

Metodologías de Pronóstico del Precio de la Energía

Seminario DERIVEX

Octubre 24 de 2012

Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación
- 3 Hacia un Modelo Dinámico de Factores
- 4 Pronósticos Intermedios
- 5 Pronóstico del Precio de la Energía
- 6 Validación de Modelos

Teoría Básica de Series de Tiempo: Box - Jenkins

Esta metodología procura el mejor modelo ARIMA para una serie de tiempo por medio de los siguientes pasos

- 1 Verificar si la serie es estacionaria
 - Graficar la serie
 - Revisar el autocorrelograma, un autocorrelograma que decae lentamente se considera evidencia de falta de estacionariedad en la serie.
 - Prueba de Dickey-Fuller
 - Para solucionar la falta de estacionariedad se diferencia la serie.
- 2 Verificar si la serie tiene un componente estacional.
 - La estacionalidad puede observarse como patrones que se repiten al graficar la serie.
 - Las series con estacionalidad también presentan patrones en su autocorrelograma.
 - Se soluciona aplicando diferencias estacionales y agregando términos AR y MA estacionales en el modelo.
- 3 Identificar el orden p del proceso autoregresivo
 - La herramienta principal para este fin es la gráfica del autocorrelograma parcial.
 - Un AR(1), tien un autocorrelograma que va decreciendo (exponencial u oscilando).
 - Para un AR(p) su autocorrelograma parcial deja de ser significativo después del rezago p .
- 4 Identificar el orden de q del proceso MA
 - Se utiliza el autocorrelograma (grafica la correlación entre las observaciones de un proceso en diferentes momentos del tiempo).
 - En un MA(q), su autocorrelograma no presenta rezagos significativos después del q .
- 5 Después de modelar la serie al mirar el autocorrelograma de los residuos no se debe encontrar ningún rezago significativo.

Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación**
- 3 Hacia un Modelo Dinámico de Factores
- 4 Pronósticos Intermedios
- 5 Pronóstico del Precio de la Energía
- 6 Validación de Modelos

Se recolectó información con una periodicidad mensual (de acceso público) sobre:

- La demanda de energía
- Los caudales en energía de los ríos que aportan agua a algún embalse del SIN
- Nivel de embalses (reserva de energía)
- Precio de la energía
- La anomalía en la temperatura en la región del niño 3.4 (5° N - 5° S, 120° W - 170° W)
- El índice de precios al productor¹

La información sobre la demanda, el nivel del embalse, y el precio de la energía se encuentran en el sistema NEON². La anomalía en la región del niño 3.4 es publicada por *The National Oceanic and Atmospheric Administration* (NOAA)³.

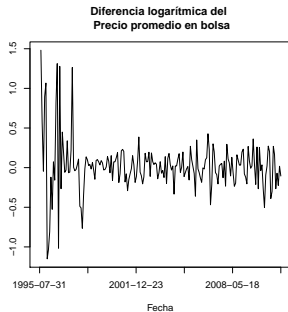
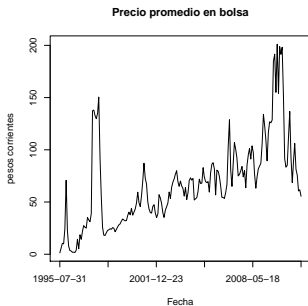
¹ Fuente Banco de la República

² <http://sv04.xm.com.co/neonweb/>

³ <ftp://ftp.cpc.ncep.noaa.gov/wd52dg/data/indices/sstoi.indices>

Precio de la Energía

- La serie del precio de la energía carece de estacionaridad.
- Al realizar un test se comprueba que la serie tiene una raíz unitaria.
- Se trabaja con la diferencia logarítmica de la serie.
- Las gráficas corresponden al precio de la energía en Kwh y a su diferencia logarítmica.



En todo el análisis que sigue se trabaja con el precio de la energía deflactado por el IPP.

Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación
- 3 Hacia un Modelo Dinámico de Factores**
- 4 Pronósticos Intermedios
- 5 Pronóstico del Precio de la Energía
- 6 Validación de Modelos

Hacia un Modelo Dinámico de Factores

- El precio de la energía se modelará a partir de modelos lineales que tienen como variables independientes
 - La anomalía
 - El nivel del embalse
 - El caudal en los ríos
 - La demanda de energía
 - Rezagos del precio de la energía
- La estrategia para generar el modelo consistirá en:
 - Generar pronósticos independientes para las variables: anomalía, ríos, demanda y embalse.
 - A partir de estos pronósticos generados (llamada pronósticos intermedios) se alimentará el modelo del precio de la energía.
 - El precio de la energía se pronosticará recursivamente apoyándose en los pronósticos intermedios.

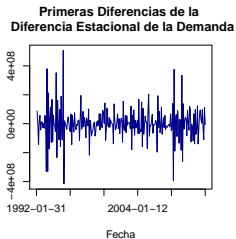
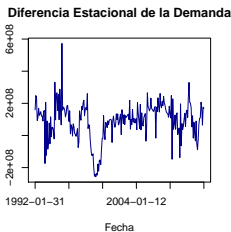
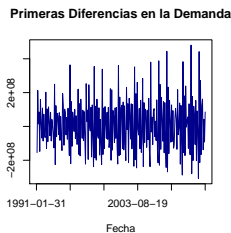
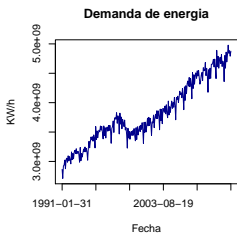
Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación
- 3 Hacia un Modelo Dinámico de Factores
- 4 Pronósticos Intermedios**
- 5 Pronóstico del Precio de la Energía
- 6 Validación de Modelos

- Para cada una de las series se calculó un modelo SARIMA (p,d,q,P,D,Q) .
- Para determinar el orden de los modelos se utilizaron los auto-correlogramas y los auto-correlogramas parciales que ayudan a dictaminar el orden auto-regresivo y de media móvil de los modelos.
- Los modelos SARIMA se usaron para modelar el nivel del embalse, el caudal en los ríos, la demanda de energía y la parte tendencial de la anomalía.

Pronóstico de la Demanda de Energía

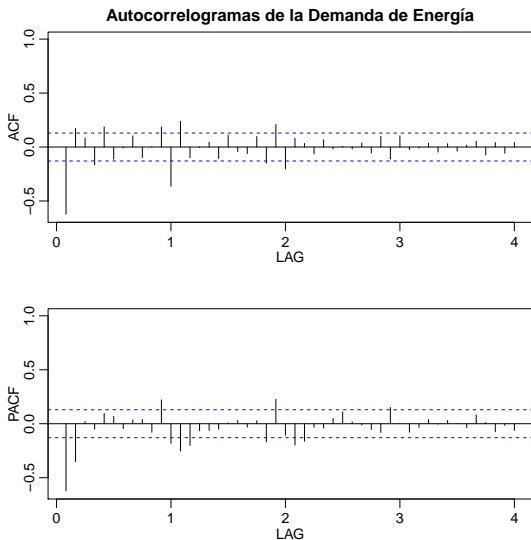
Variable: Demanda del SIN en Neon ⁴



⁴ Demanda del SIN, la cual se calcula con base en la generación neta de las plantas e incluye: hidráulicas, térmicas, plantas menores, cogeneradores, demanda no atendida, limitación del suministro e importaciones.

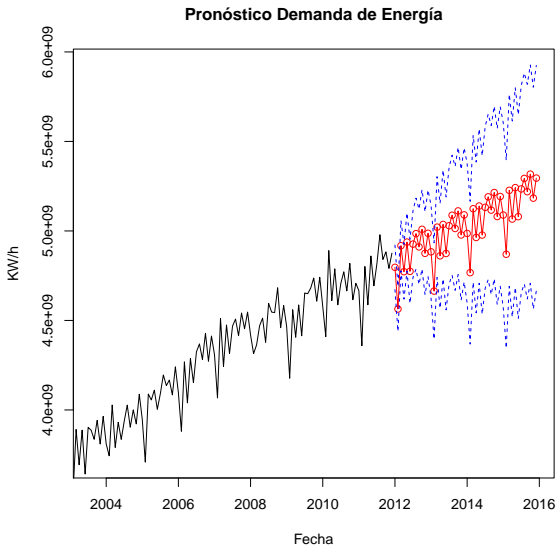
Pronóstico de la Demanda de Energía

Modelo Escogido: $SARIMA(0, 1, 7) \times (0, 1, 1)_{12}$



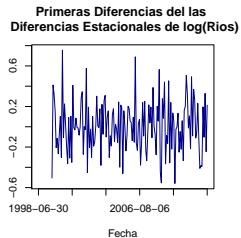
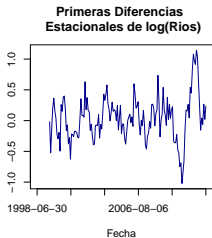
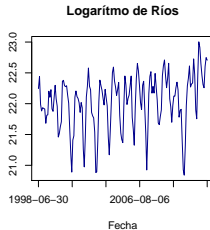
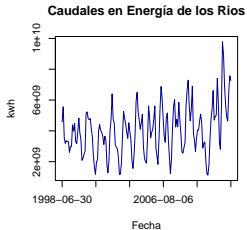
Pronóstico de la Demanda de Energía

Modelo Escogido: $SARIMA(0, 1, 7) \times (0, 1, 1)_{12}$



Pronóstico de la Variable Ríos

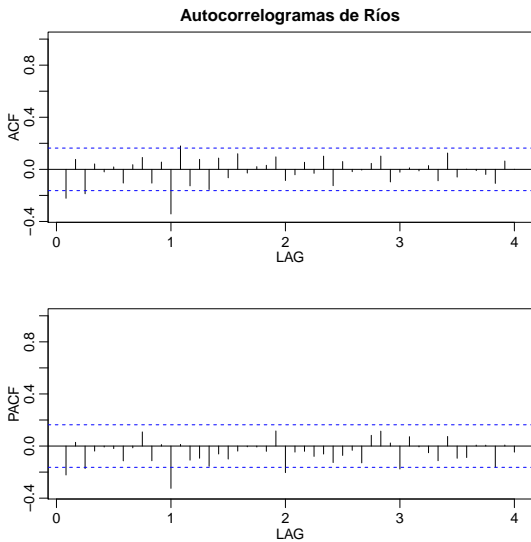
Variables: Aportes hídricos (Energía) en NEON.⁵



⁵ Caudales en energía de los ríos que aportan agua a algún embalse del SIN

Pronóstico de la Variable Ríos

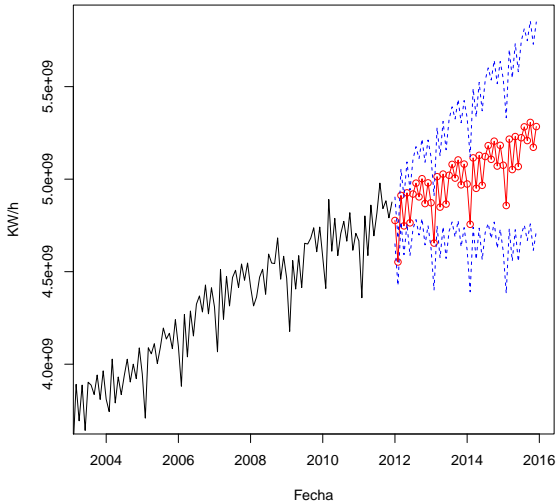
Modelo Escogido: $SARIMA(0, 1, 3) \times (0, 1, 1)_{12}$



Pronóstico de la Variable Ríos

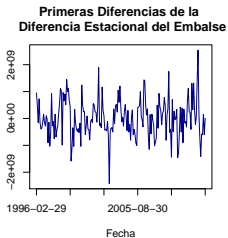
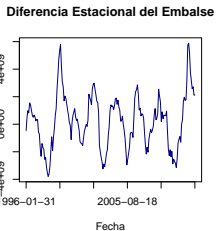
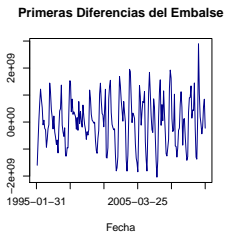
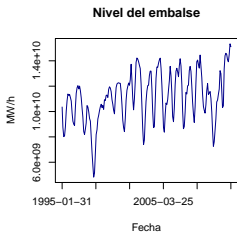
Modelo Escogido: $SARIMA(0, 1, 3) \times (0, 1, 1)_{12}$

Pronóstico nivel de Ríos



Pronóstico del Nivel del Embalse

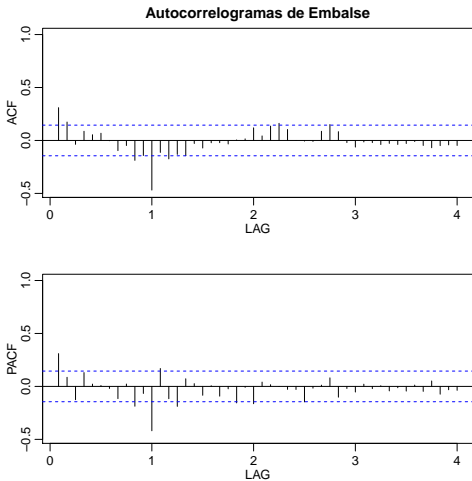
Variables: Volumen Embalse (Energía) en NEON.⁶



⁶ Es la reserva de energía de un embalse de acuerdo con la cantidad de agua almacenada en el mismo

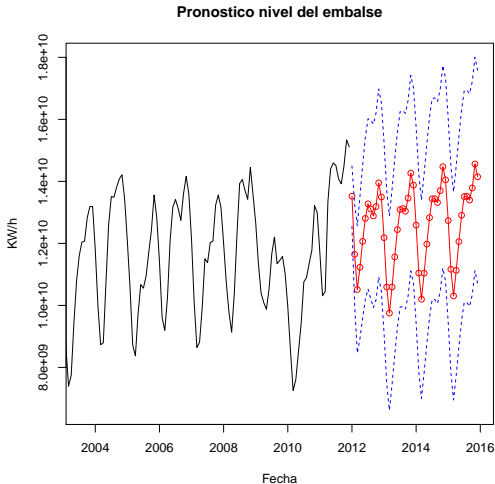
Pronóstico del Nivel del Embalse

Modelo Escogido: $SARIMA(9, 1, 2) \times (0, 1, 1)_{12}$



Pronóstico del Nivel del Embalse

Modelo Escogido: $SARIMA(9, 1, 2) \times (0, 1, 1)_{12}$



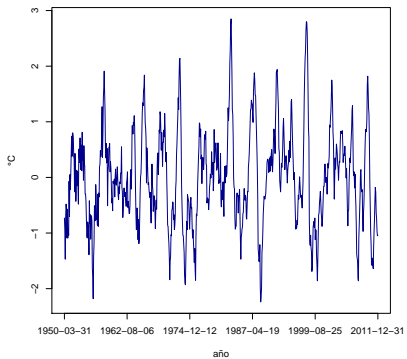
Pronóstico de la Anomalía

- La Anomalía es una serie que presenta un comportamiento bastante estacional y complejo, ampliamente documentado en la literatura especializada.
- En lugar de medir la temperatura de la superficie del mar, mide las anomalías en la temperatura de la superficie marítima las cuales están asociadas con los fenómenos del Niño y la Niña⁷.
- Para modelar su comportamiento el análisis se dividió en tres partes:
 - Se filtró la serie con una media móvil para eliminar parte de la estacionalidad.
 - Para terminar de eliminar el componente estacional de la serie se le aplicó una descomposición espectral
 - La parte no estacional de la serie se modeló por medio de un modelo ARIMA

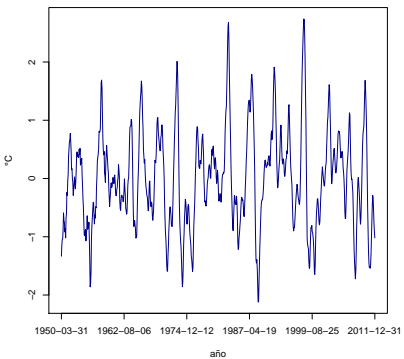
⁷ Una anomalía mayor a uno se asocia con el fenómeno del Niño. Una anomalía menor a menos uno se asocia con el fenómeno de la Niña.

Media Móvil de la Anomalía

Anomalia en la region del niño 3.4



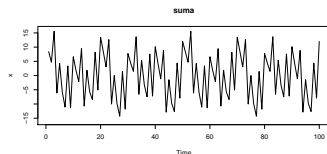
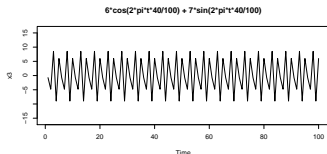
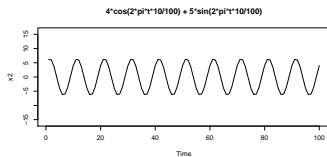
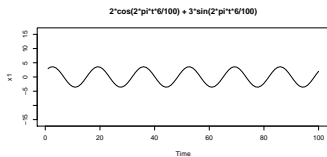
Media Móvil de la Anomalia



Descomposición Espectral

La idea detrás de esta técnica es descomponer una serie en una sumatoria de términos sinusoidales:

$$x_t = \sum_{k=1}^q U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t)$$



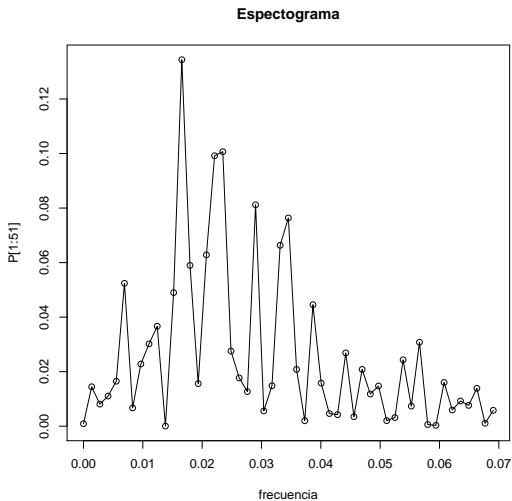
Para esto se usa el periodograma:

$$P(j/n) = \left(\frac{2}{n} \sum_{t=1}^n x_t \cos(2\pi tj/n)\right)^2 + \left(\frac{2}{n} \sum_{t=1}^n x_t \sin(2\pi tj/n)\right)^2$$

- La idea detrás del periodograma es correr una regresión de x_t con respecto a los sinusoides variando las frecuencias fundamentales j/n
- Esto se hace usando la Transformación Rápida de Fourier

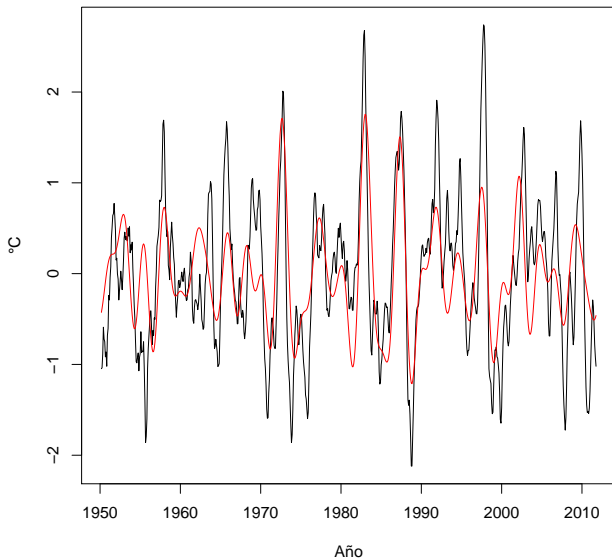
Descomposición Espectral

Sugiere los componentes: 6, 10, 12, 13, 14, 16, 17, 18, 22, 25, 26 y 29



Descomposición Espectral

Media Movil de la Anomalia
y su Descomposición Espectral



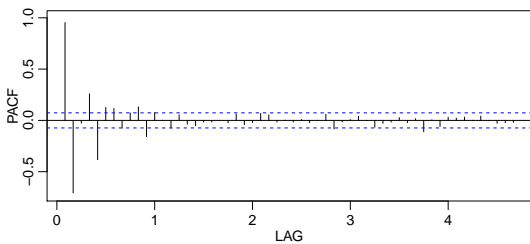
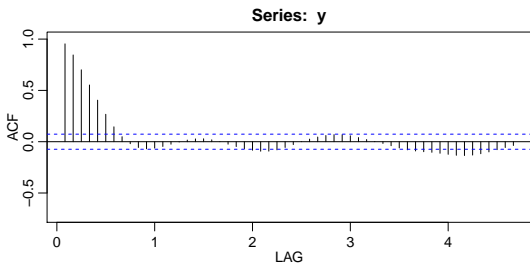
Pronóstico de la Anomalía

- Esto sugiere que la serie de la anomalía se puede descomponer en funciones sinusoidales que tienen las frecuencias indicadas en el espectograma.
- Esta función sinusoidal se estima y se utiliza para filtrar (linealmente) la serie original de la Anomalia.
- Por lo tanto la serie que se estima con el modelo Arima es:

$$y_t = MA(Anomalia) - DescEspectral(MA(Anomalia))$$

Pronóstico de la Anomalía

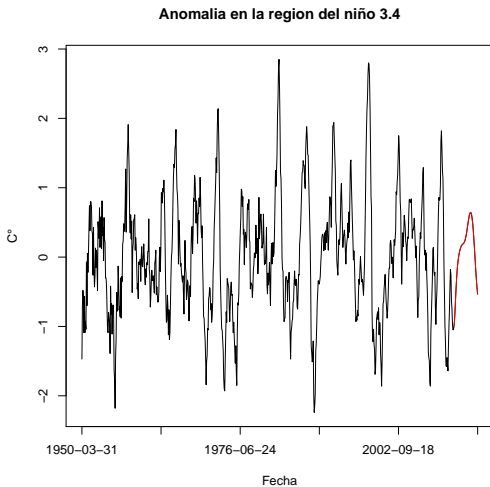
Sugiere un ARIMA(4,0,0)



Pronóstico de la Anomalía

El pronóstico final se calcula como

$$Anomalia_F = ARIMA_Anomalia_F + Seasonal_Anomalia_F$$



Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación
- 3 Hacia un Modelo Dinámico de Factores
- 4 Pronósticos Intermedios
- 5 Pronóstico del Precio de la Energía**
- 6 Validación de Modelos

Pronóstico del Precio de la Energía

- Se utilizaron tres modelos combinando técnicas de *backward selection* y componentes principales:
 - 1 Modelos de regresión lineal.
 - 2 Modelo de factores.
 - 3 Combinación de pronósticos.
- Este tipo de modelos permite reducir en gran medida la dimensionalidad pues se usan las 5 variables descritas en la sección Base de Datos y diferentes rezagos de cada variable.

Backward Selection

- La técnica de backward selection es usada en estadística para reducir la dimensionalidad de un modelo.
- Ésta es una de las modalidades de *stepwise regressions*:
 - Se empieza con todas las variables que se cree podrían incluirse en el modelo.
 - En cada paso de la selección se elimina la variable que aporte menos a la explicación de la variable objetivo.
 - En este caso se eliminará la variable que tenga el mayor p-valor correspondiente a la prueba t de significancia.
 - La selección elimina una sola variable por paso hasta que todas las variables que se incluyan en el modelo sean significativas a un nivel predeterminado.

Modelo Backward Selection Regresión Lineal

- Se toman como variables predictoras la anomalía, el nivel de los embalses y la demanda, doce rezagos de cada una de estas tres variables y doce rezagos del precio de la energía en bolsa (X_t).
- Se pronostica el precio de la energía en bolsa usando backward selection con p-valores 0.01 y 0.05 para disminuir el número de variables predictoras.

$$Y_t^1 = \beta^1 \widehat{X}_t^1 + \mu_t^1 \quad (1)$$

$$Y_t^5 = \beta^5 \widehat{X}_t^5 + \mu_t^5 \quad (2)$$

- Los superíndices 1 y 5 hacen referencia al p-valor usado en el backward selection
- \widehat{X}^1 y \widehat{X}^5 son las variables explicativas que quedan después de hacer backward selection a todas las variables contenidas en X .

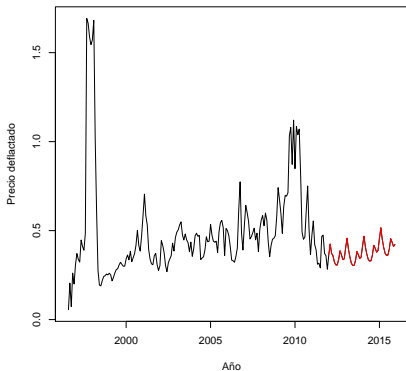
Los pronósticos obtenidos con los dos modelos estimados anteriormente se promedian y este es el pronóstico final:

$$\widehat{Y}_t = \frac{1}{2}(\widehat{Y}_t^1 + \widehat{Y}_t^5) + \epsilon_t. \quad (3)$$

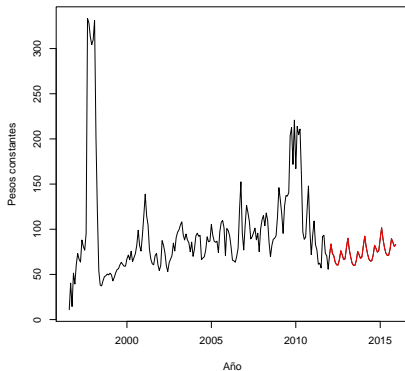
Modelo Backward Selection Regresión Lineal

Tomando como última fecha observada diciembre 2011, el pronóstico obtenido 48 meses hacia adelante por medio de este modelo sería:

Precio promedio en bolsa deflactado



Precio promedio en bolsa



Bootstrapping para la obtención de Intervalos de Confianza

Se tiene una serie de observaciones y_1, \dots, y_T . Para obtener un intervalo de confianza por medio de *bootstrapping* se sigue el siguiente procedimiento⁸:

- 1 Sean \hat{u}_t $t = 1, \dots, T$ los residuos de estimación del modelo.
- 2 Se calculan los residuos centrados $\hat{u}_1 - \bar{u}, \dots, \hat{u}_T - \bar{u}$ y se obtienen los residuos de bootstrapping u_1^*, \dots, u_T^* tomando una muestra aleatoria con remplazo de los residuos centrados.
- 3 Se obtiene la serie de tiempo de bootstrapping recursivamente:

$$y_t^* = \hat{v} + \hat{A}_1 y_{t-1}^* + \dots + \hat{A}_p y_{t-p}^* + u_t^*$$

donde se toman los mismo valores iniciales para la serie de bootstrapping,
 $(y_{-p+1}^*, \dots, y_0^*) = (y_{-p+1}, \dots, y_0)$

- 4 Basándose en la serie de bootstrapping se vuelven a estimar los parámetros A_1, \dots, A_p .
- 5 Se generan pronósticos de $y_{T+1}^*, \dots, y_{T+h}^*$ a partir del modelo estimado en el paso anterior.
- 6 El procedimiento de los pasos 3 a 5 se repite n veces y se obtiene así una distribución de los pronósticos a partir de la cual se puede calcular el intervalo de confianza.

⁸ Lütkepohl, H. *New Introduction to Multiple Time Series Analysis*. Springer, 2005.

Intervalos de Confianza: Modelo Backward Selection

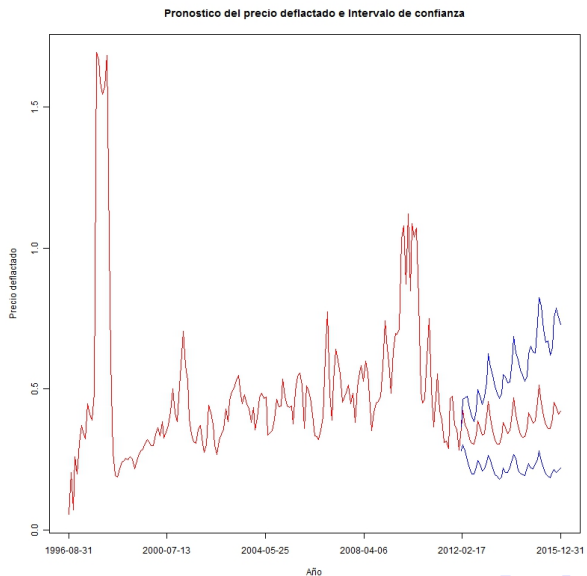
Regresión Lineal

- La obtención del intervalo de confianza del pronóstico se hizo por medio de bootstrapping.
- Se toma una muestra aleatoria de los errores centrados del modelo, ϵ_t .
- Los errores de bootstrapping se denominarán ϵ_t^* .
- Se construye el precio de la energía de bootstrapping por medio de los dos modelos de backward selection con p-valores 0.01 y 0.05 usando los errores de bootstrapping.
- Se comienza con las doce primeras observaciones del precio de la energía, Y_1, \dots, Y_{12} . a partir de éstas se van estimando recursivamente los demás valores de Y_t por medio de las ecuaciones:

$$\begin{aligned}\widehat{Y}_t^{1*} &= \beta^1 \widehat{X}_t^{1*} \\ \widehat{Y}_t^{5*} &= \beta^5 \widehat{X}_t^{5*} \\ \widehat{Y}_t^* &= 1/2(\widehat{Y}_t^{1*} + \widehat{Y}_t^{5*}) + \epsilon_t^*.\end{aligned}$$

- Basado en la serie del precio de la energía de bootstrapping, Y_t^* , se estiman de nuevo los modelos de backward selection con p-valores 0.01 y 0.05 y se pronostica a partir de estos modelos Y_t^* .
- Este procedimiento se lleva a cabo N veces, donde N es un número grande.
- A partir de este procedimiento se obtienen los intervalos de confianza del pronóstico.

Intervalos de Confianza: Modelo Backward Selection Regresión Lineal



Componentes Principales

- El análisis de Componentes Principales (PC) consiste en llevar a cabo una transformación ortogonal para convertir las observaciones de un grupo de variables, posiblemente correlacionadas, en un grupo de variables independientes a las cuales se les llama componentes principales.
- La transformación se define de manera que los componentes principales queden ordenados de manera que del primero al último cada componente explique la mayor varianza de los datos originales.

Modelo Dinámico de Factores

- La metodología usada se basa en el trabajo de Stock y Watson (2002)⁹
- El objetivo es utilizar método de componentes principales para mitigar el problema de dimensionalidad utilizando la covariabilidad de la serie en términos de un número menor de factores no observados.
- El pronóstico se lleva acabo estimando primero los factores a partir de las variables usadas para predecir la variable independiente.
- Posteriormente, se estima por medio de una regresión lineal la relación entre los factores y la variable independiente.

⁹Forecasting Using Principal Components from a Large Number of Predictors. American Statistical Association. December 2002, Vol 97, No. 460

Modelo Dinámico de Factores

- Supóngase que Y_t es la serie a predecir y sea X_t un objeto de N series de tiempo las cuales se cree predicen Y_t .
- El modelo para predecir Y_t usando factores calculados por medio de componentes principales asume que (X_t, Y_t) se puede representar de la siguiente manera:

$$X_t = \Lambda F_t + e_t \quad (4)$$

$$Y_{t+h} = \beta_F F_t + \beta_w w_t + \epsilon_{t+h} \quad (5)$$

donde

- Λ es una matriz
- β_F y β_w son vectores
- e_t es un vector de $N \times 1$ errores indiosincráticos
- w es un vector de m rezagos de Y_t
- h es el horizonte del pronóstico
- ϵ_{t+h} es el error del pronóstico.

Modelo Dinámico de Factores

Para estimar los factores por medio de componentes principales se considera la función objetivo:

$$V(F) = (NT)^{-1} \sum (X_{it} - \Lambda_i F_t)^2 \quad (6)$$

donde Λ_i es la i -ésima fila de Λ . Este es el problema clásico de componentes principales.

- Se soluciona tomando Λ como los eigenvectores de $X'X$ correspondientes a los eigenvalores más grandes.
- El estimador de F por componentes principales correspondiente sería, $F = X'\Lambda/N$.
- Cuando $N > T$ se puede estimar este problema de una manera equivalente tomando F como los eigenvectores con eigenvalores más grandes de XX' .

Modelo Dinámico de Factores

- En este modelo se toman como variables predictoras la anomalía y 24 rezagos de ésta,
- El nivel de los embalses
- Los caudales en energía de los ríos que aportan agua a algún embalse del SIN
- La demanda
- Las últimas tres con tres rezagos cada una y doce rezagos del precio de la energía en bolsa
- Se llamará X_t el conjunto de todas las variables predictoras.

- Se calculan los 10 primeros componentes principales de todas las variables en X_t menos los rezagos del precio de la energía; estos se llamarán F_t .
- Los 10 componentes principales y los 12 rezagos del precio se usan como variables independientes en el modelo lineal para el precio de la energía.
- La dimensionalidad del modelo se reduce por medio de backward selection con un p-valor de 0.05.

Modelo Dinámico de Factores

- Siendo Y_t el precio de la energía en bolsa este modelo es el modelo de Stock y Watson (2002)
- Se usa adicionalmente backward selection para disminuir la dimensionalidad del modelo.
- Las ecuaciones del modelo son las siguientes:

$$X = \Lambda F_t + e_t \quad (7)$$

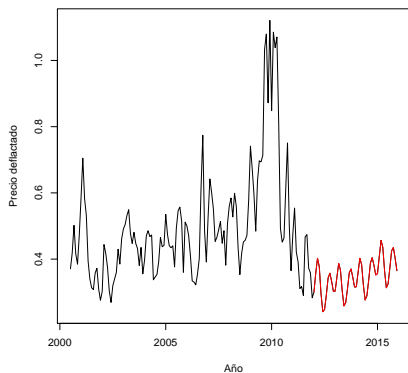
$$Y_{t+h} = \beta_F \tilde{F}_t + \beta_w \tilde{w}_t + \tilde{\epsilon}_{t+h} \quad (8)$$

- En la ecuación 7 tan solo se está llevando a cabo la estimación de los factores por medio de componentes principales.
- En la ecuación 8 se lleva a cabo el backward selection a los primeros 10 componentes principales F_t y a los doce rezagos del precio de la energía en bolsa w_t
- Siendo \tilde{F}_t y \tilde{w}_t las variables que quedan de F_t y w_t después de aplicar backward selection.

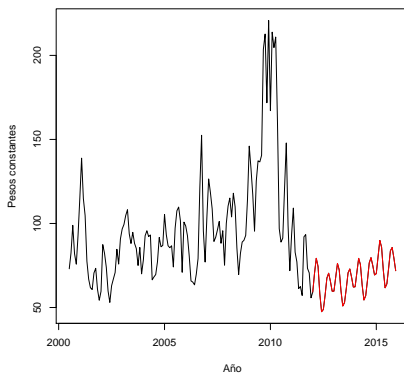
Modelo Dinámico de Factores

Tomando como última fecha observada diciembre 2011, el pronóstico obtenido 48 meses hacia adelante por medio de este modelo sería:

Precio promedio en bolsa deflactado



Precio promedio en bolsa



Combinación Óptima de Modelos

- Este modelo es una combinación del Modelo Dinámico de Factores con la parte del modelo de Backward Selection Regresión Lineal que hace el backward selection con p-valor 0.01.
- Una vez se tienen los respectivos valores pronosticados por estos dos modelos para Y_t , \widehat{Y}_t^{p1} y \widehat{Y}_t^{DFM} respectivamente, se combinan de manera óptima.
- Para combinar los modelos se quiere encontrar el γ óptimo tal que:

$$Y_t = \gamma Y_t^{p1} + (1 - \gamma) Y_t^{DFM} \quad (9)$$

con $t = 1, \dots, T$. Para esto se transforma la ecuación 9 restando Y_t^{DFM} a ambos lados y se usan los valores pronosticados para Y_t^{p1} y Y_t^{DFM} obteniendo:

$$Y_t - \widehat{Y}_t^{DFM} = \gamma(\widehat{Y}_t^{p1} - \widehat{Y}_t^{DFM}) \quad (10)$$

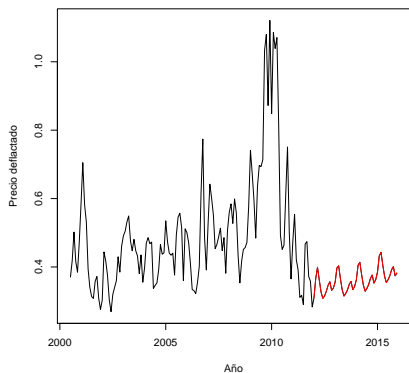
Así se obtiene el valor estimado para γ por medio de una regresión.

- Una vez estimado γ se pueden combinar los pronósticos \widehat{Y}_t^{p1} y \widehat{Y}_t^{DFM} , $t = (T + 1), \dots, (T + h)$ siendo h el horizonte del pronóstico, usando la ecuación 9.

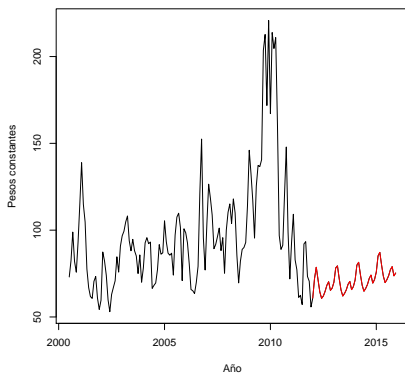
Combinación Óptima de Modelos

Tomando como última fecha observada diciembre 2011, el pronóstico obtenido 48 meses hacia adelante por medio de este modelo sería:

Precio promedio en bolsa deflactado



Precio promedio en bolsa



Intervalos de Confianza: Combinación Óptima de Modelos

Aplicando la técnica de *bootstrapping* los intervalos de confianza obtenidos para el pronóstico son:

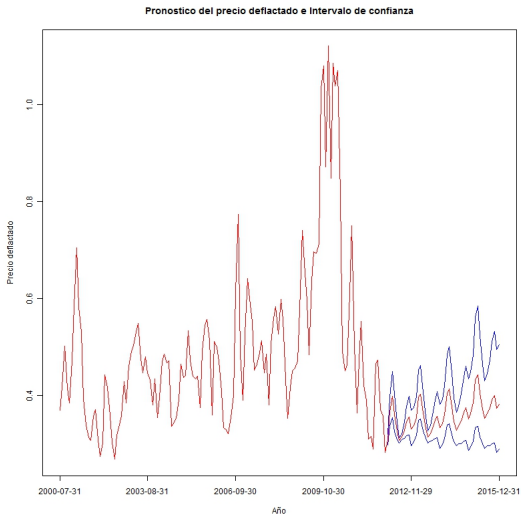


Figura: Pronostico del precio de la energía en Kw/h

Contenido

- 1 Teoría Básica de Series de Tiempo: Box - Jenkins
- 2 Aplicación
- 3 Hacia un Modelo Dinámico de Factores
- 4 Pronósticos Intermedios
- 5 Pronóstico del Precio de la Energía
- 6 Validación de Modelos**

Validación de los Modelos de Pronóstico

- Los criterios de comparación usados son:
 - Root Mean Square Error (RMSE)
 - Root Mean Square Percentage Error (RMSPE)
 - Mean Absolute Error (MAE)
 - Mean Absolute Percentage Error (MAPE)
 - Mean Percentage Error (MPE).
- Estos criterios se calculan de la siguiente manera:

$$RMSE = \frac{1}{N} \sum_t \sqrt{(Y_t - \widehat{Y}_t)^2} \quad (11)$$

$$RMSPE = \frac{1}{N} \sum_t \sqrt{\frac{(Y_t - \widehat{Y}_t)^2}{Y_t}} \quad (12)$$

$$MAE = \frac{1}{N} \sum_t |Y_t - \widehat{Y}_t| \quad (13)$$

$$MAPE = \frac{1}{N} \sum_t \left| \frac{Y_t - \widehat{Y}_t}{Y_t} \right| \quad (14)$$

$$MPE = \frac{1}{N} \sum_t \frac{Y_t - \widehat{Y}_t}{Y_t} \quad (15)$$

- El RMSE es una medida usada frecuentemente para medir la precisión de un modelo, compara las diferencias entre los valores estimados y los valores reales.
- Al no ser una medida relativa, sino absoluta, no permite comparar entre modelos para diferentes series.

$$RMSE = \frac{1}{N} \sum_t \sqrt{(Y_t - \widehat{Y}_t)^2}$$

- El RMSPE es la versión relativa del RMSE donde la diferencia entre el valor estimado y el real se divide por el valor real.

$$RMSPE = \frac{1}{N} \sum_t \sqrt{\frac{(Y_t - \widehat{Y}_t)^2}{Y_t}}$$

MAE y MAPE

- El MAE es el promedio de las desviaciones absolutas entre el valor predicho y el valor real, es una medida de que tan cerca están los pronósticos de los valores reales
- Al no ser una medida relativa, sino absoluta, no permite comparar entre modelos para diferentes series.

$$MAE = \frac{1}{N} \sum_t |Y_t - \widehat{Y}_t|$$

- Su versión relativa es el MAPE, donde las desviaciones entre el valor estimado y el real son divididas por el valor real presentando así la bondad de ajuste del modelo como un porcentaje.

$$MAPE = \frac{1}{N} \sum_t \left| \frac{Y_t - \widehat{Y}_t}{Y_t} \right|$$

$$MPE = \frac{1}{N} \sum_t \frac{Y_t - \widehat{Y}_t}{Y_t}$$

- Finalmente, el MPE al igual que es MAPE es un porcentaje de la desviación de los valores estimados frente a los reales.
- En el caso del MPE al no tomarse valor absoluto al sumar las desviaciones relativas éste representa el sesgo del modelo.
- Si el modelo se desvía siempre en la misma dirección, por encima o por debajo de los valores reales, estos errores se van a ir sumando.
- Mientras que si lo hace igual cantidad de veces en la misma magnitud en ambas direcciones el MPE debería ser aproximadamente cero.