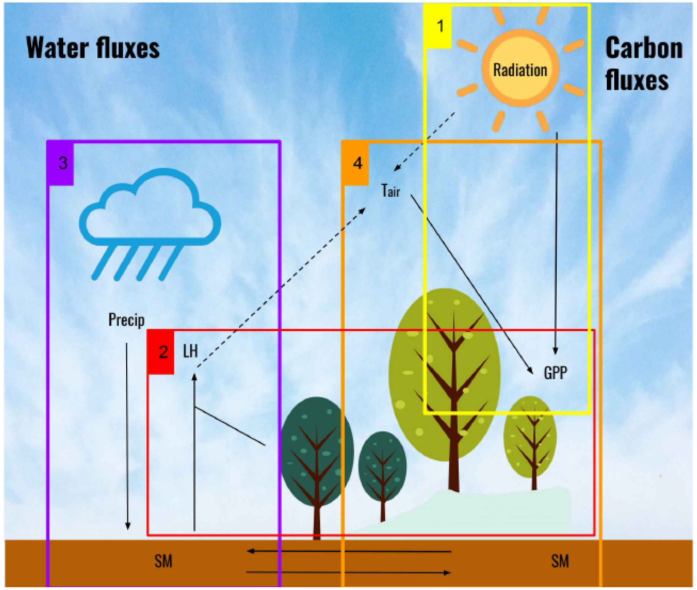
Causality in carbon and water cycle using convergent cross mapping (CCM)



Emiliano Díaz, Jose E. Adsuara, Álvaro Moreno Martínez, María Piles, and Gustau Camps-Valls

Image Processing Laboratory (IPL)
Universitat de València, Spain
emiliano.diaz@uv.es | http://isp.uv.es

USMILE
Understanding and Modelling the Earth System with Machine Learning

CCM - aknowledgements & agenda

Acknowledgements, References & Agenda:

- **CCM** methodology, figures and examples from: Sugihara, G. et al. “Detecting causality in complex ecosystems”. Science 338, 496–500 (2012).
- **Extended CCM** methodology, figures and examples from: Ye, H., Deyle, E., J. Gilarranz, L. & Sugihara, G. “Distinguishing time-delayed causal interactions using convergent cross mapping.” Sci. Reports 5, 14750, DOI: 10.1038/srep14750 (2015).

Part 1: Introduction to CCM and some local examples based on work of Sugihara et al

- **Robust CCM** methodology, figures and results from IPL manuscript “Inferring causal relations from observational long-term carbon and water fluxes records” currently under revision by Sci. Reports. Authors (Emiliano Díaz, Jose E. Adsuara, Álvaro Moreno Martínez, María Piles, and Gustau Camps-Valls)

Part 2: Our extended method and global maps for carbon & water cycle relationships

Part1: CCM and examples

Context:

- Causal inference method for time series/dynamic systems

- Intended for data from:
  - Deterministic-ish systems
  - No strong forcings
  - No “instantaneous” processes

Principle

For deterministic dynamical systems two variables are causally related by definition if one of the (ODE for ex) equations expresses the dynamics of the variable in terms of the state of the other.

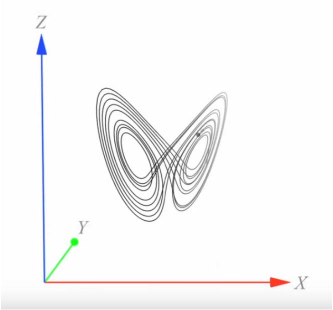
Eg. Lorenz attractor

system

$$\frac{dX}{dt} = \sigma Y - \sigma X$$

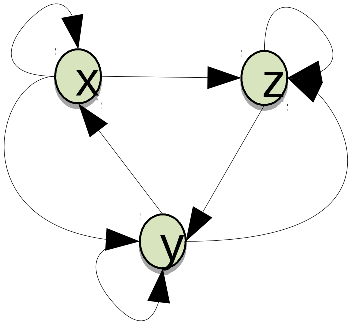
$$\frac{dY}{dt} = -XZ + \rho X - Y$$

$$\frac{dZ}{dt} = XY - \beta Z$$



Data

CCM



causality

¿how do we check two variables form part of the same ODE system from data? → CCM

Takens' Theorem

Takens' Theorem (1981)

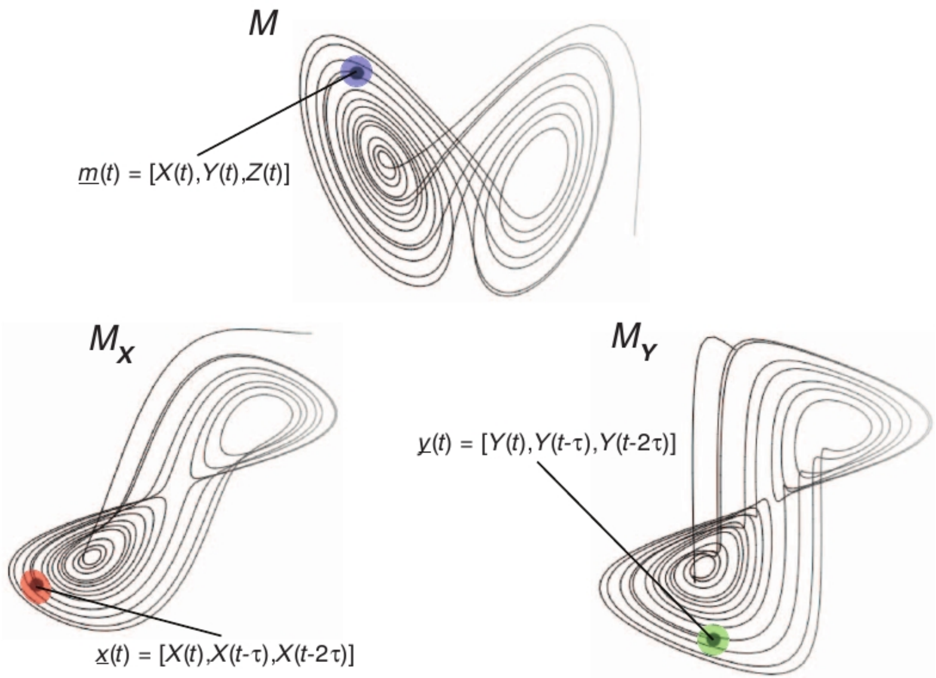
Let $\boldsymbol{M}$ be a compact manifold of dimension d, ϕ a smooth vector field, and X a smooth function on $\boldsymbol{M}$. It is a generic property that

$$\boldsymbol{\Phi}_{(\phi,X)}(\underline{m}) : \boldsymbol{M} \to \mathbb{R}^{2d+1}$$

is an embedding, where ϕ is the flow on $\boldsymbol{M}$ and

$$\boldsymbol{\Phi}_{(\phi,X)}(\underline{m}) = \langle X(\underline{m}), X(\phi(\underline{m})), X(\phi^2(\underline{m})), \ldots, X(\phi^{2d}(\underline{m}))\rangle.$$

Takens' Theorem



Key (informal) take(ns)-away for CCM:

- If M is the manifold that represents the systems' state-space through time, then...
- We can construct a "shadow manifold" (Mx and /or My) using each time series in the system which retains important mathematical properties of original Manifold such as its topology.
- Points close on M are also close on Mx and My
- We can essentially reconstruct the original state space using the time series of only one variable
- **CCM in a nutshell**: to check if two variables causally related (belong in same system), check if you can rebuild the state-space from the variables' (embedded/lagged) individual time series

Takens' Theorem

Intuition

- In a deterministic ODE system information is shared perfectly among all the variables.
- Ex. discrete-time difference equations:

$$X(t+1) = 3.9\, X(t)\, [1 - X(t) - \beta\, Y(t)]$$
$$Y(t+1) = 3.7\, Y(t)\, [1 - Y(t) - 0.2\, X(t)]$$

$$X(t) = \frac{3.9}{0.2}\left\{(1-\beta Y(t-1))\left(1 - Y(t-1) - \frac{Y(t)}{3.7Y(t-1)}\right) - \frac{1}{0.2}\left(1 - Y(t-1) - \frac{Y(t)}{3.7Y(t-1)}\right)^2\right\}$$
$$Y(t) = \frac{3.7}{\beta}\left\{(1-0.2X(t-1))\left(1 - X(t-1) - \frac{X(t)}{3.9X(t-1)}\right) - \frac{1}{\beta}\left(1 - X(t-1) - \frac{X(t)}{3.9X(t-1)}\right)^2\right\}$$

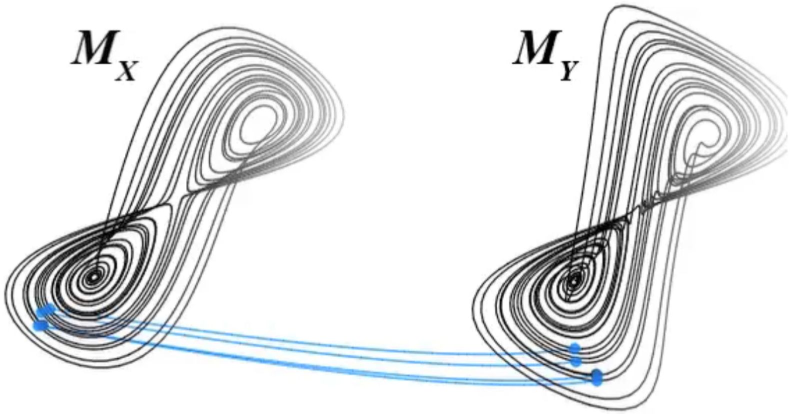
- With regular time sampled time series data although we might not be able to recover the exact manifold, I.e. the **ODE equations**, by Takens' theorem, if we choose the right embedding (the right number of lags) we can recover "similar" equations, ie a simlar manifold that retains, for example, metric properties: point close on shadow manifold also close on manifold.

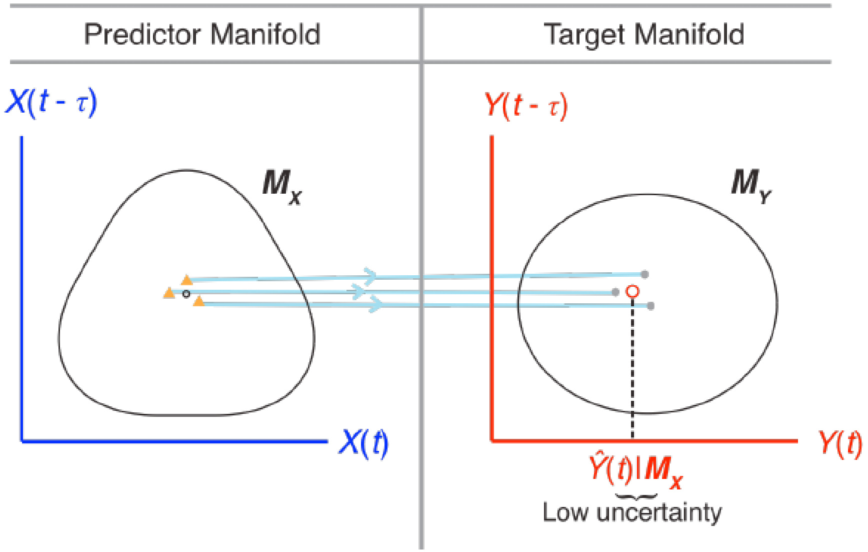
CCM algorithm

1. choose embedding dim E (number of lags),

2. construct shadow manifold Mx

3. Assume topology property to estimate distance of E+1 points (similarity of states) on M based on distance of points on Mx



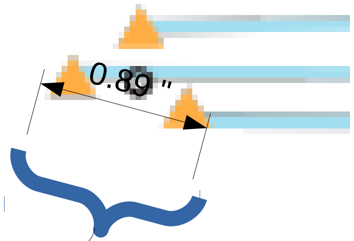


CCM algorithm

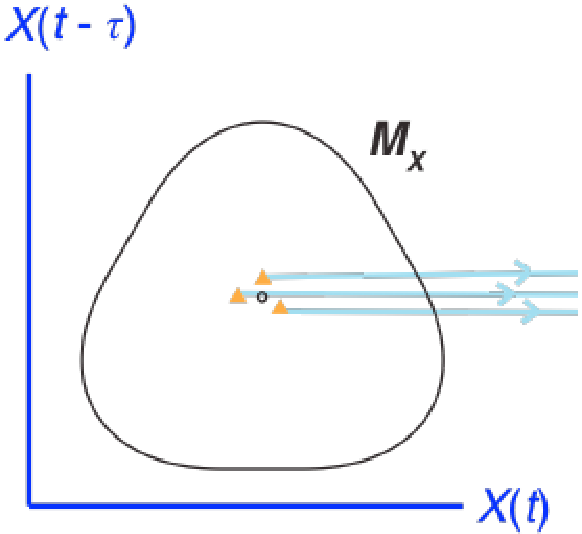
4. Check how “good” shadow manifold is: cross-map (estimate) state Y(t) using weights estimated with shadow manifold

$$\hat{Y}(t) \mid \boldsymbol{M_X} = \sum w_i \, Y(t_i)$$

$$i = 1 \ldots E{+}1$$

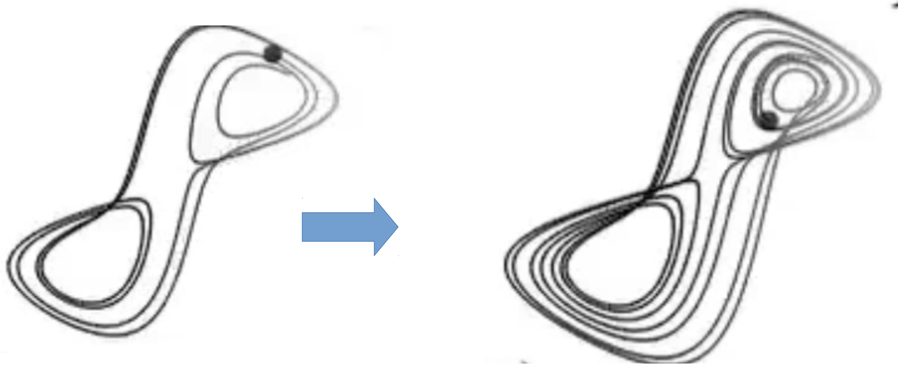


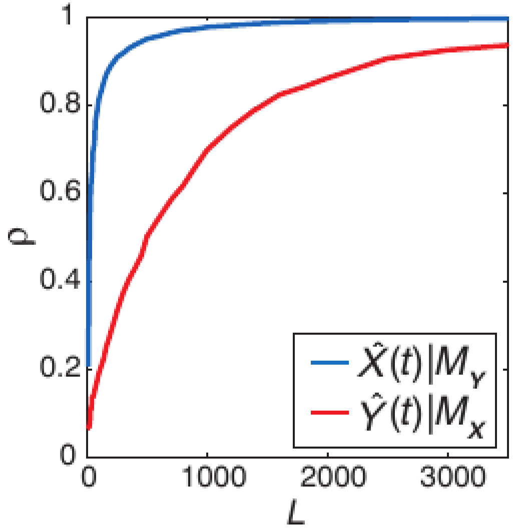
Weights inversely proportional
to distance on Mx



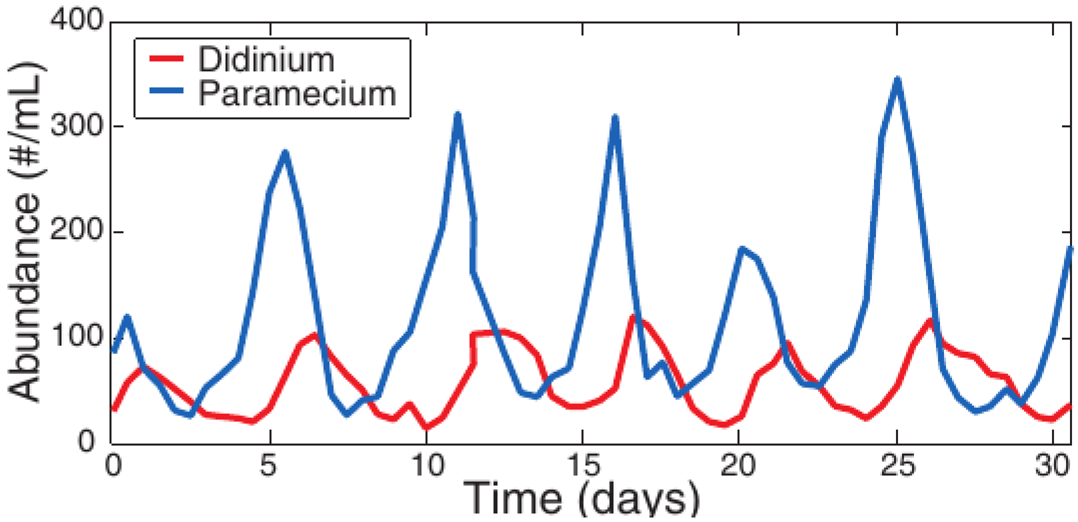
CCM algorithm

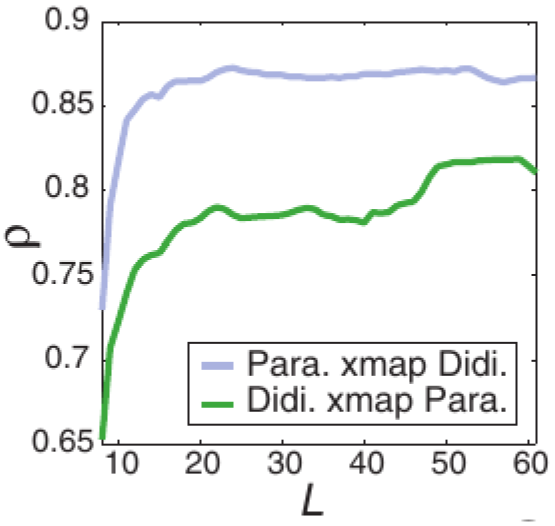
5. Check cross map skill Corr(y, y|Mx)… as number of points (not lags) in time series increases, the estimation of the manifold should be denser so cross map skill should improve and converge. If so x and y in the same system and so causally related.





Predator(Dinidium)-prey(Paramecium) example





Bidirectional causality by prey and predator. Top-down control by predator, Didinium stronger than bottom up control by prey Paramecium.

Directionality

Previous example variables causally related bidirectionally. What happens if we have convergence in cross map skill for only one of the two cross maps (y|Mx or x|My)?

Going back to system equations what does unidirectional causal relationship look like?

Forcing variable x: one of the ODE/difference equations does not involve the other variables. For ex:

$$X(t+1) = 3.9\, X(t)\, [1 - X(t) - \beta\, Y(t)]$$
$$Y(t+1) = 3.7\, Y(t)\, [1 - Y(t) - 0.2\, X(t)]$$

$$\beta = 0$$

- In this case dynamics of X only depend on its own state: x → y but NOT y→ x. How does this translate to cross-maps?
- Asymmetry is generated, Y recieves dynamics of X and its internal dynamics but X only has its internal dynamics.
- Shadow manifold Mx wont have all information of the system: predicting y (effect) with x (cause) “won’t work”

Directionality

$$X(t+1) = 3.9\, X(t)\, [1 - X(t) - \beta\, Y(t)]$$
$$Y(t+1) = 3.7\, Y(t)\, [1 - Y(t) - 0.2\, X(t)]$$

$$\beta = 0$$

$$X(t) = \frac{3.9}{0.2}\left\{(1-\beta Y(t-1))\left(1 - Y(t-1) - \frac{Y(t)}{3.7Y(t-1)}\right) - \frac{1}{0.2}\left(1 - Y(t-1) - \frac{Y(t)}{3.7Y(t-1)}\right)^2\right\}$$

$$Y(t) = \frac{3.7}{\beta}\left\{(1-0.2X(t-1))\left(1 - X(t-1) - \frac{X(t)}{3.9X(t-1)}\right) - \frac{1}{\beta}\left(1 - X(t-1) - \frac{X(t)}{3.9X(t-1)}\right)^2\right\}$$

singularity

- Singularity arises since now information on Y is not present in X: cross map y(t)|M_x will not have good cross map skill since y doesn't cause x.
- We say x → y if x(t)|M_y has good and converging cross map skill.
- Notice that (perhaps counter intuitively) we say x → y if we can predict x using y.
- So back to algorithm to add directionality...

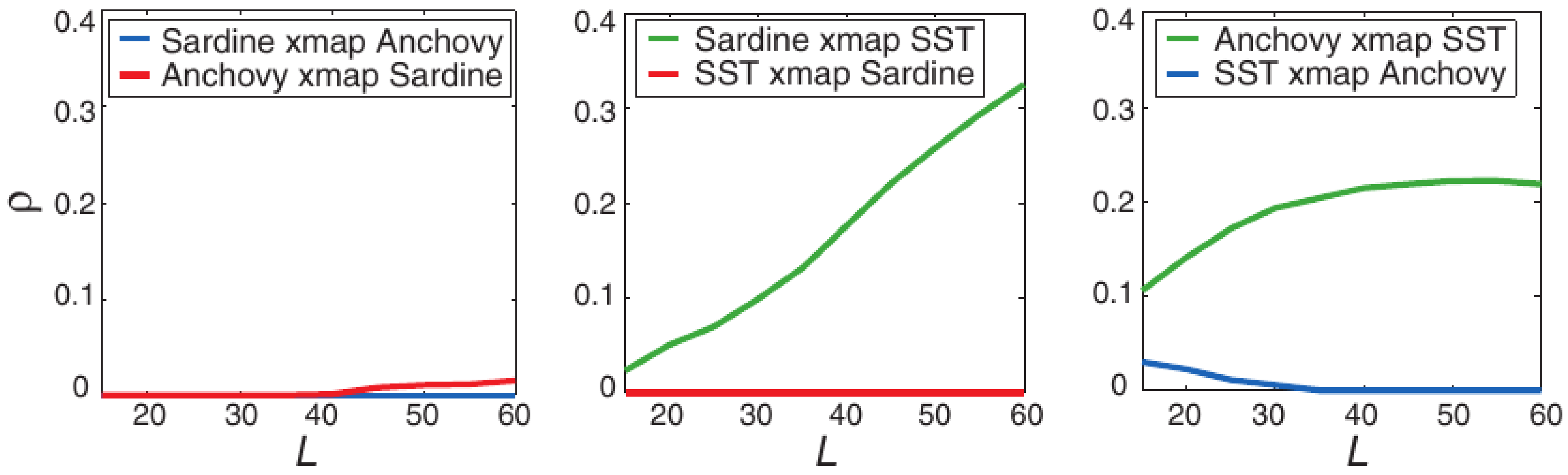
Add directionality to algorithm

First 4 steps the same:

1. choose embedding dim E,
2. Construct shadow manifold Mx
3. Estimate distance of E + 1 lagged points on Mx
4. cross-map (estimate) state Y(t)

5. Check cross map skill Corr(y, y|Mx)… as number of points (not lags) in time series increases, the estimation of the manifold should be denser so cross map skill should improve and converge: x has information of y so: $y \rightarrow x$.

CCM methodology – unidirectional example



- Sea surface temperature in Anchovy and Sardine population time series: so SST→sardine, SST→ anchovies. No feedback relationship as expected.
- Anchovy and Sardine populations don’t interact

Caveats

Already mentioned **deterministic**: For stochastic systems not clear how much information of the cause is transferred to the effect.

General synchrony: **strong unidirectional forcing** looks like bidirectional causality…

$$X(t+1) = 3.9\, X(t)\, [1 - X(t) - \beta\, Y(t)]$$
$$Y(t+1) = 3.7\, Y(t)\, [1 - Y(t) - 0.2\, X(t)]$$

$\beta = 0$

What happens if we remove internal dynamics from Y?

Caveats

Already mentioned deterministic: For stochastic systems not clear how much information of the cause is transferred to the effect.

General synchrony: strong unidirectional forcing looks like bidirectional causality…

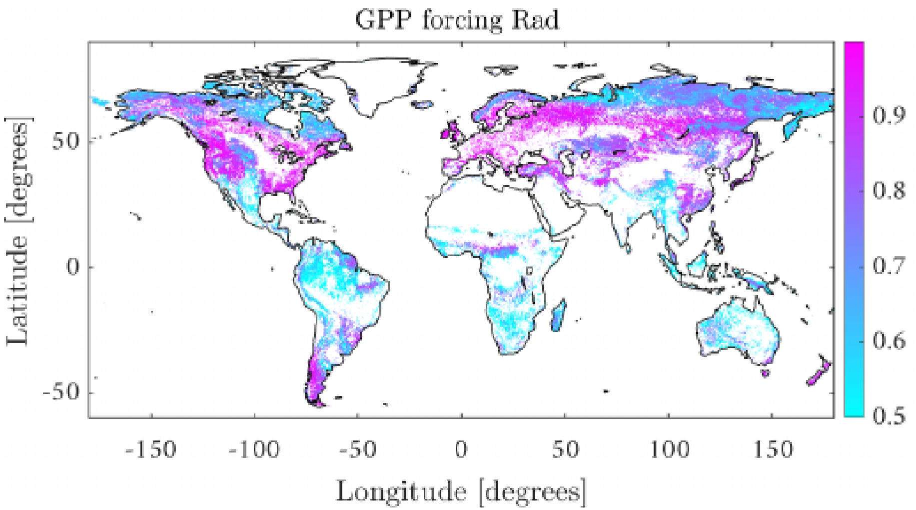
$$X(t+1) = 3.9\, X(t)\, [1 - X(t) \qquad\qquad]$$
$$Y(t+1) = 3.7 \qquad [1 \qquad\qquad - 0.2\, X(t)]$$

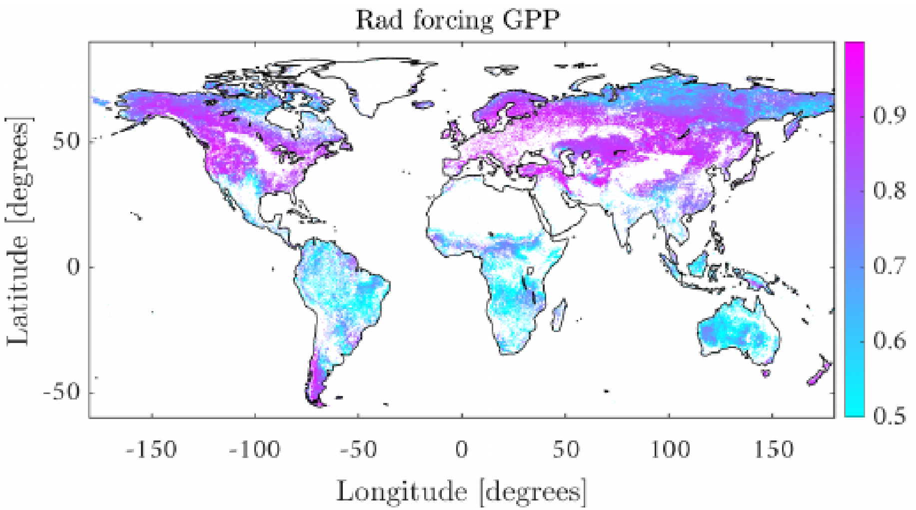
$$0.2\, x(t) = 1 - y(t+1)/3.7$$

What happens if we remove internal dynamics from Y? Asymmetry is broken: X has all the information in the system again since Y no longer has additional information (its internal dynamics are gone or weak)

Caveats

Strong “instantaneous” unidirectional forcing especially problematic





Part2: Our contribution – Robust CCM (RCCM)

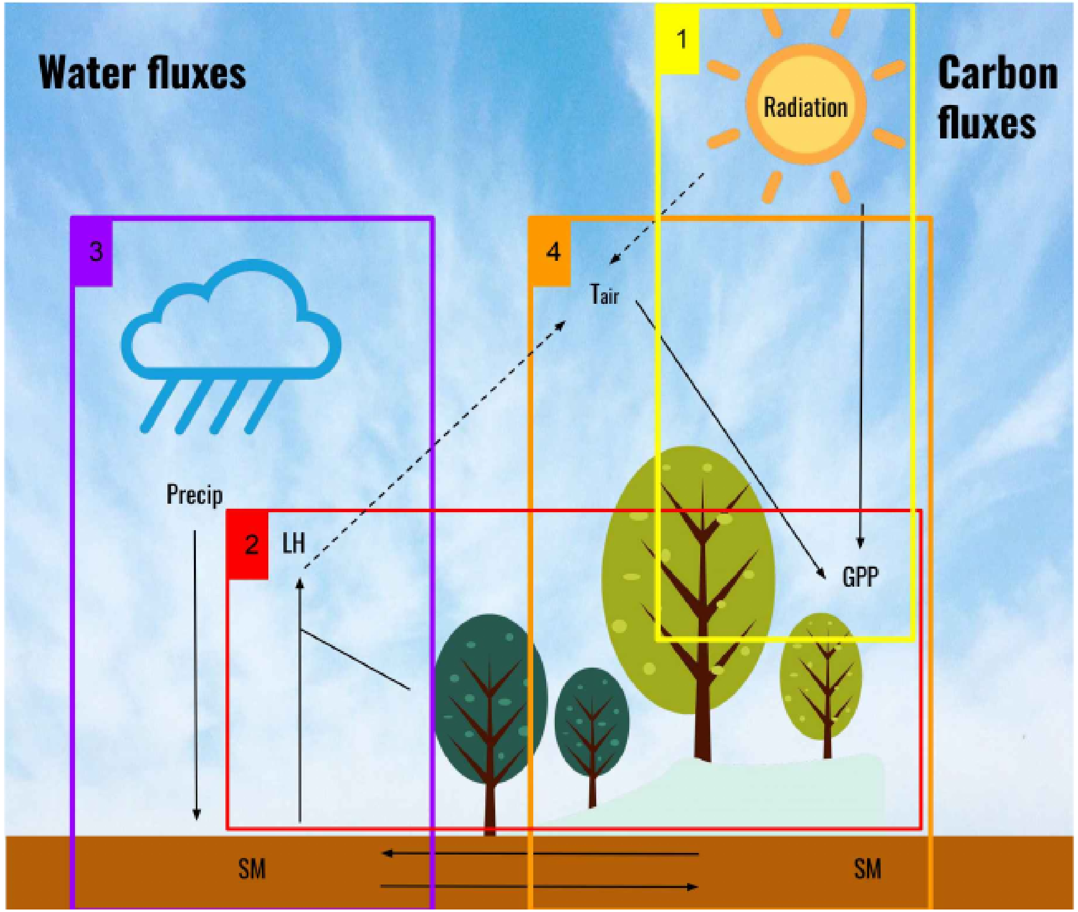
- Systematic robust estimation of critical embedding dimension E: this allows us to pool data across time, but apply algorithm “pixel-wise” to obtain spatial maps of causality

- Detect strong “instantaneous” unidirectional forcing using Information Geometric Causal Inference (IGCI). This allows us to mask spurious bidirectional causality outputs by CCM such as Radiation<->Photosynthesis due to processes occurring in time-scales not captured by sampling frequency.

- IGCI: Causal inference method for instantaneous causal relationships that works well in deterministic settings so compatible with CCM.

- IGCI: $X \rightarrow Y$ if entropy(X)>entropy(Y)

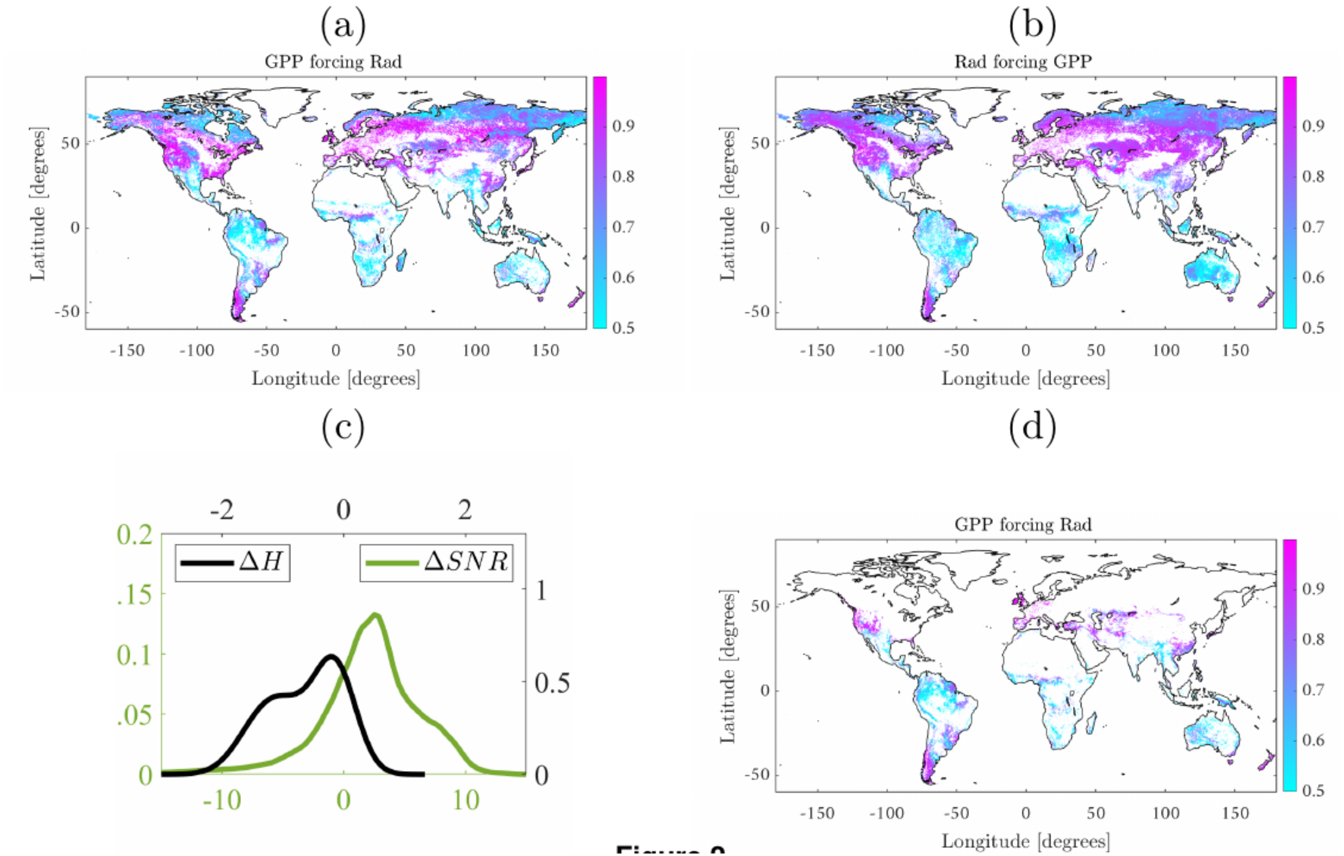
RCCM – Carbon and Water Cycle



RCCM – Data

- 6 different biosphere and atmosphere global gridded products collected and curated in the Earth System Data Lab
- GPP, SM, Tair, LH, Precip, Rad.
- 8 day temporal resoluton
- 2001-2011
- 0.0833 degrees spatial resolution

RCCM – Radiation & Photosynthesis

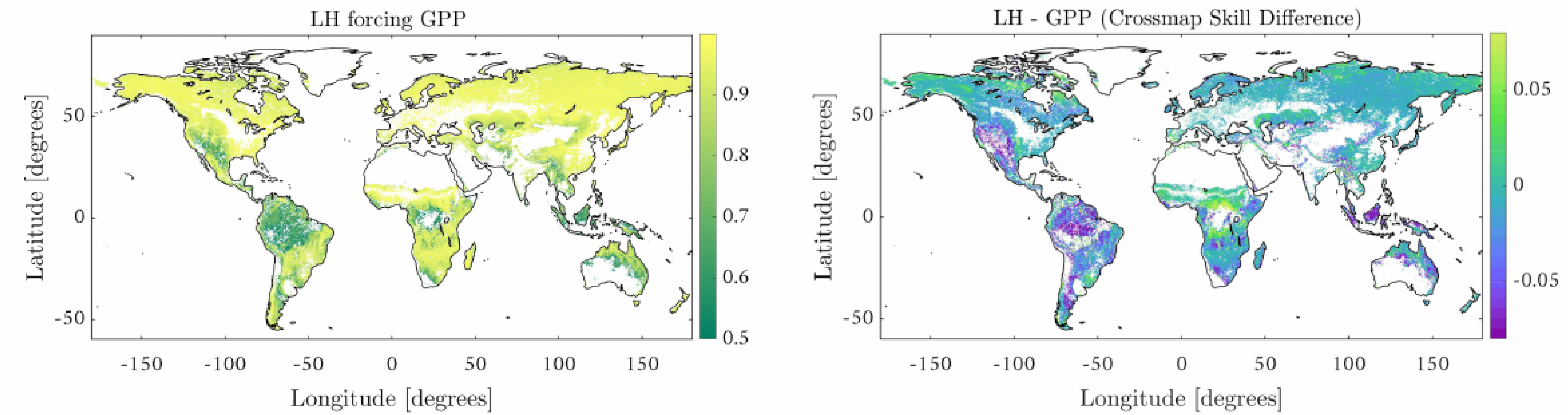


More reasonable pattern for GPP → Rad

Still GPP → Rad in tropical and cloudy regions.

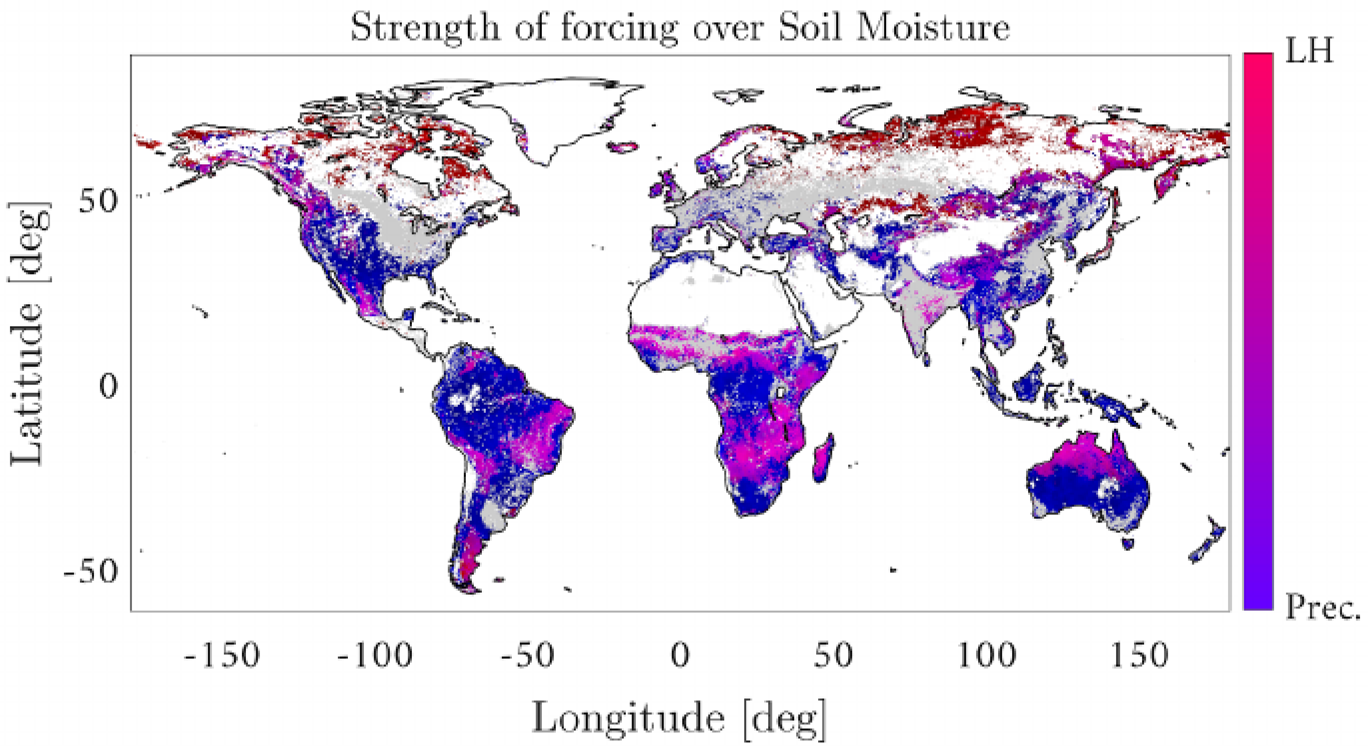
Possible mechanism: increase in GPP increases Latent Heat, moistens atmosphere affecting cloud cover

RCCM – Carbon & Latent Heat fluxes



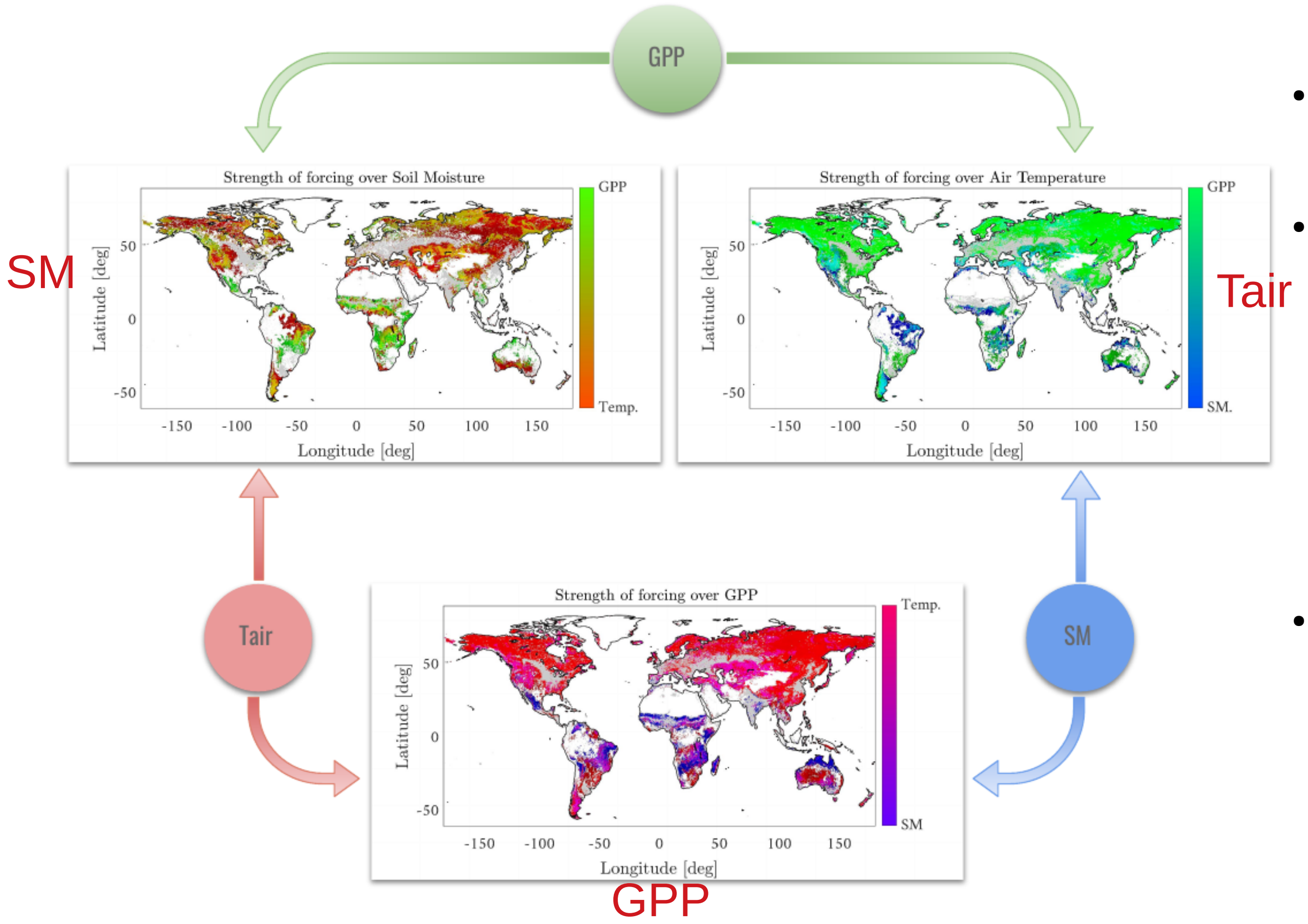
- Strong bidirectional causal influence between LH and GP
- Stronger forcing of GPP → LH in high water availability regions (eg. Amazonia)
- Stronger forcing LH → GPP in cold ecosystems and transition areas (eg African Sahel)

RCCM – Latent Heat, Precipitation & Soil Moisture



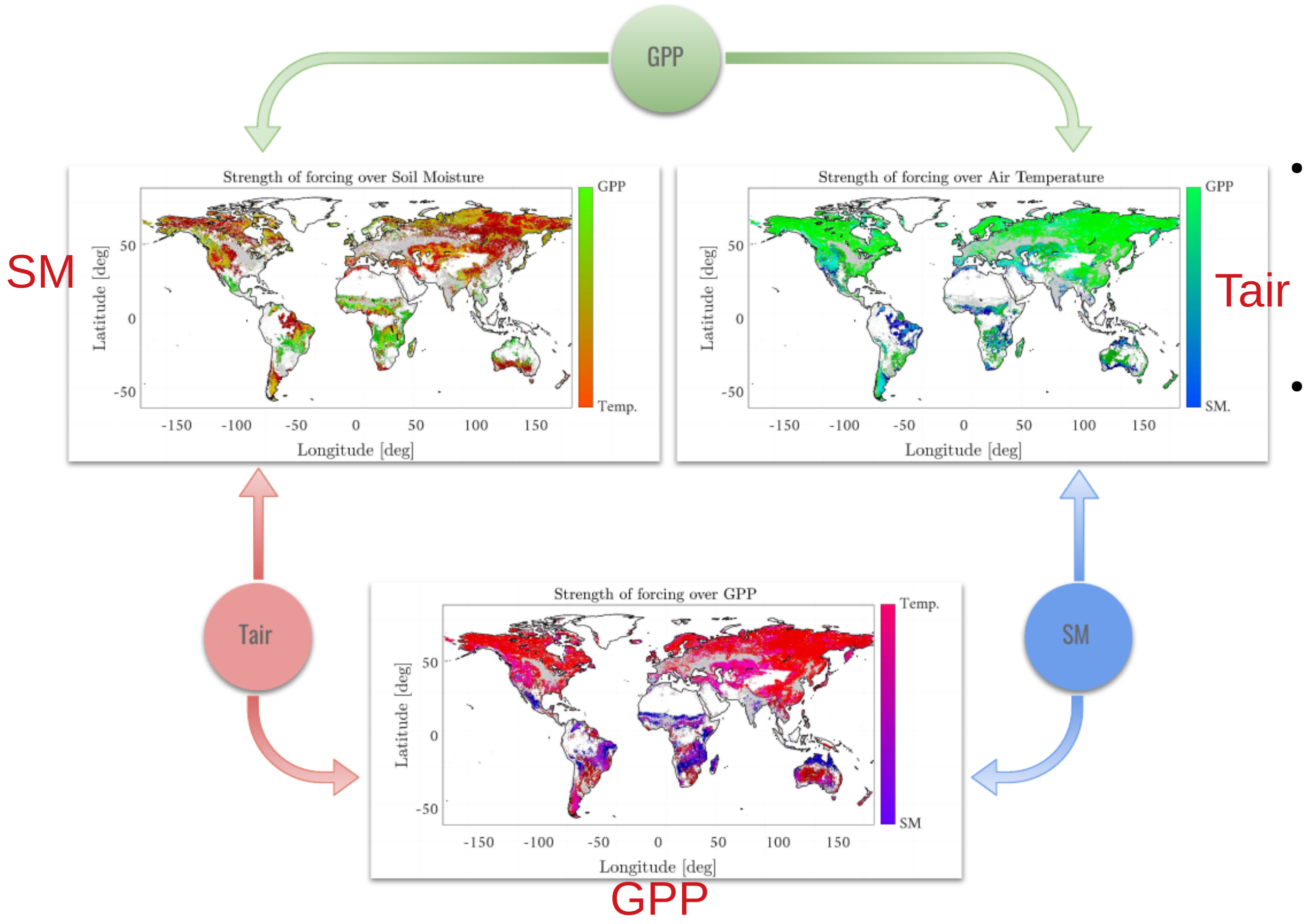
- Identify dominance of LH/Precipitation as causes of SM
- Precipitation dominant in wet tropical forests; arid & semi-arid regions such as south western US, South Africa & central Australia (bluish)
- Joint forcing in dry tropical forests (pinkish)
- LH dominant in boreal/cold ecosystems (reddish). Possibly due to soil thawing and freezing causing absorption and release of LH

RCCM – Photosynthesis, temperature & soil moisture



- Identify dominant causes in GPP-Tair-SM triad
- GPP drives Tair in many areas. In cold ecosystems possibly due to changes in land surface albedo such as snow/ice & vegetation. In warmer and drier ecosystems due to turbulent energy fluxes (enhancement of latent exchange and subsequent cooling effect)
- SM mostly controlled by Tair except in water-limited regions where GPP dominates.

RCCM – Photosynthesis, temperature & soil moisture



- Identify dominant causes in GPP-Tair-SM triad

.

- GPP dominated by Tair especially northern ecosystems where cold temp constrains photosynthesis. SM dominates in transitional regions from wet to dry climates. No strong forcings in tropical areas possibly because limiting factor is available radfiation.

Thanks for listening!

Acknowledgements & References :

- **CCM** methodology, figures and examples from: Sugihara, G. et al. "Detecting causality in complex ecosystems". Science 338, 496–500 (2012).

- **Extended CCM** methodology, figures and examples from: Ye, H., Deyle, E., J. Gilarranz, L. & Sugihara, G. "Distinguishing time-delayed causal interactions using convergent cross mapping." Sci. Reports 5, 14750, DOI: 10.1038/srep14750 (2015).

- **Robust CCM** methodology, figures and examples from IPL manuscript "Inferring causal relations from observational long-term carbon and water fluxes records" currently under revision by Sci. Reports. Authors (Emiliano Díaz, Jose E. Adsuara, Álvaro Moreno Martínez, María Piles, and Gustau Camps-Valls)

Supplementary material

CCM vs Granger Causality (GC)

GC

Causal criterion: If we can predict Y(t) better with X than without X: X→Y

Principle: Stochasticity in the process assumed so info of x imperfectly transmitted to y

Algorithm: to test x→y predict y with and without x

CCM

Causal criterion: If we can predict X(t) convergently using topology of My: X->Y

Principle: Deterministic process assumed so info of x perfeclty transmitted to y

Algorithm: to test x→y predict x using topology of My

Extended CCM

Fix General Synchrony problem

If there is a strong unidirectional forcing it “looks like” bidirectional causality using CCM

Idea: use arrow of time (past causes future) to check if bidirectional caualiity is actually unidirectional strong forcing.

Recall synchrony example where x $\rightarrow$ y:

$$X(t+1) = 3.9\, X(t)\, [1 - X(t) \qquad]$$
$$Y(t+1) = 3.7 \qquad [1 \qquad - 0.2\, X(t)]$$

$\Rightarrow$

$$0.2\, x(t) = 1 - y(t+1)/3.7$$

Note that:

- Y depends on the past of X
- X depends on the future of Y.

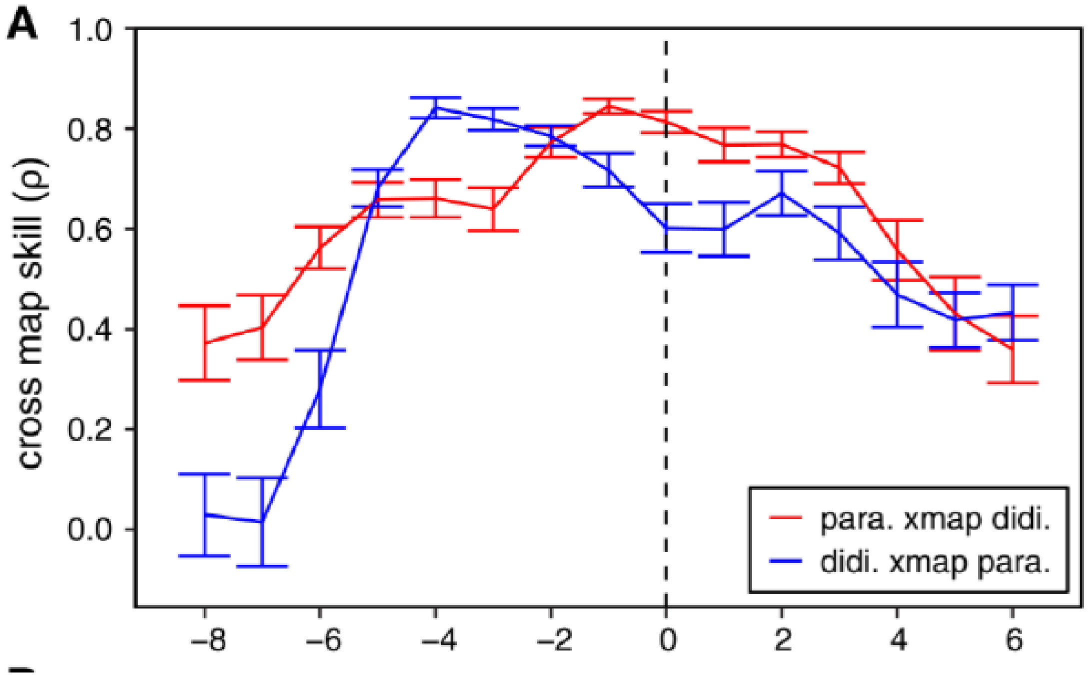
Extended CCM

Algorithm. If CCM outputs bidirectional causality repeat this time predicting y(t-L)| Mx and x(t-L)| My using different lags positive and negative lags L

For true **bidirectional** causality optimal lag L* will be negative in both cases.

For true **unidirectional** optimal lag L* will be positive (best prediction using future to predict past) for one of cases

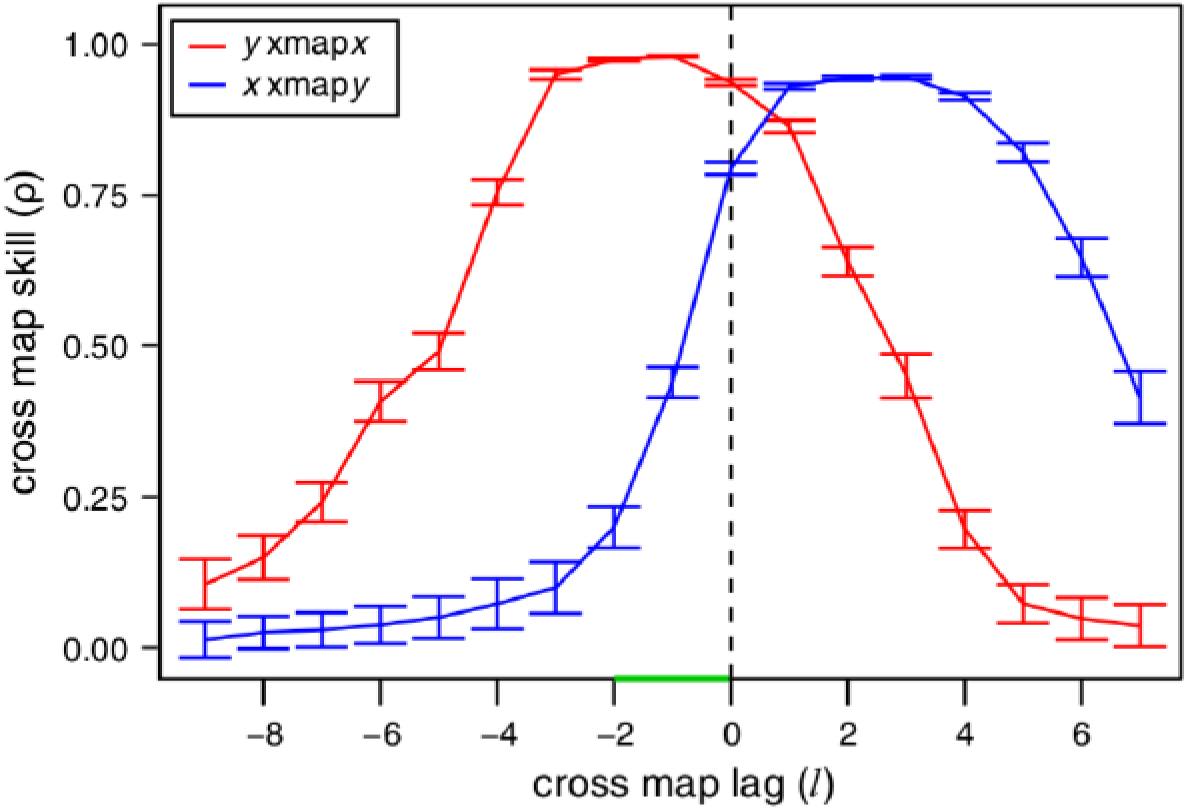
Extended CCM



As CCM suggested predator (Dinidium) causes prey (Paramecium) and vice versa.

Effect of predators on prey is immediate while that of prey on predators lagged.

Extended CCM



$$x(t+1) = x(t)[3.8 - 3.8x(t)]$$
$$y(t+1) = y(t)[3.1 - 3.1y(t) - 0.8x(t)]$$

- x → y
- Y xmap X checks x → y; negative lag implies past → future; RIGHT
- X xmap Y checks y → x; positive lag implies future → past; WRONG