www.cya.unam.mx/index.php/cya

Contaduría y Administración 65 (4), 2020, 1-19

Mean reversion in Mexico's real oil price series

Reversión a la media en las series de precios reales del petróleo en México

José Carlos Ramírez Sánchez[1]*, Fernando Cruz Aranda[2]
Agustín Cabrera Llanos[3]

[1]Profesor-investigador del Departamento de Estudios Económicos
de El Colegio de la Frontera Norte, Tijuana, México
[2]Universidad Panamericana, México
[3]Unidad Profesional Interdisciplinaria de Biotecnología
del Instituto Politécnico Nacional, México

Received February 7, 2019; accepted June 4, 2019
Available online June 4, 2019

Abstract

This paper aims to show a mean reversion pattern in the series of real prices of oil exported by Mexico to the American Continent between January 1999 and June 2017. For this purpose, we use a stochastic difference-equation to make forecasts with a window of six and twelve months. The main results drawn from the best-fit model show that there is a reversion to the long-term mean in prices initially assumed to be rational. Other statistical tests confirm that this reversion is persistent because the shocks produced on real prices do not involve permanent changes.

JEL Code: C01, C13, C65, G12, G23
Keywords: Mexic's oil prices; mean reversion; efficient-market hypothesis; impulse-response functions; price forecasting

---

* Corresponding author.
E-mail address: jcramirez@colef.mx (J. C. Ramírez Sánchez).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

http://dx.doi.org/10.22201/fca.24488410e.2020.2453
0186- 1042/©2019 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. This is an open access article under the CC BY-NC-SA (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Resumen

El objetivo de este documento es mostrar la existencia de un patrón de reversión a la media en la serie de precios reales del petróleo exportado por México al continente americano entre enero de 1999 y junio de 2017. Con ese fin adaptamos una ecuación en diferencias estocástica a la serie de precios de la variedad Maya para hacer pronósticos dentro y fuera de la muestra, con una ventana de seis y doce meses. Los principales resultados obtenidos muestran que, en efecto, hay una reversión a la media de largo plazo en los precios inicialmente supuestos como racionales. Otras pruebas estadísticas confirman que esta reversión a la media es persistente en virtud de que los shocks producidos sobre los precios reales no involucran cambios permanentes.

Código JEL: C01, C13, C65, G12, G23
Palabras clave: precios de petróleo de México; reversión a la media; hipótesis de mercados eficientes; funciones de impulso respuesta; pronóstico de precios

---

Introduction

The hypothesis that crude oil prices follow a random walk without drift is a reference point for any forecast. Its rejection or acceptance decides the usefulness of the predictive exercise since, without testing, the model used would be questionable. In order to understand the behavior of crude oil prices, it is important to know the reasons for the adoption of this hypothesis. For this purpose, it is convenient to divide these reasons into theoretical and statistical ones. Among the first reasons, two stand out. These two reasons are related to equilibrium conditions derived from the possibility of crude oil storage and futures contracting by agents (Hamilton, 2009).

Concerning the possibility of storage, the producer will always be forced to comply with condition (1). This condition establishes that the expected value of the future price $E\left(P_{t+1}\right)$ is equal to the spot price P_t plus the present value of the net carrying cost or any other premium on the risk C_t, that will induce investors to hold higher or lower levels of crude oil inventories.

$$E\left(P_{t+1}\right)=P_t+C_t \tag{1}$$

In cases where the equilibrium in (1) is lost, the difference will tend to be corrected once the agent's actions, guided by expectations, reflect the adjustments in P_t. Thus, if investors believe that

$E\left(P_{t+1}\right) > P_t + C_t$ ($E\left(P_{t+1}\right) < P_t + C_t$), they will seek to buy (sell) larger volumes of crude oil to store (release) it and sell it in the future (present), with a consequent increase (decrease) in the spot price.

A similar situation arises when there is the possibility of contracting a future F_t to acquire a certain volume of crude oil one year later. If the oil price expectations are such that $E\left(P_{t+1}\right) > F_t$ ($E\left(P_{t+1}\right) < F_t$), then it is clear that agents will find it more profitable to contract (not contract) futures and wait a year to initiate (start now) operation. Correspondingly, an increase (decrease) in demand for futures will push up (down) F_t values to restore equilibrium in (2).

$$F_t = E\left(P_{t+1}\right) + R_t \tag{2}$$

where R_t is the expected value of any risk premium or transaction cost

The equilibriums in (1) and (2) are two different ways of paraphrasing the weak efficient market hypothesis (EMH). According to this hypothesis, crude oil prices describe a random walk with independent increments, whose predictions over time $t+1$ depend entirely on the information flowing in that period. That is why there is no possibility of arbitrage using the information of the period t unless there are imperfections in the marketplace.

It is possible to see the implications of the EMH more formally by considering the ex-post rational prices P_t^* or prices that include the discounted present value of risk premiums and transaction costs, i.e.:

$$P_t^* = E(P_{t+1}) - C_t = E(P_{t+1}) + R_t \tag{3}$$

According to the EMH, the real prices P_t and F_t are the best ex-post rational price forecast, and any surprise movements in the oil market will depend exclusively on P_t^* . However, it is impossible for the condition $P_t^* = P_t$ always to hold due to random deviation throughout the period. Hence it is more feasible to expect that:

$$P_t^* = P_t + u_t \tag{4}$$

where u_t is the forecast error caused by the issuance of debt or repurchase of companies, among other explanations.

Due to the assumption of independent increments of the random walk $Var(P_t^*) = Var(P_t) + Var(u_t)$ since u_t and P_t are also independent (Shiller, 2003). Therefore, the maximum possible variance of an optimal forecast cannot be greater than the variance of the forecasted variable, and the following inequality must limit the volatilities of the rational and real price series:

$$Var(P_t) \leq Var(P_t^*) \tag{5}$$

The empirical tests of (4) and (5) have produced an uneven amount of literature that falls into two categories. Here it is necessary to go into the statistical reasons. The first category argues that, in effect, real prices follow a random walk without drift and are therefore unpredictable using past information, whether expressed in levels or growth rates (Hamilton, 2009). The rationale lies in the fact that u_t will grow at a rate $\sigma\sqrt{t}$, which is precisely the standard deviation of a random walk in any simulation designed to predict rational prices based on past prices.

On the contrary, the second category considers that crude oil prices have other determinants that make their prediction possible. Among these determinants are the elasticities of oil product prices and consumer income, the supply and price banding policies, the scarcity of crude oil, the depletion of the main oilfields, and the emergence of new consumers (Hamilton, 2009). Other determinants often mentioned by the literature are (i) changes in the supply of hydrocarbons due to political events, discoveries, and technological improvements in the extraction process; and (ii) abrupt changes in demand or expectations about the future rate of oil supply (Baumeister & Lilian, 2016). Each determinant has a different weight depending on the regions and periods considered.

The theoretical ground for the forecasts rests precisely on the fact that these determinants divert real oil prices from their rational prices on a temporary (regression toward the mean) or permanent (regime shift) basis. The search for statistical patterns includes methods dealing with linear relationships, such as ARIMA, GARCH, and ECM models, and nonlinear ones, such as neural network models, between crude oil prices and the variables associated with the determinants of oil prices (Safari & Davallou, 2018). The results provide mixed evidence to support both types of patterns.

This study falls into the second category and aims to show that the oil price series of the Maya variety exhibits long-term mean reversion. In other words, volatilities are such that they significantly violate the limits of equation (5) and, therefore, make it possible to associate crude oil prices with a predictable behavior pattern in the short term. The results obtained with the model simulation and Andrews' test for stationarity reveal that the mean reversion pattern is persistent between 1999 and 2017.

This article has two additional sections. The second offers a brief literature review, followed by the derivation of the stochastic price difference equation and the statistical analysis of the database. The third section presents the simulation and different forecast scenarios. The conclusions summarize the main results.

Review of the literature and methodology

There are relatively few studies investigating mean reversion patterns in real crude oil prices compared to the abundant literature on forecasting techniques. The reasons seem to lie in the strong association of these patterns with the behavior of variables such as interest rates and the overwhelming predominance of articles using new heuristic and hybrid methods in forecasting.

In any case, existing articles provide evidence of mean reversion in specific cases, such as in the price banding policy used by OPEC in its oil supply management strategy (Hammoudeh, 1996), and in general cases, such as in the commodity price series used by the World Bank for several countries between 1958 and 1997 (Cashin et al., 2000). These articles confirm that despite oil price shocks being of long duration, the series recovers its original level after some time (see Choi & Hammoudeh, 2009; Li & Thompson, 2010). In other words, they find that oil price series are not unit root processes and that shocks can shift prices by more than 100 units, recovering half of their original level in three years (Cashin et al., 2000).

In Mexico, most of the literature is focused on analyzing the effects of crude oil price volatilities on the exchange rate or stock market performance but not on investigating mean reversion patterns (see, e.g., Bermúdez et al., 2018; Valdés et al., 2016). This study proposes a dynamic equation based on the Cox, Ingersoll, and Ross model (CIR, 1985) and a simulation strategy using the series of real prices of the Maya variety exported to the American continent to fill this gap.

Mean reversion equation

This study's dynamic equation derives from the CIR model. This model has some advantages for analyzing oil time series, such as its condition of non-negativity (not guaranteed, for example, in Vasicek, 1977) and its ductility in discretizing the price parameters. Both features make the equation here used an ideal means to detect mean reversion in autoregressive models, whose simulation requires the computation of persistence measures. Specifically, the equation assumes that prices adjust the following dynamics:

$$dP_{t,x} = a_x\left(b_x - P_{t,x}\right)dt + \sigma_x\sqrt{P_{t,x}}dW_{x,t}, \tag{6}$$

where $P_{t,x}$ is the price of crude oil at the time t, a_x is the speed of adjustment of crude oil prices for x =Maya variety, $\left(b_x - P_{t,x}\right)$ is the correction to the mean b_x in the long term, σ_x is the standard deviation of prices, and $t = t_1, t_2, t_3, \dots, t_{n_c}$ the periods considered; $\{W_x\}_{t\geq 0}$ is a Brownian motion defined over a fixed space of probability $(\Omega, \mathcal{F}, \mathbb{P})$ in which Ω is the sample space, $\mathcal{F}$ a sigma-algebra, and $\mathbb{P}$ the probability measure.

Since, in its current state, equation (6) is very general and not useful for any simulation, it is necessary to discretize it first to treat it with statistical methods. Therefore, in order to properly calculate the values of the parameters of interest, (a_x, b_x and ${\sigma_x}^2$) in a time series, it is advisable to apply the change of variable $y_{t;x} = y\left(P_{t;x}, t\right) = 2\left[P_{t;x}\right]^{1/2}$. Thus, the first- and second-order derivatives of $y_{t;x}$ concerning $P_{t,x}$take the form:

$$\frac{\partial y_{t;x}}{\partial P_{t,x}} = \frac{1}{P_{t,x}} = \frac{2}{y_{t;x}}$$

and

$$\frac{\partial^2 y_{t;x}}{\partial P_{t,x}^2} = -\frac{1}{2P_{t,x}\sqrt{P_{t;x}}} = -\frac{1}{P_{t;x}y_{t;x}}.$$

Once Ito's lemma has been applied, $dy_{t,x}$ is obtained as a stochastic differential equation with constant variance, that is:

$$dy_{t,x} = \left[\left(2a_x b_x - \frac{{\sigma_x}^2}{2}\right)\frac{1}{y_{t;x}} - \left(1 - \frac{a_x}{2}\right)y_{t;x}\right]dt + \sigma_x dW_{x,t}. \tag{7}$$

Whose discrete expression is:

$$y_{t+1,x} = \beta_{1;x}\left[\frac{1}{y_{t;x}}\right] + \beta_{2;x}y_{t;x} + \varepsilon_t. \tag{8}$$

with $\beta_{1;x} = 2a_x b_x - \frac{{\sigma_x}^2}{2}, \beta_{2;x} = 1 - \frac{a_x}{2}$, $x = 0,1, \ldots$ and $t = t_1, t_2, t_3, \ldots, t_{n_c}$.

After some manipulation, the conclusion is that $\widehat{a_x} = 2\left(1 - \widehat{\beta_2}\right)$ and $\widehat{b_x} = \frac{\sigma_x^2 + 2\widehat{\beta_1}}{8\left(4 - \widehat{\beta_2}\right)}$. With these results, the estimated price is given by:

$$\widehat{P}(x,t) = \left(\frac{\widehat{y_x}}{2}\right)^2. \tag{9}$$

Equations (8) and (9) are now suitable for this study because the estimation of their parameters makes it possible to dynamically analyze the future behavior of crude oil prices and their volatility (Cruz, 2007; Venegas, 2008). Notice that there are other alternative methods of discretizing equation (7) over uniform and irregular intervals, such as the sets of nodes with adjacency relations, which are sensitive to the meshing employed in the partition of the domain of (7) (see Salt et al., 2016). This study decided to use the method described above because of the possibility of obtaining closed forms that facilitate the computation of the parameters employing statistical estimations.

Database

Figure 1 illustrates the monthly evolution of prices of the three types of crude oil that Mexico exports to the Americas (Isthmus, Maya, & Olmeca) during the sample period. According to data from the Pemex website, observations amount to 227 and are expressed in dollars per barrel (see references). Of the total observations, 215 (94.74%) and 221 (97.37%) are used to calibrate the statistical models. In contrast, the remaining ones are addressed to compute different forecast error metrics, such as those suggested by Hansen (2005). Thus, the analysis in this section refers only to the 221 observations comprised between January 1999 and June 2017.

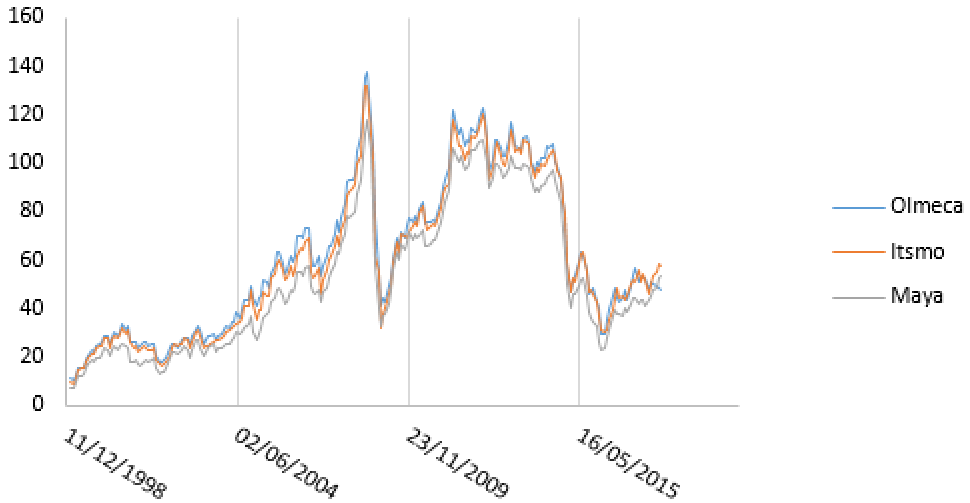

Figure 1. Monthly prices of the Olmeca, Isthmus, and Maya varieties between January 1999 and June 2017.
Source: Created by the authors based on data from http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx, and http://ebdi.pemex.com/bdi/bdiController.do?action=cuadro&subAction=applyOptions
Note: Six periods can be distinguished in the figure:[1] January 1999 to July 2006; [2] August 2006 to June 2008; [3] July 2008 to December 2008; [4] January 2009 to April 2011; [5] May 2011 to January 2015 and [6] February 2015 to June 2017.

According to Figure 1, prices of the three varieties move very closely and with the same trend throughout the period. In particular, the Olmeca variety recorded the highest average price ($61.52), followed by the Isthmus ($58.01) and Maya ($52.42) ones, although with great volatility in all three cases (32.29%, 31.65 %, and 30.10%, respectively). This last aspect becomes more evident when considering the volatility of yields for each of the varieties since it is clear that it reaches a maximum of 24.26% in the Isthmus variety and a minimum of 2.62% in the Maya variety (see Table 1).

Table 1
Volatility ranges for Mexican crude oil yields in its three varieties

	Yield volatility		
	Olmeca	Isthmus	Maya
Maximum	18.32%	24.26%	19.58%
Minimum	3.61%	2.82%	2.62%
Average	8.52%	9.09%	8.49%

Source: Created by the authors based on data from http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx
http://ebdi.pemex.com/bdi/bdiController.do?action=cuadro&subAction=applyOptions
Note: The rolling window length is one year with monthly adjustments.

This volatility, however, is not only very irregular by variety but also by time sub-period. Figure 2 shows that volatility ranges grow from 2008, then fall back to their previous trend in February 2010, fall beyond the average (8.49%) between 2011 and 2014, and, finally, exceed that value between 2015 and 2016. In other words, there is inconsistent behavior within the observation period, which makes it necessary to justify the periodization adopted.

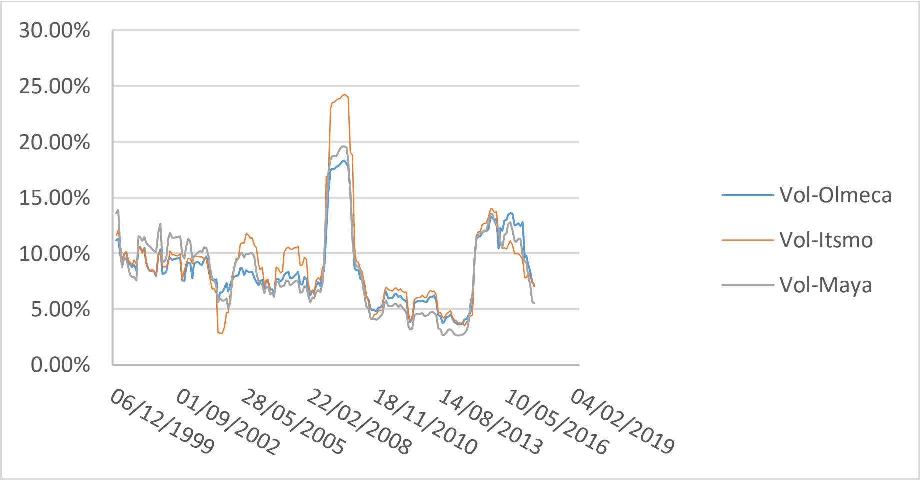

Figure 2. Evolution of the annual moving volatility of Mexican crude oil yields in its three varieties (January 1999 to June 2017).
Source: Created by the authors based on data taken from http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx, and http://ebdi.pemex.com/bdi/bdiController.do?action=cuadro&subAction=applyOptions
Note: The first yield volatility data is for January 2000. The length of the rolling window is one year with monthly adjustments.

As has been widely documented in the specialized literature, different factors have played an important role in the volatility of international crude oil prices (see Baumeister & Lilian 2016). Thus, the sustained rise in prices between 1999 and 2008 (covering the first two sub-periods in Figure 1) is largely due to the fast-growing oil demand in China and India (see Figure 1). Similarly, the sharp fall in the price of a barrel between June and December 2008 (third period) is explained by the 2008-2009 financial crisis, as is the rebound in the price between January 2009 and April 2011 due to the decision of OPEC to maintain a regulated supply among its member countries (fourth period). The abandonment of these measures by OPEC, which kept the price of a barrel stable at levels above $100 in the following three years (fifth period), together with the presence of a strong dollar and the resurgence of the US as an oil-producing power, results in an impressive price drop between July 2014 and March 2016 (sixth period).

http://dx.doi.org/10.22201/fca.24488410e.2020.2453

The combined result of these swings is a marked pattern of moving away from the trend, followed by a not necessarily long-lasting return to the price level of previous periods, as can be seen in Table 2 (which takes the US producer price index for some representative years of the six sub-periods). This Table illustrates that, for the Maya variety, there is a repeated return to its original level of real prices (2005) in the periods around December 2008, January 2015, and June 2017. Meanwhile, for the Olmeca variety, this return is earlier (January 2007), although not as recurrent as for the Maya. Finally, in the Isthmus variety, the return takes, on average, a longer time interval than for the other varieties (10 years).

The information from Table 3 can differentially explain the behavior of the mean reversion pattern in the short term. After reviewing the statistics in Table 3, one can realize the variety of means, volatilities, and distributions (normal and non-normal) associated with each sub-period and, therefore, the great analytical possibilities that a separate study can offer. The problem is that dividing the entire observation span into sub-periods can accentuate the upward or downward trend in Figure 1 and eliminate the mean reversion pattern.

Table 2
Real prices in dollars per barrel for the three varieties of oil in Mexico during some relevant dates in the period from January 2005 to April 2017 (January 1999=100)

Date	Olmeca	Isthmus	Maya	Date	Olmeca	Isthmus	Maya
January 2005	43.76	33.60	25.90	May 2010	53.88	54.88	47.38
July 2006	58.26	53.75	44.98	April 2011	80.73	78.29	70.41
January 2007	43.31	37.27	34.49	April 2014	68.65	64.77	59.99
July 2008	93.58	89.49	80.26	January 2015	32.20	32.01	27.59
December 2008	31.92	24.79	25.32	February 2016	18.86	21.30	16.76
April 2010	60.81	59.62	52.08	June 2017	35.19	35.14	27.71

Source: created by the authors with data from Pemex http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx and from the Producer Price Index (Total Manufacturing Index) statistics of the US Bureau of Labor Statistics at https://www.bls.gov/ppi/
Note: nominal prices were deflated after a base change from Dec 1984 to Jan 1999 in the series deployed by https://fred.stlouisfed.org/series/PCUOMFGOMFG.

It should bear in mind that measures of persistence require that the autoregressive processes are not unit root or that the shocks to the price series are not permanent. A periodization comprising the months from January 1999 to June 2008, or the first two sub-periods of Figure 1, is meaningless. In that period, there is no visible return of real prices to a hypothetical proposed level, and, therefore, it is highly likely that a mean reversion pattern (such as impulse response functions) has no meaning at all. Similarly, taking another large split separately, such as the one that includes the steep rises and falls of the July 2008-June 2017 period (or of the remaining four sub-periods in Figure 1), may be inappropriate. Certainly, the trend has not yet absorbed price shocks with a long half-life of survival. The analysis of reversal patterns requires long periods to observe the exhaustion of the effects of shocks on prices. Any excessive breakdown of periodization is risky, especially if the series exhibits distinct reversals in price levels (see Table 2) and a smoothed trend at the end, indicating the stabilization of shocks. The analysis of the sub-periods highlighted in Figure 1 should include other statistical methods suitable for small samples that complement the understanding of the mean reversion pattern but do not replace it.

The statistical analysis focuses on the Maya crude oil variety and the American market as the most important product and region for Mexican crude oil exports abroad (80.29 %). Although the share of that market has been decreasing significantly in Mexican crude oil sales (going from 81.2% in 1999 to 52% in 2017), its importance in the study period is greater than that of any other variety, so the conclusions obtained here do not lose generality.

Table 3
Descriptive statistics of crude oil price series in Mexico for the three varieties in six-time sub-periods

Crude oil	Mean	Variance	Price volatility	Bias	Kurtosis	Jarque Bera	Probability
January 1999 to July 2006							
Olmeca	33.9229	220.7480	14.8576	0.9985	3.1473	15.2036	0.0005
Isthmus	31.7375	186.8394	13.6689	0.9680	3.1170	14.2633	0.0008
Maya	26.4386		11.3985	0.9952	3.3629	15.5216	0.0004
August 2006 to June 2008							
Olmeca	80.1745	496.7527	22.2880	0.9139	2.9832	3.2020	0.2017
Isthmus	75.0514	518.4428	22.7693	0.9110	3.0350	3.1825	0.2036
Maya	66.3336	379.6535	19.4847	0.7767	2.6896	2.4046	0.3005
July 2008 to December 2008							
Olmeca	89.3832	1406.8673	37.5082	-0.0523	1.5606	0.5207	0.7707
Isthmus	82.3397	1480.4318	38.4764	-0.0143	1.5292	0.5410	0.7630
Maya	73.2120	1208.7860	34.7676	0.1083	1.4631	o.6022	0.7399
January 2009 to April 2011							
Olmeca	76.0438	309.5819	17.5949	0.3721	3.7626	1.3245	0.5157
Isthmus	73.4325	307.9639	17.5489	0.1692	3.7879	0.8578	0.6512
Maya	68.1441	246.4429	15.6985	0.1377	3.6544	0.5882	0.7452
May 2011 to Janaury 2015							
Olmeca	103.8125	197.6060	14.0572	-2.3535	9.3587	117.3552	0.0000
Isthmus	101.4669		13.1273	-2.3579	9.8224	129.0048	0.0000
Maya	98.8672		13.5287	-2.3836	9.0753	111.8153	0.0000
February 2015 to June 2017							
Olmeca	47.9396	68.9507	8.3037	-0.3857	2.9568	0.7213	0.6972
Isthmus	47.2895	75.6678	8.6987	-0.2681	2.6646	0.7853	0.7853
Maya	39.1513	58.2755	7.6338	-0.4103	2.7485	0.8901	0.6408

Source: Created by the authors based on data from http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx

Results

The second stage of the methodology consists of estimating the parameters of equations (8) and (9) a_x, b_x and $\widehat{P}(x,t)$by selecting the best statistical fit, applying Andrews' test, using some sample versions of loss functions, and making some model forecast error metrics created with 94.74% and 97.37% of the sample. Among these metrics, this study considers those that are already common in the specialized literature, such as the Mean Squared Error (MSE), the Root Mean Square Error (RMSE), the Mean Absolute Deviation (MAD), and the Mean Absolute Percentage Error (MAPE) (see Hansen & Lunde, 2005; Safari & Davallou, 2018).

Concerning the Andrews test, it is important to mention that it is an unbiased estimate of the median of a first-order, unit-root, autoregressive process (AR/UR), used to determine the persistence of shocks in the price series. If the series has a unit root, the impulse-response function (FIR) used in the test will never fade or, in other words, the shock in the price series will be permanent. Conversely, if the series is stationary, the IRF will fade in a finite time, and the shocks will be temporary. The FIR is defined as:

$$FIR(\tau) = \alpha^{\tau} \text{ for } \tau = 0,1,2,... \tag{10}$$

where α is the unbiased estimator of the median of the autoregressive parameter and τ is the time horizon where the shock is measured. Two other scalar measures of persistence linked to the FIR are the accumulated impulse response function (AIR) and the half-life of a shock unit (VMS), whose formal expressions are:

$$AIR = \sum_{t=0}^{\infty} FIR(\tau) = (1-\alpha)^{-1} \tag{11}$$

$$VMS = ABS\left(\frac{\ln\left(\frac{1}{2}\right)}{\ln \alpha} \right) \tag{12}$$

The advantage of using Andrews' test over the standard Dickey-Fuller or Philipps-Perron unit root measures lies in its ability to correct biases introduced by asymmetries in the distributions of the estimators of the autoregressive parameters, which are not accounted for by these traditional measures. The distributions are generally left-skewed. They also have heavy tails, as with the oil price series. The above makes the median a better estimator than the mean as a measure of central tendency in the least-squares estimation of the AR/UR models. Moreover, the test combines the unbiasedness property with point and interval estimates, which reduce the probability of making type I and II errors in determining the persistence of shocks in the price series (Cashin et al., 2000; Andrews, 1993).

In the estimation of the $AR(1)$ models, this study considers that the values of $\widehat{\mathrm{P}}(\mathrm{x},\mathrm{t})$ that tend toward $\mathrm{b_x}$ are the proxies of rational prices P_t^*, and the values of P_t are nominal prices deflated by the US producer price index based on January 1999.

The best-fit model for the period that includes 97.37% of the sample

When considering the price series between January 1999 and June 2017, it is possible to observe that the fitting model of equation (8) deployed in (13), although it presents a high R^2 (97.95%), is affected by serial autocorrelation and heteroscedasticity. Indeed, computing the Durbin-Watson (DW), Breusch-Godfrey (BG), and White tests, the finding is that the null hypotheses of no serial autocorrelation and homoscedasticity are consistently rejected in (13). Specifically, for the serial autocorrelation test, the finding was that $\mathrm{BG} = 55.08826 > \chi^2 = 5.99$, while for homoscedasticity, the statistic yielded the value $\mathrm{TR}^2_{\mathrm{errores}} = 11.0217 > \chi^2 = 5.99$, with $T = 221$ and DW=1.0132.

$$\mathrm{y}_{\mathrm{t+1,Maya}} = 1.8119\mathrm{y}^{-1}_{\mathrm{t;Maya}} + 0.9919\mathrm{y}_{\mathrm{t;Maya}} + \mathrm{u}_{\mathrm{t}}. \tag{13}$$

In order to face the serial autocorrelation and heteroscedasticity problems, the study initially applied the correction ρ=1-DW/2=0.4933775 to each of the variables in equation (13) to extract their first difference. According to the White-Huber minimum variance test, the resulting new model in (14) is non-heteroscedastic. Unlike the White test, which uses a correlation matrix and the cross-term of the explanatory variables to reject the null hypothesis of non-heteroscedasticity, the White-Huber test uses the correlation matrix to minimize the variance of the errors in the model used. Therefore, it is more exhaustive than the White test. After obtaining a p value of 0.0843 and a statistical $\mathrm{t} = -1.7345$ in the coefficient of the cross term, the new results yield a variance of the squared residuals equal to zero. Other statistics in (14) confirm that there is no serial autocorrelation since $\mathrm{BG} = 0.3750 < \chi^2 = 5.99$ and DW=1.9173, in addition to displaying a high $\mathrm{R}^2 = 94.03\%$.

$$\Delta y_{t+1,\mathrm{Maya,rt}} = 3.7557\ \Delta y_{t;\mathrm{Maya}}^{-1} + 0.9810\Delta y_{t;\mathrm{Maya,rt}} + u_t. \tag{14}$$

The two parameters estimated in (14) are not necessarily pointwise but can also be included in intervals with different confidence levels, as illustrated in Table 4. Thus, for example, the value of the unbiased estimator of the median of $\Delta y_{t;\mathrm{Maya,rt}}$ can be contained within the minimum (965313) and maximum values (0.996754) of a 95% interval or, depending on the sample size's quantile, within intervals bounded on the left or right presented below.

Table 4
Interval of the parameter values of (14) with different confidence levels

Coefficient	90%		95%		99%	
	minimum	maximum	minimum	maximum	minimum	maximum
3.756792	1.587996	5.925588	0.169145	6.344440	0.345121	7.168464
0.981034	0.967858	0.994210	0.965313	0.996754	0.960307	1.001760

Source: created by the authors

An important aspect is that (14) makes it possible to calculate the values of the parameters quickly $\widehat{a_x} = 0.037932$ and $\widehat{b_x} = 37.825686$ after obtaining $\hat{\beta}_1$ and $\hat{\beta}_2$ (see equation 8) and equation (9) price series, displayed in Figure 4. It is possible to see that P_t^* moves toward a mean $\widehat{b_x}$ of 37.8257 long-term dollars per barrel with a speed of $\widehat{a_x}$ 0. 0379 dollars/barrel/month. Andrews' test confirms this tendency to mean reversion $\mathrm{P}_\alpha = \widehat{\alpha_{\mathrm{mco}}} = q_p(\alpha) = \mathrm{p} = \mathrm{P}_{0.98}(= 0.98103 \leq 0.998) = 0.05$, which indicates that the unbiased estimator of the median of the auto-regressive parameter is contained in the 95% confidence interval. It is important to note that by (10), the FIR will tend to fade over time as τ grows, as in that case $\lim_{\tau\to\infty} \alpha^\tau = 0$ because $\alpha = 0.98103$. Therefore, while it is true that any unit shock to Maya crude oil prices causes a significant shift in its overall level of 52.71 units (per equation 11) during this period, it is also true that in a finite number of months, 36.18 months (per equation 12), half of the impulse response of that shock will tend to dissipate.

http://dx.doi.org/10.22201/fca.24488410e.2020.2453

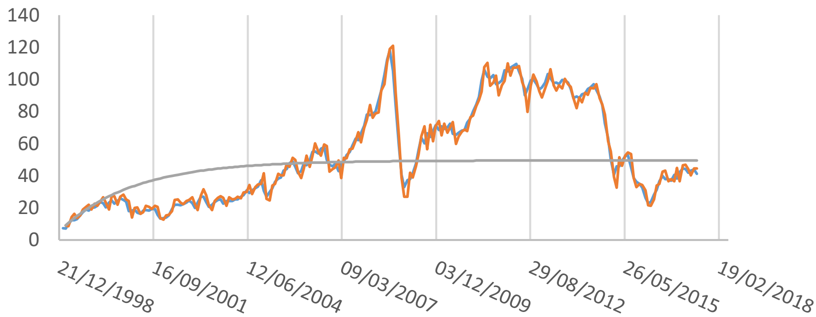

Figure 3. Dynamics of Maya crude oil prices in Mexico between January 1999 and June 2017.
Source: created by the authors.

The best-fit model for the period that includes 94.74 % of the sample

For the shorter period of 216 months between January 1999 and December 2016, the model (15) presents a better fit ($R^2 = 97.96$) than (13) because its test statistics do not reject the null hypothesis of homoscedasticity ($TR^2_{errors} = 3.376541 < \chi^2 = 5.99$; with $T = 216$) or of no serial autocorrelation (DW=1.990609).

$$y_{t+1,Maya} = 1.817518\, y_{t;Maya}^{-1} + 0.991987 y_{t;Maya} + u_t \tag{15}$$

Moreover, the point ($P_{0.99}(= 0.991987 \leq 0.998) = 0.05$) and interval estimation for the coefficients of the autoregressive parameters are unbiased concerning the median at different confidence levels (see Table 5). The above thus confirms that the process is not a unit root process.

Table 5
Interval of parameter values of (15) with different confidence levels

Coefficient	90%		95%		99%	
	minimum	maximum	minimum	maximum	minimum	maximum
1.817518	0.441626	3.193410	0.175855	3.459182	-0.347088	3.982124
0.991987	0.983900	1.000073	0.982338	1.001635	0.979265	1.004708

Source: created by the authors

From the above, the autoregressive process converges to the long-run mean $\widehat{b_x} = 38.7484784$ with an adjustment speed of $\widehat{a_x} = 0.016026$ through the trajectory of $\hat{P}(x,t) = \left(\frac{\widehat{y_x}}{2}\right)^2$ as illustrated in

Table 4. This result means that the IRF tends to disappear in a finite period provided that $\lim_{\tau \to \infty} \alpha^{\tau} = 0$, and $\alpha = 0.98103$. Specifically, the trend tends to move 123.41 units after price shocks (per equation 11) and then recover half of its original level in 86.08 months (equation 12).

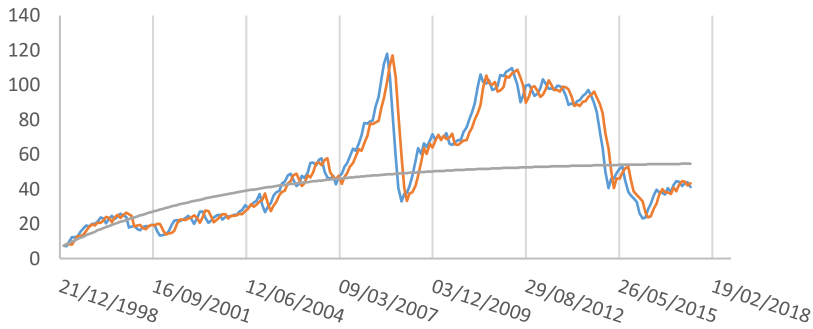

Figure 4. Dynamics of Maya crude oil prices in Mexico from January 1999 to December 2016.
Source: created by the authors.

Forecast

The price behavior generated by the mean reversion trend displays an unfinished process in Figures 3 and 4. The average of their rational prices is a limit to which the real price series converges if the pattern described by the data between January 1999 and June 2017 holds. It is not a given price to be fulfilled in the sample period (as there are still shocks not yet absorbed by the trend). For that reason, it is explainable that the figures reveal price levels above their averages on recent dates, as is clear in the following forecasts for six (July-December 2017) and twelve-month (January-December 2017) windows in which the price hovers around a value above $50 per barrel.

Tables 6 and 7 present the observed and predicted prices for these windows with different sample proportions (97.37% and 94.74%). Both tables indicate that the differences between the two series decrease significantly at the end of the forecast period. Particularly noteworthy is that specification (14) illustrates a more accelerated convergence than (15), as demonstrated by the lower values of MSE, RMSE, MAD, and MAPE in Table 8 for the forecast that includes 97.37% of the sample.

Hence, the more recent data the AR model incorporates or, in other words, the more long-term shocks the trend absorbs, the more accurate the forecast and the lower the variability between predicted and observed data. This study obviates the widespread practice of training the models with smaller sample percentages than those suggested here and evaluates the in-sample and out-of-sample forecast with the remaining proportions. The procedure used to evaluate model forecasts adjusts the percentages to the sample size and the nature of the information processed and takes the error metrics as an overall benchmark measure of model performance.

Table 6
Forecast of crude oil prices using the mean-reversion model (14)

		Maya	
Sample percentage	2017	Observed	Forecast
97.37%	July	43.877714	54.659868
	August	45.554842	54.680660
	September	48.162395	54.701115
	October	48.891212	54.721238
	November	52.824919	54.741036
	December	53.871941	54.760513

Source: created by the authors

Table 7
Forecasting oil prices using the mean-reversion model (15)

Sample percentage	Maya					
	2017	Observed	Forecast	2017	Observed	Forecast
94.74%	January	44.500073	55.213072	July	43.877714	55.350944
	February	44.174861	55.236989	August	45.554842	55.372651
	March	41.932469	55.260522	September	48.162395	55.394010
	April	43.221241	55.283678	October	48.891212	55.415026
	May	43.850742	55.306463	November	52.824919	55.435706
	June	41.151130	55.328883	December	53.871941	55.456054

Source: created by the authors

Table 8
Evaluation criteria for the robustness of the two mean reversion models

Sample percentage	Evaluation criteria					
	Period (2017)		MSE	RMSE	MAD	MAPE
97.37%	From July to December	Maya	39.4920	6.2843	5.1911	11.07%
94.74%	From January to December	Maya	102.0606	10.1025	9.3367	21.11%

Source: created by the authors

Conclusions

This study provides evidence of a reversion pattern to the long-run mean in the price series of Maya crude oil exported by Mexico to the Americas between January 1999 and June 2017. In this period, the real prices of this variety have predictable patterns in the short term and, therefore, do not follow the behavior of a random walk with independent increases as assumed by the EMH.

The two AR(I) models were developed to display a long-term monthly average of 37.8 (with 97.37% of the sample) and 38.7 (94.7% of the sample) dollars per barrel with different convergence speeds (lower in the second case). In a price shock, the above means that the FIRs of their autoregressive parameters will tend toward zero but with different speeds. In the model with the higher sample proportion, the return to its original level will take less than half the time of the one with the lower sample proportion (36.18 versus 86.08 months) because the unit shifts caused by the shocks are longer lasting in the latter than in the former (123.4 versus 52.71). Because of these differences, the predictions in both cases also diverge. However, although the two models give very acceptable results between observed and predicted prices, the one that includes a larger sample proportion presents smaller forecast errors.

In any case, the conclusion does not change. Real prices exhibit a pattern of mean reversion. It remains a pending task for further studies to determine the factors that make it possible to characterize the series forecast in the short term. In this task, the periodization suggested in Figure 1 and the determinants outlined for each sub-period must be considered.

References

Andrews, D. (1993). Exactly median-unbiased estimation of first-order autoregressive/Unit root models. Econometrica, Vol. 61 (1), 139-165. https://doi.org/10.2307/2951781

Baumeister, C., y Kilian, L. (2016). Forty years of oil price fluctuations: Why the oil price may still surprise us. The Journal of Economic Perspectives, Vol 30 (1), 139-160. HTTP:// dx.doi.org/10.2139/ssrn.2714319

Bermúdez, N. Bermúdez, E, y Saucedo, E. (2018). The relationship between oil prices, the stock market and exchange rate: Evidence from Mexico. North American Journal of Economics and Finance, Vol.45, 266-275. https://doi.org/10.1016/j.najef.2018.03.006

Cashin, P., Liang, H. y McDermott, J. (2000). How persistent are shocks to world commodity prices? IMF Staff Papers, Vol.47 (2), 177-217. http://dx.doi. org/10.5089/9781451850284.001

Cruz A. F. (2007). Valor en riesgo de bonos cupón cero en el mercado mexicano con los modelos Vasicek y CIR: Simulación Monte Carlo con saltos de Poisson y valores extremos, Tesis de Doctorado, ITESM.

Choi, K., y Hammoudeh, S. (2009). Long memory in oil and refined products markets. The Energy Journal, Vol. 30 (2), 97-116. http://dx.doi.org/10.5547/ISSN0195-6574-EJVol30-No2-5

Cox, J., Ingersoll, E. y Ross, S. (1985). A Theory of the Term Structure of Interest Rates. Econometrica, 53(2), 385-408. https://doi.org/10.1142/9789812701022_0005

Hamilton, J. (2009). Understanding oil prices. The Energy Journal, Vol. 30 (2), 179-206. https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No2-9

Hammoudeh, S. (1996). Mean reversion and zone readjustments. Southern Economic Journal, Vol. 62 (4), 916-929. https://doi.org/10.2307/1060937

Hansen, P. (2005). A test for superior predictive ability. Journal of Business & Economic Statistics, 23, 365-380. http://dx.doi.org/10.1198/073500105000000063

Hansen, P. R. y Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH (1,1)?. Journal of Applied Econometrics, 20, 873-889. http://dx.doi.org/10.1002/jae.800

http://www.pemex.com/en/about-pemex/reports-and-publications/Paginas/default.aspx

http://ebdi.pemex.com/bdi/bdiController.do?action=cuadro&subAction=applyOptions

Li, J. y Thompson, H. (2010). A note on the oil price trend and Garch shocks. The Energy Journal, Vol.31 (3), 159-165: http://mpra.ub.uni-muenchen.de/20654/

Safari, A., y Davallou, M. (2018). Oil price forecasting using a hybrid model. Energy, 148, 49-58. https://doi.org/10.1016/energy.2018.01.007

Salt et al. (2016)- Control Automático, Tiempo Continuo y Tiempo Discreto. Serie de Textos Universitarios. Ed. Reverte, S. A. ISBN: 978-84-291-4753-7.

Shiller, R. (2003). Efficient markets theory to behavioral finance. The Journal of Economic Perspectives, Vol.17 (1), 83-104. https://doi.org/10.1257/089533003321164967

Vasicek, O. A. (1977). An Equilibrium Characterization of the Term Structure. Journal of Financial Economics, 5, 177-188. https://doi.org/10.1016/0304-405X(77)90016-2

Valdés, A., Armenta, L., y Duran, R. (2016). A copula T-Garch approach of conditional dependence between oil price and stock market index: The case of Mexico. Estudios Económicos, Vol. 31 (1), 47-63. Disponible en: https://estudioseconomicos.colmex.mx/archivo/EstudiosEconomicos2016/47-63.pdf

Venegas M. F. (2008). Riesgos Financieros y Económicos. Cengage Learning. Segunda edición. ISBN-13: 9789708300087; eBook: 9786074813692