

METODOLOGÍAS DE ANÁLISIS DE FRECUENCIA EN VARIABLES HIDROLÓGICAS Y EL FENÓMENO DEL NIÑO

María Josefina Tito⁽¹⁾, Eduardo Zamanillo^(1,2) y Eduardo Díaz⁽³⁾

⁽¹⁾Facultad de Ciencias de la Administración, UNER. Avda. Monseñor Tavella 1424, Concordia, Entre Ríos.

⁽²⁾Facultad Regional Concordia, UTN. Salta 277, Concordia, Entre Ríos.

⁽³⁾Facultad de Ciencias Agropecuarias, UNER. Ruta Pcial. 11 Km 10.5, Oro Verde, Entre Ríos.
e-mail: ezamanillo@gmail.com

RESUMEN

Se analiza la relación de dependencia de la frecuencia de ocurrencia de eventos extremos de variables hidrológicas con el fenómeno de El Niño y la variabilidad climática. Se utilizaron niveles máximos diarios del río Paraná, caudales máximos diarios del río Uruguay y precipitaciones máximas en estaciones pluviométricas localizadas en la provincia de Entre Ríos. Se consideró el enfoque clásico bajo hipótesis de estacionariedad, se evaluó la posibilidad de que los valores máximos sean generados por distribuciones formadas por dos o más poblaciones, mediante la aplicación de una distribución de probabilidad mixta para modelar la probabilidad de excedencia considerando la presencia del fenómeno de El Niño para caracterizar la pertenencia poblacional. Se consideró el fenómeno de la no-estacionariedad de la(s) serie(s) a partir del supuesto de que el comportamiento extremo de la serie depende de una variable exógena. Se consideran tres casos para la distribución GEV: (i) el parámetro de posición depende linealmente de una covariable, (ii) el parámetro de posición es una función cuadrática de una covariable y (iii) el parámetro de posición y de escala dependen de una covariable. Se discute la extensión del concepto de período de retorno al ajuste no estacionario.

Palabras clave: análisis de frecuencia, distribuciones mixtas de probabilidad, series no estacionarias.

ABSTRACT

The dependence of the frequency of occurrence of extreme events of hydrological variables with El Niño and climate variability is analyzed. Daily maximum levels of the Parana river, maximum daily flows of the Uruguay river and maximum daily rainfall at several gages located in Entre Rios were used. The classical approach under assumptions of stationarity was considered. Also the possibility that the maximum values are generated by distributions of two or more populations. This was done applying a distribution of mixed probability to model the probability of exceedance considering the presence of ENSO phenomenon to characterize the population belonging. The phenomenon of non-stationarity on the data series was considered from the assumption that the extreme behavior of the series depends on an exogenous variable. Three cases for the GEV distribution are considered: (i) the position parameter depends linearly on a covariate, (ii) the position parameter is a quadratic function of a covariate and (iii) the position and scale parameters depends on a covariate. The extension of return period concept to nonstationary process is discussed.

Keywords: frequency analysis, joint probability distributions function, non-stationary series.

INTRODUCCIÓN

Para describir el comportamiento estadístico de variables hidrológicas se utilizan diferentes funciones de distribución de probabilidades, fundamentalmente en la estimación de la frecuencia de ocurrencia de eventos extremos.

En el enfoque clásico, se asume la estacionariedad del fenómeno y la recurrencia de los máximos se estiman a partir de una función de distribución de probabilidad que se selecciona entre varias candidatas a partir de la aplicación de distintos criterios de bondad de ajuste.

Hay situaciones donde los valores extremos son generados por distribuciones formadas por dos o más poblaciones, considerando en estos casos la aplicación de una distribución de probabilidad mixta para modelar la probabilidad de excedencia. Cada población queda caracterizada por alguna(s) diferencia(s) en la génesis del fenómeno en cuestión y el modelo probabilístico pasa a ser mixto, con los consecuentes problemas asociados a la estimación de parámetros y selección del modelo.

A las fuentes de incertidumbre precipitadas, se le adiciona la que introduce la no-estacionariedad de la(s) serie(s) debido a diferentes causas, como la acción antrópica o los efectos del cambio climático a largo plazo. En estos casos, las hipótesis que se realizan en el análisis clásico sufren desviaciones y el modelo estándar pierde utilidad y representatividad.

No existe una teoría general establecida para los procesos no estacionarios (Coles, 2001). Es habitual en estos casos dar una solución práctica al problema, tomando los esquemas clásicos desarrollados para valores extremos con modificaciones. Así, por ejemplo, las variaciones a través del tiempo en un proceso observado se pueden modelar como una tendencia lineal en el parámetro de posición del modelo de valor extremo apropiado.

Este trabajo aborda el fenómeno de la no-estacionariedad de la(s) serie(s) a partir del supuesto de que el comportamiento extremo de la serie está relacionado con el de otra variable exógena. Es decir, se puede modelar a través de una distribución clásica de extremos, expresando uno o más parámetros de la distribución como función de una covariable (El Adlouni et al., 2007).

Dado que la utilidad práctica usual de los resultados del análisis de frecuencia es la estimación de probabilidades de excedencias futuras, y con el propósito de comparar y evaluar diferentes metodologías, se aplicaron distintas técnicas que contemplan los enfoques arriba enunciados a diferentes casos de estudio de variables hidrológicas.

Se utilizaron los niveles máximos del río Paraná, los caudales máximos del río Uruguay y precipitaciones máximas diarias en estaciones pluviométricas localizadas en la Provincia de Entre Ríos.

Los niveles máximos del río Paraná corresponden a la serie anual de los niveles máximos diarios medidos en la estación hidrométrica Rosario, y los caudales corresponden a la serie anual de máximos medios diarios de aporte del río Uruguay al embalse de Salto Grande. En ambos casos los máximos diarios para cada año del registro histórico conforman el conjunto de valores extremos que se analiza estadísticamente.

Las precipitaciones máximas en las estaciones pluviométricas fueron utilizadas con el objeto de analizar la dependencia de la frecuencia de ocurrencia de los valores máximos diarios, del cambio climático y del fenómeno de El Niño. Las mismas fueron divididas en tres subregiones. La primera subregión está vinculada al tramo final de la cuenca del Río Paraná situada al sur de la ciudad de Diamante (ER) hasta la desembocadura del río Gualaguay, con una longitud aproximada a los 300 km. Una segunda subregión está localizada al este de la provincia, sobre el río Uruguay. La tercera subregión está localizada en el noroeste de la provincia sobre la cuenca del Arroyo Feliciano.

METODOLOGÍA Y RESULTADOS

Selección del período de análisis de las series

Para la serie de niveles máximos diarios del río Paraná medidos en la estación Rosario, la longitud inicial de la serie considerada es de 128 años (1884 – 2011).

El período de análisis que se tomó para la serie de caudales de aportes máximos diarios medidos en la represa de Salto Grande es de 43 años (1970 – 2012).

Para la información pluviométrica diaria, la Tabla 1 presenta las estaciones y los años de registro de precipitaciones diarias en cada una, y la Figura 1 muestra la localización de las mismas.

Tabla 1: Estaciones utilizadas y longitudes de registro.

Estación	Registros		Años registrados
	Desde	Hasta	
Paraná	1961	2013	52
Isletas	1958	2013	54
Crucesitas III	1956	2013	55
Oliveros	1962	2013	52
Rosario	1936	2012	77
Zavalla	1973	2013	41
Pergamino	1961	2013	53
San Pedro	1967	2013	47
Concordia	1969	2012	43
Concepción del Uruguay	1968	2013	43
Banderas	1958	2013	54
Pedro Vallejos	1958	2013	54
Colonia Avigdor	1958	2013	52
San Gustavo	1958	2013	56

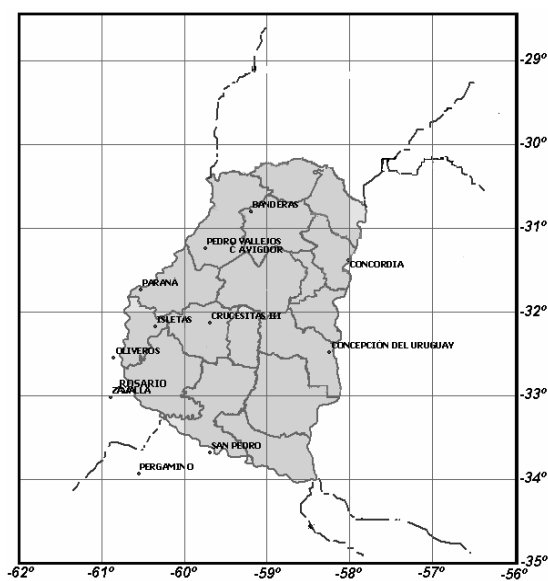


Figura 1. Localización de las estaciones seleccionadas.

Series de niveles máximos

Para definir el período de análisis de los niveles máximos medidos en la estación Rosario, se realizaron pruebas paramétricas para analizar la hipótesis de que la media y la varianza son estacionarias. Se exploró una división de las series en el año 1970, donde otros estudios realizados con anterioridad en la Cuenca del Plata (García y Vargas, 1998; Jaime y Menéndez, 2002) han detectado no-homogeneidades en los caudales del río Paraná.

Los resultados del test de no homogeneidad de la varianza, muestran que la serie es homogénea en cuanto a su varianza. Y los resultados del test paramétrico de no homogeneidad en la media (García, 2004) muestra que la serie no es homogénea en cuanto a su media. Ambos tests se hicieron con un nivel de significación del 5%. Por otra parte, se utilizó como ensayo no-paramétrico para verificar si dos muestras independientes fueron extraídas de la misma población, la prueba de suma de rangos de Mann Whitney (Fernández Loureiro de Pérez, 2011). Las hipótesis de contrastes son: las muestras provienen de la misma población versus las muestras provienen de poblaciones distintas. Se realizó esta prueba, verificándose la pertenencia a poblaciones distintas al considerar como momento de cambio el año 1970.

Para considerar la señal del fenómeno El Niño (ENSO) como variable que contribuye a estimar la recurrencia de los eventos, se tuvo en cuenta que a partir de 1950 se cuenta con una metodología consistente de medición del ENSO, por lo que el test no-paramétrico de Mann Whitney se repitió considerando la serie desde el año 1950 hasta 2011. Para esta serie (1950–2011) se adoptó, tal como en los casos anteriores, el punto de cambio en el año 1970 y se verificó que efectivamente la serie es no-homogénea.

Serie de precipitaciones máximas diarias

Se realizaron pruebas no-paramétricas para analizar la hipótesis de que la media de los máximos diarios no es estacionaria en las series de precipitaciones máximas diarias. Se exploró una división de las series en el año 1970 con el mismo criterio enunciado anteriormente para los niveles máximos (García y Vargas, 1998; Jaime y Menéndez, 2002; Zamanillo et al., 2013).

También se exploró una división de las series en el año 1980 para analizar la sensibilidad a la década de quiebre.

Se utilizó la prueba de Wilcoxon (Fernández Loureiro de Pérez, 2011) como test no paramétrico para verificar si las dos sub-muestras son independientes o fueron extraídas de la misma población. Los resultados obtenidos muestran que en general las medias de las sub-series posteriores a 1970 y 1980 (Media2) son mayores a los de las sub-series anteriores a dichos años. Sin embargo, las diferencias son estadísticamente significativas al 5% sólo en 3 de los casos.

Algo similar ocurre con respecto a la existencia de no homogeneidades en la varianza con anterioridad y posterioridad a 1970 y 1980. De hecho, las diferencias son estadísticamente significativas al 5% sólo en 5 de los casos.

Los resultados, descriptos anteriormente, muestran que no se puede considerar que las precipitaciones máximas diarias provengan de poblaciones distintas con anterioridad y posterioridad a 1970 (1980). Por lo tanto el cambio climático solo presenta evidencia aislada en cuanto a la generación de eventos máximos diarios que provengan de poblaciones diferentes en función a cuál sub-serie pertenecen.

Análisis de frecuencia clásico y distribuciones mixtas

Para estimar la frecuencia de los valores máximos registrados para las variables consideradas, se consideró en primera instancia el enfoque clásico, donde los niveles máximos se estiman a partir de una función de distribución de frecuencias. Las distribuciones de frecuencia candidatas que se consideraron para el análisis clásico son (a) General de Valores Extremos (GEV); (b) Gumbel; y (c) Log-Normal de dos parámetros (LN2).

En todos los casos los parámetros de las distribuciones se estimaron por el método de L-momentos propuesto por Hosking (1990). Las ventajas que presentan los L-momentos, respecto de los momentos convencionales, son que pueden caracterizar un número mayor de distribuciones y los estimadores de los coeficientes adimensionales son siempre insesgados aún para muestras pequeñas. Los L-momentos son más robustos frente a la presencia de valores atípicos en los datos.

Por otra parte, se analizó también la posibilidad de que los valores máximos sean generados por distribuciones formadas por dos o más poblaciones, considerando la aplicación de una distribución de probabilidad mixta para modelar la probabilidad de excedencia considerando la presencia del fenómeno de El Niño.

Se utilizó para la clasificación de un año en Niño y No Niño, el ONI (Oceanic Niño Index) publicado por la NOAA.

La distribución mixta es el resultado de la sumatoria de la función de distribución de probabilidad de cada muestra, afectada por un factor de ponderación que depende del tamaño de cada sub-muestra. Este modelo puede expresarse mediante la ecuación (1).

$$F_T(x \geq t) = \rho_1 F_1(x \geq t) + \rho_2 F_2(x \geq t) + \dots + \rho_n F_n(x \geq t) \quad (1)$$

Donde, $\sum_{i=1}^n \rho_i = 1$, F_T representa la función de distribución mixta, F_i es la función de distribución de probabilidad de cada sub-muestra y ρ_i representa el factor de ponderación de cada sub-muestra.

Se consideró como factor de influencia el fenómeno de El Niño, por lo que la serie de máximos se dividió en dos sub-series: una afectada por el fenómeno El Niño y otra no afectada por dicho fenómeno.

Para ello se utilizó la clasificación de un año en Niño y No Niño según la tabla de episodios fríos, neutrales y cálidos publicados por la NOAA. La NOAA (National Oceanic and Atmospheric Administration) clasifica en episodios cálidos y episodios fríos si el ONI (Oceanic Niño Index) alcanza un umbral de $\pm 0.5^\circ\text{C}$, en la región 3-4 (5°N - 5°S , 120° - 170°W), calculados a partir de períodos de base de 30 años centrados y actualizados cada 5 años.

En las series de niveles y de caudales máximos diarios anuales observados, los años donde el máximo se produjo en un período de episodio cálido, se clasificaron como años Niño, en caso contrario como años No-Niño.

La Tabla 2 muestra cómo se clasificaron, siguiendo este criterio, los años correspondientes a cada máximo anual observado.

Tabla 2: Años El Niño y No-Niño.

Rosario			Salto Grande		
No-Niño		Niño	No-Niño		Niño
1971	1991	1972	1970	1991	1972
1974	1993	1973	1971	1993	1973
1975	1994	1977	1974	1994	1976
1976	1996	1978	1975	1995	1982
1979	1997	1982	1977	1996	1983
1980	1999	1983	1978	1999	1987
1981	2000	1987	1979	2000	1992
1984	2001	1992	1980	2001	1997
1985	2005	1995	1981	2003	1998
1986	2006	1998	1984	2005	2002
1988	2008	2002	1985	2007	2004
1989	2009	2003	1986	2008	2006
1990	2011	2004	1988	2011	2009
		2007	1989	2012	2010
		2010	1990		

Con el objeto de verificar que efectivamente las dos sub-muestras corresponden a poblaciones diferentes, se efectuó a ambas series (Niño, No-Niño) el test no-paramétrico de Kolmogorov Smirnov (K-S).

La Tabla 3 muestra los resultados del test de K-S, para los niveles máximos de Rosario, permitiendo verificar con una significación de al menos el 5% que la hipótesis es correcta, por lo que se consideró

como alternativa válida para modelar la frecuencia el uso de distribuciones mixtas.

La Tabla 4 muestra los resultados del test de K-S para los caudales máximos medidos en la represa de Salto Grande, permitiendo verificar con una significación de al menos el 5% que la hipótesis es correcta, por lo que se consideró como una alternativa válida usar el modelo mixto en este caso.

Tabla 3: Resultados del test de Kolmogorov-Smirnov (estación Rosario).

Grupo 1	Grupo 2	n 1	n 2	Media 1 (cm)	Media 2 (cm)	DE 1 (cm)	DE 2 (cm)	p-valor
Niño	No-Niño	15	26	512	437	79	59	<0.02

Tabla 4: Resultados del test de Kolmogorov-Smirnov (estación Salto Grande).

Grupo 1	Grupo 2	n 1	n 2	Media 1 (m ³ /s)	Media 2 (m ³ /s)	DE 1 (m ³ /s)	DE 2 (m ³ /s)	p-valor
Niño	No-Niño	14	29	23844	17693	8709	5206	<0.02

El modelo general de función de distribución mixta considerado para estas dos poblaciones se presenta en (2):

$$F_T(x \geq t) = \rho_1 F_1(x \geq t) + \rho_2 F_2(x \geq t) \quad (2)$$

Donde F_1 y F_2 son distribuciones GEV o F_1 y F_2 son distribuciones Log-Normal de dos parámetros. Y ρ_1

es la frecuencia relativa de ocurrencia del fenómeno Niño en las series 1971-2011.

Para las distribuciones mixtas (2), los parámetros se estimaron, en cada estación, en forma individual a partir de cada una de las sub-series por el método de L momentos. Los resultados gráficos del análisis de frecuencia para la estación Rosario se muestran en la Figura 2.

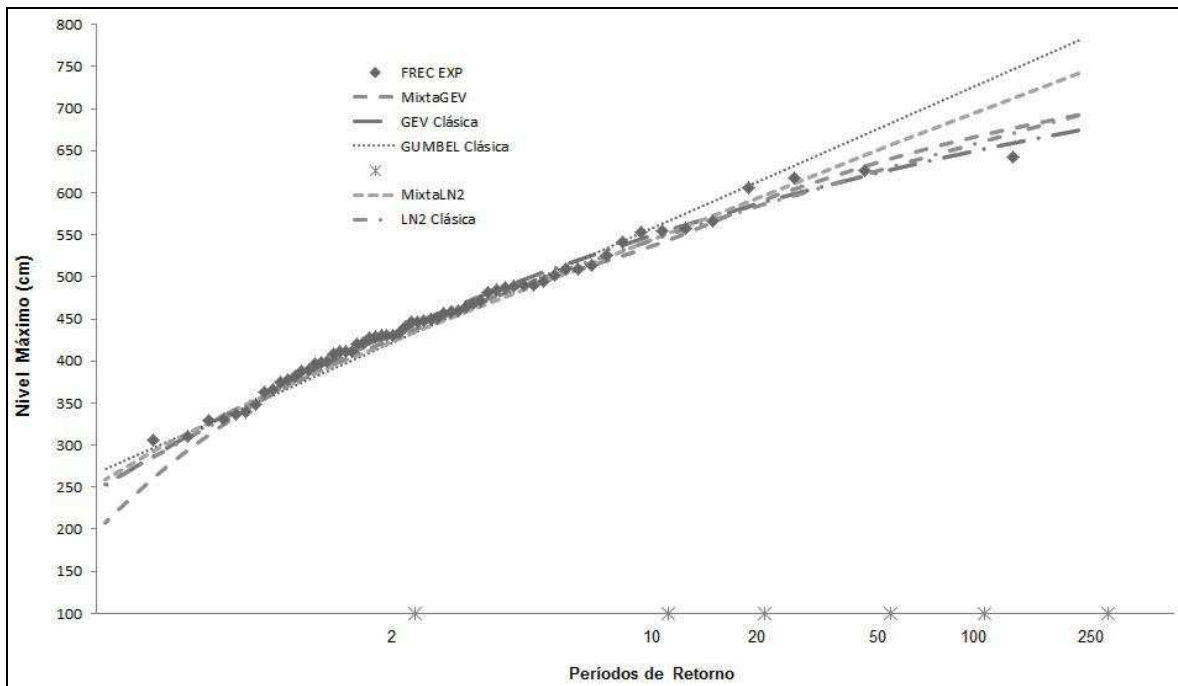


Figura 2. Análisis de frecuencia, estación Rosario.

Como indicadores para medir la bondad del ajuste se consideraron el Error Cuadrático Medio de la Variable, el Error Cuadrático Medio de la Frecuencia y el Criterio de Información de Akaike.

En la tabla 5 se muestra la distribución de mejor ajuste, para la serie de niveles máximos medidos en Rosario, de acuerdo a cada criterio cuantitativo de bondad de ajuste considerado. En la tabla se puede apreciar que hay predominio de la distribución

Mixta GEV en comparación con GEV, LN2, GUMBEL, Mixta LN2.

En la tabla 6 se muestra la distribución de mejor ajuste, para la serie de caudales de aportes máximos medidos en Salto Grande, de acuerdo a cada criterio cuantitativo de bondad de ajuste considerado. Se puede apreciar que hay predominio de la distribución Mixta GEV en comparación con GEV, LN2, GUMBEL, MixtaLN2.

Tabla 5: Valores obtenidos para ECM de la variable, ECM de la frecuencia y AIC(p).

	Criterio	GEV	Mixta GEV	LN2	Mixta LN2	Gumbel
Rosario	ECMF	0.0325	0.0259	0.0309	0.0373	0.0493
	ECMV (cm)	12.79	11.69	12.96	14.86	24.22
	AIC	215	214	214	229	265

Tabla 6: Valores obtenidos para ECM de la variable, ECM de la frecuencia y AIC(p).

	Criterio	GEV	Mixta GEV	LN2	Mixta LN2	Gumbel
Salto Grande	ECMF	0.0420	0.0380	0.0392	0.0597	0.0382
	ECMV(m³/seg)	1420	971	1184	2137	1226
	AIC	601	576	584	637	587

Análisis de las precipitaciones máximas diarias

En las series de precipitaciones máximas anuales observadas, los años donde el máximo se produjo en un período asociado a un episodio cálido, se clasificaron como años Niño, en caso contrario como años No-Niño. La Tabla 7 muestra cómo se clasificaron, siguiendo este criterio, los años correspondientes.

Tabla 7: Años El Niño posteriores a 1950, clasificados según intensidad de la señal.

Débil	Moderado	Fuerte
1952	1951	1957
1953	1963	1965
1958	1968	1972
1969	1986	1982
1976	1987	1997
1977	1991	
2004	1994	
2006	2002	
	2009	

Sin embargo, al intentar clasificar los máximos diarios anuales de precipitación utilizando solo el valor del ONI en su mes de ocurrencia surgen una serie de dificultades que distorsionan los resultados.

Para solucionar este problema, se utilizaron los lineamientos establecidos por Pisciotto et al. (1994) para determinar qué meses del año se pueden considerar que están influidos por la señal El Niño en un año Niño. Pisciotto et al. (1994), comprobaron, utilizando información anterior a 1986, que en la región correspondiente al sudeste de Sudamérica

(SSA) existe una dependencia de la magnitud de las precipitaciones mensuales con la magnitud de la señal de El Niño. Basándose en estos resultados y utilizando información pluviométrica mensual hasta 2010, se comprobó que existe una dependencia entre los valores de precipitaciones mensuales medidos en la región de estudio y el fenómeno El Niño. El criterio establecido es que, dado un año Niño, la influencia de la señal afecta los meses de septiembre, octubre, noviembre y diciembre del año declarado Niño y los meses de enero a julio del año siguiente con excepción de febrero.

Siguiendo estos criterios se definieron dos sub-series para cada estación pluviométrica y se utilizó la prueba de Wilcoxon como test no-paramétrico para verificar si las dos sub-muestras son independientes o fueron extraídas de la misma población.

La Tabla 8 presenta los resultados del test no-paramétrico para la división de las series en Niño y no-Niño. Los resultados obtenidos muestran que en general las medias de las precipitaciones máximas diarias de la sub-serie correspondiente a los años Niño (Media1) son mayores a los de la sub-serie correspondiente a los años no-Niño. Sin embargo, las diferencias son estadísticamente significativas al 5% sólo en 2 de los casos considerados.

Los resultados de los tests paramétricos de igualdad de medias y varianzas para la división de las series en Niño y no-Niño, muestran que, en general, las diferencias de medias no son estadísticamente significativas al 5% (sólo lo son en 2 de los casos). Lo mismo ocurre con la diferencia de varianzas (sólo son estadísticamente significativas al 5% en 3 de los casos).

Tabla 8: Resultados del test de Wilcoxon, años Niño – No Niño.

Estación	n 1	n 2	Media 1	Media 2	DE 1	DE 2	W	p(2 colas)
Paraná	18	34	100.5	102.1	27.7	42.3	491	0.795
Isletas	16	39	100.7	93.7	37.8	33.2	475	0.617
Crucesitas III	14	27	98.1	98.1	30.5	30.6	292	0.945
Oliveros	15	37	102.1	90.6	21.3	25.3	473	0.127
Rosario	17	46	85.5	94.8	28.3	34.4	472	0.265
Zavalla	16	25	80.9	92.6	17.6	31.3	296	0.279
Pergamino	15	38	96.1	82.0	62.9	37.0	452	0.353
San Pedro	15	32	100.4	92.5	33.5	31.0	392	0.465
Concordia	17	25	119.3	104.1	37.9	33.8	416	0.191
Banderas	12	30	132.8	119.8	55.3	41.9	272	0.697
Pedro Vallejos	19	35	132.8	96.2	65.6	32.3	634	0.044
Cnia. Avigdor	17	37	113.4	91.7	32.0	29.3	605	0.010
San Gustavo	17	27	126.4	111.4	44.2	38.4	428	0.273

Por lo que no se puede considerar que las precipitaciones máximas diarias de las estaciones consideradas provengan de poblaciones distintas en años Niño que en años No-Niño.

Si bien se comprobó que en los años Niño existe una tendencia a que los totales pluviométricos mensuales sean mayores que en los años no-Niño, pero al mismo tiempo se verificó que no hay una tendencia definida de que las precipitaciones máximas diarias sean consistentemente mayores, por lo que se considera necesario estudiar la frecuencia de eventos extremos.

En este sentido, se estudió la frecuencia de eventos por encima del evento anual ($Tr = 1.01$ años) en años Niño y No-Niño, dividiendo el total de estaciones en dos subgrupos. En ambas se verificó un aumento del promedio anual de eventos extremos en los años Niño, pero sólo en una de ellas se verificó que regionalmente este incremento es estadísticamente significativo.

Los resultados del test no-paramétrico de Wilcoxon de comparación del número medio de cantidad de eventos que superan la lluvia correspondiente al evento $Tr = 1.01$ en años Niño y No-Niño, a un Nivel de significación del 5%, muestran que en la región del Delta entrerriano no se verifican diferencias estadísticamente significativas en el número medio de eventos por año de recurrencia $Tr = 1.01$ o superiores en años Niño. En la región centro norte hay una tendencia consistente a que estas diferencias sean significativas.

Los resultados del test no-paramétrico de Wilcoxon de comparación del número medio de cantidad de eventos que superan la lluvia correspondiente al evento $Tr = 1.01$, antes y después de 1980, a un Nivel de significación del 5%, muestran que en la región del Delta entrerriano no se verifican diferencias estadísticamente significativas en el número medio de eventos por año de recurrencia $Tr = 1.01$ o superiores en las sub-series anteriores a 1980 con relación a las posteriores a dicho año. En la región norte hay una tendencia consistente a que estas diferencias sean significativas. Por lo que puede concluirse que, con la evidencia analizada, sólo en el Norte entrerriano hay un aumento consistente de la frecuencia de eventos extremos con posterioridad a 1980.

Análisis de modelos no estacionarios

En general los métodos clásicos de análisis de frecuencia de eventos extremos utilizados en hidrología se basan en la suposición de estacionariedad (es decir, un clima que no cambia en sentido estadístico).

En este trabajo se aborda el fenómeno de la no-estacionariedad de la(s) serie(s) a partir del supuesto de que el comportamiento extremo de la serie está relacionado con el de otra variable exógena. Es decir, se puede modelar a través de una distribución clásica de extremos, expresando uno o más parámetros de la distribución como función de una covariable (El Adlouni et al., 2007).

Se consideró que el máximo diario del nivel del río Paraná o el caudal máximo diario del río Uruguay, se pueden modelar a través de una distribución GEV con parámetros que son funciones de una covariable que denota la influencia del fenómeno de El Niño en cada año.

Se consideran tres casos para la distribución GEV: el parámetro de posición depende linealmente de una covariable, el parámetro de posición es una función cuadrática de una covariable y el parámetro de posición y de escala dependen de una covariable.

En primer lugar el modelo denominado GEV1, donde el máximo sigue una distribución $GEV1(\mu(t), \sigma, \xi)$ donde $\mu(t) = \mu_0 + \mu_1 \cdot SOI(t)$ y $SOI(t)$ indica la influencia del fenómeno El Niño en el año t .

Un segundo modelo denominado GEV2 donde el máximo sigue una distribución $GEV2(\mu(t), \sigma(t), \xi)$ donde $\mu(t) = \mu_0 + \mu_1 \cdot SOI(t)$ y $\sigma(t) = \exp(\sigma_0 + \sigma_1 \cdot SOI(t))$ y $SOI(t)$ indica la influencia del fenómeno El Niño en el año t .

Y el modelo denominado GEV3 donde el máximo sigue una distribución $GEV3(\mu(t), \sigma, \xi)$ donde: $\mu(t) = \mu_0 + \mu_1 \cdot SOI(t) + \mu_2 \cdot SOI(t)^2$ y $SOI(t)$ indica la influencia del fenómeno El Niño en el año t .

El test de razón de verosimilitud para modelos anidados proporciona un procedimiento que permite seleccionar un modelo sobre otro. Dados dos modelos $M_0 \subset M_1$, la deviance estadística se define como (3):

$$D = 2\{\ell_1(M_1) - \ell_0(M_0)\} \quad (3)$$

donde, $\ell_1(M_1)$ y $\ell_0(M_0)$ son los máximos obtenidos para el logaritmo de la función de verosimilitud para los modelos M_0 y M_1 , respectivamente. Valores grandes para D indican que el modelo M_1 explica mejor la variación de los datos que M_0 , valores pequeños de D sugieren que un aumento en los parámetros del modelo no trae mejoras en la capacidad del mismo para explicar la variación de los datos. El test de razón de verosimilitud establece que el modelo M_0 se rechaza con un nivel de significación α si $D > c\alpha$, donde $c\alpha$ es el cuantil $(1 - \alpha)$ de la distribución χ^2 con k grados de libertad, donde k es la diferencia en la dimensionalidad de M_1 y M_0 .

Las Tablas 9 y 10 muestran los resultados de aplicar el test de razón de verosimilitud para la serie de niveles máximos medidos en Rosario y los caudales máximos diarios medidos en la represa de Salto Grande. En todos los casos el modelo M_0 corresponde al modelo GEV clásico y el modelo M_1 es el indicado en la primera columna.

Tabla 9: Resultados del test de razón de verosimilitud aplicado a la serie de niveles máximos medidos en Rosario.

Modelo	D	c_α	p-valor
GEV1	7.865	3.8415	0.00504
GEV2	8.13	5.991	0.01716
GEV3	7.9174	5.991	0.01909

Tabla 10: Resultados del test de razón de verosimilitud aplicado a la serie de caudales máximos medidos en Salto Grande.

Modelo	D	c_α	p-valor
GEV1	4.013	3.841	0.04515
GEV2	8.3649	5.991	0.01526
GEV3	0	5.991	1

Las Tablas 11 y 12 muestran los resultados de aplicar el test de razón de verosimilitud para la serie de niveles máximos medidos en Rosario y los caudales máximos

diarios medidos en Salto Grande.

En ambos casos el modelo M_0 corresponde al modelo GEV1 y el modelo M_1 es el indicado en la primera columna.

Tabla 11: Resultados del test de razón de verosimilitud aplicado a la serie de niveles máximos medidos en Rosario.

Modelo	D	c_α	p-valor
GEV2	0.3016	3.841	0.5829
GEV3	0.0657	3.841	1

Tabla 12: Resultados del test de razón de verosimilitud aplicado a la serie de caudales máximos medidos en Salto Grande.

Modelo	D	c_α	p-valor
GEV2	4.3519	3.841	0.03697
GEV3	-4.013	3.841	1

Del análisis de la Tabla 9 se concluye que, en el caso de los niveles máximos diarios medidos en Rosario, los modelos GEV1, GEV2 y GEV3 explican mejor la variación de los datos que el modelo GEV clásico. Sin embargo, de acuerdo a los resultados mostrados en la Tabla 11, el modelo GEV2 y el modelo GEV3 no explican mejor la variación de los niveles máximos diarios que el modelo GEV1. Del análisis de la Tabla 10 se concluye que, en el caso de los caudales máximos medidos en Salto Grande, solo los modelos GEV1, GEV2 explican mejor la variación de los datos que el modelo GEV clásico. Y de acuerdo a los resultados mostrados en la Tabla 12, el modelo GEV2 explica mejor la variación de los caudales máximos diarios que el modelo GEV1.

La Figura 3 muestra el gráfico de la distribución GEV1 para distintos valores del índice SOI medio y la distribución GEV clásica para los niveles máximos medidos en Rosario.

La Figura 4 muestra el gráfico de la distribución GEV2 para distintos valores del índice SOI medio y la distribución GEV clásica para los caudales máximos diarios medidos en la represa de Salto Grande.

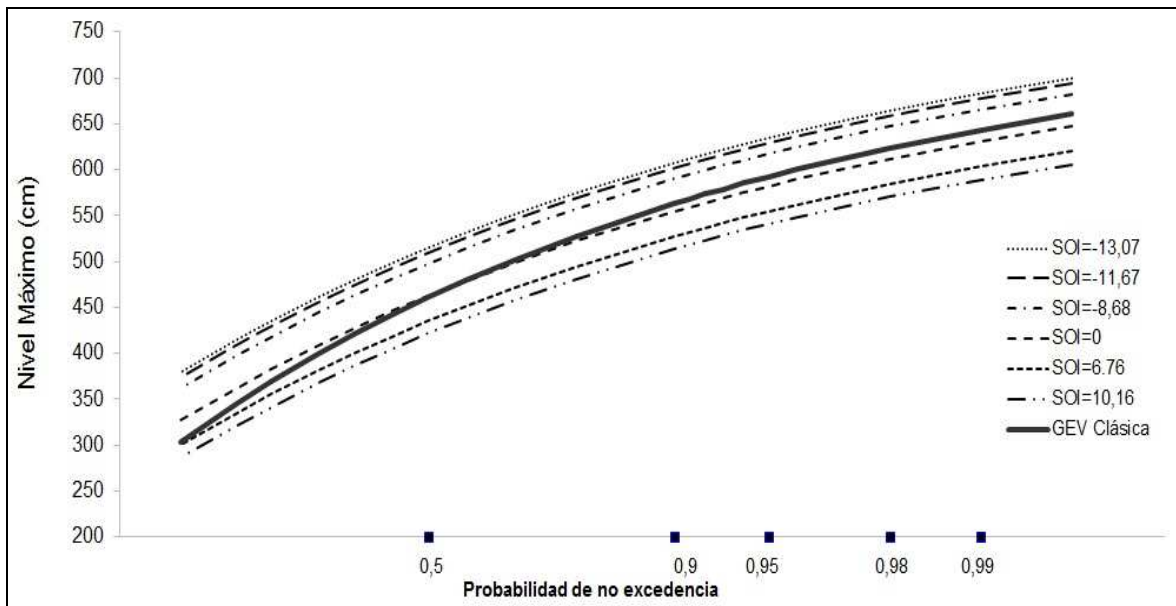


Figura 3. Análisis de frecuencia de niveles máximos en Rosario para distintos valores del índice *SOI* comparados con la distribución *GEV* clásica.

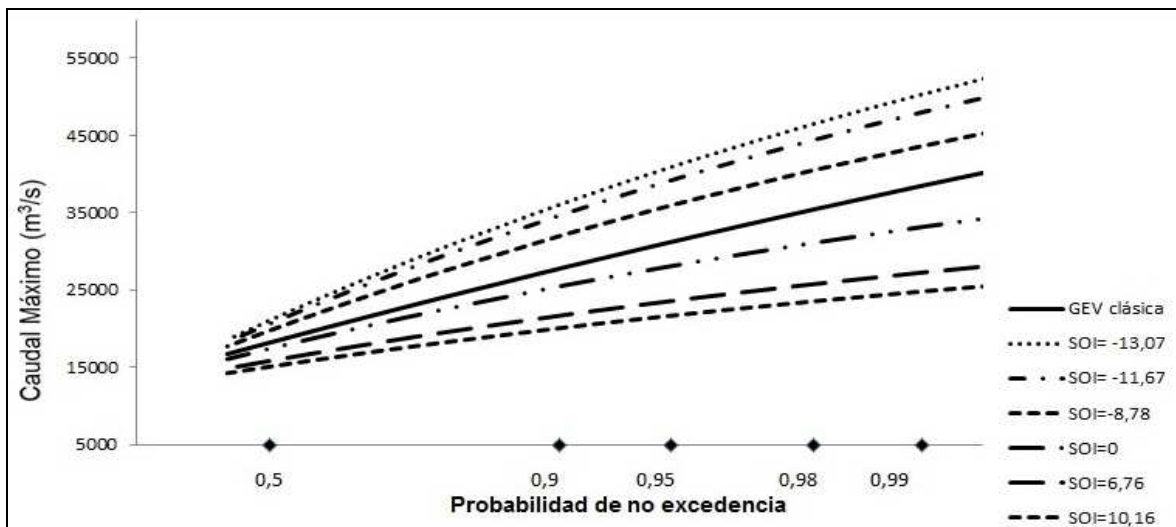


Figura 4. Análisis de frecuencia de caudales máximos en Salto Grande para distintos valores del índice *SOI* comparados con la distribución *GEV* clásica.

Al abordar el fenómeno de la no estacionariedad, uno de los problemas que se plantea es el de extender el concepto de período de retorno. Se plantean dos definiciones: la primera, (Olsen et al., 1998), define el nivel de retorno de m años como el valor del extremo para el cual el tiempo de espera esperado hasta la superación es m años.

La segunda, (Parey et al., 2010) define el período de retorno de m años como el valor extremo para el cual el número esperado de excedencias en m años

es uno. Si se considera que los parámetros dependen únicamente del tiempo (Cooley, 2012), es decir, del año de ocurrencia de cada evento máximo, entonces estas definiciones tienen sentido.

En cambio, al considerar que los parámetros dependen de una covariable, como el *SOI*, no es posible aplicar ninguna de las dos definiciones enunciadas anteriormente. Sí, es posible, calcular el nivel de retorno de m – años condicionado a un valor del promedio anual del *SOI*.

En la Figura 5 se muestra como varían los niveles de retorno condicionados por los distintos valores que alcanzó el promedio anual del SOI en los años considerados.

En la Figura 6 se muestra como varían los caudales de retorno condicionados por los distintos valores que alcanzó el promedio anual del SOI en los años considerados.

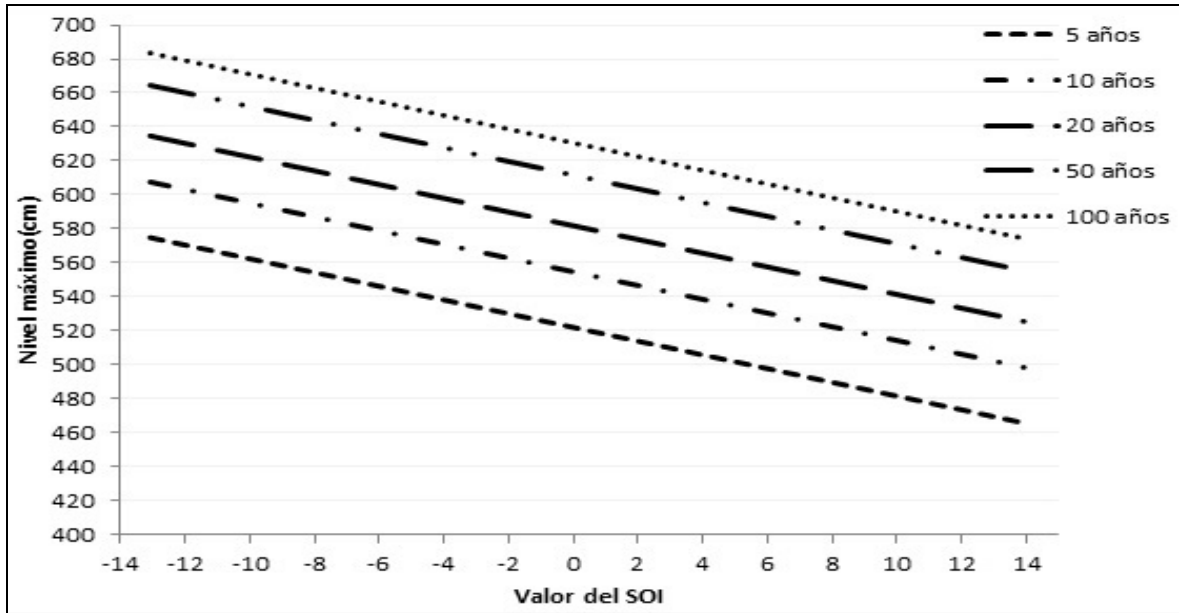


Figura 5. Niveles máximos estimados para Rosario y tiempo de retorno de 5 a 100 años mediante el modelo GEV1, en función de distintos valores del promedio anual del SOI.

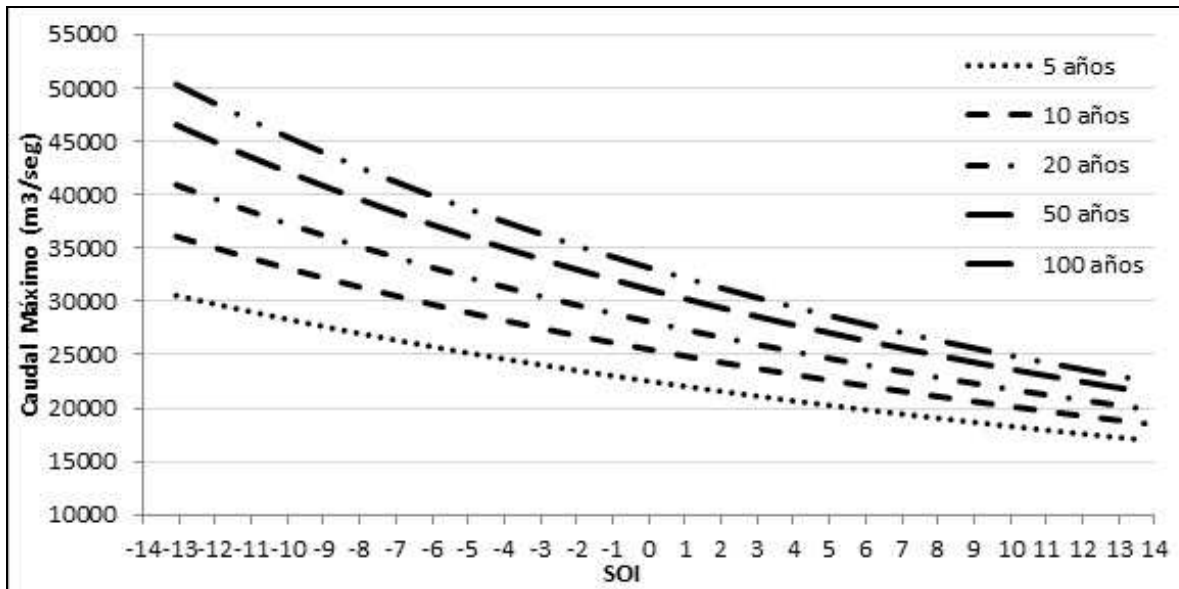


Figura 6. Caudales máximos estimados para el Río Uruguay y tiempo de retorno de 5 a 100 mediante el modelo GEV2, en función de distintos valores del promedio anual del SOI.

CONCLUSIONES Y RECOMENDACIONES

A partir de los resultados obtenidos, se concluyó que: El modelo mixto-GEV, proporciona un buen ajuste a las distribuciones de los niveles máximos diarios observados en Rosario.

El modelo mixto-GEV, proporciona un buen ajuste a las distribuciones de los caudales de aportes máximos diarios observados en la represa de Salto Grande.

Para el conjunto de estaciones pluviométricas analizadas, no se puede considerar que las precipitaciones máximas diarias provengan de poblaciones distintas con anterioridad y posterioridad al año 1970, que se considera como año característico de quiebre en cuanto a la influencia del cambio climático en la cuenca del Plata. Por lo tanto, sobre este conjunto de estaciones y por ende sobre la provincia de Entre Ríos el cambio climático sólo presenta evidencia aislada en cuanto a haber afectado la magnitud de los eventos pluviométricos máximos diarios.

Si bien las medias de las precipitaciones máximas diarias correspondiente a los años Niño son mayores a las medias correspondientes a los años no-Niño en la mayoría de las estaciones analizadas (9 de los casos), esta evidencia no es estadísticamente significativa al 5%. Por lo tanto, no se puede considerar con la información disponible, que la magnitud de las precipitaciones máximas diarias analizadas pertenezcan a poblaciones distintas en los años Niño que en los años No-Niño.

En coincidencia con la evidencia que presentaron Pisciotano et al. (1994) para la región del Noroeste del Uruguay, en esta región de estudio existe una tendencia a que los totales pluviométricos mensuales en los años Niño sean mayores que en los años no-Niño.

En la región del delta Entrerriano no se verifican diferencias estadísticamente significativas en el número medio de eventos por año de recurrencia superior a $T_r = 1.01$ años en años Niño con respecto a años No-Niño. En la región centro norte hay una tendencia consistente a que estas diferencias sean estadísticamente significativas. Por lo que, sólo en el Norte entrerriano hay un aumento consistente de la frecuencia de eventos extremos con posterioridad a 1980.

Se muestra la utilidad de establecer modelos de análisis de frecuencia no-estacionarios para considerar la posibilidad de que los parámetros varíen en dependencia a una variable exógena relacionados con la señal del fenómeno El Niño.

Si bien en el caso no estacionario no se puede aplicar el concepto clásico de período de retorno, es posible estimar valores extremos en función de la variable exógena considerada para distintos períodos de retorno.

En el caso de los niveles del río Paraná el modelo que explica mejor la variación de estos niveles máximos anuales es el de una distribución GEV donde su parámetro de posición depende linealmente del SOI.

En el caso de los caudales de aporte del río Uruguay el modelo que explica mejor la variación de estos caudales máximos anuales es el de una distribución GEV donde tanto su parámetro de posición, como el parámetro de escala, dependen del SOI.

REFERENCIAS

- Coles, S. 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. Springer Verlag London. 208p.
- Cooley, D. 2012. Return Periods Under Climate Change, Extremes in a Changing Climate: Detection, Analysis & Uncertainty. Springer, Netherlands, Dordrecht. *Water Science and Technology Series*, Vol 65.
- El Adlouni, S.; T. B. M. J. Ouarda; X. Zhang; R. Roy and B. Bobée. 2007. Generalized maximum likelihood estimators for the non-stationary GEV model. *Water Resour. Res.*, 43, W03410, doi:10.1029/2005WR004545.
- Fernández Loureiro de Pérez, E. 2011. *Estadística noparamétrica*. 2da.Edición. Ediciones Cooperativas.
- García, N. y Vargas W. 1998. The temporal climatic variability in the Río de la Plata basin displayed by the river discharges. *Climate Change*, 38, 359-379.
- García, R. M. 2004. *Inferencia estadística y diseño de experimentos*. 1ra. Edición. Ed. Eudeba.
- Gilleland, E. and Katz, R. W. 2011. New software to analyze how extremes change over time. *Eos*, 92(2), 13-14.

- Hosking, J. R. M. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B* 52: 105–124.
- Jaime, P. y Menéndez, A. 2002. Análisis del Régimen Hidrológico de los Ríos Paraná y Uruguay. Informe INA-LHA 05-216-02, Comitente: Proyecto Freplata.
- NOAA. National Oceanic and Atmospheric Administration. http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml.
- Olsen J.; Lambert J.H. and Haines Y. 1998. Risk of extreme events under nonstationary conditions. *Risk Anal* 18:497–510.
- Parey S; Hoang TTH and Dacunha-Castelle D .2010. Different ways to compute temperature return levels in the climate change context. *Environmetrics* 21:698–718.
- Pisciottano, G.; Díaz, A.; Cazes, G. and Mechoso, C. 1994 El Niño - Southern Oscillation Impact on Rainfall in Uruguay. *J. Climate*; 7: 1286-1302.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>.
- Zamanillo, E.; Tito, M. J. y Pérez, M. 2013 Análisis de frecuencia de niveles máximos en el delta del río Paraná. XXIV° Congreso Nacional del Agua 2013.

Artículo recibido el 04/2016 y aprobado para su publicación el 12/2016.