



DIPLOMA DE POSTTULO

PREPARACIÓN Y EVALUACIÓN DE PROYECTOS

MODULO: METODOS DE ESTIMACION Y DEMANDA

PROFESOR: RICHARD WEBER

MATERIAL DOCENTE

- AÑO ACADEMICO 2009 -

PRONÓSTICO DEL PRECIO ANUAL DEL COBRE MEDIANTE REDES NEURONALES

CRISTIAN FOIX*
RICHARD WEBER**

Resumen

El alto nivel que ha alcanzado el precio del cobre en los últimos años, sumado a su gran impacto en la actividad minera nacional motivan a la aplicación de nuevas técnicas para su modelamiento y pronóstico. Un pronóstico ajustado del precio tiene un gran valor no sólo para las empresas de la industria, quienes se podrían beneficiar gracias a la correcta evaluación de proyectos y negocios mineros, sino que también para la autoridad económica en la definición del presupuesto fiscal. En el presente trabajo se entregan evidencias respecto a la potencia de las redes neuronales como herramienta para el pronóstico del precio anual del cobre. Los resultados conseguidos se contrastaron con los generados mediante la aplicación de los más tradicionales y exitosos modelos de series de tiempo. Adicionalmente, se construyeron modelos híbridos combinando modelos ARIMA y redes neuronales. Los resultados revelaron un mejor desempeño de los modelos de pronóstico basados en redes neuronales en el periodo de evaluación considerado (1977-2006), especialmente en pronósticos a más de dos años. Se concluye que las redes neuronales, aunque requieren un mayor esfuerzo en su diseño, pueden ser una herramienta valiosa para el pronóstico del precio del cobre.

Palabras Clave: Pronóstico, redes neuronales, series de tiempo.

*Dirección de Estudios, Codelco

**Departamento de Ingeniería Industrial, Universidad de Chile

1. Introducción

A partir del año 2003, el precio del cobre ha mostrado un extraordinario repunte, llegando a niveles no vistos en más de 30 años. Esto se ha reflejado en nuestra economía donde el cobre ha llegado a representar sobre el 55 % del valor de nuestras exportaciones y sólo los aportes de CODELCO y ENAMI a los ingresos fiscales ya superan el 15 %.

El desconocimiento de los futuros valores de esta importante variable motiva la búsqueda de metodologías de pronóstico eficientes que permitan contar con estimaciones de la mayor precisión posible para su uso tanto en la industria como en la definición del presupuesto de la nación.

En este trabajo se evaluó el desempeño predictivo de diferentes redes neuronales de tipo multilayer perceptron, construidas sobre la base del precio rezagado del cobre y variables derivadas del mismo, tales como la última variación del precio y la desviación estándar de los últimos periodos. Los resultados conseguidos se contrastaron con los generados mediante la aplicación de los más tradicionales y exitosos modelos de series de tiempo (ARIMA, caminata aleatoria y promedio móvil), estableciéndose así el tipo de modelamiento que es capaz de aprovechar mejor la información contenida en los precios históricos considerados en el estudio.

El capítulo 2 muestra antecedentes del presente estudio. En el capítulo 3 se encuentra el modelamiento empleado con redes neuronales y modelos híbridos. El capítulo 4 exhibe los resultados obtenidos. Las conclusiones del presente estudio se encuentran en el capítulo 5 mientras el capítulo 6 describe posibles trabajos futuros.

2. Antecedentes del estudio

2.1. El impacto del precio del cobre en la actividad minera y en la economía chilena

La minería del cobre se ve directamente afectada por los vaivenes del precio. En los periodos de altos precios, aumentan las utilidades de las faenas en operación y se reactivan las exploraciones, los nuevos proyectos, las expansiones y las reaperturas. Por el contrario, en los periodos de bajos precios, disminuyen las utilidades de las operaciones existentes, disminuyen las inversiones en exploraciones, se postergan los proyectos y se cierran las faenas de

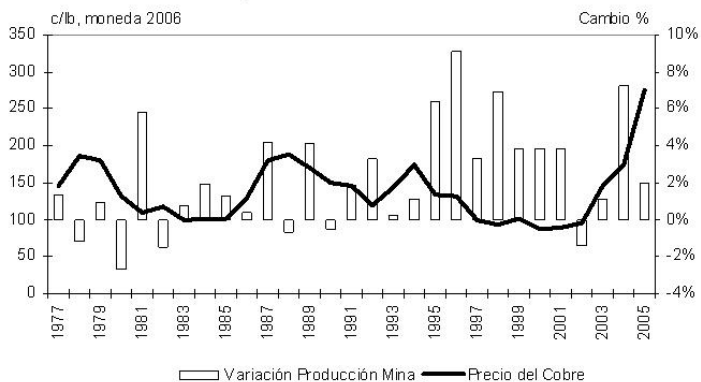
altos costos.

De este modo las fluctuaciones del precio del cobre se traducirán en variaciones de la producción minera, aunque con un desfase dado por la capacidad de reacción de la industria (Gráfico N°1).

En Chile, estas variaciones de la actividad minera repercuten en las inversiones, en el tipo de cambio vía exportaciones, y en el presupuesto fiscal a través de los impuestos, transmitiéndose al resto de la economía gracias al efecto de estas variables sobre la demanda agregada, la inflación, los precios relativos y el tipo de cambio real [19][24]. Así, en los últimos 20 años, mientras el precio registró un mínimo anual de 87,4 c/lb y un máximo anual de 188,5 c/lb (valores en moneda 2006), la minería del cobre representó entre el 34 % y el 50 % de las exportaciones y los aportes de la Corporación del Cobre (Codelco) llegaron a significar entre el 2 % y el 25 % de los ingresos fiscales [6].

Por lo anterior, contar con proyecciones precisas del precio anual del cobre tiene atractivo no sólo para los productores del metal, quienes resultarían favorecidos por el menor error en un parámetro clave del proceso de planificación minera y evaluación de proyectos, sino que también para el gobierno de la nación. Por ejemplo, el ajuste del gasto fiscal o el endeudamiento en un escenario de bajos precios del cobre dependerá de la duración esperada de la fase depresiva del precio [10].

Gráfico N°1: Precio del Cobre y Variación Porcentual de la Producción Minera Mundial



Nota: Precio del Cobre entre los años 1977 y 2006, c/lb moneda 2006. 2006, promedio a junio.
Fuente: Cochilco [7] y World Bureau of Metal Statistics.

2.2. ¿Es pronosticable el precio del cobre?

Una variable será pronosticable en la media que su comportamiento histórico nos revele patrones o relaciones que nos permitan estimar su evolución futura. De este modo, el éxito de los pronósticos basados en modelos depende de [5]:

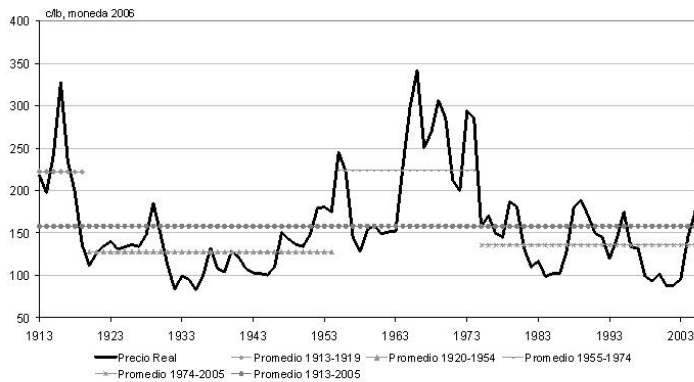
- La existencia de regularidades a ser capturadas.
- El que las regularidades sean informativas acerca del futuro.
- La adecuada captura de estas regularidades por parte del modelo construido.

Con esta definición como referente podemos examinar el caso del cobre. La evolución del precio del cobre desde 1913 en adelante está marcada por su alta volatilidad (Gráfico N°2). Las causas que provocan estas variaciones son diversas y de distinta naturaleza [18]:

- Fuertes cambios en la industria derivados ya sea de shocks en el precio de los factores productivos, grandes innovaciones tecnológicas o del desarrollo de nuevos mercados, cuyos efectos pueden extenderse por decenas de años. A estos trastornos se les denomina shocks con efecto permanente.
- Desajustes transitorios entre la oferta y la demanda de cobre provocados por los ciclos de la actividad económica mundial y la falta de flexibilidad de los productores para alterar sus tasas de operación, cerrar/reabrir faenas o adelantar el desarrollo de proyectos.
- Fallas de equipos, accidentes, huelgas, terremotos, conflictos políticos y, desde los años noventa, cada vez con mayor fuerza, operaciones financieras, tanto de cobertura como de inversión o especulación, que podrían afectar el precio por días, semanas o meses.

La suma de todos estos eventos determinará el precio del cobre en cada momento. La serie histórica de precios refleja esta descripción: el precio del cobre se mueve por decenas de años, sujeto a shocks temporales, en torno a un nivel definido por distintos shocks permanentes (Gráfico N°2).

La dificultad en el pronóstico del precio tiene que ver con lo complejo que resulta adelantarse a los shocks mencionados. Por otra parte, el potencial de pronóstico del precio del cobre se basa en el decaimiento esperado de los shocks temporales recientemente producidos [21]. La predictibilidad del decaimiento de los shocks temporales obedece al comportamiento de la producción de cobre que, siguiendo a los precios, siempre terminará ajustándose a la demanda en el mediano y largo plazo. A modo de ejemplo, desde fines del año 2003, la demanda mundial de cobre ha superado con creces a la oferta, lo que ha elevado considerablemente los precios. La oferta ha reaccionado lentamente a esta mayor demanda, esperándose que recién supere al consumo el año 2007 o 2008.

Gráfico N°2: Precio Promedio del Cobre a través del Tiempo, c/lb moneda 2006.

Fuente: Precio del Cobre: CRU y Cochilco. Deflactor del Precio: U.S. Department of Labor, Bureau of Labor Statistics.

Distinta podría ser la situación de más corto plazo, donde la mayor participación de los agentes financieros en las bolsas de metales, realizando operaciones que involucran cobre en función de la rentabilidad de otros activos, dificulta la proyección de la evolución del precio en los próximos minutos, semanas o meses [18].

En síntesis, estando el precio del cobre en el mediano y largo plazo ligado fuertemente al mercado físico, y evidenciando éste un patrón de comportamiento, las regularidades resultantes podrían ser captadas por modelos de pronóstico.

3. Construcción de modelos para pronosticar el precio del cobre

3.1. ¿Por qué pronosticar el precio del cobre utilizando redes neuronales?

El precio del cobre comúnmente se pronostica recurriendo a la econometría [4][10][14][21][29][30][31]. Para ello se utilizan ya sea grandes modelos estructurales con numerosas ecuaciones y variables, o los más sencillos modelos de series de tiempo construidos únicamente sobre la base del precio rezagado. Existen pruebas de un mejor desempeño de los modelos lineales de series de tiempo más simples (caminata aleatoria, AR1) [10], hecho que les ha ganado la categoría de benchmark en el pronóstico del precio.

Sin embargo, la existencia de evidencias de no linealidad en el comportamiento del precio [10], sumada a la falta de estudios sobre el pronóstico

del precio del cobre mediante redes neuronales motivan la realización de este trabajo.

3.2. Diseño general del ejercicio de pronóstico

En este trabajo se generaron pronósticos fuera de muestra, es decir, pronósticos referidos a datos ajenos a la muestra que sirve para la construcción de los modelos, lo que se traduce en una medida válida del real desempeño predictivo [27].

Dado que la evaluación del error de pronóstico será más representativa en la medida que se tengan más mediciones, sobre la base de distintos orígenes (años de altos precios, años de bajos precios) y de distintos alcances (pronósticos a 1 año, a 2 años, etc.), se optó por la generación de pronósticos dinámicos (pronósticos sobre pronósticos), a partir de modelos recalibrados (reestimados) periódicamente, utilizando una muestra de tamaño creciente y con un origen móvil [27]. Se optó por el uso de una muestra de tamaño creciente y no una de tamaño constante debido a que, a la luz de la coyuntura actual de precios altos, se consideró que las observaciones históricas podrían ser de utilidad [21].

En cuanto al alcance de los pronósticos, en el presente trabajo, el foco estará puesto en las proyecciones de corto y mediano plazo, en un horizonte de 1 a 6 años por sus potenciales aplicaciones en la industria del cobre (evaluación y control de los resultados de la gestión y de negocios, planificación de la producción y evaluación de proyectos) y en las finanzas públicas (prepuesto fiscal y Fondo de Compensación del Cobre o similar). No se evaluarán pronósticos de largo plazo, dado que los modelos ARIMA, utilizados como referente, no tienen un buen desempeño en proyecciones de largo alcance [10].

3.3. Datos utilizados

Las redes neuronales utilizadas en este estudio fueron construidas y entrenadas sobre la base del precio histórico del cobre y variables derivadas del mismo. Los pronósticos fuera de muestra generados con dichas redes se contrastaron con los obtenidos con la aplicación de los más exitosos modelos de series de tiempo. Dado que las redes neuronales pueden alimentarse con variables distintas al precio, la comparación puede parecer injusta, al, eventualmente, subestimarse la capacidad de pronóstico de las redes. Sin embargo, si las redes neuronales demuestran tener un mejor desempeño que los modelos de series de tiempo, se habrá establecido que las redes neuronales pueden sacar mayor provecho de los precios históricos y se tendrá una evaluación conservadora de su eficiencia predictiva que podrá alentar a la realización de investigaciones adicionales.

Por otra parte, la generación de proyecciones del precio empleando otras variables explicativas tiene el potencial problema de requerir el pronóstico de

dichas variables, lo que complica su utilización práctica.

Los datos de precios fueron sometidos a dos transformaciones para la construcción de modelos y la realización de pronósticos:

- Escalamiento de valores para eliminar el efecto de la inflación, expresando el precio en moneda constante.
- Aplicación de logaritmo natural a los datos en moneda constante para reducir su dispersión y atacar problemas de heterocedasticidad [16][17].

Los datos utilizados en este estudio corresponden al precio anual del cobre refinado de la Bolsa de Metales de Londres durante el periodo comprendido entre los años 1913 y 2006 (2006, promedio a junio). La elección de este periodo se fundamenta en tres razones:

- Si bien existen datos para el precio cobre en los Estados Unidos durante el siglo XIX, dichos valores corresponden a un precio de productores, no a un precio de mercado y obedecen a un tipo de minería muy distinta a la gran minería que comenzó a desarrollarse a partir de las primeras décadas del siglo XX (flotación, minería a rajo abierto, mecanización). Por este motivo se excluyeron los datos previos a 1910.
- El deflactor empleado por la Comisión Chilena del Cobre (COCHILCO) para expresar el precio del metal rojo en moneda constante es el Índice de Precios de Productores de los Estados Unidos (PPI All Commodities, Not Seasonally Adjusted) que registra valores desde 1913 en adelante [28].
- En el apartado "¿Es pronosticable el precio del cobre?" se describió al precio como una variable que presenta reversión a la media, en particular, a distintas medias a través del tiempo. Este comportamiento de mediano/largo plazo se observa en la ventana seleccionada, donde el logaritmo natural del precio no presenta tendencia y los test practicados rechazan la presencia de raíz unitaria.

Para medir el error de los pronósticos, se optó por un periodo de evaluación de 30 años, similar al empleado en otros trabajos [10][21] que permitió contar con un número de pronósticos de entre 30 (proyecciones a un año) y 25 (proyecciones a 6 años), abarcando tanto periodos con precios ascendentes como descendentes.

3.4. Modelamiento

Se mostrará el modelamiento con redes neuronales, modelos de series de tiempo (en particular modelos de la familia ARIMA) y modelos híbridos.

3.4.1. Redes Neuronales

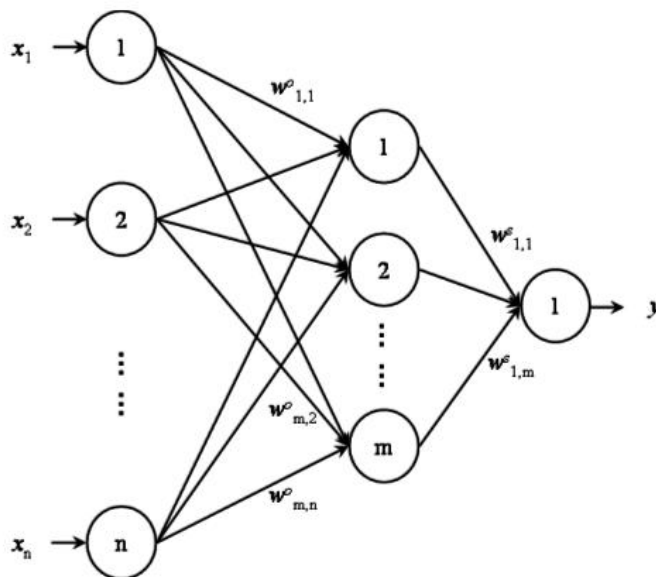
Una red neuronal puede ser descrita como un modelo de regresión no lineal cuya estructura se inspira en el funcionamiento del sistema nervioso. En términos generales, una red consiste en un gran número de unidades simples de proceso, denominadas neuronas, que actúan en paralelo y están conectadas mediante vínculos ponderados [13][23].

Cada neurona recibe inputs desde otras neuronas y genera un resultado que depende sólo de la información localmente disponible, ya sea almacenada internamente o plasmada en los ponderadores de las conexiones. El output generado por la neurona servirá de input para otras neuronas.

Mediante la adecuada modificación de los ponderadores de la red, en un proceso denominado aprendizaje, la red mejorará su desempeño en el desarrollo de la tarea para la cual fue construida.

El diseño de una red neuronal es una tarea compleja por la gran cantidad de decisiones que involucra tanto a nivel de su arquitectura como de su mecanismo de aprendizaje. Por los buenos resultados conseguidos en numerosas aplicaciones [15], en este estudio se emplearon redes MLP con una sola capa oculta, una sola neurona en la capa de salida y conexiones hacia delante (Figura N°1).

Figura N°1: Red Multilayer Perceptron.



Fuente: Hilerá, J. y Martínez, V. Redes Neuronales Artificiales. Fundamentos, Modelos y Aplicaciones. [13]

En cada una de las neuronas utilizadas, los datos de entrada se combinaron a través de una suma ponderada. Dicho resultado se transformó en las neuronas de la capa oculta, mediante la aplicación de la función tangente hiperbólica, antes de ser transferido a la capa de salida.

La red así construida, al recibir el vector de entrada (x_1, x_2, \dots, x_n) , genera el siguiente output:

$$y = w^{sc} + \sum_{j=1}^m w_{1,j}^s \cdot \tanh(w_j^{oc} + \sum_{i=1}^n w_{j,i}^o \cdot x_i) \quad (1)$$

Donde:

w^{oc}_j : Ponderador de la conexión con entrada de valor unitario y la neurona j de la capa oculta.

w^{sc} : Ponderador de la conexión con entrada de valor unitario y la neurona de la capa de salida.

$w^o_{j,i}$: Ponderador de la conexión entre la neurona i de la capa de entrada y la neurona j de la capa oculta.

$w^s_{1,j}$: Ponderador de la conexión entre la neurona j de la capa oculta y la neurona 1 de la capa de salida.

El mecanismo de aprendizaje aplicado fue off-line, donde se distinguen las fases de entrenamiento y aplicación de red, y por corrección de error, donde el ajuste de los ponderadores obedece al error respecto de la respuesta deseada. El método de aprendizaje elegido fue el de Levenberg-Marquardt por la mayor velocidad de convergencia de este último método en comparación con el método de descenso del gradiente. Adicionalmente, evaluaciones preliminares mostraron la obtención de un menor error de pronóstico.

Para el entrenamiento de la red los datos muestrales se dividieron en tres conjuntos disjuntos:

- Entrenamiento: conjunto que contiene los ejemplos que servirán para la modificación de los conectores neuronales.
- Testeo: fracción de los datos muestrales que no participa directamente en el entrenamiento de la red. Durante el proceso de aprendizaje, a intervalos regulares, se evaluó el desempeño de la red en este conjunto para verificar el cumplimiento de la meta en algún indicador crítico, como por ejemplo máximo error cuadrático medio deseado. Una vez alcanzado el valor meta se detuvo el entrenamiento (early-stopping).
- Evaluación: conjunto de datos que no participan ni directa ni indirectamente del proceso de aprendizaje. Los datos del conjunto de evaluación son lo que se utilizan para la evaluación de las proyecciones fuera de muestra.

Para el resto de los parámetros de diseño se siguió una estrategia de evaluación de distintos de valores, mediante el uso de grillas [8].

La aplicación de early-stopping requiere definir el tamaño del conjunto de testeo. Se seleccionó una fracción variable de los últimos ejemplos disponibles dentro de la muestra: 10 %, 20 % o 30 % [16], que no participó, directamente, en el entrenamiento de las redes neuronales.

Respecto al número de neuronas en la capa de entrada, se construyeron y entrenaron redes con 3, 6 o 9 neuronas en la capa de entrada, alimentadas con los correspondientes rezagos del precio, mientras que para la capa oculta se utilizaron entre 1 y 6 neuronas. Las combinaciones seleccionadas fueron determinadas por el número de conexiones asociadas y el número de ejemplos del conjunto de entrenamiento. Si bien sólo existen teoremas que relacionan el número de conexiones de la red con el número de ejemplos de entrenamiento para ciertos tipos de redes (multicapas, con entradas y salidas binarias y función de activación escalón) [25], al menos debe respetarse que el número de ejemplos supere, con alguna holgura, el número de coeficientes a estimar (ponderadores) para así evitar el overfitting. Las combinaciones presentadas en la Tabla N°1 cumplen con este requisito, ya que, en promedio, el número de ejemplos de entrenamiento quintuplica el número de coeficientes [16] (el número de ejemplos de entrenamiento depende del tamaño de la muestra, de la fracción de datos muestrales destinada al conjunto de testeo y del número de neuronas de la capa de entrada, moviéndose en un rango de entre 55 y 90 ejemplos).

Tabla N°1: Neuronas en la Capa de Entrada, en la Capa Oculta y Número de Conexiones.

Neuronas en la capa de entrada	Neuronas en la capa oculta	Número de conexiones
3	1	6
3	2	11
3	3	16
3	4	21
3	5	26
3	6	31
6	1	9
6	2	17
6	3	25
6	4	33
9	1	12
9	2	23
9	3	34

Nota: El número de conexiones considera las conexiones entre capas (incluyendo la capa de salida) y las conexiones correspondientes a las entradas de valor unitario.

Dado que los ponderadores iniciales de las conexiones determinan el punto de origen del proceso de aprendizaje (minimización de error del modelo), se recomienda repetir el entrenamiento de cada arquitectura unas 15 veces, redefiniendo aleatoriamente dichos pesos [8][26], obteniéndose así 15 redes dis-

tintas para una misma arquitectura. Si a esto agregamos que, como ya fue mencionado, cada arquitectura se entrenó utilizando conjuntos de testeo definidos de 3 maneras distintas, el total de redes construidas para una misma combinación de neuronas en la capa de entrada y neuronas en la capa oculta llega a 45 y el total de redes a entrenadas y evaluadas llega a 585.

Las redes neuronales se recalibraron para incorporar la información adicional disponible a medida que se realizaban las evaluaciones. En evaluaciones preliminares, se observó que un mayor número de recalibraciones no siempre se traducían en un mejor pronóstico. Por este motivo, cada red se recalibró 3, 6 y 10 veces.

Para entrenar y evaluar las distintas redes, se empleó el lenguaje de programación MATLAB y su toolbox de redes neuronales (MATLAB 6.1, The MathWorks Inc.). Si bien el entrenamiento y evaluación de cada red tomó pocos segundos, el gran número de redes estimadas requirió de tiempos de cálculo de hasta 6 horas (Pentium IV, 3.0 GHz, 1 GB RAM).

3.4.2. Series de tiempo

En este trabajo se emplearon modelos estocásticos y deterministas de series de tiempo. Los modelos estocásticos utilizados forman parte de la familia ARIMA [22]:

$$\Delta^d Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t + \Theta_1 \epsilon_{t-1} + \dots + \Theta_q \epsilon_{t-q} \quad (2)$$

Donde:

$$\Delta y_t = y_t - y_{t-1}$$

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$$

ϕ_i, θ_j : coeficientes de las porciones autorregresivas de media móvil del proceso.

El examen de la función de autocorrelación parcial de la serie en niveles, evaluaciones preliminares realizadas y los resultados conseguidos en otros estudios [10][21], llevaron a la construcción de un set de modelos tentativos cuyos coeficientes son significativos para los datos en niveles en el periodo de interés¹:

Se realizaron pruebas con modelos de varianza condicional constante y modelos GARCH (Heterocedasticidad Condicional Autorregresiva Generalizado) [9]:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_p \epsilon_{t-p}^2 + \lambda_1 \sigma_{t-1}^2 + \dots + \lambda_q \sigma_{t-q}^2 \quad (3)$$

¹Para los modelos de datos en primeras diferencias sólo se obtienen coeficientes significativos con los modelos AR2, MA1 y MA2.

Tabla N°2: Modelos ARIMA Utilizados en el Ejercicio de Proyección.

AR1	MA3	MA1MA2
AR2	MA4	AR3MA1MA2
AR3	AR1MA1	AR1AR2AR3
AR4	AR1MA2	AR1AR2AR3MA1MA2MA3
MA1	AR2MA1	MA1MA2MA3MA4
MA2	AR3MA1	MA1MA2MA3MA4MA5

Adicionalmente, y sólo con fines comparativos, se modeló el precio del cobre como un camino aleatorio con y sin drift.

$$y_t = \delta + y_{t-1} + \epsilon_t \quad (4)$$

Por último, en la línea de los modelos deterministas, se modeló el precio del cobre como un promedio móvil de 3, 6 y 9 años. Es interesante incluir el promedio móvil, dado que, hasta hace algunos años, fue el método utilizado para determinar el precio de referencia del Fondo de Compensación del Cobre (promedio móvil de 6 periodos) [3][10].

$$y_{t+1} = \frac{1}{6} \cdot (y_t + y_{t-1} + \dots + y_{t-5}) \quad (5)$$

Para la construcción de modelos estocásticos se utilizó el software EViews 4.1 (Quantitative Micro Software). Las rutinas programadas en EViews para la generación y evaluación de proyecciones fuera de muestra sólo necesitaron un par de segundos para ejecutar los cálculos correspondientes a cada modelo, y la evaluación del set completo de modelos tentativos fue cosa de pocos minutos. Para el resto de los modelos se empleó la planilla de cálculo Microsoft Excel 2002 (Microsoft Corporation).

Las evaluaciones realizadas mostraron que los mejores resultados se conseguían recalibrando el modelo cada vez que un nuevo dato histórico se incorporaba al conjunto de entrenamiento.

3.4.3. Modelos Híbridos

Adicionalmente se combinaron los modelos de series de tiempo y las redes neuronales de la siguiente manera [1], [2]:

- Con modelos de series de tiempo:
 - Se realizaron proyecciones fuera de muestra.
 - Se calcularon los errores de proyección fuera de muestra.
- Con redes neuronales:

- A partir de los errores calculados con series de tiempo, se entrenó una red neuronal.
- Se pronosticaron los errores de proyección fuera de muestra.
- Se corrigieron las proyecciones fuera de muestra vía modelos de series de tiempo con los errores pronosticados mediante redes neuronales.

3.5. Evaluación del desempeño predictivo: medición del error

Para la evaluación del desempeño predictivo se emplean diferentes indicadores que cuantifican qué tan cerca está la variable pronosticada de su serie de datos correspondiente [22]. Una de las medidas más utilizadas es el Promedio del Error Porcentual Absoluto (MAPE) [12][22]:

$$MAPE = \frac{1}{T} \left(\sum_{t=1}^T APE_t \right) = \frac{1}{T} \left(\sum_{t=1}^T \frac{|Y_t^s - Y_t^a|}{Y_t^a} \right) \cdot 100 \quad (6)$$

Donde:

APE : error porcentual absoluto

Y_t^s : valor pronosticado de Y_t

Y_t^a : valor real de Y_t

T: número de periodos

El MAPE mide el valor medio del error absoluto en términos porcentuales al valor real de la variable.

Para evaluar la dispersión de los errores se puede calcular el Desvío Estándar del Error porcentual absoluto (APE).

$$\text{Desvío Estándar APE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (APE_t - MAPE)^2} \quad (7)$$

Otra medida del error de pronóstico comúnmente empleada es la Raíz Cuadrática Media del Error (RMSE):

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t^s - Y_t^a)^2} \quad (8)$$

Donde:

Y_t^s : Valor pronosticado de Y_t

Y_t^a : Valor real de Y_t

T: número de periodos

El RMSE mide la dispersión de la variable simulada en el curso del tiempo, penalizando fuertemente los errores grandes al elevarlos al cuadrado. Esta característica hace que el RMSE se recomiende cuando el costo de cometer un error es aproximadamente proporcional al cuadrado de dicho error [20].

No siempre el modelo que genere pronósticos con un menor MAPE generará los pronósticos con el menor RMSE y viceversa, por lo que en la selección de los mejores modelos de pronóstico se hace necesario establecer la medida de error a utilizar para la elaboración del ranking de desempeño.

Dado que una mala estimación del precio futuro del cobre se traduce en una pérdida de ingresos proporcional al tamaño del error, el MAPE, y no el RMSE, parece ser la medida de desempeño más adecuada. A esto se suma la ventaja práctica del MAPE de no requerir ser acompañado por la media para dimensionar la magnitud del error. Luego, la medida de error que se empleará para identificar los modelos de mejor desempeño será el MAPE.

Sin perjuicio de lo anterior, para los mejores modelos identificados se presentarán tanto el MAPE, el desvío estándar del APE y el RMSE de sus pronósticos.

4. Resultados obtenidos

4.1. Resultados y comparación de modelos

En esta sección se comparan los errores de pronóstico, a distintos plazos, arrojados por los mejores modelos de series de tiempo, las mejores redes neuronales y la caminata aleatoria. En todos los alcances, las redes neuronales entregaron mejores pronósticos (menor MAPE) que los modelos de series de tiempo, y estos últimos, a su vez, superaron a la caminata aleatoria. La mayor exactitud que registran las redes neuronales respecto a los modelos de series de tiempo es marginal en pronósticos a 1 año, con una reducción del MAPE de 4%, pero se hace más significativa en pronósticos a 2 y más años, con una reducción de 30%. En cuanto a la dispersión del error, las redes que generan pronósticos de menor MAPE muestran una dispersión en sus errores similar a la de los mejores modelos de series de tiempo.

En la medición del error aplicando RMSE, nuevamente los mejores modelos de redes neuronales tuvieron un mejor desempeño que los modelos ARIMA y la caminata aleatoria, con reducciones del RMSE de entre 1% y 23% y de entre 4% y 37% respectivamente.

La notación empleada para identificar las redes es $\text{Red}(a_1; a_2; a_3; a_4; a_5)$,

Tabla N°3: Comparación de Resultados. MAPE (%) en Pronósticos a Distintos Alcances.

a 1 año	
Red(3; 5; 80/20; 14; 3)	13,48%
AR2MA1	14,08%
Caminata Aleatoria	15,26%
% Reducción Error	
Redes v/s ARIMA	-4,20%
Redes v/s Caminata	
Aleatoria	-11,70%

a 2 años	
Red(6; 3; 70/30; 8; 10)	16,13%
AR1.GARCH(0,2)	22,98%
Caminata Aleatoria	23,82%
% Reducción Error	
Redes v/s ARIMA	-29,80%
Redes v/s Caminata	
Aleatoria	-32,30%

a 3 años	
Red(3; 2; 80/20; 4; 3)(dep)	18,66%
AR2MA1.GARCH(0,2)	27,02%
Caminata Aleatoria	29,20%
% Reducción Error	
Redes v/s ARIMA	-30,90%
Redes v/s Caminata	
Aleatoria	-36,10%

a 4 años	
Red(3; 6; 80/20; 11; 3)(dep)	20,63%
AR2MA1.GARCH(0,2)	28,87%
Caminata Aleatoria	33,30%
% Reducción Error	
Redes v/s ARIMA	-28,50%
Redes v/s Caminata	
Aleatoria	-38,00%

a 5 años	
Red(3; 1; 80/20; 13; 3)(dep)	21,54%
AR2MA1.GARCH(0,2)	31,08%
Caminata Aleatoria	37,37%
% Reducción Error	
Redes v/s ARIMA	-30,70%
Redes v/s Caminata	
Aleatoria	-42,40%

a 6 años	
Red(6; 3; 70/30; 9; 6)(uvp)	20,33%
AR2MA1.GARCH(0,2)	31,07%
Caminata Aleatoria	40,14%
% Reducción Error	
Redes v/s ARIMA	-34,60%
Redes v/s Caminata	
Aleatoria	-49,30%

donde:

a1: Número de neuronas de la capa de entrada.

a2: Número de neuronas de la capa oculta.

a3: Razón entre el número de ejemplos efectivamente utilizados en el entrenamiento y el número de ejemplos empleados en el testeo (early-stopping).

a4: Número correlativo de la inicialización de la red.

a5: Número de recalibraciones efectuadas.

uvp: Última variación del precio. Res incluye la diferencia entre los dos últimos rezagos del precio en logaritmo natural.

dep: Desviación estándar del precio en los últimos periodos. Red incluye la desviación estándar de los precios rezagados que alimentan la capa de entrada.

En las Tablas N°4 y N°5, se observa que la ventaja conseguida por las mejores redes neuronales por sobre los modelos ARIMA al evaluar mediante MAPE, se hace menos significativa al emplear una mediada como el RMSE

Tabla N°4: Comparación de Resultados. Disminución del MAPE Conseguido con las Mejores Redes Neuronales Seleccionadas.

Reducción del MAPE respecto a:	a 1 año	a 2 años	a 3 años	a 4 años	a 5 años	a 6 años
ARIMA	-4,2%	-29,8%	-30,9%	-28,5%	-30,7%	-34,6%
Caminata Aleatoria	-11,7%	-32,3%	-36,1%	-38,0%	-42,4%	-49,3%

Tabla N°5: Comparación de Resultados. Disminución del RMSE Conseguido con las Mejores Redes Neuronales Seleccionadas.

Reducción del RMSE respecto a:	a 1 año	a 2 años	a 3 años	a 4 años	a 5 años	a 6 años
ARIMA	-1%	-23%	-6%	-4%	-13%	-18%
Caminata Aleatoria	-4%	-30%	-20%	-22%	-31%	-37%

Tabla N°6: Comparación de Resultados. Disminución Desvío Estándar del Error Porcentual Absoluto Conseguido con las Mejores Redes Neuronales Seleccionadas.

Reducción del MAPE respecto a:	a 1 año	a 2 años	a 3 años	a 4 años	a 5 años	a 6 años
ARIMA	0,4%	-8,0%	13,5%	-0,3%	-1,5%	-25,4%
Caminata Aleatoria	-2,4%	-23,3%	-18,5%	-23,5%	-22,7%	-39,6%

que penaliza los errores de mayor tamaño. Sin embargo, se debe recordar que los mejores modelos de redes neuronales fueron seleccionados por su menor MAPE y no por su menor RMSE.

En cuanto a la desviación estándar del error porcentual absoluto, mediante el uso de redes neuronales se observa una reducción de la dispersión en la mayor parte de los alcances.

Utilizando un modelo híbrido se generaron proyecciones a un año, buscando determinar la mayor exactitud que se podría conseguir con el uso combinado de series de tiempo y redes neuronales. Para la aplicación del modelo híbrido se trabajó con el modelo AR2MA1, por ser éste el de mejor desempeño en pronósticos a un año, mientras que, a nivel de las redes se aplicó la misma estrategia de búsqueda en torno a una grilla descrita en la sección de redes neuronales. Sin embargo, dado que se debe trabajar con residuos de proyecciones fuera de muestra y que para construir el primer modelo AR2MA1 se utilizan los datos históricos del periodo 1913 a 1942, sólo se cuenta con 64 datos de residuos para entrenar y evaluar redes. Debido a esto se trabajó con una grilla más pequeña cuidando mantener una adecuada razón entre el número de datos disponibles y el número de conexiones.

A diferencia de lo ocurrido con las redes empleadas para proyectar el precio, en la proyección de los residuos, los mejores resultados no se obtuvieron actualizando los ponderadores con el método de Levenberg-Marquardt, sino que con el método de descenso del gradiente. Esto podría indicar que la función de error para este problema es menos suave, por lo que el método de

Tabla N°7: Neuronas en la Capa de Entrada, en la Capa Oculta y Número de Conexiones.

Neuronas en la capa de entrada	Neuronas en la capa oculta	Número de conexiones
3	1	6
3	2	11
3	3	16
3	4	21
6	1	9
6	2	17
6	3	25

Nota: El número de conexiones considera las conexiones entre capas (incluyendo la capa de salida) y las conexiones correspondientes a las entradas de valor unitario.

Levenberg-Marquardt no converge al mínimo global. La mejor red resultante fue Red(3; 2; 80/20; 9; 3).

La siguiente tabla compara los errores conseguidos con los mejores modelos identificados.

Tabla N°8: Comparación de Resultados. MAPE (%) en Pronósticos a 1 Año.

	a 1 año
Método Híbrido	13,284%
Red(3; 5; 80/20; 14; 3)	13,481%
AR2MA1	14,078%
Caminata Aleatoria	15,264%
% Reducción Error	
Método Híbrido v/s Red	-1,5%
Método Híbrido v/s ARIMA	-5,6%
Método Híbrido v/s Caminata Aleatoria	-13,0%

En el conjunto de 30 pronósticos evaluados, el modelo híbrido entregó mejores resultados que el más exitoso modelo ARIMA en 16 oportunidades, logrando una disminución del MAPE de 5,6 %.

Otro dato interesante al momento de elegir una herramienta de pronóstico es la capacidad de respuesta frente a grandes cambios de precios. Entre 1977 y 2006, en 14 oportunidades el cambio del precio respecto al año anterior superó el 15

Si la ventana de pronósticos se divide en 3 subperiodos de igual duración, se aprecia que el MAPE del modelo híbrido es menor al MAPE del modelo ARIMA durante los primeros 20 años. En el tercer subperiodo, la comparación favorece al modelo ARIMA, gracias a los menores errores conseguidos en los últimos 3 años.

A pesar de estos buenos resultados, la comparación del modelo híbrido con la mejor red neuronal muestra sólo una mejora marginal del MAPE y un deterioro, también marginal, del RMSE.

Tabla N°9: MAPE (%) por subperiodo.

	Pronóstico AR2MA1	Método Híbrido
1977-1986	13,030%	10,453%
1987-1996	13,290%	12,985%
1997-2006	15,914%	16,414%

4.2. Comparación con proyecciones de expertos

Es interesante comparar los resultados conseguidos con las proyecciones publicadas en el pasado por analistas del mercado. Lamentablemente, no fue posible encontrar pronósticos de expertos para el periodo completo entre los años 1977 y 2006, por lo que la comparación se restringió a los últimos 20 años. Este es un elemento que debe ser considerado en la comparación, dado que los mejores modelos de pronóstico identificados se seleccionaron por la calidad de su desempeño en el periodo de 30 años mencionado y no en la ventana de tiempo 1987 - 2006, por lo que el contraste puede subestimar la precisión de los modelos híbridos y de las redes neuronales.

La Tabla N°10 muestra los errores de pronóstico conseguidos con los mejores modelos identificados y los errores de proyección del consultor Brook Hunt & Associates, una de las empresas más reputadas de la industria.

En el conjunto de 20 pronósticos a 1 año, el modelo híbrido entregó mejores resultados que el analista en 15 oportunidades, consiguiendo disminuir el MAPE en 33 % y disminuir el desvío estándar del APE en 14 %.

En proyecciones a 2 años (19 evaluaciones), el modelo híbrido entregó mejores resultados que el analista en 16 oportunidades, consiguiendo disminuir el MAPE en 43 % y el desvío estándar del APE en 22 %.

En cuanto al pronóstico en situaciones con grandes cambios de precios, entre 1987 y 2006, en 10 oportunidades el cambio del precio respecto al año anterior superó el 15 %, entregando el modelo híbrido los mejores pronósticos, en términos de MAPE, en 8 ocasiones.

Finalmente, en términos de RMSE, los modelos de Inteligencia Computacional entregaron menores errores con reducciones de 16 % en proyecciones a 1 año y de 40 % en proyecciones a 2 años.

Tabla N°10: Precio, Pronósticos y Error Porcentual Absoluto (APE).

	Precio Histórico	Pronósticos a 1 año		Pronósticos a 2 años	
		APE Método Híbrido	Proyección Analista	APE Red	Proyección Analista
1987	129,2	14,120%	21,441%		
1988	180,8	24,734%	33,015%	37,414%	39,714%
1989	188,5	5,718%	14,628%	11,768%	36,490%
1990	170,3	5,923%	22,744%	3,698%	34,645%
1991	149,2	2,925%	27,533%	9,405%	38,021%
1992	144,8	11,035%	35,748%	7,800%	43,536%
1993	119,5	6,848%	9,712%	18,976%	20,265%
1994	142,8	30,341%	34,611%	4,327%	38,038%
1995	175	12,549%	18,167%	20,231%	42,942%
1996	133,3	15,657%	20,315%	10,137%	12,411%
1997	132,5	1,923%	7,967%	20,919%	1,721%
1998	98,7	36,032%	6,704%	11,381%	13,373%
1999	93,2	6,515%	4,735%	46,606%	6,472%
2000	101,6	4,369%	1,572%	15,780%	14,939%
2001	87,4	19,739%	39,731%	0,340%	24,361%
2002	88,3	5,832%	4,746%	17,570%	55,704%
2003	95,8	1,522%	6,483%	0,673%	15,193%
2004	145,3	31,039%	21,208%	21,782%	35,871%
2005	173,9	14,547%	23,920%	29,818%	31,506%
2006	275,3	42,624%	43,568%	36,949%	63,820%
MAPE 1987-2006		13,284%	19,928%	17,083%	29,310%
Desvío Estándar APE		10,657%	12,463%	12,527%	15,991%

Fuente: Brook Hunt & Associates. Datos corresponden a proyecciones realizadas en Diciembre, salvo para los años 1990, 1991 y 1994, donde las proyecciones fueron realizadas en Septiembre.

Así los mejores modelos de pronóstico identificados generaron proyecciones que, en promedio y en desviación estándar del error, son más precisas que las proyecciones publicadas por Brook Hunt and Associates.

5. Conclusiones

En las evaluaciones realizadas, los modelos no lineales de pronóstico, basados en redes neuronales multilayer perceptron, superaron la exactitud de los modelos lineales de series de tiempo más comúnmente usados. De esta manera, las redes neuronales demostraron ser capaces de aprovechar mejor la información contenida en los precios históricos considerados en el estudio.

El modelo con mejores pronósticos a un año, para el periodo 1977 - 2006, fue un modelo híbrido que, utilizando conjuntamente modelos AR2MA1 y re-

des neuronales entrenadas con residuos, promedió un error porcentual absoluto (MAPE) de 13,284 %. Esta cifra significa una reducción del error, en términos relativos, de 5,6 % respecto al mejor modelo ARIMA, y de 13 % respecto a la caminata aleatoria.

Dicho modelo híbrido también mostró un mejor desempeño que ARIMA en situaciones con grandes cambios de precios (variaciones de precios superiores a 15 % de un año a otro), generando proyecciones con menor MAPE en 8 de 14 eventos ocurridos entre 1977 y 2006.

La ventaja de los modelos basados en redes neuronales por sobre los modelos ARIMA se acrecienta en pronósticos a 2 y más periodos, llegando a una reducción del MAPE del orden de 30 % en términos relativos.

Aún cuando los mejores modelos basados en redes neuronales fueron seleccionados por su menor MAPE y no por su menor RMSE, estos también consiguieron una reducción del RMSE respecto a los mejores modelos de series de tiempo, aunque de menor tamaño (entre 1 % y 23 %).

Así como la estimación con modelos GARCH permitió mejorar los resultados en los modelos ARIMA, la inclusión de variables derivadas del precio, en especial su desviación estándar, también mejoró las proyecciones de las redes neuronales.

En cuanto a la arquitectura de las redes neuronales más exitosas, éstas presentaron 3 o 6 neuronas alimentadas por rezagos del precio en la capa de entrada, y, en proyecciones a 3 y más años, una neurona adicional alimentada con la desviación estándar del precio o su última variación. En la capa oculta, las redes neuronales más exitosas mostraron un diverso número de neuronas (1, 2, 3, 5, 6).

Se observa que, con un modelamiento no lineal, los últimos 6 rezagos del precio aportan información útil para la generación de pronósticos, hecho que no se aprecia en los modelos ARIMA, donde modelos de ese orden fueron descartados por tener coeficientes no significativos.

La comparación de los pronósticos publicados por expertos en el mercado del cobre versus los resultados conseguidos con modelos de Inteligencia Computacional reveló la mayor precisión de estos últimos con reducciones del MAPE de 33 % y 43 % en proyecciones a 1 y 2 años respectivamente.

El mejor modelo híbrido también mostró un mejor desempeño que las

proyecciones de expertos en situaciones con grandes cambios de precios (variaciones de precios superiores a 15 % de un año a otro), generando proyecciones con menor MAPE en 8 de 10 eventos ocurridos entre 1987 y 2006.

La mayor exactitud de los pronósticos conseguidos utilizando redes neuronales va acompañada por una mayor complejidad en el diseño. En esta tesis se optó por sensibilizar sólo algunos parámetros del diseño de una red y, aún así, el número de cálculos y el tiempo requerido fue muy superior al de los modelos de series de tiempo (horas versus minutos).

Si bien las redes neuronales superaron a los modelos de series de tiempo, los resultados conseguidos con la aplicación de modelos híbridos muestran que, mediante el trabajo conjunto con modelos de series de tiempo y redes neuronales, es posible conseguir pronósticos más precisos para el precio del cobre. Así la pregunta planteada en el título esta tesis, en vez de tener una respuesta excluyente, tiene una respuesta integradora que alienta nuevos estudios en el área.

Las evidencias entregadas avalan a las redes neuronales como una herramienta atractiva para el pronóstico del precio del cobre, animando el desarrollo de futuros estudios que aporten más antecedentes sobre su potencial y motivando la investigación de metodologías de diseño que posibiliten el aprovechamiento de dicha capacidad de generalización.

6. Trabajos Futuros

Continuando con la evaluación del potencial de las redes neuronales como herramienta de pronóstico para el precio del cobre, podría evaluarse el uso de otras estrategias de aprendizaje: otros algoritmos, aplicación de momentum y pruning. Así mismo, también resulta interesante probar el efecto de una definición aleatoria de los conjuntos de entrenamiento y testeo en la capacidad de generalización.

A nivel de la arquitectura de las redes, se podría explorar el uso de n neuronas en la capa de salida, siendo n el número de periodos hacia adelante que se desea pronosticar.

En el área de los modelos híbridos, sería un aporte contar con una evaluación de su desempeño en pronósticos a 2 y más periodos.

Otro trabajo interesante para la entrega de evidencias sobre el atractivo de las redes neuronales consistiría en el entrenamiento de redes incluyendo variables adicionales al precio y sus derivaciones, como por ejemplo: stocks en semanas de consumo, indicadores de actividad económica y expectativas.

En la línea del diseño de redes neuronales, futuros estudios deberían explorar algunas de las heurísticas propuestas, evaluándolas y proponiendo variantes que aseguren la construcción de redes con elevada capacidad de generalización.

Una mayor comprensión del funcionamiento de las redes también puede facilitar su diseño. La extracción de reglas a partir de su funcionamiento, la sensibilización de los parámetros de la red y de las variables de entrada, junto con el análisis detallado de los residuos son áreas de interés para futuros trabajos.

Finalmente, los resultados conseguidos en este trabajo pueden ser un referente para la evaluación del potencial de otras herramientas de pronóstico, como las máquinas de soporte vectorial (support vector regression); ver por ejemplo [11].

Agradecimientos: Los autores les agradecen al Instituto Científico Milenio "Sistemas Complejos de Ingeniería" P04-066-F (www.sistemasdeingenieria.cl), por el apoyo brindado en la elaboración y financiamiento de este estudio.

Referencias

- [1] Aburto, L., Weber, R. (2007a): Improved Supply Chain Management based on Hybrid Demand Forecasts. *Applied Soft Computing* 7, No. 1, 136-144
- [2] Aburto, L., Weber, R. (2007b): A Sequential Hybrid Forecasting System for Demand Prediction. In: Petra Perner (Ed.): *Machine Learning and Data Mining in Pattern Recognition*. LNAI 4571, Springer Verlag, Berlin, Heidelberg, 518-532
- [3] CEBALLOS, J. y TILTON, J. 2005. Análisis del Fondo de Compensación del Cobre de Chile. En: LAGOS, G (Edit.). *Minería y Desarrollo. Foro en Economía de Minerales*, vol. III.
- [4] CIUDAD, J. 2005. Determinantes del Precio Spot del Cobre en las Bolsas de Metales. En: COCHILCO (Edits). *Mercado del Cobre y Desarrollo Sustentable en la Minería*. Santiago, Comisión Chilena del Cobre. pp. 145-187.

- [5] CLEMENTS, M. y HENDRY D. 2001. Forecasting Non-Stationary Economic Time Series. The MIT Press.
- [6] COMISION CHILENA DEL COBRE. 2006. [en línea]Anuario: Estadísticas del Cobre y Otros Minerales 1986-2005. <http://www.cochilco.cl/anm/articlefiles/456-ANUARIO2005-PDF%2811-AGO-06%29.pdf>.
- [7] COMISION CHILENA DEL COBRE. [en línea] Precio del Cobre. <http://www.cochilco.cl>.
- [8] CRONE, S. 2005. Stepwise Selection of Artificial Neural Network Models for Time Series Prediction. *Journal of Intelligent Systems* 14 (2-3). pp. 99-122.
- [9] ENDERS, W. 2004. Applied Econometric Time Series. 2^aed., John Wiley & Sons, Inc.
- [10] ENGEL, E. y VALDES, R. 2002. Prediciendo el Precio del Cobre: ¿Más Allá del Camino Aleatorio?. En: MELLER, P. (Edit.). Dilemas y Debates en Torno al Cobre. Santiago, Dolmen Ediciones. pp. 269-290.
- [11] Guajardo, J., Weber, R., Miranda, J. (2006): A Forecasting Methodology Using Support Vector Regression and Dynamic Feature Selection. *Journal of Information & Knowledge Management* 5, No. 4, 329-335
- [12] HANKE, J. y REITSCH, A. 1995. Estadística Para Negocios. Irwin Professional Publishing.
- [13] HILERA, J. y MARTINEZ, V. 1995. Redes neuronales Artificiales. Fundamentos, Modelos y Aplicaciones. Madrid, RA-MA, Addison-Wesley Iberoamericana, S.A.
- [14] INTERNATIONAL MONETARY FUND. 2006. World Economic Outlook September 2006, Financial Systems and Economic Cycles. Washington, D.C., USA..
- [15] ISASI, P. y GALVAN, I. 2004. Redes de Neuronas Artificiales. Un Enfoque Práctico. Madrid, Pearson Educación, S.A.
- [16] KAASTRA, I. y BOYD, M. 1996. Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing* 10(3): 215-236.
- [17] MADDALA, G. S. 1996. Introducción a la Econometría. 2^aed., Prentice-Hall Hispanoamericana, S. A.

- [18] MARSHALL, I. y SILVA, E. 1998. Fluctuaciones del Precio del Cobre. Informe Macroeconómico para la Empresa. N°35: 38-60. Instituto de Economía, Pontificia Universidad Católica de Chile.
- [19] MELLER, P. 2002. El Cobre Chileno y la Política Minera. En: MELLER, P. (Edit.). Dilemas y Debates en Torno al Cobre. Santiago, Dolmen Ediciones. pp. 11-77.
- [20] NAU, R. 2006. What's the Bottom Line? How to Compare Models. [en línea]. <http://www.duke.edu/~rnau/compare.htm>.
- [21] PHILLIPS, S. y SWISTON, A. 2002. Forecasting Copper Prices in the Chilean context. Chile: Selected Issues. IMF Country Report No.02/163. International Monetary Fund.
- [22] PINDYCK, R. S. y RUBINFELD D. L. 2001. Econometría: Modelos y Pronósticos. 4ªed. México, D.F., McGraw-Hill/Interamericana Editores, S. A.
- [23] REED, R. D. y MARKS, R. J. II. 1999. Neural Smithing. The MIT Press.
- [24] ROMAGUERA, P. 1991. Las Fluctuaciones del Precio del Cobre y su Impacto en la Economía Chilena. CIEPLAN. Notas técnicas N°143.
- [25] SILIPO, R. 2003. Neural Networks. En: BERTHOLD, M. y HAND, D. (Edits.). Intelligent Data Analysis. Nueva York, Springer-Verlag New York, Inc. pp: 269-320.
- [26] TANG, Z. y FISHWICK, A. 1993. Feed-Forward Neural Nets as Models for Time Series Forecasting. ORSA Journal of Computing, 5(4): 374-386.
- [27] TASHMAN, L. 2000. Out-of-Sample Tests of Forecasting Accuracy: An Analysis and Review. International Journal of Forecasting (16): 437-450.
- [28] U.S. DEPARTMENT OF LABOR, Bureau of Labor Statistics. Producer Price Index. Commodity Data. All Commodities Not Seasonally Adjusted. <http://data.bls.gov/cgi-bin/surveymost>. [Consulta: 5 de agosto de 2006].
- [29] ULLOA, A. 2002. Tendencia y Volatilidad del Precio del Cobre. En: MELLER, P. (Edit.). Dilemas y Debates en Torno al Cobre. Santiago, Dolmen Ediciones. pp. 291-337.
- [30] VIAL, J. 1988. An Econometric Study of the World Copper Market, Ph.D. Dissertation. University of Pennsylvania. CIEPLAN. Notas Técnicas N°112.

- [31] VIAL, J. 2004. Modeling Commodity Markets in the Global Economy: Familiar Finding and New Strategies. The Earth Institute at Columbia University. Center on Globalization and Sustainable Development. Working Paper N°18.

