

24021
4



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS
PROFESIONALES "ACATLAN"

"ANALISIS DE INTERVENCION: APLICACION A UN PROBLEMA AMBIENTAL"

T E S I S

QUE PARA OBTENER EL TITULO DE
LICENCIADO EN MATEMATICAS APLICADAS Y
COMPUTACION

PRESENTA
GERARDO CEJA PEÑA

ASESOR: MARIA DEL CARMEN GONZALEZ VIDEGARAY



FEBRERO, 2003

TESIS CON
FALLA DE ORIGEN

1



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.

A mis Padres y Hermanas

Todo verdadero regalo es recíproco. Dios, de Quien recibimos el mundo, recibe de sus criaturas el mundo. ¿Qué es una dedicatoria, qué es esta página? No es el don de esa cosa entre las cosas, un libro, ni de los caracteres que lo componen; es, de algún modo mágico, el don del inaccesible tiempo en que se escribió y, lo que sin duda no es menos íntimo, del mañana y del hoy. Sólo podemos dar el amor, del cual las cosas son símbolos.

Tuyo es el libro. ¿A qué agregar vanas y laboriosas palabras a lo que sentimos los dos?

Jorge Luis Borges

Buenos Aires 13/Jun/1968

Autorizo a la Dirección General de Bibliotecas de UNAM a difundir en formato electrónico e imp. contenido de mi trabajo recapitulativo.

NOMBRE: Cecilia Ceyra

FECHA: 27 de Febrero 2003

FIRMA: [Firma]

TESIS CON
FALLA DE ORIGEN

A todos mis amigos,

en especial a:

Adolfo, Alejandro, Amelia, Ariel,
Carlos, Claudia, Donovan, Fabiola,
Fanny, MariCarmen, Mónica, Oscar,
Ricardo, Rodrigo, Rubén, Tania, Víctor

Sólo una cosa no hay. Es el olvido.
Dios, que salva el metal, salva la escoria
y cifra en Su profética memoria
las que serán y las que han sido.
Ya todo está. Los miles de reflejos
que entre los dos crepúsculos del día
tu rostro fue dejando en los espejos
y los que irá dejando todavía.
Y todo es una parte del diverso
cristal de esa memoria, el universo;
no tiene fin sus arduos corredores
y las puertas se cierran a tu paso;
sólo del otro lado del ocaso
verás los Arquetipos y Esplendores

Everness

Jorge Luis Borges

TESIS CON
FALLA DE ORIGEN

ÍNDICE

Introducción	1
I. Series de Tiempo	5
1.1 Las series de tiempo y sus principales características	6
1.2 Metodología Box-Jenkins para series de Tiempo	10
1.3 Modelos de Series de Tiempo	15
1.3.1 Modelos Estacionarios	15
1.3.1.1 Modelos Autorregresivos AR(p) y de Medias Móviles MA(q)	15
1.3.1.2 Dualidad entre los procesos AR(p) y MA(q)	23
1.3.1.3 Modelos Mixto General ARMA(p,q)	26
1.3.2 Modelos No Estacionarios	28
1.3.2.1 No Estacionaridad en la media y varianza	28
1.3.2.2 Modelo General ARIMA (p,d,q)	29
1.3.3 Modelos Estacionales	30
1.3.3.1 Modelos Estacionales Puros	30
1.3.3.2 Modelo General Multiplicativo Estacional ARIMA(p,d,q)X(P,D,Q)s	31
1.4 Estimación y Selección de los Modelos	34
1.4.1 ¿Cómo se estima el valor de los parámetros de un modelo?	34
1.4.2 Ajuste de los Modelos	40
1.4.2.1 Análisis de Estacionaridad y Residuales	40
1.4.2.2 Sobre y Subespecificación de modelos	42
1.4.3 Criterio de Selección del modelo	43
1.4.4 Pronósticos a partir de un modelo seleccionado	46
II. El análisis de intervención y su aplicación en las series de tiempo	49
2.1 ¿Qué es el Análisis de Intervención?	50
2.2 Tipos de Intervenciones en las series de tiempo	52
2.3 Posibles respuestas a una intervención	53

Índice

2.4 Antecedentes del tema y su importancia; estudios relevantes	59
III. El problema ambiental	65
3.1 Planteamiento del problema	66
3.1.1 Límites del estudio	66
3.1.2 Condiciones actuales de la ZMVM	66
3.1.3 El Programa "Hoy No Circula"	68
3.1.4 Índices de Ozono en la ZMVM	69
3.2 Datos históricos del problema	73
3.2.1 La RAMA y su importancia	73
3.2.2 El PSI	75
3.2.3 El IMECA	76
3.3 Obtención y solución del modelo matemático	81
IV. Resultados	97
4.1 Interpretación de los resultados del modelo matemático	98
4.2 Aportación a la disciplina y posibles trabajos posteriores	99
Conclusiones	101
Referencias bibliográficas	103
Anexo I	105

TESIS CON
FALLA DE ORIGEN

INTRODUCCIÓN

En la actualidad existen muchos fenómenos de la naturaleza que no se pueden predecir. Algunos por la complejidad de su estudio, otros porque no se tiene información que se pueda analizar y realizar pruebas. Un ejemplo de este tipo es la contaminación atmosférica. Es conocido que una de las ciudades con los mayores índices de contaminación atmosférica es la Ciudad de México y su Zona Metropolitana. Este es un problema que se ha ido gestando durante décadas. Anteriormente no se registraban los niveles de contaminación y por tanto no se podían realizar evaluaciones. Obviamente, tampoco se tenían normas de calidad de aire que regularan los valores máximos permisibles de concentración de contaminantes. No existían los recursos ni la infraestructura para realizar estudios epidemiológicos, toxicológicos y de exposición, ni en animales ni en seres humanos, por lo que las normas se establecieron fundamentalmente tomando en cuenta los criterios y estándares adoptados en otros países del mundo¹.

En la Zona Metropolitana del Valle de México (ZMVM) se han realizado varios esfuerzos para medir la concentración de los contaminantes aéreos. En 1972 surgió la Subsecretaría de Mejoramiento del Ambiente (SMA), de la entonces Secretaría de Salubridad y Asistencia, que inició los trabajos tendientes a normar los contaminantes atmosféricos y propuso lo que se denominó el Índice Mexicano de la Calidad del Aire (IMEXCA), sumando los diversos contaminantes en una ecuación que no era fácil de desarrollar. Para 1982, surgió la Secretaría de Desarrollo Urbano y Ecología (SEDUE), originándose el Instituto Nacional de Ecología (INE) en su interior. De estas instancias gubernamentales surgió el Índice Metropolitano de la Calidad del Aire (IMECA). Se empezó

¹ Las normas vigentes de calidad del aire fueron publicadas por la Secretaría de Salud en el Diario Oficial de la Federación en diciembre de 1994.

Análisis de Intervención: Aplicación a un problema ambiental.

a seguir el criterio de expresar un solo contaminante como guía o centinela, destacando el ozono, ya que las concentraciones de este contaminante eran muy altas.

En el INE-SEDUE se dio prioridad a las condiciones ambientales, relegándose los estudios sobre la salud humana durante un buen tiempo. Sin embargo, se hicieron algunos ajustes en las concentraciones de contaminantes y surgió la necesidad de tener una instancia gubernamental diferente para vigilar el cumplimiento de las normas mexicanas (NOM) para la calidad del aire, creándose así la Procuraduría Federal de Protección al Ambiente (PROFEPA), dentro de la Secretaría de Desarrollo Social (SEDESOL).

La SMA instaló una red para la cuantificación de contaminantes atmosféricos, constituida por 16 casetas o monitores automáticos y las consecutivas estaciones con procesos manuales. A la primera se le denominó Red Automática de Monitoreo Atmosférico (RAMA), y a la segunda se le conoció como Red Manual, nombres que persisten hasta el momento actual.

La SEDUE reemplazó los medidores por equipo más moderno y eficaz, ampliando la red y mejorándola, quedando bajo la administración del INE. Poco después, ambas redes, la manual y la automática, quedaron bajo la administración de la Comisión Metropolitana para la Prevención y Control de la Contaminación Ambiental (CMPICA), que al principio dependía del Departamento del Distrito Federal.

En la actualidad, la RAMA tiene más de 32 estaciones, distribuidas a lo largo y ancho de la ZMVM. De cada uno de los aparatos de medición automática continua, se envían datos por vía telefónica preferencial codificada a un centro de acopio de información; de aquí se envía al Sistema de Información del Medio Ambiente (SIMA), que difunde la información a diversos lugares considerados como prioritarios, ya sea cada una, ocho o 24 horas, por lo general utilizando medios electrónicos de comunicación masiva.

Se miden los cinco contaminantes atmosféricos centinela: ozono (O₃), óxido de nitrógeno (NO), óxidos de nitrógeno (NO_x), bióxido de azufre (SO₂) y monóxido de carbono

TESIS CON
FALLA DE ORIGEN

(CO), junto con algunas condiciones ambientales y meteorológicas (humedad relativa, temperatura, velocidad y dirección del viento) y partículas menores de diez micras (PM10).

Las principales estaciones de monitoreo atmosférico, o estaciones centinela, se encuentran en cinco puntos clave de la ZMVM; dos en el norte (Xalostoc y Tlalnepantla), una en el centro (Merced), y dos en el sur (Pedregal y Cerro de la Estrella).

El gobierno federal, así como el local, han implementado medidas con el propósito de disminuir estos índices de forma gradual y permanente. Una de las primeras acciones fue la aplicación del programa "Hoy No Circula". El Hoy No Circula fue concebido como una medida regulatoria para controlar la demanda de gasolinas y reducir las emisiones vehiculares. Fue presentado como un programa temporal buscando que la gente se desplazara en transporte público o compartiendo el auto con vecinos o compañeros de trabajo.

Posteriormente, en 1990, se implementó el "Programa Integral Contra la Contaminación Atmosférica en la Zona Metropolitana del Valle de México (PICCA)". El Departamento del Distrito Federal, Petróleos Mexicanos, el Instituto Mexicano del Petróleo, el gobierno del Estado de México y la Secretaría de Desarrollo Urbano y Ecología, reunieron una serie de medidas para ser aplicadas tanto en el transporte como en la industria, en los servicios y en las actividades de producción de energéticos. El PICCA incluyó la continuación del Programa Hoy No Circula. De esta manera, el programa se hizo permanente.

Actualmente, existe el "Programa para Mejorar la Calidad del Aire en el Valle de México". Este programa, principalmente, ha dado continuidad a las medidas anteriormente tomadas e incluye otras como: el mejoramiento e incorporación de nuevas tecnologías en industria, servicios, etc.

Análisis de Intervención: Aplicación a un problema ambiental.

Durante estos programas, el "Hoy No Circula" ha estado presente. El propósito de este estudio es analizar que impacto ha tenido el programa "Hoy no circula" en la reducción de los principales índices de contaminación, en particular el Ozono (O_3). Para ello, se realizará un modelo de series de tiempo utilizando datos registrados por la RAMA en el periodo 1986-2000. Una vez hecho esto, se aplicará la metodología del "análisis de intervención" para conocer dicho impacto. Se describirán algunos casos, en los cuales está metodología ya ha sido aprovechada. Al mismo tiempo, se dará una orientación sobre los paquetes de computación que se utilizarán para los distintos cálculos.

TESIS CON
FALLA DE ORIGEN

CAPÍTULO I

SERIES DE TIEMPO

En este capítulo se definen las series de tiempo y se describen sus principales características. Se explica la metodología Box-Jenkins y las condiciones que debe cumplir una serie para poder utilizar esta metodología. Se detallan los principales modelos matemáticos, desde el más simple hasta el modelo general que puede ser útil para representar diferentes series de tiempo. Algunas veces, una serie de tiempo puede tener más de un modelo tentativo que represente sus características; por ello se dan criterios de selección de modelos y los métodos para poder estimar sus parámetros.

TESIS CON
FALLA DE ORIGEN

I. Series de Tiempo

1.1 Las series de tiempo y sus principales características

Este trabajo está relacionado con consecuencias dinámicas de eventos sobre el tiempo. La variable que se estudia puede ser denotada por Y . Cuando se estudia esta variable en el tiempo, su valor está denotado por Y_t donde t indica el tiempo. Por ejemplo, una ecuación dinámica da un valor Y en el instante t , a partir de otra variable W_t y del valor Y en un periodo previo.

$$Y_t = \phi Y_{t-1} + W_t \quad (1.1.1)$$

La ecuación (1.1.1) es una ecuación lineal de primer orden. Esta ecuación es una expresión que relaciona una variable Y_t a sus valores previos. W_t puede ser vista como una variable aleatoria o bien como un número determinado. La suposición es que una ecuación dinámica gobierna el comportamiento de y para todos los momentos t . Por lo tanto, para cada instante se tiene una ecuación relacionando el valor de y del instante previo y valor actual de w :

Tiempo	Ecuación
0	$Y_0 = \phi Y_{-1} + W_0$
1	$Y_1 = \phi Y_0 + W_1$
2	$Y_2 = \phi Y_1 + W_2$
:	:
T	$Y_t = \phi Y_{t-1} + W_t$

Si se conoce el valor inicial de y en el instante $t = -1$ y el valor de w para $t=0,1,2,\dots$, entonces es posible simular este sistema dinámico para encontrar el valor de y para cualquier instante. Por ejemplo, si se conoce el valor de y para $t = -1$ y el valor de w

TESIS CON
FALLA DE ORIGEN

1.1 Las series de tiempo y sus principales características

para $t=0$, podemos calcular y para $t=0$. Obtenido este valor de Y_0 y el valor de w para $t=1$, podemos calcular el valor de y para $t=1$:

$$Y_1 = \phi Y_0 + W_1 = \phi(\phi Y_{t-1} + W_0) + W_1,$$

O bien

$$Y_1 = \phi^2 Y_{-1} + \phi W_0 + W_1$$

Obtenido el valor de Y_1 y el valor de W para $t = 2$, podemos calcular de la misma manera el valor de y para $t = 2$. De esta manera continuando recursivamente, el valor que y toma en el tiempo t puede ser descrito como una función de su valor inicial Y_{-1} y la historia (datos históricos) de w entre el tiempo 0 y el tiempo t :

$$Y_t = \phi^{t+1} Y_{-1} + \phi^t W_0 + \phi^{t-1} W_1 + \phi^{t-2} W_2 + \dots + \phi W_{t-1} + W_t \quad (1.1.2)$$

Este procedimiento es conocido como método de sustitución recursiva.

La ecuación (1.1.2) expresa Y_t como una función lineal del valor inicial Y_{-1} y de valores históricos de w . Esto hace fácil calcular el efecto de W_0 sobre Y_t . Si W_0 cambiara con Y_{-1} y $W_1, W_2, W_3, \dots, W_t$ no se vieran afectados, el efecto sobre Y_t , estaría dado por:

$$\frac{\partial Y_t}{\partial W_0} = \phi^t$$

Se puede notar que los cálculos serían los mismos si la simulación dinámica comenzara en el instante t (tomando Y_{t-1} como dado); entonces Y_{t+j} podría ser descrita como una función de Y_{t-1} y W_{t+1}, \dots, W_{t+j}

$$Y_{t+j} = \phi^{j+1} Y_{t-1} + \phi^j W_t + \phi^{j-1} W_{t+1} + \phi^{j-2} W_{t+2} + \dots + \phi W_{t+j-1} + W_{t+j}$$

El efecto de W_t sobre Y_{t+j} está dado por:

$$\frac{\partial Y_{t+j}}{\partial W_t} = \phi^j$$

TESIS CON
FALLA DE ORIGEN

Análisis de Intervención: Aplicación a un problema ambiental.

A ϕ_j se le llama multiplicador dinámico y en este caso depende sólo de j , de la longitud de separación la entrada (W_t) y del valor observado de la salida (Y_{t+j}). El multiplicador no depende de t ; es decir, no depende de los instantes de las observaciones. Esto es cierto en cualquier modelo de ecuación lineal.

Una **serie de tiempo** es una colección de observaciones indexadas por la fecha (momento) de cada observación². Usualmente se tiene registrada información de datos iniciales en algún periodo en particular (esto es, $t=1$) y datos finales en otro ($t=T$):

$$(Y_1, Y_2, \dots, Y_T) \quad (1.1.3)$$

La muestra de observaciones (1.1.3) puede ser vista como un segmento finito de una secuencia doblemente infinita, denotada como:

$$\{Y_t\}_{t=-\infty}^{\infty} :$$

$$\{Y_t\}_{t=-\infty}^{\infty} = \{ \dots, Y_{-1}, Y_0, \underbrace{Y_1, Y_2, \dots, Y_T}_{\text{Muestra observada}}, Y_{T+1}, Y_{T+2}, \dots \}$$

Una serie de tiempo puede tener uno o más de los siguientes componentes:

- Tendencia: creciente o decreciente
- Variación Estacional: que cada cierto periodo de tiempo, menor a un año, se vuelva repetitiva la serie.
- Ciclo: los datos son repetitivos, pero el periodo de tiempo es mayor a un año
- Fluctuaciones Aleatorias (Random shocks): observaciones que son inesperadas.

² Los intervalos de tiempo entre las observaciones tienen que ser iguales. Algunas veces es necesario corregir los datos originales (Por ejemplo, años de 365 días)

1.1 Las series de tiempo y sus principales características

Las series de tiempo pueden ser usadas para alcanzar diferentes metas, por ejemplo:

- a) **Análisis descriptivo:** Consiste en determinar las tendencias (crecientes o decrecientes) y/o patrones (efectos estacionales) de una serie de tiempo usando gráficas o algunas otras técnicas. Asimismo se pueden identificar posibles valores erróneos.
- b) **Pronóstico:** Si una serie de tiempo tiene cierto comportamiento (el pasado), se pueden estimar valores futuros con el fin de tomar decisiones.
- c) **Análisis de Intervención (Control):** Si se desea modificar el comportamiento del proceso y/o examinar sus consecuencias.

Tomando en cuenta sus componentes principales, las series de tiempo puede clasificarse en:

a) Discretas o Continuas

Discretas: las observaciones se realizan en ciertos tiempos, por ejemplo, la cantidad de bacterias contenidas en un recipiente medida cada 6 horas.

Continuas: el tiempo en que se realizan las observaciones es continuo, por ejemplo; nivel del mar medido con un sensor automático.

b) Estacionarias y No Estacionarias

Estacionarias: Los datos fluctúan o varían alrededor de un valor constante.

No Estacionarios: La longitud o amplitud de la serie cambia sobre el tiempo.

Las series de tiempo con tendencia son no estacionarias.

c) Determinísticas o Estocásticas:

Determinísticas: Se puede predecir los datos pueden con exactitud.

Estocásticas: Los valores futuros de la serie tienen que ser descritos con una distribución de probabilidad.

1.2 Metodología Box-Jenkins para Series de Tiempo

El método de Box-Jenkins consiste en extraer movimientos predecibles de los datos históricos. La idea principal es elegir un modelo sencillo e ir mejorándolo progresivamente, aplicando conocimiento y experiencia sobre el fenómeno en cuestión. Por tanto, este es un proceso iterativo que construye modelos lineales, y consiste en cuatro pasos fundamentales:

- 1) Identificar características de la serie y especificaciones preliminares para el modelo. Generalmente, el primer paso es graficar las observaciones contra el tiempo y notar las características principales tales como: tendencia, varianza no constante (heteroscedasticidad), estacionalidad, etc.
- 2) Plantear un modelo y estimar sus parámetros.
- 3) Evaluación del modelo.
- 4) Pronóstico de valores futuros

Para poder utilizar la metodología Box-Jenkins en forma adecuada, la serie de tiempo estudiada debe tener ciertas características:

- a) **Estacionaridad**
- b) **No debe tener fluctuaciones estacionales (o periódicas)**
- c) **Continuas y sin datos discrepantes**

Algunas veces no se cumplen; para tales casos, se muestran métodos con los cuales es posible tenerlas.

a) Estacionaridad

Para que una serie sea estacionaria debe cumplir con las siguientes condiciones:

➤ **Media Constante.**

La *tendencia* es una característica que puede presentarse en las series de tiempo. Si no hay tendencia se puede asumir que la media es constante y que el valor observado para cada periodo puede representarse por esta media, es decir:

$$E(Y_t) = E(Y_{t+n}) = \mu \quad (1.2.1)$$

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t \quad \text{media estimada} \quad (1.2.2)$$

En caso necesario, existen varios métodos para eliminar la tendencia. Si la tendencia es fija y determinística se puede utilizar un modelo de regresión. La forma más usual de eliminar la tendencia es utilizar diferencias ya que es muy difícil saber si la tendencia es determinística o estocástica. El método de diferencias consiste en restar los valores de las observaciones en un orden establecido (Ver expresión 1.2.3). En general, si la tendencia es lineal (polinomio de grado uno), la tendencia desaparece con una diferencia; si la tendencia indica un polinomio de grado dos desaparecerá al aplicar dos diferencias y así sucesivamente. Es difícil que aparezcan series que requieran más de dos diferencias para eliminar su tendencia.

$$\begin{aligned} \Delta Y_t &= Y_t - Y_{t-1} \\ \Delta^2 Y_t &= \Delta Y_t - \Delta Y_{t-1} \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned} \quad (1.2.3)$$

Una vez realizadas las diferencias, en caso de que sea necesario, se obtiene una nueva serie de datos que puede ser analizada con mayor facilidad, por ejemplo:

$$Z_t = \Delta Y_t = Y_t - Y_{t-1}$$

➤ **Varianza constante**

$$\text{Var}(Y_t) = \text{Var}(Y_{t+n}) = \sigma^2_Y \quad (1.2.4)$$

Algunas veces se tiene que inducir una varianza constante ya que la serie no cumple con esta condición. Para realizar esto se ocupan diferentes transformaciones, de tal forma que las curvas tiendan a rectas y al mismo tiempo hacer la varianza constante. El cuadro 1.2.1 remarca las principales transformaciones usadas:

Transformación	Cuando se puede usar
$\ln(Z_t)$	Si la desviación estándar de la serie es proporcional al nivel medio
$\sqrt{Z_t}$	Si la varianza de la serie es proporcional al nivel medio
$\frac{1}{Z_t}$	Si la desviación estándar de la serie es proporcional al cuadrado del nivel medio

Cuadro 1.2.1 Principales transformaciones usadas para estabilizar la varianza

De forma más general, para estabilizar la varianza, se puede usar la transformación introducida por Box y Cox (1964):

$$T(Z_t) = Z_t^{(\lambda)} = \frac{Z_t^\lambda - 1}{\lambda} \quad (1.2.5)$$

λ es llamado el parámetro de transformación. Algunos valores usados comúnmente y sus transformaciones asociadas se muestran en el cuadro 1.2.2.

Se debe tomar en cuenta que las transformaciones enunciadas están definidas, en su mayoría, para series de valores positivos; si existen valores negativos se puede sumar una constante a cada valor de la serie sin afectar la estructura de correlación de la serie. Además, en caso de que sea necesario aplicar una transformación a la serie, se debe de realizar antes de efectuar diferencias. Este proceso se utiliza para quitar tendencia y/o estacionalidad. Frecuentemente, la transformación no sólo estabiliza la varianza, sino que también mejora la aproximación de la serie a la normal.

TESIS CON
FALLA DE ORIGEN

λ	Transformación
0	$\ln(Z_t)$
0.5	$\sqrt{Z_t}$
-0.5	$\frac{1}{\sqrt{Z_t}}$
-1	$\frac{1}{Z_t}$
1	Z_t

Cuadro 1.2.2. Transformación Box-Cox para distintos valores de lambda

➤ La relación entre dos variables sólo depende del intervalo que las separa. Esta condición influye sobre la correlación de los datos. La autocorrelación mide la correlación entre un valor observado Y_t y otro Y_{t+k} , que están separados por un intervalo de longitud k :

$$\rho_k = \text{corr}(Y_t, Y_{t+k})$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t+k})}} = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sigma^2_Y} \quad (1.2.6)$$

ρ_k (coeficiente de correlación) debe de ser tal que $-1 \leq \rho_k \leq 1$. Este valor dependerá sólo de k (intervalo). Si el proceso es estacionario estos valores pueden estimarse tomando la serie de tiempo como una muestra de tamaño n , donde n es el número de observaciones, es decir:

$$\overline{\rho_k} = \frac{\overline{\gamma_k}}{\overline{\gamma_0}} = \frac{\frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2} = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (1.2.7)$$

La gráfica de la ecuación 1.2.7 para cada $k = 0, 1, 2, 3, \dots$, se le llama diagrama de correlación o correlograma (en este caso, dado que es una misma variable en distinto t , se le llama autocorrelación). Es de notar que a medida que k es más grande, el coeficiente de correlación es más pequeño.

b) No debe tener fluctuaciones estacionales

Cuando una serie se repite por periodos marcados, tiene estacionalidad. Esta es generalmente obvia en la gráfica de los datos. La estacionalidad se identifica con mayor facilidad después de eliminar la tendencia (en caso necesario). Existen varios métodos para el manejo de las fluctuaciones estacionales basados en la información contenida en la serie. Para realizar el ajuste, primero se debe de detectar la longitud de periodo de la fluctuación estacional (es llamada s y denotado por s), y después realizar diferencias consecutivas, es decir:

$$\Delta_s Y_t = Y_t - Y_{t-s} \quad (1.2.8)$$

Cada vez que se realizan este tipo de diferencias se pierden s datos de la serie.

c) Continuas y sin datos discrepantes

Una serie es continua si para cada t (tiempo) se tiene un Y_t (observación). Algunas veces no se cuenta con observaciones para cada t , haciendo que la serie no sea continua. Los motivos pueden ser muchos, pero para tales casos se puede tomar el promedio de la serie. De esa forma, la serie sería continua. Otras veces existen datos discrepantes, que en su mayoría se deben a errores de captura. La forma más fácil de detectarlos es gráficamente. Una vez detectados se estudia el tiempo en el que se dieron y se investiga si realmente ocurrieron o es un error de captura.

En algunos casos no se altera la serie Y_t , es decir, no se aplican diferencias de ningún orden, ni se aplican transformaciones. Sin embargo, se maneja una serie Z_t . Esto es sólo por nomenclatura; para tales casos se puede hacer una aclaración para evitar confusiones, tal como:

$$Y_t = Z_t$$

1.3 Modelos de Series de Tiempo

Existen varios métodos para analizar las series de tiempo. Los modelos de series de tiempo pueden ser lineales o no lineales, univariados o multivariados. En este último caso se manejan diferentes conjuntos de series para formar modelos causales y usarlos para encontrar una función. La metodología estadística univariada permite al analista, entre otras cosas, formular predicciones, determinar sucesos atípicos, extraer señales de evaluación, examinar perfiles cíclicos o estacionales, etc. Este capítulo presenta las principales características de los modelos lineales univariados que resultan muy útiles al describir una gran variedad de series de tiempo.

1.3.1 Modelos Estacionarios

1.3.1.1 Modelos Autorregresivos AR(p) y de Medias Móviles MA(q)

Una serie de tiempo Z_t puede expresarse en función de dos cosas:

- a) Su historia: Z_{t-1}, Z_{t-2}, \dots
- b) Errores Aleatorios $e_t, e_{t-1}, e_{t-2}, \dots$

Para el primer caso, se dice que una serie de tiempo Z_t está gobernada por un proceso autorregresivo si los valores actuales de la serie pueden representarse como una función lineal de los valores anteriores y un choque aleatorio e_t . Esto es:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + e_t$$

Este modelo no es usado ya que contiene un número infinito de parámetros que es imposible estimarlos con un número finito de observaciones. En su lugar, se construyen modelos con un número finito de parámetros, es decir:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + e_t \quad (1.3.1)$$

La expresión (1.3.1) constituye el llamado modelo Autorregresivo de orden p AR(p). Algunos autores incluyen el operador de salto hacia atrás³ para representar este modelo (Ver expresión 1.3.2)

³ back shift operator

$$BZ_t = Z_{t-1}$$

$$B^2 Z_t = Z_{t-2}$$

⋮

$$B^p Z_t = Z_{t-p}$$

Sustituyendo en (1.3.1)

$$Z_t - \phi_1 BZ_t - \phi_2 B^2 Z_t - \dots - \phi_p B^p Z_t = e_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = e_t$$

$$\phi(B) Z_t = e_t \quad (1.3.2)$$

Los valores e_t son conocidos como ruido blanco y tiene las siguientes características:

$$E(e_t) = 0 \quad E(e_t Z_{t-k}) = 0$$

$$E(e_t e_k) = \begin{cases} \sigma^2 & \text{si } t = k \\ 0 & \text{otro caso} \end{cases}$$

$$e_t \sim N(0, \sigma^2)$$

Así pues, un modelo de autorregresivo de segundo orden AR(2) estará dado por:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t$$

El valor de los parámetros ϕ_1 y ϕ_2 está restringido debido a las condiciones de estacionaridad descritas en el subcapítulo 1.2. Es decir, la media y varianza tienen que ser constantes y la relación entre los valores de la variable sólo dependa del intervalo que los separa. Primero, se revisa si la media es constante:

$$\text{Si } E(Z_t) = E(Z_{t-k}) = \mu$$

Entonces

$$E(Z_t) = E(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t)$$

$$= E(\phi_1 Z_{t-1}) + E(\phi_2 Z_{t-2}) + E(e_t)$$

$$= \phi_1 E(Z_{t-1}) + \phi_2 E(Z_{t-2}) + 0$$

$$= \phi_1 E(Z_t) + \phi_2 E(Z_t)$$

$$= (\phi_1 + \phi_2) E(Z_t)$$

Por lo tanto, la media $\mu = 0$ es constante sin importar ϕ_1 y ϕ_2

Ahora se revisa si la varianza es constante:

$$\text{Si } \text{Var}(Z_t) = \sigma^2$$

Entonces

$$\begin{aligned} \gamma_0 = \text{Var}(Z_t) &= E(Z_t Z_t) \\ &= E[(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t) Z_t] \\ &= \phi_1 E(Z_{t-1} Z_t) + \phi_2 E(Z_{t-2} Z_t) + E(e_t Z_t) \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma^2 \end{aligned} \quad (1.3.3)$$

De donde se deriva que

$$\phi_2 + \phi_1 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$-1 < \phi_2 < 1$$

Se puede ver más claramente la obtención de estas condiciones utilizando en el modelo AR(2) el operador de salto hacia atrás, es decir:

$$(1 - \phi_1 B - \phi_2 B^2) Z_t = e_t$$

y el problema se reduce a encontrar las raíces B1 y B2 que satisfacen la expresión:

$$1 - \phi_1 B - \phi_2 B^2 = 0$$

o bien

$$\phi_2 B^2 + \phi_1 B - 1 = 0$$

Con ello se tiene que:

$$B_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \quad \text{y} \quad B_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}$$

Además

$$\frac{1}{B_1} = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \quad \text{y} \quad \frac{1}{B_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

La expresión anterior implica que $|B_i| > 1$ y esta a su vez $|1/B_i| < 1$, para $i=1,2$.

Entonces:

Análisis de Intervención: Aplicación a un problema ambiental.

$$\left| \frac{1}{B_1} \cdot \frac{1}{B_2} \right| = |\phi_2| < 1 \quad \text{y} \quad \left| \frac{1}{B_1} + \frac{1}{B_2} \right| = |\phi_1| < 2$$

Con ello se tienen condiciones necesarias para estacionaridad para raíces reales y complejas:

$$-1 < \phi_2 < 1 \quad \text{y} \quad -2 < \phi_1 < 2$$

Para el caso de las raíces reales, se necesita que:

$$\phi_1^2 + 4\phi_2 \geq 0$$

lo cual implica que:

$$-1 < \frac{1}{B_1} = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \leq \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} = B_2 < 1$$

que a su vez es equivalente a

$$\phi_2 + \phi_1 < 1$$

$$\phi_2 - \phi_1 < 1$$

La condición $-1 < \phi_2 < 1$ se deduce tomando en cuenta los casos de las raíces complejas y las expresiones anteriores. En general, un proceso AR(p) cumple las condiciones de estacionaridad si el valor de las p raíces del vector:

$$1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p = \Phi(B) = 0$$

están fuera del círculo unitario, es decir, $|B| > 1$.

Ahora bien, queda por analizar la función de autocorrelación ACF y la función de autocorrelación parcial PACF. LA ACF se obtiene calculando:

$$E(Z_{t-k} Z_t) = \phi_1 E(Z_{t-k} Z_{t-1}) + \phi_2 E(Z_{t-k} Z_{t-2}) + E(e_t Z_{t-k})$$

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \quad k \geq 1$$

Entonces para el caso AR(2) y tomando en cuenta la expresión (1.3.3)

TESIS CON
 FALLA DE ORIGEN

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}}{\gamma_0} = \frac{\phi_1 \gamma_{k-1}}{\gamma_0} + \frac{\phi_2 \gamma_{k-2}}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$$

$$k = 1 \quad \rho_1 = \phi_1 + \phi_2 \rho_1$$

$$k = 2 \quad \rho_2 = \phi_1 \rho_1 + \phi_2$$

Se tienen que encontrar solución al sistema de ecuaciones anterior, es decir, encontrar valores ρ_1 y ρ_2 que satisfagan simultáneamente ambas ecuaciones. Esto es:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \quad \text{y} \quad \rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$$

En general, para un proceso AR(p) la expresión:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k \geq 1$$

ayuda a obtener las ecuaciones⁴, a partir de las cuales se calculan las autocorrelaciones. Estas últimas serán decrecientes infinitas, lo cual, es útil para detectar el posible modelo a usar para representar los datos.

Los coeficientes ρ_{kk} de la función de autocorrelación parcial PACF⁵, también resultan útiles al construir un modelo AR(p). Estos coeficientes miden la correlación no tomada en cuenta por un modelo AR(k-1). Es decir, si se tiene un modelo AR(k-1) y se desea saber si un modelo AR(k) es más adecuado, se tendrá que incluir un parámetro más. Si el valor del coeficiente ρ_{kk} es significativo, entonces se deberá incluir un parámetro más.

La expresión:

$$\rho_j = \rho_{k1} \rho_{j-1} + \rho_{k2} \rho_{j-2} + \dots + \rho_{kk} \rho_{j-p} \quad j = 1, 2, 3, \dots, k$$

⁴ Estas ecuaciones son llamadas "ecuaciones de Yule-Walker"

⁵ A diferencia de los coeficientes de ACF ρ_k , los coeficientes de la PACF tendrán doble subíndice ρ_{kk}

genera el sistema de ecuaciones con el que se pueden calcular los valores de los coeficientes de la PACF, esto es,

$$\begin{aligned} \rho_1 &= \rho_{k1}\rho_0 + \rho_{k2}\rho_1 + \dots + \rho_{kk}\rho_{k-1} \\ \rho_2 &= \rho_{k1}\rho_1 + \rho_{k2}\rho_0 + \dots + \rho_{kk}\rho_{k-2} \\ &\vdots \\ \rho_k &= \rho_{k1}\rho_{k-1} + \rho_{k2}\rho_{k-2} + \dots + \rho_{kk}\rho_0 \end{aligned}$$

Utilizando la regla de Cramer sucesivamente para $k=1,2,\dots$ se tiene:

$$\begin{aligned} \rho_{11} &= \rho_1 & \rho_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} & \rho_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \dots \\ \dots \rho_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}} \end{aligned} \quad (1.3.4)$$

Así entonces, para un proceso AR(2) se tiene que:

$$\rho_{11} = \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

$$\rho_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\left(\frac{\phi_1^2 + \phi_2 - \phi_2^2}{1 - \phi_2}\right) - \left(\frac{\phi_1}{1 - \phi_2}\right)^2}{1 - \left(\frac{\phi_1}{1 - \phi_2}\right)^2} = \frac{\phi_2 \left[(1 - \phi_2)^2 - \phi_1^2 \right]}{(1 - \phi_2)^2 - \phi_1^2} = \phi_2$$

$$\rho_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho_1 & \Phi_1 + \Phi_2 \rho_1 \\ \rho_1 & 1 & \Phi_1 \rho_1 + \Phi_2 \\ \rho_2 & \rho_1 & \Phi_1 \rho_2 + \Phi_2 \rho_1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = 0$$

Así pues, $\rho_{33}=0$ ya que la última columna del determinante del numerador es una combinación lineal de las primeras dos columnas. De esa manera, se puede notar que para un proceso AR(2), su PACF se trunca en el segundo coeficiente. Usando este hecho, se generaliza para un proceso AR(p), es decir, al calcular ρ_{kk} ($k > p$), la última columna de la matriz del numerador es una combinación lineal de las columnas anteriores. Con lo cual $\rho_{kk} = 0$ y la PACF de un AR(p) se trunca en el coeficiente p. Esta es la forma en que la PACF ayuda a identificar el modelo AR(p).

Como se había mencionado, otra manera de expresar las series de tiempo es utilizando los errores aleatorios e_t . En este caso, la serie de tiempo Z_t se expresa como una combinación lineal de errores aleatorios no correlacionados, es decir:

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

o bien

$$Z_t = \theta(B)e_t$$

$$\text{donde } \theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \quad (1.3.5)$$

La expresión (1.3.5) es el llamado modelo de medias móviles (Moving Averages) de orden q MA(q). $\{e_t\}$ es ruido blanco con media igual a cero y varianza σ^2 . Esto es,

$$E(Z_t) = 0$$

$$\text{Var}(Z_t) = \gamma_0 = \sigma^2 \sum_{j=0}^q \theta_j^2 \quad \text{con } \theta_0 = 1 \quad (1.3.6)$$

$$\gamma_k = E(Z_t Z_{t+k}) = E\left(\sum_{i=0}^q \sum_{j=0}^q \theta_i \theta_j e_{t-i} e_{t-j}\right) = \sigma^2 \sum_{i=0}^q \theta_i \theta_{i+k} \quad (1.3.7)$$

y

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (1.3.8)$$

Se puede notar claramente, que para un proceso MA(q) la media y varianza son constantes y la función de correlación parcial ACF sólo depende de k. Por lo tanto, los procesos MA(q) no tienen condiciones de estacionaridad. No obstante, estos procesos tienen que cumplir con condiciones de invertibilidad.

Por ejemplo, el proceso MA(2):

$$Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

de las expresiones (1.3.6) y (1.3.7) se sabe que:

$$\begin{aligned} E(Z_t) &= 0 \\ \text{Var}(Z_t) &= \gamma_0 = \sigma^2 (1 + \theta_1^2 + \theta_2^2) \\ \gamma_1 &= -\theta_1 (1 - \theta_2) \sigma^2 \\ \gamma_2 &= -\theta_2 \sigma^2 \\ \gamma_k &= 0 \quad k > 2 \end{aligned}$$

y las autocorrelaciones obtenidas con la expresión (1.3.8) son:

$$\rho_1 = \frac{-\theta_1 (1 - \theta_2)}{(1 + \theta_1^2 + \theta_2^2)}, \quad \rho_2 = \frac{-\theta_2}{(1 + \theta_1^2 + \theta_2^2)} \quad \text{y} \quad \rho_k = 0 \quad k > 2$$

Dadas éstas, la ACF de un MA(2) se trunca después de la segunda autocorrelación. Además, para que ρ_1 y ρ_2 se puedan calcular el denominador tiene que ser distinto de cero. Esto es lo que constituye las condiciones de invertibilidad de un MA(2), que son:

$$\theta_2 + \theta_1 < 1 \quad \theta_2 - \theta_1 < 1 \quad -1 < \theta_2 < 1$$

Nótese que estas condiciones son muy parecidas a las condiciones de estacionaridad para un AR(2). La dualidad entre los procesos MA(q) y AR(p) se explica más adelante con mayor detalle.

Usando la expresión (1.3.4), y el hecho de que $\rho_k=0$ para $k>2$, se puede obtener los valores de la función de autocorrelación parcial PACF del MA(2):

$$\rho_{11} = \rho_1 \quad , \quad \rho_{22} = \frac{\rho_2 - \rho_1^2}{(1 - \rho_1^2)} \quad , \quad \rho_{33} = \frac{\rho_1^3 - \rho_1 \rho_2 (2 - \rho_2)}{1 - \rho_2^2 - 2\rho_1^2(1 - \rho_2)} \quad , \dots$$

Como se puede notar, la PACF será decreciente infinita.

En general, la ACF de un proceso MA(q) se trunca en el intervalo q y su PACF es decreciente infinita. Esta es una importante propiedad en el reconocimiento y selección de modelos. Un MA(q) no tendrá condiciones de estacionaridad, pero sí de invertibilidad. Para saber cuáles son éstas el modelo MA(q) también se puede escribir como:

$$\begin{aligned} Z_t &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \\ Z_t &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \\ Z_t &= \theta(B) e_t = \frac{1}{\Pi(B)} e_t \end{aligned}$$

Para que ello sea válido se debe cumplir que las raíces de:

$$(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) = 0$$

estén fuera del círculo unitario, esto es, $|B|>1$.

1.3.1.2 Dualidad entre los procesos AR(p) y MA(q)

Las características mostradas en el subcapítulo anterior dan a entender que puede existir dualidad entre los modelos AR(p) y MA(q). La ACF de un MA(q) se comporta como la PACF de un AR(p) y la ACF de un AR(p) se comporta como la PACF de un MA(q). Además, los AR(p) deben cumplir condiciones de estacionaridad, pero no de invertibilidad, mientras que los MA(q) tienen condiciones de invertibilidad, pero no de estacionaridad. El cuadro (1.3.1) muestra algunas de estas características.

Modelo	Condiciones	ACF	PACF
AR(p)	Estacionaridad	Decreciente Infinita	Se trunca en p
MA(q)	Invertibilidad	Se trunca en q	Decreciente Infinita

Cuadro 1.3.1 Características de los modelos AR(p) y MA(q)

Un proceso estacionario AR(p) de la forma:

$$\phi(B)Z_t = e_t$$

$$\text{donde } \phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

puede escribirse como:

$$Z_t = \frac{1}{\phi(B)} e_t = \psi(B) e_t$$

$$\text{donde } \psi(B) = (1 + \psi_1 B + \psi_2 B^2 + \dots) \text{ tal que } \phi(B)\psi(B) = 1$$

Por ejemplo, se puede escribir un proceso AR(2) como

$$Z_t = \frac{1}{(1 - \phi_1 B - \phi_2 B^2)} e_t = (1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots) e_t$$

implicando

$$(1 - \phi_1 B - \phi_2 B^2) (1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots) = 1$$

o bien

$$\begin{aligned} & 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots \\ & - \phi_1 B - \psi_1 \phi_1 B^2 - \psi_2 \phi_1 B^3 - \dots \\ & - \phi_2 B^2 - \psi_1 \phi_2 B^3 - \dots = 1 \end{aligned}$$

De esa forma se pueden obtener los valores para ψ_j :

$$\begin{aligned} B : & \quad \psi_1 - \phi_1 = 0 \Rightarrow \psi_1 = \phi_1 \\ B^2 : & \quad \psi_2 - \psi_1 \phi_1 - \phi_2 = 0 \Rightarrow \psi_2 = \psi_1 \phi_1 + \phi_2 = \phi_1^2 + \phi_2 \\ B^3 : & \quad \psi_3 - \psi_2 \phi_1 - \psi_1 \phi_2 = 0 \Rightarrow \psi_3 = \psi_2 \phi_1 + \psi_1 \phi_2 \\ & \quad \vdots \\ B^j : & \quad \psi_j = \psi_{j-1} \phi_1 + \psi_{j-2} \phi_2 \quad j \geq 2 \text{ y } \psi_0 = 1 \end{aligned}$$

Un caso especial es cuando $\phi_2=0$, ya que se tiene:

$$\Psi_j = \phi_1^j \quad j \geq 0 \quad \text{y} \quad \Psi_0 = 1$$

entonces

$$Z_t = \frac{1}{(1 - \phi_1 B)} e_t = (1 + \phi_1 B + \phi_1^2 B^2 + \phi_1^3 B^3 + \dots) e_t$$

Esto demuestra que un proceso AR(p) estacionario y finito es equivalente a un proceso MA infinito.

Ahora bien, tomando un proceso MA(q) invertible:

$$Z_t = \theta(B) e_t \quad \text{donde} \quad \theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

se puede escribir como

$$e_t = \frac{1}{\theta(B)} Z_t = \pi(B) Z_t$$

$$\text{donde} \quad \pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots \quad \text{tal que} \quad \theta(B)\pi(B) = 1$$

Por ejemplo un MA(2) estará dado por:

$$e_t = \frac{1}{(1 - \theta_1 B - \theta_2 B^2)} Z_t = (1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots) Z_t$$

implicando

$$(1 - \theta_1 B - \theta_2 B^2) (1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots) = 1$$

o bien

$$\begin{aligned} 1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots \\ - \theta_1 B + \pi_1 \theta_1 B^2 + \pi_2 \theta_1 B^3 + \dots \\ - \theta_2 B^2 + \pi_1 \theta_2 B^3 + \dots = 1 \end{aligned}$$

Ahora se obtienen los valores para π_j :

$$\begin{aligned} B : \quad & -\pi_1 - \theta_1 = 0 \Rightarrow \pi_1 = -\theta_1 \\ B^2 : \quad & -\pi_2 + \pi_1 \theta_1 - \theta_2 = 0 \Rightarrow \pi_2 = \pi_1 \theta_1 - \theta_2 = -\theta_1^2 - \theta_2 \\ B^3 : \quad & -\pi_3 + \pi_2 \theta_1 + \pi_1 \theta_2 = 0 \Rightarrow \pi_3 = \pi_2 \theta_1 + \pi_1 \theta_2 \\ & \vdots \\ B^j : \quad & \pi_j = \pi_{j-1} \theta_1 + \pi_{j-2} \theta_2 \quad j \geq 3 \end{aligned}$$

Cuando $\theta_2=0$ se tiene un MA(1) y

TFSIS CON
FALLA DE ORIGEN

$$\pi_j = -\theta_1^j \quad j \geq 1$$

entonces

$$e_t = \frac{1}{(1-\theta_1 B)} Z_t = (1 + \theta_1 B + \theta_1^2 B^2 + \theta_1^3 B^3 + \dots) Z_t$$

Por lo tanto, un proceso MA(q) invertible y finito es equivalente a un proceso AR infinito.

Como se ha demostrado existe una dualidad entre los procesos de medias móviles MA y los procesos Autorregresivos AR. Esta dualidad también se nota en las funciones de autocorrelación ACF y autocorrelación parcial PACF (Ver Cuadro 1.3.1).

1.3.1.3 Modelos Mixto General ARMA(p,q)

Los procesos estacionarios e invertibles pueden ser representados mediante modelos autorregresivos AR o de medias móviles MA. Sin embargo, un problema de estas representaciones es que pueden contener demasiados parámetros y, en general, un número grande de parámetros reduce la eficiencia del modelo⁶. Es por ello, que al construir modelos, lo mejor es incluir términos autorregresivos y de medias móviles, es decir,

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

o bien

$$\phi(B)Z_t = \theta(B)e_t$$

$$\text{donde } \theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q) \quad \text{y} \quad \phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$$

Este es llamado modelo mixto general ARMA(p,q). Dado que contiene tanto términos de medias móviles, como autorregresivos, debe cumplir condiciones de invertibilidad y de estacionaridad. Para lograr cumplir esas condiciones se requiere que las raíces de $\phi(B)=0$ y $\theta(B)=0$ estén fuera del círculo unitario.

Para obtener la función de autocorrelación ACF de un ARMA(p,q) se parte de:

$$Z_{t-k}Z_t = \phi_1 Z_{t-k}Z_{t-1} + \dots + \phi_p Z_{t-k}Z_{t-p} + Z_{t-k}e_t - \theta_1 Z_{t-k}e_{t-1} - \dots - \theta_q Z_{t-k}e_{t-q}$$

⁶ Este es llamado "principio de parsimonia". Más adelante, en el Subcapítulo 1.4.3, se da más detalle de este principio.

$$E(Z_{t-k}Z_t) = \gamma_k = \Phi_1\gamma_{k-1} + \dots \\ \dots + \Phi_p\gamma_{k-p} + E(Z_{t-k}e_t) - \theta_1E(Z_{t-k}e_{t-1}) - \dots - \theta_qE(Z_{t-k}e_{t-q})$$

dado que

$$E(Z_{t-k}e_{t-i}) \quad \text{para } k > i$$

se tiene:

$$\gamma_k = \Phi_1\gamma_{k-1} + \dots + \Phi_p\gamma_{k-p} \quad \text{para } k > (q+1)$$

$$\rho_k = \Phi_1\rho_{k-1} + \dots + \Phi_p\rho_{k-p} \quad \text{para } k > (q+1)$$

Así entonces, la función de autocorrelación de un ARMA(p,q) será decreciente infinita, debida a la parte autorregresiva y el efecto de las medias móviles desaparece después de q periodos ($k > q+1$). El valor de las primeras q autocorrelaciones $\rho_q, \rho_{q-1}, \dots, \rho_1$ depende tanto de los parámetros autorregresivos, como los de medias móviles. Por ejemplo, un ARMA(1,1) sería:

$$Z_t = \Phi_1 Z_{t-1} + e_t - \theta_1 e_{t-1}$$

o bien

$$(1 - \theta_1 B)Z_t = (1 - \Phi_1 B)e_t$$

Para cumplir con invertibilidad $|\theta_1| < 1$ y con estacionaridad $|\Phi_1| < 1$

$$\gamma_0 = \frac{(1 + \theta_1^2 - 2\Phi_1\theta_1)}{(1 - \Phi_1^2)} \sigma^2$$

$$\gamma_1 = \frac{(\Phi_1 - \theta_1)(1 - \Phi_1\theta_1)}{(1 - \Phi_1^2)} \sigma^2$$

$$\gamma_k = \Phi_1\gamma_{k-1} = \Phi_1^{k-1}\gamma_1 \quad k \geq 2$$

Por lo tanto, para un ARMA(1,1), la ACF sería:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{(\Phi_1 - \theta_1)(1 - \Phi_1\theta_1)}{(1 + \theta_1^2 - 2\Phi_1\theta_1)} & k = 1 \\ \Phi_1\rho_{k-1} = \Phi_1^{k-1}\rho_1 & k \geq 2 \end{cases}$$

Análisis de Intervención: Aplicación a un problema ambiental.

Dado que un ARMA(p,q) contiene términos de medias móviles, la función de autocorrelación parcial PACF será decreciente infinita.

1.3.2 Modelos No estacionarios

1.3.2.1 No estacionaridad en la media y varianza

Hasta ahora se ha mostrado los modelos de las series estacionarias, pero en la mayoría de los casos no se cumple con la estacionaridad. Las series de tiempo pueden ser no estacionarias en la media o en la varianza. Si no se cumple con estacionaridad, se utilizan diferentes métodos para hacer estacionario el proceso. La aplicación de éstos dependerá del tipo de no estacionaridad que se tenga:

a) No Estacionaridad en la media

Este caso la serie de tiempo tiene tendencia y se puede eliminar diferenciando como se mostró en el apartado (1.2). Una vez realizadas las diferencias se obtiene una nueva serie de tiempo. Esto es, una serie de tiempo $\{Z_t\}$ no estacionaria en la media puede reducirse a una serie estacionaria $\{(1-B)^d Z_t\}$ mediante la aplicación de $d > 0$ diferencias apropiadas.

b) No Estacionaridad en la varianza

Un proceso que es estacionario en la media no es necesariamente estacionario en la varianza. Existen muchas series de tiempo que se comportan así. No todos los problemas de no estacionaridad se pueden remediar usando diferencias. La forma de hacer estacionaria una serie de tiempo en la varianza es mediante el uso de transformaciones. Estas transformaciones también se mostraron en (1.2). Generalmente, estas transformaciones están definidas para valores positivos. Es importante hacer notar que si se pretende aplicar una transformación para estabilizar la varianza, se debería realizar antes de aplicar diferencias.

1.3.2.2 Modelo General ARIMA (p,d,q)

Si se ha modificado la serie de tiempo inicial aplicando diferencias, entonces el modelo general ARMA(p,q):

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

o bien

$$\Phi(B)Z_t = \theta(B)e_t$$

será de la forma:

$$\Phi(B)(1-B)^d Z_t = \theta(B)e_t \quad (1.3.9)$$

donde d es el número de diferencias ordinarias aplicadas al modelo ARMA(p,q) para hacerlo estacionario. Al modelo de la expresión (1.3.9) se le conoce como el modelo autorregresivo integrado y de medias móviles ARIMA (p,d,q). Otra forma de expresar este tipo de modelo es la siguiente:

$$\Phi(B)W_t = \theta(B)e_t$$

donde

$$\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$$

$$W_t = \begin{cases} (1-B)^d Z_t & d > 0 \\ Z_t & d = 0 \end{cases}$$

Así por ejemplo, un modelo ARIMA(1,1,1) y tomando (1.3.9) estará dado por:

$$\text{Modelo ARIMA(1,1,1)} \left\{ \begin{array}{l} (1 - \phi_1 B)(1 - B)Z_t = (1 - \theta_1 B)e_t \\ (1 - \phi_1 B)(Z_t - Z_{t-1}) = e_t - \theta_1 e_{t-1} \\ Z_t - Z_{t-1} - \phi_1 Z_{t-1} + \phi_1 Z_{t-2} = e_t - \theta_1 e_{t-1} \\ Z_t = Z_{t-1} + \phi_1 Z_{t-1} - \phi_1 Z_{t-2} + e_t - \theta_1 e_{t-1} \\ \text{si } W_t = (1 - B)Z_t = Z_t - Z_{t-1} \text{ entonces} \\ Z_t - Z_{t-1} = \phi_1 Z_{t-1} - \phi_1 Z_{t-2} + e_t - \theta_1 e_{t-1} \\ W_t = \phi_1 W_{t-1} + e_t - \theta_1 e_{t-1} \end{array} \right.$$

CON
FALLA DE ORIGEN

1.3.3 Modelos Estacionales

1.3.3.1 Modelos Estacionales Puros

Muchas series de tiempo contienen fenómenos que se repiten en ciertos periodos. A las series con esta característica se les llama series de tiempo estacionales. Los modelos que se construyen con este tipo de serie son llamados modelos estacionales y toman en cuenta la correlación entre los mismos meses, trimestres, etc. (según sea el periodo estacional s) de años sucesivos, así como la correlación entre meses, trimestres sucesivos. Los modelos estacionales puros pueden ser:

a) Modelos Estacionales Autorregresivos AR(P)s

Si el valor actual Z_t se puede expresar como una función lineal de los valores de la serie cada s periodos atrás, Z_{t-Ps} , y e_t , es decir:

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_P Z_{t-Ps} + e_t$$

o bien

$$\left(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}\right) Z_t = e_t$$

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

La función de autocorrelación de un SAR(P) o AR(P)s es, en general, similar a un AR(p), sólo que este caso, los valores de las autocorrelaciones aparecen en múltiplos de s . Por tanto, La ACF es decreciente infinita y distinto de cero sólo en los intervalos $s, 2s, 3s, \dots$. La PACF de un AR(P)s sólo tendrá un valor significativo en $k=s$.

b) Modelos Estacionales de Medias Móviles MA(Q)s

Si el valor actual Z_t se puede expresar como una función lineal de e_t y los choques aleatorios ocurridos cada s periodos antes, e_{t-s} , y e_t , esto es:

$$Z_t = e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \dots - \Theta_Q e_{t-Qs}$$

o bien

$$Z_t = \left(1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}\right) e_t$$

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

La ACF de MA(Q)s o SMA(Q) es similar a la de un MA(q), pero los valores de las autocorrelaciones aparecen en múltiplos de s y se truncan después de s intervalos. La PACF será decreciente infinita con valores significativos en múltiplos de s .

c) Modelos Estacionales Mixtos ARIMA(P,D,Q)s

Con los modelos estacionales mostrados en los incisos anteriores se puede construir un modelo mixto ARMA(P,Q)s de la forma:

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_P Z_{t-Ps} + e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \dots - \Theta_Q e_{t-Qs}$$

o bien

$$\left(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}\right) Z_t = \left(1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}\right) e_t$$

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

A los modelos estacionales también se les pueden aplicar diferencias como a los modelos no estacionales. En este caso, las diferencias que se aplican no son simples, es decir, diferencias con respecto de uno o dos observaciones atrás, sino que son estacionales, es decir, diferencias con respecto a s observaciones atrás.

De ese modo, se construye el modelo ARIMA(P,D,Q)s:

$$\left(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}\right) (1 - B^s)^D Z_t = \left(1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}\right) e_t$$

$$\Phi(B^s) (1 - B^s)^D Z_t = \Theta(B^s) e_t$$

Nótese que los parámetros autorregresivos y de medias móviles para modelos estacionales están representados con mayúsculas.

1.3.3.2 Modelo General Multiplicativo Estacional ARIMA(p,d,q)X(P,D,Q)s

Después de haber mostrado toda la variedad de modelos que se tienen, lo más lógico a seguir, es combinarlos en un solo modelo general, que pueda ajustarse a gran

variedad de series de tiempo. Así entonces, surge el Modelo General Multiplicativo Estacional ARIMA(p,d,q)X(P,D,Q)s :

$$\begin{aligned} & \left(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p\right) \left(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}\right) (1-B)^d (1-B^s)^D Z_t = \\ & = \left(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q\right) \left(1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}\right) e_t \end{aligned}$$

o bien

$$\phi(B) \Phi(B^s) (1-B)^d (1-B^s)^D Z_t = \theta(B) \Theta(B^s) e_t$$

Este modelo es el más general, ya que involucra parámetros autorregresivos y de medias móviles no estacionales y estacionales. De igual forma, toma en cuenta las diferencias de tipo estacional y no estacional. Algunos ejemplos de modelos multiplicativos generales son:

a) ARIMA(1,0,0)X(1,0,0)₁₂ o bien AR(1)XSAR(1)₁₂

$$(1 - \phi_1 B) (1 - \Phi_1 B^{12}) Z_t = e_t$$

b) ARIMA(0,0,1)X(0,0,1)₄ o bien MA(1)XSMA(1)₄

$$Z_t = (1 - \theta_1 B) (1 - \Theta_1 B^4) e_t$$

c) ARIMA(1,1,1)X(1,1,1)₆

$$(1 - \phi_1 B) (1 - \Phi_1 B^6) (1-B) (1-B^6) Z_t = (1 - \theta_1 B) (1 - \Theta_1 B^6) e_t$$

El cuadro 1.3.2 muestra las principales características de los modelos multiplicativos.

Modelo	ACF	PACF
AR(p) X SAR(P)	Decreciente Infinita	Se trunca después de p+Ps
MA(q) X SMA(Q)	Se trunca después de q+Qs	Decreciente Infinita
Modelos Mixtos	Decreciente Infinita	Decreciente Infinita

Cuadro 1.3.2 Características de los modelos multiplicativos

Nuevamente, se ve la importancia de la función de autocorrelación ACF y de la función de autocorrelación parcial PACF en la detección del tipo de modelo. Sin embargo, sólo se han mostrado la ACF y PACF teóricas. La expresión que se usa para obtener la ACF muestral es:

TESIS CON
FALLA DE ORIGEN

$$\rho_k = \frac{\hat{\rho}_k}{\hat{\rho}_0} = \frac{\frac{1}{n} \sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\frac{1}{n} \sum_{t=1}^n (z_t - \bar{z})^2} = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2} \quad k = 0, 1, 2, \dots$$

$$\text{donde} \quad \bar{z} = \frac{1}{n} \sum_{t=1}^n z_t$$

y la PACF muestral se obtiene con un método recursivo dado por Durbin (1960):

$$\rho_{1,1} = \rho_1$$

$$\rho_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^k (\rho_{k,j}) (\rho_{k+1-j})}{1 - \sum_{j=1}^k (\rho_{k,j}) \rho_j}$$

y

$$\rho_{k+1,j} = \rho_{k,j} - (\rho_{k+1,k+1})(\rho_{k,k+1-j}) \quad j = 1, 2, \dots, k$$

Otra forma para obtener la PACF es utilizar la regla de Cramer y resolver determinantes como se mostró en la expresión (1.3.4).

Ya que la ACF y PACF utilizan estimadores para cada valor de sus coeficientes, se obtiene un error llamado un *error estándar* definido por:

$$S_{\rho_k} = \sqrt{\frac{1}{n} (1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_q^2)} \quad k > q \quad (1.3.10)$$

Este último dato es importante, ya que es útil para determinar si un coeficiente es estadísticamente significativo. Se dice que un coeficiente es significativamente distinto de cero si el valor absoluto del estimador puntual es al menos dos veces su error estándar.

TESIS CON
FALLA DE ORIGEN

1.4 Estimación y Selección de Modelos

1.4.1 ¿Cómo se estima el valor de los parámetros?

Una vez que se ha identificado un modelo tentativo ARMA(p,q), el siguiente paso es estimar los parámetros del modelo. Es decir, se tiene que estimar el valor de los parámetros $\phi=(\phi_1,\phi_2,\dots,\phi_p)$, $\theta=(\theta_1,\theta_2,\dots,\theta_q)$ en el modelo

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

o bien

$$\phi(B)Z_t = \theta(B)e_t$$

Existen varios métodos, que pueden ser usados para estimar los parámetros de un modelo ARMA. Se pueden mencionar los siguientes:

a) Método de los momentos

Este método consiste en sustituir los momentos de la muestra tales como la media y la varianza muestrales por sus contrapartes teóricas y resolver el sistema de ecuaciones resultantes. Por ejemplo, en un proceso AR(p)

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

Para estimar el vector ϕ de parámetros, se usa la expresión (1.4.1) para obtener el sistema de ecuaciones Yule-Walker (1.4.2)

$$\rho_t = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (1.4.1)$$

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1}$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_2 + \dots + \phi_p \rho_{p-2}$$

⋮

$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \quad (1.4.2)$$

Entonces, reemplazando ρ_k por su estimación, se obtiene la estimación de cada parámetro al resolver el sistema de ecuaciones (1.4.3).

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-3} & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & \rho_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad (1.4.3)$$

Estos son los comúnmente llamados estimadores Yule-Walker y generalmente se usan como estimaciones o aproximaciones iniciales. Este método no es recomendable para resultados de estimación final y no se debe usar si el proceso está muy cerca de ser no estacionario o no invertible.

b) Métodos de Mínimos Cuadrados y Máxima Verosimilitud

Quizás, la estimación con el método de mínimos cuadrados, desarrollada para modelos de regresión, es la más usada en la estadística. Como en el caso de regresión lineal, se eligen aquellos valores de los parámetros que minimicen la suma de los cuadrados de las diferencias entre la serie real y la serie ajustada, es decir:

$$e_t = \theta^{-1}(B)\phi(B)Z_t \quad (1.4.4)$$

La ecuación (1.4.4) representa el valor de los *residuales*. El objetivo es encontrar un vector de parámetros autorregresivos ϕ y un vector de parámetros de medias móviles θ que minimicen la suma de los errores al cuadrado:

$$S(\phi, \theta) = \sum_t e_t^2 \quad (1.4.5)$$

Esta estimación se dificulta si están presentes términos de medias móviles, ya que la ecuación (1.4.4) es no lineal. Por ello, se usa un método iterativo de estimación no lineal para estimar (1.4.5). Además se debe inicializar la serie antes de aplicar el proceso de estimación no lineal ya que el primer valor de la serie, e_t , depende de los valores no observados $Z_0, Z_{-1}, \dots, Z_{-p+1}$ y $e_0, e_{-1}, \dots, e_{-q+1}$. Una posible solución es utilizar los valores esperados incondicionales que serán todos iguales a cero.

Para estimar los vectores de parámetros ϕ y θ se utiliza el hecho de que los términos de los errores et son independientes y distribuidos normalmente con media cero y varianza σ^2 . Entonces, la función de verosimilitud condicional asociada con los parámetros $(\phi, \mu, \theta, \sigma)$ está dada por :

$$F(\phi, \mu, \theta, \sigma) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{S(\phi, \theta)}{2\sigma^2}\right]$$

o bien

$$L(\phi, \mu, \theta, \sigma) = -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{S(\phi, \theta)}{2\sigma^2} \quad (1.4.6)$$

La ecuación (1.4.6) es la función de verosimilitud logarítmica condicional porque la suma de los errores cuadrados $S(\phi, \theta)$ depende de los valores pasados no observables Z_0, Z_1, \dots, Z_{p+1} y e_0, e_1, \dots, e_{q+1} . Los estimadores de máxima verosimilitud estarán dados por la minimización de la suma de los residuales al cuadrado. Entonces, y suponiendo normalidad, los estimadores de mínimos cuadrados y máxima verosimilitud serán los mismos. (Cabe notar que el estimador de mínimos cuadrados no contiene al parámetro σ^2).

c) Estimación No Lineal

Tomando en cuenta la suposición de que se inicializa la serie en los valores esperados incondicionales (igual a cero) para Z_0, Z_1, \dots, Z_{p+1} y e_0, e_1, \dots, e_{q+1} , el problema es minimizar:

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

donde Z es el vector de las observaciones Z_t . Si el proceso es puramente autorregresivo, la estimación se realiza como si fuese una regresión lineal múltiple. Esto es,

$$\phi(B)Z_t = e_t$$

o bien

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + e_t \quad (1.4.7)$$

La ecuación (1.4.7) es de la forma general:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + e_t$$

Por lo tanto, el vector ϕ de (1.4.7) se puede estimar como:

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

donde

$$X = \begin{bmatrix} Z_p & Z_{p-1} & \dots & Z_1 \\ Z_{p+1} & Z_p & \dots & Z_2 \\ \vdots & \vdots & \dots & \vdots \\ Z_{T-1} & Z_{T-2} & \dots & Z_{T-p} \end{bmatrix}$$

y

$$Y = \begin{bmatrix} Z_{p+1} \\ Z_{p+2} \\ \vdots \\ Z_T \end{bmatrix}$$

TESIS CON
FALLA DE ORIGEN

El problema es más difícil si el modelo contiene un componente de medias móviles:

$$\theta^{-1} \phi(B)Z_t = e_t \quad (1.4.8)$$

La ecuación (1.4.8) es una ecuación de regresión no lineal en los parámetros que se puede estimar con una rutina iterativa no lineal. Este proceso usa los dos primeros términos de la expansión de serie de Taylor para linealizar (1.4.8) alrededor de un valor inicial de ϕ , θ . Entonces, se realiza una regresión lineal y se obtienen estimadores de mínimos cuadrados para ϕ , θ y se hace una nueva linealización alrededor de ellos. Este

proceso continúa hasta que ocurre la convergencia. El proceso iterativo, con más detalle, es el siguiente: Sea β el vector que representa los $p+q$ parámetros ϕ , θ que se desean estimar. Se desea elegir valores para β que minimicen:

$$S(\beta) = \sum_{t=1}^T [e_t | \beta, z]^2$$

Expandiendo $[e_t]$ en una serie de Taylor alrededor de un valor inicial β_0 :

$$[e_t] = [e_t | z, \beta_0] + \sum_{i=1}^{p+q} (\beta_i - \beta_{i,0}) \left. \frac{\partial [e_t]}{\partial \beta_i} \right|_{\beta=\beta_0} + \frac{1}{2} \sum_{i=1}^{p+q} (\beta_i - \beta_{i,0})^2 \left. \frac{\partial^2 [e_t]}{\partial \beta_i^2} \right|_{\beta=\beta_0} + \dots \quad (1.4.9)$$

$[e_t]$ se aproxima con los dos primeros términos de la expansión de Taylor. Es decir,

$$X_{i,t} = \left. \frac{\partial [e_t]}{\partial \beta_i} \right|_{\beta=\beta_0}$$

y

$$[e_{t,0}] = [e_t | z, \beta_0]$$

Entonces,

$$[e_t] \approx [e_{t,0}] - \sum_{i=1}^{p+q} (\beta_i - \beta_{i,0}) X_{i,t}$$

$$[e_{t,0}] + \sum_{i=1}^{p+q} (\beta_{i,0}) X_{i,t} = [e_t] + \sum_{i=1}^{p+q} (\beta_i) X_{i,t} \quad (1.4.10)$$

Los valores β_i en la expresión (1.4.10) pueden estimarse por regresión lineal:

$$Y = \begin{bmatrix} [e_{1,0}] + \sum_{i=1}^{p+q} \beta_{i,0} X_{i,1} \\ \vdots \\ [e_{T,0}] + \sum_{i=1}^{p+q} \beta_{i,0} X_{i,T} \end{bmatrix}$$

$$X = \begin{bmatrix} X_{1,1} & X_{2,1} & \dots & X_{p+q,1} \\ X_{1,2} & X_{2,2} & \dots & X_{p+q,2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1,T} & X_{2,T} & \dots & X_{p+q,T} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{p+q} \end{bmatrix}$$

Se realiza una regresión ordinaria por mínimos cuadrados para producir un nuevo estimador de β (β_1). Este proceso se repite hasta que:

$$\beta_k - \beta_{k-1} \approx 0$$

o bien

$$|\beta_k - \beta_{k-1}| \leq \varepsilon$$

donde k es el número de convergencia (número de iteraciones). β_k será el estimador final de los parámetros ϕ , θ . También se puede utilizar, como criterio de convergencia, la suma de los cuadrados de los residuales. Algunos programas estadísticos utilizan este criterio. Es importante señalar que no siempre se puede obtener un estimador final de los parámetros, es decir, es posible que el proceso diverja o que existan soluciones múltiples. Si existe divergencia, se debe estimar el modelo nuevamente con valores iniciales diferentes. Si aún así no se converge, entonces se debe especificar nuevamente el modelo. También es importante señalar que se deben analizar diferentes soluciones con valores iniciales diferentes.

TESIS CON
FALLA DE ORIGEN

1.4.2 Ajuste de los Modelos

Teniendo el valor de cada uno de los parámetros del modelo se procede a dar un ajuste con el fin de mejorarlo. Para evaluar el ajuste del modelo se aplican, principalmente, cuatro pruebas:

- Análisis de Estacionaridad
- Análisis de Residuales
- Falta de Parámetros (subespecificación)
- Omisión de Parámetros (sobrespecificación)

Todas estas pruebas se deben de efectuar para asegurarse del ajuste del modelo. Si el modelo es cambiado se deben de efectuar todas las pruebas de nueva cuenta. Los desajustes producen errores no aleatorios

1.4.2.1 Análisis de Estacionaridad y Residuales

Una de las principales suposiciones iniciales es que la serie de tiempo es estacionaria, es decir, que la varianza de los datos es constante. El análisis de estacionaridad inicia examinando la gráfica de los datos. Cuando un proceso no es estacionario se pueden utilizar transformaciones sobre los datos o diferenciar el modelo, con el fin de obtener un proceso estacionario. El Cuadro 1.2.1 muestra las principales transformaciones que se usan. De igual forma, en la Ecuación 1.2.3 se ve la forma en que se pueden aplicar diferencias a una serie. Tanto las transformaciones, como las diferencias sólo se deben aplicar en casos necesarios; de lo contrario se pueden complicar los modelos.

Otra prueba para analizar el ajuste del modelo es analizar los errores estimados y sus propiedades. A esto se le conoce como el análisis de residuales⁷. Si los residuales tienen media cero y varianza constante (ruido blanco), entonces se sabe que el modelo planteado tiene un buen ajuste. Se espera que los residuales se parezcan lo más posible a los verdaderos errores los cuales no están correlacionados supuestamente. Graficando la ACF (función de autocorrelación) de los residuales se puede saber si realmente éstos se

⁷ Los residuales son la estimación de los errores reales.

comportan como ruido blanco. Si es así, su ACF no tendrá picos y las autocorrelaciones muestrales serán pequeñas. Otra forma es evaluar las autocorrelaciones de las primeras diferencias de los residuales. Esto es:

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

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

Finalmente, se debe saber si existe o no periodicidad en los residuales⁸. El periodograma integral (también llamado periodograma acumulado normalizado) nos ayuda en ello, y se define como:

$$C(f_j) = \frac{\sum_{i=1}^j R^2(f_i)}{ns^2}$$

donde

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

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

Los resultados del periodograma Integral se grafican contra las frecuencias o contra el periodo. Si los residuales son ruido blanco, entonces la gráfica del periodograma integral contra las frecuencias quedaría como una línea recta (45 grados), ya que todas las frecuencias tendrán la misma importancia en el proceso. En la mayoría de las veces, no se obtiene una línea recta (dado que se calcula con los residuales y no con los errores reales). Por lo tanto, sólo deben tomarse en cuenta las desviaciones muy grandes. Si existen saltos en dicha gráfica y son cercanos a los periodos s , $2s, 3s$, entonces se modelaron, posiblemente, los efectos estacionales.

⁸ Especialmente en los modelos estacionales

Algunas veces, el análisis de los residuales sirve para encontrar parámetros no especificados que deben ser agregados al modelo.

1.4.2.2 Sobre y Subespecificación de Modelos

Algunas veces necesitamos mejorar el ajuste del modelo quitando o poniendo parámetros. Cuando se incluye un parámetro adicional con el fin de mejorar el modelo existe el riesgo de redundar parámetros. Esta redundancia ocurre cuando se agrega, al mismo tiempo, un parámetro autorregresivo y otro de medias móviles. Para evitar ello, se debe de agregar un parámetro a la vez. Para saber que parámetro se debe de agregar (AR o MA) nos podemos ayudar con el análisis de los residuales como se mencionó con anterioridad. El agregar parámetros permite si el modelo contiene los parámetros apropiados.

Los parámetros excesivos o redundantes pueden detectarse usando el estadístico t de los coeficientes y la correlación entre los estimadores:

$$T = \frac{\beta_j - \beta_j}{\sigma_{\beta_j}} \quad \text{donde} \quad \begin{cases} \beta = (\beta_1, \beta_2, \dots, \beta_{p+q}) \\ \beta_j \text{ es el estimador de } \beta_j \\ \sigma_{\beta_j}^2 \text{ es la varianza estimada del estimador } \beta_j \end{cases}$$

para probar la hipótesis:

$$H_0 : B_j = 0$$

$$H_a : B_j \neq 0$$

Si el valor en tablas t -de student con un nivel de significancia de α y $(n-p-q)$ grados de libertad es mayor que T , hay evidencia de que el parámetro puede ser cero. Un parámetro insignificante es indicación de que el modelo está sobrespecificado. Si dicho parámetro es el de mayor orden, entonces el parámetro se elimina. Si el parámetro no es del de mayor orden, entonces se debe revisar las correlaciones entre los estimadores para saber cuál parámetro se debe de eliminar. Si la correlación es alta entre dos estimadores entonces, existe la posibilidad de que alguno de los dos se tenga que eliminar. Por último,

si el parámetro no es de mayor orden, pero está fuertemente correlacionado con el de mayor orden, se debe de evaluar el modelo sin el parámetro de mayor orden.

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 (Ver expresión 1.3.10).

1.4.3 Criterio(s) de Selección de Modelos

En series de tiempo, algunas veces existen varios modelos que son adecuados para representar a un conjunto de datos. A veces no es fácil decidir cuál de todos los modelos se debe de usar. Por ello, se han desarrollado numerosos criterios para comparar y facilitar la selección de modelos. Los principales criterios de selección de modelos son:

a) Criterio de Parsimonia

Este criterio está basado en el principio de parsimonia: "Escoger el modelo más simple y que, al mismo tiempo, describa o represente mejor el fenómeno". El modelo más simple será aquel con menor número de parámetros estimados. Este Principio es recomendado por Tukey (1967) y Box y Jenkins (1976)

b) Criterios Akaike (AIC y BIC)

Akaike (1973,1974a) introdujo un criterio de información AIC (Akaike's Information Criterion) definido como:

$$AIC(M) = n \ln \hat{\sigma}_a^2 + 2M$$

donde

n es el número de datos de la serie

$M = p + q$ (número de parámetros del modelo)

$\hat{\sigma}_a^2$ es la varianza estimada del modelo

De los modelos adecuados que se tengan, se escoge aquel cuyo AIC sea mínimo. Este criterio es una herramienta usada en muchos programas de cómputo para series de tiempo.

Akaike (1978,1979) ha desarrollado una extensión Bayesiana del AIC, llamada BIC (Bayesian Information Criterion) :

$$BIC(M) = n \ln \hat{\sigma}_a^2 - (n - M) \ln \left(1 - \frac{M}{n} \right) + M \ln n + M \ln \left[\frac{\hat{\sigma}_z^2 - 1}{\hat{\sigma}_a^2} \right] / M$$

donde

n es el número de datos de la serie

$$M = p + q$$

$\hat{\sigma}_a^2$ es la varianza estimada del modelo

$\hat{\sigma}_z^2$ es la varianza estimada de la serie

Obviamente, aquel modelo cuyo BIC sea el mínimo será el modelo escogido.

c) Criterio Schwartz (SBC)

El SBC (Schwartz's Bayesian Criterion) es muy similar al BIC de Akaike. El SBC está definido por la siguiente expresión (Schwartz, 1978):

$$SBC(M) = n \ln \hat{\sigma}_a^2 + M \ln n$$

d) Criterio Parzen (CAT)

El CAT (Criterion for Autoregressive Transfer functions) fue sugerido por Parzen (1977). Este criterio, a diferencia de los anteriores, sólo se puede usar si los modelos a comparar son modelos AR(p). El orden óptimo de p será aquel cuyo CAT(p) sea mínimo:

$$CAT(p) = \begin{cases} -\left(1 + \frac{1}{n}\right) & p = 0 \\ \frac{1}{n} \sum_{j=1}^p \frac{1}{\hat{\sigma}_j^2} - \frac{1}{\hat{\sigma}_p^2} & p = 1, 2, 3, \dots \end{cases}$$

donde

n es el número de observaciones de la serie

$\hat{\sigma}_j^2$ es la varianza estimada del modelo AR(j) que se ajusta a la serie

e) Criterio basado en errores de pronóstico.

Si el principal propósito del modelo deseado es pronosticar valores, entonces este criterio sería uno de los más útiles. Para conocer el error de pronóstico e_f del modelo se calcula la expresión:

$$e_f = Z_{n+f} - \hat{Z}_n(f)$$

donde:

n es el número de datos de la serie

f es el número que indica el periodo de tiempo futuro

Además, se deben comparar las siguientes estadísticas:

- Porcentaje de Error Medio (Mean Percentage Error, MPE)

$$\text{MPE} = \left(\frac{1}{M} \sum_{f=1}^M \frac{e_f}{Z_{n+f}} \right) 100\%$$

- Error Cuadrático Medio (Mean Square Error, MSE)

$$\text{MSE} = \frac{1}{M} \sum_{f=1}^M e_f^2$$

- Error Absoluto Medio (Mean Absolute Error, MAE)

$$\text{MAE} = \frac{1}{M} \sum_{f=1}^M |e_f|$$

- Porcentaje de Error Absoluto Medio (Mean Absolute Percentage Error, MAPE)

$$\text{MAPE} = \left(\frac{1}{M} \sum_{f=1}^M \left| \frac{e_f}{Z_{n+f}} \right| \right) 100\%$$

Se escogerá aquel modelo que, en general, tenga el valor mínimo de estas estadísticas, así como, su e_f sea pequeño.

1.4.4 Pronósticos a partir de un modelo seleccionado.

Uno de los principales usos de las series de tiempo es el pronóstico de valores. Esto es posible si ya se tiene seleccionado un modelo que represente de forma adecuada la serie.

En el tiempo $t=n$, denominado origen, se tienen $Z_n, Z_{n-1}, Z_{n-2}, \dots$ observaciones y se desea conocer un valor Z_{n+f} pronosticado⁹ como resultado de una combinación lineal de los valores anteriores. Es decir, $(Z_{n+f} | Z_n, Z_{n-1}, Z_{n-2}, \dots)$.

Anteriormente se definió el error al hacer un pronóstico como:

$$e_f = Z_{n+f} - \hat{Z}_n(f)$$

entonces su error cuadrático medio sería:

$$E[(e_f)^2] = E[(Z_{n+f} - \hat{Z}_n(f))^2]$$

donde

n es el origen

f es horizonte

Algunos autores no manejan la expresión anterior. En su lugar utilizan una función de costo cuadrática definida por $C(e) = \alpha e^2$. Esta función es equivalente a la anterior cuando $\alpha=1$. Esta función tiene la propiedad de que la solución para el pronóstico óptimo es la media de la distribución del pronóstico.

Por lo tanto, los pronósticos para $t=n+f$ y un proceso ARMA(p,q) estacionario e invertible se puede calcular con la expresión:

$$\begin{aligned} \hat{Z}_n(f) &= \hat{Z}_{n+f} \\ &= \phi_1 Z_{n+f-1} + \phi_2 Z_{n+f-2} + \dots + \phi_p Z_{n+f-p} \\ &\quad + e_f - \theta_1 e_{n+f-1} - \theta_2 e_{n+f-2} - \dots - \theta_q e_{n+f-q} \end{aligned}$$

donde:

⁹ El intervalo f se le denomina horizonte ($f > 0$)

ϕ_a y θ_b = valores estimados del modelo $1 \leq a \leq p$, $1 \leq b \leq q$

Z_{n+k} = valor observado $si k \leq 0$

Z_{n+k} = valor pronosticado $si 0 < k \leq f$

e_{n+k} = valor residual $si k \leq 0$

$e_{n+k} = 0$ (valor esperado) $si 0 < k \leq f$

Los pronósticos se pueden actualizar tan pronto como se conoce el punto siguiente

Z_{n+1} . El ajuste se puede realizar de dos formas:

- Secuencial: En este caso se reestiman los parámetros del modelo incluyendo el nuevo punto.
- Adaptativo: Se conservan las estimaciones de los parámetros, pero el origen n se cambia.

TESIS CON
FALLA DE ORIGEN

CAPÍTULO II

El Análisis de Intervención y su aplicación en las series de tiempo

El análisis de intervención es una metodología que resulta útil para estudiar series de tiempo que son afectadas por factores externos, tales como: políticas de empresa, promociones, huelgas, etc. En este capítulo, se explica esta metodología, los principales tipos de intervenciones, así como, las posibles respuestas o impactos que puede tener una serie a una intervención dada. Además, se mencionan algunos estudios relevantes en donde se ha utilizado el análisis de intervención.

II El Análisis de Intervención y su aplicación en las series de tiempo.

2.1 ¿Qué es el Análisis de Intervención?

Las series de tiempo son frecuentemente afectadas por ciertos eventos externos como promociones, los días feriados, huelgas y cambios de políticas. A este tipo de eventos externos a las series se les llama intervenciones. En algunas ocasiones se sabe con precisión en que momento ocurrirá la Intervención, pero en otros casos se desconoce.

A veces, al realizar el análisis de una serie de tiempo, se puede desear medir la influencia de las intervenciones, ya sea porque se juzgue importante su influencia o porque se tema que un fenómeno anormal, (errores de captura, eventos desconocidos, etc.) afecten la estimación de los parámetros y los resultados.

El análisis de intervención es una metodología usada para modelar el impacto de una intervención, la cual interrumpe el comportamiento estable de una serie de tiempo. Esas intervenciones pueden ser una nueva regulación (nuevas normas), la construcción de una planta, reducción de precios, etc. Esta metodología se puede aplicar en distintos campos como lo muestra el cuadro 2.1:

Ejemplo	Intervención dada	Efecto esperado
Contaminación en el aire o en el agua	Nuevas normas	Reducción de la contaminación
Enfermedad en un área específica	Campaña de vacunación	Reducir el número de casos
Negocios	Las promociones	Incrementar la demanda
Negocios	Descuentos	Incrementar las ventas

Cuadro 2.1 Algunos ejemplos de intervenciones

Sabiendo que una intervención de cualquier tipo y en cualquier problema, ocurre en un cierto tiempo T , el analista debe preguntarse: ¿existe evidencia de un cambio en la serie de tiempo? Y si es así, ¿cuánto o en qué medida ha cambiado la serie? ; si la intervención es planeada, ¿ha sido exitosa y se ha obtenido lo que se deseaba?

2.1 ¿Qué es el Análisis de Intervención?

En el análisis de series de tiempo (Box et al. 1994), una intervención es una entrada que indica la presencia o ausencia de un evento. Una intervención causa que una serie de tiempo se desvíe de su patrón de evolución esperado.

Las intervenciones tienen las siguientes características:

- Ocurren en un tiempo específico:

El tiempo de la intervención comienza en el instante en el cual causa una desviación.

- Tienen una duración específica:

La duración de la intervención es el periodo en el cual se causa dicha desviación.

- Son de algún tipo en particular:

El tipo de intervención es la manera de influir sobre la serie de tiempo.

- Producen una respuesta:

La respuesta a la intervención (efecto) es la manera en que la serie se comporta después de la intervención.

Uno podría pensar en realizar el análisis de la serie, antes y después de la intervención, utilizando otros métodos tales como la prueba t de Student. Con este método se estimaría y probaría un cambio en la media de la serie¹⁰. Por ejemplo, se propondría la siguiente prueba de hipótesis, con intervalos de confianza del 95% ($\alpha=0.05$):

$$H_0: \mu_1 - \mu_2 = \delta$$

$$H_1: \mu_1 - \mu_2 > \delta$$

ó

$$H_2: \mu_1 - \mu_2 < \delta$$

¹⁰ También se podrían realizar pruebas tomando en cuenta las desviaciones estándar, antes y después de la intervención.

Análisis de Intervención: Aplicación a un problema ambiental.

donde $\delta \approx 0$ (diferencia pequeña); μ_1 representa el valor de la media de la serie antes de la intervención, y μ_2 su valor posterior a la misma.

Sin embargo, sería válido utilizar la prueba t de Student, siempre y cuando las medias μ_1 y μ_2 , tuvieran una distribución normal, con varianza constante y, principalmente, fueran independientes, cosa que no siempre se cumple.

El análisis de intervención puede verse como una extensión del análisis de los procesos ARIMA, ya que un modelo para una serie Z_t que contenga los efectos de una intervención, puede expresarse como:

$$Z_t = \varepsilon_t + N_t$$

donde N_t es un modelo ARIMA, estacionario e invertible, mientras que ε_t es una función que permite representar los efectos de la intervención. Dicha función debe de permitir apreciar los cambios ocurridos en la serie no sólo en el momento en el que ocurrió, sino en los posteriores y atribuibles a ella. Más adelante, en la expresión (2.3.6), se muestra con detalle el significado de estas variables.

2.2 Tipos de Intervenciones en las series de tiempo

Existen dos **tipos** comunes de intervención:

- a) Representa una intervención que ocurre en el tiempo T y permanece de ahí en adelante. Esta intervención es una **función escalón** (step function)

$$S_t^{(T)} = \begin{cases} 0, & t < T, \\ 1, & t \geq T \end{cases} \quad (2.2.1)$$

- b) Representa una intervención que tiene lugar únicamente en un periodo de tiempo, es decir, es una **función pulso** (pulse function)

$$P_t^{(T)} = \begin{cases} 1, & t = T, \\ 0, & t \neq T \end{cases} \quad (2.2.2)$$

Esta segunda función puede ser producida haciendo una diferencia en la función escalón, es decir:

$$P_t^{(T)} = S_t^{(T)} - S_{t-1}^{(T)} = (1-B)S_t^{(T)}$$

Por lo tanto, una intervención puede ser representada igualmente bien con una función escalón o con una función pulso. El uso de una forma en específico debe basarse en la conveniencia de su interpretación.

2.3 Respuestas Posibles a una Intervención.

Hay muchas posibles *respuestas*¹¹ a una intervención escalón o pulso. Algunas comúnmente encontradas son:

1. Un impacto fijo de una intervención es sentido b periodos después de la intervención.

Entonces, dependiendo del tipo de la intervención, el impacto es:

$$wB^b S_t^{(T)} \quad (2.3.1)$$

o bien,

$$wB^b P_t^{(T)} \quad (2.3.2)$$

2. Un impacto de una intervención es sentido b periodos después de la intervención, pero la respuesta es gradual. Para una función escalón se tiene:

¹¹ Comúnmente se les denomina respuestas (response), aunque algunos autores también las llaman salidas (output) o impactos.

$$\frac{\omega B^b}{(1-\delta B)} S_t^{(T)}, \quad (2.3.3)$$

y para una función pulso,

$$\frac{\omega B^b}{(1-\delta B)} P_t^{(T)} \quad (2.3.4)$$

donde $0 \leq \delta \leq 1$. Para $\delta=0$, (2.3.3) y (2.3.4) se reducen a (2.3.1) y (2.3.2) respectivamente. Si $\delta=1$, el impacto se incrementa linealmente sin límite. Para la mayoría de los casos, se tiene $0 < \delta < 1$, y la respuesta es gradual. b indica cada cuantos periodos de tiempo se repite el efecto de intervención¹². δ mide el comportamiento del efecto permanente de la intervención.

Por lo general, y tomando en cuenta lo anterior, se encuentran tres tipos de impactos o respuestas:

a) Impacto Brusco Permanente:

Este impacto implica que la media general de la serie de tiempo cambia después de la intervención; el cambio general es denotado por ω . La ecuación (2.3.1) indica este impacto (Figura 2.1 (a)).

b) Impacto Gradual Permanente:

Este impacto implica que el incremento (o decremento) debido a la intervención es gradual, y que el impacto permanente definitivo llega a ser evidente sólo después de algún tiempo. La ecuación (2.3.3) indica este impacto. Si es $\delta=0$, la cantidad definitiva permanente de impacto podrá ser evidente después de algunas observaciones, y si $\delta=1$ este impacto sólo podrá ser evidente después de muchas observaciones. Para la mayoría de los casos, se tiene $0 < \delta < 1$, y la respuesta es gradual (Figura 2.1 (c)).

¹² Con frecuencia a la variable b se le denomina grado.

c) Impacto Brusco Temporal

Implica incremento (o decremento) brusco inicial debido a la intervención, el cual después decae (o sube), fuera de cambiar permanentemente la media de la serie. La ecuación (2.3.4) indica este impacto. Si δ es cercano a cero decae (o sube) rápidamente. Sucede lo contrario si $\delta \approx 1$. Si $0 < \delta < 1$ el cambio es gradual. (Figura 2.1 (d)).

Cuando $\delta=1$ en la ecuación de la figura 2.1(c) se deriva 2.1(e). A esta última se le conoce como respuesta "rampa". Algunas veces, las intervenciones se repiten de forma estacional, por ejemplo, efectos de la primavera o de invierno, etc. La figura 2.1(f) muestra un tipo de Intervención estacional.

Se puede ver que varias respuestas se pueden producir a través de diferentes combinaciones de entradas de escalón o impulso. Generalmente, una respuesta puede ser representada como una función racional:

$$\frac{\omega(B)B^b}{\delta(B)} I_t \quad (2.3.5)$$

donde:

$$\omega(B) = \omega_0 - \omega_1 B - \dots - \omega_s B^s$$

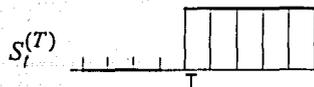
$$\delta(B) = 1 - \delta_1 B - \dots - \delta_r B^r \quad \text{son polinomios en } B$$

- b (de nueva cuenta) indica cada cuantos periodos de tiempo se repite el efecto de intervención.
- Los pesos ω_j en el polinomio $\omega(B)$ representan, la mayoría de las veces, los efectos iniciales esperados de la intervención.
- El polinomio $\delta(B)$ mide el comportamiento del efecto permanente de la intervención.
- I_t representa el tipo de intervención, escalón o pulso.

A continuación se muestran gráficas de las principales respuestas a una intervención dada:

Entradas o tipos de intervención

Función escalón



Función Pulso



Salidas o Respuestas

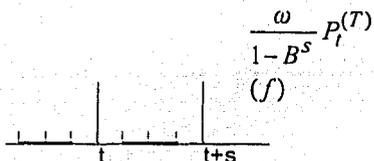
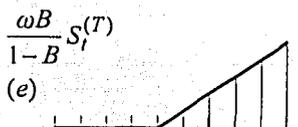
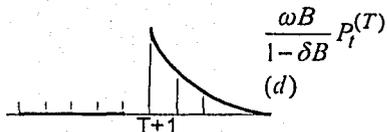
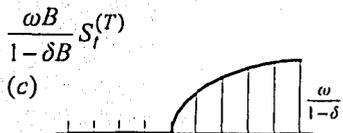
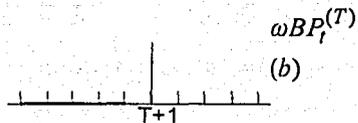


Figura 2.1 Tipos de Intervenciones y sus principales funciones de respuesta
Fuente: Box y Tiao, 1975; Wei, 1990.

Si no se conoce el tipo de respuesta, puede ser determinado analizando los datos históricos.

2.3 Respuestas posibles a una Intervención

Se asume (Box y Tiao, 1975) que las raíces del polinomio $\alpha(B)=0$ (ver expresión 2.3.5) están sobre o afuera del círculo unitario. La raíz unitaria representa un impacto que se incrementa linealmente, y la raíz afuera del círculo unitario representa un fenómeno que tiene una respuesta gradual. Así pues, se tiene el siguiente modelo general, que combina el proceso ARIMA y múltiples intervenciones¹³:

$$Z_t = \sum_{j=1}^k \frac{\omega_j(B)B^{b_j}}{\delta_j(B)} I_{jt} + \frac{\theta(B)}{\psi(B)} e_t \quad (2.3.6)$$

La forma $\omega_j(B)B^{b_j}/\delta_j(B)$ para la j -ésima intervención es postulada basándose en la respuesta esperada (salida), dado que se conoce la intervención (entrada). El principal propósito de los modelos de intervención es medir los efectos de las intervenciones. La serie de tiempo libre de intervenciones es llamada *serie de ruido*¹⁴ y se denota con N_t y su modelo es conocido como *modelo de ruido*: $[\theta(B)/\psi(B)]e_t$. Este último es usualmente reconocido usando el procedimiento de identificación para el modelo univariado, mostrado en el capítulo I, basado en la serie de tiempo Z_t antes de la fecha de intervención. Una vez que se completa y se obtiene un modelo adecuado, se pueden hacer inferencias apropiadas acerca de la intervención. En este estudio se procederá de esta manera. En caso de que el modelo obtenido no sea el adecuado, se deben realizar modificaciones y repetir el proceso de estimación y diagnóstico.

Se pueden obtener los impactos de las intervenciones mediante otras aproximaciones que manejan la serie de tiempo en diferente forma. Esto es:

¹³ Usar un número excesivo de intervenciones en un solo modelo no es recomendable. En un caso extremo, se puede usar una intervención en cada punto de tiempo, lo cual definiría otra serie.

¹⁴ Noise series

Análisis de Intervención: Aplicación a un problema ambiental.

- a) Tomar datos antes y después de la intervención, pero excluyendo aquellos dentro del intervalo $[t_0, t_0+k]$, el cual representa un "cordón sanitario". En éste, los efectos de la intervención son juzgados y se trabaja con ellos.
- b) Estimar los residuales

$$e_t = Y_t - \frac{\omega(B)B^b}{\delta(B)} I_t$$

Ignorando cualquier error de estructura, y graficar la función de autocorrelación.

- c) Comparación de resultados (parámetros obtenidos). En este caso se realiza el análisis en tres etapas: la de la serie completa, la de los datos antes de la intervención y la de los datos posteriores a la intervención. Una vez obtenidos los resultados de cada etapa, se comparan.
- d) Usar la serie completa, ignorando los efectos de la intervención.

Una vez completo el modelo que representa a la serie Z_t , es decir, se han calculado todos sus parámetros y se han desechado aquellos que son no son estadísticamente significativos, se puede hacer una evaluación de estos resultados. En dicha evaluación se sabrá si la(s) intervención(es) tiene(n) un efecto sobre la serie de tiempo o no.

TESIS CON
FALLA DE ORIGEN

2.4 Antecedentes del tema y su importancia: estudios relevantes.

Frecuentemente, los datos están organizados en forma de serie de tiempo y tienen, por tanto, un valor potencial. Por esta razón, no es difícil imaginar que existen muchos campos en donde se puede aplicar la metodología Box-Jenkins¹⁵ y el análisis de intervención. Por ejemplo; economía, demografía, ecología, etc. Con esta metodología se construyen modelos matemáticos que representan el comportamiento y las principales características de las series¹⁶. Los modelos pueden dar muchas respuestas de los problemas planteados. Con ellos se puede saber, en muchos casos, qué factores afectan una serie de tiempo y cuándo sucede ello. De esta manera, el analista puede interpretar estos datos y dar una explicación. Además, con este tipo de modelos se pueden realizar pronósticos de los fenómenos que se estudian. Un modelo es mejor, siempre y cuando describa con mayor exactitud a los datos reales.

Se han realizado trabajos con modelos matemáticos, como series de tiempo, que intentan dar explicación o respuesta a algunos fenómenos naturales tales como los meteorológicos y el clima, así como algunas aplicaciones en Astronomía. Por ejemplo, Eric D. Feigelson (1997) ha aplicado las series de tiempo en problemas de astronomía tales como los ciclos de las manchas solares, estrellas binarias, pulsares y otros. Quizás pueda confundir esto último, pero la razón que se desea dar a conocer es que, fenómenos que tal vez parecen muy difíciles de predecir, se pueden modelar y poder obtener pronósticos.

El análisis de intervención en una serie de tiempo ha sido estudiado en otros contextos por diferentes investigadores. Box y Tiao (1975) fueron los pioneros, que incorporaron la posibilidad de un cambio en el proceso de una serie de tiempo debido a

¹⁵ En el Capítulo 1 se presentan las principales características, que debe tener una serie de tiempo para poder utilizar Box-Jenkins.

¹⁶ Muchas veces es preferible estudiar fenómenos de la realidad como modelos matemáticos

una o varias intervenciones. Describieron los tipos de intervenciones más comunes en las series de tiempo. Estos investigadores aplicaron esta metodología y modelaron el problema de contaminación del aire en Los Angeles, E.U., así como, el cambio de políticas económicas tomadas por la Administración del presidente Nixon. En el primer caso, modelaron el impacto que tuvo la aplicación de normas para controlar emisiones contaminantes en el Estado de California; notaron que, debido a tal intervención, se había reducido el nivel de contaminación. La Figura 2.1 (a) muestra el tipo de respuesta que utilizaron para analizar este caso. En el segundo caso, analizaron el impacto que tuvo sobre la inflación, la aplicación de controles en el índice de precios al consumidor. Nixon aplicó estos controles en dos fases; Box y Tiao notaron que en la primera fase la intervención había tenido efectos, pero en la segunda fase ese efecto se perdía.

Keum-Rok Yoo (1995) notó que la política fiscal, además de tener un importante efecto sobre la economía, también tiene una influencia importante en el comportamiento político de los ciudadanos. Algunos gobiernos actuales tratan de manipular las políticas fiscales con el fin de alcanzar metas políticas, tales como reelecciones. En este artículo, se analiza y compara el impacto producido por las políticas fiscales sobre las elecciones en Corea, Estados Unidos y Japón, tres países con diferentes sistemas políticos entre 1960 y 1980. Se utiliza un modelo de intervención ARIMA para analizar las administraciones manipuladas con políticas fiscales para obtener reelección.

Montgomery y Weatherby (1980) también usaron, de forma exitosa, el análisis de intervención para demostrar el impacto que tuvo el embargo de petróleo árabe en 1973, en el consumo de energía en los Estados Unidos. Demostraron que el embargo indujo un cambio permanente en el nivel de consumo de energía. El efecto del embargo que encontraron, fue la reducción en un 7% del consumo de energía.

2.4 Antecedentes del tema y su importancia: estudios relevantes

Incluso, existe un estudio que utiliza el análisis de intervención para saber si la causa de nacimientos en Agosto de 1966 en Nueva York, fue un apagón que sucedió en Noviembre de 1965 (Izenman y Zabell, 1981). El resultado de este estudio fue que el nivel de nacimientos de esa ciudad no se incrementó, en forma significativa, debido a dicho apagón¹⁷.

En otros casos, el análisis de intervención se ha utilizado para estudiar la eficiencia de un mercado. Se estudia la respuesta dinámica ante la imposición de precios y su proceso de ajuste (El-Sheshai y Morin, 1984); Wichern y Jones (1977) analizan el impacto de distintos disturbios en el mercado; Deakin (1976) examina el comportamiento de las respuestas obtenidas y que han sido resultado de divulgar situaciones contables (en este caso esa divulgación se toma como la intervención).

Bhattacharyya y Layton (1979) mostraron que no sólo en economía se puede utilizar el análisis de intervención, ya que ellos estudiaron la efectividad de la legislación del uso obligatorio del cinturón de seguridad en Australia. Cohen (1981) realizó un estudio semejante al anterior. Él evaluó los beneficios de incorporar estándares en los botes y en productos de salvamento. Dichos estándares incluían, por ejemplo, suficiente flotación de los botes, no obstante que tuvieran alguna cantidad de agua en su interior. Cohen analizó el impacto sobre la seguridad, que se logró al implantar estas normas, en términos de vidas salvadas. Este investigador utilizó el número de muertes mensuales en botes como datos (serie de tiempo), que le fueron proporcionados por la Guardia Costera de los Estados Unidos. Lassarre y Tan (1982) estudiaron el efecto de medidas de seguridad sobre la frecuencia y gravedad de los accidentes de tránsito en Francia.

Las promociones o campañas publicitarias constituyen otra intervención en las series de tiempo, que ha sido estudiada. La figura 2.1 (d) muestra la respuesta que se espera tener al intervenir una serie con una promoción, de poca duración. (ω indica el

¹⁷ En este caso, el apagón se toma como una intervención en el nivel de nacimientos de Nueva York en 1965

incremento en ventas inmediatamente después de la intervención (promoción) y δ representa el grado de decaimiento).

Varias respuestas pueden ser obtenidas si se combinan diferentes entradas, es decir, si existen varias tipos de intervenciones. Por ejemplo:

$$\frac{\omega_0 B}{(1-\delta B)} P_t^{(T)} + \omega_1 B S_t^{(T)} \quad (2.4.1)$$

Tomando en cuenta

$$P_t^{(T)} = S_t^{(T)} - S_{t-1}^{(T)} = (1-B) S_t^{(T)}$$

Se puede reescribir 2.4.1

$$\left[\frac{\omega_0 B}{(1-\delta B)} + \frac{\omega_1 B}{(1-B)} \right] P_t^{(T)} \quad (2.4.2)$$

La expresión (2.4.2) se ha utilizado para representar el impacto de una intervención como promoción en ventas, siempre y cuando $\omega_0 > 0$ y $\omega_1 > 0$. Ese impacto se muestra en la figura 2.2 (a). Si en la expresión (2.4.2) $\omega_0 < 0$ y $\omega_1 < 0$, se puede representar el impacto del incremento de un precio o un impuesto. La figura 2.2 (b) muestra este impacto.

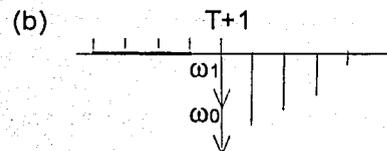
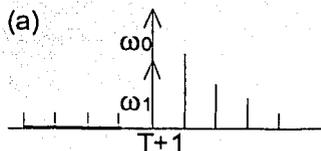


Figura 2.2 Respuestas a entradas combinadas
Fuente: Box y Tiao, 1975. Wei, 1990.

2.4 Antecedentes del tema y su importancia: estudios relevantes

Como se puede ver las series de tiempo y el análisis de intervención son herramientas muy útiles en diferentes campos de estudio. Es importante hacer saber, que se han podido realizar estos estudios porque se cuenta con datos registrados. Muchas veces los datos no están al alcance de los investigadores; incluso puede ser que no se tengan registros.

TESIS CON
FALLA DE ORIGEN

CAPÍTULO III

EL PROBLEMA AMBIENTAL

En este capítulo se da el planteamiento del problema y los límites de estudio. Se describen las condiciones ambientales de la Zona Metropolitana del Valle de México (ZMVM), así como, el programa "Hoy No Circula" (HNC). El índice de contaminación por Ozono constituye la serie de tiempo estudiada. A partir de esta, se obtiene un modelo matemático que represente sus principales características y que tome en cuenta la intervención que constituye el programa HNC.

TESIS CON
FALLA DE ORIGEN

3.1 PLANTEAMIENTO DEL PROBLEMA.

3.1.1 Límites del estudio

Los siguientes puntos son límites propuestos para este trabajo y es importante tomarlos en cuenta:

- En este trabajo se utilizan los modelos ARIMA univariados para describir las características del fenómeno.
- La serie de tiempo está constituida por datos de contaminación ambiental por Ozono en la Zona Metropolitana de la Ciudad de México (ZMCM). Estos datos corresponden a los registros máximos diarios del contaminante, y fueron obtenidos de la RAMA, y se muestran en el Anexo I.
- Los registros diarios corresponden al periodo 01/Ene/1986 a 31/Oct/1999
- Se obtendrá un modelo matemático de series de tiempo que represente los datos mencionados, así como, tome en cuenta la intervención que constituye la aplicación del programa Hoy No Circula.
- No se utilizarán otros métodos de pronósticos más que el modelo de series de tiempo.

3.1.2 Condiciones actuales de la ZMVM

La tarea de mejorar la calidad del aire en la ZMVM no es fácil. Dicha calidad es el reflejo de profundas implicaciones estructurales, funcionales y territoriales, vinculadas con la forma en que se usa y se maneja la cuenca atmosférica. La calidad del aire de una cuenca atmosférica depende, en primera instancia, del volumen de contaminantes emitidos, del comportamiento físico-químico de éstos y de la dinámica meteorológica que determina su dispersión, transformación y remoción en la atmósfera.

El Valle de México posee una serie de características fisiográficas y climáticas únicas que contribuyen de manera determinante en la severidad de los problemas de contaminación:

TESIS CON
FALLA DE ORIGEN

3.1 Planteamiento del problema

- Se encuentra a una altura de 2240 metros, por lo que el contenido de oxígeno del aire es menor que al nivel del mar. Esto hace que los procesos de combustión interna sean menos eficientes y produzcan por tanto una mayor cantidad de contaminantes.
- Está rodeado por las montañas de las sierras del Ajusco, Chichinautzin, Nevada, Las Cruces, Guadalupe y Santa Catarina, las que constituyen una barrera física natural para la circulación del viento, impidiendo el desalojo del aire contaminado fuera del Valle.
- Presenta con frecuencia inversiones térmicas que provocan el estancamiento de los contaminantes. Por las mañanas, la capa de aire que se encuentra en contacto con la superficie del suelo adquiere una temperatura menor que las capas superiores, por lo que se vuelve más densa y pesada. Las capas de aire que se encuentran a mayor altura y que están relativamente más calientes actúan, entonces, como una cubierta que impide el movimiento ascendente del aire contaminado.
- Recibe una abundante radiación solar debido a su latitud de 19° N, lo que hace que su atmósfera sea altamente fotorreactiva. En presencia de la luz solar, los hidrocarburos y los óxidos de nitrógeno reaccionan fácilmente para formar ozono y otros oxidantes.

3.1.3 El programa "Hoy No Circula"

El Programa Hoy No Circula (HNC) se implementó de manera obligatoria el 20 de Noviembre de 1989 en la ZMVM. Primero, se puso en marcha de manera temporal para el periodo que comprendía el Invierno 88-89, pero después se anunció su permanencia.

Actualmente, dicho programa se realiza con base en un marco normativo federal y local, que cuenta con procedimiento y normas técnicas claramente definidas. Por ejemplo:

- Los vehículos en la ZMVM están dentro de un programa permanente restricción vehicular para el control de emisiones de contaminantes HNC.
- Los automóviles, con base al último dígito de la placa, están distribuidos por colores a lo largo de toda la semana, quedando así sin circular un día a la semana¹⁸.
- El HNC se aplica de Lunes a Viernes de las 5 hrs. a las 22 hrs.

Asimismo, existen procedimientos que permiten supervisarlo en forma permanente y sistemas de información que son los instrumentos necesarios para su control y adecuado funcionamiento. El objetivo principal del HNC, es regular las emisiones contaminantes de vehículos automotores que circulan en la ZMVM.

En caso de que se presenten altos índices de contaminación se llevará a cabo la Fase I del Programa de Contingencias Ambientales Atmosféricas. En esta fase se activa el llamado "Doble Hoy No Circula". En dicha fase dejan de circular alternadamente las terminaciones de placa en números pares, nones y cero, de aquellos vehículos que tengan la calcomanía de verificación número 2.

¹⁸ Se aplica a los vehículos con Calcomanía 1 y Calcomanía 2. No aplica a vehículos con calcomanía 0 (generalmente, coches nuevos)

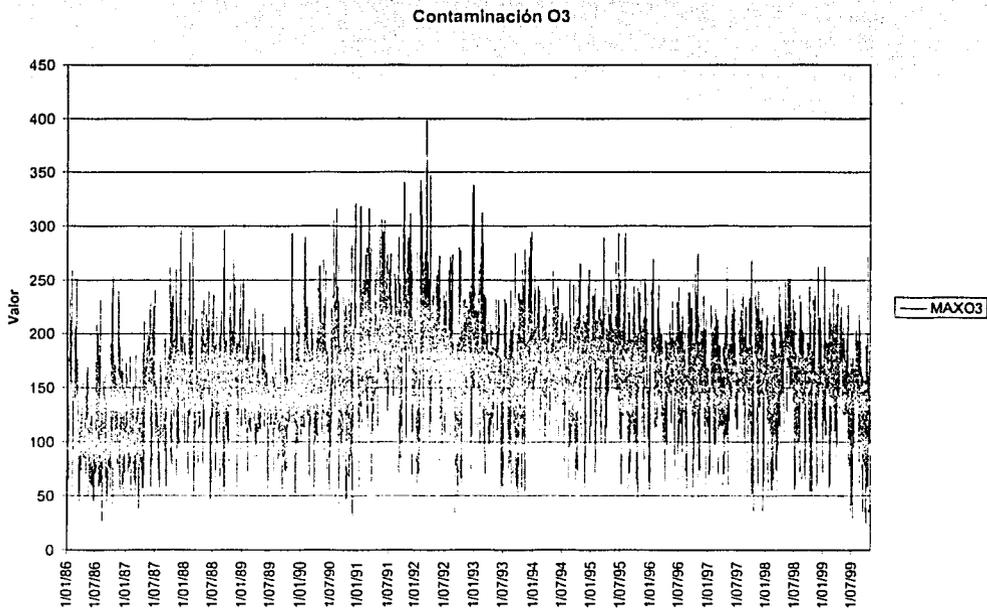
3.1.4 Índices de Ozono en la ZMVM

La evolución de los niveles de ozono es el resultado de la interacción de un conjunto de factores en donde destacan el comportamiento de sus precursores y la evolución del clima. Parece existir también una relación con la evolución y composición del parque vehicular.

El ozono es un contaminante que no se emite en los escapes o chimeneas, sino que se forma en la atmósfera a partir de reacciones muy complejas.

Existen dos ciclos generales de reacciones fotoquímicas en la formación del ozono troposférico, en los que participan el oxígeno molecular y dos de los denominados precursores del ozono: los óxidos de nitrógeno (NOx) y los hidrocarburos (HC).

Como se aprecia en la gráfica 3.1.4.1, las concentraciones de ozono alcanzaron en 1991 y 1992 los niveles más críticos registrados hasta ahora; se alcanzaron valores de hasta 398 IMECA (punto más alto en la gráfica).



Análisis de Intervención: Aplicación a un problema ambiental.

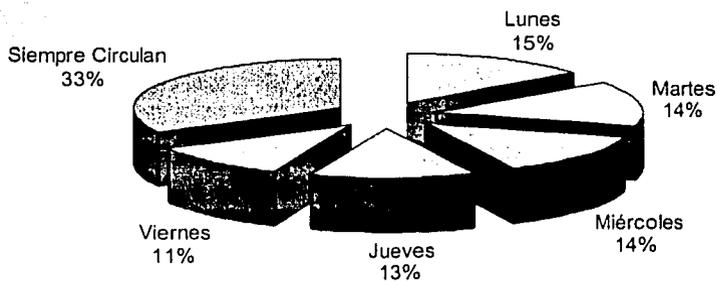
Según datos del gobierno de la ciudad, los vehículos automotores son responsables del 71% de los óxidos de nitrógeno emitidos a la atmósfera¹⁹.

Una estimación del parque vehicular de la ZMVM por parte del gobierno de la ciudad es de 3,450,000 unidades²⁰. El resultado directo de la aplicación del HNC, es la limitación a la circulación de alrededor del 14 % del parque vehicular total, cabe señalar que dicho porcentaje varía según el día de la semana. (Cuadro 3.1.4.1 y Gráfica 3.1.4.2)

	Lunes	Martes	Miércoles	Jueves	Viernes
Eliminación del Programa HNC	3450000	3450000	3450000	3450000	3450000
Aplicación del Programa HNC	2945525	2952019	2969204	2998071	3062885
Vehículos que dejan de circular aplicando el programa HNC	504475	497981	480796	451929	387115

Cuadro 3.1.4.1 No. de vehículos que dejan de circular diariamente en la ZMVM debido al HNC

Es importante señalar que una de las características fundamentales del parque vehicular en la ZMVM es que aproximadamente el 38 % tiene más de 10 años de antigüedad²¹.



Gráfica 3.1.4.2 Porcentaje de vehículos que no circulan en la ZMVM aplicando el HNC

Las altas concentraciones de ozono constituyen hoy en día el principal problema de contaminación atmosférica en el Valle de México. Las concentraciones de ozono exceden frecuentemente la norma de calidad del aire, alcanzando niveles que superan en más de un 100% el límite establecido.

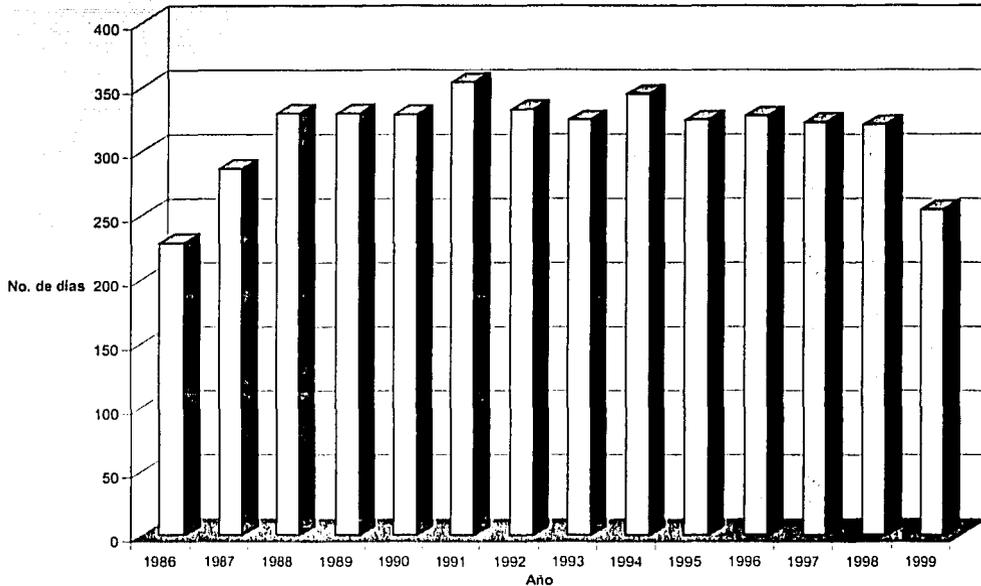
¹⁹ Secretaría del Medio Ambiente del D.F. (<http://sma.df.gob.mx>)

²⁰ Idem

²¹ Idem.

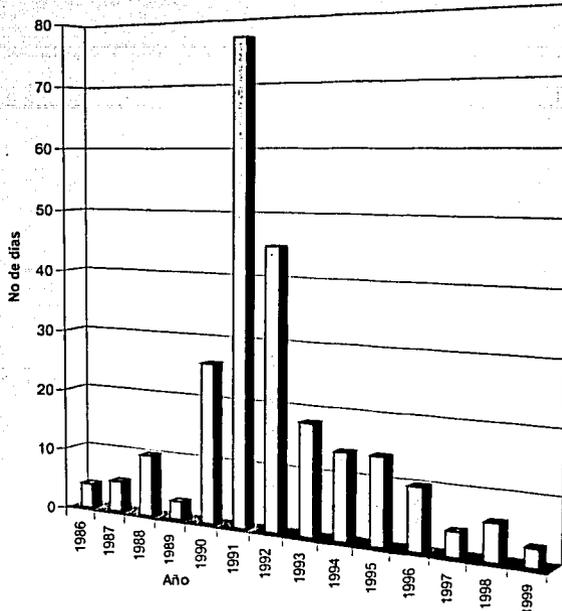
3.1 Planteamiento del problema

La Norma Internacional es de 100 puntos y lo que se recomienda es que se rebase ésta, una vez cada tres años. Actualmente, la norma de ozono se excede en más de un 90% de los días del año en la mayor parte de la Zona Metropolitana. (Ver gráfica 3.1.4.3)



Gráfica 3.1.4.3 No. de Días que se rebasó la Norma Internacional (100 puntos)

Con frecuencia se han aplicado contingencias atmosféricas con el propósito de disminuir niveles de contaminación superiores a 240 puntos IMECA (que determinan la activación de la Fase I del Programa de Contingencias Ambientales). La gráfica 3.1.4.4 muestra el número de contingencias que han sido aplicadas por cada año.



Gráfica 3.1.4.4 No. de Contingencias Ambientales

La gráfica 3.1.4.4 muestra que, a partir de 1991, el número de contingencias ambientales ha disminuido. Se afirma que es debido a las medidas tomadas por el gobierno de la Ciudad y el Estado de México. También se dice que el programa "Hoy No Circula" es una medida muy importante para la reducción de los contaminantes, principalmente óxidos de nitrógeno (precursores del ozono), y es por ello que no se puede eliminar.

En el subcapítulo 3.3 se realizará el Análisis de Intervención de este programa en la serie de tiempo representada en la figura 3.1.4.1. Se obtendrá el impacto del HNC y con ello, se podrá saber si realmente este programa ha hecho que los niveles de contaminación se reduzcan y si su aplicación sigue siendo importante.

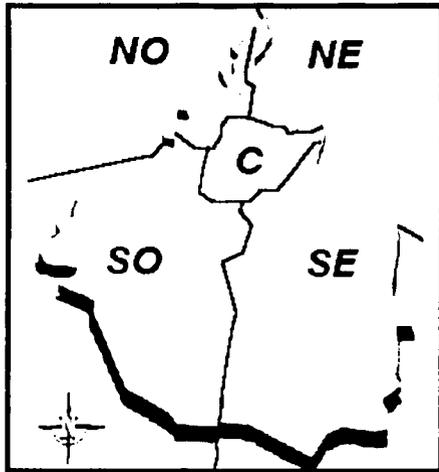
TESIS CON
FALLA DE ORIGEN

3.2 DATOS HISTÓRICOS DEL PROBLEMA.

3.2.1 La RAMA y su importancia.

Los datos históricos de la contaminación por Ozono (O_3) en la ZMVM, constituyen la serie de tiempo que se analizará. Esos datos se muestran en el Anexo I y fueron obtenidos de la Red Automática de Monitoreo Atmosférico (RAMA). La RAMA está integrada actualmente por 32 estaciones remotas de monitoreo, 21 de éstas se encuentran en el Distrito Federal y 11 en el Estado de México. Cada sitio de monitoreo cuenta con el equipamiento necesario para efectuar las mediciones de los contaminantes derivados de la actividad y uso del suelo más representativo de la región.

El diseño actual del sistema de monitoreo atmosférico de la ZMVM, que también cuenta con unidades móviles, tiene el propósito de vigilar y evaluar el estado de la calidad del aire, respecto de los contaminantes criterio y otros elementos tóxicos, así como variables meteorológicas (Cuadro 3.2.1).



3.2.1 Mapa de Zonas en el
D.F. y municipios del
Estado de México

La RAMA proporciona información sobre diversos contaminantes como son: ozono, óxidos de nitrógeno, bióxido de azufre, monóxido de carbono, ácido sulfhídrico y partículas suspendidas fracción respirable (Cuadro 3.2.2).

Símbolo	Definición	Unidades de Reporte ²²
O3	Ozono	ppm
CO	Monóxido de Carbono	ppm
SO2	Bióxido de Azufre	ppm
H2S	Ácido Sulfhídrico	ppm
NO2	Bióxido de Nitrógeno	ppm
NOX	Óxidos de Nitrógeno	ppm
PM10	Partículas Menores a 10 micras	µg/m ³

Cuadro 3.2.2 Simbología de los principales contaminantes

Esta red opera continuamente durante las 24 horas del día, todos los días del año, por lo que es posible mantener una vigilancia constante del comportamiento de los contaminantes e informar de manera oportuna la situación prevaleciente, así como poner en marcha el programa de contingencia ambiental cuando los niveles de contaminación son críticos.

La información histórica, que ha proporcionado esta red, es un elemento valioso para evaluar de manera indirecta, el impacto de las acciones de prevención y control que se han instrumentado para controlar fuentes específicas de contaminación de la ZMVM. La RAMA cuenta con equipos de medición, principalmente, analizadores de gases específicos. El principio de operación con que operan estos equipos se muestra en el cuadro 3.2.3.

CONTAMINANTE	PRINCIPIO DE OPERACIÓN
Ozono	Fotometría en el rango de ultravioleta
Óxidos de nitrógeno	Quimiluminiscencia
Bióxido de azufre	Fluorescencia pulsante
Monóxido de carbono	Espectroscopía no dispersiva por correlación de filtro gaseoso
Ácido sulfhídrico	Ionización de Flama
PM10	Atenuación de radiación beta y Balanza de Oscilación

Cuadro 3.2.3 Modo de detección de los principales contaminantes.

²² Para determinar la concentración de una sustancia química en un volumen se utilizan las partes por millón (ppm). Se divide el volumen en un millón de partes iguales. Cada millonésima parte de este volumen, correspondiente a la sustancia de nuestro interés, se considera una parte por millón de la sustancia. Las ppm se utilizan para determinar concentraciones muy pequeñas de gases en la atmósfera.; µg/m³: microgramos/metro cúbico.

3.2.2 El PSI

Anteriormente, no se contaba con un indicador de la contaminación que informara, de forma clara, a la población. Por tanto, se empezaron a proponer indicadores e índices que relacionaran las concentraciones de los contaminantes y la contaminación que se percibía.

Así pues, se propuso en E.U. el Índice de los Contaminantes Estandarizados (Pollutant Standards Index, PSI). Este un índice de tipo "máximo", que informa solamente el valor numérico más alto de todos los valores de subíndice para cada contaminante. Los valores del índice se informan a través de palabras descriptoras de la calidad del aire, tales como: bueno, satisfactorio, malo, muy malo, etc.

Los criterios utilizados en la construcción del PSI fueron:

1. Que sea fácilmente accesible al público. El PSI transforma las unidades científicas de concentración a un número que el público puede entender con facilidad; los valores del índice se relacionan fácilmente con los criterios de estándares y episodios para cada valor del contaminante.
2. No tiene inconsistencias con los niveles percibidos de contaminación del aire. Las inconsistencias se originan solamente con los índices combinados o con los agregados. Este problema se supera, ya que el PSI calcula los valores del índice por separado para cada contaminante.
3. Tiene significancia espacial. La significancia espacial del PSI (o de cualquier índice) depende de cómo se seleccionan o manejan los datos antes de usarse para computar el índice. Por ejemplo, puede usarse la concentración promedio de un contaminante medido en varias estaciones de monitoreo, o bien, se puede elegir la concentración más alta del contaminante medida en alguna de varias estaciones de una ciudad para calcularlo. Se recomienda usar este último método

para evitar disfrazar las concentraciones altas de contaminantes que resultan al promediar los datos.

4. Incluye los principales contaminantes actuales del aire, con capacidad de incluir contaminantes futuros.
5. Se calcula de una manera sencilla usando supuestos razonables.
6. Se apoya en bases científicas razonables. El PSI convierte los valores de la concentración de contaminantes, usando el método de cálculo de interpolación lineal simple.
7. Está relacionado con criterios de episodios y urgencias ambientales. Para cada contaminante, el PSI lo relaciona con los criterios de episodios y urgencias ambientales como: Alerta, 200; Precaución, 300; Urgencia, 400; Riesgo Sustancial, 500.
8. Muestra la variación día a día. Debido a que el PSI se calcula para cada contaminante por separado y se informa el valor máximo, las variaciones diarias se presentarán tanto en el PSI como en el contaminante del que se informa.
9. Esos datos permiten predecir valores del índice con 18 horas anticipadas por medio del uso de las siguientes palabras descriptoras: "permanece sin cambios", "disminuye" o "aumenta". Estos descriptores sólo son cualitativos, y, como tales, sólo indican la tendencia esperada del valor del índice durante el periodo pronosticado.

3.2.3 El IMECA

En México se han seguido las ideas planteadas en el PSI, con algunas variaciones. En 1972 surgió la Subsecretaría de Mejoramiento del Ambiente (SMA), que normó los contaminantes atmosféricos y propuso lo que se denominó el Índice Mexicano de la Calidad del Aire (IMEXCA), basado en lo propuesto por el PSI, sumando los diversos contaminantes en una ecuación que no era fácil de desarrollar.

TESIS CON
FALLA DE ORIGEN

Para 1982, surgió el Índice Metropolitano de la Calidad del Aire (IMECA). La Secretaría de Desarrollo Urbano y Ecología (SEDUE), y el Instituto Nacional de Ecología (INE) en su interior, crearon este índice. El IMECA tiene una gran semejanza con el PSI, aunque las concentraciones de los contaminantes no son exactamente iguales. Se empezó a seguir el criterio de expresar un solo contaminante como guía o centinela, destacando el ozono, por sus altas concentraciones. El IMECA es un valor de referencia para que la población conozca los niveles de contaminación prevalecientes en su zona de residencia, de manera precisa y oportuna, para que tome las medidas pertinentes de protección.

Cuando el IMECA de cualquier contaminante rebasa los 100 puntos, significa que sus niveles son perjudiciales para salud y en la medida en que aumenta el valor del IMECA se agudizan los síntomas (Cuadro 3.2.4).

IMECA	CALIDAD DEL AIRE	EFFECTOS SOBRE LA SALUD
0-100	Satisfactoria	Situación favorable para la realización de todo tipo de actividades
101-200	No Satisfactoria	Aumento de molestias, en nariz, garganta y ojos en personas sensibles
201-300	Mala	Evitar actividades al aire libre. Posibles problemas respiratorios
301-500	Muy mala	Se agudizan los síntomas anteriores en personas sensibles

3.2.4 Interpretación del IMECA

El cálculo del IMECA ha tenido una serie de cambios a lo largo del tiempo. Actualmente, y dependiendo del contaminante, en la ZMVM se instrumentan dos algoritmos para el cálculo del IMECA. El primero para O₃, NO_x, SO₂, CO es el algoritmo simplificado²³ y el segundo para PM₁₀ toma la estructura del PSI (Cuadro 3.2.5). La diferencia fundamental es que el PSI define más de un punto de quiebre mientras que el método simplificado actual utiliza solo uno, que es el valor de la norma oficial mexicana para cada contaminante. El primer punto de quiebre corresponde al valor de la norma

²³ El algoritmo simplificado se utiliza desde el 23 de diciembre de 1994 en la ZMVM.

Análisis de Intervención; Aplicación a un problema ambiental.

mexicana de calidad del aire de 0.11 ppm como promedio horario, homologado con un valor de 100 puntos. El segundo punto en 0.600 ppm corresponde a 500 puntos del IMECA.

Los valores que aparecen en el intervalo que indica el cuadro 3.2.5, se refiere al límite que señala la Norma Oficial Mexicana (NOM) de cada contaminante. A su vez, el cuadro 3.2.6 muestra estas normas.

Contaminante	Intervalo	Ecuación
PST *	0 -- 260 $\mu\text{g}/\text{m}^3$ 260 -- 1,000 $\mu\text{g}/\text{m}^3$	$IPST = 0.38461538CPST$ $IPST = (0.54054CPST) - 40.540541$
PM10	0 -- 50 $\mu\text{g}/\text{m}^3$ 51 -- 350 $\mu\text{g}/\text{m}^3$ 351 -- 420 $\mu\text{g}/\text{m}^3$ 421 -- 500 $\mu\text{g}/\text{m}^3$ 501 -- 600 $\mu\text{g}/\text{m}^3$	$IPM10 = CPM10$ $IPM10 = (1/2) CPM10 + 25$ $IPM10 = (10/7) CPM10 - 300$ $IPM10 = (10/8) CPM10 - 225$ $IPM10 = CPM10 - 100$
SO ₂	0 -- 0.13 ppm 0.13 -- 1.00 ppm	$ISo2 = 769.230769Cso2$ $ISo2 = (459.770114Cso2) + 40.22989$
CO	0 -- 11 ppm 11 -- 50 ppm	$Ico = 9.090909Cco$ $Ico = (10.25641Cco) - 12.820512$
NO ₂	0 -- 0.21 ppm 0.21 -- 2.00 ppm	$INo2 = 476.190476CNO2$ $INo2 = (223.463687CNO2) + 53.07264$
O ₃	0 -- 0.11 ppm 0.11 -- 0.60 ppm	$Io3 = 909.090909 Co3$ $Io3 = (816.32653 Co3) + 10.20409$

Cuadro 3.2.5 Cálculo del IMECA para los principales contaminantes en la ZMVM.

La función principal del IMECA es mantener informada a la población sobre la calidad del aire en la Ciudad de México, así como observar el comportamiento de los distintos contaminantes y comparar la calidad del aire entre zonas que utilicen índices similares.

TESIS CON
FALLA DE ORIGEN

* PST: Partículas suspendidas totales.

I: representa el índice para cada contaminante

C: Cantidad de cada contaminante. C esta dentro del intervalo dado

Contaminante	Concentración normada y Tiempo promedio	Frecuencia Máxima Aceptable
Ozono	0.110 ppm 1 hora	1 cada tres años
Bióxido de azufre	0.130 ppm 24 horas	1 vez al año
Bióxido de nitrógeno	0.210 ppm 1 hora	1 vez al año
Monóxido de carbono	11 ppm 8 horas	1 vez al año
PST	260 µg/m ³ 24 horas	1 vez al año
PM10	150 µg/m ³ 24 horas	1 vez al año

Cuadro 3.2.6 Normatividad para los principales contaminantes en la ZMVM.

La atención que el público le había prestado al IMECA era muy baja, hasta 1996, en que se declararon contingencias ambientales²⁴ en las cuales dejaron de circular automóviles durante dos días seguidos. A partir de ese momento, la atención se ha dirigido más y más sobre el IMECA, en especial para tener la seguridad de que el automóvil particular puede o no circular determinado día de la semana, o varios días, dependiendo de las condiciones ambientales.

Existen varias fases de contingencia ambiental en la ZMCM. En cada una de las fases, se implementan medidas restrictivas para disminuir los precursores del ozono, que cada vez son más estrictas.

1. Se declara la *fase I* de contingencia por altos niveles de contaminación atmosférica cuando se presentan:

- Valores superiores a 240 puntos IMECA en cualquiera de las estaciones de la RAMA.

TESIS CON
FALLA DE ORIGEN

²⁴ Una Contingencia Ambiental o Emergencia Ecológica es la situación eventual y transitoria declarada por las autoridades competentes cuando se presenta o se prevé con base en análisis objetivos o en el monitoreo de la contaminación ambiental, una concentración de contaminantes o un riesgo ecológico derivado de actividades humanas o fenómenos naturales que afectan la salud de la población o al ambiente de acuerdo con las normas oficiales mexicanas."

Análisis de Intervención; Aplicación a un problema ambiental.

- Niveles generalizados de ozono superiores a 100 IMECA en todas las zonas, o, cuando menos, en tres de las estaciones centinela²⁵.
2. La **fase II** se declara cuando:
 - La contaminación fluctúa alrededor de los 300 puntos del IMECA.
 - Cuando la Fase I este activa por más de 3 días
 3. La **fase III** se declara cuando los índices de contaminación fluctúan alrededor de 450 puntos IMECA.

Como se ha mencionado, el Ozono es uno de los contaminantes que presenta una concentración muy alta. Por tanto, en este estudio se toman los registros de este contaminante. Además, el Ozono ocasiona mayor número de contingencias ambientales en la ZMVM. Los datos del anexo corresponden a los registros máximos diarios de este contaminante.

Se debe tener cuidado al manejar la muestra de datos en el estudio. Esto es muy importante, ya que pueden existir datos erróneos (errores de captura, principalmente) o eventos desconocidos que afecten el desarrollo de la serie. Lo ideal es poder identificarlos, y corregirlos.

TESIS CON
FALLA DE ORIGEN

²⁵ Las estaciones centinela son: Xalostoc, Tlalnepantla, Merced, Pedregal y Cerro de la Estrella

3.3 Obtención y solución del modelo matemático

Las gráficas mostradas en los capítulos anteriores provocan cuestionamientos como: ¿Se han disminuido los índices de contaminación con la aplicación del programa "Hoy No Circula"? Y si es así, ¿en qué grado se han disminuido?

Los datos mostrados en el anexo I constituyen la serie de tiempo que será examinada. Se utilizará la metodología explicada en los capítulos I y II para obtener un modelo matemático. La obtención de un modelo matemático y los resultados que éste arroje, pueden dar respuesta a las preguntas planteadas y posiblemente indicar las rutas a seguir para poder disminuir en mayor grado la contaminación. La Figura 3.3.1 muestra la estrategia que se usará para poder encontrar un modelo ARIMA adecuado.

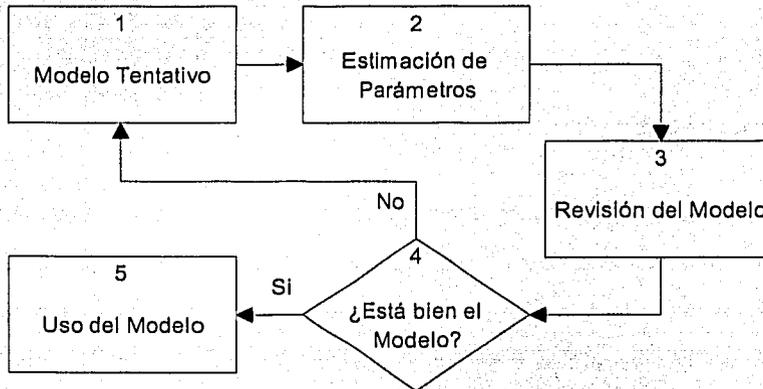


Figura 3.3.1 Diagrama de flujo para la obtención de un modelo adecuado

El primer paso es tener un modelo tentativo y para ello es necesario graficar la serie de tiempo. La figura 3.3.2 muestra el índice diario de contaminación por Ozono en la ZMVM graficado contra el tiempo. El motivo principal de realizar esta gráfica es detectar características, tales como: estacionalidad, varianza constante, tendencia, etc.

Ya que se ha analizado la gráfica de la serie se decide si es necesario o no aplicar algunas transformaciones a la serie para hacer varianza constante o realizar diferencias para eliminar tendencia o estacionalidad.

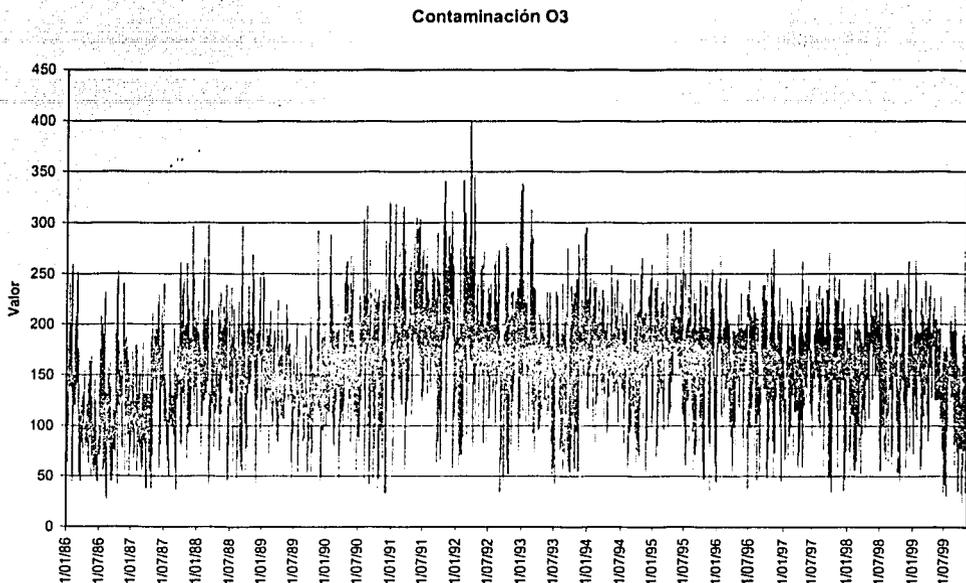


Figura 3.3.2 Índice de Contaminación por Ozono en la ZMVM

La figura 3.3.2 comprende el periodo entre el 01/Ene/1986 y 31/Oct/1999 con un total de 5052 observaciones. En adelante, esta serie se denominará O₃ (Ozono). La media de O₃ es 159.30 y su desviación estándar es 50.84. Con el fin de estabilizar la varianza y en consecuencia disminuir la desviación estándar, a la serie O₃ se le aplicará una transformación. Dado que existe un rango amplio de valores del Índice [25,398], todos los valores son no negativos y no se nota ninguna curva exponencial o logarítmica, se aplicará raíz cuadrada a la serie O₃. El resultado se muestra en la figura 3.3.3. Al realizar esta transformación se obtiene una nueva serie, llamada SQRT(O₃)²⁶, que conserva las principales características de la original, pero con media igual a 12.44, desviación estándar igual a 2.12 y rango de valores reducido a [5, 19.95]

TESIS CON
FALLA DE ORIGEN

²⁶ SQRT es la abreviación de Square Root (Raíz Cuadrada)

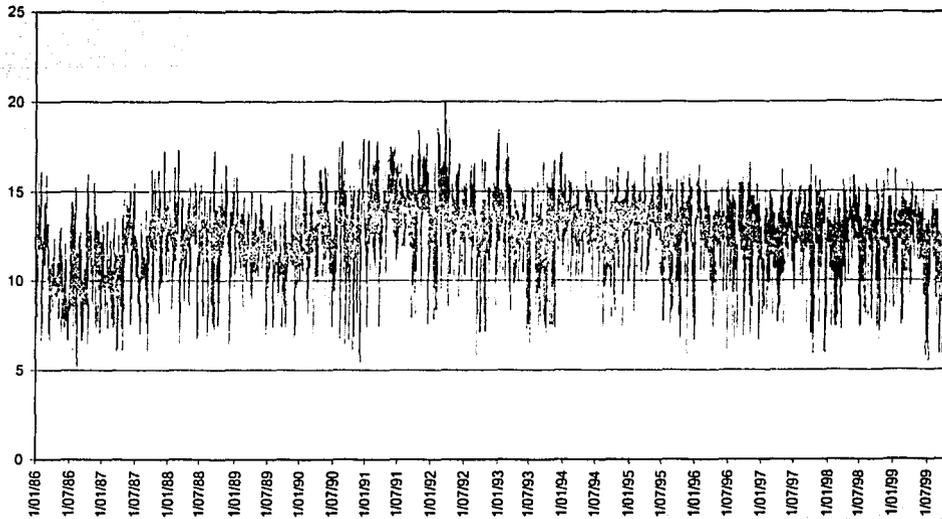


Figura 3.3.3 Raíz Cuadrada aplicada a la serie del Índice de Ozono

No obstante la reducción de media, varianza y rango, los datos de la serie $SQRT(O_3)$ conservan cierta tendencia. Por ejemplo, en la figura 3.3.3 se nota una tendencia ascendente entre 1990 y 1991. Para eliminar tendencias lo mejor es aplicar diferencias sencillas a la serie $SQRT(O_3)$ (Ver Capítulo I). La Figura 3.3.4 muestra la serie $SQRT(O_3)$ con una diferencia:

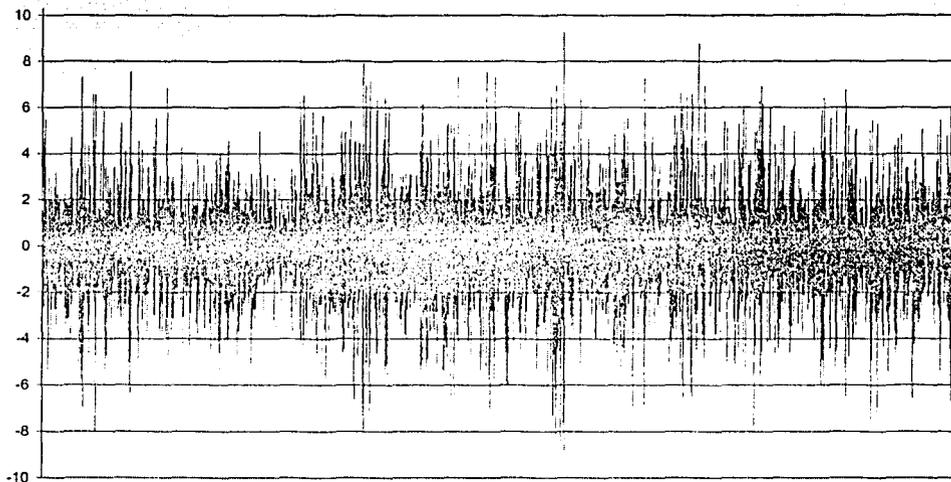


Figura 3.3.4 Raíz Cuadrada y una diferencia del índice de Ozono

Una vez hechas las transformaciones y diferencias necesarias, el siguiente paso es encontrar un modelo tentativo para representar los datos. Es importante hacer notar que no siempre se deben de aplicar transformaciones y diferencias a una serie de tiempo. Sólo se realiza esto cuando se necesita estabilizar la varianza y quitar tendencias con el fin de que se cumplan las condiciones de estacionaridad y de invertibilidad.

Ya que se va aplicar la metodología del análisis de intervención, se tiene que encontrar un modelo $ARIMA(p,d,q)$ que represente el conjunto de datos anteriores a la intervención. En este caso, el programa "Hoy No Circula" (HNC) representa la intervención en la serie del Índice de Ozono. Dado que la aplicación obligatoria del programa HNC empezó el 20 Noviembre de 1989, la serie mostrada en la figura 3.3.4 se ve reducida en número de observaciones (Ver Figura 3.3.5).

Raíz Cuadrada(O3) y 1 diferencia

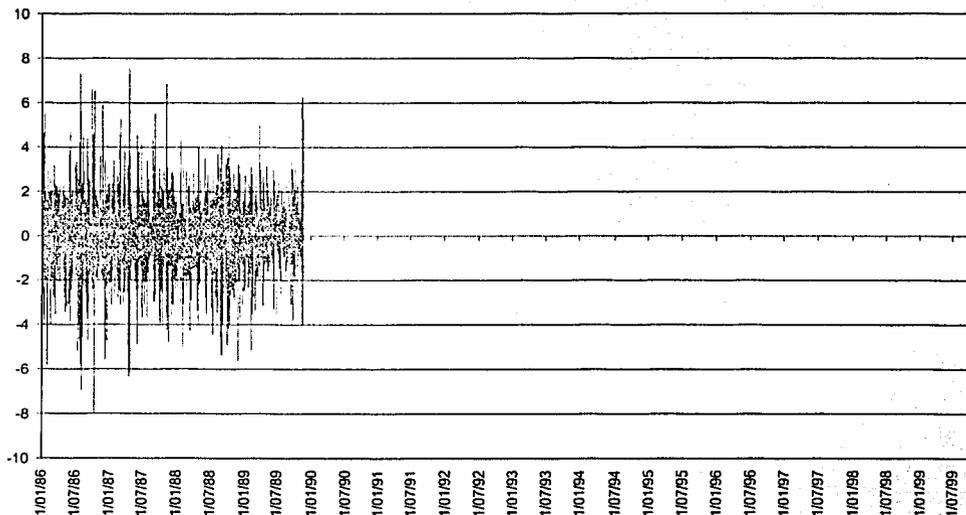
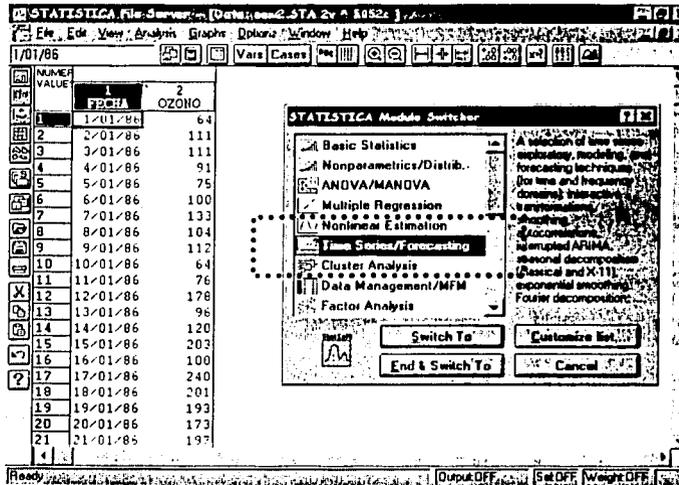


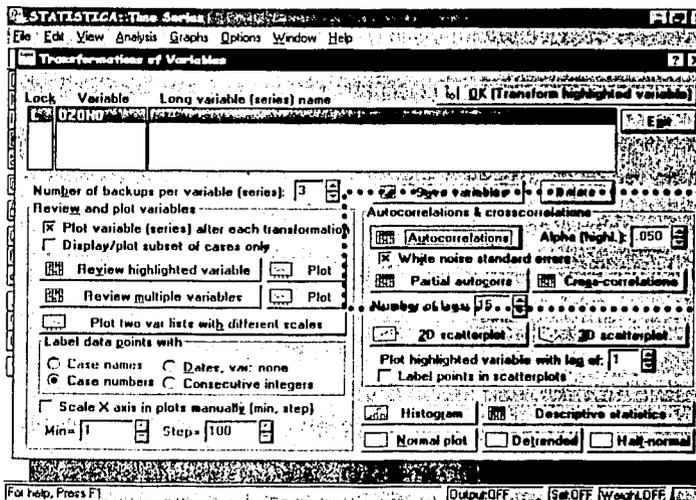
Figura 3.3.5 Serie anterior a la intervención del HNC (20/Nov/1989)

La función de autocorrelación (ACF) y la función de autocorrelación parcial (PACF) de la serie original ayudará a obtener el primer modelo tentativo. Estas gráficas se obtienen de las expresiones dadas en el Capítulo I y con la ayuda de un programa llamado "STATISTICA for Windows".

Una de las primeras pantallas que se observa al abrir el programa STATISTICA es la siguiente:



Esta pantalla contiene las variables que se analizan, así como el módulo que se va a utilizar, que en este caso es: series de tiempo (Time Series/Forecasting). En caso que se tengan capturados los datos en un archivo de otro tipo (excel, dbase, etc), STATISTICA permite migrar esos datos. Dentro del módulo de series de tiempo se pueden obtener las funciones de autocorrelación: ACF y PACF.



La ACF y PACF del índice de ozono se obtienen de esta manera y se muestran en las figuras 3.3.6 y 3.3.7, respectivamente.

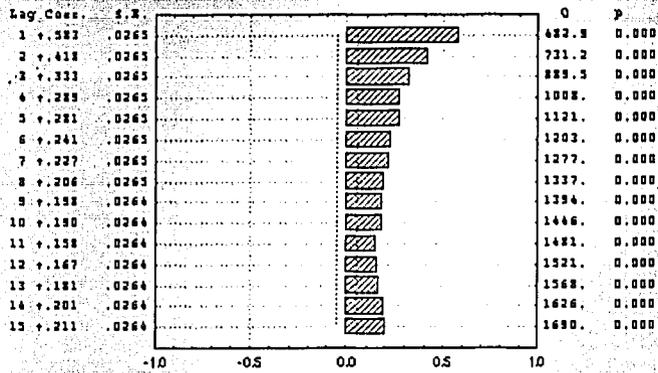


Figura 3.3.6 Función de Autocorrelación del índice de Ozono

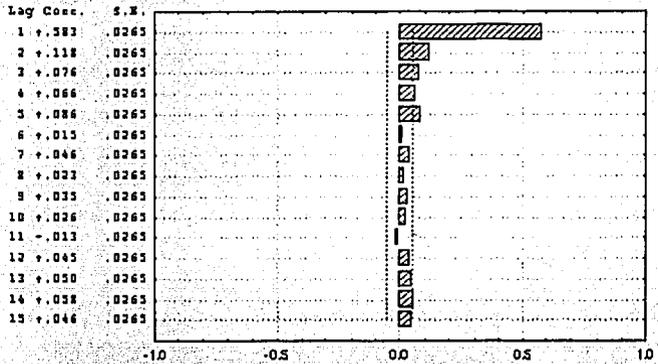


Figura 3.3.7 Función de Autocorrelación Parcial del índice de Ozono

Como se puede ver en las figuras 3.3.6 y 3.3.7, la ACF es decreciente y la PACF muestra el primer coeficiente muy grande con respecto de los demás. Por lo tanto, y de acuerdo al Cuadro 1.3.1, se puede tomar como primer modelo tentativo un ARIMA(1,1,0), es decir:

$$(1 - \phi_1 B)(1 - B)\sqrt{Z}_t = k + e_t \quad (3.3.1)$$

Para que el proceso sea estacionario las raíz de $(1 - \phi_1 B) = 0$ tiene que estar fuera del círculo unitario, y por tanto, $|\phi_1| < 1$. k es una constante que se calcula en el modelo como si fuese parámetro.

Para la estimación de los parámetros de los modelos que se presenten en este trabajo se utilizará STATISTICA (Ver Figura 3.3.8)

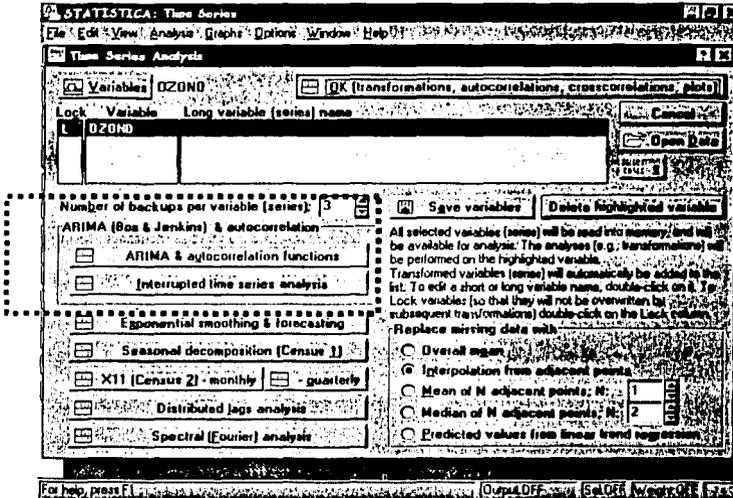


Figura 3.3.8 Pantalla de selección de los modelos ARIMA y Análisis de Intervención en STATISTICA

En la pantalla que corresponde a los modelos ARIMA se indican las variables a tomar en cuenta (Figura 3.3.9A) y los parámetros a estimar (Figura 3.3.9B). Además, se indica las transformaciones y diferencias que se deben tomar en cuenta (Figura 3.3.9C).

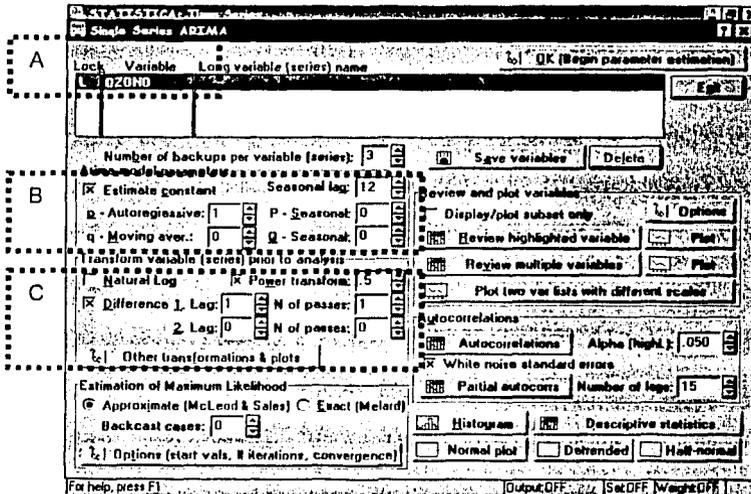


Figura 3.3.9 Modelos ARIMA en STATISTICA

El programa STATISTICA arroja los resultados en una nueva pantalla (Figura 3.3.10). En ésta se observa el valor de cada parámetro y de su correspondiente error estándar. Además se puede obtener la ACF y PACF de los residuales para saber si el modelo propuesto es el adecuado o necesita modificarse.

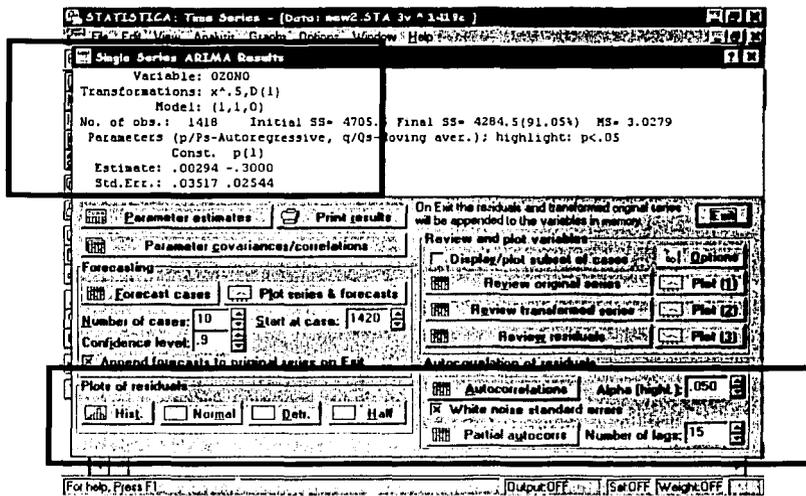


Figura 3.3.10 Pantalla de resultados de estimación

De esta forma se obtuvieron las estimaciones para el modelo ARIMA(1,1,0):

$$\begin{matrix} (1 + 0.3B)(1 - B)\sqrt{Z_t} = 0.00294 + e_t \\ (0.02544) \qquad \qquad (0.03517) \end{matrix} \quad (3.3.2)$$

El valor por debajo de cada estimación en la expresión (3.3.2) es el error estándar respectivo. Tomando en cuenta este valor, se puede decir que la constante estimada no es significativa ($k=0.0294$ no es al menos dos veces el error estándar 0.03517). En consecuencia, se vuelve a estimar el modelo ARIMA(1,1,0) de la expresión (3.3.1) pero sin incluir la constante k . El resultado se presenta en la expresión (3.3.3)

$$\begin{matrix} (1 + 0.3B)(1 - B)\sqrt{Z_t} = e_t \\ (0.02544) \end{matrix} \quad (3.3.3)$$

En este caso el valor del parámetro $\Phi_1 = -0.3$ no cambió al reestimar el modelo.

3.3 Obtención y solución del modelo matemático

El siguiente paso es revisar el modelo y saber si es el más adecuado. Para ello, se analizan la ACF y PACF de los residuales (Figuras 3.3.11 y 3.3.12).

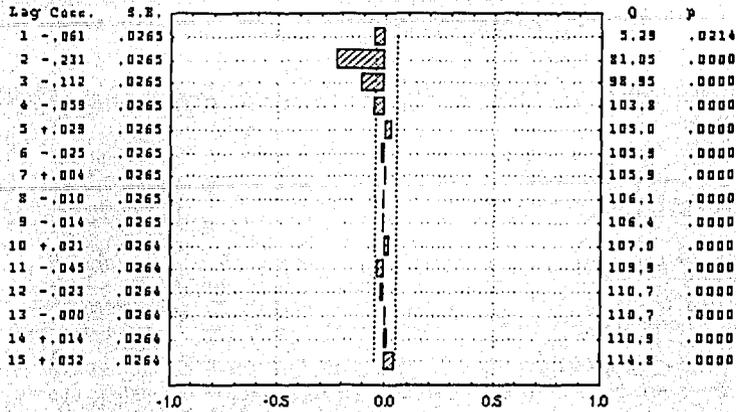


Figura 3.3.11 ACF de los residuales para el modelo ARIMA(1,1,0)

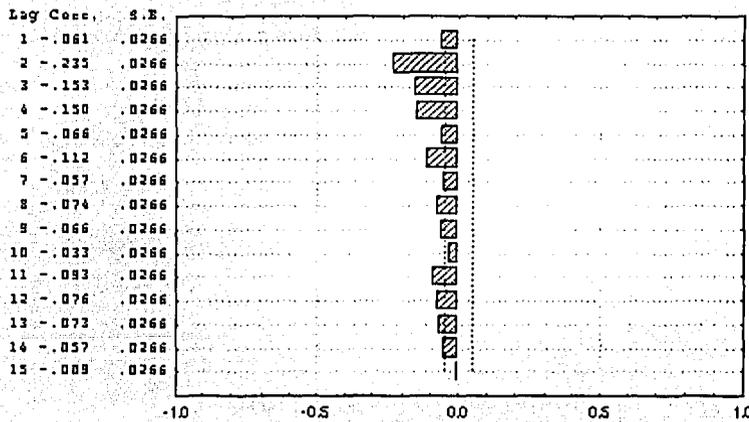


Figura 3.3.12 PACF de los residuales para el modelo ARIMA(1,1,0)

La ACF en la figura 3.3.11 muestra un pico en el segundo coeficiente, mientras que la PACF parecería ser decreciente infinita. Con base en este análisis residual, el modelo ARMA(1,1,0) resulta inadecuado. La serie de residuales no se comporta como ruido blanco, sino como un proceso MA(2). Por lo tanto, el modelo ARIMA (1,1,0) de la expresión 3.3.1 se puede redefinir como:

TESIS CON
FALLA DE ORIGEN

$$(1 - \phi_1 B)(1 - B)\sqrt{Z_t} = k + a_t$$

si $a_t = (1 - \theta_1 B - \theta_2 B^2)e_t$, entonces:

$$(1 - \phi_1 B)(1 - B)\sqrt{Z_t} = k + (1 - \theta_1 B - \theta_2 B^2)e_t \quad (3.3.4)$$

La expresión (3.3.4) es un modelo ARIMA(1,1,2). Este nuevo modelo constituye otra etapa iterativa en la búsqueda de un modelo adecuado. Las estimaciones que se obtienen para este nuevo modelo son:

$$(1 - 0.553258B)(1 - B)\sqrt{Z_t} = 0.001774 + (1 - 1.059588B + 0.103977B^2)e_t \quad (3.3.5)$$

(0.063983) (0.004269) (0.072265) (0.062408)

Nuevamente, el error estándar en la expresión (3.3.5) indica que la constante $k=0.001774$ no es significativa. De igual forma, la estimación del parámetro $\theta_2=-0.103977$ no es significativa ya que su error estándar es igual a 0.062408.

Cabe señalar que si alguna de las estimaciones significativas de los parámetros es mayor o igual a uno, en valor absoluto, el modelo podría no ser estacionario o invertible.

Dos de los parámetros que se incluían en la expresión (3.3.4) no son significativos ocasionando que el modelo ARIMA(1,1,2) está sobre-especificado. De acuerdo a lo descrito en el subcapítulo 1.4.2.2, el parámetro θ_2 se puede eliminar sin revisar las correlaciones con las estimaciones de otros parámetros, ya que es el de mayor orden. Así pues, el nuevo modelo es un ARIMA(1,1,1):

$$(1 - \phi_1 B)(1 - B)\sqrt{Z_t} = (1 - \theta_1 B)e_t \quad (3.3.6)$$

cuyas estimaciones son:

$$(1 - 0.451347B)(1 - B)\sqrt{Z_t} = (1 - 0.938495B)e_t \quad (3.3.7)$$

(0.030313) (0.013669)

Las estimaciones de ϕ_1 y θ_1 son significativas y cumplen con las condiciones de estacionaridad e invertibilidad ya que $|\phi_1| < 1$ y $|\theta_1| < 1$.

TESIS CON
FALLA DE ORIGEN

Solamente falta por revisar la ACF y PACF de los residuales (Figuras 3.3.13 y 3.3.14) para saber si el modelo ARIMA(1,1,1) es adecuado para representar la serie en cuestión.

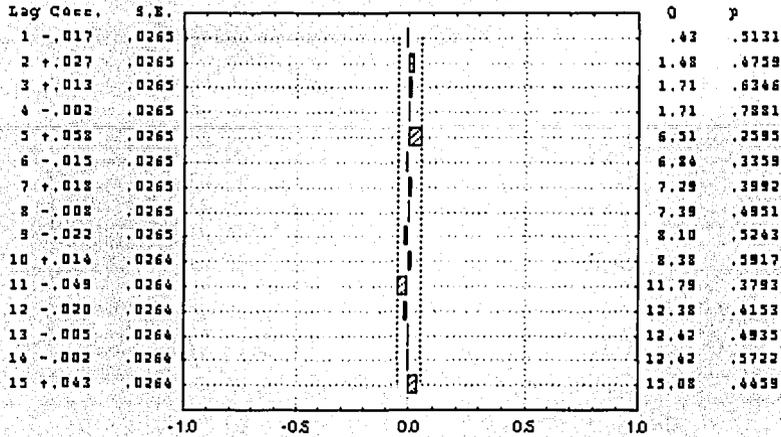


Figura 3.3.13 ACF de los residuales para el modelo ARIMA(1,1,1)

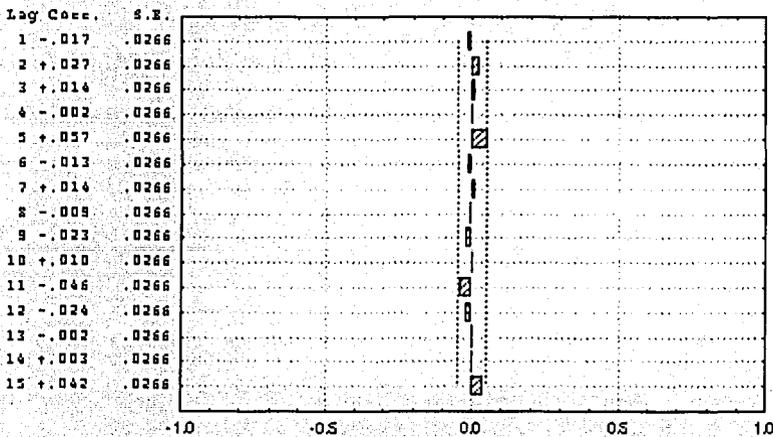


Figura 3.3.14 ACF de los residuales para el modelo ARIMA(1,1,1)

Las Figuras 3.3.13 y 3.3.14 no muestran ningún coeficiente significativo, por lo que se puede concluir que los residuales se comportan como ruido blanco. El periodograma

Integral (Figura 3.3.15) coincide con estos resultados ya que, en general, no muestra ninguna desviación de la línea de 45 grados.

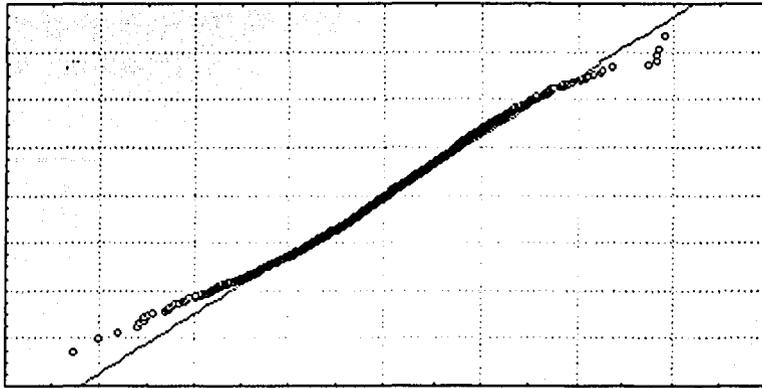


Figura 3.3.15 Periodograma Integral de los residuales del modelo ARIMA (1,1,1)

Por los resultados obtenidos al analizar las estimaciones de los parámetros y los residuales, el modelo ARIMA(1,1,1) es el adecuado para representar a los datos del índice de contaminación por ozono anteriores al programa "Hoy No Circula". Para analizar toda la serie se tiene que incluir una función que represente la intervención gradual y permanente del HNC. Esto es,

$$\sqrt{Z_t} = \frac{\omega_0}{(1-\delta_1 B)} I_t + \frac{(1-\theta_1 B)}{(1-\phi_1 B)(1-B)} e_t \quad (3.3.8)$$

donde ω_0 representa el impacto inicial del HNC en el índice de contaminación por ozono en la ZMVM y $0 < \delta_1 < 1$ representa una respuesta gradual. Si $\delta_1 = 0$, entonces representa una respuesta abrupta a la intervención y si $\delta_1 = 1$ el impacto se incrementa en forma lineal. La función I_t puede ser una función escalón o pulso dependiendo del problema. En este caso, se utiliza una función escalón, ya que la intervención del HNC no se aplica a un solo instante, sino a todo un periodo de tiempo. Así entonces,

$$I_t = \begin{cases} 1 & \text{si } t \geq 1420 \quad (20/Nov/1989) \\ 0 & \text{si } t < 1420 \end{cases}$$

Se toma en cuenta $t=1420$ ya que es la observación que corresponde a la fecha en que el HNC empezó a ser obligatorio.

A la expresión (3.3.8) se le denomina "Modelo de Intervención". Algunos autores ya no mencionan si este modelo tiene o no componentes de medias móviles o autorregresivos. Simplemente le denominan "Función de Transferencia" a todo el conjunto de términos y parámetros.

Utilizando STATISTICA se obtienen las estimaciones para el modelo de intervención planteado.

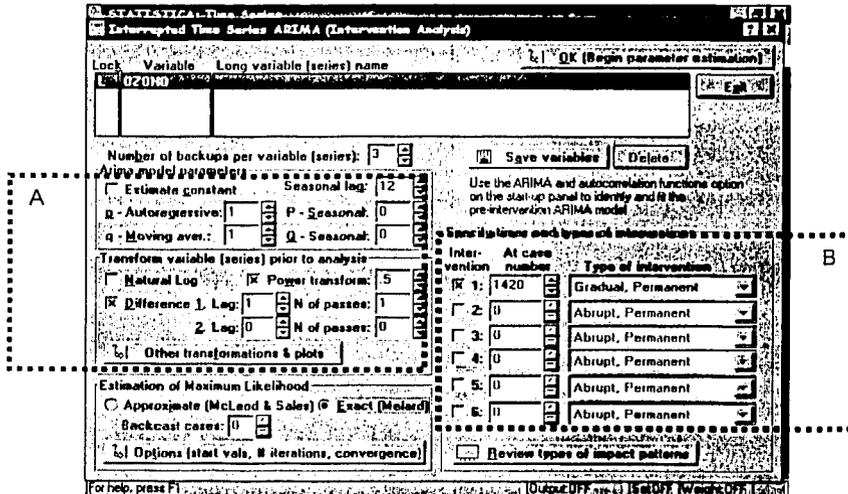


Figura 3.3.16 Pantalla para obtener las estimaciones del modelo de intervención

En la Figura 3.3.16A se indican los parámetros del modelo anteriores a la intervención, mientras que en 3.3.16B se indica el tipo de intervención. Las estimaciones y los errores estándar asociados obtenidos son:

TESIS CON
FALLA DE ORIGEN

Parámetro	Estimación	Error Estándar	Significativo
Φ_1	0.371140011	0.01570241	si
θ_1	0.963586317	0.006559171	si
ω_0	-0.700388955	1.058618426	no
δ_1	-0.776308525	0.280950856	si

La ACF, la PACF y el periodograma de los residuales (Figuras 3.3.17 a 3.3.19) indican que el modelo de intervención planteado es el adecuado.

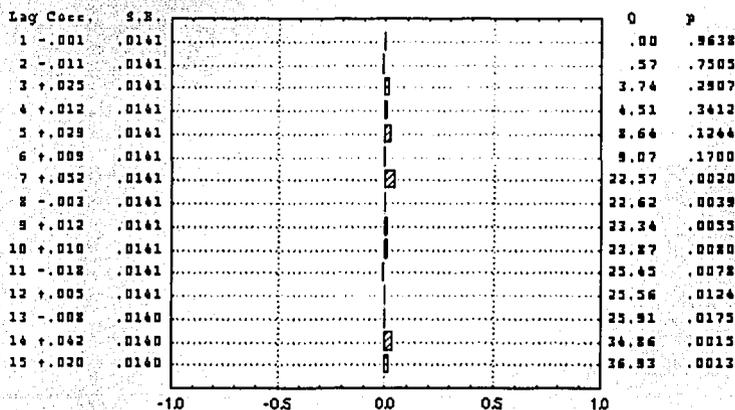


Figura 3.3.17 ACF de los residuales para el modelo de intervención

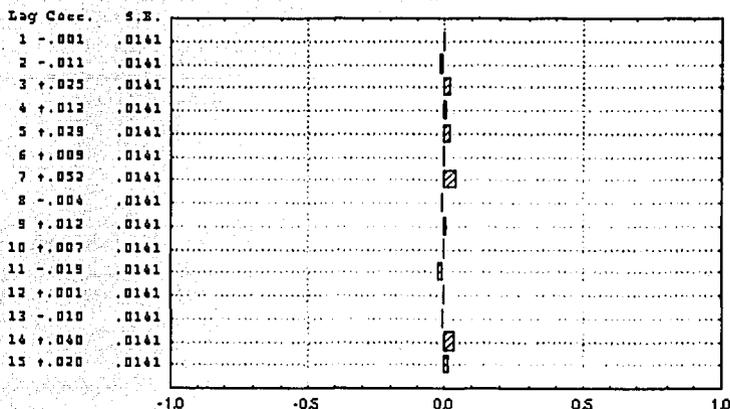


Figura 3.3.18 PACF de los residuales para el modelo de intervención

TESIS CON
FALLA DE ORIGEN

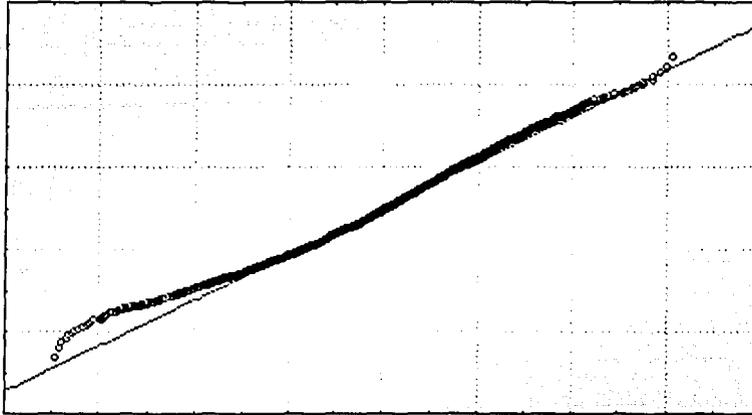


Figura 3.3.19 Periodograma Integral del Modelo de Intervención

Dados estos resultados, el impacto del programa HNC en los índices de contaminación por ozono, representado por ω_0 , no es significativo tomando la serie del 01/Ene/1986 al 31/Oct/1999.

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

CAPÍTULO IV

RESULTADOS

Algunas veces, los resultados que proporciona un modelo matemático no son claros de entender. Por ello, en este capítulo se da una interpretación de tal forma, que no importando la experiencia que una persona pueda tener en el análisis de series de tiempo, sean comprendidos los resultados

TESIS CON
FALLA DE ORIGEN

IV. Resultados

4.1 Interpretación de los resultados del modelo matemático.

En este trabajo se mostró una forma de evaluar el impacto del programa HNC en los índices de contaminación. Se estudiaron los datos históricos del índice de ozono en la ZMM, anteriores y posteriores a dicho programa, usando la metodología del análisis de intervención en las series de tiempo. Esta metodología es útil para saber de qué manera afectan algunas acciones aplicadas en el pasado sobre todo un conjunto de datos.

La serie original del índice de ozono, O_3 , fue sujeta a varios cambios para cumplir con condiciones de estacionaridad e invertibilidad. Es necesario que la serie en estudio cumpla con estas condiciones para poder utilizar la metodología Box-Jenkins y el análisis de intervención. A la serie se le aplicó una transformación (raíz cuadrada) con el fin de estabilizar la varianza y reducir su rango. En otras palabras, al realizar la transformación se evitó que los datos estuvieran muy dispersos.

Al principio del estudio, se podría pensar que el índice de contaminación por ozono se repite cada semana, dado que el programa HNC sólo se aplica de Lunes a Viernes. Es decir, que se debería tomar en cuenta un componente estacional con $s=7$. Sin embargo, la ACF y PACF de la serie (Figuras 3.3.6 y 3.3.7) no mostraban ningún componente estacional de forma clara (por ejemplo, para un componente estacional semanal, se notarían picos en el primero, séptimo y decimocuarto coeficiente).

Además, al aplicar una diferencia, no sólo se eliminó tendencia, sino que se redujo la media. Todo esto indica que el índice de ozono para un día en particular depende, en cierta medida, del índice reportado el día anterior.

El resultado obtenido del modelo matemático en el subcapítulo 3.3 mostró que el impacto del HNC, ω_0 , en el índice de ozono no era significativo. En otras palabras, no hay suficiente evidencia para poder asegurar que el programa HNC ha servido para disminuir los índices de contaminación por ozono en la Ciudad de México.

4.2 Aportación a la disciplina y posibles trabajos posteriores

La tendencia ascendente que se mostraba en los datos a partir de 1990 (Figura 3.3.2) podría hacer pensar que el HNC no sólo no redujo, sino que aumento los índices de contaminación. Sin embargo, el resultado no indicó eso. Tal vez esa tendencia se mostró por el crecimiento del volumen vehicular, ya que muchas familias ante la insatisfacción de no poder utilizar su auto un día, adquirirían otro.

4.2 Aportación a la disciplina y posibles trabajos posteriores.

Muchas de las aplicaciones de series de tiempo están relacionadas con economía. Existen pocos casos en donde estos métodos se han aplicado para estudiar fenómenos parecidos al presentado en este trabajo.

La importancia del estudio radica, principalmente, en el resultado obtenido. Se demostró que, por sí solo, el HNC no disminuye los índices de contaminación. El siguiente paso sería tomar en cuenta otras intervenciones aplicadas en la serie. Algunas intervenciones que se podrían tomar en cuenta para trabajos posteriores son:

- Mejoramiento e incorporación de nuevas tecnologías en vehículos automotores. Por ejemplo, volumen de automóviles que circulan en la ZMVM con convertidor catalítico.
- Cierre o clausura de grandes fuentes de contaminación. Por ejemplo, el cierre de la refinería 18 de Marzo.
- Mejoramiento y sustitución de energéticos en la industria y vehículos, tales como: introducción de gasolina sin plomo, o uso de gas natural.
- Oferta amplia de transporte público seguro y eficiente. Por ejemplo, inauguración de la línea B del metro.

Cada una de estas intervenciones se podría introducir al modelo planteado, como se hizo con el HNC en la expresión (3.3.8). En el modelo se representarían con un coeficiente ω_k , y un tipo de intervención, gradual o abrupta. Por ejemplo, el cierre de la

Análisis de Intervención: Aplicación a un problema ambiental.

refinería 18 de Marzo y la puesta en marcha de la línea B del metro serían intervenciones abruptas y permanentes sobre la serie, mientras que el uso de gas natural o de gasolina sin plomo se tomarían como intervenciones graduales y permanentes.

Ahora bien, en otro posible trabajo se podrían tomar en cuenta elementos de contaminación atmosférica, tales como: óxidos de azufre SO_x, óxidos de nitrógeno NO_x, etc. También se podrían realizar trabajos que analizaran la correlación entre el aumento del volumen de vehículos en la ZMVM y los índices de contaminación.

CONCLUSIONES

En este trabajo se analizó, a nivel univariante, las principales características del índice de contaminación atmosférica por ozono en la ZMVM. La aplicación de la metodología de series de tiempo con análisis de intervención a este tipo de problemas es viable. Esta metodología comprende la separación de los datos históricos en dos partes: observaciones anteriores a la intervención y observaciones que incluyen el efecto de la misma.

Usando juicios técnicos mostrados en los primeros capítulos, fue posible identificar y estimar un modelo ARIMA con análisis de intervención para recoger los efectos del programa "Hoy No Circula" en el índice de ozono. Los resultados indicaron que, por sí solo, el HNC no ha reducido la contaminación por ozono en la Ciudad de México y Área Metropolitana.

Sin embargo, el programa "Hoy No Circula" limita la circulación vehicular entre 500,000 y 390,000 unidades por día, lo que representa un importante ahorro de emisiones contaminantes precursoras de ozono. Por ello, se deberían de tomar en cuenta otro tipo de medidas para la reducción de contaminantes junto con el HNC, tales como: sustitución de energéticos, medios de transporte más eficientes, cierre de fuentes de contaminación, etc. Cada una de estas medidas representa una intervención en el índice de contaminación por ozono, es decir, comprende un periodo de tiempo y un tipo de respuesta. Además, el modelo ARIMA planteado puede resultar útil para evaluar intervenciones que aún no se aplican, pero que son similares a las pasadas. Es decir, se tendría la ventaja de poder estimar resultados de otras medidas (que no se apliquen en la realidad), disminuyendo costos económicos y sociales.

Cabe señalar que no se encontró evidencia de posibles periodos estacionales (semanal, mensual, anual, etc.) Sin embargo, se encontró que el índice de contaminación por ozono depende, en cierta medida, del reportado el día anterior.

Análisis de Intervención: Aplicación a un problema ambiental.

Finalmente, la contaminación en la ZMVM es grave y su evolución en el tiempo puede ir acumulando cada vez mayores complicaciones. En los últimos años el número de contingencias ambientales ha disminuido. No obstante, la norma internacional de ozono se excede en más de un 90% de los días del año en la mayor parte de la Zona Metropolitana. Esto es, el índice por ozono excede regularmente los 100 puntos de ozono, pero no llega a un nivel de 240. Por ello, es necesario implementar medidas eficientes y viables para la reducción de este contaminante. La sociedad y los gobiernos federal y local habrán de decidir conjuntamente sobre las intensidades y los tiempos de las acciones de combate a la contaminación atmosférica, en el entendido de que este problema no podrá resolverse ni en lo inmediato ni con paliativos recurrentes.

TESIS CON
FALLA DE ORIGEN

Referencias Bibliográficas

- Akaike, H. (1973). Information Theory and an extension of the maximum likelihood principle, Proc. 2nd International Symposium on Information Theory, Eds B.N. Petrov and F. Csaki), 267-281, Akademiai Kiado, Budapest.
- Akaike, H. (1974a) A new look at the statistical model identification. IEEE Transactions on Automatic Control.
- Akaike, H. (1978) A Bayesian analysis of the minimum AIC procedure, Ann. Inst. Statist. Math., 30A, 9-14
- Akaike, H. (1979) A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, Biometrika, 66,237-242
- Applications of Time Series Analysis in Astronomy and Meteorology. T. Subba Rao, M.B. Priestley, Eric D. Feigelson and O. Lessi Ed. Chapman and Hall. Gran Bretaña, 1997. P. 161-186
- Bhattacharyya, M.N. y Layton, A.P. (1979). Effectiveness of Seat Belt Legislation on the Queensland Road Toll. An Australian Case Study in Intervention Analysis. Journal of American Statistical Association, vol. 74, p. 596-603
- Box, G.E.P and Jenkins, G.M. (1976) Time Series Analysis Forecasting and Control, 2nd Edition, Holden Day, San Francisco.
- Box, G.E.P. y Cox D.R. (1964). An analysis of transformations, J. Roy. Stat. Soc. , Se. 26, 211-252
- Box, G.E.P. y Tiao, G.C. (1975) "Intervention analysis with applications to economic and environmental problems", *Journal of the American Statistical Association*, vol. 70, n. 349, p. 177-193.
- Cohen, Stephen B. (1981). Application of Intervention Analysis to Product Standards. Time Series Analysis. Ed. O.D. Anderson y M.R. Perryman, North-Holland, Amsterdam y Nueva York, p. 97-108.
- Deakin, E.B. (1976). Accounting Reports, Policy Interventions and the Behavior of Securities Returns. Accounting Review 51, p. 591-603

TESIS CON
FALLA DE ORIGEN

Análisis de Intervención: Aplicación a un problema ambiental

- Durbin, J. (1960) The fitting of the Time Series Models, Review of the Institute of International Statistics, 28, 233-244.
- El-Sheshal, Kamal y Morin, Roger M. (1984). Intervention Analysis and the dynamics of market efficiency. Time Series Analysis: Theory and Practice 5. Ed. O.D. Anderson, North-Holland, p. 145-157.
- González Videgaray, M.C. "Modelos de Decisión con Procesos Estocásticos II". Ed. UNAM. México, 1990.
- Instituto Nacional de Ecología – www.ine.gob.mx
- Izenman, A.J. y Zabell, S.L. (1981). Babies and the Blackout: The genesis of a misconception, Social Science Research, 10, p. 282-299.
- Kendall, Maurice y Keith Ord J. Time Series, 3a Edición, Ed. Edward Arnold. E.U., 1990.
- Keum-Rok Yoo. "Electoral Tax Cycle: A Comparative Analysis of Korea, the United States, and Japan." May 1995. 40 pp. Taxation Research Series.
- Lassarre, S. y Tan, S.H. (1982). Evaluation of Safety Measures on the Frequency and Gravity of Traffic Accidents in France by Means of Intervention Analysis. Time Series Analysis: Theory and Practice 1. Ed. O.D. Anderson, North-Holland, Amsterdam y Nueva York, p. 297-306.
- Montgomery, D.C. y Weatherby, G. (1980). Modeling and forecast time series using transfer function and intervention methods, AIIE Transactions, p. 289-307.
- Schwartz, G. (1978). Estimating the dimension of a model, Ann. Statist., 6, 461-464
- Secretaría del Medio Ambiente del Distrito Federal - <http://sma.df.gob.mx>
- Tukey, J.W. (1967) An Introduction to the Calculations of Numerical Spectrum Analysis, in advanced Seminar on Spectral Analysis. (De B. Harris), 25-46) Wiley, New York.
- Wei, William, W. S. Time Series Analysis: Univariate and Multivariate Methods. Ed. Addison Wesley. E.U., 1990.
- Wichern, Dean W. y Jones, Richard H. (1977). Assessing the Impact of Market Disturbances Using Intervention Analysis. Management Science 24:3, p. 329-337

ANEXO I

Datos Históricos del Índice de Contaminación por Ozono en la Zona
Metropolitana del Valle de México*

TESIS CON
FALLA DE ORIGEN.

* Fuente: Secretaría del Medio Ambiente del Distrito Federal - <http://sma.df.gob.mx>

TESIS CON
FALLA DE ORIGEN

106

1986

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	64	130	252	125	142	88	97	206	162	52	170	144
2	111	141	197	106	115	55	158	231	138	190	175	148
3	111	157	147	73	64	127	174	231	69	182	127	165
4	91	45	173	109	63	55	144	68	45	131	164	110
5	75	62	143	107	64	146	102	69	73	123	142	118
6	100	116	172	95	90	95	69	63	55	75	122	111
7	133	150	162	112	98	116	83	106	138	152	154	121
8	104	108	101	105	138	74	155	133	50	112	121	110
9	112	115	116	132	164	73	169	129	45	228	103	136
10	64	92	176	104	88	104	208	46	101	177	102	177
11	76	133	129	154	108	60	200	27	97	242	80	60
12	178	120	74	106	143	69	185	36	94	213	93	124
13	96	162	51	93	156	107	191	54	45	178	85	162
14	120	154	106	86	142	115	75	41	76	253	77	109
15	203	204	95	105	124	124	97	118	66	209	153	91
16	100	195	115	104	169	85	130	72	50	42	197	135
17	240	160	89	98	106	76	134	85	55	170	205	160
18	201	141	65	82	51	50	200	146	100	171	238	164
19	193	140	104	77	75	48	157	129	147	173	240	160
20	173	207	104	91	100	45	168	110	107	137	187	117
21	197	108	45	89	95	45	147	121	76	183	156	141
22	189	151	63	121	131	55	56	156	85	153	188	51
23	139	144	104	108	121	55	136	131	89	163	158	81
24	190	136	111	109	107	45	74	84	91	119	120	85
25	185	147	120	95	120	65	148	116	95	147	117	150
26	179	162	111	105	62	88	143	142	76	133	126	101
27	173	173	115	83	86	79	118	116	76	95	81	99
28	209	192	84	122	63	79	190	67	84	81	55	100
29	259		129	85	101	71	65	95	113	116	61	72
30	182		87	93	78	134	59	124	85	109	187	96
31	136		102		72		224	144		136		84

1987

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	132	158	96	106	160	154	132	124	160	187	147	99
2	101	98	55	158	169	167	102	151	113	160	158	146
3	147	78	142	102	122	229	145	166	59	83	164	165
4	90	55	154	118	139	220	121	175	107	123	197	202
5	146	74	88	108	160	154	120	138	86	199	168	223
6	173	113	67	116	149	156	105	102	88	228	178	208
7	162	100	182	85	151	58	148	160	157	125	260	296
8	126	100	106	74	188	147	130	134	91	147	224	184
9	139	120	59	94	153	168	157	111	76	191	157	184
10	94	129	91	120	150	178	112	107	37	186	156	170
11	113	130	108	131	146	121	152	142	134	127	68	164
12	117	128	119	108	170	115	105	142	169	79	227	249
13	121	144	117	99	192	110	109	136	156	124	144	161
14	105	115	132	55	184	82	146	116	133	173	180	133
15	159	79	105	98	204	164	81	106	170	235	138	151
16	90	139	105	58	182	209	102	101	181	195	93	197
17	107	151	91	107	147	181	133	102	151	132	160	157
18	68	150	82	182	141	160	107	103	157	169	136	131
19	91	148	120	72	135	164	116	104	152	188	215	183
20	75	88	131	73	153	225	76	101	122	247	98	107
21	101	63	128	96	187	207	105	111	169	139	130	169
22	84	111	83	84	167	173	104	130	177	187	102	179
23	108	64	37	156	198	139	66	110	167	166	99	209
24	77	107	73	38	143	147	91	95	140	150	145	159
25	123	78	120	187	161	183	65	109	191	122	150	146
26	120	70	86	191	187	170	76	94	190	147	152	156
27	99	80	46	147	193	152	97	91	140	129	186	202
28	178	89	39	211	207	129	115	103	177	116	162	138
29	164		64	70	222	240	49	93	196	187	141	139
30	162		147	142	166	231	108	195	261	200	110	148
31	180		92		171		133	136	161		169	

Análisis de Intervención: Aplicación a un problema ambiental

1988

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	137	211	155	125	75	159	151	141	107	134	173	42
2	98	217	128	102	159	157	155	125	149	153	156	56
3	129	233	99	79	182	212	119	164	138	56	146	113
4	156	266	67	136	174	169	139	175	170	121	190	192
5	174	262	42	166	215	193	119	176	59	151	231	140
6	222	124	76	145	231	144	133	158	138	97	156	142
7	158	158	105	175	172	239	188	192	202	127	171	178
8	176	176	68	146	179	221	236	123	262	104	184	147
9	191	195	100	113	172	222	219	179	180	142	269	143
10	171	204	92	140	192	148	218	198	218	214	266	193
11	166	202	83	136	123	146	132	164	297	118	180	104
12	175	125	139	130	111	159	135	194	175	170	179	168
13	182	202	150	173	171	212	174	97	124	79	168	189
14	150	168	200	137	159	172	146	48	73	180	230	182
15	178	184	214	147	179	105	189	59	53	201	244	223
16	156	205	107	129	174	45	151	130	80	162	212	207
17	150	209	124	115	198	71	132	68	130	151	156	175
18	153	202	122	152	171	61	170	130	150	204	207	171
19	128	192	151	198	155	104	176	211	195	238	158	202
20	151	145	180	186	158	75	78	181	171	131	173	168
21	174	130	213	149	191	47	65	173	170	213	121	189
22	157	174	137	167	217	67	105	215	165	141	171	200
23	131	205	153	173	190	87	168	187	168	211	139	248
24	71	206	166	221	218	145	189	173	194	142	147	161
25	124	298	162	134	189	208	166	223	184	169	158	134
26	169	274	162	164	156	194	216	145	123	186	130	120
27	164	227	119	156	162	224	197	136	102	201	144	94
28	109	162	156	204	162	222	181	175	178	177	130	126
29	216	158	152	174	201	211	182	76	204	194	191	147
30	185		166	162	169	169	160	139	198	133	146	151
31	187		130		156		98	119		170		150

1989

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	175	186	150	136	114	156	171	146	95	180	147	129
2	115	166	141	143	125	127	92	144	87	205	138	133
3	140	174	179	152	155	141	123	137	90	207	117	45
4	155	162	101	163	127	125	169	170	89	159	135	175
5	204	152	107	120	112	135	153	157	119	178	122	187
6	252	187	134	82	107	178	150	140	183	139	137	163
7	183	133	145	144	106	180	150	151	189	124	142	80
8	161	188	192	154	179	154	148	114	176	83	141	49
9	168	187	125	129	149	137	143	102	156	55	205	131
10	246	73	121	146	176	128	126	112	135	109	164	171
11	181	136	193	169	164	142	116	151	107	130	122	144
12	163	119	140	142	164	164	130	144	139	121	118	216
13	170	115	126	150	127	164	101	123	127	177	293	173
14	146	126	153	121	124	134	155	145	136	157	170	160
15	127	194	188	138	130	133	177	129	112	144	195	154
16	121	214	141	109	181	104	151	137	127	128	196	157
17	153	123	151	115	140	173	189	135	67	141	159	89
18	168	181	118	180	156	147	204	164	59	125	74	152
19	162	142	152	195	220	78	164	119	52	74	191	143
20	141	128	134	117	184	62	140	125	112	88	140	158
21	125	180	146	143	149	48	131	141	111	139	227	165
22	183	178	130	140	194	73	149	143	107	109	175	187
23	126	113	138	143	204	121	161	108	111	160	146	98
24	116	120	124	107	167	144	116	84	86	133	134	214
25	128	118	86	128	157	134	122	137	150	119	115	117
26	138	110	101	188	133	134	95	121	172	138	120	172
27	158	111	128	154	118	157	100	134	86	127	159	136
28	196	137	100	146	114	155	55	132	97	151	114	157
29	132		224	154	134	145	64	126	75	141	137	116
30	102		196	128	181	140	85	127	121	147	143	128
31	128		169		169		126	102		133		109

TESIS CON
FALLA DE ORIGEN

TESIS CON FALLA DE ORIGEN

1990

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	89	228	150	135	179	137	98	218	42	222	158	166
2	137	172	124	145	198	158	115	168	120	235	142	165
3	163	145	191	196	238	114	228	229	104	171	137	160
4	198	178	116	147	210	138	110	133	73	117	136	30
5	131	108	183	193	133	193	184	150	57	97	180	159
6	117	157	136	106	114	165	156	84	107	77	203	159
7	101	179	167	164	106	158	154	46	230	145	177	207
8	123	197	138	181	175	185	85	129	174	221	68	113
9	133	212	221	185	189	168	100	252	115	174	232	161
10	118	225	153	242	198	123	166	285	110	125	209	236
11	176	145	143	98	206	75	179	210	185	189	120	247
12	120	192	160	157	210	133	223	151	97	197	180	180
13	121	184	181	119	140	121	110	316	177	163	117	236
14	126	117	194	144	127	114	237	268	188	154	216	236
15	128	82	131	121	209	116	174	246	145	154	222	280
16	113	193	143	136	220	183	124	230	138	226	87	267
17	192	126	178	188	268	110	186	217	149	181	66	250
18	144	123	162	186	243	109	203	207	170	38	146	265
19	128	166	143	205	173	182	139	170	166	86	179	320
20	208	68	47	262	191	119	196	243	133	161	190	256
21	180	97	156	191	144	160	159	171	226	185	228	209
22	235	106	149	210	84	158	158	140	106	159	226	205
23	289	98	165	182	117	136	154	173	93	234	211	215
24	186	183	169	188	187	107	181	99	45	163	174	228
25	109	185	162	158	224	55	224	133	128	48	163	164
26	265	195	150	196	183	152	276	123	178	220	282	189
27	198	163	172	262	160	141	304	178	128	48	194	165
28	137	204	155	145	175	146	218	211	179	53	101	188
29	217	153	151	216	130	256	265	122	138	34	173	
30	206	102	165	258	89	233	236	110	128	93	136	
31	133	136	175	182	172	95						183

1991

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	118	181	230	164	164	253	138	184	156	158	195	220
2	157	249	208	169	212	214	197	226	202	230	258	226
3	127	172	245	201	187	295	134	207	174	223	150	242
4	128	208	106	139	235	185	149	182	225	218	106	311
5	154	159	280	165	167	238	186	143	159	131	114	262
6	177	236	267	145	184	162	217	178	196	67	235	190
7	177	143	315	126	239	163	144	164	171	83	176	195
8	183	225	250	134	237	195	203	206	186	88	93	276
9	133	181	316	216	286	154	182	216	170	102	201	190
10	54	148	159	190	289	238	259	208	140	178	179	259
11	174	176	232	256	191	304	181	138	253	112	235	237
12	266	273	244	219	207	232	246	166	236	175	279	231
13	137	250	161	208	228	251	218	218	177	200	215	238
14	119	210	201	144	239	226	168	245	289	284	238	161
15	238	218	226	192	240	156	174	255	178	135	285	58
16	240	190	222	216	255	210	228	220	63	147	270	119
17	262	147	167	210	181	200	170	232	91	257	186	257
18	133	136	280	195	183	178	260	167	246	239	189	207
19	318	206	241	216	200	188	274	253	198	298	211	147
20	259	127	138	183	251	225	183	232	175	188	221	109
21	204	148	185	189	305	242	155	195	126	185	289	203
22	223	280	56	262	286	267	229	172	109	303	284	215
23	188	173	194	199	261	157	200	163	95	340	195	159
24	243	153	169	193	294	129	132	251	139	185	210	153
25	219	209	224	200	260	210	212	233	85	227	197	111
26	206	182	171	206	178	124	212	216	94	233	270	167
27	113	122	150	176	259	129	213	169	251	208	165	180
28	160	246	198	109	268	194	232	179	116	204	234	175
29	205	151	141	176	273	177	142	185	248	245	133	
30	205	164	209	266	129	204	169	120	291	282	160	
31	240	110	263	198	246	280						186

1992

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	143	71	159	250	214	272	192	164	160	177	169	217
2	206	134	238	144	211	152	175	149	192	280	133	178
3	226	91	252	128	199	83	114	218	157	195	169	232
4	182	103	231	81	179	199	107	257	105	158	200	239
5	198	234	214	74	190	195	126	244	183	211	124	222
6	217	342	276	128	126	198	230	180	121	152	130	210
7	188	208	227	347	101	197	173	271	34	169	81	244
8	113	88	218	240	160	122	109	124	178	132	136	209
9	118	163	197	318	176	159	111	103	176	113	210	183
10	167	332	270	187	139	183	156	102	118	197	232	111
11	115	281	178	141	209	184	192	240	73	105	231	231
12	82	191	184	158	212	153	188	231	66	67	198	204
13	149	235	140	160	247	158	240	199	106	149	138	191
14	187	192	288	176	194	157	193	149	162	276	152	85
15	173	258	229	166	254	233	182	142	181	246	146	75
16	62	166	398	180	159	155	210	141	146	206	124	165
17	107	204	165	134	197	196	153	156	129	133	160	266
18	91	310	192	108	221	86	174	152	190	74	165	300
19	188	277	343	124	160	93	167	213	187	52	212	222
20	208	213	360	220	218	166	201	253	167	211	170	218
21	275	206	250	179	184	120	196	266	163	75	175	331
22	123	143	225	235	258	141	182	138	160	79	225	194
23	156	179	229	196	189	211	182	169	51	60	142	279
24	156	165	151	180	151	162	183	231	156	52	175	107
25	71	232	196	169	219	227	134	273	181	157	222	148
26	71	267	267	159	234	216	131	218	193	195	213	141
27	121	89	267	183	229	236	140	203	47	193	169	178
28	97	217	182	240	194	195	195	183	62	207	177	235
29	130	212	188	164	217	149	224	196	89	187	188	254
30	200		247	172	196	203	248	157	96	231	177	338
31	95		210		172		259	169		227		333

1993

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	151	181	182	136	134	142	167	153	65	72	166	177
2	182	182	105	154	100	175	159	118	211	237	94	271
3	150	161	92	96	168	197	117	180	154	50	177	231
4	156	179	236	136	167	231	111	151	147	95	236	188
5	164	187	146	121	154	172	148	75	173	182	244	166
6	221	128	127	107	188	180	101	190	173	167	243	155
7	147	139	116	97	181	145	43	157	113	238	160	214
8	114	131	160	184	157	177	122	95	111	206	206	150
9	175	98	157	144	125	218	207	123	110	181	226	187
10	215	121	236	128	126	166	180	80	151	78	55	226
11	182	266	165	123	162	138	145	97	182	139	278	168
12	243	280	70	155	172	146	232	204	158	138	213	183
13	173	216	199	209	181	156	222	156	275	187	191	289
14	218	201	143	104	165	127	148	147	86	121	158	235
15	223	145	110	203	173	84	193	135	82	208	193	223
16	211	278	233	159	169	169	203	240	65	232	266	144
17	182	222	134	192	186	178	204	88	67	141	55	236
18	167	312	176	170	183	146	169	126	102	231	184	218
19	191	226	200	160	151	157	172	186	60	204	192	222
20	213	216	172	185	231	53	178	115	65	147	190	279
21	221	173	192	116	200	144	212	89	160	58	170	184
22	160	227	124	137	154	153	228	56	64	152	178	295
23	162	238	156	168	110	176	196	133	135	231	180	219
24	102	181	169	185	207	200	177	168	118	169	174	244
25	151	121	185	187	166	114	158	203	191	200	164	149
26	158	284	154	181	180	97	156	167	200	191	200	189
27	190	235	158	206	179	59	191	204	57	170	226	225
28	169	170	107	169	202	119	111	151	66	259	178	181
29	187		118	178	150	96	138	193	85	58	165	148
30	264		120	188	161	87	143	142	55	169	137	187
31	147		211		111		143	172		169		168

FALTA DE ORIGEN
TESIS CON

TESIS CON FALLA DE ORIGEN

1994

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	118	244	233	147	169	169	133	133	101	177	186	151
2	132	176	188	140	182	141	222	178	101	193	149	170
3	206	191	204	132	209	207	144	162	146	180	173	229
4	116	209	187	194	196	170	141	160	118	82	207	180
5	228	194	176	183	182	180	190	212	135	75	142	226
6	236	186	154	184	258	244	213	159	121	200	103	180
7	210	210	190	181	147	235	218	183	191	214	151	209
8	191	236	153	149	178	224	140	178	156	195	187	221
9	169	226	143	200	230	204	135	164	197	86	184	184
10	149	173	192	152	169	131	178	137	200	196	173	157
11	254	187	216	176	234	173	206	56	115	183	219	204
12	241	201	160	154	236	138	193	75	182	215	210	172
13	215	83	149	223	129	109	197	111	171	132	164	243
14	240	240	144	178	188	103	199	93	245	70	173	244
15	205	139	100	158	191	185	138	124	239	138	157	228
16	175	122	142	92	187	196	124	177	129	213	210	196
17	177	189	234	99	171	154	154	198	93	239	207	164
18	156	184	203	127	170	171	82	94	136	249	151	169
19	204	198	200	184	145	161	179	103	105	190	198	203
20	186	165	141	133	175	153	196	225	196	237	167	200
21	209	195	160	133	226	222	166	207	201	146	183	259
22	175	178	175	140	187	148	145	216	210	165	211	214
23	163	133	206	150	134	213	125	156	196	146	53	246
24	194	130	197	129	207	161	155	249	187	136	100	94
25	204	177	222	148	200	187	184	165	63	163	200	143
26	219	222	158	190	230	147	212	95	142	116	212	155
27	199	166	105	149	213	154	184	143	182	265	156	142
28	169	147	226	187	143	159	202	186	166	147	182	102
29	138		173	170	117	157	164	217	244	190	164	157
30	96		196	215	207	149	194	103	204	169	86	200
31	227		155		158		169	166		172		158

1995

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	160	182	200	187	153	196	99	200	72	176	249	254
2	171	213	226	139	218	155	161	234	180	165	168	228
3	115	190	174	98	186	160	249	225	103	159	196	206
4	187	134	187	118	207	187	208	172	89	184	188	174
5	182	150	151	129	194	194	194	129	123	185	186	192
6	187	211	128	182	176	266	108	124	81	65	197	145
7	210	164	96	152	191	215	73	157	189	79	164	201
8	191	165	211	232	215	157	71	183	177	147	99	200
9	236	187	213	154	250	244	61	115	194	198	190	168
10	172	189	169	113	214	187	60	195	134	173	207	143
11	173	116	167	149	213	165	67	132	172	137	84	199
12	187	138	129	201	236	148	156	95	216	160	128	107
13	193	164	175	197	196	177	79	103	186	169	142	220
14	213	222	201	133	151	194	79	178	187	46	151	232
15	191	177	177	146	200	203	211	134	173	88	205	212
16	199	191	208	117	210	237	158	143	172	119	213	118
17	234	227	178	205	181	108	145	177	126	243	215	69
18	173	195	233	186	174	120	194	234	92	232	169	100
19	195	178	194	199	202	123	124	98	167	181	171	120
20	235	138	163	143	173	151	159	57	108	48	107	147
21	234	182	178	196	169	253	239	200	85	183	35	63
22	164	211	181	197	189	192	178	214	244	185	215	95
23	187	123	289	177	206	292	171	172	235	237	142	83
24	229	184	186	225	195	167	161	162	143	147	84	144
25	69	222	211	144	169	148	164	213	227	185	117	135
26	242	170	181	175	187	176	85	220	215	197	180	46
27	210	220	218	169	191	184	203	163	196	198	247	188
28	188	221	173	182	156	104	140	153	99	183	207	156
29	171		188	193	188	143	169	71	201	164	200	108
30	155		181	151	218	163	177	79	209	204	240	44
31	81		182		137		295	226		159		58

1996

Dia/Mes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	73	108	133	195	176	163	204	178	166	111	263	191
2	76	147	147	160	201	91	243	164	238	68	111	139
3	180	161	151	127	179	89	216	172	209	173	48	155
4	179	197	156	126	208	173	203	168	219	159	192	179
5	205	151	205	85	145	179	202	151	221	192	217	174
6	240	158	196	86	202	194	223	128	137	91	187	236
7	226	177	186	55	179	194	139	171	238	155	172	123
8	174	158	186	71	210	77	98	151	185	133	74	101
9	222	162	194	191	171	124	191	46	134	85	216	191
10	214	152	124	193	118	198	169	146	167	102	193	170
11	199	151	129	135	157	180	209	207	146	88	133	160
12	182	115	198	119	127	167	187	149	181	207	112	139
13	199	203	166	133	160	142	139	203	211	196	74	115
14	245	199	151	98	196	183	178	131	238	199	168	204
15	231	223	200	133	183	168	118	178	191	256	182	149
16	218	215	104	194	173	179	203	129	136	188	159	151
17	191	112	101	200	161	164	187	123	175	156	135	143
18	244	155	187	154	147	182	143	127	193	124	129	63
19	269	225	188	147	112	193	91	111	224	235	218	45
20	236	167	188	178	135	228	189	117	220	153	204	75
21	164	137	169	169	157	229	165	95	168	236	147	213
22	182	176	154	189	166	161	208	91	124	133	154	159
23	160	226	111	77	197	89	178	53	139	117	164	185
24	191	245	101	192	230	97	218	121	191	138	158	125
25	229	141	137	192	173	110	157	142	144	85	164	148
26	190	183	129	178	155	107	142	121	179	239	145	138
27	113	177	89	147	169	116	172	210	248	189	184	163
28	199	218	128	172	196	109	62	152	61	154	203	168
29	162	182	92	142	177	37	55	231	48	274	161	135
30	173		142	169	157	81	181	198	59	262	148	114
31	161		147		151		212	162		223		156

1997

Dia/Mes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	158	138	130	161	168	156	111	156	149	50	153	164
2	124	71	202	187	164	208	90	171	137	71	118	145
3	184	142	182	123	120	186	133	155	59	156	155	211
4	170	207	213	59	59	155	163	104	83	168	216	174
5	173	99	149	82	160	166	141	186	103	156	186	194
6	111	178	66	140	173	237	138	237	144	189	223	212
7	85	173	177	198	186	196	159	178	152	173	173	192
8	70	129	145	169	156	222	215	194	187	166	112	225
9	80	139	162	178	191	180	199	166	158	45	90	179
10	66	107	190	71	171	213	150	163	186	172	147	194
11	89	160	184	103	156	204	173	150	164	117	59	75
12	85	90	156	178	183	196	164	134	233	67	112	37
13	126	141	130	77	171	174	147	173	171	73	102	36
14	103	199	174	75	119	123	182	95	143	35	244	142
15	111	180	114	142	144	177	190	197	164	112	218	193
16	151	129	142	216	150	180	193	194	118	116	133	200
17	176	150	173	179	179	223	204	98	164	107	176	93
18	74	224	189	164	142	123	133	182	160	190	196	47
19	168	173	129	88	192	139	194	225	233	201	142	107
20	185	98	156	141	143	159	182	204	116	194	170	64
21	159	216	130	130	139	133	222	125	162	207	178	130
22	226	157	95	216	175	174	156	170	133	250	156	85
23	194	149	164	185	226	124	180	167	204	187	115	76
24	176	195	119	139	199	112	178	136	169	235	173	75
25	149	209	127	112	195	155	231	156	49	200	115	114
26	157	146	169	107	204	176	202	167	166	211	214	182
27	201	165	126	109	218	123	151	160	270	145	177	153
28	144	147	124	194	129	133	151	227	165	173	131	151
29	128		116	262	126	151	149	167	196	202	168	156
30	144		127	227	200	150	163	217	143	225	170	205
31	196		147		136	148	170		207			147

FALTA DE ORIGEN
TESIS CON

TESIS CON
FALLA DE ORIGEN

1998

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	160	71	120	149	204	234	190	204	170	125	187	130
2	147	62	126	133	213	208	221	122	201	178	147	205
3	153	168	109	185	175	219	55	144	107	165	178	262
4	173	195	104	214	176	174	120	151	70	144	188	200
5	176	56	196	192	198	251	160	206	131	209	194	180
6	118	202	129	109	229	158	137	236	148	243	121	186
7	74	167	75	102	223	139	153	67	76	103	196	169
8	89	187	142	98	199	213	182	181	62	54	144	166
9	137	174	153	224	248	175	73	108	143	59	173	138
10	218	193	195	142	192	166	57	126	179	126	218	146
11	193	147	178	158	182	172	169	183	81	55	124	214
12	148	160	173	177	196	207	172	171	91	67	240	172
13	129	121	152	211	209	220	213	180	138	138	176	140
14	120	56	178	167	196	153	221	229	159	111	234	75
15	176	76	87	190	197	200	164	181	160	193	113	182
16	162	114	53	202	186	168	99	181	128	218	129	161
17	126	168	124	202	126	204	136	174	173	194	158	153
18	114	115	99	145	197	170	160	198	160	168	166	216
19	55	208	115	177	191	173	166	133	151	107	148	214
20	122	165	182	152	169	171	139	65	177	96	138	183
21	172	65	95	166	169	184	163	114	207	52	95	158
22	129	177	92	194	192	141	180	162	225	45	150	178
23	140	147	149	166	219	136	129	136	183	65	144	168
24	111	220	166	192	236	173	161	203	149	66	59	197
25	139	101	169	182	251	199	128	191	169	54	144	125
26	157	103	173	134	220	232	72	195	191	81	85	138
27	171	106	227	146	207	179	106	136	120	108	76	128
28	138	127	205	165	194	84	173	179	158	184	179	124
29	155	184	107	212	201	201	141	169	160	169	134	114
30	124	116	129	151	138	198	136	102	231	132	156	
31	181	244	151	156	187	214	171					

1999

DiaMes	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
1	96	177	184	156	164	195	132	138	192	52		
2	131	192	207	128	137	131	102	123	117	189		
3	72	159	133	123	147	131	55	164	97	83		
4	140	159	191	95	126	159	42	196	170	31		
5	191	173	177	187	173	172	78	147	74	25		
6	188	172	163	182	187	157	104	209	80	40		
7	135	151	178	157	160	84	173	180	138	178		
8	120	140	157	238	177	132	160	145	108	190		
9	165	213	132	182	125	180	177	191	160	153		
10	164	200	196	192	189	227	163	148	35	57		
11	262	200	242	170	204	151	118	129	143	62		
12	236	57	172	168	191	141	30	147	147	68		
13	214	91	96	175	170	129	119	190	36	66		
14	170	158	157	150	206	108	151	194	36	131		
15	185	156	197	132	161	116	113	133	45	272		
16	108	190	187	127	164	95	178	99	61	185		
17	107	226	183	192	160	75	161	156	59	173		
18	112	206	143	145	125	84	94	136	164	163		
19	194	203	160	133	141	164	127	82	156	157		
20	170	145	209	129	146	149	101	191	97	35		
21	186	61	128	202	196	161	124	132	78	134		
22	153	134	149	215	173	179	148	169	96	148		
23	104	149	110	202	124	178	165	218	156	157		
24	85	150	175	196	142	109	157	203	118	103		
25	143	148	160	196	174	56	157	124	101	99		
26	175	210	168	182	178	34	188	123	105	185		
27	192	181	111	196	164	114	196	62	85	199		
28	169	110	85	188	169	163	168	127	55	218		
29	159	204	173	150	173	165	113	35	220			
30	191	170	226	126	157	203	131	35	142			
31	151	227	185	202	193	184						

Análisis de Intervención: Aplicación a un problema ambiental