

00321



# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

13

Facultad de Ciencias

REGRESION DINAMICA Y LA RELACION ENTRE  
PRECIOS DE CRUDOS LIGEROS Y PESADOS

T E S I S

Que para obtener el título de

A C T U A R I A

presenta:

LETICIA EUGENIA CORONA CALLEJAS



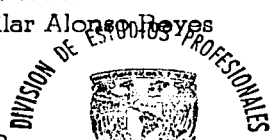
FACULTAD DE CIENCIAS  
UNAM

Director de tesis:

Dr. José Antonio Rojas Nieto

Asesora de tesis:

M. en A. P. Ma. del Pilar Alonso Reyes



2003

FACULTAD DE CIENCIAS  
SECCION ESCOLAR

TESIS CON  
FALLA DE ORIGEN

A



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# **PAGINACION DISCONTINUA**



**DRA. MARÍA DE LOURDES ESTEVA PERALTA**  
**Jefa de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

Regresión dinámica y la relación entre precios de crudos ligeros y pesados.

realizado por Leticia Eugenia Corona Callejas

con número de cuenta 09531553-3 , quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

**A t e n t a m e n t e**

Director de Tesis Propietario Dr. José Antonio Rojas Nieto

Asesor de Tesis Propietario M. en A. P. Ma. del Pilar Alonso Reyes

Propietario M. en C. José Antonio Flores Díaz

Suplente Act. Jaime Vázquez Alamilla

Suplente Act. Gerardo Chávez Heredia

*J. Antonio Rojas Nieto*  
*Del Pilar Alonso Reyes*  
*José Antonio Flores Díaz*  
*Jaime Vázquez Alamilla*  
*Gerardo Chávez Heredia*

**Consejo Departamental de Matemáticas**

*José Antonio Flores Díaz*  
M. en C. José Antonio Flores Díaz

B

*Ahora que he terminado esta etapa tan importante en mi vida, me gustaría hacerle saber a todos lo mucho que agradezco lo que han hecho por mí, sin embargo una hoja nunca es suficiente para expresarlo todo.*

*Este trabajo representa más que un simple requisito para titularme, es el esfuerzo conjunto de muchas personas plasmado en un simple libro, que si bien es cierto que yo lo realicé, también lleva en él una parte de cada una de ellas.*

*A lo largo de todo este trayecto he tenido a mi lado muchas personas, algunas han actuado como impulsos permanentes, han estado conmigo en todo momento alentándome a seguir aun a pesar de las dificultades, sabiendo escuchar cuando lo necesité, dando un ejemplo inigualable o un simple consejo, haciendo amenos los momentos difíciles con su manera tan peculiar de ver las cosas, confiando en mí a pesar de todos los errores cometidos y es definitivamente a ellos, a quienes les debo el mayor de los agradecimientos.*

*Sin embargo, no puedo dejar de agradecer a todas las personas que de un modo u otro han tocado mi vida y han dejado huellas importantes en ella, a los que han sabido ayudarme para llegar a mi destino siguiendo el camino correcto aun cuando éste no sea siempre el más corto, a quienes sin darse cuenta me han impuesto retos, a aquéllos que han contribuido en este logro y a los que indiscutiblemente siempre han estado ahí. A todos un agradecimiento desde el fondo de mi corazón, espero en verdad no haberles fallado, ni fallarles nunca.*

*Y ahora es momento de continuar el camino, porque esa niña que un día llegó a la escuela con carita de ratoncito asustado sabe que aún hay muchos retos esperándole, y a pesar de que algunas veces los nervios se interpongan, es necesario obtener fuerzas y afrontar todas las situaciones de la vida.*

*Sinceramente, muchas gracias.*

*Con todo cariño  
Lety*

*C*

# ÍNDICE

INTRODUCCIÓN.....	1
Capítulo 1	
CONCEPTOS BÁSICOS DEL PETRÓLEO .....	4
1.1 Composición del petróleo.....	6
1.2 Tipos de Petróleo .....	7
1.3 El petróleo en México.....	8
1.4 Características básicas de los principales crudos mexicanos.....	9
1.5 Proceso de Refinación .....	10
1.6 Mercado petrolero .....	11
Capítulo 2	
OBTENCIÓN Y DESCRIPCIÓN ESTADÍSTICA DE LOS DATOS .....	16
2.1 Caracterización del problema.....	16
2.2 Números Índices y deflación de las series.....	18
2.3 Análisis Descriptivo de los Datos.....	19
2.4 Series de Tiempo .....	28
2.4.1 Tendencia, Estacionalidad y Ciclos .....	27
2.4.2 Estacionariedad.....	30
2.5 Análisis de Regresión para el caso específico de los crudos.....	35
2.5.1 Cambio estructural en los datos .....	37
Capítulo 3	
REGRESIÓN DINÁMICA.....	43
3.1 Descripción de las funciones de transferencia.....	43
3.2 Parametrización de la función de transferencia.....	45
3.3 Parametrización del proceso de inercia.....	48
3.4 Generalización a más de una variable explicativa.....	49
3.5 Función de correlación cruzada.....	50
3.6 Identificación de la Función de Transferencia.....	52
3.7 Identificación del Proceso de Inercia .....	54

3.8 Pruebas de Diagnóstico .....	54
Capítulo 4	
APLICACIÓN AL CASO DE LOS CRUDOS .....	56
4.1 Especificación del modelo .....	56
4.2 Análisis de las series explicativas.....	59
4.3 Identificación de las funciones de transferencia .....	62
CONCLUSIONES.....	80
Anexos	
A1. Precios de los crudos Maya y WTI .....	84
A2. Análisis de Regresión.....	88
A2.1 Coeficiente de determinación .....	89
A2.2 Supuestos sobre los residuales.....	90
A2.3 Heteroscedasticidad.....	91
A2.4 Autocorrelación.....	93
A2.5 Multicolinealidad.....	95
A3. Modelos AR, MA, ARMA y ARIMA .....	98
A3.1 Modelo Autorregresivo (AR).....	98
A3.2 Modelos de media móvil.....	100
A3.3 Modelo Autorregresivo-Media Móvil (ARMA).....	102
A3.4 Modelo Autorregresivo-Integrado-Media Móvil (ARIMA).....	104
BIBLIOGRAFÍA.....	105
Índice de Tablas .....	106
Índice de Gráficos.....	107

E

TESIS CON  
FALLA DE ORIGEN

## INTRODUCCIÓN

A fines del siglo XIX se descubrió que la mezcla de hidrocarburos denominada petróleo podía sustituir al carbón en los procesos de combustión. Desde entonces, la comercialización del petróleo cobró gran relevancia. El carácter no renovable del petróleo, su limitada y relativamente concentrada disponibilidad, y su utilización cada vez mayor para algunos procesos fundamentales de la vida (fuerza electro-motriz, movimiento y transporte) hacen del mercado petrolero un espacio significativo desde varios puntos de vista: económico, social y político.

El anhelo permanente de control de esta materia prima estratégica (38 por ciento de las necesidades de energía en el mundo se satisfacen con petróleo) por parte de los grandes consumidores de crudo, ha generado una dinámica de conflictos prácticamente permanentes. Así se explican, por ejemplo, las pugnas entre los tres grandes bloques de consumidores (Estados Unidos, Europa y Asia-Pacífico) y los tres grandes bloques de productores (Medio Oriente, América Latina y Mar del Norte y Mar Caspio).

Parte de la pugna entre estos bloques se origina en la existencia de diferentes calidades de crudos. Con base en esta diferencia se tiene una gama variada de posibilidades para los procesos de refinación y consumo. La refinación del petróleo caracterizado como ligero, produce un porcentaje mayor de los petrolíferos llamados también ligeros, como gasolinas utilizadas para el transporte. Estas gasolinas son las de mayor precio en el mercado de combustibles. Y la refinación del petróleo caracterizado como pesado produce un porcentaje mayor de petrolíferos llamados también pesados, como residuales utilizados para generación de electricidad y transportación marítima. Estos pesados, a su vez, son los de menor precio en el mercado de combustibles.

Un problema muy interesante del mercado petrolero es, precisamente, la evolución de los precios de crudos ligeros y crudos pesados, así como de su relación. Este trabajo de tesis se orienta al estudio de esa relación. Su finalidad es construir un modelo que permita estimar la evolución de la relación que guardan los precios de los crudos ligeros con los precios de los crudos pesados.

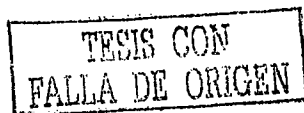


A decir de los especialistas petroleros, esta relación de precios experimenta un comportamiento cíclico. Hay periodos en los que el precio de los crudos pesados tiende a acercarse al precio de los crudos ligeros. Pero, asimismo, hay otros en los que estos tienden a distanciarse. Estas son, precisamente, características propias de un ciclo que, en este caso, se explica por la evolución de la demanda de los crudos pesados. Una reducción o un incremento de la relación entre los precios de los crudos pesados y de los crudos ligeros, sería expresión de una elevación o una reducción relativa de la demanda de crudos pesados respecto a la demanda de crudos ligeros.

Para construir un modelo de la relación que guardan los precios de estos dos tipos de crudo —pesados y ligeros— es necesario caracterizar su evolución histórica. Lo primero a considerar es su alta volatilidad. Difícilmente el precio del crudo coincide con lo que en economía clásica se llama su *precio natural*, es decir, el que permite al productor de mayor costo de producción (marginal), cuyo producto es requerido por el mercado, obtener su ganancia deseada. Normalmente el precio de mercado oscila en torno a ese nivel. Diversos factores lo explican: cambios en la demanda por alteraciones en la dinámica económica; movimientos estacionales; disponibilidad en almacenes e inventarios; presiones especulativas; o amenazas al suministro por guerras, como acontece en estos primeros meses del año 2003.

El objetivo de esta tesis es caracterizar y construir un modelo que explique el comportamiento histórico real de la relación entre las cotizaciones de mercado de los crudos pesados y los crudos ligeros. Esto supone, en primer término, identificar los periodos en los que las cotizaciones se acercan o se alejan. Pero también, de ser posible, tratar de explicar las razones de ello, vinculando el comportamiento con algunas otras variables e indicadores. La alta volatilidad de precios sugiere la conveniencia de hacer pruebas para identificar la existencia o no de momentos de cambio estructural en el comportamiento de los precios.

Así pues, el análisis de la relación de los precios, requiere la consideración de diferentes factores como lo son: la volatilidad de los precios, su tendencia histórica y cambios estructurales; además de considerar los factores de tipo económico, social o militar, que pudiesen haber afectado a los precios.



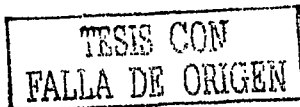
La **primera parte** de esta tesis describe de manera general los conceptos básicos referentes al petróleo, así como las características de su mercado y de su industria. Con ello se busca comprender más adecuadamente el entorno económico, social, y cultural en el que se está trabajando.

En la **segunda parte** se presenta el estudio específico de las series de precios de los crudos considerados en el análisis. Se realiza un estudio descriptivo de los datos y se analizan los componentes –tendencia, estacionalidad y ciclicidad– para el caso específico de la relación entre los precios de los crudos ligeros y los crudos pesados. Sin embargo dado que el principal objetivo de la tesis es el estudio y la caracterización del ciclo que experimenta dicha relación de precios, el análisis más detallado de este ciclo se presenta en una sección posterior.

La **tercera parte** contiene el sustento teórico de la tesis. En ésta se desarrolla de manera general la teoría de la regresión dinámica. Se establecen los principios en los que se fundamenta y los métodos de estimación e identificación necesarios para su desarrollo. Con relación a este capítulo, se incluyen dos anexos de gran importancia: uno referente al análisis de regresión lineal simple, y otro relativo a los procesos autorregresivos o modelos univariados. Ambos anexos contribuyen a la presentación del desarrollo teórico de la regresión dinámica. La unión de ambas metodologías –regresión simple y autorregresión– fundamenta la técnica de la regresión dinámica.

En la **cuarta parte** de este trabajo se aplica la teoría de la regresión dinámica en el caso de la relación de precios de crudos ligeros y pesados. Es aquí donde se estudia con más detenimiento la presencia de los ciclos. Y se analiza cada una de las variables a incluir en el modelo que se construye, para garantizar el cumplimiento de los supuestos teóricos de la regresión dinámica. Finalmente se indican los problemas que se presentan al tratar de aplicar esta metodología.

En la **última parte** se explican las conclusiones a las que se llega después de realizado el análisis.



# Capítulo 1

## CONCEPTOS BÁSICOS DEL PETRÓLEO

Desde fines del siglo pasado en que comenzó la utilización comercial del petróleo, su comercialización adquirió una importancia creciente. Actualmente el mercado petrolero es uno de los más importantes del mundo. Sin petróleo el mundo prácticamente no se moviliza, ni tiene calor para muchos procesos, ni logra generar electricidad, ni tiene calefacción o aire acondicionado, entre otras cosas. En 1999 la comercialización internacional del petróleo representó aproximadamente el 7% de las transacciones internacionales, con un volumen diario cercano a los 40 millones de barriles<sup>1</sup>, de un total consumido próximo a los 75 millones. En ese mismo año la industria de los alimentos representó el 8% del valor de las mismas transacciones. Se trata de dos de los ámbitos en los que las transacciones internacionales resultan fundamentales. México participó en esa comercialización mundial de petróleo con un millón 553 mil barriles diarios con un valor de 8 mil 859 millones de dólares. Un año después, por cierto, con un volumen que sólo creció en 100 mil barriles al día, los ingresos por este concepto alcanzaron un valor de 14 mil 887 millones de dólares. En el 2002, con exportaciones diarias de un millón 864 mil barriles, se lograron ingresos por 13 mil 087 millones de dólares.

Si esto no fuera suficiente, hay otros factores que justifican la importancia del petróleo. A diferencia de los alimentos, el petróleo no es un producto que cualquier país pueda exportar pues no dispone de él, la existencia de reservas localizadas en regiones específicas, hace que el petróleo sea altamente demandado por países con escasas o nulas reservas petrolíferas, como es el caso de Japón. En el medio oriente en las regiones de Arabia Saudita e Irak se concentran las reservas más importantes, esto hace de esa región una de las más codiciadas por los países desarrollados. En lo que se refiere a México, éste está considerado dentro de los principales países que poseen reservas petroleras, ocupó en el 2001 el noveno lugar en reservas probadas de petróleo crudo, las cuales se calcularon en 26,941<sup>2</sup> millones de barriles y produjo en ese mismo año 3

<sup>1</sup> 1 barril = 159 litros = 0.159 m<sup>3</sup>

<sup>2</sup> Fuente: Oil and Gas Journal y PEMEX

millones 122 mil barriles diarios ocupando el octavo lugar a nivel mundial. De la producción total de petróleo, el 83.86% fue producción de pesados, es decir, se produjeron 1 millón 997 mil barriles al día de petróleo pesado.

Más aún, cuando un país posee importantes reservas pero no cuenta con la infraestructura necesaria para su explotación o producción entonces es necesario recurrir a países que ofrezcan la posibilidad de extraer el crudo, así el petróleo es comercializado mundialmente de manera significativa dando a este recurso natural una importancia tanto económica como política.

En el caso particular de México, éste cuenta con seis refinерías, las cuales tienen una capacidad de refinación de aproximadamente el 80% de la correspondiente a Canadá o a Australasia, los cuales cuentan con 21 y 19 refinерías respectivamente. Por tanto es posible decir que la capacidad de refinación en México es importante considerando el número de refinерías con las que cuenta. México tuvo una capacidad de refinación de aproximadamente 1 millón 559 mil barriles por día en el 2000.

Se puede entonces concluir que el petróleo es uno de los más importantes productos que se negocian en el mercado mundial de materias primas. Las bolsas de Nueva York (NYMEX) y de Londres (IPC) son los principales centros donde se comercia, pero también tiene un mercado "spot" o al momento. Los precios se regulan por unos marcadores o "precios de referencia", entre los que sobresalen el West Texas Intermediate (WTI), el Brent y el Dubai.

México como uno de los principales países en la industria petrolera, no está exento de los cambios políticos y económicos que se generan en torno a esta industria, aun cuando no forma parte de la Organización de Países Exportadores de Petróleo (OPEP). Es por ello que es importante comparar el mercado petrolero mexicano con el internacional, y una manera de hacerlo es estudiando la relación entre los precios de crudo mexicano y los crudos extranjeros. Una manera de comparar es estudiar la manera en que afectan las variaciones en el precio de los crudos internacionales considerados como crudos marcadores en el comportamiento que tienen los precios de los crudos nacionales.

Estados Unidos de América es el principal consumidor de petróleo a nivel mundial, sin embargo no cuenta con una gran cantidad de reservas, en el 2001 sus reservas se estimaron en 21 mil 756 millones de barriles; sin embargo, es considerado como el principal país en cuanto a capacidad de refinación se refiere, en el año 2000 su capacidad de destilación primaria se calculó en 16 millones 539 mil barriles diarios, mientras que para México en ese mismo año su capacidad se estimó en solamente 1 millón 559 mil barriles diarios. Este hecho repercute evidentemente en las exportaciones que México realiza a los EUA, una muy importante proporción de las exportaciones que México realiza las hace a dicho país, de hecho es posible considerar a México como el único exportador de crudo pesado para los Estados Unidos, por lo cual es razonable sugerir una importante dependencia del precio del petróleo mexicano del petróleo estadounidense.

Para comprender mejor el entorno petrolero es necesario conocer los conceptos básicos relacionados con el tema, por lo que, el primer paso será estudiarlos de manera general. El petróleo crudo es una mezcla en la que coexisten en fases sólida, líquida y gaseosa, compuestos denominados hidrocarburos, los cuales son cadenas largas de átomos constituidas por carbono e hidrógeno y pequeñas proporciones de heterocompuestos con presencia de nitrógeno, azufre, oxígeno y algunos metales, ocurriendo en forma natural en depósitos de roca sedimentaria, como níquel, vanadio, fierro y cadmio. Su color varía entre ámbar y negro. La palabra petróleo significa aceite de piedra.

Los hidrocarburos en el petróleo crudo tienen más de cuatro átomos de carbono y están divididos en cuatro clases, parafinas (alcanos), naftalenos (cicloalcanos), aromáticos (con estructuras de benceno) y asfálticos.

## 1.1 Composición del petróleo

Dependiendo del número de átomos de carbono y de la estructura de los hidrocarburos que integran el petróleo, se tienen diferentes propiedades que los

caracterizan y que determinan su comportamiento como combustibles, lubricantes, ceras o solventes.

Además, hay hidrocarburos con presencia de azufre, nitrógeno y oxígeno formando familias bien caracterizadas, y un contenido menor de otros elementos. Al aumentar el peso molecular de los hidrocarburos las estructuras se hacen verdaderamente complejas y difíciles de identificar químicamente con precisión. Un ejemplo está dado por los asfaltenos que forman parte del residuo de la destilación al vacío.

## 1.2 Tipos de Petróleo

Son miles los compuestos químicos que constituyen el petróleo, entre muchas otras propiedades, estos compuestos se diferencian por su volatilidad, que no es más que la tendencia del petróleo a vaporizarse (dependiendo de la temperatura de ebullición). Al calentarse el petróleo, se evaporan preferentemente los compuestos ligeros (de estructura química sencilla y bajo peso molecular), de tal manera que conforme aumenta la temperatura, los componentes más pesados van incorporándose al vapor.

Las curvas de destilación TBP (del inglés "true boiling point", temperatura de ebullición real) distinguen a los diferentes tipos de petróleo y definen los rendimientos que se pueden obtener de los productos por separación directa. Por ejemplo, mientras que en el crudo Istmo se obtiene un rendimiento directo de 26% volumétrico de gasolina, en el Maya sólo se obtiene 15.7%.

La industria mundial de hidrocarburos líquidos clasifica al petróleo de acuerdo a su densidad medida utilizando el parámetro internacional del Instituto Americano del Petróleo, este parámetro se denota de acuerdo a las siglas en inglés de este instituto, es decir, densidad API (American Petroleum Institute). Los grados API sirven para medir la calidad del petróleo crudo. Esta unidad de medida está relacionada inversamente con la densidad del crudo, es decir, entre mayor sea el valor del crudo en grados API, el crudo es menos denso y por tanto de mejor calidad.



Tabla 1.1 Tipos de Petróleo<sup>3</sup>

Crudo	Densidad (g/ cm <sup>3</sup> )	Densidad/ grados API
Extra pesado	>1.0	10.0
Pesado	1.0 - 0.92	10.0 – 22.3
Mediano	0.92 - 0.87	22.3 – 31.1
Ligero	0.87 - 0.83	31.1 - 39
Super ligero	< 0.83	> 39

### 1.3 El petróleo en México

Actualmente, México produce crudos de cuatro diferentes calidades: el crudo *Maya*, considerado como un crudo pesado, el cual aporta más de la mitad de la producción nacional total; el crudo *Istmo*, un crudo ligero con bajo contenido de azufre que proporciona una tercera parte de la producción total; otro crudo es el *Olmeca*, el cual es de todos el de más alta calidad al ser considerado como super ligero y constituye una quinta parte de la producción total; y finalmente el pesado de Altamira un crudo de menor importancia, debido a su calidad. Alrededor de tres cuartas partes de la producción mexicana proviene de la costa de Campeche en el Golfo de México. El Maya es el crudo mexicano más importante en la exportación.

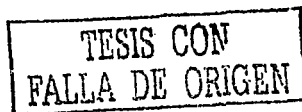
Como una combinación de los tres principales crudos de México, el Olmeca, el Istmo y el Maya, surge lo que se conoce como la mezcla mexicana. En la siguiente tabla se muestran las características principales de los crudos que conforman la mezcla mexicana:

Tabla 1.2 Crudos Mexicanos<sup>4</sup>

CRUDO	CALIDAD	GRADOS API	% DE AZUFRE	% DE LA PRODUCCION TOTAL (2002)
Maya	Pesado	22	3.3	66.2
Istmo	Ligero	33.6	1.3	17.4
Olmeca	Super ligero	39.3	0.8	14.4
Pesado de Altamira	Extra pesado			

<sup>3</sup> Fuente: Refining Industry of Mexico. Science and Technology , RICHMAC Magazine, Mayo 2002

<sup>4</sup> Mexico and the Petroleum Market, by Bruce Currie-Alder, Febrero 1998



El petróleo mexicano es materia prima de calidad que se encuentra presente en toda la industria nacional e internacional como lo es en: transporte, alimentos, fármacos, fertilizantes, pinturas, textiles, etc. Petróleos Mexicanos, en 1999, se colocó en el tercer lugar dentro de las empresas petroleras con mayor nivel de producción, en ese año produjo alrededor de 3343 miles de barriles diarios.

Aún cuando México es uno de los productores de petróleo más importantes en el mundo, cerca de un cuarto de la gasolina consumida es de origen extranjero, debido principalmente a la insuficiente capacidad de las refinerías. México cuenta con seis refinerías, con una capacidad de 1.5 millones de barriles por día; el gobierno planea un aumento en la capacidad de las refinerías de 350,000 barriles al día para el 2006 por cada una, haciendo un total de 2.1 millones de barriles por día; por otra parte debido al extremadamente limitado fondo que existe para las actividades de exploración de los yacimientos, las reservas de petróleo crudo en México están disminuyendo. Los cálculos hechos para enero del 2002, están por debajo de las hechas para enero del 2001 en aproximadamente 1.4 billones de barriles<sup>5</sup>.

#### 1.4 Características básicas de los principales crudos mexicanos

El crudo Maya se caracteriza por su alta viscosidad y contenido de azufre, metales y asfaltenos, además de un bajo rendimiento de fracciones ligeras en la destilación. Los retos derivados de la refinación de mezclas de crudos con alta proporción de Maya se hacen más complejos debido a la creciente demanda de combustibles ligeros (gasolina y diesel) con especificaciones cada vez más estrictas para proteger el ambiente.

En el 2002, este crudo representó el 68.2% de la producción total de hidrocarburos líquidos (2 millones 168 mil barriles al día sobre un total de 3 millones 177 mil). Además el volumen de las exportaciones de este crudo fue, en ese mismo año, de un millón 370 mil barriles por día, equivalentes al 82.3% de las exportaciones totales de petróleo crudo. La proporción de crudo Maya en la mezcla de crudos procesada en las refinerías aumentará

<sup>5</sup> Cálculos obtenidos de la página web de Energy Information Administration [www.eia.doe.gov](http://www.eia.doe.gov)



a 58% después de la realización de los proyectos de reconfiguración en este rubro que iniciaron en 1998.

El crudo Olmeca representa el crudo de mejor calidad que produce México. Es considerado como un crudo superligero con una densidad de 39.3 grados API y un 0.8% de azufre en peso. A pesar de que la calidad de este crudo supera a los otros dos crudos producidos en México, el volumen de las exportaciones de éste representaron en el 2002 un 14.8%, alcanzando en promedio los 264 mil barriles diarios. Es decir, de los 458 mil barriles diarios que se produjeron en México en ese año, poco más de la mitad fueron destinados a las exportaciones.

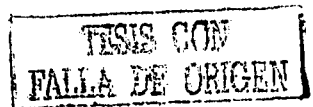
El crudo Istmo tiene una densidad de 33.8 grados API, por lo cual está considerado como un crudo ligero, contiene un 1.3% de azufre en peso. Este crudo ocupó, en el 2002, el segundo lugar en nivel de producción en México, se produjeron en promedio 552 mil barriles diarios (17.4% del total), sin embargo, el volumen de las exportaciones de este crudo es el menor en comparación con los otros dos, en promedio se exportaron 48 mil barriles diarios, que representan el 3% del total de las exportaciones.

## 1.5 Proceso de Refinación

El petróleo crudo al ser refinado, es decir, al descomponer los hidrocarburos en varias fracciones, es convertido en diferentes productos. Los productos finales más comunes pueden consistir en gas natural comprimido, gases de petróleo licuados, butano, gasolina, kerosén, diesel, residuales de crudo, lubricantes, parafinas y ceras.

La destilación es un proceso fundamental en la industria de refinación del petróleo, pues permite hacer una separación de los hidrocarburos aprovechando sus diferentes puntos de ebullición (temperatura a la cual hierve una sustancia). El primer proceso que se realiza en una refinería es la destilación atmosférica y al vacío.

El petróleo se separa en fracciones que después de procesamiento adicional, darán origen a los productos principales que se venden en el mercado: el gas LP



(utilizado en estufas domésticas), gasolina para los automóviles, turbosina para los aviones jet, diesel para los vehículos pesados y combustóleo para el calentamiento en las operaciones industriales.

Las características del crudo, así como la cantidad y la calidad de productos que se desea obtener, determinan los procesos que deben incorporarse a la refinería:

- La mayor parte de los productos obtenidos en el proceso de destilación primaria se someten a hidrot ratamiento para eliminar principalmente azufre y nitrógeno.
- Para la generación de las gasolinas se incorporan procesos como reformación catalítica, síntesis de éteres (MTBE, del inglés Methyl Tertiary-Butyl Ether), alquilación e isomerización de pentanos-hexanos, balanceados de tal forma que la mezcla resultante cumpla con la especificación establecida.
- Los gasóleos de vacío se someten a desintegración catalítica fluida para generar mayor cantidad de destilados ligeros, principalmente gasolina.
- El residuo de vacío puede también someterse a hidrodeseintegración o a coquización para aumentar el rendimiento de destilados, o a procesos de hidrot ratamiento o reducción de viscosidad para generar combustóleo.

Existen muchas operaciones en los procesos de la industria del petróleo basadas en la separación física de componentes aprovechando diversos principios.

## 1.6 Mercado petrolero

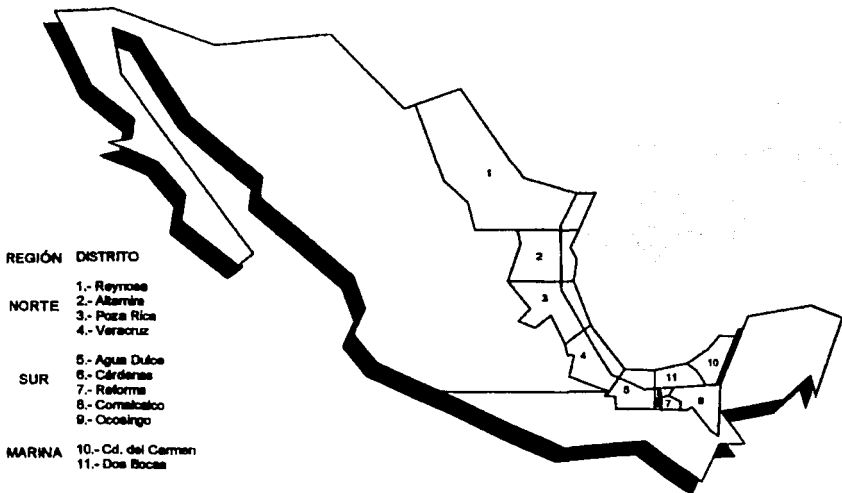
En general, hay tres parámetros claves en el mercado de los crudos: la calidad de éste, la infraestructura de producción y las reservas existentes.

El precio del petróleo crudo depende en parte a su calidad en comparación con una lista de otros crudos estándares, y por otro lado a los cargos por transportarlo a un

determinado centro de comercio. Los volúmenes de crudo son medidos en número de barriles, siendo cada uno de ellos equivalente a 35 galones imperiales o bien a 42 galones estadounidenses, equivalente a 159 litros.

Un segundo parámetro clave en el mercado del petróleo es la infraestructura de producción de un país. Además de la cantidad de petróleo que pueda existir en el subsuelo, se debe tomar en cuenta la capacidad de los yacimientos, oleoductos, tanques y refinerías, los cuales establecen límites para el suministro de petróleo crudo en cualquier momento.

**Gráfico 1.1 Áreas de producción y exploración del petróleo**



En México la producción promedio diaria de petróleo crudo en el 2001, fue de 3 millones 127 mil barriles diarios, de acuerdo a los datos publicados por PEMEX. La producción correspondiente al crudo pesado fue de 1,997 miles de barriles diarios, para el

ligero de 659 miles de barriles diarios y 471 de superligero. En ese mismo año, la región marina aportó el 81.23% (equivalente a 2,540 miles de barriles diarios) de la producción total de crudo, siendo con ello el área principal de producción en el país. La segunda región en importancia es la sur que produjo en promedio 509 mil barriles diarios, correspondientes al 16.28% del total. Finalmente está la región del norte que contribuyó con 78 miles de barriles diarios equivalentes al 2.49 % de la producción total en el 2001

El último parámetro a considerar es la localización de las reservas, que son lugares en el subsuelo donde se ha identificado la existencia de crudo y que deben ser explotados. Es importante notar que las reservas petroleras son independientes de la producción, que está en función de la inversión de capital en la infraestructura. Muchos países poseen reservas de petróleo y por ello producen para sí mismos. Las reservas probadas de petróleo crudo en México, en el 2001, alcanzaron un total de 26 mil 941 millones de barriles.

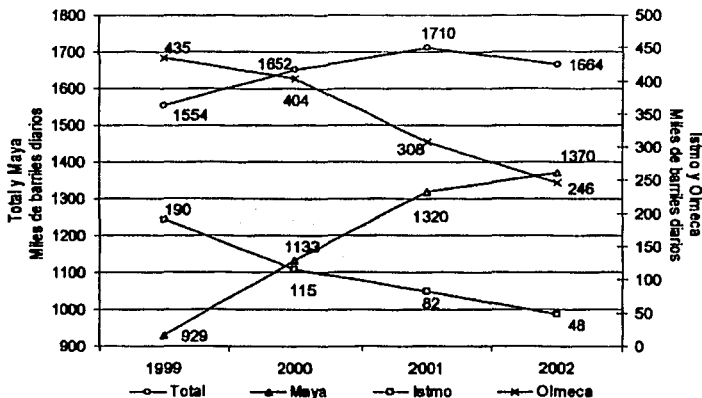
El comercio de petróleo crudo es la transferencia de un excedente inmediato de petróleo de las naciones productoras a las naciones consumidoras, las cuales no cuentan con suficientes recursos propios o bien que no tienen suministros.

En México, el valor de las exportaciones de petróleo crudo se ubicó, en el 2001, en 11,592.2 millones de dólares; referente al 2002 se sabe que hasta abril se ha acumulado un total de 3,597.5 millones de dólares. Por otra parte el volumen promedio de las exportaciones de petróleo crudo se ubicó en 1,710 miles de barriles diarios en el 2001, de los cuales el 77.19% correspondieron a exportaciones de crudo Maya, seguido por el crudo Olmeca con el 18.01%, el resto correspondió a exportaciones del crudo Istmo.

El siguiente gráfico muestra el comportamiento de las exportaciones de petróleo crudo de los últimos años, se consideran las exportaciones totales, pero además se consideran las exportaciones por separado de cada uno de los crudos mexicanos que se exportan. Se observa que el volumen de las exportaciones más significativo está dado por las correspondientes al crudo Maya, las cuales han aumentado de manera significativa desde 1999. En contraste, las exportaciones del crudo ligero, Istmo, y del súper ligero,

Olmecca, han disminuido drásticamente y además se hace notar que están muy por debajo de las correspondientes al Maya.

Gráfico 1.2 Volumen de las exportaciones de petróleo crudo



El destino de las exportaciones de petróleo crudo de México es principalmente a países de América. Del total de éstas en el 2001, se exportaron a países americanos un total de 1,485 miles de barriles diarios (aproximadamente el 86.84%). México ha sido un suministrador estratégico de los Estados Unidos, haciendo notar que es posible considerarlo como su único suministrador de crudos pesados. Además hay otras exportaciones de petróleo crudo a Europa y en menor escala al lejano oriente, ambas con un 13.16%.

En 1993, algunos países desarrollados suministraron el 79% (equivalente a 140 billones de dólares) de los crudos comerciados internacionalmente. Durante ese mismo año los quince países más importantes en la exportación de petróleo fueron, Arabia Saudita, Irán, Noruega, Emiratos Árabes Unidos, Venezuela, Nigeria, Reino Unido, México, Indonesia, Canadá, Algeria, Omán, Malasia y China. Dentro de las importaciones de ese año, el 35% (aproximadamente 60 billones de dólares) fueron realizadas por

TESIS CON  
FALLA DE ORIGEN

países desarrollados, dentro de dichos países estaban los Estados Unidos, Japón, Alemania, Italia, Francia, Corea del Sur, España, Holanda, Brasil, Reino Unido, Singapur, Canadá, India y Turquía.

En el mercado de crudos los participantes se dividen en productores, consumidores y compañías intermediarias. Dentro de los productores existe otra división dependiendo si éstos son miembros de la OPEP (Organización de Países Exportadores de Petróleo), que fue formada en 1960 y que tiene su sede en Viena. La OPEP controla más del 77% de las reservas mundiales, y por ello se asegura que seguirá influyendo de manera importante en el mercado mundial del petróleo.

Desde 1983, el NYMEX (New York Mercantile Exchange) ha sido el centro donde se establecen los precios mundiales del petróleo, y donde se introducen futuros contratos de petróleo como suma de los precios spot. Cuando el NYMEX toma dicho control, el crudo ligero de Arabia se sustituye por el West Texas Intermediate (WTI), logrando con ello que el precio de este crudo se convierta en el precio indicador mundial.

La calidad del WTI varía entre 38 y 40 grados API, que lo caracteriza como ligero, con un contenido de azufre de aproximadamente 0.3%. El cálculo del precio del WTI es para barriles entregados tanto en Cushing, Oklahoma como en Midland, Texas.

México es uno de los países productores que no forman parte de la OPEP, y por ello, no tiene cuotas previamente establecidas, además de que la producción solamente depende de la existencia de infraestructura en el país. La existencia de reservas de poco menos de 49 billones de barriles es un claro indicio de que México continuará produciendo por lo menos otros 25 años.

TESIS CON  
FALLA DE ORIGEN

## Capítulo 2

# OBTENCIÓN Y DESCRIPCIÓN ESTADÍSTICA DE LOS DATOS

### 2.1 Caracterización del problema

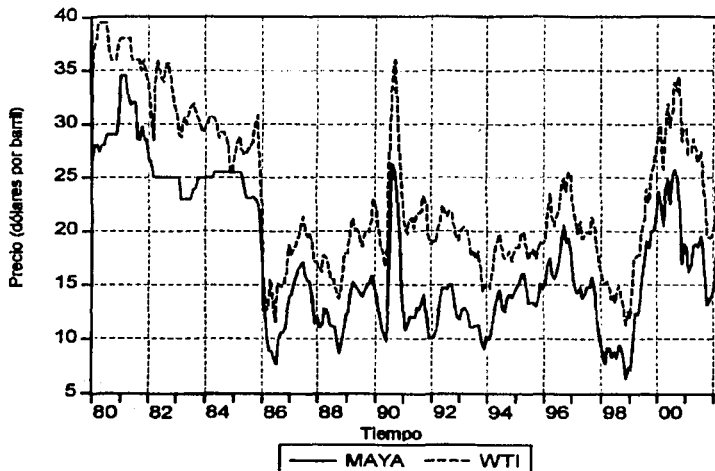
La existencia de diferentes calidades de crudos origina precios diferentes entre ellos; sin embargo, el comportamiento de los precios de los crudos pesados parece estar relacionado con el de los ligeros o viceversa, es decir, existe una relación entre ambos la cual puede deberse a diferentes factores como lo son la demanda de crudo pesado, las importaciones de éste, o el costo de refinación, por decir algunos ejemplos.

Para el análisis de dicha relación se consideran dos series temporales de precios de crudos. Una de ellas es la del crudo pesado Maya y la otra es la del crudo ligero de origen estadounidense, el WTI. Se utiliza el precio del crudo Maya, ya que éste es el crudo mexicano de mayor nivel de producción en el país, cuyo nivel de exportaciones es el más alto en comparación con los otros crudos mexicanos, y además porque es exportado esencialmente a los Estados Unidos. Para el precio del ligero se decidió utilizar el correspondiente al WTI, pues, como se mencionó en el capítulo anterior, este crudo es considerado como el crudo marcador del mercado. Las fuentes de obtención de cada una de las series fueron diferentes, pero en ambos casos se obtuvieron las series en precios corrientes, es decir, los precios a los que se cotizó el crudo en la fecha respectiva.

Para el caso del crudo mexicano (Maya) la fuente de obtención fue directamente PEMEX (Petróleos Mexicanos). Se consideró esta información por ser la más completa, además de que PEMEX tiene los derechos exclusivos para la explotación y producción del petróleo en México; para el WTI, la fuente de obtención fue el Banco de la Reserva Federal de Saint Louis, dicha información es de acceso gratuito vía Internet, lo que facilita la obtención de los datos. Así pues las series resultantes, se muestran en el anexo A1, y la gráfica 2.1 muestra el comportamiento de los datos a través del tiempo.

El periodo de estudio se considera a partir de enero de 1980, dado que el crudo Maya comenzó su comercialización a partir de 1978 y por tanto no existen datos anteriores a esta fecha.

Gráfico 2.1 Datos Maya y WTI (precios corrientes)



Hay que recordar que los datos obtenidos para las series de precios de ambos crudos están en precios corrientes. Esto es una limitante al tratar de compararlos pues es necesario que éstos estén referidos a precios de un mismo año, o de lo contrario el efecto de la inflación estaría presente en ellos, para ello, es necesario convertir las series a precios constantes.

TESIS CON  
FALLA DE ORIGEN



## 2.2 Números Índices y deflación de las series

Parte del análisis de series temporales es la construcción de números índices para reflejar cambios en los precios (o en las cantidades) a lo largo del tiempo. Existen diferentes tipos de índices, diferenciados de acuerdo a la información que utilizan.

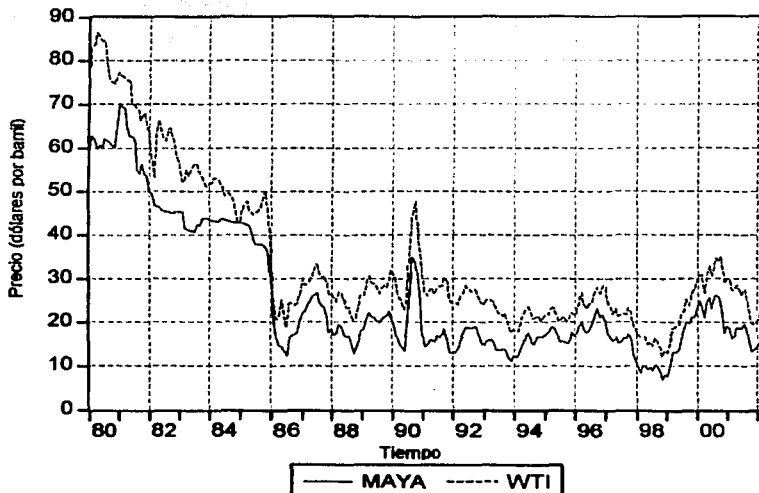
Hay esencialmente dos tipos de índices: **índices de precios** e **índices de cantidades**. Un índice de precios mide el cambio en los precios de un grupo de artículos a lo largo del tiempo; un índice de cantidades mide el cambio en la cuantía de un grupo de bienes o servicios producidos a lo largo del tiempo. Los índices del mercado de valores tales como el promedio industrial *Dow Jones* o el *Índice Nacional de Precios al Consumidor (INPC)* son índices de precios; los índices de producción de automóviles, acero y similares son índices de cantidades.

Dado que ambas series están expresadas en dólares por barril, el índice de precios al consumidor de los Estados Unidos, CPI (del inglés, *Consumers Price Index*) es adecuado para la deflación de las series. Este índice es un conjunto de datos mensuales que reflejan los cambios en los precios pagados por los consumidores urbanos por un conjunto de bienes y servicios contemplados en una canasta básica. Es el equivalente al índice mexicano INPC (Índice Nacional de Precios al Consumidor).

Se consideró como año base el 2001 por la utilidad de tener dimensiones recientes, aunque cabe mencionar que resulta indiferente el año seleccionado como base, porque el comportamiento de las series es el mismo utilizando cualquier año, lo realmente importante es que ambos estén referidos a un mismo año. Las series en precios constantes del 2001 se muestran en el anexo A1 y la gráfica 2.2 muestra el comportamiento de éstas.



Gráfico 2.2 Datos Maya y WTI (precios constantes 2001)



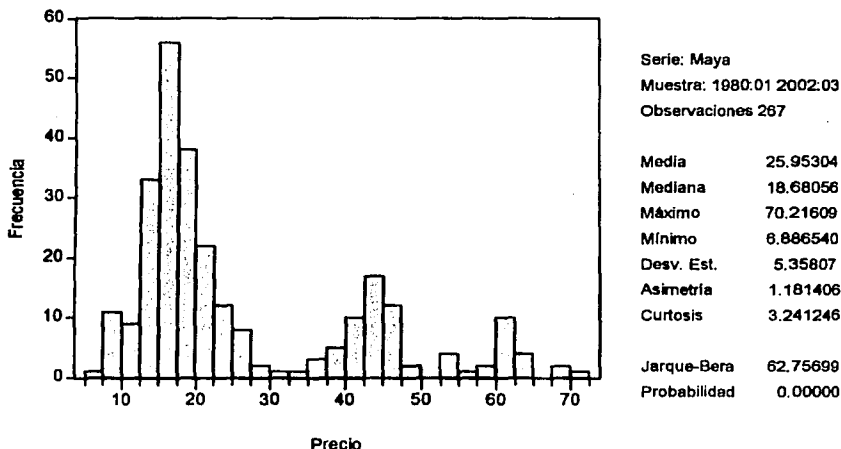
### 2.3 Análisis Descriptivo de los Datos

Un primer paso en el análisis fue elaborar la estadística descriptiva de las series de precios; lo cual permitiría conocer más a fondo el comportamiento de los datos y así poder obtener resultados más apegados a la realidad, en el momento de establecer los modelos estadísticos que los vinculen.

Para los datos del Maya, los resultados descriptivos indican que, durante el periodo de estudio, el precio promedio ha sido de aproximadamente 25.95 dólares del 2001 por barril, con una desviación estándar de poco más de 15 dólares, lo que quiere decir que en promedio ha variado aproximadamente entre los 11 y los 41 dólares, se ha tenido estadísticamente, en este periodo, un valor mínimo de 6.89 dólares y un valor máximo de 70.22 dólares.

En el gráfico 2.3 se muestra la distribución de frecuencias de los precios mensuales del crudo Maya a precios constantes del 2001. Los datos se distribuyen con un sesgo a la izquierda, debido a que los precios observados correspondientes a los meses de los primeros años están muy por arriba de la media muestral, originando pocos precios altos y por tanto menor frecuencia para clases mayores. Si se consideran las observaciones comprendidas entre los precios 10 y 20, se tiene que en este rango se acumula casi el 51% de las observaciones.

Gráfico 2.3 Histograma Maya (dólares 2001)

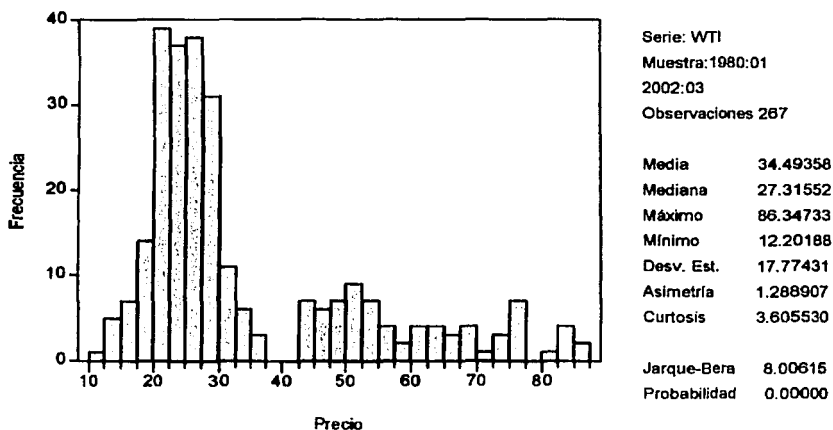


Para el caso del WTI se tiene que éste presentó un precio promedio de 34.49 dólares, y se ha desviado en promedio casi 18 dólares, y por lo tanto su fluctuación promedio ha sido entre los 17 y los 53 dólares, alcanzando estadísticamente un valor mínimo de 12.20 dólares y un máximo de 86.34 dólares.

En el histograma siguiente (gráfico 2.4) se presenta la distribución de frecuencias de los precios mensuales del WTI. Igual que en el caso del Maya, la distribución está

sesgada a la izquierda, sin embargo la mayor acumulación de las observaciones se concentra entre los 20 y los 30 dólares. Igual que como ocurrió con el Maya, si se consideran únicamente las observaciones a partir de enero de 1986, éstas no excederán los 40 dólares y por lo tanto al analizar la distribución de frecuencias se tendría un histograma centrado. Se observa que entre los 15 y los 25 dólares se encuentran 97 observaciones que representan el 36.33% del total, sin embargo al considerar las observaciones entre 25 y 35 dólares, se tienen 86 de ellas, equivalente al 32.21% de las observaciones, con lo cual es posible suponer que estos dos rangos son de igual importancia, y es relevante remarcar que entre ambos involucran la mayoría de los datos, que representan el 68.54% del total.

Gráfico 2.4 Histograma WTI (dólares 2001)



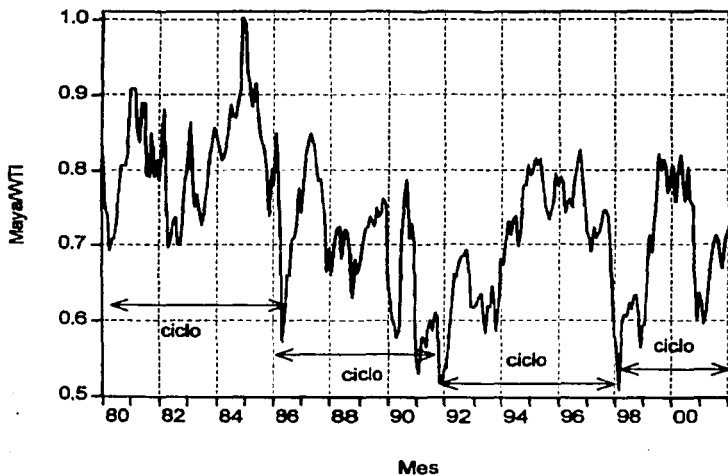
Cabe hacer notar que el precio promedio del crudo Maya está casi 9 dólares por abajo del correspondiente al WTI, pero que ambos presentan desviaciones estándares parecidas, lo que es un indicador primario de que la variabilidad del precio del crudo Maya es un reflejo de lo que sucede con el WTI o viceversa. Por otra parte en los histogramas se puede vislumbrar que sus distribuciones son parecidas, con la salvedad de estar recomidas, ambas están sesgadas a la izquierda, y para precios mayores que 40 la distribución está indefinida.

El objetivo del análisis es estudiar la relación que tienen el crudo Maya y el WTI que permita establecer un modelo o relación funcional que indique el comportamiento de dicha relación ante los posibles cambios que se presenten.

Surge la interrogante de por qué se desea modelar la relación entre los precios de crudos pesados y ligeros, y no directamente el precio de los pesados, si en ambos casos se tomará como conocido el precio del crudo ligero (en este caso WTI). Una explicación a esto puede darse considerando que, lo que se busca no es saber que pasa con el precio de los pesados al cambiar el precio de los ligeros, sino que se desea saber lo que ocurre con la relación entre los precios de los crudos cuando el precio de los ligeros cambia.

Hay dos maneras de estudiar dicha relación, una es considerando la diferencia de las series (Maya - WTI), o bien considerar la relación dada por el cociente (Maya / WTI).

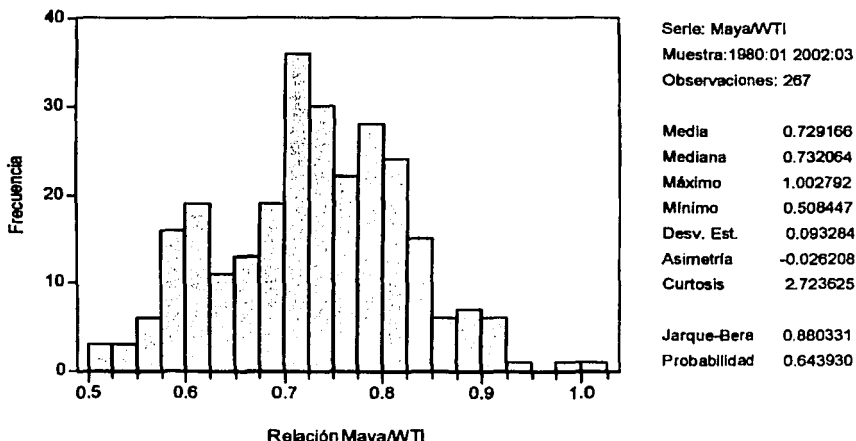
Gráfico 2.5 Relación Maya/WTI



Del estudio descriptivo de la relación dada por el cociente Maya/WTI, lo primero que se debe hacer notar es la aparente presencia de ciclos, es decir, se tienen periodos para los cuales el comportamiento de la serie es creciente y otros en los cuales decrece, este proceso es repetitivo en un cierto número de meses. En el gráfico 2.5 es claro notar el comportamiento descrito anteriormente, las flechas indican de manera intuitiva los ciclos que se presentan, sin embargo, es necesario realizar antes la identificación de patrones estacionales o tendenciales.

Si se realiza para este cociente su correspondiente histograma (gráfico 2.6), se observa como ha cambiado su distribución al involucrar ambos crudos en una relación, que al considerarlos cada uno por separado. Esto se debe al tipo de variabilidad presentado por cada uno de ellos. El histograma es más central y pudiera aproximarse a la distribución normal, con media en 0.72 y una desviación estándar de 0.09.

**Gráfico 2.6 Histograma Maya/WTI**

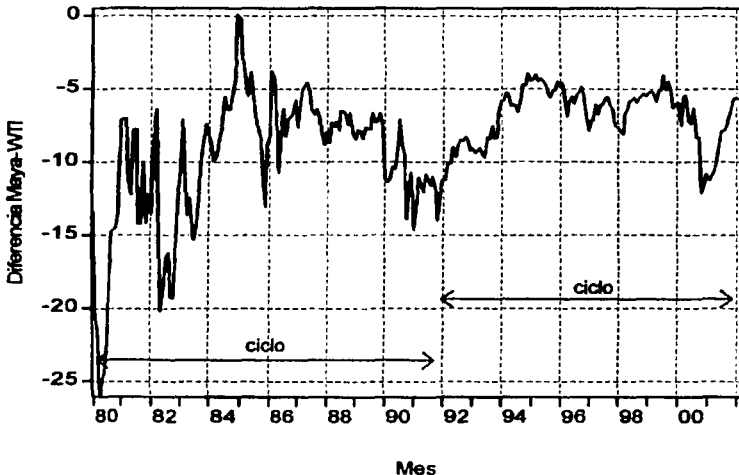


Considerando ahora la diferencia de las series, es decir, la serie obtenida a través de Maya-WTI, se puede observar en la gráfica 2.7 un comportamiento similar al de la

gráfica correspondiente al cociente, es decir, tiene una forma cíclica, aun cuando tal vez resulta menos claro. Como la característica mostrada en ambas relaciones es de una serie periódica, indica que se debe hacer un análisis detallado de esta situación, cosa que se analizará más adelante.

El valor promedio de la diferencia es de 8.54 (considerando el valor absoluto, dado que lo que interesa conocer es la magnitud y no el valor real). Se observa que el precio del WTI ha estado por encima del Maya, la mayoría de las veces, alcanzando la diferencia un valor máximo de 26.50 aproximadamente, sin embargo los precios de ambos crudos se han acercado hasta ser prácticamente iguales teniendo como valor mínimo de la diferencia 0.12 dólares, registrados en diciembre de 1984, mes en el cual el valor del Maya estuvo por encima del WTI, siendo ésta la única vez en el periodo de estudio.

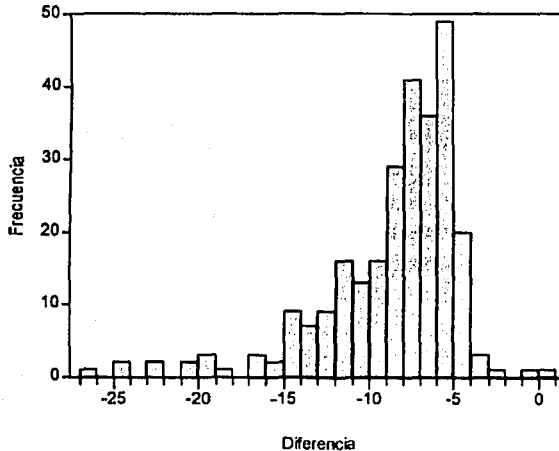
Gráfico 2.7 Diferencia Maya-WTI



En el siguiente histograma (gráfico 2.8) se presenta la distribución de frecuencias de la diferencia de las series y es claro observar que el comportamiento de la distribución

ya no se asemeja al de una distribución normal, pero es posible apreciar que la mayoría de las diferencias se acumulan formando un sesgo, el cual se presenta del lado derecho, sin embargo, resulta irrelevante hacia que lado tiende pues esto sólo depende de la manera en que se consideró la diferencia, si se hubiese tomado  $WTI - Maya$  entonces el sesgo sería del lado izquierdo, lo realmente importante es apreciar la presencia del sesgo. Además se observa acumulación de aproximadamente 64% de las observaciones entre los 10 y los 5 dólares de diferencia, remarcando que la diferencia entre el crudo ligero y el pesado, la mayor parte de las veces, ha sido de 5.5 dólares si se considera el punto medio de la clase con mayor frecuencia.

Gráfico 2.8 Histograma Maya-WTI



Serie: Maya-WTI  
 Muestra:1980:01 2002:03  
 Observaciones 267

Media -8.540535  
 Mediana -7.572630  
 Máximo 0.119390  
 Mínimo -26.53188  
 Desv. Est. 4.005219  
 Asimetría -1.660362  
 Curtosis 6.723931

Jarque-Bera 276.9554  
 Probabilidad 0.000000

Después de haber estudiado las dos posibilidades para considerar la relación entre los crudos y debido a que el cociente  $Maya/WTI$  tiene un mejor comportamiento que la diferencia, conservando una mejor distribución de las observaciones sin presentar el efecto del sesgo, entonces se considerará el cociente como el objetivo a modelar.



## 2.4 Series de Tiempo

Los datos referentes a los precios de los combustibles Maya y WTI forman series de tiempo, por ello es importante definir lo que es una serie de tiempo y algunas de sus características.

Una *serie de tiempo* se define como una secuencia de observaciones ordenadas por una unidad de tiempo, es decir, son observaciones realizadas sobre algún factor o evento en particular, las cuales serán obtenidas de manera discreta o continua, con la restricción de que el tiempo entre observaciones debe ser constante durante todo el periodo de estudio, ya que no es recomendable tener observaciones que, por ejemplo, en un principio sean mensuales, después semestrales y luego anuales.

Cuando se habla de observaciones de tipo discreto se está considerando que éstas se realizan de manera periódica y aislada, por ejemplo series de tiempo anuales, mensuales, semanales o diarias. Existen también las series de tiempo de tipo continuo, que son observaciones instantáneas del evento, estas series son menos frecuentes, ejemplos de éstas serían las mediciones realizadas por un oscilógrafo el cual registra todas las oscilaciones de un amplificador de sonido, o bien mediciones hechas por un electrocardiograma.

Lo que se espera de una serie de tiempo es que las observaciones estén igualmente espaciadas y que no presenten datos faltantes. Cuando existen datos faltantes y éstos no pueden ser obtenidos, lo que se hace es completarlos utilizando un método teórico. Existen muchos métodos que pueden ser utilizados para este fin, por ejemplo, uno de ellos consiste en reemplazar el dato por el valor promedio del resto de la serie o bien utilizar el valor promedio del periodo en el cual el dato es faltante. Otro método consiste en calcularlo como el promedio de las dos observaciones adyacentes, lo que matemáticamente es conocido como punto medio.

### 2.4.1 Tendencia, Estacionalidad y Ciclos

La utilidad de estudiar y analizar una serie de tiempo es variada y no sólo es de carácter estadístico; su estudio permite visualizar el comportamiento de la variable a través del tiempo, considerando diferentes aspectos que la caracterizan, como son: tendencia, ciclicidad, comportamiento estacional y los movimientos irregulares; siendo su último fin pronosticar a futuro valores de la serie que permitan planear políticas económicas o sociales que involucren a la variable.

Existen esencialmente dos maneras de vincular los factores que conforman una serie de tiempo una de ellas es el método multiplicativo y otra es el método aditivo. Si se denota como  $T_t$ ,  $C_t$ ,  $S_t$ ,  $I_t$  a las componentes de tendencia, ciclicidad, estacionalidad y movimientos irregulares en el tiempo  $t$  respectivamente, entonces en el modelo multiplicativo se supone que  $Y_t = T_t C_t S_t I_t$ , y en el modelo aditivo  $Y_t = T_t + C_t + S_t + I_t$ .

Resulta importante estudiar cada una de estas componentes por separado para entender el comportamiento general de la serie.

La *tendencia* puede definirse como la evolución de la serie a lo largo del tiempo. Lo más común es suponer que la *tendencia* es lineal, es decir que depende linealmente del tiempo, sin embargo es posible que la *tendencia* sea de tipo exponencial en cuyo caso se supone que la *tendencia* es un porcentaje de crecimiento casi constante.

La mayoría de las veces es fácil observar si una serie sigue determinada tendencia con sólo observar su gráfica. Del gráfico 2.5, donde se muestra la relación de los precios de los crudos, se puede observar que la serie Maya/WTI sigue una tendencia ligeramente decreciente. Existe otra posibilidad para determinar de un modo más formal si existe tendencia en la serie, se corre una regresión simple en la cual se haga depender a la serie del tiempo, entonces para el caso de los crudos se obtuvo como resultado la siguiente ecuación, los valores entre paréntesis indican, el valor del estadístico  $t$  correspondiente a cada uno de los parámetros y el  $p$ -value correspondiente, en ese orden.



$$\begin{aligned}(\text{Maya}/\text{WTI})_t &= 0.794013 - 0.0004839 t \\ &(75.5427, 0) \quad (-7.11735, 0)\end{aligned}$$

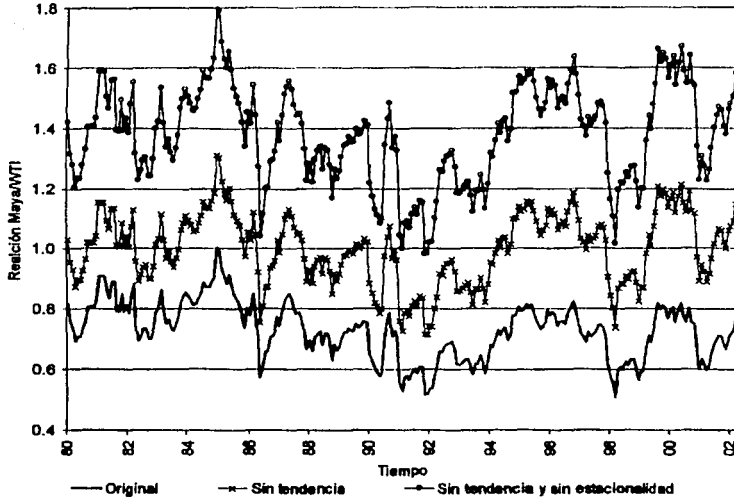
efectivamente se puede concluir que la serie tiene una tendencia negativa aunque muy pequeña, pues el coeficiente asociado a la variable tiempo es estadísticamente significativo.

El otro factor importante es la *estacionalidad* de la serie, es decir, si ésta sigue un comportamiento de acuerdo con períodos fijos, o dicho de otra manera, si se presentan patrones repetitivos a lo largo del año en ciertas estaciones por ejemplo. Para determinar si la serie es estacional se puede tomar una especie de índice, el cual se forma promediando todos los valores correspondientes al mes y después utilizando dicho índice se ponderen los datos, esto elimina la posible estacionalidad de los datos. Al realizar este procedimiento para los datos de la relación de crudos, el cambio en los valores no fue muy significativo por tanto se concluye que la serie no presenta un patrón estacional.

Para observar si existía un patrón estacional en los datos de la relación Maya/WTI se formaron las variables  $D_t$  que funcionarán como índices, se formaron doce variables de este tipo, una referente a cada mes, las cuales consisten en arreglos formados por ceros y unos. Por ejemplo, la variable correspondiente al mes de enero consiste en una variable con unos en las posiciones correspondientes a dicho mes y ceros en los lugares restantes. Una vez formadas las variables estacionales, se corre una regresión que las contenga como variables independientes. Los coeficientes obtenidos en la regresión se utilizan para desestacionalizar a la serie, ya con los datos sin la tendencia lo único que se hace es dividir cada dato entre el coeficiente correspondiente.

En el gráfico 2.9 se muestra los cambios en la serie de la relación Maya/WTI después de haber eliminado la tendencia y la estacionalidad de la serie. En efecto del gráfico es observable que la serie original seguía una tendencia decreciente, pero al comparar la serie sin tendencia y aquella a la cual además se le eliminó la estacionalidad, ambas series son sumamente parecidas, la diferencia consiste en que se marcan más las pendientes de los períodos de crecimiento y de decrecimiento.

**Gráfico 2.9 Relación Maya/WTI sin tendencia y sin estacionalidad**



Finalmente el otro factor que caracteriza a una serie es la presencia de **ciclos**, a este factor se le prestará mayor atención, pues como se dijo antes es una de las características más importantes a modelar, por tanto su análisis se realizará más adelante. Por el momento sólo se dirá que para el caso de los crudos la ciclicidad de la serie es clara, sin embargo la duración de los ciclos varía de acuerdo a la época, es decir, no se conservan de la misma longitud durante todo el periodo de estudio.

Por ejemplo, en la gráfica de la relación Maya/WTI se observan ciclos de duración indefinida, es posible considerar que el primer ciclo se presenta de 1980 a 1992, periodo en el cual los datos crecen y decrecen de manera apreciable, después se puede considerar un ciclo de menor duración que abarcaría de 1992 a 1998 y finalmente se tiene el inicio del próximo ciclo que va de 1998 al final del periodo de estudio. Sin embargo, la apreciación antes descrita de los ciclos es solamente una de entre otras que

se pueden apreciar, por ejemplo, se puede considerar que se forman ciclos de duración de seis años, los cuales están definidos por un crecimiento durante los cuatro primeros años para decrecer durante los dos últimos.

Si se observa el gráfico de la diferencia Maya-WTI en este caso los ciclos que se observan son algo diferentes. Durante los primeros cinco años los datos presentan un periodo de crecimiento y desde 1985 hasta 1992 los datos siguen una tendencia decreciente, con lo cual se forma el primer ciclo que duraría de ese modo doce años; después de 1992 la serie vuelve a crecer y comienza el decrecimiento a partir de entonces, pero en 1998 frena el decrecimiento y vuelve a crecer pero sólo por un periodo muy corto, alrededor de dos años, y después decrece nuevamente.

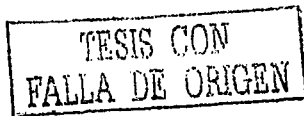
El último factor que compone a una serie es lo que se conoce como el factor irregular, en él se engloban todos los efectos aleatorios que afectan a la serie, pero que no pueden ser explicados por efectos determinados y por lo tanto son impredecibles.

## 2.4.2 Estacionariedad

Al trabajar con series de tiempo una de las primeras consideraciones que se hacen es que la serie con la que se está trabajando es estacionaria, es decir, que tanto su media como su varianza son constantes y que además la covarianza entre observaciones únicamente depende del tiempo entre ellas y no del momento en que se calcula. Sin embargo, en la realidad es frecuente encontrarse con series que no lo son.

Una serie de tiempo puede ser concebida como una sucesión de observaciones generadas por un proceso estocástico, esto lleva a definir lo que es un proceso estocástico. Un *proceso estocástico* es un conjunto de variables aleatorias asociadas a un conjunto índice de números reales, de tal forma que a cada elemento de dicho conjunto le corresponda una y sólo una variable aleatoria. En el caso de las series de tiempo el conjunto índice considerado se toma en relación al tiempo.

Uno de los procesos estocásticos que más se ha estudiado es el conocido como *proceso estocástico estacionario*. Un proceso estocástico es estacionario si su media y su



varianza son constantes en el tiempo y si además el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos y no del tiempo en el cual se ha calculado la covarianza. Es decir, lo que se espera es que no importando el instante en el cual se midan la media, la varianza y la covarianza, éstas deben permanecer iguales. Matemáticamente se dice que si  $Y_t$  es una serie de tiempo estocástica con las siguientes propiedades,

$$E[Y_t] = \mu \text{ (cte.)}$$

$$\text{Var } [Y_t] = E[(Y_t - \mu)^2] = \sigma^2 \text{ (cte.)}$$

$$\text{Cov } [Y_t, Y_{t+k}] = E[(Y_t - \mu) ( Y_{t+k} - \mu )] = \gamma(k) = \gamma_k$$

entonces el proceso estocástico del cual proviene es estacionario.

Para probar la estacionariedad de la serie es posible considerar la gráfica de ésta a lo largo del tiempo y observar si su comportamiento conserva las propiedades que debe cumplir para ser estacionaria. Una manera más precisa para determinar si existe o no estacionariedad es una prueba sencilla basada en la función de autocorrelación (FAC), la cual para el rezago  $k$  está denotada por  $\rho_k$ , y se define como

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Este cociente de correlación, al igual que los demás se encuentra entre -1 y 1. Al graficar  $\rho_k$  frente a  $k$ , se obtiene lo que se conoce como *correlograma*. En la práctica el cálculo del cociente se realiza con los datos muestrales y por lo tanto el correlograma obtenido recibe el adjetivo muestral.

Si se observa la gráfica 2.5 de la relación Maya/WTI, se puede establecer que aun cuando durante ciertos periodos la serie se aleja de su media, la cual es igual a 0.73, regresa a una vecindad de la misma, esto indica que la media de las observaciones permanece constante a lo largo del periodo, faltaría entonces verificar si la varianza y las covarianzas cumplen con las propiedades deseadas.

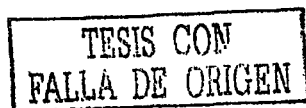
En la tabla 2.1 se presenta el correlograma muestral de la relación Maya/WTI, llegando hasta el rezago 20. En éste, se observa que la función de autocorrelación comienza en un valor muy grande y va disminuyendo paulatinamente conforme la separación entre observaciones se va alejando, es decir, cuando el orden del rezago aumenta la correlación entre las observaciones es menor. Sin embargo dicho decrecimiento se puede considerar que es rápido, este patrón en el comportamiento de la serie es una indicación de que ésta es estacionaria.

Tabla. 2.1 Correlograma de Maya/WTI

Autocorrelación	Correlación Parcial	AC	PAC	Q-Estad.	Prob
*****	*****	1 0.917	0.917	227.15	0.000
*****	. .	2 0.818	-0.148	408.31	0.000
*****	. .	3 0.732	0.043	553.93	0.000
*****	. .	4 0.653	-0.024	670.24	0.000
****	. .	5 0.585	0.026	764.08	0.000
****	. .	6 0.529	0.019	841.05	0.000
****	. .	7 0.488	0.058	906.82	0.000
****	. .	8 0.469	0.108	967.90	0.000
***	. .	9 0.453	-0.010	1025.0	0.000
***	. .	10 0.423	-0.071	1075.1	0.000
***	. .	11 0.390	-0.004	1117.8	0.000
***	. .	12 0.354	-0.035	1153.1	0.000
**	. .	13 0.300	-0.117	1178.5	0.000
**	. .	14 0.244	-0.022	1195.4	0.000
.*	. .	15 0.181	-0.087	1204.7	0.000
.*	. .	16 0.143	0.113	1210.6	0.000
.*	. .	17 0.132	0.078	1215.6	0.000
.*	. .	18 0.138	0.070	1221.1	0.000
.*	. .	19 0.154	0.067	1228.0	0.000
.*	. .	20 0.163	-0.066	1235.7	0.000

Sin embargo este criterio resulta un tanto subjetivo para determinar si la serie es estacionaria o no, y aunque es adecuado en algunos casos para otros el comportamiento no es tan claro y es necesario utilizar otros métodos. Una prueba alternativa para determinar la estacionariedad de una serie es la *prueba de la raíz unitaria*, de manera global esta prueba consiste en efectuar el siguiente modelo de regresión

$$Y_t = \rho Y_{t-1} + \varepsilon_t$$



si como resultado de éste se obtiene que estadísticamente  $\rho = 1$ , entonces surge lo que se conoce como el problema de raíz unitaria y se dice que la variable  $Y_t$  tiene una raíz unitaria. En algunas ocasiones en lugar de considerar la regresión anterior se efectúa la siguiente que es equivalente,

$$\begin{aligned}\nabla Y_t &= (\rho - 1)Y_{t-1} + \varepsilon_t \\ &= \delta Y_{t-1} + \varepsilon_t\end{aligned}$$

Entonces para determinar si una serie es estacionaria de acuerdo con esta prueba lo que se hace es efectuar la regresión correspondiente y determinar si  $\rho$  es estadísticamente igual a 1, o alternativamente, si  $\delta$  es igual a cero, utilizando por ejemplo el estadístico  $t$  de Student. El valor de  $t$  así obtenido no sigue una distribución  $t$  de Student, sin embargo se han elaborado tablas de los valores críticos de este estadístico, el cual se conoce como el estadístico  $\tau$  de Dickey-Fuller y que fueron mejorados por MacKinnon. Entonces se tiene que si el valor absoluto del estadístico  $\tau$  excede los valores absolutos críticos de Dickey-Fuller o de MacKinnon, entonces no se rechaza la hipótesis de que la serie de tiempo es estacionaria. Pero si por el contrario el valor es menor que el valor crítico, la serie de tiempo es no estacionaria.

En algunas ocasiones la prueba se aplica a regresiones en las cuales se incluyen los términos de la constante y de la tendencia, es decir, se consideran modelos con alguna de las siguientes formas,

$$\begin{aligned}\nabla Y_t &= \beta_1 + \delta Y_{t-1} + \varepsilon_t \\ \nabla Y_t &= \beta_1 + \beta_2 t + \delta Y_{t-1} + \varepsilon_t\end{aligned}$$

Finalmente al aplicarte la prueba a la variable de la relación Maya/WTI se obtiene que el valor absoluto del estadístico  $\tau$  calculado es 3.8358, que es mayor que los valores críticos correspondientes, por lo tanto la serie es estacionaria, y por lo tanto no es necesario diferenciarla. La tabla 2.2 muestra los resultados obtenidos en la prueba, así como la regresión realizada. En la regresión se consideró el modelo en el cual se incluye la constante pues al incluir la tendencia, ésta resultaba estadísticamente no significativa.

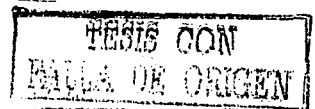




Tabla 2.2 Prueba de la Raíz Unitaria al Maya/WTI

Hipótesis Nula : RELMAYAWTI tiene una raíz unitaria

Exógena: Constante

Longitud del rezago: 1 (Automático, Rez. Máx.=15)

	t-Estadístico	Prob.*
Prueba Estadística Dickey-Fuller	-3.835893	0.0029
Valores críticos:		
1% nivel	-3.454906	
5% nivel	-2.872244	
10% nivel	-2.572547	

\*MacKinnon (1996) one-sided p-values.

Prueba Dickey-Fuller Aumentada

Variable Dependiente: D(RELMAYAWTI)

Método: Mínimos Cuadrados

Muestra (ajustada): 1980:03 2002:03

Observaciones Incluidas: 265 después de ajustar extremos

Variable	Coefficiente	Error Std.	t-Estadístico	Prob.
RELMAYAWTI(-1)	-0.093566	0.024392	-3.835893	0.0002
D(RELMAYAWTI(-1))	0.168175	0.060651	2.772825	0.0060
C	0.068264	0.017918	3.809753	0.0002
R-cuadrada	0.067451	Media var. depend.		2.43E-05
R-cuadrada ajustada	0.060332	D.S. var. depend.		0.037461
E. S. de regresión	0.036313	Akaike criterio		-3.782018
Suma resid. Cuad.	0.345485	Schwarz criterio		-3.741491
Log likelihood	504.1172	F-estadístico		9.475140
Durbin-Watson	1.983351	Prob(F-estadístico)		0.000106

El hecho de que una serie sea no estacionaria no es tan grave como parece, existen métodos sencillos para solucionar este problema y convertirla en una serie con la condición deseada. En la mayoría de las ocasiones la no estacionariedad de la serie se debe únicamente a una afectación en el nivel de la serie, es decir la serie cuenta con una tendencia propia.

En estos casos es posible eliminar la tendencia mediante la aplicación del operador diferencia un número apropiado de veces. El orden del operador diferencia es de gran importancia, si se aplican diferencias a series que ya son estacionarias, éstas

seguirán siendo estacionarias, lo cual significa que puede sobre diferenciarse una serie generando así problemas, pues al diferenciar la serie, la varianza de ésta se incrementa, lo que puede provocar problemas en la identificación del modelo y además de que se pierden observaciones innecesariamente. Entonces para conocer cuantas veces es adecuado diferenciar la serie para volverla una serie estacionaria, es necesario analizar para qué rezago el coeficiente de autocorrelación es prácticamente nulo, de forma tal que la diferencia del orden de ese rezago es la adecuada.

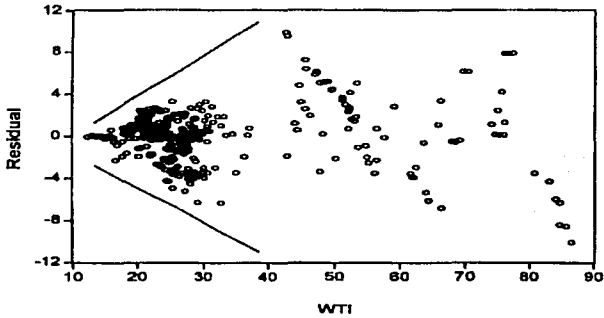
## 2.5 Análisis de Regresión para el caso específico de los crudos

Como segundo punto importante en el análisis fue elaborar un análisis de regresión para ambas series. Se considera una regresión simple en la cual la variable dependiente sea el Maya y el regresor o variable independiente el WTI, esto con el objetivo de encontrar una posible relación entre ambas variables.

Los resultados obtenidos condujeron a suponer que existe correlación serial en los errores, pues aun cuando el valor de la  $R^2$  es significativo, 0.96, resulta que el valor del estadístico Durbin-Watson es bajo, indicando la posibilidad de autocorrelación, el resultado del estadístico es de 0.28, sin embargo lo más probable es que las variables estén cointegradas.

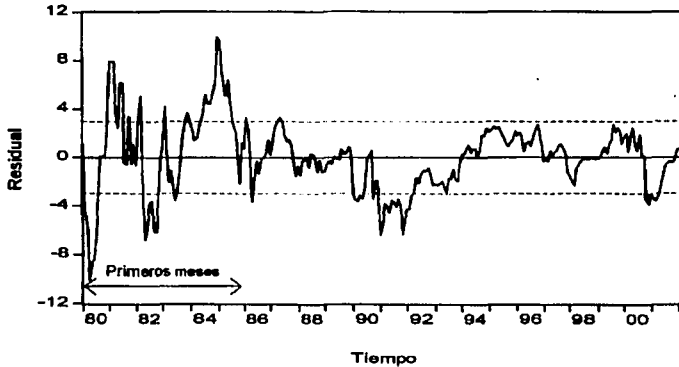
Además el gráfico de los residuales (Gráfico 2.10) indica la posible presencia de heteroscedasticidad, ya que a mayor precio del WTI la varianza es mayor, violando el supuesto de varianza constante, es decir, los residuales se acumulan formando una especie de cono, cuya cúspide se encuentra del lado izquierdo donde los precios son bajos, para los cuales el error cometido es pequeño. De nuevo, se presentan diferencias entre los precios del WTI menores y mayores que 40. Los residuales correspondientes a precios mayores tienen errores grandes, pero no tienen ningún comportamiento apreciable a simple vista, lo cual si ocurre para el otro grupo.

Gráfico 2.10 Residuales & WTI



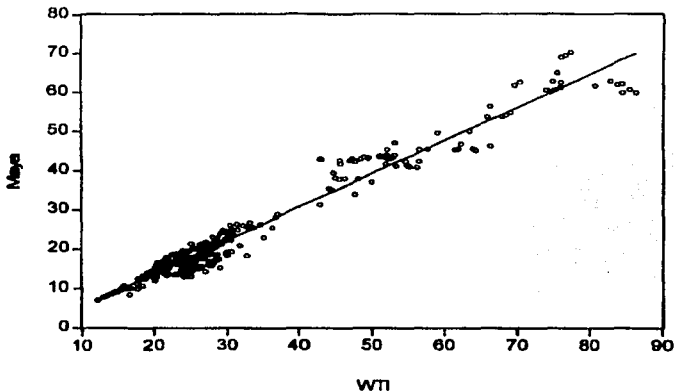
Si se analiza el comportamiento de los residuales a través del tiempo, como se muestra en el gráfico 2.11, entonces es posible observar que los residuales mayores se presentan en los primeros meses del estudio, y para observaciones posteriores el valor de los residuales disminuye. Otro aspecto importante de ver, es la presencia de ciclos en el comportamiento de éstos, pues se aprecia que los residuales crecen y decrecen de manera periódica, con ciclos de duración indefinida.

Gráfico 2.11 Residuales & Tiempo



Al observar el gráfico de dispersión de los datos (gráfico 2.12) se puede apreciar que la relación entre ellos es efectivamente lineal, no obstante es notable que para precios del WTI mayores a 40 dólares, el comportamiento estructural cambia, variando la dispersión de las observaciones.

Gráfico 2.12 Dispersión WTI & Maya



### 2.5.1 Cambio estructural en los datos

En los análisis que se han hecho hasta este momento, se intuye sobre la posible diferencia en el comportamiento para las observaciones correspondientes a precios del WTI mayores que 40 y las menores, lo cual es posible explicarlo con un enfoque socio-político, ya que cuando los precios aumentan significativamente, éstos deben ser controlados para lograr que estos bajen a niveles aceptables.

Para verificar si efectivamente existe diferencia entre ambos bloques se puede utilizar la *prueba de Chow*, es una prueba sencilla que utiliza las sumas de cuadrados de los residuos para determinar si existe diferencia estructural entre ambas. Por tanto es necesario correr las regresiones por separado de cada uno de los bloques, después, utilizando los resultados de las sumas de cuadrados de los residuales correspondientes a cada regresión, así como el correspondiente a la regresión en la cual no se hace

distinción alguna, se construye un estadístico con distribución F de Fisher y se compara con su correspondiente valor en tablas para rechazar la hipótesis de que ambas series tengan el mismo comportamiento estructural.

Entonces considerando la regresión para precios del WTI mayores que 40 se obtuvo que la suma de cuadrados debida a los residuales es  $S_2 = 1186.4029$ , teniendo  $73 = n_1 - k$  grados de libertad y para la regresión con precios menores que 40 la suma de cuadrados es de  $S_3 = 742.3284$  con  $190 = n_2 - k$  grados de libertad. Siguiendo el algoritmo de la prueba, se suma los valores de dichas sumas de cuadrados, de donde se obtiene  $S_4 = 1928.7314$  que tiene  $263 = n_1 + n_2 - 2k$ . Finalmente se calcula  $S_5 = 394.5394$ , la cual resulta de restarle a la suma de cuadrados de la regresión sin distinciones  $S_1 = 2323.2708$  el valor de  $S_4$ .

Se forma entonces el estadístico F de la siguiente manera,

$$F = \frac{\frac{S_5}{k}}{\frac{S_4}{n_1 + n_2 - 2k}} = 26.8995$$

En la fórmula anterior k corresponde al número de parámetros estimados, que en este caso particular es igual a 2. Dado que el valor de F obtenido es mayor que el correspondiente valor en tablas de  $F_{(2,263)}$ , se rechaza la hipótesis de que ambas regresiones tienen el mismo comportamiento estructural, es decir, no hay estabilidad estructural y resulta conveniente separar los datos de acuerdo al precio del WTI.

A pesar de que los resultados así obtenidos parecen ser lógicos, existen problemas al utilizar la prueba de Chow, ya que los supuestos en los que se basa esta prueba están referidos a que los términos de error de ambos bloques están distribuidos con la misma varianza (homoscedástica) y que son independientes. Dado que ambos bloques no tienen varianza constante, se violan los supuestos de la prueba y por lo tanto los resultados obtenidos no son válidos. A pesar de esto la presencia de un cambio estructural es posible, pero será necesario encontrar otra manera de detectarla.

Los resultados así obtenidos indican la diferencia entre ambos bloques, sin embargo no es posible saber si dicha conclusión se debe a la diferencia en los valores de las intersecciones al origen, a la diferencia en los valores de las pendientes, o a ambos. Con el fin de detectar cual es el origen de la diferencia se realizaron regresiones incluyendo una variable dicotómica  $D_1$  que diferenciara entre los precios del WTI mayores que 40 y los menores, es decir, la regresión propuesta es,

$$\text{Maya} = \alpha_1 + \beta_1 \text{WTI} + \beta_2 D_1 + u_4$$

donde la variable dicotómica está definida como,

$$D_1 = \begin{cases} 0 & \text{si WTI} > 40 \\ 1 & \text{si WTI} \leq 40 \end{cases}$$

Al considerar esta regresión lo que se está probando es si la diferencia estructural varía en la ordenada al origen, es decir, se prueba si la recta ajustada para ambos bloques es la misma o varía únicamente en su ordenada al origen, pero conservando la misma pendiente. Por lo tanto el coeficiente de la variable dicotómica es el que interesa, si éste resultará significativamente diferente de cero, se concluiría que las rectas ajustadas son diferentes pero paralelas, o en caso contrario las rectas serían las mismas. De los resultados de la regresión se obtiene un estadístico  $t$  significativo correspondiente a la variable dicotómica, y por lo tanto las intersecciones en el origen son diferentes en ambas rectas. El modelo resultante es el siguiente, entre paréntesis se muestran los estadísticos  $t$  correspondientes.

$$\text{Maya} = 5.57135 + 0.71382 \text{WTI} - 5.89694 D_1 \\ (4.2319, 0) \quad (33.4869, 0) \quad (-7.0082, 0)$$

Como el estadístico  $t$  referente a la variable dicotómica es significativo, se concluye que existe diferencia entre las intersecciones al origen de las rectas de los bloques. Entonces, como para precios del WTI mayores que 40 la variable dicotómica vale cero, en este caso la recta ajustada tiene un intercepto igual a 5.57; y para menores que 40 la variable dicotómica vale uno y por lo tanto el intercepto es igual a -0.32; pero en ambos

casos el valor de la pendiente se conserva igual, es decir se tienen los siguientes resultados,

$$\begin{aligned} E [ \text{Maya} | \text{WTI}, \text{Dicot.}=1 ] &= - 0.32559 + 0.71382 \text{ WTI} && \text{si } \text{WTI} < 40 \\ E [ \text{Maya} | \text{WTI}, \text{Dicot.}=0 ] &= 5.57135 + 0.71382 \text{ WTI} && \text{si } \text{WTI} > 40 \end{aligned}$$

Con lo anterior se ha probado que entre las rectas que ajustan a cada bloque la diferencia se da únicamente en la intersección, pero puede ocurrir que además difieran en la pendiente, es decir, que el coeficiente de la variable independiente WTI sea significativamente diferente de cero. Una manera de probar esto es mediante una regresión que involucre como variable Independiente el producto del WTI por la variable dicotómica, si además se incluye la variable dicotómica como tal, será posible, igual que antes, distinguir entre las intersecciones. Entonces la regresión propuesta es:

$$\text{Maya} = \alpha_1 + \beta_1 \text{ WTI} + \alpha_2 D_i + \beta_2 (\text{WTI} * D_i) + u_i$$

Considerando el supuesto básico  $E[u_i] = 0$  al calcular las esperanzas condicionales se obtiene:

$$\begin{aligned} E [ \text{Maya} | \text{WTI}, D_i = 0 ] &= \alpha_1 + \beta_1 \text{ WTI} && \text{cuando } \text{WTI} > 40 \\ E [ \text{Maya} | \text{WTI}, D_i = 1 ] &= (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \text{ WTI} && \text{cuando } \text{WTI} \leq 40 \end{aligned}$$

Los resultados de dicha regresión están dados a continuación y es claro observar que el coeficiente correspondiente a la variable dicotómica resultó significativamente diferente de cero y el coeficiente del producto entre WTI y la variable dicotómica también es diferente de cero, aun cuando su estadístico t es cercano a 2. Por tanto se puede concluir que ambos bloques tienen diferentes rectas ajustadas, con ordenadas al origen y pendientes diferentes.

$$\begin{aligned} \text{Maya} &= 7.10998 + 0.68816 \text{ WTI} - 9.19963 D_i + 0.09754 (\text{WTI} * D_i) \\ &(4.6988, 0) \quad (27.8768, 0) \quad (-5.0217, 0) \quad (2.0284, 0.0437) \end{aligned}$$

Además si se calculan las esperanzas condicionales que permiten diferenciar las dos rectas resultantes se tiene,

$$E [ \text{Maya} | \text{WTI}, D_i = 0 ] = 7.1099 + 0.6881 \text{ WTI} \quad \text{cuando } \text{WTI} > 40$$

$$E [ \text{Maya} | \text{WTI}, D_i = 1 ] = -2.0896 + 0.7857 \text{ WTI} \quad \text{cuando } \text{WTI} \leq 40$$

De donde resulta clara la diferencia entre las rectas resultantes para los bloques, además se observa que la inclinación de la recta cuando el precio del WTI es mayor que 40 es de poco más de 34° y cuando el precio es menor que 40 el ángulo de inclinación es de aproximadamente 38°, esto explica porque el valor del estadístico t correspondiente al producto de la variable dicotómica por el WTI, es poco significativo, ya que si bien es cierto que las pendientes son diferentes lo son solamente por una diferencia pequeña.

Además podría ser conveniente realizar una regresión por tramos estableciendo como punto umbral el precio de 40 dólares para el WTI, éste punto donde existe un cambio estructural en los datos. Utilizando la variable dicotómica, se puede realizar una regresión por tramos, diferenciados de acuerdo al precio del WTI, en la cual las variables explicativas sean el WTI tal cual y una variable calculada como  $(\text{WTI} - X^*) D_i$ , en donde  $X^*$  se refiere al valor en el cual se considera existe un cambio estructural, es decir, se propone el siguiente modelo

$$\text{Maya} = \alpha_1 + \beta_1 \text{ WTI} + \beta_2 (\text{WTI} - X^*) D_i + u_i$$

Si  $\beta_2$  no es estadísticamente significativo, es decir que puede considerarse igual a cero, entonces la regresión por tramos no es conveniente, ya que en ese caso las rectas ajustadas en cada tramo son iguales. Este hecho resulta más claro al analizar las esperanzas condicionadas del modelo.

$$E [ \text{Maya} | \text{WTI}, D_i = 0, X^* ] = \alpha_1 + \beta_1 \text{ WTI}$$

$$E [ \text{Maya} | \text{WTI}, D_i = 1, X^* ] = \alpha_1 - \beta_2 X^* + (\beta_1 + \beta_2) \text{ WTI}$$



En este caso particular de los precios de crudos, el valor de  $X^*$  es 40, pues es el valor en el cual se marca el cambio, no se consideró la constante  $\alpha_1$ , ya que al correr la regresión resultó ser estadísticamente insignificante, así los resultados de la regresión son los siguientes:

$$\text{Maya} = 0.795835 \text{ WTI} + 0.137443 (\text{WTI} - 40) D_1 \\ (153.2219, 0) \quad (9.352538, 0)$$

Como el valor de  $\beta_2$  es significativamente diferente de cero, se concluye que una regresión por tramos es conveniente, reafirmando así que el comportamiento estructural en cada tramo o bloque es diferente.

A partir del análisis anterior, se muestra una clara diferencia en el comportamiento, sin embargo se debe observar que los precios del WTI que cumplen con ser mayores que 40 dólares son aquellos observados al principio del periodo, es decir, los precios desde enero de 1980 hasta diciembre de 1985; después de esta fecha el precio del WTI no ha excedido los 40 dólares salvo en tres ocasiones que fueron a finales de 1990, fecha cercana a la guerra del Golfo Pérsico.

La elección del valor donde se da el cambio estructural se designó por observación del diagrama de dispersión de las variables, donde efectivamente se observan dos grupos. El primero correspondiente a precios bajos presenta un comportamiento denso, es decir los datos están más agrupados, por otro lado el grupo para precios mayores, presenta mayor dispersión. Así el valor umbral del cambio estructural igual a 40, se consideró éste como un precio intermedio entre ambos grupos. Sin embargo, si se cambia dicho valor umbral y se realiza la correspondiente prueba de Chow, por ejemplo, se obtiene que al considerar como valor umbral un precio menor a 28 dólares entonces los resultados obtenidos dicen que no existe cambio estructural, pero cuando el valor umbral es mayor que 28 entonces los resultados son cada vez más significativos y muestran que hay cambio estructural. Esto muestra que para valores mayores del comportamiento de los datos no sigue un comportamiento que sea posible caracterizar por una recta determinada, es decir, cada vez que se varía el precio es posible hacer un ajuste diferente.

## Capítulo 3

# REGRESIÓN DINÁMICA

### 3.1 Descripción de las funciones de transferencia

En los anexos 2 y 3 se resumió de manera general los conceptos y la teoría del análisis de regresión y del análisis de series de tiempo. Ambos métodos permiten modelar de manera satisfactoria el comportamiento de un conjunto de datos, que en el caso de las series de tiempo consisten en observaciones a lo largo de un determinado periodo. Sin embargo, existe una manera de que ambos métodos se complementen logrando con ello una mejor aproximación al comportamiento real.

En el enfoque de regresión clásico se busca modelar a la variable Y en función de la o las variables X pudiendo considerar los valores rezagados o actuales de ellas, con la restricción de que las variables X no estén relacionadas entre ellas y que expliquen adecuadamente el comportamiento de Y, sin embargo se sabe que los modelos obtenidos no son del todo perfectos y que por tanto, se producen errores. Dichos errores contendrán la parte de variabilidad de Y que no fue explicada por las variables X involucradas en el modelo.

El análisis univariado de series de tiempo permite encontrar modelos que describan el comportamiento de la serie en estudio, haciendo que la variable se explique por sí misma utilizando sus valores rezagados para ello. Entonces es posible vincular el enfoque de regresión clásico con el análisis de series de tiempo con el objetivo de construir modelos que permitan hacer estimaciones más precisas. A esta combinación entre los modelos de regresión clásicos y el análisis de series de tiempo, se le conoce como regresión dinámica, funciones de transferencia o modelos MARMA (modelos multivariados autoregresivos de medias móviles).

Por lo tanto una forma efectiva de explicar la parte de la variabilidad que no fue explicada por las variables incluidas en el modelo, es utilizando el análisis de series de

tiempo para ajustar un modelo ARIMA a los errores resultantes, de modo tal que se pueda explicar dicha variabilidad. Los resultados así obtenidos mejoran los modelos y los pronósticos realizados, debido a que será posible estimar los valores de los errores de acuerdo con el modelo ARIMA correspondiente.

Se concluye entonces que las funciones de transferencia relacionan la variable dependiente con sus propios valores rezagados, con los valores actuales o rezagados de una o más variables independientes, y con un error que está parcialmente explicado por un modelo de series de tiempo.

Para desarrollar la teoría de la regresión dinámica se dirá que los modelos de este tipo para una serie de tiempo  $Y_t$ , pueden interpretarse como una representación matemática que descompone a la serie en dos partes, la primera comprendida por todas las variables explicativas incluidas en el modelo y la segunda parte determinada por el efecto de las variables excluidas del modelo. Lo anterior se puede escribir matemáticamente como,

$$Y_t = Y_t^* + N_t$$

donde  $Y_t^*$  es la parte explicada por las variables explicativas incluidas en el modelo y  $N_t$  representa la parte no explicada.

Para el desarrollo teórico del modelo se considerará, por simplicidad, solamente una variable explicativa,  $X_t$ , y se denotará por  $N_t$  al proceso de inercia.

Los principios básicos para la construcción de un modelo de regresión dinámica son:

- a) La variable explicativa  $X_t$  influye sobre  $Y_{t+k}$ , para  $k \geq 0$  pero no al contrario, es decir, no existe una relación entre  $Y_t$  y  $X_{t+k}$ , con lo cual se establece una relación unidireccional entre las variables.
- b) La relación entre la variable explicativa  $X_t$  y la dependiente  $Y_t$  es constante en el periodo muestral analizado.

- c) La respuesta de  $Y_t$  a las variaciones de  $X_t$  puede aproximarse de forma lineal. Es decir,

$$Y_t^* = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots$$

donde los coeficientes  $v_i$  son constantes a determinar. La representación de los coeficientes  $v_i$  en función del retardo  $i$  se denomina *función de respuesta a impulsos* o *función de transferencia*.

- d) La parte de  $Y_t$  no explicada por  $X_t$  contendrá el efecto de otras variables relevantes excluidas del modelo, este factor se llamará proceso de inercia  $N_t$  y podrá descomponerse en dos partes,

$$N_t = f_t + \alpha_t$$

donde  $f_t$  representa su parte predecible y  $\alpha_t$  es un proceso de ruido blanco.

Es posible, que si todas las variables relevantes fueron incluidas en el modelo, entonces la parte  $f_t$ , la cual representa la parte predecible del proceso de inercia, sea igual a cero, es decir, el proceso de inercia sea ruido blanco.

Cada parte del modelo de regresión dinámica se ajusta por separado, integrando los resultados individuales posteriormente, para formar así el ajuste completo. Por ello es necesario determinar la manera en que se ajusta cada una de las partes, estimando los parámetros de las ecuaciones que los ajustan.

### 3.2 Parametrización de la función de transferencia

La primera parte que se identificará será la que representa el efecto de toda la historia de la variable  $X_t$  sobre la variable  $Y_t$ , cuya ecuación está dada por:

$$\begin{aligned} Y_t^* &= v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots \\ &= (v_0 + v_1 B + v_2 B^2 + \dots) X_t \\ &= v(B) X_t \end{aligned}$$



El objetivo aquí, es estimar los coeficientes  $v_t$  que determinen la ecuación anterior. Si se supone que  $X_t$  sufre un impulso unitario, es decir, que toma el valor uno en  $t$  y cero en los demás periodos entonces cada uno de los coeficientes  $v_t$  representará la respuesta en el periodo  $t+i$ , ante la presencia de dicho impulso.

La representación de  $v(B)$ , el polinomio de retraso, contiene una infinidad de parámetros, por tanto no es posible estimarlos a todos. Para solucionar dicho problema se puede considerar un polinomio  $w(B)$  como una aproximación de  $v(B)$ , donde  $w_m(B)$  está definido como:

$$w_m(B) = w_0 + w_1 B + w_2 B^2 + \dots + w_m B^m$$

o bien, se pueden considerar una infinidad de ellos pero con una estructura de amortiguamiento, como lo es:

$$\frac{w_0}{(1 - \delta B)} = w_0 (1 + \delta B + \delta^2 B^2 + \dots)$$

pero una representación general es suponer que  $v(B)$  sigue una estructura como la que se muestra a continuación, que consiste en un cociente de polinomios de tipo autorregresivo y de medias móviles,

$$v(B) = \frac{w_m(B)}{\delta_a(B)} = \frac{w_0 + w_1 B + w_2 B^2 + \dots + w_m B^m}{1 - \delta_1 B - \dots - \delta_a B^a}$$

Algunas veces la respuesta a un impulso no es inmediata y por lo tanto el sistema reacciona a dicho impulso después de cierto tiempo, si se supone que el tiempo transcurrido desde que ocurrió el impulso y el momento en que se inicia el efecto es  $b$  entonces será necesario añadir este nuevo parámetro a la estructura de  $v(B)$ , obteniendo lo siguiente,

$$v(B) = \frac{w_m(B)}{\delta_a(B)} B^b$$

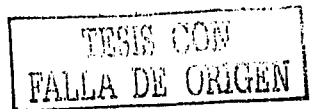
donde, si  $b=0$  entonces la respuesta al impulso es instantánea, pero si es diferente de cero comienza  $b$  periodos después. Así la estructura de  $v(B)$  queda caracterizada por los parámetros de media-móvil, autorregresivo y de respuesta,  $(m, a, b)$ .

Para especificar los parámetros que determinarán la estructura de  $v(B)$  será de utilidad examinar las funciones de correlación total y parcial, ya que existe una analogía entre los procesos ARMA y las funciones de transferencia, lo cual se debe a que la función de respuesta a impulsos (FRI) de una función de transferencia con parámetros  $(m, \alpha, b)$  es similar a la función de autocorrelación simple (FAS) de un proceso ARMA con parámetros  $(m, \alpha)$ , la única diferencia es que la función de respuesta a impulsos comienza con un valor  $w_0$  en el retardo  $t + b$  mientras que el correlograma comienza con un valor de uno en el retardo cero. En la FAS de un proceso ARMA  $(m, \alpha)$  se presentan los primeros  $\alpha$  valores sin pauta fija y a partir del rezago  $\alpha+1$  sigue una estructura senoidal o exponencial. El comportamiento en la FRI de una función de transferencia es muy semejante al de la FAS, con la única diferencia debida al tiempo que se tarda en responder al impulso. En resumen, se concluye que al identificar una función de transferencia con los parámetros  $(m, \alpha, b)$ , se tienen:

- a) Los primeros  $b$  coeficientes iguales a cero
- b) Los siguientes  $m$ , es decir, de  $b$  a  $b+m$  no siguen una estructura fija de variación
- c) Y a partir de  $b+m+1$  decrecen con un comportamiento exponencial o senoidal.

*Relación de la FAS de un proceso ARMA y la FRI de una función de transferencia ( $b=0$ )*

ARMA (p,q)		Función de Transferencia	
$Y_t = \phi_p^{-1}(B)\theta_q(B)a_t$ Estacionario (estable) si las raíces de $\phi_p(B) = 0$ están fuera del círculo unitario		$Y_t = \delta_\alpha^{-1}(B)w_m(B)X_t$ Estacionario (estable) si las raíces de $\delta_\alpha(B) = 0$ están fuera del círculo unitario	
FAS	Si $p=0$ y $q \neq 0$ $q$ coeficientes distintos de cero el resto son nulos	FRI	Si $\alpha=0$ y $m \neq 0$ $m$ coeficientes distintos de cero el resto son nulos
	Si $p \neq 0$ y $q=0$ Decrecimiento geométrico y/o senoidal		Si $\alpha \neq 0$ y $m=0$ Decrecimiento geométrico y/o exponencial
	Si $p \neq 0$ y $q \neq 0$ Decrecimiento con valores iniciales sin pauta fija		Si $\alpha \neq 0$ y $m \neq 0$ Decrecimiento con valores iniciales sin pauta fija



Así se concluye que la parte  $Y_t$  explicada por las variables incluidas en el modelo se parece a un proceso ARMA( $m, \alpha$ ) pero con una diferencia, es necesario considerar el parámetro  $b$  para marcar el periodo de retraso en la respuesta, tomando en consideración que este parámetro puede ser igual a cero, es decir, se puede escribir

$$Y_t^* = \frac{w_m(B)B^b}{\delta_\alpha(B)} X_t$$

### 3.3 Parametrización del proceso de inercia

Después de haber ajustado un modelo a la función de transferencia es posible modelar el proceso de inercia de la serie, o la parte no explicada por las variables independientes, simplemente se supondrá que dicha parte sigue el comportamiento de un proceso ARIMA ( $p, d, q$ ), es decir, que tiene la siguiente estructura,

$$\phi_p(B)\nabla^d N_t = \theta_q(B)a_t$$

Por lo tanto para su estimación se utilizarán los métodos de identificación utilizados en el análisis univariado de series de tiempo, para encontrar el modelo que mejor se ajuste al proceso de inercia.

Con estos últimos resultados se concluye que el modelo de regresión dinámica se puede expresar como una combinación de un proceso ARMA asociado a las variables explicativas y un proceso ARIMA asociado a los residuales, es decir, tiene la siguiente estructura<sup>6</sup>,

$$Y_t = \frac{w_m(B)B^b}{\delta_\alpha(B)} X_t + \frac{\theta_q(B)}{\phi_p(B)\nabla^d} a_t$$

<sup>6</sup> Ver anexo A3 para más detalles de los modelos ARMA y ARIMA



### 3.4 Generalización a más de una variable explicativa

Todo lo anterior se ha explicado considerando sólo una variable independiente, sin embargo, la generalización para varias series explicativas es sencilla. Si estas variables a incluir no están correlacionadas entre sí, se calcula la función de transferencia correspondiente a cada una de las variables por separado y después se estima conjuntamente el modelo global

$$Y_t = v_1(B) X_{1t} + v_2(B) X_{2t} + \dots + v_k(B) X_{kt} + N_t$$

Las funciones de respuesta  $v_i(B)$  son el resultado de los análisis de cada una de las variables y se considerará como estructura de  $N_t$ , la más compleja de las obtenidas en los análisis por separado.

Cuando las series explicativas que se desean considerar no son ortogonales o están correlacionadas, un procedimiento que es posible utilizar cuando se consideran pocas variables, es utilizar los principios de regresión. Por ejemplo, si se consideran sólo dos variables explicativas el proceso para construir el modelo conjunto es:

- a) Construir la función de transferencia para una de ellas

$$Y_t = v_1(B) X_{1t} + N_t$$

- b) Construir la función de transferencia entre ambas variables explicativas  $X_{1t}$ ,  $X_{2t}$

$$X_{2t} = v_{21}(B) X_{1t} + N_t^*$$

- c) Utilizar  $N_t^*$  como regresor de  $N_t$ , y estimar la función de transferencia correspondiente

$$N_t = v_3(B) N_t^* + A_t$$

- d) El modelo global, sustituyendo los resultados anteriores estaría dado por

$$Y_t = v_1(B) X_{1t} + v_3(B) [X_{2t} - v_{21}(B) X_{1t}] + A_t$$

$$\Rightarrow Y_t = [v_1(B) + v_3(B) v_{21}(B)] X_{1t} + v_3(B) X_{2t} + A_t$$





Finalmente se estima el modelo de acuerdo a la última ecuación obtenida y se realizan las pruebas, con el propósito de verificar los supuestos del modelo, y en caso de ser necesario se realizan los cambios para mejorarlo.

### 3.5 Función de correlación cruzada

Para la identificación de la función de transferencia será de mucha utilidad la examinación de la función de correlación cruzada, ya que como se ha mencionado antes, existe una relación con las funciones de correlación que es muy importante y útil. En el caso particular de la función de correlación cruzada, ésta permitirá medir la relación causal entre las variables, para verificar que se cumple el principio de causalidad en el que se basa la teoría de regresión dinámica.

Para definir la función de correlación cruzada, se define primero la función de covarianzas cruzadas. Dados dos procesos estacionarios  $x_t$ ,  $y_t$ , con medias  $\mu_x$ ,  $\mu_y$ , se define su *función de covarianzas cruzadas* como,

$$\gamma_{xy}(k) = E[(x_t - \mu_x)(y_{t+k} - \mu_y)] \quad \forall t \quad k = 0, \pm 1, \pm 2, \dots$$

se dice que ambos procesos son conjuntamente estacionarios si esta función sólo depende del retardo  $k$  y no del instante considerado.

La función de covarianzas cruzadas tiene las siguientes propiedades:

- a)  $\gamma_{xy}(k) = \gamma_{yx}(-k)$
- b) Para  $k > 0$ ,  $\gamma_{xy}(k)$  mide la relación causal de  $y$  hacia  $x$ .
- c) Para  $k < 0$  mide la relación causal de  $x$  hacia  $y$ .
- d) La función no será en general simétrica.

Se dice que  $x_t$  influye linealmente sobre  $y_t$ , si existe algún valor  $\gamma_{xy}(k)$  no nulo para  $k > 0$ . Análogamente se dirá que  $y_t$  influye linealmente sobre  $x_t$  si existen valores no nulos en  $\gamma_{xy}(-k)$  para  $k > 0$ .



Al estandarizar la función de covarianzas cruzadas se obtiene la función de correlación cruzada,  $\rho_{xy}$ , que está definida como:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}$$

Esta función de correlación al ser la estandarización de la función de covarianzas cruzadas cumple con propiedades análogas. Se tiene que  $|\rho_{xy}(k)| \leq 1$ ;  $\rho_{xy}(k) = \rho_{yx}(-k)$ ; y la función no es simétrica alrededor del origen. Además para  $k > 0$  mide la relación desde  $x$  hacia  $y$ , para  $k < 0$  desde  $y$  hacia  $x$ . Para los modelos de regresión dinámica se supone que la función de correlación cruzada es nula para  $k < 0$ .

Para determinar la relación entre la función de covarianzas cruzadas y la función de respuestas impulsos, se suponen dos procesos estacionarios de media cero relacionados por,

$$y_t = v(B)x_t + n_t$$

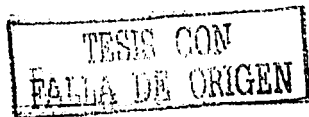
donde  $n_t$  es un proceso estacionario. Resulta intuitivo que las covarianzas cruzadas entre las dos variables deben estar relacionadas con la función  $v(B)$ . Para encontrar dicha relación, se escribe

$$y_t = v_0 x_t + v_1 x_{t-1} + v_2 x_{t-2} + \dots + n_t$$

se multiplica por  $x_{t-k}$  y se toman esperanzas, para encontrar la covarianza

$$\gamma_{xy}(k) = v_0 \gamma_{xx}(k) + v_1 \gamma_{xx}(k-1) + v_2 \gamma_{xx}(k-2) + \dots$$

por lo tanto el comportamiento de la función de covarianzas está alterada por la dinámica de la serie explicativa  $x_t$ . Ahora, si  $x_t$  fuese un proceso de ruido blanco con  $\gamma_{xx}(k) = 0, \forall k \neq 0$ , entonces:



$$\gamma_{xy}(k) = v_k \gamma_{xx}(0)$$

$$v_k = \frac{\gamma_{xy}(k)}{\gamma_{xx}(0)}$$

El resultado anterior surge al suponer que  $x_t$  es ruido blanco, pues en ese caso los valores de la variable  $x$  en diferentes tiempos son independientes uno del otro. Por tanto, solamente si la variable explicativa es ruido blanco, la covarianza cruzada de orden  $k$  entre las series es proporcional a la función de respuesta a impulsos de orden  $k$ .

### 3.6 Identificación de la Función de Transferencia

Supóngase dos procesos, no necesariamente estacionarios, relacionados por la ecuación:

$$Y_t = v(B)X_t + N_t$$

Ahora si las series  $Y$  y  $X$  no son estacionarias y es necesario diferenciarlas, entonces si se supone que  $d_x$  es el orden de diferenciación necesario para convertir a  $X$  en estacionaria, y  $d_y$  el correspondiente para  $Y$ . Si se aplica el mismo número de diferencias,  $d = \max(d_x, d_y)$  a ambas series, éstas serán estacionarias, y si se denota como  $y_t = \nabla^d Y_t$ ,  $x_t = \nabla^d X_t$ ,  $n_t = \nabla^d N_t$ ; se obtiene después de aplicar el operador diferencia  $d$  veces

$$y_t = v(B)x_t + n_t$$

con esto se concluye que las series diferenciadas el mismo número de veces están relacionadas por la misma función de transferencia que las series originales. Por lo tanto el primer paso para la identificación es diferenciar las series el mismo número de veces para que ambas sean estacionarias.

Una vez diferenciadas, se reintentará estimar la función de respuesta a impulsos, calculando las covarianzas cruzadas entre  $y_t$  y  $x_t$ , se obtendrá, como se vio en la sección anterior, que cada término  $\gamma_{xy}(k)$  está en función de todas las  $v_i$ , y además de la propia

estructura de autocorrelación de  $x_t$ . Para evitar eso, es necesario que la variable explicativa sea ruido blanco, para ello, se supone que  $x_t$  sigue un proceso ARMA de la forma

$$\phi_x(B)x_t = \theta_x(B)\alpha_t$$

donde  $\alpha_t$  es un proceso de ruido blanco, con variancia  $\sigma_\alpha^2$ , que se obtiene con:

$$\begin{aligned}\alpha_t &= \theta_x^{-1}(B)\phi_x(B)x_t \\ &= \psi^{-1}(B)x_t\end{aligned}$$

Entonces al aplicar el operador  $\Psi^{-1}(B)$  a los dos miembros de la ecuación con las series diferenciadas, se tiene que:

$$\Psi^{-1}(B)y_t = v(B)\Psi^{-1}(B)x_t + \Psi^{-1}(B)\eta_t$$

y llamando  $\beta_t = \Psi^{-1}(B)y_t$ ,  $\varepsilon_t = \Psi^{-1}(B)\eta_t$  entonces

$$\beta_t = v(B)\alpha_t + \varepsilon_t$$

con lo que se concluye que existe la misma función de transferencia entre  $\beta_t$  y  $\alpha_t$  que entre las series originales. La diferencia es que en esta última ecuación, la variable explicativa es ruido blanco. Por lo tanto, dados los resultados de la sección anterior se puede calcular  $v_k$  como,

$$v_k = \frac{\gamma_{\alpha\beta}(k)}{\sigma_\alpha^2}$$

y utilizando la función de correlación cruzada entre ambas series se tiene

$$v_k = \rho_{\alpha\beta}(k) \frac{\sigma_\beta}{\sigma_\alpha}$$

El procedimiento para la identificación de la función de transferencia que se describió es conocido como el *método de preblanqueo* de las series. Tiene el inconveniente que no puede extenderse cuando existe más de un regresor. Un modo alternativo que es conocido como *aproximación infinita*, se puede aplicar a todos los casos, consiste en estimar la ecuación considerando un número arbitrario  $h$  (entre 8 y 12) de rezagos de la variable  $x$ , esto es

$$y_t = v_0 x_t + v_1 x_{t-1} + \dots + v_h x_{t-h} + n_t$$

y con base a esta estimación identificar de manera sencilla los órdenes ( $m$ ,  $a$ ,  $b$ ) del proceso.

### 3.7 Identificación del Proceso de Inercia

Un procedimiento sencillo para identificar la función de transferencia consiste en calcular

$$\begin{aligned}\hat{N}_t &= Y_t - \hat{Y}_t \\ &= Y_t - \hat{v}_0 X_t - \hat{v}_1 X_{t-1} - \dots - \hat{v}_h X_{t-h}\end{aligned}$$

ya calculada la estimación para el proceso de inercia, utilizar la metodología del análisis univariadas de series de tiempo, para asignarle un modelo a  $\hat{N}_t$ .

### 3.8 Pruebas de Diagnóstico

Cuando se ha terminado la estimación de los parámetros, es necesario realizar contrastes diagnósticos para comprobar la adecuación del modelo estimado. Se elaboran contrastes sobre los parámetros estimados y sobre los residuales.

En las pruebas sobre los parámetros estimados debe comprobarse que no existen correlaciones altas entre ellos, pues esto sugeriría una mala definición de la situación de

estimación, y además se debe verificar que cada uno de ellos es significativamente diferente de cero.

Igual que en los modelos tradicionales de regresión, los supuestos sobre los residuales se deben verificar, es decir, éstos deben cumplir que siguen una distribución normal, que tienen media cero, varianza constante y que no hay correlación entre ellos para ningún retardo.

TESIS CON  
FALLA DE ORIGEN

## Capítulo 4

### APLICACIÓN AL CASO DE LOS CRUDOS.

#### 4.1 Especificación del modelo

El primer paso en la modelación es especificar las variables a incluir en el modelo, es decir, determinar cuales serán las variables explicativas o independientes que se involucrarán. Este paso no es del todo sencillo, es necesario encontrar las variables adecuadas que tengan algún tipo de relación con la variable dependiente, en el caso que concierne a este trabajo el cociente Maya/WTI, pero que entre ellas no exista una correlación alta.

Al especificar el modelo es importante no incurrir en la sobre especificación de éste, es decir, se deben tener involucradas únicamente a las variables que realmente aporten algo significativo al modelo. En un modelo es posible involucrar una gran cantidad de variables pero la presencia de algunas de ellas será innecesaria pues la información que pudieran aportar es irrelevante. Además, se debe tomar en cuenta que al involucrar demasiadas variables se puede caer en problemas de multicolinealidad, pues se consideraron dos variables entre las cuales existe una alta dependencia, y por tanto tener consideradas ambas resulta problemático.

Para comenzar con la modelación de la relación Maya/WTI es necesario investigar cuales variables serán convenientes para incluirlas en el modelo. Se consideran variables como: el consumo de crudo pesado en los Estados Unidos (CONSUMRESID), los insumos de las refinerías (REFINPUT), las importaciones de crudo, tanto las totales (IMPORT) como las provenientes exclusivamente de México (IMPORTMEX), el índice de la producción Estados Unidos (PRODIND), y también los costos de refinación, considerando en éstos el costo total (REFACQCO) y aquel relacionado sólo con el crudo importado (IMPOIL).

Sin embargo aún cuando estas variables parecen ser adecuadas, desde un punto de vista intuitivo, si se observa la matriz de correlaciones (tabla 4.1) de las variables antes mencionadas, con el objetivo de establecer una relación inicial entre ellas que permita conocer cuales serán las variables que mejor sirvan para modelar la relación, los resultados indican que las únicas variables que parecen ser más relevantes o estar más correlacionadas con la relación Maya/WTI son el insumo de las refinерías (REFINPUT); los costos de refinación, tanto para el crudo importado (IMPOIL) como para el crudo total (REFACQCO) y las importaciones de crudo (IMPORT).

Pero es importante notar que las series de los costos de refinación de los crudos importados y los totales arrojan los mismos resultados y por tanto su correlación es igual a 1, por tanto la inclusión de ambas variables en el modelo no es adecuada. También es importante notar que la correlación entre el insumo de las refinерías y las importaciones de crudo es de poco más de 0.95, por lo tanto será conveniente seleccionar solamente una de estas variables y no ambas.

**Tabla 4.1 Matriz de correlaciones (datos mensuales)**

	RELMAYAWTI	IMPOIL	REFACQCO	REFINPUT	PRODIND
RELMAYAWTI	1	0.5441 <sup>a</sup>	0.5674 <sup>a</sup>	-0.3848 <sup>b</sup>	-0.2894
IMPOIL	0.5441	1	1 <sup>c</sup>	-0.5857 <sup>b</sup>	-0.6710
REFACQCO	0.5674	1	1	-0.6226 <sup>b</sup>	-0.6739
REFINPUT	-0.3848	-0.5857	-0.6226	1	0.6646
PRODIND	-0.2894	-0.6710	-0.6739	0.6646	1
CAPUT	-0.0125	-0.3421	-0.3756	0.3837	0.2713
IMPORT	-0.4078	-0.6256	-0.6532	0.9545	0.9265
IMPORTMEX	-0.2731	-0.5897	-0.5860	0.7618	0.8957
IMPORTRESID	0.2499	0.7110	0.8957	-0.6207	-0.7492

	CAPUT	IMPORT	IMPORTMEX	IMPORTRESID
RELMAYAWTI	-0.0125	-0.4078 <sup>a</sup>	-0.2731	0.2499
IMPOIL	-0.3421	-0.6256 <sup>b</sup>	-0.5897	0.7110
REFACQCO	-0.3756	-0.6532 <sup>b</sup>	-0.5860	0.8957
REFINPUT	0.3837	0.9545 <sup>b</sup>	0.7618	-0.6207
PRODIND	0.2713	0.9265	0.8957	-0.7492
CAPUT	1	0.3225	0.0664	-0.2158
IMPORT	0.3225	1	0.8426	-0.7040
IMPORTMEX	0.0664	0.8426	1	-0.7429
IMPORTRESID	-0.2158	-0.7040	-0.7429	1

<sup>a</sup>Las variables que representan más alta correlación con la relación Maya/WTI

<sup>b</sup>Las variables con correlaciones más altas, entre las variables que están altamente correlacionadas con la relación.





Para examinar un poco más a fondo si estas variables serán significativas en la modelación, se probó para cada una de ellas un análisis de regresión simple, considerando además la información que aportan los gráficos de dispersión correspondientes. En la siguiente tabla se muestra un resumen de los datos obtenidos para cada una de las variables.

**Tabla 4.2 Regresiones Simples**

	R <sup>2</sup>	DW	T	Error estándar
IMPOIL	0.29802	0.22109	10.55609	0.00027
REFINPUT	0.14803	0.19413	-6.78566	0.08627
IMPORT	0.16633	0.19300	-7.27136	0.08533
D(IMPOIL)	0.01215	0.16726	1.80211*	0.00308
D(REFINPUT)	0.00963	0.18100	-1.60184*	0.09302
D(IMPORT)	0.01998	0.17949	-2.31410	0.09254
IMPOIL(-1)	0.27687	0.21843	10.05391	0.00027
REFINPUT(-1)	0.12792	0.21116	-6.22301	0.08729
IMPORT(-1)	0.14550	0.20501	-6.70463	0.08641
D(IMPOIL(-1))	0.02558	0.18853	2.62764	0.00310
D(REFINPUT(-1))	0.00261	0.16475	-0.82996*	0.09351
D(IMPORT(-1))	0.00759	0.17933	-1.41807*	0.09328

\*Resultados del estadístico t estadísticamente no significativos

Como es observable en los resultados de las regresión<sup>7</sup>, con la utilización de estas variables se obtienen coeficientes de determinación R<sup>2</sup> por debajo de 0.3, lo que es indicio de un mal ajuste, incluso, en los casos en los que se considera la variable diferenciada, ésta deja de ser estadísticamente significativa, de acuerdo con la información proporcionada por el estadístico t. Además es observable que los resultados correspondientes al estadístico Durbin-Watson (DW) muestran que en todos los casos existe el problema de autocorrelación. De la información correspondiente al error estándar, el cual nos proporciona una medida de dispersión de los coeficientes estimados con respecto a su media, es posible observar que este no es muy grande en ningún caso. Sin embargo, uno de los propósitos de este proyecto es tratar de vincular la relación de

<sup>7</sup> Más detalles sobre los modelos de regresión lineal simple se muestran en el anexo A2

los precios de crudo con alguna otra variable, de forma tal que permita explicar el comportamiento que sigue la relación de precios entre los crudos ligeros y los crudos pesados, por ello se considera que una de las variables más adecuadas para incluir en el modelo es el costo de refinación de los crudos importados, que para efectos prácticos se ha denotado como IMPOIL.

## 4.2 Análisis de las series explicativas

Uno de los primeros pasos para comenzar a trabajar con las series es verificar que éstas sean estacionarias. La importancia de la estacionariedad de las series es uno de los aspectos más importantes al trabajar con series de tiempo; si la serie en estudio no fuese estacionaria y se hiciera caso omiso de esto, entonces dado que la media o la varianza no son constantes o las covarianzas de las series no dependen sólo del tiempo entre observaciones, esto repercute en los estadísticos t y F, los cuales dejarían de ser válidos.

Tabla 4.3 Correlograma IMPOIL

Muestra: 1980:01 2002:12

Observaciones incluidas: 267

Autocorrelación	Correlación Parcial	AC	PAC	Q-Stat	Prob
1	0.986	0.986	262.24	0.000	
2	0.964	-0.253	514.08	0.000	
3	0.940	-0.025	754.64	0.000	
4	0.917	0.029	984.40	0.000	
5	0.896	0.034	1204.3	0.000	
6	0.877	0.064	1416.0	0.000	
7	0.860	0.010	1620.3	0.000	
8	0.844	-0.013	1817.7	0.000	
9	0.826	-0.049	2007.8	0.000	
10	0.808	-0.009	2190.4	0.000	
11	0.789	-0.043	2385.0	0.000	
12	0.768	-0.035	2531.1	0.000	
13	0.744	-0.097	2687.8	0.000	
14	0.722	0.074	2835.8	0.000	
15	0.702	0.042	2976.2	0.000	
16	0.684	0.030	3110.2	0.000	
17	0.670	0.051	3239.0	0.000	
18	0.656	-0.019	3363.1	0.000	
19	0.642	-0.014	3482.7	0.000	
20	0.628	-0.006	3597.5	0.000	
21	0.613	-0.007	3707.3	0.000	
22	0.598	-0.020	3812.0	0.000	
23	0.581	-0.035	3911.3	0.000	
24	0.564	0.025	4005.5	0.000	

Del correlograma correspondiente al IMPOIL se observa que la autocorrelación disminuye lentamente conforme el rezago aumenta, además de que la autocorrelación parcial en el primer rezago es altamente significativa. De lo anterior se concluye que la serie en cuestión muy probablemente sea no estacionaria, para verificar la veracidad de esta observación se aplica la *prueba de la raíz unitaria*. Al correr la prueba, sin considerar la tendencia ni la constante, se obtienen resultados contrapuestos, ya que el valor del estadístico  $t$  es ligeramente mayor para los valores críticos correspondientes al 5% y al 10%, lo cual propone estacionariedad en la serie, pero para el valor correspondiente al 1% el valor es menor, en cuyo caso la serie es no estacionaria. Ahora al considerar únicamente la constante, como se muestra en la tabla 4.4, los resultados indican que la serie es no estacionaria.

**Tabla 4.4 Prueba de la raíz unitaria del IMPOIL**

Hipótesis Nula: IMPOIL tiene una raíz unitaria

Exógena: Constante

Longitud Rezag: 2 (Automático, Rez. Máx. =15)

	t-Estadístico	Prob.*
Prueba Estadística Dickey-Fuller Aumentada	-2.400957	0.1425
Valores críticos:		
1% nivel	-3.455001	
5% nivel	-2.872288	
10% nivel	-2.572570	

\*Mackinnon (1996)

Prueba Dickey-Fuller Aumentada

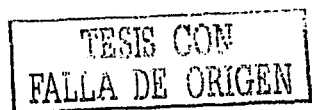
Variable Dependiente: D(IMPOIL)

Método: Mínimos Cuadrados

Muestra (ajustada): 1980:04 2002:03

Observaciones incluidas: 264 después de ajustar los extremos

Variable	Coefficiente	Error Est.	t-Estadístico	Prob.
IMPOIL(-1)	-0.013434	0.005595	-2.400957	0.0171
D(IMPOIL(-1))	0.558910	0.061474	9.091851	0.0000
D(IMPOIL(-2))	-0.144526	0.061234	-2.360213	0.0190
C	0.322037	0.204274	1.576494	0.1161
R-cuadrada	0.270800	Media var. depend.	-0.194387	
R-cuadrada ajustada	0.262387	D.S. var. depend.	1.847845	
E. S. de regresión	1.587011	Akaike criterio	3.778617	
Suma resid. Cuad.	654.8369	Schwarz criterio	3.830798	
Log probabilidad	-494.5135	F-estadístico	32.18511	
Durbin-Watson	1.980731	Prob (F-estadístico)	0.000000	



Sin embargo, los resultados obtenidos rechazan la inclusión de la constante en la prueba y al no considerarla, se obtienen, como ya se mencionó antes, resultados en los cuales la estacionariedad es válida a niveles aceptables de confiabilidad (5 y 10%) pero no lo es en un nivel más exigente como lo es el 1%. Concluir que la serie es no estacionaria con los resultados obtenidos, sería como tratar de forzar la serie para concluir que ésta es no estacionaria. Pero aun a pesar de los resultados numéricos obtenidos en la prueba, el comportamiento que se presenta en el correlograma es propio de una serie no estacionaria.

Por lo tanto, ahora se debe decidir entre la necesidad de diferenciar o no la serie para volverla estacionaria, ya que, si bien es cierto que con la inclusión de la constante los resultados indican que es necesario diferenciar, el estadístico  $t$  correspondiente a la constante es poco significativo. Al diferenciar la serie se asegura que ésta es estacionaria, sin embargo, con esto la correlación entre IMPOIL diferenciado y la relación entre los crudos disminuye significativamente, además de que es importante remarcar que al diferenciar las series puede provocar la pérdida de información valiosa sobre la relación entre las variables en el largo plazo. En casos como éste en el que la necesidad de diferenciación no es muy evidente, lo más recomendable es diferenciar pues se incurre en menos errores, de los que se cometerían al no diferenciar una serie.

Por otro lado, aun cuando la serie es la más adecuada de todas, el valor de  $R^2$  es muy bajo y además el estadístico Durbin-Watson es muy pequeño, lo cual lleva a suponer que la serie está autocorrelacionada. Esta suposición también se verifica al examinar el correlograma correspondiente a la relación, en donde la autocorrelación parcial del primer rezago es altamente significativa y por lo tanto se concluye que es necesario utilizar en la modelación los procesos autorregresivos.

Más aún si después de haber incluido en el modelo dichas variables los resultados no son del todo satisfactorios entonces es posible que existan variables que han sido excluidas del modelo y que estén alterando el comportamiento aleatorio de los residuales, a esto se le conoce como *proceso de inercia*, si esto último fuese cierto entonces se estaría considerando una regresión dinámica o función de transferencia.

Al tomar un proceso autorregresivo de orden 1 en la modelación de la relación los resultados obtenidos se mejoran significativamente, el coeficiente de determinación es alto superando el 0.84 y el valor del estadístico Durbin-Watson aumenta aproximándose a 2. Sin embargo, si se considera un proceso AR(2), éste incrementa el coeficiente  $R^2$ , pero con un incremento muy pequeño, alcanzando un valor poco mayor a 0.85. El cambio realmente significativo se presenta en el incremento en el Durbin-Watson el cual pasa de 1.67 a 1.96, solucionando el problema de la autocorrelación satisfactoriamente.

Hasta este momento se ha establecido que la relación entre los precios de los crudos es una serie que se explica a si misma en un porcentaje alto, sin embargo dado que los objetivos del proyecto consisten en vincular el comportamiento con otra variable, es necesario tratar de explicar la variación que no ha sido explicada por la misma serie con otra variable.

De los resultados anteriores, se establece que lo más conveniente es modelar la relación de los precios de crudos en función del precio del WTI y el costo de refinación de los crudos importados (IMPOIL).

$$(Maya/WTI)_t = v_1(B) WTI_t + v_2(B) IMPOIL_t + N_t$$

### 4.3 Identificación de las funciones de transferencia

Lo que se realiza, ahora, es un análisis de las series para determinar la función de transferencia de cada una de ellas y después juntarlas en el modelo global. Es decir, se analizará por separado cada una de las variables explicativas para que utilizando el método del preblanqueo de las series, explicado en el capítulo anterior, sea posible determinar la estructura correspondiente a cada una de las funciones de transferencia,  $v_i(B)$ .

Antes de comenzar a trabajar con las series para ajustar un modelo de regresión dinámica es importante verificar los principios en los que se basa la construcción de un modelo de regresión dinámica. Uno de dichos principios indica que debe existir una

relación de causalidad unidireccional entre las variables explicativas y la dependiente, es decir, en el caso específico de los combustibles se debe verificar que el IMPOIL influye sobre la relación Maya/WTI pero no al contrario, y análogamente con el WTI.

Desafortunadamente, éste es uno de los problemas en los que se incurre al tratar de modelar la relación de los precios de los crudos bajo esta metodología. Al calcular las correlaciones cruzadas entre las variables, por ejemplo, el IMPOIL y la relación MAYA/WTI (tabla 4.5), éstas son significativas para rezagos negativos, específicamente para el rezago de orden -1, lo cual no debe ocurrir si la relación de causalidad solamente ocurre de manera unidireccional, es decir, cuando el IMPOIL influye sobre la relación y no al contrario.

**Tabla 4.5 Correlaciones Cruzadas D(IMPOIL) y D(RELMAYAWTI)**

Muestra: 1990:01 2002:12

Observaciones Incluidas: 266

Correlaciones son aproximaciones consistentes asintóticamente

D(IMPOIL),D(RELMAYAWTI)(-)	D(IMPOIL),D(RELMAYAWTI)(+)	i	Lag	lead
.***	.***	0	0.2032	0.2032
.***	.**	1	0.3492	0.1272
.*	.**	2	0.1028	0.1087
.**	.-	3	0.0873	-0.0290
.**	.***	4	0.0641	-0.1908
.*	.***	5	-0.0683	-0.1573
.*	.*	6	-0.0686	-0.0593
.*	.-	7	-0.0993	-0.0396
.*	.-	8	-0.0444	0.0271
.-	.-	9	-0.0030	0.0317
.-	.-	10	0.0997	-0.0022
.**	.**	11	0.1792	0.0745
.***	.**	12	0.1632	-0.0971
.-	.*	13	0.0084	-0.1140
.*	.***	14	-0.1122	-0.1665
.*	.***	15	-0.1273	-0.1853
.*	.-	16	-0.1186	-0.0357
.*	.-	17	-0.0478	-0.0021
.*	.**	18	-0.0764	0.0503
.-	.**	19	-0.0315	0.1130
.-	.-	20	-0.0023	0.0365

Es importante notar que al estar trabajando con las series anteriores, se consideraron las diferencias en ambas, esto debido principalmente al problema sobre la necesidad de diferenciar o no la serie correspondiente al IMPOIL; y además el correlograma de correlaciones cruzadas de las series originales presenta una estructura muy complicada sugiriendo con ello la necesidad de diferenciar.

Así, de manera gráfica, parece haber un problema importante debido a la mala causalidad, pero dado que los métodos gráficos son siempre cuestionables es más conveniente utilizar una prueba estadística, en este caso se utilizó la *prueba de Granger*<sup>8</sup>, la cual sirve para determinar si la causalidad es unidireccional y en que sentido, si existe causalidad bilateral, o bien si las series son independientes. De la prueba aplicada sobre las series IMPOIL y RELMAYAWTI, se obtienen los resultados mostrados en la tabla 4.6.

**Tabla 4.6 Prueba de causalidad Granger RELMAYAWTI & IMPOIL**

Muestra: 1980:01 2002:03

Rezagos: 5

Hipótesis Nula:	Obs.	F-Estadístico	Prob.
RELMAYAWTI no Granger causa IMPOIL	262	8.28407	2.9E-07
IMPOIL no Granger causa RELMAYAWTI		3.87319	0.00213

Estos resultados confirman lo ya antes observado en la gráfica de las correlaciones cruzadas, existe causalidad en el sentido contrario al esperado, de manera significativa, es decir, RELMAYAWTI causa a IMPOIL; pero también se da la causalidad en el otro sentido, por lo tanto se tiene una retroalimentación o causalidad bilateral. Cabe destacar que al utilizar estas pruebas se consideraron las series originales y no diferenciadas, pero al diferenciarlas los resultados obtenidos coinciden con los obtenidos con las series originales, los resultados se muestran en la tabla 4.7

**Tabla 4.7 Prueba de causalidad Granger D(RELMAYAWTI) & D(IMPOIL)**

Muestra: 1980:01 2002:03

Rezagos: 5

Hipótesis Nula:	Obs.	F-Estadístico	Prob.
D(RELMAYAWTI) no Granger causa D(IMPOIL)	261	7.70107	9.4E-07
D(IMPOIL) no Granger causa D(RELMAYAWTI)		3.22908	0.00783

<sup>8</sup> Referencia: *Econometría Básica*, Damodar N. Gujarati

Este problema encontrado en la causalidad de las variaciones en las series es de suma importancia, si se hiciera caso omiso de esta violación y se continuara el análisis de las series de acuerdo con la teoría, el trabajo realizado se detendría en el momento de tratar de identificar la estructura  $v_t(B)$ .

Si se considera que la causalidad es bilateral, pero con la observación de que la causalidad en el sentido IMPOIL causa RELMAYAWTI es menos significativa que en el otro sentido, sería posible continuar con el proceso para encontrar los parámetros de la función de transferencia. Si se considera entonces la posibilidad, lo primero por hacer es encontrar el mejor modelo univariado para la serie del IMPOIL.

El modelo se construye para la serie diferenciada y se elabora el correlograma correspondiente. De éste se identifica que los rezagos de orden uno, dos y cuatro son significativos. Y después de varias alternativas se obtiene como un modelo adecuado, el modelo autorregresivo de orden dos, los resultados obtenidos se muestran en la tabla 4.8

**Tabla 4.8 Modelo univariado para el IMPOIL (ARIMA (2,1,0))**

Variable Dependiente: D(IMPOIL)

Método: Mínimos Cuadrados

Muestra (ajustada): 1980:04 2002:03

Observaciones incluidas: 264 después de ajustar extremos

Convergencia alcanzada después de 3 iteraciones

Variable	Coficiente	Error Std.	t-Estadístico	Prob.
AR(1)	0.571992	0.061806	9.254705	0.0000
AR(2)	-0.148793	0.061592	-2.415775	0.0164
R-cuadrada	0.251283	Media var. dependiente		-0.194387
R-cuadrada ajustada	0.248425	D.S. var. dependiente		1.847845
E.S. de regresión	1.601960	Akaike info. criterio		3.787880
Suma resid. cuad.	672.3642	Schwarz criterio		3.814970
Log probabilidad	-498.0001	Durbin-Watson		1.990388
AR Raíces invert.	.29 -.26i	.29+.26i		

Una vez obtenido este ajuste se deben calcular los residuales correspondientes que formarán la variable de ruido blanco  $\alpha_t$  y después utilizando los parámetros del

TESIS CON  
FALLA DE ORIGEN



modelo se obtienen los valores para  $\beta_i$ , aplicando los parámetros en la serie de la RELMAYAWTI. El polinomio  $\psi^{-1}(B)$  que se utilizará es

$$\psi^{-1}(B) = (1 - 0.571992 B + 0.148793 B^2)$$

Después de aplicar el modelo a ambas series y una vez obtenidas las nuevas,  $\alpha_i$  y  $\beta_i$ , debe ser posible realizar la identificación inicial de los parámetros de la función de transferencia. Para ello se calculan las correlaciones cruzadas para estas nuevas variables y se elabora el correlograma cruzado correspondiente. El gráfico que muestra el comportamiento de las correlaciones cruzadas para la identificación de la función de transferencia entre IMPOIL y RELMAYAWTI, se muestra en la tabla 4.9.

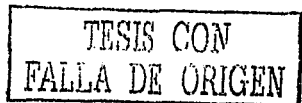
**Tabla 4.9 Correlograma Cruzado ALFA, BETA**

Muestra: 1980:01 2002:03

Observaciones Incluidas: 264

Correlaciones son aproximaciones asintóticamente consistentes

ALFA, BETA(-)	ALFA, BETA(+)	i	lag	Lead
. .	. .	0	-0.0003	-0.0003
. .	. .	1	0.3291	0.0176
. .	. .	2	-0.1067	0.0984
. .	. .	3	0.0288	0.0050
. .	. .	4	0.0994	-0.1343
. .	. .	5	-0.0865	-0.0684
. .	. .	6	-0.0125	0.0225
. .	. .	7	-0.0628	-0.0621
. .	. .	8	0.0105	0.0413
. .	. .	9	-0.0285	0.0326
. .	. .	10	0.0372	-0.0930
. .	. .	11	0.1028	0.1544
. .	. .	12	0.0936	-0.1320
. .	. .	13	-0.0213	-0.0028
. .	. .	14	-0.0738	-0.0425
. .	. .	15	-0.0237	-0.1432
. .	. .	16	-0.0851	0.0456
. .	. .	17	0.0277	-0.0182
. .	. .	18	-0.0710	0.0083
. .	. .	19	-0.0170	0.0965
. .	. .	20	0.0484	-0.0181
. .	. .	21	-0.0728	-0.0278
. .	. .	22	0.0452	0.0685
. .	. .	23	0.0305	-0.0595
. .	. .	24	-0.0782	0.0756
. .	. .	25	0.0361	-0.0335



Como se ignoró el principio básico de la causalidad, no es posible establecer una identificación inicial de los parámetros de la función de transferencia, a partir del correlograma cruzado. El problema de la causalidad entre las series, no permite realizar la identificación, ya que para rezagos positivos las correlaciones son nulas, es decir, estadísticamente no son significantes. Pero debe observarse que para el rezago negativo de orden 1, la correlación es significativa. Teóricamente cuando los supuestos del modelo se cumplen, esto no debe ocurrir, pues en los modelos de funciones de transferencia se supone que para los rezagos negativos la función de correlación cruzada es nula. Por lo tanto no es posible establecer los parámetros de la función de transferencia.

Es entonces necesario conseguir una variable que tenga una influencia adecuada sobre la variable dependiente, sin embargo esto no es sencillo. La variable dependiente con la que se está trabajando, es decir, la relación dada por el cociente de los precios de crudos, es una variable que como ya se vio antes se explica a sí misma, esto quiere decir que es una serie altamente autocorrelacionada, y por lo tanto la información que pudiese aportar alguna otra serie sería mínima.

Se intentó con algunas de las otras variables con las que se contaba, para tratar de aplicar la metodología de las funciones de transferencia. Pero, los resultados obtenidos no son satisfactorios, al igual que como ocurrió con el IMPOIL, las causalidades no cumplen con el supuesto de la regresión dinámica. Los resultados obtenidos al aplicar la prueba de causalidad de Granger, se muestran en la tabla 4.10, de ésta, se observa que o bien la causalidad es bilateral, o se da en el sentido o contrario, o ni siquiera existe causalidad entre ellas. Como ejemplos se pueden comentar los resultados obtenidos al usar la serie del WTI junto con la serie de la relación, en donde la relación es bilateral; de manera opuesta si consideramos la relación con el consumo de residuales, entonces de los resultados se concluye que no existe ningún tipo de causalidad entre dichas series; y de otro de los casos que se dan es cuando sí existe causalidad unidireccional pero ésta no cumple con los principios teóricos necesarios, como ocurre al considerar la serie de las importaciones (IMPORT) igual que con el IMPOIL.



**Tabla 4.10 Pruebas de causalidad**

Muestra: 1980:01 2002:03	Rezagos:5	Obs.	F-Estadístico	Probabilidad
Hipótesis Nula:				
WTI no Granger Causa RELMAYAWTI		262	3.06539	0.01051
RELMAYAWTI no Granger Causa WTI			11.2636	8.2E-10
REFACQCO no Granger Causa RELMAYAWTI		262	3.83986	0.00227
RELMAYAWTI no Granger Causa REFACQCO			10.3727	4.7E-09
REFINPUT no Granger Causa RELMAYAWTI		262	1.85516	0.10275
RELMAYAWTI no Granger Causa REFINPUT			1.15233	0.33334
IMPORT no Granger Causa RELMAYAWTI		262	2.27226	0.04797
RELMAYAWTI no Granger Causa IMPORT			3.94156	0.00185
IMPORTMEX no Granger Causa RELMAYAWTI		262	0.99885	0.41898
RELMAYAWTI no Granger Causa IMPORTMEX			3.17077	0.00855
PRODRESID no Granger Causa RELMAYAWTI		262	0.53774	0.74760
RELMAYAWTI no Granger Causa PRODRESID			1.06993	0.37752
IMPORTRESID no Granger Causa RELMAYAWTI		262	0.86036	0.50836
RELMAYAWTI no Granger Causa IMPORTRESID			0.95481	0.44620
CONSUMRESID no Granger Causa RELMAYAWTI		262	0.61793	0.68625
RELMAYAWTI no Granger Causa CONSUMRESID			0.70148	0.62280
PRODIND no Granger Causa RELMAYAWTI		262	0.90624	0.47754
RELMAYAWTI no Granger Causa PRODIND			1.01971	0.40649
CAPUT no Granger Causa RELMAYAWTI		262	1.17381	0.32250
RELMAYAWTI no Granger Causa CAPUT			0.97249	0.43514

La obtención de la variable adecuada para la modelación utilizando la regresión dinámica es complicada, aunque esto no significa que sea imposible. En muchos casos la utilización de modelos complicados es adecuado, pero para otros no es así. Modelos sencillos son buenos en el sentido que proporcionan buenos ajustes. En este caso particular de la relación Maya/WTI, un modelo univariado sencillo es un ajuste bueno y no requiere de mayores complicaciones. Con esto no se está sugiriendo que un modelo univariado sea el mejor modelo para la relación, lo que se concluye es que será un modelo adecuado, en función de los resultados estadísticos obtenidos y el ajuste que se obtenga.

Si se trabaja para encontrar un modelo univariado de la relación Maya/WTI entonces de acuerdo con el correlograma correspondiente, la correlación parcial para el rezago de orden uno es altamente significativa estadísticamente y la correspondiente al rezago de orden dos es estadísticamente significativa pero en menor proporción. Por lo tanto, una posibilidad es modelar la relación con un proceso autorregresivo de orden dos, como ya se había mencionado anteriormente. Al correr la regresión utilizando este modelo se obtienen resultados adecuados, como se muestra en la tabla 4.11, se tiene un  $R^2$  grande y estadísticos significativos.

**Tabla 4.11 Modelo AR(2) para MAYA/WTI**

Variable Dependiente: RELMAYAWTI

Método: Mínimos Cuadrados

Muestra (ajustada): 1980:03 2002:03

Observaciones incluidas: 265 after adjusting endpoints

Convergencia después de 3 iteraciones

Variable	Coefficiente	Error Std.	t-Estadístico	Prob.
C	0.729579	0.023843	30.59981	0.0000
AR(1)	1.074809	0.060756	17.68731	0.0000
AR(2)	-0.168175	0.060651	-2.772825	0.0060
R-cuadrada	0.850178	Media var. Dependiente	0.728726	
R-cuadrada ajustada	0.849034	D.S. var. Dependiente	0.093460	
E.S. de la regresión	0.036313	Akaike criterio	-3.782016	
Suma resid. cuad.	0.345485	Schwarz criterio	-3.741491	
Log likelihood	504.1172	F-estadístico	743.3689	
Durbin-Watson	1.983351	Prob (F-estadístico)	0.000000	
AR Raíces invert.	.66	.19		

En el gráfico 4.1, se muestra el ajuste realizado con un modelo autorregresivo de orden dos, y además se muestra la gráfica de los residuales contra el tiempo. Es posible observar que el ajuste es adecuado y los residuales no tienen un comportamiento que sugiera problemas de autocorrelación o alguna violación a los supuestos del modelo de regresión lineal.

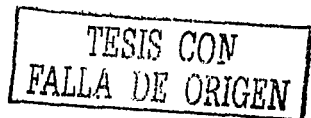
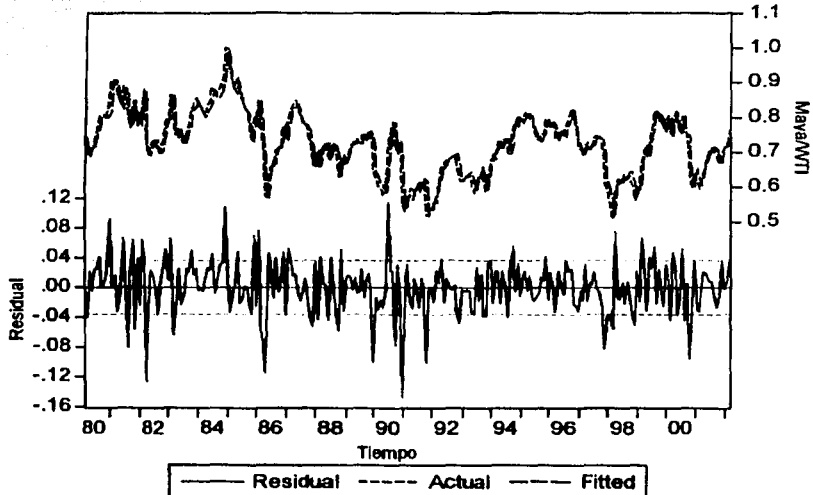
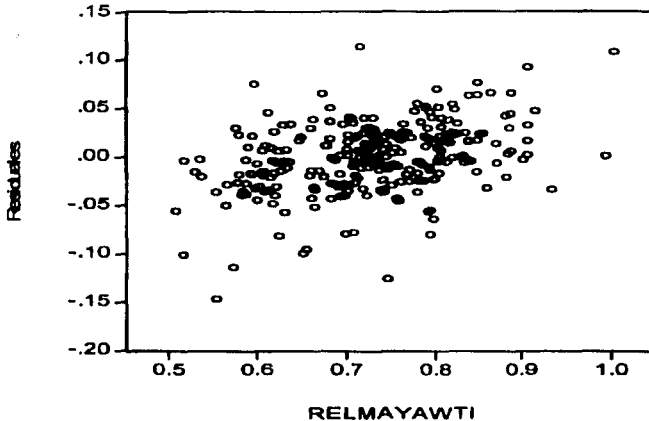


Gráfico 4.1 Ajuste AR(2) y Residuales del modelo univariado de MAYA/WTI



Para verificar que éste sea un modelo adecuado es necesario estudiar el comportamiento de los residuales del modelo. A primera vista, los residuales obtenidos no presentan algún tipo de violación a los supuestos, es decir, se puede considerar que siguen un proceso de ruido blanco. Por lo tanto si se tratara de mejorar el modelo utilizando la segunda parte de la teoría de regresión dinámica, es decir, ajustando un modelo ARIMA a los residuales obtenidos, resultaría inútil pues estos no siguen un comportamiento determinado, el cual pueda ser representado con algún modelo de este tipo, parece pues que la variación que no es explicada por la misma serie es puramente aleatoria.

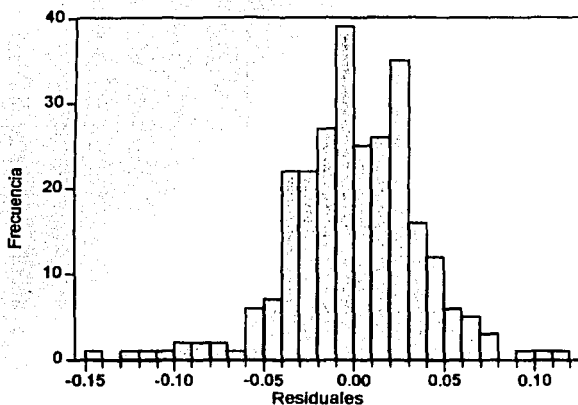
Gráfico 4.2 RELMAYAWTI &amp; Residuales del modelo AR(2)



Si se estudia con más detalle los residuales de este modelo es posible observar que, como lo muestra el gráfico 4.2, no presentan violación alguna al supuesto de la varianza constante, ya que el gráfico no muestra la formación de figuras claras que sugieran problemas de heteroscedasticidad. Además se puede observar que los residuales se distribuyen alrededor del cero de manera dispersa.

Ahora para verificar la normalidad de los residuales, es de utilidad graficar su histograma el cual se muestra en el gráfico 4.3, de éste se concluye que el comportamiento de los residuales es normal, pero presenta un exceso en la curtosis y además no es del todo simétrico. En términos generales se puede decir que estos siguen una distribución normal, aun cuando el estadístico Jarque Bera es poco significativo pues la probabilidad relacionada con éste es casi nula.

Gráfico 4.3 Histograma Residuales del modelo AR(2)



Serie: Residuales	
Muestra: 1980:01 2002:03	
Observaciones: 265	
Media	3.20E-15
Mediana	-0.000203
Máximo	0.113938
Mínimo	-0.147084
Dev. Est.	0.036175
Asimetría	-0.440391
Curtosis	4.804253
Jarque-Bera	44.51014
Probabilidad	0.000000

Al analizar el correlograma de los residuales no se presenta ningún tipo de estructura con lo que se concluye que no existe correlación entre ellos, y por lo tanto estos son independientes unos de otros. Lo anterior permite concluir que el modelo es adecuado pues los residuales siguen un proceso de ruido blanco y los estadísticos obtenidos son significativos.

Ahora, el interés principal es encontrar un modelo de la relación que explique el comportamiento de los ciclos de ésta, es decir, sus periodos de crecimiento y de decrecimiento en los cuales los precios de los crudos ligeros y pesados se acercan y se alejan. Dado que la mayoría de las variables con las que se ha trabajado son variables relacionadas hasta cierto punto con el mercado petrolero, es posible que la inclusión de otro tipo de variables resulte más adecuada para la modelación.

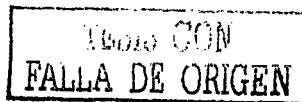
Variabes que proporcionen información sobre cuales han sido los cambios en la economía que han afectado o repercutido en los cambios de los precios de crudo, como por ejemplo la demanda de crudo mundial o los cambios en los procesos de refinación, o

también los posibles impactos originados por sucesos históricos como guerras o cambios en la oferta mundial de petróleo, son variables que pueden ser estadísticamente útiles.

**Tabla 4.12 Correlograma Residuales del modelo AR(2)**  
 Muestra: 1980:03 2002:03  
 Observaciones Incluidas: 285  
 Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelación	Correlación Parcial	AC	PAC	Q-Est.	Prob
. .	. .	1	0.007	0.007	0.0116
. .	. .	2	-0.053	-0.053	0.7663
. .	. .	3	0.042	0.043	1.2520 0.263
. .	. .	4	-0.014	-0.018	1.3048 0.521
. .	. .	5	-0.002	0.003	1.3055 0.728
. .	. .	6	-0.029	-0.032	1.5290 0.821
* .	* .	7	-0.088	-0.087	3.6638 0.599
. .	. .	8	0.014	0.012	3.7186 0.715
. *	. *	9	0.111	0.105	7.1267 0.416
. .	. .	10	0.027	0.034	7.3323 0.501
. .	. .	11	0.039	0.047	7.7553 0.559
. *	. *	12	0.097	0.092	10.398 0.406
. .	. .	13	0.005	0.005	10.406 0.494
. .	. *	14	0.064	0.066	11.552 0.482
* .	* .	15	-0.106	-0.108	14.735 0.324
* .	* .	16	-0.104	-0.079	17.819 0.215
* .	* .	17	-0.071	-0.088	19.242 0.203
* .	* .	18	-0.073	-0.079	20.753 0.188
. *	. *	19	0.101	0.111	23.660 0.129
. .	. .	20	-0.020	-0.030	23.775 0.163

La metodología de regresión dinámica requiere que las variables utilizadas estén relacionadas, pero que además esa relación tenga implícita una causalidad unidireccional y que ésta sea en el sentido adecuado, por tanto para su aplicación en el caso de los crudos es necesario encontrar la variable adecuada. Se ha intentado con algunas variables, pero aún a pesar de que la relación que existe entre éstas y el cociente Maya/WTI, es pequeña, dichas variables no parecen ser útiles para la aplicación de la regresión dinámica.





Sin embargo, aún cuando la conclusión de que la relación entre los precios de crudos es una serie que no requiere de otras variables para explicarse, es decir, que se explica a sí misma con sus valores rezagados, es necesario encontrar un modelo que vincule el comportamiento de los precios de los crudos con una variable, pues al trabajar con los datos anuales la inclusión de procesos autorregresivos resulta un tanto inadecuada y por lo tanto es necesario explicarlo de manera diferente.

Otro de los problemas que se presentan es el análisis de las series anuales. Este análisis es necesario ya que la información necesaria para la elaboración de los pronósticos, se obtiene para datos anuales y por lo tanto su estudio es de importancia. Sin embargo, los datos con los que se cuenta son relativamente pocos, se cuenta únicamente con 22 datos que comprenden desde 1980 hasta el 2001, el pequeño tamaño de la muestra puede repercutir en los estimadores, los cuales no permanecerán constantes durante la estimación, es decir, podrían cambiar significativamente al cambiar el período muestral o bien al agregar más datos. Para la mayoría de las series utilizadas es posible anexar datos para ampliar el tamaño de la muestra, pero particularmente para el precio del crudo Maya, la información con la que se cuenta inicia en enero de 1980 y esto impide trabajar con una muestra más grande.

Para comenzar a estudiar las series anuales se analiza el comportamiento de éstas, así como las correlaciones entre las variables a considerar. Al analizar el correlograma de los datos anuales se observa que las correlaciones son pequeñas para todos los rezagos, no obstante la correlación parcial para el primer rezago es relativamente grande y por lo tanto es necesario tener cuidado con este aspecto, aunque en realidad el valor es pequeño y por lo tanto no forzosamente indica la existencia de correlación.

De la matriz de correlaciones (tabla 4.14) se concluye que, igual como ocurría con los datos mensuales, la serie de la relación está altamente vinculada con el IMPOIL, el REFACQCO y con el WTI, pero estos también tienen correlaciones altas entre ellos y por lo tanto no es conveniente la inclusión simultánea de dos o más de ellas.

**Tabla 4.13 Correlograma Relación Maya/WTI (datos anuales)**

Muestra: 1980 2002

Observaciones incluidas: 22

Auto correlación	Correlación Parcial	AC	PAC	Q-Stat	Prob	
		1	0.477	0.477	5.7293	0.017
		2	0.247	0.025	7.3388	0.025
		3	0.188	0.080	8.3235	0.040
		4	0.045	-0.101	8.3821	0.079
		5	-0.018	-0.023	8.3902	0.136
		6	-0.143	-0.165	9.0608	0.170
		7	-0.162	-0.028	9.9903	0.189
		8	-0.175	-0.077	11.150	0.193
		9	-0.219	-0.086	13.080	0.159
		10	-0.093	0.091	13.468	0.189
		11	-0.064	-0.024	13.665	0.252
		12	-0.057	-0.019	13.836	0.311
		13	-0.095	-0.126	14.368	0.349
		14	-0.125	-0.081	15.390	0.352
		15	0.032	0.124	15.468	0.418
		16	-0.050	-0.134	15.686	0.475
		17	-0.170	-0.184	18.730	0.344
		18	-0.069	0.048	19.358	0.370
		19	0.019	0.100	19.420	0.430
		20	-0.054	-0.154	20.178	0.447

Todas las regresiones lineales realizadas con estas variables dan como resultado valores para el coeficiente de correlación menores a 0.75 y para el estadístico Durbin Watson valores pequeños que caen en el intervalo de indecisión sobre la existencia de correlación. Una manera rápida de eliminar este problema es la inclusión de un proceso autorregresivo, pero este aspecto no es del todo adecuado cuando se trabaja con series anuales.

**Tabla 4.14 Matriz de correlaciones (datos anuales)**

	MAYA/WTI	WTI	IMPOIL	REFACQCO	REFINPUT
MAYA/WTI	1.000	0.580	0.627	0.634	-0.472
WTI	0.560*	1.000	0.930 <sup>#</sup>	0.867 <sup>#</sup>	-0.600
IMPOIL	0.627*	0.930	1.000	0.986 <sup>#</sup>	-0.478
REFACQCO	0.634*	0.867	0.986	1.000	-0.469
REFINPUT	-0.472	-0.600	-0.478	-0.469	1.000
CAPUT	-0.035	-0.410	-0.450	-0.486	0.458
PRODRESID	0.234	0.825	0.616	0.495	-0.453
CONGRESID	0.348	0.877	0.673	0.566	-0.591
IMPORTMEX	0.006	0.070	0.266	0.303	0.511
IMPORTRESID	0.380	0.814	0.633	0.561	-0.743

\*Las variables que representan más alta correlación con la relación Maya/WTI

<sup>#</sup>Las variables con correlaciones más altas, entre las variables que están altamente correlacionadas con la relación.

Tabla 4.14 (Continuación) Matriz de correlaciones

	CAPUT	PRODRESID	CONGRESID	IMPORTMEX	IMPORTRESID
MAYAWTI	-0.035	0.234	0.348	0.006	0.380
WTI	-0.410	0.825	0.877	0.070	-0.814
IMPOIL	-0.450	0.816	0.873	0.266	0.633
REFACOCO	-0.486	0.495	0.566	0.303	0.561
REFINPUT	0.458	-0.453	-0.591	0.511	-0.743
CAPUT	1.000	-0.103	-0.218	-0.325	-0.343
PRODRESID	-0.103	1.000	0.962	-0.186	0.815
CONGRESID	-0.218	0.962	1.000	-0.174	0.930
IMPORTMEX	-0.325	-0.186	-0.174	1.000	-0.215
IMPORTRESID	-0.343	0.815	0.930	-0.215	1.000

Para estudiar la relación de los precios en datos anuales, se realiza una regresión simple entre las series de los precios (tabla 4.15). En dicha regresión se considera como variable dependiente a la serie del precio del Maya en datos anuales, y la variable independiente la serie del WTI. Los resultados obtenidos indican una alta correlación entre ambos, sin embargo el valor del estadístico Durbin Watson está por debajo del límite superior  $d_u$ , y por lo tanto es posible que exista presencia de autocorrelación en los residuales.

Del gráfico 4.4, donde se presenta la serie ajustada junto con la serie real, se observa que el ajuste es bueno porque sigue el comportamiento de la serie de modo muy general. Sin embargo, se puede observar en el gráfico de los residuales, que éstos no son del todo ruido blanco, parece que siguen un comportamiento definido, lo cual es reflejo del valor pequeño del estadístico Durbin Watson. Pero al referirse estrictamente a los valores superior e inferior del estadístico con un nivel de significancia  $\alpha = 0.05$ , éstos valen 1.541 y 1.147 respectivamente, por tanto el valor obtenido para la regresión cae dentro de la región de indecisión, pero está demasiado próximo al valor del límite superior, esto sugiere que los residuales considerando un nivel de significancia menor, permitirían que el valor del estadístico caiga en la región de no autocorrelación. Por ejemplo si se considera  $\alpha = 0.01$  los valores correspondientes son  $d_u = 1.284$  y  $d_L = 0.914$ , es claro ahora que el valor es mayor a  $d_u$  y por lo tanto no hay evidencia de correlación.

**Tabla 4.15 Regresión Maya & WTI (datos anuales)**

Variable Dependiente: MAYA

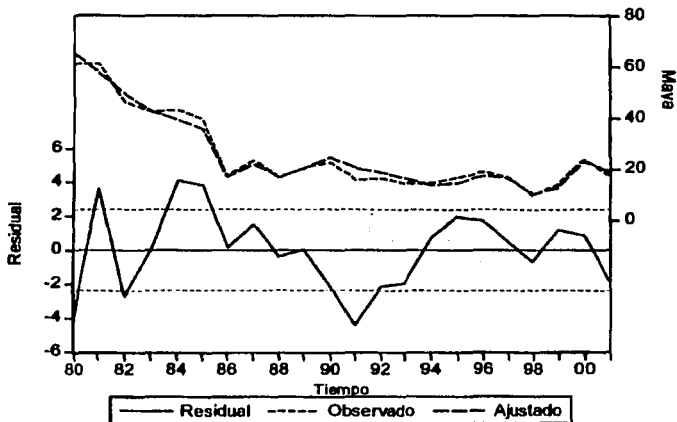
Método: Mínimos Cuadrados

Muestra (ajustada): 1980 2001

Observaciones incluidas: 22 después de ajustar los extremos

Variable	Coefficiente	Error Std.	t-Estadístico	Prob.
C	-3.484376	1.139884	-3.065555	0.0061
WTI	0.853067	0.028361	29.05396	0.0000
R-cuadrada	0.976855	Media var. dependiente	26.07319	
R-cuadrada ajustada	0.975698	D.S. var. dependiente	15.44970	
E. S. de la regresión	2.408480	Akaike criterio	4.682360	
Suma resid. cuasad.	116.0136	Schwarz criterio	4.781546	
Log probabilidad	-49.50596	F-estadístico	844.1325	
Durbin-Watson	1.540364	Prob (F-estadístico)	0.000000	

**Gráfico 4.4 Ajuste y residuales del modelo Maya = f(WTI) (datos anuales)**



Para asegurarse de la no existencia de correlación entre los residuales se obtiene su correlograma, el cual se muestra en la tabla 4.12. Los resultados indican que no existen correlaciones significativas en los rezagos, es decir, los residuales no están relacionados con sus propios valores rezagados. No es posible identificar un comportamiento que pueda ser ajustado con un modelo univariado. Se puede concluir

TESIS CON  
FALLA DE ORIGEN

que si los residuales no están correlacionados entonces el comportamiento de los residuales debe cumplir con los supuestos de la regresión clásica, o bien, parte de la varianza no ha sido explicada.

**Tabla 4.16 Correlograma Residuales MAYA = f (WTI)**

Muestra: 1980 2001

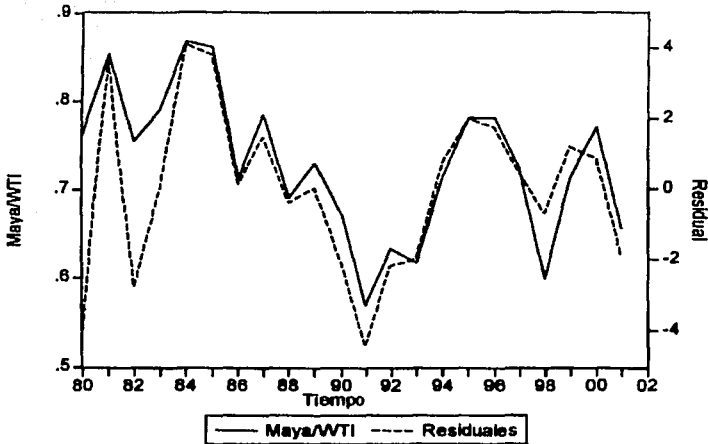
Observaciones incluidas: 22

Autocorrelación	Correlación Parcial	AC	PAC	Q-Est.	Prob	
. **	. **	1	0.149	0.149	0.5562	0.456
. *	. **	2	0.117	0.097	0.9179	0.632
.	.	3	0.026	-0.003	0.9390	0.816
**	**	4	-0.196	-0.217	2.0631	0.724
***	***	5	-0.360	-0.333	6.0834	0.298
***	**	6	-0.283	-0.205	8.7355	0.189
***	**	7	-0.321	-0.258	12.373	0.089
**	.	8	-0.074	-0.041	12.582	0.127
.	.	9	-0.006	-0.089	12.584	0.182
.	.	10	0.139	-0.039	13.440	0.200
. **	.	11	0.279	0.051	17.166	0.103
. *	**	12	0.062	-0.221	17.523	0.131

A pesar de que las pruebas realizadas indican que no existe autocorrelación, el comportamiento que siguen los residuales al graficarlos contra el tiempo, no sugieren que éstos representen un proceso de ruido blanco. Esto se debe en parte a que los residuales presentan la dinámica que sigue la relación Maya/WTI, para verlo en el gráfico 4.5 se comparan los residuales obtenidos con la relación Maya/WTI, es claro ver que las series están afectadas por la misma dinámica y que por lo tanto el modelo de ajuste lineal que se ha utilizado en los datos anuales no ha explicado este comportamiento. Al tratar de explicar ese comportamiento en los datos anuales lo que se logró fue explicarlo utilizando un modelo autorregresivo, sin embargo esta idea no es adecuada.

TESIS CON  
FALLA DE ORIGEN

Gráfico 4.5 Comparación entre Maya/WTI y los residuales



Aún cuando se presentan problemas con la dinámica de los residuales este modelo tiene el único objetivo de establecer una relación entre los dos precios de crudos para su pronóstico en el largo plazo. Las pruebas sobre normalidad y homocedasticidad realizadas a los residuales demuestran que éstos no violan los supuestos del modelo clásico de regresión lineal, y que por lo tanto los resultados obtenidos son confiables, sin embargo el limitado tamaño de la muestra hace que éstos no sean completamente confiables. Con el modelo será posible calcular el precio anual que en promedio tendrá el Maya en los siguientes años, una vez que se conozca el precio anual del WTI. Ya pronosticados los valores correspondientes al crudo Maya es posible calcular el valor de la relación para cada año.

TESIS CON  
FALLA DE ORIGEN

ESTA TESIS FUE DEPOSITADA EN LA BIBLIOTECA DE LA SECRETARÍA DE ECONOMÍA

## CONCLUSIONES

El trabajo realizado permite concluir que la relación entre los precios de los crudos sigue patrones cíclicos, que requieren ser estudiados con gran detenimiento. El trasfondo de dichos ciclos implica la existencia de una relación importante entre los precios de crudos de diferentes calidades, específicamente ligeros y pesados. Es indiscutible la existencia de factores que propician que los precios de los crudos se acerquen o que éstos se alejen de manera casi predecible; sin embargo, se han considerado diferentes tipos de variables con el fin de representar dichos factores, pero aun cuando éstas parecen ser adecuadas intuitivamente, los resultados estadísticos obtenidos no permiten determinar de manera satisfactoria como es que los cambios en dichas variables repercuten en la relación de los precios de crudos, dando lugar a los ciclos.

Es por ello que se concluye, después del análisis realizado, que la relación entre los precios de crudos de diferentes calidades está altamente determinada por el propio comportamiento de ella misma, y no es necesaria la inclusión de ningún otro tipo de variable. A pesar de lo sugerido, es importante notar que, en efecto, estadísticamente un modelo autorregresivo ofrece un buen ajuste, pero de igual manera es bien sabido que en el mercado petrolero existen muchos factores que determinan los precios de los crudos.

El petróleo, con el paso del tiempo y con los cambios tecnológicos obtenidos, se ha convertido en uno de los recursos naturales más preciados en la sociedad, la necesidad de algunos países para importar y de otros para exportar hace que el petróleo sea uno de los principales recursos naturales que se comercian a nivel mundial. La disponibilidad de petróleo en el mundo hace que el precio de este recurso varíe. Como es bien sabido, los países que cuentan con reservas pueden afectar la oferta de petróleo, reduciendo el número de barriles que se comercian en el mercado petrolero, y por lo tanto al existir una reducida cantidad, es decir, al disminuir la oferta, los precios se elevan significativamente, como ocurrió a finales del año de 1991 cuando tuvo lugar la guerra del Golfo Pérsico.

Los precios de petróleo crudo están en cierta medida, determinados por la oferta y demanda de éste, pero además se debe tener en cuenta la enorme influencia que pueden ejercer algunas organizaciones como la OPEP, la cual es la responsable de los cambios en precios y oferta de los crudos para los países que pertenecen a ella.

Otro aspecto importante a considerar es la relación entre los países consumidores y productores de crudos. Por ejemplo, existe una enorme relación entre México y los Estados Unidos con relación al suministro de petróleo pesado. La mayoría de las exportaciones mexicanas realizadas, se hacen al país del norte y eso hace que este vínculo entre ambos países sea de suma importancia. Además de ese factor, el desarrollo tecnológico en lo que se refiere a la refinación del petróleo es un factor importante. Si existe un proceso de refinación que reduzca los costos de ésta, entonces es razonable suponer que los países preferirán consumir crudos pesados y someterlos a refinación en lugar de consumir crudos ligeros, siempre y cuando, éstos últimos no representen una ventaja económica.

Como se ve, existen muchos factores que intervienen en la determinación de los precios de crudos de diferentes calidades y que por tanto afectarán a la relación entre éstos, pero algunos de estos factores no es posible considerarlos en la modelación porque con ellos se obtienen resultados que violan los supuestos en los que se basa la teoría de regresión dinámica.

Otro problema es que existen factores que no es posible predecir, por ejemplo, si la OPEP realiza un corte en la oferta de crudos entonces los precios de estos se elevarán alterando así la relación, ya que dada la falta de oferta de crudos, los precios de éstos aumentarán. Este tipo de factores pudieran ser incluidos en el modelo utilizando variables dicotómicas que permitan marcar los momentos de cambio ocasionados por políticas específicas o conflictos sociales, pero resultaría poco conveniente pues dichos momentos no son predecibles y por tanto se desconoce el valor de la variable dicotómica en el futuro. Es por tanto evidente que el mercado del petróleo es complicado y tratar de vincular el comportamiento de la relación entre los precios de crudos no es sencillo.



A pesar de que intuitivamente muchas de las variables utilizadas en el estudio deberían influir en la relación, estadísticamente éstas no fueron del todo útiles. La alta autocorrelación que se presenta en la serie del cociente de los precios de crudos de diferentes calidades, hace que prácticamente sea imposible vincularla con alguna otra variable de tipo económico, concluyendo con ello que un modelo autorregresivo es adecuado para la modelación mensual. Más específicamente se encontró que el modelo AR(2) es un modelo adecuado y estadísticamente aceptable, el cual ajusta a la serie de manera satisfactoria.

El resultado de elegir como ajuste un modelo univariado de tipo autorregresivo no sugiere que éste sea el mejor modelo de todos los posibles, simplemente es un modelo bueno que ha surgido después de haber intentado utilizar alguna otra metodología que proporcionara un mejor ajuste, pero como se obtuvo, no es posible utilizar la metodología de la regresión dinámica con las variables consideradas debido a los problemas que se presentan en la causalidad de las variables.

Sin embargo, uno de los problemas resulta al tratar de obtener un modelo anual que vincule los datos de ambos crudos en un ajuste. El modelo, resultado de modelar el precio del crudo Maya en función del precio del WTI, proporciona un ajuste adecuado con estadísticos buenos, pero se presenta un problema en los residuales pues éstos, aun cuando su correlograma no indique correlación, siguen el comportamiento de la relación y por tanto no pueden considerarse ruido blanco. Esto último hace que el modelo obtenido no sea válido y por lo tanto sea necesario construir otro tipo de modelo.

Es por lo tanto, posible encontrar otro modelo que no precisamente ocupe la teoría de la regresión dinámica y que permita resolver el problema de la modelación de ciclos más satisfactoriamente. Métodos alternativos pueden aportar mejores resultados, sin embargo, dado que el objetivo de la tesis es específicamente tratar de modelar los ciclos mediante la utilización de la regresión dinámica, no se profundiza en los métodos alternos y se concluye exclusivamente que no es posible la utilización de la teoría de regresión dinámica pues no se cuenta con la o las variables adecuadas para su aplicación.

Uno de los métodos alternativos es el análisis espectral que es un método para modelar los ciclos económicos en las series de tiempo. De manera general, se dirá que este método consiste en tratar de seleccionar entre diferentes curvas senoidales, aquella que mejor se ajuste a los ciclos de la serie en cuestión. Si bien éste pudiera ser un método adecuado para la solución de la modelación de los ciclos presentes en la relación de los precios de crudos, su utilización no es sencilla y requiere de un estudio detallado. Además de que esto no proporciona el resultado esperado de vincular el comportamiento de la relación con alguna otra variable, simplemente permite encontrar una estructura senoidal que ajuste los datos de manera satisfactoria.

## ANEXOS

## A1. Precios de los crudos Maya y WTI

Mes	Precios Corrientes		Precios constantes (dii.2001)		CPI
	WTI	Maya	WTI	Maya	
Ene-80	32.50	28.59	73.97	80.52	0.43938
Feb-80	37.00	28.00	83.04	82.84	0.44558
Mar-80	36.00	28.00	84.00	81.90	0.45237
Abr-80	39.50	27.36	86.35	59.82	0.45745
May-80	39.50	28.00	85.50	60.61	0.46197
Jun-80	39.50	29.00	84.57	59.95	0.46706
Jul-80	39.50	29.00	84.57	62.09	0.46706
Ago-80	38.00	29.00	80.77	61.64	0.47044
Sep-80	36.00	29.00	75.89	61.13	0.47440
Oct-80	36.00	28.99	75.17	60.54	0.47892
Nov-80	36.00	29.00	74.55	60.06	0.48287
Dic-80	37.00	30.39	75.92	62.35	0.48739
Ene-81	36.00	34.50	77.34	70.22	0.49134
Feb-81	38.00	34.50	76.55	69.50	0.49642
Mar-81	36.00	34.50	76.03	69.03	0.49961
Abr-81	36.00	32.86	75.52	64.91	0.50320
May-81	36.00	31.80	74.93	62.69	0.50715
Jun-81	36.00	32.00	70.38	62.54	0.51167
Jul-81	36.00	32.00	69.59	61.86	0.51732
Ago-81	36.00	28.58	69.06	54.82	0.52127
Sep-81	36.00	28.50	68.39	54.15	0.52626
Oct-81	35.00	29.73	66.35	56.37	0.53148
Nov-81	36.00	26.50	66.03	53.86	0.52618
Dic-81	35.00	26.50	65.93	53.69	0.53087
Ene-82	33.85	26.61	63.56	49.96	0.53257
Feb-82	31.56	26.50	59.07	49.60	0.53426
Mar-82	28.48	25.06	63.38	47.00	0.53370
Abr-82	33.45	25.00	62.41	46.85	0.53698
May-82	35.93	25.00	66.41	46.21	0.54104
Jun-82	35.07	25.00	64.02	45.84	0.54787
Jul-82	34.16	25.00	62.04	45.40	0.55064
Ago-82	33.95	25.00	61.53	45.31	0.55177
Sep-82	35.63	25.00	64.44	45.22	0.55290
Oct-82	35.88	25.00	64.34	45.09	0.55459
Nov-82	34.15	25.00	61.70	45.17	0.55346
Dic-82	31.72	25.00	57.55	45.38	0.55120
Ene-83	31.19	25.02	56.47	45.30	0.55233
Feb-83	28.95	25.01	52.36	45.24	0.55398
Mar-83	28.82	23.00	52.13	41.80	0.55290
Abr-83	30.61	23.00	54.97	41.30	0.55685
May-83	30.00	23.00	53.55	41.05	0.56024
Jun-83	31.00	23.00	55.17	40.93	0.56184
Jul-83	31.86	23.00	56.12	40.77	0.56419
Ago-83	31.91	23.90	56.39	42.23	0.56589
Sep-83	31.11	24.00	54.70	42.20	0.56871
Oct-83	30.41	24.87	53.31	43.77	0.57041
Nov-83	29.84	26.00	52.21	43.74	0.57154
Dic-83	28.24	25.00	51.11	43.70	0.57210
Ene-84	29.69	25.00	51.59	43.44	0.57549
Feb-84	30.15	25.00	52.13	43.23	0.57831
Mar-84	30.76	25.00	53.09	43.14	0.57944
Abr-84	30.62	25.00	52.58	42.94	0.58227
May-84	30.52	25.46	52.25	43.61	0.58396
Jun-84	29.97	25.50	51.17	43.54	0.58566
Jul-84	28.75	25.50	48.91	43.37	0.58791
Ago-84	29.25	25.50	49.56	43.21	0.59017
Sep-84	29.31	25.50	49.43	43.00	0.59300
Oct-84	28.77	25.50	48.37	42.88	0.59489
Nov-84	28.10	25.50	47.25	42.88	0.59469
Dic-84	25.43	25.50	42.78	42.88	0.59469

Anexos A1. Precios de los crudos Maya y WTI

Mes	Precios Corrientes		Precios constantes (al 2001)		CPI
	WTI	Maya	WTI	Maya	
Ene-85	26.64	25.50	43.03	42.80	0.56682
Feb-85	27.27	25.50	45.55	43.60	0.56864
Mar-85	26.24	25.50	46.99	42.44	0.60090
Abr-85	28.81	25.50	47.71	42.24	0.60373
May-85	27.82	25.29	45.56	41.73	0.60569
Jun-85	27.14	23.95	44.67	39.42	0.60766
Jul-85	27.33	23.16	44.89	38.04	0.60961
Ago-85	27.76	23.08	45.90	37.65	0.60994
Sep-85	28.29	23.16	46.25	37.66	0.61027
Oct-85	28.54	23.15	46.12	37.71	0.61398
Nov-85	30.81	22.77	50.05	36.99	0.61559
Dic-85	27.23	21.80	44.11	35.32	0.61726
Ene-86	22.85	17.90	37.07	28.92	0.61898
Feb-86	15.44	13.10	25.02	21.23	0.61726
Mar-86	12.82	10.02	20.54	16.30	0.61446
Abr-86	12.95	8.98	20.95	14.65	0.61333
May-86	15.44	8.84	25.11	14.17	0.61502
Jun-86	13.47	8.24	21.76	13.32	0.61841
Jul-86	11.56	7.63	18.72	12.34	0.61841
Ago-86	15.09	9.96	24.36	18.06	0.61954
Sep-86	14.91	10.53	23.96	16.93	0.62236
Oct-86	14.85	10.55	23.84	16.93	0.62293
Nov-86	15.21	11.02	24.39	17.68	0.62349
Dic-86	16.29	12.48	25.76	18.86	0.62495
Ene-87	18.83	13.84	29.67	22.04	0.62801
Feb-87	17.79	14.03	26.23	22.26	0.63027
Mar-87	18.30	15.10	28.91	23.85	0.63309
Abr-87	18.85	15.64	28.30	24.56	0.63046
May-87	19.42	16.48	30.40	25.80	0.63674
Jun-87	20.08	18.73	31.33	28.10	0.64100
Jul-87	21.30	17.13	33.14	26.65	0.64270
Ago-87	20.38	15.98	31.47	24.70	0.64598
Sep-87	19.46	15.36	29.99	23.63	0.64647
Oct-87	19.86	15.30	30.53	23.50	0.65117
Nov-87	19.03	13.77	29.20	21.12	0.65773
Dic-87	17.19	11.42	26.36	17.53	0.65173
Ene-88	17.08	11.87	26.14	18.17	0.65343
Feb-88	16.75	11.07	25.57	18.89	0.65512
Mar-88	16.22	11.39	24.85	17.31	0.65784
Abr-88	17.77	12.76	26.87	18.30	0.65953
May-88	17.44	12.63	26.26	19.03	0.66399
Jun-88	16.92	11.27	24.79	18.92	0.66642
Jul-88	15.52	11.17	23.16	16.70	0.66924
Ago-88	15.51	11.11	23.07	16.53	0.67206
Sep-88	14.26	10.03	21.54	14.83	0.67656
Oct-88	13.80	8.70	20.32	12.82	0.67984
Nov-88	14.15	9.64	20.83	14.18	0.67941
Dic-88	16.45	10.80	24.18	16.82	0.68302
Ene-89	17.96	12.17	26.26	17.76	0.68362
Feb-89	17.97	12.60	26.17	18.34	0.68875
Mar-89	19.51	14.07	26.25	20.38	0.69070
Abr-89	21.19	15.32	30.47	22.03	0.69522
May-89	20.25	14.81	28.97	21.33	0.69917
Jun-89	20.03	14.63	26.56	20.86	0.70087
Jul-89	19.81	14.36	26.20	20.44	0.70256
Ago-89	19.45	13.91	26.41	19.77	0.70385
Sep-89	19.54	14.42	27.86	20.42	0.70595
Oct-89	20.06	14.93	28.33	21.04	0.70834
Nov-89	19.90	15.14	27.69	21.29	0.71103
Dic-89	21.07	15.86	29.99	22.30	0.71216
Ene-90	22.80	14.80	31.83	20.71	0.71890
Feb-90	22.11	13.86	30.59	19.17	0.72289
Mar-90	20.41	12.30	28.08	16.92	0.72684
Abr-90	19.43	10.93	25.32	15.01	0.73197
May-90	18.10	10.45	24.80	14.32	0.73267
Jun-90	16.75	9.82	22.83	13.38	0.73362
Jul-90	16.44	13.19	25.05	17.91	0.73945
Ago-90	27.36	20.83	36.84	26.02	0.74322
Sep-90	33.34	26.24	44.49	35.01	0.74944

TESIS CON  
FALLA DE ORIGEN

Mes	Precios Corrientes		Precios constantes (diciembre 2001)		CPI
	WTI	Maya	WTI	Maya	
Oct-80	35.85	25.46	47.68	33.77	0.75395
Nov-80	32.43	23.59	42.92	31.22	0.75565
Dic-80	27.36	18.22	36.20	25.44	0.75565
Ene-81	24.91	13.80	32.77	18.15	0.76017
Feb-81	20.47	10.85	28.89	14.25	0.76130
Mar-81	18.90	11.37	25.97	14.81	0.76242
Abr-81	20.87	12.07	27.33	15.81	0.76255
May-81	21.22	11.99	27.71	15.66	0.76581
Jun-81	20.26	11.93	26.37	15.53	0.76807
Jul-81	21.43	12.65	27.86	16.71	0.76920
Ago-81	21.86	12.72	28.11	16.49	0.77148
Sep-81	21.91	13.36	28.27	17.24	0.77485
Oct-81	23.25	14.09	29.95	18.16	0.77546
Nov-81	22.53	11.84	28.95	14.86	0.77824
Dic-81	19.48	10.06	24.90	12.95	0.77880
Ene-82	18.82	10.06	24.13	12.92	0.77963
Feb-82	19.05	10.22	24.33	13.08	0.78276
Mar-82	18.91	10.83	24.03	13.99	0.78671
Abr-82	20.23	12.76	25.86	15.87	0.78784
May-82	21.00	13.88	26.61	17.80	0.78897
Jun-82	22.37	14.72	28.26	18.59	0.79179
Jul-82	21.74	14.63	27.40	18.52	0.79348
Ago-82	21.35	14.59	26.80	18.34	0.79575
Sep-82	21.86	16.00	27.39	18.80	0.79800
Oct-82	21.71	15.05	27.12	18.79	0.80083
Nov-82	20.32	13.50	25.30	16.83	0.80186
Dic-82	19.41	12.00	24.23	14.97	0.80139
Ene-83	19.12	11.80	23.74	14.86	0.80535
Feb-83	20.09	12.51	24.89	15.47	0.80617
Mar-83	20.35	12.81	25.09	15.80	0.81089
Abr-83	20.26	12.87	24.81	15.83	0.81325
May-83	19.87	12.22	24.52	15.01	0.81436
Jun-83	19.05	11.11	23.34	13.62	0.81561
Jul-83	17.82	11.02	21.65	13.52	0.81561
Ago-83	18.04	11.15	22.05	13.64	0.81777
Sep-83	17.46	11.28	21.31	13.76	0.81947
Oct-83	18.11	11.22	22.01	13.83	0.82285
Nov-83	16.59	9.72	20.11	11.80	0.82342
Dic-83	14.48	8.10	17.80	11.05	0.82342
Ene-84	15.01	10.23	18.18	12.39	0.82588
Feb-84	14.76	9.85	17.82	12.02	0.82650
Mar-84	14.70	10.31	17.88	12.40	0.83133
Abr-84	16.35	11.94	19.65	14.35	0.83245
May-84	17.67	12.62	21.57	15.39	0.83302
Jun-84	19.08	14.01	22.83	16.75	0.83584
Jul-84	19.70	14.53	23.90	17.33	0.83610
Ago-84	18.29	12.78	21.74	15.17	0.84148
Sep-84	17.44	12.53	20.88	14.83	0.84375
Oct-84	17.77	13.83	21.04	16.36	0.84431
Nov-84	16.10	14.10	21.41	16.68	0.84544
Dic-84	17.16	13.83	20.30	16.36	0.84544
Ene-85	18.02	14.29	21.23	16.83	0.84880
Feb-85	18.53	14.80	21.74	17.37	0.85222
Mar-85	18.85	15.12	21.89	17.88	0.85505
Abr-85	19.86	16.03	23.16	18.89	0.85787
May-85	19.74	16.05	22.98	18.87	0.85956
Jun-85	18.42	14.61	21.39	16.97	0.86126
Jul-85	17.30	13.28	20.09	15.39	0.86126
Ago-85	16.03	13.45	20.86	15.57	0.86352
Sep-85	16.21	13.39	21.05	15.44	0.86521
Oct-85	17.44	12.99	20.99	14.97	0.86803
Nov-85	16.00	13.56	20.74	15.85	0.86747
Dic-85	19.02	15.12	21.94	17.44	0.86981
Ene-86	18.80	14.64	21.56	16.79	0.87188
Feb-86	19.09	15.10	21.83	17.26	0.87481
Mar-86	21.33	16.66	24.25	18.97	0.87803
Abr-86	23.51	17.46	26.83	19.76	0.88272
May-86	21.24	16.07	24.02	18.17	0.88441
Jun-86	20.45	15.56	23.11	17.56	0.88458

TESIS CON  
FALLA DE ORIGEN

Anexos A1. Precios de los crudos Maya y WTI

Mes	Precios Corrientes		Precios constantes (diciembre 2001)		CPI
	WTI	Maya	WTI	Maya	
Jul-96	21.32	15.98	24.04	18.02	0.88867
Ago-96	21.93	17.31	24.68	18.63	0.88837
Sep-96	24.00	19.31	26.93	21.87	0.88119
Oct-96	24.90	20.56	27.85	23.00	0.88401
Nov-96	23.72	18.91	26.48	21.12	0.89571
Dic-96	25.41	19.32	28.37	21.57	0.89571
Ene-97	25.13	18.06	27.97	20.10	0.89453
Feb-97	22.19	15.71	24.61	17.43	0.89136
Mar-97	20.95	14.48	23.19	16.03	0.90381
Abr-97	19.75	14.26	21.83	15.78	0.90474
May-97	20.91	14.83	23.13	16.41	0.90418
Jun-97	19.28	13.90	21.29	15.24	0.90531
Jul-97	19.63	14.16	21.86	15.62	0.90644
Ago-97	19.93	14.82	21.95	16.32	0.90613
Sep-97	19.78	14.75	21.72	16.20	0.91039
Oct-97	21.27	15.70	23.31	17.20	0.91265
Nov-97	20.18	14.32	22.12	15.70	0.91209
Dic-97	18.30	11.44	20.08	12.56	0.91088
Ene-98	16.89	9.71	18.29	10.64	0.91285
Feb-98	16.07	8.88	17.58	9.71	0.91434
Mar-98	15.10	7.88	16.48	8.38	0.91604
Abr-98	15.32	9.14	16.89	9.98	0.91773
May-98	14.93	9.08	16.24	9.86	0.91843
Jun-98	13.99	8.30	14.87	9.01	0.92059
Jul-98	14.12	8.80	15.32	9.56	0.92189
Ago-98	13.38	8.24	14.51	8.87	0.92252
Sep-98	14.97	8.47	16.21	10.24	0.92385
Oct-98	14.42	8.15	15.57	9.86	0.92620
Nov-98	12.95	7.87	13.98	8.48	0.92820
Dic-98	11.29	6.37	12.20	6.89	0.92864
Ene-99	12.48	7.42	13.45	7.89	0.92780
Feb-99	12.00	7.13	12.91	7.87	0.92903
Mar-99	14.88	9.88	15.74	10.58	0.93185
Abr-99	17.34	12.35	18.47	13.16	0.93363
May-99	17.74	12.24	18.90	13.04	0.93383
Jun-99	17.90	13.07	19.07	13.93	0.93883
Jul-99	20.08	15.32	21.30	16.27	0.94145
Ago-99	21.27	17.43	22.53	18.47	0.94571
Sep-99	23.88	19.08	25.19	20.12	0.94823
Oct-99	22.69	16.43	23.89	19.41	0.94982
Nov-99	24.95	19.97	26.17	21.01	0.95049
Dic-99	26.11	20.09	27.47	21.14	0.95049
Ene-00	27.28	21.55	28.59	22.80	0.95331
Feb-00	28.39	23.86	30.64	24.67	0.95689
Mar-00	26.88	22.81	30.59	23.36	0.95987
Abr-00	25.78	20.45	28.65	21.14	0.96743
May-00	26.80	23.57	29.73	24.33	0.96856
Jun-00	31.86	24.81	32.74	25.46	0.97384
Jul-00	29.71	22.51	30.44	23.08	0.97580
Ago-00	31.33	25.18	32.10	25.78	0.97580
Sep-00	33.89	25.71	34.54	28.21	0.98089
Oct-00	33.02	24.88	33.80	26.29	0.98288
Nov-00	34.40	22.50	34.98	22.85	0.98325
Dic-00	28.35	17.01	28.85	17.31	0.98288
Ene-01	29.58	18.82	29.89	18.03	0.98489
Feb-01	29.58	18.37	29.78	16.50	0.98285
Mar-01	27.18	16.22	27.32	16.30	0.98511
Abr-01	27.40	16.84	27.42	16.86	0.98906
May-01	26.81	18.57	26.51	18.51	1.00358
Jun-01	27.57	18.78	27.42	18.68	1.00637
Jul-01	26.43	16.90	26.37	18.45	1.00745
Ago-01	27.40	19.52	27.33	19.47	1.00745
Sep-01	26.08	18.45	25.90	18.32	1.00697
Oct-01	22.08	15.05	22.00	15.00	1.00386
Nov-01	19.67	13.15	19.63	13.12	1.00186
Dic-01	19.33	13.87	19.37	13.70	0.99793
Ene-02	19.87	14.08	19.67	14.06	1.00019
Feb-02	20.74	15.08	20.85	15.02	1.00414
Mar-02	24.24	18.50	24.01	18.32	1.00876

TESIS CON  
FALLA DE ORIGEN

## A2. Análisis de Regresión

El análisis de regresión sirve para establecer una relación entre variables, de modo que sea posible expresar el comportamiento de una determinada variable en función del comportamiento de otra(s) variable(s), con el objetivo de estimar o predecir el valor promedio o media poblacional de la variable dependiente en términos de los valores observados de la variable(s) explicativa(s). Uno de los primeros supuestos del modelo de regresión lineal es la existencia de una relación funcional entre la variable dependiente y las independientes. Si la variable dependiente está en función de solamente una variable, entonces se dice que es una *regresión lineal simple*; pero si por el contrario, está en función de varias variables entonces es una *regresión lineal múltiple*.

En términos matemáticos, si se considera a  $Y$  como la variable dependiente, y a  $X_1, X_2, \dots, X_i$  las variables explicativas, como ya se dijo lo que se busca es expresar de manera lineal la variable  $Y$  en términos de las  $X$ , es decir,

$$Y = A + B_1 X_1 + B_2 X_2 + \dots + B_i X_i \quad (\text{A2.1})$$

Pero este modelo sólo es el modelo teórico, en la realidad el modelo contendrá un término aleatorio originado por diferentes circunstancias; a este término se le denomina el error o residual, y sobre el se harán importantes supuestos que permitirán hacer un desarrollo teórico más profundo. Entonces el modelo que se ajustará es el siguiente

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + u \quad (\text{A2.2})$$

Cuando se habla de linealidad, es necesario diferenciar entre linealidad en las variables y linealidad en los parámetros. La linealidad en las variables se considera cuando las variables explicativas están elevadas a la primera potencia, o bien no tienen ningún tipo de transformación, el caso de la regresión anterior es un ejemplo claro de una regresión lineal, pues existe linealidad en las variables. Sin embargo en la siguiente regresión no existe linealidad en las variables,

$$Y_i = \alpha + \beta_1 X_i^2 + \beta_2 e^X + u_i$$

La linealidad en los parámetros, de manera análoga, se da cuando los parámetros involucrados en la función se encuentran elevados a la primera potencia, como ejemplo de linealidad en los parámetros se tiene la ecuación (A2.2) en la cual existe linealidad tanto en los parámetros como en las variables. Si se considera la siguiente ecuación es claro notar que se viola la linealidad de los parámetros,

$$Y = \alpha + \beta_1^2 X + (1/\beta_2) e^X + u_i$$

Para efecto de explicar los conceptos e ideas básicas del análisis de regresión, se considerará el caso de la regresión lineal simple, en el cual solamente se involucra solamente una variable explicativa, es decir, se considera el siguiente modelo,

$$Y_i = \alpha + \beta X_i + u_i$$

donde el término  $u_i$  es una variable aleatoria que representa el error de estimación cometido en la observación  $t$ , también conocido como residual. Entonces, el objetivo del modelo clásico de regresión lineal es estimar los parámetros  $\alpha$  y  $\beta$ , de modo tal que sea posible inferir sobre los verdaderos valores.

El modelo clásico de regresión lineal está basado en diferentes supuestos, los cuales son de importancia, pues dado que la variable  $Y$  depende de  $X_i$  y de  $u_i$ , es necesario especificar la forma como se generan las  $X_i$  y las  $u_i$  para poder hacer inferencia estadística sobre las  $Y_i$  y también sobre los parámetros, además de que la efectividad de los supuestos asegura que la interpretación de los valores estimados de la regresión es válida.

### A2.1 Coeficiente de determinación

Parte importante de los modelos de regresión es considerar la *bondad de ajuste* de la línea de regresión ajustada a un conjunto de datos, es decir, saber que tan bien se



ajusta dicha recta a las observaciones. Generalmente los residuales de la regresión no serán cero, pues en ese caso se tendría un ajuste perfecto, lo cual sucede muy raras veces; en realidad se tendrán valores de  $u_i$  positivos y otros negativos, y lo que se espera es que dichos residuales sean lo más pequeños posibles y que además sean aleatorios.

El coeficiente de determinación  $R^2$  es una medida resumen que dice qué tan bien se ajusta la línea de regresión muestral a los datos. Entre menor sea la variación del valor de los residuales alrededor de la recta ajustada, mejor será la estimación realizada. Entonces esta medida indica la proporción de la variación en la variable dependiente que está explicada por la variación en las variables independientes. Por tanto es claro notar que el valor de la  $R^2$  fluctuará entre 0 y 1, pues entre mayor sea la proporción de la varianza explicada por los regresores mayor será el valor del coeficiente, siendo 1 cuando el total de la variación de  $Y$  está explicada por las variables independientes y 0 cuando no existe relación alguna entre las variables.

## A2.2 Supuestos sobre los residuales

Algunos de los supuestos en los que se basa el modelo clásico están referidos al comportamiento de los residuales. Estos son de suma importancia ya que existe una estrecha relación entre los residuales y la variable dependiente, de modo tal que la violación de estos afectaría directamente a la variable que deseamos modelar. Los cuatro supuestos de los residuales son:

- a) El error tiene esperanza nula, o el valor esperado de los errores es cero, esto significa que el valor medio del error es igual a cero,  $E[u_i]=0$ . Lo que significa que todos los factores que no están incluidos en el modelo y que, por consiguiente, están incorporados en el residual, no afectan el valor de la media de  $Y$ , es decir, los valores positivos de  $u_i$  se cancelan con los valores negativos, de tal manera que el efecto del promedio es nulo.
- b) La varianza del error es siempre constante, y no depende de  $X$ , a este supuesto se le conoce como el supuesto de la homocedasticidad,  $\text{Var} [ u_i ] = \sigma^2$ . Este supuesto significa que sin importar el valor que tome la variable independiente  $X$ , la dispersión

de los residuales alrededor de la recta ajustada en este valor, debe ser igual a la correspondiente al resto de los valores de la variable X.

- c) Los errores son independientes entre sí, es decir, el conocimiento de alguno de ellos no implica el conocimiento de ningún otro, esto se expresa como,  $E[u_i u_j] = 0$  para  $i \neq j$ . Este supuesto también es conocido como el supuesto de no autocorrelación entre los residuales, lo cual implica que el valor del residual no depende de los valores de él mismo en el pasado, y que por lo tanto no están relacionados entre ellos.
- d) Dado que el error es una variable aleatoria, es necesario hacer un supuesto sobre su distribución, por lo tanto se supondrá que el error tiene distribución normal,  $u_i \sim N(0, \sigma^2)$ .

### A2.3 Heteroscedasticidad

Uno de los supuestos más importantes en el modelo de regresión lineal entorno a las perturbaciones o residuales  $u_i$  es que éstas son homocedásticas, es decir, todas tienen la misma varianza. Esto significa que no importa el valor de las variables explicativas o el tiempo en el que se calcule, los errores conservarán la misma varianza. Sin embargo, es frecuente encontrar que las varianzas de los residuales cambian cuando las variables explicativas cambian, a esto se le conoce como *heteroscedasticidad*.

El supuesto considerado de la varianza constante en los residuales, es uno de los supuestos más importantes y por ende requiere de un poco más de detenimiento. La heteroscedasticidad no elimina las propiedades de insesgamiento<sup>9</sup> y consistencia<sup>10</sup> de los

<sup>9</sup> Se dice que un estimador  $\hat{\theta}$  es un estimador insesgado de  $\theta$  si el valor esperado de  $\hat{\theta}$  es igual al verdadero valor de  $\theta$ , es decir,  $E[\hat{\theta}] = \theta$ .

<sup>10</sup> Un estimador es consistente si éste se aproxima al verdadero valor a medida que el tamaño de la muestra se hace más grande.



estimadores, sin embargo estos estimadores dejan de tener la varianza mínima, y por lo tanto ya no son eficientes.

Existen métodos informales para saber si existe heteroscedasticidad en el modelo trabajado, el más fácil de ellos es el método gráfico, el cual consiste en que una vez realizada la regresión se calculan los residuales y se elabora un gráfico de dispersión entre la variable dependiente y los residuales al cuadrado. Si en el gráfico se percibe algún patrón en el comportamiento, es decir, si los datos no se distribuyen de manera aleatoria, si no que forman figuras, como curvas, conos, o a mayor valor de  $Y$  los residuales son mayores, entonces se intuye la presencia de heteroscedasticidad. Los gráficos pueden realizarse utilizando la variable dependiente o bien alguna de las variables explicativas.

Además del método gráfico existen otras pruebas más formales para determinar la presencia de heteroscedasticidad, el estudio de cada una de ellas requiere de mayor detenimiento. Existen pruebas que suponen que la varianza  $\sigma^2$  depende del valor de alguna de las variables independientes  $X_i$ , por lo tanto, para cerciorarse de la presencia de heteroscedasticidad, corren regresiones simples de los residuales sobre las variables explicativas, ejemplos de estas pruebas son la *prueba de Park* y la *prueba de Glesjer*.

Existe otra prueba más, cuya simplicidad la hace más conocida y utilizada, ésta es la *prueba general de heteroscedasticidad de White*, la cual corre una regresión de los residuales estimados al cuadrado utilizando como variables independientes las mismas variables utilizadas en la regresión original, los cuadrados de éstas, y sus productos cruzados. Si el valor obtenido para la  $R^2$  multiplicado por el número de observaciones incluidas, es mayor que el valor crítico de una Ji-cuadrada al nivel de confianza especificado, se concluye que hay heteroscedasticidad. Esta prueba es la más común entre los paquetes de estadística.

Una vez que se ha detectado la existencia de heteroscedasticidad en los datos es necesario tratar de remediarla. Para ello existen dos enfoques, uno cuando la varianza es conocida, en cuyo caso bastará con ponderar las observaciones utilizando como

ponderador el inverso de la varianza correspondiente, es decir, se considera la siguiente regresión,

$$\frac{Y_i}{\sigma_i} = \alpha \left( \frac{1}{\sigma_i} \right) + \beta \left( \frac{X_i}{\sigma_i} \right) + \frac{u_i}{\sigma_i}$$

pero si la varianza es desconocida entonces una manera de remediar el problema es hacer transformaciones en las variables de modo tal que el problema desaparezca. Ahora el problema es elegir la transformación más adecuada. Una manera de saber cual transformación es la más adecuada consiste en examinar como se comporta la varianza de los residuales y descubrir si ésta es proporcional a alguna forma funcional ya sea de la variable explicativa o de la dependiente. Una transformación logarítmica de las variables frecuentemente ayuda a reducir el problema de heteroscedasticidad cuando se compara con el modelo original. Otros tipos de transformaciones que se pueden realizar dependerán de la relación que tengan los residuales con la variable dependiente. Por ejemplo, si se considera el modelo simple de regresión siguiente

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

y se concluye que

$$E(u_i^2) = \sigma^2 X_i^2 \Rightarrow \frac{Y_i}{X_i} = \beta_1 + \beta_2 \frac{1}{X_i} + v_i$$

$$E(u_i^2) = \sigma^2 X_i \Rightarrow \frac{Y_i}{\sqrt{X_i}} = \beta_1 \frac{1}{\sqrt{X_i}} + \beta_2 \sqrt{X_i} + v_i$$

#### A2.4 Autocorrelación

Si se viola el supuesto del modelo clásico de regresión lineal de que los errores son aleatorios o no correlacionados, surge el problema de *autocorrelación* o correlación serial. Este problema puede surgir por diversas razones, tales como la inercia o lentitud de las series de tiempo económicas, la exclusión de variables importantes del modelo o la utilización de una forma funcional incorrecta.

La autocorrelación conduce a tener estimadores insesgados y consistentes pero que dejan de ser eficientes<sup>11</sup>, y por tanto las pruebas de significancia  $t$  y  $F$ <sup>12</sup> ya no son del todo válidas. Por tanto, es necesario corregir dicho error; sin embargo, el remedio depende de la naturaleza de los errores, y como éstos no son observables sino estimados, lo más común es suponer que el error cometido en el tiempo actual está linealmente relacionado con el error en el periodo anterior, lo cual es conocido como un modelo *autorregresivo de primer orden*.

Para detectar el problema de autocorrelación existen diferentes métodos, de los cuales el más conocido es el estadístico  $d$  de Durbin-Watson; sin embargo, este estadístico es indicador de la presencia de un sesgo de especificación y no de autocorrelación pura. Una manera sencilla de detectar la autocorrelación es elaborando una gráfica de los residuales contra el tiempo, si la gráfica muestra un patrón de crecimiento, decrecimiento o presencia de ciclos, y no un comportamiento aleatorio entonces esto indica que el problema está presente.

En la práctica hay métodos comunes para detectar la presencia de autocorrelación son los correlogramas. Existen métodos más formales para detectar la autocorrelación, desde los más sencillos como lo es la prueba no paramétrica de "rachas", y otras un poco más complicadas como la prueba asintótica o de grandes muestras.

Una de las pruebas más comúnmente utilizada es la prueba de Breusch-Godfrey sobre autocorrelación de orden superior, la cual sugiere correr una regresión auxiliar con los residuales obtenidos, utilizando de variables independientes las mismas consideradas en el modelo original y los residuales rezagados, es decir, se considera el siguiente esquema

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t$$

<sup>11</sup> Un estimador  $\hat{\beta}$  es eficiente si para un determinado tamaño de muestra la varianza de  $\hat{\beta}$  es menor que para cualquier otro estimador de determinado parámetro, es decir, es el estimador con varianza mínima.

<sup>12</sup> Las pruebas  $t$  y  $F$  son métodos estadísticos que sirven para determinar la significancia de las variables incluidas.

una vez obtenidos los resultados se utiliza el coeficiente de determinación  $R^2$  y se multiplica por  $(n-p)$  que es el número de observaciones incluidas. Si dicho producto excede el valor crítico de una Ji-cuadrada al nivel de confianza deseado, entonces por lo menos uno de los  $p_i$  es significativamente diferente de cero.

### A2.5 Multicolinealidad

Un supuesto más es la no-multicolinealidad de las variables explicativas, el cual se presenta cuando estamos trabajando con modelo multivariado. Este supuesto se viola cuando alguna de las variables explicativas puede ser calculada a partir de las otras, es decir, la existencia de una relación casi perfecta o casi exacta entre algunas o todas las variables explicativas, sin embargo, pueden existir variables que estén intercorrelacionadas pero no en forma perfecta, lo cual también presenta un problema de multicolinealidad.

Existen diferentes factores que pueden originar la multicolinealidad en las variables, por ejemplo, el método de recolección de información empleado o las restricciones sobre la población que es objeto del muestreo. Además una mala especificación del modelo, por ejemplo la consideración de términos polinomiales, o bien si el modelo tiene más variables explicativas que número de observaciones, pueden ser causas de multicolinealidad. Matricialmente esto se puede entender si al considerar la matriz de observaciones tenemos columnas linealmente dependientes, entonces el modelo tendrá problemas de multicolinealidad.

Varias son las consecuencias de la multicolinealidad. Si existe colinealidad perfecta entre las variables, sus coeficientes de regresión son indeterminados y sus errores estándar no están definidos. Si la colinealidad es alta pero no perfecta, es posible estimar los coeficientes de la regresión pero sus errores estándar tienden a ser grandes, por tanto los valores poblacionales de los coeficientes no pueden ser estimados de manera precisa.

No existen métodos seguros para detectar la multicolinealidad, sin embargo existen algunas formas que indican la presencia de este problema:

- a) Uno de los indicios más claros de este problema es cuando el coeficiente de determinación  $R^2$  es muy elevado, pero la mayoría de los coeficientes de regresión no son estadísticamente significativos.
- b) Una posibilidad es regresar cada una de las variables  $X_i$  sobre el resto de las variables explicativas en el modelo y encontrar los coeficientes de determinación correspondientes. Un  $R^2$  elevado sugerirá que  $X_i$  está altamente correlacionada con el resto de las  $X$ , y por lo tanto puede ser recomendable eliminarla del modelo.
- c) Gráficamente al realizar un diagrama de dispersión entre las variables explicativas es posible encontrar si entre éstas existe colinealidad de acuerdo al comportamiento que presenta la dispersión. Si ésta no sugiere algún tipo de relación, es decir, los datos están dispersos entonces no existe correlación entre las variables, pero si los datos se distribuyen con cierto comportamiento tendencial o polinómico por ejemplo, entonces existe una relación lineal entre ellas y la presencia de ambas puede estar originando la violación del supuesto.

No existe en realidad una manera segura de deshacerse del problema pero existen algunas medidas que pueden ayudar: obtener información adicional nueva, transformar los datos, omitir una variable si es altamente colineal, entre otros.

TESIS CON  
FALLA DE ORIGEN

En el siguiente cuadro se da un resumen de los supuestos en los que se basa el modelo clásico de regresión lineal.

Supuesto	Expresión matemática		Violaciones
	Modelo Bivariado	Modelo Multivariado	
La variable dependiente se puede expresar como una función lineal de un grupo específico de variables independientes, más un error	$Y_t = \alpha + \beta X_t + u_t$ $t = 1, \dots, T$	$Y = X\beta + \epsilon$	Maños regresores No linealidad Parámetros cambiantes
El valor esperado del error es cero	$E[u_t] = 0 \forall t$	$E[\epsilon] = 0$	Intercepto sesgado
Los errores tienen varianza constante y no están correlacionados.	$E[u_i u_j] = 0$ para $i \neq j$ $= \sigma^2$ para $i = j$	$E[\epsilon \epsilon'] = \sigma^2 I$	Heteroscedasticidad Errores autocorrelacionados
Las observaciones de las variables independientes se pueden considerar fijas en muestras repetidas	$X_t$ fija en muestras repetidas	$X$ fija en muestras repetidas	Errores en las variables Autoregresión Ecuaciones simultáneas
No existe relaciones lineales entre las variables independientes y hay más observaciones que variables.	$\sum_{i=1}^T (x_i - \bar{x})^2 \neq 0$	Rango (X) = K ≤ T	Multicolinealidad perfecta

TESIS CON  
FALLA DE ORIGEN



### A3. Modelos AR, MA, ARMA y ARIMA

Existen modelos que se utilizan para explicar las series de tiempo a través de sus propios valores en el pasado y también con ponderaciones de los residuales ya sean rezagados o no. Estos procesos son de mucha importancia en el desarrollo de la regresión dinámica pues en ellos se basa su teoría y por tanto es necesario examinarlos de manera general.

#### A3.1 Modelo Autorregresivo (AR)

En el modelo autorregresivo, a diferencia de los modelos de regresión clásicos donde se supone la dependencia de otras variables aleatorias, se supone que la variable dependiente  $Y_t$  depende de sus propios valores en el pasado.

El modelo autorregresivo más sencillo es el modelo de orden 1, denotado por AR(1), éste relaciona linealmente  $Y_t$  con su primer rezago  $Y_{t-1}$ , es decir, considera la ecuación dada por:

$$Y_t = \delta + \phi Y_{t-1} + \varepsilon_t$$

si se toman esperanzas y se considera el supuesto que la serie es estacionaria y por lo tanto se cumple que  $E[Y_t] = E[Y_{t-1}] = \mu$ , de donde se obtiene lo siguiente,

$$\mu = \frac{\delta}{1 - \phi}$$

por lo tanto si se resta a la variable  $Y_t$  su media y se denota como  $\tilde{Y}_t = Y_t - \mu$  entonces se puede escribir  $\tilde{Y}_t - \phi \tilde{Y}_{t-1} = \varepsilon_t$ .

Parte importante al trabajar con los modelos univariados es la utilización del operador retardo denotado con B (del inglés backward), este operador está definido como  $Y_{t-k} = B^k Y_t$ . Con este operador, se puede escribir la última expresión como,

$$(1 - \phi B)Y_t = \varepsilon_t$$

si se denota como  $\phi(B)$  al polinomio sobre B, es decir,  $\phi(B) = 1 - \phi B$ , se tiene

$$\phi(B)Y_t = \varepsilon_t$$

Interesa saber si la serie en el largo tiempo se estabilizará, o mejor dicho, si la serie es estacionaria. Para que dicha serie sea estacionaria se requiere que la raíz de la ecuación  $\phi(B) = 0$  se encuentre fuera del círculo unitario, es decir, se necesita que  $|\phi| < 1$  para asegurar la estacionariedad del proceso.

Como  $Y_t$  y  $\varepsilon_t$  son independientes, se tiene que la varianza del proceso estará dada por la siguiente expresión, en la cual se denota como  $\sigma_Y^2$  a la varianza del proceso,

$$\sigma_Y^2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2}$$

de nuevo aparece la condición  $|\phi| < 1$  que ahora es necesaria para que la varianza sea finita y positiva. La función de autocovarianza se obtiene al multiplicar la ecuación del proceso por  $Y_{t-k}$  y tomando esperanzas, de donde se obtiene que

$$\gamma_k = \phi \gamma_{k-1} \quad \text{para } k = 1, \dots ; \quad \gamma_0 = \sigma_Y^2$$

en donde  $\gamma_k = E[(Y_{t-k} - \mu)(Y_t - \mu)]$

De lo anterior se concluye que la función de autocorrelación está dada por

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi \gamma_{k-1}}{\gamma_0} = \phi \rho_{k-1}$$

pero como  $\rho_1 = \phi$  entonces, de aquí se deduce que cuando  $k$  es grande, la correlación tiende a cero.

Si se considera en el caso general un modelo autorregresivo de orden  $p$ , denotado por  $AR(p)$ , lo que se está suponiendo es que las variables de las que depende son todos sus rezagos de orden menor o igual a  $p$ , y la ecuación está dada por

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

en dicha ecuación el parámetro  $\delta$  es una constante la cual estará relacionada con la media del proceso y  $\varepsilon_t$  es un proceso de ruido blanco. La constante involucrada  $\delta$  es igual a  $(1 - \phi_1 - \dots - \phi_p)\mu$ , de modo tal que se puede escribir la ecuación del proceso autorregresivo como,

$$\phi(B)\tilde{Y}_t = \varepsilon_t \quad \text{donde} \quad \tilde{Y}_t = Y_t - \mu; \quad \phi(B) = 1 - \phi_1 - \phi_2 - \dots - \phi_p$$

### A3.2 Modelos de media móvil (MA)

En el caso de los modelos autorregresivos se considera la posibilidad de que la variable dependiente, esté representada en función de sus valores rezagados, sin embargo, es posible que el número de rezagos de los que dependa sea infinito y por lo tanto resulte imposible la estimación de los coeficientes correspondientes a cada uno de ellos. Una forma equivalente de resolver este problema es considerar que la variable no depende de una infinidad de sus rezagos sino de un número finito de los residuales rezagados, así como del residual en el período.

Los modelos de media móvil suponen que cada valor de la variable independiente,  $Y_t$  está generado por un promedio ponderado de los residuales en los últimos periodos

Con fines de describir más a fondo el modelo se considera el proceso de promedios móviles de orden uno, es decir, el modelo MA(1) que corresponde a la siguiente ecuación

$$\begin{aligned} Y_t &= \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} \\ \tilde{Y}_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} \end{aligned}$$

En la última expresión se denotó, igual que antes,  $\tilde{Y}_t$  al resultado de restarle a la variable su media, es decir,  $Y_t - \mu$ . Haciendo uso del operador retardo B, se obtiene,

$$\tilde{Y}_t = (1 + \theta_1 B) \varepsilon_t$$

de donde se obtiene de inmediato que  $E[\tilde{Y}_t] = 0$  y  $\text{Var}[\tilde{Y}_t] = \sigma_\varepsilon^2(1 + \theta^2)$ . Además las autocovarianzas están dadas por

$$\gamma_k = E[\tilde{Y}_t \tilde{Y}_{t-k}] = \begin{cases} \theta \sigma_\varepsilon^2 & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

de donde se tiene que las autocorrelaciones están dadas por

$$\rho_k = \begin{cases} \frac{\theta}{1 + \theta^2} & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

Como las correlaciones para rezagos mayores a uno son cero, se concluye que el proceso de media móvil de orden uno, a diferencia del proceso autorregresivo, tiene una memoria limitada, recordando solamente lo ocurrido en el periodo anterior. Sin embargo la primera autocorrelación no puede ser muy grande, pues esto indicaría una alta dependencia de la observación actual con la anterior, y de ésta con su anterior y así sucesivamente, por lo cual sería más adecuado considerar un modelo autorregresivo. Esto último está relacionado con que es posible demostrar que  $|\rho_1| \leq 0.5$ , entonces si se obtiene un valor mayor a 0.5 lo más adecuado es utilizar un proceso autorregresivo.

De manera general, si se supone un modelo de media móvil de orden  $q$ , lo cual significa que  $Y_t$  depende de los residuales correspondientes a las últimas  $q$  observaciones, entonces el modelo se denota por MA( $q$ ) y su ecuación está dada por,

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

en donde los valores de los parámetros  $\theta_i$ , pueden ser valores positivos o negativos y no forman precisamente un promedio pues su suma no debe ser forzosamente igual a 1, además se asume que los residuales son un proceso de ruido blanco.

En este caso general, se tiene que

$$E[\tilde{Y}_t] = 0$$

$$\text{Var}[\tilde{Y}_t] = \sigma_\varepsilon^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)$$

$$\gamma_k = \begin{cases} (\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q)\sigma_\varepsilon^2 & \text{si } k = 1, \dots, q \\ 0 & \text{si } k \geq q+1 \end{cases}$$

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{si } k = 1, \dots, q \\ 0 & \text{si } k \geq q+1 \end{cases}$$

En esta última parte se ha concluido que existe una relación importante entre los procesos autorregresivos y de medias móviles. De manera simple se tiene que considerar un proceso autorregresivo de orden finito es equivalente a considerar un proceso de medias móviles de orden infinito; de manera equivalente considerar un proceso de medias móviles de orden finito es igual a considerar un proceso autorregresivo de orden infinito.

Como se ha visto las funciones de correlación son de mucha utilidad para la identificación del orden de los procesos tanto autorregresivos como de promedios móviles.

### A3.3 Modelo Autorregresivo-Media Móvil (ARMA)

En este modelo se juntan los dos anteriores se supone que la variable dependiente depende tanto de sus valores rezagados como de los residuales en los últimos periodos. Un proceso autorregresivo con media móvil que depende de sus  $p$  últimos valores rezagados y los últimos  $q$  residuales se denota por ARMA  $(p, q)$ , y su ecuación es de la forma

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

El modelo más sencillo de este tipo es el modelo ARMA(1,1), sin embargo el ser el modelo de menor orden no le quita importancia, al contrario, en la práctica este modelo

permite representar adecuadamente a muchas series de la realidad. Está definido por la siguiente ecuación

$$(1 - \phi B) \tilde{Y}_t = (1 + \theta B) \varepsilon_t$$

donde  $\phi \neq \theta$ ,  $|\phi| < 1$  para que el proceso sea estacionario, y  $|\theta| < 1$  para que sea invertible.

Para obtener la función de autocorrelación del proceso, se multiplica por  $\tilde{Y}_{t-k}$  y se toman esperanzas, obteniendo así:

$$\gamma_k = \phi \gamma_{k-1} + E[\varepsilon_t \tilde{Y}_{t-k}] + \theta E[\varepsilon_{t-1} \tilde{Y}_{t-k}]$$

para  $k > 1$ , el ruido  $\varepsilon_t$  es independiente de la historia de la serie y por lo tanto se tiene,

$$\gamma_k = \phi \gamma_{k-1} \quad \text{si } k > 1$$

para  $k = 0$ , se tiene que

$$E[\varepsilon_t \tilde{Y}_t] = \sigma^2 \quad \text{y}$$

$$E[\varepsilon_{t-1} \tilde{Y}_t] = E[(\phi \tilde{Y}_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}) \varepsilon_{t-1}] = \phi \sigma^2 + \theta \sigma^2$$

por lo tanto  $\gamma_0 = \phi \gamma_0 + \sigma^2 + \theta \sigma^2 (\phi - \theta)$

y finalmente para  $k = 1$  se tiene que  $\gamma_1 = \phi \gamma_0 + \theta \sigma^2$

Ahora para obtener la función de autocorrelación simple, se divide entre  $\gamma_0$  con lo que se obtienen los siguientes resultados que especifican la función de autocorrelación de un proceso ARMA(1,1)

$$\rho_1 = \phi + \theta \frac{\sigma^2}{\gamma_0}$$

$$\rho_k = \phi \rho_{k-1}$$

De las expresiones que determinan la función de correlación simple es posible concluir que tiene un decrecimiento que comienza a partir de  $\rho_1$  y no de  $\rho_0=1$  como ocurre en el AR(1), esto debido a la inclusión del parámetro de promedios móviles.

**A3.4 Modelo Autorregresivo-Integrado-Media Móvil (ARIMA)**

Los modelos anteriores suponen que la serie es estacionaria, sin embargo en la realidad es muy común trabajar con series que no lo son, pero que después de diferenciarlas un determinado número de veces éstas resultan ser estacionarias. Si para volver estacionaria a una serie es necesario diferenciarla  $d$  veces entonces se dice que es integrada de orden  $d$ , y se denota como  $I(d)$ .

Lo anterior lleva a los modelos *autorregresivos integrados con medias móviles*, los cuales están denotados por ARIMA( $p$ ,  $d$ ,  $q$ ). Estos son necesarios para establecer modelos en los cuales la serie es no estacionaria. Los parámetros del modelo se refieren, igual que antes, al número de rezagos a incluir en el modelo para formar el polinomio autorregresivo de orden ( $p$ ), y el número de residuales involucrados en el polinomio de promedio móviles de orden ( $q$ ), pero además está el nuevo parámetro que denota el número de veces que es necesario diferenciar la serie para volverla estacionaria ( $d$ ).

Entonces la expresión matemática para estos procesos está dada por

$$\phi(B)\nabla^d Y_t = \theta(B)\varepsilon_t$$

TESIS CON  
FALLA DE ORIGEN

## BIBLIOGRAFÍA

- *Estadística, Modelos y Métodos*; Daniel Peña Sánchez  
Tomo II, Alianza Universidad Textos, España
- *Econometría Básica*; Damodar N. Gujarati  
Ed. Mc. Graw Hill, Tercera Edición, Colombia 1997.  
Traducción Gladys Arango Medina
- *Econometric Models & Economic Forecasts*; Robert S. Pindyck, Daniel L. Rubinfeld  
Ed. Mc. Graw Hill, Tercera Edición, Singapur 1991
- *Análisis Estadístico de Series de Tiempo Económicas*; Victor M. Guerrero  
Colección CBI, UAM Unidad Iztapalapa, México 1991
- *Estadística Aplicada a la administración y a la economía*; David K. Hildebrand, R. Lyman Ott  
Ed. Addison-Wesley Iberoamericana, Wilmington, EUA 1997  
Traducción Carlos Torres Alcaraz
- Páginas web:
  - [www.pemex.gob.mx](http://www.pemex.gob.mx)
  - [www.geocities.com/currie\\_alder/oad3.pdf](http://www.geocities.com/currie_alder/oad3.pdf)
  - [www.platts.com](http://www.platts.com)
  - [www.api.org](http://www.api.org)
  - [www.eia.doe.gov/emgu/cabs/mexico.html](http://www.eia.doe.gov/emgu/cabs/mexico.html)
  - [www.imp.mx/petroleo](http://www.imp.mx/petroleo)
  - [www.imp.mx/investigación/maya.html](http://www.imp.mx/investigación/maya.html)

TESIS CON  
FALLA DE ORIGEN



## Índice de Tablas

- 2.1 Correlograma Maya/WTI
- 2.2 Prueba raíz unitaria Maya/WTI
  
- 4.1 Matriz de correlaciones (datos mensuales)
- 4.2 Regresiones simples
- 4.3 Correlograma IMPOIL
- 4.4 Prueba raíz unitaria IMPOIL
- 4.5 Correlaciones cruzadas D(IMPOIL) y D(RELMAYAWTI)
- 4.6 Prueba de causalidad de Granger RELMAYAWTI & IMPOIL
- 4.7 Prueba de causalidad de Granger D(RELMAYAWTI) & D(IMPOIL)
- 4.8 Modelo univariado para el IMPOIL (ARIMA(2,1,0))
- 4.9 Correlograma cruzado ALFA, BETA
- 4.10 Pruebas de causalidad
- 4.11 Modelo AR(2) para Maya/WTI
- 4.12 Correlograma Residuales del modelo AR(2)
- 4.13 Correlograma Relación Maya/WTI
- 4.14 Matriz de correlaciones (datos anuales)
- 4.15 Regresión Maya & WTI (datos anuales)
- 4.16 Correlograma Residuales Maya = f(WTI)

TESIS CON  
FALLA DE ORIGEN

## **Índice de Gráficos**

- 1.1 Volumen de la Exportaciones de Petróleo Crudo
  
- 2.1 Datos Maya y WTI (precios corrientes)
- 2.2 Datos Maya y WTI (precios constantes)
- 2.3 Histograma Maya (dólares 2001)
- 2.4 Histograma WTI (dólares 2001)
- 2.5 Relación Maya/WTI
- 2.6 Histograma Maya/WTI
- 2.7 Diferencia Maya-WTI
- 2.8 Histograma Maya-WTI
- 2.9 Residuales & WTI
- 2.10 Residuales & Tiempo
- 2.11 Dispersión WTI & Maya
  
- 4.1 Ajuste AR(2) y Residuales del modelo para Maya/WTI
- 4.2 RELMAYAWTI & Residuales del modelo AR(2)
- 4.3 Histograma Residuales del modelo AR(2)
- 4.4 Ajuste y residuales del modelo  $\text{Maya} = f(\text{WTI})$  (datos anuales)
- 4.5 Comparación entre Maya/WTI y los residuales