

SOCIEDAD CHILENA DE INGENIERIA HIDRAULICA
VIII CONGRESO NACIONAL

GENERACION DE SERIES DE VARIABLES HIDROLOGICAS
USANDO MODELO ARMA

Ximena Vargas M. (1)
Ernesto Brown F. (1)
Julio Miranda P. (2)

RESUMEN

Se investiga la aplicabilidad de Modelos ARMA en la generación de caudales medios mensuales, comparando la generación directa de éstos, con su simulación a través de series de precipitación generadas, usadas como entradas de un modelo de simulación hidrológica.

Los resultados se analizan mediante distintas pruebas de significancia estadística. También, se examinan las series generadas desde el punto de vista de su posible aplicación en el diseño de volúmenes de regulación de embalses, a través del análisis de rangos, magnitudes de déficit y duración de estos periodos de déficit.

(1) Profesor e Investigador Centro de Recursos Hidráulicos.

(2) Ayudante de Investigación

1. INTRODUCCION

Este trabajo corresponde a una extensión de aquel presentado por los autores en el V Congreso de Metodologías en Ingeniería de Sistemas (Vargas y Brown, 1985). Aquí se presenta un análisis complementario al anterior, con respecto al uso de modelos de tipo ARMA para generar series, tanto de caudales medios mensuales directamente, como de precipitaciones mensuales para usarlas como entrada a un modelo de simulación hidrológica que, a su vez, genera las series de caudales medios mensuales en la sección de salida de una cuenca. Adicionalmente, en este trabajo se examina las características de estas series en cuanto a su empleo en la determinación de volúmenes de regulación de embalses. El modelo de simulación hidrológica utilizado, corresponde a una versión levemente modificada de aquel desarrollado por Ferrer et al (1973).

2. MODELACION

2.1 Selección del Modelo ARMA

Un modelo no estacional ARMA (p, q) puede expresarse como

$$\phi(B) \bar{z}_t = \theta(B) a_t \quad (1)$$

en que $\phi(B)$ y $\theta(B)$ son polinomios autorregresivos (AR) de orden p, y de promedio móvil (MA) de orden q, respectivamente; B es el operador de retroceso definido como $B^i \bar{z}_t = \bar{z}_{t-i}$; $\bar{z}_t = Z_t - \bar{Z}$, donde Z_t es la serie de tiempo a modelar y \bar{Z} su media; a_t es una serie de impactos aleatorios o residuos normalmente distribuidos, con media cero y varianza σ_a^2 .

La serie \bar{z}_t debe ser estacionaria y poseer una distribución normal, para lo cual esta debe ser estandarizada y transformada apropiadamente, de modo de preservar los estadígrafos de interés

de la serie histórica (Vargas y Brown, 1985). Estos estadígrafos, para las series que se intenta modelar, pueden variar significativamente de un mes a otro, por lo que, al estandarizar dichas series, debe tenerse en cuenta esta característica. Si se considera que la variable X representa la serie histórica, e Y representa la serie transformada, la serie estandarizada Z_t queda entonces definida como:

$$Z_t = \frac{Y_{t, \tau} - \bar{Y}_\tau}{S_{Y_\tau}} \quad (2)$$

siendo \bar{Y}_τ la media y S_{Y_τ} la desviación estándar del mes τ de los valores transformados, calculados de modo de preservar al menos los estadígrafos \bar{X}_τ y S_{X_τ} .

Considerando que en el proceso de generación deben mantenerse los valores \bar{Y}_τ y S_{Y_τ} , es necesario que la serie Z_t corresponda a una densidad de probabilidades normal con media cero y varianza uno. Para lograr esto, los residuos a_t en el modelo dado por la ecuación (1), deben ser generados con un valor apropiado de su varianza (σ_a^2) que dependerá del tipo de modelo.

Para identificar el orden de los polinomios $\phi(B)$ y $\theta(B)$, y calcular los parámetros del modelo ARMA se usa la metodología descrita por los autores (Vargas et al 1983, 1984) para modelos ARIMA, de los cuales estos modelos son un caso especial.

2.2 Experimentación

Para seleccionar las estadísticas históricas a utilizar como base del proceso de generación, se tuvo en cuenta que su longitud fuese en lo posible superior a 20 años; correspondieran a ríos con regímenes fundamentalmente de tipo pluvial con estadísticas concurrentes de precipitaciones; y, presentaran diferencias apreciables en cuanto a magnitudes absolutas de los estadígrafos

y característicos.

En la Tabla 1 se presentan las estadísticas seleccionadas, y sus características fundamentales. Es conveniente hacer notar que para el modelo de simulación hidrológica, además se usó estadísticas evaporimétricas representativas de las respectivas cuencas.

TABLA 1
Características de estadísticas seleccionadas

ESTADÍSTICA	TIPO	PERIODO	MEDIA ANUAL (m ³ /s)	DESVIACION ESTANDAR (m ³ /s)	COEFICIENTE DE VARIACION (ANUAL)
- Purapel en Nirivilo	Q	5/57-4/73	2,0	1,0	0,502
Constitución	P	5/59-4/77			
- Putagán en Yerbas Buenas	Q	5/47-4/81	20,5	6,3	0,309
Linares	P	5/30-4/69			
Panimávida	P	5/39-4/69			
- Bureo en Mulchén	Q	5/41-4/81	47,2	13,9	0,295
Cerro el Padre	P	5/43-4/85			
- Mulchén en Mulchén	Q	5/37-4/72	24,0	7,7	0,321
Mulchén	P	5/38-4/85			
- Tucapel en Cañete	Q	5/54-4/71	12,8	4,1	0,319
Cañete	P	5/62-4/85			
- Purén en Tranamán	Q	5/49-4/81	11,6	4,3	0,370
Traiguén	P	5/37-4/65			
- Chillán en Esperanza	Q	5/46-4/75	14,4	4,5	0,313

NOTA : Q indica estadística de caudales.

P indica estadística de precipitaciones.

Con el objeto de no tomar en cuenta en el análisis de resultados las deficiencias propias del modelo de simulación hidrológica, por no corresponder esto a los objetivos de este trabajo, para fines comparativos, se optó por usar como serie "Histórica", la serie sintetizada a través del modelo de simulación hidrológica usando como entrada la precipitación histórica. En el caso del

rio Chillán en Esperanza, no se procedió de la forma descrita, puesto que sólo se usó para hacer una comparación de la generación directa, con los resultados obtenidos de la aplicación de un modelo mensual del tipo Markov de orden uno (Brown y Torretti, 1977).

Para cada estadística de caudales y precipitaciones se identificaron los modelos ARMA apropiados, que se indican en la Tabla 2.

TABLA 2
Modelos ARMA ajustados.

ESTADÍSTICA	M O D E L O	σ_a^2	TRANSF.
Purapel en Nirivilo	$Z_t = 0,364Z_{t-1} + 0,28Z_{t-2} + a_t$	0,676	Ln(x)
Constitución	$Z_t = a_t + 0,140a_{t-2} + 0,134a_{t-5}$	0,964	\sqrt{x}
Putagán en Yerbas Buenas	$Z_t = 0,144Z_{t-2} + 0,16Z_{t-3} + 0,19Z_{t-4} + a_t$	0,868	\sqrt{x}
Índice Linares-Panimávida	$Z_t = a_t$	1,0	\sqrt{x}
Bureo en Mulchén	$Z_t = 0,354Z_{t-1} + 0,169Z_{t-2} + a_t$	0,795	\sqrt{x}
Cerro El Padre	$Z_t = a_t$	1,0	\sqrt{x}
Mulchén en Mulchén	$Z_t = 0,308Z_{t-1} + 0,229Z_{t-2} + a_t$	0,796	Ln(x)
Mulchén	$Z_t = a_t$	1,0	\sqrt{x}
Tucapel en Cañete	$Z_t = 0,247Z_{t-1} + 0,236Z_{t-2} + 0,13Z_{t-4} + a_t$	0,787	\sqrt{x}
Cañete	$Z_t = a_t - 0,1a_{t-4} - 0,14a_{t-5}$	0,972	\sqrt{x}
Purén en Tranamán	$Z_t = 0,311Z_{t-1} + 0,25Z_{t-2} + 0,1Z_{t-3} + 0,107Z_{t-4} - 0,153Z_{t-5} + a_t$	0,697	\sqrt{x}
Traiguén	$Z_t = a_t + 0,1a_{t-4} - 0,137a_{t-5}$	0,972	\sqrt{x}
Chillán en Esperanza	$Z_t = 0,517Z_{t-1} + 0,154Z_{t-2} + a_t$	0,564	Ln(x)

Con respecto a los modelos identificados, vale la pena destacar que en el caso de la variable caudal, todos los modelos son del tipo

autorregresivo, exhibiendo distintas órdenes según el caso; por otra parte, los modelos de precipitación son todos de promedio móvil y más aún, varios son directamente ruidos blancos; en aquellos en que existen componentes de promedio móvil propiamente tales, estas siempre tienen parámetros de poca importancia. Esto último sugiere que el fenómeno de la precipitación mensual probablemente no difiere mucho de un ruido blanco puro; sin embargo para esta aplicación, se ha preferido dejar aquel modelo que resultaba más apropiado desde el punto de vista matemático.

Los modelos identificados fueron utilizados para generar 50 series de caudales medios mensuales y precipitaciones mensuales de 50 años de longitud cada una. Al comienzo de cada serie se generaron siempre 8 años de estadística que se eliminaban de los análisis posteriores, para minimizar la influencia de los valores iniciales arbitrarios.

En la generación de los impactos aleatorios a_t , se usó la subrutina GAUSS de IBM.

En el proceso de generación se acepta que la variable Z_t resulte con valores negativos, y sólo se adopta el procedimiento convencional de asignar de un valor nulo a la variable original, si luego del proceso de desestandarización-transformación, ésta resulta negativa.

Con el fin de detectar posibles anomalías en las series generadas, se las sometió a diversas pruebas estadísticas, tales como las señaladas por Brown y Torretti (1977): prueba de diferencias de medias (D.M), prueba F para varianzas, prueba para los coeficientes de autocorrelación (R) y prueba de Wilcoxon modificada (M.W), además de la prueba bimuestral de Kolmogorov-Smirnov (K.S) (Hollander y Wolfe 1973) que permite decidir si dos muestras independientes poseen la misma función densidad de frecuencias. En todos los casos, se adoptó un nivel de significancia de 95% para aplicar estas pruebas. Si para

algún valor mensual generado, al test de diferencia de medias u otras tres pruebas simultáneamente indicaban que éste valor no pertenecería a la misma población, para el nivel de significancia dado, se rechazaba la serie completa, generándose series adicionales que permitieran llevar a cabo los análisis de rangos para un número de 50 series.

2.3 Análisis de las Series.

Con el fin de determinar cuan diferentes resultarían los diseños de volúmenes de regulación de embalses, calculados con las series "históricas" y con las series generadas correspondientes, se sometió a estas a un análisis de su rango, definido como la máxima diferencia entre superávits y déficits con respecto a la media, acumulados. Luego, se hizo un análisis de la regulación de un embalse hipotético, planteando diversos valores de la capacidad de regulación de este y diversos valores de demanda. Los primeros fueron planteados expresados en tanto por uno del gasto medio anual de la estadística simulada, los últimos en tanto por uno de la capacidad de regulación correspondiente. La distribución mensual se adoptó como una única curva de variación estacional de la demanda. En estos aspectos se consideraron los mismos valores dados por Brown y Torretti (1977). Los déficits totales (en volumen) y los períodos de déficit (en meses), permitieron calcular los parámetros δ_D y δ_T que indican la mayor o menor severidad de las magnitudes y períodos de déficit generados, en relación a los que se obtienen con la serie histórica. Estos parámetros se definen como:

$$\delta_D = \frac{D_H - \bar{D}_G}{\sigma_D} \quad \delta_T = \frac{T_H - \bar{T}_G}{\sigma_T} \quad (30)$$

en que D_H : déficit total calculado según la serie histórica.

\bar{D}_G : promedio de los déficits generados para el mismo río.

T_H : Duración total de período de déficit histórico.

\bar{T}_G : Duración promedio de los períodos de déficit en las series generadas.

σ_D y σ_T : Desviaciones estándar de magnitudes y períodos de déficit generados respectivamente.

3. RESULTADOS

3.1 Características de las series generadas

En las Figuras 1 y 2 se muestran las características principales de las series generadas para dos de las estadísticas analizadas, que son representativas del comportamiento general observado en todas ellas. Los resultados respaldan las conclusiones anteriores de los autores (Vargas y Brown 1985), en cuanto a que para reproducir en forma apropiada los estadígrafos históricos de las series de caudales, resulta más atractiva la generación de series de precipitaciones, para que éstas sean utilizadas como entrada a un modelo de simulación de caudales. En esta forma, se preservan también mejor los coeficientes de correlación mensuales y estacional, observándose esto último en la Tabla 3.

En cuanto a las pruebas estadísticas, se observa en la Tabla 4, que el test de autocorrelaciones es el que provoca un mayor porcentaje de rechazos en el caso de generación directa, mientras que cuando se usa el modelo de simulación con precipitaciones generadas, es la prueba de varianzas (test F) la que motiva un mayor número de rechazos, aunque estos no son sustancialmente superiores a los del primer método. Al comparar la generación directa con los resultados del modelo de Markov de orden uno planteado para cada mes por Brown y Torretti (1977) para la estadística del río Chillán en Esperanza, se observa que al menos el test F indica un porcentaje inferior de rechazos con la metodología aquí planteada, sin embargo el test de autocorrelación indica una clara superioridad del modelo de Markov en la reproducción de este estadígrafo.

3.2 Análisis de Rangos, magnitudes y duración de déficits.

En la Tabla 5 se presentan los resultados obtenidos del análisis de rangos. Se observa que el rango histórico (R_H) se sitúa aleatoriamente sobre o bajo el rango promedio generado (R_G), los

Figura 1.
Purén en Tranamán.
CARACTERISTICAS SERIES GENERADAS

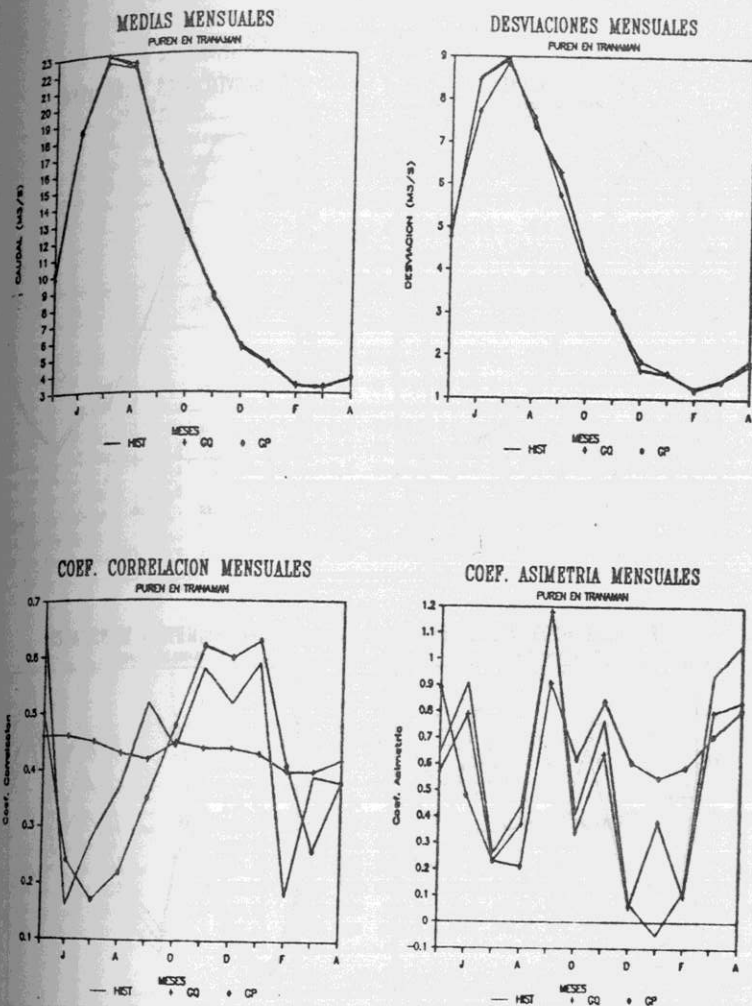


Figura 2.
Bureo en Mulchén
CARACTERISTICAS SERIES GENERADAS

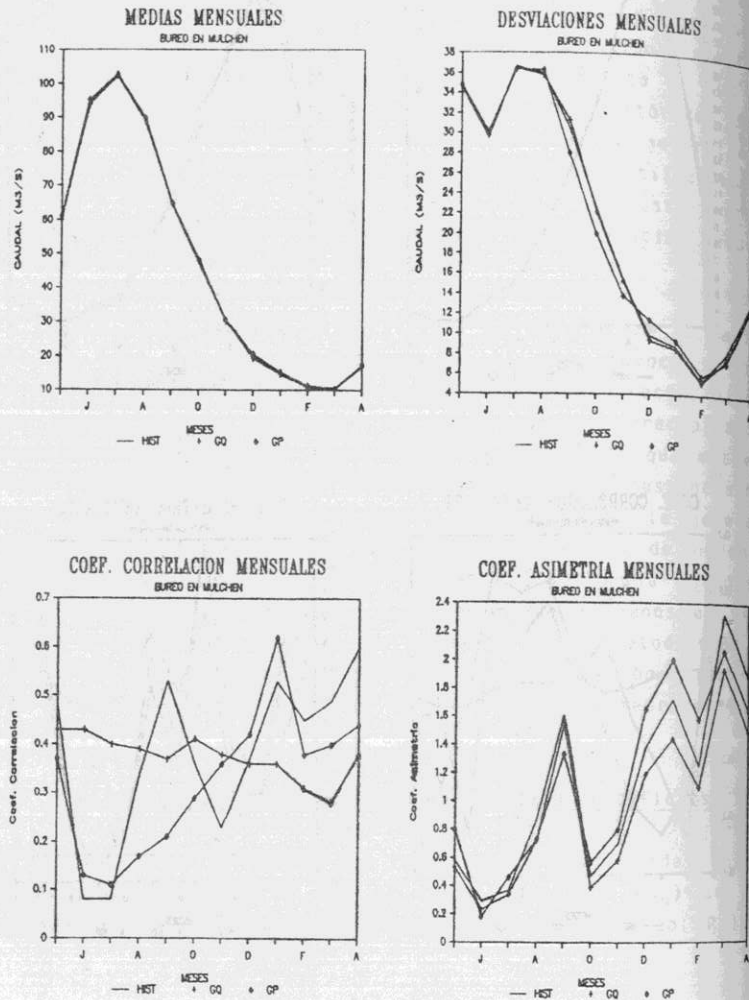


TABLA 3
COEFICIENTES DE CORRELACION ESTACIONAL

ESTADISTICA	r estacional		
	HIST.	GEN. DIRECTA	GEN. INDIRECTA
Purapel en Nirivilo	0,64	0,37	0,65
Putagán en Y. Buenas	0,66	0,33	0,72
Bureo en Mulchén	0,15	0,23	0,17
Mulchén en Mulchén	0,21	0,23	0,13
Tucapel en Cañete	0,70	0,40	0,52
Purén en Tranamá	0,61	0,33	0,58
Chillán en Esperanza	0,56	0,36	0,21

TABLA 4
RECHAZOS EN PRUEBAS ESTADISTICAS

ESTADISTICA	GENERACION DIRECTA					GENERACION INDIRECTA													
	D.M.		F		K.S.	M.W.		R		D.M.		F		K.S.	M.W.		R		
	#	%	#	%	#	#	%	#	%	#	%	#	%	#	%	#	%		
Purapel en Nirivilo	-	-	34	5,7	-	-	10,2	35	5,8	10,2	61	10,2	2	0,3	-	-	14	2,3	
Putagán en Y. Buenas	1	0,2	8	1,3	5	0,8	10,2	52	8,7	10,2	25	4,2	17	2,8	10,2	38	6,3		
Bureo en Mulchén	1	0,2	32	5,3	7	1,2	40,7	61	10,2	-	-	56	9,3	9	1,5	2	0,3	38	6,3
Mulchén en Mulchén	3	0,5	43	7,2	2	0,3	40,7	21	3,5	10,2	47	7,8	8	1,3	3	0,5	12	2,0	
Tucapel en Cañete	1	0,2	30	5,0	9	1,5	10,2	181	30,2	-	-	39	6,5	2	0,3	-	-	12	2,0
Purén en Tranamá	1	0,2	9	1,5	2	0,3	10,2	17	2,8	10,2	16	2,7	4	0,7	10,2	5	0,8		
Chillán en Esperanza	1	0,2	44	7,3	25	4,2	9	1,5	120	20,0	-	-	130	12,0	-	-	11	1,0	

TABLA 5
ANALISIS DE RANGOS GENERADOS

ESTADISTICA	Qh		Rh		GENERACION DIRECTA					GENERACION INDIRECTA				
	m3/s	m3	mill. m3	mill. m3	Rg	r _R	Δ _R	Rh-Rg	r	Rg	r _R	Δ _R	Rh-Rg	r
Purapel en Nirivilo	1,8	136,4	259,6	64	311,0	-1,94	253,8	64	330,7	-1,83				
Putagán en Y. Buenas	21,5	980,1	1627,5	397	2141,9	-1,62	1695,6	432	1918,8	-1,66				
Bureo en Mulchén	47,1	3408,3	3282,5	790	4075,4	0,16	2844,0	562	2431,9	1,00				
Mulchén en Mulchén	22,3	2102,1	1850,6	444	2192,9	0,57	1508,7	282	1204,2	2,11				
Tucapel en Cañete	11,6	878,2	968,8	236	1260,1	-0,38	818,0	172	964,4	0,35				
Purén en Tranamá	11,0	520,1	763,4	180	941,2	-1,35	678,2	157	912,6	-1,09				
Chillán en Esperanza	15,3	683,6	1537,0	359	1788,5	-1,88	740,0	189	1015,5	-0,46				

† Indica que en generación indirecta se han colocado los valores correspondientes a la generación con modelos de Markov de orden uno (Brown y Torretti 1977).

que presentan amplitudes (D_R) mayores cuando se generan directamente los caudales con un modelo ARMA, aún cuando los valores medios y las desviaciones (σ_R) son semejantes.

En la Tabla 6 se muestran los resultados del análisis mediante los parámetros δ_D y δ_T definidos por la ecuación (3). En ella se indican los valores de los parámetros para cada volumen de regulación del embalse (α) en tanto por uno del caudal medio anual y, para cada demanda (β) en tanto por uno del volumen de regulación. Se observa en la Tabla que para demandas menores o iguales que la mitad del caudal medio anual, sólo en Purapel en Nirivilo que posee el valor más bajo, se presentan períodos de déficit. Estos en general son menos severos en magnitud y con duraciones mayores que los correspondientes valores históricos. Se aprecia también que en la generación directa, en general las magnitudes de déficit son menores que los históricos, al igual que lo indicado por otros investigadores (Askew et al 1970, 1971) para modelos Markovianos; sin embargo las duraciones de estos déficit resultan en general mayores que los correspondientes históricos.

Con la generación indirecta en cambio, se observa que pueden producirse en algunos casos magnitudes de déficit bastante más severas que las históricas, existiendo una leve tendencia a que las duraciones de estos períodos sean también mayores que las históricas.

4. CONCLUSIONES

Tanto los resultados del análisis de las series, como los resultados de la utilización de éstas para determinar eventuales volúmenes de regulación de embalses, indican que la generación indirecta, vale decir generando series de precipitaciones que sirven como entrada a una relación precipitación-escorrentía, es más adecuada como procedimiento, que generar directamente series de caudales.

TABLA 6
ANÁLISIS DE MAGNITUDES Y DURACIONES DE DÉFICIT

α		β		ESTADÍSTICA															
				Purapel en Nirivilo				Putagán en Y. Buenas				Bureo en Mulchén				Mulchén en Mulchén			
				Gen.Dir.		Gen.Ind.		Gen.Dir.		Gen.Ind.		Gen.Dir.		Gen.Ind.		Gen.Dir.		Gen.Ind.	
δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T				
0.25	1.0	0.46	-0.51	0.35	-0.38														
	1.5	-0.39	-0.29	-1.22	0.53														
0.50	1.0	0.13	-0.33	-0.30	0.02														
	1.5	-1.48	-0.26	-1.97	0.02	0.39	-0.16	-0.11	0.52	0.66	-0.59	-0.02	0.29	0.29	-0.44	-0.84	0.59		
0.75	0.5	0.15	-0.18																
	1.0	0.39	-0.66	0.09	-0.49	0.28	-0.32	0.23	-0.25	0.29	-0.31	0.14	-0.14	0.46	-0.51	0.41	-0.79		
1.00	0.5	0.26	-0.29	0.28	-0.29														
	1.0	-2.56	-0.05	-2.60	-0.03	-0.05	-0.72	-1.12	-0.09	0.65	0.32	-0.04	0.79	0.06	0.42	-0.50	0.90		
1.50	0.5	0.72	-0.66	0.66	-0.52														
	2.00	1.06	-1.24	0.95	-1.14	1.11	-0.90	0.69	-0.52	0.57	-0.65	0.15	-0.24	-0.64	0.10	-0.94	0.76		

α		β		ESTADÍSTICA											
				Tucapel en Canete				Purén en Tranahán				Chillán en Esperanza			
				Gen.Dir.		Gen.Ind.		Gen.Dir.		Gen.Ind.		Gen.Dir.		Gen.Ind.	
δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T	δ_D	δ_T				
0.50	1.50	1.00	-0.99	0.40	-0.42	0.74	-0.72	0.49	-0.58	1.04	0.98	-0.28	0.22		
	0.75	1.00	0.49	-0.47	0.30	-0.33	0.25	-0.25	0.14	-0.14	0.67	-0.62	-0.20	0.21	
1.00	1.00	-1.30	0.34	-1.71	0.77	0.61	-0.51	-0.13	0.08	-0.83	-0.76	-0.68	0.51		
	2.00	0.50	0.12	-0.58	-0.40	-0.26	1.02	-0.79	0.58	-0.46	1.28	-1.09	-0.71	0.36	

Lo anterior sugiere la necesidad de tener desarrolladas relaciones precipitación-escorrentía o modelos simulación hidrológica, que sean capaces de simular el comportamiento histórico de diferentes tipos de cuencas hidrográficas, de tal modo de poder ampliar éstas, si se quiere generar series sintéticas de caudales.

REFERENCIAS BIBLIOGRAFICAS

- Askew A.J., W.W-G. Yeh y W.A.Hall. 1970 "Stream flow Generating Techniques: A comparison of Their Abilities to Simulate Critical Periods of Drought". Water Resources Center Contribution N°131, University of California, Los Angeles.
- Askew A.J., W.W-G. Yeh y W.A. Hall. 1971 "A Comparative Study of Critical Drought Simulation". Water Resources Research, Vol 7 N°1.
- Brown E. y E. Torretti. 1977 "Algunas Consideraciones Sobre la Aplicación del Modelo de Markov en Generación Estocástica de Caudales". III Coloquio Nacional de Ingeniería Hidráulica. Sociedad Chilena de Ingeniería Hidráulica. Santiago pp. 375-403.
- Ferrer P., E. Brown y L. Ayala. 1973. "Simulación de Gastos Medios Mensuales en una Cuenca Pluvial". II Coloquio Nacional de Ingeniería Hidráulica. Sociedad Chilena de Ingeniería Hidráulica. Santiago. Vol.2 pp. 3.49-3.65.
- Hollander M. y D. Wolfe. 1973. "Non Parametric Statistical Methods". Wiley Interscience.
- Vargas X., E. Brown y J.P. Salas. 1983. "Modelos ARIMA. El Proceso de Identificación". VI Congreso Nacional de Ingeniería Hidráulica. Sociedad Chilena de Ingeniería Hidráulica. Santiago. Vol. 2 pp. 513-530.

- Vargas X., E. Brown y P. Parot. 1984. "Análisis del Proceso de Selección de Modelos ARIMA para Series de Gastos Medios Mensuales". XI Congreso Latinoamericano de Hidráulica. Buenos Aires, Argentina. Vol 1 pp. 196-206.
- Vargas X. y E. Brown. 1985. "Análisis de Procedimientos Alternativos para generar Series de Caudales Usando Modelos ARMA". V Congreso de Metodologías. Santiago. pp. 35-51.

RECONOCIMIENTO

Los autores desean expresar su reconocimiento al Departamento de Investigación y Bibliotecas de la Universidad de Chile, por el financiamiento parcial de este trabajo a través del proyecto DIB N° 2461.