

CARACTERIZACIÓN DE SERIES CLIMÁTICAS USANDO EL EXPONENTE DE HURST

CHARACTERIZATION OF TIME SERIES USING HURST EXPONENT

María B. Arouxet^a y Verónica E. Pastor^b

^a*Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina, belen@mate.unlp.edu.ar*

^b*Facultad de Ingeniería, Universidad Nacional de Buenos Aires, Argentina, veronicapastor@gmail.com*

Palabras clave: Series temporales, Exponente de Hurst, Transformada Wavelet, Rango Rescalado.

Resumen. El exponente de Hurst (H) es un parámetro que varía entre 0 y 1 e indica el nivel de persistencia de una serie temporal. Valores de H entre 0 y 0,5 indican que la serie es antipersistente mientras que valores de H entre 0,5 y 1 indican que la serie es persistente, si $H=0,5$ se dice que la serie es aleatoria. En estudios previos realizamos una mejora del método Promedio de los Coeficientes Wavelet (AWC, *average wavelet coefficient*) para el cálculo del exponente de Hurst utilizando diferentes estimadores para la varianza de los coeficientes de los detalles wavelet. A fin de realizar las comparaciones numéricas de las distintas implementaciones, previamente hemos generado series artificiales con distintos valores de H utilizando transformada wavelet. En este trabajo aplicamos dicha herramienta a series climáticas de la República Argentina. En la bibliografía consultada hemos encontrado métodos que involucran wavelets pero no especifican la familia más apropiada para tal fin ni si se requiere un análisis previo de la serie a estudiar.

Keywords: Time Series, Hurst Exponent, Wavelet Transform, Rescaled Range.

Abstract. The Hurst exponent (H) is a parameter that varies between 0 and 1 and indicates the persistence level of a time series. Values of H between 0 and 0.5 indicate that the series is antipersistent, while values of H between 0.5 and 1 indicate that the series is persistent, if $H = 0.5$ the series is said to be random. In previous studies we performed an improvement of the Average Wavelet Coefficient (AWC) method for the calculation of the Hurst exponent using different estimators for the variance of the wavelet details coefficients. In order to make the numerical comparisons of the different implementations, we have previously generated artificial series with different values of H using wavelet transform. In this work we apply this tool to climatic series of Argentine Republic. In the consulted bibliography we have found methods that involve wavelets but do not specify the most appropriate family for that purpose or if a previous analysis of the series to be studied is required.

1. INTRODUCCIÓN

La predicción de sequías e inundaciones es uno de los temas importantes en cuanto a investigación climatológica en el último tiempo, incluso en la República Argentina. Saber cuándo podrían ocurrir, o no, las precipitaciones significaría poder alertar a agricultores, ganaderos, y a la población en general a fin de resguardar la producción o resguardarse. En los últimos años, debido al cambio climático se han producido estos fenómenos en distintos puntos del país, incluso han ocurrido sequías o inundaciones en lugares donde nunca habían ocurrido antes. Además de ser este tipo de información de conocimiento público, hay grupos y centros de investigación dedicados a ello, pero por lo general se basan en estudios estadísticos y no consideran la naturaleza fractal de la series.

Para caracterizar la fractalidad de una serie temporal se puede hacer por medio del exponente de Hurst, (Dimensión fractal=2-Exponente de Hurst). Este parámetro tuvo origen en el área de la hidrología (Hurst (1951)), sin embargo su uso se extendió a diversas áreas, incluso las de origen humano, como la financiera en Serinaldi (2010).

Dicho exponente es un índice en el intervalo $[0; 1]$ que mide la persistencia de una serie temporal. Si $H \in (0,5; 1]$ se dice que la serie es **persistente**, es decir está caracterizada por efectos de memoria de largo plazo. Teóricamente lo que suceda hoy impactará en el futuro por siempre, como por ejemplo cambios mensuales de ahora están correlacionados con los cambios mensuales futuros. Por ejemplo, si $H = 0,7$ nos indica que existe un 70% de probabilidades que el siguiente período siga la misma tendencia que en el período anterior. Si $H \in [0; 0,5)$ se dice que la serie es **antipersistente**, con frecuencia se les denomina con reversión a la media¹. Es decir, si la serie ha estado arriba de un determinado valor que hace las veces de media de largo plazo en el período anterior, es más probable que esté abajo en el período siguiente y viceversa. Si $H = 0,5$, la serie es completamente aleatoria, los datos son independientes y decimos que no hay memoria.

Existen en la actualidad distintos métodos para estimar el exponente de Hurst, entre los que podemos citar el rango reescalado propuesto por Hurst (1951), el análisis de fluctuación sin tendencia (DFA, *detrended fluctuation analysis*) de Peng et al. (1995), y el análisis wavelet de Simonsen et al. (1998), entre otros. Una revisión de los métodos se encuentra en L. et al. (2017) donde los autores recopilan la información de los distintos métodos aplicados a series de precipitaciones y temperatura a nivel mundial donde se aprecia que las mayores concentraciones de investigación están en Europa y Medio Oriente. Los autores Rangarajan y Sant (2004) han estudiado las series de precipitaciones en India y han introducido un índice de predicción (PI) basados en la dimensión fractal, D , o el exponente de Hurst, H :

$$PI = 2 \times |D - 1,5| = 2 \times |0,5 - H|.$$

Coincidiendo con lo reportado en L. et al. (2017), en nuestra búsqueda bibliográfica a nivel nacional sólo hemos encontrado publicado el trabajo de Pérez et al. (2009) quienes calculan el exponente de Hurst por medio del rango rescalado.

En este trabajo estudiaremos series de precipitaciones mensuales de distintas regiones climáticas de Argentina, con el fin de caracterizarlas según su persistencia o no por medio del promedio de los coeficientes wavelet, considerando las mejoras introducidas en Arouxet y Pastor (2017).

¹Se usa para describir cómo las rentabilidades en una serie temporal financiera pueden ser extremas a corto plazo, pero más estables a largo plazo.

2. MÉTODOS

2.1. Series de precipitaciones

Una serie temporal es una sucesión de observaciones de una variable, la precipitación es una serie que evoluciona tanto en el tiempo como en el espacio. Por este motivo, no sólo analizamos la evolución temporal sino también su comportamiento en diferentes regiones climáticas de la República Argentina. Dichas series fueron obtenidas de la Base de Datos Climáticos de la [3CN-CIMA/CONICET-UBA \(2014\)](#), eligiendo el período que va de 1960 a 2010 de datos grillados, no los simulados. La frecuencia aquí presentada es mensual, pero también se han analizado series diarias.

La República Argentina es un país extenso el cual abarca una gran diversidad de climas y se lo divide en 5 regiones: andina, centro o chacopampeana, húmeda o mesopotámica, patagónica y territorios insulares en el Atlántico sur. En la Tabla 1 se muestran las 32 series temporales seleccionadas: las siete primeras series corresponden a la región andina, las once siguientes al centro del país, las once que siguen a la región húmeda y las últimas tres a la región patagónica. De esta última región no se seleccionaron más series por no ser lo suficientemente abarcativa la base de datos por debajo del paralelo 40° S, como así tampoco se han seleccionado series de los territorios insulares en el Atlántico sur. En la Fig. 1 se muestra la ubicación de las series seleccionadas.

Si bien se considera que Argentina tiene clima templado, tiene regiones muy húmedas como Misiones y regiones muy áridas como San Juan. En la Fig. 2 se puede observar las series de precipitaciones de dos locaciones de las provincias de Misiones y San Juan, El Dorado y Campamento Filo del Sol, respectivamente.



Figura 1: Ubicación de las series de precipitaciones seleccionadas en el mapa de la República Argentina, para latitudes menores a $40^\circ 5'$. ([3CN-CIMA/CONICET-UBA \(2014\)](#))

2.2. Método wavelet

En la literatura existen variados métodos para el cálculo del exponente de Hurst, desde el rango rescalado hasta teoría de fractales pasando por DFA ([Peng et al. \(1995\)](#)), etc. Dos de los

Provincia	Localidad	Ubicación de la serie	Altitud snm
Jujuy	Susques	Lat. 23°5' long. 66°5'	4225
Catamarca	Cerro El Cóndor	Lat. 26°5' long. 68°5'	5399
San Juan	Campamento Filo del Sol	Lat. 28°5' long. 69°5'	4087
La Rioja	Tama	Lat. 30°5' long. 66°5'	730
Mendoza	Puente del Inca	Lat. 32°5' long. 69°5'	4200
Mendoza	San Rafael	Lat. 34°5' long. 68°5'	696
Mendoza	Malargüe	Lat. 36°5' long. 68°5'	1316
Formosa	Comunidad La Rinconada	Lat. 23°5' long. 61°5'	178
Formosa	Formosa	Lat. 26°5' long. 58°5'	72
Santiago del Estero	Alberdi	Lat. 26°5' long. 62°5'	185
Santiago del Estero	Mailín	Lat. 28°5' long. 63°5'	116
Chaco	Santa Sylvina	Lat. 27°5' long. 61°5'	61
Tucumán	La Cocha	Lat. 27°5' long. 65°5'	397
Córdoba	Leones	Lat. 32°5' long. 62°5'	113
Córdoba	Obispo Trejo	Lat. 30°5' long. 63°5'	102
San Luis	La Pampa Grande	Lat. 32°5' long. 66°5'	611
San Luis	Arizona	Lat. 35°5' long. 65°5'	317
La Pampa	Colonia Santa Teresa	Lat. 37°5' long. 63°5'	155
La Pampa	Santa Rosa	Lat. 36°5' long. 64°5'	217
La Pampa	Gdor. Duval	Lat. 38°5' long. 66°5'	206
Misiones	El Dorado	Lat. 26°5' long. 54°5'	217
Corrientes	Gral. Alvear	Lat. 28°5' long. 56°5'	79
Santa Fé	Puerto Ocampo	Lat. 28°5' long. 59°5'	54
Entre Ríos	San Gustavo	Lat. 30°5' long. 59°5'	50
Entre Ríos	Gdor. Mansilla	Lat. 32°5' long. 59°5'	46
CABA	CABA	Lat. 34°5' long. 58°5'	30
Buenos Aires	Blaquier	Lat. 34°5' long. 62°5'	109
Buenos Aires	Junín	Lat. 35°5' long. 60°5'	81
Buenos Aires	Tandil	Lat. 37°5' long. 59°5'	216
Buenos Aires	Bahía Blanca	Lat. 38°5' long. 62°5'	62
Buenos Aires	Viedma	Lat. 40°5' long. 62°5'	44
Neuquén	Añelo	Lat. 38°5' long. 69°5'	481
Neuquén	Achico	Lat. 40°5' long. 70°5'	969
Río Negro	Sierra Colorada	Lat. 40°5' long. 67°5'	520

Tabla 1: Ubicación de las series de precipitaciones mensuales de distintas regiones de la Argentina.

métodos más reconocidos son el de rango rescalado y el método de wavelet. En trabajos previos, hemos observado que dada una serie de tiempo artificial cualquiera, el método wavelet no sólo encuentra un valor de H más aproximado que el rango rescalado, sino que no es necesario saber *a priori* la naturaleza de la serie, esto es, si la serie es estacionaria o no estacionaria. Por lo dicho anteriormente, en este trabajo utilizamos sólo el método wavelet.

La transformada wavelet permite descomponer la serie de tiempo en el dominio tiempo-frecuencia, y se define mediante:

$$W(a, b) = \int_{-\infty}^{\infty} f(s)\psi_{a,b}(s)ds \quad \text{para } a > 0 \quad (1)$$

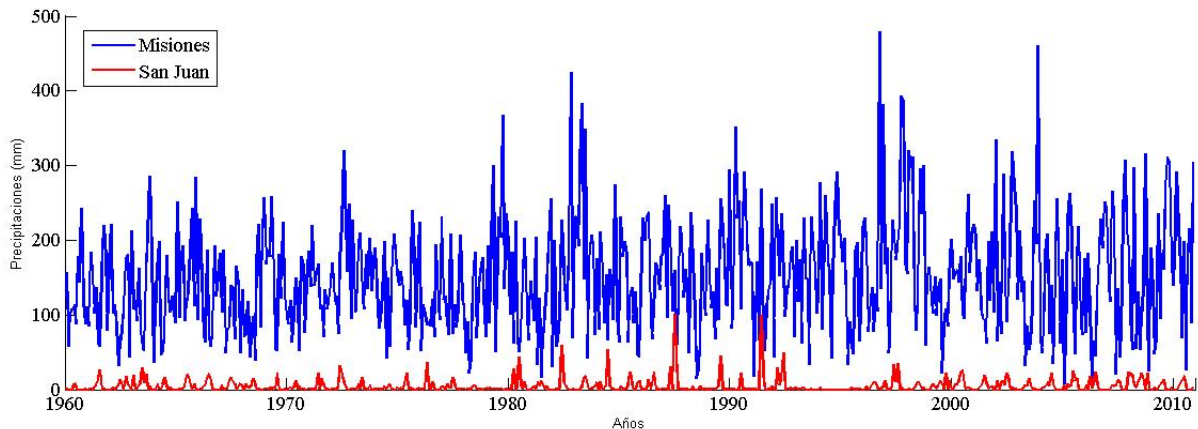


Figura 2: Series de precipitaciones, en milímetros, de Misiones y San Juan entre los años 1960 y 2010.

donde a es el parámetro de escala, b el de traslación y $\psi(t)$ es la wavelet madre.

Si la serie de tiempo posee H-autoafinidad, esto es, que la serie satisface una ley de potencia del tipo $X(ct) \approx c^H X(t)$, la varianza de la transformada wavelet (Ec. (1)) será afectada por un parámetro de escala de manera asintótica:

$$Var(a) = \mathbb{E}(W(a, b))^2 - (\mathbb{E}(W(a, b)))^2 \approx a^\beta \quad (2)$$

donde el exponente $\beta \in [-1, 3]$. Malamud y Turcotte (1999) encontraron una relación entre los parámetros H y β para series autoafines. Por lo que el exponente de Hurst queda definido como:

- $H = \frac{\beta+1}{2}$, para $\beta \in [-1, 1)$, si la señal es del tipo fGn,
- $H = \frac{\beta-1}{2}$, para $\beta \in [1, 3]$, si la señal es del tipo fBm,

donde fBm es movimiento Browniano fraccionario y fGn es ruido Gaussiano fractal.

El método wavelet que hemos utilizado es un método basado en el método AWC de Simonsen et al. (1998) el cual propone hacer los siguientes pasos:

- 1 se aplica la transformada wavelet de los datos en el dominio wavelet, $W(a, b)$;
- 2 para una escala fija, a , se calcula el promedio de los coeficientes wavelet;
- 3 se grafica en doble logaritmo los coeficientes vs. la escala a .

En trabajos previos hemos presentado mejoras al método wavelet utilizando estimadores más robustos para el cálculo en el punto [2]. Basados en la Ec. 1, en Arouxet y Pastor (2017), hemos propuesto hacer una estimación de la varianza de los coeficientes de los detalles de la serie estudiada utilizando dos estimadores: el primero es el conocido estimador insesgado de la varianza $\hat{V}ar$ en el cual dada una serie de datos x_1, \dots, x_m usando el estimador:

$$\hat{V}ar(x) = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2,$$

y el segundo es un estimador más robusto de la varianza como es la desviación mediana absoluta (MAD, *median absolut deviation*) (Maronna et al. (2006)):

$$MAD(W(a, b)) = Med(|W(a, b) - Med(W(a, b))|) \quad (3)$$

En este trabajo utilizamos el método AWC y el estimador MAD, ya que por los resultados obtenidos en los trabajos previos, usando series artificiales, hemos mostrado que así es más eficiente que al usar los otros estimadores de la varianza de los coeficientes wavelet.

3. RESULTADOS Y CONCLUSIONES

Las series de precipitaciones seleccionadas de [3CN-CIMA/CONICET-UBA \(2014\)](#) son de frecuencia mensual entre los años 1960 y 2010, con 612 datos cada una.

El método AWC utilizando el MAD para estimar la varianza de los coeficientes wavelet fue implementado en MATLAB, con la instrucciones `wavedec.m`, para la descomposición con el máximo nivel permitido para la wavelet Haar, y `wnoisest.m` para estimar el MAD. La elección de la instrucciones `wavedec.m` se debe a que si bien la serie temporal es discreta, proviene de una variable, las precipitaciones en este caso, que son continuas, y la mejor aproximación es esta transformada discreta, que considera los multiniveles.

En la Tabla 2 se muestran los valores obtenidos para las 32 series temporales tanto del exponente de Hurst, H , como el índice de predicción, PI .

De la Tabla 2 se observa que de las 32 series observadas, 17 son antipersistentes, de las cuales 13 con $H < 0,45$, y 15 persistentes, de las cuales $H > 0,55$ en 9 de ellas. Además, en 10 series el exponente de Hurst toma valores entre $0,45$ y $0,55$, para las cuáles se podría considerar que tienen un comportamiento casi aleatorio, incluso el promedio de los valores de H es $0,48$. En cuanto al índice de predicción, los valores obtenidos son menores que $0,5$ con un promedio de $0,23$. Todas las series, salvo la de Leones, en la Provincia de Córdoba en la región centro, tienen PI menor que $0,5$, con $H = 0,2$. En nuestro trabajo se refleja que las series resultan ser algunas persistentes, aún cuando las series elegidas son series en donde las precipitaciones son muy abundantes como en Misiones o muy escasas como en San Juan (Fig. 2) y en otros casos son antipersistentes en lugares como Entre Ríos y Jujuy en donde en la primera las lluvias son abundantes y en la segunda son escasas (Fig. 3) o Córdoba y Mendoza (Fig. 4).

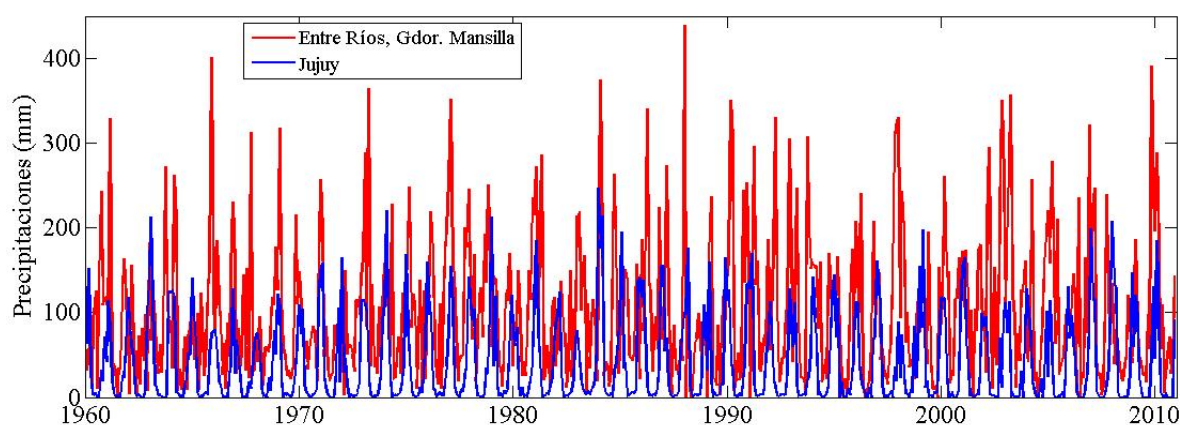


Figura 3: Series de precipitaciones, en milímetros, de Jujuy y Entre Ríos (Gdor. Mansilla) entre los años 1960 y 2010.

Los resultados obtenidos demuestran algunas diferencias con los análisis preliminares realizado por otros autores en otras regiones del mundo. Por ejemplo, India ([Rangarajan y Sant \(2004\)](#)) o en México ([Valle et al. \(2013\)](#)) donde estos últimos autores muestran que los índices

Provincia	Localidad	H	PI
Jujuy	Susques	0,4314	0,1372
Catamarca	Cerro El Cóndor	0,4419	0,1162
San Juan	Campamento Filo del Sol	0,7175	0,4350
La Rioja	Tama	0,4945	0,0110
Mendoza	Puente del Inca	0,6757	0,3514
Mendoza	San Rafael	0,4225	0,1551
Mendoza	Malargüe	0,6589	0,3178
Formosa	Comunidad La Rinconada	0,2668	0,4664
Formosa	Formosa	0,3023	0,3954
Santiago del Estero	Alberdi	0,4500	0,1001
Santiago del Estero	Mailín	0,2721	0,4559
Chaco	Santa Sylvina	0,5365	0,0730
Tucumán	La Cocha	0,4320	0,1360
Córdoba	Leones	0,2063	0,5875
Córdoba	Obispo Trejo	0,2673	0,4654
San Luis	La Pampa Grande	0,4695	0,0610
San Luis	Arizona	0,3455	0,3090
La Pampa	Colonia Santa Teresa	0,5345	0,0691
La Pampa	Santa Rosa	0,6277	0,2554
La Pampa	Gdor. Duval	0,5471	0,0941
Misiones	El Dorado	0,6651	0,3302
Corrientes	Gral. Alvear	0,5602	0,1204
Santa Fé	Puerto Ocampo	0,4810	0,0380
Entre Ríos	San Gustavo	0,4305	0,1390
Entre Ríos	Gdor. Mansilla	0,2978	0,4045
CABA	CABA	0,3317	0,3367
Buenos Aires	Blaquier	0,5494	0,0988
Buenos Aires	Junín	0,5391	0,0783
Buenos Aires	Tandil	0,3784	0,2432
Buenos Aires	Bahía Blanca	0,5654	0,1307
Buenos Aires	Viedma	0,5324	0,0647
Neuquén	Añelo	0,5935	0,1870
Neuquén	Achico	0,5018	0,0037
Río Negro	Sierra Colorada	0,6374	0,2748

Tabla 2: Valores del Exponente de Hurst y del índice de predicción para series de precipitaciones mensuales de distintas regiones de la Argentina.

de predicción de las series de precipitaciones el estado de Zacatecas son valores cercanos a 1, con valores de H menores a 0,3 demostrando que dichas series resultan ser antipersistentes.

4. DISCUSIONES Y TRABAJO FUTURO

Este estudio preliminar de las series de precipitaciones del territorio argentino es muy generalizado ya que la región abarca zonas climáticas diferentes entre sí, desde zonas de montaña y lejanas al mar (Jujuy) hasta zonas cercanas al mar y de poca altitud (Río Negro). En los resultados podemos observar que hay series de ambos tipos, en las distintas regiones consideradas. Por ejemplo, la serie de Tandil es antipersistente y se encuentra a una altitud sobre el nivel del mar

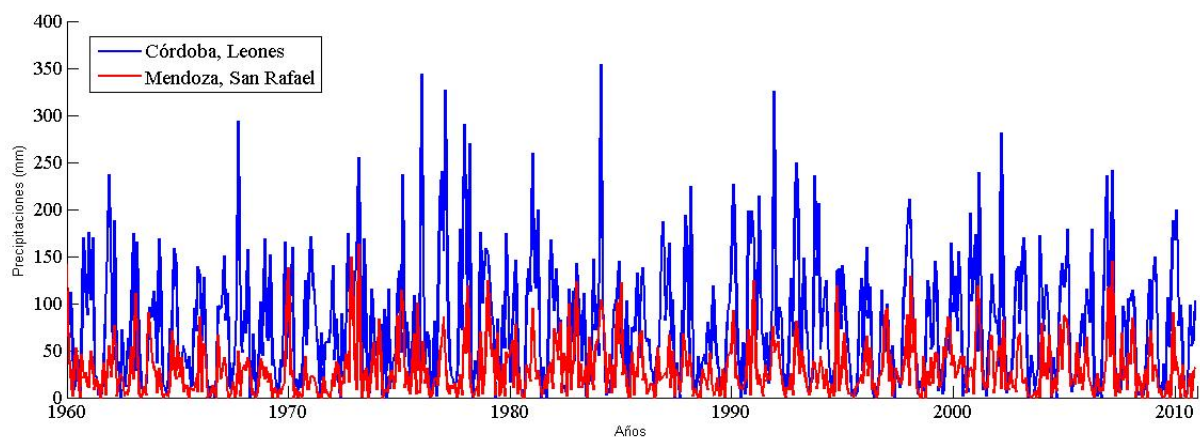


Figura 4: Series de precipitaciones, en milímetros, de Córdoba (Leones) y Mendoza (San Rafael) entre los años 1960 y 2010.

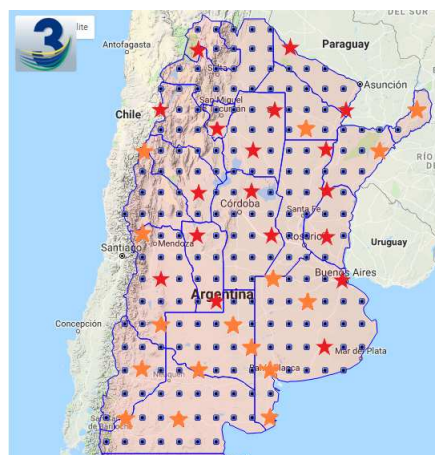


Figura 5: Distinción geográfica de las series persistentes (naranja) y antipersistentes (rojo).

superior respecto a las series cercanas. Análogamente, en el caso de la provincia de Mendoza, la serie de San Rafael es antipersistente y se encuentra a una altitud menor que las series cercanas persistentes. Luego, consideramos que para poder afirmar si el cambio de persistencia en una determinada región se debe a la ubicación geográfica, necesitaríamos focalizar más el estudio. Como se puede observar en la Fig. 5, se deberá realizar un estudio más localizado para poder concluir sobre la persistencia de las series en cada región. Incluso, analizaremos las series de localidades donde se han producido sequías o inundaciones en el período estudiado (1960-2010).

También, como hemos dicho en la Sección 2.1, hemos analizado las series de precipitaciones diarias. Los resultados no se reportan ya que en muchas regiones las series tienen muchos días en donde no hay precipitaciones, por ejemplo San Juan. De aquí que los coeficientes de la descomposición wavelet de la serie sean nulos y, por lo tanto, no se puede estimar la pendiente con la regresión lineal log-log de los detalles de los coeficientes vs. la escala. A fin de analizar las series diarias, y dado que utilizamos la wavelet de Haar para la descomposición de la serie con la mayor cantidad de niveles posibles, estamos estudiando usar otras wavelets, como así también menos niveles en dicha descomposición.

REFERENCIAS

- 3CN-CIMA/CONICET-UBA. Agradecemos al CIMA/CONICET-UBA por la provisión de los datos climáticos de la 3ra. Comunicación Nacional sobre Cambio Climático. <http://3cn.cima.fcen.uba.ar>, 2014.
- Arouxet M. y Pastor V. Estudio del exponente de Hurst. *Mecánica Computacional*, 35:2501–2508, 2017. ISSN 2591-3522.
- Hurst H. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–808, 1951.
- L. A., C. E., F. C., L. A., y L. M. Una revisión de los métodos para estimar el exponente de Hurst y la dimensión fractal en series de precipitación y temperatura. *Revista Mexicana de Física*, 63:244–267, 2017. ISSN 0035-001X.
- Malamud B. y Turcotte D. Self-affine time series: measures of weak and strong persistence. *Journal of Statistical Planning and Inference*, 80:173 – 196, 1999. ISSN 0378-3758. doi: [https://doi.org/10.1016/S0378-3758\(98\)00249-3](https://doi.org/10.1016/S0378-3758(98)00249-3).
- Maronna R.A., Martin R.D., y Yohai V.J. *Robust Statistics: Theory and Methods*. John Wiley & Sons, Ltd, 2006. ISBN 9780470010945. doi:10.1002/0470010940.
- Peng C., Havlin S., Stanley H.E., y Goldberger A.L. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87, 1995. doi:10.1063/1.166141.
- Pérez S., Sierra E., Massobrio M., y Momo F. Análisis fractal de la precipitación anual en el este de la provincia de La Pampa. 9:25–31, 2009.
- Rangarajan G. y Sant D.A. Fractal dimensional analysis of Indian climatic dynamics. *Chaos, Solitons & Fractals*, 19(2):285 – 291, 2004. ISSN 0960-0779. doi:[https://doi.org/10.1016/S0960-0779\(03\)00042-0](https://doi.org/10.1016/S0960-0779(03)00042-0). Fractals in Geophysics.
- Serinaldi F. Use and misuse of some Hurst parameter estimators applied to stationary and non-stationary financial time series. *Physica A: Statistical Mechanics and its Applications*, 389(14):2770 – 2781, 2010. ISSN 0378-4371. doi:<https://doi.org/10.1016/j.physa.2010.02.044>.
- Simonsen I., Hansen A., y Nes O. Determination of the Hurst exponent by use of wavelet transforms. *Physical Review E*, 58:2779, 1998.
- Valle M.A.V., García G.M., Cohen I.S., Klaudia Oleschko L., Ruiz Corral J.A., y Korvin G. Spatial variability of the Hurst exponent for the daily scale rainfall series in the state of Zacatecas, Mexico. *Journal of Applied Meteorology and Climatology*, 52(12):2771–2780, 2013. doi:10.1175/JAMC-D-13-0136.1.