

SOCIEDAD CHILENA DE INGENIERÍA HIDRÁULICA

XVIII CONGRESO DE INGENIERÍA HIDRÁULICA

LA INCERTIDUMBRE EN LA MODELACIÓN MULTIVARIADA DE SERIES ESTOCÁSTICAS DE CAUDALES EN ZONAS CON ESCASA INFORMACIÓN.

**MARTÍN VALENZUELA S.¹
XIMENA VARGAS M.²**

RESUMEN

Se estudia la incertidumbre estadística asociada a modelos multivariados de generación de caudales en lugares con escasa información. La aplicación se desarrolla para las cuencas que constituyen las tres principales fuentes superficiales del sistema de abastecimiento de agua potable de la ciudad de Punta Arenas, Chile. Se utilizan dos modelos estocásticos, ARMAX y CARMA, los que se calibran a base de la escasa información disponible, obteniéndose intervalos de confianza asociados a la estimación de cada parámetro. Se generan 100 series sintéticas de caudales medios mensuales en cada punto de interés, utilizando distintos valores de los parámetros escogidos aleatoriamente al interior de sus intervalos de confianza. Se evalúa el desempeño de los modelos a través de la reproducción de los estadísticos y de la estructura de interdependencia de las series históricas, y se compara la incertidumbre en la generación de las series sintéticas mediante los anchos de los intervalos de confianza de los caudales asociados a probabilidades de excedencia de 50% y 90%. Con ambos modelos se obtienen similares resultados, de lo que se concluye que el mejor modelo, en este caso, es el CARMA, ya que su aplicación es mucho más sencilla.

¹ Ingeniero Civil Hidráulico, DOH Región de Los Lagos, MOP – mail: martin.valenzuela@mop.gov.cl

² Profesor Asociado, Departamento de Ingeniería Civil, Universidad de Chile – mail: xvargas@ing.uchile.cl

1. INTRODUCCIÓN

Las series de valores de parámetros hidrológicos pueden considerarse como una muestra finita de las infinitas realizaciones posibles de un proceso estocástico, ya que están conformadas por variables aleatorias que cambian en el tiempo. Por lo tanto, los estadísticos de las muestras son también variables aleatorias, cuya incertidumbre es necesario tener en cuenta para la modelación.

Los principales estadísticos para la modelación estocástica de series hidrológicas son la media, desviación estándar, coeficiente de asimetría y autocorrelación. Los dos primeros son los menos inciertos, por lo que son razonablemente reproducidos por todos los modelos. El coeficiente de asimetría muestral tiene asociada una alta incertidumbre, y el que un modelo sea capaz de reproducirla depende principalmente de la longitud de la muestra. La autocorrelación también está sujeta a una gran incertidumbre, especialmente en muestras breves (Salas et al., 1980). Una vez que se ha determinado qué estadísticos se desea preservar con el modelo, debe escogerse aquel que lo permite utilizando el menor número de parámetros que sea posible, lo que se conoce como el principio de parsimonia. Lo ideal es que el cociente entre el tamaño muestral y el número de parámetros del modelo sea, al menos, de 15 (Salas et al., 1980).

Las n series que componen una serie multivariada pueden analizarse por separado, considerando que las componentes estocásticas de las series son variables aleatorias mutuamente dependientes. Por lo tanto, al generar las series, no sólo interesa preservar las características estadísticas, sino también su estructura de dependencia mutua. Así, un estadístico de importancia es la correlación cruzada entre las series, con un desfase k , r_k . Por ejemplo, si se tienen dos series, X_t^1 y X_t^2 :

$$r_k^{12} = \frac{\sum_{t=1}^{N-k} (X_t^1 - \bar{X}_t^1)(X_t^2 - \bar{X}_t^2)}{\sqrt{\sum_{t=1}^{N-k} (X_t^1 - \bar{X}_t^1)^2 \cdot \sum_{t=1}^{N-k} (X_{t+k}^2 - \bar{X}_{t+k}^2)^2}} \quad (1)$$

Si bien existen múltiples modelos estocásticos para series hidrológicas, todos ellos tienen ciertas técnicas estadísticas y metodologías a seguir comunes. Así, cuando un modelo tiene parámetros α y β , hay que encontrar sus estimadores, $\hat{\alpha}$ y $\hat{\beta}$. Gran parte de la teoría de la hidrología estadística ha sido desarrollada suponiendo que las variables se distribuyen normalmente, por lo cual muchas veces es necesario transformar las series de modo de cumplir con dicha condición, lo cual se puede hacer centrando la serie (restando la media a todos los valores) o estandarizándola.

Gran parte de los modelos estocásticos se basan en los modelos autorregresivos (AR). A partir de éstos se han derivado modelos autorregresivos con componentes de media móvil (ARMA), variaciones de éstos para la modelación de series periódicas (PAR y PARMA), multivariadas, multivariadas contemporáneas (CAR y CARMA) y aquellos que incluyen variables exógenas (ARMAX).

El modelo ARMA, del que derivan los modelos ARMAX y CARMA, considera que el valor actual de una variable aleatoria es una combinación lineal de sus p anteriores valores, más un residuo aleatorio y una combinación lineal de los q anteriores valores del residuo. Matemáticamente, el modelo ARMA(p,q) se expresa como:

$$Z_t = \sum_{i=1}^p f_i \cdot Z_{t-i} + e_t - \sum_{j=1}^q q_j \cdot e_{t-j} \quad (2)$$

En que se tienen p parámetros autorregresivos ϕ_i y q parámetros de media móvil θ_j , que ponderan a los residuos ε_t .

El modelo ARMAX (autorregresivo con media móvil y variable exógena) considera lo mismo que el modelo ARMA, pero agregando una combinación lineal de los valores previos de variables exógenas, con cierto desfase (Zavala, 1987), por lo que puede ser considerado como un modelo de función de transferencia (FT) (Box y Jenkins, 1976), al que se le ha agregado una componente estocástica, quedando como un modelo de transferencia – ruido (Naim, 1989). Aplicándolo a la modelación de caudales, los retardos se deben a que, dependiendo del tipo de cuenca, la influencia de variables exógenas como la precipitación o la temperatura, se manifiesta en un tiempo posterior. Matemáticamente, se expresa como:

$$Z_t = \sum_{i=1}^r d_i \cdot Z_{t-i} + \sum_{i=1}^m v_{i,0} \cdot P_{i,t-b_i} - \sum_{i=1}^m \sum_{j=1}^{s_i} v_{i,j} \cdot P_{i,t-b_i-j} + \sum_{i=1}^p f_i \cdot n_{t-i} + a_t - \sum_{i=1}^q q_i \cdot a_{t-i} \quad (3)$$

En que se tienen r parámetros autorregresivos δ_i y m variables exógenas P_i , cada una ponderada por s_i parámetros $\omega_{i,j}$ con un desfase b , y los residuos n se obtienen con un modelo ARMA(p,q) con ruidos aleatorios a_t . El modelo se calibra con el algoritmo IVAML (variables instrumentales por aproximación de máxima verosimilitud) (Moore y O'Connell, 1978).

El modelo CARMA (Autorregresivo con media móvil y contemporáneo) permite modelar n series a través de n modelos ARMA en los que las n series de residuos no están autocorrelacionadas, pero sí contemporáneamente correlacionadas entre ellas. Así, el modelo corresponde a una simplificación de los modelos multivariados comunes, ya que considera que las matrices de parámetros son diagonales (Matalas, 1967; Salas et al., 1980; Camacho et al., 1985). Matemáticamente, se expresa como:

$$Z_t = \sum_{j=1}^p a_j \cdot Z_{t-j} + e_t - \sum_{j=1}^q c_j \cdot e_{t-j}; \quad i = 1, \dots, n \quad (4)$$

En que los parámetros a_j y c_j corresponden a los ϕ_i y θ_i del modelo ARMA, y por lo tanto se estiman de la misma manera. Los residuos ε_t se modelan como:

$$\mathbf{e}_t = \mathbf{B} \cdot \mathbf{x}_t \quad (5)$$

Donde la serie de residuos ξ_t es independiente y normalmente distribuida y \mathbf{B} es una matriz que es solución de la siguiente ecuación matricial:

$$\hat{\mathbf{B}}\hat{\mathbf{B}}^T = \hat{\mathbf{G}} \quad (6)$$

Donde \mathbf{G} es la matriz de varianza-covarianza de los residuos ϵ_t (Salas et al., 1980), que es simétrica y semidefinida positiva, lo que permite que la matriz \mathbf{B} sea única y triangular inferior (Haltner y Salas, 1988). De este modo, la estimación de \mathbf{G} se basa en los valores de los estimadores de los parámetros a y c .

2. INCERTIDUMBRE EN MODELOS ESTOCÁSTICOS

En la utilización de modelos, existen básicamente dos tipos de incertidumbre, la que es inherente al modelo, ya que no se conocen modelos reales de las series hidrológicas y los modelos escogidos son sólo aproximaciones; y la otra que corresponde a los parámetros, ya que éstos se estiman en base a una cantidad limitada de datos disponibles (Salas et al., 1980).

El análisis de la incertidumbre se enfoca a encontrar expresiones que permitan evaluar la precisión con que se puede generar caudales medios mensuales sintéticos dado el nivel de información existente. Por lo tanto, la idea no es cuantificar el error que se comete en los cálculos al modelar la estructura de dependencia espacio-temporal de las diversas fuentes de un sistema complejo (Harambour y Vargas, 2002), sino determinar la precisión con que estos modelos reproducen lo observado.

El intervalo de confianza para la media en ambos modelos es (Salas et al., 1980):

$$\left[\bar{Z} - s(\bar{Z}) \cdot t(N-1)_{1-\frac{\alpha}{2}}, \bar{Z} + s(\bar{Z}) \cdot t(N-1)_{1-\frac{\alpha}{2}} \right] \quad (7)$$

En que s es la desviación estándar y $t(N-1)_{1-\frac{\alpha}{2}}$ es el cuantil con nivel de confianza $\left(1 - \frac{\alpha}{2}\right)$ de la distribución t de Student, con $(N-1)$ grados de libertad.

La teoría de la estimación estadística indica que, en muestras grandes, los estimadores de máxima verosimilitud de los parámetros se distribuyen en forma normal, con media igual al valor verdadero del parámetro (Box y Jenkins, 1976; Stedinger et al., 1993). Por lo tanto, es necesario encontrar expresiones para las varianzas de los parámetros.

Para el modelo ARMAX, entonces, basta con determinar intervalos de confianza para los parámetros δ_i , ω_i , ϕ_i y θ_i , lo que se reduce a encontrar sus respectivas varianzas, las que se obtienen en la última iteración del algoritmo IVAML con que se calibran los modelos.

Para el modelo CARMA, la incertidumbre de los parámetros autorregresivos y de media móvil se obtiene del mismo modo que en los modelos ARMAX, y la de los componentes de la matriz B se transfieren a través de la incertidumbre de los mismos parámetros, y asimismo a los de los residuos ε_t (Salas, 2004).

3. DESCRIPCIÓN DE LA ZONA EN ESTUDIO

La zona de estudio corresponde a tres cuencas hidrográficas ubicadas en la vertiente oriental de la Península de Brunswick, cuyo referente es la ciudad de Punta Arenas. Punta Arenas se encuentra ubicada en los 53°08' de latitud sur y en los 70°55' de longitud oeste, en el margen occidental del Estrecho de Magallanes, ocupando una franja costera de 3,5 por 8 [Km] de largo. Es la capital de la XII Región de Magallanes y la Antártica Chilena, y cuenta en la actualidad con 121.533 habitantes, según el censo del año 2002.

El clima de la zona es del tipo templado frío lluvioso, sin estación seca. La precipitación media anual en Punta Arenas es de 492,5 [mm], los que se reparten en forma relativamente uniforme en el año. La temperatura media anual es de 6,4 [°C], las máximas y mínimas anuales alcanzan los 25 y -16 [°C]. El clima está condicionado por la interacción de masas de aire frío del polo y cálido del norte (Harambour, 2001). En el invierno, el frente polar se desplaza hacia una zona comprendida entre las latitudes 30° y 45° Sur, en cambio, durante el verano, el frente se instala entre los 40° y 50°, quedando hacia el norte la masa de aire subtropical y hacia el sur la polar, lo que produce una gran inestabilidad por efecto del gradiente barométrico, originándose fuertes vientos en primavera y verano, cuyo valor medio es del orden de los 25 [Km/h], pudiendo alcanzar rachas de hasta 150 [Km/h] en la dirección oeste (Zamora y Santana, 1979).

El relieve de la zona está dominado, al este, por la Cordillera Patagónica Occidental (prolongación de la Cordillera de Los Andes), y al oeste por la Planicie de la Estepa Fría Magallánica (pampas patagónicas que se extienden hasta la costa atlántica). Es debido a esto que existe una gran variación de la precipitación en el sentido oeste-este, pasando de magnitudes del orden de los 5.000 [mm] en el sector occidental del estrecho hasta precipitaciones bajas en el sector oriental, del orden de los 200 [mm] (Zamora y Santana, 1979). En invierno, las precipitaciones son en forma de lluvia, nieve y granizo, y la nieve que se acumula en las partes altas de las cuencas generalmente comienza a derretirse en septiembre (Harambour, 2001).

Las 3 cuencas que se muestran en la figura 1, son las de los ríos Las Minas, Leñadura y Tres Brazos (en orden de preferencia de su aprovechamiento para agua potable), que constituyen las 3 principales fuentes fluviales del sistema de abastecimiento de agua potable de la ciudad de Punta Arenas.

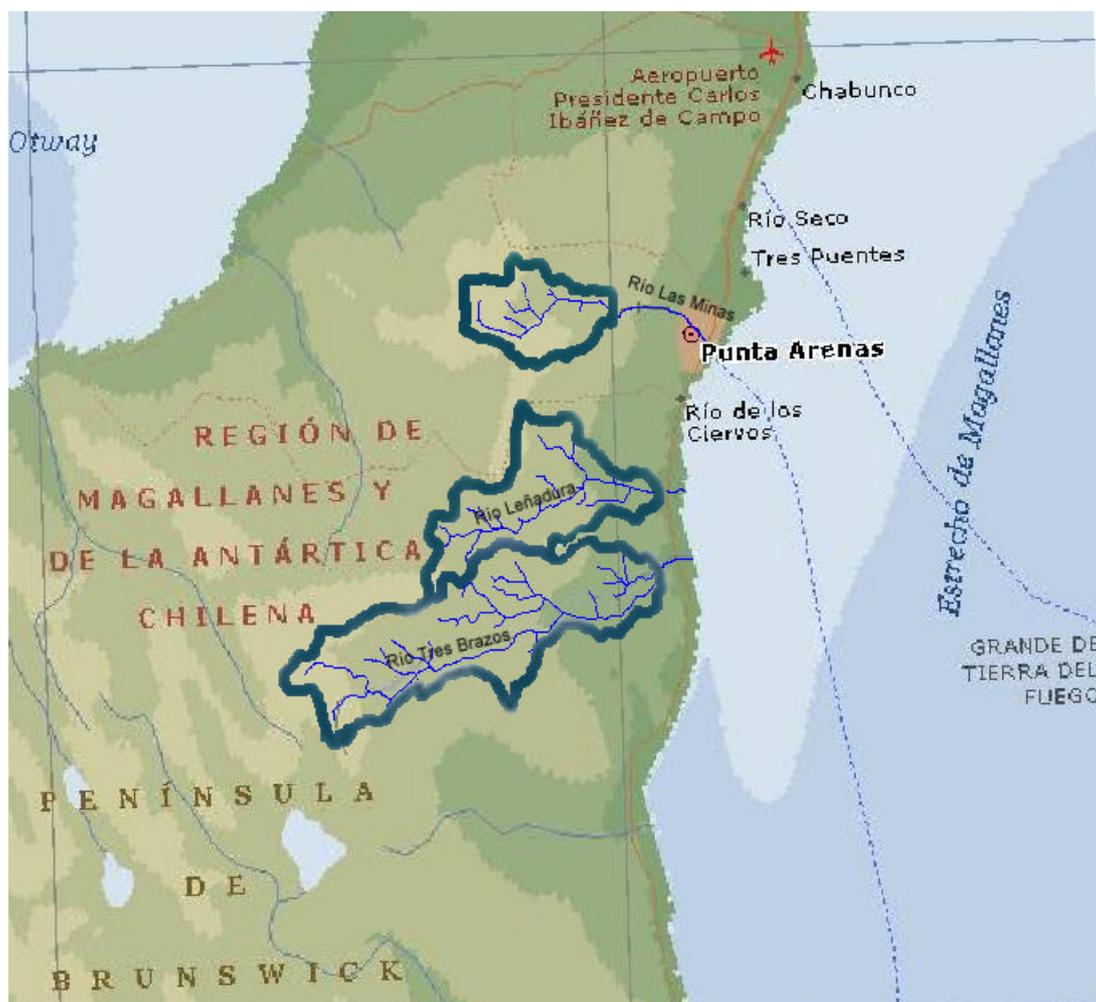


Figura N° 1: Ubicación de las Cuencas en Estudio en la Península de Brunswick

Las estaciones de control fluviométrico de las cuencas en estudio, así como su área aportante, altura e inicio y término de la estadística de caudales medios mensuales disponibles, se muestran en la Tabla N° 1.

La Figura N° 2 muestra los caudales medios mensuales de los tres ríos en estudio.

Tabla N° 1: Estaciones Fluviométricas.

Estación	Área [Km ²]	Altura [m.s.n.m.]	Fecha Inicio	Fecha Término
Las Minas en BT SENDOS	33,6	200	Junio 1982	Febrero 2003
Leñadura antes de BT SENDOS	60,8	30	Abril 1982	Diciembre 1991
Tres Brazos antes de BT SENDOS	113,6	21	Junio 1982	Febrero 2003

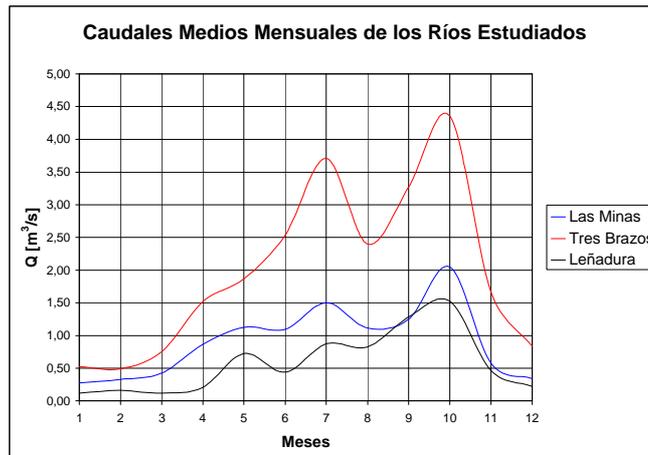


Figura N°2: Caudales Medios Mensuales Ríos Las Minas, Leñadura y Tres Brazos.

Del gráfico se desprende que los tres ríos presentan un régimen hidrológico similar, con caudales máximos en octubre, correspondientes al deshielo, y un segundo máximo en julio.

Debido a que el modelo ARMAX considera la precipitación y temperatura como variables exógenas, es necesario recurrir a registros de éstas. Tras hacer un análisis de las estaciones existentes en la zona, y aplicando el método de los Polígonos de Thiessen, se concluyó que para las cuencas de los ríos Las Minas y Leñadura, la estación más representativa es “Laguna Lynch”, en cambio, para la cuenca del río Tres Brazos, se consideró un índice de precipitaciones con una ponderación del 52,1% para la estación “Lago Parrillar” y un 47,9% para “Laguna Lynch”. Las estaciones pluviométricas, su altura y las fechas de inicio y término de sus registros, se muestran en la Tabla N° 2. Las precipitaciones mensuales medias en “Laguna Lynch” y “Lago Parrillar” se muestran en las figuras N° 3 y N° 4.

Tabla N° 2: Estaciones Pluviométricas.

Estación	Altura [m.s.n.m.]	Fecha Inicio	Fecha Término
Laguna Lynch	210	Diciembre 1979	Marzo 2003
Lago Parrillar	335	Enero 1991	Marzo 2003

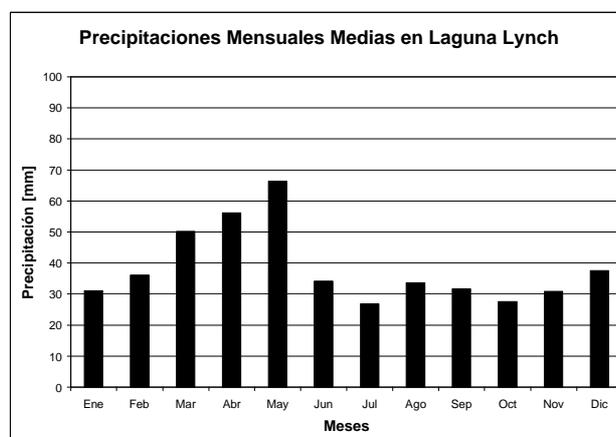


Figura N° 3: Precipitaciones Mensuales en Laguna Lynch

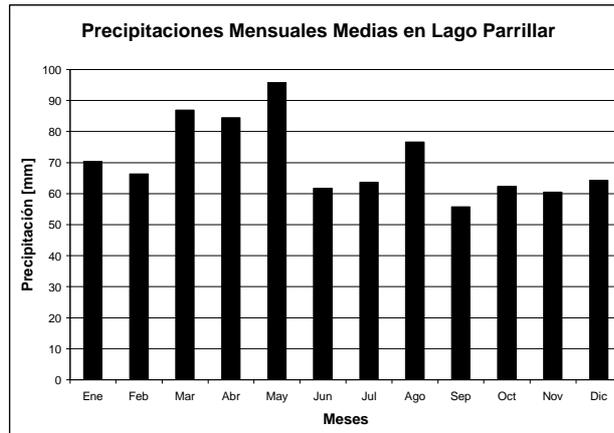


Figura N°4: Precipitaciones Medias en Lago Parrillar

Cabe destacar que la precipitación es bastante pareja a lo largo del año, siendo los meses más lluviosos los comprendidos entre marzo y mayo.

La temperatura tiene incidencia como variable causal del caudal al ser el principal factor de incidencia en los deshielos que, como quedó reflejado en la figura N° 2, son los que producen los máximos anuales de caudal medio mensual en octubre.

La temperatura media mensual en la ciudad de Punta Arenas se obtiene de la estación del mismo nombre, ubicada a una altura de 10 [m.s.n.m.] y considerando registros desde Enero de 1974 a Marzo de 2003. Ésta se muestra en la Figura N° 5.

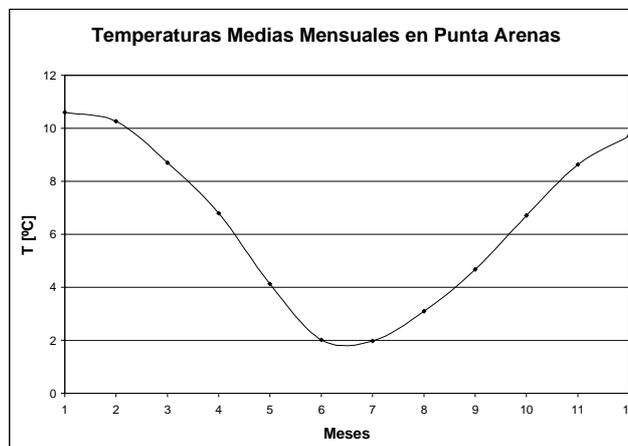


Figura N°5: Temperaturas Medias Mensuales en Punta Arenas

4. CALIBRACIÓN Y ANÁLISIS DE RESULTADOS

La utilización de modelos ARMAX requiere contar previamente con modelos ARMA calibrados para las series de temperaturas y precipitaciones centradas. De este modo, se seleccionaron modelos ARMA (0,6) y ARMA (0,4) para la precipitación en “Laguna

Lynch” y la temperatura en Punta Arenas respectivamente, los cuales cumplieron con los supuestos correspondientes. Además, con la expresión (7) se determinaron los intervalos de confianza, con un nivel del 95%, para los estimadores de los parámetros de ambos modelos.

Con estos modelos, se procede a generar 100 series sintéticas de precipitación y temperatura. Pero para que en esta generación quede representada la incertidumbre de los modelos, al generar cada serie se escogen aleatoriamente distintos valores para cada parámetro, al interior del intervalo de confianza respectivo, suponiendo una distribución normal dentro de ellos. Además, para que estas series de valores centrados representen valores de precipitaciones y temperaturas propiamente tales, es necesario realizar la transformación inversa, para lo cual se considera la incertidumbre en la media de la serie, escogiendo aleatoriamente un valor perteneciente al intervalo de confianza de ella, determinado con la expresión (7). Así, se obtienen las series, las que deben superar las pruebas pertinentes para considerar que son estadísticamente similares a la histórica.

Para obtener las series sintéticas de la estación “Lago Parrillar”, a utilizarse para el modelo ARMAX del río Tres Brazos, se recurrió a la correlación existente entre la serie histórica de esta estación y la de “Laguna Lynch”. Esto debido a que la utilización del índice de precipitaciones presupone una concurrencia entre los datos de ambas precipitaciones, la que no se da al considerar series sintéticas generadas independientemente.

El modelo ARMAX se aplicó a las series de los ríos en estudio entre junio de 1982 y diciembre de 2002 en el río Las Minas; entre abril de 1982 y diciembre de 1991 en el río Leñadura; y entre enero de 1990 y diciembre de 2002 en el río Tres Brazos, considerando la distinta disponibilidad de datos para cada uno. Los modelos escogidos presentan las siguientes estructuras:

Modelo ARMAX Río Las Minas:

$$Z_t = d_1 \cdot Z_{t-1} + w_0 \cdot P_t + y_0 \cdot T_{t-3} + h_t \quad (8)$$

Modelo ARMAX Río Leñadura:

$$Z_t = d_1 \cdot Z_{t-1} + w_0 \cdot P_t + y_0 \cdot T_{t-4} + h_t \quad (9)$$

Modelo ARMAX Río Tres Brazos:

$$Z_t = d_1 \cdot Z_{t-1} + v_0 \cdot P_t - v_1 \cdot P_{t-1} + y_0 \cdot T_{t-3} - y_1 \cdot T_{t-4} + h_t \quad (10)$$

Una vez calibrados los modelos, fue posible obtener sus respectivas series de residuos η_t , a las que se ajustaron modelos ARMA de órdenes (0,2) para el río Las Minas, (0,7) para el río Leñadura y (1,5) para el río Tres Brazos.

En la calibración de cada modelo fueron determinados también los respectivos intervalos de confianza, con un nivel del 95% para los estimadores de cada parámetro, del mismo modo que para los modelos de precipitación y temperatura. De este modo fue posible la

generación de las 100 series sintéticas de residuos. Con estas series, más las de precipitación y temperatura, y escogiendo aleatoriamente valores para los parámetros de los modelos ARMAX al interior de sus respectivos intervalos de confianza, como también para los valores de las medias, fue posible generar las 100 series sintéticas de 360 valores de caudales medios mensuales, equivalentes a 30 años.

En el caso del modelo CARMA, considerando los disímiles niveles cuantitativos de información disponibles para los 3 ríos, se decidió aplicar dos modelos CARMA. El primero, o C1, considera sólo a los ríos Las Minas y Tres Brazos, tomando en cuenta información entre junio de 1982 y febrero de 2003, es decir, 249 valores de caudal medio mensual por serie. El segundo, o C2, considera a los tres ríos, pero como este modelo requiere de la utilización de series históricas concurrentes, el registro utilizado es menor, sólo entre junio de 1982 y diciembre de 1991, es decir, 115 datos por serie.

Para calibrar los modelos CARMA, se aprovecha su estructura en base a matrices diagonales, lo que permite descomponerlos en modelos ARMA.

Para el caso C1 se ajustó un modelo ARMA (0,1) para el río Las Minas (M) y un ARMA (0,2) para el Tres Brazos (B), presentando el modelo CARMA la siguiente estructura matricial:

$$\begin{bmatrix} Y_M \\ Y_B \end{bmatrix} = \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \end{bmatrix} - \begin{bmatrix} \mathbf{q}_{1M} & 0 \\ 0 & \mathbf{q}_{1B} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \end{bmatrix}_{t-1} - \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{q}_{2B} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \end{bmatrix}_{t-2} \quad (11)$$

Para el caso C2 se ajustaron modelos ARMA (0,2) para los ríos Las Minas (M) y Tres Brazos (B) y un ARMA (0,1) para el río Leñadura (L), presentando el modelo CARMA la siguiente estructura matricial:

$$\begin{bmatrix} Y_M \\ Y_B \\ Y_L \end{bmatrix} = \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \\ \mathbf{e}_L \end{bmatrix} - \begin{bmatrix} \mathbf{q}_{1M} & 0 & 0 \\ 0 & \mathbf{q}_{1B} & 0 \\ 0 & 0 & \mathbf{q}_{1L} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \\ \mathbf{e}_L \end{bmatrix}_{t-1} - \begin{bmatrix} \mathbf{q}_{2M} & 0 & 0 \\ 0 & \mathbf{q}_{2B} & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_M \\ \mathbf{e}_B \\ \mathbf{e}_L \end{bmatrix}_{t-2} \quad (12)$$

En ambos casos, C1 y C2, los residuos ε_t están contemporáneamente correlacionados a través de la matriz B, la que es de orden 2 para C1 y de orden 3 para C2.

Luego, para la obtención de las series sintéticas, se generan primero las series de residuos ξ_t , luego se escogen aleatoriamente valores para los parámetros de media móvil al interior de los respectivos intervalos de confianza, y con ellos se encuentran las componentes de cada matriz B, con las que es posible obtener los residuos ε_t y generar las 100 series sintéticas de 360 valores de caudales medios mensuales con los modelos C1 y C2, considerando también la incertidumbre en las medias, del mismo modo que con los otros modelos.

El análisis de los resultados puede hacerse a través de las series de 360 valores de caudales medios mensuales o considerando las series de 30 caudales medios obtenidos para cada uno

de los 12 meses del año, más la serie de caudales medios anuales. En este trabajo, para evaluar ambos modelos, se estudió la capacidad de cada uno de ellos para reproducir los estadísticos observados y la estructura de interdependencia de los caudales de los ríos, además de la incertidumbre inherente a las series sintéticas generadas con cada modelo.

La Tabla N° 3 exhibe las medias y desviaciones estándar históricas y los valores medios de las series sintéticas para los modelos utilizados, considerando las series de 360 valores de caudales medios mensuales. Notar que un mismo río puede tener estadísticos históricos distintos para modelos diferentes, debido a que los períodos de información utilizados dependen de la información disponible para las otras series consideradas en cada modelo.

Tabla N° 3: Medias y Desviaciones Históricas y Sintéticas para cada Serie y Modelo.

Modelo	Río	Media Hist. [m ³ /s]	Media Sint. [m ³ /s]	Desv. Hist. [m ³ /s]	Desv. Sint. [m ³ /s]
ARMAX	Las Minas	0,92	0,95	0,77	0,76
ARMAX	Tres Brazos	2,08	2,09	1,79	1,43
ARMAX	Leñadura	0,59	0,61	0,60	0,56
C1	Las Minas	0,91	0,94	0,77	0,76
C1	Tres Brazos	2,00	2,06	1,73	1,56
C2	Las Minas	0,93	0,93	0,69	0,67
C2	Tres Brazos	1,94	1,94	1,58	1,44
C2	Leñadura	0,60	0,62	0,60	0,60

Se observa que en general la reproducción de estos estadísticos es satisfactoria con todos los modelos, constituyendo la excepción la desviación estándar de la serie de caudales medios mensuales del río Tres Brazos, la que no pudo ser correctamente reproducida ni por el modelo ARMAX ni por los dos escenarios del modelo CARMA.

Desde la óptica de las series de caudales para cada mes, los modelos ARMAX, C1 y C2 reproducen las medias satisfactoriamente, y las desviaciones estándar sólo exhiben diferencias importantes en los meses de caudales extremos (máximos y mínimos). La excepción nuevamente está en el río Tres Brazos, en el que los valores de las desviaciones estándar son sustancialmente diferentes de los observados en prácticamente todos los meses, para cada modelo.

Los valores de la correlación cruzada histórica y los promedios de las series sintéticas, correspondientes a cada modelo y par de ríos, se muestran en la Tabla N° 4:

Tabla N° 4: Correlaciones Cruzadas Históricas y Sintéticas para cada Par y Modelo.

Modelo	Par de Ríos	Correlación Hist.	Correlación Sint.
ARMAX	Las Minas – Tres Brazos	0,72	0,52
ARMAX	Las Minas – Leñadura	0,52	0,47
ARMAX	Tres Brazos – Leñadura	0,58	0,66
C1	Las Minas – Tres Brazos	0,72	0,62
C2	Las Minas – Tres Brazos	0,72	0,54
C2	Las Minas - Leñadura	0,52	0,48
C2	Tres Brazos - Leñadura	0,58	0,63

Se observa que la reproducción de este estadístico es plenamente satisfactoria sólo para los pares en los que no está involucrado el río Tres Brazos, aunque la del par “Tres Brazos – Leñadura” se comporta mejor que la del par “Las Minas – Tres Brazos”.

Para comparar el grado de incertidumbre con que cada modelo es capaz de generar caudales medios mensuales sintéticos, se utilizó el ancho del intervalo de confianza con un nivel del 90% a escala mensual. Como para cada río se tienen 100 series sintéticas de valores, primero se determinó el valor de los caudales asociados a probabilidades de excedencia del 50% y 90% (Q_{50} y Q_{90}) para cada serie de promedio mensual más la media anual. Así se obtienen conjuntos de 100 valores de Q_{50} y Q_{90} para cada mes en cada río. A estos conjuntos se les hizo un análisis de frecuencia, determinando los valores asociados a probabilidades de excedencia del 5% y 95%, obteniendo así los anchos de los intervalos de confianza buscados. Mientras menor sea el ancho, más certero será el modelo en la generación sintética de caudales medios mensuales.

Los valores medios anuales de los anchos de los intervalos de confianza para el Q_{50} y Q_{90} de cada serie y cada modelo se muestran en la tabla N° 5.

Tabla N° 5: Anchos Medios Anuales de los Intervalos de Confianza para Q_{50} y Q_{90} para cada Serie y Modelo.

Río	Modelo	Q_{50}	Q_{90}
Las Minas	ARMAX	0,31	0,28
Las Minas	C1	0,32	0,23
Las Minas	C2	0,32	0,30
Tres Brazos	ARMAX	0,31	0,48
Tres Brazos	C1	0,68	0,60
Tres Brazos	C2	0,53	0,80
Leñadura	ARMAX	0,19	0,30
Leñadura	C2	0,23	0,25

Nuevamente, sólo se aprecian diferencias sustanciales para el caso del río Tres Brazos, en el que se obtienen anchos de intervalos de confianza considerablemente menores para las series sintéticas generadas a través del modelo ARMAX.

5. CONCLUSIONES

Al observar los resultados aquí desplegados, se puede concluir que, desde el punto de vista de la reproducción que los estadísticos de las series sintéticas logran hacer de los estadísticos de las series históricas u observadas, no existen grandes diferencias entre lo obtenido con cada modelo, ya que todas las fallas están referidas al Río Tres Brazos.

Por otra parte, al observar los resultados que permiten comparar la incertidumbre en la generación sintética con los distintos modelos, el único caso que permite decir que el modelo ARMAX se comporta sustancialmente mejor que el modelo CARMA, es el que corresponde a la modelación de los caudales medios mensuales del río Tres Brazos.

¿Es suficiente eso para concluir que, a la luz de los resultados de esta investigación, el modelo ARMAX permite hacer una modelación estocástica multivariada más precisa que la que se obtiene con el modelo CARMA? La tentación es grande, pero la respuesta es que no, ya que no es posible perder de vista lo sustancial que debe exigírsele a un modelo, que es la reproducción satisfactoria de las características de las series históricas. Y en este caso, los modelos ajustados a las series de caudales medios mensuales del río Tres Brazos no lo lograron de forma satisfactoria, independientemente de las causas de ello. De todos modos, es interesante recalcar que, en los procesos de calibración mediante el algoritmo IVAML, las varianzas de los residuos de este río, siempre fueron del orden del doble de las de los otros dos ríos.

Hecho este alcance, se puede concluir que, para los ríos cuyos modelos estocásticos sí lograron reproducir satisfactoriamente los estadísticos de las series observadas, no existen diferencias considerables entre el grado de incertidumbre con el que cada modelo es capaz de generar series sintéticas de caudales medios mensuales.

La interrogante que surge entonces es ¿Si ambos modelos, ARMAX y CARMA, son igualmente imprecisos al generar series sintéticas, cuál de ellos debo usar al modelar? La respuesta está condicionada por el nivel de complejidad asociado a la utilización de cada modelo. Para usar el modelo ARMAX, es necesario recurrir previamente a la modelación de series de variables exógenas, que son las que le otorgan a este modelo el carácter de multivariado, al considerar que la estructura de interdependencia se mantiene debido a la correlación existente entre los mecanismos de generación física de la escorrentía. Por otro lado, para utilizar el modelo CARMA sólo se necesitan las series de caudales medios mensuales, ya que el modelo considera que toda la información relevante para la mantención de la estructura de correlaciones de las series está ya incluida en las mismas series de valores. Por lo tanto, es claro que el modelo CARMA es más conveniente de utilizar al requerir la modelación multivariada de series, la que es necesaria en sistemas complejos, como lo puede ser la modelación de la oferta de agua para abastecer a una población que recurre a diversas fuentes superficiales, como es el caso de varias ciudades.

De todos modos es necesario desarrollar mayor investigación en este ámbito, ya que este trabajo se aplicó en una zona de características especiales, en donde la información es muy escasa y donde la gran variabilidad espacio-temporal del clima dificulta enormemente la calibración de los modelos.

Resultaría de interés la aplicación de investigaciones similares a ésta en sistemas de ríos con otras características climáticas o de disponibilidad de información. Otra opción radica en la utilización de modelos periódicos, lo cual sólo es posible en cuencas con mayor nivel de información, de modo de no violar el principio de parsimonia.

REFERENCIAS

- Box, G. y Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden – Day.
- Camacho, F.; McLeod, I. y Hipel, K. (1985). Contemporaneous Autorregresive – Moving Average (CARMA) Modeling in Water Resources. *Water Resources Bulletin*, Vol 21, 4, 709 – 720.
- Haltmier, J.P. y Salas, J.D. (1988). Development and Testing of a Multivariate, Seasonal ARMA (1,1) Model. *Journal of Hydrology*, 104, 247 – 272.
- Harambour, F. (2001). Determinación de Caudales de Estiaje en Fuentes de Abastecimiento de Agua de Punta Arenas. ESMAG.
- Harambour, F. y Vargas, X. (2002). Evaluación de la Disponibilidad de Agua en Sistemas Hidrológicos Complejos. XX Congreso Latinoamericano de Hidráulica, La Habana, Cuba.
- Matalas, N.C. (1967). Mathematical Assesment of Synthetic Hydrology. *Water Resources Research*, Vol 3, 4, 937 - 945.
- Moore, R.J. y O'Connell, P.E. (1978). Real-Time Forecasting of Flood Events Using Transfer Function Noise Models. Institute of Hydrology, Wallingford, Oxon, UK.
- Naim, A. (1989). Un Modelo Estocástico de Simulación de Caudales Medios Mensuales. Memoria para Optar al Título de Ingeniero Civil. Universidad de Chile.
- Salas, J.D.; Delleur, J.W.; Yevjevich, V. y Lane, W.L. (1980). *Applied Modeling of Hydrologic Time Series*. Water Resources Publications.
- Salas J.D. (2004). Comunicación personal.
- Stedinger, J.R.; Lettenmaier, D.P. y Vogel, R.M. (1985). Multisite ARMA (1,1) and Disaggregation Models for Annual Streamflow Generation. *Water Resources Research*, vol. 21, 4, 497 – 509.
- Zamora, E. y Santana, A. (1979). Características Climáticas de la Costa Occidental de la Patagonia entre las Latitudes 46°40' y 56°30' S. Apartado de los Anales del Instituto de la Patagonia.
- Zavala, H. (1987). El Uso del Filtro de Kalman en el Rastreo de Crecidas en Tiempo Real. Memoria para Optar al Título de Ingeniero Civil. Universidad de Chile.