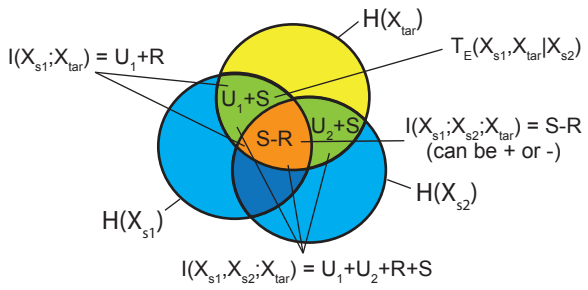

Figure 3.1: Venn diagram of entropy and information measures. Circle areas represent Shannon Entropy and overlapping areas represent reductions in uncertainty. Information partitioning components and traditional measures such as Transfer Entropy (T_E) are labeled. All regions are non-negative except for the center overlapping area that represents interaction information, which can be negative or positive depending on the relative dominance of synergy or redundancy.

$$I(X_{s1}, X_{s2}; X_{tar}) = U_1(X_{tar}; X_{s1}) + U_2(X_{tar}; X_{s2}) + R(X_{tar}; X_{s1}, X_{s2}) + S(X_{tar}; X_{s1}, X_{s2}) \tag{3.3}$$

where U_1, U_2, R, and S are non-negative quantities. R is information that both sources share with the target *redundantly*, U_1 and U_2 are information that only X_{s1} and X_{s2}, respectively, share with the target *uniquely*, and S is information that is provided to the target only when both sources are known together, or *synergistically*. While R can be interpreted as overlapping shared information, S is a cooperative provision of shared information that causes the information "whole" to be greater than the union of the "parts" [Griffith and Koch, 2014] in that the joint provision of information from sources may be greater than the sum of their individual contributions.

Individual mutual information between each source and the target decompose as [Williams and Beer, 2010]:

$$I(X_{s1}; X_{tar}) = U_1(X_{s1}; X_{tar}) + R(X_{tar}; X_{s1}, X_{s2}) \tag{3.4}$$

$$I(X_{s2}; X_{tar}) = U_2(X_{s2}; X_{tar}) + RX_{tar}; X_{s1}, X_{s2}). \tag{3.5}$$

[Williams and Beer, 2011] apply information partitioning to T_E to show that while the

conditioning omits the redundant quantity $R(X_{tar}; X_{s1}, X_{s2})$, it adds in the synergistic component $S(X_{tar}; X_{s1}, X_{s2})$ if it exists, so that $T_E \equiv I(X_{s1}; X_{tar}|X_{s2}) = U_1 + S$ (Figure 3.1). This can be explained by the interaction information II, which is defined as:

$$II \equiv I(X_{s1}; X_{s2}; X_{tar}) = I(X_{s1}; X_{tar}|X_{s2}) - I(X_{s1}; X_{tar}). \tag{3.6}$$

II can be either positive or negative (Figure 3.1), and it can be shown [Williams and Beer, 2011] that $II = S - R$, such that $II > 0$ indicates a dominantly synergistic relationship and $II < 0$ indicates dominant redundancy. Interaction information is also known as the co-information [Bell, 2003, Rosas et al., 2016] and the redundancy-synergy index (RSI) [Timme et al., 2014]. It can be interpreted as the information "whole" ($I(X_{s1}, X_{s2}; X_{tar})$) minus the "sum of the individuals" ($I(X_{s1}; X_{tar}) + I(X_{s2}; X_{tar})$). Therefore, $II < 0$ indicates that redundancy is greater than synergy, i.e. that the "sum of the individuals" ($I(X_{s1}; X_{tar}) + I(X_{s2}; X_{tar})$) is greater than the "whole" ($I(X_{s1}, X_{s2}; X_{tar})$). In other words, the whole is greater than the sum of the parts (i.e. $II > 0$) when synergy is greater than redundancy. Since Equation (3.6) allows the direct computation of II, the difference between S and R can be estimated directly from data without partitioning $I(X_{s1}, X_{s2}; X_{tar})$ directly. In cases where $S > R$,$I(X_{s1}; X_{tar}|X_{s2}) > I(X_{s1}; X_{tar})$. In general, while the ratio of T_E (a special case of Equation (3.2) where $X_{s2} = X_{tar}(t - \tau)$) to $I(X_{s1}, X_{s2}; X_{tar})$ (Equation (3.3)) gives a relative indication of synergy, uniqueness, and redundancy [Goodwell and Kumar, 2015], a direct partitioning would provide a more clear interpretation of relationships between variables.

From Equations (3.3)-(3.5), it is evident that an additional formula for any single U, R, or S component will enable the computation of all others, since the mutual information values are directly computable from the pdfs. To this end, various methods of computing U [Bertschinger et al., 2014], S [Griffith and Koch, 2014, Olbrich et al., 2015], and R [Williams and Beer, 2010, Harder et al., 2013, Griffith and Ho, 2015] components have been proposed, although there is no universal agreement on the appropriate method. The most common approach is to first compute R, and several proposed measures have been shown to reduce to simply the minimum shared information between the target X_{tar} and either source as follows [Barrett, 2015]:

$$R_{MMI} = \min[I(X_{s1}; X_{tar}), I(X_{s2}; X_{tar})] \tag{3.7}$$

where MMI denotes "minimum mutual information". This formulation actually represents a maximum bound for redundancy, since it assumes that all information provided by the weaker

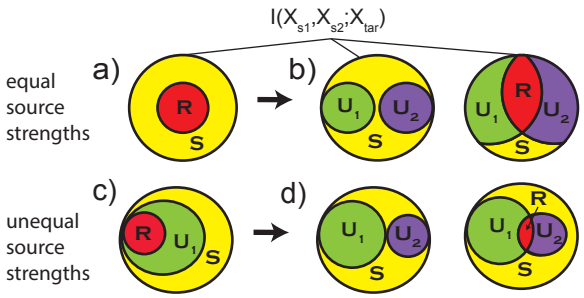

Figure 3.2: Illustration of information partitioning into U, R, and S components. Circle areas represent information partitioning of $I(X_{s1}, X_{s2}; X_{tar})$ into U_1, U_2, R and S components. *(a)* If two sources share equal amounts of information with a a target, R_{MMI} results in only S and R components. *(b)* If redundancy is defined based on source correlation, this case could result in many different combinations of U_1, U_2, R and S. *(c)* Similarly, if one source is stronger than another, R_{MMI} results in only three possible information components since all information from the weaker source is defined as redundant. *(d)* An alternate partitioning method could enable a wider range of information partitioning results depending on source correlations.

source is redundant [Barrett, 2015, Griffith and Ho, 2015]. The use of R_{MMI} as a redundancy metric greatly decreases the number of ways in which information can be partitioned into unique, synergistic, and redundant components. For example, for two sources of equal strength ($I(X_{s1}; X_{tar}) = I(X_{s2}; X_{tar})$), a partitioning based on R_{MMI} can only result in S and R (Figure 3.2a) components rather than any combination of U, R, and S (Figure 3.2b). Similarly, when two sources have different strengths, only three information components are possible (Figure 3.2c,d) since all information provided by the weaker source is defined as redundant. In the next subsection, we argue that this estimate of $R = R_{MMI}$ is inadequate for studies of environmental signals, and present an alternate approach.

3.2.3 Redundancy and source dependency: A die-rolling example

Environmental signals represent different variables, and detected dependencies between them may reflect multiple processes related to energy, water, or nutrients. Additionally, the functional form of relationships may not be known. We first focus understanding how multiple sources provide unique information. We present a simple example case, where X_{s1} and X_{s2} are independent rolls of six-sided dice and $X_{tar} = X_{s1} + X_{s2}$ (Figure 3.3a). Based on the *pdfs*, each die shares 0.689 *bits* of information with the sum X_{tar} ($I(X_{s1}; X_{tar}) = I(X_{s2}; X_{tar}) = 0.689$),

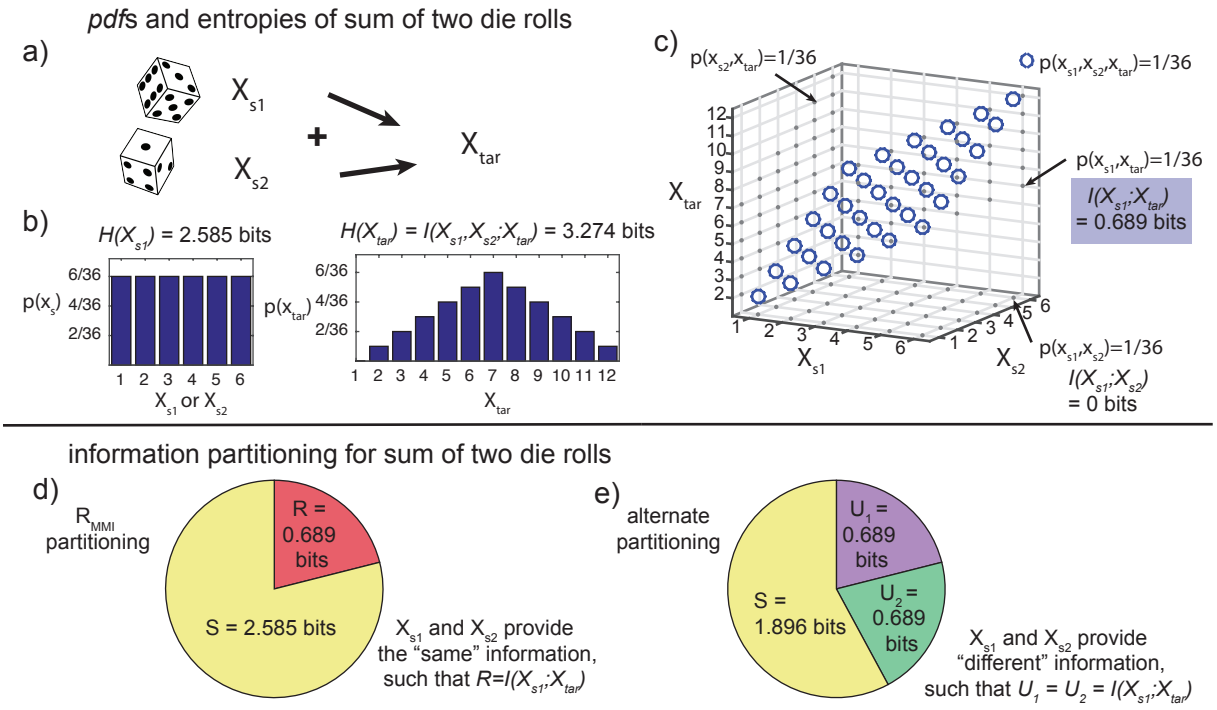

Figure 3.3: Illustration of information partitioning applied to a sum of two die rolls. In this case, information partitioning varies depending on the partitioning method used. *(a)* The system is composed of two dice, X_{s1} and X_{s2} and their sum X_{tar}. *(b) Pdf*s of X_{tar}, X_{s1}, and X_{s2} and corresponding values of Shannon Entropy H. *(c)* 3d *pdf* (open circles) and corresponding marginal 2d *pdf*s (gray dots). *(d)* Based on partitioning of shared information with R_{MMI}, only R and S are detected. *(e)* We approach the partitioning from the perspective that since the dice are independent, the information they provide to the sum should be unique.

and jointly they share $H(X_{tar}) = 3.274$ *bits* of information with X_{tar} (Figure 3.3b,c). R_{MMI} results in the detection of 0.689 *bits* of redundant information ($R_{MMI} = 0.689$), no unique information ($U_1 = U_2 = 0$), and $S = H(X_{tar}) - R_{MMI} = 2.585$ *bits* (Figure 3.3d). This partitioning considers that all information provided to the target "looks the same", such that the unique information provided by the stronger source is equal to the difference in strength between the sources. The argument for this perspective is that the knowledge of either source (in this case, either die) effectively provides the same amount of information to the target (the sum), so that the information should be redundant.

We take the contrasting view that the *independent* die rolls should provide unique rather than redundant information to the sum, and that an information partitioning should result

in $R = 0$, thereby leading to $U_1 = U_2 = 0.689$, and $S = 1.896$ *bits* (Figure 3.3e). From our perspective, each of the multiple sources may represent a different variable or process that uniquely influences the target. For a general case where $X_{tar} = f(X_{s1}, X_{s2})$ and the functional forms of relationships between targets and sources may be unknown, we wish to distinguish between scenarios where the information is maximally redundant ($X_{s1} = f(X_{s2})$ or vice versa) or maximally unique (X_{s1} and X_{s2} are independent). In other words, if a target receives information from multiple sources, each source may uniquely provide some information that no other source provides, regardless of its relative strength. This requires an alternate definition of redundancy that involves an estimation of shared information that is due to source correlations versus individual influences. In a system of interacting time-series variables, a prevalence of highly correlated sources indicates that they are partially lag-synchronized due to either feedback or common forcing. In contrast, independent sources indicate that the sources influence the target via different mechanisms even though the strength (in *bits*) of their influences may be similar. For example, while air temperature and wind speed may influence relative humidity uniquely due to different forcing mechanisms, air temperature and solar radiation may provide redundant information due to their partial synchronization.

3.2.4 Re-scaled Redundancy shifts with source dependency

To account for source dependency in information partitioning, we develop a measure called Re-scaled Redundancy, R_s, as follows:

$$R_s = R_{\min} + I_s(R_{MMI} - R_{\min}) \tag{3.8}$$

where

$$R_{\min} = \max(0, -II) \tag{3.9}$$

and

$$I_s = \frac{I(X_{s1}; X_{s2})}{\min[H(X_{s1}), H(X_{s2})]}. \tag{3.10}$$

$R_{\min}$ represents a minimum boundary for R. For cases where $II = S - R$ (Equation (3.6)) is negative, indicating that $R > S$, any chosen metric for R must be greater than or equal to the positive value of $R - S$ so that S is non-negative. For example, if two sources are

detected to both share 1 *bit* of information to a target individually and 1.5 *bits* together ($I(X_{s1}; X_{tar}) = I(X_{s2}; X_{tar}) = 1$ and $I(X_{s1}, X_{s2}; X_{tar}) = 1.5$), $II = -0.5$ *bits*. In this case, the value of an R metric must be at least 0.5 *bits*, regardless of the detected correlation between the sources. A lower R would cause $S < 0$, which violates the non-negativity of information partitioning [Williams and Beer, 2010]. While $R_{\min}$ and R_{MMI} are minimum and maximum bounds, respectively, I_s represents the source dependency, so that R_s varies between $R_{\min}$ for independent sources and R_{MMI} for maximally dependent sources. In the die rolling example, $I_s = 0$ and $R_{\min} = 0$, so Re-scaled Redundancy results in desired partitioning where all information is synergistic and unique.

When considering information partitioning and other information theoretic measures in a dynamical system, we necessarily assume stationarity within the time period of interest. The violation of this assumption can lead to misleading interpretations of information partitioning. For example, air temperature may be driven by different mechanisms before and after sunrise, due to the presence or absence of solar radiation. This represents the emergence of a new source of influence (solar radiation) within a time window. Another situation arises if source dependencies change over time. For example, vegetation activity may switch between water-limited and energy-limited conditions as moisture conditions change. In these situations, synergistic or redundant information from two sources may be detected even if they do not always influence the target simultaneously. Additionally for the R_s metric, source dependencies that shift over time would influence the determination of U, R, and S components, regardless of whether the sources influence the target when they are more or less dependent. In a natural system where dependencies between energy and water components are changing with weather conditions and seasons, we partition data into short time windows to capture these different types of influences. Specifically in the weather station application we present in Section 4, we define windows based on night and daytime conditions in order to capture separate types of influences for each period. However, without auxiliary knowledge, it is generally not possible to fully distinguish a threshold type of behavior from true redundant or synergistic relationships within a system.

In the following subsection, we explore properties of the R_s metric over a range of source correlations using a Gaussian sum example case. In this case and following example cases, we note that the target variable is purely a function of the two sources (i.e. $H(X_{tar}) = I(X_{s1}, X_{s2}; X_{tar})$). We also consider the influences of a range of source variances (σ^2) and noise levels on our information partitioning R_s metric versus R_{MMI}. In a natural system, external or unmeasured sources influence the target such that it is not completely predictable (i.e.,

$H(X_{tar}) > I(X_{s1}...X_{sn}; X_{tar})$). However, these simple examples enhance our understanding of how properties of source dependencies, noise, and the functional relationships between sources and targets influence information partitioning.

3.2.5 Information partitioning of Gaussian sums

Here we seek to test our R_s metric and subsequent information partitioning for a simple case where the functional form of the relationship is known, but a variety of dependencies between source variables may exist. In the context of environmental signals, these test cases address the characterization of joint dependencies that may vary depending on the relative strength of interaction between each source and target, in addition to the synchronization between sources. For example, vegetation growth in an ecosystem may be either water or energy limited, and vegetation properties such as *NDVI* (Normalized Difference Vegetation Index) may depend on the phase difference between annual cycles of rainfall and potential evaporation [Feng et al., 2015]. We expect our method to reveal high uniqueness and synergy when sources are independent, and increasing redundancy as sources become correlated or synchronized. We study the specific case of $X_{s1} = N[0, \sigma^2_{x_1}]$ and $X_{s2} = N[0, \sigma^2_{x_2}]$ with correlation $\rho \equiv \rho_{x_{s1},x_{s2}}$, and $X_{tar} = X_{s1} + X_{s2}$. [Barrett, 2015] studied the partitioning of mutual information for a range of Gaussian cases based on $R = R_{MMI}$ in Equation (B.7) and found that net synergy ($S > R$) is a common occurrence in these systems. This indicates that $R_{\min} = 0$ for many typical scenarios, but source dependencies as measured by I_s cause R_s to differ from R_{MMI}. We assess these differences and their influences on S and U over a range of σ and ρ parameters. Here, V_1 and V_2 are the measured sources which consist of the sources X_{s1} and X_{s2}, respectively, with some level of uncorrelated added noise as follows:

$$V_1 = X_{s1} + N[0, \sigma^2_{n1}] \tag{3.11}$$

$$V_2 = X_{s2} + N[0, \sigma^2_{n2}] \tag{3.12}$$

$$X_{tar} = X_{s1} + X_{s2}. \tag{3.13}$$

The entropy and mutual information measures for V_1, V_2, and X_{tar} can be derived as functions of $\sigma_{x_{s1}}$, $\sigma_{x_{s2}}$, $\sigma_{x_{n1}}$, $\sigma_{x_{n2}}$, and ρ, and are listed in the Appendix as Equations (A.1)-(A.6). We note that only the source entropies (Equations (A.1)-(A.2)) depend on the noise components of V_1 and V_2. We first focus on a noiseless case ($\sigma^2_{n1} = \sigma^2_{n2} =$

0), then extend the analysis to assess the influence of source noise. We note that this example case is based on a continuous distribution, for which information theoretic measures are characterized as differential entropies [Lazo and Rathie, 2006], and rules differ between discrete and discontinuous cases. For examples, the rule that $I(V_1;V_2) \leq \min[H(V_1), H(V_2)]$ does not apply [Cover and Thomas, 2006], and measures are not bounded by a number of possible categories as in a discrete case. However, we use this Gaussian example to introduce a range of scenarios where sources influence targets with different strengths and correlations by considering a domain and parameter range where $I(V_1;V_2) \geq \min[H(V_1), H(V_2)]$ holds.

Redundancy and synergy for a noiseless Gaussian sum

A threshold on $(\sigma_{x_{s1}}, \sigma_{x_{s2}}, \rho)$ combinations for the Gaussian sum case occurs when $R_{\min} \geq R_{MMI}$. This statement is equivalent to $I(V_1;V_2) \geq \min[H(V_1), H(V_2)]$, which can occur for continuous distributions. For the Gaussian sum case, we omit any $(\sigma_{x_{s1}}, \sigma_{x_{s2}}, \rho)$ combinations for which $R_{\min} \geq R_{MMI}$ (dark gray shading in Figure 3.4).

For a noiseless case ($\sigma_{x_{s1}} = \sigma_{x_{s2}} = \sigma$), we insert Equations (A.1)-(A.6) into Equation (3.9) to derive the following condition for $R_{\min} > 0$:

$$\ln\left(\frac{c\sigma^2 b}{a}\right) > 0. \tag{3.14}$$

in which the terms are as follows: $a = 2 + \rho$, $b = 1 - \rho^2$, and $c = 2\pi e$. This represents the condition for which redundancy must be non-zero regardless of source correlation (hatched regions in Figure 3.4). As noted by [Barrett, 2015], most (σ, ρ) combinations result in $S > R$ (i.e. $II > 0$, Equation (3.6)) such that $R_{\min} = 0$, but large $|\rho|$ and small σ combinations result in $S < R$.

We solve for U_1, U_2, and S components from Equations (3.3)-(3.5) to arrive at the following expression for this noiseless $\sigma_{x_{s1}} = \sigma_{x_{s2}} = \sigma$ case:

$$S(\textit{Gaussian sum}) = R + \frac{1}{2}\ln\left(\frac{c\sigma^2 b^2}{a}\right) \tag{3.15}$$

Here, any redundancy measure may be inserted for R. For example, R_s varies based on the sign of $II = S - R$ (Equation 3.6) as follows:

$$R_s(\textit{Gaussian sum}) = \begin{cases} \frac{1}{2} I_s \ln\left(\frac{a}{b}\right), & \text{if } II \geq 0. \\ \frac{1}{2}\ln\left(\frac{a}{c\sigma^2 b^2}\right) + \frac{1}{2} I_s \ln\left(c\sigma^2 b\right), & \text{otherwise.} \end{cases} \tag{3.16}$$

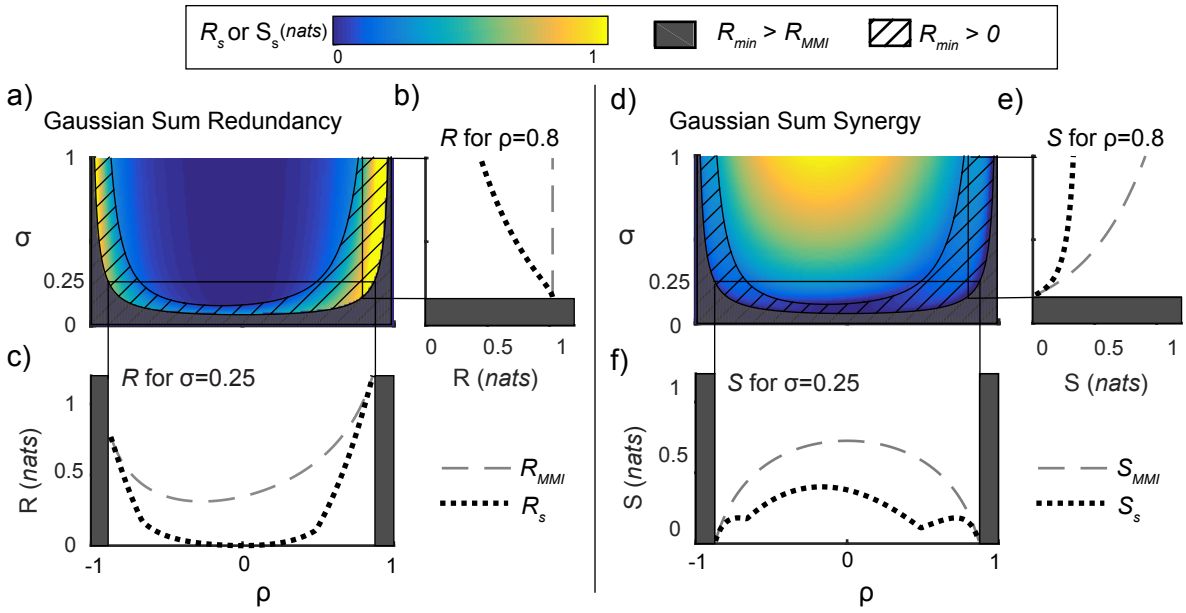

Figure 3.4: Illustration of R and S for Gaussian sum cases with varying parameter values. R and S are shown for different partitioning methods for a noiseless case with equal source variances ($\sigma_{x_1} = \sigma_{x_2} = \sigma$). *(a)* R_s computed from Equation (3.8) for Gaussian sum case ($\sigma_{x_1} = \sigma_{x_2}$). Hatching indicates R_1 crossing minimum boundary corresponding to $S - R < 0$ (Equation (3.14)), and gray shading indicates $R_{min} > R_{MMI}$ corresponding to $I(X_1; X_2) > \min[H(X_1), H(X_2)]$. Redundancy measures R_{MMI} (Equation (3.7)) and R_s (Equation (3.8)) vs. σ for constant $\rho = 0.8$ show decreasing R with increasing σ. Redundancy measures vs. ρ for constant $\sigma_{x_1} = \sigma_{x_2} = \sigma = 0.25$ show decreased R for decreasing source correlation ρ. *(b)* S computed from Equations (3.3) and (3.8) for Gaussian sum case corresponding to *(a)*. Synergy measures computed based on R_{MMI}, and R_s for constant $\rho = 0.8$, and constant $\sigma = 0.25$.

where the source dependency

$$I_s(Gaussian\ Sum) = -\frac{\ln(b)}{\ln(c\sigma^2)} \tag{3.17}$$

increases with $|\rho|$ as expected, but decreases with σ since the scaling is relative to the source entropy $H(V)$. For cases where $II < 0$, R is bounded by $R_{\min}$ and is a higher order function of ρ. We see for the noiseless $\sigma_{x_{s1}} = \sigma_{x_{s2}}$ case that S is maximized for large σ (Figure 3.4d) and a negative ρ. The breaks in smoothness of the S curve (Figure 3.4f) correspond to points at which $II = 0$ ($S = R$) so that $R_{\min}$ shifts from $R_{\min} = 0$ to a positive value.

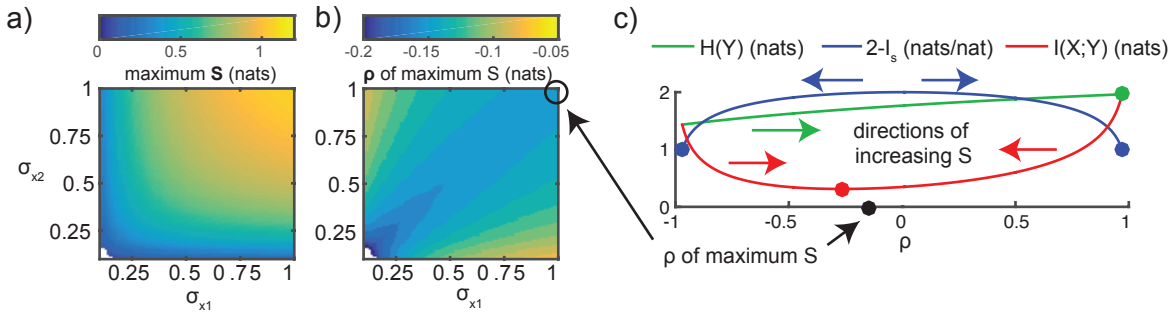

Figure 3.5: Illustration of S for Gaussian sum cases, which is maximized for a negative correlation ρ. Here, maximum S occurs for a combination of high target entropy, low source correlation, and high shared information between sources and their sum. *(a)* Maximum S $(nats)$ over range of $-1 < \rho < 1$ for $(\sigma_{x_1}, \sigma_{x_2})$ pairs. Increased σ leads to increased S. *(b)* Value of ρ associated with maximum S from *(a)* is always negative, and most negative for the $\sigma_{x_1} = \sigma_{x_2}$ cases. *(c)* For the case where $\sigma_{x_1} = \sigma_{x_2} = 1$, individual terms of Equation (3.18) vary with ρ. Dots indicate minimum and maximum values for each term. Synergistic information S $(nats)$ from Equation (3.18) is maximized for a negative ρ.

We do not expand Equation (3.16) for cases where $\sigma_{x_1} \neq \sigma_{x_2}$, but we note that maximum S occurs for $\rho < 0$ for all $(\sigma_{x_1}, \sigma_{x_2})$ scenarios (Figure 3.5b) and that the value of the maximum S $(nats)$ increases with σ (Figure 3.5a). In the following discussion, we illustrate this heuristically for the $\sigma_{x_{s1}} = \sigma_{x_{s2}}$ case where $II \geq 0$. Equation (3.3) can be solved for S to obtain the following relationship:

$$S(Gaussian\ Sum) = H(X_{tar}) - (2 - I_s)I(V; X_{tar}). \tag{3.18}$$

S is maximized when $H(X_{tar})$ is maximized and $(2 - I_s)$ and $I(V; X_{tar})$ are minimized simultaneously (Figure 3.5c). $I(V; X_{tar})$ is minimized for $\rho^* = \sqrt{3} - 2 = -0.268$, $(2 - I_s)$ is

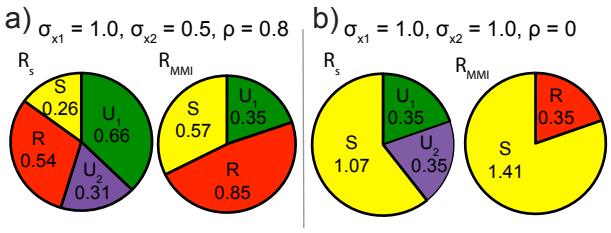

Figure 3.6: Pie charts of information partitioning results for noiseless Gaussian sum cases. For $\sigma_{x_1} = 1$, information partitioning for two cases with varied σ_{x_2} and source correlation ρ. Circle areas represent total shared information $I(Y; X_1, X_2) = H(Y)$ (*nats*), and results for U_1, U_2, R, and S decomposition shown based on previous R_{MMI} metric and R_s metric introduced here. *(a)* When $\sigma_{x_2} < \sigma_{x_1}$, R_{MMI} omits any U from the smaller source, whereas R_s retains two U components since the variables are only partially redundant. *(b)* When $\sigma_{x_2} = \sigma_{x_1}$ and $\rho = 0$, R_s results in only S, U_1, and U_2 since sources are independent.

minimized for $|\rho| = 1$, and $H(X_{tar})$ increases with ρ. We see that the second term dominates, resulting in a maximum S at a small negative ρ (Figure 3.5c). Qualitatively, maximum S occurs for a balance between the following components:

1. A high total information content (a large "whole" $= H(X_{tar}) = I(V_1, V_2; X_{tar})$).

2. Low individual shared information (a small "sum of parts", given as $I(V_1; X_{tar}) + I(V_2; X_{tar})$).

3. Highly correlated sources (a large I_s) so that R approaches R_{MMI}.

For the Gaussian addition example, this balance is achieved when sources are slightly negatively correlated. This demonstrates how both source→target dependencies and source→source dependencies impact shared information. Even when the functional relationship is very simple, such as addition, different $(\sigma_{x_{s1}}, \sigma_{x_{s2}}, \rho)$ combinations result in a variety of U_1, U_2, R, and S components (Figure 3.6). For cases where sources are unequal in strength due to different variances (Figure 3.6a) or equal in strength and uncorrelated (Figure 3.6b), R_s identifies both U_1 and U_2 components when sources are not completely synchronized.

This Gaussian sum analysis shows that our proposed measure R_s captures the balance of information components in a way that reveals not only the quantity of shared information in a source→target relationship, but its synergy, uniqueness, or redundancy given another source.

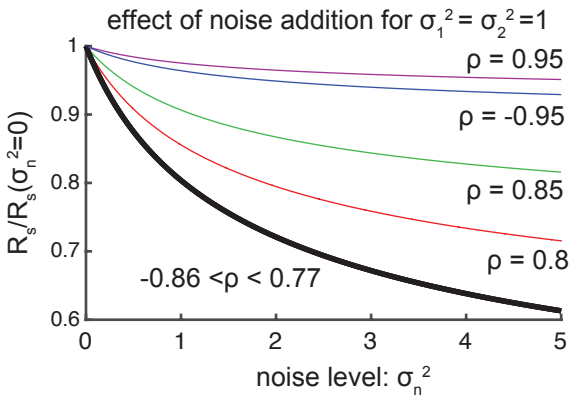

Figure 3.7: The influence of noise on R_s for a Gaussian sum case. When noise ($\sigma_{n1} = \sigma_{n2} = \sigma n > 0$) is added to the sources, R_s decreases relative to the R_s for the noiseless case for a constant value of ρ.

Redundancy for a Gaussian sum with noise

When variables represent diverse quantities such as solar radiation, temperature, rainfall, or windspeed, we expect shared information to consist of all components. In addition to multiple processes, unobserved drivers and measurement errors introduce noise to our measurements of complex systems. To understand this, we extend the Gaussian sum example to explore the influence of uncorrelated noise on information partitioning.

The addition of noise in the form of $\sigma_{n1} > 0$ and/or $\sigma_{n2} > 0$ in Equations (3.11)-(3.12) to the Gaussian sum case does not influence $R_{\min}$ or R_{MMI} since these measures are only dependent on the information shared between each source and target pair. However, as shown in Equations (A.1)-(A.2), $H(V_1)$ and $H(V_2)$ increase with increasing σ_{n1} and σ_{n2}, respectively, and this causes a decrease in I_s even though the dependency between V_1 and V_2 remains constant in terms of strength (*bits*). This decreased I_s results in decreased R_s as noise increases. Here we explore the extent of this noise-induced bias in R_s for the Gaussian sum case, and discuss implications for time-series data analysis.

The ratio of R_s with $\sigma_{n1} > 0$ to R_s for a noiseless case is as follows:

$$\frac{R_s(\sigma_n > 0)}{R_s(\sigma_n = 0)} = \frac{R_{\min} + \frac{I(X_{s1};X_{s2})}{H(V_1)}(R_{MMI} - R_{\min})}{R_{\min} + \frac{I(X_{s1};X_{s2})}{H(X_{s1})}(R_{MMI} - R_{\min})} \tag{3.19}$$

For cases in which $R_{\min} > 0$, this ratio depends on the values of $R_{\min}$, R_{MMI}, $H(X_{s1})$, $H(X_{s2})$, $H(V_1)$, $H(V_2)$ and ρ as it affects $I(X_{s1}; X_{s2})$. However, for cases when $R_{\min} = 0$ (Equation

(3.9)), Equation (3.19) reduces to a minimum value, which is a simple ratio of entropies:

$$\frac{R_s(\sigma_n > 0)}{R_s(\sigma_n = 0)} = \frac{H(X_{s1})}{H(V_1)}. \tag{3.20}$$

For a $\sigma_1 = \sigma_2 = 1$ case, this holds for $-0.86 < \rho < 0.77$, for which we observe the maximum difference between R_s with and without noise. For values of ρ outside this range, R_s becomes less biased due to the noise component (Figure 3.7). However, even when σ_n^2 is five times greater than σ_x^2 ($\sigma_n^2 = 5$ so that SNR=0.2), we find that R_s is never reduced by more than 40%.

The above result shows that the addition of source noise does bias our proposed R_s metric through an underestimation of R. However, R_{MMI} results in a larger overestimation of R when sources are not correlated ($\rho = 0$) (Figure 3.4a). Even when source correlation is high, such as for $\rho = 0.8$ (right panel of Figure 3.4a), R_{MMI} is twice the value of R_s.

While source noise in this example case leads to a bias due to I_s, the addition of target noise would also influence information partitioning. For the Gaussian sum, a component of target noise would lead to increased entropy $H(X_{tar})$, and the target uncertainty would no longer be completely reduced due to the knowledge of the sources. In this paper, we consider example cases where two sources completely inform the target, but noise in environmental systems could exist in all variables due to external or unobserved influences or measurement errors.

3.2.6 Behavior of R_s for an autoregressive process

While the previous cases involved simple addition with no time lag component, environmental process occur on multiple timescales and involve feedbacks between variables. For example, a species population may be driven by a logistic growth function in addition to external environmental perturbations. Here we explore the implications of our proposed R_s measure versus R_{MMI} for a dynamical system. This illustrates the effectiveness of our measure in detecting more complex interactions within a dynamical system with time dependencies, as opposed to cases of simple addition of signals as discussed earlier.

We replicate an example provided by [Barrett, 2015] of a multivariate autoregressive (MVAR) process in which a target variable X_{tar} receives information from both its own history and a source X_{s1} as follows:

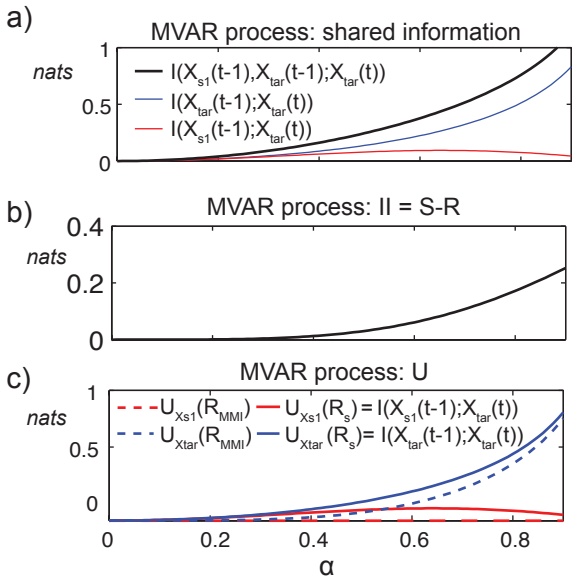

Figure 3.8: Information partitioning for a multivariate autoregressive (MVAR) process (Equations (3.21)-(3.22)). *(a)* $X_{tar}(t-1)$ provides more information to $X_{tar}(t)$ than $X_{s1}(t-1)$. *(b)* $II = S - R > 0$ for all α values, indicating net synergy in the system and that $R_{min} = 0$. *(c)* For information partitioning based on R_{MMI}, no unique information is detected from the weaker source (dotted red line), while our method based on R_s (solid lines) results in U equal to the individual shared information quantities from both sources since the sources are independent and $R_{min} = 0$.

$$X_{tar}(t) = \alpha X_{tar}(t-1) + \alpha X_{s1}(t-1) + N[0,1] \tag{3.21}$$

$$X_{s1}(t) = N[0,1] \tag{3.22}$$

where α is a coefficient that may range between 0 and 1. Here we note that the sources $X_{s1}(t-1)$ and $X_{tar}(t-1)$ are not correlated, and that although X_{tar} is a stronger source of information relative to X_{s1} due to the self-feedback, $X_{tar}(t)$ cannot be fully determined based on its own history. [Barrett, 2015] derives the information quantities of $I(X_{tar}(t-1); X_{tar}(t))$, $I(X_{s1}(t-1); X_{tar}(t))$, and $I(X_{tar}(t-1), X_{s1}(t-1); X_{tar}(t))$ from the covariance matrix (Equations (A.7)-(A.9)). We note that the equations listed in the appendix and resulting information partitioning components only depend on the parameter α.

We see that $X_{s1}(t-1)$ is the weaker source of information (Figure 3.8a) as expected and that the interaction information is positive, indicating that $S > R$ (Figure 3.8b). Accordingly,

a partitioning based on R_{MMI} results in a large S, a small R that increases with α, and no unique contribution from $X_{s1}(t-1)$ since it is the weaker source (Figure 3.8c). Meanwhile, a partitioning based on R_s better reflects the unique contribution of $X_{s1}(t-1)$. $R_s = 0$ for all α values, reflecting that the two sources are independent and $S - R > 0$, resulting in unique components that are equal to individual mutual information quantities (Figure 3.8c). This MVAR example is illustrative of a range of situations occurring in nature in which a variable is driven by its own history in addition to an external influence. For example, a simple model can be devised in which native mussel abundance is predicted as a logistic function of species population, in addition to fluctuations in streamflow, sediment concentration, and phytoplankton [Hansen et al., 2016, Foufoula-Georgiou et al., 2015]. In a scenario such as this, the Transfer Entropy T_E captures the combination of unique and synergistic information provided by a given source along with the history of mussel population. Although the external influence or perturbation may be a weaker influence, it is still valuable to distinguish its role as either a unique driver or an influence that is correlated with logistic population growth.

This section provides a context for information partitioning as a method to characterize the nature of interactions between signals in multivariate systems. While the previously introduced R_{MMI} measure overestimates redundancy at the expense of uniqueness, our proposed R_s measure effectively captures these unique influences.

3.3 Information partitioning applied to time-series data

We next move toward a practical application of information partitioning to the analysis of time-series signals. We explore the previously discussed noise-induced bias, in addition to other issues that arise when information measures are applied to observed or synthetic time-series datasets rather than analytical cases. These issues include data pre-processing and filtering, estimating *pdf*s from data, accounting for effects of zero-values in mixed distributions, and determining the statistical significance of detected information measures. Since environmental time-series datasets from different sources such as field observations or model simulations vary widely in length, time interval, and distribution, these issues are important to consider in combination. The goal of this section is to establish "best practices" for *pdf* estimation and move from a theoretical interpretation of S, R, and U to a practical application to datasets that involve noisy and non-linear interactions. This analysis will enable us to perform meaningful information partitioning to environmental time-series data.

First we discuss *pdf* estimation techniques. We then introduce a logistic sum case that

illustrates the influence of noise, non-linear dependencies, and aspects related to *pdf* estimation on information partitioning.

3.3.1 *Pdf* estimation from data

To compute information partitioning measures from time-series data, we must estimate 1d, 2d and 3d *pdf*s. Here we briefly discuss our approaches to estimate these *pdf*s from sparse data, test for statistical significance of shared information, and deal with issues common in ecohydrologic data such as the dominance of the diurnal cycle and mixed distributions. We refer the reader to the Appendix for more detailed information on these approaches, and examples based on solar radiation (Rg) and air temperature (Ta) recorded from a weather station.

We use Kernel Density Estimation (KDE) [Silverman, 1986, Lee et al., 2012] in order to compute robust *pdf*s given as few as $n_{steps} = 200$ data points [Goodwell and Kumar, 2015]. While KDE requires more computations than fixed-binning (FB) *pdf* estimation, it is advantageous for high-dimensional *pdf*s based on sparse data since it becomes independent of bin size above a certain number of bins. Regardless of the *pdf* estimation method used, any detected value of information $I(X_s; X_{tar})$ must be associated with some level of statistical significance. We employ a shuffled surrogates method [Ruddell and Kumar, 2009a, Goodwell and Kumar, 2015] in which the source X_s is shuffled to destroy time correlations and shared information $I(X_{shuff}; X_{tar})$ is computed. If shuffled iterations result in $I(X_{shuff}; X_{tar}) < I(X_s; X_{tar})$ at a 99% or greater confidence level, the value $I(X_s; X_{tar})$ is defined as statistically significant.

In addition to using robust *pdf* estimation methods and statistical significance testing to accurately detect information measures, we address specific characteristics of environmental time-series data that may inhibit detection of certain dependencies. Namely, the diurnal cycle is a dominant trend for many environmental variables, and the resulting synchronization is often much stronger than other types of interactions that occur on faster timescales. For example, Ta is tightly lag-synchronized with Rg as Earth's surface gains and loses heat over the course of a day. For variables such as these, we filter out the diurnal cycle to enhance the detection of other interactions that may be weaker or occur at different timescales. Another characteristic of many environmental signals such as rainfall, ephemeral streamflow, or shortwave solar radiation is that they have mixed distributions in which many values are zero. This alters detection of shared information, particularly when the KDE method is

used since it will smooth zero values over a continuous range. [Gong et al., 2014] discusses a solution to this for a 1d case which involves considering a *pdf* as a mixed distribution. We expand this to 2d and 3d *pdf* estimations. The Appendix and Figure A.1 contain a more detailed description of these issues of filtering diurnal cycles and accounting for mixed distributions, in addition to the handling of outliers in the data.

3.3.2 A logistic sum case

We next explore how non-linear relationships and noise influence detections of shared information and subsequent S, R and U components based on data. In the Gaussian sum example, we found that added noise causes a bias in our R_s metric. Here, we expect to detect a similar bias based on generated datasets with different levels of noise. To address this influence of noise on information measures as detected from data, we generate source variables X_{s1} and X_{s2} that are uniformly distributed over $[0, 1]$. We generate $n_{steps} = 1000$ data points and introduce a correlation between X_{s1} and X_{s2} by replacing values of $X_{s2}(t)$ with $X_{s1}(t)$ with probability p (i.e. $p(X_{s1}(t) = X_{s2}(t)) = p$). This results in approximately pn_{steps} identical data points and $(1-p)n_{steps}$ uncorrelated points. The target variable X_{tar} is a chaotic logistic function of X_{s1} and X_{s2} as follows:

$$X_{tar}(t) = \frac{1}{2}f(X_{s1}(t)) + \frac{1}{2}f(X_{s2}(t)) \tag{3.23}$$

where $f(X(t)) = 4X(t)(1 - X(t))$ so that each term in Equation (3.23) is a chaotic map. We introduce a noise component to X_{s1} and X_{s2} to represent source noise that does not participate in forcing of the target. As in the Gaussian sum case, we assume that the "measured" source variables are V_1 and V_2 as follows:

$$V_1(t) = X_{s1}(t) + N[0, \sigma_n^2] \tag{3.24}$$

$$V_2(t) = X_{s2}(t) + N[0, \sigma_n^2] \tag{3.25}$$

where σ_n^2 represents the level of noise present in each measured source variable, and relationships between X_{s1}, X_{s2}, and X_{tar} do not change with noise. For V_1 and V_2 in Equations (3.24)-(3.25), the SNR is a function of the noise variance σ_n^2 as follows:

$$SNR = \frac{\sigma_x^2}{\sigma_n^2} = \frac{1}{12\sigma_n^2} \tag{3.26}$$

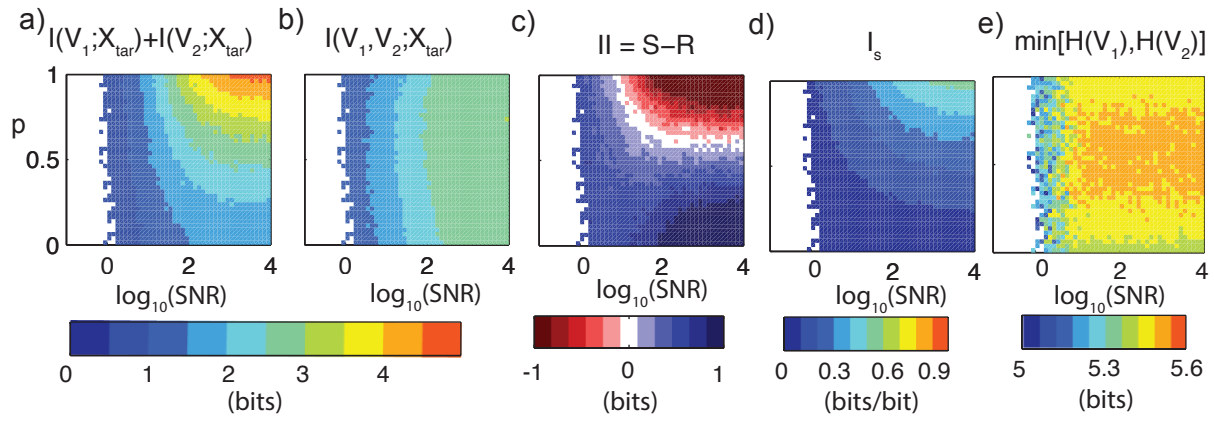

Figure 3.9: Information partitioning measures for chaotic logistic datasets with a range of correlations (p) and noise levels (indicated as $\log_{10} SNR$) show that all detected measures are influenced by noise. *(a)* Sum of source→target mutual information. *(b)* Total shared information $I(X_{s1}, X_{s2}; X_{tar})$. *(c)* $S - R$, equivalent to quantities in *(b)* subtracted from *(a)*. *(d)* I_s measure used to gage source dependency for computation of R_s. *(e)* Denominator of I_s term is minimum of source entropies $H(V_1)$ and $H(V_2)$.

since $\sigma^2_{x_{s1}} = \frac{1}{12}$ for the uniform random variable X_{s1}. In the Gaussian sum case, uncorrelated source noise only influences the entropies of the measured sources V_1 and V_2. However, our shared information detections will be influenced due to our inability to distinguish between noise and forcing components within V_1 and V_2. Specifically, we find that for pdfs estimated from time-series data, the noise (σ^2_n) results in decreased estimates of shared information quantities (Figure 3.9a,b) because the noise causes the links to be more weakly detected. At approximately $\log_{10}(SNR) = 0$ ($\sigma^2_n = \sigma^2_x$), shared information ceases to be detected as statistically significant. While individual shared information quantities $I(V_1; X_{tar})$ and $I(V_2; X_{tar})$ show dependency on both SNR and p, total information $I(V_1, V_2; X_{tar})$ is only weakly dependent on p. These features combine to influence $II = S - R$ (Equation 3.6) as noise and source dependencies shift (Figure 3.9c). While this dependency of $S - R$ on noise does not exist in the theoretical Gaussian case, it significantly influences the partitioning of information for observed data.

We find that I_s depends on both the SNR and source dependency p (Figure 3.9d). While the Gaussian example would predict that I_s should decrease with p due to an increase in its denominator ($\min[H(V_1), H(V_2)]$), we find that this term actually decreases slightly for high noise levels (Figure 3.9e). This indicates that I_s decreases with added noise (decreasing SNR) because of a decrease in the detected $I(V_1; V_2)$ rather than an increase in source entropy.

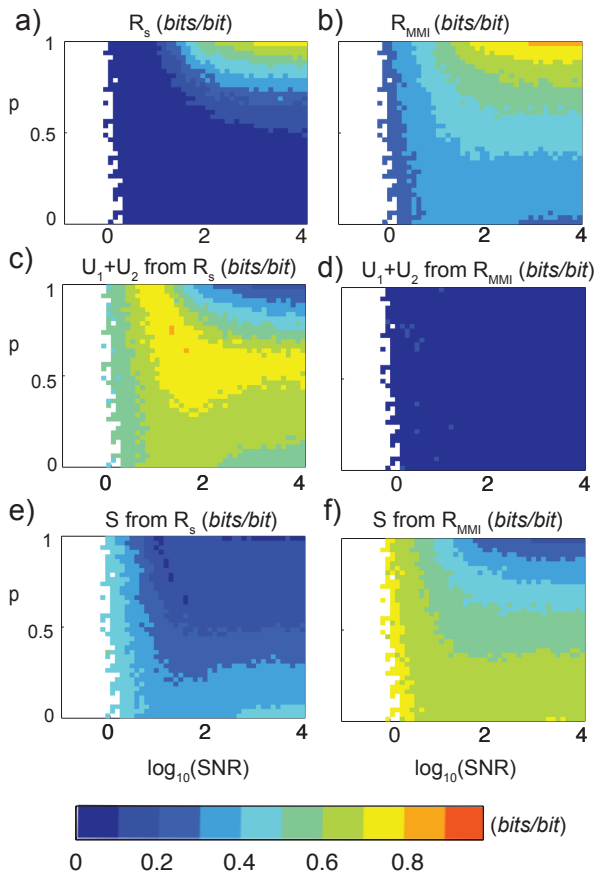

Figure 3.10: Information partitioning components for chaotic logistic datasets based on R_s and R_{MMI} with a range of source correlations (p) and noise levels ($\log_{10} SNR$). *(a)* R computed based on our R_s metric and *(b)* R_{MMI}. *(c)* The sum of unique components ($U_1 = U_2$ since sources are of equal influence to the target) based on partitioning with R_s and *(d)* R_{MMI}. *(e)* S computed based on partitioning with R_s and *(f)* with R_{MMI}.

In other words, the dependency of I_s on noise is more related to *pdf* estimation than the bias discussed in the Gaussian sum example. Aside from this noise influence, I_s decreases as expected as source dependency (p) decreases. We also performed this analysis for linear data ($X_{tar}(t) = X_{s1}(t) + X_{s2}(t)$, where X_{s1} and X_{s2} are uniformly distributed) and obtained similar patterns in terms of detected I_s and $S - R$.

Since shared information measures directly computed from the Shannon Entropy measures shift with noise and source dependency in contrast to a Gaussian case, any information partitioning will also be influenced. From a comparison of partitioning based on R_s and R_{MMI} measures, we observe non-linear dependencies of S, R, and U on p and SNR for both partitioning methods (Figure 3.10) As expected, R_{MMI} results in $U_1, U_2 = 0$ regardless of

source dependency since the sources provide the same quantity of information (Figure 3.10e). For U based on R_s at the highest SNR, we see that the proportion of U increases slightly as p goes from 0 to about 0.5, then decreases to nearly 0 as p goes to 1, when sources $X_{s1} = X_{s2}$ become completely redundant. The maximum $U_1 + U_2$ proportion for a non-zero source dependency corresponds to the area where $S - R = 0$ (Figure 3.9c). For increasing noise (decreasing $\log_{10} SNR$), we find that all information components reach a nearly constant proportion and no longer depend on p.

Since I_s results in an underestimation of R, situations where $S = R$ will lead to the highest U. In other words, when R is underestimated and $S = R$, S will be similarly underestimated so that all remaining information contained in $I(X_{s1}; X_{s2}, X_{tar})$ is defined as unique from either source. The source dependency $I_s = 1$ ($R_s = R_{MMI}$) only when the estimated 3d *pdf* shows complete dependency between the two source variables. This does not occur even when the sources are copies of each other ($X_{s1} = X_{s2}$) due to *pdf* estimation methods. A potential improvement for this measure would be an alternate empirical scaling of $I(X_{s1}; X_{s2})$ to define I_s. This would involve computing a 3d *pdf* in which the larger source (e.g. X_{s1}) replaces the smaller source (e.g. $X_{s1} = X_{s2}$ so that the *pdf* is $p(x_{s1}, x_{s1}, x_{tar})$). We would compute $I_{s,max}$ based on this "maximally redundant" *pdf*, and set $I_s = \frac{I(X_{s1};X_{s2})}{I_{s,max}}$. This would not resolve noise-induced bias as presented in the theoretical Gaussian sum example, but would increase R_s and the corresponding detected S. In general however, we find from this analysis that the R_s measure as currently defined detects all four information components (U_1, U_2, R, S), while R_{MMI} cannot be used to capture U.

This analysis establishes the feasibility of applying information partitioning to time-series datasets. For generated chaotic datasets, we detect statistically significant shared information for $SNR > 1$, although the strength of information measures and resulting partitioning are influenced at lower noise levels, in the range of SNR = 10 to 100. Although many interactions between environmental signals are relatively weak, we apply filtering, outlier-removal, and consideration of mixed distributions in order to enhance their detection and subsequent partitioning into unique, synergistic, and redundant components.

3.4 Environmental signals: unique influences to relative humidity

In this section, we apply our information partitioning approach based on the estimation techniques described previously to analyze an illustrative weather station dataset. Here we consider τ-lagged air temperature (Ta) and wind speed (WS) as potential sources of

information to relative humidity (RH) as measured from a 1-minute weather station dataset (weather station details are provided in the Appendix). As in the autogressive example, we know that the two sources are not completely redundant with each other, since WS and Ta influence RH via differing mechanisms. Specifically, a change in Ta changes the capacity of air to hold water vapor, and a change in WS influences the rate of evapotranspiration. From information partitioning with R_s, we expect to detect unique information from both sources. We segment three days (DOY 170-172 of 2015) into multiple time windows and analyze how S, R, and U based on both the R_{MMI} and R_s metrics vary between windows.

For this analysis, we segment data into four contiguous time windows per day based on the sunrise and sunset time as measured from shortwave radiation Rg, so that there are two day-time windows and two night-time windows for each 24 hour period. We remove outliers, filter the diurnal cycle from Ta and RH as discussed previously, and normalize each variable (Figure 3.11) to a $[0, 1]$ range based on minimum and maximum values. We use $N_{bins} = 75$ and the KDE method to compute all pdfs.

Generally for all windows, we observe from visual inspection of the time-series that WS varies on a fast timescale compared to fluctuations in RH and Ta, indicating its larger variability and the prevalence of fluctuations on a 1-minute timescale (Figure 3.11a-f). We also see that RH and Ta tend to be inversely correlated, due to the influence of air temperature on the atmospheric water vapor holding capacity. During certain windows, we observe some correlation between WS and RH, but this does not occur at all times, and the relationship appears to switch between a positive and negative correlation. Additionally, some windows exhibit more variability in RH than others. In particular, we find that windows in the afternoon and evening of DOY 171 (Figure 3.11d,e) exhibit more variable RH, while RH is more stable in late DOY 170 and early DOY 172 (Figure 3.11c,f). While higher target (RH) uncertainty implies the potential for more shared information from source variables, we will see that this is not necessarily the case for environmental signals due to external or unobserved influences. For example, RH could be highly variable during a certain window due to variability in other drivers such as wind direction, rainfall, or evapotranspiration.

We consider the maximum statistically significant $I(Ta(t-\tau); RH(t))$ and $I(WS(t-\tau); RH(t))$ over time delays between $\tau = 1 - 10$ minutes as the dominant strengths of the interactions. This range of τ is chosen to capture interactions occurring at the fastest timescales that can be measured from 1-minute resolution data. For each time window, the total information $I(Ta(t-\tau), WS(t-\tau); RH(t))$ is partitioned into unique information from Ta and WS, U_{Ta} and U_{WS}, respectively, and R and S components (Figure 3.11g,h). For

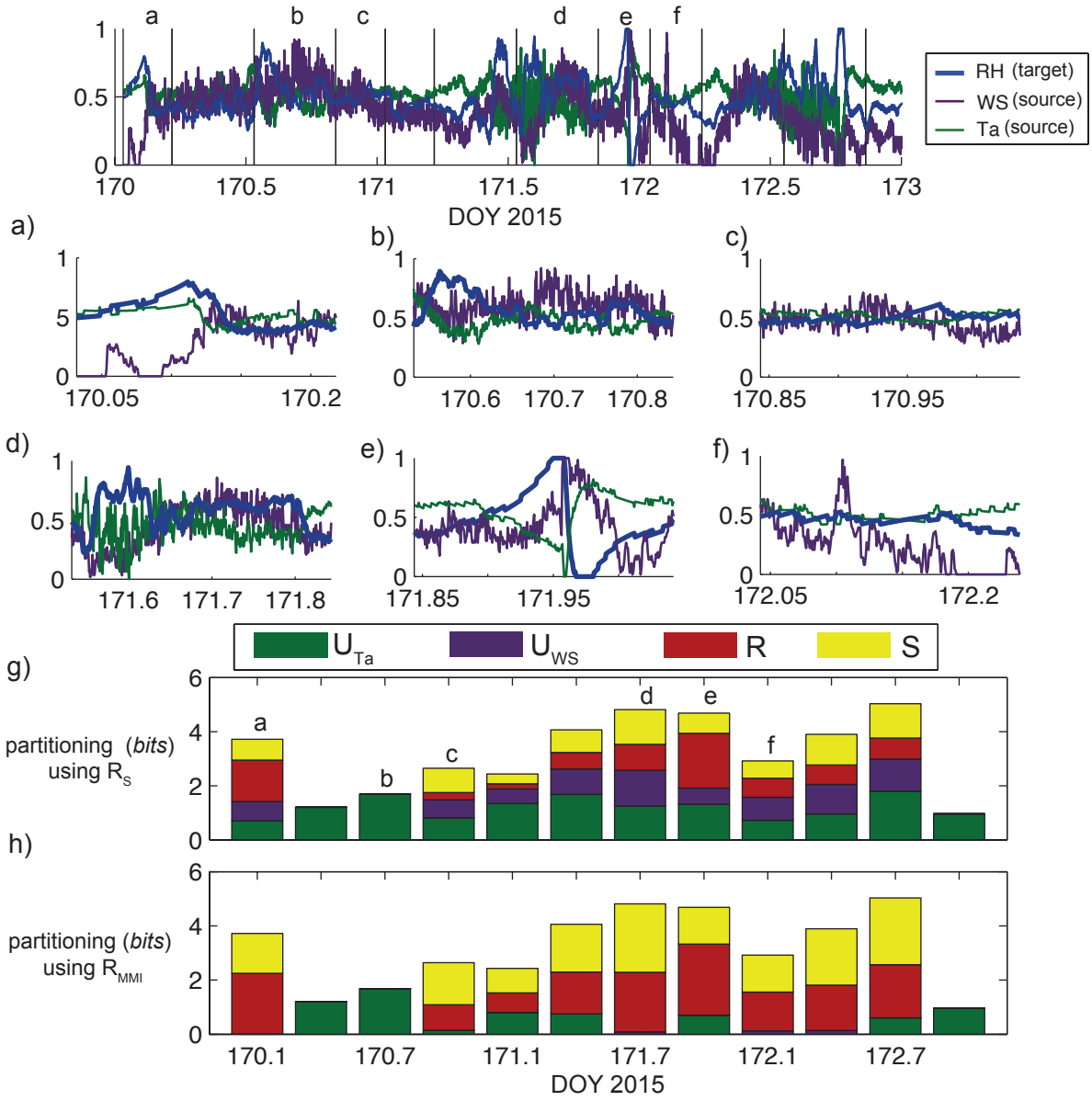

Figure 3.11: Example of information partitioning applied to weather station data. Time series data are segmented into several-hour windows over a three-day period (*top*), based on sunrise and sunset times, and information partitioning components for each window. *(a-f)* Several selected time-series windows exhibiting varying relationships between relative humidity (RH), windspeed (WS), and air temperature (Ta). *(g)* Information partitioning is performed for each window using $R = R_s$, and *(h)* $R = R_{MMI}$. In *(g)* and *(h)*, bar heights represent total information shared between WS and Ta at some time lag τ between 1 and 10 minutes to RH, and windows *(a-f)* are labeled.

several time windows (such as Figure 3.11d), no shared information is detected between *WS* and *RH*, so the only information component is U_{Ta}. Otherwise, as discussed in the several example cases presented in this paper, R_{MMI} does not allow U to be detected from the weaker source, which in this case is typically *WS*. From information partitioning with R_s, we find that the two sources are indeed not so redundant as to exclude a U_{WS} component in any of the time windows which involve a link between *WS* and *RH*.

From this set of 12 time windows in which various strengths of total information and proportions of U_{WS}, U_{Ta}, R, and S are detected, we describe several different categories of relationships as follows:

1. Low information content due to absence of *WS* as a source: For some time windows, *WS* is not detected as a statistically significant source, and no information partitioning is possible since all shared information is attributed to *Ta*. For these time windows, (such as *b* in Figure 3.11g) we see from the time-series that while *RH* and *Ta* appear tightly inversely linked, *WS* is varying independently (Figure 3.11b). This detected behavior indicates that although *WS* is highly variable, it does not influence humidity levels strongly enough for a dependency to be detected.

2. Low information content with all information components: For some time windows, both *WS* and *Ta* are detected as sources of information to *RH*, but the total information content is relatively low (*c* and *f* in Figure 3.11g). These time windows typically have lower entropy $H(RH)$, thus a lower potential for shared information. However, even for the window in Figure 3.11c when all variables are relatively stable, some shared information is still detected. For these periods, the relative contributions of U, R, and S information are more likely to be influenced by noise due to the weak detection of information, but still show that the time dependency consists of multiple types of information.

3. High information content, dominance of R: Time windows such as *a* and *e* in Figure 3.11g exhibit high total shared information from the two sources, but the largest information component is R. This indicates that both *WS* and *Ta* provide information regarding the future of *RH*, but the two sources are themselves dependent. From inspection of the time-series in Figures 3.11a and e, we see that *WS* and *Ta* appear somewhat negatively correlated, which leads to the detection of R. These are both night-time windows during which *WS* varies more smoothly than during the day. Here,

it is likely that changes in WS are induced by temperature changes, so that Ta influences both RH and WS.

4. High information content, dominance of U and S: There are other time windows for which information content is high, and contributions of U and S are larger than R. For example, in time window d in Figure 3.11g, unique components are nearly equal, and U_{Ta}, U_{WS}, and S are all higher than R. This indicates that both WS and Ta influence RH but via different mechanisms. From inspection of Figure 3.11d, we see that all variables are fluctuating on a fast timescale compared to other windows. For windows such as these, it is difficult to visualize the ways in which the source variables influence the target, but the detected information partitioning shows they are synergistic. Ta is linked to RH via the water vapor holding capacity, and WS is simultaneously causing changes in RH through evaporation and introducing mixing of air with different states of Ta and RH.

This analysis illustrates the potential for information partitioning to reveal aspects of interactions within natural systems that could not be detected based on traditional measures. Additionally, the ability to segment long time-series data into smaller windows is valuable to detect how environmental dynamics may evolve over time or shift due to perturbations or stresses.

3.5 Discussion

In complex systems where multiple variables interact non-linearly and at varying timescales and strengths, information partitioning can elucidate the nature of time dependencies beyond the understanding gained from traditional information measures such as Transfer Entropy or interaction information. In this study, we build the case for an information partitioning framework through several examples where two sources drive a target via different mechanisms and with different source correlations and noise levels. While these examples are relatively simple compared to natural systems with many non-linearly interacting components and external influences, they illustrate the depth of understanding concerning system behavior that can be gained from information partitioning.

Even for straightforward operations such as addition, deterministic non-linear relationships, and multivariate autoregressive processes, the estimation of synergistic, unique, and redundant influences enable us to separate driving interactions from feedback and quantify weak or

strong influences. An appropriate partitioning of information shared by two sources to a target is important in order to not mistakenly classify two potential sources as redundant when they actually provide unique information. Although from a purely information theoretic perspective, information shared between one source and a target may "look the same" as information shared between another source and the same target, we wish to distinguish between individual and joint influences due to physical mechanisms that can be detected based on observed or simulated time-series data. Therefore, we introduce a measure that retains the "uniqueness" of shared information to the extent that sources are not synchronized or otherwise correlated.

This is particularly relevant when attributing shared information to physical processes occurring between environmental signals. For example, if we detect that wind speed and air temperature provide information to relative humidity, we expect that these two sources provide some combination of unique, synergistic, and redundant information that depends on the forcing and feedback relationships between all three variables. In our information partitioning framework, a completely redundant relationship indicates that one source is a function solely of the other source. In this context, we would only expect maximal redundancy if wind speed was a function of only temperature and temperature was driving both wind speed and relative humidity. However, in a natural system, more complex relationships exist between the three variables and external influences such as solar radiation and precipitation.

When information measures are applied to time-series data, it is important to consider several issues that influence the detection of shared information and subsequent partitioning. In environmental time-series datasets, zero-values, sparse data, dominant diurnal and seasonal cycles become part of an estimated *pdf*, in addition to the *pdf* estimation method. The formulation and examples presented here provide significant insight into information partitioning and can be applied to continuous or discrete cases, and analytically derived or empirically estimated *pdf*s.

An additional aspect of ecohydrologic behavior is the capacity for many components to simultaneously influence overall system dynamics. In the example where wind speed and temperature are found to provide varying types and magnitudes of information to relative humidity during different time periods, the consideration of other variables such as wind direction [Sendrowski and Passalacqua, 2017] or moisture conditions could explain some of these shifting properties. In a companion paper [Goodwell and Kumar, 2017b], we develop a network-based approach that uses information partitioning to reveal the nature of time dependencies and ecohydrologic behaviors over the course of a growing season.

In a network of more than two interacting components, the detection of synergistic, unique, and redundant information from subsets of sources to targets can reveal important behaviors that may not be observable otherwise. The framework and methodologies presented here have broad implications to reveal the nature of complex relationships in ecohydrologic or other types of natural systems. The identification of forcing and feedback relationships, their shifts over time in terms of strength and composition, and their timescales can enhance process understanding and modeling efforts.

CHAPTER 4

TEMPORAL INFORMATION PARTITIONING NETWORKS (TIPNETS): A PROCESS NETWORK APPROACH TO INFER ECOHYDROLOGIC SHIFTS

4.1 Introduction [1]

Ecohydrologic systems can be conceptualized as complex networks in which interactions occur between and within the atmospheric, root-soil, and canopy systems. These interactions result in net exchanges of information, energy, mass and momentum within subsystems and with the surrounding environment [Jenerette et al., 2012, Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b, Kumar and Ruddell, 2010]. Some interacting components may be non-separable, in that processes occur synchronously [Sugihara et al., 2012] or near the time scale of observations. Additionally, couplings can involve multiple driving forces, feedbacks, threshold behaviors, and non-linearity, and may break down or shift in association with permanent or short-term changes in overall ecosystem behavior and health. In ecohydrologic systems, time scales can range from the order of seconds, such as for the case of stomatal response to water availability, to years, such as for the case of co-evolution of landscapes and ecosystems.

Characterization of these dependencies in ecohydrologic systems can quantify strengths of known process interactions, detect previously unknown influences, and also help to reveal emergent properties of ecosystems. Emergent properties are novel behaviors that arise from interactions between individual components, such as the temporal synchronization of chirping crickets or blinking fireflies [Strogatz, 2001]. In these situations, knowledge of an individual entity cannot explain the resulting large-scale effect. This aligns with the notion that an "ecosystem is more than the sum of its parts" [Jørgensen et al., 1992] and motivates us to approach analysis of systems holistically, or in terms of relationships as opposed to individual behaviors [Kumar, 2007].

A process network is defined as a network of time dependent interactions that depict the magnitude and direction of the flow of matter, energy, or information between compo-

[1] This chapter is in review for Water Resources Research [Goodwell and Kumar, 2017b]

nents [Ruddell and Kumar, 2009a, Sendrowski and Passalacqua, 2017]. Information theory measures based on Shannon Entropy ($H(X)$) [Shannon, 1948] enable the detection of non-linear and weak dependencies that occur frequently in natural systems. Metrics such as entropy, mutual information, and Transfer Entropy (T_E) [Schreiber, 2000] have been used to address various topics in Earth science [Balasis et al., 2013, Gong et al., 2014], including the measurement of complexity [Balasis et al., 2009], identification of critical transitions [Saco et al., 2010, Ferri et al., 2012], and couplings between variables that relate to feedback or causal relationships [Ruddell and Kumar, 2009a, Kumar and Ruddell, 2010]. T_E, a special case of conditional mutual information, has been used to infer the nature of time dependencies in terms of forcing, feedback, and synchronization in various fields [Ruddell and Kumar, 2009a, Lee et al., 2012, Hlaváčková-Schindler et al., 2007]. However, T_E does not address issues that arise when multiple sources influence a single target [Runge, 2015].

In these cases, information provided from a source to a target may be redundant (R), synergistic (S) or unique (U). R is overlapping information that two or more sources share with a target *redundantly*, S is information obtained only when two sources are known together, or *synergistically*, and U is information that only a single source provides, *uniquely*. When two lagged sources are found to share information with a target, the relative proportions of S, R, and U components can provide a deeper understanding of the nature of the time dependency. The detection of R indicates that sources are partially dependent, which could occur due to a common driver or feedback between sources. Meanwhile, U indicates that a source provides information not provided by any other source. In the absence of unmeasured or missing variables within the system, U could be inferred as a potential forcing relationship. S implies that two sources jointly influence the target, which could either indicate that sources are independent or that there is a conditional relationship where the knowledge of both sources is needed to predict the target. Several measures of S, R, and U have been proposed and applied to reveal aspects of time dependencies within various types of systems [Williams and Beer, 2010, Griffith and Koch, 2014, Griffith and Ho, 2015, Barrett, 2015, Faes et al., 2015, Goodwell and Kumar, 2017a].

In this paper, we build on and substantially expand the process network approach for analysis of process connectivity in ecohydrologic systems. We develop an approach to construct Temporal Information Partitioning Networks (TIPNets) that characterize the nature of time dependencies between multiple signals using multivariate information metrics. In our TIPNet approach, we implement a new method to partition information into S, R, and U components

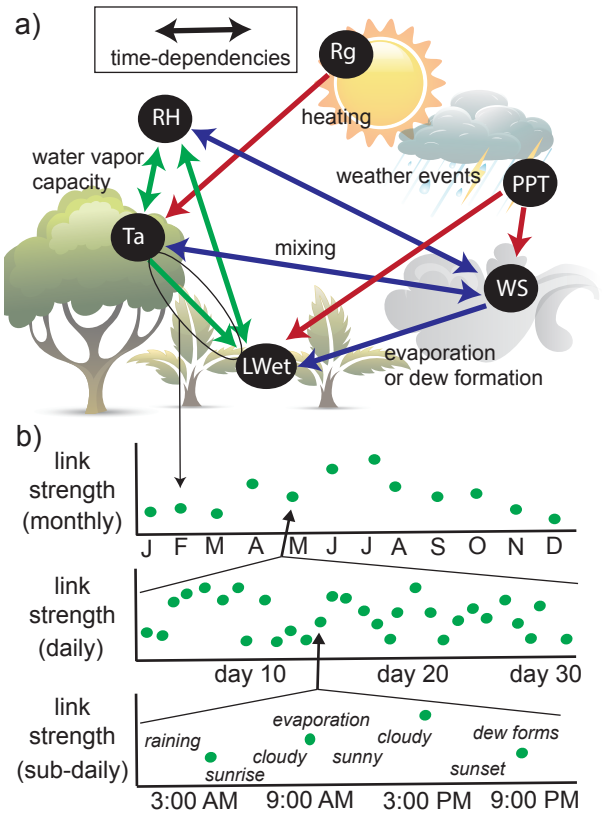

Figure 4.1: Illustration of complex network behavior in an ecohydrologic system. *(a)* Time dependent interactions occur between solar radiation (Rg), precipitation (PPT), leaf wetness ($LWet$) or moisture condition, wind speed (WS), relative humidity (RH), and air temperature (Ta). We characterize these dependencies as information transfers within a network that are associated with properties of time scale, strength, uniqueness, redundancy, and synergy. *(b)* Network properties that may be detected on a seasonal timescale (top) result from an accumulation of interactions that vary on much shorter timescales such as daily or sub-daily (bottom).

[Goodwell and Kumar, 2017a] and apply it in a network context to analyze the temporal evolution of system dynamics. We define links as statistically significant detections of shared information between τ-lagged source and target time-series variables. We further analyze each link in terms of its unique, synergistic, or redundant relationship with other links in the network to reveal the nature of dominant interactions. This approach constitutes a novel framework and toolbox for constructing process networks, and is distinctive from T_E based approaches which only allow for binary comparisons between time-series signals [Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b]. To capture the aspect of temporal

evolution of dependencies within the system, we consider process networks in the context of temporal networks [Holme and Saramäki, 2012], which recognize time as an additional dimension to capture network evolution, since links may not be continuously active.

We apply the TIPNet framework to study short-term variability of interacting environmental signals as measured from a weather station in Central Illinois. To identify short term shifts in network structure, we consider variables over several-hour time windows of 1-minute data that encompass individual weather events such as rainfall, dew formation, and evaporation. For these ecohydrologic process networks, interactions that vary on the order of hours could inform longer timescale behaviors such as response to accumulating drought conditions or changes in rainfall variability (Figure 4.1). These issues regarding the behavior of the network as a whole and its evolution over time include:

1. How do feedbacks and forcing shift under different moisture and radiation conditions? Do these shifts occur at event timescales or gradually over many days or a season?

2. Are there persistent network patterns associated with seasonality, weather conditions, or short-term moisture variability?

In an ecohydrologic system, we hypothesize that emergence, breakdowns, or changes in strength of interactions will reflect weather conditions and seasonal transitions. For example, a period of high rainfall during a growing season may result in higher connectivity as processes of evapotranspiration and nutrient fluxes participate in more feedbacks with atmospheric conditions. In contrast, a dry time period may relate to lower variability and fewer connections between variables. In terms of the nature of information transfers, we expect that variables such as air temperature and solar radiation tend provide redundant information, such that a driving or "causal" forcing can only be inferred, and variables such as wind speed should provide more synergistic and unique information since it is less likely to be lag-synchronized with any other variable.

This paper is organized as follows. In Section 4.2, we introduce the TIPNet framework as an approach to construct and analyze process networks based on time-series signals. In Section 4.3, we present our TIPNet analysis to a 1-minute dataset of environmental signals from a weather station in Central Illinois. We discuss typical network behaviors over a 180-day study period in terms of time scale, interaction strength and type, and variability with weather conditions, and analyze short-term shifts in interactions that occur over several-hour periods. In Section 4.4, we discuss implications of a network approach for the study of ecohydrologic interactions and broader applications.

4.2 TIPNets: Information in an evolving network context

Our Temporal Information Partitioning Network (TIPNet) framework involves aspects of information theory, information partitioning and moving time windows (Figure 4.2). To convey the advantages and nuances of this framework compared to previous approaches, we briefly introduce information metrics and their application in a network context. Next, we outline the TIPNet framework that uses information partitioning to construct networks in which each link is associated with dominant timescale and a strength that can be partitioned into unique, redundant, and synergistic components (Figure 4.2).

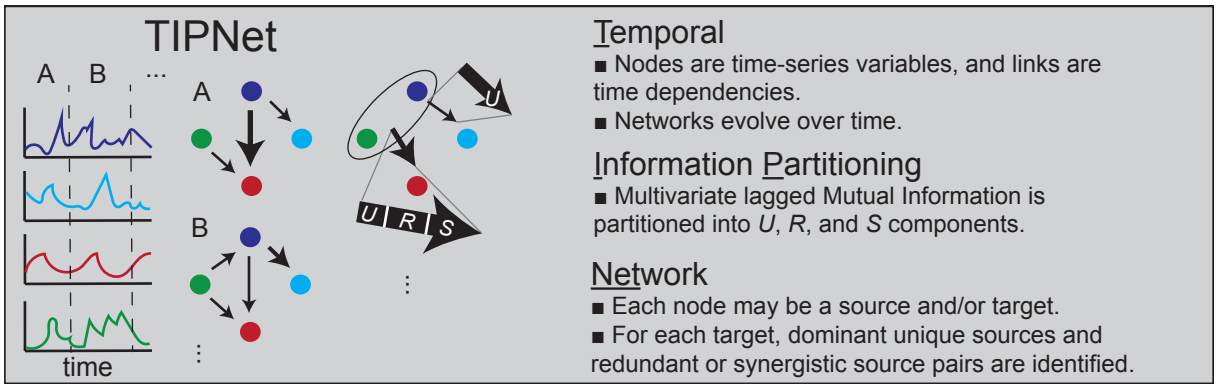

Figure 4.2: The TIPNet framework characterizes process networks based on time-series signals by identifying unique (U), synergistic (S), and redundant (R) transfers of information between interacting variables. *(left)* Time series signals are segmented into windows (e.g. A and B) and a network of pairwise links is constructed from lagged mutual information associated with a time scale and strength (arrow thickness). *(middle)* The joint information from each pair of sources is partitioned into U, R, and S. We see that for time window A, a pair of sources provides all types of information to one target node, and only one source provides U to another target node since it is the only detected source for that target.

4.2.1 Information Partitioning

For a system composed of a set of variables $Y = \{Y_1, Y_2, Y_3...\}$ for which we have measured time-series signals, we seek to characterize strengths and time scales of dependencies to reveal the nature of interactions within the system. Each variable is considered as a possible target node X_{tar} that receives information, and a lagged source node X_{s1} that provides information to a target. Source nodes are τ-lagged values, i.e. $X_{s1} = Y_1(t - \tau)$. For each time-series signal, the Shannon Entropy (*bits*), $H(X_{tar}) = -\sum p(x_{tar}) \log_2 p(x_{tar})$, quantifies its degree

of uncertainty. Lagged Mutual Information $I_\tau = I(X_{s1}; X_{tar})$ is the reduction in uncertainty of X_{tar} due to the knowledge of X_{s1}:

$$\begin{aligned} I(X_{s1}; X_{tar}) &= H(X_{tar}) - H(X_{tar}|X_{s1}) \\ &= \sum p(x_{s1}, x_{tar}) \log_2 \left(\frac{p(x_{s1}, x_{tar})}{p(x_{s1}) p(x_{tar})} \right). \end{aligned} \tag{4.1}$$

I_τ between X_{s1} and X_{tar} may be computed for any of a range of time lags $\tau = \tau_1, \tau_2 ... \tau_{n_\tau}$ between them. For each pair of source and target variables (X_a, X_{tar}), we define the dominant time scale of the link between them as the τ that maximizes $I(X_a; X_{tar})$ as follows:

$$\tau = \underset{\tau}{\text{argmax}}[I(X_a(t - \tau); X_{tar})]. \tag{4.2}$$

We use Kernel Density Estimation (KDE) to compute all *pdf*s [Lee et al., 2012, Silverman, 1986, Sharma and Mehrotra, 2014, Goodwell and Kumar, 2015] from data. Although this method can be used to detect dependencies given fewer than 50 data points, we use a minimum of $n_{steps} = 200$ data points per environmental signal to account for noise, weak links, and damping due to multiple influences. We use shuffled surrogates to test for statistical significance of each $I(X_{s1}; X_{tar})$ [Goodwell and Kumar, 2015, Ruddell and Kumar, 2009a], where X_{s1} is randomly shuffled to destroy its time correlations. We refer the reader to a companion paper for a detailed description of these techniques in the context of ecohydrologic time-series analyses [Goodwell and Kumar, 2017a], and for brevity do not repeat the details here.

Information partitioning [Williams and Beer, 2010] is concerned with categorizing shared information from multiple sources to a target as unique (U), redundant (R), or synergistic (S). When two sources provide information to a target, R is overlapping information that both sources provide, U is information that a source provides only individually, and S is information that the sources provide only when both are known together. The total shared information from X_{s1} and X_{s2} to a target X_{tar} and the individual mutual information quantities are combinations of U, R, and S components as follows [Williams and Beer, 2010]:

$$I(X_{s1}, X_{s2}; X_{tar}) = U_1(X_{s1}; X_{tar}) + U_2(X_{s2}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}) + S(X_{s1}, X_{s2}; X_{tar}) \quad (4.3)$$

$$I(X_{s1}; X_{tar}) = U_1(X_{s1}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}) \quad (4.4)$$

$$I(X_{s2}; X_{tar}) = U_2(X_{s2}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}) \quad (4.5)$$

where U_1, U_2, R, and S are non-negative quantities (Figure 4.3).

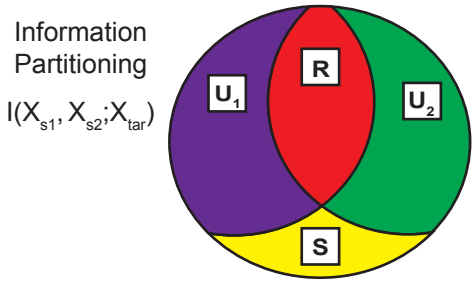

Figure 4.3: Information partitioning diagram showing partitioning of total information into four components. Circle area represents total shared information $I(X_{s1}, X_{s2}; X_{tar})$ that is partitioned into four non-negative components of $U_1 = U_1(X_{s1}; X_{tar})$, $U_2 = U_2(X_{s2}; X_{tar})$, $R = R(X_{s1}, X_{s2}; X_{tar})$, and $S = S(X_{s1}, X_{s2}; X_{tar})$.

From Equations (4.3)-(4.5), we see that an estimation of any of U_1, U_2, R, or S enables the computation of all components, since the mutual information values are directly computable from *pdf*s estimated from data. A number of methods have been proposed for direct estimation of R to enable this [Griffith and Ho, 2015]. These existing approaches consider that redundancy is the difference in strength between the stronger and weaker source. In other words, information provided by two sources to a target is redundant if both sources provide the same *quantity* of information, regardless of the level of correlation between sources. In contrast, we consider that two sources may provide the same quantity of information to a target uniquely if the sources are independent of each other. To achieve this, we implement a Rescaled Redundancy measure, R_s, as follows [Goodwell and Kumar, 2017a]:

$$R_s = R_{\min} + I_s(R_{MMI} - R_{\min}) \quad (4.6)$$

where $I_s = \frac{I(X_{s1};X_{s2})}{\min[H(X_{s1}),H(X_{s2})]}$ is the normalized source dependency, and R_{MMI} and R_{min} are upper and lower bounds of redundancy, respectively, as follows:

$$R_{MMI} = \min[I(X_{s1}; X_{tar}), I(X_{s2}; X_{tar})] \tag{4.7}$$

$$R_{\min} = \max(0, -II) \tag{4.8}$$

where $II \equiv I(X_{s1}; X_{s2}; X_{tar})$ is the interaction information, and equivalent to $S - R$. R_{MMI} is a previously proposed R metric that is equivalent to the "minimum mutual information" that either source shares with the target [Williams and Beer, 2010, Barrett, 2015], and is an upper bound for any measure of R. R_{min} defines a lower bound for R that ensures that all information decomposition components are non-negative. In Equation (4.6), $R_s = R_{min}$ when sources are completely independent ($I_s = 0$) and $R_s = R_{MMI}$ when sources are redundant or completely dependent ($I_s = 1$). For a detailed justification of the use of R_s in Equation (4.6) as an appropriate redundancy measure for environmental signals, please refer to [Goodwell and Kumar, 2017a]. In general, R_s enables the detection of unique information from multiple sources. Henceforth, R will be used to represent redundancy as defined by R_s, i.e. $R \equiv R_s$. While information measures such as Transfer Entropy (T_E) have been used previously in a network context, we introduce a more comprehensive framework for process network construction that utilizes information partitioning to detect and interpret interactions between multiple sources and targets. This framework goes beyond pairwise identification of source and target variables and approaches a more holistic definition of system behavior. Although this study is limited to information transfers between one or two sources to a target node, the same framework could be applied to analyze shared information between many sources to a target simultaneously if data availability supported the construction of higher dimensional *pdf*s.

4.2.2 Information within a network

Information partitioning enables us to define shared information in terms of U, R, or S components. However, in the context of a system with multiple drivers and feedbacks, a target node may receive information from more than two sources, and a source may be synergistic or redundant with any other source. Moreover, a source may provide R with one source and S with another. In a system with multiple interactions, we must apply information partitioning to each pair of sources to determine the most influential unique, redundant, and synergistic contributors to a given target. For example, given time-series of relative humidity

(RH), shortwave radiation (Rg), wind speed (WS), and temperature (Ta), we may find that τ-lagged Ta, Rg, and WS all share information with RH. Lagged I and T_E may be similarly high for all couplings ($Rg \rightarrow RH$, $WS \rightarrow RH$, $Ta \rightarrow RH$), but it would be deceptive to interpret all three links as equally "causal" or "influential" to RH. In this example case, we may find that the source pair $\{Ta,\ WS\}$ provides the most S to RH out of all source pairs, while the pair $\{Rg,\ Ta\}$ provides the most R. To address such situations, we first apply Equation (4.6) in combination with Equations (4.3)-(4.5) to compute U, R, and S components for each source pair to a given target. Then, for each detected source→target link in the network, we define information components as follows:

$$R(X_{s1}, X_{tar}) = \max_{X_{s2}} [R_s(X_{s1}, X_{s2}; X_{tar})] \tag{4.9}$$

$$U(X_{s1}, X_{tar}) = I(X_{s1}; X_{tar}) - R(X_{s1}, X_{tar}) \tag{4.10}$$

$$S(X_{s1}, X_{tar}) = \max_{X_{s2}} [S(X_{s1}, X_{s2}; X_{tar})] \tag{4.11}$$

This requires a pairwise evaluation of sources for each target node in a network. We choose the maximum R_s over all possible sources since we are interested in determining whether any two source nodes are redundant. If a source X_{s1} is not completely redundant with any other source, the remaining information in $I(X_{s1}; X_{tar})$ is defined as U according to Equation (4.4). Similarly, the maximum S over all possible sources identifies the source pair that most synergistically informs the target node. This approach assumes that for a given target node, there is no redundant (synergistic) information beyond that provided by the largest pair of redundant (synergistic) sources. While a higher dimensional approach could detect additional redundancy and synergy, we limit our analysis to the two source case. With this formulation, a source node X_{s1} may be redundant with one source node and synergistic with another. The node X_{s2} that maximizes either Equation (4.9) or (4.11) forms a "source pair" with X_{s1} in terms of R or S shared with X_{tar}. In following sections, we refer to a source pair with the notation $\{X_{s1}, X_{s2}\}$.

We introduce a hypothetical network example to illustrate how the the balance of U, R and S is defined for a network composed of multiple time dependencies (Figure 4.4 and Table 4.1). In this scenario, we detect that target X_{tar} receives information from three sources (Figure 4.4a). There are then three source node pairs (Figure 4.4b), and total shared information from any pair ($I(X_{s1}, X_{s2}; X_{tar})$) may be greater than, less than, or equal to the sum of the individual shared information quantities $I(X_{s1}; X_{tar}) + I(X_{s2}; X_{tar})$, indicating the prevalence

of S versus R. Sources may be independent, such as $\{X_1$,$X_2\}$, or partially dependent, such as $\{X_1$,$X_2\}$ (Figure 4.4c). We apply Equation (4.6) to determine R, and Equations (4.3)-(4.5) to determine U and S for each pair of sources (Figure 4.4d-f and Table 4.1). We see that the source pair $\{X_1, X_2\}$ provides the largest S to X_{tar}, while $\{X_1, X_3\}$ provides the largest R. Since X_1 is the strongest individual source and only provides 1 *bit* of redundant information along with X_3, it is the strongest source of U to X_{tar}. This pairwise and triplet-based analysis of information links places information partitioning into a network context.

Table 4.1: Hypothetical network example based on Figure 4.4, where three sources are found to provide information to a target node. The second three columns denote hypothetically computed information quantities, while the right-side columns are the resulting measures of uniqueness, redundancy, and synergy.

Pair $\{X_{s1}, X_{s2}\}$	$I(X_{s1}; X_{tar})$+ $I(X_{s2}; X_{tar})$	$I(X_{s1}, X_{s2}; X_{tar})$	$I_s(X_{s1}; X_{s2})$	R_{min}	R_{MMI}	R_s	U_{s1}	U_{s2}	S
$\{X_1, X_2\}$	5	7	0	0	2	0	3	2	2
$\{X_1, X_3\}$	4	3	0.5	1	1	1	2	0	0
$\{X_2, X_3\}$	3	3	0.25	0	1	0.25	1.75	0.75	0.25

4.2.3 Characterizing network behaviors

While some network properties may remain stable over long time periods, others may shift as a system evolves or responds to varying conditions. Here we introduce the temporal aspect of TIPNets and validate the utility of our approach for time-series applications compared to the use of Transfer Entropy ($T_E = I(X_{s1}; X_{tar}|X_{tar}(t-\tau))$). In terms of information partitioning, T_E is equal to a unique component of information from a source to a target plus any synergistic information provided by the source and target histories together [Williams and Beer, 2010].

As a surrogate for a network for which connectivity changes with time, we generate a series of five-node networks ($X_i, i = 1...5$) with increasing levels of connectivity. Although we generate a separate dataset for each pre-defined connectivity structure, we use this example to illustrate a network in which links are appearing over time. We connect nodes using the chaotic logistic mapping $X_i(t) = f(X_j(t-\tau))$ where $f(X(t)) \equiv 4X(t)(1-X(t))$. We use the chaotic logistic equation to test the method over highly variable and non-linear data to ensure robustness [Goodwell and Kumar, 2015]. Signals are generated as follows:

$$X_i(t) = \begin{cases} \frac{1}{k_i}\sum_{j=1}^{N} w_{j,i}[f(X_j(t-\tau_{j,i}))], & \text{if } k_i > 0 \\ U[0,1], & \text{if } k_i = 0 \end{cases} \tag{4.12}$$

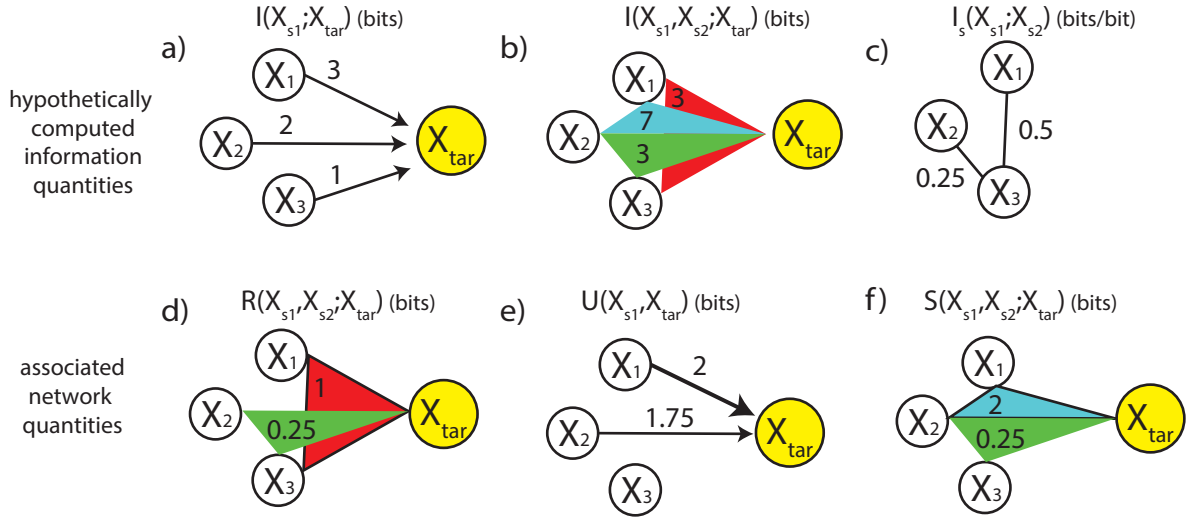

Figure 4.4: Hypothetical network example where multiple sources inform a target. *(a)* In a hypothetical network, target X_{tar} receives information (*bits*) from three sources X_1, X_2, and X_3. *(b)* Each source pair (e.g. $\{X_1, X_2\}$) provides lagged $I(X_{s1}, X_{s2}; X_{tar})$. *(c)* I_s for each source pair indicates the level of source dependency. *(d-f)* Based on *(a-c)*, we compute S, R, and U components from each source pair in Table 4.1. We see that $\{X_1, X_3\}$ is the dominant source of R *(d)*, X_1 is the strongest source of U *(e)*, and $\{X_1, X_2\}$ is the dominant source of S *(f)*.

where N is the number of nodes ($N = 5$), k_i is the in-degree (number of incoming links) of node X_i, w is the adjacency matrix, and τ is the time lag matrix. Each node X_i is uniformly distributed $X_i \sim U[0, 1]$ unless it is linked to another node ($k_i > 0$), in which case it is driven by the chaotic logistic equation. The initial network only contains one link $X_1 \to X_4$ at a specified time lag of $\tau = 1$ (Figure 4.5a), for which we generate a time-series dataset of $n_{steps} = 2000$. For each subsequent case, we add a single link until there are five links (Figure 4.5b-e). As shown in Figure 4.5, each additional link directly or indirectly influences X_4, and we focus on X_4 as the target node. If many links were to be added, nodes would begin to synchronize to a steady state due to accruing feedback [Goodwell and Kumar, 2015]. However, these cases with five links or fewer do not exhibit any synchronization or constrained behavior.

We detect statistically significant $T_E = I(X_s(t - \tau), X_4(t)|X_4(t - 1))$ for all imposed links, and observe that T_E decreases only slightly as links are added (Figure 4.5f). This would lead us to correctly infer that each source provides information to X_4 that is not also provided by

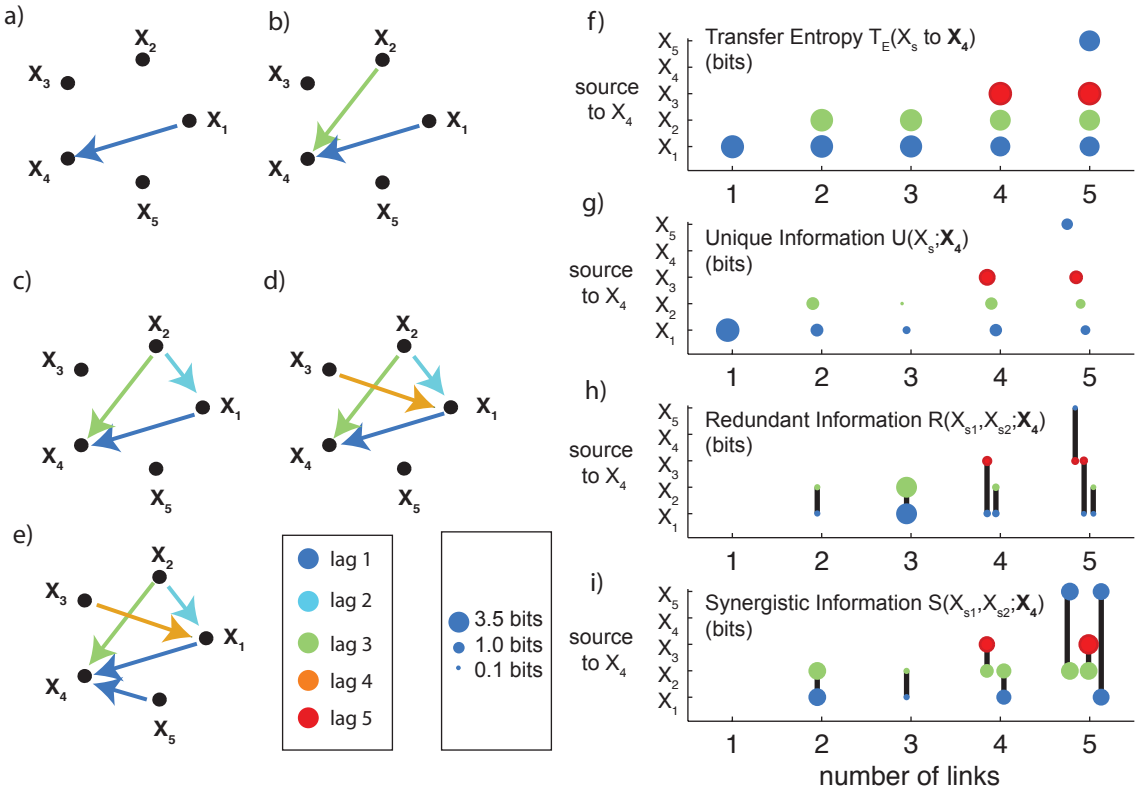

Figure 4.5: Information partitioning for chaotic logistic networks with increasing connectivity. Chaotic logistic links are imposed on uniform random nodes $X_i, i = 1...5$ for five cases as shown in *(a-e)*. Each added link influences the chosen target node X_4 at different time lags as indicated by arrow color. In *(f-i)*, circle color indicates dominant detected time lag τ and size indicates strength (*bits*). In *(h,i)*, black lines indicate source pairs that together provide R or S to X_4. *(f)* T_E is similar for each detected link as links are added. *(g)* U decreases as multiple links are added but increases when independent sources are added. *(h)* R is small when sources are independent and large when sources become linked. Source pairs are linked by black lines. *(i)* S arises as independent sources contribute and decreases as sources become redundant. Source pairs are linked by black lines.

$X_4(t-1)$. In fact, we see that X_4 is never linked as a source to another node, so no feedbacks are ever induced [Goodwell and Kumar, 2015] that would cause the history of X_4 to directly or indirectly inform its own future. However, we would be incorrect to infer from T_E that each source is an equally strong and unique driver of X_4.

An analysis of U, R, and S reveals the dependencies that actually exist *between* the various sources (Figure 4.5). For the first case, the single link is detected as completely unique (Figure 4.5a, Figure 4.5g) since no S or R is possible. When the second link $X_2 \rightarrow X_4$ is added, we

see that the two sources provide mainly U and S, while R is much lower since X_1 and X_2 are independent (Figure 4.5b, g-i). However, some small R is still detected either because $R_{\min} > 0$ (Equation (5.7)) or $I_s > 0$ based on the *pdf*s estimated from the data. The addition of the third link $X_2 \to X_1$ (Figure 4.5c) causes the two previously independent sources to X_4 to become redundant (Figure 4.5c). In fact, we know they are completely redundant since X_2 is the principle driver and X_1 acts as an intermediate influence, i.e. $X_2 \to X_1 \to X_4$. However, the fourth link $X_3 \to X_1$ at $\tau = 4$ introduces a new independent driver so that X_1 and X_2 are no longer completely redundant (Figure 4.5d). We note that this fourth link results in the detection $X_3 \to X_4$ at $\tau = 5$ due to the $\tau = 1$ link from $X_1 \to X_4$ (Figure 4.5g-i). This link re-introduces U and S components, while R is more weakly detected. This increase in S is related to a decrease in U (Figure 4.5g) because of the increased number of joint relationships. Finally, the addition of the fifth link $X_5 \to X_4$ (Figure 4.5e) further increases S as it introduces a direct independent source to X_4 (Figure 4.5i). Some R is falsely detected for the independent source pair $\{X_5, X_3\}$ (Figure 4.5h), although this quantity is weak compared to the stronger U and S. A similar analysis with X_1, X_2, X_3, or X_5 as targets could be performed to define the entire network in terms of S, R, and U.

In summary, this example shows that while our base measure of I_τ (Equation (4.1)) captures time dependencies at correct timescales, more nuanced characteristics can be obtained when uniqueness, synergy, and redundancy are considered as entities in a network context. We next apply the TIPNet framework (Figure 4.2) to reveal frequently shifting patterns in ecohydrologic interactions as measured from weather station data.

4.3 Ecohydrologic networks

In an ecohydrologic system, precipitation, wind, evapotranspiration, and radiation influence each other through forcing and feedback on timescales on the order of minutes. These interactions, as characterized by our TIPNet framework, constitute the ecohydrologic process network structure. As weather conditions vary, network behavior may shift from one time window to the next in terms of altered connection strengths, connection patterns, timescales, and partitioning between uniqueness, synergy, and redundancy. In this section, we use a 1-minute resolution weather station data set to study process network behaviors over several-hour time windows. Dependencies that appear, disappear, or otherwise shift in response to individual weather events may not be captured based on other datasets with averaging intervals of 15 minutes or longer. We look for similarities and differences between networks

for day-time and night-time windows, and windows with rainfall, dew formation, and surface evaporation compared to dry periods. With 1-minute data, we illustrate the assessment of ecohydrologic responses to a multitude of ephemeral conditions (Figure 4.1).

We use data from the Sangamon Forest Preserve (SFP), a site located in Mahomet, Illinois (40.28°N, 88.355°W). The site is a restored prairie, surrounded by corn and soybean agricultural fields and the riparian corridor of the Sangamon River. A weather station has been operated through the Intensively Managed Landscape Critical Zone Observatory (IML-CZO) since June 2014. It collects radiation (Rg, W/m^2), wind speed (WS, m/s), precipitation (PPT, mm/min), relative humidity (RH), air temperature (Ta, ^oC), and leaf wetness ($LWet$, dielectric counts) at 1-minute temporal resolution. A detailed description of weather station instrumentation is provided in the Appendix.

We analyze data over a 180-day time period from April to October (DOY 120-300) of 2015. We separate day-time and night-time periods by the sunrise and sunset as detected from Rg, and further partition day-time and night-time periods into equal length windows so that each day is comprised of four total time windows. Therefore, the lengths of the day-time and night-time windows change throughout the season as day-light hours vary, but are always between $n_{steps} = 250 - 500$ data points. Variables that exhibit a strong diurnal and seasonal cycle (Rg, RH, and Ta) are filtered to retain short term fluctuations while omitting these cycles. This omits synchronization forced by the diurnal cycle, but retains and enhances the detection of other dependencies. Rg is filtered, then night-time periods (when orginally $Rg = 0$) are reset to zero values to omit spurious fluctuations. Details on data pre-processing such as normalization, outlier removal, and consideration of mixed distributions are discussed in detail in a companion paper [Goodwell and Kumar, 2017a] and not repeated here for brevity. We compute lagged I for time lags $\tau = 1 - 60$ minutes, and apply the TIPNet framework to evaluate how network components shift between these several hour time windows and between different weather conditions.

In the next two subsections, we analyze typical network behaviors for rainy, dry, and dew formation and evaporation periods for both day-time and night time networks in terms of dominant time scales of interactions and the partitioning of information into unique, synergistic, and redundant components.

4.3.1 Weather-dependent time scales of interactions

With our 1-minute resolution dataset, we seek to characterize interactions that occur on timescales of minutes to an hour over the set of 720 day-time and night-time windows during the 180-day study period. This is the range of timescales at which we expect to capture interactions shaped by rapid fluctuations that are related to surface heating and cooling, evapotranspiration, rainfall events, and atmospheric mixing (Figure 4.1). While the timescale of an interaction does not provide information regarding its strength (e.g. a longer τ could be associated with either higher or lower information content), it reveals the shifting responsiveness of network components. An increase in dominant τ indicates a more lagged relationship or "delayed" reaction time that could be due to storage or intermediate influences. Correspondingly, a decrease in τ indicates a variable pair that has become more rapidly covarying. When we compare average dominant τ for each of the weather classifications (Figure 4.6a) to the overall average dominant τ (Figure 4.6b-d), we find that different dominant timescales are detected depending on whether it is night-time or day-time in addition to wetness condition.

Out of the 360 day-time windows, we observe 31 for which PPT is detected for at least 5 time steps during the window (Figure 4.6b, top), 88 that are completely dry ($PPT = 0$ and $LWet = 0$) (Figure 4.6c, top), and 127 for which the surface is wet for at least an hour but no rainfall was recorded, indicating wetness from dew or previous rainfall (Figure 4.6d, top). Similarly, for night-time networks we observe 30 rainy, 38 dry, and 227 wet time windows that indicate night-time dew formation. The remaining 179 windows out of the 720 total windows either showed $LWet > 0$ between 1 and 60 minutes or $PPT > 0$ for 1 to 5 minutes, and we do not include them in this analysis.

Average timescales of interactions

For each source→target variable pair, we define the dominant time scale of interaction as the τ at which maximum lagged mutual information I_τ is detected. For all 720 networks, the τ associated with the maximum I_τ is on average less than 40 minutes for all source→target node pairs (Figure 4.6a). This tendency for $\tau < 40$ minutes indicates that feedback and forcing reactions occur on very short timescales and may not be captured from other data sources with lower temporal resolution.

While most variables are sources and targets of information at similar timescales, $LWet$ is a source of information at a shorter average τ than the τ at which it receives (Figure 4.6a).

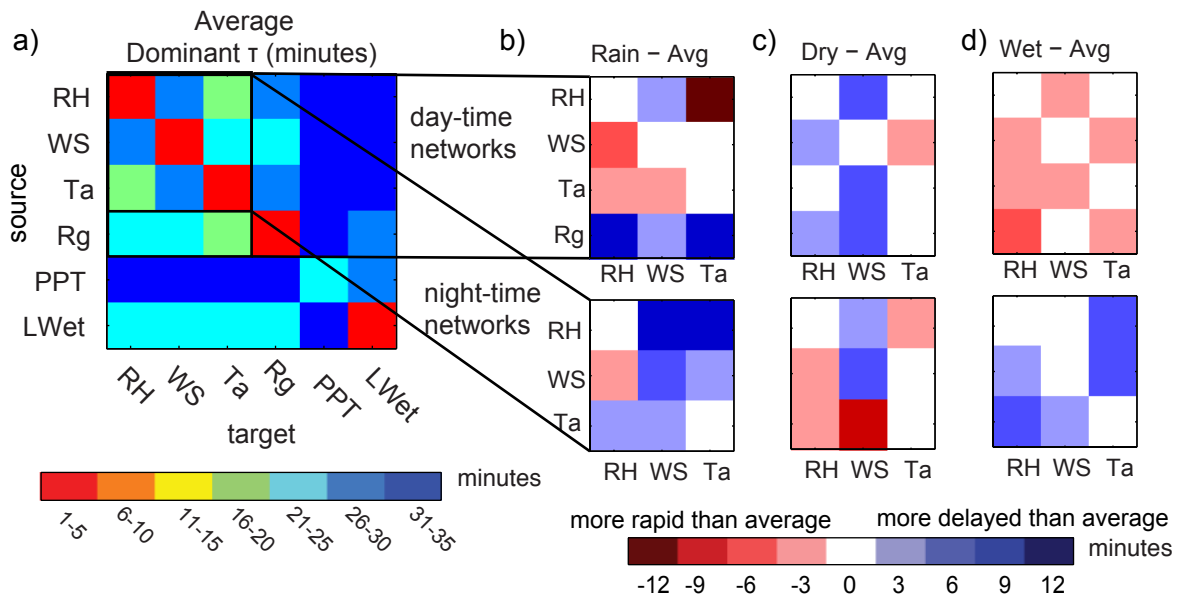

Figure 4.6: Variation of dominant time scales τ (*minutes*) with weather conditions. *(a)* Average τ for each coupling over 720 total network windows. *(b)* Difference between average τ and average of 31 rainy day-time windows (top) and 30 rainy night-time windows (bottom). *(c)* Difference between average τ and average of 88 dry day-time windows (top) and 38 dry night-time windows (bottom). White X marks denote couplings that are detected for fewer than 25% of dry daytime networks. *(d)* Difference between average τ and average of 127 dry day-time windows (top) and 227 dry night-time windows (bottom).

This indicates that the surface wetness condition may influence the rest of the network more rapidly than wetness is itself influenced. We also find that some links such as $RH \leftrightarrow Ta$ and $Rg \rightarrow Ta$ peak at approximately $\tau = 20$ minutes, while others such as $RH \leftrightarrow WS$ peak closer to $\tau = 30$ minutes. Subsets of nodes such as $\{Ta,\ RH,\ Rg\}$ that tend to interact on a faster time scale than others such as $\{RH,\ WS,\ LWet\}$ may indicate that fluctuations in certain nodes may propagate on different timescales than others. Dominant τ is shortest for self-information (e.g. $RH \rightarrow RH$), for which peak I_τ is detected at $\tau < 5$ minutes for all nodes except PPT. Since PPT exhibits a mixed distribution where many values tend to be zero due to intermittent rainfall [Gong et al., 2014], it generally has little information content with which to participate in node interactions. For other nodes, short timescale self-interactions imply that their own most recent histories most reduce their uncertainties. When average timescales for all windows are compared to average timescales for wet, dry, or rainy conditions during day-time or night-time windows (Figure 4.6b-d), we find that

self-interaction timescales tend to be stable, but other links are detected on faster or slower than average timescales according to weather condition.

Fast interactions on rainy days and dry nights

For night-time windows, rainy and dry periods show opposite trends in timescales, in that couplings occur on longer delays for rainy periods and shorter delays for dry periods (Figure 4.6b-c, bottom). This indicates that moisture from night-time rainfall causes *RH* to respond less rapidly to variability in *WS* and *Ta* than for a dry night-time window. During a day-time period in which rainfall occurs (Figure 4.6b, top), we see that *RH*, *Ta*, and *WS* generally interact on a faster than average timescale, indicating their rapid responsiveness to each other during rainy weather conditions. Meanwhile, *Rg* is a source of information at a longer time delay. This reflects the wet and likely cloudy conditions during these periods, for which variability in *Rg* does not cause immediate changes in *RH*.

Fast wet day *Rg* forcing and *RH* response

The most frequent condition for both day and night-time windows is wet surfaces with no rainfall. This indicates regular dew formation overnight and its subsequent evaporation in the morning, in addition to gradual evaporation of intercepted rainfall from a previous time window. For day-time wet windows, interactions occur on slightly faster than average timescales, particularly forcing from *Rg* and responses of *RH* (Figure 4.6d, top), in contrast to rainy day-time windows with delayed forcing from *Rg*. From this we infer that variability of *Rg* leads to relatively rapid responses in *RH* during these evaporative periods. This could be related to increased latent heat flux during these periods, which may intensify connectivity between radiation and atmospheric water vapor. Alternately, slower connectivity during rainfall could be due to the high moisture in the system, which maintains *RH* at a stable high value largely regardless of variability in *Rg*. For night-time windows during which dew or moisture from intercepted rainfall is present (Figure 4.6d, bottom), we find a tendency toward more delayed timescales of connectivity. This could relate to the absence of solar radiation that would otherwise induce mixing and evaporation.

This analysis enables us to define network connectivity in terms of "delayed reactions" or "heightened reflexes" that relate to weather conditions. We next add to this the aspects of information quantity and type for the six weather categories as determined from TIPNet

analysis.

4.3.2 The synergistic, unique, or redundant nature of interactions

For environmental time-series signals, S, R, and U information components arise from forcing and feedback interactions that may or may not occur depending on conditions. In an extreme case, a variable that is not a source or target of any type of shared information is either de-coupled due to an unobserved influence or perturbation, or exhibit low variability (low $H(X)$) during the time window of interest. If a target only receives information from one source, all detected I_τ must be unique since there is no possibility for R or S to exist. Otherwise, each source→target I_τ link contains some balance of S, R, and U. In this section, we discuss salient features of TIPNets for the different weather conditions as categorized in the previous timescale analysis: rainy days, rainy nights, wet days, wet nights, dry days, and dry nights. Here we expect that certain nodes such as *WS* should be more influential in terms of information transfers during rainy (stormy) conditions and day-time windows compared to night-time windows where *WS* is low, while *LWet* should be an important factor during periods of dew formation or evaporation.

Based on the template presented in Figure 4.7 for visualization of network properties, Figure 4.8 illustrates the average information partitioning between sources and targets for each of the six categories. Only links which are active for at least 25% of networks in each category are shown, so that some individual time windows may exhibit additional links but these are relatively infrequent. Additionally, we omit detections of *Rg* as a target of information since we interpret *Rg* as a driving or source variable. Although cloud conditions related to *WS* and *RH* influence *Rg*, we expect this physically based interaction to occur on a longer timescale than the timescale considered here at which *Rg* influences local atmospheric and surface moisture conditions. In the following discussion, we explain prevalent differences in network structure across weather categories.

Common behaviors across networks

For day-time windows, the source pair {*Rg, Ta*} tends to provide the most R to many target variables (Figure 4.8a,b,c). This indicates that fluctuations of *Rg* and *Ta* tend to be correlated (at an average time lag of around τ = 15-20 minutes as shown in Figure 4.6a), even though the diurnal cycle has been filtered. Since *Rg* drives *Ta* through heating of the Earth surface, any

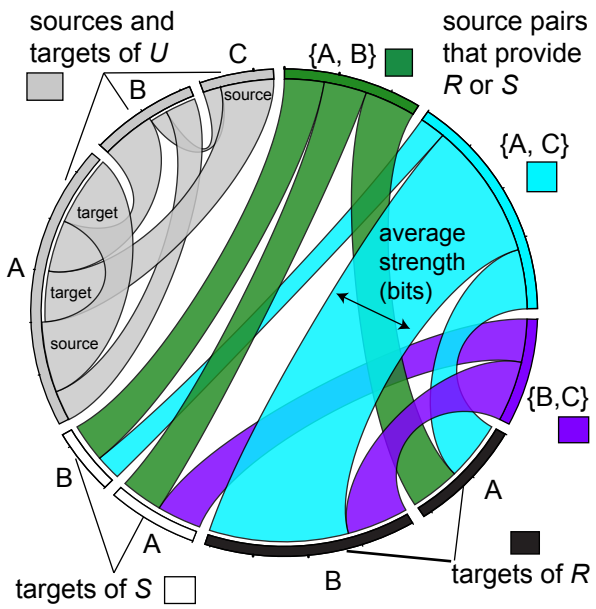

Figure 4.7: Example network visualization of source pairs that provide R and S and individuals that provide U. Link widths between nodes A, B, and C indicate average strengths (*bits*), and colors indicate source pairs that may provide either R or S or both. Gray links indicate transfers of U between individual nodes (note that a node, e.g. A in this figure, can provide U to itself, indicating that the history of A uniquely predicts its current state. Black and white colors indicate targets of R and S, respectively.

provision of R from the pair $\{Rg,\ Ta\}$ to a target variable is likely attributable to variability in Rg. Any further U or S that Ta provides to a target is then related to variability in Ta that is unrelated to Rg.

For night-time windows, the source pair $\{Ta,\ RH\}$ largely dominates both S and R transfers within the network (Figure 4.8d,e). While Ta and RH are directly linked through the atmospheric water vapor holding capacity, they also respond differently to other influences like wind gusts or evaporation, so they are not completely synchronized at any time scale.

While $\{Rg,\ Ta\}$ and $\{Ta,\ RH\}$ on average provide higher R than S, the source pair $\{WS,\ RH\}$ provides a higher proportion of S for both day-time and night-time windows. We see that for all conditions and particularly strongly for dry periods (Figure 4.8b,e), $\{WS,\ RH\}$ provides S to both Ta and WS. WS is also frequently a synergistic source of information along with other variables. These trends reflect the high variability of WS and its nature as a more external source that cannot necessarily be predicted based on the knowledge of the history of another variable.

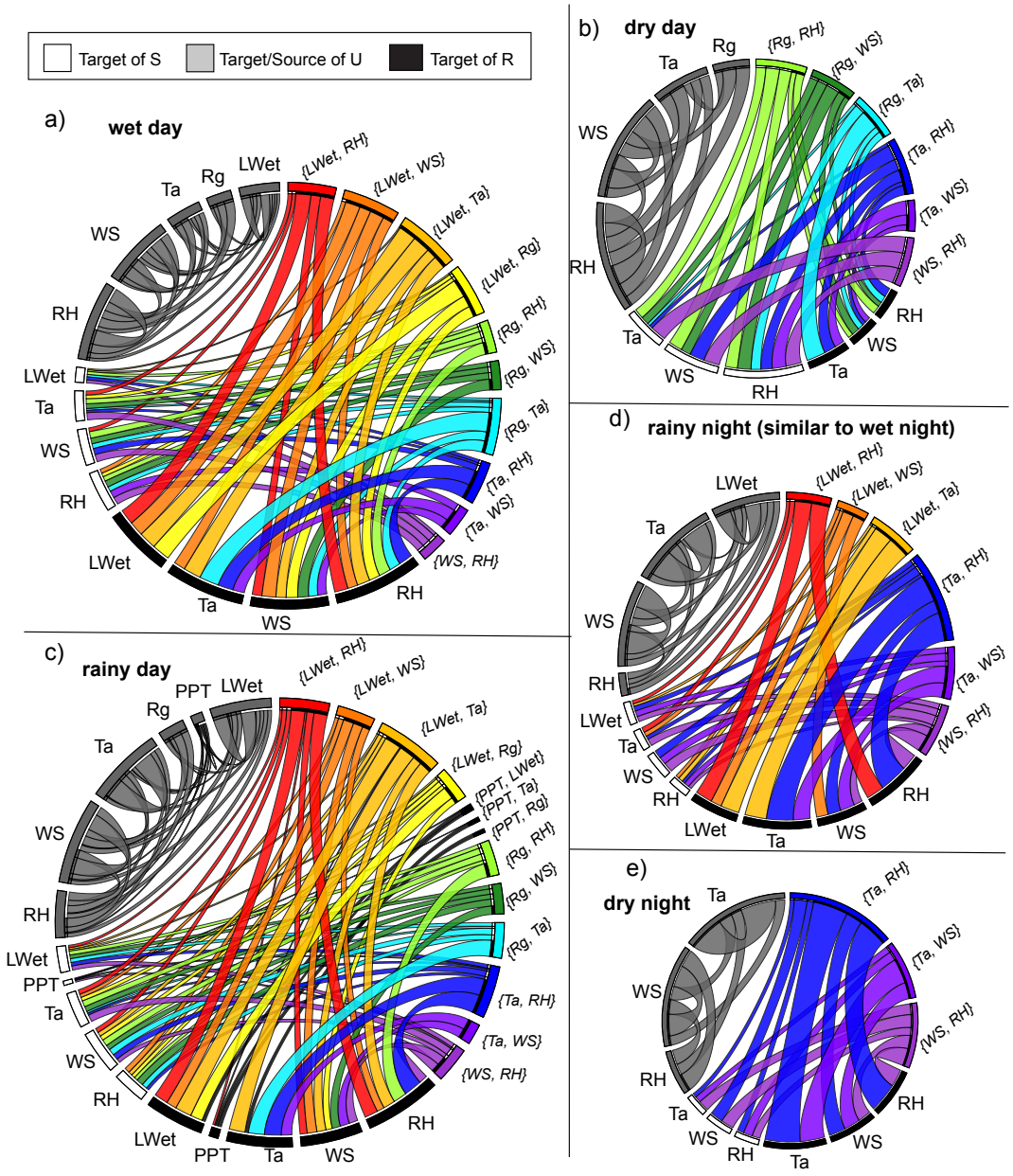

Figure 4.8: Network visualizations for weather categories for different radiation and moisture conditions. Strengths (*bits*) of information transfers that occur for at least 25% of all *(a)* wet day-time, *(b)* dry day-time, *(c)* rainy night-time, *(d)* rainy day-time, and *(e)* dry night-time windows. Target nodes have gaps between links and the outer circumference, while sources are directly attached. Colors indicate source pairs of R and S, and gray links indicate U. Link width indicates average strengths (*bits*), and circle sizes are normalized so that link widths represent the same strength for each weather classification.

Influences of surface moisture and rainfall

For both wet and rainy day networks (Figure 4.8a,b), *LWet* paired with other sources typically provide a large amount of R, and *LWet* is also a target of much higher R than S. These trends are similar for rainy night windows (Figure 4.8d), for which {*LWet, Ta*} is a strong source of R to *LWet* and *Ta*. This tendency towards strong redundancy for *LWet* contributes to larger information content for wet and rainy time windows (larger circle sizes in Figure 4.8) compared to dry windows. This R provided to and from *LWet* indicates that the presence of surface wetness leads to corresponding fluctuations in other nodes such as *RH*, *WS*, and *Ta*, so that any information they provide is redundant.

While wet and dry time windows contrast due to variability in *LWet*, fewer differences can be observed in overall links between wet and rainy day-time windows. While PPT is a source and a target of very small amounts of information for day-time networks (Figure 4.8c), we find that is does not tend to transmit or receive information during rainy night-time windows (Figure 4.8d). This weak or absent influence of PPT is due it its low variability, even during most windows with rainfall. We find that *LWet* is a better indicator of moisture conditions within the network than PPT. While *LWet* is a target of R from its own history paired with other sources, it receives weak S from a variety of sources, leading us to infer that it is not completely redundant with any other variable. Other differences between wet and rainy day-time windows include a higher proportion of U on rainy days, higher influence of source pairs that include *WS* on rainy days, and higher influence of source pairs that include *Rg* on wet days. As discussed previously, dominant timescales also differ between wet and rainy days. During wet day-time windows, *Rg* interacts on a faster timescale as a source of information, while on raining day-time windows, *RH* acts at a faster timescale.

These subtle differences between wet and rainy periods distinguish between different levels of moisture conditions. Moreover, the connectivity of nodes during wet conditions reveals the multivariate dependencies and feedbacks induced by moisture conditions. Namely, precipitation and dew formation are not solely external drivers but both provide and receive information on short timescales.

Wet and dry night-time networks

Night-time networks are characterized by lower information content due to the absence of *Rg* influence, but also by several other features. Mainly, the influence of {*Ta, RH*} is much stronger and provides mainly R (Figure 4.8d,e). We also observe a dominance of self-transfers

of U from WS and Ta at night. This higher uniqueness reflects the variability of fewer nodes such that it is more likely for an individual source to influence a target.

Wet nights and rainy nights are very similar (rainy nights shown in Figure 4.8d, wet nights not shown) in terms of typical structure. Minor differences include stronger S and U influences from WS on rainy nights, and higher U from RH on wet nights. These aspects distinguish between rainy nights that are likely to have more wind influence and nights with dew formation where wind is less variable and surface moisture from condensation is dependent on atmospheric water vapor and temperature.

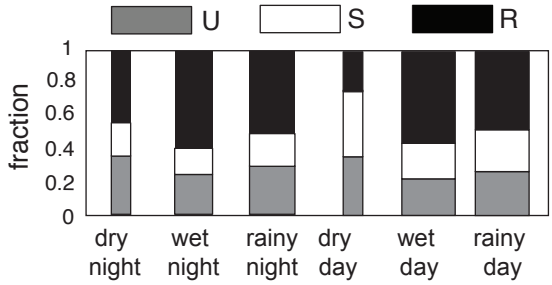

Figure 4.9: Proportions of U, R, and S components show relative differences in information partitioning for weather categories. Bar widths reflect total information strength (proportional to circumference of network circles in Figure 4.8).

We note from Figure 4.8 the wide variety of prevalent linkages for different moisture and radiation conditions. However, the proportions of S, R, and U information components with regards to the total information flow within each network type is relatively similar (Figure 4.9). Dry time windows are characterized by low total information content and a high proportion of U, due to a lack of participating nodes. While dry nights also have a high proportion of R due to redundant influences from $\{Ta,\ RH\}$, dry days have higher S that is partially attributable to more synergistic influences from WS along with other sources. We that information content increases from wet to rainy conditions, and from night to day-time conditions. This may be expected due to the additional nodes that become variable and active within the network for these periods (Rg for day-time conditions, $LWet$ for wet conditions, and PPT and $LWet$ for rainy conditions). However, the shifting proportions of S, R, and U indicate that the added variables exhibit different types of time dependent interactions, and also lead to changes in behaviors of other variables. Namely, the influence of $LWet$ introduces more R during wet periods, more highly variable WS introduces S during rainy periods, and variability in Rg leads to higher S during dry days as compared to dry nights.

This analysis indicates that the system, as characterized by shared information through the network of interactions, is continuously changing through shifts in total shared information as well as the relative proportions of S, R, and U components. This is further explored in the next subsection.

4.3.3 Temporal evolution of network behaviors

Although the analysis of average network structure over a range of weather conditions reveals some prominent TIPNet features, a temporal network approach [Holme and Saramäki, 2012] is needed to gage the timing and extent of forcing and feedback variability as weather conditions change. For example, in our weather station TIPNet application, short-term network behaviors may accumulate over a longer time period to define an ecosystem's response to larger transitions such as drought, altered precipitation variability, or potential temperature increases due to climate change.

The following temporal network analysis can be compared to the previously discussed chaotic logistic network example (Figure 4.5) in that we consider a series of networks that evolve as connectivity changes. In contrast to the chaotic logistic networks for which we validated detected network characteristics based on imposed links, here we must interpret detected links in terms of physical influences on measured surface wetness ($LWet$) and humidity (RH). For a given day of the the year (DOY), we denote the four time periods as DOY_1 (predawn), DOY_2 (morning), DOY_3 (afternoon), and DOY_4 (post-sunset). First we look at leaf wetness ($LWet$) as a target of information over the course of a day during which rainfall occurs and surface moisture subsequently evaporates. Next, we compare sources to relative humidity (RH) for three selected days that exemplify the previously defined categories of rainy, wet, and dry conditions. We present these as illustrations of the understanding that can be gained from a temporal process network approach. A similar analysis on any chosen variable as a target of information could similarly be done to quantify other aspects of system behavior and to obtain a complete picture of network behavior for a given time window.

Moisture shifts due to rainfall and evaporation

Here we analyze how sources to $LWet$ shift over the four time windows of DOY 124, 2015 (Figure 4.10a). On this day, rainfall begins in the middle of 124_1 (Figure 4.10a) and ends at the beginning of 124_2, which is followed by a gradual drying period. Two very short rain

events occur during the latter half of the day.

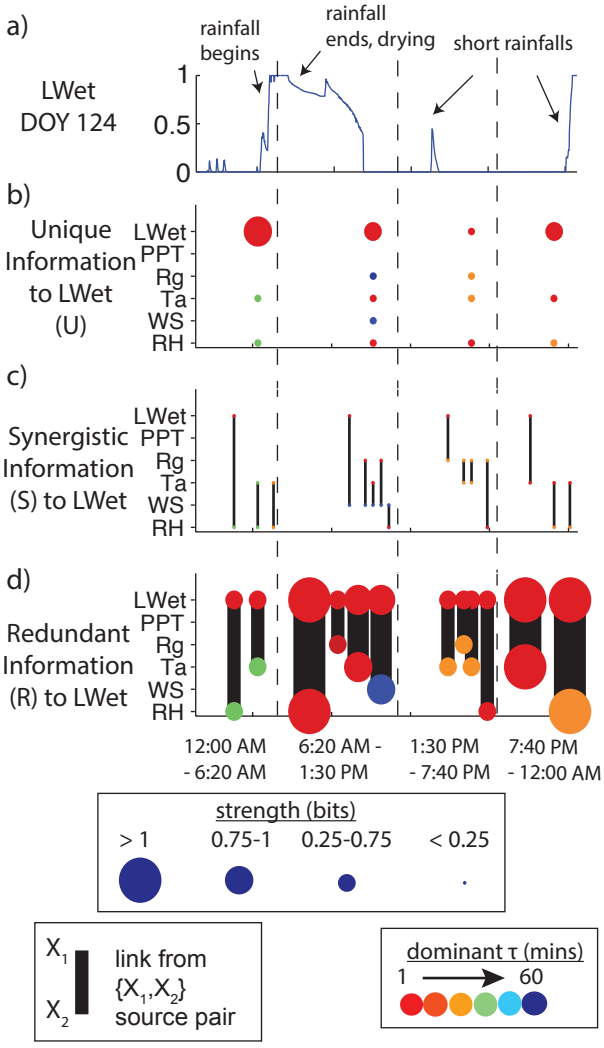

Figure 4.10: Temporal networks of sources to leaf wetness (*LWet*) for a 1-day period. S, R, and U provided to *LWet* shift between time windows due to processes of rainfall, dew, and evaporation of surface moisture. *(a)* Information provided to *LWet* changes over DOY 124 due to a rainfall event that occurs in the early morning, and two very short rainfall events that occur later in the day. *(b)* U is highest in the early morning (124_1) when the rain event begins and leaves become wet. *(c)* As the day progresses, various source pairs provide S to *LWet*, but relatively weakly (strength $U < 0.25$ *bits*). *(d)* Source pairs provide R to *LWet* more strongly than S. During the drying period after the rainfall in 124_2, the pair $\{LWet,RH\}$ is a strong source of R.

Throughout DOY 124, *LWet* is frequently a target of R from its own history paired with

that of other variables (Figure 4.10d). This corresponds to the typical behavior of *LWet* as a redundant source and target, as discussed in the previous section. Compared to R, sources of U and S (Figure 4.10b,c) are relatively weak except for self-provision of U. While timescales of *LWet*→*LWet* information transfers are always very fast (τ between 1 and 5 minutes), other sources provide information at timescales up to an hour.

During 124_1, *LWet* is mainly a target of U from its own history (Figure 4.10b-d). We see that little to no dew formation occurred during this night (Figure 4.10a), and all surface wetness is due to rainfall. During 124_2 when weather conditions shift from raining to drying, *LWet* receives much higher R from a greater variety of sources. While variability in *LWet* in 124_1 is due to rainfall, variability in 124_2 is related to both rainfall and evaporation such that other variables in the network are influential. During this period, *Rg* and *Ta* are redundant sources at short timescales, and *WS* is a source at a longer time delay. This could indicate that evaporation occurs more rapidly due to radiation than wind speed during this period.

In 124_3, all sources of information become weaker and transfers involving *Ta* and *Rg* are detected on a slower timescale. This low information content is due to relatively low *H(LWet)* since $LWet = 0$ for most of 124_3. DOY 124_4 is a night-time window with rainfall, and the strongest links are R to *LWet* from {*LWet, Ta*} and {*LWet, RH*}.

RH interactions over a rainy, dry, and wet day

To evaluate how other network behaviors change with weather conditions, we compare network connectivity for DOY 189, DOY 197, and DOY 293 (Figure 4.11). These three days are representative of rainy (DOY 189), wet (DOY 197), and dry (DOY 293) conditions. On DOY 189, rainfall occurs from 189_2 to 189_3 (Figure 4.11a). During DOY 197_1 and 197_2, *LWet* variability indicates dew formation, evaporation, and a small amount of rainfall (Figure 4.11b). DOY 293 is a dry day during the drought of Fall 2015 (Figure 4.11c). Here we focus on *RH* as a target of information from other variables in the network, since *RH* exhibits very different behaviors depending on radiation and wetness condition (Figure 4.8).

We note the following features that are common across these three days:

1. *RH* receives the highest U from its own history at a fast timescale.

2. During windows with surface moisture, some information that *RH* provides to itself is redundant with a more delayed influence from *LWet*.

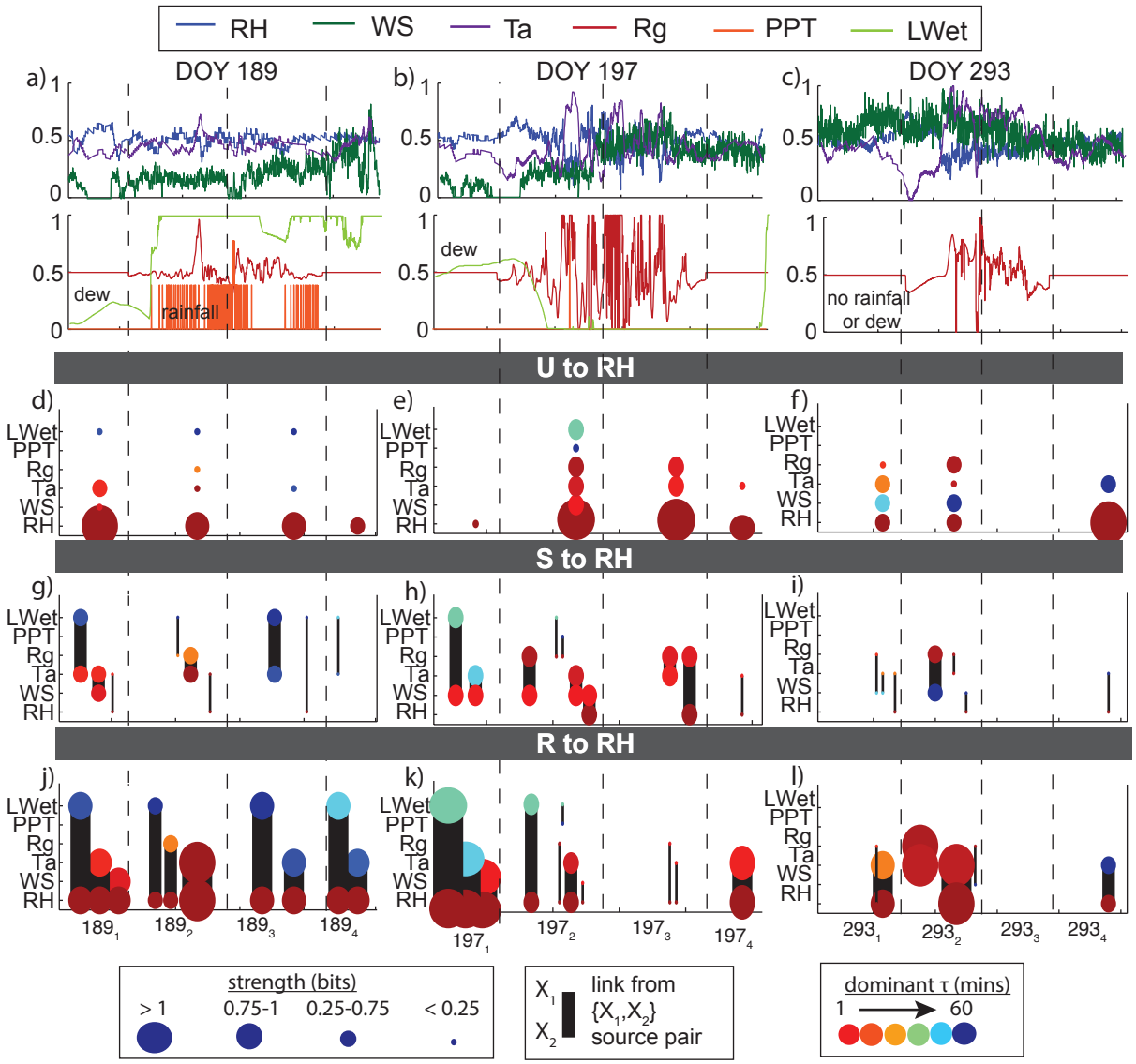

Figure 4.11: Temporal networks of sources to relative humidity (RH) for varying weather conditions over 3 days: DOY 189, 197, and 293 of 2015. *(a)* DOY 189 is a day with rainfall during much of the morning and afternoon windows. *(b)* DOY 197 is a period with dew formation and evaporation, and high variability of other variables. *(c)* DOY 293 is a dry day in the middle of the Fall 2015 drought. Normalized time-series data are plotted for each of the six variables. In *(g-l)* black lines indicate source pairs that provide S or R to RH. *(d-f)* U sources to RH. *(g-i)* S sources to RH. *(j-l)* R sources to RH.

3. *RH* tends to *S* from source pairs that do not include its own history, such as {*LWet, Ta*}, {*Ta, WS*}, and {*Rg, WS*}.

4. During periods with rainfall, *PPT* provides very little or no information to *RH*, due to the low information content of *PPT* relative to other variables.

These features correspond to the trends determined in the previous analyses of dominant timescales and network structures, but we find that other connections and their strengths vary widely even between adjacent time windows and for similar moisture and radiation conditions. For example, although 189_1 and 197_1 are both periods of early morning dew formation, *LWet* and *Ta* provide high *R* to *RH* during 197_1, while *RH* is a stronger unique source to itself for 189_1. Although these two windows are similar in terms of increasing wetness condition, differences in other variables such as *WS* in addition to aspects of soil moisture, wind direction, or vegetation that we do not consider here may lead to different network structures.

For DOY 189, information transfers to *RH* are similar throughout the day as rainfall and evaporation occur intermittently. Sources to *RH* are weakest during 189_4, after rainfall has stopped and surfaces are wet. During this window, *RH* is very high due to saturated air as cooling occurs at night. The tendencies for {*Ta, RH*} and {*LWet, RH*}to provide *R* to *RH* reflects a strong synchronized connection between temperature, surface moisture, and moisture in the atmosphere.

The driest windows of the three days, DOY 197_3 and all of DOY 293, are marked by fewer information links to *RH*. While the afternoon of DOY 197 (197_3) exhibits very weak *R* (Figure 4.11k) and stronger *S* and *U* (Figure 4.11e,h) from different sources, *RH* is completely disconnected during DOY 293_3 (Figure 4.11f,i,l).

From an in-depth look at individual networks under various weather conditions, we see that dominant sources of *S*, *R*, and *U* shift depending on moisture condition, solar radiation, antecedent conditions, and wind. Although the networks often exhibit similar average link strength or number of connected nodes, the composition of links shifts between many possible scenarios.

4.3.4 Seasonal patterns of shared information on link and network scales

Here we consider network evolution over the entire 180-day study period to address the issue of stable or shifting seasonal patterns in couplings. We seek patterns in constructed TIPNets

that reveal whether trends in network metrics can be attributed to gradual seasonal changes or only to short-term changes due to weather conditions. A change in link strengths that persists over many time windows could indicate that the network behavior has altered due to accumulating stresses or seasonal changes. These longer term patterns could indicate healthy or stressed ecosystem conditions, or mark periods where sudden and potentially irreversible changes have occurred.

Our study period of DOY 120-300 of 2015 consists of a wet spring with decreasing rainfall through late summer, and a very dry autumn with no rainfall after DOY 260 (Figure 4.12a). From a 15-day moving average of total information flow (sum of U, R, and S components in Figure 4.12b), we can see trends in proportions of information types and total information strengths over the course of the season. Total information peaks in the early summer (DOY 180) and has a more gradual peak in the early fall (DOY 240). These peaks can mainly be attributed to changes in R, although S also exhibits similar behavior on a smaller scale. The first peak occurs during a high rainfall period of peak vegetation growth, while the second peak corresponds to lower average rainfall during the fall. Total information flow is lowest in the early spring and late fall. For the early spring, weaker information flows are likely related to little vegetation growth as leaves begin to emerge and fewer associated feedbacks between temperature, moisture, and radiation. For the late fall, the similarly weak information flow is likely to be related to a lack of rainfall and both senescence and drought-responses of prairie vegetation during the time period.

The seasonal trend somewhat distinguishes summer and fall transitional periods, throughout which vegetation, radiation, and moisture conditions jointly determine feedbacks between interacting variables. To analyze seasonal patterns in individual interactions such as that between Ta and RH, we can assess each individual interaction in terms of its S, R, and U patterns over the course of the season. We focus on the Ta to RH information link (Figure 4.12c) since it is a relationship that is understood in terms of the influence of temperature on atmospheric water vapor holding capacity, but changes in strength depending on other influences such as WS [Goodwell and Kumar, 2017a]. This source→target link was chosen for illustrative purposes, but any other source→target link can be similarly analyzed.

The overall trend for the $Ta \rightarrow RH$ linkage generally follows that of the entire network. The largest discrepancy occurs at the end of the study period, when total information flow is decreasing but the strength of the $Ta \rightarrow RH$ link is stable (Figure 4.12c). While Ta always drives RH though the atmospheric water vapor holding capacity as discussed in the previous sections, vegetation transpiration, evaporation from surfaces, and atmospheric mixing create

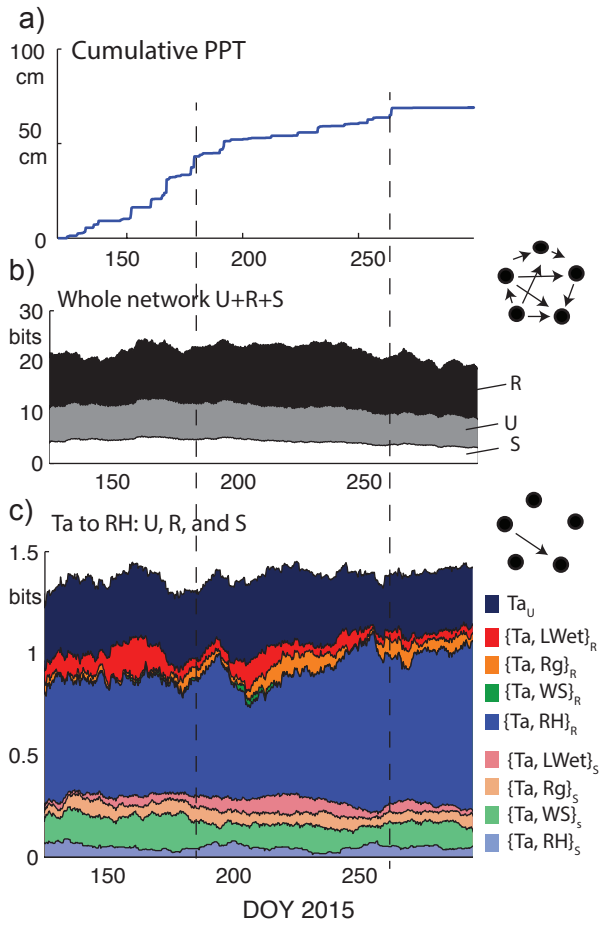

Figure 4.12: Trends in TIPNets for weather station networks over a season. *(a)* Cumulative precipitation (PPT) over the study period. *(b)* 15-day moving average of total information within process network (sum of all network links) and relative contributions of S, R, and U components show that R is largest and shows most variability at the seasonal timescale. *(c)* The 15-day moving average of information related to the link from lagged Ta to RH separated into U (unique information from Ta), R (Ta and another source provide R to RH), and S (Ta and another source provide S to RH).

further feedback between the variables.

We note the following features of the seasonal patterns in information transfers from *Ta* to *RH*:

1. R from the pair $\{Ta,\ RH\}$ increases through the season, indicating increased lagged synchronization between them. This balances with decreasing U from *Ta* to *RH*, so that the increased R does not tend to lead to greater total information flow to *RH*. In other words, the history of *Ta* provides a similar quantity of information to *RH* through the season, but the history of *RH* itself increasingly provides the same information.

2. The peak in total information flow from *Ta* to *RH* in the early spring (DOY 170-180) is largely due to high R from $\{Ta,\ LWet\}$ and high S from $\{Ta,\ WS\}$. This corresponds to the previous findings that *LWet* typically provides R and *WS* provides S.

3. The spring peak in total information flow occurs during a period of relatively high rainfall, and but is followed by a sharp decrease in information flow just after a very large rainfall event. This could indicate a period during which *RH* is very high due to wet conditions and not varying according to fluctuations in other variables.

4. In mid-summer, there is a sharp decrease in R from $\{Ta,\ RH\}$, but this is balanced by an increase in R from $\{Ta,\ LWet\}$ and U from *Ta*. This leads us to infer that during this period (DOY 210-230), *RH* is less predictable based on its own history compared to early spring and fall periods. This period is characterized by lower rainfall and peak vegetation activity, so that components of *Ta* and *LWet* may be more highly variable within each time period and provide more information to *RH*.

While the increasing R from $\{Ta,\ RH\}$ at the expense of U from *Ta* indicates the increasing lagged synchronization between *Ta* and *RH*, we find that *Ta* also provides information to *RH* jointly with other variables. These joint provisions of information indicate that the connection between *Ta* and *RH* is not completely stable throughout the season. Rather, it is mediated by other influences including moisture, radiation, and wind. For example, the larger influence of *LWet* in terms of both S and R in the late summer could indicate the ecosystem's heightened responses to wetness during a dry period with high biomass. An extended study with observations of evapotranspiration, carbon fluxes, and heat fluxes could further inform many of these shifting dependencies.

This seasonal analysis of network-scale information transfer and components of a single link illustrates the understanding of seasonal changes that can be gained from a network

approach in which multivariate interactions are considered. While the relationship between Ta and RH is relatively well-understood and can be simulated, we find in our analysis using observed data that the dependency is actually quite complex and varies over time on both daily and seasonal time-scales as conditions change. In general, a simple lagged relationship between two variables can be greatly altered when other influencing factors become active.

4.4 Discussion

The partitioning of information transfer between environmental signals into unique (U), synergistic (S) and redundant (R) components reveals a deeper of understanding of ecohydrologic interactions than would be obtained using linear correlation measures or pairwise information theory metrics. While measures such as Transfer Entropy (T_E) partially gage the uniqueness of a time dependency, they do not include information that is gained from multiple sources or place individual links into the context of the network as a whole. Based on a 1-minute weather station dataset, the TIPNet framework enables us to make several conclusions about the behavior of this particular ecohydrologic network:

1. On average when time lags τ =1-60 minutes are considered, dominant time scales of interactions are $\tau < 40$ minutes. While RH influences other variables faster than average for both dry nights and raining days, Rg has the shortest timescale of influence on days when surface evaporation occurs.

2. Strengths of U, R, and S interactions and their relative proportions of total information flow vary with moisture and radiation conditions. Generally, total information increases when solar radiation and moisture-related variables are active in the network, and day-time networks have more synergistic interactions than night-time network.

3. While certain links remain stable over the growing season in terms of their strength, others tend to peak in strength at different times in the season.

From an ecohydrologic perspective, TIPNets from high resolution weather station data reveal the constantly shifting structure of interactions that could not be inferred given a more sparse data set or different statistical analysis technique. We implement this framework to address the notion that for an ecosystem, the "whole" is greater than the union of its "parts". In other words, synergistic and redundant information components reveal aspects of

interactions that cannot be defined simply as driver-response mechanisms because multiple interrelated components are involved. These components do not arise unless at least three variables are considered simultaneously. A network of more than three variables, such as the network of environmental signals presented here, can have a complex network structure in which the predictability of a single target is maximized only when a number of sources and source pairs are considered.

We show that forcing and feedback occur in this system on time scales of less than an hour and can be detected over several-hour time windows. This temporal resolution is appropriate to capture individual storm events, dew formation and evaporation, or changes in cloudiness. Although these findings are specific to the temporal resolution, study period, and location of this particular dataset, our novel approach has implications for the broader study of ecosystem vulnerability and resilience in addition to modeling. By removing the diurnal cycle and detecting lagged time dependencies on short timescales, we can make inferences regarding what aspects of these dependencies and their thresholds for activation are important drivers within the system that are not included in models. Given a more sparse resolution dataset such as hourly flux tower data, this approach can still reveal interactions that occur on hourly to daily timescales over several-day time windows. A flux tower or simulated dataset such as from North American Land Data Assimilation System (NLDAS) forcing or model data [Mitchell et al., 2004] would also contain a greater variety of variables such as CO_2 concentrations, water vapor fluxes, evapotranspiration, or long wave radiation that would capture more interactions than those discussed here.

This framework could lead to better understanding of how ecosystems respond to perturbations at various timescales. Land cover change and intensive management impacts soil moisture, evapotranspiration, and reflectivity, all of which influence local or regional climate dynamics such as wind or rainfall that feed back into the ecohydrologic system at different time scales. Prolonged drought results in vegetation stress, and interactions between the plant, root, and atmospheric systems may shift as an ecosystem becomes degraded. Evaluations of these multi-faceted interactions requires a robust network approach.

The framework presented here can also deepen understanding of other types of systems that involve complex time-dependent interactions. The application of temporal network and information measures can reveal emergent properties related to system level behavior. The ability to identify and characterize shifts in connectivity can help us understand how changes in a single component, such as land use or rainfall, may propagate through the network of interactions.

CHAPTER 5

LOCAL AND NON-LOCAL PROCESS CONNECTIVITY ALONG AN ELEVATION AND CLIMATE GRADIENT

5.1 Introduction

In an ecohydrologic system, components of the atmospheric, soil, and canopy subsystems are connected through a variety of nonlinear forcing and feedback interactions. Although the states of variables such as heat fluxes, nutrient concentrations, and moisture content are constrained by conservation laws and rate limits, their behaviors and variability are otherwise complex and not fully predictable. For example, energy inputs from solar radiation drive heat and water vapor fluxes between the soil and the atmosphere, but these fluxes also participate in feedbacks with atmospheric, vegetation, and soil states. Many of these dependencies also involve an aspect of spatial connectivity, in that fluctuations in variability propagate both between ecohydrologic processes and spatially with weather patterns, flow paths, and atmospheric mixing. While some interactions may occur dominantly locally, based on conditions at a site, others may exist over a spatial gradient. The entire set of local and non-local interactions constitutes a complex network in which the knowledge of individual variables or dependencies between pairs of variables cannot explain larger scale behaviors such as heat flux partitioning, vegetation responses to climatic variability, or the influence of directional weather patterns on ecosystem behaviors.

Since interactions within an ecohydrologic system tend to be multivariate, non-linear and vary with time, their characterization necessitates an appropriate network framework. A process network is defined as the set of time-dependent "links" detected between measured time-series variables that act as "nodes" [Ruddell and Kumar, 2009a]. Process networks of time dependencies in ecohydrologic systems have characterized the influence of weather condition on interactions between variables measured at a weather station [Goodwell and Kumar, 2017a, Goodwell and Kumar, 2017b], and seasonal connectivity between ecosystem fluxes measured at an eddy covariance flux tower [Ruddell and Kumar, 2009a, Ruddell and Kumar, 2009b]. Within a delta system, process networks involving tides, wind events, and water levels

at different sites addressed both spatial and temporal components of process connectivity [Sendrowski and Passalacqua, 2017].

In previous ecohydrologic studies that employ the process network approach, unexplained variability, or behavior in a variable that cannot be explained by influences within the network, must be attributed to external or observed drivers or "noise" [Goodwell and Kumar, 2015]. A portion of this unexplained variability may be due to low quality or missing observations at the study site. For example, if solar radiation is not measured at a weather station, we would expect some fluctuations in temperature or heat flux variables to be unexplained. As more connected variables are considered, more joint variability can be detected. However, some unexplained variability may be due to non-local influences. In this case, observations at neighboring or distant locations are needed to construct a more complete network that represents both connectivity between processes at a location and over space. For example, connections between surface temperature anomalies in global climate networks have been detected over large spatial gradients [Deza et al., 2015, Wiedermann et al., 2016], and rainfall patterns can be studied in terms of spatio-temporal synchronization [Rheinwalt et al., 2016, Boers et al., 2013].

In this paper, we extend the process network framework to address aspects of local and nonlocal connectivity in ecohydrologic systems. Specifically, we seek to (i) characterize how forcing and feedback interactions differ along climate, elevation, and vegetation gradients, and (ii) determine the extent of process connectivity between sites along these gradients. The first problem is concerned with the identification of shifts in dominant processes related to precipitation type and amount, climate, or vegetation and soil properties along the gradient. The second problem is concerned with distinguishing local connectivity from variability at a site that can only be explained when spatial connectivity is accounted for. This involves the identification of dependencies between uphill and downhill neighboring locations as sources of non-local interactions.

To achieve these goals, we apply a Temporal Information Partitioning Network (TIPNet) methodology, in which nodes are time-series variables and links between them are information theoretic metrics. We consider mutual information detected between τ-lagged "source" and "target" nodes as time-dependent links associated with a strength and a dominant timescale. Moreover, we compute joint dependencies from two sources to a target, and partition joint shared information into unique, synergistic, or redundant components (Figure 5.1a). Redundant information (R) is overlapping information that both sources provide to a target, and indicates a degree of lag-synchronization between sources. Unique information (U)

is information that only a single source provides to a target, and synergistic information (S) is only provided when both sources are known together. A common example of S is the binary exclusive-OR (XOR) operation, for which the individual sources appear to be independent of each other and the target unless both sources are known together. This partitioning into U, R, and S components enables us to interpret the nature of a time dependency in a multivariate context. Here we perform information partitioning for three-node cases (two sources and a target), but this could be extended given the ability to compute higher than 3-dimensional probability density functions (pdfs). For a given time window, each variable is considered as a potential source and target node, and a TIPNet is characterized as the set of source pairs that provide R and S, and individual sources that provide U. In this study, we consider 10-day and 1-month time windows to capture seasonal shifts in network behaviors. This method enables us to identify multivariate interactions within and between sites that would not be observable based on an analysis of pairwise interactions or individual nodes.

We construct TIPNets from flux towers located along two transects at the Southern Sierra and Reynolds Creek Critical Zone Observatories (CZOs) located in California and Idaho, respectively. Each transect is composed of four eddy covariance towers approximately equally spaced along an elevation gradient. Along both transects, precipitation increases, temperature decreases, and vegetation composition shifts as elevation increases, but the sites are close enough to experience similar weather and seasonal patterns. These datasets allow us to address issues related to connectivity between energy and water balance components, in addition to spatial information transfer between neighboring locations (Figure 5.1b). Specifically, we address several questions regarding influences to ecohydrologic behaviors along these transects:

1. What is the partitioning of information provided to latent, sensible, and ground heat fluxes (LE, H, and G) from radiation, atmospheric, and soil variables (Figure 5.2a), and how does this vary along the elevation gradient?

2. How does the forcing relationship from incoming solar radiation (SW_{in}) to carbon flux (Fc) depend on seasonality, and the suite of other interacting variables related to soil and vegetation processes (Figure 5.2b)?

3. To what extent does unique shared information exist locally, or at a single site, versus non-locally, or between sites (Figure 5.1b)?

This paper is organized as follows. In Section 5.2, we describe the Temporal Information Partitioning Network (TIPNet) methodology. In Section 5.3, we describe the study sites and

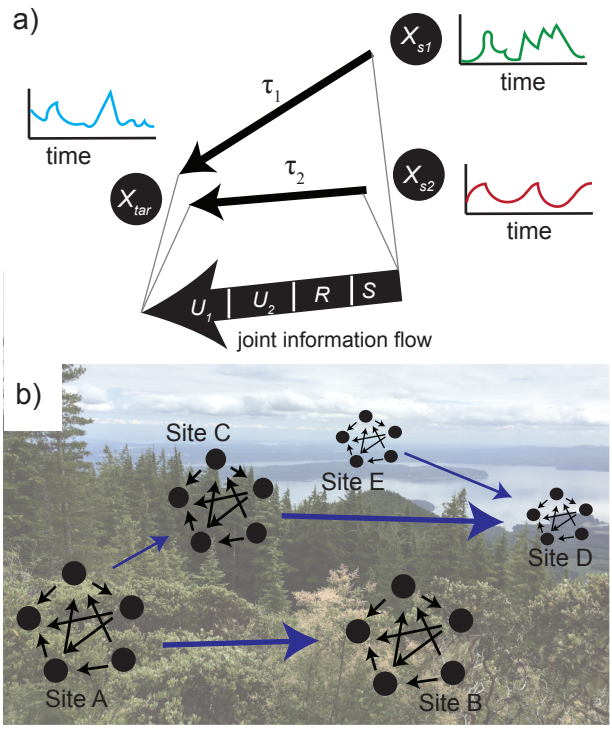

Figure 5.1: Illustration of information partitioning and networks in a spatial context. *(a)* Time-series variables are nodes in a network. Lagged mutual information constitutes a link between source nodes (X_{s1} and X_{s2}) and target nodes (X_{tar}). The joint mutual information $I(X_{tar}; X_{s1}, X_{s2})$ can be partitioned into unique (U), synergistic (S), and redundant (R) information components. *(b)* The Temporal Information Partitioning Network (TIPNet) is the set of S, R, and U interactions detected between all nodes. A TIPNet could be constructed from variables at a single site to capture local process connectivity, or from variables at multiple sites to capture both local and non-local time dependencies.

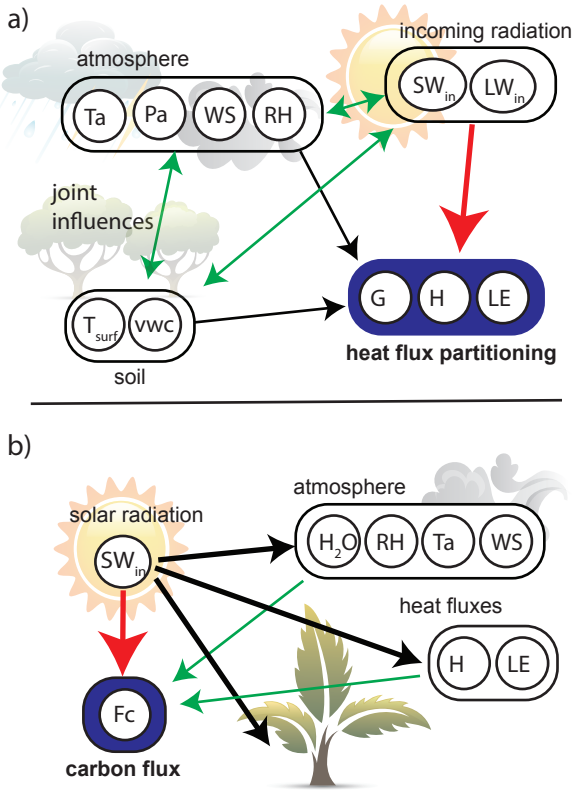

Figure 5.2: Illustration of information flows from environmental drivers to heat and carbon fluxes. *(a)* We address the partitioning of information provided to energy fluxes (G, H, and LE) from radiation, soil, and atmospheric influences, jointly or individually. *(b)* We also address the influence of radiation and atmospheric states on carbon flux (Fc). Variables are flux tower observations as described in Table 5.2.

our TIPNet. In Section 5.4, we discuss the flow of information from radiation and atmospheric variables to carbon and heat flux partitioning at the different sites. In Section 5.5, we address the spatial component of information flow between sites in terms of directionality and time scales. In Section 5.6, we discuss these findings in the broader context of complex networks.

5.2 Review of Temporal Information Partitioning Networks

Here we outline the TIPNet methodology [Goodwell and Kumar, 2017b] that we use to reveal local and non-local process connectivity along each flux tower transect. We consider time-series variables as nodes that may be "sources" or "targets" of information. Mutual information is a measure of a time-dependency in terms of a reduction in uncertainty of a target node X_{tar} due to the knowledge of a τ-lagged source X_{s1}:

$$I(X_{s1}; X_{tar}) = \sum p(x_{s1}, x_{tar}) \log_2 \left(\frac{p(x_{s1}, x_{tar})}{p(x_{s1}) p(x_{tar})} \right). \tag{5.1}$$

Although I (in units of *bits* due to natural logarithm) is symmetric, the lagged node X_{s1} is regarded as the source of information, while the current node is the target. We employ Kernel Density Estimation (KDE) [Silverman, 1986, Lee et al., 2012] to estimate *pdf*s from observed data, and test estimated values of $I(X_{s1}; X_{tar})$ for statistical significance using the shuffled surrogates method [Ruddell and Kumar, 2009a, Goodwell and Kumar, 2017a]. X_{s1} and X_{tar} may be different variables, or current and lagged versions of the same node (e.g. $X_{tar} = X_{s1}(t - \tau)$). The maximum statistically significant value of $I(X_{s1}; X_{tar})$ and its associated time lag τ are defined as the strength and timescale, respectively, of the source→target link between the two variables.

Based on this pairwise network of shared information strengths and dominant timescales, we apply information partitioning to further characterize multivariate linkages. For a given target node, we seek to quantify the shared information that is redundant between source nodes, unique to a single node, or synergistic between source nodes. If only one source node is detected to provide information to a target, this information is defined as U. Otherwise, information provided from two sources to a target could be unique to either source, redundant or overlapping between sources, or synergistic between sources. We employ information partitioning, in which total mutual information from two sources to a target is partitioned into these unique (U), synergistic (S), or redundant (R) components [Williams and Beer, 2010, Barrett, 2015]:

$$I(X_{s1}, X_{s2}; X_{tar}) = U_1(X_{s1}; X_{tar}) + U_2(X_{s2}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}) + S(X_{s1}, X_{s2}; X_{tar}) \quad (5.2)$$

$$I(X_{s1}; X_{tar}) = U_1(X_{s1}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}) \quad (5.3)$$

$$I(X_{s2}; X_{tar}) = U_2(X_{s2}; X_{tar}) + R(X_{s1}, X_{s2}; X_{tar}). \quad (5.4)$$

While the left-hand-side terms of Equations (5.2)-(5.4) can be computed from Equation (5.1), one component out of U_1, U_2, R, S, must be directly computed in order to obtain the information partitions. We compute R based on Rescaled Redundancy R_s, as follows [Goodwell and Kumar, 2017a]:

$$R_s = R_{\min} + I_s(R_{MMI} - R_{\min}) \quad (5.5)$$

where $I_s = \frac{I(X_{s1};X_{s2})}{\min[H(X_{s1}),H(X_{s2})]}$ is the normalized source dependency, and R_{MMI} and R_{min} are upper and lower bounds of redundancy, respectively, as follows:

$$R_{MMI} = \min[I(X_{s1}; X_{tar}), I(X_{s2}; X_{tar})] \quad (5.6)$$

$$R_{\min} = \max(0, -II) \quad (5.7)$$

where $II \equiv I(X_{s1}; X_{s2}; X_{tar})$ is the interaction information, and equivalent to $S - R$. This formulation for R is based on the concept that independent sources should be maximally unique and highly dependent sources maximally redundant. A detailed discussion of advantages and limitations the R_s metric as compared to other information partitioning techniques can be found in [Goodwell and Kumar, 2017a], and further application in [Goodwell and Kumar, 2017b]. In general, R_s allows for the detection of multiple unique information components, and is appropriate for application to environmental datasets where nodes are time-series variables that influence each other via different processes. While existing information measures like Transfer Entropy [Schreiber, 2000] and the interaction information are combinations of information partitioning components, a measure for R provides a way to directly partition information and distinguish between different types of information flow.

For each pair of source variables that provides information to a given target, we compute S, R and U components. We then define unique information for each source-to-target pair as the minimum U that the source provides with the target along with any other source. In

Table 5.1: Site details for Reynolds Creek (RC) and Southern Sierra (SS) flux towers, including location, vegetation types, and study period.

Transect	Flux tower name (abbrev)	Short name	Elev (m)	Latitude	Longitude	vegetation	study period
RC	Wyoming Big Sage (wbsec)	RC_1	1425	43.1675	-116.7132	Wyoming sage	2015
RC	Lower Sheep (losec)	RC_2	1608	43.1439	-116.7356	low sage	2015
RC	Upper Sheep (138h08ec)	RC_3	1878	43.1207	-116.7231	post-fire sage	2015
RC	Mountain Big Sage (mbsec)	RC_4	2111	43.0645	-116.7486	mtn big sage	2015
SS	San Joaquin Exp. Range (SJER)	SS_1	405	37.1086	-119.7314	grasses, pine, oak	2011
SS	Soaproot Saddle	SS_2	1160	37.0306	-119.2562	pine and oak	2011
SS	Providence Creek P301	SS_3	2015	37.067432	-119.1935	pine and fir	2011
SS	Shorthair Creek	SS_4	2700	37.0691	-118.9823	subalpine	2011

other words, $U(X_{tar}, X_{s1})$ is defined to be the information within the $X_s \rightarrow X_{tar}$ link that is not provided by any other link. The TIPNet is then defined as the set of source pairs that provide R and S to each target node, the individual sources of U, and their associated strengths (*bits*) and time scales (τ).

5.3 Site description

We use data from two flux tower transects to analyze network behavior in terms of seasonal differences in dominant interactions, changes along the temperature, precipitation, or vegetation gradients, and the nature of information flow between sites. The study transects are located within the Southern Sierra (SS) and Reynolds Creek (RC) Critical Zone Observatories (CZOs) in California and Idaho, respectively (Figure 5.3). Each CZO maintains four towers that are approximately equally spaced along elevation gradients (site details in Table 5.1). Variables related to radiation, weather, and heat and carbon fluxes are measured at each tower at a half hourly time step (Table 5.2). We analyze data from SS during 2011, and RC during 2015, based on data availability. At all sites, gaps in data due to instrument failure are most frequent during winter months (November - February), and we omit these months from further analysis.

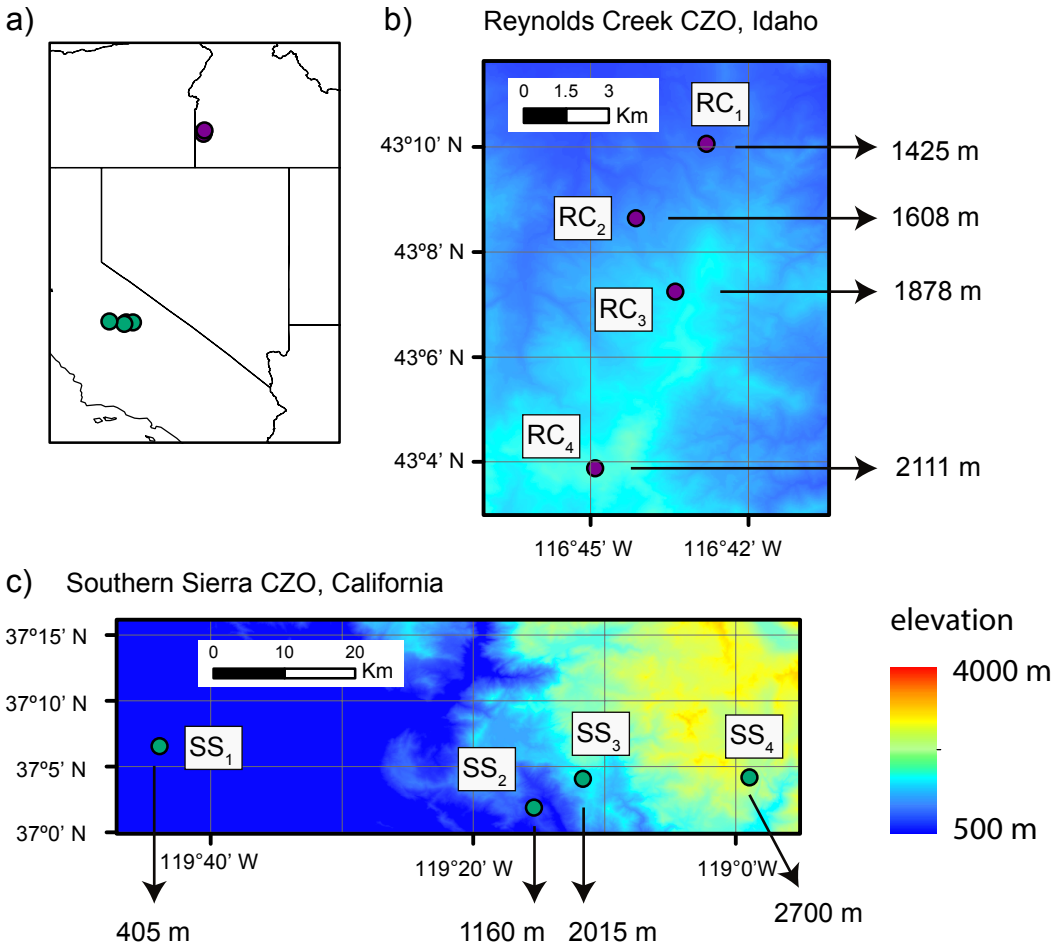

Figure 5.3: Maps of flux tower transects at Reynolds Creek (RC) and Southern Sierra (SS) CZOs. *(a)* Eddy covariance flux tower transects are located at Southern Sierra (SS) CZO in California, and the Reynolds Creek (RC) CZO in Idaho. *(b)* At SS, tower elevations increase about 800 m between each site. *(c)* At RC, sites are more closely spaced and elevations increase about 200 m between sites.

5.3.1 Southern Sierra

A detailed description of the SS flux tower data set and climate-dependent patterns of evapotranspiration (ET) and carbon uptake can be found in [Goulden et al., 2012]. The region has a Mediterranean climate with wet winters and a dry season between May and October. Conditions become wetter and colder as elevation increases from east to west, with decreasing vegetation cover and a shift from annual grasses (SS_1) to subalpine forest (SS_4). Along the transect, ET is limited by dry summer conditions at the lowest elevation site (SS_1) and by cold winters at the highest elevation site (SS_4) where vegetation is dormant during the winter. ET peaks in May at SS_1 and in August at SS_4. Net CO_2 uptake is negative at SS_4 during the winter, indicating a flux of CO_2 into the atmosphere. The two mid-elevation sites (SS_2 and SS_3) are characterized by relatively higher annual ET, a mix of snow and rain precipitation, and vegetation activity that persists throughout the year [Goulden et al., 2012]. At the flux tower sites, ET is quantified by the latent heat flux variable, $LE = \lambda ET$, where λ is the latent heat of evaporation.

5.3.2 Reynolds Creek

The Reynolds Creek CZO has a long history of soil and weather observations dating from the 1960s [Seyfried et al., 2011, Hanson et al., 2001, Nayak et al., 2010]. The climate is characterized as typical of much of the intermountain region of the Western U.S. As at SS, precipitation increases with elevation, and winters are wet compared to summers. Although temperatures generally decrease with elevation, a strong temperature inversion in the valley bottom during the winter can cause day-time minimum temperatures to be colder at the lower elevation site (RC_1) compared to the middle elevation sites (RC_2). A temperature increase in this region of about 2°C has been detected over the past 40 years, which has led to a decreasing proportion of snow to rain but no statistically significant change in precipitation volume [Nayak et al., 2010]. This shift in temperature over the past several decades is causing the higher elevation sites to become more like the mid-elevation sites as they existed in the 1960s. While SS sites are up to 40 km apart, each RC site is only several km away from the other sites.

Table 5.2: List of time series variables measured at RC and SS flux towers.

variable	description	measurement sites	filter method
SW_{in}	downward shortwave radiation (W/m^2)	RC and SS	5-day anomaly
SW_{out}	upward shortwave radiation (W/m^2)	RC and SS	5-day anomaly
LW_{in}	downward longwave radiation (W/m^2)	RC	5-day anomaly
LW_{out}	upward longwave radiation (W/m^2)	RC	5-day anomaly
PAR	photosynthetically active radiation (W/m^2)	RC and SS	5-day anomaly
LE	latent heat flux (W/m^2)	RC and SS	Butterworth filter
H	sensible heat flux (W/m^2)	RC and SS	Butterworth filter
G	ground heat flux (W/m^2)	RC and SS	Butterworth filter
T_{surf}	surface temperature ($°C$)	RC and SS	Butterworth filter
T_{soil}	soil temperature at 5 cm ($°C$)	RC and SS	5-day anomaly
VWC	volumetric water content at 5 cm depth	RC and SS	none
Ta	air temperature ($°C$)	RC and SS	Butterworth filter
Pa	air pressure (kPa)	RC	Butterworth filter
RH	relative humidity (%)	RC and SS	Butterworth filter
H_2O	water vapor density (g/m^3)	RC and SS	Butterworth filter
WS	wind speed (m/s)	RC and SS	none
WD	wind direction (° from N)	RC and SS	none
CO_2	carbon dioxide concentration (mg/m^3)	RC	Butterworth filter
Fc	carbon dioxide flux ($mg/m^2/s$)	RC and SS	Butterworth filter
PPT	precipitation (mm)	RC	none

5.3.3 TIPNet application to flux data

We take several pre-processing steps before applying our TIPNet methodology to the set of observations from each flux tower. For radiation variables that exhibit a diurnal cycle, but are zero during the night, we take the 5-day anomaly. For other variables such as air and surface temperatures, humidity, and carbon and heat fluxes, we use a Butterworth filter to omit the diurnal cycle. Both of these techniques enhance the detection of links between nodes at shorter than daily time scales (Table 5.2). We omit extreme outliers, and normalize each dataset to a $[0, 1]$ range based on the minimum and maximum before estimating *pdfs*. We segment the 1-year dataset into 10-day or monthly windows to construct networks. Since many flux tower datasets contain intermittent gaps due to instrument failure or extreme weather, we omit time windows for a given variable as a possible source or target if more than 20% of data points (e.g. 100 out of 500 points) are missing. We seek to detect information measures at lags between $\tau = 30$ minutes (1 time step) and $\tau = 1$ day (48 time steps).

To address questions regarding the connectivity between the energy balance, water balance, and carbon dynamics, we segment the flux tower datasets into 10-day windows after pre-processing. Each window overlaps the previous window by 5 days, so that there are about six windows within a given month during 2015 at RC and 2011 at SS. To study levels of

connectivity within and between sites, we segment the datasets into monthly windows that do not overlap. We compute TIPNets for each window, and then focus on specific source and target nodes to answer related questions.

5.4 Connectivity between energy, water, and carbon fluxes

A characterization of influences to carbon and heat flux components improves our understanding of the relative dominance of various drivers and mechanisms that act at different locations and times over a season. First we address the influences behind the partitioning of energy from solar radiation into latent, sensible, and ground heat fluxes. For this analysis, we focus on only the four RC sites, as there are fewer gaps for the variables of interest. Second, we assess how energy fluxes in addition to soil, moisture, and temperature properties jointly influence carbon flux along with solar radiation at both RC and SS sites.

5.4.1 Shifting influences to energy partitioning

Heat flux components of latent (LE), sensible (H), and ground heat (G) are constrained by the energy balance relationship involving shortwave (SW) and longwave (LW) radiation as follows:

$$SW_{in} + LW_{in} = SW_{out} + LW_{out} + H + LE + G + \Delta G. \tag{5.8}$$

where ΔG is a storage term of ground heat. Although this balance constrains the participating variables, the variables on the right hand side of Equation (5.8) also depend on interactions with surface wetness, temperature, atmospheric conditions, and vegetation activity. For example, the proportion of incoming energy partitioned to LE versus H is constrained by soil moisture availability (VWC), in that the presence of water enables some solar energy to be directed towards evaporating moisture that would otherwise go toward H, which leads to a temperature increase. Additionally, wind speed and direction (WS, WD) induce mixing and advection of moisture and heat in the atmospheric boundary layer, which in turn influences heat flux partitioning. G depends on soil thermal properties and the gradient of temperatures within soil layers. Generally, soil heat flux is downward during the day as deeper soil layers are heated from the surface, and upward at night when when deeper soil layers are heating the surface.

From spring to summer at both RC and SS sites, temperature increases, vegetation activity increases, and moisture availability decreases. As elevation increases along each transect, temperature decreases and precipitation increases. Here we identify influences to the partitioning between *LE*, *H*, and *G* at each site, and how this connectivity shifts within the growing season (Figure 5.2a).

We consider heat fluxes *LE*, *H*, and *G* as target nodes, and combinations of the following 8 variables as potential sources and source pairs: SW_{in}, LW_{in}, *Ta*, *WS*, *RH*, T_{surf}, *VWC*, and *Pa*. The resulting network for each window then consists of up to 28 potential source pairs that may provide 2 types of information (*R* and *S*), and 8 individual sources that provide *U* to 3 targets, which is equivalent to $(28 \times 2 \times 3) + (8 \times 3) = 192$ potential linkages.

We compare behaviors at each site between early (May-July) and late (August-October) growing season periods. May-July captures the end of the wet season and beginning of vegetation growth, while August-October captures the peak growing season and the driest part of the year. We hypothesize that each component of *LE*, *H*, and *G* should receive similar amounts of information from solar radiation variables, but otherwise respond to different sources. For example, we expect *LE* to receive more information from water-related variables (*RH*, *VWC*) that indicate the dependence of *LE* on water availability. Similarly, we expect *G* and *H* to be highly dependent on soil variables such as T_{surf} and *VWC*. Additionally, we expect water variables to exert more influence on heat flux variability in May, and for temperature variables to exert more influence in August.

For each of the sites at RC, the sum of *U*, *R*, and *S* information flow from the different sources to *LE* is higher from May-July than August-October (Figure 5.4a). Between the different sites along the transect, transfers to *LE* are slightly higher at RC_1 and RC_4 compared to RC_2 and RC_3. For *LE*, the higher information content from May-July for each of the sites can be linked to slightly higher precipitation (Figure 5.4c) and higher average soil moisture (*VWC*, Figure 5.4d) compared to August-October. The high total information to *LE* at RC_4 can be linked to higher accumulated precipitation before May and high soil water content (Figure 5.4c,d). However, similarly high information for RC_1 may be mainly related to a single large storm event that occurred in July that involved more rainfall than at RC_1 than at the other sites (Figure 5.4c), and the resulting large and lasting increase in *VWC*. This trend in decreasing information flow to *LE* as the dry season progresses, and high information flow as wetness increases indicates that the connectivity of various states to *LE* is constrained by available moisture.

In contrast to the behavior of information flow to *LE*, for information transfers to *G* and

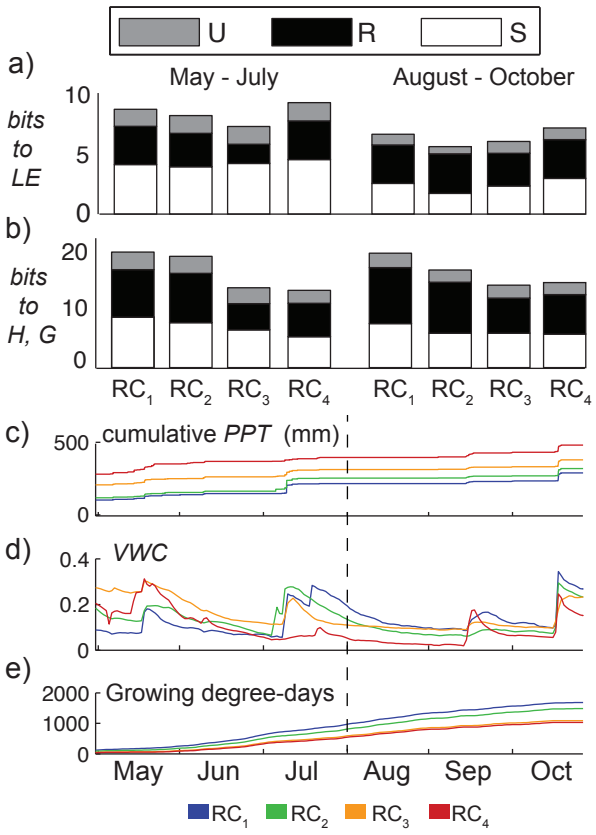

Figure 5.4: Total information transfers to LE, H, and G at Reynolds Creek during the 2015 growing season. *(a)* Total S, R, U transferred from atmospheric, soil, and radiation variables to LE at Reynolds Creek between May-July and Aug-Oct of 2015 for each of the four RC sites. *(b)* Total S, R, U transferred from atmospheric, soil, and radiation variables to H and G. *(c)* Cumulative precipitation (PPT) between May-Oct of 2015 for the four RC sites. *(d)* Volumetric water content (VWC) for the four sites. *(e)* Cumulative growing degree-days (GDD) for the four sites, computed for each day as $GDD = \frac{T_{max}-T_{min}}{2} - T_{base}$ where $T_{base} = 10°C$.

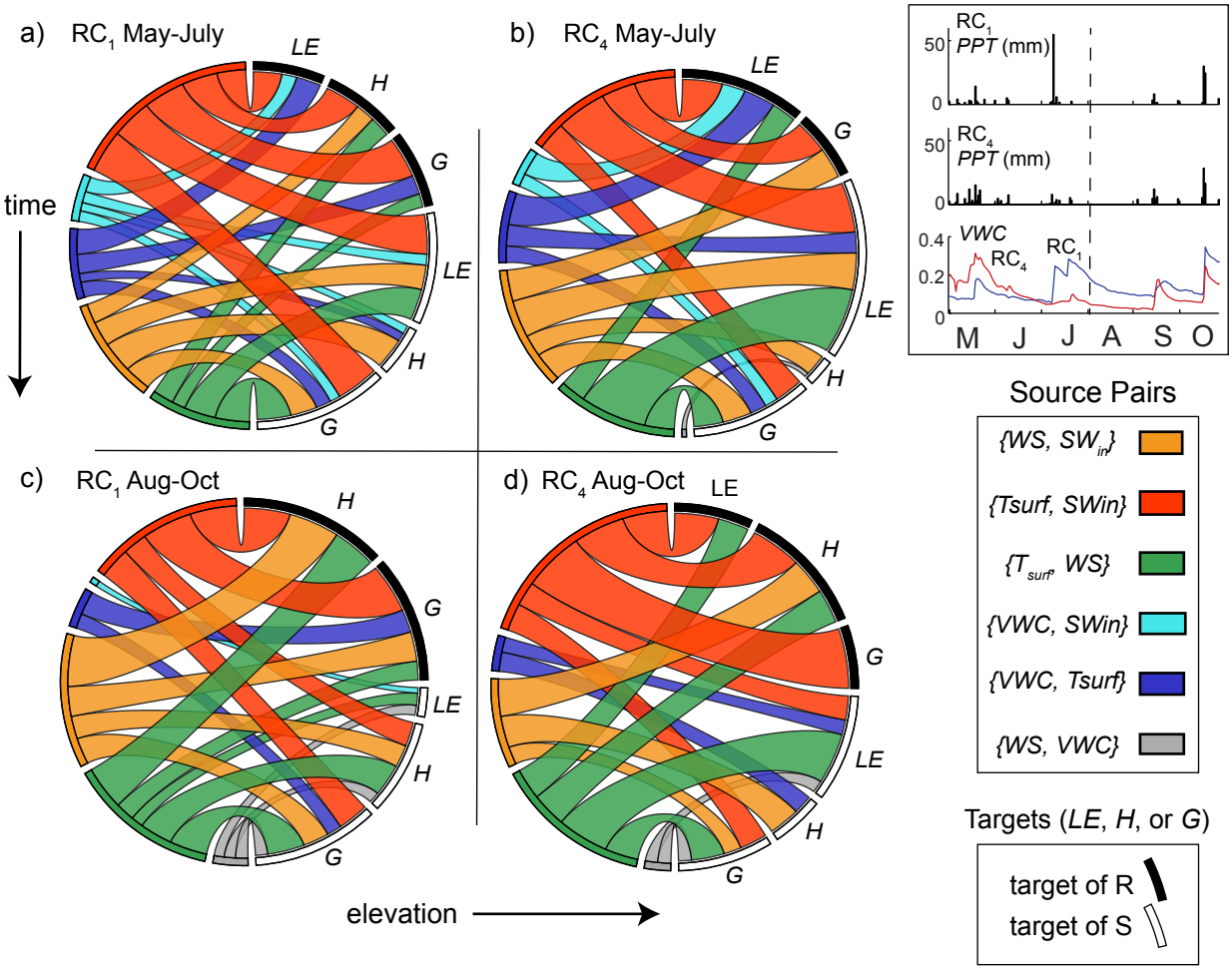

Figure 5.5: Network visualization of specific sources of information to LE, H, and G. We focus on specific transfers of R and S from pairs involving T_{surf}, VWC, SW_{in}, and WS to targets LE, H, and G for site *(a)* RC_1 during May-July, *(b)* RC_4 during May-July, *(c)* RC_1 during August-October, and *(d)* RC_4 during August-October. Colors indicate source pairs, and black and white indicate targets of R and S, respectively. *(inset)*: Daily precipitation and volumetric water content for each site between May and October. These network visualizations were developed using the Circos online tool [Krzywinski et al., 2009].

H, no trend is observed within the growing season, but total information flow decreases with elevation (Figure 5.4b). In particular, we observe a step change between sites RC_2 and RC_3, particularly in terms of R. This can be related to the accumulated growing degree-days ($GDD = \frac{T_{max}-Tmin}{2} - T_{base}$ where $T_{base} = 10°C$) over the year at each site (Figure 5.4e), where the lower elevation sites, RC_1 and RC_2, accumulate many more GDD than the higher sites, RC_3 and RC_4, through the period. This pattern where information flows to G and H correlate to temperatures over the season indicates that connectivity in this case is mediated by heat rather than moisture availability. This indicates that process connectivity between environmental variables and heat fluxes fundamentally differ when different heat fluxes are considered. While constrained LE for dry conditions reduces the amount of information it receives, information flows H and G are more dependent on the temperature at a given site.

To gage whether specific sources to LE, H, or G shift over the season or the elevation gradient, we must consider the network of individual source-pair to target interactions. To simplify this network of up to 192 potential links, we focus on only the source pairs that include SW_{in}, T_{surf}, VWC, and WS, so that there are only 8 possible source pairs. These source variables are chosen to represent solar radiation, temperature, moisture, and atmospheric conditions within the system, respectively, all of which are expected to influence heat flux partitioning. We compare the highest (RC_4, Figure 5.5b,d) and lowest (RC_1, Figure 5.5a,c) elevation sites during May-July and August-October of 2015.

There are several similarities in behaviors between the four cases, such as the relatively large flow of R from $\{T_{surf}, SW_{in}\}$ to G, and the tendency for $\{T_{surf}, WS\}$ to provide more S than R. The source pairs $\{T_{surf}, WS\}$, $\{WS, SW_{in}\}$, and $\{T_{surf}, SW_{in}\}$ are tend to be detected as strong influences, while pairs involving VWC are weaker. This can be attributed to the lower variability in VWC on the sub-daily time scales we consider here. In other words, while temperature, wind, and radiation fluctuate rapidly within a 10-day window, soil water content is likely to increase or decrease more gradually and only respond rapidly to rainfall events. However, particularly during May-July for both sites (Figure 5.5a,b), we see that $\{VWC, T_{surf}\}$ and $\{VWC, SW_{in}\}$ are significant sources of information. Since WS is typically a dominantly synergistic source due to its large variability on a fast time scale, its redundancy with SW_{in} in August indicates a close and lag-synchronized relationship between atmospheric mixing and radiation fluctuations in the late summer.

As discussed previously, we find that information flow to LE is related to seasonal moisture conditions, while information flow to H and G is associated with temperature differences. From a more detailed analysis of source pairs, we see that the decrease in information flow to

LE between seasons (Figure 5.5a,b versus 5.5c,d) is due to weaker links overall and a lack of detected information from pairs involving VWC. In comparison, the absence of R provided to H leads to the decrease in information flow to H and G in May-July between RC_1 and RC_4. In August-October, the decrease is mainly due to lower R provided to G (Figure 5.5d). From this we infer that while LE is constrained by available moisture, H and G are variably influenced by the level of constraint on LE in addition to differences in temperature with elevation.

5.4.2 Characterizing influences to carbon flux

While heat fluxes participate in forcing and feedback interactions with radiation, atmospheric, and soil components on similar time scales, we expect that carbon dynamics are forced by similar drivers at the sub-hourly to sub-daily timescales that we consider here, but participating in feedbacks at much longer timescales. For example, while temperature fluctuations result in variability in carbon fluxes (Fc) due to microbial or vegetation activity at a sub-daily time scale, changes in carbon fluxes influence temperature through longer-term or cumulative mechanisms such as the greenhouse effect.

Here we detect shifting influences to Fc over the course of the study period for each flux tower site. Specifically, we assess the dependency between SW_{in} and Fc, and the influence of intermediate sources LE, H, Ta, RH, H_2O, and WS from month to month at the different sites. Here we omit soil variables due to gaps in data for several sites during months of interest. While we expect that SW_{in} regularly drives variations in Fc, other factors may act as redundant or synergistic influences depending on conditions. For example, while vegetation is a sink of CO_2 during photosynthesis, soil is often a source of CO_2 due to microbial respiration, such that moisture and temperature variations can have an indirect but large effect on Fc in addition to the constraints imposed by energy availability.

For each month between March and October, we analyze the average strength of S and R information transfers between each of LE, H, Ta, RH, $H2O$, and WS along with SW_{in} to Fc. We separate time windows by month between March-October (approximately 6 overlapping 10-day time windows per month), and assess shifts in average behavior between months at each site. Here we hypothesize that influences to carbon dynamics shift according to the time of peak vegetation activity, which is lagged in time between the lowest and highest elevation sites. We also expect that a greater number of interactions occur early in the growing season as vegetation is establishing and is more responsive to variability.

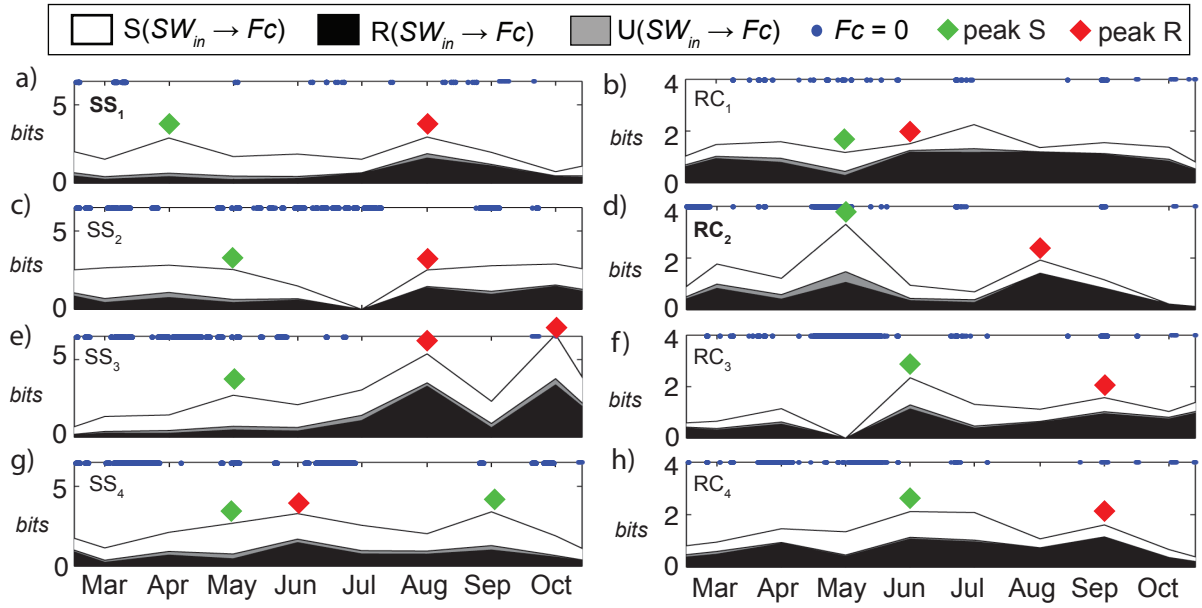

Figure 5.6: Information transfers to carbon flux (Fc) from solar radiation and other variables. Total strengths of S, R, and U indicated by bar width, proportions indicated by shading. For March to October, *(a,c,e,g)* Sierra sites SS_1-SS_4, and *(b,d,f,h)* Reynolds sites RC_1-RC_4. Blue dots indicate data points where $Fc = 0$, indicating potential gaps in the dataset.

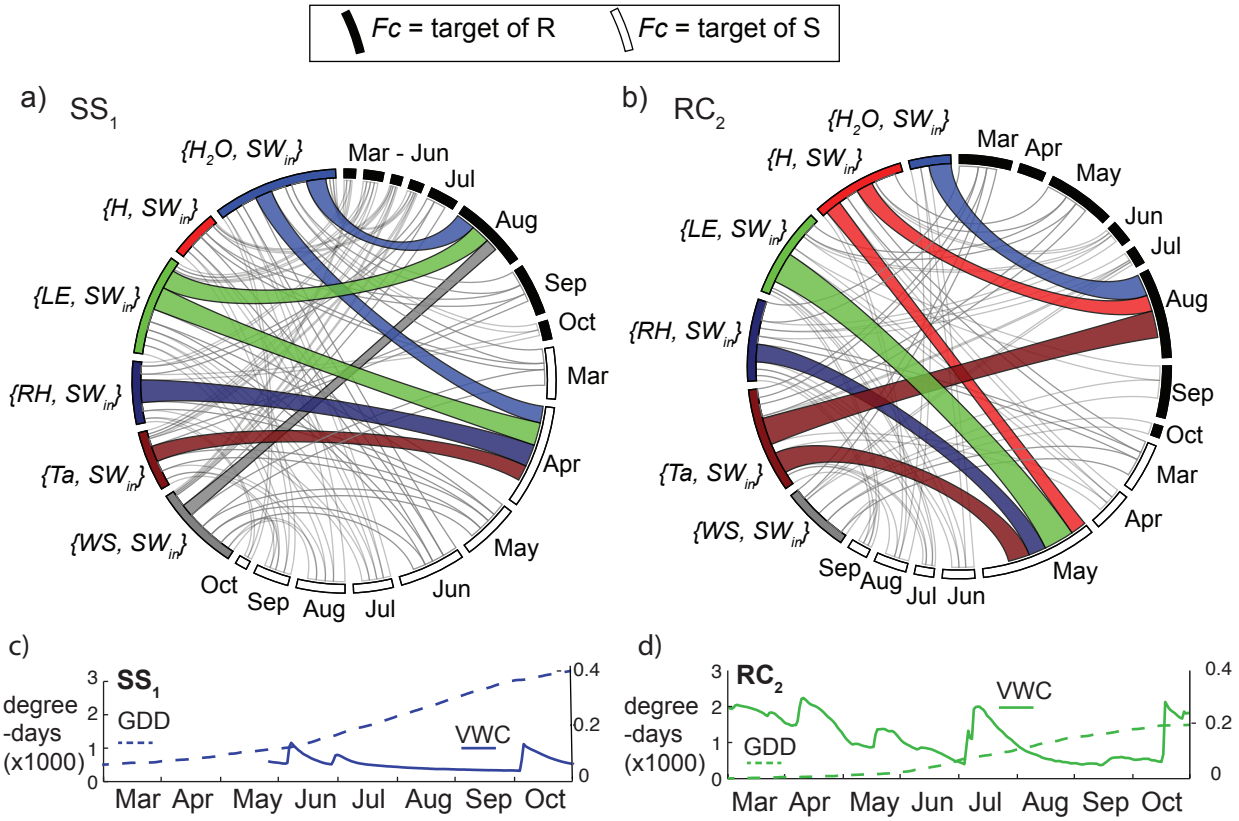

Figure 5.7: Network of information transfers to carbon flux (Fc) for two selected sites. Specific source pairs that include $SWin$ that transfer information to Fc between March and October. Links for April (for SS) or May (for RC) and August, corresponding to peaks in S and R, respectively, indicated by colors. *(a,b)* The high proportion of S during spring months is due to the influence of $\{LE,\ SW_{in}\}$, while heightened R later in the season is due to $\{LE,\ SW_{in}\}$ and $\{H_2O,SW_{in}\}$ at SS *(a)* and $\{H_2O,SW_{in}\}$ and $\{Ta,SW_{in}\}$ at RC *(b)*. *(c,d)* Cumulative growing degree-days (GDD) and daily averaged soil volumetric water content (VWC) for sites SS_1 and RC_2. These network visualizations were developed using the Circos web tool [Krzywinski et al., 2009].

When dependencies of all types are taken together, we find that the overall strength of influence from SW_{in} to Fc varies between months for each site (Figure 5.6). In the spring to early summer, total information strength peaks in April for SS_1 (Figure 5.6a), in May for SS_2 and SS_3 (Figure 5.6c,e), and in June for SS_4 (Figure 5.6g). This shift in spring-time peak information flow from April to June as elevation increases reflects the shifting start of the growing season along the elevation gradient. At RC sites, the early summer peak is the highest information flow from $SW_{in} \rightarrow Fc$ of the study period, except for RC_1 which exhibits a peak in July due to the large rainfall event during that month. Each site at SS exhibits a second peak later in the summer, in August for SS_1, October for SS_2, August and October and SS_3, and September for SS_4. Overall, total information flow from $SW_{in} \rightarrow Fc$ is higher at SS than at RC. This could indicate either that Fc is responding more to other variables at RC, or that its variability is overall lower.

At all 8 sites, the proportion of S in the $SW_{in} \rightarrow Fc$ link is high during the spring, and R becomes dominant during the summer and fall (black and white colors in Figure 5.6a-h). In comparison, the proportion of U from SW_{in} to Fc (gray in Figure 5.6) is always relatively low. The seasonal pattern of S and R indicates SW_{in} and the other drivers more independently influence Fc during the spring, and that drivers become more synchronized with SW_{in} during the summer such that they provide R. Upon closer inspection of the specific joint interactions at sites SS_1 (Figure 5.7a) and RC_2 (Figure 5.7b) behind this shift from high S in the spring to high R in the fall, we find that different source pairs are causal. We choose these two sites because they both exhibits peaks in S in the spring (April and May) and peaks in R during August.

At SS_1, the source pair $\{LE,\ SW_{in}\}$ is a mainly synergistic influence in April, May, and June, and redundant in August (Figure 5.7a). Similarly, $\{H_2O,\ SW_{in}\}$, $\{RH,\ SW_{in}\}$, and $\{WS,\ SW_{in}\}$ provide S between March-June, and R between July-September. In other words, the variables H_2O, RH, WS, and LE all provide information to Fc along with SW_{in} throughout the growing season, but they become more redundant with SW_{in} in July. This could indicate that as the summer progresses, SW_{in} exerts an influence such that all variables become more closely synchronized to SW_{in}. In contrast, during the spring as the wet season is ending, variables related to water availability and weather events (H_2O, RH, WS, LE) are more independently influencing carbon dynamics.

At RC_2 , $\{H,\ SW_{in}\}$ and $\{Ta,\ SW_{in}\}$ are relatively strong joint influences to Fc. $\{H,\ SW_{in}\}$ provides similar R and S during May, but R is greater than S during August. $\{Ta,\ SW_{in}\}$ provides higher S in May, and R in August. This is similar to the behavior of WS, H_2O, and

RH at SS_1, in that these variables appear to act more synergistically during the spring to influence Fc and more redundantly in the summer and fall. Additionally, the large proportion of S in May is mainly due to $\{LE, SW_{in}\}$, while the large proportion of R in August is mainly due to $\{Ta, SW_{in}\}$. This coincides with the finding in the previous analysis that LE receives more information earlier in the growing season. Due to the many feedbacks and in the system, a variable that is a strong source of information is also likely to be detected as a target.

In this analysis, we characterized the complex network of time dependencies that exist at individual sites as they relate to fluxes of carbon and energy. This enables us to infer dominant processes and their shifts through the season and compare between the different sites. However, it does not allow us to infer the dependencies between sites that represent physical forcing mechanisms that span multiple locations, such as moving weather fronts and atmospheric mixing. In the following section, we analyze the flow of local and non-local information between the neighboring sites at each transect.

5.5 Flow of information between sites

Previously, each flux tower site comprised a set of individual TIPNets, and network behaviors were compared between sites and time periods. Here we consider a single network that includes variables from multiple sites, namely the four flux towers that make up each transect. In this configuration, the set of nodes from each site forms a subsystem. A subsystem can participate in local interactions, which are transfers of information detected between nodes within the subsystem, in addition to non-local interactions, which are transfers of information to or from nodes at another site. There may also be joint local and non-local dependencies (e.g. a given target receives joint information from one source node locally and another at a different site), but we do not consider them here.

We construct a multi-site network from subsystems of 11 variables at each transect: WS, WD, H_2O, Ta, LE, H, G, T_{surf}, Pa, RH, and Fc. We omit radiation variables from this analysis on the assumption that they should be relatively synchronized between the sites, particularly at Reynolds at which sites are only several km apart. For this analysis we focus on unique transfers of information. While the presence of local U would indicate purely local forcing and feedbacks at a site, non-local U would indicate forcing emanating *uniquely* from neighboring locations. For this study, we take monthly time windows that do not overlap, and consider time lags τ between 30 minutes and 24 hours. We hypothesize that sites that are closer together geographically should exchange more information, but that local drivers

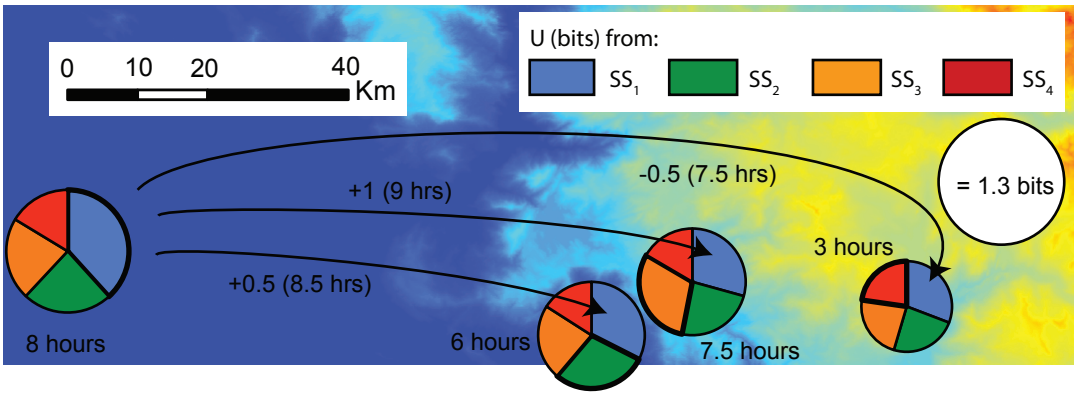

Figure 5.8: Unique information transfers and time scales within and between SS sites, as detected based on 11 time-series nodes at each site. The size of each circle represents the monthly averaged total received U (*bits*) for each target site, and the sections represent the proportion of information provided by variables from each site (SS_1 , SS_2 , SS_3 , or SS_4). Arrows indicate the average dominant time scales of detected interactions.

likely dominate dynamics such that the largest detections of U should be between variables at a single site. We also predict that timescales of interactions in the form of dominant τ should increase with distance, such that the most distant sites (e.g. SS_1 and SS_4) should exchange information at the longest timescales.

At SS, on average over all months, we find that SS_1 is the largest source and target of U overall (Figure 5.8). Variables in the SS_1 subsystem tend to provide U in proportion similar to or greater than local U transfers at SS_2 , SS_3 , and SS_4 (pie charts in Figure 5.8). While SS_1 subsystem nodes provide high U to sites towards the east, SS_4 nodes provide a much smaller proportion of U to sites towards the west. From this trend, we infer that the directionality of non-local U is from west to east, corresponding with predominant weather patterns moving into the mountains from the Pacific Ocean [Goulden et al., 2012]. Moreover, SS_1 appears to be the most "locally-driven" site, which can mainly be attributed to a lack of observations from any westward location. In terms of average dominant time scales, we find that local U transfers at SS_1, SS_2, and SS_3 range from 6 to 8 hours, and transfers between sites occur at longer timescales that increase with distance (Figure 5.8 arrows). For example, while SS_1 nodes interact with each other at a time delay of $\tau = 8$ hours, the same variables are sources of information to SS_2 at $\tau = 8.5$ hours and SS_3 at $\tau = 9$ hours. This trend differs for SS_4, which exhibits very fast ($\tau = 3$ hour) local connectivity. All non-local transfers to and from this site are detected at timescales longer than 3 hours but shorter than the 6-8

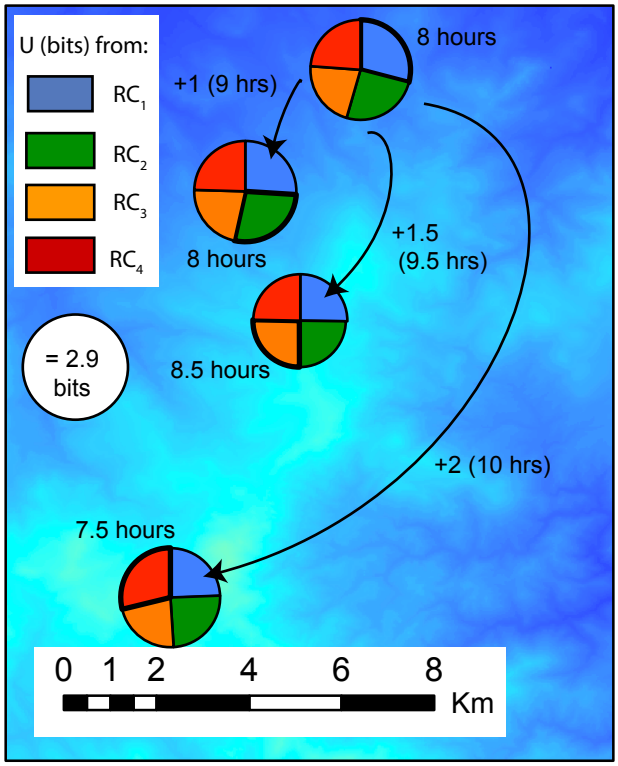

Figure 5.9: Unique information transfers and time scales within and between RC sites, as detected based on 11 time-series nodes at each site. Circle sizes represent the monthly averaged total received U $(bits)$ for each target site, and the sections represent the proportion of information provided by variables from each site. RC_1 is the largest receiver and provider of U overall but each site is dominated by local U.

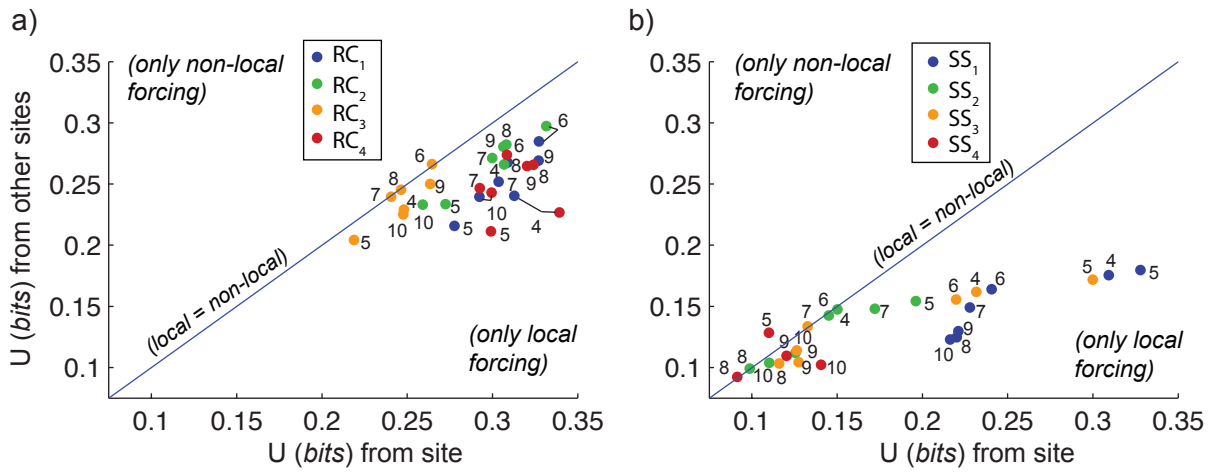

Figure 5.10: Local versus non-local transfers of unique information for each month at RC and SS. The within-site U (local U) and U received from other sites (non-local U) varies between months (Mar=3, Apr =4, May=5...) for each flux tower location at *(a)* Reynolds *(b)* Sierra. At both sites, local U dominates, but several mid-summer (6,7,8) time periods exhibit nearly equal local and non-local drivers. Local U at SS sites shows more spread between sites and throughout the year.

hour dominant timescales at the other sites.

Similar to sites along the SS transect, we find that RC_1, the lowest elevation site at RC, is the largest overall source and target of U, locally and non-locally (Figure 5.9). However, each site receives the most U locally, indicating less dominance in directionality from low to high elevation as observed at SS. Additionally, strengths of U (*bits*) are much greater at RC sites than at SS. Since RC sites are more closely spaced compared to SS, we infer that they have more potential to be strongly linked. Timescales of local U at RC range from $\tau = 7.5 - 8.5$ hours, and these timescales become longer as the distance between sites increases. Although sites at RC are much closer together than sites at SS, we note similar timescales of information flow between sites. At SS, this could indicate that the timescales between sites are associated with faster-moving weather fronts that are oriented more directly along the transect.

When we consider individual monthly time windows, we find contrasting behaviors in terms of local and non-local U (Figure 5.10). While most months exhibit a dominance of local U compared to non-local U (U averaged over other three sites), there are several periods, such as June-August at RC_3 (Figure 5.10a), May at SS_4 (Figure 5.10b), and June-August at SS_4 for which local and non-local U are approximately equal. This indicates that there

are periods where information flow due to directional weather forcing or atmospheric mixing provides more predictability than does information flow from variables at the site of interest.

In general, RC subsystems exhibit regularly higher U (*bits*) from both local and non-local sources compared to SS, and sites and months are similar to each other in terms of these characteristics (clustering in Figure 5.10a). At SS, overall U is lower than at RC (Figure 5.10b), and there is significantly more spread in local U between sites. For example, at SS_1, all months exhibit higher local U than do all months for SS_3 and SS_4. As mentioned previously, this is related to the location of SS_1 as the western site along a transect in which weather fronts move from west to east, such that SS_1 is a dominant source of U to itself and other sites, but receives low U from other sites.

5.6 Discussion

The analysis presented here improves our understanding of process connectivity within an ecohydrologic system as it exists both at a location and over a spatial gradient. While local drivers of fluctuations in energy and carbon fluxes vary in strength and type throughout a season, additional drivers of ecohydrologic variability may emanate from neighboring or distant locations. Locally, we observe a transition from synergistic to redundant intermediate drivers of carbon flux along with solar radiation from late spring to early summer. This indicates a shift during the growing season, in that multiple drivers that act independently early in the season later become more synchronized. At Southern Sierra, this can be linked to increasing carbon uptake in early spring until June or July, and decreasing uptake through late summer and fall. As atmospheric states and heat fluxes become more redundant with incoming solar radiation in mid-summer when vegetation activity (in terms of evapotranspiration) peaks, carbon fluxes begin to decrease in magnitude. We also see shifts between early and late summer in terms of influences to latent, sensible, and ground heat flux. While soil, atmospheric, and radiation variables provide less information to H and G at sites with cooler temperatures, they provide less information to LE as conditions become drier during the summer dry season. From this, we infer that environmental conditions constrain information flow to variables in different ways.

The concept of spatial information transfer enables us to gage the level of connectivity of processes between sites. While behaviors in regions like Southern Sierra display a dominant directional connectivity from west to east according to prevailing weather patterns, behaviors at Reynolds are subject to a higher proportion of local forcing and less directionality of

non-local forcing. This could relate to the orientation of the transect from north to south, or the proximity of the observation sites that causes sites to appear nearly equally connected. This type of study extended to a greater range of locations or specific processes may be constructive in diagnosing missing links between observations and modeled behaviors in terms of both dominant process links and spatial connectivity.

The patterns and behaviors determined using our process network approach aid in interpreting the connectivity between ecohydrologic behaviors within and between sites that could not be discerned from an analysis of individual or pairwise relationships. Given concerns regarding the impact of shifting temperatures and carbon concentrations on ecohydrologic dynamics, this type of approach works towards a holistic picture of dominant multivariate interactions and their shifts over climate gradients.

CHAPTER 6

CONCLUSIONS

This thesis focused on the characterization of process networks based on novel information theoretic measures. The suite of concepts, methods, and applications presented here help us understand forcing and feedback interactions that occur within complex ecohydrologic systems, and characterize system behaviors more holistically.

6.1 Complex networks and information flow

Multivariate information theoretic measures are useful to infer connectivity in networks where interactions may be non-linear and of varying strengths. In a synchronizing chaotic logistic network, for example, induced feedbacks cannot be distinguished from imposed drivers after a certain level of connectivity has been surpassed. Information measures such as lagged mutual information (I_τ), transfer entropy (T_E), and a new measure (T/I) are shown to be useful to distinguish between high and low levels of connectivity and between random, or "external", drivers and local, or "internal", drivers. The first proposed measure (T/I), indirectly introduces information partitioning in a network context since it is a measure of relative synergy and uniqueness from one of two sources relative to the total joint information provided by both sources. The second measure, Rescaled Redundancy (R_s), adds the capacity to directly distinguish redundant, synergistic, and unique shared information rather than infer them indirectly from combined measures. The methodological contributions of this thesis are (i) a framework with which to study complex systems of interacting components, (ii) a deeper understanding of information theory measures as they can be interpreted in a network context. The concepts and methods presented here are broadly applicable to a range of systems that can be studied as complex networks of interacting time-series variables.

Although the utility of information measures to interpret system dynamics and influences is shown, several pitfalls are noted in terms of *pdf* estimation issues, characteristics of ecohydrologic datasets, and potential misinterpretations of information theoretic measures as

"causal". As examples, a detected unique information flow from one variable to another may be interpreted as causal or "forcing", and some unexplained variability may be attributed to "noise". In both these cases, there may also be an unmeasured or missing aspect of the system, such as a non-local source of information flow or gaps in observations. If this missing aspect were included in the network, the previously detected uniqueness may be determined to be redundant, or the unexplained variability may be captured as a time-dependent link. In this way, temporal information partitioning in a network context allows us to reconstruct certain but not all system dynamics.

Besides the inherently incomplete nature of process network connectivity as detected from observed data, other issues arise from our limited approach where we consider sources only in a pairwise context (e.g. $I(X_{s1}, X_{s2}; X_{tar})$ as opposed to $I(X_{s1}, X_{s2}...X_{sn}; X_{tar})$ as the total information provided to a target). A more complete understanding of information flow would be gained from partitioning information from many sources to a target into multiple unique, redundant, and synergistic components [Williams and Beer, 2010]. However, here we limit ourselves to 3d pdfs to ensure somewhat robust estimations given few data points. Although we integrate our pairwise approach into a network context, we are not able to classify aspects such as synergy between more than two variables, or the uniqueness of a single link with regards to all other links.

Cognizant of these issues, this study introduces a network framework that facilitates the understanding of time dependencies as influences within the network associated with physical process connectivity. We explore theoretical examples, synthetic chaotic datasets, and ecohydrologic applications based on weather station and flux tower datasets.

6.2 Process connectivity in ecohydrologic systems

This thesis contributes several findings related to dynamics in ecohydrologic systems. While water-related properties such as relative humidity and latent heat flux are constrained in variability or total information content by moisture availability, they otherwise participate in a variety of forcing and feedback interactions. Interactions between variables are detected at a range of timescales, and information flow exists between spatial locations as well as at a site, indicating the presence of non-local connectivity due to weather and mesoscale forcing.

For the construction of ecohydrologic process networks, multiple nodes are considered as targets and sources of unique, synergistic, and redundant information at different time scales. The TIPNet framework is applied to a high temporal resolution weather station dataset

and along a flux tower transect dataset to understand how different aspects of the process network shift between time windows, sites, and through different climatic conditions. From the analysis of 1-minute weather station data, it is found that time dependencies are detected at timescales faster than the usual 30-minute or hourly temporal resolution of observed data, and strengths of interactions shift quickly as conditions change. While a wet or rainy day-time period is characterized by a highly complex and synergistic network of interactions, a dry period exhibits many fewer linkages and high redundancy between air temperature and solar radiation.

The use of eddy covariance flux towers on flux tower transects along two elevation gradients enables the study of (*i*) differences in process connectivity due to climate and vegetation regime, and (*ii*) the extent of information flow that occurs over a spatial gradient. The application of the TIPNet framework to study a wider range of processes and include the aspect of local and non-local information flows demonstrates the utility of the approach for various types of applications. When information transfers to heat and carbon fluxes are considered at individual sites in California and Idaho, a shift from dominant synergy in the spring to redundancy in the fall is detected for several joint interactions. This increasing redundancy between driving processes during the dry season indicates an increased level of synchronization between radiation and other variables as moisture becomes limited and less influential as an independent driver. Across the spatial gradient, the dominant direction of information flow at the Southern Sierra, California site is coherent with typical weather forcing, while directionality is much weaker at Reynolds Creek, Idaho.

6.3 Avenues for future research

This thesis includes several applications to ecohydrologic time-series data that reveal a deeper layer of process interactions than would be detected based on other methods. The framework proposed here and related findings suggest possibilities for further research in hydrology and other fields, such as the following:

- The study of species resilience to shifting dynamics under climatic variability and human influence. For example, populations of species such as mussels have internal population dynamics in addition to responses to feedbacks and forcing with variables such as sediment concentration and flow [Hansen et al., 2016, Foufoula-Georgiou et al., 2015].
- An analysis of local and non-local interactions at a site to determine vulnerability to

perturbations such as land use change at different locations.

- Diagnosing the differences between models and observed data in terms of multivariate interactions. For example, it has been shown that sources of uncertainty result from different aspects of models such as parameters, forcing data, or model structure [Nearing et al., 2016].

The value of our proposed framework is exhibited here in combination with existing knowledge of processes. For example, the the known constraints imposed by the energy and water balances, the relationship between temperature and atmospheric water vapor, and influences of directional weather patterns and atmospheric mixing are all linked to detected network behaviors. However, even in a system where the functional forms of relationships are completely unknown, we can still gain an understanding about levels of connectivity, time scales, synchronization, and joint interactions. This perspective of connectivity between variables in a network of forcing and feedback contributes to a holistic interpretation of a system and potentially to the prediction of large-scale emergent properties such as ecosystem resilience.

REFERENCES

[Aguirre et al., 2014] Aguirre, J., Sevilla-Escoboza, R., Gutierrez, R., Papo, D., and Buldu, J. M. (2014). Synchronization of interconnected networks: The role of connector nodes. *Physical Review Letters*, 112.

[Albert and Barabasi, 2002] Albert, R. and Barabasi, A. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97.

[Alizad-Rahvar and Ardakani, 2012] Alizad-Rahvar, A. and Ardakani, M. (2012). Finding weak directional coupling in multiscale time series. *Physical Review E*, 86.

[Atay et al., 2004] Atay, F., Jost, J., and Wende, A. (2004). Delays, connection topology, and synchronization of coupled chaotic maps. *Physical Review Letters*, 92.

[Balasis et al., 2009] Balasis, G., Daglis, I. A., Papadimitriou, C., Kalimeri, M., Anastasiadis, A., and Eftaxias, K. (2009). Investigating dynamical complexity in the magnetosphere using various entropy measures. *Journal of Geophysical Research: Space Physics*, 114(6):1–13.

[Balasis et al., 2013] Balasis, G., Donner, R. V., Potirakis, S. M., Runge, J., Papadimitriou, C., Daglis, I. a., Eftaxias, K., and Kurths, J. (2013). Statistical Mechanics and Information-Theoretic Perspectives on Complexity in the Earth System. *Entropy*, 15(11):4844–4888.

[Barrett and Seth, 2011] Barrett, A. and Seth, A. (2011). Practical measures of integrated information for time-series data. *PLOS Computational Biology*, 7(1).

[Barrett, 2015] Barrett, A. B. (2015). Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. *Physical Review E*, 91(5).

[Bell, 2003] Bell, A. (2003). The co-information lattice. In Bell, A., Wickerhauser, M., and Szu, H., editors, *Independent Component Analyses, Wavelets, and Neural Networks*, volume 5102 of *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)*, pages 383–388. Chapman & Hall, London-New York.

[Bertschinger et al., 2014] Bertschinger, N., Rauh, J., Olbrich, E., Jost, J., and Ay, N. (2014). Quantifying unique information. *Entropy*, 16(4):2161–2183.

[Boba et al., 2015] Boba, P., Bollmann, D., Schoepe, D., Wester, N., Wiesel, J., and Hamacher, K. (2015). Efficient computation and statistical assessment of transfer entropy. *Computational Physics*, 3:1–9.

[Boers et al., 2013] Boers, N., Bookhagen, B., Marwan, N., Kurths, J., and Marengo, J. (2013). Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System. *Geophysical Research Letters*, 40(16):4386–4392.

[Bollt, 2012] Bollt, E. M. (2012). Synchronization as a process of sharing and transferring information. *International Journal of Bifurcation and Chaos*, 22.

[Cakan et al., 2014] Cakan, C., Lehnert, J., and Scholl, E. (2014). Heterogeneous delays in neural networks. *The European Physical Journal B*, 87(3).

[Cover and Thomas, 2006] Cover, T. and Thomas, J. (2006). *Elements of Information Theory 2nd Edition*. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, 2 edition.

[Deza et al., 2015] Deza, J. I., Barreiro, M., and Masoller, C. (2015). Assessing the direction of climate interactions by means of complex networks and information theoretic tools. *Chaos*, 25(3).

[Duan et al., 2013] Duan, P., Yang, F., Chen, T., and Shah, S. (2013). Direct causality detection via the transfer entropy approach. *IEEE Transactions on Control Systems Technology*, 21(6).

[Eichler, 2012] Eichler, M. (2012). Graphical modelling of multivariate time series. *Probability Theory and Related Fields*, 153(1-2):233–268.

[Erramuzpe et al., 2015] Erramuzpe, A., Ortega, G. J., Pastor, J., de Sola, R. G., Marinazzo, D., Stramaglia, S., and Cortes, J. M. (2015). Identification of redundant and synergetic circuits in triplets of electrophysiological data. *Journal of neural engineering*, 12(6):066007.

[Faes et al., 2015] Faes, L., Porta, A., and Nollo, G. (2015). Redundant and synergistic information transfer in cardiovascular and cardiorespiratory variability. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 4033–4036. IEEE.

[Feng et al., 2015] Feng, X., Porporato, A., and Rodriguez-Iturbe, I. (2015). Stochastic soil water balance under seasonal climates. *Proceedings of the Royal Society. Mathematical, physical, and engineering sciences*, 471(2174):20140623.

[Ferri et al., 2012] Ferri, G. L., Figliola, A., and Rosso, O. A. (2012). Tsallis' statistics in the variability of El Niño/Southern Oscillation during the Holocene epoch. *Physica A: Statistical Mechanics and its Applications*, 391(5):2154–2162.

[Foufoula-Georgiou et al., 2015] Foufoula-Georgiou, E., Takbiri, Z., Czuba, J. A., and Schwenk, J. (2015). The change of nature and the nature of change: Hydrologic regime shifts modulate ecological transitions. *Water Resources Research*, 51(1):6649–6671.

[Frenzel and Pompe, 2007] Frenzel, S. and Pompe, B. (2007). Partial mutual information for coupling analysis of multivariate time series. *Physical Review Letters*, 99(20):1–4.

[Friedman et al., 2008] Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441.

[Gong et al., 2014] Gong, W., Yang, D., Gupta, H. V., and Nearing, G. S. (2014). Estimating information entropy for hydrological data: One-dimensional case. *Water Resources Research*, 50(6):5003–5018.

[Goodwell and Kumar, 2015] Goodwell, A. and Kumar, P. (2015). Information Theoretic Measures to Infer Feedback Dynamics in Coupled Logistic Networks. *Entropy*, 17(11):7468–7492.

[Goodwell and Kumar, 2017a] Goodwell, A. and Kumar, P. (2017a). Temporal Information Partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *in review at Water Resources Research.*

[Goodwell and Kumar, 2017b] Goodwell, A. and Kumar, P. (2017b). Temporal Information Partitioning Networks (TIPNets): A process network approach to infer eco-hydrologic shifts. *in review at Water Resources Research.*

[Goulden et al., 2012] Goulden, M. L., Anderson, R. G., Bales, R. C., Kelly, A. E., Meadows, M., and Winston, G. C. (2012). Evapotranspiration along an elevation gradient in California's Sierra Nevada. *Journal of Geophysical Research*, 117(1):1–13.

[Griffith and Ho, 2015] Griffith, V. and Ho, T. (2015). Quantifying redundant information in predicting a target random variable. *Entropy*, 17(7):4644–4653.

[Griffith and Koch, 2014] Griffith, V. and Koch, C. (2014). Quantifying synergistic mutual information. In *Guided Self-Organization: Inception*, pages 159–190. Springer.

[Hansen et al., 2016] Hansen, A. T., Czuba, J. A., Schwenk, J., Longjas, A., Danesh-Yazdi, M., Hornbach, D. J., and Foufoula-Georgiou, E. (2016). Coupling freshwater mussel ecology and river dynamics using a simplified dynamic interaction model. *Freshwater Science*, 35(1):200–215.

[Hanson et al., 2001] Hanson, C. L., Marks, D., and Van Vactor, S. S. (2001). Long-term climate database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research*, 37(11):2839–2841.

[Harder et al., 2013] Harder, M., Salge, C., and Polani, D. (2013). Bivariate measure of redundant information. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(1):1–14.

[Hlaváčková-Schindler et al., 2007] Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., and Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports*, 441(1):1–46.

[Holland, 2000] Holland, J. H. (2000). *Emergence: From chaos to order*. OUP Oxford.

[Holme and Saramäki, 2012] Holme, P. and Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3):97–125.

[James et al., 2016] James, R. G., Barnett, N., and Crutchfield, J. P. (2016). Information Flows? A Critique of Transfer Entropies. *Physical Review Letters*, 116(23):1–5.

[Jenerette et al., 2012] Jenerette, G. D., Barron-Gafford, G. a., Guswa, A. J., McDonnell, J. J., and Villegas, J. C. (2012). Organization of complexity in water limited ecohydrology. *Ecohydrology*, 5(2):184–199.

[Jørgensen et al., 1992] Jørgensen, S. E., Patten, B. C., and Straškraba, M. (1992). Ecosystems emerging: toward an ecology of complex systems in a complex future. *Ecological Modelling*, 62(1-3):1–27.

[Krzywinski et al., 2009] Krzywinski, M. I., Schein, J. E., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., and Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*.

[Kumar, 2007] Kumar, P. (2007). Variability, feedback, and cooperative process dynamics: elements of a unifying hydrologic theory. *Geography Compass*, 1(6):1338–1360.

[Kumar and Ruddell, 2010] Kumar, P. and Ruddell, B. L. (2010). Information driven ecohydrologic self-organization. *Entropy*, 12(10):2085–2096.

[Lazo and Rathie, 2006] Lazo, A. V. and Rathie, P. (2006). On the entropy of continuous probability distributions (corresp.). *IEEE Trans. Inf. Theor.*, 24(1):120–122.

[Lee et al., 2012] Lee, J., Nemati, S., Silva, I., Edwards, B. A., Butler, J. P., and Malhotra, A. (2012). Transfer entropy estimation and directional coupling change detection in biomedical time series. *Biomedical Engineering Online*, 11.

[Marti et al., 2008] Marti, A. C., Ponce, M., and Masoller, C. (2008). Dynamics of delayed-coupled chaotic logistic maps: Influence of network topology, connectivity and delay times. *Pramana-Journal of Physics*, 70(6):1117–1125.

[Masoller and Atay, 2011] Masoller, C. and Atay, F. M. (2011). Complex transitions to synchronization in delay-coupled networks of logistic maps. *European Physical Journal D*, 62(1).

[Mitchell et al., 2004] Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. a., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., Tarpley, J. D., Lettenmaier, D. P., Marshall, C. H., Entin, J. K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B. H., and Bailey, A. a. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, 109(D7).

[Nayak et al., 2010] Nayak, A., Marks, D., Chandler, D. G., and Seyfried, M. (2010). Long-term snow, climate, and streamflow trends at the reynolds creek experimental watershed, Owyhee Mountains, Idaho, United States. *Water Resources Research*, 46(6):1–15.

[Nearing et al., 2016] Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V., and Xia, Y. (2016). Benchmarking NLDAS-2 soil moisture and evapotranspiration to separate uncertainty contributions. *Journal of Hydrometeorology*, (2013):160113112628008.

[Niso et al., 2013] Niso, G., Bruna, R., Pereda, E., Gutierrez, R., Bajo, R., Maestu, F., and del Pozo, F. (2013). Hermes: Towards an integrated toolbox to characterize functional and effective brain connectivity. *Neuroinformatics*, 11.

[Olbrich et al., 2015] Olbrich, E., Bertschinger, N., and Rauh, J. (2015). Information decomposition and synergy. *Entropy*, pages 3501–3517.

[Paredes et al., 2013] Paredes, G., Alvarez-Llamoza, O., and Cosenza, M. G. (2013). Global interactions, information flow, and chaos synchronization. *Physical Review E*, 88.

[Rheinwalt et al., 2016] Rheinwalt, A., Boers, N., Marwan, N., Kurths, J., Hoffmann, P., Gerstengarbe, F. W., and Werner, P. (2016). Non-linear time series analysis of precipitation events using regional climate networks for Germany. *Climate Dynamics*, 46(3-4):1065–1074.

[Rosas et al., 2016] Rosas, F., Ntranos, V., Ellison, C., Pollin, S., and Verhelst, M. (2016). Understanding Interdependency Through Complex Information Sharing. *Entropy*, 18(2):38.

[Rosenblum et al., 1997] Rosenblum, M., Pikovsky, A., and Kurths, J. (1997). From phase to lag synchronization in coupled chaotic oscillators. *Physical Review Letters*, 78(22):4193–4196.

[Ruddell and Kumar, 2009a] Ruddell, B. L. and Kumar, P. (2009a). Ecohydrologic process networks: 1. identification. *Water Resources Research*, 45.

[Ruddell and Kumar, 2009b] Ruddell, B. L. and Kumar, P. (2009b). Ecohydrologic process networks: 2. analysis and characterization. *Water Resources Research*, 45.

[Runge, 2015] Runge, J. (2015). Quantifying information transfer and mediation along causal pathways in complex systems. *Physical Review E*, 92(6):062829.

[Runge et al., 2012] Runge, J., Heitzig, J., Petoukhov, V., and Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Physical Review Letters*, 108(25):1–5.

[Saco et al., 2010] Saco, P. M., Carpi, L. C., Figliola, A., Serrano, E., and Rosso, O. A. (2010). Entropy analysis of the dynamics of El Niño/Southern Oscillation during the Holocene. *Physica A: Statistical Mechanics and its Applications*, 389(21):5022–5027.

[Schreiber, 2000] Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2):461.

[Sendrowski and Passalacqua, 2017] Sendrowski, A. and Passalacqua, P. (2017). Process Connectivity in a Naturally Prograding River Delta. *Water Resources Research*, 53.

[Seyfried et al., 2011] Seyfried, M., Chandler, D., and Marks, D. (2011). Long-Term Soil Water Trends across a 1000-m Elevation Gradient. *Vadose Zone Journal*, 10(4):1276.

[Shannon, 1948] Shannon, C. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(I):379–423.

[Sharma and Mehrotra, 2014] Sharma, A. and Mehrotra, R. (2014). An information theoretic alternative to model a natural system using observational information alone. *Water Resources Research*, 50(1):650–660.

[Silverman, 1986] Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, volume 26. CRC press.

[Smirnov, 2013] Smirnov, D. a. (2013). Spurious causalities with transfer entropy. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(4):1–12.

[Stramaglia et al., 2012] Stramaglia, S., Wu, G. R., Pellicoro, M., and Marinazzo, D. (2012). Expanding the transfer entropy to identify information subgraphs in complex systems. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 066211:3668–3671.

[Strogatz, 2001] Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268–276.

[Sugihara et al., 2012] Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106):496–500.

[Sun and Bollt, 2014] Sun, J. and Bollt, E. M. (2014). Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings. *Physica D: Nonlinear Phenomena*, 267:49–57.

[Timme et al., 2014] Timme, N., Alford, W., Flecker, B., and Beggs, J. M. (2014). Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. *Journal of Computational Neuroscience*, 36(2):119–140.

[Vejmelka and Paluš, 2008] Vejmelka, M. and Paluš, M. (2008). Inferring the directionality of coupling with conditional mutual information. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 77(2):1–12.

[Vlachos and Kugiumtzis, 2010] Vlachos, I. and Kugiumtzis, D. (2010). Nonuniform state-space reconstruction and coupling detection. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 82(1):1–16.

[Wibral et al., 2014] Wibral, M., Vicente, R., and Lindner, M. (2014). Transfer entropy in neuroscience. In Wibral, M., Vicente, R., and Lizier, J., editors, *Directed Information Measures in Neuroscience*, Understanding Complex Systems, Springer Complexity, pages 3–36.

[Wiedermann et al., 2016] Wiedermann, M., Radebach, A., Donges, J. F., Kurths, J., and Donner, R. V. (2016). A climate network-based index to discriminate different types of El Niño and La Niña. *Geophysical Research Letters*, 43(13):7176–7185.

[Williams and Beer, 2010] Williams, P. L. and Beer, R. D. (2010). Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515*.

[Williams and Beer, 2011] Williams, P. L. and Beer, R. D. (2011). Generalized Measures of Information Transfer. *ArXiv e-prints*.

APPENDIX A

SUPPLEMENTARY INFORMATION

A.1 Information measures for theoretical example cases

A.1.1 Gaussian sum information measures

The entropy and mutual information measures for V_1, V_2, and X_{tar} are easily derived and given as follows (in units of *nats* due to the natural logarithm):

$$H(V_1) = \frac{1}{2}\ln(2\pi e(\sigma_{x_{s1}}^2 + \sigma_{n1}^2)) \tag{A.1}$$

$$H(V_2) = \frac{1}{2}\ln(2\pi e(\sigma_{x_{s2}}^2 + \sigma_{n2}^2)) \tag{A.2}$$

$$I(V_1; V_2) = I(X_{s1}; X_{s2}) = -\frac{1}{2}\ln(1-\rho^2) \tag{A.3}$$

$$H(X_{tar}) = \frac{1}{2}\ln(2\pi e(\sigma_{x_{s1}}^2 + \sigma_{x_{s2}}^2 + \rho\sigma_{x_{s1}}\sigma_{x_{s2}})) \tag{A.4}$$

$$\begin{aligned} I(V_1; X_{tar}) &= I(X_{s1}; X_{tar}) \\ &= \frac{1}{2}\ln\left(\frac{\sigma_{x_{s1}}^2 + \sigma_{x_{s2}}^2 + \rho\sigma_{x_{s1}}\sigma_{x_{s2}}}{\sigma_{x_{s2}}^2(1-\rho^2)}\right). \end{aligned} \tag{A.5}$$

$$\begin{aligned} I(V_2; X_{tar}) &= I(X_{s2}; X_{tar}) \\ &= \frac{1}{2}\ln\left(\frac{\sigma_{x_{s1}}^2 + \sigma_{x_{s2}}^2 + \rho\sigma_{x_{s1}}\sigma_{x_{s2}}}{\sigma_{x_{s1}}^2(1-\rho^2)}\right). \end{aligned} \tag{A.6}$$

A.1.2 Multivariate autoregressive example cases

[Barrett, 2015] derives the following information quantities from the covariance matrix of the MVAR system that depend only on α (Figure 3.8a):

$$I(X_{tar}(t-1); X_{tar}(t)) = \frac{1}{2}\ln\left(\frac{1}{1-\alpha^2}\right) \tag{A.7}$$

$$I(X_{s1}(t-1); X_{tar}(t)) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1+\alpha^4}\right) \tag{A.8}$$

$$I(X_{tar}(t-1), X_{s1}(t-1); X_{tar}(t)) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2}\right) \tag{A.9}$$

A.2 Estimation of *pdf*s from data

To illustrate *pdf* estimation that accounts for typical issues encountered in environmental signals, we use an example dataset of 1-minute resolution solar radiation (Rg) and air temperature (Ta) recorded from a weather station during DOY 193 of 2015 (Figure A.1). The site is a restored prairie, surrounded by corn and soybean agricultural fields and the riparian corridor of the Sangamon River. A weather station has been operated through the Intensively Managed Landscape Critical Zone Observatory (IML-CZO) since June 2014. A Decagon PYR Solar Radiation Sensor collects Rg data, a Decagon VP-4 collects Ta and RH, and a Decagon Davis Cup Anemometer collects WS. In the example provided in this Appendix for illustration of *pdf* estimation issues, we consider Rg as a source variable that provides information to Ta. We seek to detect the lagged dependency at an arbitrarily chosen $\tau = 10$ minutes between the variables ($I(Rg(t-\tau); Ta(t))$) given as few as $n = 200$ data points. In Section 4 of this paper, we consider Ta and WS as sources to RH at time lags τ between 1 and 10 minutes.

A.2.1 KDE and fixed-bin *pdf* estimation methods

We use Kernel Density Estimation (KDE) [Silverman, 1986, Lee et al., 2012] in order to compute robust *pdf*s given as few as 200 data points [Goodwell and Kumar, 2015]. We employ the multivariate Epanechnikov kernel, for which an optimal window width is chosen based on the number of data points, dimension of the *pdf*, and variance of the data as described in [Silverman, 1986]. While KDE requires more computations than fixed-binning (FB) in addition to the tuning of a smoothing parameter, it is advantageous for high-dimensional *pdf*s based on sparse data since it becomes independent of bin size above a certain number of bins. For the two segments of DOY 193 (Figure A.1b), we see that

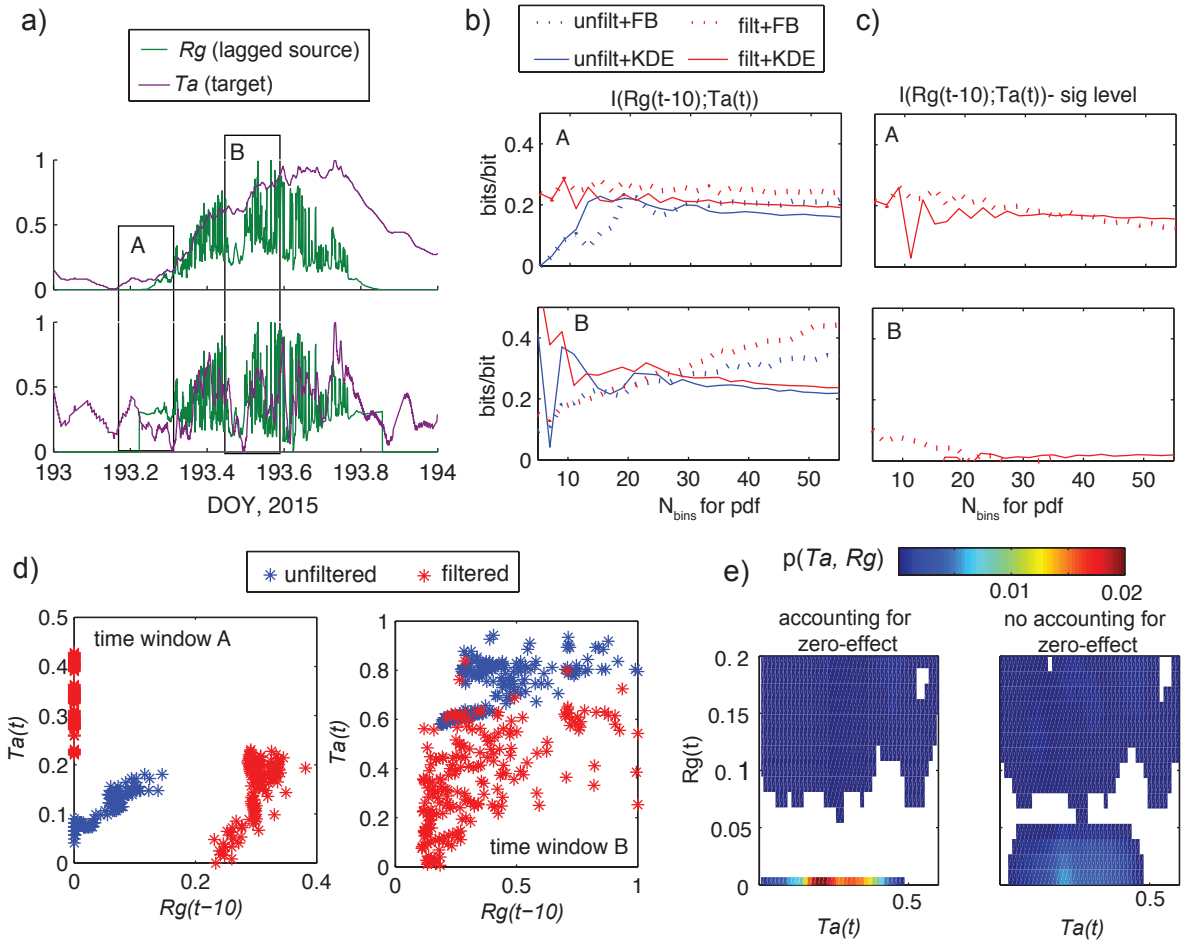

Figure A.1: *Pdf* estimation issues illustrated with weather station data. *(a)* 1-minute time-series of Rg and Ta over DOY 193 in 2015 as measured at the SFP weather station, normalized to scale between 0 and 1 by minimum and maximum values (top), and filtered and normalized (bottom). *(b)* Shared information $I(Rg(t-10); Ta(t))$ depends on the number of *pdf* bins N_{bins}, *pdf* estimation method (KDE or FB), and whether the data are filtered or unfiltered. *(c)* Statistical significance testing eliminates non-robust detections of shared information. Fixed-binning (FB) causes measure to be detected with decreasing statistical significance as N_{bins} increases, while KDE leads to more stable detected measures since the method becomes independent of N_{bins} for many bins. *(d)* Data from segments A and B in *(a,b)* shows differences between filtered and unfiltered data. *(e)* Accounting for the atom-at-zero effect (left) improves *pdf* estimation for hybrid discrete-continuous data with zero values, while otherwise zero-values are smoothed by the KDE method (right).

both data filtering and *pdf* estimation methods influence the detected shared information $I(Rg(t-10);Ta(t))$. Particularly for the afternoon segment, we see that while the KDE method results in stabilizing information detection as the number of intervals N_{bins} increases, FB leads to increased information quantities with higher N_{bins}.

A.2.2 Accounting for the atom-at-zero effect

Environmental signals such as rainfall, ephemeral streamflow, or shortwave solar radiation often have many zero values which can skew a *pdf* and alter detections of shared information. [Gong et al., 2014] discusses a solution to this for a 1d case which involves considering a *pdf* as a mixed (discrete-continuous) distribution. If we define the bin intervals with $i = 1...N_{bins}$, a 1d *pdf* can be defined as follows:

$$\begin{cases} p(x_i) = 1 - k_x, & \text{if } i = 1. \\ p(x_i) = k_x f(x_i) h_i, & \text{otherwise.} \end{cases} \tag{A.10}$$

in which k_x is the proportion of non-zero values in the variable X, h_i is the width of bin i, and $f(x_i)$ represents the *pdf* resulting from any estimation method. We expand this formulation of [Gong et al., 2014] to the 2d and 3d case. For example, a 2d equivalent is as follows:

$$\begin{cases} p(x_i, y_j) = p_{00}, & \text{if } i = 1, j = 1 \\ p(x_i, y_j) = p_{x0} f_{x0}(x_i) h_i, & \text{if } i > 1, j = 1 \\ p(x_i, y_j) = p_{0y} f_{0y}(y_j) h_j, & \text{if } i = 1, j > 1 \\ p(x_i, y_j) = p_{xy} f_{xy}(x_i, y_j) h_i h_j, & \text{if } i > 1, j > 1 \end{cases} \tag{A.11}$$

in which $f_{x0}(x)$ and $f_{0y}(y)$ indicate marginal *pdf*s computed for X and Y values for which corresponding Y and X are zero-values, respectively, and the coefficients p_{xx} indicate the proportion of data points in each category obtained via direct count from the data points. A 3d *pdf* is constructed with a similar expansion of Equation A.10 with marginal *pdfs* determined for 1d cases (e.g. $i = 1$, $j = 1$, and $k > 1$) and 2d cases (e.g. $i = 1$, $j > 1$, and $k > 1$, in addition to the 3d *pdf* estimated for data points with no zero values.

Figure A.1e shows an inset of the computed KDE 2d *pdf* $p(Rg(t), Ta(t))$ for all of DOY 193 for cases that do *(left)* and do not *(right)* account for the fact that many *Rg* values are zero. We see that while zero-effect accounting categorizes all $Rg = 0$ values into a single bin of negligible width, a blind method smooths these values over a range of small *Rg* values.

A.2.3 Statistical significance of shared information

While pre-processing and *pdf* estimation methods influence the detection of shared information, any detected value of information (denoted as $I(X_s; X_{tar})_d$) must be associated with some level of statistical significance. We employ a shuffled surrogates method [Ruddell and Kumar, 2009a, Goodwell and Kumar, 2015] in which the source X_s is shuffled to destroy time correlations and shared information $I(X_{shuff}; X_{tar})$ is computed. This is repeated for $N_{tests} = 100$ iterations, and $\bar{x}$ and s_x are the mean and standard deviations of the shuffled values, respectively. We define the statistical significance of $I(X_s; X_{tar})$ as follows:

$$I(X_s; X_{tar}) = \begin{cases} I(X_s; X_{tar})_d, & \text{if } I(X_s; X_{tar})_d > \bar{x} + cs_x. \\ 0, & \text{otherwise.} \end{cases} \tag{A.12}$$

for which we set $c = 3$ to obtain a 99% significance level. Since all computed values of shared information are non-zero, this test for statistical significance eliminates many detections of shared information that are not robust. While this test for statistical significance may omit true links in some cases, lowering the significance threshold may introduce "false" links. In this way, a balance must be reached between excluding weak but physical relationships and retaining falsely detected relationships. Values of mutual information from shuffled significance tests were found to be generally normally distributed based on chaotic logistic generated cases.

We perform information partitioning only if significant shared information is detected between at least two sources and a target. We do not compute further statistical significance tests for information partitioning measures, since together they must sum to equal the total shared information according to Equation (4.3). Statistical significance testing could be performed for the 3d measure of total mutual information, and may result in the omission of some of these relationships due to the more sparse *pdf*. Due to this, we assume that if the individual mutual information values between two sources and a target is statically significant based on 2d measures, their joint information is also relevant to study.

A.3 Filtering and pre-processing

The diurnal cycle is a dominant pattern for many environmental variables, and the resulting synchronization is often much stronger than other types of interactions between variables

that occur on faster timescales. For example, air temperature (Ta) is tightly lag-synchronized with Rg as Earth's surface gains and loses heat over the course of a day. This is indeed the case for DOY 193 (Figure A.1a, top panel), where synchronization results in a nearly linear relationship between the two variables. To capture interactions besides those related to the diurnal cycle that may be weaker or occur at different timescales, we must filter out the diurnal cycle to enhance their detection. For example, instead of the synchronization of Ta and Rg through the day, we may seek lagged responses of Ta to fluctuations in Rg due to changing cloud cover or wetness condition. Variables that exhibit a strong diurnal and seasonal cycle (Rg, RH, and Ta) are filtered using a high-pass Butterworth filter to retain short term fluctuations while omitting these cycles (Figure A.1a, bottom panel). This omits synchronization forced by the diurnal cycle, but retains and enhances the detection of other dependencies. Rg is filtered, then night-time periods (when orginally $Rg = 0$) are reset to zero values to omit spurious fluctuations. Figure A.1d shows the filtered and unfiltered data points for two chosen subsets of DOY 193 (A and B in FigureA.1a). We see that the filtering generally leads to more spread in the values since it enhances small fluctuations, and weakens the linear trend.

Statistical outliers (except for PPT and $LWet$ which are often zero) below the value of $X_{min} = P_{25} - 1.5\text{IQR}$, where P_i is the i^{th} percentile, and IQR is the interquartile range, are set to the value of X_{min}. Similarly, outliers above $X_{max} = P_{75} + 1.5\text{IQR}$ are set to the value of X_{max}. Finally, all nodes are normalized based on their updated minimum and maximum values to lie in the $[0, 1]$ range.

A.4 Threshold behavior

While examples in Chapter 3 involve two sources that drive a target via the same function throughout the time period of interest, information partitioning may also be applied to time windows during which forcing mechanisms are shifting. If the data cannot be partitioned based on these shifts, information partitioning reflects the combination of mechanisms that act at different times. Over a time period that include such combinations, R_s has an advantage over R_{MMI} of identifying unique contributing sources of information. In a simple example, variables X_{s1} and X_{s2} are uniformly distributed and uncorrelated variables for a time window of length $n_{steps} = 1000$. For this time window, let us assume that one source drives X_{tar} for the first half of the period and the other drives for the second half as follows:

$$X_{tar}(t) = \begin{cases} f(X_{s1}(t)), & \text{if } t < 500. \\ f(X_{s2}(t)), & \text{otherwise.} \end{cases} \tag{A.13}$$

Examples of this switching or threshold behavior from an eco-hydrologic context include intermittent rainfall, evaporation, heating, and cooling that may occur over short periods of time. Over longer time periods, this could represent shifts in forcing due to land use change, other human influences, or climate. Based on R_{MMI}, a dataset generated from Equation (A.13) would result in the detection of only R, since X_{s1} and X_{s2} provide the same quantity of total information over the entire period. In contrast, R_s would estimate equal amounts of U from each source variable. In this way, $U > 0$ may indicate both that the sources are not correlated and that there is a switching in driving mechanisms during the time window of consideration.

A.5 Weather station data

Table A.1: List of Sangamon Forest Preserve weather station instruments.

Variable	Units	Instrument	resolution	range
RH	%	Decagon VP-4	1%	[0 100%]
Ta	$^{\circ}C$	Decagon VP-4	0.1 deg	[-40 80]
WS	*m/s*	Decagon Davis Cup Anemometer	.45 m/s	[0, 1.3]
WD	degrees from N	Decagon Davis Cup Anemometer	1 deg	[0 360]
Rg	*W/m2*	Decagon PYR Solar Radiation Sensor	N/A	[0 1750]
PPT	*mm*	Decagon ECRN-100 Rain Gage	0.2 mm	[0]
LWet	raw counts	Decagon LWS	1	[455 1400]

The Decagon weather station at the Sangamon Forest Preserve collects radiation (*Rg*, W/m^2), wind speed (*WS*, m/s), precipitation (*PPT*, *mm*), relative humidity (*RH*), air temperature (*Ta*, ^{o}C), and leaf wetness (*LWet*, dielectric counts) at 1-minute temporal resolution. The raw *LWet* variable consists of dielectric counts on the synthetic leaf sensor surface, where higher values (> 1000) typically represent very wet conditions due to rainfall, low values (< 450) indicate dry surface conditions, and intermediate values indicate the presence of dew or frost. Instrumentation details are located in Table A.1 below.

APPENDIX B

TIPNET MATLAB SOFTWARE

As part of this thesis, a Graphical User Interface (GUI) has been developed for the TIPNet measures used in Chapters 2-5. The program takes inputs of time-series datasets as "nodes" in a network, and computes information measures to identify and characterize time dependencies between nodes. The latest version of this GUI and more comprehensive user manual can be found at *https://github.com/HydroComplexity/TIPNet*.

B.1 Quick start

Run the file called EntropyGUI_mainwindow.m. Click **Load New Data** option, and load either a .mat or .xls file containing columns of time series data. A completed .mat project file can also be loaded for immediate viewing of results in the **Load Project file** option. For a .xls file, variable names should be the top row of the file and the first column should be a time step. A .mat file should include a (no. variables x no. timesteps) matrix called "data" and a (1 x no. variables) cell called *varnames* with variable names. A vector called "timestep" is optional (for future version). For any processing or *pdf* options, see the appropriate section. To compute a single network using all data with all default options, click on Compute Links. All results are stored in the *entropy* structure that is saved in the project file. Results can then be viewed by clicking Plot Results.

B.2 Information measures

B.2.1 Entropy and mutual information

$$H(X) = -\sum p(x)\log_2(p(x)) \tag{B.1}$$

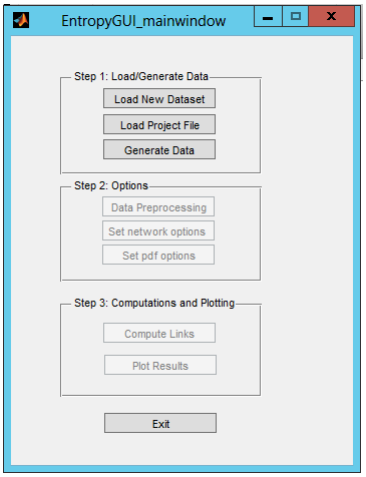

Figure B.1: Main screen of GUI. One of first three buttons must be chosen to load data or project file, or generate test data.

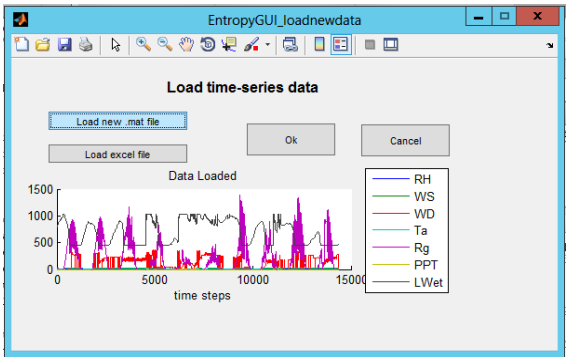

Figure B.2: Example of loading a weather station data set as a .mat file.

$$I(X;Y) = H(Y) - H(Y|X) = \sum p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right) \tag{B.2}$$

where X and Y are time-series variables that may be simultaneous or involve some time lag between them. When we consider X to be a "source" node and Y to be a "target" node, the quantity $I(X;Y)$ indicates the strength of a link from X to Y in that X reduces the uncertainty of the Y. For a range of lag times τ, $I(X(t-\tau);Y)$ is computed. Transfer Entropy $T_E(X(t-\tau) \to Y)$, which is equivalent to the conditional information $I(X(t-\tau);Y|Y(t-1))$ is also computed as follows:

$$I(X;Y|Y_1) = \sum_{x,y,y_\tau} p(x,y,y_1) \log \left[\frac{p(x,y,y_1)}{p(y,y_1)} \right] \tag{B.3}$$

where abbreviated symbols are $x = x(t-\tau)$, $y = y(t)$, and $y_1 = y(t-1)$. As discussed in [Goodwell and Kumar, 2015], T_E omits a redundant component (overlapping information shared to target $Y(t)$ by both $X(t-\tau)$ and $Y(t-1)$) but adds in a synergistic component (information shared to the target $Y(t)$ due to knowledge of both sources together). The dominant time scale of the link from X to Y is the $\tau > 0$ corresponding either to the maximum $I(X(t-\tau);Y)$ (bits) or the normalized value $\frac{I(X(t-\tau);Y)}{min(H(X),H(Y))}$ (bits/bit), depending on the **mi.NormOpt** parameter (see next section).

B.2.2 *Pdf* estimation and statistical significance

Computation of these measures involves estimating joint probability density functions (*pdf*) for lagged X and Y. We employ a fixed bin method [Ruddell and Kumar, 2009a, Lee et al., 2012] or a Kernel Density Estimation method [Lee et al., 2012, Silverman, 1986] to estimate *pdf*s from data. While the fixed binning method tends to be faster, the *KDE* method can be advantageous for sparse data sets since it smooths the *pdf* based on the sample size. For any detected $I(X;Y)$ value, we test for statistical significance using a shuffled-surrogate hypothesis test in which the time-series data are shuffled randomly to destroy any time correlations. Mutual information is then computed for $N_{tests} = 100$ (default) surrogates of shuffled data, and a 99% significance test is performed to assess whether the computed measure is significantly stronger than links detected from the shuffled surrogates [Goodwell and Kumar, 2015, Ruddell and Kumar, 2009a].

B.2.3 Information partitioning measures

Once the dominant links are detected based on lagged mutual information, we further assess each link in terms of its uniqueness, synergy, or redundancy by analyzing its relationship with other links to the same target. As introduced in [Williams and Beer, 2010] and discussed in [Goodwell and Kumar, 2015, Barrett, 2015, Goodwell and Kumar, 2017a], the total information shared between 2 source nodes X_1 and X_2 to a target Y can be partitioned into four components as follows:

$$I(X_1, X_2; Y) = U_1(Y; X_1) + U_2(Y; X_2) + R(Y; X_1, X_2) + S(Y; X_1, X_2) \tag{B.4}$$

where U_1, U_2, R, and S are non-negative quantities. R is information that both sources share with the target *redundantly*, U_1 and U_2 are information that only X_1 and X_2, respectively share with the target *uniquely*, and S is information that is provided to the target only when both sources are known together, or *synergistically*. Individual mutual information terms decompose as [Williams and Beer, 2010]:

$$I(Y; X_1) = U_1 + R \tag{B.5}$$

$$I(Y; X_2) = U_2 + R. \tag{B.6}$$

The proposed redundancy measure R_{MMI} [Williams and Beer, 2010, Barrett, 2015] is actually an upper bound for redundant information:

$$R_{MMI} = \min[I(X_1; Y), I(X_2; Y)] \tag{B.7}$$

The minimum bound of redundant information is as follows [Goodwell and Kumar, 2017a]:

$$R_{\min} = \max[0, I(X_1; Y) + I(X_2; Y) - I(X_1, X_2; Y)] \tag{B.8}$$

We implement a scaled version of R:

$$R = R_{\min} + I_s(R_{MMI} - R_{\min}) \tag{B.9}$$

where $I_s = \frac{I(X_1;X_2)}{\min[H(X_1),H(X_2)]}$ is the scaled source dependency, so that independent sources X_1 and X_2 result in minimum redundancy and highly dependent sources result in maximum redundancy. After R is computed for a given two sources to a target, the quantities U_1, U_2, and S can be computed directly. For a network of multiple interacting nodes, we consider each pair of sources to a target and evaluate the redundancy, uniqueness and synergy of each source pair.

We compute the measure T/I as follows [Goodwell and Kumar, 2015]:

$$\frac{T}{I}(X_{s1}|X_{s2} \to X_{tar}) = \frac{U_{s1} + S_{s1,s2}}{U_{s1} + U_{s2} + S_{s1,s2} + R_{s1,s2}} = \frac{I(X_{tar}; X_{s1}|X_{s2})}{I(X_{tar}; X_{s1}, X_{s2})} \tag{B.10}$$

For each source link X_{s1}, we define $T/I(X_{s1} \to X_{tar})$ as the minimum value of Equation

(B.10) given any other source node X_{s2} as follows:

$$\frac{T}{I}(X_{s1} \to X_{tar}) = \min_{X_{s2}} \left[\frac{T}{I}(X_{s1}|X_{s2} \to X_{tar}) \right] \tag{B.11}$$

In this way, T/I for a source to target link indicates the relative uniqueness and synergy of that link with respect to each other source to the same target.

We apply Equation B.9 to compute U_1, U_2, R, and S components for every pair of sources to a target. Similarly to T/I, we define the components for each link as follows:

$$R(X_{s1} \to X_{tar}) = \max_{X_{s2}} [R(X_{s1}, X_{s2}; X_{tar})] \tag{B.12}$$

$$U(X_{s1} \to X_{tar}) = \min_{X_{s2}} [U(X_{s1}, X_{s2}; X_{tar})] \tag{B.13}$$

$$S(X_{s1} \to X_{tar}) = \max_{X_{s2}} [S(X_{s1}, X_{s2}; X_{tar})] \tag{B.14}$$

The matrices $R(X, Y)$ and $R_pair(X, Y) = Z$ in the **entropy** results structure identifies the redundancy (in bits) that source node X shares with target Y along with source node Z. The source nodes X and Z are the most highly redundant sources to Y. Similarly, $S(X, Y)$ and $S_pair(X, Y) = W$ in the *entropy* results structure identifies the synergistic information (in bits) that source node X shares with target Y along with source node W, and X and W are the most strongly synergistic sources to Y.

B.3 Guide

B.3.1 Getting started

Important! First Time Use Only If you choose to use the KDE method for pdf computations, you must compile 3 C-mex files in matlab as follows: Go the the Functions folder, then type in the command line *mex -mdKDE_1d.c*. If an error occurs, you may need to choose a C compiler. Do the same operation for *mdKDE_2d.c* and *mdKDE_3D.c*. This only needs to be done the first time you use the program.

Run the file called *EntropyGUI_mainwindow.m*. Click **Load New Data** option, and load either a .mat or .xls file containing columns of numeric time series data. Examples of .mat files and .xls files are provided in the folder projects_datasets. For a .xls file, variable names

should be the top row of the file. A .mat file must include a (# variables x # timesteps) matrix called *data* and a (1 x # variables) cell called *varnames* with variable names. Once a data set is loaded, click **OK** to save the file as a project file. This project file will contain the **mi** (**m**odel **i**nformation) structure with all default parameters to run the temporal network program. When parameters are altered in the **pre-processing**, **network option**, or **pdf options**, they are updated in the **mi** structure in the project file. To reset all parameters to their default values, load the data as a new data set. To re-load a project file with any parameters that have been previously altered from default values, choose the load project option on the main screen.

B.3.2 Generating test data

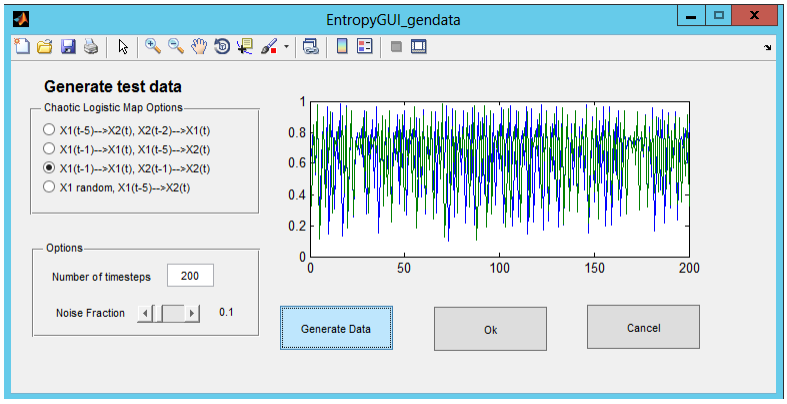

Figure B.3: Example of generated data in GUI. Chaotic logistic test data with 0.1 noise (random component).

Alternatively to loading a time series data set, the **Generate Data** option generates a 2-node chaotic logistic time series data set for one of four different forcing cases:

1. Feedback forcing, where $X1$ and $X2$ drive each other:

$$X2(t) = 4X1(t-5)[1 - X1(t-5)] \tag{B.15}$$

$$X1(t) = 4X2(t-2)[1 - X2(t-2)] \tag{B.16}$$

2. $X1$ drives itself via the chaotic logistic equation and also drives $X2$:

$$X2(t) = 4X1(t-5)[1 - X1(t-5)] \tag{B.17}$$
$$X1(t) = 4X1(t-1)[1 - X1(t-1)] \tag{B.18}$$

3. $X1$ and $X2$ are independent, each driven by the chaotic logistic equation:

$$X2(t) = 4X2(t-1)[1 - X2(t-1)] \tag{B.19}$$
$$X1(t) = 4X1(t-1)[1 - X1(t-1)] \tag{B.20}$$

4. $X1$ is a uniform random variable, and drives $X2$ through the chaotic logistic equation:

$$X2(t) = 4X1(t-5)[1 - X1(t-5)] \tag{B.21}$$
$$X1(t) = U(0,1). \tag{B.22}$$

For any case, the noise fraction slider bar for $0 \leq \epsilon_z \leq 1$ can be altered to add a degree of randomness into every node. For example, $\epsilon_z = 1$ generates 2 independent uniform random nodes.

B.3.3 Options

After loading a project file, new data file, or generated test data, there are three buttons to alter network parameters and properties from default values. These options include *pdf* estimation methods, network run options, and time-series pre-processing.

B.3.4 Pre-processing options

Each timeseries variable X is automatically normalized between (0,1) as follows:

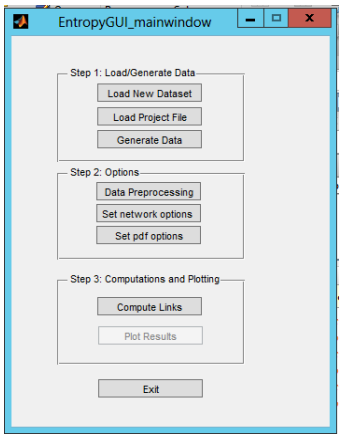

Figure B.4: Main TIPNet screen after loading data. A dataset, project file, or generated data must be loaded before continuing with options.

$$X_norm = \frac{(X - X_{min})}{X_{max} - X_{min}} \tag{B.23}$$

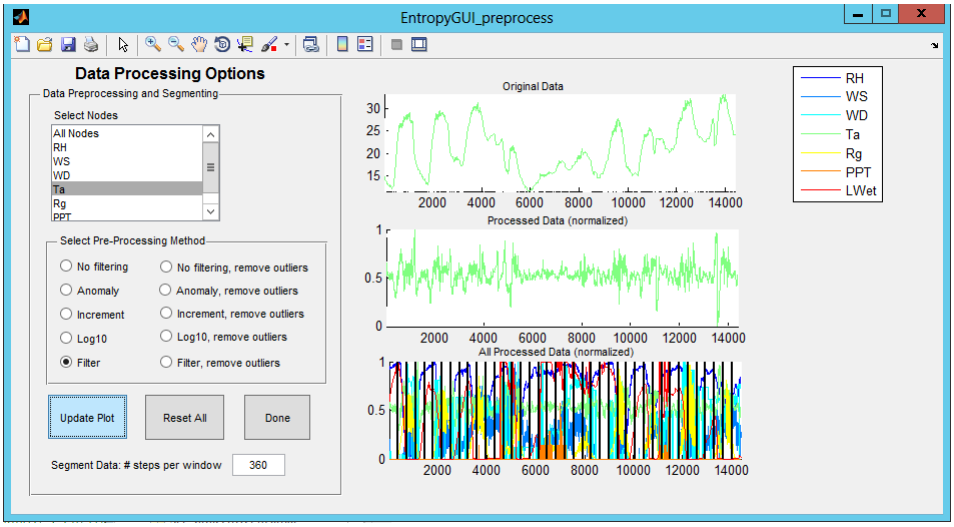

Figure B.5: TIPNet data preprocessing screen. Weather station data here has been segmented into 360 minute (6 hour) time segments as shown by black lines in bottom plot. Nodes can be pre-processed individually or as a group.

Then, there are 5 types of data filtering or altering. For each type, there is an option to remove or not remove outliers.

No Filtering This option reverts the data to the original normalized data set.

Anomaly For data that exhibit diurnal or seasonal cycle, the X-day anomaly is the difference between the value at a certain time (e.g. 12:00 noon on Day 100) and the mean value at that time on the X surrounding days (e.g. 12:00 noon on Days 95-105 for a 10-day anomaly). The anomaly can only be computed for 1 variable at a time, and the user must check on the time step and units of the data (minutes, days) and units of the desired anomaly (days, years). The anomaly of the originally loaded data is then normalized to a (0,1) range.

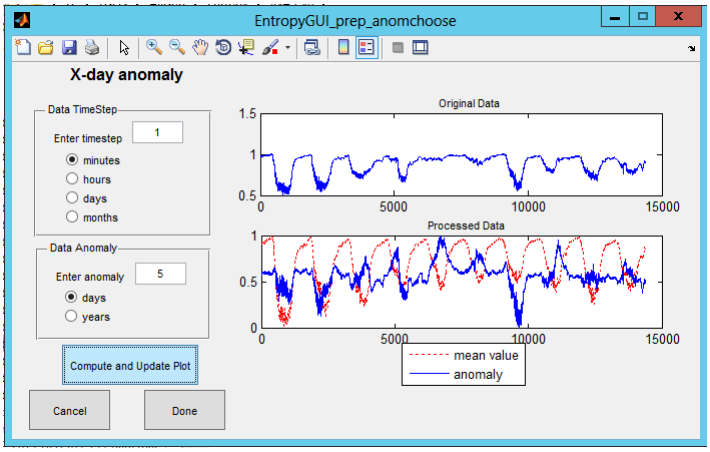

Figure B.6: Example of a 5-day anomaly applied as filtering method, for relative humidity data.

Increment For data where an increase or decrease may be more relevant than an actual value (e.g. a population variable). This changes the data as follows

$$X(t) = X(t) - X(t-1) \tag{B.24}$$

Log 10 : This takes the base 10 logarithm for skewed input data (e.g. flow rate data)

Filter For a single variable at a time, this option applies a Butterworth filter to the data for a high-pass or low-pass filter to preserve or omit short-term fluctuations. This can be used to (a) omit the diurnal and/or seasonal cycle with a high-pass filter (b) omit noise with a low-pass filter.

For each option, outlier removal is performed after the operation (e.g. after taking the logarithm or increment). Outliers, data points that lie above $X_{75} + 1.5IQR$ or below $X_{25} - 1.5IQR$, are set to the values $X_{75} + 1.5IQR$ or $X_{25} - 1.5IQR$, respectively rather

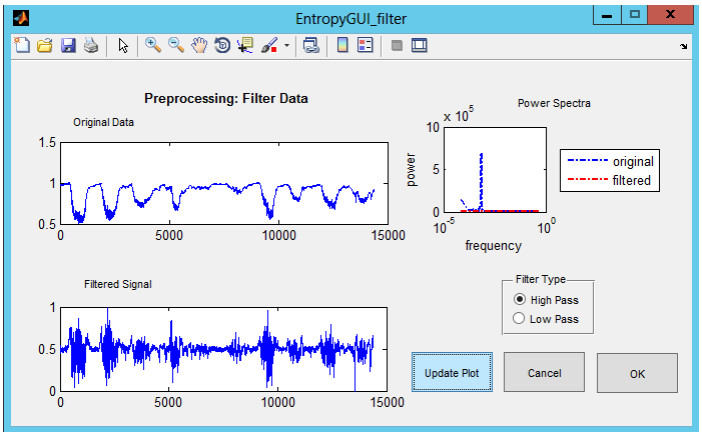

Figure B.7: High pass Butterworth filter applied to relative humidity data.

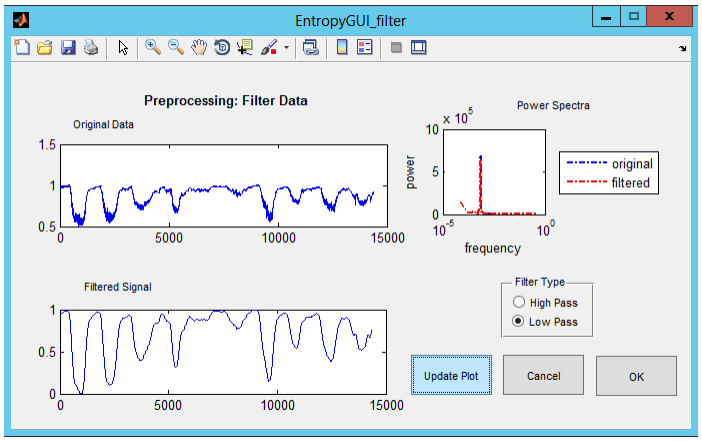

Figure B.8: Low pass Butterworth filter applied to relative humidity data.

than being removed. Removal of outliers would impact the time dependencies by removing a time-step of the specified variable. Any outlier removal via gap-filling or other methods should be done prior to loading a dataset.

Finally, to partition a long time-series data sets into multiple segments, the segment length can be changed. This option results in computation of one network for each time-series segments, and is useful to compare before-after scenarios or to consider the evolution over time of interactions.

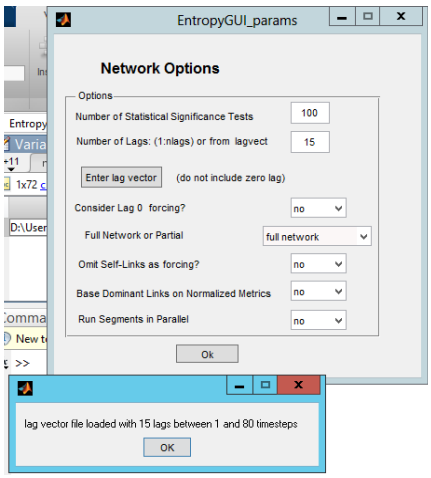

Figure B.9: TIPNet Network Options screen. When a *lagvect.mat* file is chosen, a verification message appears if the file is properly loaded. Alternatively, the number of lags can simply be entered in the text box for consecutive lag times.

B.3.5 Network options

The Network Options screen contains several options:

Statistical Sig Tests: The shuffled surrogates method is used to determine statistical significance of each computed $I(X_1; X_2)$ value. The default number of significance tests is 100.

Number of Lags: The number of lags for which lagged information measures are to be computed as $\tau = 1...$nlags.

Enter Lag Vector: Alternatively to specifying a number of consecutive lags, load a .mat file called lagvect.mat with a vector of lags named lagvect, containing lags. This can be used to compute lags at intervals, for example lagvect = [5 10 15 30 60 120] to compute network measures at only 6 time lags but for different lag times than 1-6. A *lagvect.mat* file is provided in the folder *UserData*, and should be overwritten as needed. The lag vector shoudl consist of non-negative integers, should not include zero (see next point).

Lag Zero Forcing: By default, zero-lag or instantaneous mutual information is not considered as a dominant link that can be redundant, synergistic, or unique with any other link. To include zero-lag forcing (e.g. if the time step is such that X may be expected to drive Y at a time scale much lower than the time step), change this option to *Yes*.

Network Run Option: By default, the program will perform all computations for mutual information, transfer entropy, and information decomposition as described in the previous section. To only compute individual node entropy or mutual information, change this option as appropriate. Note: the **Plot Results** viewer will not function if this option is altered.

Omit Self-Links: By default, node X is considered as a potential source to itself, and a detected link $I(X(t-\tau);X(t))$ may be unique, synergistic, or redundant when another link to X is considered. To omit these "self" links, change this option to *Yes*.

Run Segments in Parallel: If your data set is segmented into multiple time series in the **Pre-Processing Options** and your computer can run parallel code in Matlab (parfor loops), enable this to run segments in parallel.

B.3.6 PDF options

All information measures computed in this program are based on 1D, 2D, and 3D *pdf*s. This screen allows you to view these *pdf*s and alter parameters.

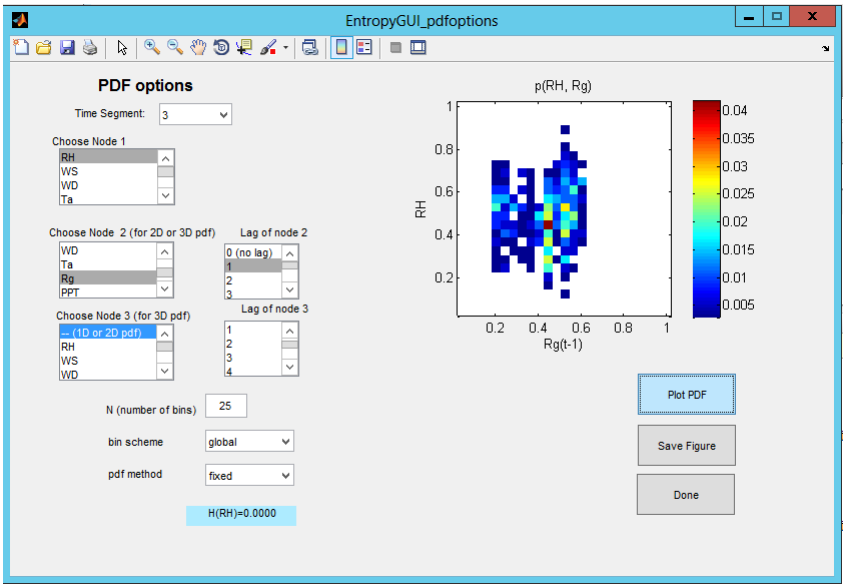

Figure B.10: TIPNet example 2d *pdf* for a segment using fixed bin method. *Pdf* of RH and lagged Ta for a specific segment using global binning, fixed bins and $N = 25$. Red color indicates higher value of $p(RH, Ta(t-1))$.

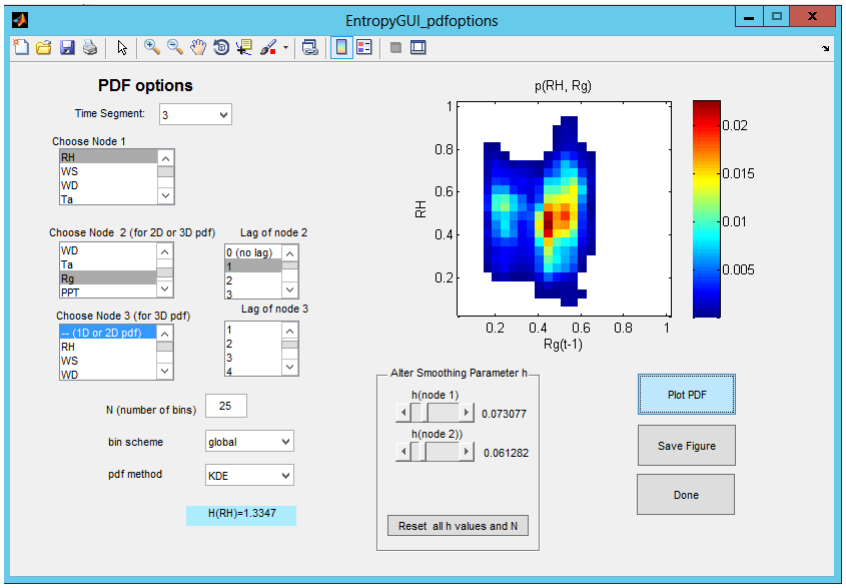

Figure B.11: TIPNet example 2d *pdf* using KDE method. *Pdf* of RH and lagged Ta for a specific segment using global binning, KDE and increasing h smoothing parameters slightly for both nodes.

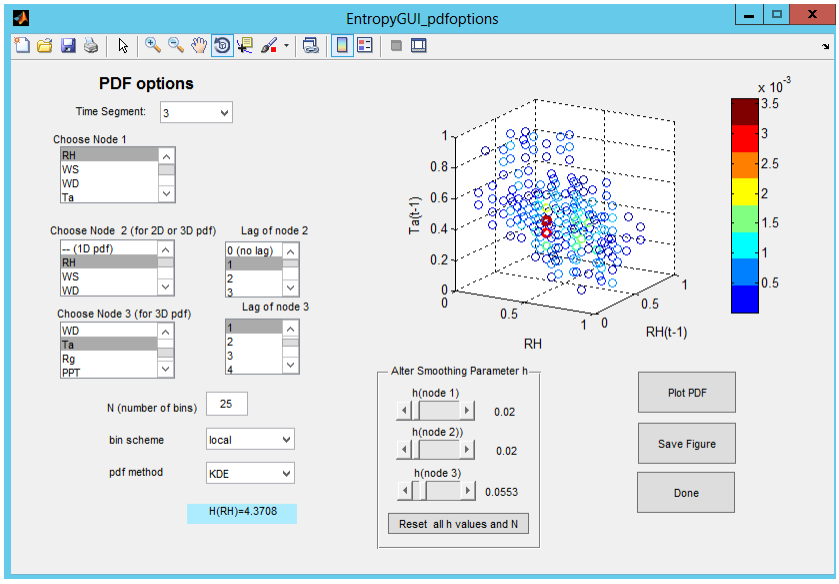

Figure B.12: TIPNet example 3d *pdf* using KDE method. 3d *pdf* of RH, lagged Ta, and lagged RH for a specific segment using KDE method. Red color indicates higher value of $p(RH, Ta(t-1), RH(t-1))$.

Time Segment For data sets that have been segmented in **Pre-Processing Options**, choose segment to view pdf.

Choose nodes and lags Choose 1,2 or 3 nodes to view 1d, 2d, or 3d *pdf*, respectively. To view lagged *pdf*, choose lag for second and third nodes. A 1d *pdf* will appear as a bar chart where the height of each bar corresponds to $p(x)$. A 2d *pdf* will appear as a color scaled image where the color corresponds to $p(x, y)$. A 3d *pdf* will appear as a 3d point cloud, where a point represents a $p(x, y, z) > 0$.

N Number of bins or locations at which to compute *pdf*. The default value is $N = 25$, and N can range up to 100.

pdf method Choose between the KDE method and fixed bin method (default). For the KDE method, a box will appear in which the smoothing parameter h can be altered. A larger h value for a node results in a smoother *pdf*. Once h is changed for a node, it is updated in the **mi.KDEparams** structure.

bin scheme For segmented data, a global bin scheme (default) scales the data between the global minimum (0) and maximum (1) values. A local bin scheme scales the data for each segment separately between the minimum and maximum values in that segment.

After selecting nodes and/or altering parameters, clicking **Plot PDF** will update the *pdf* plot accordingly. When the KDE method is selected, the **Reset all h values and N** button will reset any previously altered smoothing parameters to the default values and set $N = 25$. Clicking **Done** will save any altered parameters.

B.3.7 Network computations and plotting

Once all options have been selected as desired, click **Compute Links** to construct the temporal information networks.

If the Parallel option is turned off (default option in **Network Options**), a timer window will appear for each segment. For large data sets (typically greater than 1000 data points per segment, more than 20 nodes, or many segments), this could take several minutes to initialize and up to multiple hours to complete. When the Parallel Option is turned on, a progress bar will appear in the Matlab command window. When all computations are finished, the output is saved in the previously created project file in a structure called *entropy*.

Click **Plot Results** to view network figures.

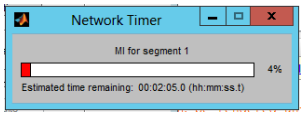

Figure B.13: TIPNet timer bar, will appear for each segment of data set.

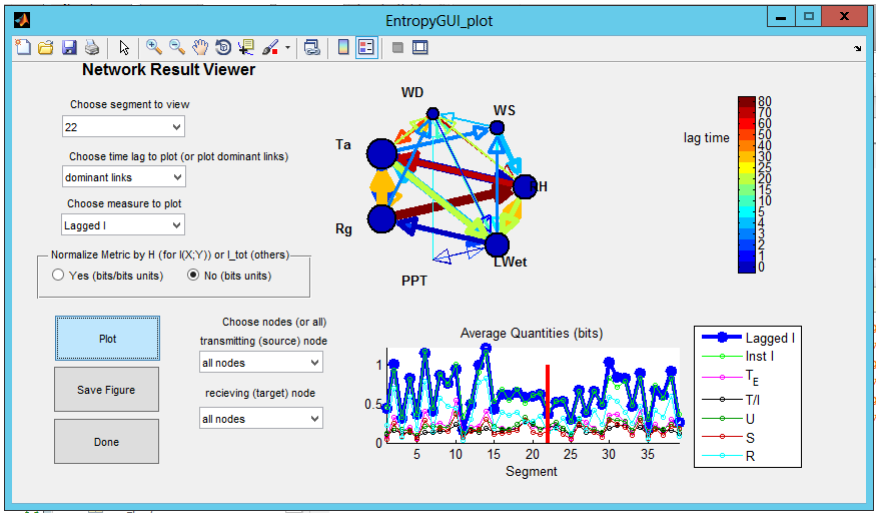

Figure B.14: TIPNet Result Viewer showing network statistics. Here we show Segment 20 lagged information network statistics for all nodes.

The network circle plots contain each node and depict several information measures and associated time lags and strengths. The arrow indicates directionality (source to target), the color indicates time lag of detected link, and the line width indicates the strength of the link. The node size and color correspond to the "self"-link properties, which may or may not be relevant depending on the selection of **Omit Self Links** in the Network Options. The time series or point plot below the circle network shows each segment (for 1 or more segments) and the total values (averages) for 6 information measures.

Choose Segment This list box is only visible if the data set has been partitioned into multiple segments in the **Pre-Processing Options**.

Choose Time Lag For lagged mutual information only, the value $I(X(t-\tau);Y(t))$ can be plotted for individual values of τ as defined in the lag vector ($mi.lagvect$). For all other measures, only the dominantly detected lags are shown in the circle network plot.

Choose Measure 6 measures can be plotted as described in the previous section

Normalize To depict links normalized by entropy $H(X)$ (for lagged I) or total information I_tot (for all other values except T/I which is already normalized), check *Yes*. Otherwise, values plotted are in units of *bits*.

Choose nodes (or all) Select a specific node pair to view only statistics for that link, or a single source or target node to view out-going or incoming links, respectively.