

# Interpretando la Descomposición Causal aplicada a Fenómenos Climáticos

## Interpreting Causal Decomposition applied to Climate Phenomena

Presentación: 8-9/10/2024

Doctorando:

**Juan Pablo MUSZKATS**

Universidad Nacional del Noroeste de la Provincia de Buenos Aires - Argentina  
[jpmuszkats@comunidad.unnoba.edu.ar](mailto:jpmuszkats@comunidad.unnoba.edu.ar)

Director:

**Miguel Eduardo ZITTO**

Codirectora:

**Rosa PIOTRKOWSKI**

### Resumen

La Descomposición Causal (CD) es una técnica novedosa que permite establecer niveles de causalidad entre series temporales. Para ello se vale de la Descomposición Empírica en Modos (EMD). En esta presentación se aplica la CD para establecer relaciones de causalidad después de un análisis preliminar con recursos derivados de la Transformada Wavelet Continua. Las series estudiadas provienen de fenómenos climáticos: el Niño – Oscilación del Sur (ENSO), precipitaciones de Buenos Aires (Argentina) y de Junín (Buenos Aires, Argentina). Como es de esperar, se observa que ENSO es un forzante de las series de precipitaciones. Sin embargo, se encuentran también presuntas relaciones de causalidad entre las series de precipitaciones que exigen una interpretación cuidadosa. Esta clase de dificultades forma parte de un debate actual en torno a la interpretación de la CD.

Palabras clave: Descomposición Causal, EMD, ENSO, Coherencia Wavelet

### Abstract

Causal Decomposition (CD) is a novel technique which establishes causality levels between time series. CD bases on the Empirical Mode Decomposition (EMD). In this presentation, CD is applied to establish causality relationships after a preliminary analysis with tools derived from the Continuous Wavelet Transform. The studied series stem from climate phenomena: el Niño-Southern Oscillation (ENSO) and precipitations in Buenos Aires (Argentina) and Junín (Buenos Aires, Argentina). As expected, ENSO is a driver of the precipitation series. However, causality relations are also found between the precipitation series, which deserve a careful interpretation. This difficulty is part of an ongoing debate about the interpretation of CD.

Keywords: Causal Decomposition, EMD, ENSO, Wavelet Coherence

## Introducción

La oscilación de los datos meteorológicos del Pacífico ecuatorial llamada fenómeno del Niño - Oscilación del Sur (ENSO) es un conocido forzante que incide sobre el clima en vastas regiones del planeta. Su influencia ha sido ampliamente estudiada con diversas técnicas que van desde la correlación hasta la medición de causalidad de Granger (Silva et al., 2021). Esta última consiste en una prueba estadística que permite establecer la capacidad predictiva de una serie sobre otra. Para ello se considera un modelo lineal y estacionario, en el que la causalidad requiere la precedencia temporal de la causa al efecto.

En el presente trabajo estudiamos y cuantificamos relaciones de causalidad entre un índice de ENSO y series de precipitaciones en la región pampeana. Como primer paso, utilizamos recursos de la Transformada Wavelet Continua para explorar posibles relaciones de causalidad a distintas frecuencias (Torrence y Compo, 1998; Grinsted et al., 2004). A continuación, aplicamos la novedosa técnica de Descomposición Causal (CD) propuesta por Yang et al. (2018). Dicha técnica está basada en la Descomposición Empírica en Modos (EMD): cada señal se descompone en Modos Intrínsecos de oscilación (IMFs por Intrinsic Mode Functions). A diferencia de las descomposiciones tradicionales como la de Fourier, esta técnica no proyecta la señal sobre una base prefijada sino que está guiada por los datos (Huang et al. 1998). Una vez que las señales se descomponen en sus respectivas IMFs se calcula la coherencia en fase de los modos correspondientes. Estudiando la coherencia antes y después de suprimir cada modo particular puede establecerse una medida y sentido de causalidad entre dos series.

Existen diversos métodos para realizar la EMD, muchos de los cuales implican el agregado de ruido a las señales. En particular, en este trabajo se aplicó la Noise Assisted - Multivariate EMD (NA-MEMD). Cada descomposición obtenida agregando ruido conlleva ligeras diferencias en el resultado. Por ende, la Descomposición Causal se repite como un experimento numérico y se valida estadísticamente con pruebas de hipótesis (Muszkats et al., 2024).

El método de CD ha sido, desde su publicación, aplicado en diversas ramas de la investigación científica. Sin embargo, sus resultados no están exentos de controversia. Chang et al. (2022) notan que dos series temporales podrían tener una influencia común a ambas y por ello exhibir causalidades espurias. A modo de ejemplo aplican la CD a un modelo de efecto Moran, en el que dos series que no están causalmente vinculadas tienen una correlación alta debido a un forzante aleatorio que comparten. Los autores del método de CD responden a la objeción señalando que la causalidad observada es esperable (Yang et al., 2022). Al haber un forzante común a ambas series, luego de un despeje de las ecuaciones se puede expresar a cada serie dependiendo de los valores anteriores de su contraparte. Más aún, la causalidad en un sentido u otro depende de cuánto influya el forzante sobre cada una de las series. Siguiendo esta línea de la discusión, en el presente trabajo se exploran relaciones de causalidad entre dos series de precipitaciones, sabiendo que el ENSO es un forzante en común.

## Métodos

Dada una serie  $x_n$ , su Transformada Wavelet Continua (CWT)  $W_n^X(s)$  se obtiene en este trabajo mediante la convolución con la wavelet de Morlet

$$\psi_0(\eta) = \pi^{-\frac{1}{4}} e^{i6\eta} e^{-\frac{1}{2}\eta^2} \quad (1)$$

normalizando la wavelet y cambiando la escala mediante  $\eta = s \cdot t$ . A partir de la CWT se calcula la Potencia Wavelet como  $|W_n^X(s)|^2$ . La Coherencia Wavelet entre dos series se define de manera análoga a la correlación tradicional, pero en el espacio de tiempo y frecuencia definido por las respectivas CWT (Grinsted et al., 2004):

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{XY}(s))|^2}{S(s^{-1}|W_n^X(s)|^2) \cdot S(s^{-1}|W_n^Y(s)|^2)} \quad (2)$$

donde  $S$  es un operador de suavizado y  $W^{XY} = W^X \cdot W^{Y*}$  es la Transformada Wavelet Cruzada (XWT).

La EMD expresa a una señal  $x(t)$  como una suma de  $n$  IMFs  $c_j(t)$  junto con un resto o tendencia  $r(t)$ :

$$x = \sum_{j=1}^n c_j + r \quad (3)$$

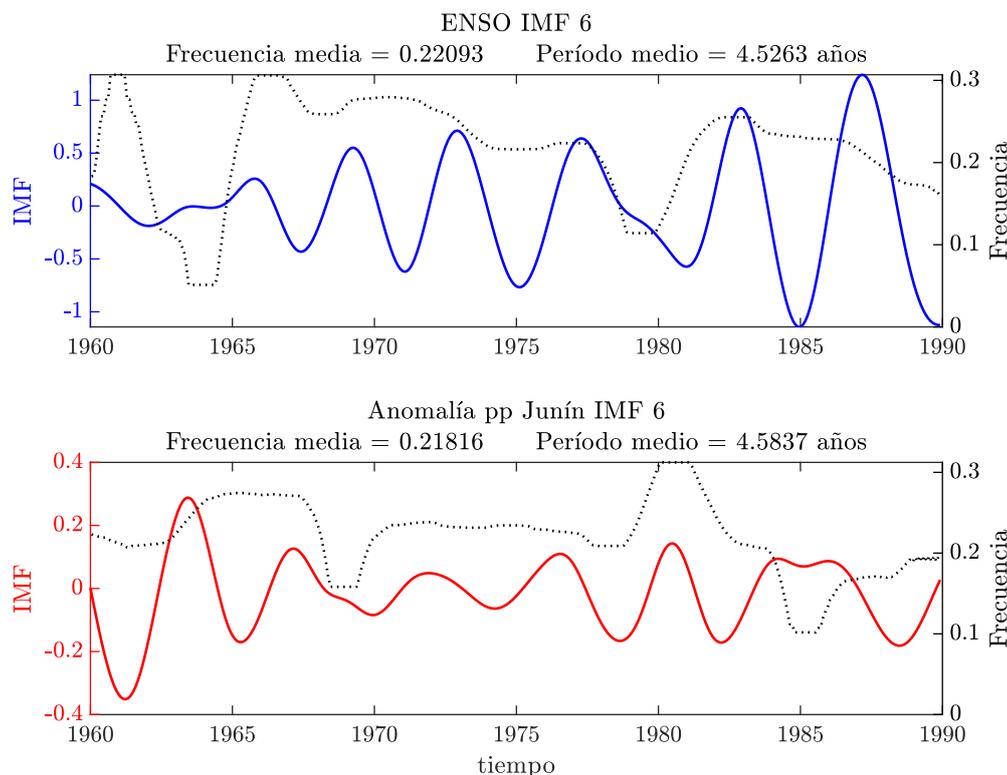
de manera tal que cada IMF es una señal oscilatoria que verifica

- el número de máximos y mínimos relativos coincide con el número de raíces o difiere, a lo sumo, en una unidad.
- en cualquier punto, el valor medio de la envolvente definida por los máximos relativos y la definida por los mínimos relativos es cero.

Esta descomposición permite expresar a cada IMF en términos de una amplitud y fase instantáneas

$$c_j(t) = a_j(t) \cos[\varphi_j(t)] \quad (4)$$

lo cual permite a su vez definir su frecuencia instantánea como  $\omega = \frac{d\varphi}{dt}$ . El algoritmo original de EMD dio origen a múltiples variantes. En el presente trabajo hemos optado por aplicar la NA-MEMD de Rehman y Mandic (2011). Este método presenta la ventaja de que descompone simultáneamente varias señales, como se



aprecia en la Figura 1. Con esto se logra una mejor alineación de los modos respectivos de las dos señales.

Figura 1: sexta IMF de las series de ENSO y de anomalías de precipitaciones en Junín obtenida mediante NA-MEMD. Se grafica también la frecuencia instantánea (suavizada con una media móvil).

A partir de la ecuación (4) puede definirse una medida de la coherencia de fase instantánea entre las respectivas IMFs de dos series  $s_1$  y  $s_2$ :

$$\text{coh}(s_{1j}, s_{2j}) = \frac{1}{T} \left| \int_0^T e^{i\Delta\varphi(t)} dt \right| \quad (5)$$

donde  $s_{1j}, s_{2j}$  son las  $j$ -ésimas IMFs de las respectivas series,  $T$  es su longitud y  $\Delta\varphi(t) = \varphi_{2j}(t) - \varphi_{1j}(t)$ . La coherencia alcanza su valor máximo 1 si la diferencia de fase permanece constante a lo largo del dominio  $[0, T]$ . Dadas dos series vinculadas causalmente, es de esperar que la coherencia disminuya si se suprime una IMF en una de ellas y se reconstruye la serie. De allí que se defina la fuerza causal mediante

$$D(s_{1j} \rightarrow s_{2j}) = \left\{ \frac{\sum_{k=1}^m \text{var}_{1k} \cdot \text{var}_{2k}}{\sum_{p=1}^m \text{var}_{1p} \cdot \text{var}_{2p}} [\text{coh}(s_{1k}, s_{2k}) - \text{coh}(s_{1k}, s'_{2k})]^2 \right\}^{1/2} \quad (6)$$

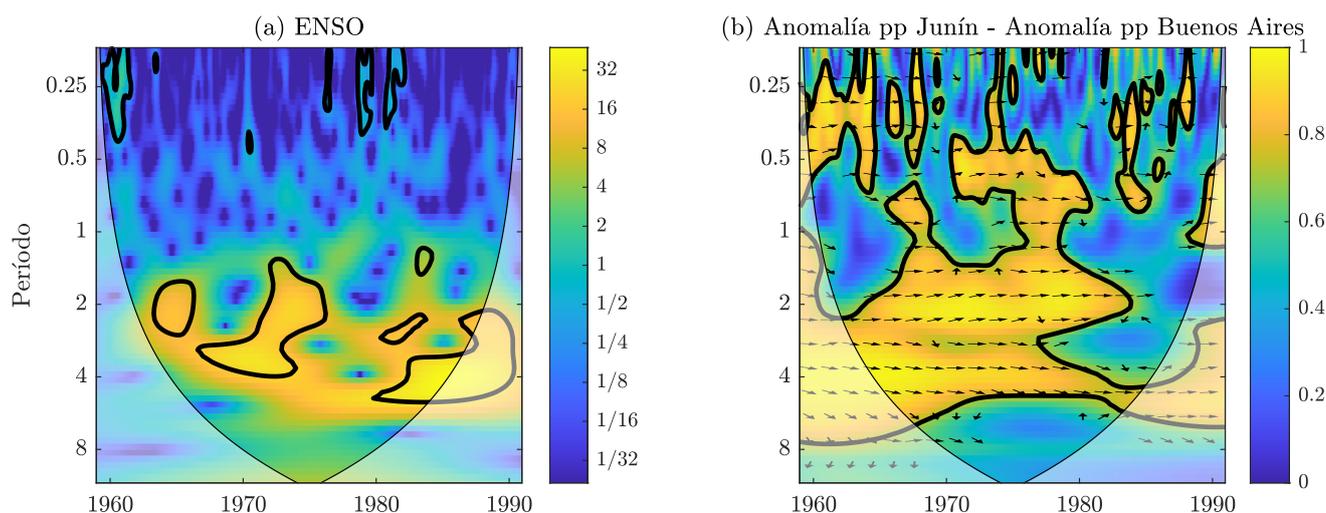
donde  $m$  es el total de IMFs y  $s'$  se refiere a la serie reconstruida después de suprimir su  $j$ -ésima IMF. Vale decir que  $D(s_{1j} \rightarrow s_{2j})$  mide cuán influyente es la serie 1 sobre la serie 2 al nivel de la  $j$ -ésima IMF. Su valor está entre 0 y 1 y si aparece una diferencia significativa sobre el valor de  $D(s_{2j} \rightarrow s_{1j})$  es indicativo de una causalidad diferencial de  $s_1$  sobre  $s_2$ . Esto es equivalente a que  $D(s_{1j} \rightarrow s_{2j}) - D(s_{2j} \rightarrow s_{1j})$  resulte positiva. Teniendo en cuenta que el valor fluctúa ligeramente por el agregado de ruido del NA-MEMD, el cálculo se repite 150 veces y se verifica estadísticamente la robustez de los valores promedio (Muszkats et al., 2024).

## Datos

Las series de precipitaciones de Buenos Aires y Junín (provincia de Buenos Aires) son las anomalías mensuales con respecto a los valores medios (Climate Explorer). En cuanto al índice de ENSO, se usó el Niño 3.4: la versión mensual de anomalías en la temperatura superficial del mar (Climate Explorer). El período de estudio elegido es 1960-1990 porque abarca episodios ENSO de gran amplitud (Torrence y Compo, 1998) y las series se encuentran completas, lo que evita la necesidad de interpolar valores.

## Resultados

En la Figura 2 se aprecia que la coherencia wavelet entre las precipitaciones en Junín y Buenos Aires alcanza



sus mayores niveles para los mismos períodos en que realiza su mayor potencia wavelet la señal de ENSO.

Figura 2: análisis con transformada wavelet. Los contornos gruesos indican regiones en que los valores tienen un nivel de significación del 5% contra la hipótesis nula de que se trata de ruido. Los datos que sufren efectos de borde se muestran con menor intensidad (cono de influencia). La fase relativa entre las series se ilustra con flechas, que apuntan hacia la derecha si el ángulo es cero y crecen en sentido antihorario. (a)

Potencia wavelet de la serie ENSO. (b) Coherencia wavelet entre las precipitaciones. Se destaca que están mayormente en fase.

En la Tabla 1 se exhiben las fuerzas causales entre ENSO y las precipitaciones en Buenos Aires y Junín. Para las IMFs 5 y 6 se verifica una marcada diferencia causal a favor de ENSO que pudo comprobarse estadísticamente con un nivel de significación  $\alpha = 0,05$ . La Tabla 2 exhibe los períodos medios correspondientes a cada IMF. Las IMFs 5 y 6 se corresponden aproximadamente con los períodos de mayor potencia wavelet de ENSO.

IMF	D(ENSO→BsAs)	D(BsAs→ENSO)	D(ENSO→Junín)	D(Junín→ENSO)
1	0,065	0,062	0,064	0,063
2	0,064	0,064	0,070	0,071
3	0,053	0,051	0,070	0,066
4	0,066	0,064	0,078	0,075
5	0,109	0,083	0,114	0,079
6	0,098	0,064	0,085	0,072

Tabla 1: Fuerzas causales entre ENSO y precipitaciones. Promedio de 150 realizaciones siguiendo la definición (6). No se exhiben las IMFs mayores que 6 por corresponder a oscilaciones de muy baja amplitud.

IMF	Período medio ENSO	Período medio Junín	Período medio Buenos Aires
1	0,29	0,29	0,30
2	0,44	0,40	0,42
3	0,72	0,73	0,68
4	1,30	1,30	1,27
5	2,69	2,03	2,05
6	4,14	4,26	4,29

Tabla 2: Períodos medios de cada IMF de ENSO y precipitaciones. Promedio de 150 realizaciones.

En la Tabla 3 se exhiben las fuerzas causales entre las precipitaciones en Junín y Buenos Aires. Si bien en las IMFs 5 y 6 no aparecen fuerzas causales notables, se observa una marcada diferencia causal en la IMF 1.

IMF	D(Junín→BsAs)	D(BsAs→Junín)
1	0,122	0,158
2	0,107	0,105
3	0,078	0,071
4	0,042	0,059
5	0,044	0,047
6	0,032	0,029

Tabla 3: Fuerzas causales entre Junín y Buenos Aires. Promedio de 150 realizaciones siguiendo la definición (6). No se exhiben las IMFs mayores que 6 por corresponder a oscilaciones de muy baja amplitud.

## Conclusiones

Los resultados de la Descomposición Causal son coherentes con las relaciones esperadas entre ENSO y las precipitaciones. Más aún, al establecer niveles de causalidad para los distintos modos de oscilación se aprecia la influencia de ENSO en sus niveles de mayor potencia wavelet. En cuanto a la presunta causalidad entre las precipitaciones de Buenos Aires y Junín, es un resultado que puede esperarse de la misma forma que ocurre en

el modelo del efecto Moran: las series tienen un forzante en común (y presumiblemente otros), lo cual también se refleja en la marcada coherencia wavelet que exhiben. Lo hecho hasta aquí sugiere profundizar la investigación acerca del aporte del análisis wavelet a la interpretación de la Descomposición Causal.

## Referencias

- Chang, C-W., Munch, S. B., y Hsieh, C-H. (2022). Comments on identifying causal relationships in nonlinear dynamical systems via empirical mode decomposition. *Nature communications*, 13(1), Artículo 2860. <https://doi.org/10.1038/s41467-022-30359-8>
- Climate Explorer. World Meteorological Organization. *Nino3.4 [K] HadISST1 Nino3.4 index*. Descargado en Julio 26, 2024 desde [https://climexp.knmi.nl/getindices.cgi?WMO=UKMODData/hadisst1\\_nino3.4a&STATION=NINO3.4&TYPE=i&id=someone@somewhere](https://climexp.knmi.nl/getindices.cgi?WMO=UKMODData/hadisst1_nino3.4a&STATION=NINO3.4&TYPE=i&id=someone@somewhere)
- Climate Explorer. World Meteorological Organization. *Monthly GHCN v2 precipitation Buenos Aires (Argentina)*. Descargado en Julio 26, 2024 desde <https://climexp.knmi.nl/getprcpall.cgi?WMO=87585>
- Climate Explorer. World Meteorological Organization. *Monthly GHCN v2 precipitation Junín Aero (Argentina)*. Descargado en Julio 26, 2024 desde [https://climexp.knmi.nl/getprcpall.cgi?id=someone@somewhere&WMO=87548&STATION=JUNIN\\_AERO&extraargs=](https://climexp.knmi.nl/getprcpall.cgi?id=someone@somewhere&WMO=87548&STATION=JUNIN_AERO&extraargs=)
- Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear processes in geophysics*, 11(5/6), 561-566. <https://doi.org/10.5194/npg-11-561-2004>
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung C.C. y Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, 454(1971), 903-995.
- Muszkats, J. P., Muszkats, S. R., Zitto, M. E. y Piotrkowski, R. (2024). A statistical analysis of causal decomposition methods applied to Earth system time series. *Physica A: Statistical Mechanics and its Applications*, 641, Artículo 129708. <https://doi.org/10.1016/j.physa.2024.129708>
- Rehman, N. y Mandic, D. P. (2011). Filter bank property of multivariate empirical mode decomposition. *IEEE transactions on signal processing*, 59(5), 2421-2426. <https://doi.org/10.1109/TSP.2011.2106779>
- Silva, F. N., Vega-Oliveros, D. A., Yan, X., Flammini, A., Menczer, F., Radicchi, F., Kravitz, B. y Fortunato, S. (2021). Detecting climate teleconnections with Granger causality. *Geophysical Research Letters*, 48(18), Artículo e2021GL094707. <https://doi.org/10.1029/2021GL094707>
- Torrence, C. y Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79 (1), 61-78.
- Yang, A. C., Peng, C-K y Huang, N. E. (2018). Causal decomposition in the mutual causation system. *Nature communications*, 9 (1), 1-10. <https://doi.org/10.1038/s41467-018-05845-7>
- Yang, A.C., Peng, C-K. y Huang, N.E. (2022). Reply To: Comments on identifying causal relationships in nonlinear dynamical systems via empirical mode decomposition. *Nature Communications* 13, Artículo 2859. <https://doi.org/10.1038/s41467-022-30360-1>