

Parte 2

Modelos de pronóstico de corto plazo para precios del petróleo WTI spot

Por **Claudio Araya Sassi**

Ingeniero de estudios, Empresa Nacional del Petróleo, Chile.

Esta segunda parte del trabajo realizado por Araya Sassi expone la Metodología utilizada por el autor en su evaluación del desempeño de pronóstico de los diferentes modelos del petróleo WTI spot presentes en la literatura. Su objetivo busca proponer un nuevo modelo que permita pronosticar los precios del crudo WTI en el corto plazo. En la próxima edición, se publicarán los resultados conseguidos y las conclusiones.

Metodología

Series de Tiempo de Variables Relevantes

Se construyeron tres diferentes sets de datos, con las siguientes frecuencias: semanal, mensual y trimestral. Se consideraron inicialmente las variables más relevantes que empírica y teóricamente influyen en el comportamiento del mercado mundial del petróleo. A continuación, se detallan las series utilizadas, el período de la muestra, la frecuencia, la fuente de información, la unidad en que se expresan y el acrónimo utilizado.

Debido a la disponibilidad limitada de variables estructurales a más altas frecuencias, el set de datos semanal sólo incluye el inventario comercial de petróleo de Estados Unidos, los precios de WTI spot, precios de futuros NYMEX, Open Interest (OI) y Posición Larga de No-Comerciales (NCLP), estas últimas dos variables informadas los martes de cada semana y contabilizadas hasta el viernes anterior por la Comisión de Comercio de Futuros de Commodity (CFTC).

Prueba de Raíz Unitaria de Dickey-Fuller Aumentada (ADF)

La prueba de Dickey-Fuller Aumentada para raíz unitaria se basa en la siguiente regresión:

Δy_t=α₀+α₁t+β₁y_{t-1}+β₂Δy_{t-1}+...+β_pΔy_{t-p}+v_t (21)

Donde:

β₁=p-1

p : rezago del término de diferencia de la variable dependiente y

1. Planteamiento de la hipótesis

Las hipótesis nula y alternativa pueden ser escritas como:

H₀: β₁ = 0 : La serie es no estacionaria:
Tiene una raíz unitaria

H₁: β₁ < 0 : La serie es estacionaria

2. Estadísticos para la prueba

Se evalúa la hipótesis usando el **t** - **ratio** convencional para β₁:

t_{β1}=β̂₁/(se(β̂₁)) (23)

Donde:

β̂₁ = es el estimado de β₁ y

se(β̂₁) : es el coeficiente de error estándar

Dickey y Fuller demostraron que bajo la hipótesis nula de una raíz unitaria, este estadístico no sigue la distribución t de Student convencional, y ellos derivaron resultados asintóticos y simularon valores críticos para varios test y tamaños de muestra. Más recientemente, McKinnon (1991,1996) implementa un set de simulaciones mucho más grande que esas tabuladas por Dickey y Fuller. Además, McKinnon estima superficies de respuesta para los resultados de la simulación, permitiendo el cálculo de los valores críticos de Dickey-Fuller y *p-values* para muestras de tamaño arbitrario. Eviews utiliza los cálculos de valores críticos más recientes de McKinnon para construir los resultados del test. Asimismo, la selección de *p* está basada en el criterio de información de Schwartz (SIC; Schwartz Information Criterion).

Variable	Muestra		Frec.	Fuente	Unidad	Acrónimo
Precio spot WTI	02/01/1994	31/12/2007	S.M.T	PLATTS	US\$/barril	WTI
Precio futuro WTI NTMEX contrato 1	02/01/1994	31/12/2007	S.M.T	EIA	US\$/barril	F1
Precio futuro WTI NTMEX contrato 2	02/01/1994	31/12/2007	S.M.T	EIA	US\$/barril	F2
Precio futuro WTI NTMEX contrato 3	02/01/1994	31/12/2007	S.M.T	EIA	US\$/barril	F3
Precio futuro WTI NTMEX contrato 4	02/01/1994	31/12/2007	S.M.T	EIA	US\$/barril	F4
Open Interest Futuros WTI NYMEX	02/01/1994	31/12/2007	S.M.T	CFTC	Nº	OI
Non-Comercial Long Position	02/01/1994	31/12/2007	S.M.T	CFTC	Nº	NCLP
Índice de Precios del Productor	01-1994	12-2007	M.T	BLS	Nº	PPI
Producto Interno Bruto USA nominal	01-1994	12-2007	M.T	EIA	MMM US\$	GDP
Deflactor Producto Interno Bruto USA	01-1994	12-2007	M.T	BL	Nº	GDPUSDEF
Inventario Comercial de Petróleo USA	02/01/1994	31/12/2007	S.M.T	EIA	MM barriles	USCS
Inventario Industrial de OECD	01-1994	12-2007	M.T	EIA	MM barriles	OECDIS
Inventario Total de OECD	01-1994	12-2007	M.T	EIA	MM barriles	IN
Consumo de Petróleo de OECD	01-1994	12-2007	M.T	EIA	MMbbl/día	OECDP
Consumo de Petróleo No-OECD	01-1994	12-2007	M.T	EIA	MMbbl/día	NOECDP
Consumo de Petróleo Mundial	01-1994	12-2007	M.T	EIA	MMbbl/día	WC
Producción de Petróleo de OECD	01-1994	12-2007	M.T	EIA	MMbbl/día	OECDP
Producción de Petróleo No-OECD	01-1994	12-2007	M.T	EIA	MMbbl/día	NOECDP
Producción de Petróleo OPEP	01-1994	12-2007	M.T	EIA	MMbbl/día	OPEP
Producción de Petróleo No-OPEP	01-1994	12-2007	M.T	EIA	MMbbl/día	NOPEP
Producción de Petróleo Mundial	01-1994	12-2007	M.T	EIA	MMbbl/día	WP

Tabla 1. Lista completa de variables utilizadas en el análisis empírico.

Notas: S = frecuencia semanal; M = frecuencia mensual; T = frecuencia trimestral; MM = Millones; MMM = Mil millones; EIA = Energy Information Administration; BLS = Bureau of Labour Statics; CFTC = Commodity Futures Trading Commission; BL = Bloomberg

3. Regla de decisión

Si $|t| \leq |valor\ crítico\ DF| \rightarrow$ Rechace $H_0 \rightarrow$
Serie estacionaria

Si $|t| > |valor\ crítico\ DF| \rightarrow$ Acepte $H_0 \rightarrow$
Serie No estacionaria

Metodología de Especificación y Corrección Econométrica de Modelos

Significancia de Parámetros y Bondad de Ajuste

- La selección de los modelos se realizó de acuerdo con la significancia de los parámetros ($p < 0.05$) y coeficiente de determinación ajustado (R_2 ajustado).

Autocorrelación

Un rasgo común en las regresiones de series de tiempo es que los residuos están correlacionados con sus propios valores rezagados. Esta correlación serial viola el supuesto estándar de la teoría de regresión de que las perturbaciones no están correlacionadas con otras perturbaciones. Los problemas primarios con la correlación serial son:

- Los estimadores MCO (mínimos cuadrados ordinarios) no son los más eficientes entre los estimadores lineales.
- Los errores estándares de los estimadores calculados usando MCO no son los correctos, y generalmente son subestimados.
- Si existen variables dependientes rezagadas en el lado derecho de la ecuación, los estimadores MCO son sesgados e inconsistentes.

Detección de autocorrelación

La autocorrelación se detectó en primera instancia con el estadígrafo Durbin-Watson, y se verificó en segunda instancia con el Correlograma de Residuos.

En el caso de modelos con variables dependientes rezagadas en el lado derecho de la ecuación se aplicó la Prueba de Autocorrelación Breush-Godfrey (LM Test). Este test es alternativo al estadístico Q (Correlograma de residuos) para prueba de correlación serial. El LM Test pertenece a la

clase de tests asintóticos (muestras grandes) conocidos como Tests de Multiplicadores de Lagrange (Eviews, 2004). Al contrario del estadígrafo Durbin-Watson para errores AR(1) (autorregresivos de primer orden), el LM Test puede ser usado para testear órdenes más altas de autocorrelación y es aplicable si existen o no variables dependientes rezagadas. La hipótesis nula del LM Test es que no existe correlación serial hasta el rezago de orden p , donde p es un entero pre-especificado. Además, en los resultados de dicho test se observa una regresión de los residuales sobre los regresores originales X y los residuales rezagados hasta el orden p . Eviews reporta dos test estadísticos desde esta prueba de regresión. El estadístico F es un test de variable omitida para la significancia global de todos los residuos rezagados. El estadígrafo Breush-Godfrey LM Test es $Obs \cdot R^2$ -squared, el cual es calculado como el número de observaciones por el R^2 desde el test de regresión. Bajo las condiciones generales totales, el estadístico LM Test está asintóticamente distribuido como un $X^2(p)$.

Corrección por autocorrelación

Se eliminó la autocorrelación mediante técnicas que estiman, en primer lugar, el coeficiente de autocorrelación de primer orden de los residuos (ρ) y luego estiman la ecuación definitiva mediante una regresión lineal con variables transformadas: la variable X^* es igual a $X(t) - \rho \cdot X(t-1)$, donde X es la variable del modelo original.

En el modelo de dos variables independientes, además del rezago de la variable dependiente, se realizó la siguiente corrección por autocorrelación:

$$\ln(Y_t) = \alpha_0 + \beta_0 \ln(X_{1t}) + \gamma_0 \ln(X_{2t}) + \delta_0 \ln(Y_{t-1}) + \mu_t \quad (24)$$

$$\mu_t = \rho \mu_{t-1} + \varepsilon_t \quad (25)$$

realizando una diferenciación generalizada y sustituyendo (20) en (19) queda:

$$\ln(Y_t) = \alpha_0(1-\rho) + \rho \ln Y_{t-1} + \beta_0(\ln X_{1t} - \rho \ln X_{1,t-1}) + \gamma_0(\ln X_{2t} - \rho \ln X_{2,t-1}) + \delta_0(\ln Y_{t-1} - \rho \ln Y_{t-2}) + v_t \quad (26)$$

Cabe destacar que el valor AR(1) que arroja Eviews en el resultado de la regresión corresponde a ρ .

Estimación de modelos

Estimación de modelo financiero semanal

Se estimó el siguiente modelo semanal tal como el señalado en (1) —véase *Petrotecnia*, año XLIX, n° 3, junio 2008, p. 55—:

$$WTI_t = \alpha(1-\rho) + \rho WTI_{t-1} + \beta(F_{t+1} - \rho F_{t+2}) + \varepsilon_t$$

Donde:

WTI_t : precio del petróleo spot en el tiempo t
 F_{t+1} : precio del petróleo futuro en el tiempo $t+1$ con vencimiento T

α, β : parámetros

ε_t : término de error aleatorio o estocástico

$\rho = AR(1)$: coeficiente de autocorrelación de orden 1

Estimación de modelos financieros y de series de tiempo

a) Modelo 1: autorregresivo

$$WTI_t = \alpha + \beta WTI_{t-1} + \varepsilon_t \quad (28)$$

b) Modelo 2: modelo propuesto 1 logarítmico

$$\ln(WTI_t) = \alpha + \beta \ln(WTI_{t-1}) + \delta \ln(NCLP_t) + \varepsilon_t \quad (29)$$

c) Modelo 3: modelo propuesto 2

$$WTI_t = \alpha + \beta WTI_{t-1} + \delta NCLP_t + \varepsilon_t \quad (30)$$

d) Modelo 4: modelo de caminata aleatoria (Random Walk)

$$WTI_t = \alpha + WTI_{t-1} + \varepsilon_t \quad (31)$$

e) Modelo 5: modelo propuesto 3

$$WTI_t = \alpha + \beta NCLP_t + \delta PP_t + \varepsilon_t \quad (32)$$

Pronósticos estáticos y dinámicos

Un pronóstico estático para el precio del petróleo se define como un pronóstico de un paso hacia adelante

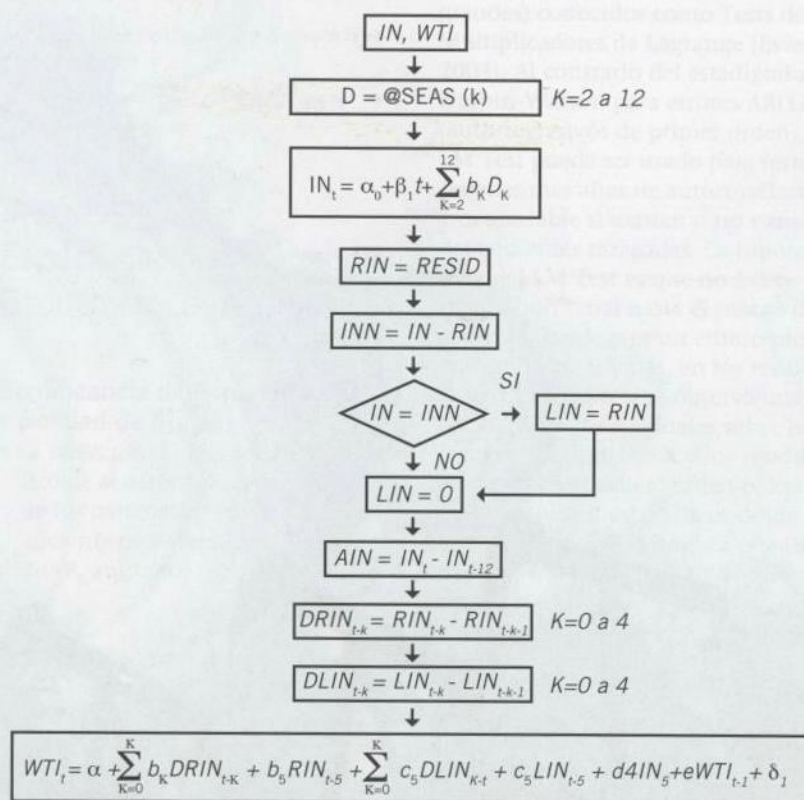


Figura 1. Diagrama de flujo de estimación de modelo de Ye *et al.* (2002). Resume las ecuaciones desde la (3) a la (8).

Fuente: Elaboración propia.

para $tS()$ $tWTI$.

Suponga que el modelo de referencia es:

$$S_t = \alpha S_{t-1} + \beta X_t + \epsilon_t \quad (33)$$

El modelo de referencia se estima sobre la muestra de estimación fija para obtener las estimaciones de los parámetros por medio de mínimos cuadrados ordinarios, es decir, $\hat{\alpha}$ y $\hat{\beta}$. Entonces, la secuencia de pronósticos estáticos se calcula como:

$$S_{t+1} = \hat{\alpha} S_t + \hat{\beta} X_{t+1}; S_{t+2} = \hat{\alpha} S_{t+1} + \hat{\beta} X_{t+2}, \text{ etc.}$$

Un pronóstico dinámico del precio del petróleo es un pronóstico de múltiples pasos hacia adelante para tS . Suponga que el modelo de referencia es idéntico al modelo descrito en (33). Al igual que en el marco de pronóstico estático, el modelo de referencia es estimado sobre la muestra de estimación fija para obtener las estimaciones de los parámetros

por medio de mínimos cuadrados ordinarios, es decir, $\hat{\alpha}$ y $\hat{\beta}$. Luego, la secuencia de pronósticos dinámicos se calcula como:

$$S_{t+1} = \hat{\alpha} S_t + \hat{\beta} X_{t+1}; S_{t+2} = \hat{\alpha} S_{t+1} + \hat{\beta} X_{t+2}, \text{ etc.}$$

etc. Es evidente que los pronósticos estáticos y dinámicos son idénticos para un paso hacia adelante, mientras que, para $n > 2$, los pronósticos estáticos de n pasos hacia adelante difieren de los correspondientes pronósticos dinámicos de n pasos hacia adelante, ya que la secuencia de valores reales $S_{t+1}, S_{t+2}, \dots$ entran en la expresión de los pronósticos estáticos, mientras que los pronósticos dinámicos dependen de la secuencia de valores predichos $S_{t+1}, S_{t+2}, \dots$

Evaluación de pronósticos

Suponga que la muestra de pronóstico es $t = T+1, \dots, T+k$ y que, en el

tiempo t , los valores reales y ajustados de la variable dependiente son, respectivamente, S_t y $\hat{S}_t$.

Las medidas de evaluación de pronóstico se pueden definir como:

a) Root Mean Square Error (raíz del error cuadrático medio):

$$RMSE = \sqrt{\frac{\sum_{t=T+1}^{T+k} (\hat{S}_t - S_t)^2}{k}} \quad (34)$$

b) Mean Absolute Error (error absoluto medio):

$$MAE = \frac{\sum_{t=T+1}^{T+k} |\hat{S}_t - S_t|}{k} \quad (35)$$

c) Mean Absolute Percentage Error (error porcentual absoluto medio):

$$MAPE = 100 \frac{\sum_{t=T+1}^{T+k} \left(\frac{|\hat{S}_t - S_t|}{S_t} \right)}{k} \quad (36)$$

d) Theil Inequality Coefficient (coeficiente de desigualdad de Theil):

$$Theil = \frac{\sqrt{\frac{\sum_{t=T+1}^{T+k} (\hat{S}_t - S_t)^2}{k}}}{\sqrt{\frac{\sum_{t=T+1}^{T+k} \hat{S}_t^2}{k}} + \sqrt{\frac{\sum_{t=T+1}^{T+k} S_t^2}{k}}} \quad (37)$$

Mientras, RMSE y MAE son dependientes de la escala de la variable dependiente y deben ser utilizados como medidas relativas para comparar pronósticos para las mismas series a través de diferentes modelos. Mientras más pequeño es el error, mejor es la capacidad de pronóstico de ese modelo de acuerdo con ese criterio.

MAPE y Theil son a escala invariante. Además, Theil va entre el rango de 0 a 1, con cero indicando ajuste perfecto.

El error cuadrático medio del pronóstico puede ser descompuesto como:

$$\sum (\hat{S}_i - S)^2 / k = ((\sum \hat{S}_i / k) - \bar{S})^2 + (\sigma_s - \sigma_s)^2 + 2(1-r) \sigma_s - \sigma_s$$

Donde:

$\sum \hat{S}_i / k$: media de $\hat{S}_i$

$\bar{S}$: media de S

σ_s : desviación estándar de $\hat{S}_i$

σ_s : desviación estándar de S (38)

enero de 1994 a diciembre de 2006, mientras que el intervalo desde enero a diciembre de 2007 es utilizado para la evaluación de los pronósticos, excepto para el modelo de inventario de Ye *et al.* (2002) cuya muestra utilizada para evaluar su pronóstico corresponde de enero a septiembre de 2007.

Este trabajo técnico continúa en la próxima edición de *Petrotecnia*:

"Modelos de pronóstico de corto plazo para precios del petróleo WTI spot". Tercera parte: Resultados y conclusiones. Bibliografía y sitios consultados.

Las proporciones pueden ser definidas como:

e) Bias Proportion (proporción de sesgo): da cuenta de cuán lejos se encuentra la media del pronóstico de la media de la serie real.

$$BP = \frac{((\sum \hat{S}_i - k) - \bar{S})^2}{\sum (\hat{S}_i - S)^2 / k} \quad (39)$$

f) Variance Proportion (proporción de varianza): da cuenta de cuán lejos se encuentra la variación del pronóstico de la variación de la serie real.

$$VP = \frac{(\hat{\sigma}_s - \sigma_s)^2}{\sum (\hat{S}_i - S)^2 / k} \quad (40)$$

g) Covariance Proportion (proporción de covarianza): Mide la permanencia del error de pronóstico no sistemático.

$$CP = \frac{2(1-r) \sigma_s - \sigma_s}{\sum (\hat{S}_i - S)^2 / k} \quad (41)$$

Note que las proporciones de sesgo, varianza y covarianza suman 1.

Si el pronóstico es "bueno", las proporciones de sesgo y varianza deberían ser pequeñas por lo que la mayoría del error debería estar concentrado en la proporción de covarianza.

Muestras de estimación y de pronósticos

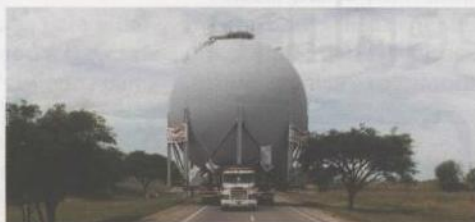
El período de estimación para todos los modelos corresponde de



grúas
transportes
técnica
rigging
seguridad
excelencia

COAMTRA
INGENIERIA DEL MOVIMIENTO

Salta 1212 (1872) Avellaneda
Buenos Aires - Argentina
Tel.: 4001-3600
www.coamtra.com.ar



COAMTRA es una compañía Argentina que opera en el país desde hace más de 25 años y se especializa en ingeniería de movimientos de componentes de gran porte.

La empresa tiene activa presencia en las principales obras de infraestructura del país con una capacidad operativa en izajes de 15 a 1.000 toneladas - con grúas telescópicas y reticuladas - y una flota de transportes compuesta por una gama de capacidad portante desde 1 hasta 700 toneladas. La filosofía de COAMTRA tiene como pilares insoslayables la excelencia en el servicio, la seguridad de todas y cada una de sus operaciones y el cuidado del medio ambiente.

