Causal Inference for Earth System Science

Wenbo Lv

Xi'an, 2025. 2. 26

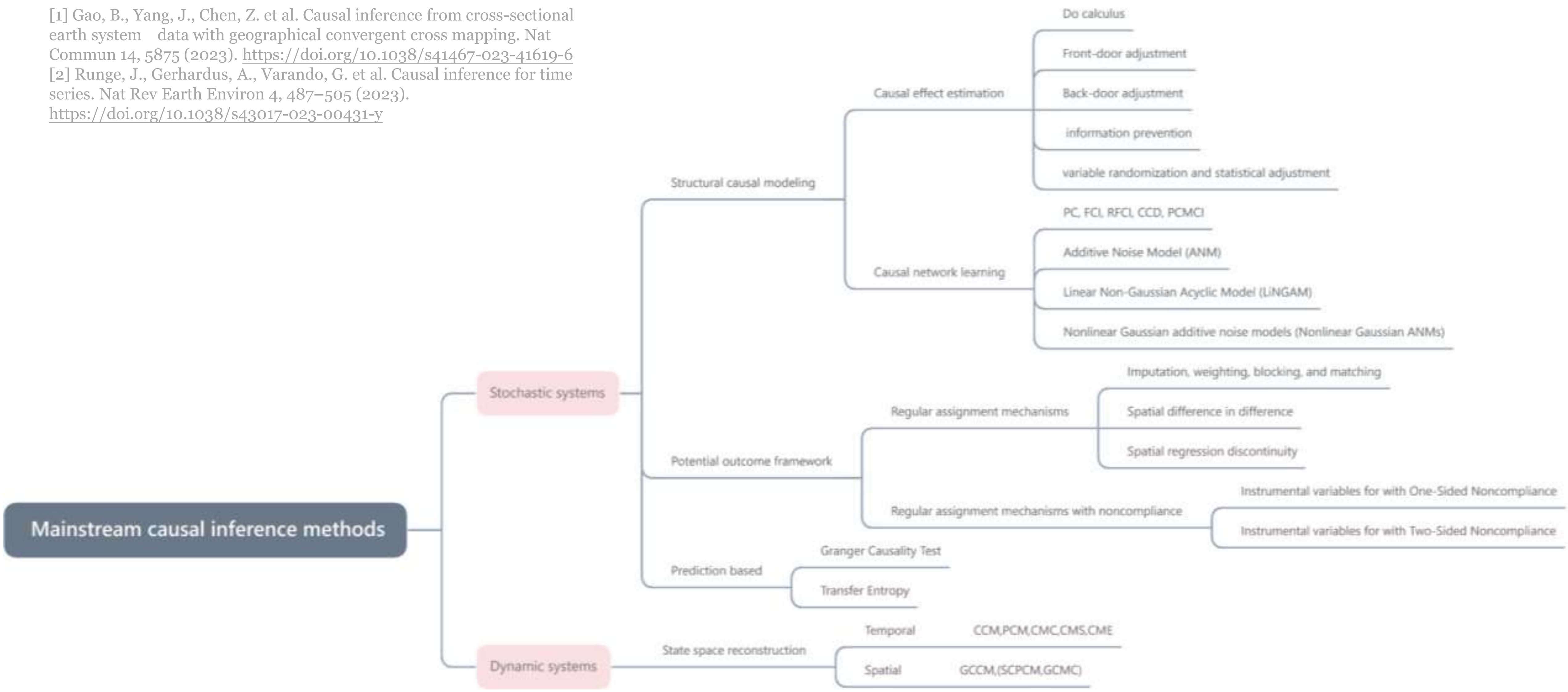
香港科技大学（广州）
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)
AI4CITY LAB
Curtin University
Geospatial Intelligence Lab

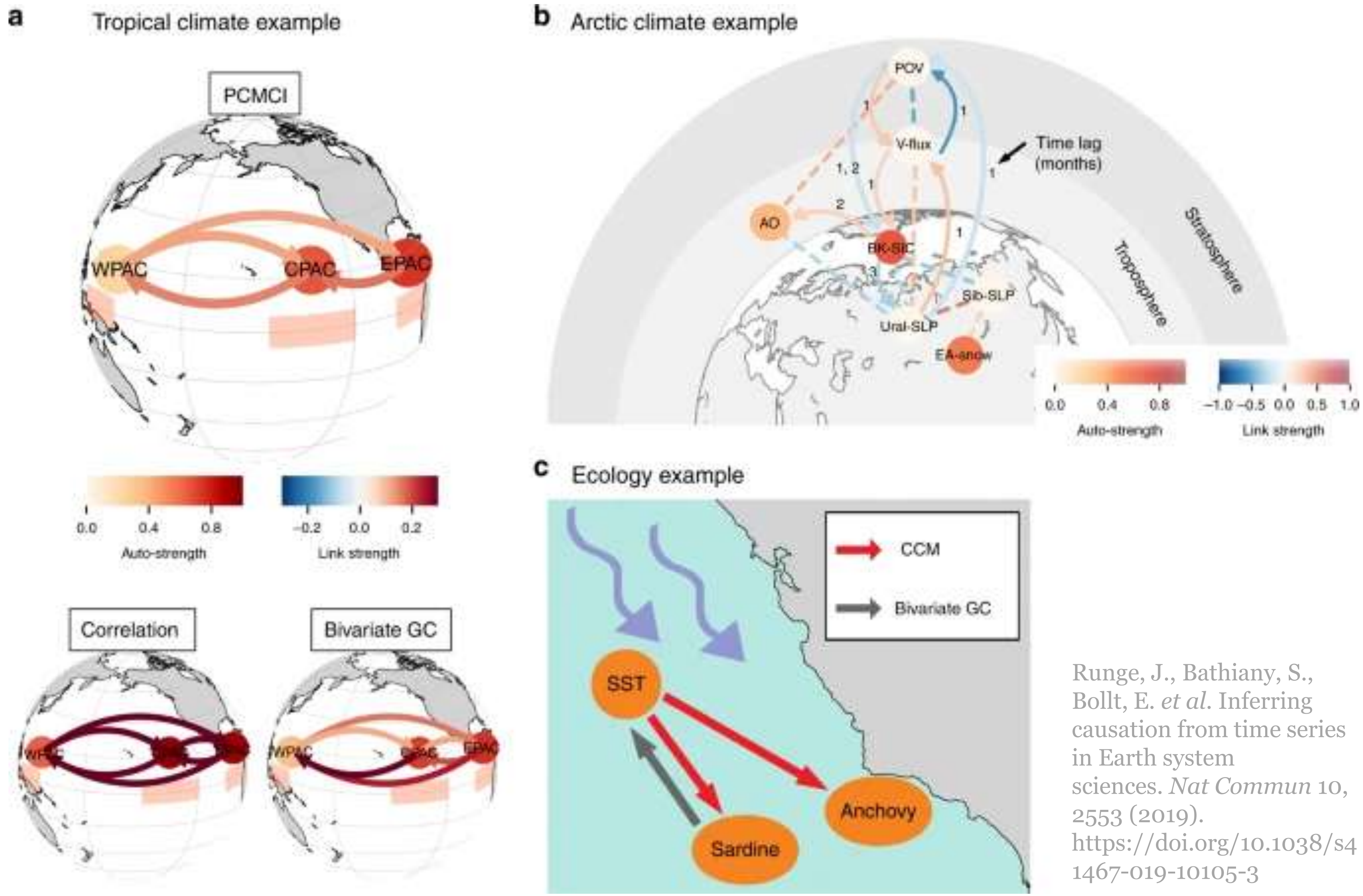
Contents:

- Data-driven algorithms vs. Earth system model simulations
- Inferring causation from time series
- Inferring causation from spatial cross sections
- Temporally or spatially?
- Spatial Empirical Dynamic Modeling

	Data-driven algorithms	Earth system model simulations
Approach	Data-centric, based on statistical methods	Mechanistic, based on physical laws
Type	Correlation-based	Causality from system dynamics
Strengths	Flexibility, applicability to complex data	Deep understanding of system behavior
Limitations	May miss underlying causal mechanisms	Computationally expensive, requires high-quality data
Example	GCT(Granger Causality Test); CCM(Convergent Cross Mapping)	CESM(Community Earth System Model);WRF(Weather Research and Forecasting Model);

[1] Gao, B., Yang, J., Chen, Z. et al. Causal inference from cross-sectional earth system data with geographical convergent cross mapping. Nat Commun 14, 5875 (2023). https://doi.org/10.1038/s41467-023-41619-6
[2] Runge, J., Gerhardus, A., Varando, G. et al. Causal inference for time series. Nat Rev Earth Environ 4, 487–505 (2023). https://doi.org/10.1038/s43017-023-00431-y

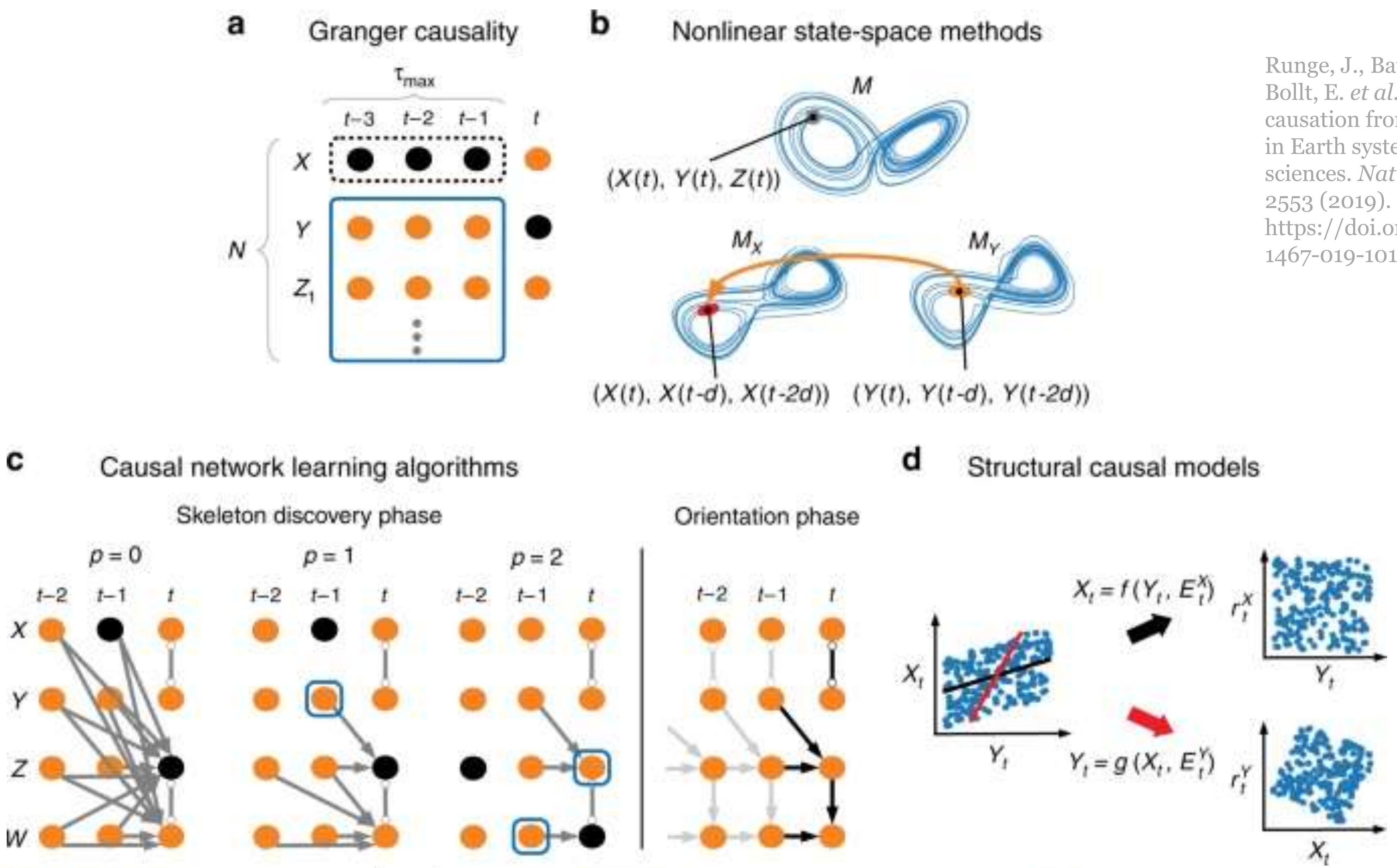




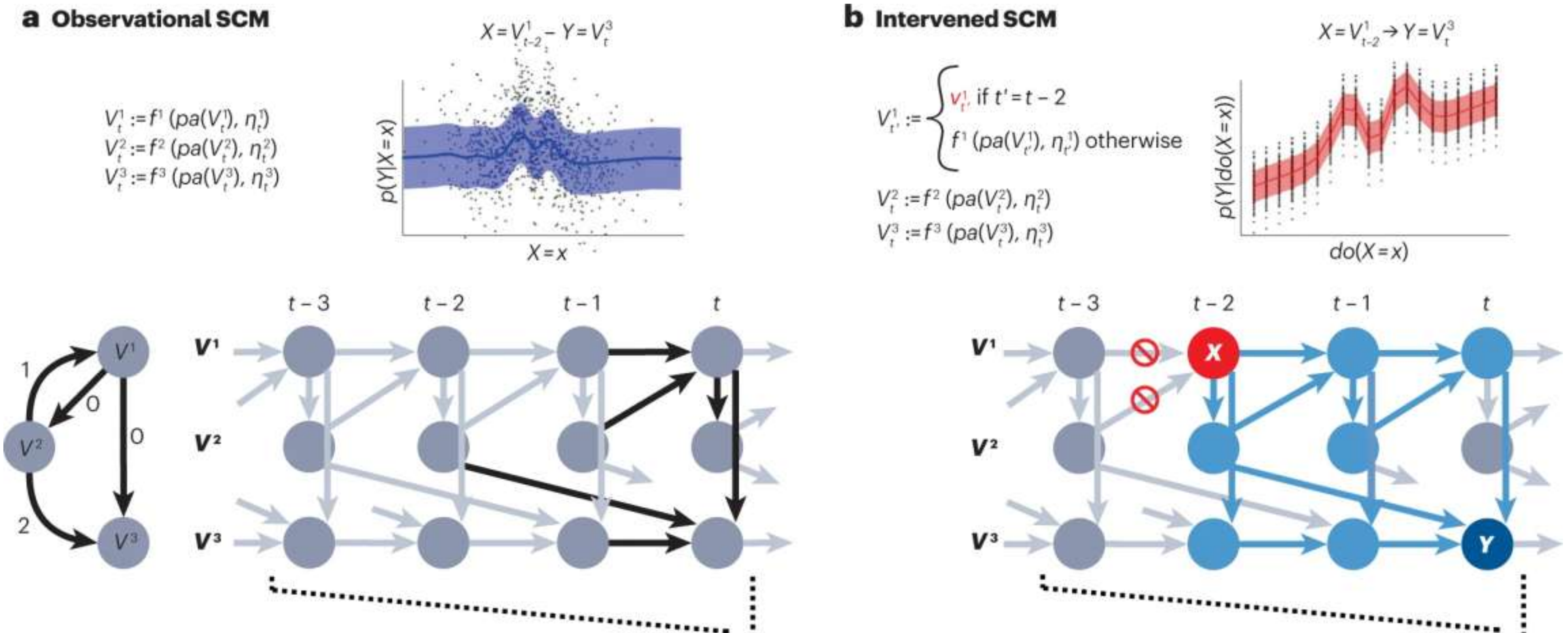
Runge, J., Bathiany, S., Bollt, E. *et al*. Inferring causation from time series in Earth system sciences. *Nat Commun* 10, 2553 (2019). https://doi.org/10.1038/s41467-019-10105-3

Inferring causation from time series

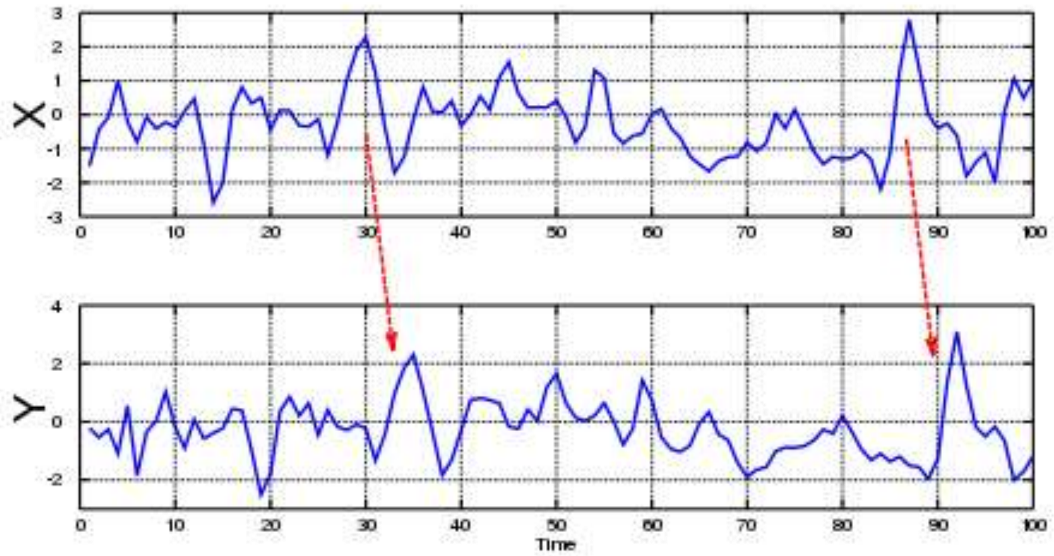
a: Tropical climate example showing dependencies between surface pressure anomalies in the West Pacific and temperature anomalies in the Central and East Pacific, with a multivariate causal method identifying the Walker circulation.
b: Arctic climate example highlighting how sea ice concentrations in the Barents and Kara seas influence the winter Arctic Oscillation via tropospheric mechanisms and the stratospheric Polar vortex.
c: Ecology example showing that sardine and anchovy abundances are influenced by sea surface temperatures, with convergent cross mapping revealing a stronger causal relationship than Granger causality.

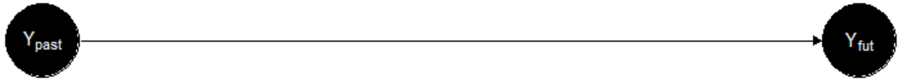


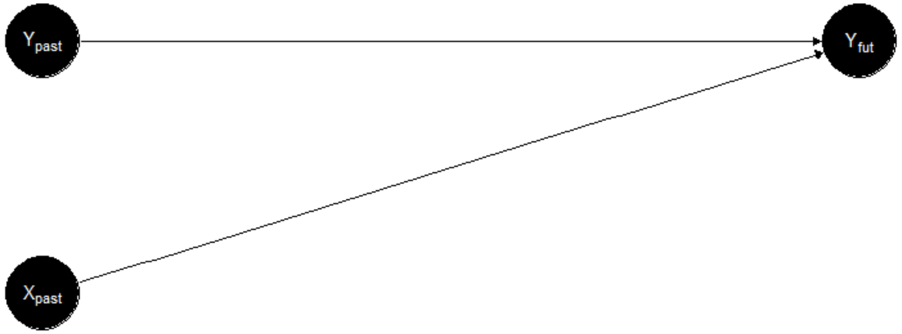
Runge, J., Bathiany, S., Bollt, E. *et al.* Inferring causation from time series in Earth system sciences. *Nat Commun* 10, 2553 (2019). https://doi.org/10.1038/s41467-019-10105-3



Runge, J., Gerhardus, A., Varando, G. et al. Causal inference for time series. Nat Rev Earth Environ 4, 487–505 (2023). https://doi.org/10.1038/s43017-023-00431-y







Granger Causality

$$Y(t) = \alpha_1 Y(t-1) + error_1(t) \quad (1)$$

$$Y(t) = \alpha_1 Y(t-1) + \beta_1 X(t-1) + error_2(t) \quad (2)$$

- A F-test is performed with *the null hypothesis* of $Y(t)$ equals to model (1) against *the alternative hypothesis* of $Y(t)$ equals to model (2).
- We say that X Granger causes Y if we reject the null hypothesis.

$$Y(t) = \sum_{j=1}^{p} \alpha_j Y(t-j) + error_1(t)$$

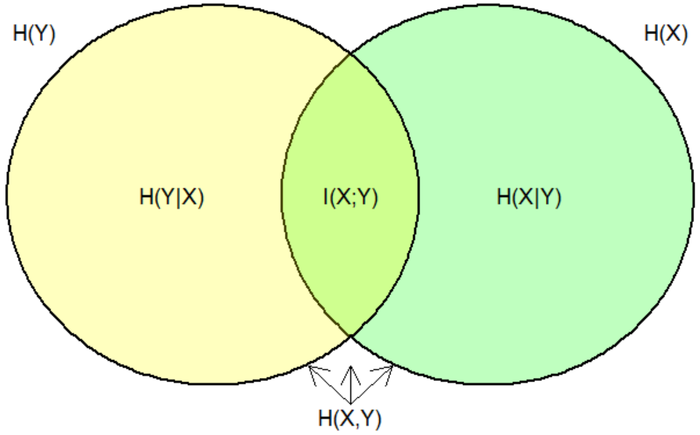
$$Y(t) = \sum_{j=1}^{p} \alpha_j Y(t-j) + \sum_{j=1}^{p} \beta_j X(t-j) + error_2(t)$$

```r
# Which came first: the chicken or the egg?
# US chicken population and egg production
# An annual time series from 1930 to 1983.
df = as.data.frame(lmtest::ChickEgg)
head(df)
#>   chicken  egg
#> 1  468491 3581
#> 2  449743 3532
#> 3  436815 3327
#> 4  444523 3255
#> 5  433937 3156
#> 6  389958 3081
```

```r
cor.test(df$chicken,df$egg)
#>
#>  Pearson's product-moment correlation
#>
#> data:  df$chicken and df$egg
#> t = -1.4327, df = 52, p-value = 0.1579
#> alternative hypothesis: true correlation
is not equal to 0
#> 95 percent confidence interval:
#>  -0.43969371  0.07689656
#> sample estimates:
#>        cor
#> -0.1948765
```

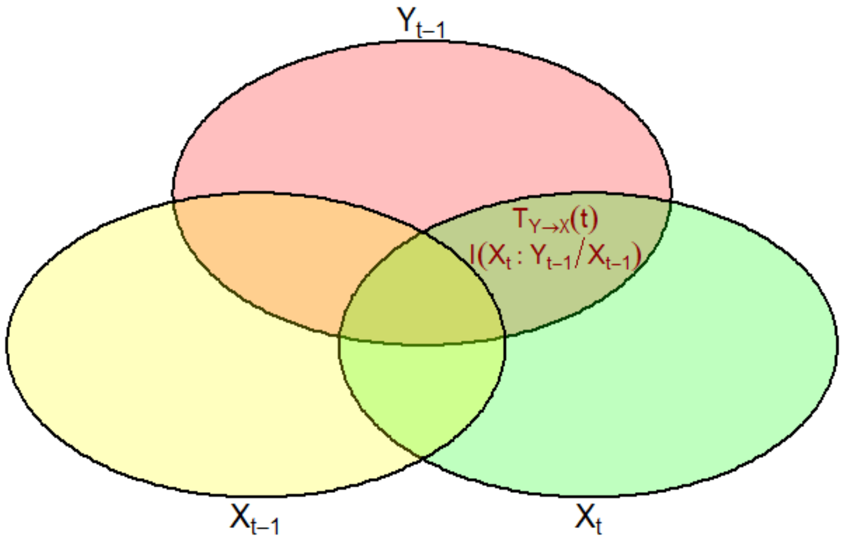
```r
#------------------- Granger Causality Test --------------------
## chickens granger-cause eggs?
lmtest::grangertest(egg ~ chicken, order = 3, data = df)
#> Granger causality test
#>
#> Model 1: egg ~ Lags(egg, 1:3) + Lags(chicken, 1:3)
#> Model 2: egg ~ Lags(egg, 1:3)
#>   Res.Df Df      F Pr(>F)
#> 1     44
#> 2     47 -3 0.5916 0.6238
## eggs granger-cause chickens?
lmtest::grangertest(chicken ~ egg, order = 3, data = df)
#> Granger causality test
#>
#> Model 1: chicken ~ Lags(chicken, 1:3) + Lags(egg, 1:3)
#> Model 2: chicken ~ Lags(chicken, 1:3)
#>   Res.Df Df     F   Pr(>F)
#> 1     44
#> 2     47 -3 5.405 0.002966 **
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Transfer Entropy



$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

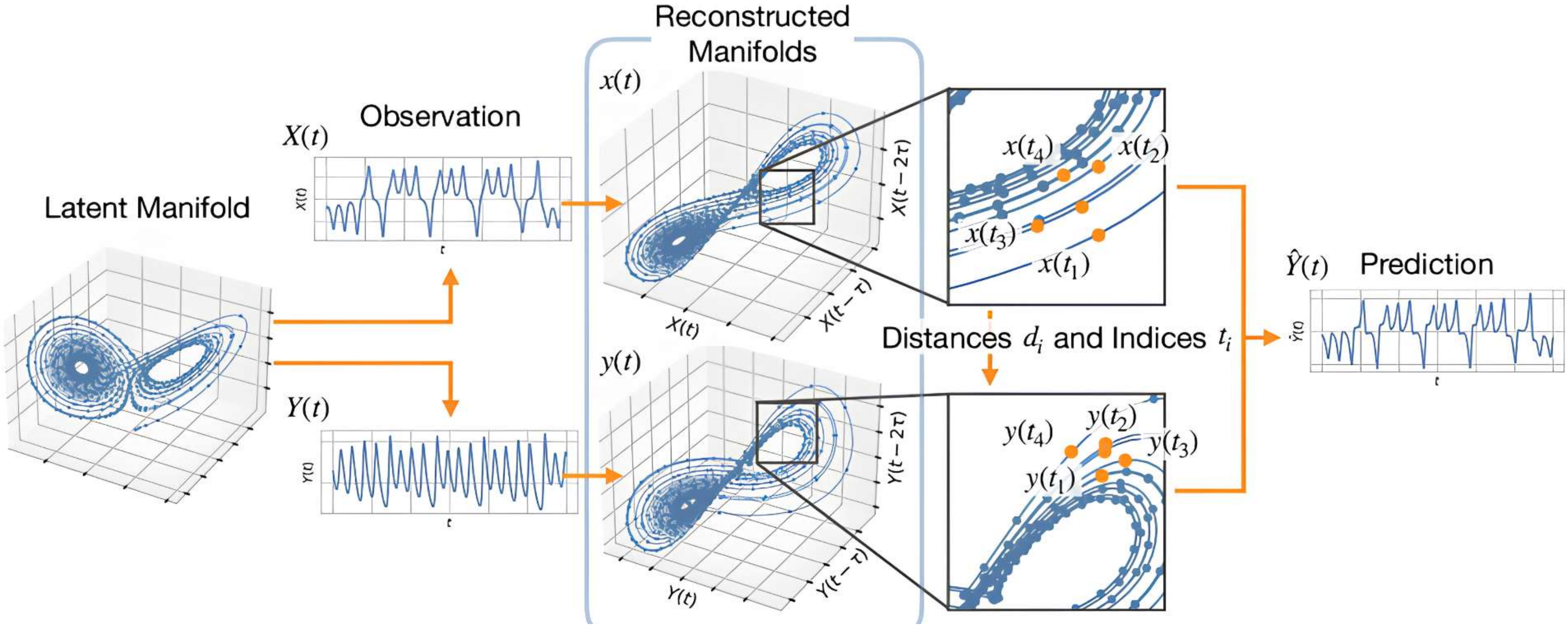
$$\begin{aligned} T_{Y \to X}(t) &= I\left(X_t : Y_{t-1} \middle| X_{t-1}\right) \\ &= H\left(X_t \middle| X_{t-1}\right) - H\left(X_t \middle| X_{t-1} Y_{t-1}'\right) \\ &= \sum p\left(x_t, x_{t-1}, y_{t-1}\right) * \log_2 \frac{p\left(x_t \middle| x_{t-1} y_{t-1}'\right)}{p\left(x_t \middle| x_{t-1}\right)} \end{aligned}$$



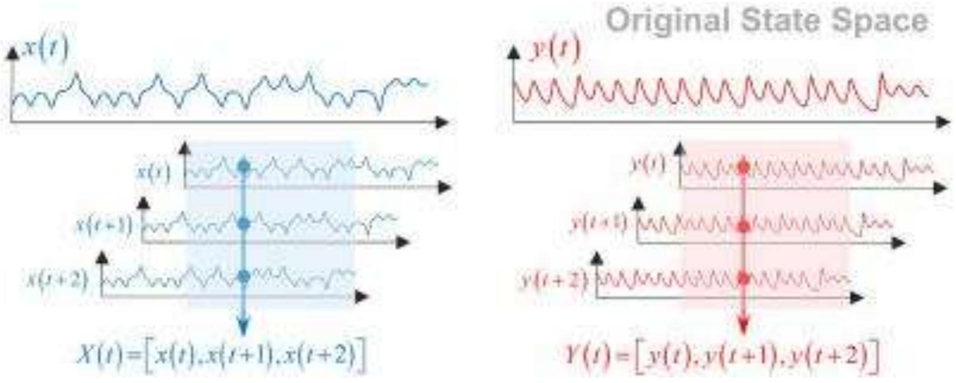
- Transfer entropy can be considered a non-parametric equivalent of Granger Causality(it also works for *nonlinear categorical variables*).
- The mutual information between both is symmetric (undirected), but the experimentally introduced time delay allows for establishing directionality.

```r
#------------------------- Transfer Entropy -----------------------------

df = as.data.frame(lmtest::ChickEgg)

# Method1: Continuous Transfer Entropy using the Kraskov estimation
## TE: chickens —> eggs
NlinTS::te_cont(df$egg, df$chicken, p = 3, q = 3, k = 6, normalize = F)
#> [1] 0.007843137
## TE: eggs -> chickens
NlinTS::te_cont(df$chicken, df$egg, p = 3, q = 3, k = 6, normalize = F)
#> [1] 0.5748366

# Method2: Pre-discretization
chicken_disc = sdsfun::discretize_vector(df$chicken,n = 5,method = 'natural')
egg_disc = sdsfun::discretize_vector(df$egg,n = 5,method = 'natural')

## TE: chickens —> eggs
NlinTS::te_disc(egg_disc, chicken_disc, p = 3, q = 3, normalize = TRUE)
#> [1] 0.2670127
## TE: eggs -> chickens
NlinTS::te_disc(chicken_disc, egg_disc, p = 3, q = 3, normalize = TRUE)
#> [1] 0.2745098
```

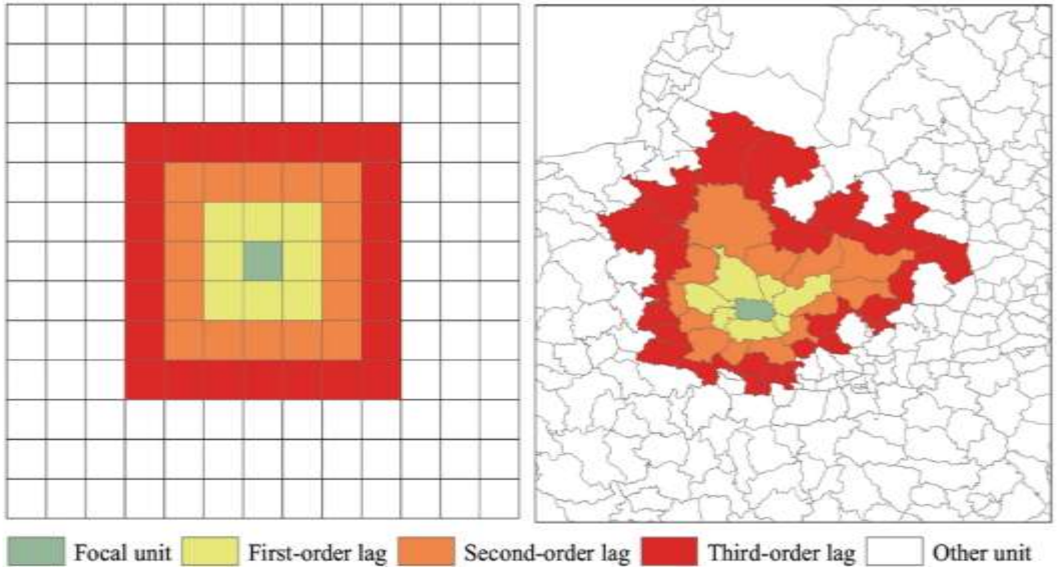
Cross Mapping



Time series and spatial cross-sectional embeddings

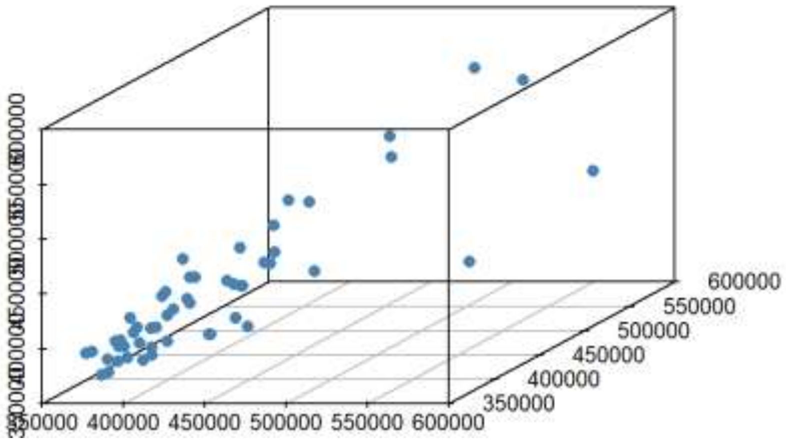


$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_E \\ x_2 & x_3 & \cdots & x_{E+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-E+1} & x_{n-E+2} & \cdots & x_n \end{bmatrix}$$

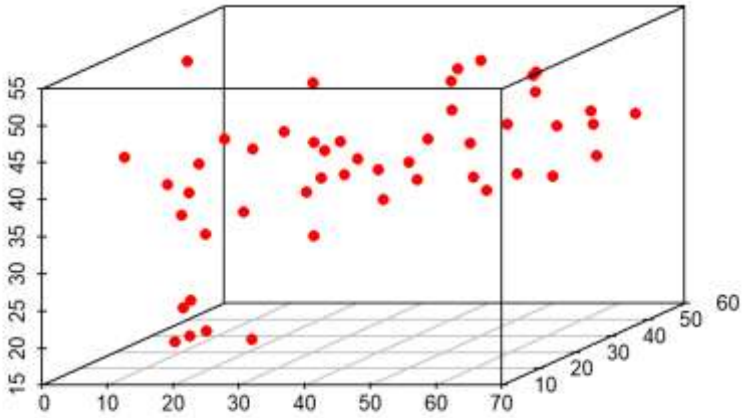


$$\mathbf{X} = \begin{bmatrix} \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j}^{(1)} & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j}^{(2)} & \cdots & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j}^{(k)} \\ \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j+1}^{(1)} & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j+1}^{(2)} & \cdots & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{i-j+1}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{n-m}^{(1)} & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{n-m}^{(2)} & \cdots & \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_{n-m}^{(k)} \end{bmatrix}$$

```r
# temporal embeddings:
chickegg = as.data.frame(lmtest::ChickEgg)
m1 = stats::embed(chickegg$chicken,dimension = 3)
head(m1,5)
#>        [,1]   [,2]   [,3]
#> [1,] 436815 449743 468491
#> [2,] 444523 436815 449743
#> [3,] 433937 444523 436815
#> [4,] 389958 433937 444523
#> [5,] 403446 389958 433937
```




```r
# spatial embeddings:
columbus = sf::read_sf(system.file("shapes/columbus.gpkg", package="spData"))
m2 = spEDM::embedded(columbus,target = "CRIME", E = 3, tau = 0)
head(m2,5)
#>          [,1]     [,2]     [,3]
#> [1,] 15.72598 24.71427 41.55964
#> [2,] 18.80175 26.24684 44.57868
#> [3,] 30.62678 29.41175 43.45135
#> [4,] 32.38776 34.64648 37.95647
#> [5,] 50.73151 40.46533 34.43484
```



Simplex Projection

$$\text{distances}_j = \sqrt{\sum_{k=1}^{E} \left(x_{i+k-1} - x_{j+k-1}\right)^2}$$

$$\text{neighbors} = \text{order}(\text{distances})[2:(E+1)]$$

$$w_j = \exp\left(-\frac{\text{distance}_j}{\text{min.distance}}\right)$$

$$\hat{x}_{i+1} = \frac{\sum_{j \in \text{neighbors}} w_j \cdot x_{j+1}}{\sum_{j \in \text{neighbors}} w_j}$$

S-Mapping

$$\text{distances}_j = \sqrt{\sum_{k=1}^{E} \left(x_{i+k-1} - x_{j+k-1}\right)^2}$$

$$w_j = \exp\left(-\frac{\theta \cdot \text{distance}_j}{\overline{\text{distance}}}\right)$$

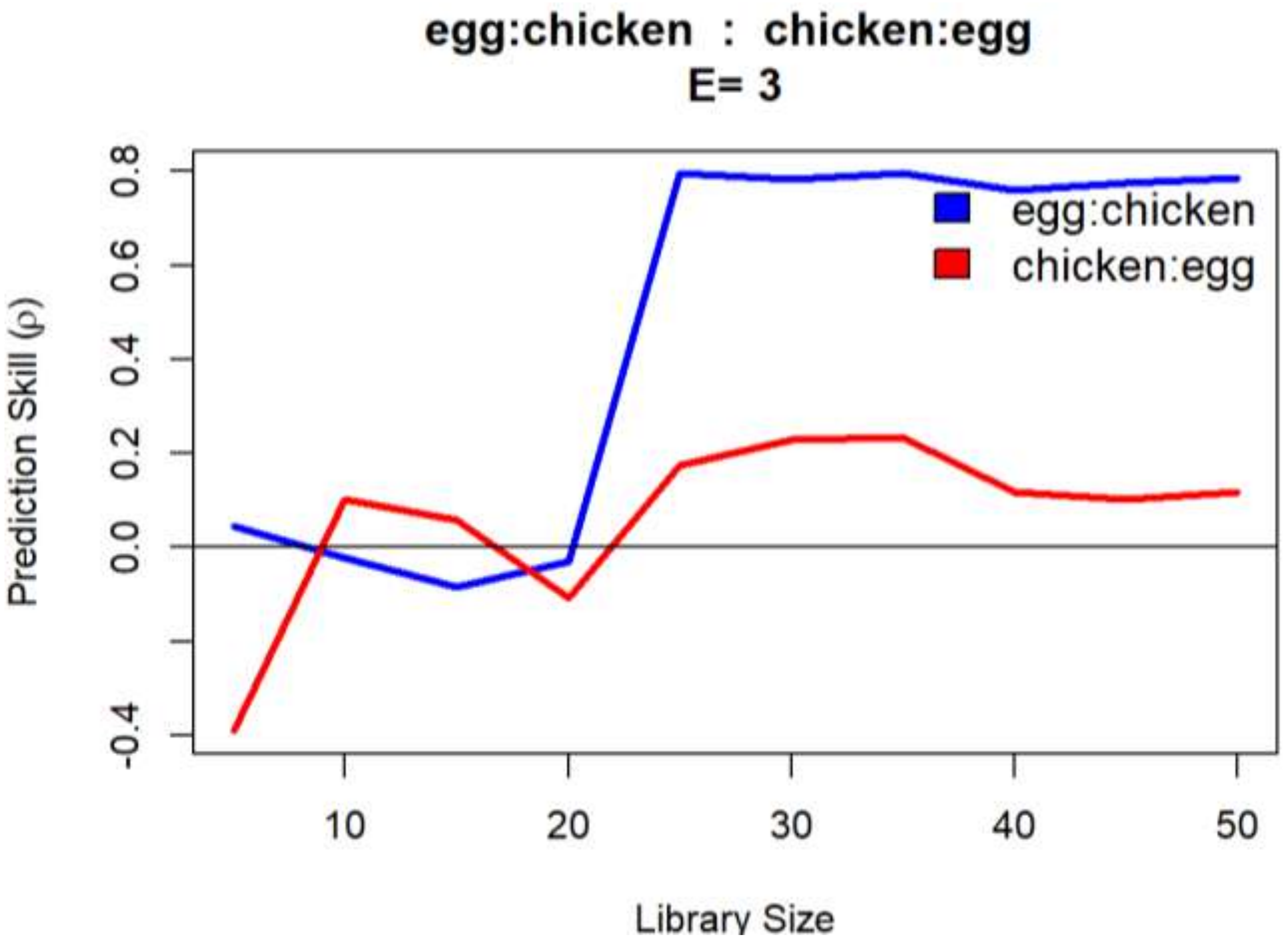
$$\mathbf{A} = [\mathbf{X}, \mathbf{w}]$$

$$\mathbf{map} = V \cdot \Sigma^{-1} \cdot U^T$$

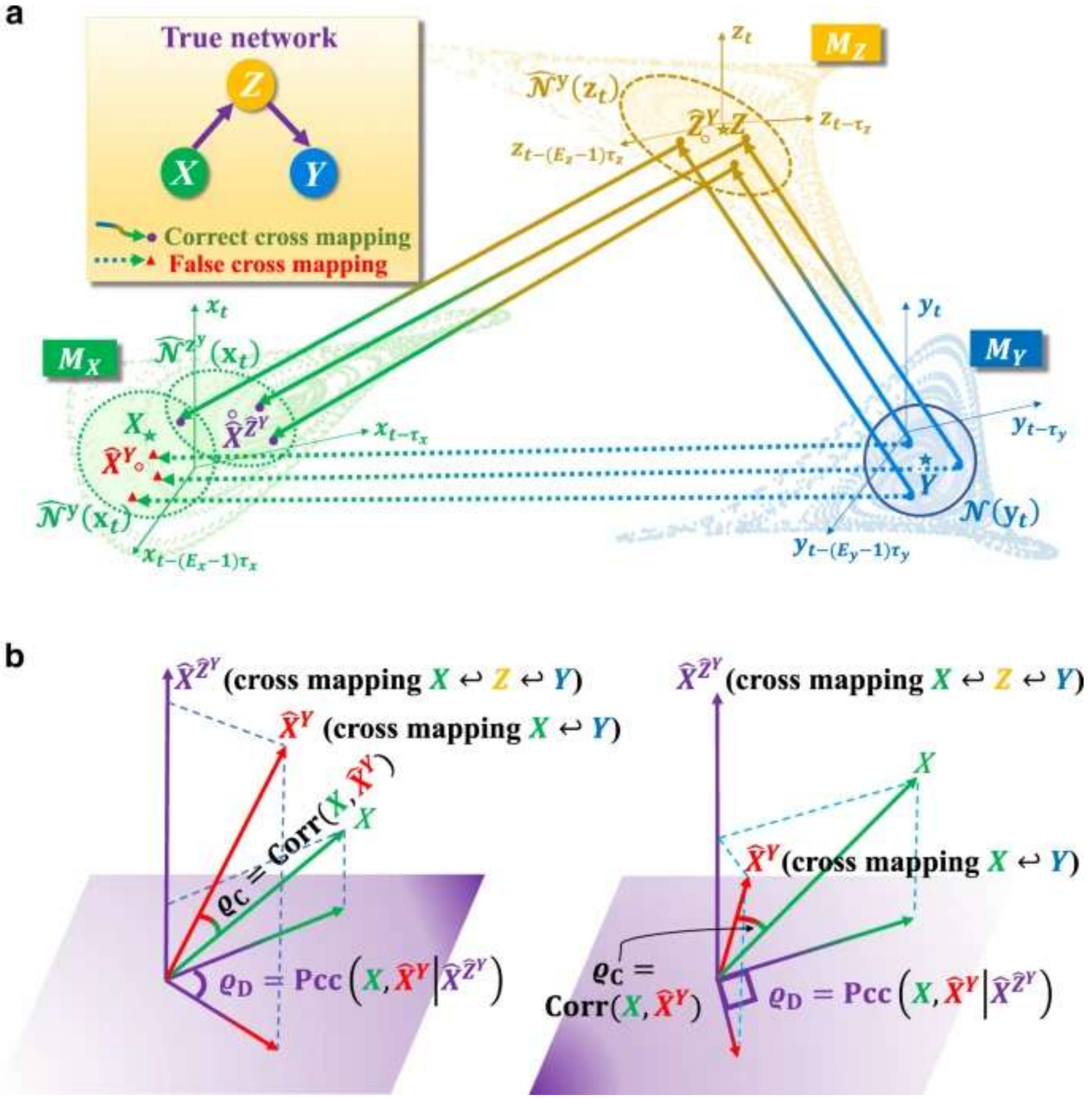
$$\hat{x}_{i+1} = \mathbf{map}^T \cdot [\mathbf{x}_i, 1]$$

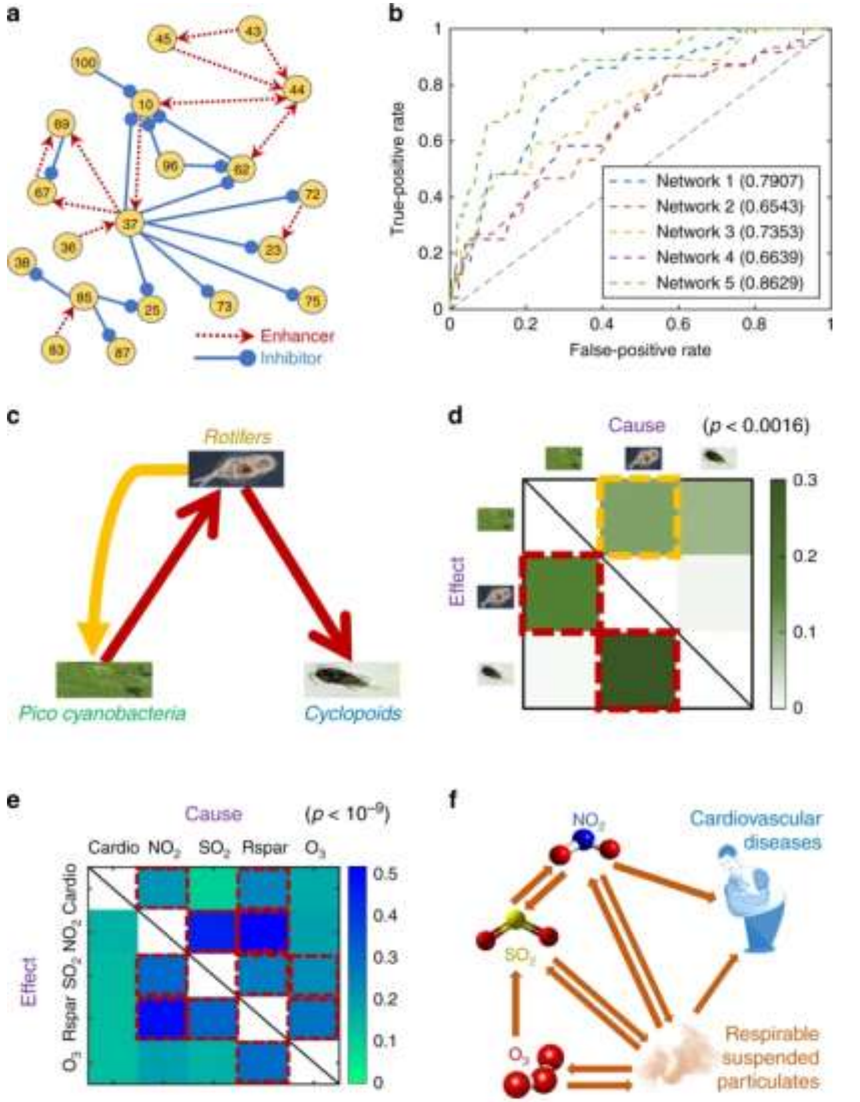
```r
# simplex proojection
spEDM::RcppSimplexForecast(m2,columbus$CRIME,1:49,1:49,4)
#>  [1] 17.70104 19.93840 30.16243 33.07348 52.18022 30.18953 13.01769 38.77808
#>  [9] 32.23513 31.48435 59.09663 56.54921 52.37962 55.85027 45.97856 56.90080
#> [17] 32.00166 42.04745 57.18469 11.72599 40.22695 33.58007 19.03443 43.43395
#> [25] 61.19662 41.27700 47.49645 57.54308 60.16271 60.75822 16.88356 18.35503
#> [33] 38.45912 19.50335 36.89147 16.98435 42.51416 53.16424 17.19591 16.21441
#> [41] 17.43940 17.79455 35.36479 30.37855 27.84116 18.27504 18.16851 26.01435
#> [49] 22.38665

# s-mapping
spEDM::RcppSMapForecast(m2,columbus$CRIME,1:49,1:49,4,theta = 0.1)
#>  [1]  1.627749e+02  1.386350e+03  1.322136e+03  1.712858e+03 -3.264970e+03
#>  [6] -4.801324e+02  4.574963e+02 -3.206982e+03 -1.089545e+03  3.597287e+02
#> [11] -6.972924e+01  4.814860e+03 -2.115798e+02  3.454839e+02 -1.318667e+03
#> [16]  6.773236e+02 -4.492509e+02  1.091786e+03 -2.059843e+02  2.330649e+02
#> [21] -1.940220e+02  7.405511e+02 -9.607234e+02 -9.933905e+02  3.984855e+03
#> [26]  4.145653e+02 -3.183077e+03 -1.776549e+03  8.340536e+02  2.323735e+03
#> [31]  3.861250e+02 -4.719258e+02 -1.072537e+03  2.790086e-02  5.645311e+03
#> [36]  2.159156e+02  2.532053e+03 -3.401927e+03 -4.756470e+02 -1.735974e+02
#> [41]  1.671329e+03 -3.355987e+01 -3.302869e+02 -1.042772e+03  9.089417e+02
#> [46] -5.821938e+02 -7.604911e+01  1.389021e+03 -1.093324e+04
```

```r
# convergent cross-mapping
chickegg = as.data.frame(lmtest::ChickEgg)
ccmres = rEDM::CCM(dataFrame = chickegg, E = 3, columns = "egg", target = "chicken",
                   libSizes = "5 50 5", random = FALSE, noTime = TRUE)
ccmres
#>    LibSize egg:chicken chicken:egg
#> 1        5  0.04364260 -0.39121891
#> 2       10 -0.02381291  0.10223896
#> 3       15 -0.08433760  0.05800941
#> 4       20 -0.03005538 -0.10681737
#> 5       25  0.79516487  0.17436527
#> 6       30  0.78193136  0.22865146
#> 7       35  0.79522593  0.23369108
#> 8       40  0.75874122  0.11634096
#> 9       45  0.77645014  0.10142964
#> 10      50  0.78570387  0.11601891
```



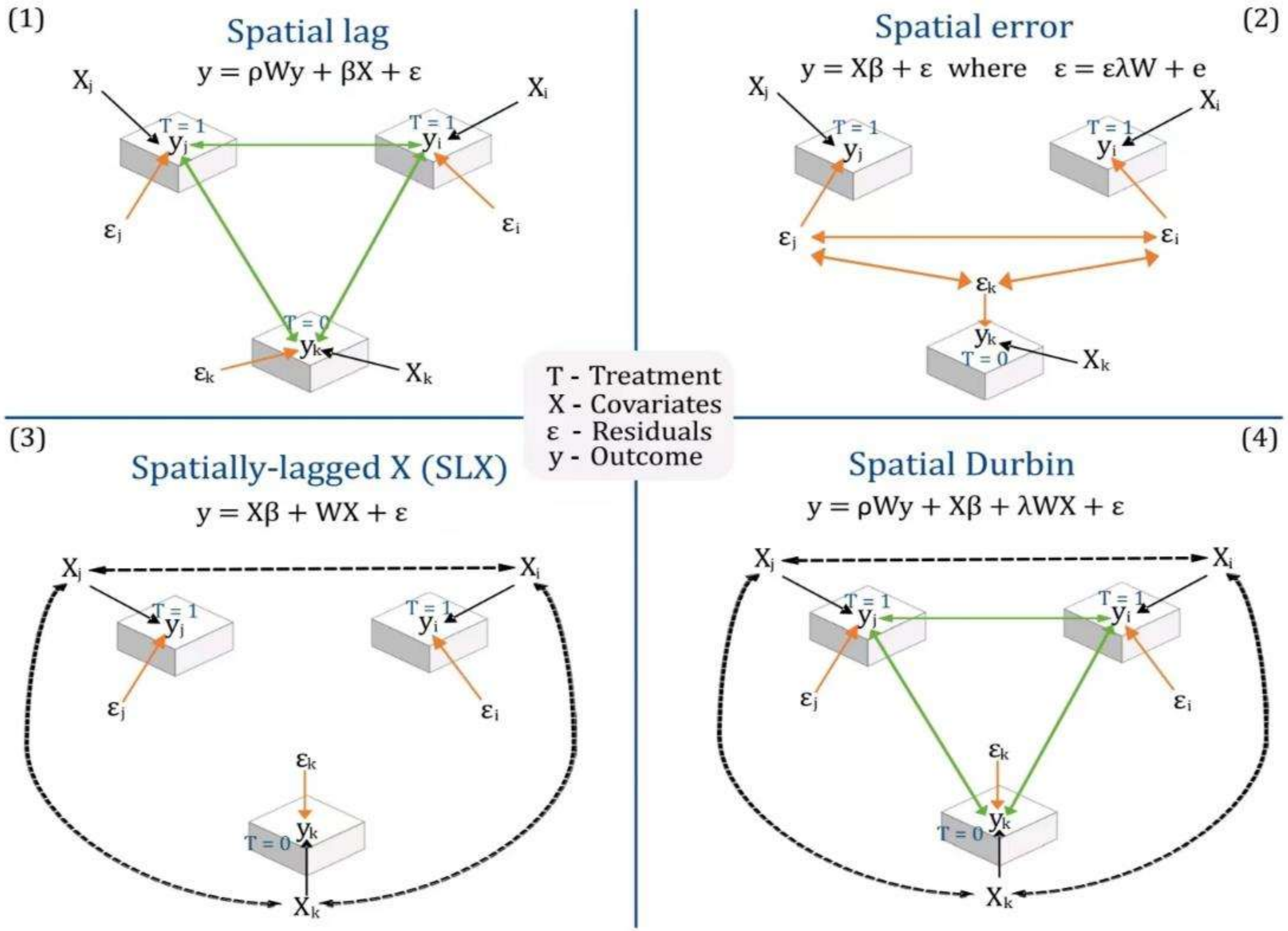
Partial Cross Mapping





Leng, S., Ma, H., Kurths, J. *et al*. Partial cross mapping eliminates indirect causal influences. *Nat Commun* **11**, 2632 (2020). https://doi.org/10.1038/s41467-020-16238-0

Inferring causation from spatial cross sections



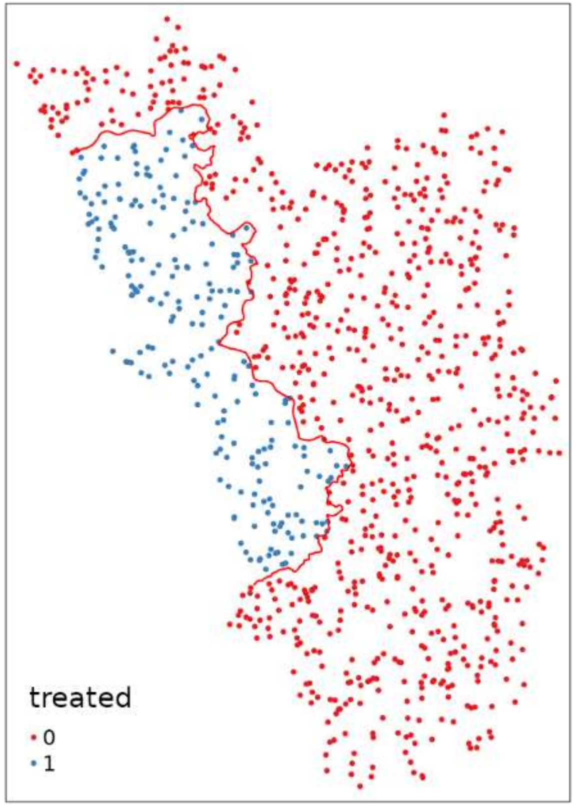
Spatial Difference-in-differences

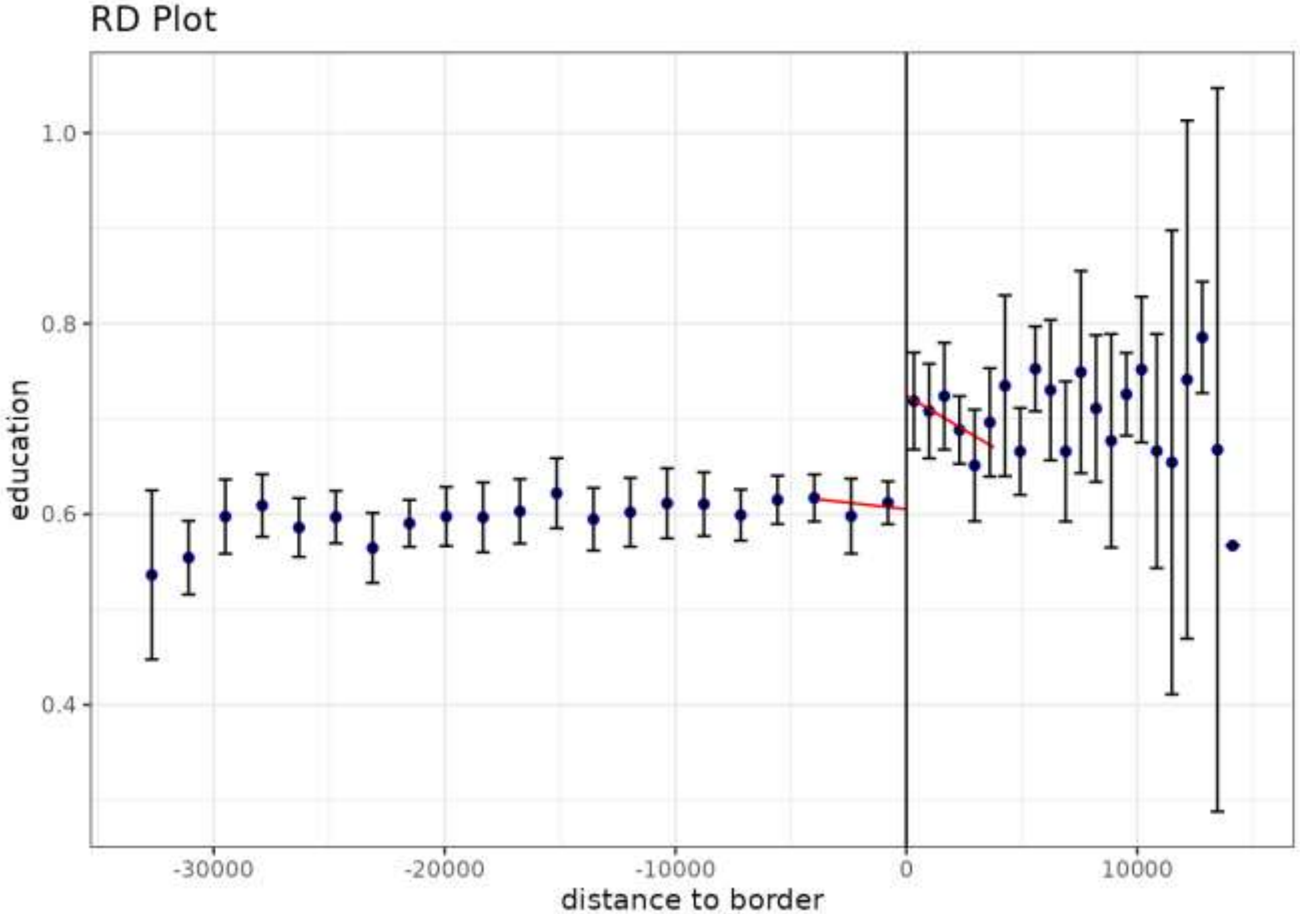
$$y_it = \alpha_0 + \alpha_1 X_it + \alpha_2 D_it + \alpha_3 T_it + \alpha_4 D_itT_it + \varepsilon_it$$

$$\begin{aligned} ATE(x) = \; & E[y|X = x, D = 1, T = 1] - \\ & E[y|X = x, D = 0, T = 1] - \\ & E[y|X = x, D = 1, T = 0] - \\ & E[y|X = x, D = 0, T = 0]. \end{aligned}$$

Akbari, K., Winter, S. and Tomko, M. (2023), Spatial Causality: A Systematic Review on Spatial Causal Inference. Geogr Anal, 55: 56-89. https://doi.org/10.1111/gean.12312

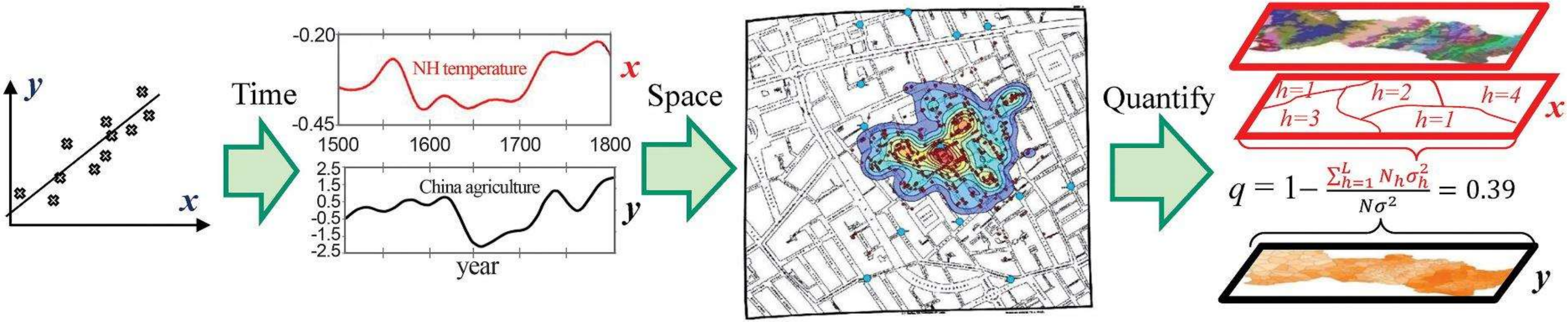
Spatial Regression Discontinuity Designs





$$y_i = \beta_0 + \beta_1 D_i + \beta_2 \text{dist}_i + \beta_3 D_i \times \text{dist}_i + \sum_{s=1}^{S} \gamma_b \, \text{seg}_i^s + v_i{}' \phi + \epsilon_i$$

Spatial Stratified Heterogeneity



a. Correlation between 2 pairs of points

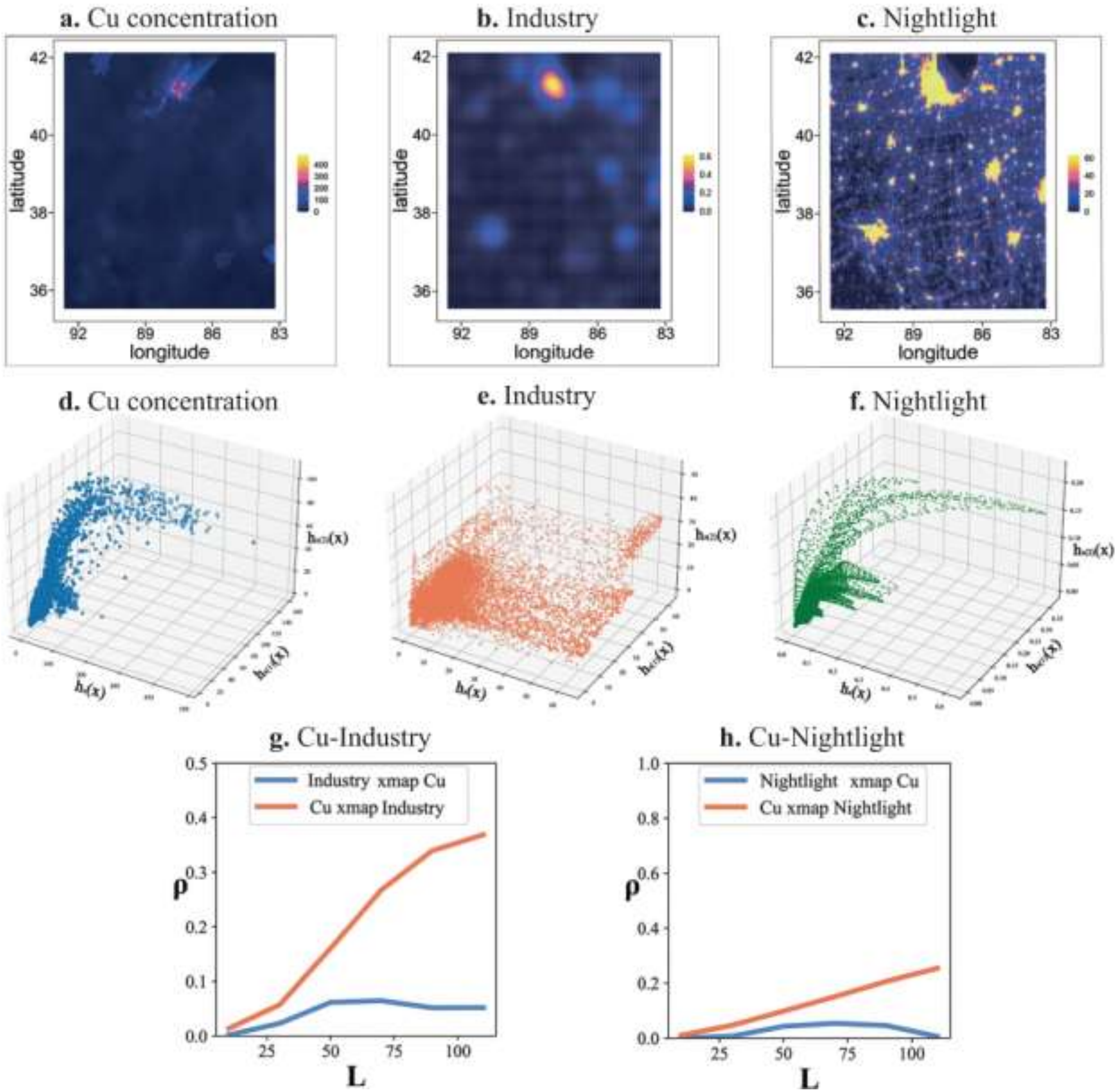
b. Consistency & causality between two time series

c. Coupling & causality between two spatial patterns

d. Coupling & causality between two spatial patterns

Wang, J., Haining, R., Zhang, T., Xu, C., Hu, M., Yin, Q., ... Chen, H. (2024). Statistical Modeling of Spatially Stratified Heterogeneous Data. *Annals of the American Association of Geographers*, *114*(3), 499–519. https://doi.org/10.1080/24694452.2023.2289982

```r
# geographical detector
columbus = sf::read_sf(system.file("shapes/columbus.gpkg", package="spData"))
gdverse::opgd("HOVAL ~ CRIME", data = columbus, discnum = 3:15)
#> ***   Optimal Parameters-based Geographical Detector
#>                      Factor Detector
#>
#> | variable | Q-statistic |  P-value   |
#> |:--------:|:-----------:|:----------:|
#> |  CRIME   |  0.5885002  | 0.03009555 |
gdverse::opgd("CRIME ~ HOVAL", data = columbus, discnum = 3:15)
#> ***   Optimal Parameters-based Geographical Detector
#>                      Factor Detector
#>
#> | variable | Q-statistic |   P-value   |
#> |:--------:|:-----------:|:-----------:|
#> |  HOVAL   |  0.627518   | 0.001022616 |
```



Geographical Convergent Cross Mapping

$$\hat{Y}_s \mid M_x = \sum_{i=1}^{L+1} (\omega_{si} Y_{si} \mid M_x)$$

$$dis(\psi(x, s_i), \psi(x, s)) = \frac{1}{L}\left(|h_{si}(x) - h_s(x)| + \sum_{k=1}^{L-1} abs[h_{si(k)}(x), h_{s(k)}(x)] \right)$$

$$weight(\psi(x, s_i), \psi(x, s)) = \exp\left(- \frac{dis(\psi(x, s_i), \psi(x, s))}{dis(\psi(x, s_1), \psi(x, s))} \right)$$

$$\omega_{si} \mid M_x = \frac{weight(\psi(x, s_i), \psi(x, s))}{\sum_{i=1}^{L+1} weight(\psi(x, s_i), \psi(x, s))}$$

$$\rho = \frac{Cov(Y, \hat{Y})}{\sqrt{Var(Y)Var(\hat{Y})}}$$

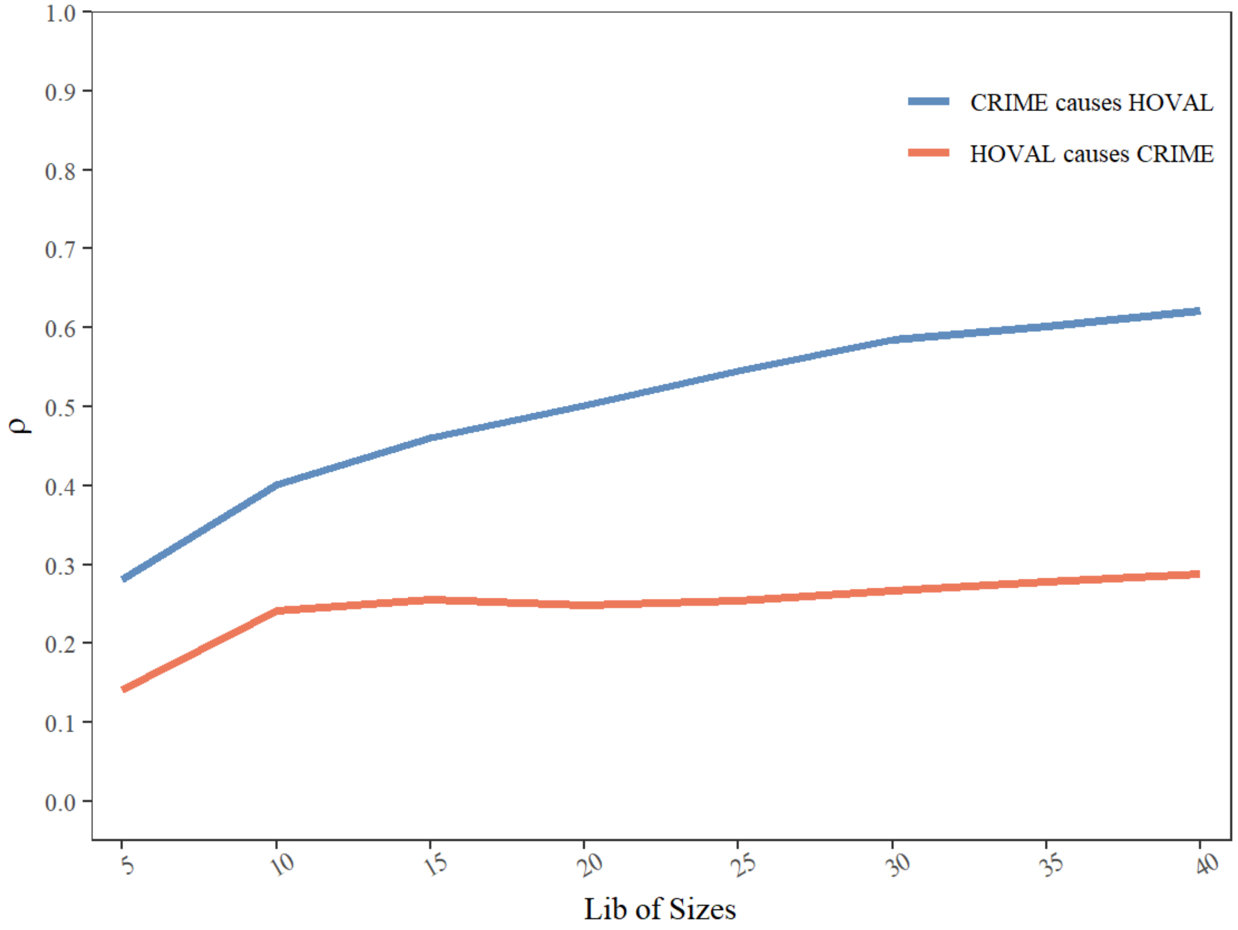
Gao, B., Yang, J., Chen, Z. et al. Causal inference from cross-sectional earth system data with geographical convergent cross mapping. Nat Commun 14, 5875 (2023). https://doi.org/10.1038/s41467-023-41619-6

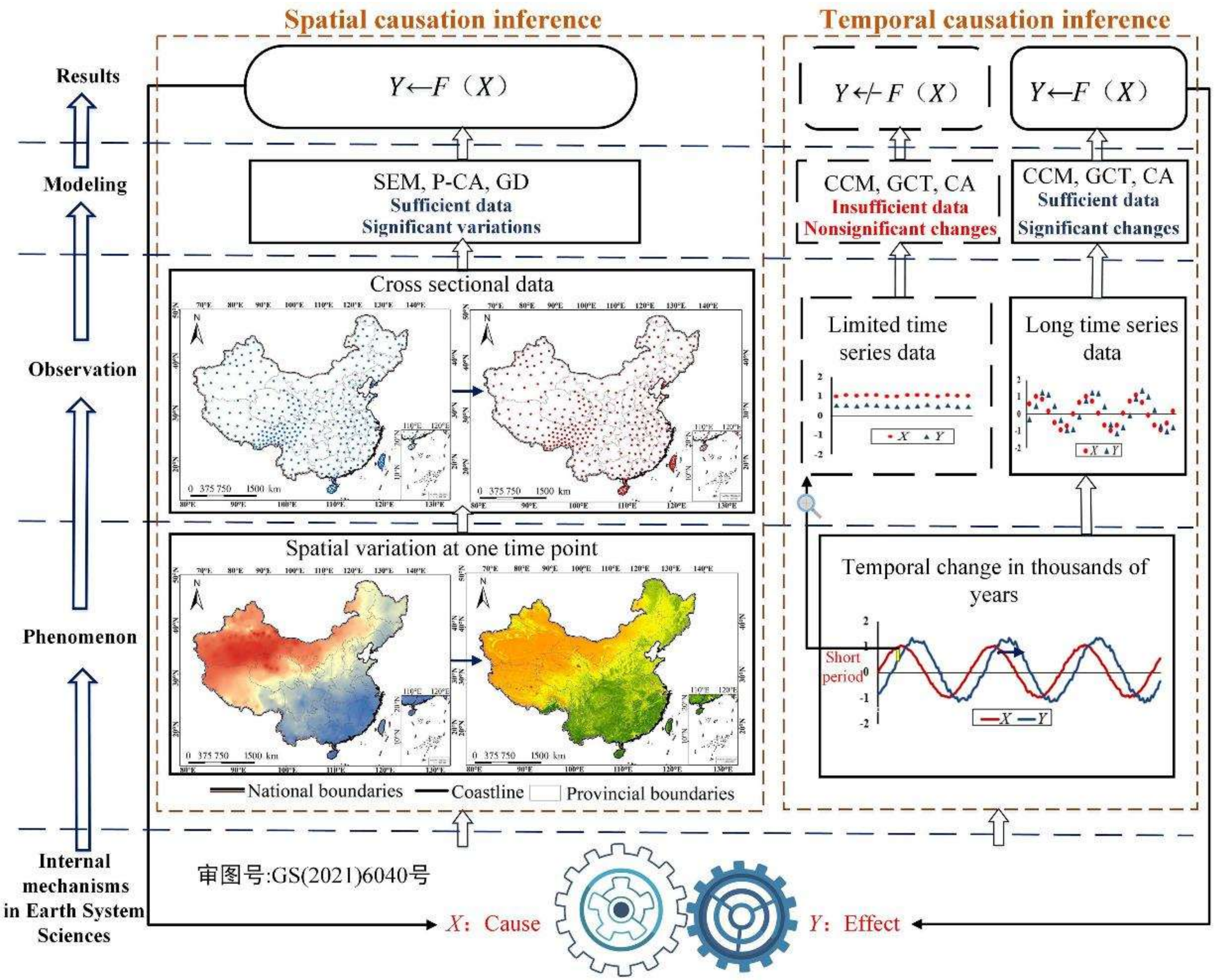
香港科技大学(广州) THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU) | AI4CITY LAB | Curtin University | Geospatial Intelligence Lab

```r
# geographical convergent cross mapping
columbus = sf::read_sf(system.file(
"shapes/columbus.gpkg", package="spData"))
spEDM::simplex(columbus,target = "HOVAL",
               lib = 1:49)
#> The suggested embedding dimension E for
variable HOVAL is 6
#>          E        rho      mae     rmse
#>  [1,]   1 0.02078897 17.72059 23.91022
#>  [2,]   2 0.13227442 14.94268 21.20952
#>  [3,]   3 0.30246001 13.70158 18.66540
#>  [4,]   4 0.14412367 15.07788 20.40852
#>  [5,]   5 0.17810605 15.14727 20.34214
#>  [6,]   6 0.33371524 13.96825 18.88074
#>  [7,]   7 0.30107433 13.43619 19.09429
#>  [8,]   8 0.29861812 13.51815 19.13343
#>  [9,]   9 0.29555999 13.55432 19.17256
#> [10,] 10 0.29555999 13.55432 19.17256
```

```r
spEDM::simplex(columbus,target = "CRIME",
               lib = 1:49)
#> The suggested embedding dimension E for
 variable CRIME is 5
#>          E       rho       mae     rmse
#>  [1,]   1 0.5064332 10.752317 15.45595
#>  [2,]   2 0.5832484  9.753875 14.06003
#>  [3,]   3 0.6026672 10.093884 13.82248
#>  [4,]   4 0.6174598 10.269995 13.71842
#>  [5,]   5 0.6313140  9.808677 13.45584
#>  [6,]   6 0.6285516  9.832107 13.58268
#>  [7,]   7 0.6299036  9.787748 13.55552
#>  [8,]   8 0.6304351  9.769181 13.54919
#>  [9,]   9 0.6303645  9.770848 13.55023
#> [10,] 10 0.6303645  9.770848 13.55023
```

```r
g = spEDM::gccm(columbus, "HOVAL", "CRIME",
                libsizes = seq(5,40,5), E = c(6,5))
#> Computing: [========================================] 100% (done)
#> Computing: [========================================] 100% (done)
g
#>   libsizes HOVAL->CRIME CRIME->HOVAL
#> 1        5    0.1405456    0.2804058
#> 2       10    0.2413240    0.4006664
#> 3       15    0.2554682    0.4609814
#> 4       20    0.2482281    0.5019578
#> 5       25    0.2541110    0.5452084
#> 6       30    0.2671786    0.5847262
#> 7       35    0.2783446    0.6019328
#> 8       40    0.2876909    0.6217028
```



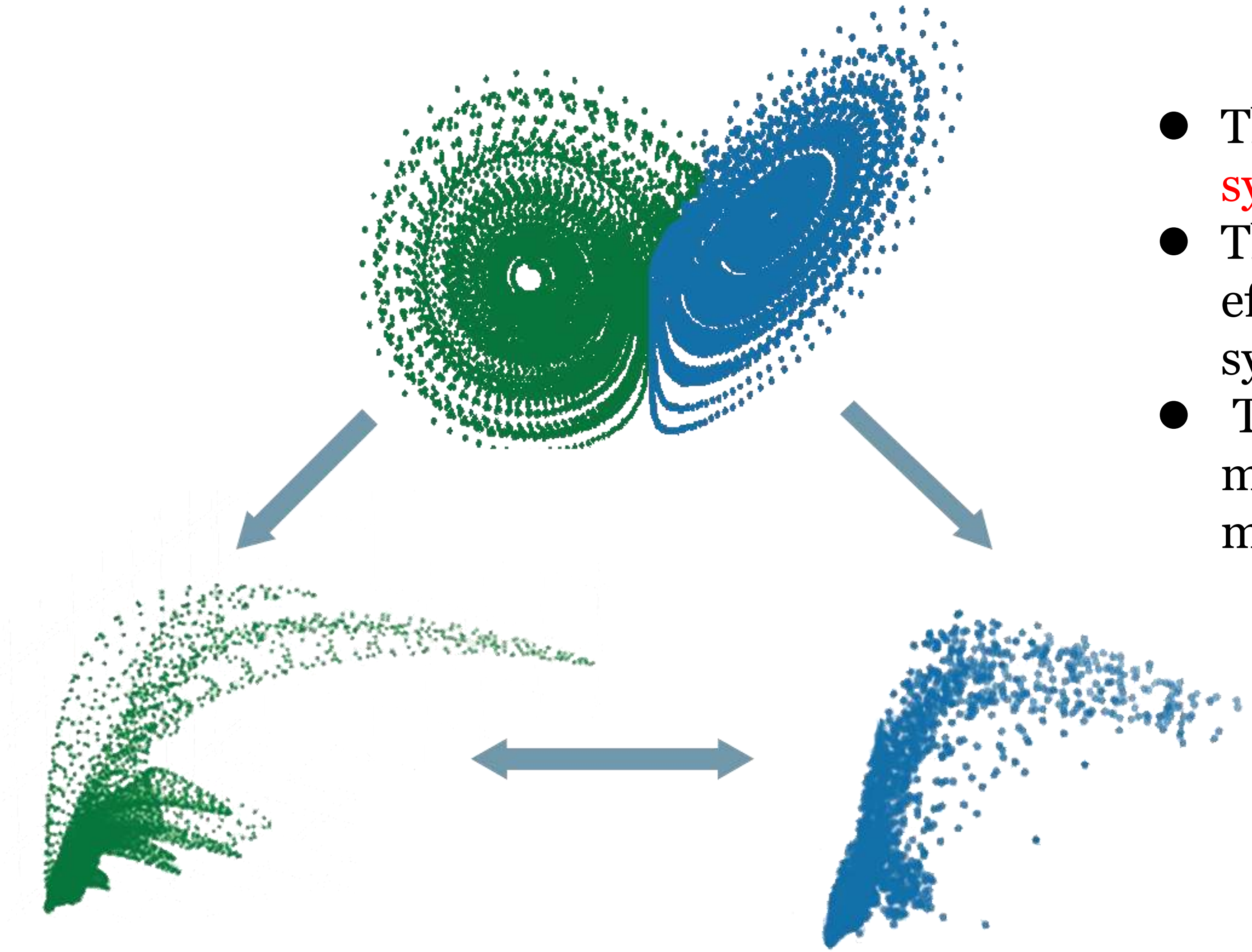


Temporally or spatially?

- Capture causality through temporal variations reflected in *time series data*.
- Capture causality through spatial variations reflected in *spatial cross-sectional data*.
- When changes are insufficient in time or space, incorporate information from the other dimension to capture causality.
- Jointly model causality using both spatial and temporal information.

Gao, B., Li, M., Wang, J., & Chen, Z. (2022). Temporally or spatially? Causation inference in earth system sciences. *Sci. Bull*, *67*(3). DOI:10.1016/j.scib.2021.10.002

Spatial Empirical Dynamic Modeling



- The Earth is a highly interconnected dynamical system.
- The state space reconstruction method can be effectively applied to causal inference in dynamic systems.
- There are many state space reconstruction models oriented towards time series, but similar models in the spatial domain are relatively few.

[1]George Sugihara et al. ,Detecting Causality in Complex Ecosystems.Science338,496-500(2012).DOI:10.1126/science.1227079

- [2]Runge, J., Gerhardus, A., Varando, G. et al. Causal inference for time series. Nat Rev Earth Environ 4, 487–505 (2023). https://doi.org/10.1038/s43017-023-00431-y

[3]Gao, B., Yang, J., Chen, Z. et al. Causal inference from cross-sectional earth system data with geographical convergent cross mapping. Nat Commun 14, 5875 (2023). https://doi.org/10.1038/s41467-023-41619-6

Package: spEDM 1.5

spEDM: Spatial Empirical Dynamic Modeling

Inferring causal associations in cross-sectional earth system data through empirical dynamic modeling (EDM), with extensions to convergent cross mapping from Sugihara et al. (2012) <doi:10.1126/science.1227079>, partial cross mapping as outlined in Leng et al. (2020) <doi:10.1038/s41467-020-16238-0>, and cross mapping cardinality as described in Tao et al. (2023)<doi:10.1016/j.fmre.2023.01.007>.

Wenbo Lv

Authors: Wenbo Lv [aut, cre, cph]

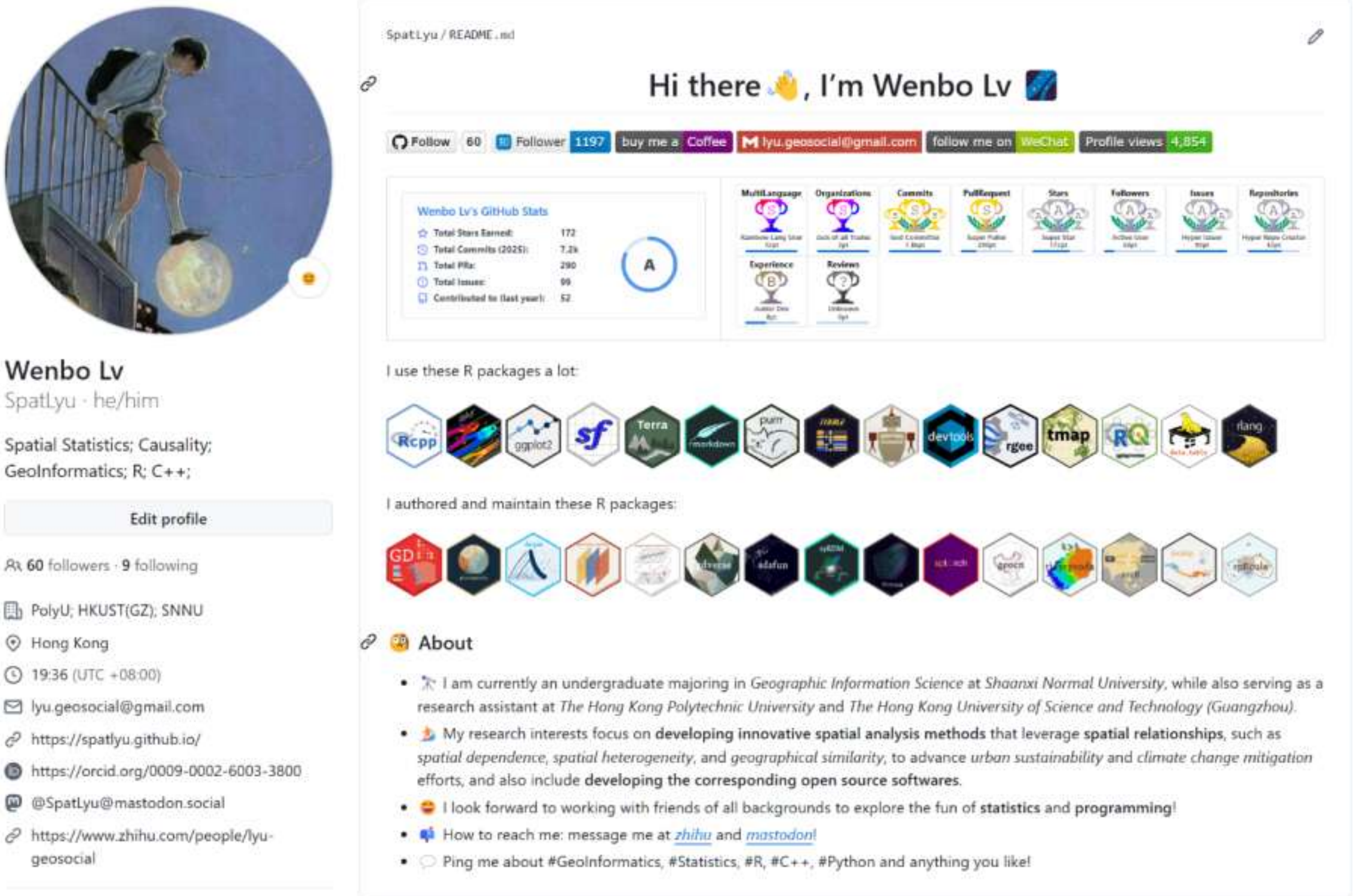
- spEDM_1.5.tar.gz
- spEDM_1.5.zip (r-4.5) spEDM_1.5.zip (r-4.4) spEDM_1.5.zip (r-4.3)
- spEDM_1.5.tgz (r-4.5-x86_64) spEDM_1.4.tgz (r-4.5-arm64) spEDM_1.5.tgz (r-4.4-x86_64) spEDM_1.5.tgz (r-4.4-arm64) spEDM_1.5.tgz (r-4.3-x86_64) spEDM_1.5.tgz (r-4.3-arm64)
- spEDM_1.5.tar.gz (r-4.5-noble) spEDM_1.5.tar.gz (r-4.4-noble)
- spEDM_1.5.tgz (r-4.4-emscripten) spEDM_1.5.tgz (r-4.3-emscripten)
- spEDM.pdf | spEDM.html ✨
- spEDM/json (API)
- NEWS

```r
# Install 'spEDM' in R:
install.packages('spEDM', repos = c('https://stscl.r-universe.dev', 'https://cloud.r-project.org'))
```

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)
AI4CITY LAB
Curtin University
Geospatial Intelligence Lab

Thanks.

https://github.com/ai4city-hkust/geocausality_workshop
https://ai4city-hkust.github.io/workshop/geocausality/geocausality_workshop.html



香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)
AI4CITY LAB
Curtin University
Geospatial Intelligence Lab