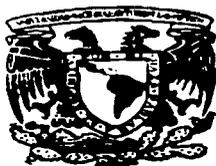


46  
2ej.



**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**

**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES**

**MODELOS DE SERIES DE TIEMPO PARA LA  
PRODUCCION DE PETRÓLEO CRUDO Y GAS  
NATURAL Y VENTAS DE PRODUCTOS  
DERIVADOS.(PERIODO 1982-1995)**

**T E S I S**  
QUE PARA OBTENER EL TITULO DE :  
LICENCIADO EN MATEMATICAS  
APLICADAS Y COMPUTACION  
**P R E S E N T A**

**MA. ASCENCIÓN OLVERA GUDIÑO**

**ASESOR : ACT. MARIA DEL CARMEN GONZÁLEZ VIDEGARAY**

**ACATLAN , EDO. MEX. 1997**

**TESIS CON  
FALLA DE ORIGEN**





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.

**GRACIAS**

**A Dios por todo lo que me ha dado.**

**A mis padres:**

**Fabián y Carmen, a quienes les debo lo que soy, por todo su amor y confianza que han depositado en mí.**

**A mis hermanos:**

**José, Martha y Juanita, por el cariño y comprensión que me han brindado en todo momento.**

**A la Escuela Nacional de Estudios profesionales (ENEP-ACATLAN):**

**Por la formación que me brindó a través de los profesores,  
gracias a todos ellos, y en especial a la Act. Ma. del Carmen González V.  
por su valiosa colaboración en este trabajo.**

**Al personal de la Gerencia de Análisis de Inversión en Explotación de  
PEMEX,  
en especial al Ing. Rodrigo Pulido Maldonado, por todo el  
apoyo recibido para la culminación de esta meta.**

**Al Fis. Primitivo Pimentel y el Lic. Ernesto Arellano:**

**Personal de PEMEX, por su valiosa  
colaboración para el acopio de información  
contenida en el presente trabajo.**

**A todas aquellas personas que de alguna forma contribuyeron al desarrollo y culminación  
de este proyecto.**

---

## INDICE

<b>INTRODUCCION.....</b>	<b>iii</b>
<b>1 Antecedentes.....</b>	<b>1</b>
1.1 Origen de la Producción de Petróleo Crudo y Gas Natural.....	2
1.2 Sistema Nacional de Producción y Estadísticas de Hidrocarburos.....	2
1.2.1 Estadísticas de la Producción de Petróleo Crudo.....	3
1.2.2 Estadísticas de la Producción de Gas Natural.....	5
1.2.2.1 Aprovechamiento del Gas Natural.....	7
1.3 Proceso para la obtención de Productos Derivados.....	8
1.4 Utilidad de los Hidrocarburos.....	10
1.4.1 Mercado Nacional.....	10
1.4.2 Exportación e Importación.....	11
1.5 Situación Actual y Perspectivas.....	12
1.5.1 Distribución geográfica y cantidad de las reservas totales de hidrocarburos en 1995.....	14
<b>2 Metodología para el Análisis de Series de Tiempo.....</b>	<b>15</b>
2.1 Descripción de la metodología de Box y Jenkins.....	16
2.2 Identificación de modelos.....	19
2.2.1 Clasificación de Modelos para Series de Tiempo.....	25
2.3 Estimación de Parámetros de modelos.....	28
2.4 Verificación del modelo.....	36
2.5 Pronósticos.....	43
<b>3 Análisis de las Series de Tiempo para la Producción y las Ventas de Productos Derivados.....</b>	<b>45</b>
3.1 Series de Tiempo de Producción de Petróleo Crudo y Gas Natural y las Ventas de Productos Derivados.....	46
3.2 Comportamiento de la Serie Original, Identificación y Estimación de los modelos.....	46
3.2.1 Serie de la Producción de Petróleo Crudo.....	47
3.2.2 Serie de la Producción de Gas Natural.....	49
3.2.3 Serie de las Ventas de Productos Petrolíferos.....	52
3.2.3.1 Serie de las Ventas de Gas Natural.....	55
3.2.3.2 Serie de las Ventas de Gasolinas.....	57
3.2.3.3 Serie de las Ventas de Gas Licuado.....	60
3.2.3.4 Serie de las Ventas de Turbosinas.....	63
3.3 Diagnóstico y Selección de los modelos con mejor ajuste a las series analizadas.....	67
3.3.1 Producción de Petróleo Crudo y Derivados.....	68
3.3.2 Producción de Gas Natural y Derivados.....	74
<b>4 Pronósticos y Simulación de las Series de Tiempo de Producción y Ventas de Productos Derivados.....</b>	<b>77</b>
4.1 Pronósticos.....	78

---

---

4.1.1 Serie de la Producción de Petróleo Crudo.....	79
4.1.1.1 Serie de las Ventas de Productos Petrolíferos .....	80
4.1.1.1.1 Serie de las Ventas de Gas Licuado .....	81
4.1.1.1.2 Serie de las Ventas de Gasolinas .....	82
4.1.1.1.3 Serie de las Ventas de Turbosinas .....	83
4.1.2 Producción de Gas Natural.....	84
4.1.2.1 Serie de las Ventas de Gas Natural.....	84
4.2 Simulación del Fenómeno .....	85
4.2.1 Serie de la Producción de Petróleo Crudo.....	88
4.2.1.1 Serie de las Ventas de Productos Petrolíferos .....	89
4.2.1.1.1 Serie de las Ventas de Gas Licuado.....	90
4.2.1.1.2 Serie de las Ventas de Gasolinas.....	91
4.2.1.1.3 Serie de las Ventas de Turbosinas .....	92
4.2.2 Serie de la Producción de Gas Natural.....	93
4.2.2.1 Serie de las Ventas de Gas Natural.....	94
<b>5 Análisis de Resultados.....</b>	<b>95</b>
5.1 Evaluación de Resultados.....	96
<b>CONCLUSIONES.....</b>	<b>103</b>
<b>ANEXO.....</b>	<b>107</b>
<b>BIBLIOGRAFIA .....</b>	<b>113</b>

---

## INTRODUCCION

El sector energético ha sido sin duda un pilar fundamental en la conformación del México moderno. El desarrollo económico del país se ha visto impulsado por una oferta cabal y confiable de insumos energéticos, a la vez que el despliegue de actividades del sector han constituido un estímulo a la oferta nacional de bienes de capital, insumos industriales y servicios.

Dentro de este sector energético se encuentra Petróleos Mexicanos que es el proveedor de los hidrocarburos, fuente principal de la producción de energía primaria y que tiene dos destinos principalmente. En el primero de ellos la energía se canaliza hacia los centros de transformación (refinerías, despuntadoras, plantas de gas y fraccionadoras). Y en el segundo se utiliza directamente por el consumidor final (sectores de transporte, industrial, agropecuario, residencial, comercial y público), es decir, utilizan la energía secundaria que es generada por los productos derivados de los hidrocarburos después de que se llevó a cabo el proceso de transformación en los centros antes mencionados.

Como vemos el consumo de energético es de vital importancia para los diferentes sectores de la sociedad. De aquí la importancia de este trabajo el cual tiene como objetivo obtener modelos de series de tiempo de la producción y volumen de las ventas internas de algunos hidrocarburos para generar pronósticos que puedan ser útiles en la planeación y toma de decisiones, ya que realizar una planeación de mediano y largo plazo implica el tener conocimiento profundo y detallado del comportamiento de la sociedad en la producción, transformación, manejo y consumo de energía.

El análisis que se realizará para el desarrollo de modelos de series de tiempo parte de la idea de que el conjunto de datos históricos de la producción y ventas de hidrocarburos tienen un patrón que puede ser explotado en la preparación de pronósticos, y para identificar este patrón es conveniente pensar que estos datos están compuestos de varias componentes.

Es precisamente la Metodología de Análisis de Series de Tiempo de Box-Jenkins la que identifica esas componentes, ya que su idea fundamental radica en la estrategia para construir modelos, los cuales no sólo deben ser adecuados para representar el comportamiento de los datos observados, sino que su elección debe ser sugerida por los datos mismos. Es decir, la metodología detecta las distintas componentes usando los filtros correspondientes hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final.

La aplicación de esta metodología se debe a que utiliza la dependencia en las observaciones más efectivamente que otros modelos y en consecuencia es posible que produzca pronósticos más exactos. Además de que ofrece un mejor enfoque sistemático para

---

---

construir, analizar y pronosticar con modelos de series de tiempo y sobre todo porque se cuenta con la cantidad suficiente de datos que son necesarios para su desarrollo.

Por lo tanto, si se llegan a obtener pronósticos aceptables quizá puedan tomarse decisiones que contribuyan al fortalecimiento de las actividades productivas donde se asegure la capacidad de respuesta del sector en el suministro de combustible y energía en condiciones competitivas en calidad y volumen.

Otra finalidad de este trabajo es que sirva de apoyo al alumno del área de Simulación en algunas materias, ya que se presenta una aplicación real de uno de tantos conocimientos adquiridos durante la carrera y que pretende contribuir para el aprendizaje y experiencia del alumno de Matemáticas Aplicadas y Computación y áreas afines.

El trabajo esta integrado por cinco capítulos. En el primero de ellos se tiene el origen, evolución, situación y perspectivas de la producción y el volumen de ventas de productos petrolíferos.

En el segundo capítulo se describe la metodología de Box y Jenkins, considerada como una de las mejores para generar pronósticos. Se definen los pasos a seguir para la identificación estimación verificación y pronósticos de los modelos.

El capítulo tres contiene el análisis de las series de tiempo de producción de Petróleo Crudo y Gas Natural y Ventas de Productos Derivados. Este análisis incluye la identificación de las series estacionarias, las series con variación estacional, identificación de modelos tentativos que se ajusten a los datos y por último la selección de los mejores modelos en base a los estadísticos obtenidos y al análisis de residuales.

En el capítulo cuatro se tienen los pronósticos generados por los modelos seleccionados en el capítulo anterior. Así mismo se realiza una simulación con dichos modelos.

El quinto y último capítulo presenta la evaluación de los resultados, agrupando toda la información de los pronósticos, la simulación y los datos reales.

---

## **CAPITULO 1**

### **ANTECEDENTES**

---

## **1.1 Origen de la Producción de Petróleo Crudo y Gas Natural**

La historia de México como productor de hidrocarburos se tiene desde 1901, cuando se extraían por compañías petroleras extranjeras, en el periodo de ese año y hasta 1916 la producción es relativamente baja y su ritmo de crecimiento es modesto. Según archivos existentes la producción pasó de sólo 50 barriles diarios en 1901 a 110 mil barriles diarios en 1916 y siguió incrementándose hasta llegar en 1921 a una producción de 530 mil barriles diarios de crudo. Después de este año, aún cuando se inicia la producción de nuevos campos, ésta sufre un descenso radical debido a la explotación intensiva e irracional de los campos.

La expropiación petrolera marca el fin de una etapa en la que se llevó a cabo una explotación irracional de éstos recursos naturales; en la que fueron frecuentes las intromisiones de los intereses extranjeros en los asuntos internos del país y donde la existencia misma de las empresas petroleras extranjeras constituyeran uno de los principales obstáculos para el desarrollo económico nacional.

Junto con la estrangulación económica a que se vio sometida la industria petrolera nacionalizada en sus primeros años, existía la convicción, de parte de las compañías extranjeras expropiadas, de que los mexicanos estaban incapacitados técnica y económicamente para operar e impulsar a la industria petrolera nacionalizada. Han sido muchos los problemas que se han tenido que vencer; sin embargo, en su oportuna y gradual solución estriba una de las claves del poderoso desarrollo iniciado al finalizar la década de los treinta.

A pesar de los múltiples obstáculos, esta industria ha alcanzado un crecimiento extraordinario. Al cambiar la orientación que tenía antes de 1938, dejó de ser un instrumento de extracción de crudos para la exportación y se convirtió en la principal fuente de abastecimiento de las necesidades energéticas del país y de la propia industria.

## **1.2 Sistema Nacional de Producción y Estadísticas de Hidrocarburos**

En la actualidad el sistema nacional de producción de hidrocarburos esta integrado por tres regiones como se muestra en la figura 1.1.



Figura 1.1 Sistema Nacional de Producción de Hidrocarburos.

La producción de hidrocarburos esta integrada principalmente por dos tipos que son el Petróleo Crudo o Aceite y el Gas Natural.

### 1.2.1 Estadísticas de la Producción de Petróleo Crudo

El petróleo crudo se considera como pesado, ligero o superligero, según los criterios de la tabla 1.1:

Tipo de Petróleo Crudo	Densidad
Crudo Pesado	Igual o Inferior a 27°
Crudo Ligero	Superior a 27° y hasta 38°
Crudo Superligero	Superior a 38°

Tabla 1.1

Los hidrocarburos líquidos producidos, o sea el crudo más los condensados del gas obtenido en campo, alcanzó el valor más alto de nuestra historia en 1993 (2,774.4 MBD) . En relación a las calidades de crudo que se producen, cabe hacer notar que si bien la producción de aceite pesado o maya ha venido en aumento hasta 1992 (para 1993 y 1994 disminuye ligeramente) la producción de ligero es mayor, y se ha mantenido en los últimos años. Esto ha permitido un equilibrio de casi 50 % de cada tipo. A partir de 1988 se ha venido produciendo, una cantidad cada vez mayor de crudo super ligero u olmeca, que proviene de la región Sur

La producción de aceite se obtiene como resultado de dos tipos de proyectos; los de desarrollo, o sea aquella producción que se obtiene de los nuevos pozos productores pertenecientes a los nuevos campos que se descubren e incorporan año con año y la que se obtiene de los campos que están en etapa de explotación.

Es evidente que el principal soporte de la plataforma de producción es el aceite proveniente de estos últimos. En los años recientes la producción obtenida de los proyectos de explotación y mantenimiento atenúa su declinación e incluso ha aumentado en los últimos años.

En la figura 1.2 se muestra la producción histórica del Aceite de 1901 a 1994

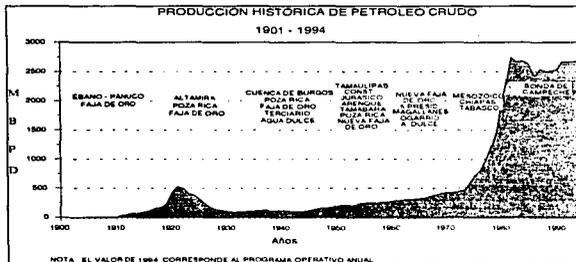


Figura 1.2 Producción Histórica de Petróleo Crudo.

La Región Norte sostiene una plataforma de aproximadamente 100 mil barriles por día en lo que se refiere al crudo.

Los Yacimientos de la Región Sur, por su estado de avance en su explotación, están en una etapa de fuerte declinación, que aunque ha sido atenuada con diversas acciones, requiere de inversiones en proyectos que permitan mantener su plataforma de producción, que actualmente es de 600 mil barriles por día de crudo.

La Región Marina, desde su incorporación a la producción nacional, ha ido incrementándose su participación, permitiendo contrarrestar su propia declinación, la de las otras dos regiones y mantener la plataforma nacional, en cifras superiores a los 2 millones 600 mil barriles por día, en los últimos años.

## 1.2.2 Estadísticas de la Producción de Gas Natural

El gas natural se obtiene de la separación del gas asociado al crudo y de la explotación de yacimientos propios de gas.

Dentro de los yacimientos, debido a las altas presiones y temperaturas existentes, los fluidos contienen un porcentaje de gas en solución, que al pasar a la superficie se libera al cambiar las condiciones a baja presión y temperatura, produciéndose de esta manera aceite y *Gas Asociado*.

En la Región Norte, las áreas productoras se encuentran en los distritos de: Altamira, Poza Rica, Veracruz, y en el sector operativo Cerro Azul, de donde se obtiene un total de 112 millones de pies cúbicos por día, mediante la explotación de 2,011 pozos en 116 campos, la aportación de esta región representa el 4 por ciento del total nacional de gas asociado.

La Región Sur aporta el 53 por ciento de la producción total de gas asociado, con un volumen diario de 1,644 millones de pies cúbicos, a través de la explotación de 1,337 pozos en 84 campos productores, localizados en el área del mesozoico Chiapas - Tabasco y en los distritos Agua Dulce y Ocosingo.

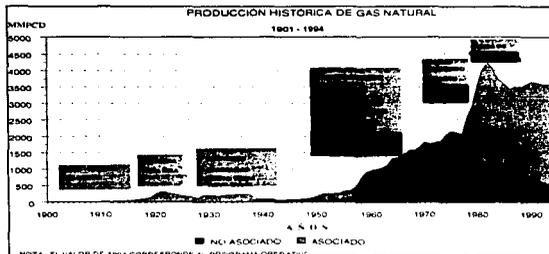
La Región Marina, se caracteriza por producir únicamente gas asociado al aceite y su aportación de 1,336 millones de pies cúbicos por día, corresponde al 43 por ciento de la producción nacional, de gas asociado, obteniéndose mediante la operación de 308 pozos localizados en 15 campos productores, pertenecientes a los distritos de Ciudad del Carmen y Dos Bocas, localizados en la Sonda de Campeche.

Se denomina *Gas No Asociado* al que se obtiene de yacimientos de gas seco y húmedo, con bajo contenido de hidrocarburos líquidos.

En la Región Norte, las áreas de gas no asociado se localizan en los distritos de Reynosa y Veracruz, de donde se obtienen 360 millones de pies cúbicos por día, de 763 pozos en 115 campos productores, este volumen representa el 68 por ciento de la producción nacional.

En la Región Sur se obtiene el 32 por ciento restante de gas no asociado, 170 millones de pies cúbicos por día, proveniente de 148 pozos de 17 campos localizados en los distritos de Agua Dulce y Ocosingo.

En la figura 1.3 se muestra la producción histórica del gas natural en el periodo de 1901 a 1994.



**Figura 1.3 Producción Histórica de Gas Natural.**

De la producción nacional el 84 por ciento corresponde a gas asociado y el restante 16 por ciento a gas no asociado, la razón principal de esta diferencia de producción es que el gas asociado se extrae junto con el crudo, hacia el cual se han dirigido la mayor parte de recursos para su búsqueda y explotación, motivo por el cual el gas no asociado ha sufrido un decremento a partir de 1982 al quedar limitado el desarrollo de los proyectos relativos a este tipo de yacimientos.

Esta orientación obedece a que los recursos de inversión que el gobierno federal ha orientado en los últimos años a la rama de la industria petrolera nacional que se encarga de la exploración y producción de hidrocarburos, se han canalizado preferentemente a los proyectos de explotación de aceite, debido a que esto hasta ahora han sido mas rentables que los proyectos de gas natural, en razón de su significativa diferencia de precios.

En apego a esta política se ha dado preferencia a la incorporación de yacimientos productores de crudos ligeros, de mayor valor comercial, que a su vez traen consigo superiores volúmenes de gas asociado.

En consecuencia, las cifras históricas reflejan que no obstante el apego a la política de mantener una plataforma de producción de crudo y dejando al gas natural en una posición relativamente complementaria, la producción de este hidrocarburo no se ha deteriorado, e incluso se ha elevado modestamente en los últimos años, compensando la declinación natural.

### 1.2.2.1 Aprovechamiento del Gas Natural

Durante los primeros años de la explotación del petróleo, el gas era considerado mundialmente como un subproducto de poco valor comercial, por lo que se desperdiciaba en su mayor parte, quemándolo a la atmósfera.

México no fue ajeno a esta práctica y desde el descubrimiento de los campos de crudo y gas asociado en 1901 hasta finales de la década de los cuarenta, el gas producido se envió a la atmósfera, pues no se disponía de los recursos necesarios para su aprovechamiento, y por otra parte, tampoco se contaba con la demanda para este producto. Es a finales de la década de los 40's cuando se inicia el aprovechamiento del gas, al entrar en operación las primeras estaciones de compresión y plantas de tratamiento en Poza Rica.

Con el descubrimiento de los campos de gas no asociado, en el distrito frontera noreste a partir de 1945 y en Ciudad Pemex a partir de 1949, y la construcción de los gasoductos de distribución, permitieron abrir importantes mercados para este producto.

Sin embargo, y con el incremento de la producción de gas a partir de 1958 y hasta el año de 1968, se enviaron a la atmósfera un promedio diario anual del orden de 600 millones de pies cúbicos por día.

No obstante los esfuerzos que se hicieron a partir de 1969 para la adquisición e instalación de equipo de compresión en campos, con el incremento en la producción de gas por el descubrimiento de los campos del mesozoico en los estados de Chiapas y Tabasco en 1973, se continuó con la quema de gas con valores del orden de los 500 millones de pies cúbicos por día; aunque en términos porcentuales el gas aprovechado pasó de un valor de 46 por ciento en 1968 a 92 por ciento en 1979.

Con el descubrimiento y desarrollo de los campos de la Sonda de Campeche en 1979, y con la máxima producción del mesozoico en la región sur, durante los años de 1980 a 1983, los volúmenes de gas quemado se incrementaron nuevamente disminuyendo el gas aprovechado en términos porcentuales; y no fue sino hasta 1984 y con el inicio de operación de las plataformas de compresión Akal C, Akal J y Nohoc A, durante los años de 1981 a 1983, cuando se inicia un incremento real del gas aprovechado, llegando a valores del 98 por ciento a partir de 1990.

Podemos observar en la gráfica el aprovechamiento del gas natural durante la historia

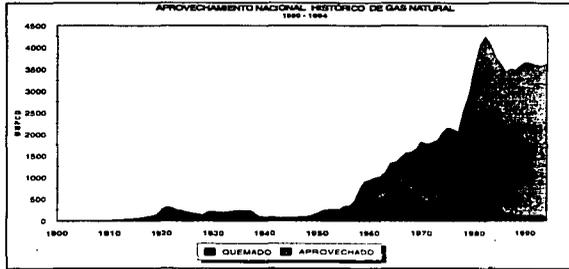


Figura 1.4 Aprovechamiento histórico de Gas Natural.

Un aspecto histórico a considerar lo constituye el crecimiento acelerado de la industria petrolera en el período 1976-1981. En esta época, la intención de aprovechar las oportunidades que ofrecía el mercado internacional de petróleo crudo, así como el descubrimiento de los yacimientos del mesozoico de Chiapas-Tabasco y de la sonda de Campeche, canalizaron cuantiosas inversiones para levantar una infraestructura destinada a la producción de aceite.

Al presentarse la presión sobre la deuda externa y modificarse las características del mercado, el cual paso de un patrón de vendedores a uno de compradores, no fue posible aprovechar en su totalidad el gas asociado que se producía en estos yacimientos, el cual se quemaba para producir el aceite a los ritmos que demandaba el país.

### 1.3 Proceso para la obtención de Productos Derivados

La refinación es el conjunto de una serie de procesos físicos y químicos a los que se somete el petróleo crudo, la materia prima, para obtener de él, por destilación, los diversos hidrocarburos o las familias de hidrocarburos con propiedades físicas y químicas bien definidas.

Después de la separación se aplican a los derivados así obtenidos diversos procesos de conversión para obtener de ellos productos más valiosos y éstos se someten finalmente a tratamientos con ácidos, álcalis, solventes extractivos, catalíticos con hidrógeno y reactivos químicos en general, a fin de eliminar las impurezas que los hacen impropios para su empleo comercial.

El aceite crudo, de muy diversa constitución según el origen (de base asfáltica, nafténica o mezclada), tiene rendimientos variables en el proceso de destilación y de fraccionamiento, a determinadas condiciones de presión y temperatura. Estos rendimientos generalmente no concuerdan con el patrón de consumo, en el cual en algunos casos, según el país de que se trate, presenta diversos requerimientos de productos ligeros, de peso molecular que no están contenidos en el aceite crudo, o por el contrario, productos residuales con alto peso molecular.

Por lo tanto, es necesario ajustar los rendimientos y características de las fracciones o cortes que constituyen los diferentes combustibles, al mencionado patrón de consumo. Este ajuste se hace sometiendo las fracciones a los diversos procesos de conversión, con objeto de obtener los productos que el mercado requiere. Tales procedimientos de conversión se aplican a las diferentes familias de hidrocarburos para obtener, mediante arreglos moleculares, productos más ligeros o de mayor cantidad de octano, o de menor viscosidad.

Es por la aplicación de estos procesos como la refinación puede poner a disposición del consumidor, una amplia gama de productos comerciales:

- a) Energéticos; combustibles específicos para los transportes, la agricultura, la industria, la generación de corriente eléctrica para uso doméstico.
  - b) Productos especiales; lubricantes, parafinas, asfaltos, grasas para vehículos, construcción y uso industrial.
  - c) Materias primas para la industria petroquímica básica.
- Para lograr lo anterior, como se ha descrito ya, es necesario someter las materias primas a una serie de pasos de transformación, los cuales se dividen en tres grupos principales.

#### *Proceso de destilación del petróleo crudo*

El petróleo crudo está formado por varios hidrocarburos que comprenden desde el gas licuado hasta el asfalto. Su separación en columnas de destilación se realiza por las diferencias de volatilidad que tienen unos y otros. El procedimiento utilizado consiste en calentar el petróleo crudo a una temperatura en que los componentes más ligeros se evaporan, para ser enseguida condensados. Esta condensación se efectúa a diferentes temperaturas. Los hidrocarburos más volátiles se condensan a menor temperatura que los menos volátiles.

De esta manera se obtienen distintos condensados, cuyas propiedades corresponden a las de gas licuado, gasolinas, kerosinas o combustible diesel.

#### *Proceso de desintegración*

El residuo de la destilación del petróleo crudo se somete a una nueva destilación al alto vacío para separar componentes menos volátiles, que de acuerdo a las propiedades del petróleo crudo de que se trate, serán destinados a lubricantes o a ser desintegrados catalíticamente. El residuo de la destilación al vacío es asfalto, o bien, carga para la planta

de coque o para la hidrodeseintegradora de residuales y la subsecuente obtención de destilados.

Los destilados al vacío que se destinan a lubricantes se someten a procesos adicionales: extracción con furfural y desparafinación con metil-etil-cetona. Al final se obtiene de ellos lubricantes básicos, que con diferentes aditivos, forman los lubricantes y las parafinas que existen en el mercado.

Los destilados al vacío que por sus características no se dedican a lubricantes, se desintegran catalíticamente para convertirse en productos comerciales: gas licuado, gasolinas de alto índice de octano y combustible diesel.

#### *Proceso de purificación*

Estos procesos eliminan de los productos obtenidos por destilación o por desintegración, algunos compuestos que les causan propiedades inconvenientes. Los principales contaminantes en estos procesos son los compuestos derivados del azufre. Los inconvenientes que se presentarían los derivados del petróleo sin estos tratamientos, serían mal olor y contaminación de la atmósfera al ser quemados.

Los trabajos de refinación transcurren continuamente durante todas las horas y todos los días. La producción obtenida tiene que ser oportunamente distribuida, tarea que se lleva a cabo por todos los medios de transporte conocidos: poliductos, buquetanques, autotanques y carrotanques.

## **1.4 Utilidad de los Hidrocarburos**

### **1.4.1 Mercado Nacional**

Tradicionalmente el crudo producido en México ha tenido como principal destino, cubrir las necesidades del área de refinación, además de la exportación de volúmenes remanentes generando divisas para el país.

La entrega de crudo a plantas ha sido en constante aumento, alcanzando en 1993 la cifra de 1 millón 306 mil barriles por día y el programa de 1994 es de 1 millón 338 mil.

La oferta de gas natural desde 1980, ha sido del rango de los 3,200 millones de pies cúbicos por día. Este volumen de gas es el que PEMEX Exploración y Producción (PEP) ha entregado a la subsidiaria Gas y Petroquímica Básica para su procesamiento y venta.

El gas natural juega un papel importante para sustentar las actividades de las subsidiarias de Petróleos Mexicanos; se utiliza como materia prima para la industria petroquímica, como combustible, y fluido de empuje en sistemas artificiales de producción de petróleo crudo.

El consumo de PEP y el total de PEMEX, desde 1988 hasta el presente has sido del orden de 1500 millones de pies cúbicos diarios, lo cual representa aproximadamente un 40% de la producción total de gas .

En el país, la venta de gas producido por PEP a los sectores industrial, eléctrico y doméstico, se han incrementado notablemente en los últimos 5 años. Actualmente, del gas comercializado en el mercado nacional, el sector industrial consume el 63%, la generación de electricidad 30% y el doméstico 7%.

En términos generales, para darnos cuenta de la gran importancia de los hidrocarburos diremos que éstos participan actualmente con el 90.7% de la producción de energía primaria, misma que es útil para la generación de energía secundaria que es aprovechada por los consumidores finales que integran los sectores industrial, del transporte, comercial, residencial, etc.

#### 1.4.2 Exportación e Importación

Respecto a la exportación, las cifras se han mantenido en un promedio de 1 millón 326 mil barriles por día durante los últimos seis años, siendo la característica principal del periodo, el constante aumento del volumen de crudo olmeca, ya que de 72 mil barriles por día exportados en 1988 se llegó a 220 mil en 1993 y el programa para 1994 es el de exportar 371 mil barriles por día.

Por ser el olmeca el tipo de crudo mexicano de mayor precio, su creciente participación en la exportación representa un incremento de divisas para el país.

Por lo que toca al gas natural se han exportado pequeñas cantidades únicamente hacia los Estados Unidos de Norteamérica, y sujetas a variaciones temporales en la producción. Se inició en la década de los 40's a través de la ciudad fronteriza de Miguel Alemán, y desde 1957 a través de Reynosa, hacia el sistema de la compañía Texas Eastern Transmission Corporation.

En 1993 y 1994 se exportaron excedentes como consecuencia del incremento en la disponibilidad de gas por el efecto combinado de aumento de la producción, la racionalización del uso de la energía en PEMEX, y de un abundante temporada de lluvias que favoreció la generación de hidroelectricidad.

PEMEX ha requerido importar pequeñas cantidades de gas natural para cubrir la demanda de los consumidores del noroeste del país, debido a la falta de infraestructura para transportarlo a dicha área geográfica, la cual resultaría costosa, y la importación es conveniente. Esta ha sido además propiciada por el incremento de la demanda de gas en los últimos años.

## 1.5 Situación Actual y Perspectivas

En 1995, el comportamiento del mercado interno de hidrocarburos fue consistente con la recesión económica del país. El producto interno bruto registró un descenso del 6.9 por ciento, los sectores industriales y de la construcción mostraron la caídas más pronunciadas. Las ventas internas de productos petrolíferos y gas natural descendieron 6 por ciento, con relación a 1994. Una contracción de esta magnitud no se había presentado en la historia de Petróleos Mexicanos. La devaluación del peso obligó a ajustar los precios internos de productos petrolíferos y gas natural. Los precios del petróleo se elevaron, permitiendo obtener un aumento de 1 146 millones de dólares en el valor de las exportaciones netas.

La comercialización de petrolíferos se enfocó a satisfacer en forma eficiente y oportuna los requerimientos de la demanda interna, optimizando sus sistemas de distribución y adecuándolos a las necesidades de la actividad económica del país; además de continuar con la introducción en el mercado de productos con mayor calidad.

La demanda interna de petrolíferos y de gas natural fue en conjunto de 1661 MBD, cifra 6 por ciento inferior a la de 1994. Las ventas internas de gas natural promediaron 229 MBD equivalentes a combustóleo, 7 por ciento más que las registradas en el año inmediato anterior, como resultado de la sustitución de combustóleo por gas natural en el sector eléctrico. La demanda de productos petrolíferos fue de 1432 MBD, registrando una contracción de 8 por ciento. Resalta la depresión de los mercados de diesel y combustóleo. El primero registró una caída de 8 por ciento, reflejando una reducción importante en la demanda para transporte. En el caso del combustóleo, la baja en su demanda por parte del sector eléctrico y de la industria cementera, así como su sustitución por gas natural dado el menor precio relativo de este último, resultaron en una disminución cercana al 14 por ciento en las ventas.

En 1995 el valor de las ventas de los productos petrolíferos valuadas a precios Pemex ascendió a 58 944 millones de pesos, superior en 28 por ciento en términos reales a las registradas en el año previo. Un mayor nivel de precios nominales compensó el menor volumen vendido. Sus ventas de gas natural fueron de 5191 millones de pesos, monto que en términos reales fue del 11 por ciento mayor al obtenido en 1994.

En el mismo año, la extracción de petróleo crudo y de gas natural se llevó a cabo cuidando el desarrollo óptimo de los yacimientos. Uno de los objetivos fue elevar la rentabilidad de la producción de hidrocarburos y aprovechar al máximo las reservas nacionales mediante la aplicación de tecnologías avanzadas y sistemas más eficientes de administración de los procesos productivos.

La producción de hidrocarburos se orientó hacia la satisfacción de la demanda interna y a la generación de excedentes para la exportación. En 1995, la producción de crudo promedió 2617 MBD, lo que reflejó una reducción de casi 3 por ciento con respecto al volumen registrado en 1994.

La región Marina aportó 74%, la Sur 22% y la Norte 4%. El tipo de crudo extraído mantuvo la misma proporción que en 1994: el 53% de crudos ligeros e intermedios y el 47% restante correspondió al pesado.

La producción de gas natural promedió 3759 MMPCD, volumen superior en 4% al observado en 1994, pese a los problemas que se presentaron por los efectos climáticos. En diciembre la producción superó los 4000 MMPCD, nivel no alcanzado desde 1983. El crecimiento en la producción de gas natural responde a la continua expansión de la demanda interna que se apoya en el desarrollo industrial del país y en la modificación de los patrones de consumo que se prevé para los próximos años, al considerar la sustitución del combustible por gas natural.

Por tipo:

El gas Asociado representó el 84% de la producción total de gas natural mientras que el no asociado fue de sólo 16%.

Por Región :

La norte contribuyó con el 14%, la Marina con el 37% y la Sur con el 49% del total de la producción de gas natural.

Con el propósito de estar en la posibilidad de responder a la demanda de gas natural y evitar la importación de volúmenes significativos de este producto en el mediano plazo, en 1995 PEP orientó sus actividades al desarrollo de yacimientos de gas no asociado y a los de alta relación de gas-aceite. En el caso de los primeros destacan los distritos de Reforma en Chiapas y de Reynosa en Tamaulipas.

### **1.5.1 Distribución geográfica y cantidad de las reservas totales de hidrocarburos en 1995.**

A fines de 1995, las reservas probadas de petróleo crudo equivalente con las que contó México fueron de 62 058 millones de barriles; de éstas, 79% corresponden a aceite y condensados y el 21% restante a gas seco. De las reservas de aceite, 52 por ciento se refieren a crudos ligeros. A los ritmos actuales de producción, la vida media de estas reservas se estima en 48 años.

En la Región Marina se localiza 46% de las reservas totales de hidrocarburos; allí se concentra 56% de las reservas probadas de aceite y el 17% de las de gas. En esta región se encuentran los yacimientos más importantes del país. La Región sur acumula 29% de las reservas de gas y 15% de las de aceite. En la Región Norte se ubica la tercera parte de las reservas de hidrocarburos, y se concentra 54% de las reservas totales de gas y 29% de las de aceite.



## 2.1 Descripción de la metodología de Box y Jenkins

Una gran cantidad de información acerca de las características económicas, tanto de individuos como empresas o países, es comúnmente recopilada con fines de análisis, para posteriormente llevar a cabo planeación y toma de decisiones. Al registro metódico de la medición u observación numérica, efectuada a intervalos de tiempos fijos, de tales características o variables económicas, es a lo que generalmente se le conoce como *Serie de Tiempo*.

Es importante notar que una serie de tiempo observada no es más que una realización de un proceso estocástico, lo cual significa que bien pudo haberse observado otra realización del mismo proceso, pero cuyo comportamiento fuese distinto del que se observó en la realidad.

El cuerpo de técnicas para el análisis de las series con observaciones dependientes es llamado *Análisis de Series de Tiempo*.

Este análisis tiene 4 objetivos básicamente:

- 1.- Descripción: simplemente enunciar el comportamiento de un proceso permitiendo una visualización clara.
- 2.- Explicación; permitiendo un análisis de tipo causa-efecto, al relacionar un fenómeno con el medio circundante, p. ej., puede explicarse la contaminación de acuerdo al día de la semana, con la cantidad de autos en circulación.
- 3.- Pronóstico; en el caso en que se desee estimar la cantidad de turismo que ingresa para preparar las condiciones de los sitios turísticos.
- 4.- Control; si no sólo se desea observar sino modificar el comportamiento para obtener algún beneficio.

Así pues, se construirán modelos estocásticos para series de tiempos discretas en el espacio tiempo, donde los modelos derivados serán muy simples<sup>1</sup> y tendrán el mínimo de parámetros con una representación adecuada. La obtención de tales modelos será de importancia debido a que:

- Ellos pueden decirnos algo sobre la naturaleza del sistema de generación de las series de tiempo;
- Pueden ser usados para obtener pronósticos óptimos de valores futuros de las series.

La exactitud de los pronósticos puede ser mejorada empleando un modelo que exprese a la serie de tiempo  $y_t$  como una función del error aleatorio actual y el anterior. Así consideramos el modelo

$$y_t = \mu + \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

<sup>1</sup>Principio de Parsimonia: el cual indica que se debe elegir el modelo con el menor número de parámetros.

donde  $\varepsilon_t, \varepsilon_{t-1}, \dots$  son los errores actual y anterior y donde  $\psi_0, \psi_1, \psi_2, \dots$  son los parámetros del modelo. En este modelo se supone que los errores se distribuyen normalmente, tienen media cero y varianzas iguales. Sin embargo, es claro que la sucesión de valores de las series de tiempo en este modelo son dependientes, porque ellas están en función de algún error en común. Box y Jenkins han desarrollado un procedimiento sistemático para analizar y pronosticar series de tiempo usando este modelo. Este procedimiento se le ha llamado la metodología de Box y Jenkins.

Su idea fundamental propuesta radica en la estrategia que ellos proponen para construir modelos, los cuales no sólo deben ser adecuados para representar el comportamiento de los datos observados, sino que su elección debe ser sugerida por los datos mismos.

La metodología de Box-Jenkins es muy útil para pronosticar los valores de las series de tiempo teniendo valores que son estadísticamente dependientes o correlacionados. Consiste de 4 pasos que son : identificación, estimación, diagnóstico, pronóstico.

El método consiste en extraer los movimientos predecibles de los datos observados. La Serie de Tiempo se descompone en varios componentes, algunas veces llamados "filtros", precisamente porque la filosofía del método consiste en detectar las distintas componentes usando los filtros correspondientes, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final. El enfoque de Box-Jenkins hace principalmente uso de 3 filtros lineales: el autorregresivo, el de integración y el de medias móviles.

El desarrollo consiste de varios pasos iterativos. El primer paso es donde se identifica el modelo tentativo analizando los datos históricos. Para llevar a cabo esta identificación se utilizarán conceptos como la función de autocorrelación (ACF), la función de autocorrelación parcial (PACF), series estacionarias y no estacionarias. Los procedimientos empleados en la identificación son inexactos y requieren de mucho razonamiento y experiencia, así como del conocimiento del fenómeno en cuestión. La idea es elegir un modelo y tratar de mejorarlo.

El segundo es la estimación de los parámetros desconocidos. El tercer paso es el diagnóstico donde se realizan las pruebas para la adecuación del modelo y si es necesario sugerir formas para mejorarlo. Para hacer el diagnóstico es importante el estudio de los residuales.

Una vez que el modelo de la serie de tiempo ha sido desarrollado, un cuarto paso, llamado pronosticar, genera las predicciones de valores futuros de la serie.

Es indispensable hacer notar que hay modelos que se ajustan a los datos pero que no generan pronósticos aceptables. En este caso hay que regresar a la identificación, o inclusive, revisar si el método es realmente adecuado. Por otro lado, de acuerdo al objetivo final, puede sacrificarse el ajuste por la precisión en los pronósticos.

En la práctica, los modelos de regresión y alisamiento exponencial son frecuentemente aplicados con buenos resultados para pronosticar valores de series de tiempo con dependencia u observaciones autocorrelacionadas. Sin embargo, puesto que la metodología de Box-Jenkins utiliza la dependencia en las observaciones más efectivamente que los modelos anteriores es posible que produzca pronósticos más exactos. Más aún, la metodología ofrece un mejor enfoque sistemático para construir, analizar y pronosticar con modelos de series de tiempo.

Los modelos que se verán aquí se basan en la idea de que una serie de tiempo, cuyos valores sucesivos pueden ser altamente dependientes, puede considerarse generada a partir de una serie de choques aleatorios independientes  $\{a_t\}$ . Estos choques aleatorios se supone que son realizaciones independientes de una variable aleatoria cuya media es constante (generalmente se le considera igual a cero) y cuya varianza es  $\sigma^2$ . A esta sucesión de variables aleatorias  $\{a_t\}$  se le conoce como *proceso de ruido blanco*<sup>2</sup> (o tan solo ruido blanco)

Los datos de series de tiempo son frecuentemente examinados esperando descubrir un patrón histórico que pueda ser explotado en la preparación de pronósticos. Para identificar este patrón es conveniente pensar que las series de tiempo están compuestas de varios componentes como:

- Tendencia
- Ciclos
- Variación Estacional
- Fluctuaciones Irregulares

La **Tendencia** se refiere a los movimientos ascendentes y descendentes que caracterizan a una serie de tiempo a través de un periodo de tiempo. Así, la tendencia refleja un gran crecimiento o decline en las series de tiempo.

Los **Ciclos** se refieren a los movimientos recurrentes y descendentes alrededor de los niveles de la tendencia, generalmente son mayores de un año.

La **Variación Estacional** son patrones periódicos en una serie de tiempo que se completan dentro del periodo de tiempo, generalmente al año.

Las **Fluctuaciones Irregulares** son movimientos erráticos en una serie de tiempo que no sigue un patrón regular o reconocible. Tales movimientos representan lo que se conoce como residuales en la serie después de eliminar la tendencia, ciclos y/o variación estacional. Muchas fluctuaciones irregulares son causadas por eventos no usuales que no pueden ser pronosticados. También pueden ser causados por errores en la partes del análisis de las series.

---

<sup>2</sup> Dicha terminología surge del hecho de que el concepto de ruido blanco se utiliza frecuentemente, en la Ingeniería Electrónica y de Comunicaciones, en estudios donde se pretende distinguir a las señales con mensaje de los simples ruidos.

## 2.2 Identificación de modelos

La identificación de un modelo tentativo usado para pronosticar series de tiempo es hecho a través del análisis de datos históricos.

Esta etapa tiene como objetivo principal determinar los órdenes de los polinomios autorregresivos y de promedios móviles, así como el número de veces que deberá aplicarse el operador diferencia para cancelar la no estacionariedad homogénea. De manera más general podría decirse que la etapa de identificación consiste en determinar, primero una serie estacionaria en función de la serie original, para la cual se pueda tener una representación con los modelos de procesos autorregresivos y de medias móviles con parámetros  $p$  y  $q$ , ARMA( $p,q$ ), y posteriormente, en fijar los valores de  $p$  y  $q$ . A continuación veremos la forma de identificar una serie estacionaria o no estacionaria

### *Series de tiempo Estacionarias y no Estacionarias*

Considérese una serie de tiempo que puede ser descrita por el modelo

$$y_t = \mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$$

Este modelo puede representar efectivamente las series de tiempo estacionarias y no estacionarias. Cuando la serie es estacionaria, sus valores fluctúan alrededor de la media. Esta identificación es llevada a cabo a través del análisis de  $n$  valores observados  $y_1, y_2, \dots, y_n$  de las series de tiempo. Si la identificación indica que la serie es no estacionaria, es importante transformar los valores de la serie no estacionaria a valores que puedan ser descritos por series de tiempo estacionarias. Si la serie no posee variación estacional, los valores de la serie estacionaria pueden modificarse por medio de una transformación que involucra tomar las primeras diferencias de la serie de tiempo original  $y_1, y_2, \dots, y_n$ . Esto es,

$$z_t = \Delta y_t = y_t - y_{t-1} \quad \text{para } t=2, \dots, n$$

A continuación se lista los valores originales y los transformados con la primera diferencia de la serie de tiempo.

Valores originales	Primera Diferencia
$y_1$	
$y_2$	$z_2 = \Delta y_2 = y_2 - y_1$
$y_3$	$z_3 = \Delta y_3 = y_3 - y_2$
.	.
$y_{n-1}$	.
$y_n$	$z_n = \Delta y_n = y_n - y_{n-1}$

Cuando se trata con datos de negocios y económicos que no poseen variación estacional, si los valores de la serie son no estacionarios, los valores transformados  $\Delta y_2, \Delta y_3, \dots, \Delta y_n$  son usualmente estacionarios. Sin embargo, si los valores transformados todavía no son estacionarios, tomando las segundas diferencias de la serie original  $y_1, y_2, \dots, y_n$  usualmente producirá valores de la serie estacionarios. Esto es,

$$z_t = \Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2} \quad \text{para } t=3, \dots, n$$

Valores originales	Segunda Diferencia
$y_1$	
$y_2$	
$y_3$	$z_3 = \Delta^2 y_3 = y_3 - 2y_2 + y_1$
$y_4$	$z_4 = \Delta^2 y_4 = y_4 - 2y_3 + y_2$
$\vdots$	$\vdots$
$y_{n-1}$	$\vdots$
$y_n$	$z_n = \Delta^2 y_n = y_n - 2y_{n-1} + y_{n-2}$

#### Autocorrelación y Autocorrelación Parcial

Se tienen los valores de  $z_a, z_{a+1}, \dots, z_n$ , los cuales se supone han sido generados por una serie estacionaria. Note que no se han escrito los valores de  $z_1$ , debido a que los valores de  $z_a, z_{a+1}, \dots, z_n$  podrían ser las primeras o segundas diferencias de una serie no estacionaria  $y_1, y_2, \dots, y_n$ . Si los valores de  $z_a, z_{a+1}, \dots, z_n$  son la primera diferencia, entonces

$$z_a = z_2 = \Delta y_2 = y_2 - y_1$$

en el caso de que  $a = 2$ ; considerando que los valores de  $z_a, z_{a+1}, \dots, z_n$  son las segundas diferencias, entonces

$$z_a = z_3 = \Delta^2 y_3 = y_3 - 2y_2 + y_1$$

en el caso de que  $a=3$ . Por supuesto si los valores de la serie original son estacionarios, entonces  $z_a = y_1$ , en el caso de que  $a=1$ .

Una implicación importante de esta suposición de que los valores de  $z_a, z_{a+1}, \dots, z_n$  son estacionarios es que las propiedades estadísticas de la serie no son afectadas por el cambio en el origen del tiempo. Esto significa, por ejemplo, que las relaciones estadísticas entre  $n$  observaciones al origen  $t$ , decir que  $z_t, z_{t+1}, \dots, z_{t+n-1}$ , son las mismas que las del origen  $t+j$   $z_{t+j}, z_{t+j+1}, \dots, z_{t+j+n-1}$ . Una de estas importantes relaciones es medida por  $\rho_k$ , la cual es la autocorrelación entre dos observaciones de la serie separadas por un intervalo de  $k$  unidades de tiempo. Se puede mostrar que  $\rho_k$  es adimensional, y  $-1 \leq \rho_k \leq 1$ ,  $\rho_k = \rho_{-k}$  lo cual implica que es necesario considerar sólo los intervalos positivos. Aunque  $\rho_k$  es un

parámetro que no puede ser conocido con certeza,  $\rho_k$  puede ser estimado usando las observaciones  $z_1, z_2, \dots, z_n$ . El estimador de  $\rho_k$  es llamado autocorrelación muestral en el intervalo  $k$ , es denotado por el símbolo  $r_k$ , y es dado por la fórmula

$$r_k = \frac{\sum_{i=1}^{n-k} (z_i - \bar{z})(z_{i+k} - \bar{z})}{\sum_{i=1}^{n-k} (z_i - \bar{z})^2}$$

donde  $\bar{z}$  es el promedio de las observaciones  $z_1, z_2, \dots, z_n$  y es dado por la fórmula

$$\bar{z} = \frac{\sum_{i=1}^n z_i}{n - a + 1}$$

La función de autocorrelación (ACF) teórica es definida por los valores de  $\rho_k$  para los intervalos de  $k=1, 2, \dots$ . La función de autocorrelación muestral son los valores de  $r_k$  para los intervalos de  $k=1, 2, \dots$ . La función de autocorrelación teórica de una serie estacionaria tiende a decrecer al incrementarse los intervalos de  $k$  o se trunca después de un intervalo particular  $k = q$ . Cuando decimos que la función de autocorrelación tiende a truncarse después de un valor de  $k = q$  significa que

$$\rho_k = 0 \quad \text{para } k > q$$

Se puede determinar cuando la ACF teórica se trunca usando la ACF muestral. Sin embargo, aunque la ACF teórica puede ser tal que

$$\rho_k = 0 \quad \text{para } k > q$$

debido a la variación mostrada, la ACF muestral  $r_k$  probablemente será pequeña, pero no igual a cero, para  $k > q$ . La pregunta es: ¿Que tan pequeña debe ser  $r_k$  para concluir que  $\rho_k = 0$ ?

Una regla de decisión para poder concluir si  $\rho_k = 0$  para  $k > q$  es

$$|r_k| \leq 2 \frac{1}{(n - a + 1)^{1/2}} \left( 1 + 2 \sum_{j=1}^q r_j^2 \right)^{1/2} \quad \text{para } k > q$$

Otro método que es usado frecuentemente para decidir si se puede concluir que  $\rho_k = 0$  es calcular el estadístico "t"

$$t_{r_k} = \frac{r_k}{s_{r_k}}$$

donde

$$s_{r_k} = \frac{1}{(n-a+1)^{1/2}} \left( 1 + 2 \sum_{j=1}^a t_j^2 \right)^{1/2}$$

Como una regla de decisión, puede ser concluido que

$$\rho_k = 0 \quad \text{si} \quad \left| t_{r_k} \right| = \left| \frac{r_k}{s_{r_k}} \right| \leq 2$$

Contrariamente, como una regla de decisión, si  $|t_{r_k}| > 2$  es razonable concluir que  $\rho_k \neq 0$ . Así, por ejemplo, si  $|t_{r_k}| > 2$  para los intervalos 1, 2, ..., q y si  $|t_{r_k}| \leq 2$  para intervalos  $q+1, q+2, \dots$ , es razonable concluir que la función de autocorrelación teórica se corta después de q intervalos.

Con esto puede mostrarse que si una serie es no estacionaria, entonces la ACF muestral no se truncará ni decrecerá rápidamente, sino más bien decrecerá muy lentamente. Por consiguiente, si se nos da un conjunto de observaciones de una serie de tiempo  $y_1, y_2, \dots, y_n$ , primero se calculará la ACF muestral para el conjunto

$$z_a = z_1 = y_1$$

$$z_{a+1} = z_2 = y_2$$

$$z_n = y_n$$

Si la ACF muestral de las observaciones originales decrece o se corta muy rápidamente, podemos asumir que la serie original es estacionaria. Sin embargo si esta decrece muy lentamente, supondremos que se trata de una serie no estacionaria. El preciso significado de los términos "decrece rápido" y "extremadamente lento" son algo arbitrario que se puede determinar mejor a través de la experiencia.

Si la ACF muestral de la serie transformada con las primeras diferencias decrece o se corta muy rápidamente, podemos decir que las primeras diferencias representan a una serie estacionaria. De otro modo no son estacionarias. Por lo que será necesario aplicar una segunda diferencia, es muy raro que una serie de tiempo se le apliquen más de dos diferencias.

Una vez que los valores de la serie  $y_1, y_2, \dots, y_n$  han sido encontrados o transformados a valores estacionarios  $z_a, z_{a+1}, \dots, z_n$ , es importante identificar un modelo de series de tiempo estacionarias que pueda asumirse que ha generado las observaciones  $z_a, z_{a+1}, \dots, z_n$ .

Otra herramienta que se utiliza para decidir si un modelo de series de tiempo estacionarias ha generado las observaciones  $z_a, z_{a+1}, \dots, z_n$  es  $\rho_{kk}$ , la cual es la autocorrelación parcial entre

2 observaciones separadas por un intervalo  $k$  unidades de tiempo. La autocorrelación parcial puede pensarse que es la autocorrelación de dos observaciones,  $z_t$  y  $z_{t+k}$ , separadas por un intervalo de  $k$  unidades de tiempo, con los efectos de las observaciones intermedias  $z_{t+1}, z_{t+2}, \dots, z_{t+k-1}$  eliminadas. Esto muestra que  $\rho_{kk} = 0$ . Aunque  $\rho_{kk}$  es un parámetro que no se conoce con certeza, puede ser estimado usando las observaciones muestrales  $z_n, z_{n+1}, \dots, z_n$ . El estimador de  $\rho_{kk}$  es llamado "la autocorrelación parcial muestral en el intervalo  $k$ ", es denotado por el símbolo  $r_{kk}$  y es dado por la fórmula

$$r_{kk} = \begin{cases} r_1 & \text{si } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k,j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{si } k = 2, 3, \dots \end{cases}$$

donde

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad \text{para } j = 1, 2, \dots, k-1$$

**Nota:** en la fórmula  $r_k$  es la autocorrelación muestral en el intervalo  $k$

La función de autocorrelación parcial (PACF) teórica es definida por la serie de datos de  $\rho_{kk}$  para los intervalos  $k=1, 2, \dots$ . La PACF muestral es una serie de datos de  $r_{kk}$  para los intervalos  $k=1, 2, \dots$ . La PACF teórica de una serie de tiempo estacionaria tiende a disminuir al incrementarse el intervalo  $k$  o se trunca después de un intervalo particular  $k=q$ . Se puede determinar cuándo la PACF teórica se trunca usando la PACF muestral. Sin embargo, aunque la PACF teórica puede ser tal que

$$\rho_{kk} = 0 \quad \text{para } k > q$$

debido a la variación muestral la PACF muestral  $r_{kk}$  probablemente será pequeña, pero no igual a 0, para  $k > q$ . Pero ahora ¿Cómo saber que tan pequeña debe ser  $r_{kk}$  para concluir que  $\rho_{kk} = 0$ ? Una regla de decisión para concluir que  $\rho_{kk} = 0$  para  $k > q$  es

$$|r_{kk}| \leq 2 \frac{1}{(n-a+1)^{1/2}} \quad \text{para } k > q$$

Otro método utilizado para decidir si  $\rho_{kk} = 0$  es calcular el estadístico  $t$

$$t_{r_{kk}} = \frac{r_{kk}}{1 / (n-a+1)^{1/2}}$$

Con la regla de decisión se puede concluir que

$$\rho_{kk} = 0 \quad \text{si} \quad t_{kk} = \frac{\Gamma_{kk}}{1/(n-a+1)^{1/2}} \leq 2$$

Contrario a la regla de decisión, si  $|r_{kk}| > 2$  es razonable concluir que  $\rho_{kk} \neq 0$ . Así pues, si  $|t_{kk}| > 2$  para los intervalos 1, 2, ..., q y si  $|t_{kk}| \leq 2$  para los intervalos q+1, q+2, ..., se concluye que la PACF se corta después del intervalo q.

Es importante saber el comportamiento de las ACF y PACF, ya que cada modelo de los autorregresivos, medias móviles o mezclados, es decir autorregresivos y medias móviles, están caracterizados por el comportamiento de su función de autocorrelación teórica y la de autocorrelación parcial.

Se requiere que el modelo que se utilizara para describir y pronosticar una serie de tiempo estacionaria debe cumplir con las condiciones de estacionaridad e invertibilidad. Estas condiciones pueden ser expresadas en términos de los parámetros del modelo.

Para ilustrar esto, considérese el modelo

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

que es el modelo de medias móviles de primer orden. Para algún valor de  $\theta_1$  el modelo describe una serie estacionaria. Por lo tanto, no hay una condición que se le ponga al parámetro para que sea estacionario. Sin embargo si es necesario poner una condición a  $\theta_1$  para que el modelo sea invertible. Esta condición es:

$$|\theta_1| < 1$$

Estas condiciones son de gran importancia en la estimación de parámetros del modelo.

Puede mostrarse que la media del modelo

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

es  $\mu$ , y que la ACF teórica decrece y la PACF se trunca en el intervalo 1.

### **Series Estacionales**

Por una serie estacional se entenderá una serie de tiempo que, aparte de contener una tendencia (y/o ciclos) de larga duración, muestre fluctuaciones que se repiten anualmente, quizás con cambios graduales a través de los años.

Una característica fundamental de las series estacionales con observaciones mensuales, es la similitud de comportamiento entre observaciones para el mismo mes en años consecutivos. Es importante notar que aún cuando la variación estacional se considera en general como un fenómeno repetitivo anual, esto no implica que no pueda existir un cierto

patrón del comportamiento periódico con duración menor al año; por ejemplo, podría existir un periodo estacional semestral, de tal forma que los meses de diciembre fuesen similares a los de junio, además de que los diciembres muestren un comportamiento similar entre sí y de que los junios se comporten también de manera parecida. Por lo anterior es conveniente hablar genéricamente de un periodo estacional que comprenda  $D$  observaciones contiguas, en donde  $D$  no es necesariamente igual a 12 para series mensuales, ni igual a 4 para series trimestrales.

#### *Operadores Puramente Estacionales*

Así como hay operadores para series no estacionales, resulta conveniente utilizar la siguiente notación para series estacionales. El operador de diferencia estacional  $\Delta_D^K$  se define como

$$\Delta_D^K z_t = (1 - B^D)^K z_t$$

### **2.2.1 Clasificación de Modelos para Series de Tiempo**

#### *Modelos Autorregresivos*

Se dice que una serie de tiempo (estacionaria) esta gobernada por un proceso autorregresivo de primer orden si los valores actuales de la serie  $z_t$ , pueden expresarse como una función lineal del primer valor anterior y un choque aleatorio  $e_t$ .

$$z_t = \phi_1 z_{t-1} + e_t$$

donde  $\phi_1$  es el parámetro autorregresivo que describe el efecto de un cambio de  $z_{t-1}$  en  $z_t$ .  
Modelo general AR(p)

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + e_t$$

#### *Modelos de Medias Móviles*

Una forma alternativa de representar el proceso es en base a los choques aleatorios ponderados en el tiempo y acumulados. Se ha llamado medias móviles por la similitud con este tipo de suavizamiento. El valor actual de  $z_t$  puede representarse, por ejemplo, como el valor del choque aleatorio actual y el valor anterior:

$$z_t = e_t - \theta_1 e_{t-1}$$

donde  $\theta_1$  describe el efecto del choque aleatorio anterior a  $z_t$ .

Una razón importante para usar las medias móviles es que el número de parámetros puede reducirse drásticamente y, como los parámetros deben estimarse con un número finito de puntos es importante, representar el proceso con el menor número posible de parámetros.

Modelo general del Ma(q)

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Modelos Mezclados ARMA(p,q)

El modelo mezclado general será de la forma:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Tendrá por supuesto condiciones tanto de estacionaridad como de invertibilidad (bastante complejas), y su función de autocorrelación será decreciente infinita (ya que incluye una parte de tipo autorregresivo), en el cual el efecto de la parte de medias móviles desaparece después de q periodos de acuerdo a la experiencia los valores de p y q deben ser menores a 2.

### Modelos Estacionales

#### Modelos Estacionales Autorregresivos

Este modelo se aplica si el valor actual de la serie  $z_t$  puede expresarse como una función lineal del valor de la serie  $s$  periodos atrás,  $z_{t-s}$ , y el choque aleatorio  $e_t$ , esto es:

$$z_t - \Phi_1 z_{t-s} = e_t ;$$

$$(1 - \Phi_1 B^s) z_t = e_t$$

donde  $\Phi_1$  es el parámetro autorregresivo estacional. Este modelo podría denotarse como AR(1)<sub>s</sub>, o puede expresarse como SAR(1) y el valor de "s" se inferirá del contexto. De la misma forma puede extenderse a un SAR(p):

$$\Phi(B^s) z_t = e_t$$

donde:

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$$

$$z_t = \Delta^d \Delta^s w_t$$

#### Modelos Estacionales de Medias Móviles

Se dice que una serie de tiempo está gobernada por un proceso de medias móviles estacionales de primer orden si el valor actual de la serie  $z_t$  puede representarse como el

choque aleatorio  $e_t$ , y el choque ocurrido exactamente "s" observaciones antes,  $e_{t-s}$ , donde "s" es igual a la longitud del período.

$$z_t = e_t - \Theta_1 e_{t-s}$$

ó

$$z_t = (1 - \Theta_1 B^s) e_t$$

Puede denotarse como MA(1)s o SMA(1).

En general un modelo SMA de orden Q:

$$z_t = \Theta (B^s) e_t$$

donde :

$$\Theta (B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}$$

$$z_t = \Delta^p \Delta^d w_t$$

#### *Modelos Estacionales Mezclados*

Puede construirse un modelo estacional mezclado autorregresivo y de medias móviles, de orden (P,Q) :

$$\Phi (B^s) z_t = \Theta (B^s) e_t$$

La única diferencia entre las autocorrelaciones de un proceso estacional mixto y uno ordinario es que las autocorrelaciones distintas de cero aparecen en los intervalos múltiplos de s.

En general, puede denotarse ARIMA(P,D,Q)s donde :

P = orden del proceso AR estacional  
 Q = orden del proceso MA estacional  
 D = orden de diferencias estacionales  
 s = longitud del período.

#### *Modelos Generales Multiplicativos Estacionales*

Finalmente pueden combinarse todos los modelos en una clase general que, para muchas series de tiempo, proporciona muy buen ajuste y genera pronósticos adecuados. Esta clase general se llama modelos ARIMA multiplicativos y se expresa como:

$$\Phi (B) \Phi (B^s) z_t = \theta (B) \Theta (B^s) e_t$$

ó

$$\text{ARIMA}(p,d,q) \times (P,D,Q)_s$$

Este es el modelo más general, que da como resultado una gran cantidad de modelos con pocos parámetros, de donde es casi seguro que se encontrará alguno adecuado al fenómeno que se este manejando.

Apartir de los resultados anteriores, puede concluirse que existe una dualidad entre los procesos AR y MA:

- la ACF de un modelo MA se comporta como la PACF de un modelo AR; y
- la PACF de un modelo MA se comporta como la ACF de un modelo AR.
- los procesos AR deben cumplir condiciones de estacionaridad, pero no de invertibilidad.
- los procesos MA tienen condiciones de invertibilidad pero no de estacionaridad.

En el cuadro siguiente se presenta las características de ACF y PACF de modelos estacionales y no estacionales:

Modelo No Estacional	ACF	PACF
AR(p)	Decae	Se trunca después de p
MA(q)	Se trunca después de q	Decae
ARMA(p,q)	Decae	Decae

Modelo Estacional	ACF	PACF
AR(p) x SAR(P)	Decae	Se trunca después de p + sP
MA(q) x SMA(Q)	Se trunca después de q	Decae
ARMA(p,q)	Decae	Decae

### 2.3 Estimación de Parámetros de modelos

Esta etapa presupone que se ha identificado ya a un modelo y que, de ser este adecuado, lo único que resta es encontrar los mejores valores de los parámetros para que dicho modelo represente a la serie en consideración.

La estimación de los parámetros será obtenida por el criterio de mínimos cuadrados. La mayoría de los programas de computadoras aplica un procedimiento de búsqueda iterativo para la función de suma de cuadrados y requiere que el usuario especifique los estimadores preliminares de los parámetros desconocidos como entrada de los programas. Esas estimaciones preliminares pueden ser obtenidas por medio de la relación entre la función de autocorrelación teórica y los parámetros del modelo. Estas relaciones son usadas para resolver los parámetros del modelo en términos de las autocorrelaciones teóricas, las cuales son estimadas por las correspondientes autocorrelaciones muestrales, así se producen un conjunto de estimadores preliminares de los parámetros del modelo.

Solo un conjunto de estimadores preliminares serán seleccionados y son aquellos que satisfacen las condiciones de estacionaridad e invertibilidad.

Sin embargo, a veces es muy complicado o consume mucho tiempo resolver la relación entre los parámetros del modelo en términos de las autocorrelaciones teóricas. Por esta razón Box y Jenkins dieron unas tablas donde ya se tienen valores de varios modelos que utilizan ellos.

Cuando se emplean modelos estacionales, también es difícil encontrar los parámetros del modelo en términos de las autocorrelaciones y en las tablas no se encuentran estos valores, entonces es un procedimiento fácil usar el número 0.1 como estimador preliminar de los parámetros excepto de  $\mu$ . Se excluye  $\mu$  pues ya se explicó anteriormente que su estimador es

$$\bar{z} = \frac{\sum_{i=1}^n z_i}{n - u + 1}$$

Usar 0.1 como estimador preliminar de los parámetros es importante por dos razones. Primero, usando 0.1 generalmente significa que los estimadores preliminares satisfacen las condiciones de estacionaridad e invertibilidad para el modelo bajo consideración. La segunda es que el estimador final obtenidos por programas de computadoras no son muy sensibles para estimadores preliminares usados como entradas. Generalmente el uso de 0.1 como estimador preliminar es satisfactorio.

Para estimarse los parámetros del modelo ARIMA, se pueden usar los métodos de los momentos, el método de máxima verosimilitud o el método de mínimos cuadrados. De acuerdo a William W. S. Wei<sup>1</sup>, lo más recomendable es utilizar máxima verosimilitud, que resultará equivalente a mínimos cuadrados. Sin embargo, como se obtiene un sistema de ecuaciones no lineales, será necesario hacer uso de otras herramientas.

#### **Métodos de Mínimos cuadrados y Máxima Verosimilitud**

Una vez que se ha hecho una especificación tentativa del modelo, i.e., elegidos los valores  $p$ ,  $d$  y  $q$  para el modelo :

$$\Phi(B)z_t = \Theta(B)e_t \quad (2.1)$$

deben obtenerse estimadores para los  $p$  parámetros autorregresivos,  $\phi_1, \dots, \phi_p$  y los  $q$  parámetros de medias móviles  $\theta_1, \dots, \theta_q$ . Como en el caso del modelo de regresión lineal, se eligen los valores de los parámetros que minimicen la suma de los cuadrados de las diferencias entre la serie real  $z_t = \Delta^d y_t$ , y la serie ajustada  $z_t$ , es decir,

$$e_t = \Theta^{-1}(B) \Phi(B) z_t \quad (2.2)$$

donde este valor representará el valor de los residuales, al momento de sustituir los valores estimados de los parámetros.

<sup>1</sup>Wei, William. Time Series Analysis. Univariate and Multivariate Methods, pp. 135-156

El objetivo de la estimación es encontrar un vector de parámetros autorregresivos  $\phi = (\phi_1, \phi_2, \dots, \phi_p)$  y un vector de parámetros de medias móviles  $\theta = (\theta_1, \theta_2, \dots, \theta_q)$  que minimicen los errores al cuadrado:

$$S(\phi, \theta) = \sum_t e_t^2 \quad (2.3)$$

Se denotará los vectores que minimizan (2.3) con

$$\phi = (\phi_1, \phi_2, \dots, \phi_p) \quad y$$

$$\theta = (\theta_1, \theta_2, \dots, \theta_q) \quad ;$$

y los residuales asociados con estos valores de los parámetros por  $e_t$ , de modo que:

$$e_t = \theta^{-1}(B) \phi(B) z_t$$

entonces

$$S(\phi, \theta) = \sum_t e_t^2$$

Esta estimación puede ser difícil si están presentes términos de medias móviles, ya que la ecuación (2.2) es no lineal. Por esta razón debe usarse un método iterativo de estimación no lineal para minimizar (2.3). Además, el primer error en la serie de tiempo,  $e_1$ , depende de los valores pasados no observados  $z_0, z_{-1}, \dots, z_{-p+1}$  y  $e_0, e_{-1}, \dots, e_{-q+1}$ . De modo que debe usarse algún método para inicializar la serie (i.e. elegir números para estos valores no observados) antes de aplicar el proceso de estimación no lineal.

Si se asume que un total de  $T+d$  observaciones están disponibles para una serie de tiempo homogénea no estacionaria de orden  $d$  (es decir, que requieren de  $d$  diferencias ordinarias para ser estacionaria),  $y_t$ , y se denotan estas observaciones como  $y_{d+1}, \dots, y_m, y_1, \dots, y_T$ ; después de diferenciar esta serie " $d$ " veces se obtiene una serie estacionaria  $z_t$  con  $T$  observaciones  $z_1, \dots, z_T$ . El problema es estimar los vectores de parámetros  $\phi$  y  $\theta$  para el modelo ARMA(p,q) especificado para la serie  $z_t$ . Para hacer esto se utiliza el hecho de que (por suposición) los términos de los errores  $e_t$  son distribuidos normalmente e independientes, con media cero y varianza  $\sigma^2$ . Entonces la función de verosimilitud condicional asociada con los parámetros  $(\phi, \theta, \sigma)$  esta dada por:

$$F(\phi, \theta, \sigma) = \frac{[(2\pi\sigma^2)^{-1/2}]^T \exp(-S(\phi, \theta)/2\sigma^2)}{(2\pi\sigma^2)^{-T/2} \exp(-S(\phi, \theta)/2\sigma^2)}$$

Obteniendo el logaritmo

$$L(\phi, \theta, \sigma) = -\frac{T(\log 2\pi\sigma^2)}{2} - \frac{S(\phi, \theta)}{2\sigma^2} \quad (2.4)$$

Se dice que  $L(\phi, \theta, \sigma)$  es la función de verosimilitud logarítmica "condicional" porque la suma de errores cuadrados  $S(\phi, \theta)$  depende de los valores pasados no observables  $z_0, z_1, \dots, z_{p+1}, e_0, e_1, \dots, e_{q+1}$ .

Esto puede verse escribiendo la ecuación para el primer error observable  $e_1$  en la forma expandida del modelo ARMA.

$$e_1 = z_1 - \phi_1 z_0 - \phi_2 z_{-1} - \dots - \phi_p z_{-p+1} + \theta_1 e_0 + \dots + \theta_q e_{-q+1} \quad (2.5)$$

Los estimadores de máxima verosimilitud de  $\phi$  y  $\theta$  estarán dados por la minimización de la suma de residuales al cuadrado  $S(\phi, \theta)$ . Entonces, bajo la suposición de normalidad, el estimador de máxima verosimilitud es igual al de mínimos cuadrados.

#### *Inicialización de la serie*

Ya que la función suma de cuadrados  $S(\phi, \theta)$  y por lo tanto la función de verosimilitud  $L(\phi, \theta, \sigma)$  son ambas condicionantes con respecto a valores pasados no observables de  $z_t$  y  $e_t$ , los estimadores de mínimos cuadrados que se obtengan dependen de la elección de los valores  $z_0, z_1, \dots$  etc.

Una posible solución es tomar  $z_0, z_1, \dots, z_{p+1}$  y  $e_0, e_1, \dots, e_{q+1}$  iguales a sus valores esperados incondicionales. Los valores esperados incondicionales de  $e_0, e_1, \dots, e_{-q+1}$  son todos cero, y si  $\delta = 0$  (suponiendo que se manejan las desviaciones con respecto a la media), los valores incondicionales de  $z_0, \dots, z_{p+1}$  también serán ceros.

Esta solución da una aproximación razonablemente buena al procedimiento correcto, si los valores reales de  $\phi_1, \phi_2, \dots, \phi_p$  no son muy cercanos a uno (a la falta de estacionaridad) y si el número de observaciones  $T$  es grande en relación con  $p$  y  $q$ . Esto último generalmente ocurre, ya que se tendrán por lo menos 50 observaciones originales, y los valores de  $p$  y  $q$  rara vez son mayores a 2.

También existe un procedimiento para encontrar los valores esperados condicionales para  $z_0, \dots, z_{p+1}$ ; esto es, valores que son condicionales con respecto a los valores observados de  $z_1, z_2, \dots, z_T$ , y los valores estimados de  $e_1, e_2, \dots, e_T$ . Esencialmente pueden inicializarse  $z_0, z_1, \dots, z_{p+1}$  y  $e_0, e_1, \dots, e_{q+1}$  con cero; estimando el modelo ARMA minimizando  $S(\phi, \theta)$  condicionados a estos valores cero, y entonces pronosticar hacia atrás el modelo para generar nuevos valores para  $z_0, z_1, \dots, z_{p+1}$ . Como la serie diferenciada  $z_t$  es estacionaria, el proceso ARMA puede regresarse en el tiempo; i.e. es equivalente (estadísticamente en una muestra grande) al proceso:

$$\phi(F) z_t = \theta(F) e_t \quad (2.6)$$

donde  $F$  es el "operador de salto hacia adelante" (forward shift operator):

$$\begin{aligned} F z_t &= z_{t+1} \\ F^2 z_t &= z_{t+2} \end{aligned}$$

$$\Rightarrow z_t = \phi^{-1}(F) \theta(F) e_t \quad (2.7)$$

y puede usarse para encontrar los valores de  $z_0, z_1, \dots, z_{p-1}$  en términos de los valores estimados de  $e_1, \dots, e_T$ . Estos valores estimados son los residuales formados al restar la serie real  $z_t$  de los valores generados por el modelo ARIMA inicialmente estimado. Los valores de  $e_t$  no observados, no podrán ser estimados, ya que se carece de las observaciones de la serie en esos períodos.

Pueden obtenerse "nuevos" estimadores de mínimos cuadrados para  $\phi$  y  $\theta$  minimizando  $S(\phi, \theta, \sigma)$  condicionada a  $z_0 = z_0, z_1 = z_1$ , etc.

Un nuevo conjunto de estimadores  $z_0, z_1, \dots, z_{p-1}$  puede calcularse con (2.7), y se repite el proceso hasta que los estimadores sean convergentes, i.e., hasta que  $\phi$  y  $\theta$  dejen de cambiar en forma significativa. Desafortunadamente no hay garantía de que el proceso sea convergente. Si no lo es, nos veremos forzados a fijar  $z_0, \dots, z_{p-1}$  con sus estimadores iniciales.

La determinación de los valores condicionales esperados de la serie  $z_0, \dots, z_{p-1}$  es difícil, y usualmente involucran demasiados cálculos, de modo que es razonable preguntarse si los estimadores de  $\phi$  y  $\theta$  serán o no muy sensibles a la inicialización de la serie. La respuesta depende de la longitud de la serie de tiempo, y en particular, de la longitud relativa a los valores de  $p$  y  $q$ . Si la serie es grande a comparación de  $p$  y  $q$ , la función de suma de cuadrados condicional será prácticamente igual a la suma incondicional, y los estimadores de  $\phi$  y  $\theta$  serán insensibles al proceso de inicialización. En este caso, simplemente se inicializan todos los valores como cero. Si la serie es pequeña (a comparación de  $p$  y  $q$ ), se ganará alguna eficiencia si se usan los valores esperados condicionales, pero no es posible determinar esta ganancia. Nuestro conocimiento de las propiedades estadísticas de los estimadores se basa en resultados asintóticos, i.e., en muestras grandes. La mayoría de los paquetes permite esta opción en el proceso, y puede compararse si existe diferencia significativa al hacer pronóstico hacia atrás. Desafortunadamente se desconocen las propiedades de los estimadores de muestras pequeñas.

Por lo general no se desea modelar una serie de tiempo como ARMA a menos que se tenga un número suficientemente grande de observaciones (más de 50). De otro modo es poco probable que pueda obtenerse un ajuste adecuado. Además la mayor parte de los modelos ARMA en aplicaciones económicas y financieras son de orden bajo ( $p$  y  $q \leq 3$ ). En estos casos la inicialización no será problema.

#### **Estimación no lineal de los parámetros del modelo**

El problema es encontrar valores de  $\phi$  y  $\theta$  que minimicen la suma de errores cuadrados:

$$S(\phi, \theta) = \sum_t e_t^2 = \sum_t [e_t / \phi, \theta, z_t]^2 \quad (2.8)$$

donde  $z$  es un vector de las observaciones  $z_t$  y los errores  $e_t$  están condicionados a  $\phi, \theta, z$ . Suponiendo que la inicialización de la serie se basa, como se sugirió, en los valores esperados incondicionales (que son cero) de  $z_0, \dots, z_{p-1}$  y  $e_0, \dots, e_{q+1}$ , los límites en la suma de (2.8) son  $t=1$  hasta  $T$ . El problema es

$$S(\phi, \theta) = \sum_{t=1}^T [e_t / \phi, \theta, z]^2 \quad (2.9)$$

Ahora, suponga que el modelo es puramente autorregresivo, i.e., de la forma:

$$\phi(B)z_t = e_t \quad (2.10)$$

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + e_t \quad (2.11)$$

observe que (2.11) es de la forma general:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + e_t \quad (2.12)$$

que pueden estimarse como regresión lineal múltiple.

$$\phi = (X^T X)^{-1} X^T Y$$

donde

$$X = \begin{bmatrix} z_p & z_{p-1} & \dots & z_1 \\ z_{p+1} & z_p & \dots & z_2 \\ \dots & \dots & \dots & \dots \\ z_{T-1} & z_{T-2} & \dots & z_{T-p} \end{bmatrix}$$

y

$$Y = \begin{bmatrix} z_{p+1} \\ z_{p+1} \\ \cdot \\ \cdot \\ z_T \end{bmatrix}$$

Aunque para un modelo puramente autorregresivo el proceso de estimación es esencialmente una regresión lineal, el problema es más difícil si el modelo contiene un componente de medias móviles. En este caso puede representarse como:

$$\theta^{-1}(B) \phi(B) z_t = e_t \quad (2.13)$$

Esta "ecuación de regresión" es no lineal en los parámetros y no pueden estimarse por una sencilla aplicación de mínimos cuadrados. Sin embargo, puede estimarse por una rutina general iterativa no lineal.

El proceso usa los dos primeros términos de la expansión en serie de Taylor<sup>4</sup> para linealizar (2.13) alrededor de un valor inicial de  $\phi$  y  $\theta$ . Se realiza una regresión lineal en esa nueva

<sup>4</sup>Burden, Faires y Reynolds. Análisis Numérico. (pp. 94-96)

ecuación linealizada, se obtienen estimadores de mínimos cuadrados para  $\phi$  y  $\theta$ , y se hace una nueva linealización alrededor de ellos. El proceso se repite en forma iterativa hasta que ocurre la convergencia.

Los errores estándar y los estadísticos  $t$  de nuestros estimadores se calculan a partir de la última linealización, de la misma forma que en cualquier procedimiento de estimación lineal. De modo que los estadísticos " $t$ " tienen poca relevancia, indicando sólo la significancia de los estimadores de mínimos cuadrados obtenidos para la linealización final de la ecuación no lineal. De la misma forma puede calcularse  $R^2$  (el coeficiente de determinación), que tendrá una importancia similar. Por esta razón no necesariamente deberá rechazarse una ecuación con una  $R^2$  pequeña

$$R^2 = 1 - \frac{\sum e_i^2}{\sum z_i^2} = \frac{\sum z_i^2 - \sum e_i^2}{\sum z_i^2}$$

Aunque se obtenga un valor pequeño no hay garantía de que el proceso de estimación descrito arriba converja siempre a un estimador final de los parámetros. Es muy posible que el proceso diverja, i.e., que los estimadores sucesivos  $\beta_1, \beta_2, \dots, \beta_k$  se alejen cada vez más. Aún más, también es posible que existan soluciones múltiples; en este caso, ocurre la convergencia pero diferentes estimadores finales de uno o más de los parámetros. El hecho de que exista divergencia o soluciones múltiples depende tanto de la especificación del modelo ARIMA como de los datos.

Si ocurre divergencia el modelo puede reestimarse una o más veces usando diferentes valores iniciales. Un nuevo valor inicial puede obtener la convergencia, aunque no siempre es éste el caso. Si resulta imposible encontrar una solución convergente, tendrá que buscarse una nueva especificación del modelo.

Aún si la convergencia ocurre el primer intento, es conveniente si hay múltiples soluciones reestimando el modelo con diferentes valores iniciales. Si ocurren soluciones múltiples el estimador final será aquel que minimice la suma de errores al cuadrado. (No todos los paquetes permiten al usuario elegir valores iniciales, de modo que esto no es siempre posible).



Supongamos que  $\rho_1 = 0.4$  en la función de autocorrelación muestral. Entonces :

$$\theta_1 = \frac{-1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1} = \frac{-1 \pm 0.6}{0.8} = \begin{cases} -2 \\ -0.5 \end{cases}$$

como la condición de invertibilidad es  $|\theta_1| < 1$ , se elegirá  $\theta_{1,0} = -0.5$ . Desafortunadamente la solución de  $\theta$  en términos de  $\rho$  se hace más complicada a medida que  $q$  se incrementa. De hecho para poder obtener aproximaciones iniciales para el modelo MA( $q$ ), es necesario resolver  $q$  ecuaciones simultáneas no lineales.

Podríamos preguntarnos por qué esta aproximación inicial no es suficiente para propósitos prácticos. Una razón es que la función de autocorrelación muestral es sólo un estimador de la función de autocorrelación real y está por lo tanto sujeta a error. De hecho para muestras pequeñas (menores de 50), la función de autocorrelación muestral será sesgada (hacia abajo) de la función real. Otra razón es que la función de autocorrelación muestral no contiene tanta información como la serie original.

Existen tablas de (Box y Jenkins, op.cit. pp. 517-520) para obtener aproximaciones iniciales de los parámetros, pero algunos autores (Bowerman y O'Connell<sup>5</sup>) consideran que pueden utilizarse el valor de 0.1 para cualquier parámetro desconocido (excepto  $\mu$ ), sin que esto afecte mayormente los resultados finales. Otros (Pindyck & Rubinfeld<sup>6</sup>) toman cero como valor inicial, con buenos resultados.

## 2.4 Verificación del modelo

La etapa de verificación de la Metodología de Box-Jenkins tiene su origen en la idea de que todo modelo es erróneo, puesto que los modelos son meras representaciones simplificadas de la realidad. Lógicamente si hay que elegir entre varios modelos, habrá que elegir aquel que presente menos fallas, o bien fallas menos importantes; por este motivo habrá que poner a todos los posibles modelos en tela de juicio para detectar sus fallas (que se miden como violaciones a los supuestos que fundamentan al modelo).

Una de las formas más claras y simples para detectar esas violaciones es a través del análisis de residuales en donde, como residual se considera aquella parte de las observaciones que no es explicada por el modelo.

Si el modelo se acerca bastante al proceso ARIMA que rige la serie, los errores se comportarán como ruido blanco: tendrán media cero, varianza constante y no estarán correlacionados.

De modo que un examen de las propiedades de los errores permitirá evaluar el ajuste del modelo.

<sup>5</sup>Bowerman & O'Connell, *Forecasting Time Series*, pp. 369,370.

<sup>6</sup>Pindyck & Rubinfeld, *Econometric Models and Economic Forecasts*, pp. 545,546.

Una vez que el modelo se ha identificado y se ha estimado sus parámetros, es necesario verificar si el modelo puede mejorarse. Por supuesto que no podrá hablarse de un modelo óptimo, sin embargo, se trata de mejorarlo hasta donde sea posible sin olvidar el principio de parsimonia.

Para el estimador de un coeficiente pueden estimarse las siguientes estadísticas de interés:

- Valor del estimador.
- Estimador del error estándar del estimador ( $1/\sqrt{n}$ ).
- El cociente del estimador entre su error estándar (estadístico t).
- Límite del 95% de confianza.

Además existen varias medidas de bondad de ajuste del modelo:

- Estimador de la desviación estándar de los residuales
- Suma de los errores al cuadrado ( $R^2 = 1 - (\sum e_i^2 / \sum z_i^2)$ )
- ACF y PACF de los residuales
- Periodograma Integral de los residuales
- Estadístico Box-Pierce de los residuales

Aunque  $R^2$  tiene la misma interpretación en el modelo ARIMA que en regresión lineal, las características especiales de los modelos de series de tiempo ofrecen la oportunidad de una vista al interior, que puede ser muy útil en el análisis. Puede mostrarse que el valor  $R^2$  poblacional está relacionado con los coeficientes  $\phi_1$  y  $\theta_1$  y no dependen de  $\sigma^2$ . De modo que no puede usarse tal como en regresión lineal. Nelson (1976) analizó algunos modelos especiales ARMA y encontró que para un modelo AR(1),  $R^2 = \phi_1^2$ , de modo que si  $\phi_1 = 0.5$ , podemos anticipar una  $R^2$  de a lo más 0.25. Para un modelo MA(1) puede demostrarse que  $R^2 = \theta_1^2 / (1 + \theta_1^2)$ , que tendrá un valor máximo de 0.50. Desafortunadamente, el caso de modelos mezclados ARMA no produce resultados tan claros.

También debe tenerse precaución con el uso de  $R^2$  para seleccionar entre varios modelos alternativos. Como en el caso de la regresión lineal, usar  $R^2$  para elegir un modelo es un error si la variable dependiente  $z_t$  no es la misma. Específicamente, no puede usarse  $R^2$  para seleccionar entre modelos con diferentes transformaciones estacionarias. Como se verá en las siguientes secciones, la varianza de los errores de pronóstico está determinada por la varianza de los residuales.

Entonces, para propósitos de pronósticos la varianza del error es el criterio más importante.

Ahora se describirán 4 grupos de pruebas de diagnóstico que pueden usarse para evaluar el ajuste del modelo. Si se detecta falta de ajuste deben buscarse diferentes especificaciones:

- Análisis de Estacionaridad
- Análisis de Residuales
- Modelo Sub-especificado (agregar parámetros)

- **Modelo Sobre-especificado (omitir parámetros)**

Estos criterios no tienen que aplicarse necesariamente en forma secuencial deben compararse en conjunto, no en forma aislada.

### **Análisis de Estacionariedad**

Debe tenerse cuidado de no sobrediferenciar el modelo. Esto genera un modelo más complicado, lo cual va en contra del principio de parsimonia. Esto ya se mencionó con anterioridad, así como la forma de detectarlo.

Si el proceso no es estacionario, la selección no funcionará, ya que ésta es una de las suposiciones originales. Sin embargo, existen distintas alternativas que pueden dar como resultado un proceso estacionario. En este caso es conveniente analizar todas y finalmente compararlas con objeto de elegir aquélla que de un mejor resultado en cuanto a ajuste y pronóstico.

### **Análisis de Residuales**

Si el modelo se acerca bastante al proceso ARIMA que rige la serie, los errores se comportarán como ruido blanco: tendrán media cero, varianza constante y estarán incorrelacionados. De modo que un examen de las propiedades de los errores permitirá evaluar el ajuste del modelo. Sin embargo, siempre debe tomarse en cuenta que el análisis de los errores se basa en estimadores.

Como siempre, el análisis se inicia observando la gráfica. Esto puede revelar que no hay estacionariedad, o valores discrepantes que puedan o no ser relevantes, así como fluctuaciones estacionales.

Es razonable suponer que las autocorrelaciones de los residuales proporcionen información, pero hay que recordar que los residuales están necesariamente correlacionados aunque los errores reales sean independientes.

Si los residuales son realmente ruido blanco, su ACF no tendrá picos y las autocorrelaciones muestrales serán todas pequeñas. El estimador del error estándar de las autocorrelaciones, para muestras grandes, es  $1/\sqrt{n}$ . Sin embargo, se ha demostrado (Box & Pierce) que este estimador sólo es bueno para autocorrelaciones con intervalos grandes. El error estándar de las primeras autocorrelaciones puede ser mucho menor que  $1/\sqrt{n}$ . Consecuentemente, los valores de  $r_k$  a intervalos pequeños que sean cercanos a los límites de confianza pueden ser significativamente distintos de cero y requerirán mayor análisis.

Ha sido encontrada una forma efectiva de medir la adecuación del modelo tentativo que es examinar una cantidad que determina si las primeras  $K$  autocorrelaciones de los residuales, considerándolos juntos, indican adecuación del modelo. Esta cantidad es llamado la estadística  $\chi^2$  de Box-Pierce, es denotado por el símbolo  $Q$ , y es calculado por la fórmula

$$Q = (n - d) \sum_{i=1}^k r_i^2(\epsilon)$$

donde

- $n$  es el número de observaciones de la serie original
- $d$  es el grado de diferenciación que se usó para transformar la serie original a una serie estacionaria
- $r_i^2(\epsilon)$  es el cuadrado de  $r_i(\epsilon)$ , la autocorrelación muestral de los residuales en el intervalo  $i$  — esto es, la autocorrelación muestral de residuales separada por un intervalo de  $i$  unidades de tiempo.

El proceso de modelado supuesto explica la relación entre las observaciones. Si esto hace contar para estas relaciones, los residuales no estarán relacionados, y así la autocorrelación de los residuales será pequeña. Por lo tanto  $Q$  será pequeña. Si el valor de  $Q$  es grande, más grandes serán las autocorrelaciones entre los residuales. Así un valor grande de  $Q$  indica que el modelo es inadecuado. En la práctica se acepta la adecuación del modelo si el valor de  $Q$  es menor que  $\chi^2(K - n_p)$ , el cual es definido para ser el punto sobre la escala de la distribución chi-cuadrada teniendo  $K - n_p$  grados de libertad tal que hay una área de .05 debajo de la curva de esta distribución anterior a este punto. Aquí,  $n_p$  es el número de parámetros que debe ser estimado en el modelo bajo consideración. La elección de  $K$ , el número de autocorrelaciones usadas en el cálculo de  $Q$ , es arbitrario, pero en la práctica se calcula para  $K=12$  (y posiblemente 24 y 36).

También se notará que la adecuación de un modelo de Box-Jenkins puede ser juzgado considerando la cantidad

$$s = \sqrt{\frac{SSE}{n - n_p}} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - n_p}} \quad ; \text{ SSE} = \text{Suma de Errores al cuadrado.}$$

donde  $n$  es el número de observaciones en la serie original, y  $n_p$  es el número de parámetros que deben ser estimados en el modelo. El valor pequeño de  $s$  indica un buen ajuste. En particular, será notorio que un valor pequeño de  $s$  producirá intervalos de confianza pequeños para los pronósticos cuando el modelo de Box-Jenkins es usado.

Una tercera herramienta para determinar si los errores son ruido blanco es evaluar las autocorrelaciones de las primeras diferencias de los residuales. Si son ruido blanco, entonces:

$$e_t - e_{t-1} = a_t - \theta_1 e_{t-1} ; \theta_1 = 1$$

La primera autocorrelación de un proceso MA(1) con  $\theta_1 = 1$  es  $-0.5$ ,  $\rho_1 = \theta_1 / (1 + \theta_1^2)$ . Entonces, si los residuales son ruido blanco, sus primeras diferencias deberán mostrar una primera autocorrelación cercana a  $-0.5$ .

Es importante reconocer que los residuales de modelos mal especificados pueden usarse para hacer alteraciones lógicas con el objeto de mejorar el modelo.

Finalmente, en algunas situaciones, particularmente aquellas que involucran modelos estacionales, es posible que no todas las características de la serie se hayan filtrado. Una forma de detectar si hay periodicidad en los residuales es evaluar el *Periodograma acumulado normalizado o Periodograma integral*.

Este periodograma integral se define como:

$$C(f_i) = \frac{\sum_{t=1}^n R^2(f_i)}{ns^2} \quad (2.15)$$

donde

$$R^2(f_i) = \frac{2}{n} \left[ (\sum_{t=1}^n e_t \cos 2\pi f_i t)^2 + (\sum_{t=1}^n e_t \sin 2\pi f_i t)^2 \right] \quad (2.16)$$

$$f_i = \frac{i}{n} = \text{frecuencia} \quad \frac{1}{f_i} = \text{período} \quad e_t = \text{errores} \quad s^2 = \sigma^2$$

La cantidad  $R^2(f_i)$  es una medida de que tanto una función trigonométrica (seno, coseno) con frecuencia  $f_i$  se ajusta a los residuales. Es un instrumento para correlacionar las  $e_t$ 's con las funciones seno y coseno de diferentes frecuencias. Por ejemplo, un residual con patrón estacional de período 12 ( $s=12$ ) o frecuencia  $1/12 = 0.08$  registrará valores grandes para  $R^2(1/12)$ ,  $R^2(1/24)$ ,  $R^2(1/36)$ , etc. Como resultado, el periodograma acumulativo  $C(f_i)$  mostrará saltos en estas frecuencias.

Un instrumento visual, manejado por la mayoría de los paquetes, es la gráfica del periodograma integral contra las frecuencias o el período. Para una serie de ruido blanco la gráfica de  $C(f_i)$  contra  $f_i$  o  $1/f_i$  quedaría como la línea recta uniendo los puntos (0,0) y (0.5,1), ya que todas las frecuencias tendrían la misma importancia dentro del proceso.

Los desajustes del modelo producirán errores no aleatorios cuyos periodogramas acumulativos mostrarán desviaciones sistemáticas de esta línea. Las diferencias pronunciadas a bajas frecuencias o períodos grandes pueden ser evidencia de un ciclo con un período largo que a su vez puede interpretarse como tendencia, lo cual indica que la diferenciación no ha sido suficiente.

Si los saltos ocurren cerca de los períodos  $s$ ,  $2s$ ,  $3s$ ,..., entonces es muy probable que no se haya modelado los efectos estacionales.

Las desviaciones de la línea recta pueden evaluarse usando los límites de confianza de Kolmogorov-Smirnov. Para el intervalo del 95% los límites están a  $\pm 1.36/\sqrt{m}$  arriba y abajo de la línea teórica a  $45^\circ$ , donde  $m = (n-2)/2$  para  $n$  par y  $m = (n-1)/2$  para  $n$  impar, y  $n$  es el número de observaciones disponibles.

Al calcular el periodograma se usan los residuales en lugar de los errores; aún si el modelo es adecuado los residuales pueden mostrar desviación de la línea, por lo tanto sólo deben tomarse en cuenta las desviaciones muy grandes.

Es importante observar que cuando se analizan las autocorrelaciones aún si alguna aparece grande en un intervalo sin significado, no deberá tomarse en cuenta.

#### **Modelos Sobrespecificado : Omitir Parámetros**

Los parámetros redundantes o excesivos pueden localizarse usando cuidadosamente el estadístico t de los coeficientes y el estimador de las correlaciones entre los estimadores.

Puede usarse la desviación estándar para evaluar la significancia estadística de un coeficiente. Como regla de dedo puede decirse que un coeficiente es significativamente distinto de cero si el valor absoluto del estimador puntual es al menos dos veces el error estándar. También puede usarse el estadístico T del estimador:

$$T = \frac{\beta_j - \beta_j}{\sigma_{\beta_j}}$$

para probar la hipótesis :

$$H_0 : \beta_j = 0$$

$$H_a : \beta_j \neq 0$$

Si el valor en tablas de T con un nivel de significancia  $\alpha$  y  $(n-p-q)$  g.l. es mayor que T, hay evidencia de que el parámetro puede ser cero.

Un parámetro insignificante es una indicación de que el modelo puede estar sobrespecificado y deberá simplificarse. Si este parámetro insignificante es el de mayor orden, entonces, en general, el modelo ARIMA puede simplificarse eliminando este parámetro del modelo.

Si el parámetro insignificante no es el de mayor orden, entonces examinar las correlaciones entre los estimadores ayudará a saber cuál parámetro eliminar. Una correlación alta entre dos estimadores indica que hay la posibilidad de que uno de los dos parámetros se omita sin cambiar el ajuste del modelo. El parámetro que sí se incluya será capaz de reflejar la contribución del otro.

Si el parámetro insignificante no es el de mayor orden pero está fuertemente correlacionado con el de mayor orden, se deberá evaluar el modelo sin el parámetro de orden mayor. Si esta correlación no existe, se reestimaré el modelo ARIMA suprimiendo el parámetro insignificante.

Si las correlaciones entre estimadores son altas para coeficientes de distintos polinomios, por ejemplo  $\phi_2$  y  $\theta_3$ , siendo o no alguno de ellos insignificante, la situación requiere un

análisis más cuidadoso. Únicamente podrán evaluarse modelos suprimiendo uno de los parámetros.

En el caso de que el parámetro insignificante no sea el de mayor orden y no exista correlación, seguramente estará reflejando un efecto de variación estacional (modelos multiplicativos para resolver el problema).

**Modelos Subespecificados: Ajustar parámetros extras**

Para verificar que el modelo tentativo contiene el número apropiado de parámetros, puede incluirse un parámetro adicional para ver si de esta adición se obtiene un mejor modelo.

Al ajustar parámetros extra el modelo de series de tiempo, existe un grave riesgo, llamado la *redundancia de parámetros*. Esta ocurre cuando se agrega al mismo tiempo un parámetro autorregresivo y uno de medias móviles. Supóngase que el modelo correcto es un AR(1):

$$\text{entonces:} \quad z_t = \phi_1 z_{t-1} + e_t \quad (2.17)$$

$$z_{t-1} = \phi_1 z_{t-2} + e_{t-1} \quad (2.18)$$

Restando (2.17) de (2.18):

$$\delta \quad z_t = (1+\phi) z_{t-1} - \phi z_{t-2} + e_t - e_{t-1} \quad (2.19)$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t - \theta_1 e_{t-1} \quad (2.20)$$

con  $\phi_1 = (1+\phi)$ ,  $\phi_2 = -\phi$ ,  $\theta_1 = 1$

El modelo ARMA(2,1) puede también derivarse de (2.17) simplemente agregando un término autorregresivo y uno de medias móviles. Como (2.20) se deriva directamente de (2.17), la diferencia entre estos dos modelos es ilusoria. De hecho, (2.17) se preferirá a (2.20) ya que es más eficiente. Primero, puede usarse las observaciones de  $z_t$  para  $t=2, \dots, n$ ; mientras que (2.20) sólo puede usar  $z_t$  para  $t=3, \dots, n$ ; además, (2.17) tiene menos parámetros y  $\phi_1$  y  $\phi_2$  están correlacionados en (2.20).

En términos más generales, el problema de la redundancia de parámetros puede explicarse como sigue reescribiendo (2.17) como:

$$(1-\phi B)z_t = e_t \quad (2.21)$$

y multiplicando por (1-B) se obtiene:

$$(1-B)(1-\phi B)z_t = (1-B)e_t \quad (2.22)$$

Esta ecuación muestra claramente que no es difícil generar modelos complicados o elaborados ya que siempre pueden multiplicarse a la izquierda y derecha de una ecuación por el mismo factor. Moraleja: jamás agregar un parámetro AR y uno MA en forma simultánea.

## 2.5 Pronósticos

Una vez que se ha encontrado un modelo adecuado para representar los datos, puede utilizarse para generar pronósticos.

Se llama fecha de *origen* al periodo actual, esto es, al período "n" y supóngase que se desea pronosticar "h" períodos de tiempo hacia adelante, hasta el período  $n+h$ , esto es, se desea conocer el valor de la observación no realizada  $z_{n+h}$ . El intervalo de tiempo "h" se llama *horizonte*, también se denota como  $z_n(h)$  al valor futuro estimado de la variable aleatoria.

Entonces  $z_{n+h}$  es una variable aleatoria cuya distribución es condicional a los datos pasados y presentes, así como a la especificación del modelo ARIMA.

Puede definirse el error al hacer un pronóstico en el momento  $n$  para un futuro  $n+h$ ,  $z_n(h)$ , como:

$$e_n(h) = z_{n+h} - z_n(h) \quad (2.23)$$

entonces el pronóstico óptimo será el valor de  $z_n(h)$  que minimice el costo esperado  $C(e_n(h))$ . En la práctica se desconoce, la mayoría de las veces, en forma precisa esta función. En general, puede usarse una función de costo cuadrática:

$$C(e) = \alpha e^2 \quad (2.24)$$

donde  $\alpha$  es una constante positiva.

Esta función define el costo en forma proporcional al cuadrado del error. También tiene la propiedad de que la solución para el pronóstico óptimo es la media de la distribución del pronóstico.

Esta función de costo implica que los costos son simétricos, si esto no se apega a la realidad puede tomarse, por ejemplo,

$$C(e) = \begin{cases} \alpha a & e > 0 \\ 0 & e = 0 \\ \beta e & e < 0 \end{cases} \quad (2.25)$$

Con  $\alpha$  y  $\beta > 0$ , siendo los costos unitarios de subestimar y sobreestimar  $z_{n+h}$  respectivamente.

Ahora se mostrará que la media de la distribución del pronóstico minimiza el valor esperado de la función de costo cuadrática.

Sea  $m_h$  el valor esperado de  $z_{n+h}$ ;  $m_h = E(z_{n+h})$ . Sea  $m$  cualquier otro pronóstico de  $z_{n+h}$ , definido como

$$m = m_h + d \quad (2.26)$$

donde "d" es la diferencia entre  $m$  y  $m_h$ . Usando el pronóstico puntual  $m$  el valor esperado del error es:

$$\begin{aligned} E[(z_{n+h} - m)^2] &= E[z_{n+h} - (m_h + d)]^2 \\ &= E[(z_{n+h} - m_h)^2] - 2d E(z_{n+h} - m_h) + d^2 \\ &= E(z_{n+h} - m_h)^2 + d^2 \end{aligned}$$

Para minimizar, "d" deberá ser igual a cero. De modo que el pronóstico óptimo de  $z_{n+h}$  se obtiene con  $m = m_h = E z_{n+h}$ .

La media de la distribución del pronóstico,  $E(z_{n+h})$ , puede calcularse como sigue: Sea  $z_t$  un proceso ARMA(p,q) estacionario e invertible. Para  $t = n+h$ :

$$z_{n+h} = \phi_1 z_{n+h-1} + \phi_2 z_{n+h-2} + \dots + \phi_p z_{n+h-p} + e_t - \theta_1 e_{n+h-1} - \dots - \theta_q e_{n+h-q}$$

- 1.- Reemplace el error actual y los pasados,  $e_{n+j}$ ,  $j \leq 0$ , con los residuales.
- 2.- Reemplace cada error futuro  $e_{n+j}$ ,  $0 < j \leq 0$  con su valor esperado, es decir cero.
- 3.- Reemplace las observaciones  $z_{n+j}$ ,  $j \leq 0$ , con los valores observados.
- 4.- Reemplace cada valor futuro de  $z_{n+j}$ ,  $0 < j < h$ , con el pronóstico adecuado  $z_n(j)$ ; de modo que hay que pronosticar  $z_{n+1}, \dots, z_{n+h-1}$  para pronosticar  $z_{n+h}$ .

Además, los parámetros del modelo se reemplazan por sus estimadores.

---

## **CAPITULO 3**

### **ANÁLISIS DE LAS SERIES DE TIEMPO PARA LA PRODUCCIÓN Y LAS VENTAS DE PRODUCTOS DERIVADOS**

---

### 3.1 Series de Tiempo de Producción de Petróleo Crudo y Gas Natural y las Ventas de Productos Derivados

El análisis que se llevará a cabo en este capítulo es relevante debido al papel tan importante que representan los hidrocarburos en la economía de México, ya que la industria petrolera contribuye en gran medida con los ingresos que genera y que se obtienen por un lado, con la producción de hidrocarburos que es la base principal de la industria y está orientada hacia la satisfacción de la demanda interna y a la generación de excedentes para la exportación, y por otro, con las ventas de los productos derivados destinados a los sectores industrial, del transporte, agropecuario, residencial, comercial, etc., que son base importante para el desarrollo del país.

Entre las series que se analizarán están las de Producción de Petróleo Crudo y de Gas Natural a nivel nacional, además del volumen de las Ventas Internas de Productos Petrolíferos.

En la tabla 3.1 se presentan las series a analizar:

PRODUCCIÓN	VENTAS
Gas Natural	Gas Natural
Petróleo Crudo	Petrolíferos
	• Gasolinas
	• Gas Licuado
	• Turbosinas

Tabla 3.1 Series de Tiempo a analizar<sup>1</sup>.

Para llevar a cabo este análisis se utilizó el paquete estadístico STATGRAPHICS, el cual maneja el análisis de series de tiempo con la metodología de Box-Jenkins.

### 3.2 Comportamiento de la Serie Original, Identificación y Estimación de los modelos.

Debido a que las series de tiempo consisten de datos numéricos, es natural hacer uso de herramientas estadísticas para describirlas y analizarlas, por lo que se construyen las gráficas antes de llevar a cabo cualquier tipo de cálculo, para verificar visualmente la congruencia de los datos.

Los datos históricos consisten de la cantidad de Petróleo Crudo y Gas Natural que se esta produciendo mensualmente a nivel nacional en miles de barriles (MBD) y millones de pies

<sup>1</sup> La datos de las Series de Tiempo se obtuvieron de la Base de Datos Institucional (BDI) de Petróleos Mexicanos.

cúbicos diarios (MMPCD) respectivamente durante el periodo de 1982 a 1995. También se tiene las ventas de algunos productos petrolíferos que se derivan de dichas producciones en el mismo periodo y también se encuentran en MBD y MMPCD.

Es importante mencionar que se probaron muchos modelos para cada serie de tiempo, pero como sólo algunos se ajustaron mejor a la serie son los que se presentan en cada tabla, y de los cuales se seleccionarán el o los que tengan un mejor ajuste a la serie.

### 3.2.1 Serie de la Producción de Petróleo Crudo

Las producciones mensuales para la serie de Producción de Crudo están graficadas en la figura 3.1

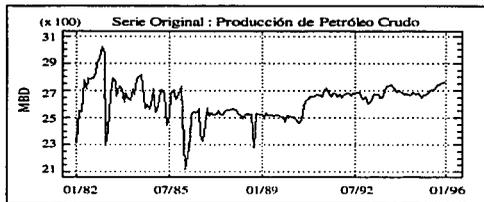


Figura 3.1 Serie de Tiempo de la Producción de Petróleo Crudo.

En la gráfica es posible apreciar las características relevantes de la Producción de Petróleo Crudo. Se advierte que hasta antes de 1989 la producción tuvo muchos altibajos muy significativos, llevando a una tendencia decreciente posiblemente ocasionados por las crisis económicas y otro tipo de acontecimientos a nivel internacional. A principios de 1982 se inicia el "Boom" Petrolero etapa en la que el crecimiento fue sin paralelo y el petróleo se convirtió en la principal fuente de divisas de México ocupando un 75% de su ingreso total anual, a su vez el valor porcentual del petróleo en el Producto Interno Bruto había escalado a principios de los 80 a 5.5. El aumento productivo de esta época de abundancia está ligado al descubrimiento de los campos de la Sonda de Campeche, descubiertos en 1976. Hasta la fecha es la provincia petrolera más importante del país y una de las más grandes del mundo. Después de 1991 se da un incremento notable para mantenerse en los siguientes años en un nivel aproximadamente constante.

En general los datos históricos de la serie despliegan poca tendencia, una varianza constante y no se ve que exista variación estacional a simple vista, en las ACF (función de Autocorrelación) y PACF (Función de Autocorrelación Parcial) podemos comprobar que la

serie no es estacionaria ya que la primera decrece lentamente mientras que la PACF se corta en el primer valor (ver figura 3.2).

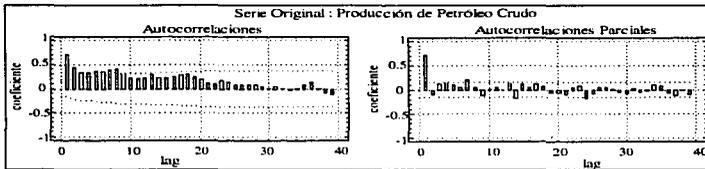


Figura 3.2 ACF y PACF de la serie "Producción de Petróleo Crudo".

Para hacer estacionaria la serie será necesario aplicarle una diferencia ordinaria y obtener nuevamente sus ACF y PACF (ver fig. 3.3) para verificar que ya lo es.

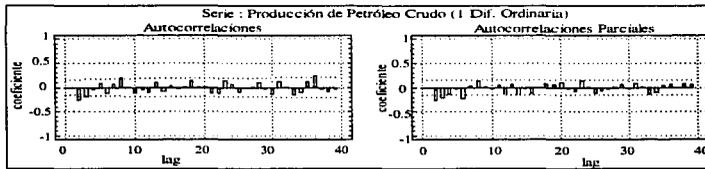


Figura 3.3 ACF y PACF de la serie "Producción de Petróleo Crudo" con una diferencia.

A partir de las siguientes series y con la ayuda de las ACF y PACF identificaremos los modelos que posiblemente se ajusten a la serie en estudio.

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

En la ACF de la Serie # 1 indica falta de estacionaridad ya que decrece lentamente, mientras que en la Serie # 2 con una diferencia se convierte en estacionaria con algunos valores significativos en 2, 3 y 8. En cuanto a la PACF la Serie # 1 tiene un valor significativo en 1 y para la # 2 observamos que decrece.

En base a estas funciones podemos *identificar* algunos modelos tentativos como:

1) ARIMA(1,0,0):

Operador : Autorregresivo no estacional de orden 1  
Diferencias : Ninguna

2) ARIMA(2,0,0):

Operador : Autorregresivo no estacional de orden 2  
Diferencias : Ninguna

3) ARIMA(0,1,2):

Operador : Promedios Móviles no estacional de orden 2  
Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

4) ARIMA(1,1,2):

Operador : Autorregresivo no estacional de orden 1  
Promedios Móviles no estacional de orden 2  
Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

En la tabla 3.2 se presentan los resultados numéricos necesarios para llevar a cabo la verificación, dicha tabla incluye los estadísticos t, el de la Chi-Cuadrada de Box-Pierce e indica los intervalos (lags) en los cuales los residuales poseen autocorrelaciones significativas. Con la ayuda de esta información y la gráfica de los residuales, será relativamente simple seleccionar un modelo dentro de varias alternativas.

Modelo	Parámetros	Estimación	Error Estnd	Error Estándar R.	Valor de T	Estadístico Q, g.l.	Residuales Significativos ACF y PACF
ARIMA(1,0,0)	AR(1)	0.72017	0.05492	90.1373	13.11245	35.9775,	8, 13, 36
						164	8, 13
ARIMA(2,0,0)	AR(1)	0.78845	0.07835	89.9715	10.0634	35.3016,	8, 13, 36
	AR(2)	-0.09617	0.07820		-1.22981	163	8, 13
ARIMA(0,1,2)	MA(1)	0.14798	0.07361	89.6874	2.01035	26.194, 161	36
	MA(2)	0.36752	0.07352		4.99866,		8, 12
ARIMA(1,1,2)	AR(1)	0.31477	0.19423	89.1889	1.62059	22.34, 160	36
	MA(1)	0.42304	0.18609		2.27339		8, 12
	MA(2)	0.29259	0.09459		3.09315		

Tabla 3.2 Modelos de la Producción de Petróleo Crudo.

### 3.2.2 Serie de la Producción de Gas Natural

En la figura 3.4 se observa que hasta julio de 1985 la producción de gas tiene un notable descenso, para después conservar un nivel constante hasta 1995 que es cuando se da un incremento debido al mayor uso que se le empezó a dar dentro de algunos sectores. No hay varianza en los datos y no se observa variación estacional.

Es notoria la falta de estacionaridad pues la ACF decrece lentamente mientras que la PACF se corta en 3, lo podemos ver en la figura 3.5.

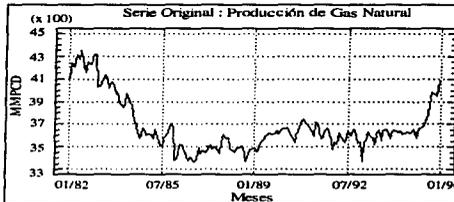


Figura 3.4 Serie de Tiempo de la Producción de Gas Natural.

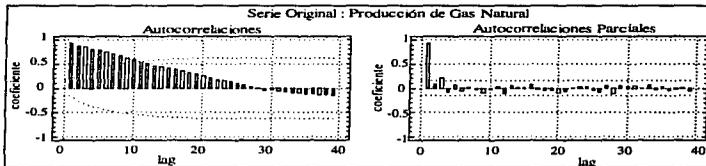


Figura 3.5 ACF y PACF de la serie "Producción de Gas Natural".

Al aplicar una diferencia ordinaria para eliminar la tendencia se observa en la ACF y PACF que es ruido blanco (ver fig. 3.6).

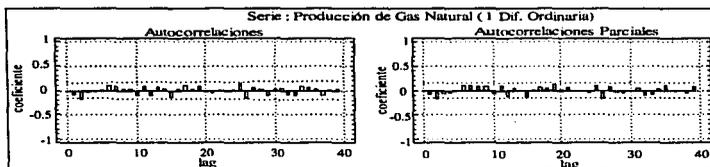


Figura 3.6 ACF y PACF de la serie "Producción de Gas Natural" con una diferencia.

A pesar de lo anterior se tratará de identificar algún modelo a partir de las siguientes series:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

En las figuras de las ACF's, la Serie # 1 indica falta de estacionaridad ya que decrece lentamente, no así en la Serie # 2 la cual podría considerarse como ruido blanco debido a que no se visualizan valores significativos, solo uno pequeño en el lag 2. En cuanto a la PACF la Serie # 1 tiene valores significativos en 1 y 3, y para la # 2 tiene picos poco significativos en 2 y 26.

Analizando estas funciones podemos *identificar* algunos modelos tentativos como:

1) ARIMA(1,0,0):

Operador : Autorregresivo no estacional de orden 1  
Diferencias : Ninguna

2) ARIMA(1,1,0):

Operador : Promedios Móviles no estacional de orden 1  
Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

3) ARIMA(1,1,0):

Operador : Autorregresivo no estacional de orden 1  
Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

4) ARIMA(1,1,1):

Operador : Autorregresivo no estacional de orden 1  
Promedios Móviles no estacional de orden 1  
Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

Al igual que la serie de Producción de Crudo se evaluarán los modelos de acuerdo a los resultados numéricos presentados en la tabla 3.3. Con estos datos y el análisis de residuales se seleccionará un modelo de entre ellos.

Modelo	Parámetros	Estimación	Error Estnd	Error Estándar R.	Valor de T	Estadístico Q. g.l.	Valores Significativos Residuales
ARIMA(1,0,0)	AR(1)	0.9901	0.01106	67.0267	89.49652	20.0804, 165	
ARIMA(0,1,1)	MA(1)	0.11595	0.07991	67.2239	1.45103	19.7389, 163	2 2, 26
ARIMA(1,1,0)	AR(1)	-0.7879	0.07968	67.3461	0.98895	19.822, 163	2 2, 26
ARIMA(1,1,1)	AR(1) MA(1)	0.40946 0.58449	0.28320 0.24634	66.7173	1.44582 2.37265	17.0592, 162	0 26

Tabla 3.3 Modelos de la Producción de Gas Natural.

Es importante hacer una observación del modelo ARIMA(1,0,0), pues el valor del parámetro es muy cercano a 1 y con esto podría pensarse que en realidad la serie de Producción de Gas Natural está siendo generada a partir del valor anterior de la serie más menos un error aleatorio y no precisamente por un modelo del tipo ARIMA.

### 3.2.3 Serie de las Ventas de Productos Petrolíferos

En la figura 3.7 se presentan los datos históricos de la serie, donde observamos un notable decremento en 1995, esto se debió a la crisis de ese año y que afectó el consumo de productos petrolíferos.

Las características de la serie son : tienen varianza constante y tendencia creciente, indicio de que la serie no es estacionaria, y esto lo podemos corroborar en la ACF que aparece en la figura 3.8 la cual decrece lentamente.

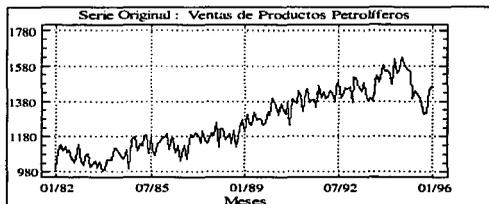


Figura 3.7 Serie de Tiempo de las Ventas de Productos Petrolíferos.

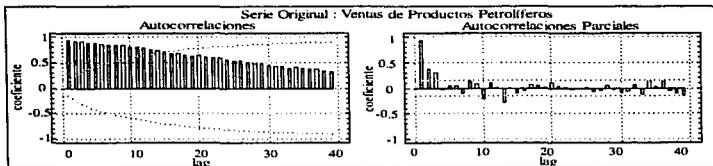


Figura 3.8 ACF y PACF de la serie "Ventas de Productos Petrolíferos".

En la figura 3.9 se presentan las ACF y PACF de la serie con una diferencia ordinaria.

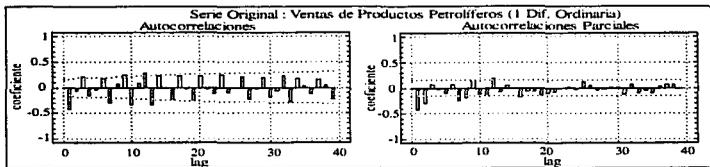


Figura 3.9 ACF y PACF de la serie "Ventas de Productos Petrolíferos" con una diferencia.

Junto con la diferencia ordinaria se considera también una estacional (ver fig. 3.10), por el comportamiento que presentan aún las ACF y PACF en la figura 3.9.

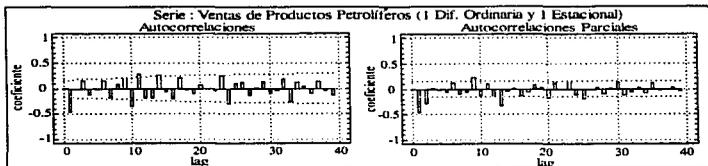


Figura 3.10 ACF y PACF de la serie "Ventas de Productos Petrolíferos" con una diferencia ordinaria y una estacional ( $s=12$ ).

A partir de las siguientes series identificaremos los modelos que se ajustan a la Ventas de Productos Petrolíferos:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

Serie # 3  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

En las ACF's observamos que la Serie # 1 indica falta de estacionaridad ya que decrece lentamente, no así en la Serie # 2 la cual tiende a decrecer, mientras que en la #3 se observan picos en 1, 10, 11 y 24. En cuanto a las PACF's la Serie # 1 se corta en el valor 13, para la # 2 se presentan picos en 1, 2, 7, 8 y 12, por último tenemos a la Serie #3 donde se observan autocorrelaciones significativas en 1, 2, 9 y 13.

Observando estas funciones podemos *identificar* algunos modelos tentativos como:

1) ARIMA(0,1,1)x(1,0,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

2) ARIMA(2,1,0)x(1,0,0)<sub>12</sub>:

Operador : Autorregresivo no estacional de orden 2

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

3) ARIMA(2,1,0)x(0,1,1)<sub>12</sub>:

Operador : Autorregresivo no estacional de orden 2

Promedios Móviles estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

4) ARIMA(0,1,1)x(0,1,1)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Promedios Móviles estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

El siguiente paso es estimar los parámetros de cada uno de ellos y examinar sus propiedades que se desglosan en la tabla 3.4, tales como el estadístico t, la Chi-Cuadrada de Box-Pierce y los intervalos (lags) en los cuales los residuales poseen autocorrelaciones significativas.

Modelo	Parámetros	Estimación	Error Estnd	Error Estándar R	Valor de T	Estadístico Q, g.l.	Residuales Significativos ACF y PACF
ARIMA(0,1,1) x (1,0,0) <sub>12</sub>	SAR(1)	0.28675	0.08211	50.6964	3.49234	55.8523,	7.10,13,16,23
	MA(1)	0.51557	0.06677		7.72180	164	7.10
ARIMA(2,1,0) x (1,0,0) <sub>12</sub>	AR(1)	-0.54312	0.07615	50.2921	-7.13260	33.0561,	7.10
	AR(2)	-0.27642	0.07816		-3.53669	163	7.10
ARIMA(2,1,0) x (0,1,1) <sub>12</sub>	SAR(1)	0.27698	0.08609		3.21736		
	AR(1)	-0.65318	0.07630	52.187	-8.5608	32.4031,	7.10
ARIMA(0,1,1) x (0,1,1) <sub>12</sub>	AR(2)	-0.37631	0.07672		-4.90474	151	7.10, 13
	SMA(1)	0.73867	0.06117		12.07613		
ARIMA(0,1,1) x (0,1,1) <sub>12</sub>	MA(1)	0.60000	0.06528	53.2592	9.19060	79.2341,	3.7,10,13
	SMA(1)	0.76248	0.05818		13.10585	152	3.7,10,13

Tabla 3.4 Modelos de la serie "Ventas de Productos Petrolíferos".

### 3.2.3.1 Serie de las Ventas de Gas Natural

Los datos históricos despliegan tendencia creciente y decreciente en diferente períodos y no se ve que exista variación estacional a simple vista como lo muestra la figura 3.11. Lo que sí es notable es que la serie no es estacionaria y esto lo podemos comprobar con las ACF y PACF, pues vemos que la primera decrece lentamente mientras que la PACF se corta en el valor 2 (ver fig. 3.12).

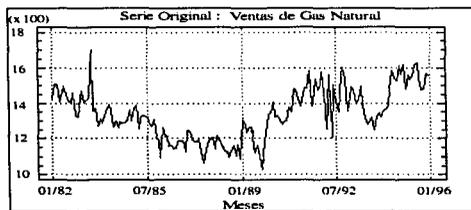


Figura 3.11 Serie de Tiempo de Ventas de Gas Natural.

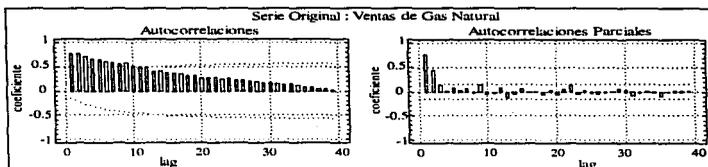


Figura 3.12 ACF y PACF de la serie "Ventas de Gas Natural".

Como se mencionó antes es notable la tendencia por lo que se decide aplicar una diferencia ordinaria. La figura 3.13 ilustra las ACF y PACF de la serie transformada.

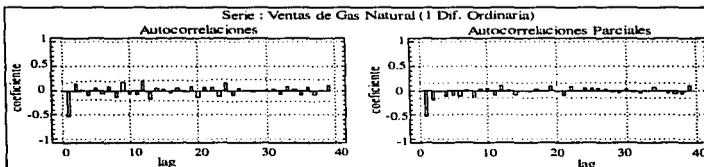


Figura 3.13 ACF y PACF de la serie "Ventas de Gas Natural" con una diferencia.

Las series a partir de las cuales se identificará los modelos de las Ventas de Gas Natural son las siguientes:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

La ACF de la Serie # 1 indica falta de estacionaridad ya que decrece lentamente, no así en la Serie # 2 la cual tiene una autocorrelación significativa en 1. En cuanto a la PACF la Serie #1 se corta en el intervalo 2, y decrece para la # 2.

En base a lo anterior podemos *identificar* algunos modelos tentativos como:

1) ARIMA(0,1,1):

Operador : Promedios Móviles no estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

2) ARIMA(1,1,1):

Operador : Autorregresivo no estacional de orden 1  
 Promedios Móviles no estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

El siguiente paso es la evaluación de los modelos que se ajustaron más y se llevará a cabo con los valores presentados en la tabla 3.5 donde se tienen los estimadores de los parámetros, algunos estadísticos y los intervalos (lags) en los cuales los residuales poseen autocorrelaciones significativas.

Modelo	Parámetros	Estimación	Error Estnd	Error Estándar R	Valor de T	Estadístico Q . g.l.	Valores Significativos Residuales
ARIMA(1,1)	MA(1)	0.53973	0.06640	77.7188	8.12861	13.2963,165	
ARIMA(1,1)	AR(1)	0.07457	0.14082	77.8913	0.52949	12.5466,	
	MA(1)	0.60171	0.11493		5.23529	164	

Tabla 3.5 Modelos de las Ventas de Gas Natural.

3.2.3.2 Serie de las Ventas de Gasolinas

En la figura 3.14 los datos históricos de la serie presentan tendencia creciente sin una variación significativa, en la ACF vemos que la serie no es estacionaria pues tiende a decrecer lentamente mientras que la PACF se corta en el intervalo 13 (ver fig. 3.15).

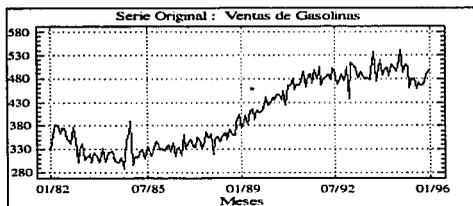


Figura 3.14 Serie de Tiempo de Ventas de Gasolinas.

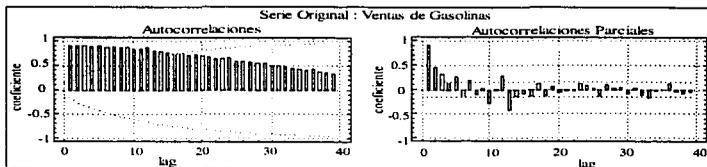


Figura 3.15 ACF y PACF de la serie "Ventas de Gasolinas".

Debido a que la serie no es estacionaria se aplicó una diferencia ordinaria quedando las ACF y PACF como se muestra en la figura 3.16.

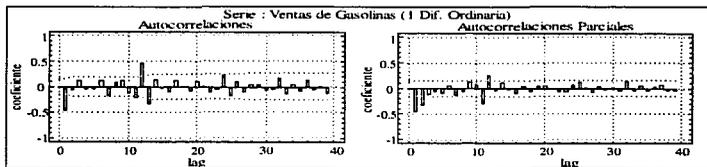


Figura 3.16 ACF y PACF de la serie "Ventas de Gasolinas" con una diferencia.

Observando las ACF y PACF con una diferencia ordinaria se comprueba que existen características estacionales en la serie por los picos significativos en 12 y 24, por lo que se decidió aplicar una diferencia estacional con  $s = 12$  (ver fig. 3.17).

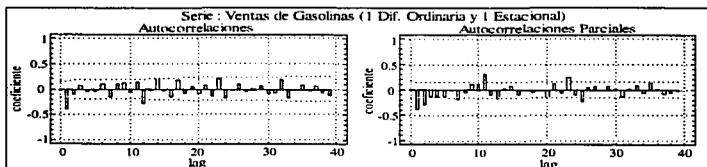


Figura 3.17 ACF y PACF de la serie "Ventas de Gasolinas" con una diferencia ordinaria y una estacional ( $s = 12$ ).

Las series que se derivaron y a partir de las cuales se identificarán los modelos son las siguientes:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

Serie # 3  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

Lo que se observó en las ACF's de las series fue que en la # 1 indica falta de estacionaridad ya que decrece lentamente, no así en la Serie # 2 la cual decrece pero con valores significativos en 12, 13 y 24, y en lo que se refiere a la Serie # 3 podría considerarse que se corta en la autocorrelación 12. En cuanto a las PACF's la Serie # 1 se corta en la autocorrelación 13 mientras que la Serie # 2 lo hace en la 12 y por último en la # 3 se diría que decrece.

Con ayuda de las ACF y PACF podemos *identificar* algunos modelos tentativos como:

1) ARIMA(0,1,1)x(2,0,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 2

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

2) ARIMA(0,1,1)x(1,0,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

3) ARIMA(0,1,1)x(0,1,1)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Promedios Móviles estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

4) ARIMA(0,1,1)x(0,1,2):

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 2

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

La evaluación de los modelos se hará con los resultados presentados en la tabla 3.6, donde se encuentran los estadísticos que se utilizarán para definir el mejor modelo.

Análisis para las Series de Tiempo de la Producción y las Ventas de Productos Derivados

Modelo	Parámetros	Estimación	Error Estnd	Error Estandar R.	Valor de T	Estadístico Q, g.l.	Residuales Significativo ACF y PACF
ARIMA(0,1,1) x (2,0,0) <sub>12</sub>	SAR(1)	0.35440	0.08050	17.677	4.40278	24.9113,	0
	SAR(2)	0.12189	0.08368		1.45656	163	0
	MA(1)	0.59568	0.06234		9.55452		
ARIMA(0,1,1) x (1,0,0) <sub>12</sub>	SAR(1)	0.39414	0.07579	17.7349	5.20020	20.7892,	0
	MA(1)	0.59210	0.06219		9.52087	164	13
ARIMA(0,1,1) x (0,1,1) <sub>12</sub>	MA(1)	0.59561	0.06527	17.7334	9.12491	29.9043,	9
	SMA(1)	0.66805	0.06902		9.67883	152	9, 14
ARIMA(0,1,1) x (0,1,2) <sub>12</sub>	MA(1)	0.59148	0.06610	17.5352	8.948339	20.6825,	9
	SMA(1)	0.53149	0.08280		6.41882	151	9
	SMA(2)	0.19170	0.08442		2.27075		

Tabla 3.6 Modelos de la Serie "Ventas de Gasolinas".

### 3.2.3.3 Serie de las Ventas de Gas Licuado

Las Ventas de Gas Licuado mensuales para 1982 a 1995 están graficadas en la figura 3.18.

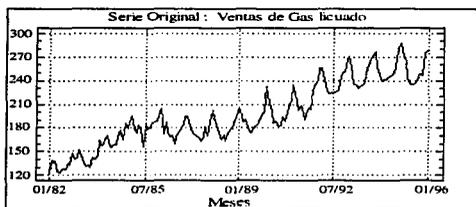


Figura 3.18 Serie de Tiempo de las Ventas de Gas Licuado.

Los datos históricos de la serie de Ventas de Gas Licuado tienen una tendencia creciente sin varianza en los datos y es notable la variación estacional, lo anterior lo podemos verificar en las ACF y PACF pues la primera decrece lentamente y tiene picos en los valores de 12, 24, etc., mientras que la PACF se corta en la autocorrelación 13 (ver fig. 3.19).

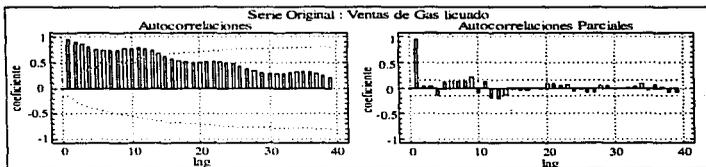


Figura 3.19 ACF y PACF de la serie "Ventas de Gas Licuado".

Con la existencia de tendencia es necesario eliminarla por lo que se aplicará una diferencia ordinaria. En la figura 3.20 se muestra las ACF y PACF de la serie transformada.

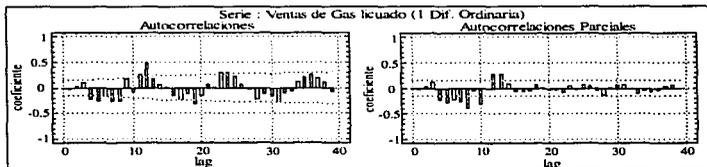


Figura 3.20 ACF y PACF de la serie "Ventas de Gas Licuado" con una diferencia.

En la figura 3.20 se visualiza que existe variación estacional, por lo que se aplicará además una diferencia estacional con periodo  $s = 12$  para eliminarla, en la figura 3.21 se tiene la transformación de la serie.

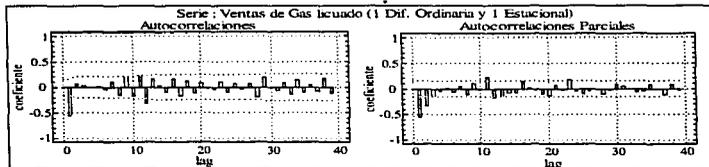


Figura 3.21 ACF y PACF de la serie "Ventas de Gas Licuado" con una diferencia ordinaria y una estacional ( $s = 12$ ).

Se consideran para la serie de Ventas de Gas Licuado los dos tipos de diferenciación, la ordinaria y la estacional a partir de aquí identificaremos los modelos de las siguientes series derivadas:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

Serie # 3  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

La Serie # 1 indica falta de estacionaridad ya que la ACF decrece lentamente, no así en la Serie # 2 la cual refleja variación estacional pues tiene algunos valores significativos en 11, 12, 23 y 24, en la Serie # 3 la ACF se corta en la autocorrelación 12. Por lo que se refiere a las PACF's, la Serie # 1 tiende a cortarse en 1, para la Serie # 2 se corta después del intervalo 13, mientras que para la #3 tiene valores significativos en 1, 2, 11, 12 y 23.

Con ayuda de las ACF y PACF podemos *identificar* algunos modelos tentativos como:

- 1) ARIMA(0,1,1):

Operador : Promedios Móviles no estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

- 2) ARIMA(0,1,1)x(0,1,1)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Promedios Móviles estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

- 3) ARIMA(0,1,1)x(1,1,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

- 4) ARIMA(1,1,0)x(1,0,0)<sub>12</sub>:

Operador : Autorregresivo no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

- 5) ARIMA(0,1,1)x(1,0,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

- 6) ARIMA(0,1,1)x(2,0,0)<sub>12</sub>:

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 2

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

En la tabla 3.7 se tiene los resultados numéricos para llevar a cabo la verificación de los modelos y poder decidir cuál es el mejor. Esta integrada por los estadísticos  $t$ , el de la Chi-Cuadrada de Box-Pierce, la estimación de los parámetros y las autocorrelaciones significativas.

Modelo	Parámetros	Estimación	Error Estnd	Error Estandar R.	Valor de T	Estadístico Q, g.l.	Residuales Significativos ACF y PACF
ARIMA(0,1,1)	MA(1)	0.66412	0.05959	8.2166	11.14475	32.0519, 153	12
ARIMA(0,1,1)	MA(1)	0.64228	0.06165	7.6166	10.41781	21.108,	12, 24
* (0,1,1) <sub>12</sub>	SMA(1)	0.49664	0.07623		6.51480	152	9
ARIMA(0,1,1)	MA(1)	0.65204	0.06064	7.84229	10.75712	21.7656,	9
* (1,1,0) <sub>12</sub>	SAR(1)	-0.31473	0.07954		-3.95688	152	9, 24
ARIMA(1,1,0)	AR(1)	-0.47808	0.07072	8.22951	-6.76054	32.7453,	0
* (1,0,0) <sub>12</sub>	SAR(1)	0.75113	0.06057		12.40053	164	2, 8, 12
ARIMA(0,1,1)	MA(1)	0.63631	0.06204	7.97715	10.25585	31.0739,	0
* (1,0,0) <sub>12</sub>	SAR(1)	0.84050	0.05362		15.67400	164	12
ARIMA(0,1,1)	MA(1)	0.63670	0.06127	7.80955	10.39166	24.2985,	9
* (2,0,0) <sub>12</sub>	SAR(1)	0.66724	0.08001		8.33947	163	9, 24
	SAR(2)	0.24068	0.08398		2.86604		

Tabla 3.7 Modelos de la serie "Ventas de Gas Licuado".

### 3.2.3.4 Serie de las Ventas de Turbosinas

Los datos históricos de la serie que se muestran en la figura 3.22 presentan tendencia creciente sin variación en los datos y no es muy notoria la variación estacional, con las ACF comprobamos que no es estacionaria pues decrece lentamente mientras que la PACF se corta en la autocorrelación 13 (fig. 3.23).

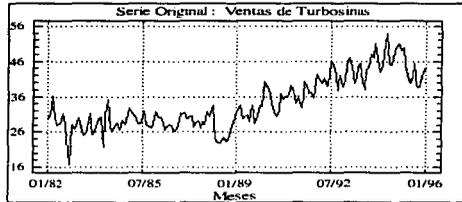


Figura 3.22 Serie de Tiempo de Ventas de Turbosinas.

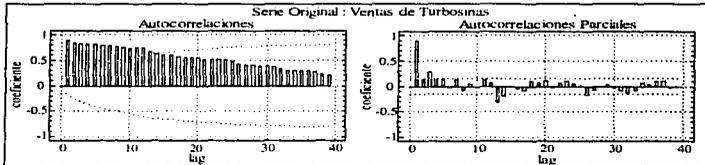


Figura 3.23 ACF y PACF de la serie "Ventas de Turbosinas".

Se aplica una diferencia ordinaria para transformar a la serie en estacionaria y se ilustra en la figura 3.24.

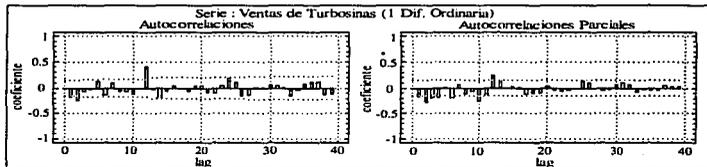


Figura 3.24 ACF y PACF de la serie "Ventas de Turbosinas" con una diferencia.

Debido al comportamiento de las ACF y PACF de la fig. 3.24 se decide aplicar también una diferencia estacional (véase figura 3.25).

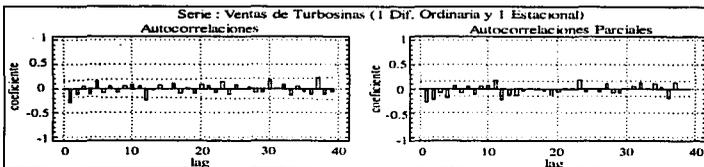


Figura 3.25 ACF y PACF de la serie "Ventas de Turbinas" con una diferencia ordinaria y una estacional ( $s = 12$ ).

Se consideran para la serie de Ventas de Turbinas los dos tipos de diferenciación, la ordinaria y la estacional a partir de aquí identificaremos los modelos de las siguientes series:

Serie # 1  $w_t = y_t$

Serie # 2  $w_t = \Delta y_t = y_t - y_{t-1}$

Serie # 3  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

Como se mostró en las ACF's, la Serie # 1 indica falta de estacionaridad ya que decrece lentamente, no así en la Serie # 2 ya que se observa un comportamiento decreciente de una serie estacional con valores significativos en 12 y 24, en la # 3 sólo hay picos en 1 y 12 donde podemos decir que se corta. En cuanto a las PACF's, la Serie # 1 se corta en el valor 13, para la # 2 se ve que es decreciente, por último en la #3 los picos están en 1, 2, 4, 11, 12, 23 y 36.

En base a éstas funciones podemos *identificar* algunos modelos tentativos como:

1) ARIMA(0,1,1)x(1,0,0)<sub>12</sub>

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta y_t = y_t - y_{t-1}$

2) ARIMA(0,1,1)x(0,1,1)<sub>12</sub>

Operador : Promedios Móviles no estacional de orden 1

Promedios Móviles estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

3) ARIMA(0,1,1)x(1,1,0)<sub>12</sub>

Operador : Promedios Móviles no estacional de orden 1

Autorregresivo estacional de orden 1

Diferencias :  $w_t = \Delta \Delta^{12} y_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$

La estimación de los parámetros de los modelos así como sus respectivos estadísticos se presentan en la tabla 3.8.

Modelo	Parámetros	Estimación	Error Estnd	Error Estandar R.	Valor de T	Estadístico Q. g.l.	Residuales Significativos ACF y PACF
ARIMA(0,1,1) $x(1,0,0)_{12}$	SAR(1) MA(1)	0.48485 0.48293	0.07270 0.06974	2.75433	6.66930 6.92475	22.8342, 164	0 2
ARIMA(0,1,1) $x(0,1,0)_{12}$	SMA(1) MA(1)	0.52000 0.50886	0.07261 0.07024	2.77464	7.16112 7.24474	14.0339, 152	0 0
ARIMA(0,1,1) $x(1,1,0)_{12}$	SAR(1) MA(1)	-0.33391 0.50954	0.07919 0.07139	2.85918	-4.21668 7.13742	16.5126, 152	24 0

Tabla 3.8 Modelos de la serie "Ventas de Turbosinas".

### **3.3 Diagnóstico y Selección de los modelos con mejor ajuste a las series analizadas**

Una vez que se tienen los modelos identificados junto con la estimación de sus parámetros, será necesario verificar si pueden mejorarse.

Una de las formas más claras y simples para detectar violaciones a los supuestos de los modelos es a través del análisis de los residuales, en donde, como residual se considera aquella parte de las observaciones que no es explicada por el modelo. Los residuales miden la discrepancia entre los valores observados y los valores estimados por el modelo. Además cuando el tamaño de muestra es grande, los errores aleatorios y los residuales (que también son variables aleatorias) son esencialmente iguales; por esta razón, al analizar los residuales observados  $\{\hat{a}_t\}$  se analiza básicamente lo que debería ser una realización del proceso de ruido blanco  $\{a_t\}$ .

Por esta razón se graficará los residuales de cada uno de los modelos que se eligieron como mejores y que se presentan en la tabla 3.9.

En las figura 3.26 a la 3.37 se presentan las gráficas de los residuales de cada uno de los modelos, en ellas podemos observar que en casi todas los residuales se puede considerar como ruido blanco, además del periodograma integral que indica un buen ajuste.

### 3.3.1 Producción de Petróleo Crudo y Derivados

Modelo	Parámetros	Estimación	Media y Varianza Residual	Error Estándar R.	Valor de T	Estadístico Q. g.l.	ACF y PACF Significativas Residuales
--------	------------	------------	---------------------------	-------------------	------------	---------------------	--------------------------------------

#### Producción de Petróleo Crudo

ARIMA(1,0,0)	AR(1)	0.72017	3.36581 8124.74	90.1373	13.11245	35.9775.	8, 13, 36 164 8, 13
ARIMA(0,1,2)	MA(1) MA(2)	0.14798 0.36752	1.43693 8043.84	89.6874	2.01035 4.99866	26.194. 161	36

#### Venta de Productos Petrolíferos

ARIMA(2,1,0)	AR(1) AR(2)	-0.54312 -0.27642	0.380234 2529.29	50.2921	-7.13260 -3.53669	33.0561. 163	7, 10 7, 10
ARIMA(2,1,0)	AR(1) AR(2)	0.27698 -0.65318	0.338217 2723.49	52.187	3.21736 -8.5608	32.4031. 151	7, 10 7, 10, 13
ARIMA(2,1,0)	AR(1) AR(2) SMA(1)	-0.37631 0.73867			-4.90474 12.07613		

#### Venta de Gas Licuado

ARIMA(1,1,0)	AR(1) SAR(1)	-0.47808 0.75113	-0.0750975 67.7248	8.22951	-6.76054 12.40053	32.7453. 164	0 2, 8
ARIMA(0,1,1)	MA(1) SMA(1)	0.64228 0.49664	-0.0493447 58.0126	7.6166	10.41781 6.51480	21.108, 152 9	9

#### Venta de Gasolinas

ARIMA(0,1,1)	MA(1) SAR(1)	0.59210 0.39414	0.400003 314.528	17.7349	9.52087 5.20020	20.7892. 164	0 0
ARIMA(0,1,1)	MA(1) SMA(1)	0.59148 0.53149	0.660422 307.482	17.5352	8.94839 6.41882	20.6825. 151	9 9
ARIMA(0,1,1)	MA(2) SMA(2)	0.19170			2.27075		

#### Venta de Turbosinas

ARIMA(0,1,1)	MA(1) SAR(1)	0.48293 0.48485	0.0497764 7.58636	2.75433	6.92475 6.66930	22.8342. 164	0 2
ARIMA(0,1,1)	MA(1) SMA(1)	0.50886 0.52000	0.03834 7.69865	2.77464	7.24474 7.16112	14.0339. 152	0 0

Tabla 3.9 Mejores Modelos de las Series en Estudio.

Serie : Producción de Petróleo Crudo

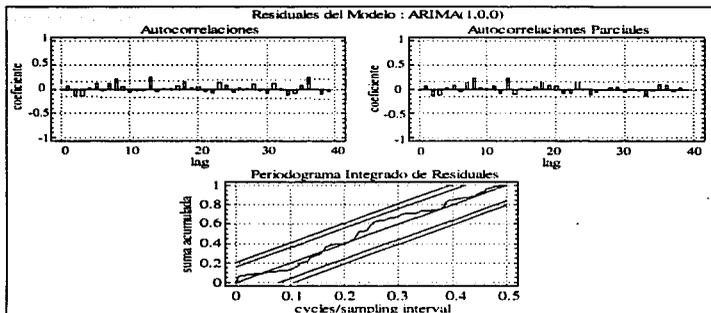


Figura 3.26 Residuales y Periodograma del modelo ARIMA(1,0,0) de la serie "Producción de Crudo".

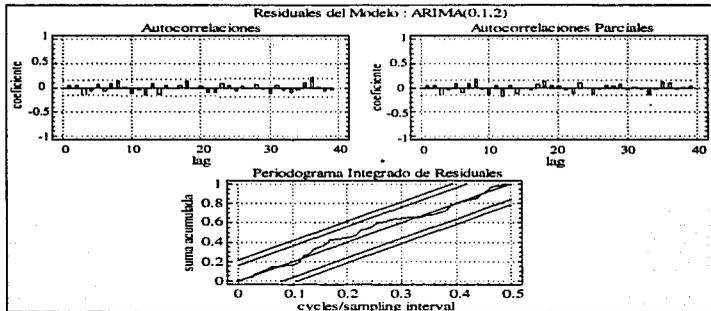


Figura 3.27 Residuales y Periodograma del modelo ARIMA(0,1,2) de la serie "Producción de Crudo".

Serie : Ventas de Productos Petrolíferos

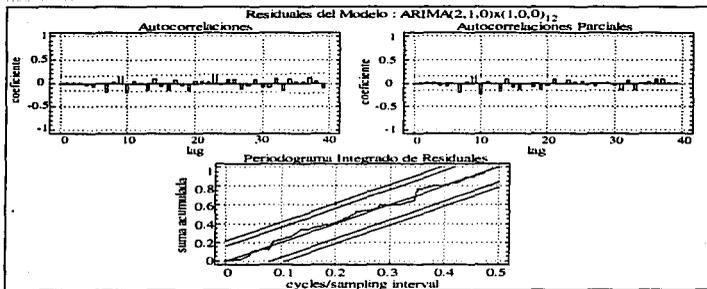


Figura 3.28 Residuales y Periodograma del modelo ARIMA(2,1,0)x(1,0,0)<sub>12</sub> de la serie "Ventas de Productos Petrolíferos".

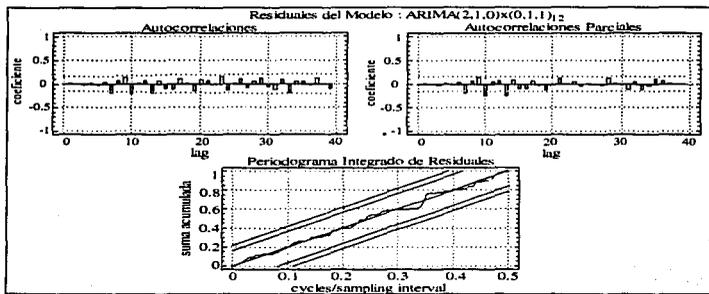


Figura 3.29 Residuales y Periodograma del modelo ARIMA(2,1,0)x(0,1,1)<sub>12</sub> de la serie "Ventas de Productos Petrolíferos".

Serie : Ventas de Gas Licuado

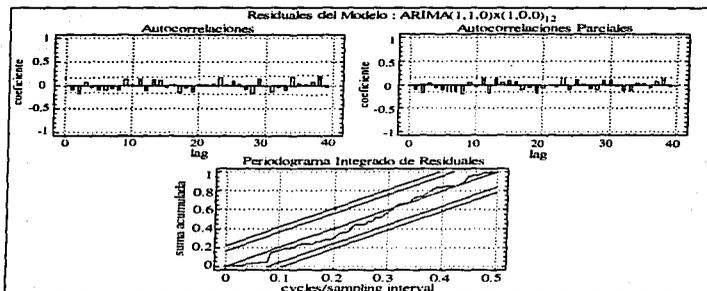


Figura 3.30 Residuales y Periodograma del modelo ARIMA(1,1,0)x(1,0,0)<sub>12</sub> de la serie "Ventas de Gas Licuado".

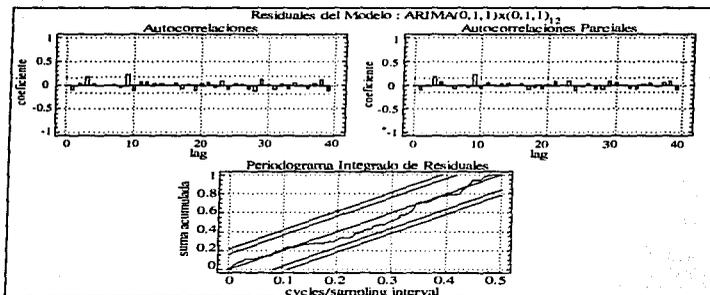


Figura 3.31 Residuales y Periodograma del modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub> de la serie "Ventas de Gas Licuado".

Serie: Ventas de Gasolinas

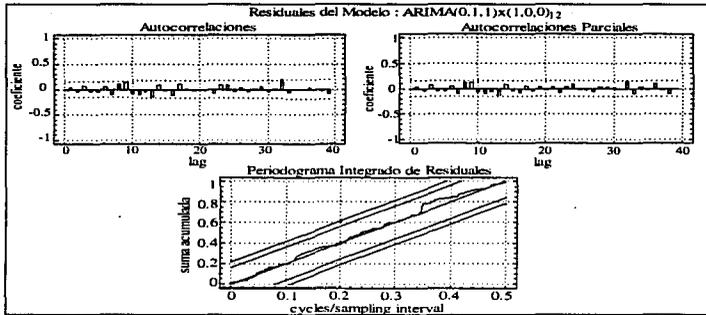


Figura 3.32 Residuales y Periodograma del modelo  $ARIMA(0,1,1) \times (1,0,0)_2$  de la serie "Ventas de Gasolinas".

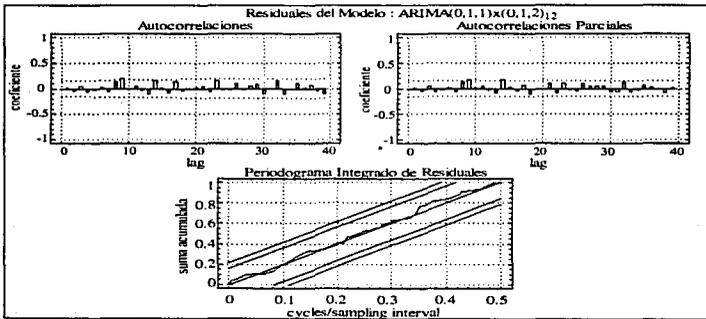


Figura 3.33 Residuales y Periodograma del modelo  $ARIMA(0,1,1) \times (0,1,2)_2$  de la serie "Ventas de Gasolinas".

Serie: Ventas de Turbinas

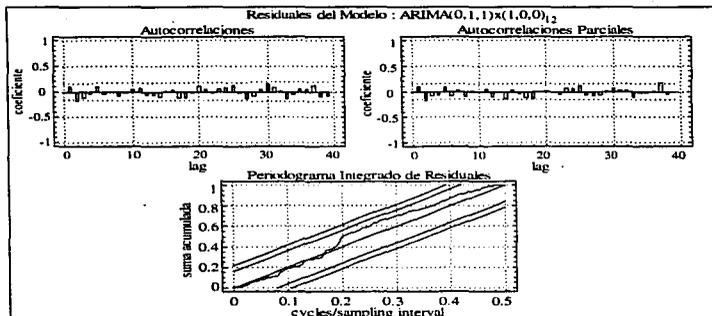


Figura 3.34 Residuos y Periodograma del modelo ARIMA(0,1,1)x(1,0,0)<sub>12</sub> de la serie "Ventas de Turbinas".

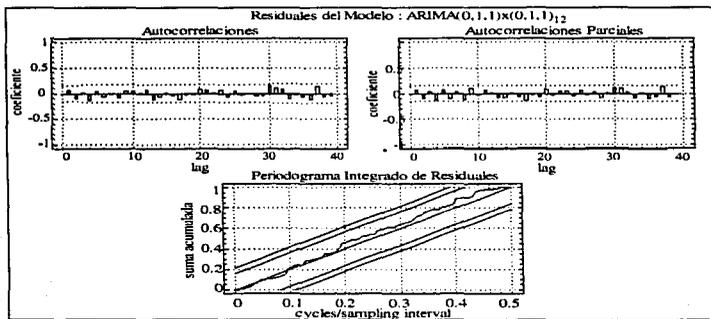


Figura 3.35 Residuos y Periodograma del modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub> de la serie "Ventas de Turbinas".

### 3.3.2 Producción de Gas Natural y Derivados

Modelo	Parámetros	Estimación	Media y Varianza Residual	Error Estándar R.	Valor de T	Estadístico Q, g.l.	Residuales Significativos ACF y PACF
<b>Producción de Gas Natural</b>							
ARIMA(1,0,0)	AR(1)	0.9901	-4.15025 4492.57	67.0267	89.49652	20.0804, 165	
<b>Ventas de Gas Natural</b>							
ARIMA(0,1,1)	MA(1)	0.53973	-813.43 6140.21	77.7188	8.12861	13.2963, 165	

Tabla 3.10 Mejores Modelos de las Series en Estudio de la Producción de Gas Natural y Ventas de Productos derivados.

A continuación tenemos los residuales de los modelos de la serie de Producción de Gas Natural.

#### Serie : Producción de Gas Natural

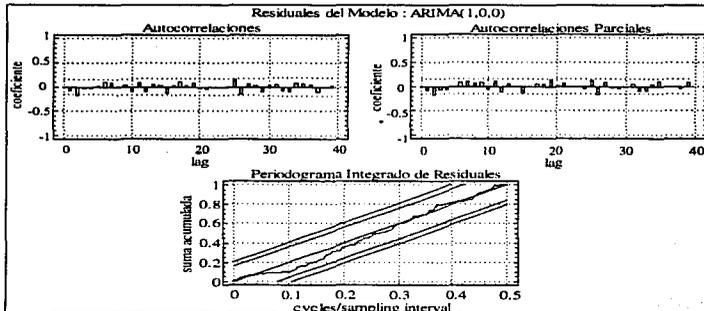


Figura 3.36 Residuales y Periodograma del modelo ARIMA (1,0,0) de la serie "Producción de Gas Natural".

Serie : Ventas de Gas Natural

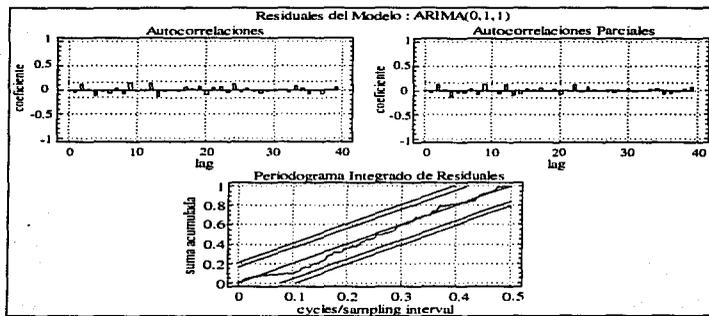


Figura 3.37 Residuales y Periodograma del modelo ARIMA (0,1,1) de la serie "Ventas de Gas Natural".

---

## **CAPITULO 4**

### **PRONÓSTICOS Y SIMULACIÓN DE LAS SERIES DE TIEMPO DE PRODUCCIÓN Y VENTAS DE PRODUCTOS DERIVADOS**

---

---

#### **4.1 Pronósticos**

La última etapa de la estrategia de Box y Jenkins de construcción de modelos viene a ser la del uso del modelo para los fines para los cuales fue construido. Ya que uno de los fines más frecuentes al construir un modelo para una serie de tiempo dada es el de pronóstico.

Una vez que ya construimos los modelos para representar a las series de tiempo en estudio, ya estamos en condiciones de obtener pronósticos, por lo que resulta conveniente y quizá necesario, analizar y comparar la capacidad de pronóstico de cada uno de los modelos en consideración.

En las gráficas 4.1 a la 4.11 se presenta los pronósticos a 2 años junto con sus intervalos de confianza al 95% generados con los modelos que se ajustaron mejor. Además de los pronósticos también se graficó la serie original para visualizar cómo era el comportamiento de la serie y cuál será el pronóstico del mismo para los próximos años.

Haciendo una evaluación de los modelos junto con el personal que utiliza la información de las series de tiempo, se llegó a la conclusión de seleccionar los que se indican en la siguiente sección.

Los modelos que generaron pronósticos más adecuados son los de las series de Gas Licuado, Gasolinas y Turbosinas.

#### 4.1.1 Serie de la Producción de Petróleo Crudo

El modelo ARIMA(0,1,2) fue el que se decidió elegir como el mejor, y se utilizará en la simulación más adelante. Se puede ver en la figura 4.1 los pronósticos con este modelo.

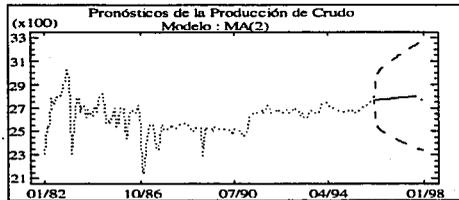


Figura 4.1 Pronósticos con el modelo ARIMA(0,1,2) de la serie "Producción de Crudo".

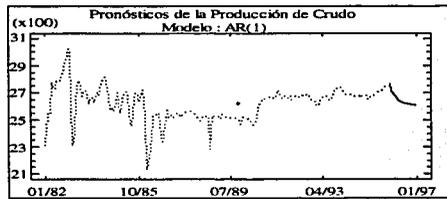


Figura 4.2 Pronósticos con el modelo ARIMA(1,0,0) de la serie "Producción de Crudo".

#### 4.1.1.1 Serie de las Ventas de Productos Petrolíferos

En la figura 4.3 se tienen los pronósticos del modelo  $ARIMA(2,1,0) \times (1,0,0)_{12}$  que se seleccionó como mejor de las Ventas de Productos Petrolíferos

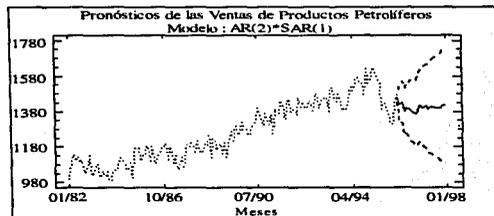


Figura 4.3 Pronósticos con el modelo  $ARIMA(2,1,0) \times (1,0,0)_{12}$  de la serie "Ventas de Productos Petrolíferos".

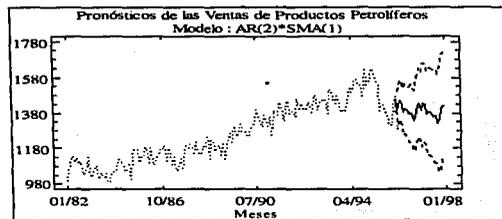


Figura 4.4 Pronósticos con el modelo  $ARIMA(2,1,0) \times (0,1,1)_{12}$  de la serie "Ventas de Productos Petrolíferos".

4.1.1.1.1 Serie de las Ventas de Gas Licuado

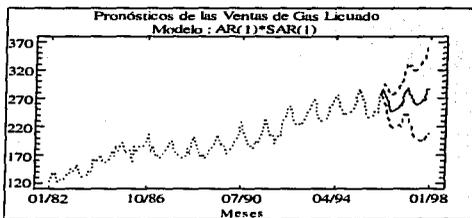


Figura 4.5 Pronósticos con el modelo  $ARIMA(1,1,0) \times (1,0,0)_{12}$  de la serie "Ventas de Gas Licuado".

El modelo  $ARIMA(1,1,0) \times (1,0,0)_{12}$  fue elegido para generar pronósticos de las Ventas de Gas Licuado que se observan en la figura 4.6. Además se utilizará para realizar una simulación en el siguiente apartado.

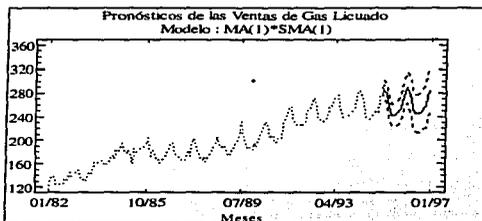


Figura 4.6 Pronósticos con el modelo  $ARIMA(0,1,1) \times (0,1,1)_{12}$  de la serie "Ventas de Gas Licuado".

4.1.1.1.2 Serie de las Ventas de Gasolinas

En la figura 4.7 se observan los pronósticos que se generaron con el modelo  $ARIMA(0,1,1) \times (1,0,0)_{12}$ , seleccionado como mejor para la serie de Ventas de Gasolinas.

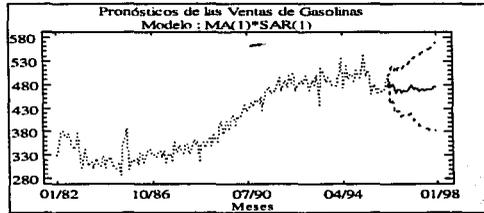


Figura 4.7 Pronósticos con el modelo  $ARIMA(0,1,1) \times (2,0,0)_{12}$  de la serie "Ventas de Gasolinas".

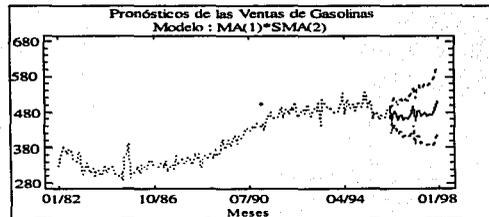


Figura 4.8 Pronósticos con el modelo  $ARIMA(0,1,1) \times (0,1,2)_{12}$  de la serie "Ventas de Gasolinas".

#### 4.1.1.1.3 Serie de las Ventas de Turbosinas

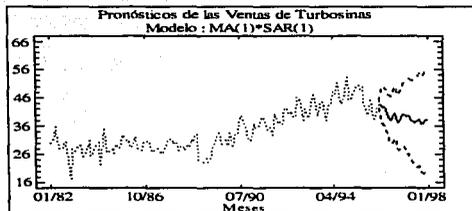


Figura 4.9 Pronósticos con el modelo ARIMA(0,1,1)x(1,0,0)<sub>12</sub> de la serie "Ventas de Turbosinas".

Con el modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub> se generaron los pronósticos de la figura 4.10, que se consideraron como buenos y por lo tanto se utilizará para generar una simulación.

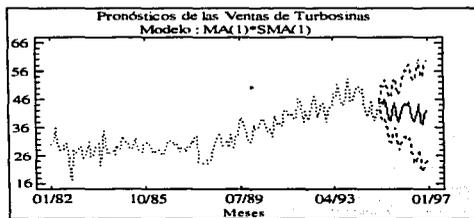


Figura 4.10 Pronósticos con el modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub> de la serie "Ventas de Turbosinas".

#### 4.1.2 Producción de Gas Natural

Con la serie de Producción de Gas Natural no se identificó un modelo adecuado, pero aún así se probó con el ARIMA(1,0,0).

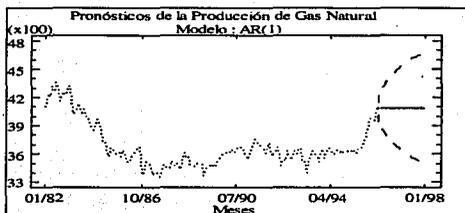


Figura 4.11 Pronósticos con el modelo ARIMA(1,0,0) de la serie "Producción de Gas Natural".

##### 4.1.2.1 Serie de las Ventas de Gas Natural

El modelo ARIMA(0,1,1) fue seleccionado para generar los pronósticos de las Ventas de Gas Natural.

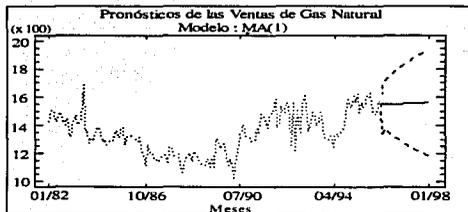


Figura 4.12 Pronósticos con el modelo ARIMA(0,1,1) de la serie "Ventas de Gas Natural".

## 4.2 Simulación del Fenómeno

En los últimos años, la simulación en computadoras se ha popularizado enormemente en muchas áreas, como la economía, la administración, etc., como medio para analizar el comportamiento de sistemas complejos. En la actualidad, la simulación en computadora abarca una amplia gama de aplicaciones que van desde aplicaciones específicas de negocios, por ejemplo, el control de inventarios y la planeación de producción, hasta las simulaciones de corporaciones completas.

Antes de seguir adelante es conveniente definir el significado de la simulación.

*"Simulación se define como una técnica numérica empleada para realizar experimentos con ciertos tipos de modelos matemáticos que describen el comportamiento de un sistema complejo, en una computadora digital y durante periodos prolongados"*<sup>1</sup>.

### *El por qué de la Simulación*

Para entender las razones que legitiman la simulación en computadoras como medio de análisis para modelos de sistemas de negocios y económicos, se debe señalar una similitud obvia, pero muy importante, entre la simulación en computadora y varias técnicas analíticas normales tales como el cálculo diferencial, la programación matemática y el cálculo de variaciones. En general, la razón principal para usar cualquiera de estos medios de análisis es la búsqueda de conocimientos científicos relacionados con el comportamiento de un sistema económico dado. Cuando se aplica a sistemas económicos, el método científico sigue el procedimiento conocido de 4 etapas: 1) la observación del sistema, 2) la formulación de un modelo matemático que trate de explicar las observaciones del sistema, 3) la predicción del comportamiento de sistema tomado como base del modelo, utilizando deducciones matemáticas o lógicas, es decir, mediante la obtención de soluciones al modelo y 4) la realización de experimentos para comprobar la validez de dicho modelo.

La simulación en computadoras se convierte en un medio esencial para analizar sistemas económicos cuando no se puede llevar a cabo uno o más de los cuatro pasos del método científico.

Los modelos de computadoras pueden hacerse tan complejos y realistas como lo permitan las teorías, ya que son innecesarias las soluciones analíticas de ellos. Sea cual fuere el grado de complejidad que tenga la formulación del modelo, las técnicas de simulación permiten encontrar sus consecuencias.

La simulación proporciona un medio para descubrir los efectos de las reglas de decisión alternativas sobre el comportamiento de la compañía dentro de los límites de un

---

<sup>1</sup> Thomas H. Naylor, Experimentos de Simulación en computadoras con Modelos de Sistemas Económicos, pp. 14.

experimento de laboratorio minuciosamente comprobado. Por supuesto, se puede argumentar que puede hacerse lo mismo con técnicas analíticas más tradicionales; pero con la simulación por computadora se puede experimentar con más variables, más reglas de decisión, con modelos más complejos y modelos que sean muy similares al comportamiento real de las empresas comerciales, y todo se puede hacer de una forma rápida.

*Otras razones que justifican el uso de la simulación son:*

- La simulación permite el estudio y la experimentación con instalaciones internas complejas de un sistema dado, ya sea que se trate de una compañía, una industria, un sistema económico o cualquier subsistema de éstas.
- Mediante la simulación se puede estudiar los efectos de ciertos cambios de información, organización y ambientales en el funcionamiento de un sistema, haciendo alteraciones en el modelo de dicho sistema y observando los efectos que los cambios tienen en sus comportamientos.
- La observación minuciosa del sistema simulado puede llevar a una mejor comprensión de dicho sistema y a hacer sugerencias que lo perfeccionen y que de otra manera no se podrían obtener.
- La experiencia en el diseño de un modelo de simulación en computadora puede ser más valiosa que la simulación propiamente dicha. Los conocimientos obtenidos al diseñar un estudio de simulación sugieren a menudo cambios en el sistema simulado. Los efectos de estos cambios se pueden comprobar después mediante la simulación, antes de aplicarlos al sistema real.
- La simulación de sistemas complejos puede ayudar a incrementar los conocimientos sobre qué variables son más importantes que otras dentro del sistema y su forma de interacción.
- La simulación se puede utilizar para experimentar con situaciones nuevas sobre las que se tenga pocos datos o ninguno a fin de estar preparados para lo que se pueda suceder.
- La simulación puede servir como una "prueba de prevención" para comprobar nuevas normas y reglas de decisión con el fin de manejar un sistema, antes de que se corra el riesgo al experimentar en el sistema real.
- Cuando se introducen elementos nuevos a un sistema, la simulación se puede emplear para anticipar estancamiento y otros problemas que puedan surgir en el comportamiento del sistema.

Como pudimos ver es grande el uso que se le puede dar a la simulación. Así pues el objetivo de la simulación de la producción de hidrocarburos y ventas de productos derivados será pronosticar valores a corto plazo que puedan ser utilizados para la planeación y toma de decisiones.

Una vez formulado el modelo y estimado sus parámetros, cosa que se llevó a cabo en el capítulo anterior, se debe tomar en cuenta la validez del mismo antes de utilizarlo para conducir experimentos de simulación de las distintas series de tiempo.

A través de la validación<sup>2</sup>, una de las etapas de un estudio de simulación, es posible detallar deficiencias en la formulación del modelo o en los datos alimentados al modelo. Las formas más comunes de validar un modelo son:

- a) La opinión de expertos sobre los resultados de la simulación.
- b) La exactitud con que se predicen datos históricos .
- c) La exactitud en la predicción del futuro.
- d) La aceptación y confianza en el modelo de la persona que hará uso de los resultados que arroje el experimento de simulación.

Por lo tanto, para cada una de las series de tiempo se efectuaron corridas de simulación, primero simulando los años de 1994 y 1995 que se presentan en las gráficas y después simulando los pronósticos de 1996 y 1997. Para llevar a cabo lo anterior se desarrollo un módulo en la herramienta de Fox-Pro para Windows versión 2.5, en el anexo se presentan más detalles del módulo.

En las figuras 4.13 a 4.19 se muestran las gráficas de las series cuyas trayectorias de tiempo se simularon y corresponden al periodo de 1994 a 1995. Las trayectorias de tiempo reales también aparecen graficadas con el objeto de establecer una comparación. Si se toman como base estas comparaciones graficadas de las producciones y ventas de productos petrolíferos simuladas del modelo y las trayectorias reales de tiempo de las series correspondientes, se llega a la conclusión de que los modelos logran simular con éxito el comportamiento de la producción de hidrocarburos y las ventas de productos petrolíferos en el periodo correspondiente.

Ahora faltaría ver hasta que grado estos modelos sirven para predecir el futuro, para esto se generaron algunas corrida para simular los pronósticos de cada una de las series en el periodo de 1996 a 1997. Los valores de las producciones y las ventas para cada mes aparecen en las tablas 4.1 a la 4.7. Donde se observó que no hubo mucha variación de una corrida a otra.

---

<sup>2</sup>Coss Bu Raúl. Simulación un Enfoque Práctico, pp. 13

#### 4.2.1 Serie de la Producción de Petróleo Crudo

La simulación se muestra en la figura 4.13, donde se tienen tanto los valores simulados como los observados reales.

Modelo ARIMA(0,1,2)

$$y_t = y_{t-1} - 0.14798c_{t-1} - 0.36752e_{t-2} + e_t$$

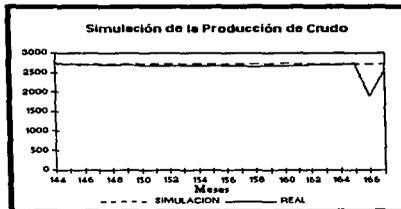


Figura 4.13 Simulación de la Serie "Producción de Crudo" (1994-1995).

Simulación de la Producción de Crudo (MBD)					
Datos	Fecha	1	2	4	
169	Ene 96	2761.5336	2762.3563	2763.6476	2763.2057
170	Feb 96	2760.6734	2761.9179	2762.979	2762.2636
171	Mar 96	2761.2202	2761.935	2763.7367	2762.5487
172	Abr 96	2760.0272	2761.2048	2763.1835	2762.0313
173	May 96	2760.0697	2761.0138	2764.3047	2760.7386
174	Jun 96	2760.842	2760.9164	2765.1657	2764.4271
175	Jul 96	2760.3611	2760.0987	2762.405	2763.9515
176	Ago 96	2759.29	2760.3844	2763.9975	2763.0303
177	Sep 96	2760.9383	2759.9051	2765.127	2763.2544
178	Oct 96	2760.3473	2758.5505	2764.2374	2762.0222
179	Nov 96	2759.052	2759.9586	2764.7801	2761.2336
180	Dic 96	2757.8997	2760.4677	2764.901	2761.7593
181	Ene 97	2759.1394	2759.0698	2765.55	2762.1415
182	Feb 97	2759.6695	2758.8171	2762.965	2761.5155
183	Mar 97	2758.3465	2758.1156	2763.6858	2761.7713
184	Abr 97	2757.6734	2759.7825	2764.6437	2761.5574
185	May 97	2756.9565	2759.8879	2762.5997	2759.787
186	Jun 97	2758.7128	2758.9156	2761.0073	2759.9385
187	Jul 97	2757.5085	2759.3189	2762.3387	2759.1424
188	Ago 97	2756.3533	2760.5764	2761.143	2760.7516
189	Sep 97	2755.646	2761.1042	2761.7364	2758.8298
190	Oct 97	2755.1966	2759.7527	2762.6255	2760.9888
191	Nov 97	2757.2066	2758.9581	2763.342	2760.6697
192	Dic 97	2756.8369	2760.3405	2762.6372	2759.6739

Tabla 4.1 Corridas de la Simulación de la Producción de Crudo (1996-1997).

#### 4.2.1.1 Serie de las Ventas de Productos Petrolíferos

La simulación es presentada en la figura 4.14, donde se tienen tanto los valores simulados como los observados reales.

Modelo ARIMA(2,1,0)x(1,0,0)<sub>12</sub> :

$$y_t = 0.45688y_{t-1} + 0.2667y_{t-2} + 0.27642y_{t-3} - 0.27698y_{t-12} - 0.12655y_{t-13} - 0.45688y_{t-14} + 0.45688y_{t-15} + e_t$$

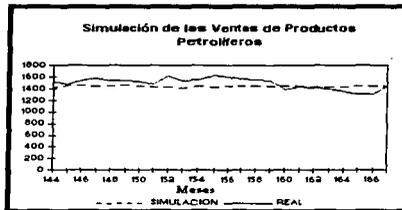


Figura 4.14 Simulación de la Serie "Ventas de Productos Petrolíferos" (1994-1995).

Simulación de las Ventas de Productos Petrolíferos (MDD)					
Datos	Fecha	1	2	3	4
169	Ene 96	1409.6849	1409.6638	1407.63	1412.4796
170	Feb 96	1422.8105	1421.8615	1423.1772	1423.0381
171	Mar 96	1422.3314	1422.3526	1422.5151	1425.4039
172	Abr 96	1372.7384	1374.8701	1372.4641	1376.1503
173	May 96	1390.5992	1390.0827	1389.6945	1392.3799
174	Jun 96	1381.6669	1382.0861	1382.0844	1384.8047
175	Jul 96	1375.9071	1376.1427	1378.5116	1382.6162
176	Ago 96	1368.0563	1368.2861	1368.498	1372.1758
177	Sep 96	1351.5982	1351.3616	1352.1109	1357.1592
178	Oct 96	1351.9033	1352.8569	1354.2481	1356.6888
179	Nov 96	1390.4669	1391.7141	1391.8201	1395.8851
180	Dic 96	1395.6044	1395.4705	1393.726	1397.8242
181	Ene 97	1380.9265	1381.2703	1380.1357	1383.4112
182	Feb 97	1383.1455	1386.0458	1385.7237	1386.4214
183	Mar 97	1384.6665	1385.7351	1383.6023	1387.872
184	Abr 97	1371.4885	1374.4687	1371.2921	1372.0919
185	May 97	1376.0522	1372.7707	1375.9746	1377.5501
186	Jun 97	1374.3283	1375.907	1375.4726	1373.642
187	Jul 97	1371.2103	1373.7568	1371.9618	1372.8451
188	Ago 97	1369.4587	1370.738	1371.1087	1370.0128
189	Sep 97	1365.3549	1366.5458	1366.0236	1365.3988
190	Oct 97	1366.267	1366.8206	1367.7392	1365.4436
191	Nov 97	1375.9063	1375.9915	1378.779	1377.2993
192	Dic 97	1375.4937	1379.4262	1378.0608	1377.661

Tabla 4.2 Corridas de la Simulación de las Ventas de Productos Petrolíferos (1996-1997).

#### 4.2.1.1.1 Serie de las Ventas de Gas Licuado

La simulación se muestra en la figura 4.15, donde se tienen tanto los valores simulados como los observados.

Modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub>:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \epsilon_t - 0.64228\epsilon_{t-1} - 0.49664\epsilon_{t-12} + 0.318982\epsilon_{t-13}$$

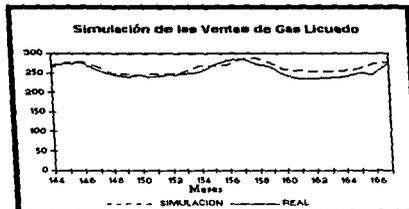


Figura 4.15 Simulación de la Serie "Ventas de Gas Licuado" (1994-1995).

Simulación de las Ventas de Gas Licuado (MBD)					
Datos	Fecha	1	2	3	4
169	Ene 96	282.9683	282.6759	286	284.8147
170	Feb 96	268.7989	267.9977	270.1949	267.3099
171	Mar 96	262.7078	261.6309	263.2491	263.5086
172	Abr 96	240.8422	241.1141	240.5265	240.1851
173	May 96	232.8906	233.5133	232.8032	235.8425
174	Jun 96	231.7251	233.9952	232.8671	233.8404
175	Jul 96	233.3734	233.9022	232.4646	236.5502
176	Ago 96	237.8489	240.2655	236.9473	239.0472
177	Sep 96	246.6176	248.7506	246.6811	248.6313
178	Oct 96	241.7565	243.3406	242.9892	246.0576
179	Nov 96	271.334	274.6881	272.0668	272.5179
180	Dic 96	275.7492	277.0488	277.7649	278.6503
181	Ene 97	279.5895	281.8817	281.5519	283.3274
182	Feb 97	263.97	267.3633	263.488	267.9154
183	Mar 97	257.5128	260.1199	260.8038	262.9567
184	Abr 97	236.1169	237.9709	237.1247	241.5645
185	May 97	229.1034	231.482	230.112	235.0556
186	Jun 97	226.7915	233.8137	228.4586	233.5945
187	Jul 97	227.9458	234.1443	229.4324	236.6881
188	Ago 97	233.7018	237.6437	232.7594	238.5073
189	Sep 97	243.9844	246.0891	242.0348	248.1467
190	Oct 97	238.3678	244.5619	238.9761	245.8309
191	Nov 97	270.1915	273.2956	267.1838	273.8705
192	Dic 97	272.3719	276.0878	273.4976	278.4658

Tabla 4.3 Corridas de la Simulación de Ventas de Gas Licuado (1996-1997).

#### 4.2.1.1.2 Serie de las Ventas de Gasolinas

La simulación se muestra en la figura 4.16, donde se tienen tanto los valores simulados como los observados.

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

$$y_t = y_{t-1} + 0.39414y_{t-12} - 0.39414y_{t-13} + e_t - 0.59210e_{t-1}$$

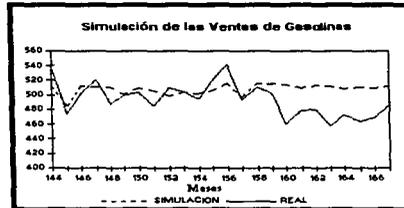


Figura 4.16 Simulación de la Serie "Ventas de Gasolinas" (1994-1995).

Simulación de las Ventas de Gasolinas (MBD)					
Datos	Fecha	1	2	3	4
169	Ene 96	477.2708	476.5483	477.8441	478.1293
170	Feb 96	484.4173	485.0063	483.5608	485.4074
171	Mar 96	480.7771	482.5852	482.7665	480.6835
172	Abr 96	466.2156	464.9183	463.4147	464.024
173	May 96	474.4365	474.2662	471.7234	472.2754
174	Jun 96	473.0636	472.3344	473.0211	473.0793
175	Jul 96	465.3064	464.7865	463.4145	465.4302
176	Ago 96	469.8772	468.0699	470.5986	469.6568
177	Sep 96	465.8904	466.5553	466.8666	464.9919
178	Oct 96	467.4635	468.0167	467.8521	468.4273
179	Nov 96	472.8473	475.0319	475.2985	476.1549
180	Dic 96	480.9511	477.8232	475.8815	480.7007
181	Ene 97	471.6669	471.2504	470.9547	472.3339
182	Feb 97	472.8231	475.6383	472.4214	475.3071
183	Mar 97	473.2048	473.2237	473.2398	474.5812
184	Abr 97	467.2369	466.3896	463.7095	467.5784
185	May 97	471.312	469.6633	465.6821	470.1968
186	Jun 97	468.8969	470.3054	468.4515	471.2748
187	Jul 97	466.453	467.1055	464.7578	468.5532
188	Ago 97	467.8252	468.0526	469.0905	471.0952
189	Sep 97	468.3165	466.8263	467.0635	467.6897
190	Oct 97	467.6381	468.2913	467.964	470.7318
191	Nov 97	470.548	471.7619	471.5077	474.176
192	Dic 97	472.6378	470.0746	471.5042	476.0908

Tabla 4.4 Corridas de la Simulación de Ventas de Gasolinas (1996-1997).

#### 4.2.1.1.3 Serie de las Ventas de Turbinas

La simulación se muestra en la figura 4.17, donde se tienen tanto los valores simulados como los reales.

Modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub>:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0.50886e_{t-1} - 0.52000e_{t-12} + 0.264607e_{t-13}$$

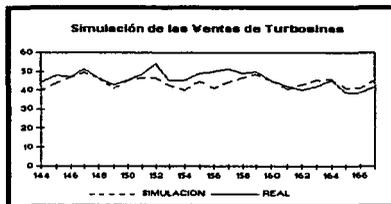


Figura 4.17 Simulación de la serie "Ventas de Turbinas" (1994-1995).

Simulación de las Ventas de Turbinas (MBD)					
Datos	Fecha	SIMU1	SIMU2	SIMU3	SIMU4
169	Ene 96	45.912	45.7595	44.878	46.8788
170	Feb 96	42.9083	43.6722	42.2685	46.7018
171	Mar 96	43.9291	42.9929	44.3611	46.65
172	Abr 96	37.1406	36.7	41.2279	41.5883
173	May 96	32.0829	36.259	38.338	38.9972
174	Jun 96	32.7545	31.953	36.4294	36.1954
175	Jul 96	34.3308	34.455	36.275	26.8385
176	Ago 96	36.1129	36.6952	40.9605	40.0684
177	Sep 96	28.8229	31.359	31.3628	35.2486
178	Oct 96	30.4768	32.1471	31.0926	34.0361
179	Nov 96	33.593	34.8769	36.4092	39.6375
180	Dic 96	34.9721	37.1721	37.6754	41.7897
181	Ene 97	36.6062	37.5623	38.7964	41.545
182	Feb 97	32.6571	37.3456	37.9268	42.7936
183	Mar 97	34.214	36.5488	38.0684	41.4779
184	Abr 97	29.6638	31.0631	35.5659	37.1569
185	May 97	24.2827	30.3428	31.5164	33.5719
186	Jun 97	24.4685	26.0662	29.0713	31.5404
187	Jul 97	26.6374	29.7721	31.8	33.6543
188	Ago 97	28.6801	31.4993	34.7287	38.142
189	Sep 97	22.6149	26.6685	24.859	31.5919
190	Oct 97	23.1032	26.937	25.6428	30.483
191	Nov 97	25.9597	29.1579	29.6706	33.1254
192	Dic 97	27.0463	31.237	30.5445	37.4839

Tabla 4.5 Corridas de la Simulación de Ventas de Turbinas (1996-1997).

#### 4.2.2 Serie de la Producción de Gas Natural

La simulación se muestra en la figura 4.18, donde se tienen tanto los valores simulados como los observados reales.

Modelo ARIMA(1,0,0)

$$y_t = 0.9901y_{t-1} + \epsilon_t$$

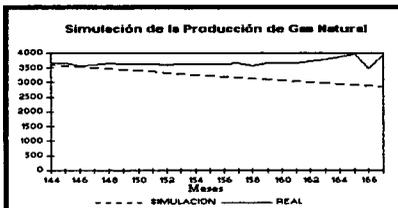


Figura 4.18 Simulación de la serie "Producción de Gas Natural" (1994-1995).

Simulación de la Producción de Gas Natural (MMPCD)					
Datos	Fecha	1	2	3	4
169	Ene 96	4041.7188	4041.5759	4041.0576	4042.2529
170	Feb 96	4001.3888	4001.1774	4002.1287	4001.9223
171	Mar 96	3961.3992	3963.3601	3965.2174	3961.9227
172	Abr 96	3923.0586	3926.1273	3927.2255	3923.144
173	May 96	3883.8003	3886.6837	3889.2625	3884.8588
174	Jun 96	3845.8696	3850.0933	3850.0843	3846.4022
175	Jul 96	3806.7169	3811.5269	3813.9039	3807.4854
176	Ago 96	3771.0144	3772.471	3776.9183	3771.0742
177	Sep 96	3734.7583	3735.6145	3740.1907	3734.5902
178	Oct 96	3697.3475	3697.6116	3702.9876	3697.2905
179	Nov 96	3659.9111	3660.6685	3666.067	3662.3418
180	Dic 96	3623.509	3624.63	3629.0572	3627.2541
181	Ene 97	3587.6344	3592.5602	3593.1591	3592.3412
182	Feb 97	3552.377	3556.4222	3557.2347	3557.5165
183	Mar 97	3516.5109	3519.7294	3520.3961	3521.9661
184	Abr 97	3480.7883	3485.4158	3486.0206	3488.0344
185	May 97	3447.5605	3449.8194	3451.1022	3453.5802
186	Jun 97	3413.8771	3416.209	3417.0483	3418.1562
187	Jul 97	3380.1796	3382.0831	3382.8023	3382.0515
188	Ago 97	3346.4212	3348.8674	3347.357	3348.3798
189	Sep 97	3312.6649	3318.0875	3313.6093	3315.2351
190	Oct 97	3280.9778	3285.3899	3280.836	3280.0005
191	Nov 97	3248.4107	3252.6084	3246.2046	3247.6168
192	Dic 97	3217.8872	3218.8101	3212.5207	3215.6328

Tabla 4.6 Corridas de la Simulación de Producción de Gas Natural (1996-1997).

#### 4.2.2.1 Serie de las Ventas de Gas Natural

La simulación se muestra en la figura 4.19, donde se tienen tanto los valores simulados como los observados reales.

Modelo ARIMA(0,1,1) :  
 $y_t = y_{t-1} + e_t - 0.53973e_{t-1}$

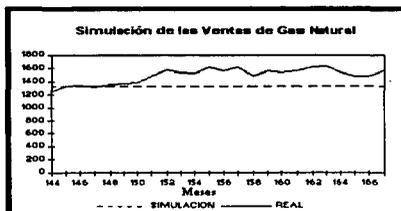


Figura 4.19 Simulación de la serie "Ventas de Gas Natural" (1994-1995).

Simulación de las Ventas de Gas Natural (MMPCD)					
Datos	Fecha	1	2	4	
169	Ene 96	1558.2815	1556.282	1555.2712	1554.5096
170	Feb 96	1558.6188	1555.7954	1556.5806	1555.8555
171	Mar 96	1557.8159	1557.1878	1557.2045	1555.2096
172	Abr 96	1557.7471	1557.6987	1556.5576	1556.3768
173	May 96	1556.1153	1557.016	1556.9886	1554.9334
174	Jun 96	1555.8771	1557.6567	1555.9536	1554.3431
175	Jul 96	1556.7249	1558.6269	1557.3403	1555.4657
176	Ago 96	1556.8018	1557.5347	1556.6323	1555.7082
177	Sep 96	1559.7414	1557.5795	1556.8034	1555.429
178	Oct 96	1557.479	1557.6748	1555.9208	1556.5786
179	Nov 96	1556.7766	1558.3451	1553.0453	1555.2427
180	Dic 96	1555.4039	1557.3385	1555.2828	1554.9502
181	Ene 97	1556.0903	1556.3008	1553.6702	1555.925
182	Feb 97	1555.8125	1558.7138	1554.1172	1555.648
183	Mar 97	1555.8037	1558.3052	1554.6252	1555.1601
184	Abr 97	1554.5431	1556.4106	1554.5172	1556.7849
185	May 97	1555.2055	1557.0701	1555.1549	1555.1148
186	Jun 97	1555.2277	1556.7747	1553.0586	1555.6601
187	Jul 97	1553.8047	1556.4796	1553.9941	1556.6452
188	Ago 97	1554.2011	1557.3569	1553.294	1555.4225
189	Sep 97	1554.1415	1557.2615	1554.2727	1556.4596
190	Oct 97	1553.5681	1556.7398	1554.4511	1553.9211
191	Nov 97	1554.139	1555.6054	1554.4711	1555.6415
192	Dic 97	1553.2454	1555.0299	1555.0756	1554.6554

Tabla 4.7 Corridas de la Simulación de Ventas de Gas Natural (1996-1997).

---

## **CAPITULO 5**

### **ANÁLISIS DE RESULTADOS**

---

## 5.1 Evaluación de Resultados

Hasta ahora se han obtenido los modelos que se ajustaron mejor a las series de tiempo.

En las tablas 5.1 a la 5.7 se encuentran los pronósticos y sus intervalos de confianza, así como las corridas de la simulación. Una vez que se tiene esta información será conveniente analizar y comparar la capacidad de pronóstico de cada modelo de las series.

Una forma de estudiar dicha capacidad radica en contrastar los valores reales con los del pronóstico. La discrepancia entre éstos y las observaciones (los errores de pronóstico) permitirá distinguir la existencia de sub o sobreestimación en los valores pronosticados, para ello se calcula la *media de los errores de pronóstico* y el *error cuadrático medio*<sup>1</sup>. Además debe estar claro que mientras más cercanas a cero sean estas medidas, mayor será la capacidad de predicción del modelo.

Una cosa importante que hay que recordar es que la validez de los pronósticos también dependerá mucho de la persona que conoce el comportamiento de las series de tiempo y que de alguna forma dirá la última palabra a cerca de lo adecuado de los valores pronosticados. Por lo tanto considero interesante presentar además de los valores obtenidos, las gráficas de los pronósticos de las series en estudio junto con los datos reales hasta agosto de 1996, y que podemos observar en las figuras 5.1 y 5.2.

### Serie de la Producción de Crudo

Los estadísticos del error reflejan un valor más pequeño en los pronósticos que en los valores simulados, mismos que no tienen mucha variación de una corrida a otra. Además podemos observar que los datos reales se encuentran dentro de los intervalos de confianza que se construyeron para el pronóstico. Por lo que podemos reafirmar que el modelo  $ARIMA(0,1,2)$  que generó estos valores pronosticados y simulados es adecuado para representar a la serie.

---

<sup>1</sup> Guerrero, Víctor M. Análisis Estadístico de Series de Tiempo Económicas. pp.220

Datos Reales (1998)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MBD		Inferior	Superior	1	2	3	4
Enero	2795.82	2764.28	2587.13	2941.44	2761.5336	2762.3563	2763.6476	2763.2057
Febrero	2844.40	2766.05	2533.31	2998.78	2760.6734	2761.9179	2762.979	2762.2636
Marzo	2868.65	2767.81	2519.75	3015.87	2761.2202	2761.935	2763.7367	2762.5487
Abril	2852.24	2769.57	2507.09	3032.06	2760.0272	2761.2048	2763.1835	2762.0313
Mayo	2874.26	2771.34	2495.17	3047.51	2760.0697	2761.0138	2764.3047	2760.7386
Junio	2879.82	2773.10	2483.91	3062.3	2760.842	2760.9164	2765.1637	2764.4271
Julio	2872.37	2774.87	2473.20	3076.53	2760.3611	2760.0987	2762.405	2763.9515
Agosto	2830.44	2776.63	2463.00	3090.27	2759.29	2760.3844	2763.9975	2763.0303
Media	Errores	81.7938			91.7479	91.0216	88.5726	89.4754
	ECM	7314.99744			9129.19606	9009.2518	8543.1234	8713.30558

Tabla 5.1 Datos Reales, Pronósticos y Simulación de la Producción de Crudo.

## Serie de las Ventas de Productos Petrolíferos

En los resultados del modelo  $ARIMA(2,1,0) \times (1,0,0)_1$  de la serie de Productos Petrolíferos, se puede observar que no hay mucha variación entre los valores de la simulación y que los valores de los estadísticos del error de los pronósticos son más pequeños.

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MBD		Inferior	Superior	1	2	3	4
Enero	1489.47	1414.70	1313.37	1512.03	1409.6849	1409.6638	1407.63	1412.4796
Febrer	1513.09	1425.54	1316.34	1534.75	1422.8105	1421.8615	1423.1772	1423.0381
Marzo	1480.94	1426.94	1307.96	1545.92	1422.3314	1422.3526	1422.5151	1425.4039
Abril	1497.02	1379.83	1246.06	1513.60	1372.7384	1374.8701	1372.4641	1376.1503
Mayo	1469.06	1397.05	1253.14	1540.96	1390.5992	1390.0827	1389.6945	1392.3799
Junio	1423.73	1390.52	1236.95	1544.09	1381.6669	1382.0861	1382.0844	1384.8047
Julio	1494.59	1387.83	1224.52	1551.14	1375.9071	1376.1427	1378.5116	1382.6162
Agosto	1435.37	1379.56	1207.45	1551.68	1368.0563	1368.2861	1368.498	1372.1758
Media	Errores	75.1625			82.4344	82.2406	82.3369	79.276875
	ECM	6338.23118			7492.38761	7443.2989	7461.1295	6957.158085

Tabla 5.2 Datos Reales, Pronósticos y Simulación de las Ventas de Productos Petrolíferos.

## Serie de las Ventas de Gas Licuado

Esta es una de las series a la que se le ajustó mejor el modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub>, ya que como observamos en la tabla 5.3 los estadísticos del error son muy cercanos a cero en los pronósticos.

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MBD		Inferior	Superior	1	2	3	4
Enero	300.23	285.657	270.606	300.708	282.9683	282.6759	286	284.8147
Febrero	284.99	278.204	262.218	294.189	268.7989	267.9977	270.1949	267.3099
Marzo	265.70	266.624	249.756	283.492	262.7078	261.6309	263.2491	263.5086
Abril	254.91	249.601	231.895	267.308	240.8422	241.1141	240.5265	240.1851
Mayo	246.41	242.974	224.467	261.480	232.8906	233.5133	232.8032	235.8425
Junio	239.81	241.955	222.721	261.269	231.7251	233.9952	232.8671	233.8404
Julio	247.57	243.111	223.099	263.123	233.3734	233.9022	232.4646	236.5502
Agosto	249.74	246.278	225.554	267.002	237.8489	240.2655	236.9473	239.0472
Media	Errores	4.3695			12.2756	11.7832	11.7884	11.032675
	ECM	41.9670785			169.7369	160.0601	157.43188	144.36531

Tabla 5.3 Datos Reales, Pronósticos y Simulación de las Ventas de Gas Licuado.

## Serie de las Ventas de Gasolinas

Esta es otra de las series de la cual se puede decir que se ajustó bien el modelo ARIMA(0,1,1)x(1,0,0)<sub>12</sub>, pues son pequeños los valores de los estadísticos de los pronósticos y más aún los de la simulación.

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MBD		Inferior	Superior	1	2	3	4
Enero	472.46	474.083	439.057	509.109	477.2708	476.5483	477.8441	478.1293
Febrero	481.61	481.415	443.587	519.243	484.4173	485.0063	483.5608	485.4074
Marzo	472.80	478.405	437.969	518.841	480.7771	482.5852	482.7665	480.6835
Abril	479.44	461.962	419.076	504.848	466.2156	464.9183	463.4147	464.024
Mayo	488.89	469.582	424.379	514.785	474.4365	474.2662	471.7234	472.2754
Junio	462.48	470.174	422.769	517.581	473.0636	472.3344	473.0211	473.0793
Julio	488.97	462.005	412.491	511.518	465.3064	464.7865	463.4145	465.4302
Agosto	476.96	467.811	416.278	519.345	469.8772	468.0699	470.5986	469.6568
Media	Errores	7.27163			4.03069	4.38686	4.65829	4.3655125
	ECM	197.797574			150.07398	163.7149	186.03668	167.77661

Tabla 5.4 Datos Reales, Pronósticos y Simulación de las Ventas de Gasolinas.

## Serie de las Ventas de Turbinas

Los estadísticos del error nos muestran que sí existe una diferencia entre los pronósticos del modelo ARIMA(0,1,1)x(0,1,1)<sub>12</sub> y las corridas de la simulación con respecto a los datos reales.

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MBD		Inferior	Superior	1	2	3	4
Enero	47.53	45.3519	39.8688	50.8350	45.912	45.7595	44.878	46.8788
Febrero	46.96	44.4777	38.3690	50.5863	42.9083	43.6722	42.2685	46.7018
Marzo	46.41	45.9882	39.3123	52.0641	43.9291	42.9929	44.3611	46.65
Abril	48.44	42.3381	35.1395	49.5367	37.1406	36.7	41.2279	41.5883
Mayo	42.92	38.9752	31.2894	46.6610	32.0829	36.259	38.338	38.9972
Junio	40.96	38.7007	30.5567	46.8446	32.7545	31.953	36.4294	36.1954
Julio	47.22	41.6831	33.1055	50.2607	34.3308	34.455	36.225	36.8385
Agosto	45.38	44.8498	35.8594	53.8402	36.1129	36.6952	40.9605	40.0684
Media	Errores	2.93191			7.58111	7.16665	5.14145	3.9877
	ECM	12.4901611			73.706136	65.914302	33.399984	27.695566

Tabla 5.5 Datos Reales, Pronósticos y Simulación de las Ventas de Turbinas.

## Serie de la Producción de Gas Natural

Aunque no tiene un buen ajuste el modelo seleccionado, debido a que como ya se había mencionado no se identificó un buen modelo del tipo ARIMA por el comportamiento que tiene la serie, se eligió el modelo ARIMA(1,0,0) con el cual se generaron pronósticos con sus intervalos de confianza, dentro de los cuales caen los datos reales.

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MMPCD		Inferior	Superior	1	2	3	4
Enero	4180.27	4081.66	3949.29	4214.03	4041.7188	4041.5759	4041.0576	4042.2529
Febrero	4270.71	4081.77	3895.49	4268.05	4001.3888	4001.1774	4002.1287	4001.923
Marzo	4256.64	4081.88	3854.86	4308.91	3961.3992	3963.3601	3965.2174	3961.9257
Abril	4284.83	4081.99	3821.13	4342.86	3923.0586	3926.1273	3927.2255	3923.144
Mayo	4328.42	4082.10	3791.86	4372.34	3883.8003	3886.6837	3889.2625	3884.8588
Junio	4334.73	4082.21	3765.81	4398.61	3845.8696	3850.0933	3850.0843	3846.4022
Julio	4204.54	4082.32	3742.22	4422.42	3806.7169	3811.5269	3813.9039	3807.4854
Agosto	3788.91	4082.43	3720.6	4444.25	3771.0144	3772.471	3776.9183	3771.0742
Media	Errores	124.086			301.76	299.504	297.906	301.247975
	ECM	42829.8738			113128.832	111412.23	110600.66	112768.688

Tabla 5.6 Datos Reales, Pronósticos y Simulación de la Producción de Gas Natural.

## Serie de las Ventas de Gas Natural

Se puede ver que los estadísticos del error no varían mucho entre las corridas de la simulación, además de que son más pequeños que los del pronóstico. A pesar de esto los datos reales se encuentran dentro de los intervalos de confianza que se obtuvieron para el pronóstico con el modelo ARIMA(0,1,1).

Datos Reales (1996)		Pronóstico	Intervalos de Confianza 95%		Corridas de la Simulación			
Mes	MMPCD		Inferior	Superior	1	2	3	4
Enero	1564.36	1545.36	1391.88	1698.85	1558.2815	1556.282	1555.2712	1554.5096
Febrero	1591.84	1545.95	1376.99	1714.92	1558.6188	1555.7954	1556.5806	1555.8555
Marzo	1585.88	1546.54	1363.41	1729.68	1557.8159	1557.1878	1557.2045	1555.2096
Abril	1571.89	1547.14	1350.85	1743.43	1557.7471	1557.6987	1556.5576	1556.3768
Mayo	1605.62	1547.73	1339.11	1756.34	1556.1153	1557.016	1556.9886	1554.9334
Junio	1703.99	1548.32	1328.07	1768.57	1555.8771	1557.6567	1555.9536	1554.3431
Julio	1621.73	1548.91	1317.61	1780.22	1556.7249	1558.6269	1557.3403	1555.4657
Agosto	1604.86	1549.50	1307.65	1791.36	1556.8018	1557.5347	1556.6323	1555.7082
Media	Errores	58.84			49.02345	49.046475	49.7051625	50.971012
	ECM	5072.37165			4131.45146	4048.32142	4141.86646	4292.9288

Tabla 5.7 Datos Reales, Pronósticos y Simulación de las Ventas de Gas Natural.

Finalmente se tiene en las figuras 5.1 y 5.2 los pronósticos y los datos reales de cada una de las series, graficadas en el periodo de Enero - Agosto de 1996. Como se observa en la figura 5.2 se tienen las series cuyos modelos generaron buenos pronósticos, es decir, se ajustaron mejor los modelos a las series de ventas que a las de producciones. Esto debido al comportamiento irregular de éstas últimas.

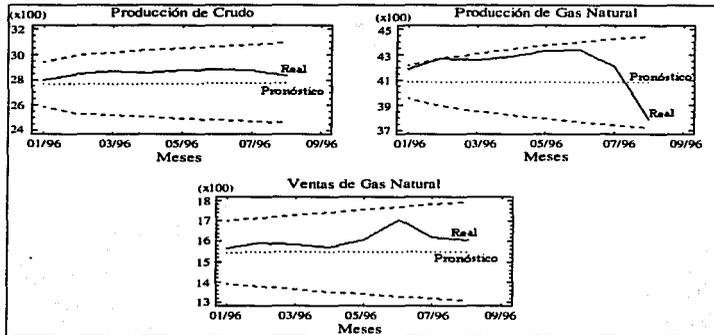


Figura 5.1 Pronósticos y Datos Reales de las Producciones y Ventas de Gas Natural en el periodo (Enero-Agosto 1996).

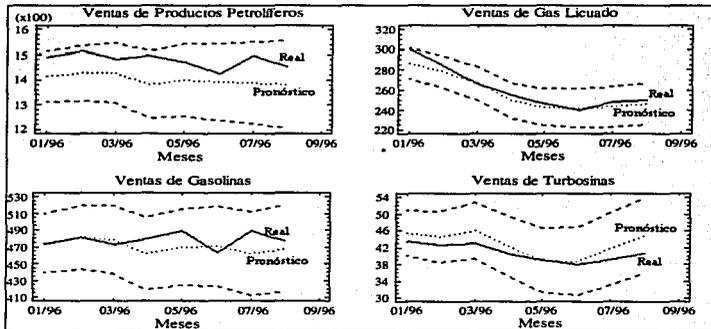


Figura 5.2 Pronósticos y Datos Reales de las Ventas de Productos Petrolíferos en el periodo (Enero-Agosto 1996).



---

## CONCLUSIONES

En general se puede concluir que se cumplió el objetivo del trabajo al obtener modelos para las series analizadas, que han servido para generar pronósticos. Además de que se aplicaron estos modelos para realizar una simulación, que dio como resultados una secuencia adecuada de valores futuros para cada una de las series.

En las series analizadas se observaron que casi todas presentaron tendencia y estacionalidad, por lo que fue necesario el uso de operadores de diferencia para eliminarlas, y así poder aplicar la Metodología de Box y Jenkins que generó modelos de series de tiempo que representan a cada una de las series, los modelos generados se evaluaron utilizando las estadísticas antes descritas, éstas junto con la opinión de aquellos que manejan las series analizadas, sirvieron para seleccionar aquellos modelos con los que se obtuvieron los pronósticos. Por lo tanto, se pudo comprobar la gran utilidad que tiene la metodología en el tipo de series que se estudiaron.

Cada uno de estos resultados se evaluaron tomando como referencia los valores observados reales para el periodo de Enero-Agosto de 1996. Concluyendo que los pronósticos resultaron ser buenos para la mayoría de las series (Ventas de Productos), mientras que los valores simulados siguieron también un comportamiento adecuado en comparación con los datos reales. Es decir, de acuerdo a los intervalos de confianza calculados para los pronósticos se observó que los valores simulados no sobrepasaron esos límites, al igual que los datos observados reales.

Es gratificante ver que este trabajo despertó el interés del personal de Petróleos Mexicanos que proporciono los datos de las series e información para el desarrollo del mismo, ya que para ellos es importante el conocer cuál será el comportamiento de las series de tiempo, pues les pueden proporcionar indicadores interesantes que les sirvan para dar opiniones o sugerencias que vayan encaminadas a contribuir al mejoramiento de políticas dentro de su área de trabajo así como tomas de decisiones que bien o mal podrían afectar en buena medida a algunos sectores que integran a la sociedad como el industrial, el de transporte, el agropecuario, el residencial, comercial y público, puesto que todos ellos dependen tanto del suministro de energía primaria como secundaria (energías producidas por los hidrocarburos) para llevar a cabo muchas de sus actividades.

Recalcamos la importancia que tienen las series analizadas, ya que del total de consumo energético de los sectores se tiene un consumo de 40.5% por parte del sector transporte, donde predomina el uso de Gasolinas y un tanto de Turbosinas. Un 34.6% se destina al sector industrial que utiliza una gran cantidad de Gas Natural y por último se tiene con un 22.5% de consumo al sector residencial, comercial y público, donde se hace uso de una gran parte de Gas Licuado.

---

---

Finalmente siempre se espera que un pronóstico sea lo más cercano a la realidad, pero en un país tan cambiante en la situación económica como lo es México, se correrá el riesgo de que lo pronosticado no sea ni siquiera parecido a lo que sucede en el comportamiento real.

Hay que mencionar que así como se pudo realizar el análisis de las series de producción y ventas de hidrocarburos, existen muchas áreas donde pueden llevarse a cabo estudios de este tipo que han demostrado ser útiles en la generación de pronósticos.

Como egresada de la preespecialidad de Simulación y Análisis de Decisiones de la carrera de Matemáticas Aplicadas, espero que los conocimientos aplicados en este trabajo para obtener los resultados que se consideraron como buenos, puedan contribuir en algo al mejoramiento de la sociedad.

## ANEXO

A continuación se presentan las tablas con los datos de las series que se analizaron en el trabajo en el periodo de 1982 a 1995. Esta información fue obtenida de la BDI (Base de Datos Institucional) que se maneja en Petróleos Mexicanos.

Datos de la Serie Original : Producción de Petróleo Crudo (MBD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	2312.65	2551.2	2542.45	2778.39	2712.39	2788.25	2786.04	2791.03	2827.74	2898.06	2937.01	3023.67
1983	2976.45	3294.36	2411.24	2667.71	2792.5	2773.51	2658.53	2731.44	2711.44	2617.43	2681.2	2640.3
1984	2628.75	2714.31	2666.53	2781.41	2800.67	2816.33	2728.24	2571.26	2599.36	2563.83	2634.19	2714.37
1985	2539.53	2581.92	2698.79	2703.52	2674	2441.44	2505.74	2679.17	2697.44	2632.27	2672.33	2734.32
1986	2510.17	2121.74	2218.09	2358.39	2527.12	2547.02	2536.35	2566.57	2373.2	2324.36	2452.34	2569.39
1987	2510.76	2541.6	2522.3	2529.17	2533.93	2530.41	2522.05	2543.68	2556.17	2552.96	2557.39	2567.58
1988	2561.35	2532.34	2517.42	2491.99	2523.79	2527.84	2528.31	2523.97	2283.75	2530.51	2524.38	2529.31
1989	2524.98	2497.02	2536.13	2520.07	2520.32	2519.32	2515.24	2516.91	2521.26	2506.47	2509.07	2468.83
1990	2517.38	2513.97	2505.16	2505.46	2479.18	2459.06	2479.34	2540.41	2620.27	2637.52	2659.1	2657.63
1991	2661.39	2674.16	2669.49	2652.92	2696.27	2717.93	2688.71	2657.68	2675.75	2683.1	2657.43	2675.12
1992	2650.18	2664.54	2680.96	2678.32	2658.52	2678.84	2681.67	2686.44	2687.1	2654.26	2638.22	2653.82
1993	2604.23	2611.45	2635.07	2673.51	2672.87	2675.08	2649.24	2649.89	2700.03	2730.87	2730.08	2745.23
1994	2724.12	2708.51	2686.79	2698.88	2688.39	2673.09	2672.04	2678.59	2667.14	2681.91	2672.54	2673.44
1995	2680.22	2647.36	2669.51	2668.43	2680.6	2649.38	2706.41	2707.69	2738.66	1898.98	2555.8	2762.52

Datos de la Serie Original : Producción de Gas Natural (MMPCD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	4101.01	4234.26	4207.39	4311.11	4273.89	4251.14	4277.19	4158.69	4252.97	4242.71	4229.96	4316.33
1983	4313.2	4025	4036.37	4086.81	4192.84	4082.27	4020.21	4070.16	4043.82	3973.78	3973.72	3881.13
1984	3851.23	3890.01	3972.26	3918.47	3864.92	3725.33	3722.33	3626.35	3569.51	3664.66	3630.35	3599.53
1985	3617.77	3607.81	3573.68	3658.04	3587.38	3508.74	3511.01	3580.36	3597.74	3631.22	3701.49	3671.05
1986	3382.27	3397.92	3517.66	3514.88	3484.78	3440.15	3379.42	3407.72	3398.93	3360.85	3393.03	3494.5
1987	3443.63	3466.22	3448.75	3481.44	3520.19	3477.57	3502.67	3478.87	3437.47	3542.82	3607.28	3571.39
1988	3579.98	3488.44	3488.66	3453.16	3494.57	3494.37	3507.61	3488.93	3369.71	3442.28	3471.16	3477.45
1989	3493.96	3458.96	3479.85	3544.03	3563.66	3597.53	3600.08	3622.74	3614.08	3618.55	3651.91	3608.13
1990	3646.92	3696.48	3664.89	3671.87	3623.68	3584.04	3542.53	3566.16	3644.31	3710.32	3751.62	3707.64
1991	3704.87	3674.03	3652.02	3595.33	3719.84	3695.27	3571.45	3601.4	3639.83	3654.75	3607.86	3479.82
1992	3541.3	3522.16	3627.91	3569.22	3538.82	3582.58	3639.8	3598.15	3660.19	3649.36	3531.82	3538.21
1993	3375.02	3531.99	3571.81	3630.89	3589.34	3582.55	3518.42	3611.38	3650.34	3554.57	3645.13	3658.27
1994	3649.13	3573.47	3609.17	3651.52	3638.15	3623.81	3640.58	3610.48	3620.07	3631.11	3615.77	3628.3
1995	3654.52	3580.3	3662.87	3668.03	3669.45	3717.37	3786.09	3872.31	3981.44	3478.34	3950.5	4081.55

**FALTA PAGINA**

No. 105a la 107

Datos de la Serie Original : Ventas de Gas Natural (MMPCD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	1426.43	1505.93	1603.83	1405.83	1466.53	1492.73	1449.83	1410.73	1400.43	1460.23	1332.03	1318.25
1983	1433.08	1468.51	1402.54	1420.06	1440.39	1701.51	1357.46	1371.38	1271.02	1310.02	1286.02	1332.02
1984	1366.15	1390.05	1365.05	1263.05	1298.22	1280.45	1293.32	1285.22	1290.75	1309.77	1362.71	1298.05
1985	1370.77	1389.02	1255.91	1313.28	1330.31	1324.3	1318.15	1296.08	1266.25	1311.29	1204	1189.99
1986	1095.02	1264.02	1210.04	1212.71	1154.36	1159.88	1140.01	1143.75	1181.73	1185.86	1185.16	1124.82
1987	1246.03	1238.43	1191.07	1180.94	1181.87	1197.54	1122.63	1064.26	1137.43	1165.66	1198.38	1200.51
1988	1141.43	1216.85	1194.17	1160.16	1129.26	1126.44	1096.08	1131.56	1161.61	1106.51	1165.93	1089.71
1989	1302.67	1281.46	1232.71	1263.52	1263.83	1156.97	1119.88	1160.49	1103.72	1024.99	1180.79	1228.18
1990	1334.19	1345.39	1409.8	1321.53	1328.75	1303.9	1277.39	1293.54	1297.38	1374.95	1350.21	1478.21
1991	1479.11	1420.8	1393.1	1454.04	1482.59	1466.3	1581.53	1384.18	1412.2	1534.17	1474.25	1506.36
1992	1576.56	1453.37	1252.2	1560.79	1209.8	1506.42	1407.73	1351.17	1563.73	1604.31	1536.5	1506.05
1993	1431.92	1493.72	1472.03	1397.02	1424.41	1494.38	1387.26	1317.75	1277.53	1303.87	1322.6	1246.58
1994	1324.74	1345.34	1319.19	1356.14	1365.53	1385.11	1481.82	1580.91	1537.92	1524.75	1614.03	1563.18
1995	1619.61	1477.42	1562.95	1531.51	1557.43	1619.5	1626.8	1540.31	1474.89	1479.96	1567.04	1555.78

Datos de la Serie Original : Ventas de Productos Petroliferos (MBD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	994.73	1099.47	1133.44	1097.36	1126.33	1089.75	1106.21	1054.17	1033.3	1056.85	1136.91	1038.12
1983	1017.81	1067.75	1084.67	1006.11	1029.59	1042.32	1002.38	1040.51	986.62	992.81	1045.23	1048.44
1984	1047.07	1114.86	1115.78	1097.66	1073.99	1055.91	1054.83	1104.45	989.22	1163.54	1175.51	1178.02
1985	1100.9	1140.82	1126.8	1182.72	1188.36	1083	1182.16	1105.73	1072.4	1140.42	1150.23	1178.3
1986	1176.59	1199.29	1111.99	1184.12	1162.29	1089.53	1129.77	1043.28	1083.9	1130.04	1053.01	1162.4
1987	1193.55	1175.36	1202.52	1185.24	1144.22	1210.48	1165.47	1138.68	1178.16	1195.36	1188.98	1258.56
1988	1119.83	1224.64	1221.61	1161.71	1181.82	1197.92	1140.45	1212.52	1119.81	1183.53	1230.56	1276.81
1989	1208.39	1303.77	1261.71	1244.54	1314.83	1276.06	1274.13	1280.44	1244.73	1262.51	1323.43	1298.13
1990	1397.06	1367.27	1330.27	1301	1365.33	1330.18	1304.75	1372.02	1243.62	1395.37	1376.03	1369.41
1991	1438.42	1414.75	1320.52	1433.97	1447.79	1373.78	1385.85	1385.6	1347.91	1472.03	1397.62	1435.27
1992	1391.77	1403.69	1438.44	1425.65	1378.03	1458.2	1489.4	1398.43	1427.06	1454.81	1448.3	1458.81
1993	1374.96	1516.15	1513.45	1454.8	1431.78	1483.54	1430.4	1398.31	1398.31	1375.8	1449.44	1526.67
1994	1485.34	1546.39	1584.31	1544.87	1550.34	1524.69	1473.45	1538.11	1538.45	1555.5	1630.97	1609.26
1995	1575.43	1555.04	1546.62	1393.25	1441	1411.59	1399.97	1363.71	1307.61	1312.9	1446.61	1462.81

Datos de la Serie Original : Ventas de Gas Licuado (MBD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	123.52	138.63	138.14	124.89	122.14	126.61	128.28	127.26	135.51	132.7	147.87	139.68
1983	142.33	163.05	144.61	134.8	130.45	132.99	130.12	142.82	139.82	145.36	164.27	157.3
1984	164.48	169.62	161.43	155.49	158.74	159.02	168.94	177.54	165.59	185.93	179.07	167.76
1985	185.13	173.1	173.92	183.5	178.68	185.15	184.94	178.53	179.97	187.37	187	187.78
1986	196.4	205.25	172.18	186.99	173.86	169.06	169.56	160.09	169	174.29	179.3	184.98
1987	193.85	194.36	183.32	173.37	171.01	170.58	167.33	163.67	168.83	181.41	170.48	189.16
1988	202.43	191.91	184.24	173.51	165.68	171.16	163.5	173.33	177.65	179.85	187.83	193.96
1989	205.51	195.17	187.12	181.28	180.17	173.07	175.98	181.73	184.13	191.81	199.37	201.14
1990	232.97	211.51	201.42	185.74	188.31	181.74	183.79	194.7	188.81	200.31	211.08	215.82
1991	234.91	224.67	202.78	207.9	205.58	190.66	202.88	205.81	204.01	227.74	235.87	239.99
1992	255.74	256	240.59	226.8	223.1	224.01	225.07	225.76	228.12	236.88	240.56	251.43
1993	267.37	271.09	257.51	236.6	234.83	230.12	232.25	233.64	238.17	257.2	258.2	268.9
1994	273.41	276.07	257.03	247.73	239.28	241.9	240.22	244.25	244.25	263.52	266.96	281.4
1995	286.54	271.29	265.24	243.41	235.99	235.33	237.25	240.81	249.3	246.58	274.97	278.21

Datos de la Serie Original : Ventas de Gasolinas (MBD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	330.65	380.87	381.86	379.56	362.96	375.72	375.39	351.07	347.94	341.86	376.35	347.74
1983	301.59	331.61	341.95	307.96	312.68	321.97	301.15	323.72	316.39	301.99	308.19	332.41
1984	303.79	323.33	323.33	327.02	305.2	302.45	301.43	313.6	287.4	349.16	363.17	380.11
1985	295.81	314.34	312.53	324.9	328.64	309.89	335.71	327.94	314.95	335.39	348.69	337.86
1986	329.63	331.77	326.68	335.76	339.8	325.14	344.84	314.56	330.42	334.97	317.73	360.82
1987	332.38	343.01	352.57	336.95	333.86	356.45	352.81	333.11	348.82	366.53	353.42	364.17
1988	318.29	351.53	358.37	345.65	358.72	365.34	350.74	372.62	362.2	360.19	386.5	403.7
1989	375.7	388.23	402.05	382.34	410.99	415.56	392.33	414.4	407.34	411.87	424.88	440.59
1990	423.68	429.35	441.04	434.45	446.98	446.52	435.66	453.22	423.03	485.93	465.23	479.16
1991	456.13	468.1	465.38	483.2	493.87	462.15	487.15	489.37	468.96	498.54	478.99	503.65
1992	466.07	478.41	482.27	488.87	476.81	501.88	492.74	467.29	479.49	498.84	478.19	504.72
1993	437.2	514.31	507.81	499.65	481.93	494.68	487.34	478.87	481.12	477.67	487.69	537.43
1994	473.48	502.44	520.07	487.41	500.05	502.78	484.34	509.67	504.69	493.98	519.75	541.61
1995	493.1	510.95	502.56	460.09	478.67	479.42	457.94	471.92	463.85	468.36	486.31	497.05

Datos de la Serie Original : Ventas de Turbosinas (MBD)												
	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1982	29.9	30.62	36.07	30.04	27.79	28.04	28.37	31.32	28.1	24.1	16.74	28.13
1983	26.86	27.87	29.93	27.46	25.11	25.81	27.3	31.17	25.39	26.46	27.76	29.58
1984	30.23	21.79	30.82	35.06	27.18	26.26	27.74	28.77	26.53	29.53	26.07	28.36
1985	32.66	31.38	30.52	30.51	28.47	28.29	30.37	32.11	28.27	27.71	27.12	30.19
1986	31.68	30.16	30.11	28.56	26.7	27.53	28.05	28.03	26.34	26.42	27.97	31.26
1987	30.87	31.42	29.67	30.49	30.35	27.37	28.59	29.07	27.2	28.95	28.18	31.8
1988	30.37	31.76	33.69	23.95	23.09	23.11	23.5	24.47	23.21	24.76	26.55	28.83
1989	31.68	32.74	33.49	29.62	30.37	30.81	28.87	33.46	28.48	30.85	33.64	33.3
1990	40.18	38.64	36.64	33.8	31.65	30.47	31.92	36.96	35.42	36.25	36.05	38.15
1991	39.27	37.51	34.19	36.24	34.28	32.88	40.24	38.99	36.78	37.26	35.69	40.22
1992	42.34	40.8	38.83	41.27	38.99	39.27	46.01	44.89	42.37	37.78	41.93	38.63
1993	40.45	45.02	47.06	44.32	38.69	41.22	44.75	45.38	39.98	37.93	43.66	44.06
1994	47.95	46.93	50.98	46.65	42.98	45.07	48.26	53.84	45.31	45.09	48.75	49.9
1995	51.13	48.86	49.85	45.15	41.84	39.86	41.92	45.38	38.76	38.64	41.87	43.77

### Módulo para la Simulación

Como se mencionó en el capítulo IV, para llevar a cabo la simulación se diseñó y programó un pequeño módulo llamado GRAFSIMU para simular los datos con los modelos que se seleccionaron de cada una de las 7 series analizadas. En la figura 1 se tiene la pantalla principal del módulo donde se introduce la serie que se va a simular, el tipo de modelo, las diferencias que se hicieron, los datos a simular y el origen de la serie.

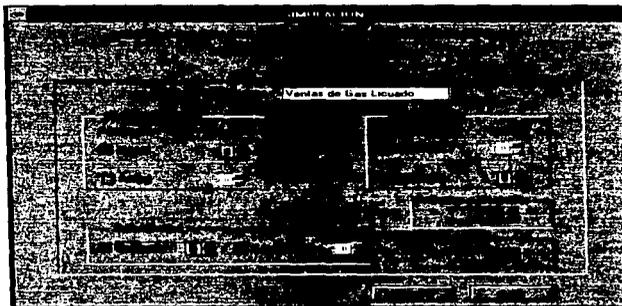


Figura 1. Pantalla Principal.

Hay que mencionar que con este módulo no sólo se pueden generar valores simulados para los modelos que se identificaron, sino también se pueden manejar diferentes modelos ajustándolos a cada serie para poder observar que comportamiento se tendría si se modifican los valores de los parámetros.

Este módulo fue desarrollado en Fox-Pro para Windows version 2.5.

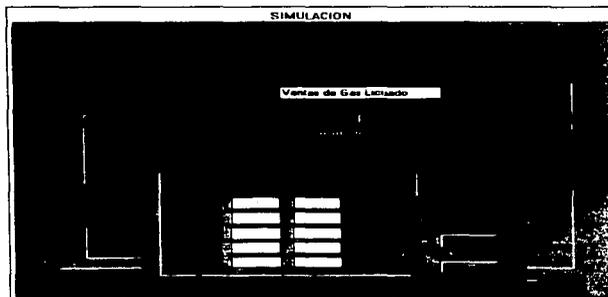


Figura 2 Pantalla donde se capturan el valor de los parámetros.

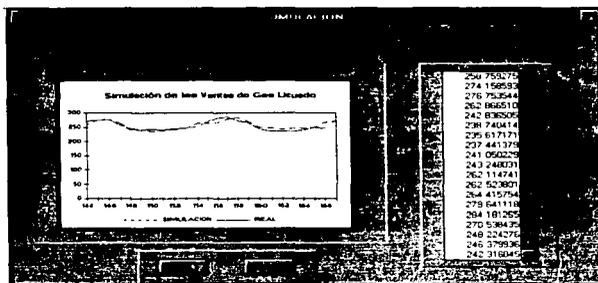


Figura 3 Gráfica y Datos simulados.

En la figura 2 se capturan los valores de los parámetros. En la figura 3 se muestra las salidas con la gráfica de la simulación, además de los datos de la misma.



---

## **BIBLIOGRAFIA**

- **Box and Jenkins**  
**TIME SERIES ANALYSIS. Forecasting and Control.**  
Holden-Day, 1976
  - **Bowerman & O'Connel**  
**Time Series and Forecasting**  
Duxbury Press, 1979
  - **Wheelwright & Makridakis**  
**Forecasting Methods for Management**  
Wiley & Sons, 1985
  - **González Videgaray, María del Carmen**  
**Modelos de Decisión con Procesos Estocásticos II**  
México. ENEP Acatlán , 1990
  - **Guerrero, Víctor M.**  
**Análisis Estadístico de Series de Tiempo Económicas**  
Universidad Autónoma Metropolitana, 1991
  - **Naylor, Thomas H.**  
**Experimentos de Simulación en Computadores con Modelos de Sistemas Económicos**  
Limusa, 1982
  - **Coss Bú, Raúl**  
**Simulación, un Enfoque Práctico**  
Limusa, 1989
  - **Petróleos Mexicanos**  
**Memoria de Labores 1995**  
México, 1996
  - **Pemex Exploración Producción**  
**"La Producción de Gas en México", Evolución. Situación Actual y Perspectivas**  
Informe, 1994
-

- **Pemex Exploración Producción**  
**"La Producción de Crudo en México", su Evolución, Situación Actual y Perspectivas**  
**Informe, 1994**
- **Secretaría de Energía**  
**Balance Nacional. ENERGIA 1994**  
**Subsecretaría de Política y Desarrollo de Energeticos, a cargo de la Dirección General de**  
**Política Energética, y de la Unidad de Comunicación Social, 1995**