

# Ajuste del caudal máximo anual con la función general de valores extremos en contexto no-estacionario

Carolina Lauro<sup>1</sup>, Alberto I J Vich<sup>1,2</sup>

**Resumen** La variabilidad climática y el cambio climático desafían los supuestos de estacionariedad, dando lugar al desarrollo de nuevos métodos para la estimación de los parámetros de las funciones de distribución. El objetivo es identificar la presencia de no-estacionariedad en una serie temporal de caudal máximo anual regional y modelar los extremos hidrológicos en un contexto no estacionario. Para ello, se emplearon diversas técnicas estadísticas de detección de tendencia y saltos abruptos. Se ajustó la función general de valores extremos mediante el método de máxima verosimilitud, considerando la anomalía de temperatura superficial del mar en la región 3.4 del océano Pacífico ecuatorial como co-variable. Se identificó una tendencia significativa y saltos en los años 1971 y 2007. El modelo no estacionario presentó un mejor ajuste de los datos observados, y los cuantiles estimados representan la variabilidad interanual de la serie. Estos resultados aportan información para la planificación de los recursos hídricos en un contexto de variabilidad climática.

**Palabras clave** Eventos hidrológicos extremos, Cuyo, modelo no estacionario, variabilidad climática.

## INTRODUCCIÓN

Los eventos extremos constituyen una de las principales líneas de investigación en hidrología; comprender las características de estos eventos, como su intensidad, duración y frecuencia, es fundamental para mejorar la gestión mediante la elaboración de planes de prevención y mitigación con información precisa.

El estudio de estos eventos se realiza por medio del análisis de series temporales de diversas variables hidroclimáticas. El análisis de frecuencia es un método estadístico que permite conocer la magnitud de un evento en términos de su probabilidad de ocurrencia. El método requiere que las series temporales sean homogéneas, independientes y estacionarias. Sin embargo, producto de la variabilidad climática y del cambio climático, estos supuestos no se cumplen en la mayoría de los casos (Milly et al., 2008). En consecuencia, se han desarrollado métodos capaces de incorporar la variabilidad climática mediante la incorporación de co-variables en la estimación de los parámetros de las funciones de distribución (Khaliq et al., 2006) que describen los eventos extremos y reducir la incertidumbre en la estimación de cuantiles.

Las variables hidroclimáticas de la región de Cuyo en Argentina presentan evidencias de tendencias y cambios abruptos, en ocasiones relacionados con fenómenos océano-atmosféricos de macro escala (Lauro et al., 2019, 2021). Se propone identificar la presencia de no-estacionariedad en una serie temporal de caudal máximo anual regional, identificar cambios en los parámetros de la función de distribución y estimar cuantiles en un contexto no-estacionario.

## ÁREA DE ESTUDIO

La zona de estudio comprende los ríos San Juan, Mendoza, Tunuyán, Diamante, Atuel, Grande y Barrancas, que integran el sistema hidrográfico del río Colorado. Los cauces se extienden desde las altas cumbres de la cordillera de Los Andes hacia el este hasta el colector del río Desaguadero; éste finalmente drena en el río Colorado. Pertenecen a un mismo sistema climático con precipitaciones

<sup>1</sup>Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, CCT-Mendoza, CONICET, Argentina. Email: [clauro@mendoza-conicet.gob.ar](mailto:clauro@mendoza-conicet.gob.ar)

<sup>2</sup>Facultad de Filosofía y Letras, Universidad Nacional de Cuyo, Argentina.

invernales provenientes de masas de aire cargadas de humedad del océano Pacífico, que precipitan en las altas cumbres de la cordillera durante el invierno. En consecuencia, el régimen de escurrimiento de estos ríos es unimodal, con máximos durante el período estival y mínimos durante el invierno. Además, presentan una alta variabilidad interanual asociada a fenómenos océano-atmosféricos de macro escala (Lauro et al., 2021). Los ríos de estas cuencas abastecen los principales oasis irrigados de la región de Cuyo.

## MATERIALES Y MÉTODOS

Se construyeron series de caudal máximo anual a partir de registros de caudal medio diario de siete sitios de aforo de los ríos mencionados anteriormente (<https://snih.hidricosargentina.gob.ar/Inicio.aspx>). Se consideró el año hidrológico de julio a junio. Los sitios analizados se encuentran aguas arriba de intervenciones antrópicas, y se emplearon registros mayores a 30 años. Dado que los caudales extremos son eventos excepcionales, es necesario contar con series hidrológicas largas para su análisis. Por ello, considerando la alta correlación entre los caudales de los siete puntos de aforo (0.74-0.94, p-valor 0.05), se construyó una serie regional para el período 1931-2020:

$$Q_{R,t} = \sum_{i=1}^{n_t} Q_{i,t}/n_t, \quad (1)$$

donde  $Q_{R,t}$  es el caudal regional promedio en el año  $t$ ,  $Q_{i,t}$  es el caudal  $i$  en el año  $t$  y  $n_t$  es el número de estaciones con datos en el año  $t$ .

La identificación de la no-estacionariedad en la serie se realizó mediante la detección de cambios graduales y abruptos. Para la detección de tendencias, se empleó el test de Mann Kendall (Hirsch, 1982; Westmacott & Burn, 1997), con correcciones por pre-blanqueo (Yue & Wang, 2002) y varianza (Hamed & Rao, 1998). Para la detección de saltos, se utilizó una prueba de saltos múltiples.

Luego, se ajustó la función general de valores extremos (GEV) para el período completo analizado y sub-períodos conforme los saltos encontrados. Los parámetros se estimaron por el método de máxima verosimilitud (Coles, 2001).

Para evaluar la existencia de cambios en la distribución entre el período completo y los sub-períodos se aplicó la prueba de Kolmogorov-Smirnov, ésta permite comparar si dos muestras provienen de la misma distribución. Las hipótesis son:  $H_0: F(x) = G(x)$  versus  $H_1: F(x) \neq G(x)$ , para al menos un  $x$ . La prueba es sensible a cualquier discrepancia entre las distribuciones (dispersión, posición, simetría, etc.). El estadístico se basa en la máxima diferencia entre las dos distribuciones.

En los modelos no estacionarios los parámetros de la distribución se asocian a co-variables. En este caso particular se empleó como co-variable las anomalías de la temperatura superficial del mar en la región niño 3.4 del océano Pacífico ecuatorial para los meses julio, agosto, septiembre (TSM 3.4-JAS; [https://psl.noaa.gov/gcos\\_wgsp/Timeseries/Nino34/](https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/)). La serie de caudal regional tiene un coeficiente de correlación de Spearman de 0.36 (p-valor < 0.001) con la serie TSM 3.4 JAS.

Para evaluar la significancia de la incorporación de co-variables al modelo ajustado inicialmente, se aplicó la prueba razón de verosimilitud. La prueba compara dos modelos anidados de modo que el modelo base, debe estar contenido dentro del modelo con más parámetros. La hipótesis nula  $D = 0$  se rechaza si  $D$  excede el  $D_{crítico}$  para un determinado nivel de significancia ( $\alpha = 0.05$ ) (Coles, 2001).

Alternativamente, para comparar la calidad relativa de un modelo se incluyen los criterios de información de Akaike (AIC) y Bayesiano (BIC). Ambos intentan contrarrestar el problema de la sobreparametrización incorporando una penalización basada en el número de parámetros. El BIC es más parsimonioso que el AIC. Entre los modelos candidatos, se prefiere el que tiene un AIC/BIC más bajo.

Todos los análisis se realizaron en el software R (<http://cran.r-project.org>), empleando los paquetes

trend (Pohlert, 2023), strucchange (Zeileis et al., 2002, 2003) y extRemes (Gilleland & Katz, 2016).

## RESULTADOS

A nivel regional, considerando la serie de caudal máximo anual del período 1931-2020, se encuentra una tendencia creciente en los últimos 90 años. En el año 1971 se registra un incremento de 107.6 % en la media de la serie temporal, en tanto en el año 2007 se registra una disminución del 46.9 % (Tabla 1).

Durante el período 1972-07 los caudales medios máximos se encuentran por encima de los valores medios de la serie completa, en tanto, durante los períodos 1931-71 y 2008-20 los caudales se encuentran por debajo del valor medio de la serie (Figura 1A). Los cambios en los valores medios evidencian la falta de estacionariedad en la serie temporal. En la Tabla 1 se presentan los parámetros de la función General de Valores Extremos ajustada para el período completo y en cada sub-período.

**Tabla 1.** Valores medios y parámetros de la función GEV para diversos períodos.

	1931-1971	1972-2007	2008-2020	1931-2020
Valor medio de $Q_{\max}$ (m <sup>3</sup> /s)	83.67	173.73	92.13	120.92
Desvío estándar	26.95	73.80	34.25	67.19
$\mu$	70.24	141.18	76.69	85.21
$\sigma$	18.34	58.81	24.92	35.23
$\kappa$	0.11	-0.02	0.04	0.36
Log verosimilitud	-182.15	-202.90	-62.64	-481.10
AIC	370.30	411.82	131.29	968.21

Los parámetros (dispersión, posición o simetría) de la función de distribución ajustada para los sub-períodos 1931-71 y 1971-07 difieren de los parámetros ajustados para el período completo (1931-20). Estas variaciones en los parámetros pueden indicar cambios en la magnitud, frecuencia y variabilidad del caudal máximo.

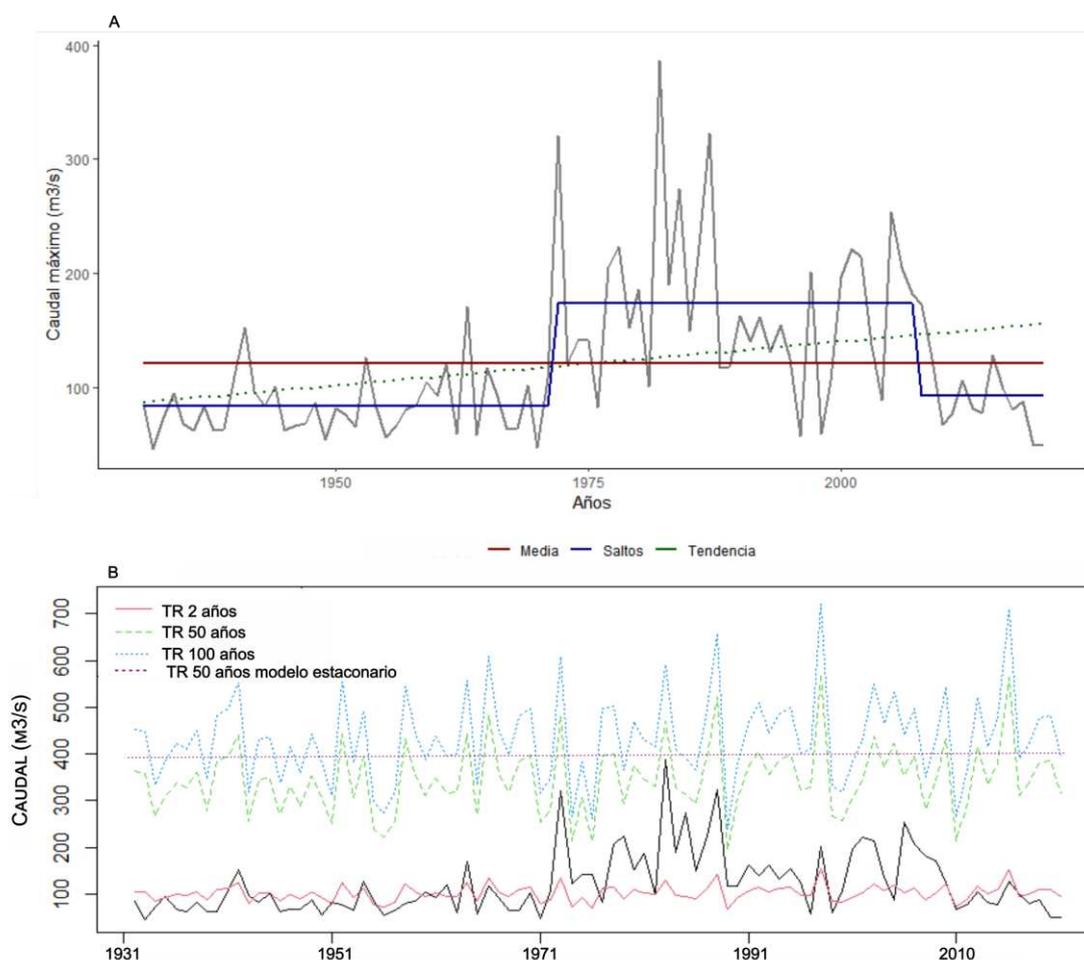
Al considerar el período completo de análisis los caudales máximos para diversos períodos de retornos, resultan superiores comparandolos con los sub-períodos 1931-71 y 2008-20 e inferiores en comparación con el sub-período 1972-07 con excepción en este último caso del caudal estimado para un TR de 50 y 100 años (Tabla 2).

**Tabla 2.** Caudales (m<sup>3</sup>/s) estimados para diversos períodos de retorno en cada sub-período.

TR (años)	1931-1971	1972-2007	2008-2020	1931-2020
2	77.1	162.6	85.9	99.0
5	100.2	227.5	115.2	155.5
20	134.9	308.8	155.3	273.8
50	160.1	358.6	181.9	389.1
100	180.8	395.5	202.5	505.1

Al identificar la presencia de tendencias y saltos en la serie temporal completa, se aplicaron diversos modelos no estacionarios empleando como co-variable explicativa de los parámetros de distribución GEV la anomalía de la temperatura superficial del mar del océano Pacífico ecuatorial. En la Tabla 3 se detallan los resultados de los criterios de información y la prueba de razón de verosimilitud.

Se encuentra que modelar el parámetro de localización en dependencia lineal con el índice TSM 3.4-JAS mejora el ajuste del modelo (modelo 2). Obteniendo un mejor ajuste cuando, además del parámetro de localización, se modela el parámetro de escala en dependencia lineal con el índice climático (modelo 3). En tanto, modelar el parámetro de forma (modelo 4) mejora el ajuste respecto al modelo estacionario (modelo1), pero no mejora respecto del modelo 3.



**Figura 1.** A) Serie temporal del caudal máximo anual regional, tendencia y valor medio en los sub-períodos comprendidos entre los saltos detectados en 1971 y 2007 y valor medio de la serie completa. B) Estimación del caudal para los períodos de retorno de 2, 50 y 100 años con el modelo no estacionario y de 50 años con el modelo estacionario.

En la Figura 1B, se representó el caudal estimado para diversos períodos de retorno con el modelo 3 y el caudal para un período de retorno de 50 años con el modelo 1. El modelo no estacionario es capaz de incorporar a los caudales estimados la variabilidad interanual observada en los datos. Ambos modelos cubren el máximo valor observado ( $387 \text{ m}^3/\text{s}$  en 1982) con los valores estimados para un TR de 50 años.

## CONCLUSIONES

El valor medio de la serie de caudal máximo anual regional, representativa de siete ríos de la región de Cuyo para el período 1931-2020, presenta cambios graduales y abruptos, con dos períodos de caudales por debajo de la media (1931-71 y 2008-20) y un período caracterizado por caudales por encima de la media (1971-07). Durante el período 1971-07, las anomalías de la TSM 3.4 JAS son mayores a 1, evidenciando que el calentamiento del océano Pacífico tropical da lugar al incremento del caudal máximo anual regional.

La presencia de saltos y tendencias interfiere en la estimación de cuantiles. La omisión de la presencia de no-estacionariedad puede ocasionar errores de sobre o sub-estimación en la extrapolación de caudales extremos cuando se producen cambios en los parámetros de la distribución del período de registro.

**Tabla 3.** Métricas de ajuste de los modelos no estacionarios analizados.

Modelo	AIC	BIC	$D$ ( $D_{\text{crítico}}$ )	P-valor	$H_0: D = 0$
1 GEV ( $X; \mu, \sigma, \kappa$ )	968.2	975.7	-	-	-
2 GEV ( $X; \mu = a_0 + a_1 \text{TSM 3.4-JAS}, \sigma, \kappa$ )	962.6	972.6	7.57 (3.84)	0.006	Rechazo
3 GEV ( $X; \mu = a_0 + a_1 \text{TSM 3.4-JAS}, \sigma = b_0 + b_1 \text{TSM 3.4 - JAS}, \kappa$ )	959.1	971.6	13.10 (5.99)	0.001	Rechazo
4 GEV ( $X; \mu = a_0 + a_1 \text{TSM 3.4-JAS}, \sigma = b_0 + b_1 \text{TSM 3.4-JAS}, \kappa = c_0 + c_1 \text{TSM 3.4-JAS}$ )	961.0	976.0	13.19 (7.81)	0.004	Rechazo

El hecho de que los parámetros de la función de distribución cambien con el tiempo indica que el caudal máximo anual está sujeto a variabilidad temporal y pueden estar influenciados por diversos factores, siendo uno de ellos los forzantes climáticos globales. El análisis del caudal máximo en la región de Cuyo muestra que los modelos no-estacionarios tienen un mejor ajuste de los datos respecto del modelo estacionario cuando los parámetros de localización y escala se modelan como dependencia lineal con el índice TSM 3.4-JAS. Siendo estos modelos de interés para la extrapolación de eventos extremos.

## REFERENCIAS

- Coles, S.** (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag. <https://doi.org/10.1007/978-1-4471-3675-0>
- Gilleland, E. & Katz, R. W.** (2016). extRemes 2.0: An extreme value analysis package in R. *Journal of Statistical Software*, 72(8), 1–39. <https://doi.org/10.18637/jss.v072.i08>
- Hamed, K. & Rao, A.** (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1-4), 182–196.
- Hirsch, R.** (1982). A comparison of four streamflow record extension techniques. *Water Resources Research*, 18(4), 1081–1088.
- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B.** (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329, 534–552.
- Lauro, C., Vich, A., & Moreiras, S.** (2019). Streamflow variability and its relationship with climate indices in western rivers of Argentina. *Hydrological Sciences Journal*, 57(1). <https://doi.org/10.1080/02626667.2019.1594820>
- Lauro, C., Vich, A. I., Otta, S., Moreiras, S., Vaccarino, E., & Bastidas, L.** (2021). Recursos hídricos superficiales de la vertiente oriental de los Andes Centrales (28°-37°s) en contexto de variabilidad hidroclimática. *Boletín De Estudios Geográficos*, 116, 45–71.
- Milly, P., Betancourt, J., Falkenmark, M., Hirsch, R., Kundzewicz, Z., Lettenmaier, D., & Stouffer, R.** (2008). Stationarity is dead: whither water management. *Science*, 319(5863), 573–574.
- Pohlert, T.** (2023). *Trend: Non-parametric trend tests and change-point detection*. R package version 1.1.5. <https://CRAN.R-project.org/package=trend>
- Westmacott, J. & Burn, D.** (1997). Climate change effects on the hydrologic regime within the Churchill Nelson River Basin. *Journal of Hydrology*, 202, 263–279.
- Yue, S. & Wang, C.** (2002). The influence of serial correlation on the mann-whitney test for detecting a shift in median. *Advances in Water Resources*, 25, 325–333.
- Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K.** (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44(1-2), 109–123. [https://doi.org/10.1016/S0167-9473\(03\)00030-6](https://doi.org/10.1016/S0167-9473(03)00030-6)
- Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C.** (2002). strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software*, 7(2), 1–38. <https://doi.org/10.18637/jss.v007.i02>