

MODELADO DEL PRECIO SPOT DE LA ELECTRICIDAD EN BRASIL USANDO UNA RED NEURONAL AUTORREGRESIVA

ELECTRICITY SPOT PRICE MODELLING IN BRASIL USING AN AUTOREGRESSIVE NEURAL NETWORK

Juan D. Velásquez¹ Isaac Dyner² Reinaldo C. Souza³

Recibido: 4 de septiembre de 2006, aceptado: 17 de agosto de 2007
Received: September 4, 2006 Accepted: August 17, 2007

RESUMEN

Una red neuronal autorregresiva es estimada para el precio mensual brasileño de corto plazo de la electricidad, la cual describe mejor la dinámica de los precios que un modelo lineal autorregresivo y que un perceptrón multicapa clásico que usan las mismas entradas y neuronas en la capa oculta. El modelo propuesto es especificado usando un procedimiento estadístico basado en el contraste del radio de verosimilitud. El modelo pasa una batería de pruebas de diagnóstico. El procedimiento de especificación propuesto permite seleccionar el número de unidades en la capa oculta y las entradas a la red neuronal, usando pruebas estadísticas que tienen en cuenta la cantidad de los datos y el ajuste del modelo a la serie de precios. La especificación del modelo final demuestra que el precio para el próximo mes es una función no lineal del precio actual, de la energía afluente actual y de la energía almacenada en el embalse equivalente en el mes actual y dos meses atrás.

Palabras clave: Red neuronal autorregresiva, procedimiento de especificación, precio spot de electricidad, modelado no lineal de series temporales.

ABSTRACT

An autoregressive neural network model is estimated for the monthly Brazilian electricity spot price, which describes the prices dynamics better than a linear autoregressive model and a classical multilayer perceptron using the same input and neurons in the hidden layer. The proposed model is specified using a statistical procedure based on a likelihood ratio test. The model passes a battery of diagnostic tests. The proposed specification procedure allows us to select the number of units in hidden layer and the inputs to the neural network based on statistical tests, taking into account the number of data and the model fitting to the price time series. The final model specification demonstrates that the price for the next month is a nonlinear function of the current price, the current energy inflow, and the energy saved in the equivalent reservoir in the current month and two months ago.

Keywords: Autoregressive neural network, specification procedure, electricity spot price, nonlinear time series modelling.

INTRODUCCIÓN

El proceso de reforma de los sectores eléctricos en el mundo ha creado nuevos mercados competitivos diseñados bajo el principio de eficiencia económica [1,2,3] que incentivan la entrada de capitales privados a los negocios de generación, transmisión y distribución. Con la llegada de la nueva estructura competitiva del mercado, se implementaron nuevos mecanismos para la formación de los precios de la electricidad, obligando al desmonte de todas las protecciones y controles existentes. Es así entonces,

como las series de precios de la electricidad poseen ahora características únicas que no han estado presentes en los mercados eléctricos tradicionales; tal como es indicado por Pilipovic [4], los precios de la electricidad en los mercados desregulados o liberalizados se ven influenciados en el corto plazo por las condiciones instantáneas de operación del sistema de generación para satisfacer la demanda, mientras que en el largo plazo dependen del crecimiento de la demanda y de la nueva capacidad instalada para atenderla; consecuentemente, la evolución del precio se ve influenciada por una gran cantidad de factores complejos

¹ Instituto de Energía. Universidad Nacional de Colombia. Medellín, Colombia. E-mail: jdvelasq@unal.edu.co. Autor para correspondencia.

² Instituto de Energía. Universidad Nacional de Colombia. Medellín, Colombia. E-mail: idyner@unal.edu.co

³ Faculdade de Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, Brasil. E-mail: reinaldo@ele.puc-rio.br

que se resumen en las condiciones físicas del sistema de generación, las decisiones de negocio de los agentes, y las decisiones y pautas del regulador.

Así, el modelado y la predicción de los precios de la electricidad han cobrado una importancia fundamental para los distintos agentes económicos, ya que están relacionados con la necesidad que tienen los generadores, los distribuidores y el regulador del mercado, de tomar decisiones tanto operativas como estratégicas [5] relacionadas con la formulación de estrategias de comercialización en el corto y en el largo plazo, la optimización de su programa de generación [6], la conformación del portafolio de activos de generación, la adquisición de nuevas plantas, y el abandono de plantas existentes [7,8]; así como un mejor manejo y planeamiento del mercado [7] desde el papel del regulador.

Particularmente en el caso brasileño, el proceso de reforma creó el Mercado Atacadista de Energía (MAE) donde la electricidad es libremente negociada a través de contratos forward bilaterales entre los agentes, y una bolsa de excedentes donde las compras y ventas se liquidan directamente con el ente administrador del sistema a un precio de mercado establecido por el gobierno. El precio spot o de corto plazo es usado como una señal económica que refleja el costo marginal de operación del sistema de generación en el corto plazo. Su modelado es particularmente difícil debido a la complejidad de la dinámica de la serie y la corta información histórica disponible. Adicionalmente, existen pocos trabajos relacionados con su modelado e interpretación [9]. En este trabajo se realiza una contribución en este aspecto, al seleccionar las variables que influyen en su formación y establecer un modelo que permite preparar predicciones condicionales a la evolución de las variables consideradas.

Por otra parte, se ha aceptado en términos generales que la dinámica que siguen los precios de la electricidad es no lineal; sin embargo, el modelado de los precios se ve afectado por dos condiciones fundamentales: en primer lugar, cada mercado posee una estructura única, dictaminada en términos de las características de la oferta, la demanda y la regulación, de tal forma que cada serie de precios es única ya que refleja dichas particularidades; y en segundo lugar, existen diferentes modelos que podrían ser adecuados para representar la dinámica de cada serie, pero no existen suficientes evidencias empíricas o teóricas que permitan seleccionar a priori alguno de ellos, por lo que dicha selección sólo puede realizarse a partir de los resultados del modelado de la serie en estudio. Este trabajo hace una contribución en este sentido, ya que realiza el modelado de la serie de precios de la electricidad en el mercado de corto plazo usando un modelo no lineal a través de un proceso ordenado y sistemático basado en pruebas estadísticas, y

compara el modelo obtenido con dos técnicas alternativas, sentando las bases para establecer en un trabajo futuro qué tipo de modelo podría ser más adecuado.

En relación al modelo utilizado, la evidencia aportada por algunos investigadores sugiere que las redes neuronales artificiales, y particularmente los perceptrones multicapa, son una de las técnicas más ampliamente utilizadas en la solución de problemas de regresión, clasificación de patrones y predicción [10]. Zhang, Patuwo y Hu [11] presentan una revisión general sobre el estado del arte, mientras que aplicaciones específicas son presentadas por Heravi, Osborn y Birchenhall [12], Swanson y White [13,14], Velásquez, Dyner y Souza [15,25] y Kuan y Liu [16], entre muchos otros. Algunos aspectos conceptuales son analizados por Sánchez, Velásquez y Castaño [17,18]. En este trabajo se usa un modelo conocido como red neuronal autorregresiva, cuyo uso ha sido poco difundido, el cual está compuesto por un modelo lineal autorregresivo más un perceptrón multicapa con una única capa oculta [19,20,21]. Su proceso de especificación es basado en pruebas de Multiplicadores de Lagrange, en donde se compara el modelo actual con H neuronas en la capa oculta con un nuevo modelo con $H+1$ neuronas en la capa oculta; para ello, se parte de una aproximación al error obtenido con el nuevo modelo que tiene $H+1$ neuronas, pero sin que realmente se realice su estimación. Si se acepta la nueva neurona adicional, el modelo es estimado, y se repite el proceso hasta que la nueva neurona sea rechazada. No obstante, este tipo de contraste no es adecuado, ya que se parte de una aproximación para aceptar la nueva neurona, pero no hay garantía de que el nuevo modelo, al ser estimado, permita obtener un error significativamente más bajo que el modelo anterior con H neuronas. Adicionalmente, las variables de entrada deben ser seleccionadas con un paso previo a la construcción del modelo de redes neuronales. En este trabajo se presenta una contribución en este aspecto, ya que se formula una estrategia basada en el contraste del radio de verosimilitud, en donde el nuevo modelo compara los modelos ya estimados y no sus aproximaciones. Adicionalmente, esta prueba permite realizar la selección de las variables de entrada para el modelo.

LA RED NEURONAL AUTORREGRESIVA Y SU PROCEDIMIENTO DE ESPECIFICACIÓN

El modelo

El objetivo de este trabajo es obtener una caracterización de las propiedades de la serie de precios mensuales de la electricidad en el mercado de corto plazo brasileño para la región Sur-Este/Centro (SE). Para ello, se utilizará una red neuronal autorregresiva (AR-NN), la cual combina un modelo lineal autorregresivo (AR) con un perceptrón

multicapa (MLP) con una única capa oculta, este es un modelo que permite combinar las ventajas de los modelos autorregresivos y de las redes neuronales, de tal forma que es más fácil capturar dinámicas complejas, tal como es el caso de los precios de electricidad.

En un modelo AR-NN, la variable dependiente y_t (en nuestro caso corresponde a los precios de la electricidad) es obtenida después de aplicar la función no lineal:

$$y_t = \beta_* + \sum_{i=1}^I \varphi_i x_t^{(i)} + \sum_{h=1}^H \beta_h G \left((2\sigma_y)^{-1} \alpha_{*,h} + \sum_{i=1}^I \alpha_{i,h} x_t^{(i)} \right) + e_t \quad (1)$$

a un conjunto de regresores $x_t^{(i)}$ (para $i=1, \dots, I$), con $e_t = \sigma \varepsilon_t$. Los regresores pueden estar conformados por los retardos de y_t (precios en los meses anteriores), así como por variables exógenas que pueden dar cuenta del comportamiento de y_t . $G()$ es la función de activación de las neuronas de la capa oculta. σ es la desviación estándar de los errores. ε_t es una variable aleatoria que sigue una distribución normal estándar. H es el número de neuronas en la capa oculta; I es el número de regresores. σ_y es la desviación estándar de y_t ; su uso evita tener que transformar y_t para restringir sus valores al rango de la función $G()$.

Los parámetros del modelo ($\Omega = [\beta_*, \beta_h, \varphi_i, \alpha_{*,h}, \alpha_{i,h}]$, para $h=1, \dots, H$; $i=1, \dots, I$) son obtenidos maximizando el logaritmo de la función de verosimilitud de los errores:

$$L = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \frac{e_t^2}{\sigma^2} \quad (2)$$

mediante alguna técnica de optimización, usualmente basada en gradientes. σ^2 es el promedio de los errores al cuadrado. (2) es obtenida al asumir que los residuales e_t siguen una distribución normal con media cero y varianza desconocida; la maximización de (2) equivale a minimizar el error cuadrático medio, que es el procedimiento común en la literatura de redes neuronales. El uso de (2) es la base del procedimiento estadístico de inferencia para la especificación del modelo, el cual se describe más adelante.

El problema de identificación (determinación del valor óptimo de los parámetros y la cantidad de unidades en la capa oculta) del modelo descrito por (1) está asociado a los siguientes aspectos:

1. El modelo puede estar mal condicionado, lo cual es síntoma de que las diferencias entre las magnitudes

de los parámetros son muy grandes; ello puede causar que las neuronas en la capa oculta estén en su nivel de saturación (diferentes entradas netas producen la misma salida), y consecuentemente que el gradiente de (2) sea cercano a cero, por lo que la optimización numérica se hace ineficiente.

- Existen múltiples configuraciones del modelo AR-NN que dan exactamente el mismo resultado a la salida. Ellas son obtenidas al permutar las neuronas de la capa oculta, de tal forma que las conexiones que entran y salen de ellas son arrastradas al realizar la permutación. Esto es, las neuronas de la capa oculta intercambian su posición generando modelos diferentes cuyo comportamiento es idéntico.
- Para algunas especificaciones de $G()$, la contribución que llega de cada neurona oculta a la salida, permanece constante, si los pesos de las conexiones que entran y salen de dicha unidad oculta son multiplicados por -1. Esto genera juegos de pesos que difieren en el signo, o sea, diferentes redes neuronales de comportamiento idéntico.

Las tres consideraciones anteriores muestran que existe una multitud de modelos diferentes con el mismo desempeño, lo que dificulta el proceso de estimación de los parámetros, ya que fácilmente el algoritmo de optimización puede quedar atrapado en un punto de máxima local. Para resolver estos problemas, (1) es especificado de la siguiente forma:

- Tal como es indicado en [22], algunos autores han sugerido por su experiencia práctica, que las funciones tipo sigmoidea que son simétricas alrededor del origen convergen más rápidamente que la función sigmoidea tradicional; adicionalmente, la adición de un término lineal puede ayudar a la convergencia, ya que se evita la saturación de la neurona o unidad de procesamiento en la capa oculta, y garantiza un gradiente mínimo cuando la salida neta de la función sigmoidea es cercana a sus valores extremos. Consecuentemente con las razones expuestas, $G()$ es especificada como:

$$G(u) = \frac{1 - \exp(-u)}{1 + \exp(-u)} + ku \quad (3)$$

Donde k es una constante pequeña. En este caso se usó $k = 0.0025$.

- Para un conjunto cualquiera de parámetros Ω , los parámetros β_h pueden restringirse a ser positivos, ya que la contribución neta de cada unidad oculta, β_h

$G()$, no cambia de signo si los parámetros $\beta_h, \alpha_{*,h}, \alpha_{i,h}$ cambian de signo puesto que $\beta_h G(u) = -\beta_h G(-u)$.

- Se sabe que diferentes configuraciones de (1) pueden ser obtenidas permutando las unidades de la capa oculta. Esta multiplicidad puede evitarse obligando a que los parámetros β_h estén siempre ordenados de forma creciente: $0 < \beta_1 < \beta_2 < \dots < \beta_h$.

Las tres restricciones enumeradas anteriormente buscan reducir la cantidad de modelos posibles, de tal forma que se facilite el proceso de estimación de la red neuronal artificial. Al imponer que los parámetros φ_i sean igual a cero, (1) se reduce a un perceptron multicapa (MLP). Si H se hace igual a cero, (1) se reduce a un modelo autorregresivo lineal con entradas exógenas (ARX).

Contraste de no linealidad y especificación del modelo

El procedimiento de especificación del modelo ARX-NN está basado en el siguiente algoritmo:

- Se estima el orden óptimo P para un modelo ARX, usando un criterio de información, y se estima el valor del logaritmo de su función de verosimilitud, L_{ARX} , usando (2).
- Usando la prueba del radio de verosimilitud, se contrasta la hipótesis de que la serie modelada (los precios) es lineal, versus la hipótesis alternativa de que sigue una dinámica no lineal. El estadístico de este contraste se calcula como:

$$S^* = 2(L_{H=1} - L_{ARX}) \quad (4)$$

donde $L_{H=1}$ es el logaritmo de la función de verosimilitud definido en (2) para un modelo ARX-NN con $H=1$. S^* sigue una distribución χ^2 con un grado de libertad. Si el valor crítico para el estadístico es suficientemente pequeño, se rechaza la hipótesis de linealidad, y se procede con el paso 3; en caso contrario, la serie es lineal y se continúa en el paso 6, con $H=0$.

- Sea H el número de unidades en la capa oculta del último modelo ARX-NN estimado ($H=1$ cuando se ejecuta este paso la primera vez). Se estima un modelo tentativo con $H+1$ unidades en la capa oculta.
- La aceptación de la nueva unidad ($H+1$) se hace a partir del valor crítico del estadístico de la prueba del radio de verosimilitud:

$$S = 2(L_{H+1} - L_H) \quad (5)$$

- Si la nueva unidad es aceptada (se rechaza la prueba) se retorna al paso 3; en caso contrario se pasa al paso 6. Cuando se acepta la prueba, quiere decir que los modelos con H y $H+1$ no son estadísticamente diferentes, por lo que el número óptimo de unidades ocultas es H . Esto se debe a que para dos modelos con desempeño similar se escoge el más simple, con menor número de parámetros.
- Se estima un modelo restringido y se calcula su radio de verosimilitud $L_H^{(i)}$. El modelo restringido se obtiene al eliminar la entrada i del modelo actual.
- Para cada uno de los modelos restringidos del paso anterior, se estima el estadístico del radio de verosimilitud como:

$$S^{(i)} = 2(L_H - L_H^{(i)}) \quad (6)$$

- Se elimina la entrada i que tenga el mayor valor crítico para el estadístico definido en (6), y para la cual se acepte el contraste del radio de verosimilitud. Si no se puede eliminar ninguna entrada se termina el proceso; en caso contrario, se retorna al paso 6.

Los pasos 1 y 2 corresponden a un contraste de no linealidad; el modelo ARX se obtiene a partir del modelo ARX-NN con $H=1$, imponiendo la restricción de $\beta_1=0$; de esta forma, el modelo ARX-NN corresponde al modelo libre, mientras que el modelo ARX corresponde al modelo restringido. Ya que el modelo ARX-NN tiene más parámetros que el modelo ARX, se cumple obligatoriamente que $L_{H=1} \geq L_{ARX}$, por lo que el estadístico S^* siempre es positivo. La hipótesis nula es que la restricción se cumple efectivamente, y que la serie es bien representada por un modelo lineal. Lo anterior implica que siempre el modelo ARX-NN debe ajustarse igual o mejor a los datos que el modelo ARX (ya que tiene más parámetros), pero la prueba exige un umbral mínimo en la mejora para que se considere que vale la pena usar el modelo no lineal.

Los pasos 4 al 6 desarrollan un proceso constructivo de la red neuronal donde se parte de un modelo mínimo (con $H=1$) y se van agregando unidades a la capa oculta hasta que se agregue una unidad irrelevante. Al igual que en la prueba de no linealidad, se contrasta la hipótesis de que $\beta_{H+1}=0$.

Una vez que se ha determinado el número óptimo de unidades en la capa oculta, los pasos 7 y 8 permiten verificar la contribución de cada entrada al modelo. La eliminación de la entrada $x_t^{(i)}$ equivale a hacer cero los parámetros $\alpha_{i,h}$ para $h=1, \dots, H$, y φ_i , de tal forma que el estadístico $S^{(i)}$ sigue una distribución χ^2 con $H+1$ grados de libertad.

Una vez se ha obtenido el modelo final, se continúa con un proceso de diagnóstico de los residuales, e_t para $t=1, \dots, T$, en los cuales se constatan tres supuestos: que sigan una distribución normal, que sean incorrelacionados, y que no existan evidencias de heterocedasticidad.

DESCRIPCIÓN DE LOS DATOS

El precio de corto plazo es fijado semanalmente, y se calcula para cada uno de los cuatro submercados (Norte-Nordeste, Sureste/Centro, Oeste y Sur) en que se encuentra dividido el sector eléctrico brasileño. Corresponde al costo marginal de demanda, resultante de la optimización de la operación mensual usando un modelo de planificación centralizada y posteriormente una planificación determinística para llevarlo a nivel semanal. No obstante, el precio puede ser modificado si se considera que no refleja adecuadamente las condiciones del sistema de generación en el corto plazo.

Los precios son una señal económica relacionada con el costo marginal del mercado, los cuales permiten a los generadores la recuperación de sus costos operacionales en

el corto plazo, y de sus inversiones en el largo plazo; más aún, el crecimiento de los precios señala la necesidad de nuevas inversiones en activos de generación. Es evidente la importancia de los pronósticos de los precios de electricidad, ya que todas las decisiones operativas y estratégicas de los participantes en el mercado son basadas en ellos.

Medeiros [9] indica que el parque de generación es predominantemente hidráulico, alcanzando aproximadamente un 85% de la capacidad instalada del sistema (unos 68 GW), mientras que la capacidad restante es suministrada por plantas térmicas (9 GW aproximadamente), importaciones (2 GW) y pequeñas plantas (1 GW). El sistema tiene un cubrimiento del 97% de la demanda total del sistema, dada la importancia del modelado y la predicción de largo plazo para las decisiones de los agentes del mercado.

A la fecha se posee la información histórica del precio, la energía almacenada en el embalse agregado (EA) y la energía afluente al embalse (ENA), entre 1996:1 y 2003:9, en la región SE, para un total de 93 datos. En la figura 1 se presenta el precio mensual de corto plazo de la región SE y sus determinantes.

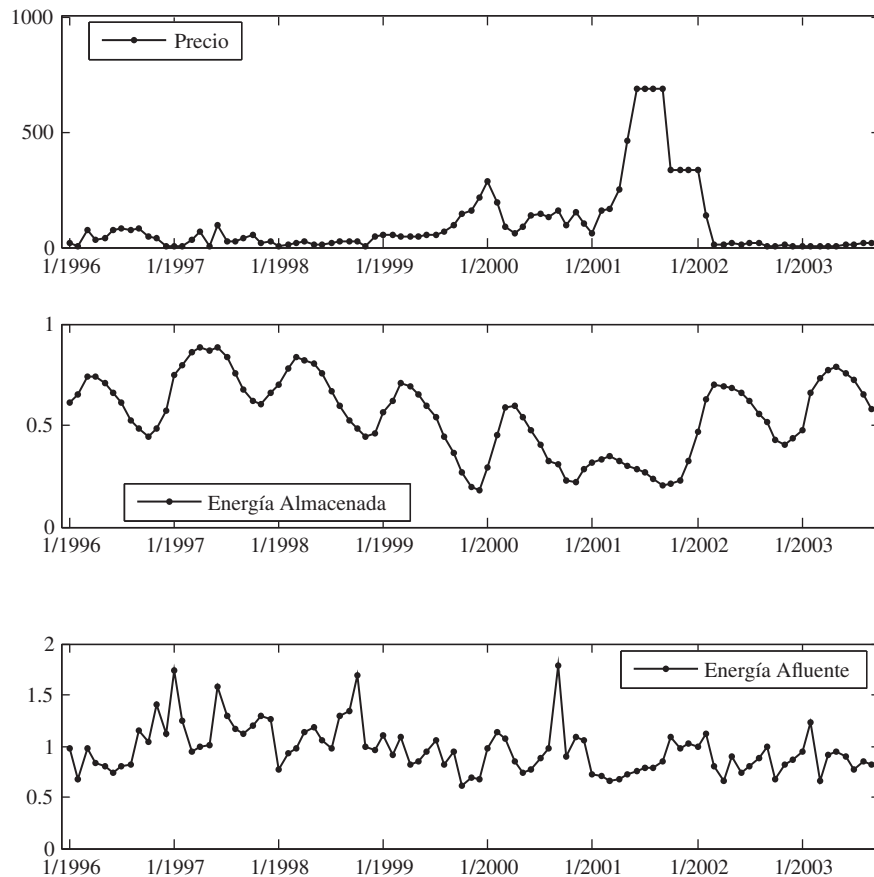


Figura 1. Precio mensual de corto plazo para la región SE, y sus determinantes.

MODELADO DE LA SERIE E INTERPRETACIÓN DE RESULTADOS

En esta sección se reportan los resultados obtenidos al estimar un modelo ARX-NN usando el algoritmo especificado en la sección anterior.

Como variable dependiente se usa el logaritmo natural de los precios, y_t . Esta transformación impone una restricción estructural de los precios, ya que evita que los valores pronosticados puedan ser negativos; adicionalmente, facilita la labor del modelado ya que la distribución de la variable dependiente se hace más semejante a una normal, minimizando la influencia de los valores extremos. Las propiedades estadísticas del logaritmo del precio son resumidas en la figura 2. La serie y_t muestra una relación inversa con la disponibilidad del recurso hídrico que es representada por las variables EA y ENA ; así, ante escasez del agua (en los veranos y especialmente durante los periodos de sequía) los precios aumentan, ya que la electricidad debe ser generada por unidades térmicas que son mucho más costosas que las plantas hidráulicas; mientras que en periodos de abundancia (inviernos y años especialmente húmedos) la electricidad es generada por las plantas hidráulicas, con un uso marginal de las plantas térmicas; la gráfica en el tiempo también muestra que hay cambios importantes en el nivel de la serie, como por ejemplo la caída de precios que se presenta en 2002:3, la cual coincide con la recuperación del embalse agregado (información que se ve reflejada en la variable EA). Los autocorrelogramas simple y parcial indican que el precio actual es influenciado prácticamente por el precio anterior, y posiblemente por el precio de hace dos meses; no hay evidencias de componentes cíclicas asociadas al ciclo anual de las estaciones. El histograma de y_t indica que los precios transformados no siguen una distribución normal.

Para iniciar el proceso de especificación, se seleccionaron como variables exógenas los tres primeros retardos de la EA y la ENA . Para ambas variables es difícil justificar en un sentido físico porque retardos más lejanos (superiores) tendrían influencia sobre el precio.

Paso 1. Se determinó el orden óptimo, P , del modelo ARX. Para obtenerlo, se estimaron modelos $ARX(p)$ para $p=1, \dots, 6$, seleccionándose aquel que tuviera el menor valor del criterio de información. Los resultados son presentados en la tabla 1. Para determinar el orden óptimo del modelo se usaron los criterios de información de Akaike, Hannan y Quinn y Schwartz. En los tres casos el orden óptimo del modelo es $p=1$, cuyo $L_{ARX} = -97.3563$.

Tabla 1. Valor de los criterios de información para el modelo $ARX(p)$.

p	CAK	CHQ	CSC
1	2.325*	2.424*	2.570*
2	2.351	2.461	2.624
3	2.389	2.510	2.688
4	2.389	2.521	2.716
5	2.463	2.606	2.817
6	2.445	2.599	2.826

CAK: es el criterio de Akaike.

CHQ: es el criterio de Hannan-Quinn.

CSC: es el criterio de Schwartz.

* Indica el valor mínimo según el criterio utilizado.

Paso 2. Se estimó un modelo ARX-NN con $H=1$. La prueba del radio de verosimilitud rechaza la hipótesis de que ambos no son significativamente diferentes:

$S^* = 2[-82.3890 - (-97.3563)] = 29.9346$ el cual es significativo al 1% de confiabilidad. Consecuentemente se acepta el modelo ARX-NN, y se demuestra que la serie sigue una dinámica no lineal.

Pasos 3 y 4. Se estimó un segundo modelo ARX-NN con $H=2$, para el cual $L_{H+1} = -81.5035$. El estadístico de la prueba del radio de verosimilitud es $S = 2[-81.5035 - (-82.3890)] = 1.7710$ el cual no es significativo para un nivel de confiabilidad del 5%, por lo que se rechaza una segunda unidad oculta. Se continúa en el paso 6.

Pasos 6, 7 y 8. Se aplicó el algoritmo descrito en estos pasos obteniéndose como variables relevantes y_{t-1} , ENA_{t-1} , EA_{t-1} , EA_{t-2} y EA_{t-3} . La tabla 2 presenta los parámetros óptimos del modelo final estimado.

A continuación se procede con el diagnóstico de los residuales e_t del modelo obtenido. Sus propiedades estadísticas son resumidas en la figura 3. El autocorrelograma simple muestra que no hay correlaciones seriales residuales. El histograma indica que los residuales siguen aproximadamente una distribución normal; para verificar esta conclusión se recurrió al contraste de Jarque-Bera [23]; el estadístico obtenido fue 40.06, el cual es significativo para una confiabilidad del 1%; ello indica que los errores no son normales. Esto es debido posiblemente, a que existen algunos errores de magnitud apreciable que corresponden a los cambios bruscos en el nivel de los precios, y que no es posible pronosticar con suficiente precisión dada la limitada información disponible. Por otra parte, se utilizó la prueba de contraste de Engle [24] para determinar los errores que tienen una varianza constante; los estadísticos obtenidos no fueron significativos para ninguno de los primeros seis rezagos.

Para establecer las bondades del modelo obtenido, se estimaron dos modelos alternativos que tienen las mismas entradas del modelo original; el primero es un modelo ARX, y el segundo es un perceptrón multicapa (MLP), el cual se obtiene al imponer la restricción $\varphi_i=0$ para todo i .

La predicción un paso adelante para los tres modelos es presentada en la figura 4. En dicha Figura, los puntos representan la serie y_t , mientras que la línea continua representa la predicción un paso adelante; el área

sombreada corresponde a la región de confiabilidad para un 95%. En la figura puede observarse cómo el modelo ARX-NN pronostica con mayor precisión la variable dependiente. Para comparar la calidad del ajuste de los modelos considerados, se utilizaron el error cuadrático medio (MSE), el error medio absoluto (MAE), y la raíz del error cuadrático medio geométrico. Los resultados obtenidos son presentados en la tabla 3, y corroboran cómo el modelo ARX-NN supera en precisión a los dos modelos alternativos considerados.

Tabla 2. Parámetros óptimos del modelo ARX-NN.

Parámetro	Valor	Error Estándar	Estadístico T
β_*	3.7853	10.446	0.3624
Componente lineal			
$\varphi_1 (y_{t-1})$	0.8939	1.4101	0.6339
$\varphi_2 (ENA_{t-1})$	-1.1286	3.6427	- 0.3098
$\varphi_3 (EA_{t-1})$	-4.8083	15.310	-0.3141
$\varphi_4 (EA_{t-2})$	6.917	27.921	0.2477
$\varphi_5 (EA_{t-3})$	-3.0513	20.094	-0.1519
Componente no lineal			
$\alpha_{1,1} (y_{t-1})$	-11.536	57.713	-0.1999
$\alpha_{2,1} (ENA_{t-1})$	5.1754	45.963	0.1126
$\alpha_{3,1} (EA_{t-1})$	2.6358	336.37	0.0078
$\alpha_{4,1} (EA_{t-2})$	1.7188	903.28	0.0019
$\alpha_{5,1} (EA_{t-3})$	14.536	687.24	0.0212
$\alpha_{*,1}$	-3.1974	43.194	-0.0740
β_*	1.8645	7.1264	0.2616
σ_e^2	0.3709		

Tabla 3. Estadísticos de ajuste

Estadístico	ARX-NN	MLP	ARX
MSE	0.3690	0.4403	0.4839
MAE	0.4254	0.4476	0.4736
GRMSE	0.0427	0.0563	0.0654

MSE: Error cuadrático medio

MAE: Error medio absoluto

GRMSE: Raíz del error cuadrático medio geométrico

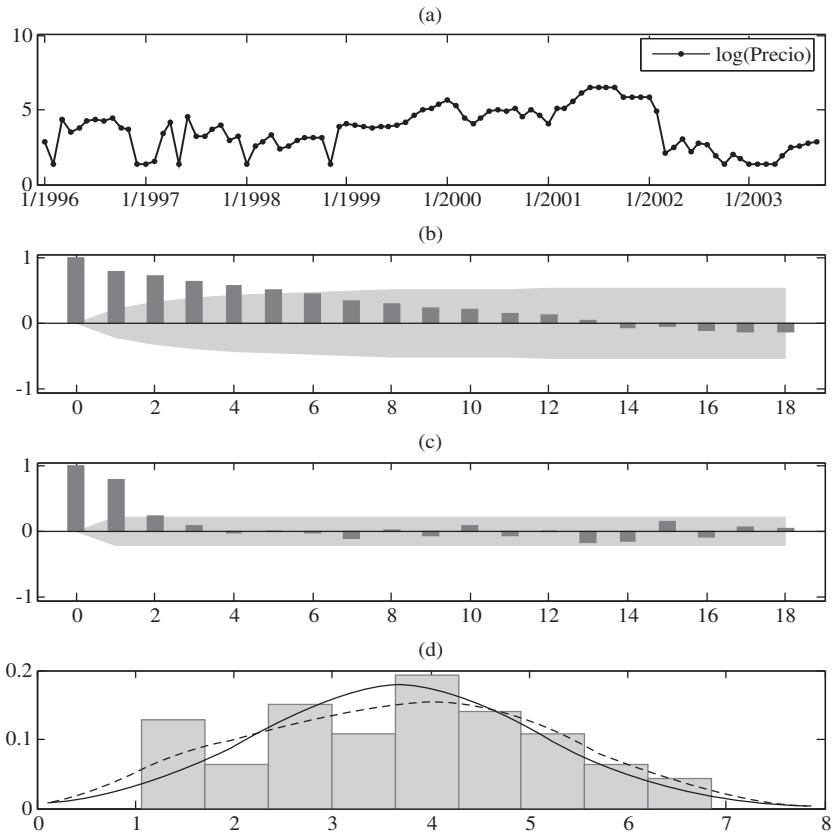


Figura 2. Propiedades estadísticas del logaritmo del precio. (a) Serie; (b) Autocorrelograma simple; (c) Autocorrelograma parcial; (d) Histograma.

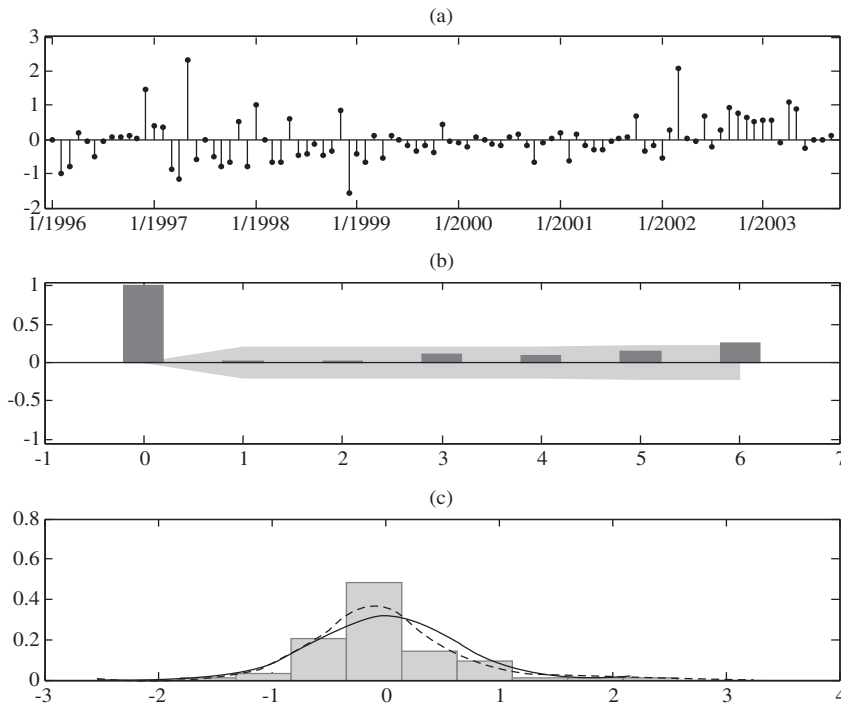


Figura 3. Propiedades estadísticas de los residuos del modelo ARX-NN. (a) Residuales; (b) Autocorrelograma simple; (c) Histograma.

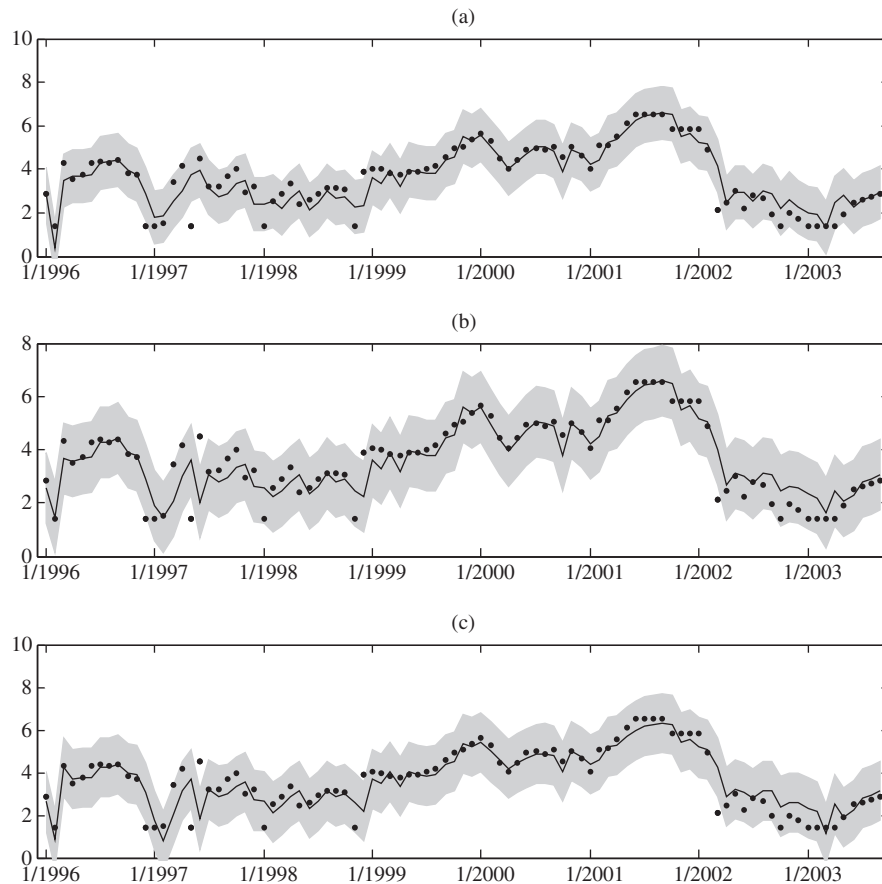


Figura 4. Predicción un paso adelante. (a) Modelo ARX-NN; (b) MLP; y (c) ARX.

CONCLUSIONES

En este trabajo se presentan los resultados obtenidos al modelar la serie de precios mensuales de la electricidad en el mercado de corto plazo de la región SE del Brasil usando un modelo ARX-NN. Las contribuciones de este trabajo están relacionadas con los siguientes aspectos:

- Se modifica la arquitectura del modelo ARX-NN, facilitando su proceso de optimización.
- Se desarrolla un algoritmo de especificación basado en el contraste del radio de verosimilitud.
- Se ejemplifica el uso del algoritmo propuesto a través de un caso real.
- Se corrobora que la dinámica de la serie es no lineal, y se determinan sus principales determinantes.

REFERENCIAS

[1] F. Steiner. "Regulation, industry structure and performance in the electricity supply industry". OECD Economic Studies. Vol. 32. 2000.

[2] M. Armstrong, S. Cowan and J. Vickers. "Regulatory Reform: Economic Analysis and British Experience". The MIT Press. Cambridge. 1994.

[3] D. Bunn. "Reflections on the progress of electricity re-structuring, privatisation and regulation in the UK during 1988-1998". XXXIII Reunión de Altos Ejecutivos C.I.E.R. 1998.

[4] D. Pilipovic. "Energy Risk. Valuing and Managing Energy Derivates". McGraw-Hill. 1998.

[5] Y. Hong and C. Lee. "A neuro-fuzzy price forecasting approach in deregulated electricity markets". Electric Power Systems Research. Vol. 73, pp. 151-157. 2005.

[6] A. Conejo, J. Contreras, R. Espinosa and M. Plazas. "Forecasting electricity prices for a day-ahead pool-based electricity energy market". International Journal of Forecasting. Vol. 21, pp. 435-462. 2005.

- [7] X. Lu, Z. Dong and X. Li. "Electricity market price spike forecast with data mining techniques". *Electric Power Systems Research*. Vol. 73 N° 1, pp. 19-29. 2005.
- [8] A. Angelus. "Electricity price forecasting in deregulated markets". *The Electricity Journal*. Vol. 14 N° 3, pp. 32-41. 2001.
- [9] L. de Medeiros. "Previsão do Preço Spot no Mercado de Energia Elétrica". Tesis para optar al grado de doctor. Pontificia Universidade Católica do Rio de Janeiro. 2004.
- [10] T. Masters. "Practical Neural Network Recipes in C++". First edition. Academic Press, Inc. 1993.
- [11] G. Zhang, B. Patuwo and M. Hu. "Forecasting with artificial neural networks: the state of the art". *International Journal of Forecasting*. Vol. 14, pp. 35-62. 1998.
- [12] S. Heravi, D. Osborn and C. Birchenhall. "Linear versus neural network forecasts for European industrial production series". *International Journal of Forecasting*. Vol. 20, pp. 435-446. 2004.
- [13] N. Swanson and H. White. "Forecasting economic time series using adaptive versus nonadaptive and linear versus non-linear econometric models". *International Journal of Forecasting*. Vol. 13, pp. 439-461. 1997a.
- [14] N. Swanson, H. White. "A model selection approach to real time macroeconomic forecasting using linear models and artificial neural networks". *Review of Economics and Statistics*. Vol. 39, pp. 540-550. 1997b.
- [15] J.D. Velásquez, R. Dyner y R.C. Souza. "Un modelo no lineal de los precios spot medios mensuales en el Mercado Colombiano". VII Seminario Internacional sobre Análisis y Mercados Energéticos & I Seminario CERES. Santa Fe de Bogotá, Colombia. 5-7 de octubre de 2005.
- [16] C. Kuan and T. Liu. "Forecasting exchange rates using feedforward and recurrent neural networks". *Journal of Applied Econometrics*. Vol. 10, pp. 347-364. 1995.
- [17] P. Sanchez, J. Velásquez and E. Castaño. "Artificial neural networks for modeling nonlinear time series with segment trends". 25th International Symposium on Forecasting. San Antonio - Texas, EE.UU. International Institute of Forecasters. 2005a.
- [18] P. Sanchez, J. Velásquez and E. Castaño. "Nonlinear time series with breaks in the seasonal pattern: a modeling approach using neural networks". 25th International Symposium on Forecasting. San Antonio - Texas, EE.UU. International Institute of Forecasters. 2005b.
- [19] H. White. "An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks". *Proceedings of the International Joint Conference on Neural Networks*. Vol. 2, pp. 451-455. IEEE Press. Washington DC., New York. 1989.
- [20] T.H. Lee, H. White and C.W.J. Granger. "Testing for neglected nonlinearity in time series models". *Journal of Econometrics*. Vol. 56, pp. 269-290. 1993.
- [21] T. Teräsvirta, C.F. Lin and C.W.J. Granger. "Power of the neural network linearity test". *Journal of Time Series Analysis*. Vol. 14, pp. 209-220. 1993.
- [22] Y. LeCun, L. Bottou, G.B. Orr and K.R. Muller. *Neural Networks - Tricks of the Trade*, Springer Lecture Notes in Computer Sciences 1524, chapter Efficient Backprop, pp. 5-50. 1998.
- [23] C. Jarque and A. Bera. "A test for normality of observations and regression residuals". *International Statistical Review*. Vol. 55, pp. 163-172. 1987.
- [24] R. Engle. "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflations". *Econometrica* Vol. 50, pp. 987-1007. 1982.
- [25] J.D. Velásquez, R. Dyner y R.C. Souza R.C. *Predicción Condicional del Precio Mensual de Bolsa basada en Escenarios de Eventos Hidrológicos Extremos*. VII Seminario Internacional sobre Análisis y Mercados Energéticos & I Seminario CERES. Santa Fe de Bogotá, Colombia. 5-7 de octubre de 2005.