Assessing causal dependencies in climate indices

Riccardo Silini[1], **Giulio Tirabassi[1]**, Marcelo Barreiro[2], Laura Ferranti[3], Cristina Masoller[1]

[1]Universitat Politècnica de Catalunya,
[2]Universidad de la República, Montevideo
[3]ECMWF, Reading

Weather and Climate Extremes and their Predictability
CAFE Final Conference

UPC

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

ECMWF

Transfer Entropy

Mutual information

$$M_{XY} = \sum_{i,j} p_{XY}(i,j)\, log\left[\frac{p_{XY}(i,j)}{p_X(i)p_Y(j)}\right]$$

Transfer entropy

$$TE_{Y \to X} = \sum_{i,j} p_{XY}\left(i_{n+1}, i_n^{(k)}, j_n^{(l)}\right) log\left[\frac{p_{XY}\left(i_{n+1} \middle| i_n^{(k)}, j_n^{(l)}\right)}{p_{XY}\left(i_{n+1} \middle| i_n^{(k)}\right)}\right]$$

lag Markovian processes

Transfer entropy with the entropy of normal distributions

$$TE_{Y\rightarrow X} = \sum_{i,j} p\left(i_{n+1}, i_n^{(k)}, j_n^{(l)}\right) log\left[\frac{p\left(i_{n+1}\middle|i_n^{(k)}, j_n^{(l)}\right)}{p\left(i_{n+1}\middle|i_n^{(k)}\right)}\right]$$

$$TE_{Y\rightarrow X} = H\left(i_n^{(k)}, j_n^{(l)}\right) - H\left(i_{n+1}, i_n^{(k)}, j_n^{(l)}\right) + H\left(i_{n+1}, i_n^{(k)}\right) - H\left(i_n^{(k)}\right)$$

Normal distribution: $H_p(x) = \frac{1}{2}\left(\cancel{p + p\,log(2\pi)} + log|\Sigma_x|\right)$

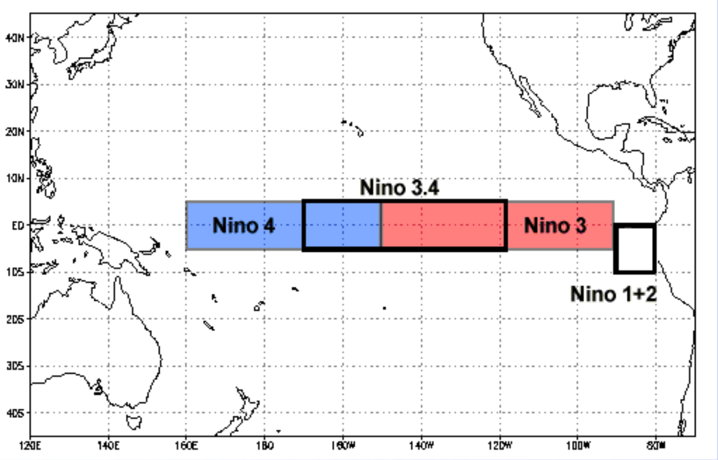
pseudo
Transfer Entropy

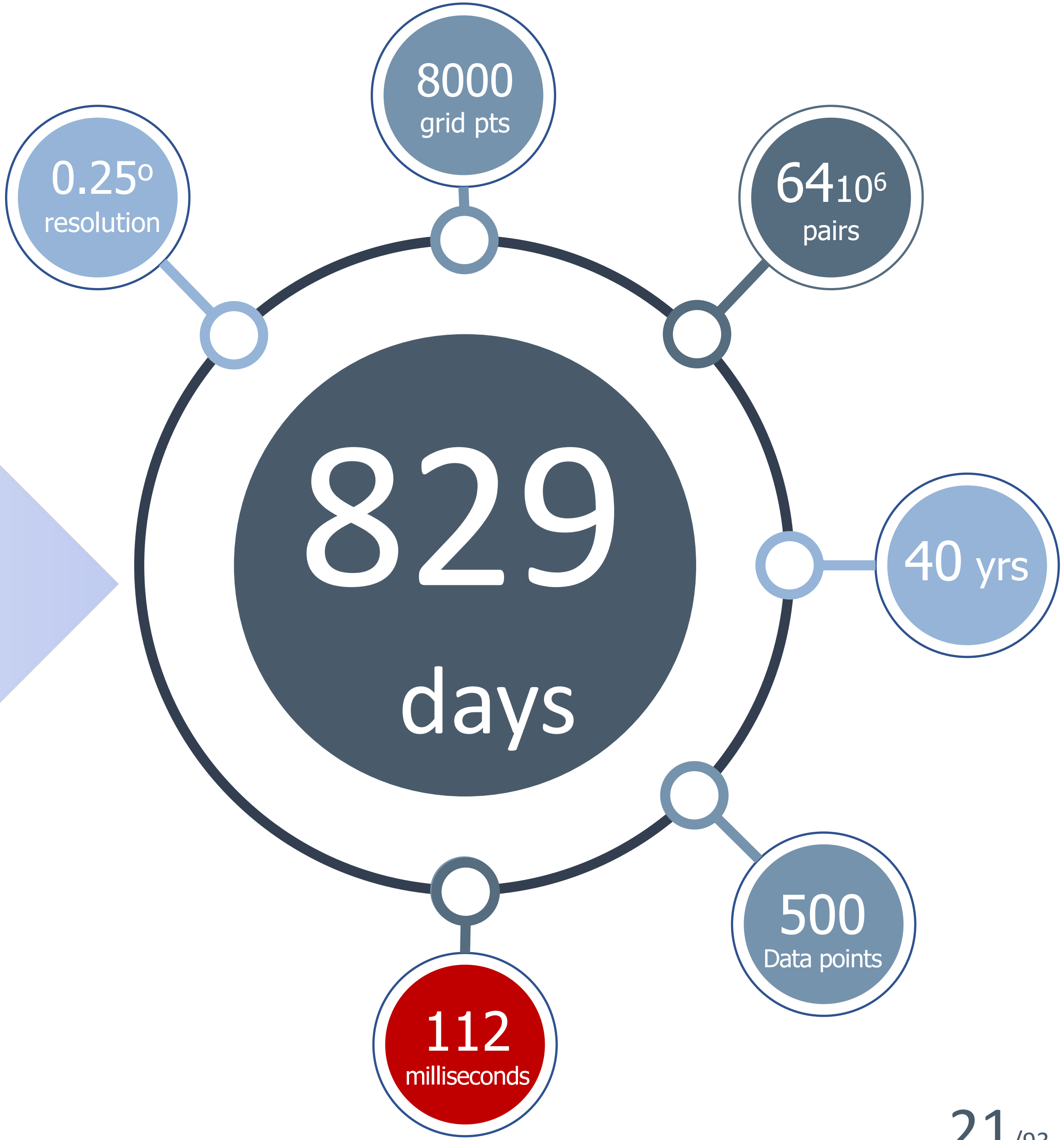
$$pTE_{Y\to X} = \frac{1}{2}\log\left(\frac{\left|\Sigma\left(I_n^{(k)}\oplus J_n^{(l)}\right)\right|\cdot\left|\Sigma\left(\boldsymbol{i}_{n+1}\oplus I_n^{(k)}\right)\right|}{\left|\Sigma\left(\boldsymbol{i}_{n+1}\oplus I_n^{(k)}\oplus J_n^{(l)}\right)\right|\cdot\left|\Sigma\left(I_n^{(k)}\right)\right|}\right)$$

Silini R & Masoller C, *Sci. Rep.* **11,** 8423 (2021)

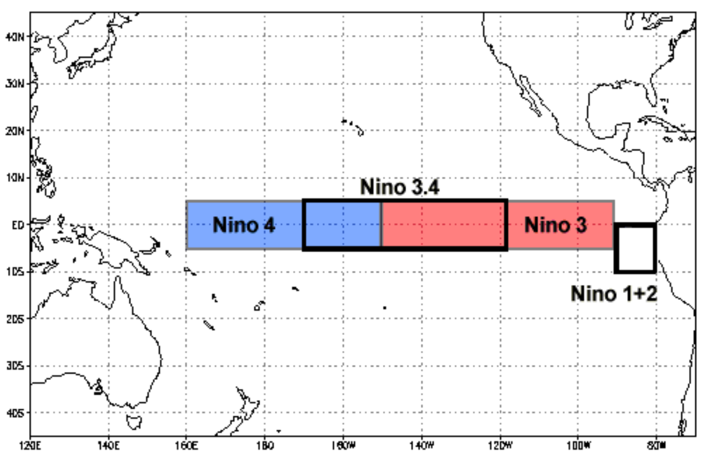
https://github.com/riccardosilini/pTE

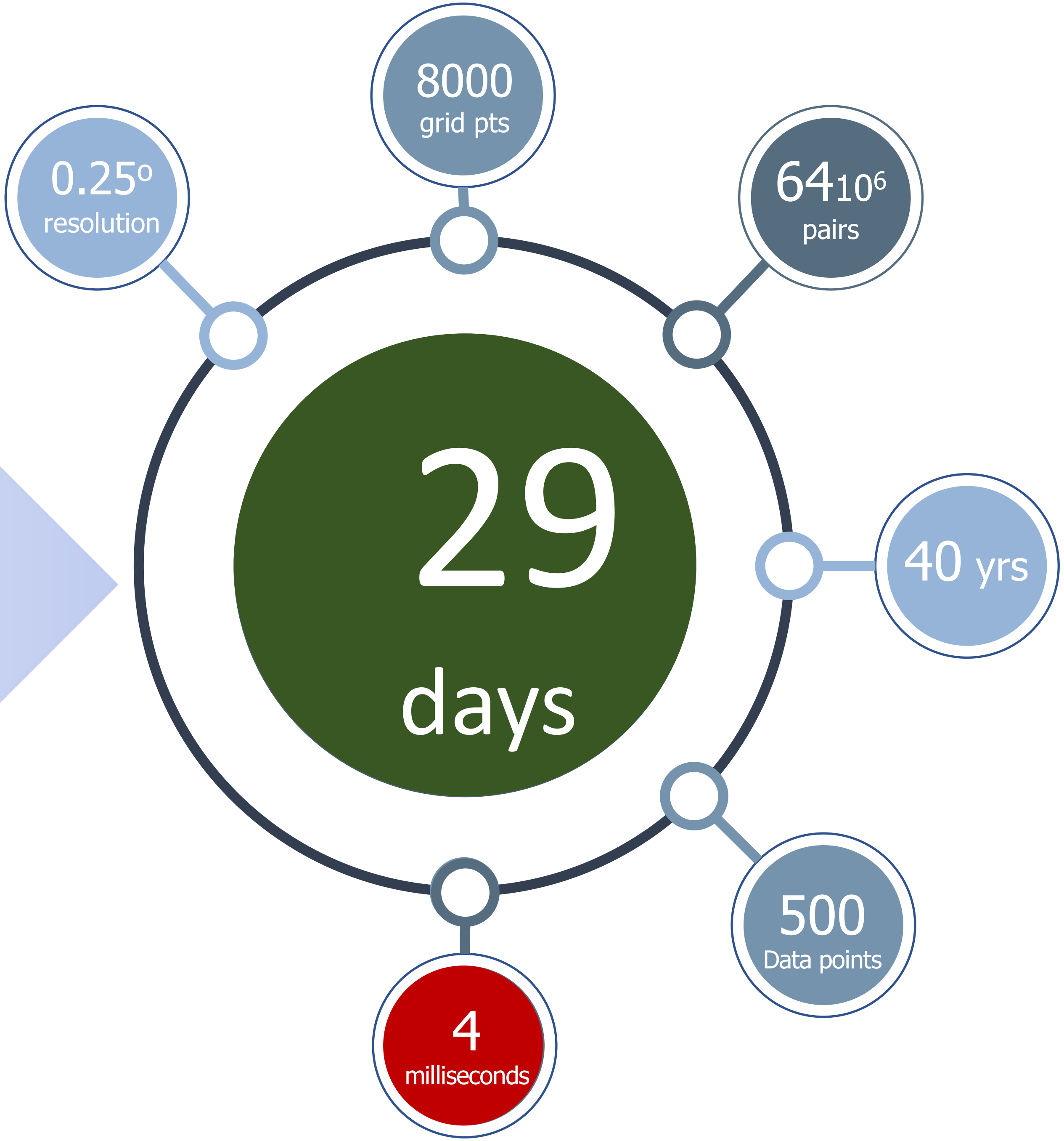
HOW MUCH TIME CAN WE SAVE?





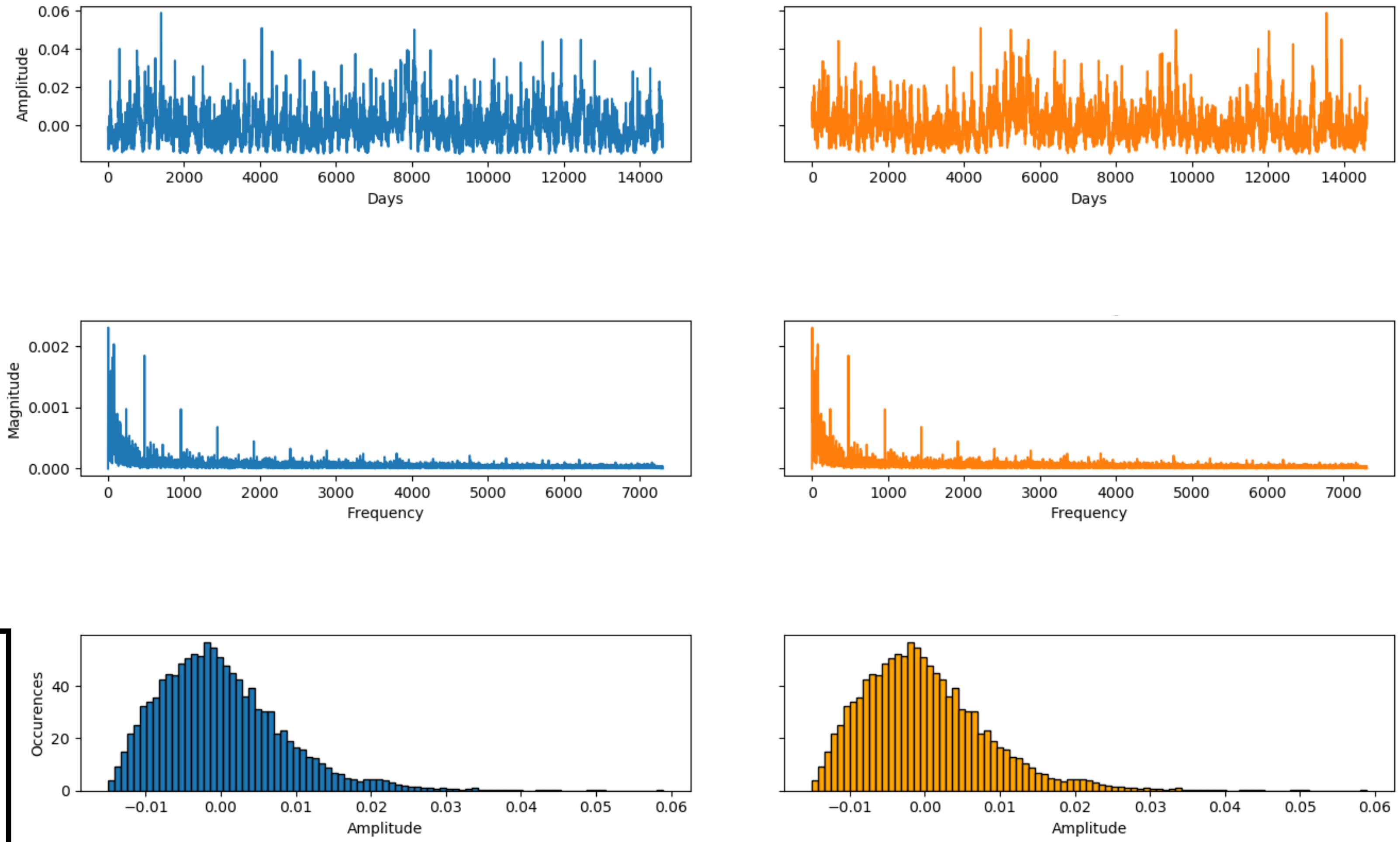
HOW MUCH TIME CAN WE SAVE?





Significance testing - surrogates

1 Original **N IAAFT**



Time series

FFT

PDF

Schreiber, T. and Schmitz, A.: *Surrogate time series*, Physica D, 142(3–4), 346–382, 2000.

Data Generating Processes

	Model
Y X	M0
	M1
	M2
Y→X	M3
	M4
	M5
	M6
	M7
	M8
	M9
	M10
	M11
	M12
Y⇆X	M13
	M14

Nonlinear: Weak, Strong, Chaotic

Some examples…

$$x_t = 0.6\, x_{t-1} + 0.5\, y_{t-1} + E_{1t} \qquad y_t = 0.5\, y_{t-1} + E_{2t}$$

$$x_t = 0.6\, x_{t-1} + 0.5\, y_{t-1}^2 + E_{1t} \qquad y_t = 0.5\, y_{t-1} + E_{2t}$$

$$x_t = 0.6\, x_{t-1} + \frac{2.4 - 0.9 y_{t-3}}{1 + e^{-4 y_{t-3}}} + E_{1t} \qquad y_t = 0.5\, y_{t-1} + E_{2t}$$

$$\begin{cases} \dot{x}_1 = -(1 + 0.015) x_2 - x_3 + 0.1(y_1 - x_1) \\ \dot{x}_2 = (1 + 0.015) x_1 + 0.15 x_2 \\ \dot{x}_3 = 0.2 + x_3 (x_1 - 10) \end{cases} \qquad \begin{cases} \dot{y}_1 = -(1 - 0.015) y_2 - y_3 \\ \dot{y}_2 = (1 - 0.015) y_1 + 0.15 y_2 \\ \dot{y}_3 = 0.2 + y_3 (y_1 - 10) \end{cases}$$

Results

	Model	pTE $Y \rightarrow X$	pTE $X \rightarrow Y$
Y X	M0	3.8	3.9
	M1	2.3	2.6
	M2	4.2	4.7
$Y \rightarrow X$	M3	100	4.5
	M4	80.7	3.8
	M5	100	2.2
	M6	100	1.8
	M7	100	2.8
	M8	100	4.5
	M9	100	0.1
	M10	62.6	3.1
	M11	46.1	43.1
	M12	99.9	1.0
$Y \leftrightarrows X$	M13	100	100
	M14	100	100

Nonlinear: Weak, Strong, Chaotic

Comparison with other methods

	Model	pTE		GC		TE	
		Y → X	X → Y	Y → X	X → Y	Y → X	X → Y
Y X	M0	3.8	3.9	5.1	5.0	4.4	4.4
	M1	2.3	2.6	3.3	3.1	100	100
	M2	4.2	4.7	5.5	5.9	4.7	4.9
Y → X	M3	100	4.5	100	4.8	70.2	5.6
	M4	80.7	3.8	84.2	4.9	96.0	4.7
	M5	100	2.2	100	3.1	100	3.8
	M6	100	1.8	100	2.8	100	4.3
	M7	100	2.8	100	3.4	100	4.0
	M8	100	4.5	100	5.6	100	100
	M9	100	0.1	100	0.1	100	100
	M10	62.6	3.1	67.3	4.3	12.2	4.5
	M11	46.1	43.1	53.1	49.8	37.8	45.0
	M12	99.9	1.0	100	0.9	100	0
Y ⇆ X	M13	100	100	100	100	100	100
	M14	100	100	100	100	100	100

Nonlinear: Weak | Strong | Chaotic

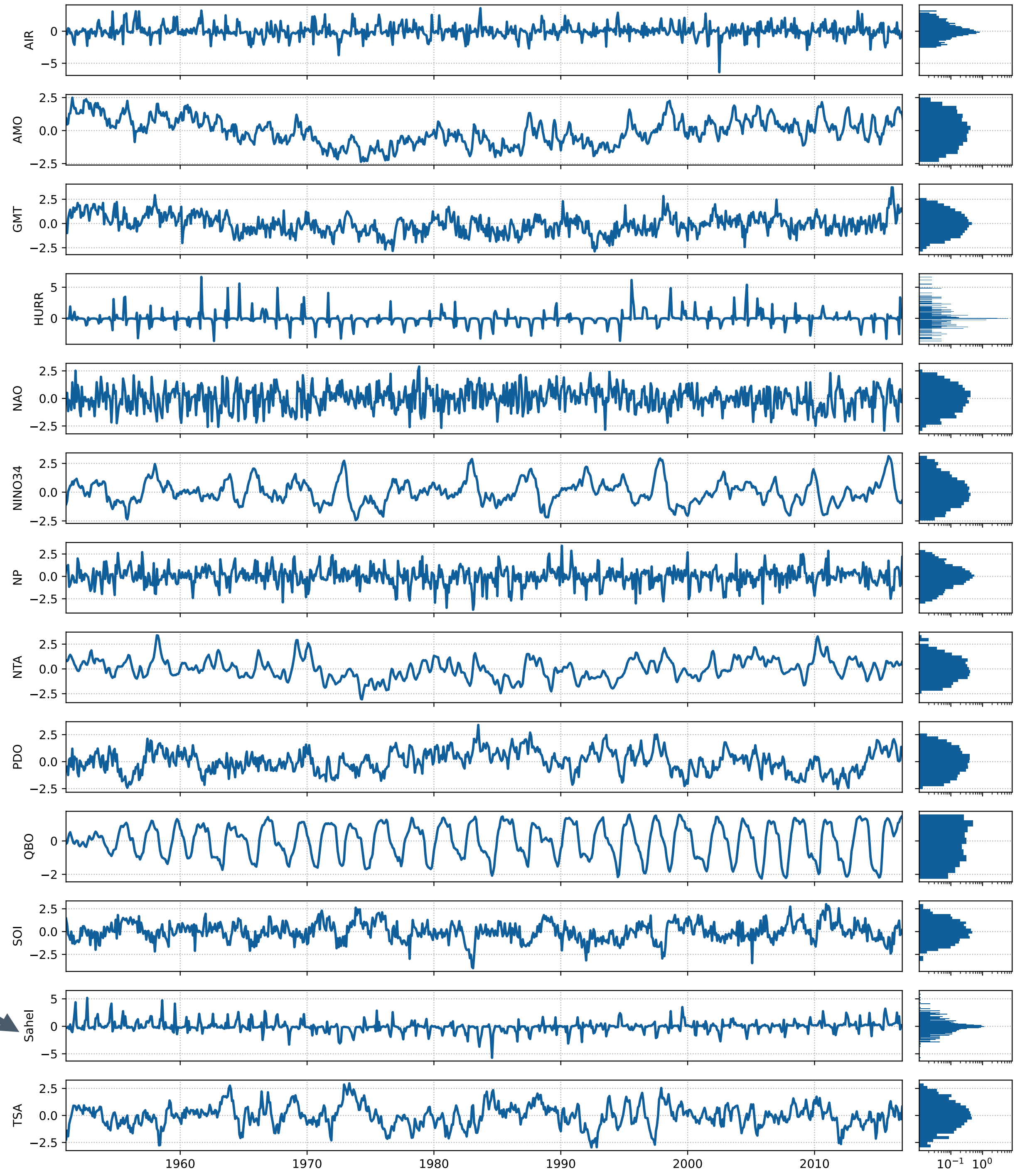
Climate Indices:

Monthly Atmospheric & Ocean Time-Series

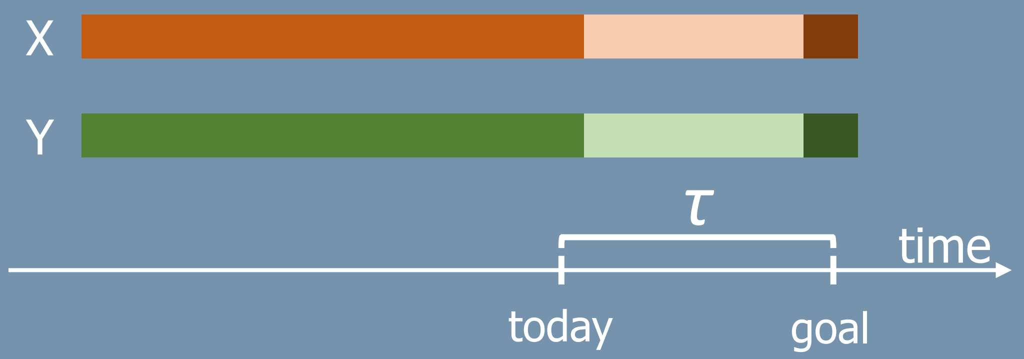
NINO34

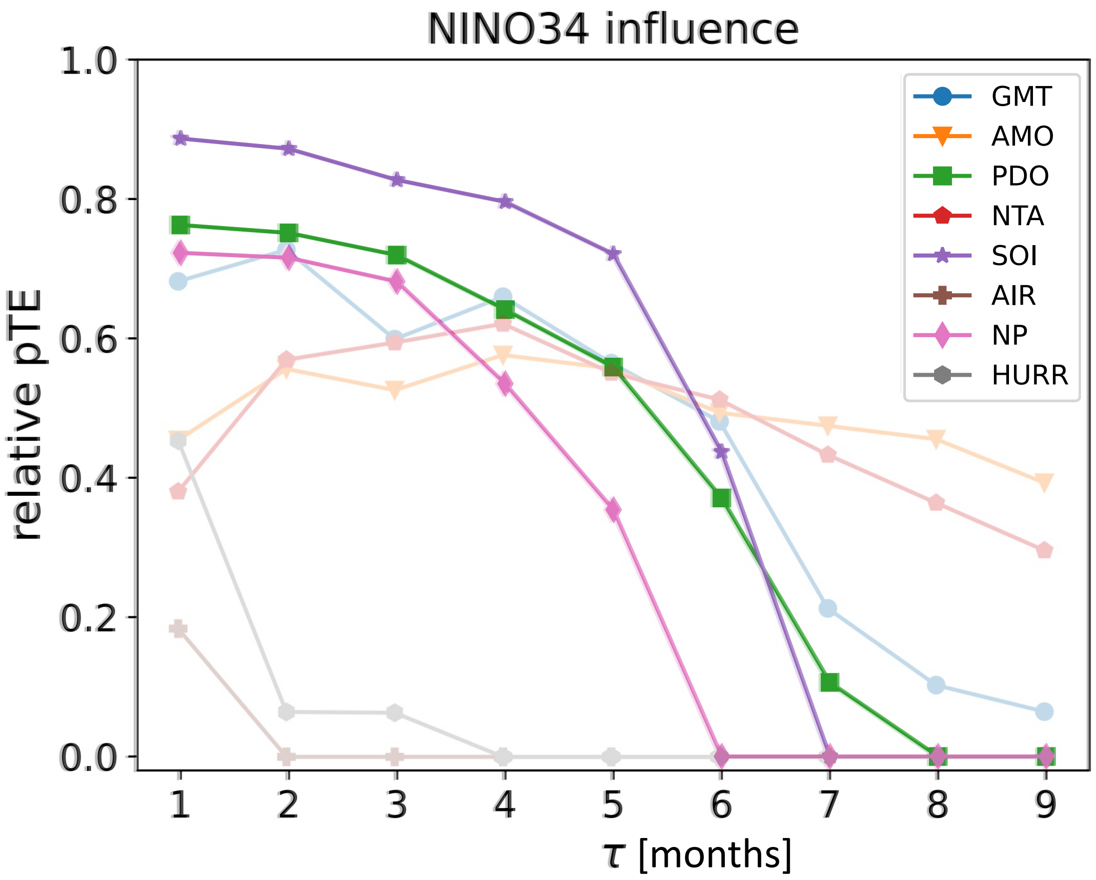
NTA

Sahel rainfall

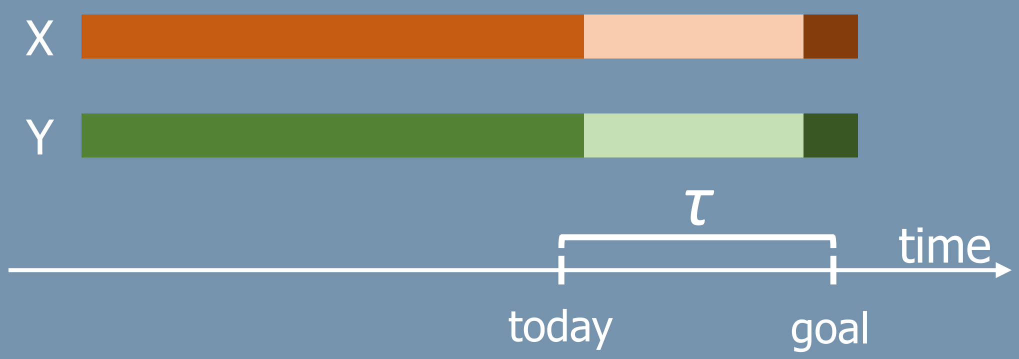


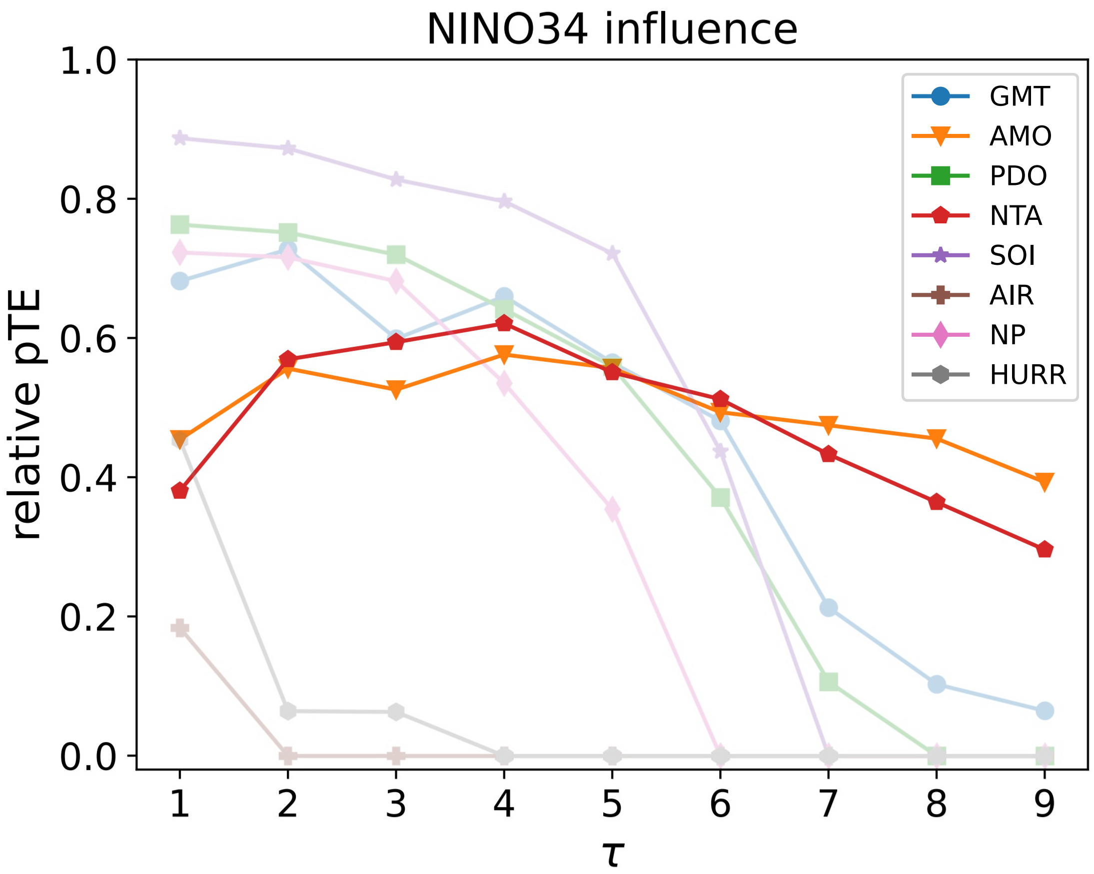
Lag τ


X
Y
τ
time
today
goal

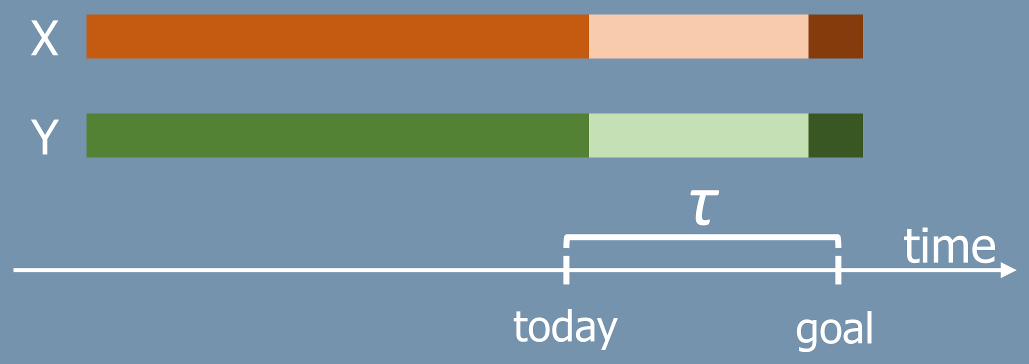

NINO34 influence
relative pTE
τ [months]
GMT
AMO
PDO
NTA
SOI
AIR
NP
HURR
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9

Lag τ

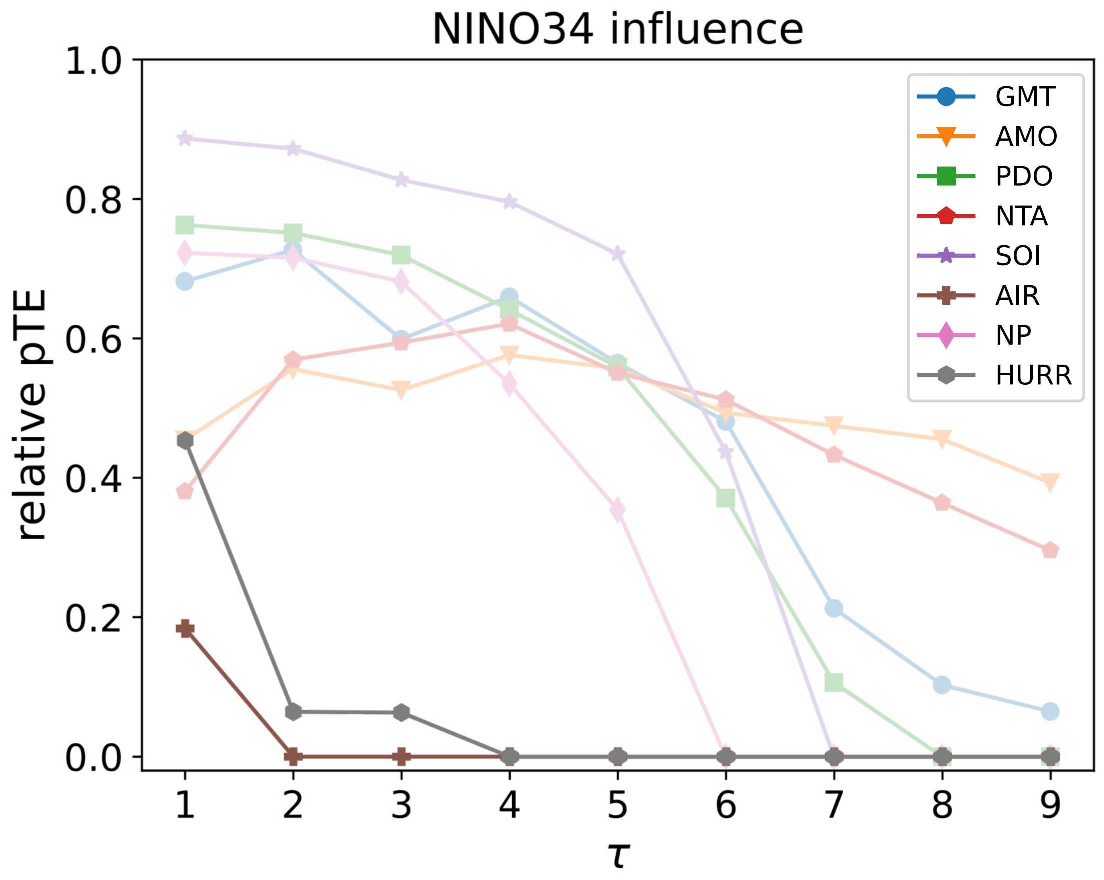

X
Y
τ
time
today
goal


NINO34 influence
relative pTE
τ
GMT
AMO
PDO
NTA
SOI
AIR
NP
HURR
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9

Lag τ



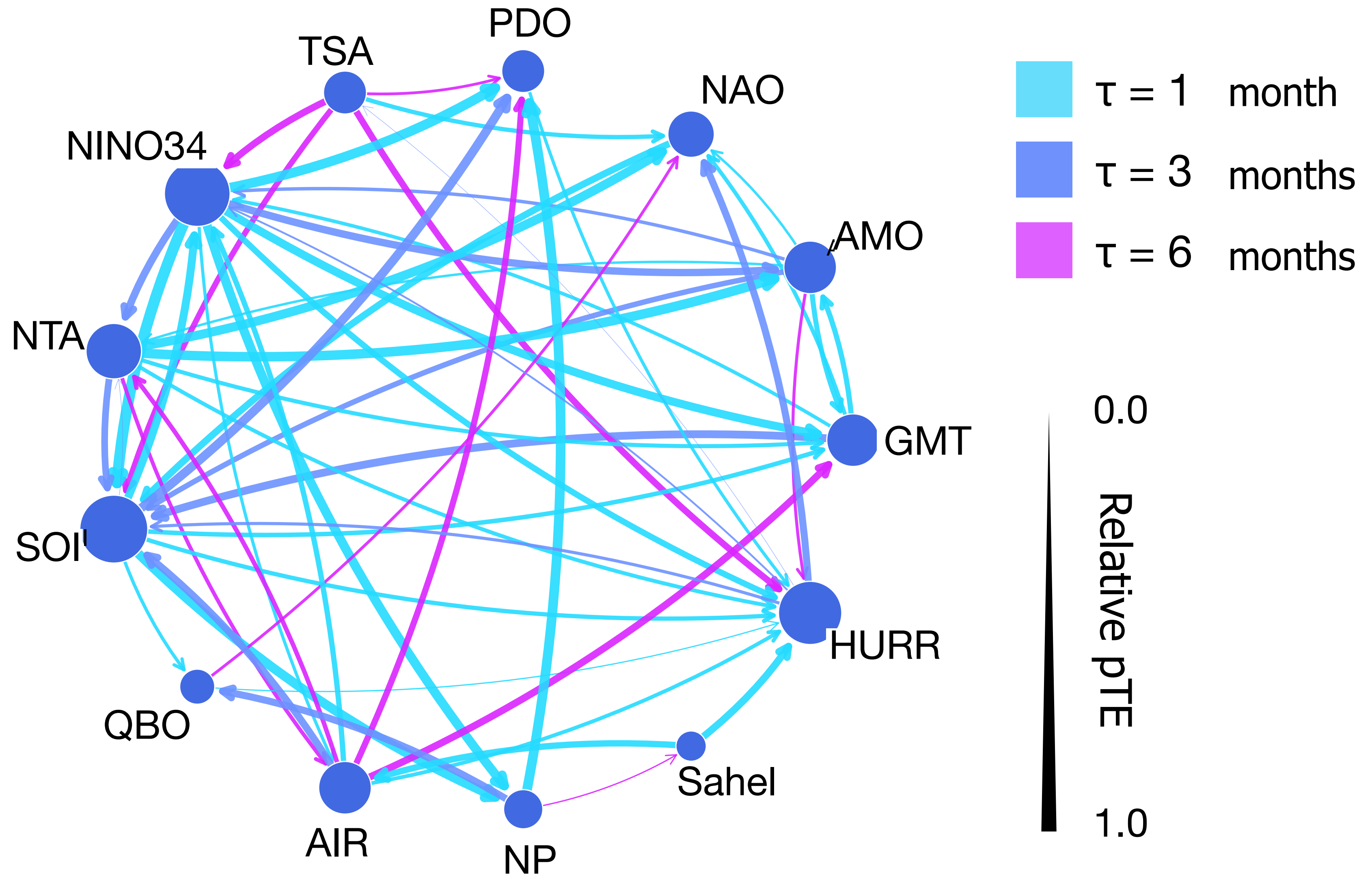
X
Y
τ
time
today
goal



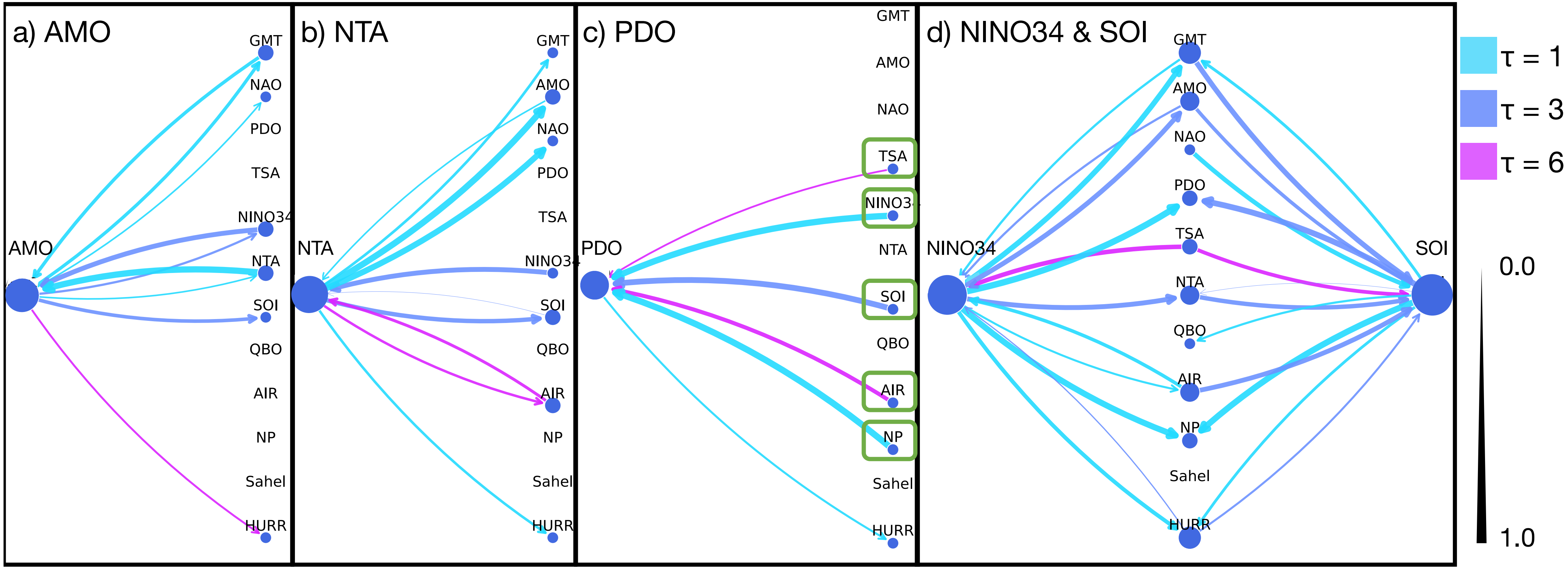
NINO34 influence
relative pTE
τ
GMT
AMO
PDO
NTA
SOI
AIR
NP
HURR
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9

Build causality networks

With lags 1, 3, 6 months



Some examples…



Contain useful **information** to improve the **prediction** of PDO

What

WHERE to get the information from

Why

Information used to predict the **future** of a variable

Problem

Information-theoretic causality measures are computationally **expensive**

Solution

Approximation with a **normal distribution**

Benefits

Fast and effective metric to find information transfer

- Good for linear and weakly nonlinear processes
- Good for relatively short time series

Publications

- Silini R, Tirabassi G., Barreiro M, Ferranti L & Masoller C, *Clim. Dyn.* (2022) , under review.