Information sharing in eco-hydrologic systems: synergy, uniqueness, and redundancy

Praveen Kumar and Allison Goodwell

Civil and Environmental Engineering, UIUC

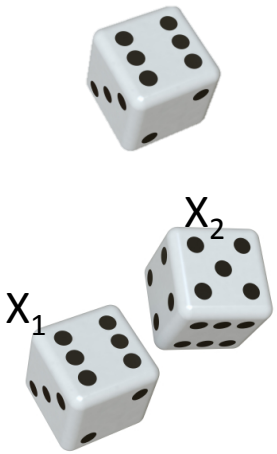
Information Theory Workshop April 2016

REACH Resilience under Accelerated Change

NASA Earth and Space Science Fellowship (NESSF)

NSF Intensively Managed Landscape Critical Zone Observatory (IML-CZO)

Complexity and Emergence



Which number will you bet on?

Which number will you bet on now?

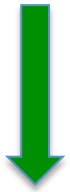
Incomplete question unless we define what interaction we are looking for?

X_1+X_2, $\text{Min}(X_1,X_2)$, others…

$$H(X1) = H(X2) = \log_2 6 = 2.58\ bits$$

$$H(Y) = H(X1+X2) = 3.274\ bits < 2*2.58$$

Randomness



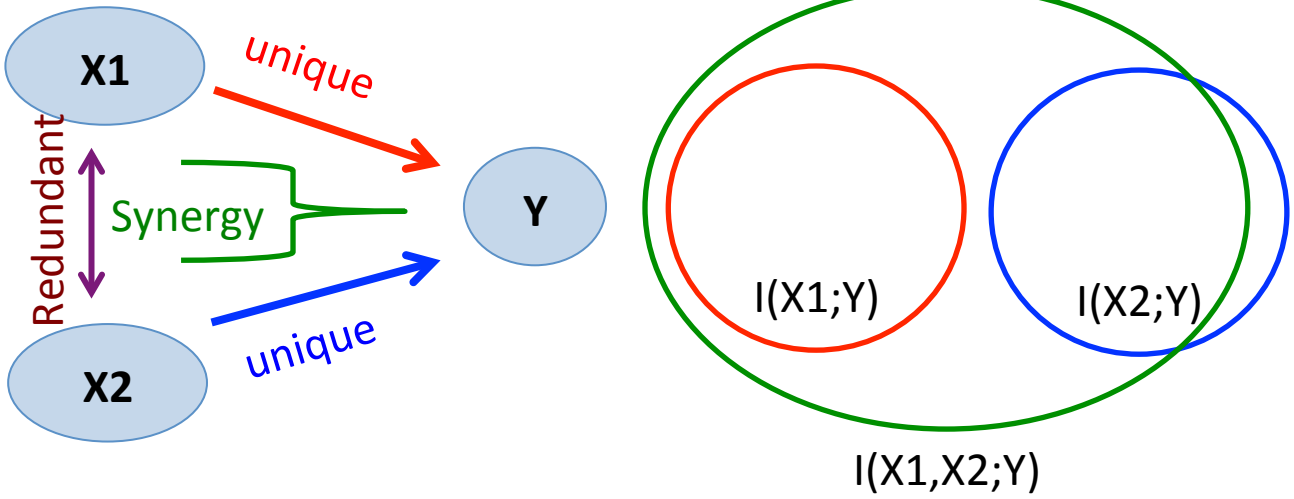
Emergence

Order

Spontaneous reduction in uncertainty due to (synergistic) Interactions of elementary behavior in an open system

Whole is greater than the sum of the parts

Partial Information Decomposition



(Williams and Beer, 2011)

$$I(Y;X1,X2) = U(Y;X1) + U(Y;X2) + R(Y;X1,X2) + S(Y;X1,X2)$$

unique information from each source individually

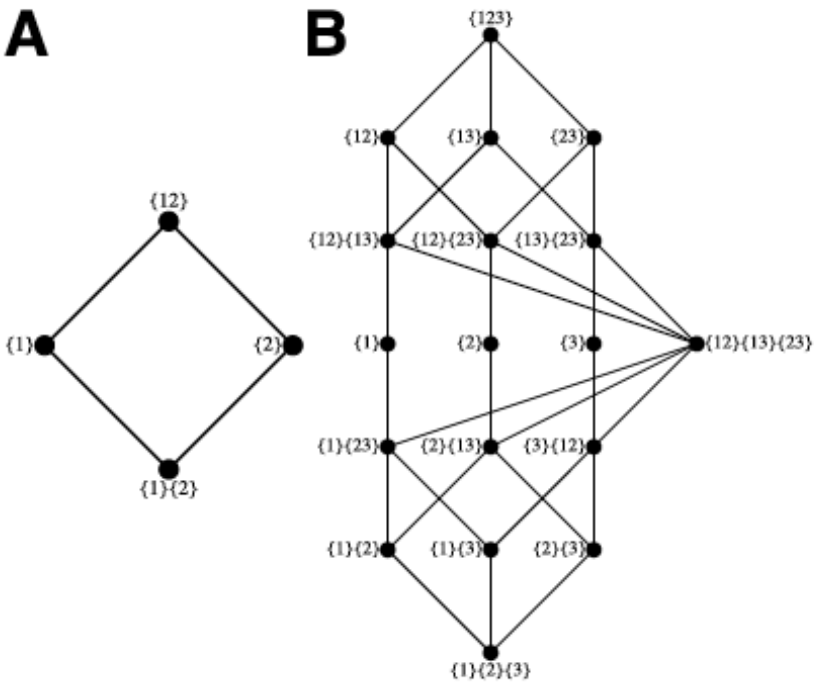
redundant information

synergistic information provided by sources together

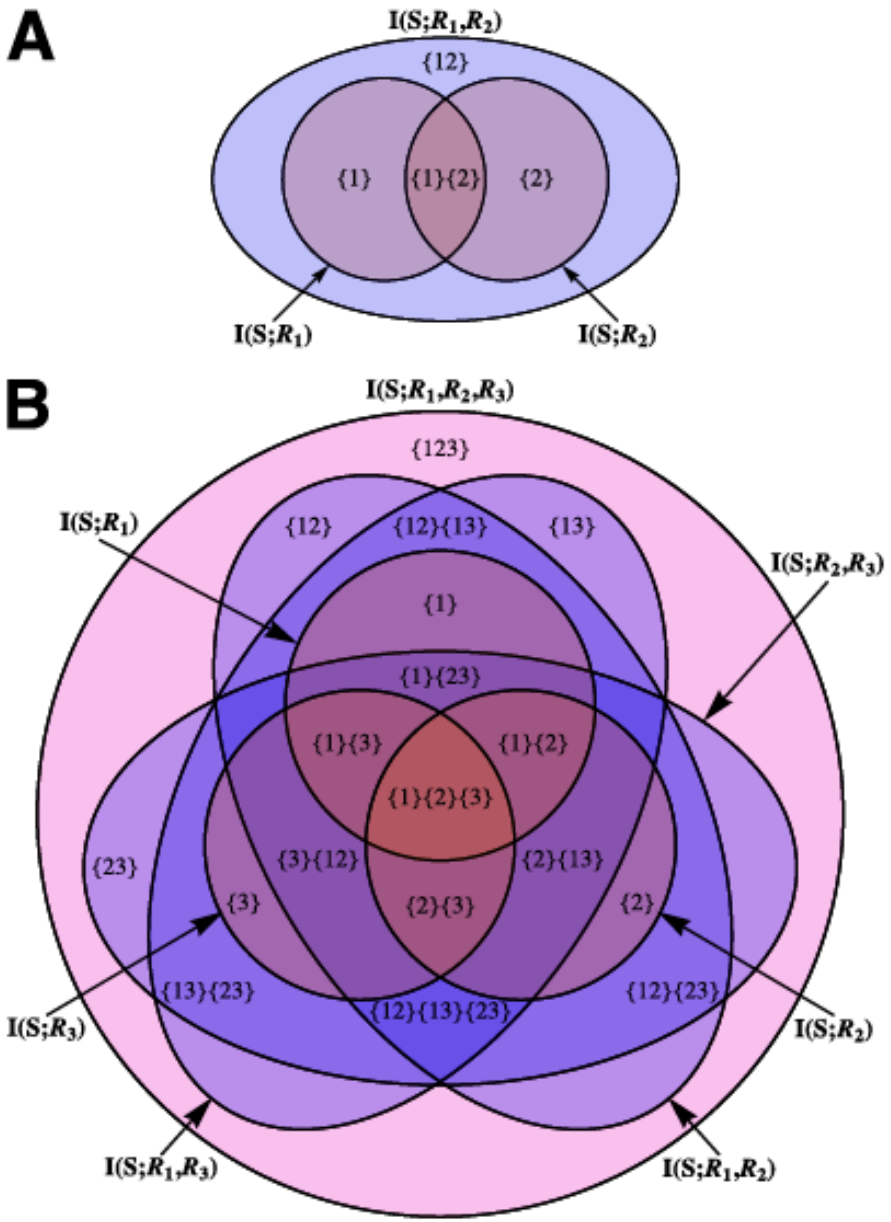
How can we estimate S-U-R components?

Question: Is synergistic information associated with emergent behavior?

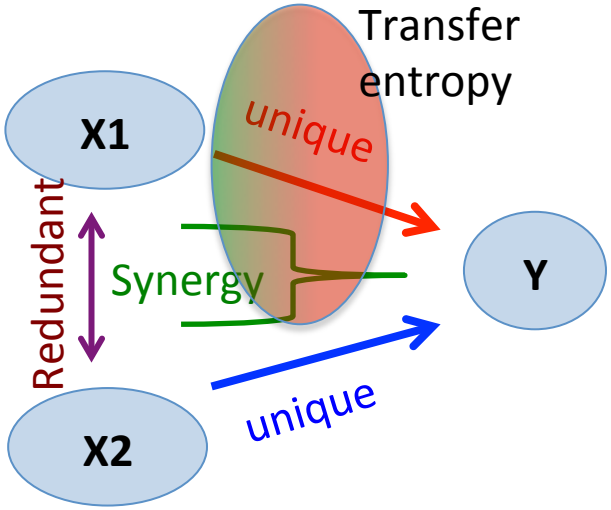
Link to Lattice Theory

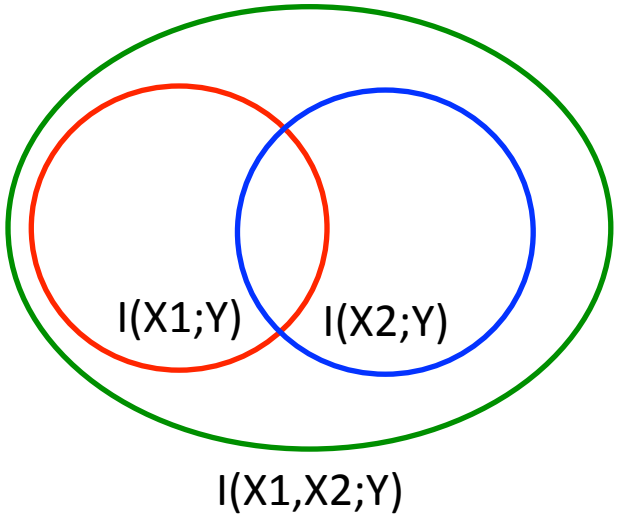


(Williams and Beer, 2011)



Partial Information Decomposition





(Williams and Beer, 2011)

$$I(Y;X1,X2) = \underbrace{U(Y;X1) + U(Y;X2)}_{\text{unique information from each source individually}} + \underbrace{R(Y;X1,X2)}_{\text{redundant information}} + \underbrace{S(Y;X1,X2)}_{\text{synergistic information provided by sources together}}$$

How can we estimate S-U-R components?

Question: Is synergistic information associated with emergent behavior?

estimating redundant information

$$I(Y;X1,X2) = U(Y;X2) + U(Y;X2) + R(Y;X1,X2) + S(Y;X1,X2)$$

*One of the proposed estimates of **R***

$$\underbrace{\max[0, I(X1;Y) + I(X2;Y) - I(X1,X2;Y)]}_{R_{min}} \leq R \leq \underbrace{\min[I(X1;Y), I(X2;Y)]}_{R_{max}}$$

$$R = R_{\min} + \underbrace{\frac{I(X1;X2)}{\min[H(X1), H(X2)]}}_{\text{normalized source dependency}} \left(R_{\max} - R_{\min}\right)$$

if sources are totally dependent (redundant) → link is maximally redundant. If sources are independent of each other, the link is minimally redundant

S (U1+U2+**S**+R) -(U1+R) -(U2+R) +R

$$S(Y;X1,X2) = I(X1,X2;Y) - [I(X1;Y) + I(X2;Y) - R]$$

Simple examples: uniform and Gaussian distribution

For Sum of Two Die Rolls:

$$H(X1) = H(X2) = \log_2 6 = 2.58\ bits$$

$$H(Y) = H(X1 + X2) = 3.274\ bits$$

$$I(X1;Y) = 0.689\ bits = U1 = U2$$

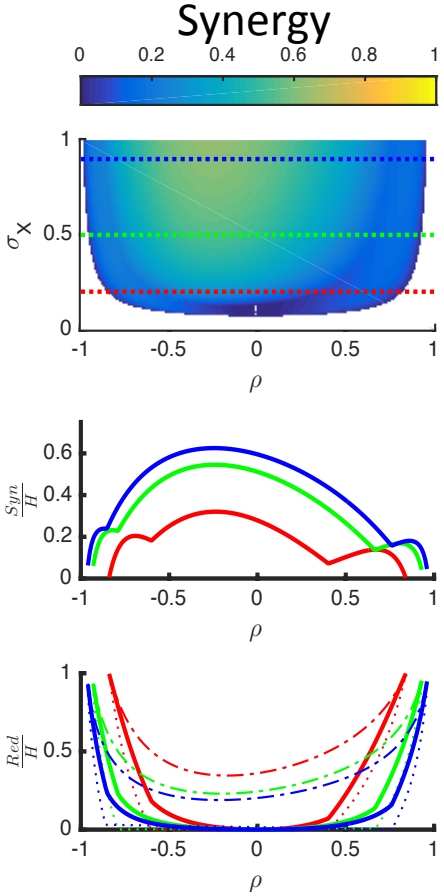
$$S = 3.27 - (0.689 + 0.689 - 0) = 1.896\ bits$$

For sum of two Gaussian variables:

$$R_{\min} = \max\left[0, \frac{1}{2}\ln\left(\frac{(\sigma_{x1}^2 + \sigma_{x2}^2 + \rho\sigma_{x1}\sigma_{x2})}{2\pi e\sigma_{x1}^2\sigma_{x2}^2(1-\rho^2)^2}\right)\right] \qquad R_{\max} = \frac{1}{2}\ln\left(\frac{(\sigma_{x1}^2 + \sigma_{x2}^2 + \rho\sigma_{x1}\sigma_{x2})}{(1-\rho^2)\max(\sigma_{x1}^2, \sigma_{x2}^2)}\right)$$

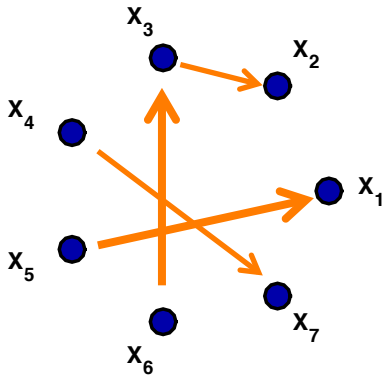
$$R = R_{\min} - \frac{\frac{1}{2}\ln(1-\rho^2)}{\ln(2\pi e\min(\sigma_{x1}^2, \sigma_{x2}^2))}(R_{\max} - R_{\min})$$

$$S = \frac{1}{2}\ln\left(\frac{2\pi e\sigma_{x1}^2\sigma_{x2}^2}{(1-\rho^2)^2(\sigma_{x1}^2 + \sigma_{x2}^2 + \rho\sigma_{x1}\sigma_{x2})}\right) + R$$



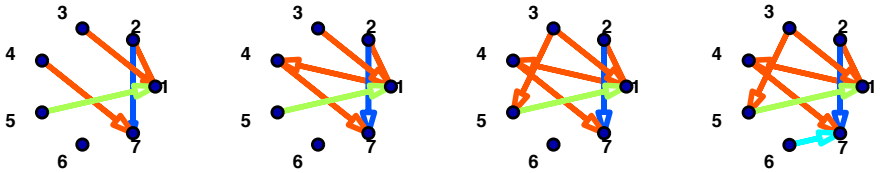
studying an ecohydrologic network with the S-R-U approach

different types of information?



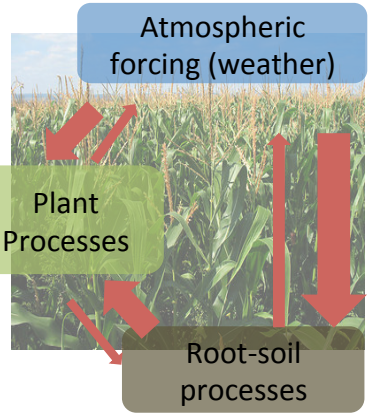
a network context

shifting links and information

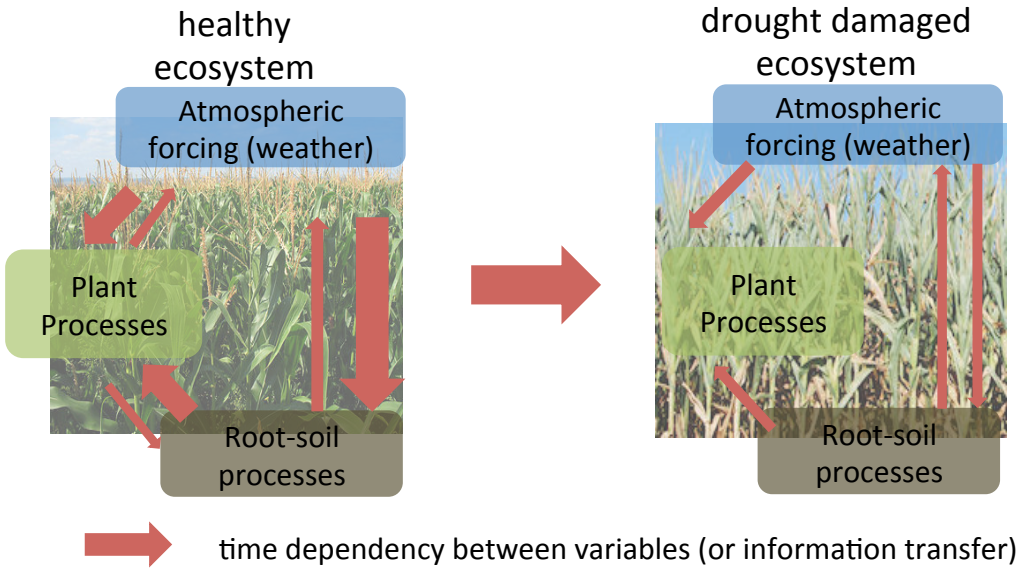


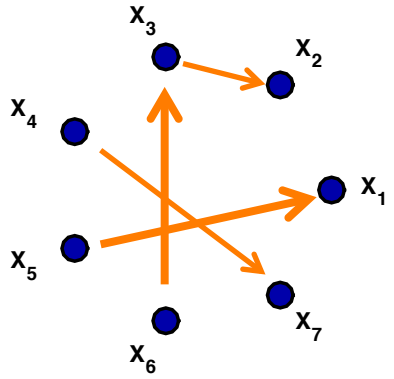
Illinois weather station case

ecosystems as networks



perturbations alter interactions and can lead to shifts in behavior





Questions: How can we identify these interactions and shifts, and are there signatures in S-R-U information?

Strength How dominant is the link?	**Time Scale** Does link occur at fast or slow timescale?
Redundancy Is the link unique, or redundant due to feedback?	**Synergy** Does the link provide additional information with another link?

network of time dependencies is established using information theory

entropy is uncertainty

$$H(X_{tar}) = -\sum p(x_{tar}) \log_2 p(x_{tar})$$

mutual information is a reduction in uncertainty

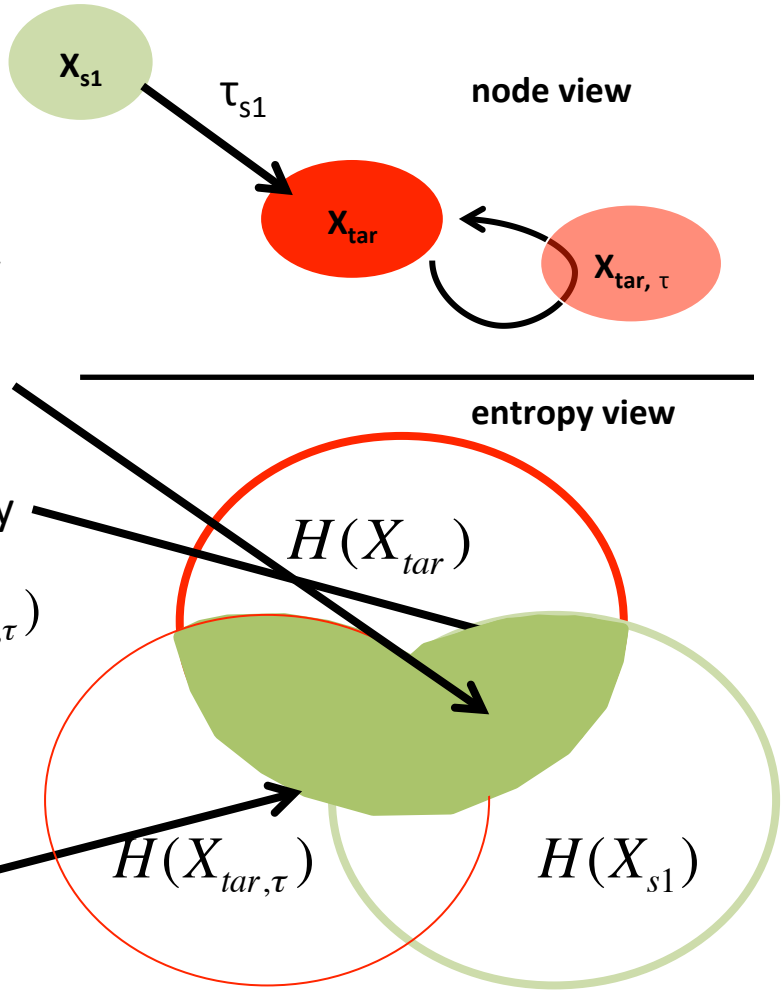
$$I(X_{s1}; X_{tar}) = H(X_{tar}) - H(X_{tar} \mid X_{s1})$$

transfer entropy is a further reduction in uncertainty

$$I(X_{tar}; X_{s1} \mid X_{tar,\tau}) = H(X_{tar} \mid X_{tar,\tau}) - H(X_{tar} \mid X_{s1}, X_{tar,\tau})$$

total information from two sources to target

$$I(X_{tar}; X_{s1}, X_{tar,\tau}) = I(X_{tar}; X_{s1} \mid X_{tar,\tau}) + I(X_{tar}; X_{tar,\tau})$$



network links have synergistic, redundant, and unique components

=

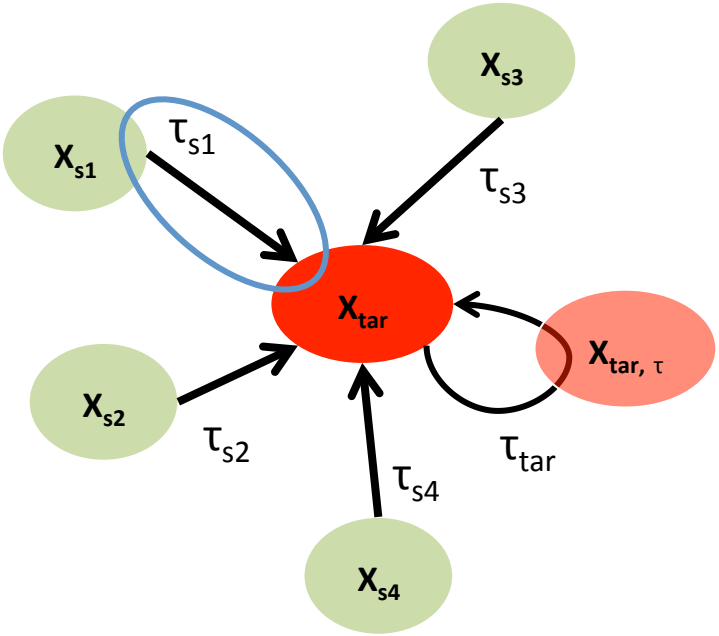
(Williams and Beer, 2011)

$$I(X_{tar};X_{s1},X_{tar,\tau}) = \underbrace{U(X_{tar};X_{tar,\tau})}_{\text{unique information from each source individually}} + U(X_{tar};X_{s1}) + \underbrace{R(X_{tar};X_{s1},X_{tar,\tau})}_{\text{redundant information}} + \underbrace{S(X_{tar};X_{s1},X_{tar,\tau})}_{\text{synergistic information provided by sources together}}$$

=

for any source → target link:
maximum R (redundancy with another source)
maximum S (synergy with another source)

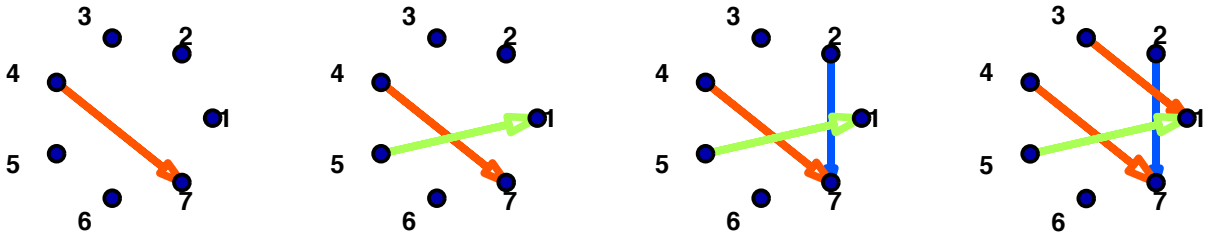
*pairwise comparison of links:
for example, X_{s3} link might be most redundant with X_{s1} link, but X_{s4} link may be most synergistic with X_{s1}



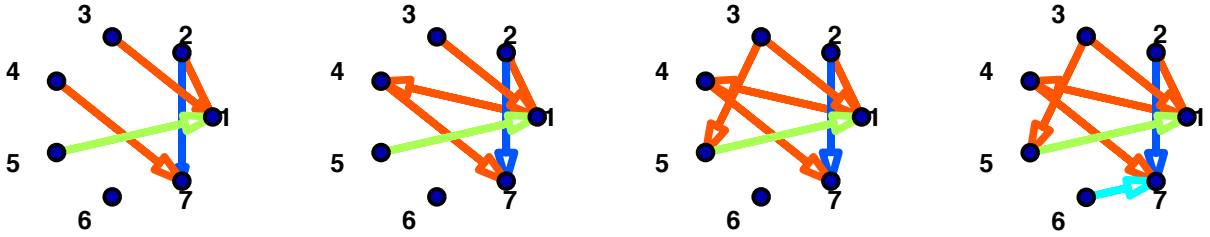
S-R-U can reveal different types of feedbacks/links in network

$f(X) = 4X(1-X)$ ← chaotic logistic equation

$$X_i(t) = \frac{\varepsilon}{k_i}\sum_{j=1}^{N} w_{j,i}\left[f(X_j(t-\tau_{j,i}))\right] + (1-\varepsilon)z$$ ← coupled chaotic logistic + noise



randomly add one link at a time…

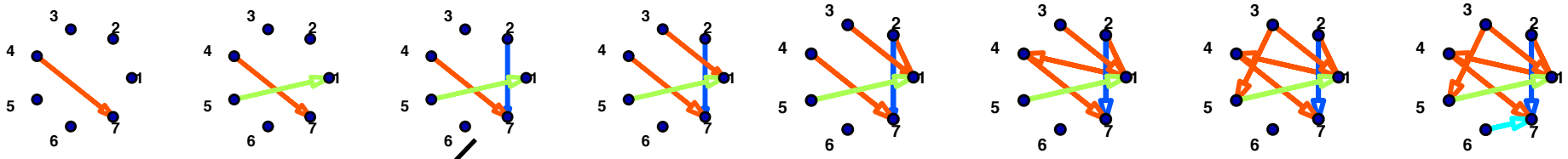


….full connectivity → synchronized nodes

information along the path to synchronization

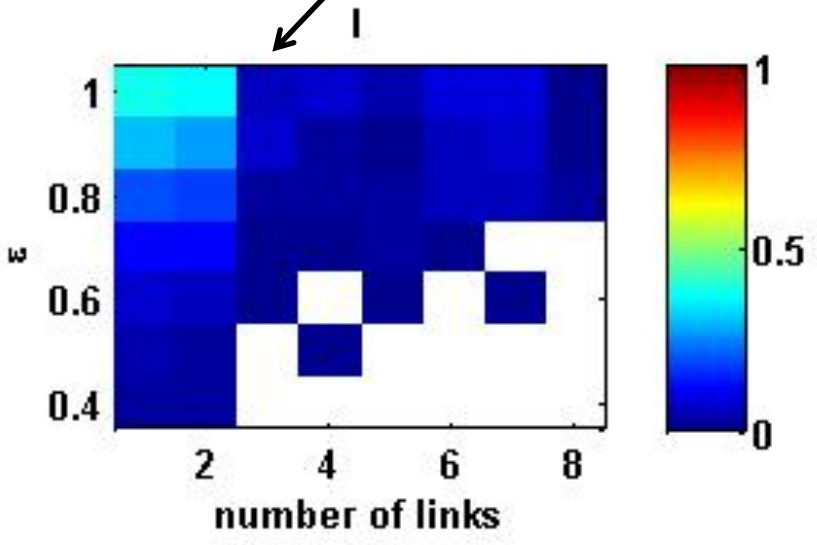
$$X_i(t) = \frac{\varepsilon}{k_i}\sum_{j=1}^{N} w_{j,i}\left[f(X_j(t-\tau_{j,i}))\right] + (1-\varepsilon)z$$

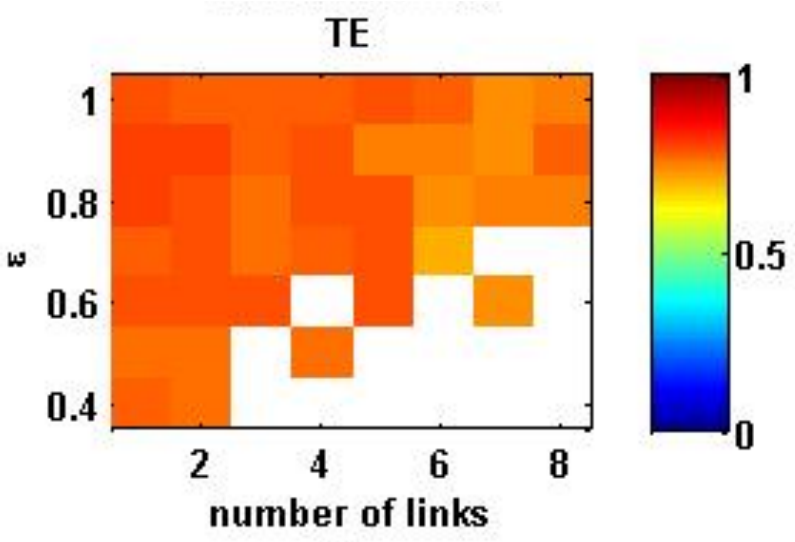
look at detected measures of lagged MI and TE for this original 4→7 link



decrease in detected strength at round 3 due to other source

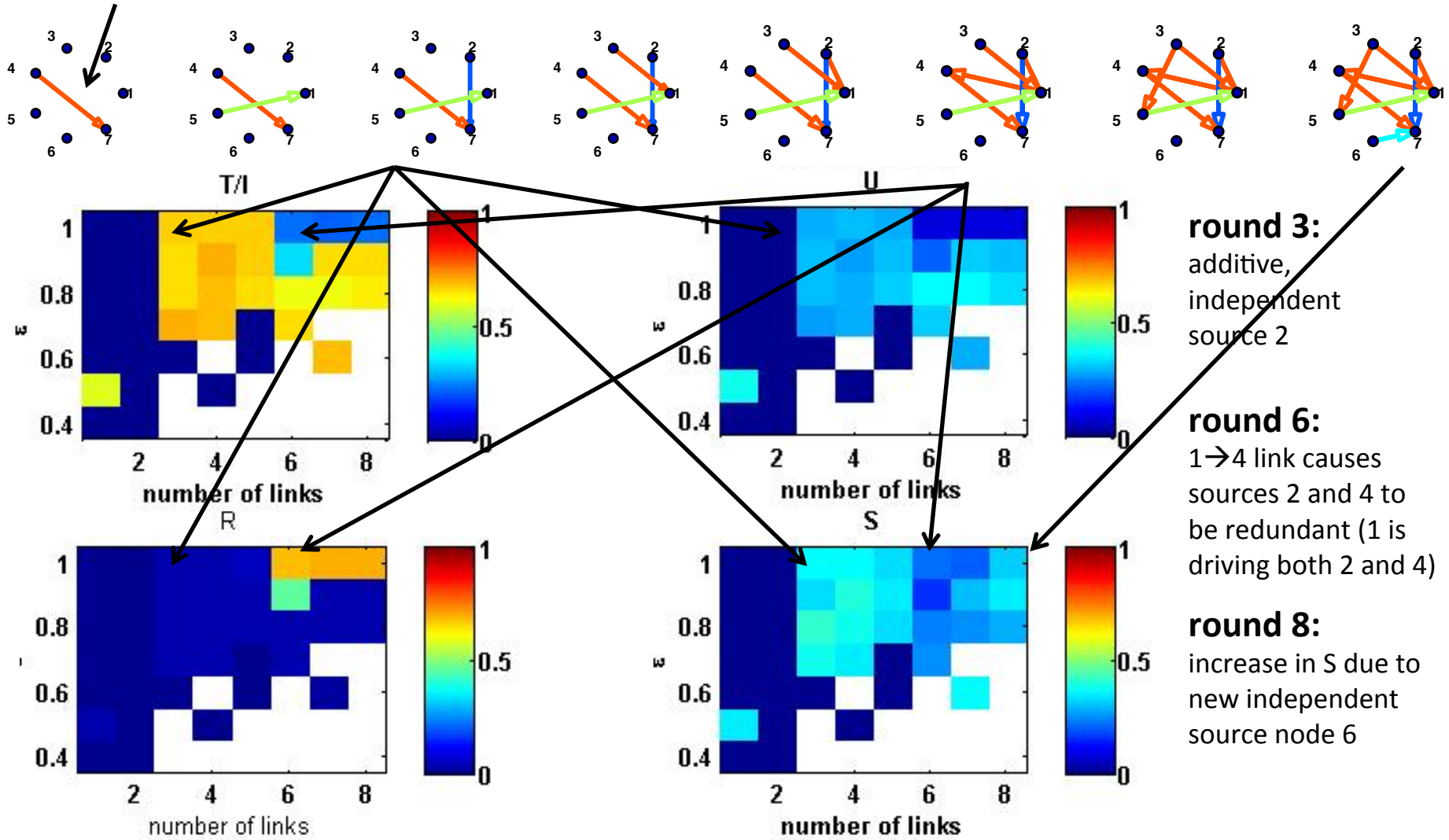
few changes in transfer entropy





information along the path to synchronization

look at detected measures of lagged MI and TE for this original 4→7 link

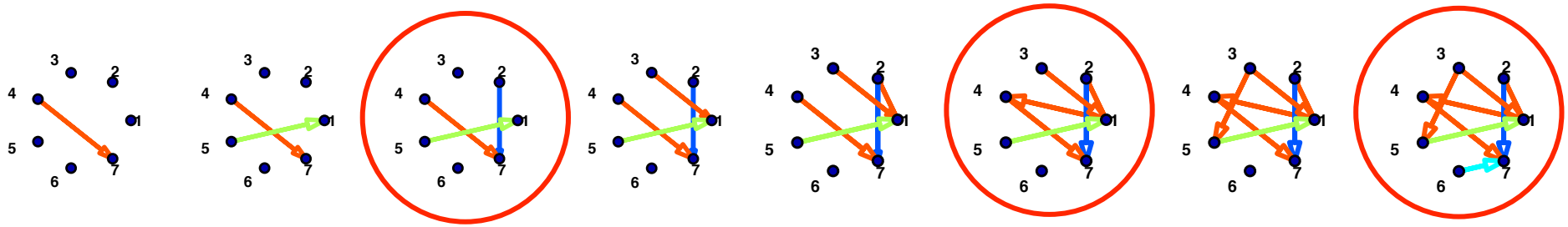


round 3:
additive, independent source 2

round 6:
1→4 link causes sources 2 and 4 to be redundant (1 is driving both 2 and 4)

round 8:
increase in S due to new independent source node 6

for test cases, changes in S-R-U shift as links are added

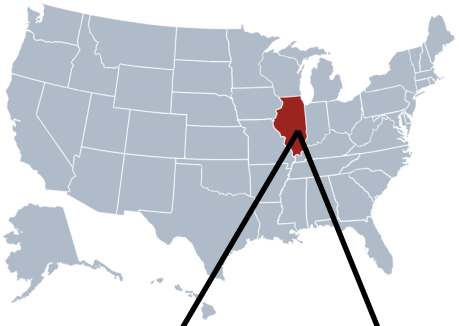


redundancy
due to feedbacks, nodes synchronize in some way

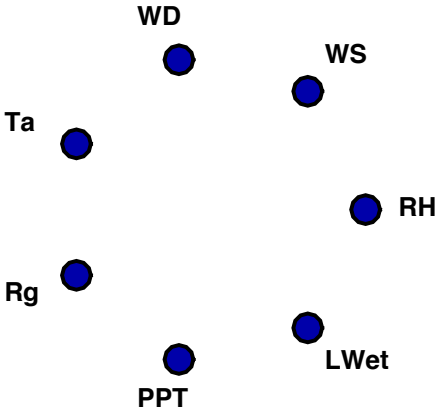
uniqueness
all sources are independent or partially dependent

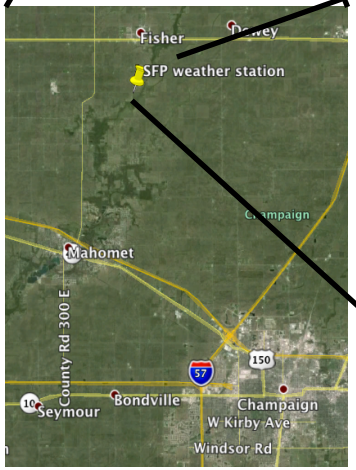
synergy
multiple sources provide more information together

we use weather station data to explore these interactions



Sangamon Forest Preserve Intensively Managed Landscapes (IML) CZO

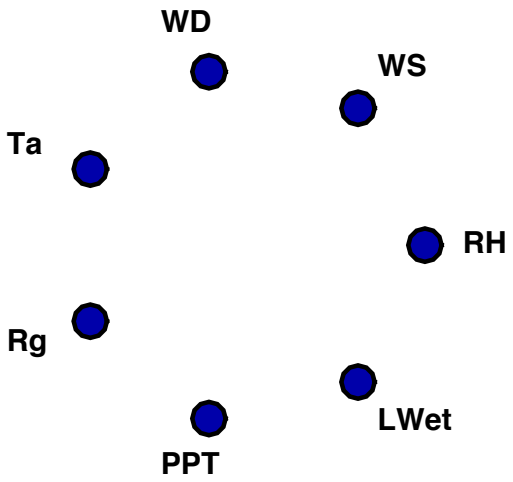




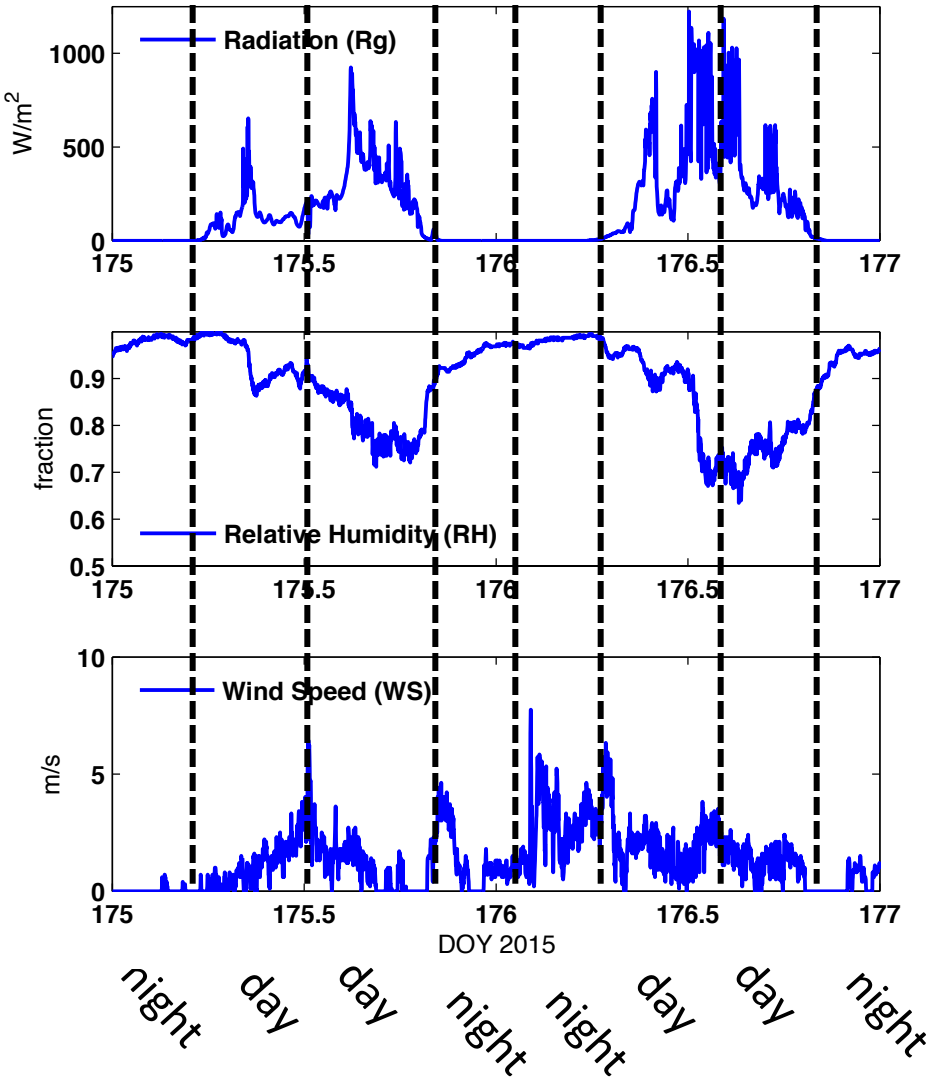
knowing our nodes:

air temperature	Ta
wind direction	WD
wind speed	WS
relative humidity	RH
leaf wetness	Lwet
rainfall	PPT
radiation	Rg

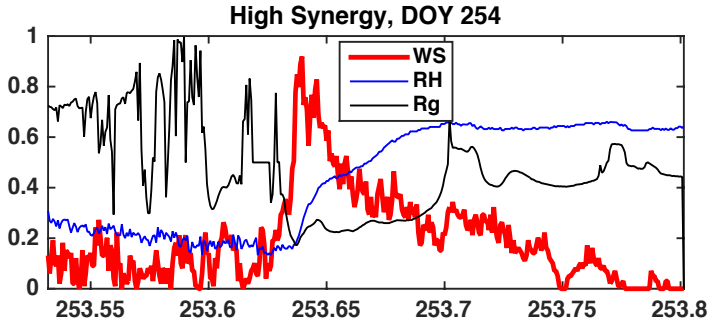
with high temporal resolution data, we identify interactions on sub-daily scale

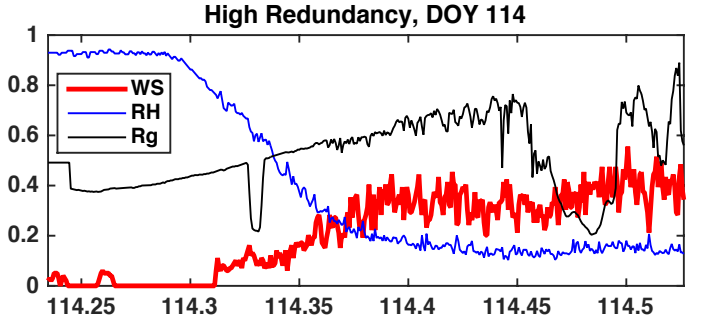


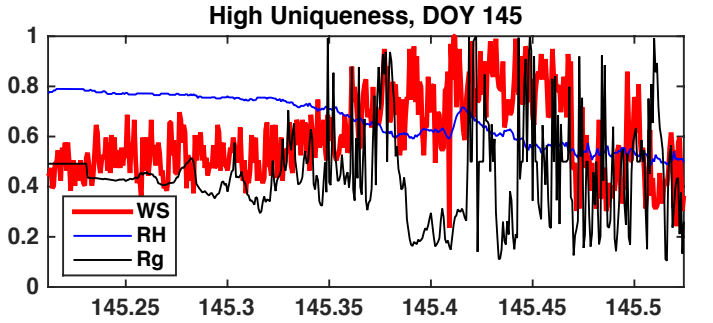
February – November 2015
~ 600 day-time network segments

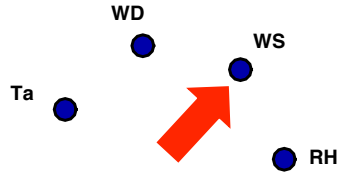


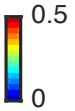
S,R,U measures indicate different types of interactions and feedbacks



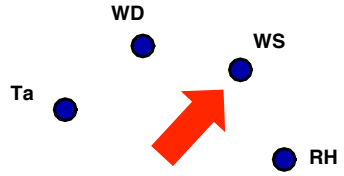


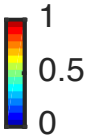




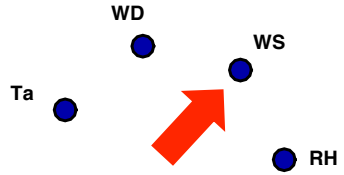


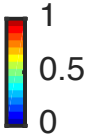
WS = synergistic target of RH and Rg



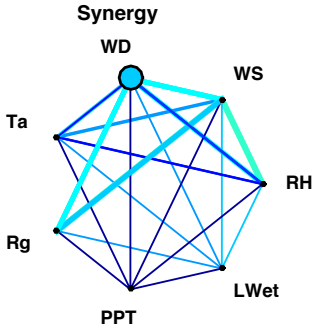


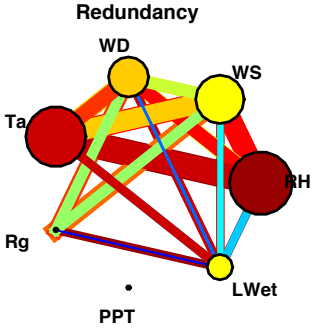
WS = redundant target of RH, Rg, WS

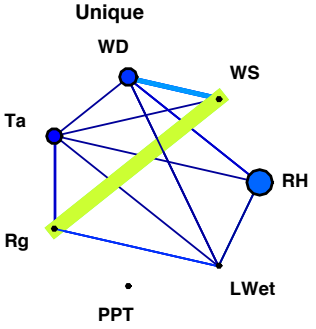




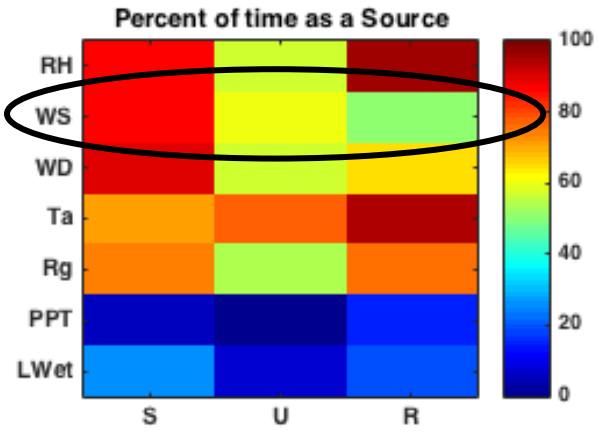
WS = unique target of Rg

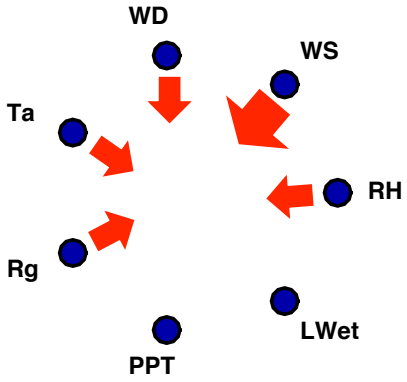


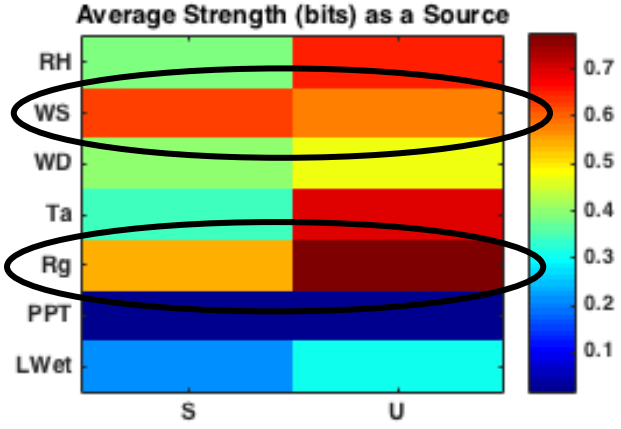


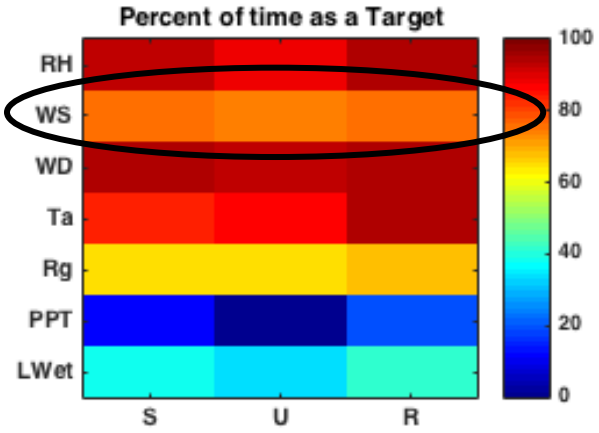


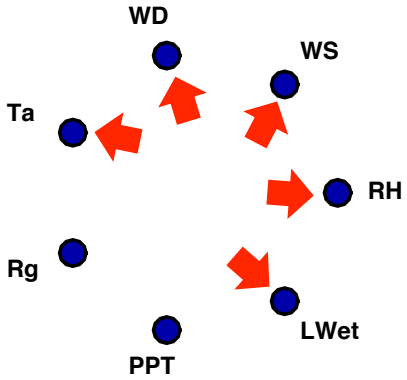
nodes are both sources and targets of S-R-U, but interaction strengths vary

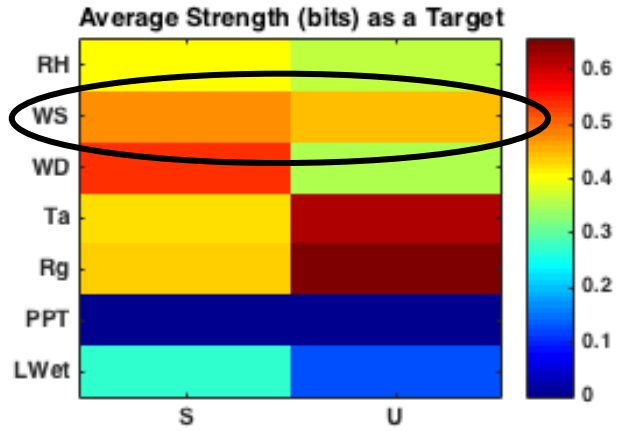








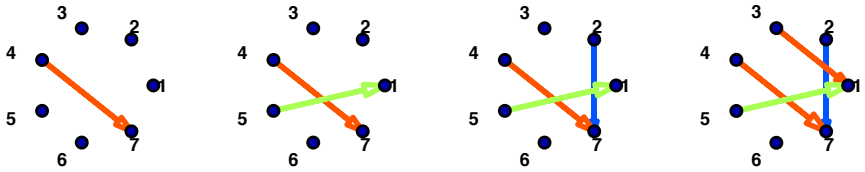




conclusion: an information approach can reveal ecosystem responses

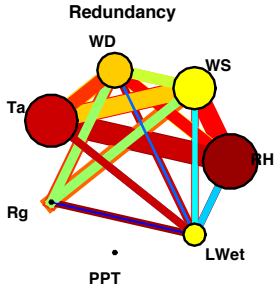
information measures used to detect time dependent interactions in terms of :

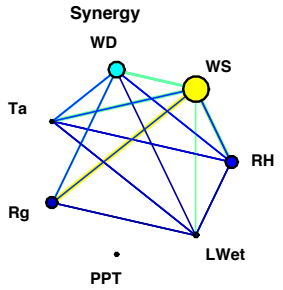
Strength	Time Scale
Redundancy	Synergy

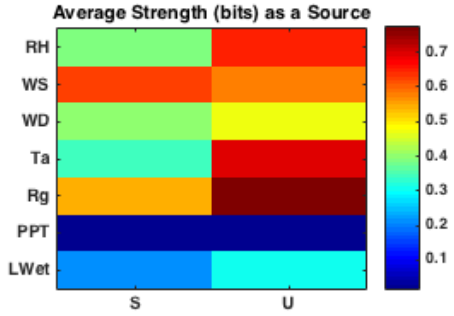


S-R-U measures detect shifting links

measures applied to weather station data show shifting links with conditions

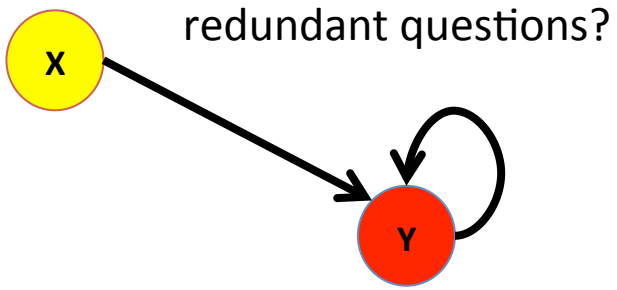






this approach can reveal feedback, self-feedback, synchronization, and multiple unique or synergistic drivers in eco-hydrologic systems

questions?



Allison Goodwell
graduate student, University of Illinois
allison@goodwell.net

REACH
Resilience under Accelerated Change

NASA
Earth and Space Science Fellowship (NESSF)

NSF
Intensively Managed Landscape Critical Zone Observatory (IML-CZO)