Detecting Causal Interactions between Atlantic and Pacific Ocean Variability

Supervisors: Prof. Paolo Tieri, Dott. Alessio Bellucci, Dott. David Docquier, Dott. Valerio Lembo, Prof. Manuela Petti

Jonas Barth

May 30, 2025

Sapienza Università di Roma

Climate Change

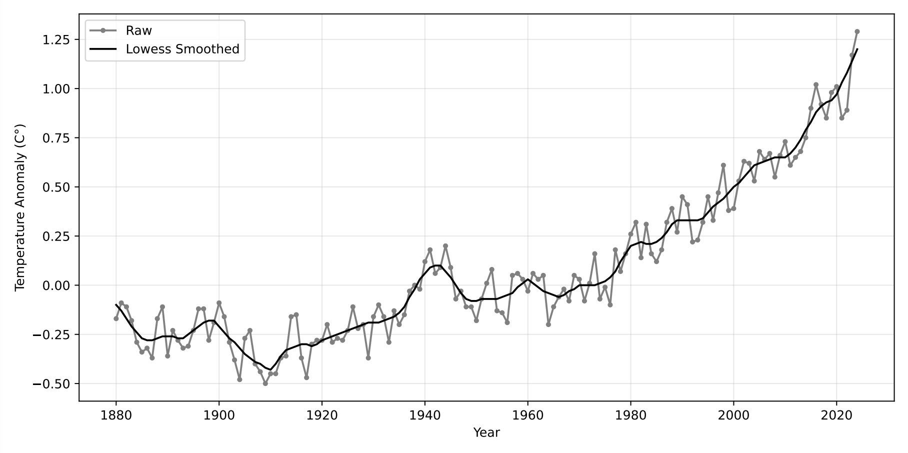
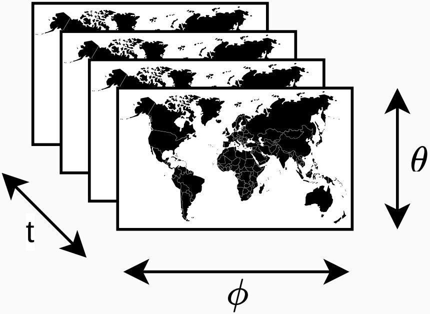


Figure 1: Change in global surface temperature compared to the long-term average from 1951 to 1980 [9].

Data

Sea Surface Temperature (SST)



Name	Resolution t	Resolution ϕ, θ	Time Range	Source
HadISST	Monthly	$\phi = [-180, 180], \theta = [-90, 90]$	1871 – present	[12]
ERSSTV5	Monthly	$\phi = [-180, 180], \theta = [-90, 90]$	1854 – present	[5]

Table 1: Summary of the HadISST and ERSSTV5 datasets.

Climate Indices

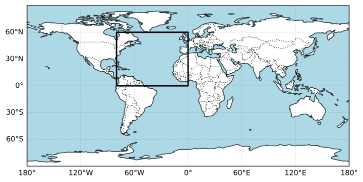


Figure 2: AMV Boundaries

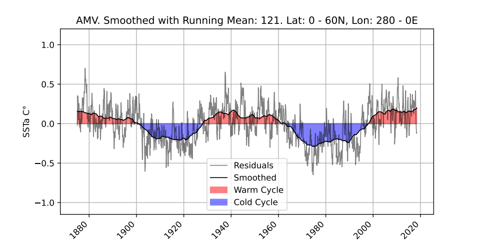


Figure 3: Atlantic Multidecadal Variability (AMV)

Climate Indices

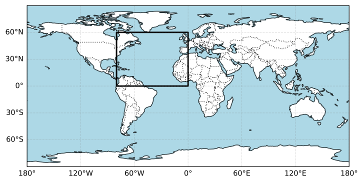


Figure 2: AMV Boundaries

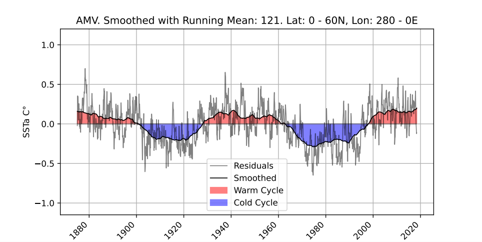


Figure 3: Atlantic Multidecadal Variability (AMV)

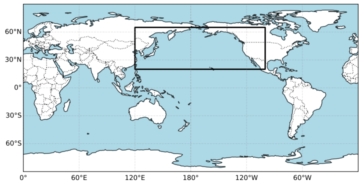


Figure 4: PDO Boundaries

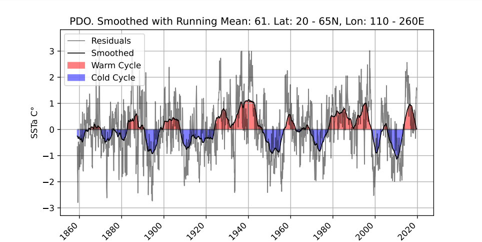


Figure 5: Pacific Decadal Oscillation (PDO)

Research Question

What is the relationship between the AMV and PDO indices?

Pacific Modes of Variability

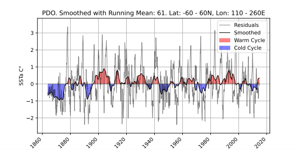


Figure 6: ENSO

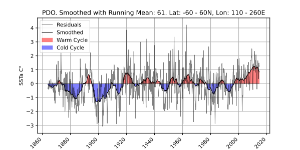


Figure 7: SEPM

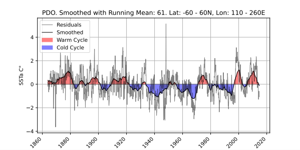


Figure 8: SPDOL

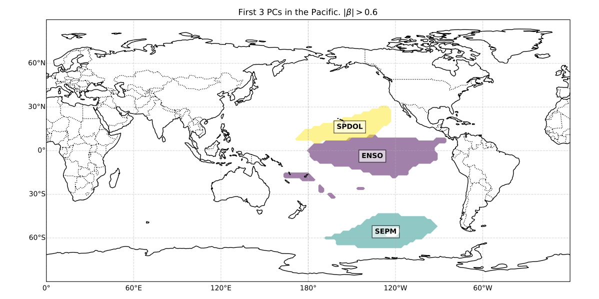
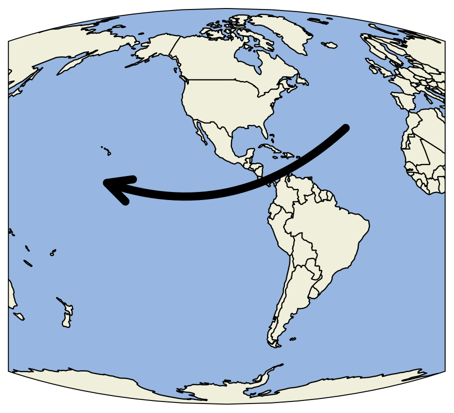


Figure 9: The first 3 varimax rotated principal components in the Pacific 60S to 60N, 110E to 100W.

Existing literature

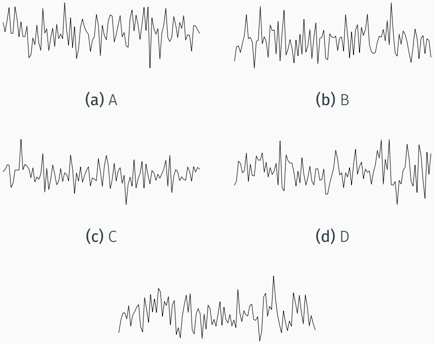
Previous research show a connection from the AMV to the Pacific [16, 10, 2, 15, 1, 3], but use **model simulations**.

We want to use **observational data**.



Causal Discovery Networks

Learn a **causal graph** from **observational** data. In the causal graph, nodes are observations, and edges are causal relationships.

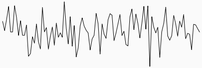


(a) A (b) B (c) C (d) D (e) E

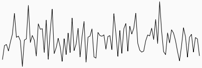
$A = f_A()$ $\quad$ $C = f_C(D)$ $\quad$ $E = f_E(A, C)$

$B = f_B(A, D)$ $\quad$ $D = f_D()$

Learn a **causal graph** from **observational** data. In the causal graph, nodes are observations, and edges are causal relationships.



(a) A

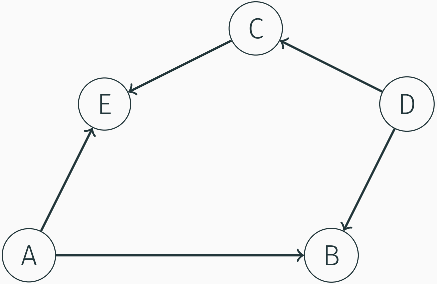


(b) B

(c) C

(d) D

(e) E



$$A = f_A()$$
$$C = f_C(D)$$
$$E = f_E(A, C)$$
$$B = f_B(A, D)$$
$$D = f_D()$$

Granger Causality

Check if f_{full} has lower prediction error than f_{red}:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t, f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \quad (1)$$

PCMCI+

Use conditional independence (CI) tests to iteratively remove links from a complete graph.

$$X \perp\!\!\!\perp Y \quad | \quad Z \quad (2)$$

LKIF

Use a maximum likelihood estimator (MLE) to estimate the rate of entropy transfer between variables in the system. A link exists if the information transfer is statistically significant at level α.

Granger Causality

Check if f_{full} has lower prediction error than f_{red}:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t, f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \tag{1}$$

PCMCI+

Use conditional independence (CI) tests to iteratively remove links from a complete graph.

$$X \perp\!\!\!\perp Y \quad | \quad Z \tag{2}$$

LKIF

Use a maximum likelihood estimator (MLE) to estimate the rate of entropy transfer between variables in the system. A link exists if the information transfer is statistically significant at level α.

Methods

Granger Causality

Check if f_{full} has lower prediction error than f_{red}:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t, f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \quad (1)$$

PCMCI+

Use conditional independence (CI) tests to iteratively remove links from a complete graph.

$$X \perp\!\!\!\perp Y \quad | \quad Z \quad (2)$$

LKIF

Use a maximum likelihood estimator (MLE) to estimate the rate of entropy transfer between variables in the system. A link exists if the information transfer is statistically significant at level α.

Experiment Setup

Apply Granger Causality, PCMCI+, LKIF to two networks, at $\alpha = 0.05$, $\tau = 12$:

Network 1

- AMV
- PDO

Network 2

- AMV
- PDO
- PC1 Varimax Rotated (ENSO)
- PC2 Varimax Rotated (SEPM)
- PC3 Varimax Rotated (SPDOL)

Result Graphs

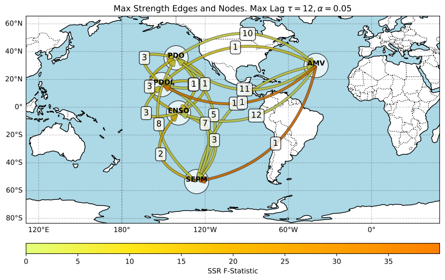


Figure 11: Granger Causality

Result Graphs

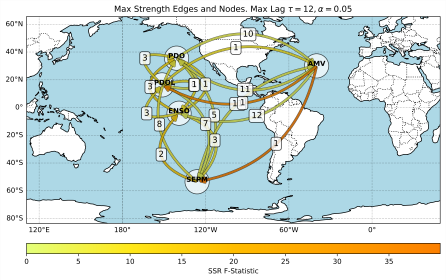


Figure 11: Granger Causality

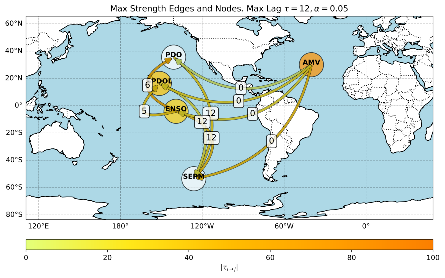


Figure 12: LKIF

Result Graphs

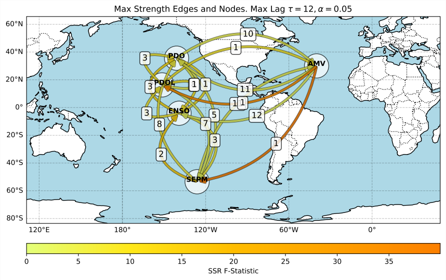


Figure 11: Granger Causality

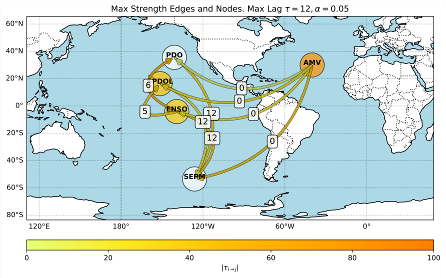


Figure 12: LKIF

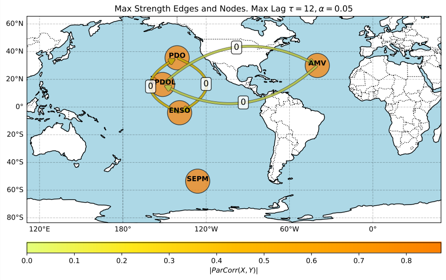


Figure 13: PCMCI+

Result Edge Strengths - from AMV

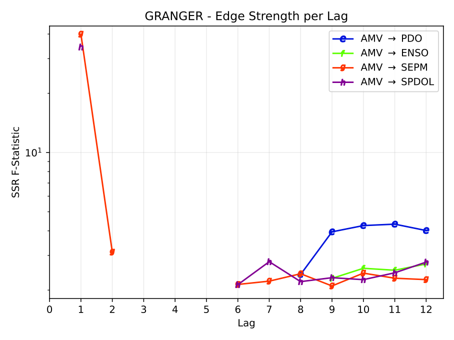


Figure 14: Granger Causality

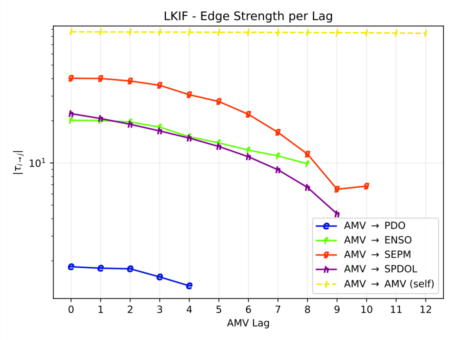


Figure 15: LKIF

Result Edge Strengths - to PDO

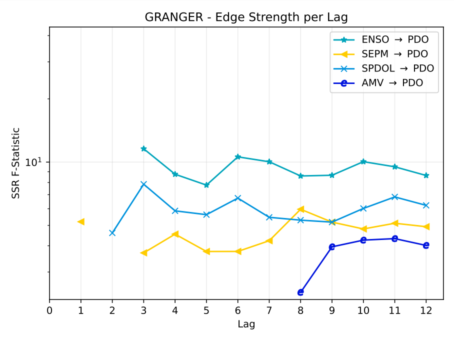


Figure 16: Granger Causality

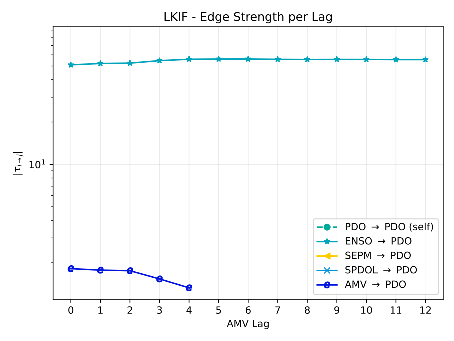


Figure 17: LKIF

Conclusion

- Granger Causality and LKIF find links from AMV to PDO $\rightarrow$ changes in the Atlantic are followed by changes in the Pacific.
- Links from AMV to PDO tend to be weaker than links from other Pacific modes to PDO (Figure 16, Figure 17) $\rightarrow$ PDO is be modulated by AMV via other Pacific components.

Conclusion

- Granger Causality and LKIF find links from AMV to PDO $\rightarrow$ changes in the Atlantic are followed by changes in the Pacific.
- Links from AMV to PDO tend to be weaker than links from other Pacific modes to PDO (Figure 16, Figure 17) $\rightarrow$ PDO is be modulated by AMV via other Pacific components.

References

[1] Xiaohe An et al. **"Atlantic Multidecadal Oscillation Drives Interdecadal Pacific Variability via Tropical Atmospheric Bridge".** In: *Journal of Climate* 34.13 (2021), pp. 5543–5553. DOI: 10.1175/JCLI-D-20-0983.1. URL: https://journals.ametsoc.org/view/journals/clim/34/13/JCLI-D-20-0983.1.xml.

[2] Marc d'Orgeville and W. Richard Peltier. **"On the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation: Might they be related?"** In: *Geophysical Research Letters* 34.23 (2007). DOI: https://doi.org/10.1029/2007GL031584. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL031584. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL031584.

[3] Zhanqiu Gong et al. **"An inter-basin teleconnection from the North Atlantic to the subarctic North Pacific at multidecadal time scales".** In: *Climate Dynamics* 54.1 (2020), pp. 807–822. ISSN: 1432-0894. DOI: 10.1007/s00382-019-05031-5. URL: https://doi.org/10.1007/s00382-019-05031-5.

[4] C. W. J. Granger. **"Investigating Causal Relations by Econometric Models and Cross-spectral Methods".** In: *Econometrica* 37.3 (1969), pp. 424–438. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912791 (visited on 07/07/2024).

[5] Boyin Huang et al. ***NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5 [indicate subset used].*** Accessed via NOAA/ESRL/PSD: https://www.esrl.noaa.gov/psd/ [2025-04-07]. 2017. DOI: 10.7289/V5T72FNM.

[6] X. San Liang. **"Normalized Multivariate Time Series Causality Analysis and Causal Graph Reconstruction".** In: *Entropy* 23.6 (2021). ISSN: 1099-4300. DOI: 10.3390/e23060679. URL: https://www.mdpi.com/1099-4300/23/6/679.

[7] X. San Liang. **“The Liang-Kleeman Information Flow: Theory and Applications”.** In: *Entropy* 15.1 (2013), pp. 327–360. ISSN: 1099-4300. DOI: 10.3390/e15010327. URL: https://www.mdpi.com/1099-4300/15/1/327.

[8] X. San Liang and Richard Kleeman. **“Information Transfer between Dynamical System Components”.** In: *Phys. Rev. Lett.* 95 (24 Dec. 2005), p. 244101. DOI: 10.1103/PhysRevLett.95.244101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.95.244101.

[9] NASA. ***Global Temperature.*** https://climate.nasa.gov/vital-signs/global-temperature/?intent=121. Accessed: 2025-04-30. 2025.

[10] Yuko M. Okumura et al. **“North Pacific Climate Response to Freshwater Forcing in the Subarctic North Atlantic: Oceanic and Atmospheric Pathways”.** In: *Journal of Climate* 22.6 (2009), pp. 1424–1445. DOI: 10.1175/2008JCLI2511.1. URL: https://journals.ametsoc.org/view/journals/clim/22/6/2008jcli2511.1.xml.

[11] Milan Paluš et al. **"Causality, dynamical systems and the arrow of time".** In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.7 (July 2018), p. 075307. ISSN: 1054-1500. DOI: 10.1063/1.5019944. eprint: https://pubs.aip.org/aip/cha/article-pdf/doi/10.1063/1.5019944/13283420/075307_1_online.pdf. URL: https://doi.org/10.1063/1.5019944.

[12] N. A. Rayner et al. **"Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century".** In: *Journal of Geophysical Research: Atmospheres* 108.D14 (2003). Accessed via https://www.metoffice.gov.uk/hadobs/hadisst/ [2025-04-07], p. 4407. DOI: 10.1029/2002JD002670.

[13] Jakob Runge. ***Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets.*** 2022. arXiv: 2003.03685 [stat.ME]. URL: https://arxiv.org/abs/2003.03685.

[14] Jakob Runge et al. **"Detecting and quantifying causal associations in large nonlinear time series datasets".** In: *Science Advances* 5.11 (2019), eaau4996. DOI: 10.1126/sciadv.aau4996. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aau4996. URL: https://www.science.org/doi/abs/10.1126/sciadv.aau4996.

[15] Cheng Sun et al. **"Western tropical Pacific multidecadal variability forced by the Atlantic multidecadal oscillation".** In: *Nature Communications* 8 (July 2017), p. 15998. DOI: 10.1038/ncomms15998.

[16] Rong Zhang and Thomas L. Delworth. **"Impact of the Atlantic Multidecadal Oscillation on North Pacific climate variability".** In: *Geophysical Research Letters* 34.23 (2007). DOI: https://doi.org/10.1029/2007GL031601. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL031601. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL031601.

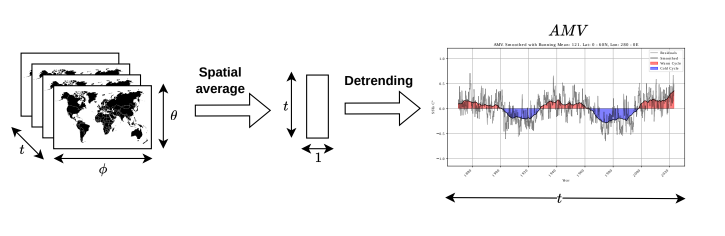


Figure 18: Atlantic Multidecadal Variability (AMV)

Climate Indices

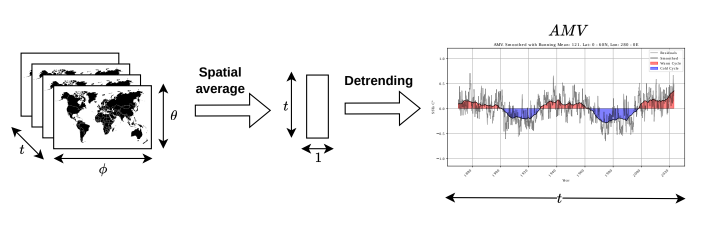


Figure 18: Atlantic Multidecadal Variability (AMV)

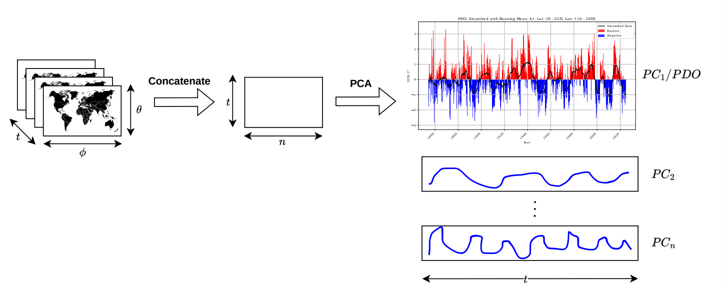
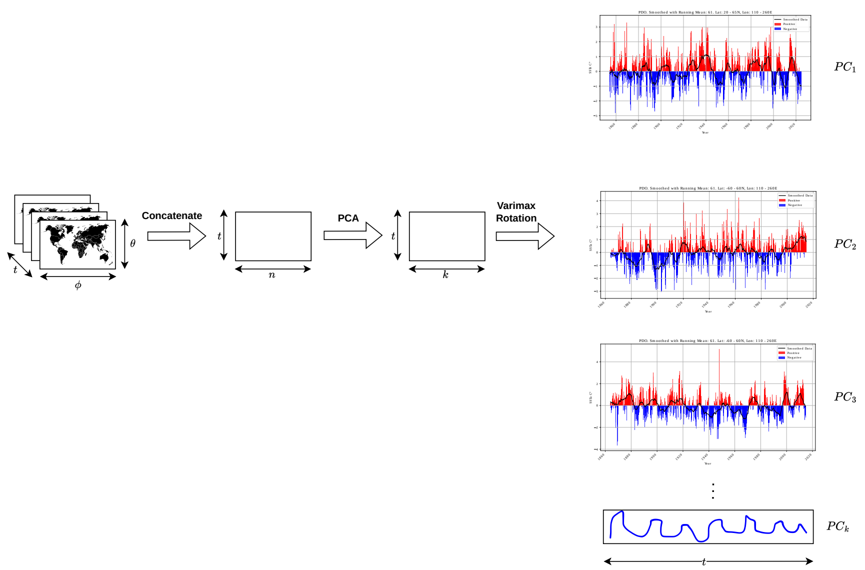


Figure 19: Pacific Decadal Oscillation (PDO)

Varimax Rotation

Enrich the system by adding more Pacific variables, computed via PCA followed by a Varimax rotation.



Granger Causality [4, 11]

Given timeseries X (cause) and Y (effect), fit **two** regression models to predict Y:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t \tag{3}$$

$$f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \tag{4}$$

Compare using F-test where F is:

$$F = \frac{SSE_{red} - SSE_{full}/(r-s)}{SSE_{full}/(T-r)} \tag{5}$$

Granger Causality [4, 11]

Given timeseries X (cause) and Y (effect), fit **two** regression models to predict Y:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t \tag{3}$$

$$f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \tag{4}$$

Compare using F-test where F is:

$$F = \frac{SSE_{red} - SSE_{full}/(r - s)}{SSE_{full}/(T - r)} \tag{5}$$

Granger Causality [4, 11]

Given timeseries X (cause) and Y (effect), fit **two** regression models to predict Y:

$$f_{red} = a_0 + \sum_{i=1}^{p} a_i Y_{t-i} + \epsilon_t \tag{3}$$

$$f_{full} = b_0 + \sum_{i=1}^{p} b_i Y_{t-i} + \sum_{j=1}^{q} c_j X_{t-j} + \eta_t \tag{4}$$

Compare using F-test where F is:

$$F = \frac{SSE_{red} - SSE_{full}/(r - s)}{SSE_{full}/(T - r)} \tag{5}$$

Granger Causality [4, 11]

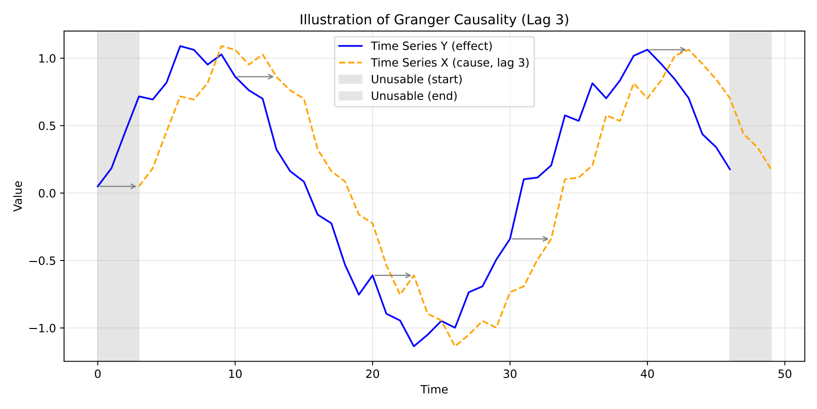


Figure 21

PCMCI+ [14, 13]

Use **Conditional Independence** (CI) tests to determine if two variables are independent **given** a set Z.

$$X_i \perp\!\!\!\perp X_j | Z \tag{6}$$

1. Start with fully connected graph.
2. Discover lagged parents $\widehat{B_t^-}(X_t^i)$.
3. Condition on lagged parents $\widehat{B_t^-}(X_t^i)$ and contemporaneous connections S to refine parent set.

$$MCI : X_i^{t-\tau} \perp\!\!\!\perp X_j^t \quad | \quad S, \widehat{B}_-^t(X_j^t) \setminus \left\{X_i^{t-\tau}\right\}, \widehat{B}_-^{t-\tau}(X_i^{t-\tau}) \tag{7}$$

PCMCI+ [14, 13]

Use **Conditional Independence** (CI) tests to determine if two variables are independent **given** a set Z.

$$X_i \perp\!\!\!\perp X_j | Z \tag{6}$$

1. Start with fully connected graph.
2. Discover lagged parents $\widehat{B_t^-}(X_t^i)$.
3. Condition on lagged parents $\widehat{B_t^-}(X_t^i)$ and contemporaneous connections S to refine parent set.

$$MCI : X_i^{t-\tau} \perp\!\!\!\perp X_j^t \quad | \quad S, \widehat{B}_-^t(X_j^t) \setminus \{X_i^{t-\tau}\}, \widehat{B}_-^{t-\tau}(X_i^{t-\tau}) \tag{7}$$

PCMCI+ [14, 13]

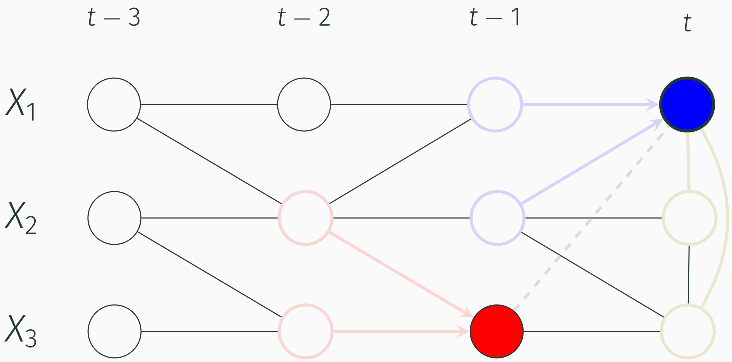
Use **Conditional Independence** (CI) tests to determine if two variables are independent **given** a set Z.

$$X_i \perp\!\!\!\perp X_j | Z \tag{6}$$

1. Start with fully connected graph.
2. Discover lagged parents $\widehat{B_t^-}(X_t^i)$.
3. Condition on lagged parents $\widehat{B_t^-}(X_t^i)$ and contemporaneous connections S to refine parent set.

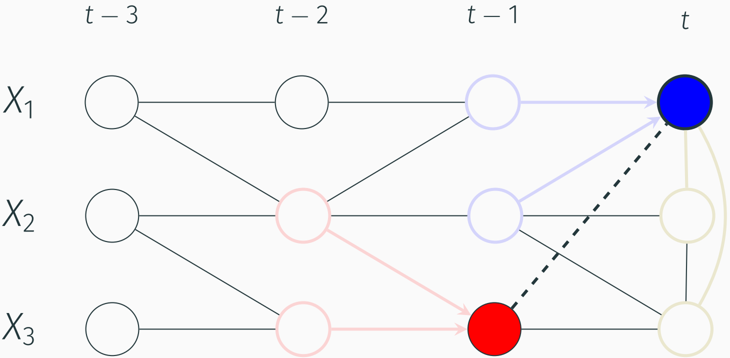
$$MCI : X_i^{t-\tau} \perp\!\!\!\perp X_j^t \quad | \quad S, \widehat{B}_-^t(X_j^t) \setminus \left\{X_i^{t-\tau}\right\}, \widehat{B}_-^{t-\tau}(X_i^{t-\tau}) \tag{7}$$

PCMCI+ Timeseries Graph



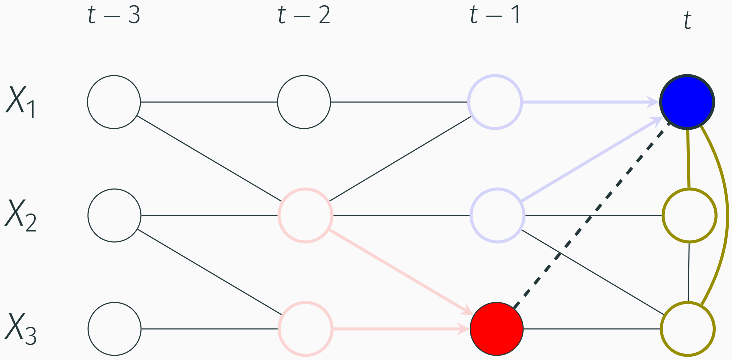
$$MCI : \mathbf{X_3^{t-1}} \perp\!\!\!\perp \mathbf{X_1^t} \quad | \quad S, \widehat{\mathcal{B}}^{t-1}_{-}(X_3^{t-1}), \widehat{\mathcal{B}}^{t}_{-}(X_1^t) \setminus \{X_3^{t-1}\} \tag{8}$$

PCMCI+ Timeseries Graph



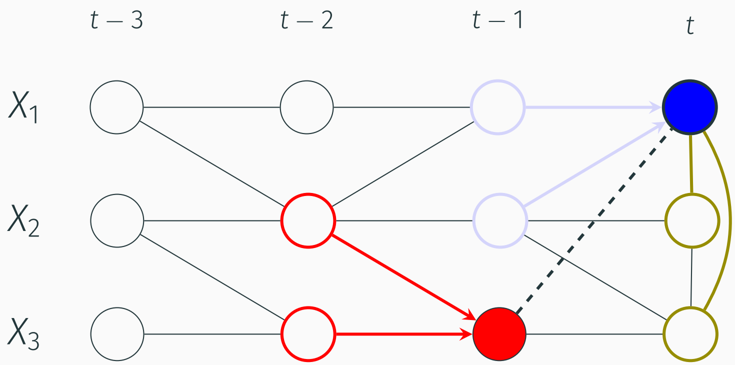
$$MCI : \mathbf{X_3^{t-1}} \perp\!\!\!\perp \mathbf{X_1^{t}} \quad | \quad S, \widehat{B}_-^{t-1}(X_3^{t-1}), \widehat{B}_-^{t}(X_1^{t}) \setminus \{X_3^{t-1}\} \tag{8}$$

PCMCI+ Timeseries Graph



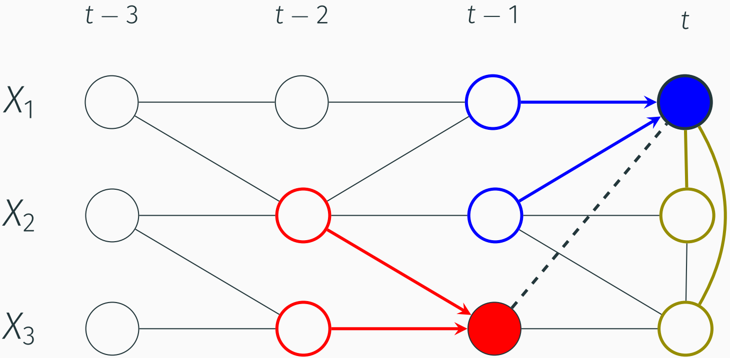
$$MCI : \mathbf{X_3^{t-1}} \perp\!\!\!\perp \mathbf{X_1^t} \quad | \quad S, \widehat{B}_-^{t-1}(X_3^{t-1}), \widehat{B}_-^t(X_1^t) \setminus \{X_3^{t-1}\} \tag{8}$$

PCMCI+ Timeseries Graph



$$MCI : \mathbf{X_3^{t-1}} \perp\!\!\!\perp \mathbf{X_1^t} \quad | \quad S, \widehat{\mathcal{B}}_-^{t-1}(X_3^{t-1}), \widehat{\mathcal{B}}_-^{t}(X_1^{t}) \setminus \{X_3^{t-1}\} \tag{8}$$

PCMCI+ Timeseries Graph



$$MCI : \mathbf{X_3^{t-1}} \perp\!\!\!\perp \mathbf{X_1^{t}} \quad | \quad S, \widehat{B}^{t-1}_{-}(X_3^{t-1}), \widehat{B}^{t}_{-}(X_1^{t}) \setminus \{X_3^{t-1}\} \tag{8}$$

LKIF [8, 7, 6]

Measure the **information flow** between variables, using a Maximum Likelihood Estimator (MLE).

$$\hat{T}_{i \to j} = \frac{1}{det\mathbf{C}} \sum_{k=1}^{d} \Delta_{jk} C_{k,di} \frac{C_{ij}}{C_{ii}} \tag{9}$$

Normalised with

$$Z_i = \sum_{k=1}^{d} \left| \hat{T}_{k \to i} \right| + \left| \frac{dH_i^{noise}}{dt} \right| \tag{10}$$