ta. Sin embargo, el proceso de tomar decisiones de correcciones requiere de apreciaciones subjetivas, al igual que el método de las CDA; si se pretende adoptar criterios muy automáticos de corrección, el método tiende a exagerar la proporcionalidad de los valores, haciendo perder a estos su aleatoriedad inherente.

- En general la tendencia que exhiba el vector regional resultante, reflejará aquella predominante en la mayoría de las estaciones en cada intervalo de tiempo. En este sentido no es factible distinguir entre tendencias mas "antiguas" y mas "recientes", lo que puede ser inconveniente.
- Según lo expresado en la conclusión anterior, si una mayoría de los valores de la matriz original pertenecen a tendencias "no deseables", el vector regional resultante puede ser poco representativo como patrón de comparación. En consecuencia debería compararse siempre el vector regional resultante para una zona, con el correspondiente a zonas vecinas, antes de aceptarse como patrón de diagnóstico.

## REFERENCIAS BIBLIOGRAFICAS


1. Benítez A., E. Brown, P. Ferrer, R. Feliú y A. Bancel. "Ampliación de Estadísticas Pluviométricas y Fluviométricas". la Mesa Redonda de Hidrología. Soc. Chilena de Ingeniería Hidráulica. 1974.
2. Brown E. e I. Merillán. "Homogeneidad de Series de Estadísticas Pluviométricas Usando Métodos Computacionales". Centro de Recursos Hidráulicos, Depto. de Ing. Civil, U. de Chile. 1983.
3. Hiez G. "L'homogénéite des Données Pluviométriques". Cahiers O. R.S.T.O.M. Sér. Hidrologie Vol. XIV, N° 2, 1977.
4. Searcy J.K. & C.H. Hardison. "Double Mass Curves". U.S. Geological Survey, Water-Supply Paper 1541-B, 1960.
5. Vicuña E. "Estudio de la Plvuiometría en la Zona Central de Chile". Memoria para optar al Título de Ing. Civil. U. de Chile, 1968.


## RECONOCIMIENTO

Este trabajo ha contado con financiamiento parcial a través del Proyecto de Investigación E 1531 del Departamento de Desarrollo de la Investigación, de la Universidad de Chile.



SOCIEDAD CHILENA DE INGENIERIA HIDRAULICA

VI CONGRESO NACIONAL

# MODELOS ARIMA. EL PROCESO DE IDENTIFICACION

Ximena Vargas (1)

Ernesto Brown (1)

Juan Pablo Salas (2)

## RESUMEN

Se presenta en este trabajo un análisis del proceso de identificación del tipo de modelo ARIMA apropiado para una determinada serie de tiempo, especialmente con vistas a su aplicación para series hidrológicas. Para esto se examina el proceso de identificación en el caso de series de gastos medios mensuales. En el trabajo se extraen conclusiones con respecto a las dificultades y grado de subjetividad que tiene implícito este proceso.

(1) Ingenieros Civiles. Profesores e Investigadores. Centro de Recursos Hidráulicos. Departamento de Ingeniería Civil. Universidad de Chile.

(2) Ingeniero Civil. Ex Ayudante de Investigación. Centro de Recursos Hidráulicos.

## 1. INTRODUCCION

Este trabajo corresponde a una etapa dentro de un proyecto de investigación más amplio que el Centro de Recursos Hidráulicos está desarrollando. El proyecto está dirigido al estudio de los modelos ARIMA(*) con vistas a su empleo para simular series de tiempo de variables hidrológicas. Para ello se ha dividido el estudio en dos etapas. La primera, que se expone parcialmente en este trabajo, pretende analizar tanto las bases teóricas y matemáticas de los modelos ARIMA, como los procesos de identificación del tipo de modelo apropiado para una determinada serie de tiempo; también comprende el estudio y desarrollo de los procesos de cálculo de los parámetros de los modelos y la creación del software computacional para lograr los objetivos anteriores (Brown et al., 1983).

En una segunda etapa los modelos ARIMA se aplicarán en forma sistemática, para pronóstico y generación de distintas clases de series de tiempo de variables hidrológicas, lo que permitirá evaluar las potencialidades y limitaciones de estos modelos.

## 2. DEFINICIONES

Previo al análisis de formación de los modelos ARIMA, se definen algunos operadores :

- Operador de retroceso no estacional (B)

$$B\,Z_t = Z_{t-1} \qquad B^m\,Z_t = Z_{t-m} \tag{1}$$

- Operador de retroceso estacional ($B^s$)

$$B^s\,Z_t = Z_{t-s} \qquad B^{ms}\,Z_t = Z_{t-ms} \tag{2}$$

(*) Nota: ARIMA es una contracción del inglés "autoregresive integrated moving averages", es decir, modelos autorregresivos integrados de promedio móvil.



en que s define el ciclo de estacionalidad.

- Operador de retroceso diferencial no estacional ($\nabla$)

$$\nabla(Z_t) = (1 - B)\,Z_t = Z_t - Z_{t-1} \qquad \nabla^d(Z_t) = (1 - B)^d\,Z_t \tag{3}$$

en que d es un grado de diferenciación no estacional superior a 1.

- Operador de retroceso diferencial estacional ($\nabla_s$)

$$\nabla_s(Z_t) = (1 - B^s)\,Z_t = Z_t - Z_{t-s} \qquad \nabla_s^D(Z_t) = (1 - B^s)^D\,Z_t \tag{4}$$

en que D es un grado de diferenciación estacional superior a 1.

Para modelar una serie de tiempo con componentes autorregresivas y de promedio móvil utilizando los modelos ARIMA, es necesario que ella sea estacionaria; es decir, que sus propiedades estadísticas no sean afectadas por un cambio en el origen de tiempo.

Si los valores sucesivos de una serie de tiempo son muy dependientes entre sí, ella puede ser generada por una serie de impactos aleatorios $a_t$ con distribución normal de media cero y varianza $\sigma_a^2$ (Yule, citado por Box y Jenkins, 1970), afectada por un filtro lineal. Es decir :

$$\tilde{Z}_t = Z_t - \mu = \psi(B)\,a_t \tag{5}$$

siendo $\mu$ la media del proceso y $\psi(B)$ un operador lineal denominado "función de transferencia de filtro": $\psi(B) = 1 + \psi_1(B) + \psi_2(B^2)\ldots$

El proceso $Z_t$ será estacionario si la secuencia $a_t$ es finita o infinita y convergente. Si el proceso es no estacionario, puede hacerse estacionario, diferenciando la serie, tantas veces como sea necesario, haciendo uso de las ecuaciones (3) y/o (4).

En general, un proceso $Z_t$ no estacional, puede expresarse como :

$$\psi(B)\ Z_t = \theta(B)\ a_t \qquad (6)$$

El polinomio $\psi(B)$ en este caso, puede descomponerse en $\phi(B)\ \nabla^d$, en que d indica el número de diferenciaciones necesarias para que el proceso sea estacionario. Adicionalmente el modelo será estacionario si el polinomio autorregresivo $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 \ldots - \phi_p B^p$, tiene sus raíces fuera del círculo unitario. El polinomio $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q$ representa la parte de promedio móvil del proceso. Este modelo recibe el nombre de ARIMA (p, d, q).

Definiendo $W_t = \nabla^d Z_t$, en que $W_t$ es un proceso estacionario, se tiene el modelo ARMA (p, q)

$$\phi(B)\ W_t = \theta(B)\ a_t \qquad (7)$$

En estos dos modelos se tiene p + q + 2 parámetros a estimar : $\mu;\ \phi_1, \phi_2, \ldots \phi_p;\ \theta_1, \theta_2 \ldots \theta_q;\ \sigma_a^2$.

Si el proceso a modelar tiene características estacionales, el modelo se denota como ARIMA (P, D, Q) o ARMA (P, Q) y su representación es :

$$\Phi(B^s)\ W_t = \Theta\ (B^s)\ a_t \qquad (8)$$

en que el polinomio autorregresivo de carácter estacional de orden P es : $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} \ldots - \Phi_p B^{ps}$ y el polinomio de promedio móvil de carácter estacional de orden Q es : $\Theta\ (B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} \ldots - \Theta_Q B^{Qs}$.

Análogamente a los casos anteriores, se tiene en este caso P + Q + 2 parámetros desconocidos : $\mu;\ \Phi_1, \Phi_2 \ldots \Phi_p;\ \Theta_1, \Theta_2, \ldots \Theta_Q;\ \sigma_a^2$.

Si la serie en estudio exhibe en forma conjunta características estacionales y no estacionales, ella puede representarse por medio de un modelo ARIMA o ARMA multiplicativo. Un modelo de este tipo se denota por ARIMA(p, d, q) x $(P, D, Q)_s$ o, ARMA (p, q) x $(P, Q)_s$ (si no se ha diferenciado la serie) y su representación más general es :



$$\phi(B)\ \Phi(B^s)\ \nabla^d\ \nabla_s^D\ Z_t = \theta(B)\ \Theta\ (B^s)\ a_t \qquad (9)$$

En este tipo de modelos el número de parámetros desconocidos es p + q + P + Q + 2. Para todos los tipos de modelos el valor de los parámetros debe determinarse a base de los datos originales.

Los casos particulares de los modelos aquí presentados se abordan en la publicación de Brown et al. (1983).

## 3. FUNCIONES QUE SE USAN EN LA IDENTIFICACION

El proceso de identificación del tipo de Modelo ARIMA más apropiado para una determinada serie de tiempo, tiene por finalidad estimar, el grado de diferenciación necesario para producir estacionariedad y también, determinar el orden de los polinomios AR y MA.

No existe una formulación precisa para resolver la etapa de identificación, recurriéndose a métodos estadísticos que dan una primera idea del tipo de modelo a ajustar. En el proceso de identificación se recurre a analizar gráficamente las siguientes funciones estadísticas :

ACF : función de autocorrelación
PACF : función de autocorrelación parcial
IACF : función de autocorrelación inversa
IPACF : función de autocorrelación parcial inversa

Estas funciones se deben analizar en conjunto, con el objeto de decidir el o los modelos a ajustar.

### 3.1 ACF. Función de Autocorrelación

Esta función $\rho(k)$ mide el grado de dependencia lineal entre los valores de una serie de tiempo, y se expresa como un cuociente de autocovarianzas $(\gamma_k)$ :

$$\rho(k) = \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{E|(Z_t - \mu)(Z_{t+k} - \mu)|}{E|(Z_t - \mu)^2|} \qquad (10)$$

$$-1 \leq \rho_k \leq 1$$

siendo $\rho_k$ el coeficiente de autocorrelación de orden k.

De acuerdo a la bibliografía (Jenkins y Watts, 1968, citado por Box y Jenkins, 1970), la mejor estimación de la ACF ($r_k$) se logra usando las estimaciones de las autocovarianzas ($C_k$), es decir :

$$r_k = \frac{C_k}{C_0} \qquad (11)$$

Con : $C_k = \frac{1}{n} \sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}) \qquad k = 0, 1, 2, \ldots K$

$$C_0 = \frac{1}{n} \sum_{t=1}^{n} (Z_t - \bar{Z})^2$$

en que : n = número de observaciones de la serie de tiempo

$\bar{Z}$ = media del proceso $Z_t$

k = orden de la autocorrelación

K = valor finito previamente definido

Para usar la ACF en la identificación de modelos, se recomienda calcular y graficar la ACF versus k (intervalo u orden) hasta n/4.

### 3.2 PACF. Función de Autocorrelación Parcial

Se denomina coeficiente de autocorrelación parcial al coeficiente de orden k de un modelo autorregresivo de orden k, AR(k), ajustado a la serie. Estos coeficientes se denotan por $\phi_{kk}$ y la función de autocorrelación parcial se define como $\phi(k,k) = \phi_{kk}$ en donde el primer subíndice indica el orden del proceso AR y el segundo indica que es el último coeficiente del proceso. Para un proceso AR(k) la relación entre los coeficientes de autocorrelación ($\rho_k$) y los coeficientes del modelo $\phi_{ki}$, se obtiene de las ecuaciones de Yule - Walker (Brown et al., 1983), lo que en forma matricial se expresa como :



$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{k-3} \\ \vdots & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{bmatrix} \qquad (12)$$

El método de cálculo de la PACF usado en la identificación de modelos, que es un método aproximado basado en las ecuaciones de Yule - Walker, da origen a las siguientes fórmulas de recurrencia (Durbin, 1960, citado por Box y Jenkins, 1970) :

$$\hat{\phi}_{p+1,j} = \hat{\phi}_{pj} - \hat{\phi}_{p+1,p+1}\, \hat{\phi}_{p,p-j+1} \qquad j = 1, 2 \ldots p$$

$$\hat{\phi}_{p+1,p+1} = \frac{r_{p+1} - \sum_{j=1}^{p} \hat{\phi}_{pj}\, r_{p+1-j}}{1 - \sum_{j=1}^{p} \hat{\phi}_{pj}\, r_j} \qquad (13)$$

en que $\hat{\phi}_{p+1,p+1}$ representa el valor estimado del coeficiente de autocorrelación parcial de orden p+1. Para iniciar el proceso de cálculo se debe calcular $\hat{\phi}_{11} = r_1$.

En la etapa de identificación de modelos se recomienda calcular y graficar la PACF hasta un orden 20 a 40, estimándose que en ese rango ella entrega toda la información que puede proporcionar acerca de la serie.

### 3.3 IACF. Función de Autocorrelación Inversa

La función de autocorrelación inversa de un modelo ARIMA (p, d, q) x $(P, D, Q)_s$, se define como la función de autocorrelación (ACF) de un proceso (q, d, p) x $(Q, D, P)_s$ (Cleveland, 1972), es decir la ACF de

$$\theta(B)\ \Theta(B^s)\ \nabla^d\ \nabla_s^D\ Z_t = \phi(B)\ \Phi(B^s)\ a_t$$

El método más usado para estimar la IACF, consiste en ajustar un proceso autorregresivo de orden L a la serie de tiem

po analizada, es decir :

$$\tilde{Z}_t = a_t + \sum_{i=1}^{L} \pi_i \tilde{Z}_{t-1} \qquad (14)$$

Los coeficiente $\pi_i$ se pueden calcular a partir de las ecuaciones de Yule - Walker (ecuaciones 12), reemplazando $\rho_k$ por sus estimadores $r_k$ y utilizando las fórmulas de recurrencia para la estimación de la PACF (ecuaciones 13).

El coeficiente de autocorrelación inversa de orden k ($ri_k$) se calcula por medio de la expresión :

$$ri_k = \frac{-\pi(k,k) + \sum_{i=1}^{L-k} \pi(k,i)\,\pi(k,i+k)}{1 + \sum_{i=1}^{L} \pi_i^2} \qquad (15)$$

El valor de L, se determina a partir del gráfico de la PACF, tomándolo igual al intervalo a partir del cual la PACF toma valores despreciables.

Para la etapa de identificación es conveniente graficar la IACF hasta un intervalo igual a L.

### 3.4 IPACF. Función de Autocorrelación Parcial Inversa

Esta función, introducida por Hipel et al. (1977) a partir del análisis de series de tiempo, se define como la función de autocorrelación parcial de la función de autocorrelación inversa, es decir :

IPACF = PACF(IACF)

Aquí se hace uso también de las ecuaciones de Yule-Walker (ecuaciones 12) que en este caso se expresan a través de la ecuación 16, siendo $\hat{\phi}i_{kj}$ la estimación del coeficiente j de un proceso autorregresivo de orden k. El coeficiente $\hat{\phi}i_{kk}$ es el último término del proceso y corresponde al valor de la función IPACF para el orden k.



$$\begin{bmatrix} 1 & ri_1 & ri_2 & \dots & ri_{k-1} \\ ri_1 & 1 & ri_1 & \dots & ri_{k-2} \\ ri_2 & ri_1 & 1 & \dots & \\ \vdots & & & & \\ ri_{k-1} & ri_{k-2} & ri_{k-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}i_{k1} \\ \hat{\phi}i_{k2} \\ \hat{\phi}i_{k3} \\ \vdots \\ \hat{\phi}i_{kk} \end{bmatrix} = \begin{bmatrix} ri_1 \\ ri_2 \\ ri_3 \\ \vdots \\ ri_k \end{bmatrix} \qquad (16)$$

Para encontrar las soluciones de este sistema se usan las fórmulas de recurrencia dadas por la ecuación 13.

Para la identificación de modelos, es conveniente calcular y graficar la IPACF hasta un orden similar al usado para la IACF.

## 4. COMPORTAMIENTO DE LAS FUNCIONES ESTADISTICAS

De las funciones estadísticas expuestas en el punto anterior, la ACF es la única que tiene un comportamiento teórico claramente identificable; el comportamiento de las otras funciones es producto principalmente de la experimentación. Box y Jenkins (1970) analizan en forma teórica la ACF y la PACF; en cambio Cleveland (1972) con la IACF e Hipel et al. (1977) con la IPACF, muestran el comportamiento de estas funciones como producto de un análisis experimental.

El comportamiento típico de todas las funciones, se ilustra gráficamente en la Fig. 1, para los modelos más simples.

### 4.1 Comportamiento de la ACF

En un proceso AR(p) el coeficiente de autocorrelación de orden k (Brown et al. 1983) puede expresarse como

$$\phi(B)\,\rho_k = 0 \qquad (17)$$

La solución general de esta ecuación es :

$$\rho_k = A_1 G_1^{\,k} + A_2 G_2^{\,k} + \dots + A_p G_p^{\,k} \qquad (18)$$

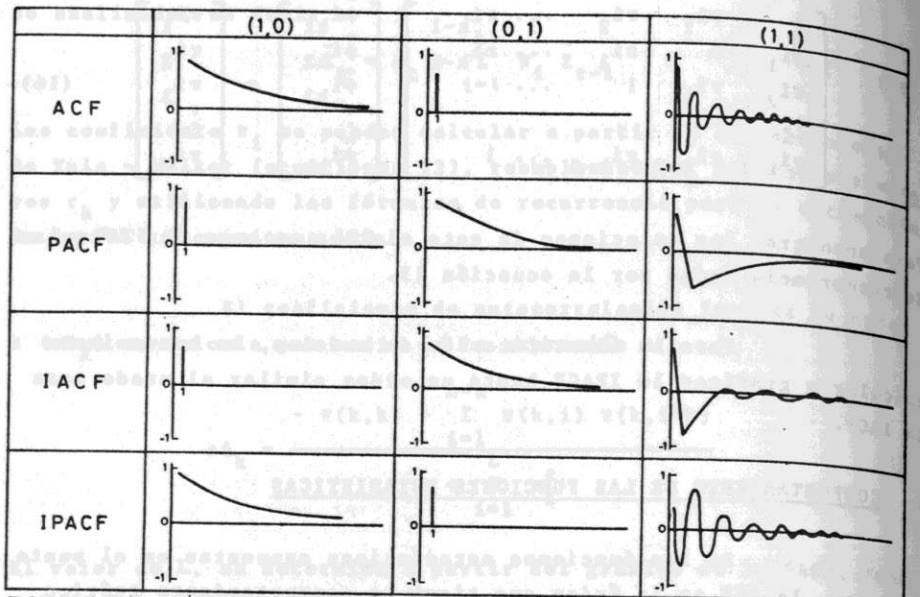


FIG. 1 (a) EJEMPLOS DE COMPORTAMIENTO DE LAS FUNCIONES ESTADISTICAS

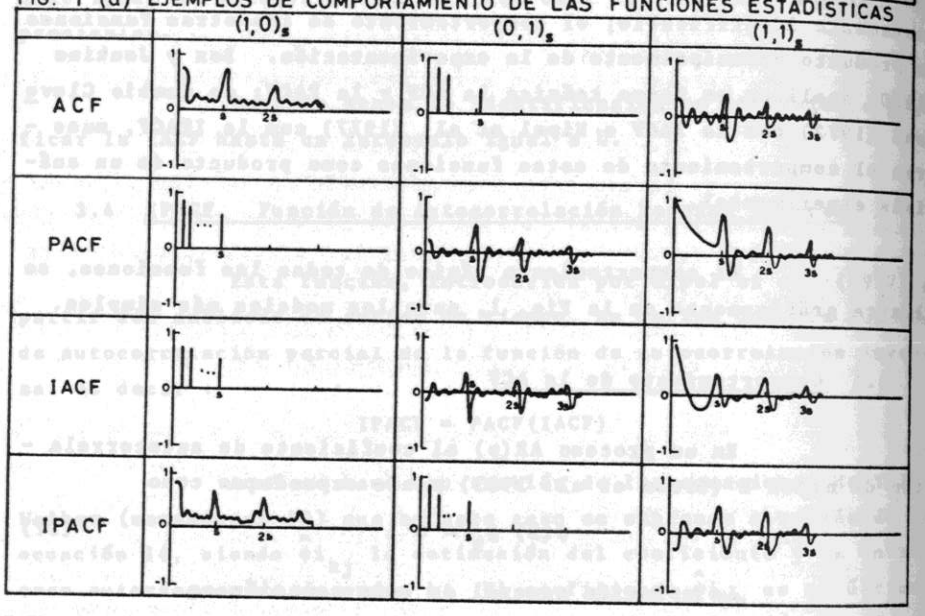


FIG. 1 (b) EJEMPLOS DE COMPORTAMIENTO DE LAS FUNCIONES ESTADISTICAS



siendo $\frac{1}{G_i}$ $i = 1, \ldots p$ las raíces del polinomio autorregresivo $\phi(B)$, las que deben cumplir la condición $|G_i| \leq 1$ para asegurar la estacionariedad, es decir, la serie $\rho_k$ debe converger dentro o sobre el círculo unitario. Las raíces pueden ser reales o complejas conjugadas, en el primer caso $\rho_k$ se comporta como una exponencial amortiguada; en el segundo caso $\rho_k$ representa una curva sinusoidal amortiguada. Lo más común es encontrar una combinación de exponenciales y sinusoides amortiguadas.

Para procesos AR de carácter estacional, la ACF se atenúa en intervalos múltiplos de s, en que s es el período de estacionalidad.

En un proceso MA(q) el coeficiente de autocorrelación de orden k (Brown et al. 1983) se expresa como :

$$\rho_k = \begin{cases} \dfrac{(-\theta_k + \theta_1\,\theta_{k+1} + \theta_2\,\theta_{k+2} + \ldots + \theta_{q-k}\,\theta_q)}{1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2} & 1 \leq k \leq q \\ 0 & k > q \end{cases} \qquad (19)$$

Esta ecuación muestra que si $k > q$ la ACF se trunca y comienza a tener valor nulo.

Si el proceso es MA de carácter estacional, se ha demostrado (Hipel et al. 1977) que la ACF se trunca en sQ.

Para un proceso mixto ARMA (p,q), la ACF a partir del orden q + 1 se expresa por la ecuación 17. En consecuencia, después del intervalo q - p, el comportamiento de esta función, en este tipo de proceso, es análogo al de un proceso AR; es decir, está dominada por exponenciales y sinusoides amortiguadas a partir del intervalo q - p + 1.

Si el proceso es MA multiplicativo la ACF se trunca en el intervalo q + sQ. En cambio, si el proceso es AR multiplicativo, la ACF se atenúa en intervalos múltiplos de s y en los intervalos intermedios (Hipel et al. 1977).

### 4.2 Comportamiento de la PACF

A partir de la definición de esta función, se puede decir que en un proceso AR(p), la PACF será no nula para intervalos menores o iguales a p y tendrá un valor nulo para órdenes superiores.

De acuerdo a la bibliografía (Box y Jenkins, 1970) en los procesos MA y ARMA la PACF está dominada por una mezcla de sinusoides y/o exponenciales amortiguadas.

En procesos multiplicativos AR, la PACF se trunca después del intervalo p + sP, en cambio, si se amortigua en intervalos múltiplos de s y no se trunca en otros intervalos, se requiere componentes MA estacionales y no estacionales respectivamente.

### 4.3 Comportamiento de la IACF y de la IPACF

Para las funciones inversas no existe un análisis teórico que sea factible hacer para visualizar su comportamiento en cada tipo de proceso. Sin embargo, el comportamiento observado experimentalmente para estas funciones es análogo al de las dos anteriores. Se tiene así que la IACF se comporta en forma similar a la PACF, en cambio la IPACF tiene un comportamiento análogo al de la ACF.

De acuerdo a su autor (Cleveland, 1972), la IACF presenta una ventaja en los modelos AR, que consiste en permitir identificar los parámetros del modelo que tienen valor nulo, ya que en esos intervalos la IACF tiene a menudo valores muy cercanos a cero.

## 5. PROCEDIMIENTOS PARA LA ETAPA DE IDENTIFICACION DEL MODELO

- Se grafica la serie original para detectar si la serie presenta características estacionales.
- Se analiza la necesidad de transformar la serie original. En series hidrológicas es frecuente ocupar la transformación logarítmica, con el fin de estabilizar la varianza de las series y para que la serie tenga una distribución normal.
- Se calcula y grafica la ACF de la serie transformada. Si la ACF no decae rápidamente dentro del rango 0 → n/4, se debe diferenciar la serie, ya sea en forma estacional y/o no estacional hasta lograr la estacionariedad.
- Se calculan y grafican las otras funciones estadísticas para la serie estacionaria y de acuerdo al comportamiento de cada una de ellas, se propone el (o los) modelo(s) adecuado(s).



## 6. APLICACION A SERIES DE GASTOS MEDIOS MENSUALES

En esta etapa del estudio se trabajó con 6 estadísticas de caudales medios mensuales correspondientes a 6 ríos chilenos: Elqui, Grande, Maipo, Tinguiririca, Maule y Chillán. En este trabajo se presenta la aplicación realizada para dos de estas series: Grande en Puntilla San Juan y Maule en Armerillo.

Al graficar la serie original para estos ríos (Brown et al. 1983), se aprecia claramente el período de estacionalidad (s = 12); en cuanto a los regímenes hidrológicos observados, ambos ríos presentan máximos en el octavo mes hidrológico, es decir, Noviembre, lo cual confirma la principal componente, de carácter nival, en ambos casos.

En un estudio anterior de estas estadísticas (Brown y Torretti, 1978) se determinó que la distribución de mejor ajuste era la log-normal, razón por la cual se adoptó en este caso el ajuste de modelos ARIMA a los logaritmos de los valores muestrales.

Siguiendo el procedimiento indicado en el punto 4.3 se calcula y grafica las ACF de cada estadística, que se muestran en la Fig. 2. En dicha figura se observa que esta función tiene en ambos casos un comportamiento de sinusoide, pero no presenta un claro decaimiento; esto último sugiere que se trata de procesos no estacionarios, por lo que se necesitaría diferenciar las series.

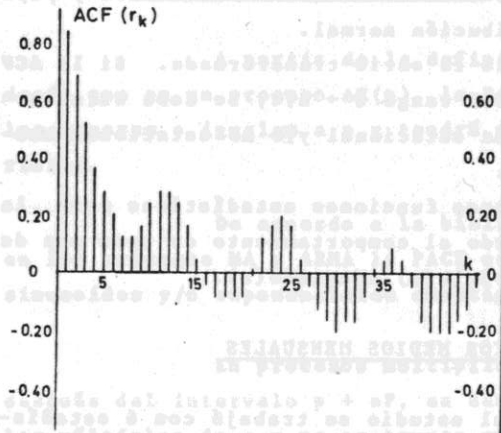


FIG. 2a FUNCION DE AUTOCORRELACION RIO GRANDE EN PUNTILLA SAN JUAN

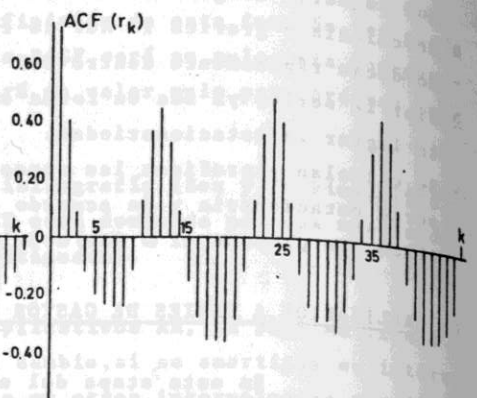


FIG. 2b FUNCION DE AUTOCORRELACION RIO MAULE EN ARMERILLO

Al realizar una diferenciación estacional (D = 1) en ambas series se observa que la ACF decae rápidamente como sinusoidal amortiguada (Ver Fig. 3a). En cambio el comportamiento de la ACF, al diferenciar las series en forma no estacional (d = 1) no produce un rápido decaimiento, persistiendo el efecto sinusoidal con máximos en intervalos múltiplos de 12; sin embargo, si en este caso se agrega una diferenciación de tipo estacional (D = 1), el comportamiento se modifica, truncándose las series (Ver Fig. 4a).

Dado que con las diferenciaciones (D = 1) y (d = 1; D = 1) se logran series estacionarias, de acuerdo al comportamiento de la ACF, se prosiguió con el cálculo de las otras funciones, para ambos tipos de series.

Aquí se presenta el análisis para dos de estas series que son: río Grande en Puntilla San Juan con 1 diferenciación estacional (Figs. 3a a 3d) y río Maule en Armerillo con 1 diferenciación estacional más 1 no estacional (Figs. 4a a 4d).

El análisis de los gráficos muestra lo siguiente :

- Río Grande en Puntilla San Juan: la ACF se amortigua en los intervalos múltiplos de 12 y entre 1 y 12, 12 y 24, etc.; la PACF



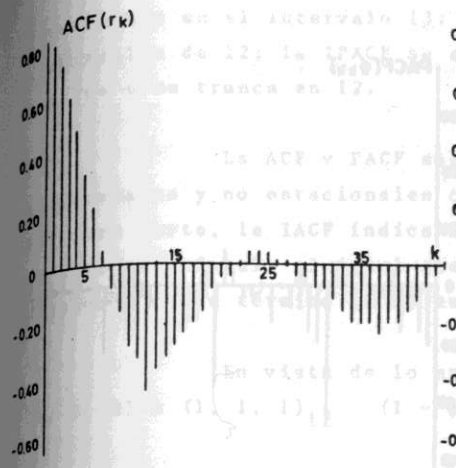


FIG. 3a FUNCION DE AUTOCORRELACION. RIO GRANDE EN PUNTILLA SAN JUAN. 1 DIFERENCIACION ESTACIONAL.

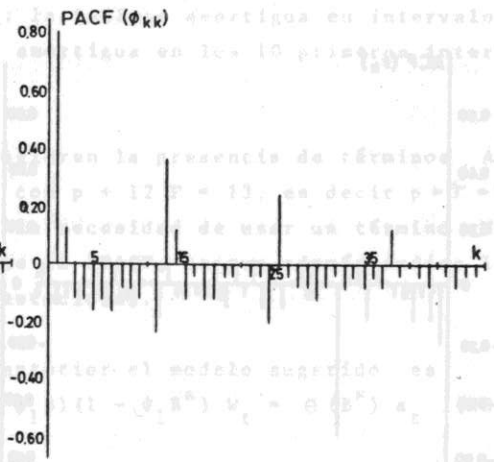


FIG. 3b FUNCION DE AUTOCORRELACION PARCIAL. RIO GRANDE EN PUNTILLA SAN JUAN. 1 DIFERENCIACION ESTACIONAL.

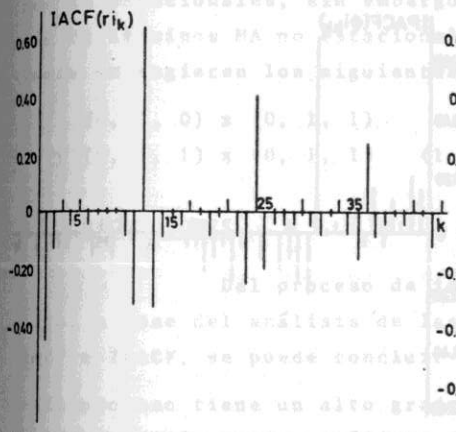


FIG. 3c FUNCION DE AUTOCORRELACION INVERSA. RIO GRANDE EN PUNTILLA SAN JUAN. 1 DIFERENCIACION ESTACIONAL.

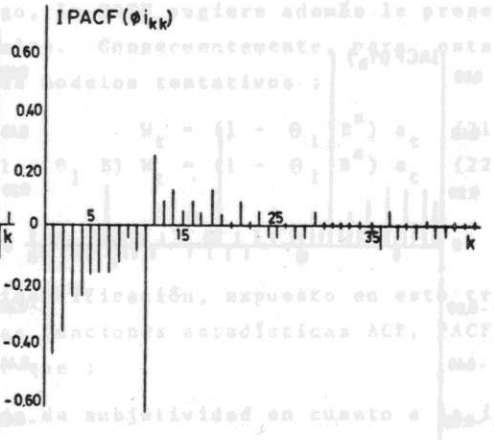


FIG. 3d FUNCION DE AUTOCORRELACION PARCIAL INVERSA. RIO GRANDE EN PUNTILLA SAN JUAN. 1 DIFERENCIACION ESTACIONAL.

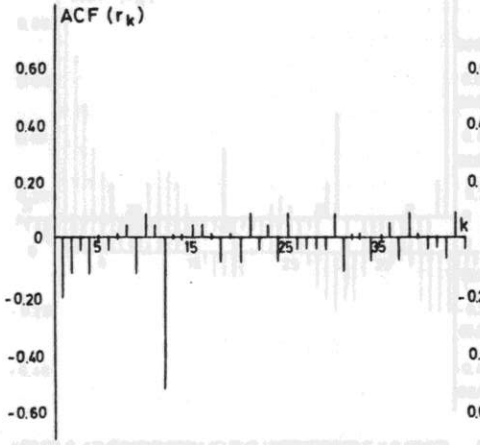


FIG. 4a FUNCION DE AUTOCORRELACION RIO MAULE EN ARMERILLO. 1 DIF. ESTACIONAL MAS 1 DIF. NO ESTACIONAL

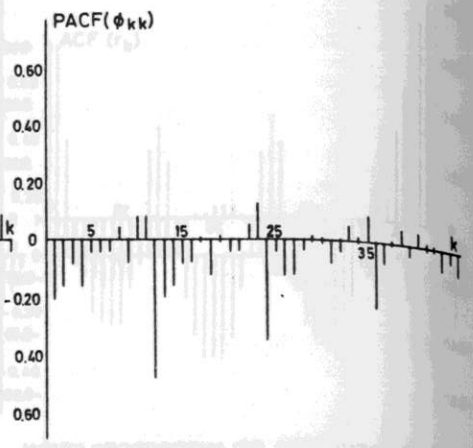


FIG. 4b FUNCION DE AUTOCORRELACION PARCIAL. RIO MAULE EN ARMERILLO. 1 DIF. ESTACIONAL MAS 1 DIF. NO ESTACIONAL

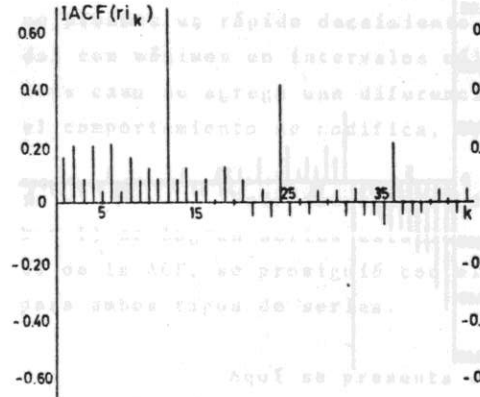


FIG. 4c FUNCION DE AUTOCORRELACION INVERSA. RIO MAULE EN ARMERILLO. 1 DIF. ESTACIONAL MAS 1 DIF. NO ESTACIONAL

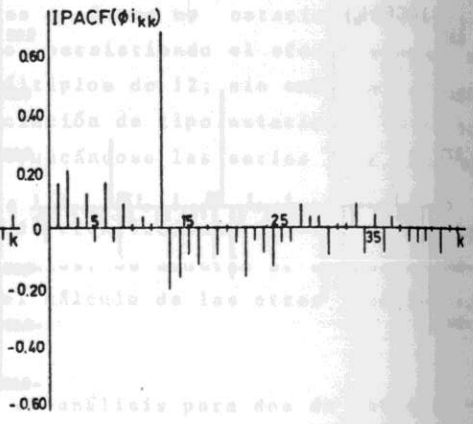


FIG. 4d FUNCION DE AUTOCORRELACION PARCIAL INVERSA. RIO MAULE EN ARMERILLO. 1 DIF. ESTACIONAL MAS 1 DIF. NO ESTACIONAL



se trunca en el intervalo 13; la IACF se amortigua en intervalos múltiplos de 12; la IPACF se amortigua en los 10 primeros intervalos y se trunca en 12.

La ACF y PACF sugieren la presencia de términos AR estacionales y no estacionales con p + 12 P = 13, es decir p = P = 1. Por otra parte, la IACF indica la necesidad de usar un término MA de tipo estacional, al igual que la IPACF, la que además indica la presencia de un término AR no estacional.

En vista de lo anterior el modelo sugerido es

$$(1, 0, 0) \times (1, 1, 1)_{12} \quad (1 - \phi_1 B)(1 - \Phi_1 B^s)\, W_t = \Theta(B^s)\, a_t \qquad (20)$$

- Río Maule en Armerillo: la ACF se trunca en el intervalo 12; la PACF se amortigua en intervalos múltiplos de 12 y además entre 1 y 12, 12 y 24, etc.; la IACF se amortigua en intervalos múltiplos de 12; la IPACF se trunca en el intervalo 12.

Todas las funciones sugieren la presencia de términos MA estacionales, sin embargo, la PACF sugiere además la presencia de términos MA no estacionales. Consecuentemente, para esta serie se sugieren los siguientes modelos tentativos :

$$(0, 1, 0) \times (0, 1, 1) \qquad W_t = (1 - \Theta_1 B^s)\, a_t \qquad (21)$$

$$\text{y } (0, 1, 1) \times (0, 1, 1) \quad (1 - \theta_1 B)\, W_t = (1 - \Theta_1 B^s)\, a_t \qquad (22)$$

## 7. CONCLUSIONES

Del proceso de identificación, expuesto en este trabajo, a base del análisis de las funciones estadísticas ACF, PACF, IACF e IPACF, se puede concluir que :

- El proceso tiene un alto grado de subjetividad en cuanto a la interpretación de los gráficos de las funciones; vale decir que, siguiendo las pautas de interpretación expuestas, personas diferentes podrían extraer distintas conclusiones acerca del tipo de modelo más apropiado para representar una determinada serie. Se detectó en este estudio además, que el criterio de interpretación del analista iba variando con el tiempo, lo que indica la

conveniencia de interpretar los gráficos en forma reiterada a modo de uniformar el criterio de interpretación.

- Al usar la ACF para determinar los grados de diferenciación no estacional y estacional (d y D respectivamente) necesarios para lograr la estacionariedad de la serie, se requiere un proceso de tanteos que también es bastante subjetivo.
  En la práctica, algunos modelos considerados como estacionarios en esta etapa, resultan posteriormente ser no estacionarios en la etapa de determinación de los parámetros (Brown et al. 1983). Dado que es más conveniente trabajar con series no diferenciadas, es necesario notar que, la causa de la no estacionariedad de una serie hidrológica, puede en general ser analizada de los registros originales y corregida si es posible.

- Las funciones inversas, de proposición relativamente reciente, para mejorar el proceso de identificación de modelos ARIMA, no resultan ser tan ventajosas y concluyentes como sus autores aducen. Sin embargo, ellas permiten, en la mayoría de los casos, respaldar las conclusiones que pueden extraerse de la ACF y de la PACF.
  En general puede decirse, que la ACF es la función que tiene el comportamiento más claramente identificable según las pautas indicadas en el punto 4.

## BIBLIOGRAFIA


Box G.E.P. y Jenkins G.M. (1970) "Time Series Analysis Forecasting and Control". Holden Day Inc., San Francisco, California.

Brown E. y Torretti E. (1978) "Experimentación con el Modelo de Markov de Primer Orden en la Generación Estocástica de Caudales" Centro de Recursos Hidráulicos, Depto. de Ing. Civil, Universidad de Chile. CRH 78-9-I.

Brown E., Vargas X. y Salas J.P. (1983) "Selección de Modelos ARIMA para Series de Tiempo Hidrológicas". Centro de Recursos Hidráulicos, Depto. de Ing. Civil, U. de Chile (en prensa).

Cleveland W.S. (1972) "The inverse autocorrelations of a time series and their applications". Technometrics, Vol. 14, N° 2.

Hipel K.W., Mc Leod A.I. y Lennox W.C. (1977) "Advances in Box - Jenkins Modeling 1. Model Construction". Water Resources Research Vol. 13, N° 3.




SOCIEDAD CHILENA DE INGENIERIA HIDRAULICA

VI CONGRESO NACIONAL

# GENERACION DE PRECIPITACIONES DIARIAS MEDIANTE UN PROCESO DE MARKOV SIMPLE

Ernesto Brown F. (1)
Ximena Vargas M. (1)
Juan Arrese L. (2)

## RESUMEN

Se presentan aquí los resultados de la generación de series de precipitaciones diarias, mediante un proceso estocástico simple de Markov. Estos resultados se comparan a nivel anual con aquellos obtenidos por Brown y Cisternas (1978) con un modelo que combinaba un proceso de Poisson para la ocurrencia de tormentas, con un proceso de Markov simple, para generar la lluvia en los días de tormenta. Se discuten los resultados obtenidos y se comparan las ventajas y desventajas de ambos procedimientos.

(1) Ingenieros Civiles; Profesores e Investigadores Centro de Recursos Hidráulicos. Departamento de Ingeniería Civil. Universidad de Chile.
(2) Memorista. Centro de Recursos Hidráulicos. Departamento de Ingeniería Civil. Universidad de Chile.