

MODELACIÓN HIDROLÓGICA Y SU EFECTO EN LOS PRECIOS SPOT

Eduardo Minder, Bonifacio Fernández, Hugh Rudnick

*Escuela de Ingeniería, Pontificia Universidad Católica de Chile
Casilla 306, Correo 22, Santiago, Chile*

Resumen: El modelo utilizado en generación eléctrica en Chile no reconoce la propiedad para el despacho de las centrales. Este trabajo contribuye con una metodología para revisar los datos y para la representación de los parámetros relevantes de la hidrología considerados en los modelos, basándose en series sintéticas. El trabajo concluye que las series hidrológicas tienen un impacto significativo en la señal de costo marginal utilizada en el sistema chileno. Las series utilizadas actualmente entregan costos marginales más altos que los obtenidos usando series sintéticas. La diferencia implica un cambio significativo en los ingresos de las generadoras eléctricas.

Abstract: The deregulation model for competitive generation in Chile assumes that all generation is centrally dispatched, irrespective of plant ownership. The paper contributes with a methodology to review the data and represent the relevant hydrology for the operational models of a system, based on synthetic streamflows. The paper concludes that the hydrological data series has a significant impact on the marginal cost signals being used in the Chilean pool. The data being used today provides higher marginal costs than those the one obtained by using built synthetic matrices. The difference implies a significant variance in terms of revenues for generation firms.

Keywords: Precios spot, costo marginal, despacho hidrotérmico, modelación hidrológica, modelos periódicos autorregresivos, matrices sintéticas.

1. INTRODUCCIÓN

En Chile se ha estimulado la creación de un mercado competitivo de generación eléctrica (Rudnick *et al.* 1997) a través de la utilización de modelos de despacho en base a costo marginal de corto plazo con costos auditados y despacho centralizado sin reconocer propiedades de las centrales generadoras. Esto impone una importante complejidad al costear la producción hidroeléctrica, ya que se deben considerar varios aspectos en la modelación: embalses interanuales (a modo de ejemplo, el Lago Laja con capacidad de regulación de 20 meses), tamaño del Sistema Interconectado Central (potencia instalada de 4.750 MW en 1996), cantidad de centrales (más de 35 centrales generadoras), y preponderancia de centrales hidroeléctricas (77% de la potencia instalada corresponde a este tipo), entre otras. Por otra parte, la

Comisión Nacional de Energía (CNE) es la que entrega los planes de obras indicativos.

A nivel de generación es posible distinguir dos tipos de mercados:

Mercado spot: en éste se transan excedentes de generación con respecto a los compromisos de cada generador. Participan sólo los generadores y las transacciones sólo se hacen entre ellos. En el mercado spot, el precio de la energía corresponde al costo marginal de corto plazo, varía horariamente y depende de la condición hidrológica presente, ubicación geográfica y del nivel de tensión. Estos precios son determinados por el Centro de Despacho Económico de Carga (CDEC). El costo marginal de potencia corresponde al costo de desarrollo de capacidad de punta.

Mercado de contratos: también denominado mercado a término, en el cual un generador establece contratos de abastecimiento. En el mercado de los contratos, los precios pueden ser libremente negociados entre las partes o regulados por la autoridad.

Resulta importante destacar que como consecuencia de lo anterior aparecen los conceptos de generador excedente y deficitario. Si bien cada empresa generadora suscribe contratos con clientes, como se indica anteriormente, el criterio que prima para la operación del sistema en su conjunto es de mínimo costo sin reconocer propiedad de las centrales. Luego, se producen diferencias entre lo aportado por un determinado generador y lo retirado por sus respectivos clientes. Lo retirado por los clientes se remunera de acuerdo a los contratos suscritos, no obstante las transferencias entre generadores son pagadas a precio spot.

Por lo anterior, la necesidad de determinar la forma de operar el sistema y, más aún, los costos marginales resultantes, es fundamental. De aquí que los modelos utilizados cobren especial importancia, en particular la metodología de cálculo utilizada (multiembalse, multinodal, etc.) para la operación del sistema, así como los modelos empleados para considerar los aportes de recursos hídricos disponibles.

Este trabajo contribuye con una metodología para revisar la información hidrológica a ser utilizada en la modelación, precisando los principales aspectos que deben ser tomados en cuenta (Minder, 1997). Contribuye también con una metodología para representar adecuadamente los aspectos relevantes para la modelación de la operación de un sistema hidrológico, basado fundamentalmente en el uso de series de caudales sintéticos. Esta metodología incorpora incertidumbre en el uso de información hidrológica, lo que permite obtener un perfil de costos marginales que reflejan las características y la variabilidad de las series de caudales.

2. DESPACHO HIDROTÉRMICO Y CÁLCULO DE COSTOS MARGINALES

Chile tiene 15 millones de habitantes y dos sistemas interconectados. El Sistema Central Interconectado (SIC) con una demanda máxima de 3.500 MW y 22.400 GWh in 1996, con una red de transmisión en 500, 220, 154 y 110 kV. Con una capacidad instalada de cerca de 4.750 MW (78% hidroeléctrico) en el SIC. La generación hidráulica varía de 60% a 95% dependiendo de la condición hidrológica. Las unidades térmicas son abastecidas por carbón principalmente y durante 1997, con la llegada del gas natural, comienzan a operar las centrales de ciclo combinado.

Las regulaciones existentes en Chile requieren que la operación de cada sistema interconectado se haga de manera de abastecer la demanda a mínimo costo, tomando en cuenta restricciones de medio ambiente, seguridad y calidad de servicio. Para este propósito, modelos de despacho son usados por los CDEC encargados de cada sistema. El modelo OMSIC es un modelo de operación del SIC (CDEC-SIC, 1991) con etapas mensuales que utiliza programación dinámica, considera un consumo centrado en una barra única, trabaja con cuatro etapas semanales al comienzo del horizonte de planificación y con etapas mensuales en el resto del periodo. Este modelo fundamentalmente optimiza las extracciones del Lago Laja, que es el de mayor volumen de regulación del sistema, de manera de minimizar el costo esperado de generación termoeléctrica y falla en todo el horizonte de planificación. En este modelo la hidrología se considera independiente entre meses en el periodo de invierno (Abril-Septiembre) y dependiente en el periodo de verano (Octubre-Marzo, meses de deshielo), y las decisiones se toman en la modalidad azar-decisión. Este modelo considera dos etapas, una de optimización y otra de simulación. En la fase de optimización se busca minimizar los costos presentes y futuros de operación del sistema, mediante la utilización de programación dinámica. En la fase de simulación del modelo se utiliza el método de Montecarlo. El modelo entrega como resultados una tabla de costos marginales esperados, tabla de cotas finales esperadas en el Lago Laja y la generación esperada por central.

3. CONSTRUCCIÓN DE BASES DE DATOS HIDROLÓGICAS

La operación de los modelos actualmente en uso para colaborar en la gestión del SIC requiere de información hidrológica que represente la oferta de agua al sistema. Esta información se representa mediante caudales medios en diferentes secciones de las cuencas que participan con su oferta al sistema de centrales y embalses que forman el SIC.

En la actualidad los modelos utilizan como información básica series cronológicas de valores medios semanales de 40 años de extensión, representativas del periodo 1941-1980, organizadas por años hidrológicos (de Abril a Marzo). Estas series de información de caudales disponibles provienen de mediciones que se han realizado en las diversas cuencas del país. Los valores considerados en esta base de datos son obtenidos fundamentalmente de tres formas:

Medición directa de la serie en estaciones fluviométricas ubicadas en el lugar considerado.

Flujos para Laja en Tucapel y Estero el Toro en Captación

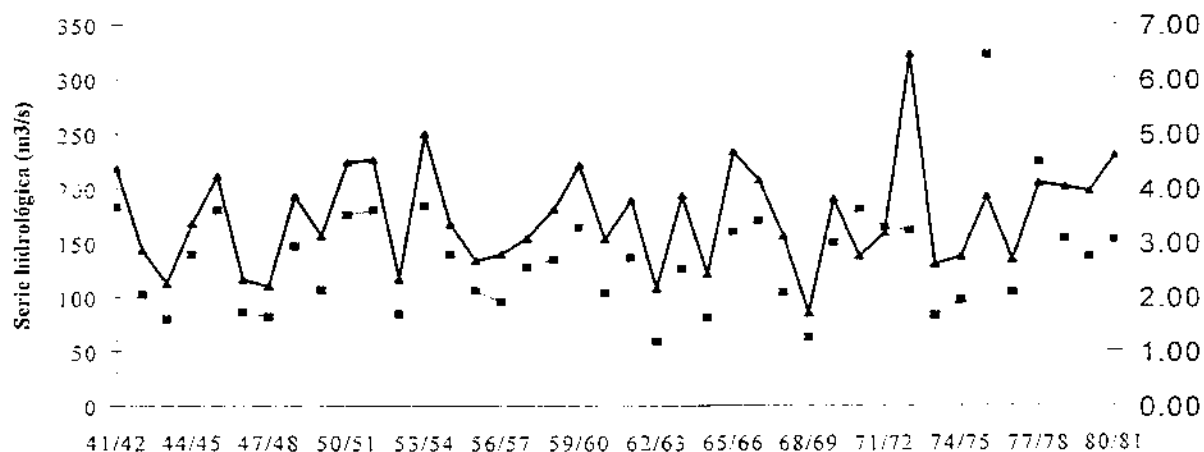


Figura 1: Comparación de series Laja en Tucapel y Estero El Toro en Captación.

Obtención de los valores a través de la adición y sustracción de flujos medidos en otros lugares de la cuenca (dado que en el lugar de interés no se cuenta con mediciones directas), de manera de completar la información necesaria, en base a relaciones de continuidad.

Estimaciones a partir de otras mediciones (Ej.: filtraciones, superficies de napas, precipitaciones, etc.), empleando para ello regresiones, o relaciones de distintos tipos.

A modo de ejemplo, es posible considerar la serie "Laja en Tucapel", que es medida en la estación fluviométrica del mismo nombre en el lugar del mismo nombre y se ha registrado directamente desde Abril de 1916. Por otra parte, la "Cuenca intermedia entre Captación Rama Laja, Central Antuco y Lago Laja" ha sido medida sólo desde Abril de 1981. No obstante éstas dos series se han medido de manera directa, la disponibilidad de registros es absolutamente distinta. Luego, para poder utilizar la segunda de las series nombradas en los modelos construidos, ha sido necesario efectuar un relleno y extensión a un período común.

Debido a la existencia de estos periodos de relleno en las series de caudales de las cuencas chilenas, en general, los datos a utilizar deben ser cuidadosamente filtrados de modo de trabajar con los datos más fidedignos posibles. Cabe destacar que el relleno de datos en si no es el problema, sino que el problema es que éste fue realizado sin la preocupación de reproducir momentos de mayor orden que el primero. En efecto, con los rellenos hechos se reprodujeron promedios, pero no desviación estándar, asimetría, curtosis, etc.; por lo que se tiene como consecuencia que las series rellenadas son menos fluctuantes (menor desviación estándar) que las reales. Además, las series rellenadas presentan una mayor

dependencia temporal y espacial debido a que las técnicas de regresión y uso de relaciones empíricas no toman en cuenta la aleatoriedad.

Para visualizar el efecto de este tipo de rellenos es posible considerar los valores que se presentan para las series Laja en Tucapel y Estero el Toro en Captación, donde los 10 primeros años de la segunda serie tienen la misma forma de los análogos de la serie Laja en Tucapel que se presenta en la figura 1.

4. GENERACIÓN DE SERIES SINTÉTICAS DE CAUDALES

La modelación de una serie cronológica consiste en seleccionar un modelo que reproduzca adecuadamente las características estadísticas de la serie. Dado que las posibilidades están restringidas a los modelos conocidos (Hipel and Mcleod, 1994), en la mayor parte de las situaciones se supone un cierto modelo y se verifica si es el adecuado (Salas *et al.* 1985). En este sentido, no se trata de encontrar un modelo que corresponda al proceso que ha generado la serie en cuestión, sino más bien elegir por comparación, entre los posibles modelos, el que mejor represente las características consideradas relevantes.

Para el caso de los modelos lineales existe una metodología de modelación propuesta por Box y Jenkins (1976) que es ampliamente aceptada. Ésta consiste en tres etapas que son: identificación del orden, estimación de parámetros y validación. A esto Salas agrega dos etapas previas: composición del modelo y selección del tipo (Salas *et al.* 1985).

La modelación general es complementada con el uso del modelo, ya sea en simulación o pronóstico. En el caso de la simulación, ésta se realiza mediante la generación de series sintéticas (matrices hidrológicas con los datos obtenidos a partir de la utilización de los modelos ajustados que se suponen equiprobables a las series empleadas en su construcción), a partir de los modelos construidos, siendo éstas utilizadas como datos de entrada para el modelo OMSIC. Esto permite obtener distintos valores de costo marginal para cada una de las hidrologías generadas a partir de los modelos.

A continuación se presentan los pasos necesarios para la generación de las matrices sintéticas (Salas *et al.*, 1985) utilizadas para alimentar el modelo OMSIC y sus posibles efectos sobre el comportamiento del sistema.

4.1 Modelación temporal de cada serie

Con la modelación temporal se busca representar la dependencia que tiene la serie consigo misma, de tal manera de obtener ruidos independientes temporalmente, aunque dependientes espacialmente. En esta etapa se abordan los siguientes pasos: normalización de los valores, estandarización, selección del modelo y representación de los efectos periódicos.

La normalización de la serie: consiste en hacer una transformación de tal modo que la función de distribución de probabilidades marginal de los valores sea lo más simétrica posible.

La estandarización consiste: en restar la media μ y dividir por la desviación estándar σ , obteniéndose una serie que debe distribuir Normal estándar ($N(0,1)$).

Orden del modelo: está dado por el número de parámetros que es necesario estimar para realizar una adecuada predicción. Para discriminar entre los distintos órdenes de un modelo del tipo AR se utilizan criterios de comparación; en este caso se utilizó el Criterio de Akaike (Salas *et al.*, 1985).

Periodicidad, modelación de la dependencia temporal y la utilización de un modelo autorregresivo con parámetros periódicos, PAR, en que todos los parámetros o algunos de ellos varían dentro de un año o tienen un determinado periodo. De esta manera, el valor del caudal, normalizado y estandarizado, para el mes τ del año v en un modelo autorregresivo, está dado por la siguiente expresión:

$$z_{v,\tau} = \phi_{1,\tau} z_{v,\tau-1} + \phi_{2,\tau} z_{v,\tau-2} + \dots + \phi_{p,\tau} z_{v,\tau-p} + \varepsilon_{v,\tau} \quad (1)$$

donde dicho valor es función de los p valores anteriores, siendo p el orden del modelo que fue identificado

anteriormente. Los ϕ son los parámetros del modelo a estimar y ε es la variable aleatoria independiente de distribución normal con media cero y desviación típica $\sigma\varepsilon$.

4.2 Modelación espacial del conjunto de series

Cuando existen varias series, en varios lugares diferentes de una región de la misma variable como es el caso de los aprovechamientos del SIC, o de diferentes variables en un mismo lugar, o combinaciones similares, que presentan dependencia entre los valores de las series, se deben considerar modelos multivariados.

Considerando los valores de los ruidos a los que ya ha sido removida la dependencia temporal (de acuerdo al metodología antes descrita), se construye la matriz de coeficientes espaciales, considerando un modelo contemporáneo.

Para ello primero se calculan las correlaciones de desfase cero, o contemporáneas, entre cada una de las series consideradas, dando lugar a la matriz de covarianzas, donde el coeficiente b_{ij} representa la correlación existente entre la serie i y la serie j .

4.3 Generación de series de tiempo sintéticas

A partir de esta etapa se debe aplicar en sentido inverso lo hecho en las etapas anteriores, verificando el cumplimiento de las hipótesis en cada paso. Para la generación de series sintéticas, lo primero es generar un conjunto de números $ITN(0,1)$, de acuerdo a algunas de las técnicas descritas en la literatura (Salas *et al.*, 1985).

Utilizando los parámetros del modelo y descartando los primeros valores para inicializar los datos se reconstruye la serie de valores que son normales de media cero y desviación estándar uno. Si se le aplicó una transformación para normalizar la serie se debe desnormalizar dicha serie aplicando la transformación inversa. A casi todas las series se les debió aplicar transformación para normalizarlas (se utilizó la transformación de Box-Cox). Una vez que las series están desnormalizadas se debe desestandarizarlas. Lo que resulta es una serie de tiempo sintética para cada uno de los lugares escogidos. Finalmente, es fundamental verificar la reproducción de los estadísticos relevantes en cada etapa.

5. SERIES SINTÉTICAS Y SU APLICACIÓN EN EL SIC

Mediante la utilización de los modelos ajustados es posible generar valores distintos de los reales, pero teniendo la certeza que ellos son igualmente válidos (equiprobables). A partir de dichos valores es posible construir matrices hidrológicas sintéticas, y en particular, de las características

de las que utiliza el modelo OMSIC, es decir, 40 años hidrológicos correspondientes a los periodos 41/42 al 80/81, que son alimentados al modelo a través del archivo de datos de entrada OMSHID.

Es importante hacer notar que en esta aplicación se considera para cada lugar sólo los datos medidos para la construcción y estimación de los modelos, aunque ello significa disponer de cantidades distintas de información para cada serie. Sin embargo, las series generadas corresponden a un período común de 40 años similar al empleado con las series históricas por el OMSIC.

La principal ventaja que presenta la utilización de las matrices sintéticas es la posibilidad de generar más de una matriz hidrológica, donde sólo existe una serie de tiempo. Esto posibilita la obtención de un perfil de costos marginales dependientes de la serie hidrológica generada de acuerdo a los modelos, que reflejan la influencia de las incertidumbres del comportamiento hidrológico en la operación del sistema.

Para ejemplificar lo antes descrito, se utiliza una corrida del modelo OMSIC realizada con las matrices de datos reales y se comparan con los resultados obtenidos de realizar diez corridas con matrices sintéticas del mismo modelo. Las series utilizadas para esta ejemplificación son todas las correspondientes a la Cuenca del Laja. El resultado obtenido se presenta en la figura 2, donde se puede apreciar que los valores de costo marginal promedio de la energía obtenidos como resultados de las corridas del modelo OMSIC son ligeramente inferiores, alrededor de 2,00 millsUS\$/KWh en promedio, dependiendo del período considerado, para el caso de las matrices sintéticas y la desviación estándar promedio es alrededor de 1,00 millsUS\$/KWh.

La razón de la diferencia que se presenta entre los valores entregados por el modelo con la estadística original y los entregados con los valores sintéticos se puede explicar fundamentalmente por dos posibles efectos. El primero de ellos es la consideración de una muestra de 30 años hidrológicos (período 1951-1980) debido a la existencia de rellenos, en lugar de los 40 años (período 1941-1980) que actualmente se utilizan. El segundo efecto es que al generar estos valores aparecen algunos datos menores que cero, lo cual no tiene sentido en términos de caudales, por lo que dichos valores son reemplazados por cero. La razón por la que se producen estos inconvenientes es por la alta asimetría de algunas series, la cual a pesar de las transformaciones aplicadas mantiene algún efecto sobre los modelos. Esto se da sólo en algunas series y, en particular, en aquellas con gastos cercanos a cero. Al llevar los valores a cero, la matriz modificada contiene más agua que la real, debido a la sustitución de los valores negativos por cero, por lo tanto la señal de costo marginal obtenida del modelo OMSIC resulta menor. Este segundo efecto tiene un menor impacto que el primero considerado por las razones explicadas.

La construcción de un modelo estocástico para representar datos hidrológicos se realiza de manera de imitar las propiedades estadísticas de las series de datos disponibles. En esta forma la solución del modelo de optimización del sistema depende de manera relevante de los datos disponibles. Es, por lo tanto, importante que esa información sea revisada cuidadosamente para asegurar que no existan valores rellenos que no aportan información o que la distorsionan. Por otra parte, la construcción del modelo debiera ser una oportunidad única para la revisión de la información. Este tipo de modelos trata de reproducir los valores esperados, las varianzas, las covarianzas o dependencias temporales en una misma serie y las espaciales entre series cercanas.

Por otra parte, es sabido que la estadística de caudales en Chile es relativamente corta y que sólo se dispone de unas pocas series completas para el período de análisis, de manera que muchas de ellas contienen datos rellenos mediante diferentes técnicas. Estas técnicas, en general, mejoran la estimación de los valores esperados, pero introducen sesgos en las estimaciones de la varianza y las dependencias. Estos dos aspectos son esenciales en el modelo estocástico de caudales planteado para ser empleados en la operación del sistema. Si esto no se revisa cuidadosamente el modelo reproducirá estas distorsiones. Es por ello importante que para la construcción del modelo se verifique y separe los datos reales de los datos rellenos y completados, y se asegure que las estimaciones de las varianzas y las dependencias temporales y espaciales sean las correctas.

Adicionalmente, es altamente recomendable agregar datos de mediciones más recientes si se busca ampliar la muestra, usar los últimos 50 años hidrológicos, o bien reemplazar los primeros 10 o 15 años de la muestra considerada en la actualidad por los datos de los años 81-82 al 90-91 o al 95-96. La ventaja fundamental de realizar esto, es trabajar con una muestra de datos efectivamente medidos y evitar los problemas antes explicados.

Para la modelación hidrológica de los aportes resulta fundamental realizar algunas consideraciones:

La tarea de modelación es muy diferente a la de simulación, debiendo el modelo en sí ser único si los datos lo son y sólo se podría pensar en modificarlo si se cuenta con nuevos antecedentes. Desde esta perspectiva disponer de un programa que construya modelos a gusto del usuario en la selección del orden y la estimación de los parámetros cada vez que hace una simulación no parece muy razonable.

Comparación de resultados de corridas con series hidrológicas sintéticas y reales

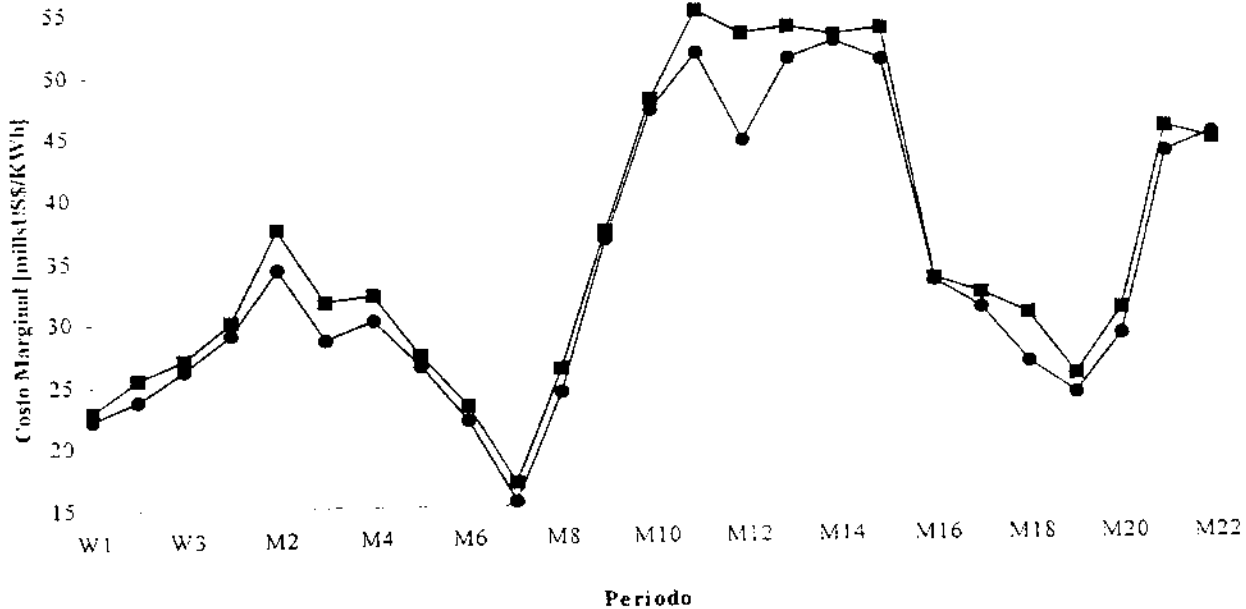


Figura 2: Resultados de las corridas realizadas al modelo OMSIC.

Dada las altas asimetrías existentes en las series de datos chilenos resulta adecuado aplicar las transformaciones sobre los datos y no sobre los residuos, ya que esto permite además tener un mayor control sobre la calidad de los datos al poder detectar en las primeras etapas del proceso la existencia de outliers o información anómala, la que al ser removida al principio no altera la estimación de parámetros ni la identificación de los modelos.

Al utilizar matrices sintéticas se puede generar más de una matriz hidrológica para cada serie de tiempo teniendo la certeza que ellas son igualmente válidas (equiprobables), por la manera en que se ha realizado su construcción. La ventaja que presenta este hecho es que, dejando el resto de las variables constantes se puede obtener información acerca de la sensibilidad del modelo OMSIC respecto de las posibles condiciones hidrológicas, a través de la obtención de un perfil de costos marginales. Cabe destacar que esta utilización en la simulación de la operación resulta imposible de realizar con las matrices hidrológicas reales, o con la reordenación de ellas.

Por otra parte, la generación de series sintéticas presenta algunos inconvenientes como la aparición de valores negativos lo que se puede deber, entre otras razones, a la gran asimetría de las series consideradas. Resulta muy importante revisar las series sintéticas de

modo de detectar dicho problema y corregirlo mediante el ajuste de otro tipo de modelo, o bien, la consideración de otros órdenes para el mismo tipo de modelo.

Finalmente, se concluye que los datos hidrológicos tienen un impacto significativo en la señal de costo marginal que es usada en SIC. Las series usadas en la actualidad entregan valores de costo marginal más altos que los que se obtendrían utilizando las matrices sintéticas construidas. La diferencia de 2 mills US\$/kWh implica una significativa varianza en términos de los ingresos para las empresas generadoras. Empresas cuya capacidad instalada sea principalmente hidroeléctrica se pueden beneficiar más que las térmicas, debido a que la cantidad de agua disponible es sobreestimada. Finalmente, los consumidores perciben un costo más alto por la energía generada.

AGRADECIMIENTOS

Los autores agradecen a Coibún y Endesa por su apoyo a esta investigación. También aprecian el apoyo del proyecto Fondecyt 1950983.

REFERENCIAS

- Box, G. and Jenkins (1976) G. *Time Series Analysis, Forecasting and Control*. Revised Edition.
- CDEC-SIC (1991). *Modelo OMSIC*, Informe Interno CDEC-SIC.
- Hipel, K. and McLeod, A. (1994) *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier; Amsterdam, London, New York, Tokyo.
- Minder, E. (1997) "Modelación hidrológica en el Sistema Interconectado Central". Tesis de Magister en Ciencias de la Ingeniería. Universidad Católica de Chile.
- Rudnick, H., Varela, H., Hogan, W. (1997) "Evaluation of alternatives for power system coordination and pooling in a competitive environment", *IEEE Trans. on Power Sys.*, Vol. **12**, N°2. pp. 605-613.
- Salas, J.; Delleur, J.; Yevjevich, V.; and Lane, W. (1985). *Applied Modeling of Hydrologic Time Series*. Water Resources Publications.