

## ANÁLISIS DE TENDENCIA EN SERIES AUTOCORRELACIONADAS

**Patricia M. López<sup>(1)</sup>, Rafael S. Seoane<sup>(2,4)</sup>, Luis A. Fornero<sup>(1)</sup>,  
José L. Rodríguez<sup>(3)</sup> y Marta S. Nuñez<sup>(1)</sup>**

Instituto Nacional del Agua, <sup>(1)</sup> Centro Regional Andino, <sup>(2)</sup> Dirección de Servicios Hidrológicos

<sup>(3)</sup> Laboratorio de Matemática Aplicada, UTN Facultad Regional Mendoza

<sup>(4)</sup> Consejo Nacional de Investigaciones Científicas y Técnicas

Belgrano 210, M5500FIF Mendoza. Telefax: 0261-4288251. E-mail: [pmlopez@ina.gov.ar](mailto:pmlopez@ina.gov.ar)

### RESUMEN

Las posibles modificaciones en las componentes del ciclo hidrológico asociadas a la ocurrencia de cambios climáticos han sido un tema de interés para los hidrólogos y planificadores de los recursos hídricos durante los últimos años. Numerosos investigadores han orientado sus estudios a la identificación de tendencia en series de variable hidrometeorológicas, siendo la prueba no paramétrica más utilizada la de Mann-Kendall (MK).

La presencia de autocorrelación en las series temporales afecta los resultados de esta prueba, aumentando (disminuyendo) la probabilidad de detección de tendencia cuando la autocorrelación es positiva (negativa). Distintos análisis estadísticos confirman la presencia de autocorrelación en series hidrológicas en diferentes escalas temporales.

Los trabajos de distintos autores confirman la presencia de autocorrelación en series hidrológicas en diferentes escalas temporales.

En este estudio se estiman, mediante simulación, las relaciones de potencia de la prueba MK modificada y la técnica de pre-blanqueo TFPW, metodologías que resuelven el problema que presenta la aplicación de la prueba a series autocorrelacionadas.

Las relaciones obtenidas muestran que la potencia crece con el nivel de significación seleccionado, la magnitud de la tendencia, el tamaño de la muestra y decrece con la variación interna de la serie. Se observó que la técnica de pre-blanqueo TFPW tiende a ser más potente para coeficientes de autocorrelación positivos, mientras que la prueba MK modificada lo es para coeficientes negativos.

Este estudio permite precisar algunas limitaciones para la aplicación de la prueba Mann-Kendall y puede favorecer a una mejor comprensión de los resultados obtenidos en trabajos anteriores que investigan la posible presencia de tendencias en series de caudales observados en ríos argentinos.

**Palabras Claves:** Prueba Mann-Kendall; análisis de tendencia; autocorrelación; potencia de una prueba estadística.

## INTRODUCCIÓN

Las modificaciones potenciales en las componentes del ciclo hidrológico, debidas a cambios climáticos, han sido un tema de interés para los hidrólogos y planificadores de recursos hídricos durante los últimos años. Numerosos autores han orientado sus estudios a la identificación de tendencia en series de variables hidrometeorológicas, detectando variaciones en el régimen de caudales de Estados Unidos (Lettenmaier *et al.*, 1994), Canadá (Westmacott y Burn, 1997; Cunderlik y Burn, 2002, 2004; Yue *et al.*, 2003) y Suiza (Birsan *et al.*, 2005). En Argentina, varios estudios analizan tendencias y cambios en el régimen fluvial de ríos con nacientes en la Cordillera de los Andes (Seoane *et al.*, 2005a, b; Vich *et al.*, 2006).

La prueba no-paramétrica más utilizada en la identificación de tendencia en series de variables hidrometeorológicas ha sido la de Mann-Kendall (MK). La autocorrelación presente en algunas de estas series, afecta los resultados de la prueba, causando problemas en la detección y evaluación del nivel de significación. La presencia de autocorrelación positiva tiende a aumentar la probabilidad de detección de tendencia cuando, en efecto, ésta no existe, mientras que la autocorrelación negativa tiende a disminuirla.

Para corregir este efecto, Hamed y Rao (1998) derivan una relación teórica para calcular la varianza del estadístico de la prueba para datos autocorrelacionados y proponen una prueba MK modificada, mientras que Yue, *et al.* (2002b) presentan una técnica de pre-blanqueo (trend-free pre-whitening, TFPW) que remueve la correlación serial sin alterar la tendencia presente en la serie.

En este estudio se estiman, mediante simulación, las relaciones de potencia de la prueba MK obtenidas de la aplicación de las dos metodologías anteriores. Estas relaciones muestran que la potencia crece con el nivel de significación seleccionado, la magnitud de la tendencia, el tamaño de la muestra y decrece con la variación interna de la serie. Se observó que la técnica de pre-blanqueo TFPW tiende a ser más potente para coeficientes de autocorrelación positivos, mientras que la prueba MK modificada lo es para coeficientes negativos.

## METODOLOGÍAS PARA SERIES TEMPORALES CON AUTOCORRELACIÓN

La prueba no-paramétrica MK (Hirsch *et al.*, 1982; Westmacott y Burn, 1997; Cunderlik y Burn, 2002, Yue *et al.*, 2002a), establece como hipótesis nula  $H_0$  que los datos  $(x_1, \dots, x_N)$  son variables aleatorias independientes e idénticamente distribuidas. Si la serie analizada tiene autocorrelación positiva la varianza estadístico de la prueba  $S$  es subestimada, aumentando así la probabilidad de detección de tendencia. El efecto inverso se presenta en series con autocorrelación negativa. Si bien varios autores han planteado soluciones a este problema, en este estudio se consideran las metodologías propuestas por Hamed y Rao (1998) y Yue *et al.* (2002b).

### Prueba MK modificada

La prueba MK modificada para series autocorrelacionadas propuesta por Hamed y Rao (1998) está basada en la corrección de la varianza de  $S$  en función del número efectivo de observaciones,  $N_s$ . La varianza modificada es:

$$V^*(S) = Var(S) \frac{N}{N_s} \quad (1)$$

La mejor aproximación a los valores teóricos derivados por los autores está dada por el factor de corrección:

$$\frac{N}{N_s} = 1 + \frac{2}{N(N-1)(N-2)} \sum_{i=1}^{N-1} (N-i)(N-i-1)(N-i-2) \rho_S(i) \quad (2)$$

donde  $N$  es el número total de observaciones y  $\rho_S(i)$  son valores significativos de la función de autocorrelación de la serie de los órdenes de las observaciones, calculados después de remover la tendencia estimada por algún método no-paramétrico robusto (Sen, 1968).

### **Metodología de pre-blanqueo TFPW**

El procedimiento de pre-blanqueo propuesto por Yue *et al.* (2002b) para eliminar el efecto de la correlación serial en la aplicación de la prueba MK es denominado *pre-blanqueo libre de tendencia* (trend-free pre-whitening, TFPW). Esta técnica propone remover de la serie estandarizada la tendencia, estimada por algún método robusto (Sen, 1968), antes de calcular el coeficiente de autocorrelación de retardo-1 ( $r_1$ ). De esta forma se resuelve el problema de la influencia de la tendencia en la estimación del coeficiente autocorrelación, mientras que la remoción de la tendencia no afecta al proceso auto-regresivo subyacente. Se asume que la tendencia es lineal y se elimina de la serie según:

$$Y_t = X_t - T_t = X_t - bt \quad (3)$$

Se estima el coeficiente  $r_1$  de la serie  $Y_t$  y si éste no es significativamente distinto de cero, la serie se considera independiente y la prueba MK se aplica a la serie original. Si  $r_1$  es significativamente distinto de cero, la serie  $Y_t$  se pre-blanquea a fin de remover el proceso AR(1) como:

$$Y_t' = Y_t - r_1 Y_{t-1} \quad (4)$$

La tendencia estimada y la serie  $Y_t'$ , compuesta por residuos que deben ser independientes, se combinan como:

$$Y_t'' = Y_t' + T_t \quad (5)$$

y la prueba MK se aplica a la serie  $Y_t''$ , evitando de esta forma el efecto negativo de la correlación serial.

## **MODELOS CONSIDERADOS**

Existen dos tipos de errores asociados a las pruebas de hipótesis: (a) Error tipo I: rechazar la hipótesis nula siendo cierta y (b) Error tipo II: aceptar la hipótesis nula siendo falsa. La función de potencia de una prueba puede ser definida como  $K(H)$  la probabilidad de rechazar la hipótesis nula  $H_0$  bajo la hipótesis  $H$ . Si  $H_a$  es una hipótesis alternativa:

$$K(H_a) = 1 - \text{Probabilidad de cometer el error tipo II} \quad (6)$$

La expresión anterior es la potencia de la prueba para  $H_a$ , o sea la probabilidad de rechazar la hipótesis nula cuando es falsa.

La hipótesis nula en una prueba de aleatoriedad es  $H_0: X_t$  ( $t = 1, \dots, N$ ) es un proceso aleatorio compuesto por variables idénticamente distribuidas y no correlacionadas.

Para investigar la potencia de las metodologías estudiadas se utilizan como hipótesis alternativas series generadas con estructura de autocorrelación retardo-1 y tendencia lineal.

Las relaciones de potencia se obtienen, mediante simulación, utilizando como mecanismo generador modelos con estructura de autocorrelación retardo-1,  $A_t$ , y tendencia lineal,  $T_t$ , de la forma:

$$X_t = A_t + T_t \quad (7)$$

El modelo AR(1) expresado en función de los momentos del proceso aleatorio  $A_t$  es:

$$A_t - \mu = \rho_1 (A_{t-1} - \mu) + \sigma_A (1 - \rho_1^2)^{0.5} \varepsilon_t \quad (8)$$

donde  $\mu$  es la media del proceso,  $\sigma_A^2$  su varianza,  $\rho_1$  coeficiente de autocorrelación de retardo-1 y  $\varepsilon_t$  es una variable aleatoria normal independiente con media cero y varianza uno. En el caso del modelo AR(1) el parámetro auto-regresivo del modelo,  $\phi_1$ , coincide con  $\rho_1$ .

Según la ecuación (7), se simularon 2000 series sintéticas de longitudes  $N = 10$  (10) 150 con media uno y varianza  $\sigma_A^2 = (0.1 \times i)^2$  ( $i = 1$  (1) 10). Los experimentos diseñados consideran procesos auto-regresivos con coeficiente  $\rho_1 = -0.6$  (0.2) 0.6 y tendencia lineal  $T_t = bt$ , siendo  $b = -0.01$  (0.002) 0.01.

## RESULTADOS

Las relaciones de potencia estimadas de la prueba MK para series con autocorrelación son semejantes a las obtenidas para el caso de series independientes (Yue *et al.*, 2002a). La potencia crece con el valor absoluto de la magnitud ( $b$ ) de la tendencia, el tamaño de la muestra ( $N$ ) y el nivel de significación prefijado ( $\alpha$ ) y decrece con la variación interna de la serie, medida por su coeficiente de variación ( $Cv$ ).

Las dos metodologías aplicadas resuelven el problema que presenta la utilización de la prueba original en series autocorrelacionadas, aumentando la probabilidad de detección de tendencias para autocorrelación negativa y disminuyéndola en los casos de autocorrelación positiva.

Los resultados muestran que la técnica de pre-blanqueo TFPW tiende a ser más potente para coeficientes de autocorrelación positivos, mientras que la prueba MK modificada lo es para coeficientes negativos.

Con el propósito de demostrar como varía la función de potencia se presentan en las Figuras 1 a 4 algunos de los resultados obtenidos. La relación potencia,  $b$  y  $\alpha$  se presenta en la Figura 1 para series de longitud  $N = 50$ ,  $Cv = 0.5$  y coeficiente de autocorrelación ( $\rho_1$ ) de 0.2, donde se observa que la potencia crece con el valor absoluto de  $b$  para un  $\alpha$  fijo, mientras que para un valor de  $b$  determinado la potencia crece al incrementar el nivel de significación de la prueba.

La Figura 2 muestra la relación potencia- $Cv$ - $b$  para  $N = 30$ ,  $\rho_1 = 0.6$  y un nivel de significación 0.05, donde se advierte la disminución de la potencia con el aumento la variación interna de la serie, medida por el  $Cv$ . Esta relación indica que es más difícil detectar una tendencia determinada cuando la variación interna de la serie crece ya que ésta enmascara la tendencia existente.

La relación potencia,  $\rho_1$  y  $N$ , para series con  $Cv = 0.5$  y tendencia lineal con pendiente  $b = 0.004$  y nivel de significación 0.05, se presenta en la Figura 3. Se puede observar que la potencia aumenta al incrementarse el número de observaciones y que para el mismo número de observaciones  $N$ , la técnica de pre-blanqueo TFPW tiende a ser más potente para coeficientes de autocorrelación positivos, mientras que la prueba MK modificada lo es para coeficientes negativos.

Las relaciones potencia,  $Cv$  y  $N$  permiten, seleccionado la potencia requerida y para un nivel de significación preasignado, determinar la cantidad mínima de observaciones necesarias para identificar una tendencia determinada en función de variación interna y de la estructura de autocorrelación de la serie. Un ejemplo de estas relaciones se muestra en la Figura 4.

En la Tabla 1 se presentan, para  $Cv$  entre 0.1 y 0.9, los valores mínimos de observaciones necesarias para obtener potencias mayores o iguales al 90% para series normalmente distribuidas independientes y autocorrelacionadas ( $0.2 \leq \rho_1 \leq 0.6$ ).

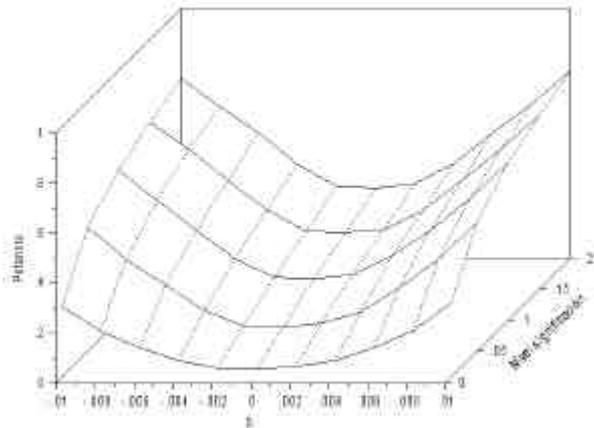
Tabla 1. Valores  $N$  necesarios para obtener potencias  $\geq 0.90$  para  $\alpha=0.05$

$Cv$	Tipo de serie	$ b $				
		0.002	0.004	0.006	0.008	0.01
0.1	SI	70	40-50	30-40	30	20-30
	SA <sup>(1)</sup>	70-80	40-50	30-40	30	30
	SA <sup>(2)</sup>	70-100	50-60	40-50	30-40	30-40
0.3	SI	140-150	90-100	70	60	50
	SA <sup>(1)</sup>	140-150	90-100	70	60	50
	SA <sup>(2)</sup>		100-130	70-100	60-80	50-70
0.5	SI		130	100	80-90	70
	SA <sup>(1)</sup>		120-130	100	80-90	70
	SA <sup>(2)</sup>		140->150	100-140	90-120	70-100
0.7	SI			120-130	100	90
	SA <sup>(1)</sup>			120-130	100	90
	SA <sup>(2)</sup>			130->150	110-140	90-130
0.9	SI			140-150	120	100-110
	SA <sup>(1)</sup>			140-150	110-120	100-110
	SA <sup>(2)</sup>				130->150	110-150

SI: Serie independiente. SA: Serie autocorrelacionada  
<sup>(1)</sup>: TFPW. <sup>(2)</sup>: MK modificada

La Tabla 1 muestra la dificultad para identificar tendencias de baja pendiente en series con escasas observaciones y gran variabilidad interna y que, en general, con 30 años de registro se pueden detectar tendencias de mediana a alta pendiente siempre solo si la variación interna de la serie es muy baja.

(a) Técnica de pre-blanqueo TFPW



(b) Prueba MK modificada

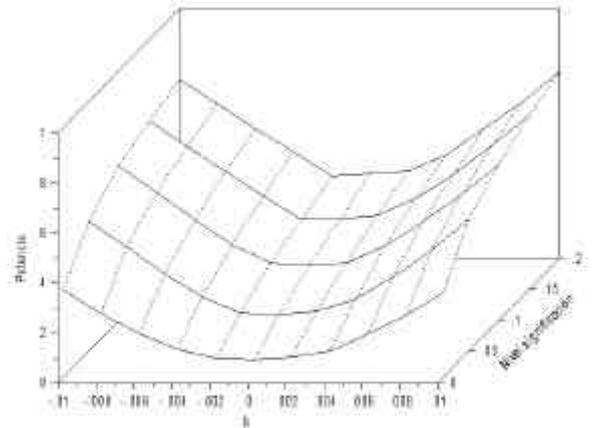
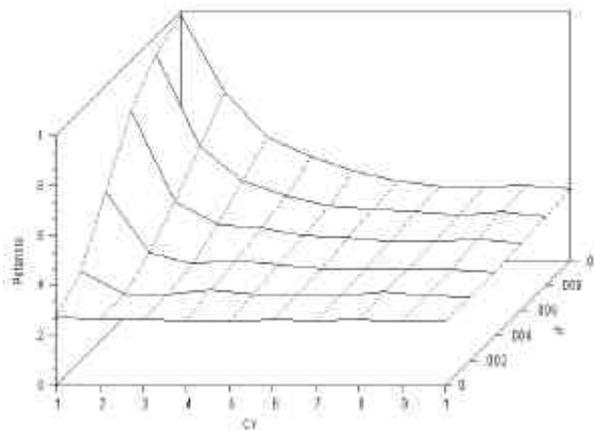


Figura 1. Potencia- $b$ -nivel de significación para  $N = 50$ ,  $Cv = 0.5$ ,  $\rho_1 = 0.2$

(a) Técnica de pre-blanqueo



(b) Prueba MK modificada

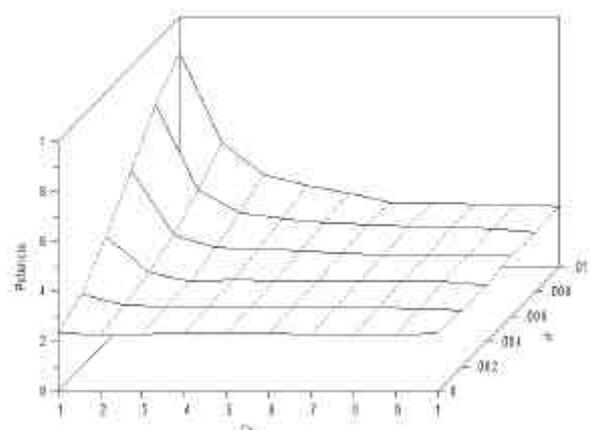
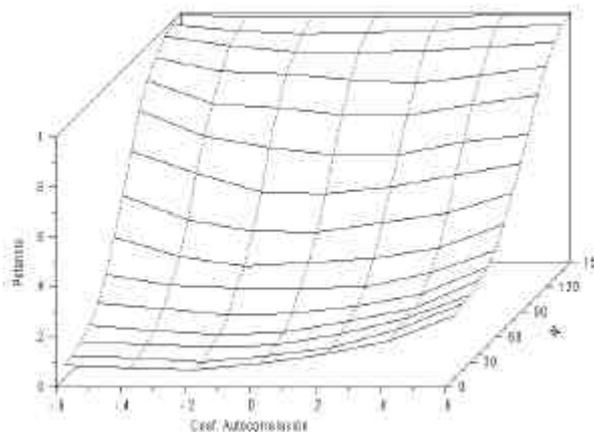


Figura 2. Potencia- $Cv$ - $b$  para  $N = 30$ ,  $\rho_1 = 0.6$ ,  $\alpha = 0.05$

(a) Técnica de pre-blanqueo



(b) Prueba MK modificada

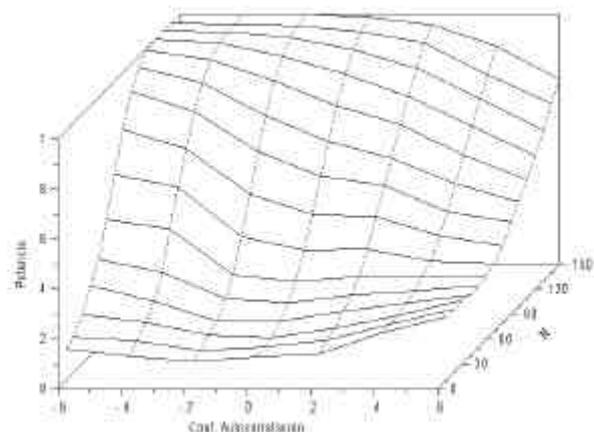
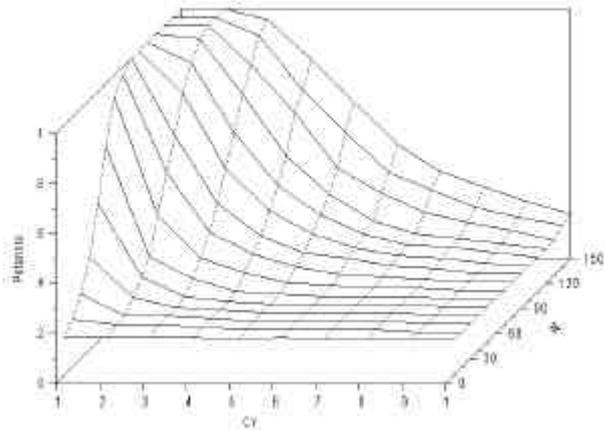


Figura 3. Potencia-coeficiente de autocorrelación- $N$  para  $Cv = 0.5$ ,  $b = 0.004$ ,  $\alpha = 0.05$

(a) Técnica de pre-blanqueo TFPW



(b) Prueba MK modificada

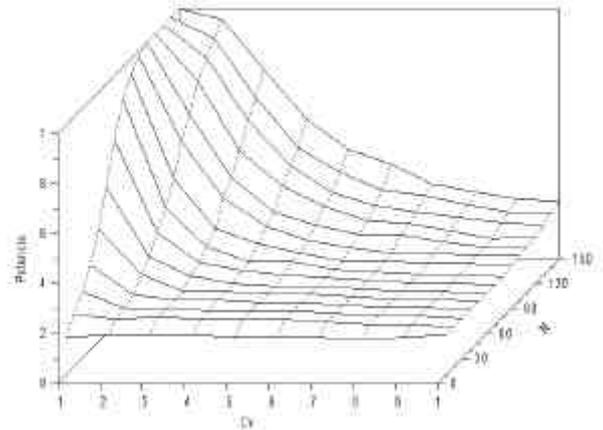


Figura 4. Potencia- $Cv$ - $N$  para  $b = 0.002$ ,  $\rho_1 = 0.4$ ,  $\alpha = 0.05$

## CONCLUSIONES

En este estudio se han estimado, mediante simulación, las relaciones de potencia de la prueba MK modificada y de un procedimiento de pre-blanqueo TFPW, metodologías utilizadas para estudios de identificación de tendencia en series temporales autocorrelacionadas.

Se verifica que las dos metodologías estudiadas resuelven el problema que presenta la aplicación de la prueba a series autocorrelacionadas y los resultados de los experimentos desarrollados demuestran que estas relaciones de potencia son semejantes a las obtenidas para el caso de series independientes. La potencia crece con el nivel de significación seleccionado ( $\alpha$ ), la magnitud de la tendencia ( $b$ ), el tamaño de la muestra analizada ( $N$ ) y decrece con la variación interna de la serie, medida por su coeficiente de variación ( $Cv$ ).

Los resultados de este estudio indican que la técnica de pre-blanqueo TFPW tiende a ser más potente para coeficientes de autocorrelación positivos, mientras que la prueba MK modificada lo es para coeficientes negativos. Este efecto asimétrico de las metodologías estudiadas con respecto al coeficiente de autocorrelación ocasiona diferencias en la significación de las tendencias detectadas.

La técnica de pre-blanqueo TFPW supone una aproximación lineal de la tendencia presente y la prueba MK modificada está basada en la corrección de la varianza del estadístico en función del tamaño efectivo de la serie analizada, valor presentado por otros autores como difícil de estimar de manera confiable. Por lo tanto, si la presencia de tendencia es evidente se recomienda utilizar la técnica de pre-blanqueo TFPW para determinar la significación de la misma.

Las consideraciones anteriores deberían ser examinadas conjuntamente con las relaciones de potencia obtenidas en este estudio, a fin de seleccionar la propuesta más adecuada para detectar tendencia en series autocorrelacionadas.

Las relaciones obtenidas de potencia- $Cv$ - $N$  permiten, seleccionando la potencia deseada, estimar la cantidad mínima de observaciones necesarias para detectar una tendencia determinada. Información que puede resultar útil para indicar el período mínimo de observación necesario para monitorear y evaluar cambios en variables hidrometeorológicas.

## AGRADECIMIENTOS

Este estudio fue financiado por el Instituto Nacional del Agua, en el marco de Subsidios de Proyectos Institucionales de Investigación y Desarrollo.

## BIBLIOGRAFÍA

- Birsan, M. V.; P. Molnar, P. Burlando y M. Pfaundler** (2005) “*Streamflow trends in Switzerland*”. *Journal of Hydrology* 314, pp. 312-329.
- Cunderlik, J. M. y D. H. Burn** (2002) “*Local and regional trends in monthly maximum flows in southern British Columbia*”. *Canadian Water Resources Journal* 27(2), pp. 191-212.
- Cunderlik, J. M. y D. H. Burn** (2004) “*Linkages between regional trends in monthly maximum flows and selected climatic variables*”. *Journal of Hydrologic Engineering ASCE*, 9(4), pp. 246-256.
- Hamed K. H. y A. R. Rao** (1998) “*A modified Mann-Kendall Trend Test for Autocorrelated Data*”. *Journal of Hydrology* 204, pp. 182-196.
- Hirsch, R. M.; J. R. Slack y R. A. Smith** (1982) “*Techniques of trend analysis for monthly water quality data*”. *Water Resources Research* 18(1), pp. 107-121.
- Lettenmaier, D. P.; E. F. Wood y J. R. Wallis** (1994) “*Hydro-climatological trends in the continental United States 1948-1988*”. *Journal of Climate* 7(4), pp. 586-607.
- Sen, P. K.** (1968) “*Estimates of the regression coefficient based on Kendall’s tau*”. *Journal of the American Statistical Association*, 63, pp. 1379-1389.
- Seoane, R. S.; J. Valdés y L. J. Mata** (2005a) “*Climate variability and climate change in Patagonian rivers*”. *Regional Hydrological Impacts of Climatic Change Impact Assessment and Decision Making*. Edited by Stewart Franks, Thorsten Wagener, Eva Bøgh, Hoshin V. Gupta, Luis Bastidas & Carlos Nobre & Carlos de Oliveira Galvão Publ. 296. ISBN 1-901502-13-9; 300 pp. 26-36.
- Seoane, R. S.; P. M. López; M. C. Moyano; N. Tomazin y M. E. Koutsovitis** (2005b) “*Análisis estadístico de series hidrológicas en la cuenca del río Limay (Argentina)*”. Libro de Resúmenes del XX Congreso Nacional del Agua y III Simposio de Recursos Hídricos del Cono Sur (ISBN 987-22143-0-1), pp. 334 y en CD (ISBN 987-22143-0-1).
- Vich, A. I.; P. M. López y E. Martínez** (2006) “*Detección de tendencias y cambios abruptos en el régimen fluvial de ríos de Argentina, con nacientes en la Cordillera de los Andes*”. Libro de Resúmenes del Simposio sobre Cambio Climático: Organizando la Ciencia para la Cordillera Americana (CONCORD), pp. 62-63.
- Westmacott, J. R. y D. H. Burn** (1997) “*Climate changes effects on the hydrologic regime within the Churchill-Nelson River Basin*”. *Journal of Hydrology* 202, pp. 263-279.
- Yue, S.; P. Pilon y G. Cavadias** (2002a) “*Power of the Mann-Kendall and Spearman’s tests for detecting monotonic trends in hydrological series*”. *Journal of Hydrology* 259, pp. 254-271.
- Yue, S., P. Pilon, B. Phinney y G. Cavadias** (2002b) “*The influence of autocorrelation on the ability to detect trend in hydrological series*”. *Hydrological Processes* 16, pp. 1807-1829.
- Yue, S.; P. Pilon y B. Phinney** (2003) “*Canadian streamflow trend detection: Impacts of serial and cross-correlation*”. *Hydrological Sciences Journal* 48(1), pp. 51-63.