U N I V E R I S I D A D D E S O N O R A

DEPARTAMENTO DE MATEMATICAS

"MODELOS MATEMATICOS PARA PRONOSTICO DE TEMPERATURAS AMBIENTALES"

T E S I S

QUE PARA OBTENER EL GRADO DE:

L I C. E N M A T E M A T I C A S

P R E S E N T A:

JOSE MARIA BRAVO TAPIA

HERMOSILLO, SONORA OCT. DE 1992

D E D I C A T O R I A

A MI ESPOSA ANA ALICIA

A MIS HIJOS JOSE IVAN Y ANA ALEJANDRA

A MIS PADRES

A MIS HERMANOS

A TODA MI FAMILIA

A TODOS MIS AMIGOS

AGRADECIMIENTOS

Quiero dejar patente mi agradecimiento a aquellas personas que hicieron posible la realización del presente trabajo. Particularmente al Ing. Carlos Jesús Arias quien me sugirió el tema y puso a mi disposición datos recabados por la Comisión Nacional del Agua; al Dr. J. Enrique Ramos quien además de brindarme la asesoría necesaria me motivo continuamente para la conclusión del mismo.

A todos mis compañeros de trabajo, tanto en la Comisión Nacional del Agua como de la Universidad de Sonora, quienes me suministraron datos e ideas y soportaron mi constante incertidumbre. Muy especialmente a los profesores Angelina Uribe Araujo, Dra. Marisa Jasa, M.C. Gerardo Gutierrez, al Ing. Carlos Anaya, Ing. Agr. Jesús Lopez Elías y al M.C. Alfredo Serrano.

INDICE

Cap 1 INTRODUCCIÓN 1

Cap 2 DISTRIBUCIÓN DE DATOS DE HELADAS Y PERÍODOS LIBRES DE HELADAS PARA SERIES CLIMATOLÓGICAS CON AÑOS SIN HELADAS

Introducción 4

2.1 **Función de distribución para helada** 5

2.2 **Estimación** 8

2.3 **Estimación a través de la distribución normal** 11

2.4 **Fecha media de helada** 15

2.5 **Media del período libre de helada** 15

Cap 3 METODOLOGIA DE BOX-JENKINS

Introducción 18

3.1 **Identificación del modelo** (datos no estacionales) 20

3.1.1 Series de tiempo estacionarias y no estacionarias 21

3.1.2 Uso de la FACM para la identificación del modelo 30

3.2 **Identificación del modelo** (datos estacionales) 35

3.2.1 Análisis de la componente estacional 35

3.2.2 Uso de la FACM y la FACPM para la identificación del modelo 36

3.3 **Estimación** 42

3.3.1 Condición de estacionariedad e invertibilidad 42

3.3.2 Estimación preliminar 45

3.3.3 Estimación 45

3.4 **Comprobación y Diagnóstico** 46 48

3.4.1 Comprobación 48

3.4.2 Diagnóstico 51

3.5 **Pronóstico** 53

Cap 4 APLICACIONES Y RECOMENDACIONES

4.1 **Adolfo Ruiz Cortínez** 58

4.1.1 Metodología de Thom 58

4.1.2 Metodología de Box-Jenkins 58

4.2 **Agua Prieta** 64

4.2.1 Metodología de Thom 64

4.2.2 Metodología de Box-Jenkins 68

4.3 **Metodología de Thom** (para las Est. restantes) 71

4.4 **Metodología de Box-Jenkins** 72

Cap 5 CONCLUSIONES Y RECOMENDACIONES

5.1 **Justificación del uso de la Temp. Min. Extrema** 78

5.2 **Recomendaciones para el uso de la metodología de Thom** 80

5.3 **Relación entre los métodos** 81

Anexo 1 82

Anexo 2 84

Tabla 1a 85

Tabla 1b 86

Bibliografía 88

CAPITULO 1

I N T R O D U C C I O N

PLANTEAMIENTO DEL PROBLEMA.

La Comisión Nacional del Agua (C.N.A.), sección de Hidroclimatología, de a la Secretaria de Agricultura y Recursos Hidráulicos (S. A. R. H), Delegación Estatal Sonora, tiene como una de sus principales funciones la recopilación de datos Climatológicos e Hidrológicos.

Dicha dependencia realiza regularmente esta función desde 1958. Actualmente cuenta con 186 estaciones climatológicas en el Estado. Si bien es cierto que se ha avanzado bastante en el registro de este tipo de información, a la fecha su análisis e interpretación ha sido prácticamente nulo.

Este trabajo pretende establecer una metodología para estudiar el comportamiento del fenómeno de las heladas, a partir de la información registrada en esas estaciones [ver el anexo 1]. Para ello se procedió de la siguiente manera:

Primero, de acuerdo a la revisión bibliográfica existe una metodología para el análisis adecuado del fenómeno de las heladas;

ésta consiste en determinar el Período Libre de Helada (P. L. H.).

Segundo, partiendo de la definición de una helada [ver la definición 2.1] estableceremos una función matemática estadística que nos permita modelar con la mayor precisión posible el comportamiento de la serie de datos de las temperaturas mínimas del mes.

Para una presentación clara de este trabajo, en el siguiente capítulo desarrollamos la metodología de análisis para el P. L. H.

En el capítulo 3 presentamos la metodología de Box and Jenkins aplicada a la serie de temperaturas mínimas.

Con el propósito de validar estas dos metodologías, en el capítulo 4 realizamos la aplicación de éstas en seis estaciones climatológicas con diferente clasificación climática.

Por último, en el capítulo 5 damos algunas conclusiones y recomendaciones.

OBJETIVO GENERAL

Incrementar el uso del análisis estadístico y matemático de la información climatológica en la toma de decisiones agrícolas.

OBJETIVOS PARTICULARES

Estimación de las Frecuencias de Distribución de las Heladas, antes y después de una fecha.

Estimación del Período Libre de Heladas Promedio para diferentes estaciones climatológica del Estado de Sonora.

Pronóstico de temperaturas mínimas para un año, en las mismas estaciones climatológicas.

CAPITULO 2

DISTRIBUCION DE DATOS DE HELADAS Y PERIODO LIBRE DE HELADA PARA SERIES CLIMATOLOGICAS CON AÑOS SIN HELADAS

INTRODUCCION

Las bajas temperaturas ambientales están consideradas como uno de los fenómenos meteorológicos que más afectan la actividad del hombre, en casos particulares suele referirse a esos fenómeno como heladas; Para efectos de este trabajo decimos:

Definición 2.1. Una helada ocurre cuando la temperatura del aire es menor o igual a 0°C.

Los estudios que se han desarrollado en torno a este fenómeno son muy variados, obteniéndose los mejores resultados cuando estos se abordan desde el punto de vista estocásticos y uno de los objetivos ha sido establecer Períodos Libres de Heladas (P. L. H.); así, es en esos términos que desarrollaremos el trabajo de este capítulo, en el cual discutiremos el método que permite calcular el P. L. H., realizado por Thom [1], y llevado a la práctica por Grassi [2], para los estados de Puebla y Tlaxcala.

En la siguiente sección se desarrolla la función de distribución de la ocurrencia de una helada. Una estimación no paramétrica de la ocurrencia de una helada se presenta en la sección 2.2. En la sección 2.3, la misma estimación se realiza usando la distribución normal; el cálculo de la fecha media de helada se presenta en la sección 2.4 y, finalmente, en la sección 2.5 se presenta la fórmula para calcular el P. L. H. promedio.

2.1. FUNCION DE DISTRIBUCION PARA HELADAS

La metodología expuesta por Thom para determinar el P. L. H., empieza por establecer dos tipos de información, que son:

Helada Temprana.- Fecha en que ocurre la primera helada en el año, una vez iniciado el otoño del año en curso.

Helada Tardía.- Fecha en que ocurre la última helada en el año, antes o durante la primavera del año en curso.

En general, definir estos tipos de helada depende mucho del lugar geográfico donde ocurren. Por eso, hemos considerado más apropiado definirlas en términos de su ocurrencia en el Estado de Sonora. Y en el mismo sentido, podemos decir que una helada se considera tárdia si se presenta del 1ero. de enero al 30 de junio y temprana si se presenta del 1ero de julio al 31 de diciembre.

Según Thom [1], las fechas de ocurrencia de una helada

son valores de una variable aleatoria, por ello podemos encontrar una función de distribución para las mismas. El modelo para determinar la probabilidad de que ocurra una helada en una fecha determinada (de cualquiera de las dos tipos de heladas), es en base a una distribución que involucra una distribución discreta, que describe la ocurrencia o no de una helada durante el año, y una distribución esencialmente continua que describe las fechas de ocurrencia de helada, dado que en el año ha ocurrido una.

Lo que resta de esta sección lo dedicaremos a construir las distribuciones antes mencionadas.

En general una función de distribución continua queda definida por:

$$F(x)=\int_{-\infty}^{x} \ell(u)\,d(u) \qquad (2.1.1)$$

donde $\ell(u)$ es la función de densidad de probabilidades con $F(-\infty) = 0$ y $F(\infty) = 1$. Aquí $F(x)$es la probabilidad de que u sea menor que x. Por lo que podemos definir:

$F_S(x)$ como la función de distribución para las fechas de heladas en el primer semestre y que indica la probabilidad de que ocurra una helada antes de la fecha x en el primer semestre, en los años que éstas ocurren.

$F_A(x)$ como la función de distribución para las fechas de

heladas en el segundo semestre y que indica la probabilidad de que ocurra una helada antes de la fecha x en el segundo semestre, en los años en que éstas ocurren.

De acuerdo al concepto de las distribuciones **mixtas** citado por Thom [1], la función de distribución G(x) para las heladas Tardía o Ultimas heladas es:

$$G(x) = Q_s + P_s F_S(x), \qquad (2.1.2)$$

que representa la probabilidad de que no ocurra una helada en el primer semestre o que ocurra una helada antes de la fecha x. Aquí Q_s representa la probabilidad de que no ocurra una helada en el primer semestre y $1 - Q_s = P_s$ representa la probabilidad de que ocurra una helada en el primer semestre. En el caso del primer semestre es de mayor interés calcular la probabilidad de que ocurra una helada después de una fecha x, por lo que conviene presentar a las últimas heladas de la siguiente manera:

$$H(x) = 1 - G(x). \qquad (2.1.3)$$

sustituyendo (2.1.2) en (2.1.3) obtenemos,

$$H(x) = 1 - (Q_s + P_s F_S(x)). \qquad (2.1.4)$$

Puesto que

$$P_s + Q_s = 1,$$

se tiene que

$$H(x) = P_s(1 - F_S(x))$$

y si

$I(x) = 1 - F_S(x)$, entonces

$$H(x) = P_s I(x). \qquad (2.1.5)$$

La expresión anterior representa la función de distribución de las últimas heladas y calcula la probabilidad de que ocurra una helada después de la fecha x.

Para las heladas Tempranas o Primeras heladas tenemos una función mixta similar

$$J(x) = P_a F_A(x) \qquad (2.1.6)$$

que representa la función de distribución de las Primeras heladas y calcula la probabilidad de que ocurra una helada antes de la fecha x. Aquí P_a es la probabilidad de que ocurra una helada en el segundo semestre.

2.2. ESTIMACION

Dada una fecha x cualquiera, para estimar las

probabilidades tratadas en la sección anterior para una estación climatolígica, las fechas observadas se separan en dos grupos: las relativos a heladas Tardías y las relativos a heladas Tempranas, en cada grupo estas se ordenan en forma creciente independientemente del año en que ocurrieron y, deacuerdo a ese orden se enumeran (Ver Cuadro 2.1).

La primera estimación de H y de J es empírica o no paramétrica, para lo cual debemos hacer la estimación de I y F_A; para estimar estas últimas, un criterio propuesto es uno que se basa en los valores K (el número correspondiente en el orden) y M (el número de años con heladas). Esa estimación se hace apartir del valor K/M, siempre que M sea lo suficientemente grande; para corregir el sesgo, con probabilidades muy grandes o muy pequeñas, se usan:

$$I^* = 1 - F_S^*$$

o bien

$$I^* = 1 - \frac{K_S}{M_S + 1} \qquad (2.2.1)$$

$$F_A^* = \frac{K_A}{M_A + 1} \; . \qquad (2.2.2)$$

Así, las ecuaciones anteriores representan las estimaciones no paramétricas de F e I. Para una completa estimación de H y J, estimaremos los valores P_a y P_S de la manera siguiente:

$$\hat{P}_S = \frac{M_S}{N_S} \qquad \hat{P}_A = \frac{M_A}{N_A}; \qquad (2.2.3)$$

Cuadro 2.1. Estimación de la probabilidad de helada después de una fecha dada en la serie de helada TARDIA y antes de una fecha en la serie de helada TEMPRANA, siguiendo un método no paramétrico y otro teórico a través de la distribución normal.

E S T A C I O N C L I M A T O L O G I C A
S A N I S I D R O

H E L A D A T A R D I A

K	FECHA	X	$\hat{I}^{*}(X)$	$\hat{I}(X)$	$H^{*}(X)$	$\hat{H}(X)$	|d|
1	Ene 6	6	0.969	0.961	0.940	0.932	0.007
2	Ene 7	7	0.938	0.958	0.909	0.929	0.019
3	Ene 8	8	0.906	0.954	0.879	0.925	0.046
4	Ene 20	20	0.875	0.880	0.849	0.854	0.005
5	Ene 22	22	0.844	0.862	0.818	0.836	0.018
6	Ene 24	24	0.813	0.843	0.788	0.818	0.029
7	Ene 24	24	0.781	0.843	0.758	0.818	0.060
8	Ene 26	26	0.750	0.822	0.728	0.797	0.070
9	Feb 2	33	0.719	0.735	0.667	0.713	0.015
10	Feb 2	33	0.688	0.735	0.667	0.713	0.046
11	Feb 3	34	0.656	0.721	0.637	0.699	0.062
12	Feb 3	34	0.625	0.721	0.606	0.699	0.093
13	Feb 6	37	0.594	0.677	0.576	0.656	0.080
14	Feb 17	48	0.563	0.498	0.546	0.483	0.063
15	Feb 18	49	0.531	0.481	0.515	0.467	0.049
16	Feb 22	53	0.500	0.415	0.485	0.402	0.083
17	Feb 24	55	0.469	0.382	0.455	0.371	0.084
18	Feb 25	56	0.438	0.366	0.424	0.355	0.069
19	Feb 25	56	0.406	0.366	0.394	0.355	0.039
20	Feb 26	57	0.375	0.350	0.364	0.340	0.024
21	Feb 27	58	0.344	0.335	0.333	0.325	0.009
22	Feb 28	59	0.313	0.320	0.303	0.310	0.007
23	Mar 6	65	0.281	0.235	0.273	0.228	0.045
24	Mar 10	69	0.250	0.187	0.243	0.181	0.061
25	Mar 10	69	0.219	0.187	0.212	0.181	0.031
26	Mar 11	70	0.188	0.176	0.182	0.170	0.012
27	Mar 15	74	0.156	0.136	0.152	0.131	0.020
28	Mar 16	75	0.125	0.127	0.121	0.123	0.002
29	Mar 22	81	0.094	0.081	0.091	0.079	0.012
30	Mar 31	90	0.063	0.038	0.061	0.037	0.024
31	Abr 2	92	0.031	0.032	0.030	0.031	0.000

En donde:

$N = 32$, $M = 31$, $P = 0.969$ y

$P(\text{Max } |d| \geq 0.093) > P(\text{Max } |d| \geq 0.240) = 0.05$ con $\alpha = 0.05$

Cuadro 2.1. Cont.

E S T A C I O N C L I M A T O L O G I C A
S A N I S I D R O

H E L A D A T E M P R A N A

K	FECHA	X	$F^*(X)$	$\hat{F}(X)$	$J^*(X)$	$\hat{J}(X)$	|d|
1	Nov 19	323	0.042	0.026	0.030	0.018	0.012
2	Nov 20	324	0.083	0.032	0.060	0.023	0.037
3	Nov 27	331	0.125	0.111	0.090	0.080	0.010
4	Nov 29	333	0.167	0.149	0.120	0.107	0.013
5	Dic 2	336	0.208	0.221	0.150	0.159	0.009
6	Dic 3	337	0.250	0.249	0.180	0.179	0.001
7	Dic 5	339	0.292	0.310	0.210	0.223	0.013
8	Dic 7	341	0.333	0.376	0.240	0.271	0.031
9	Dic 8	342	0.375	0.411	0.270	0.296	0.026
10	Dic 9	343	0.417	0.447	0.300	0.321	0.021
11	Dic 10	344	0.458	0.483	0.330	0.347	0.017
12	Dic 10	344	0.500	0.483	0.360	0.347	0.013
13	Dic 11	345	0.542	0.519	0.390	0.373	0.017
14	Dic 12	346	0.583	0.555	0.420	0.399	0.021
15	Dic 13	347	0.625	0.590	0.450	0.425	0.025
16	Dic 13	347	0.667	0.590	0.480	0.425	0.055
17	Dic 14	348	0.708	0.625	0.510	0.450	0.060
18	Dic 14	348	0.750	0.625	0.540	0.450	0.090
19	Dic 24	358	0.792	0.890	0.570	0.641	0.071
20	Dic 26	360	0.833	0.920	0.600	0.663	0.063
21	Dic 27	361	0.875	0.933	0.630	0.672	0.042
22	Dic 28	362	0.917	0.944	0.660	0.680	0.020
23	Dic 30	364	0.958	0.962	0.690	0.692	0.002

En donde:

$N = 32$, $M = 23$, $P = 0.719$ y

$P(\text{Max } |d| \geq 0.090) > P(\text{Max } |d| \geq 0.275) = 0.05$ con $\alpha = 0.05$

donde N es el número de años en que se hicieron las observaciones. Y con todas estas estimaciones obtenidas se pueden obtener las de H y J, de la siguiente manera:

$$H^*(x) = \hat{P}_s \, I^*(x) \qquad (2.2.4)$$

$$J^*(x) = \hat{P}_a \, F_A^*(x), \qquad (2.2.5)$$

que representan las estimaciones no paramétricas de las Ultimas y Primeras heladas respectivamente. Los resultados se presentan en el Cuadro 2.1.

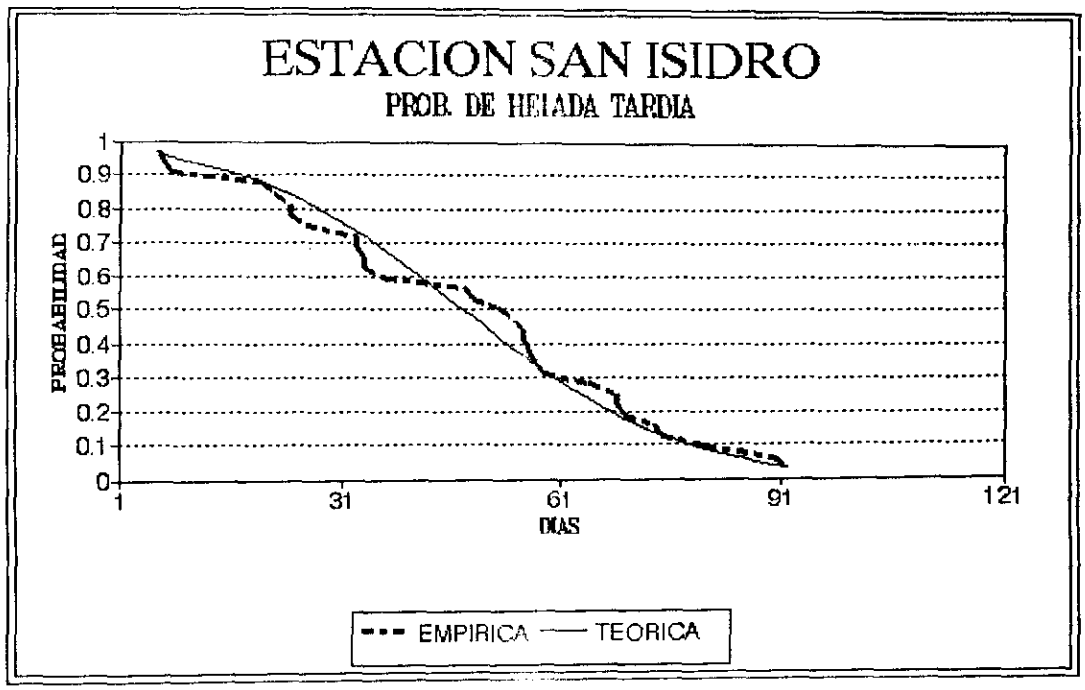
2.3. ESTIMACION A TRAVES DE LA DISTRIBUCION NORMAL

Thom [1] no recomienda las estimaciones no paramétricas para ser usadas en el cálculo de probabilidades cuando en su lugar podemos utilizar una distribución teórica. Sin embargo, las estimaciones experimentales son necesarias para evaluar el ajuste de la distribución teórica.

Thom [1] menciona que Reed [3], Thom y Sham [4], hallaron que la distribución normal provee muy buen ajuste a las series de fechas, en un amplio rango de condiciones, para valores de temperatura que se presentaron en todos los años observados (ya sean en período Tardío o Temprano). A este tipo de series se les llama Completas.

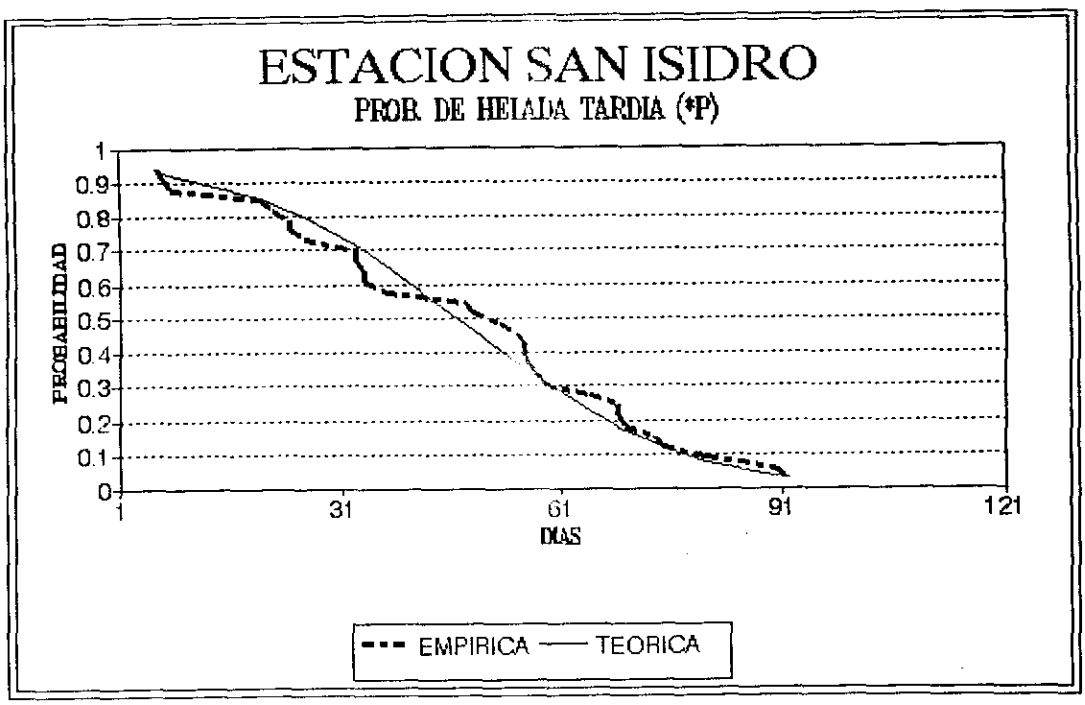
El ajuste a través de la distribución normal para las series de las helada Tardía y helada Temprana, $I(x)$ y $F(x)$ se desarrolla de la manera usual: por estimación de la media y la desviación estándar. Estas distribuciones, estimadas teóricamente, respectivamente se denotan como $\hat{I}$ y $\hat{F}$ (Ver el Cuadro 2.1). Con éstos valores y $\hat{P}$ se calculan $\hat{H}$ y $\hat{J}$.

Fig 2.1



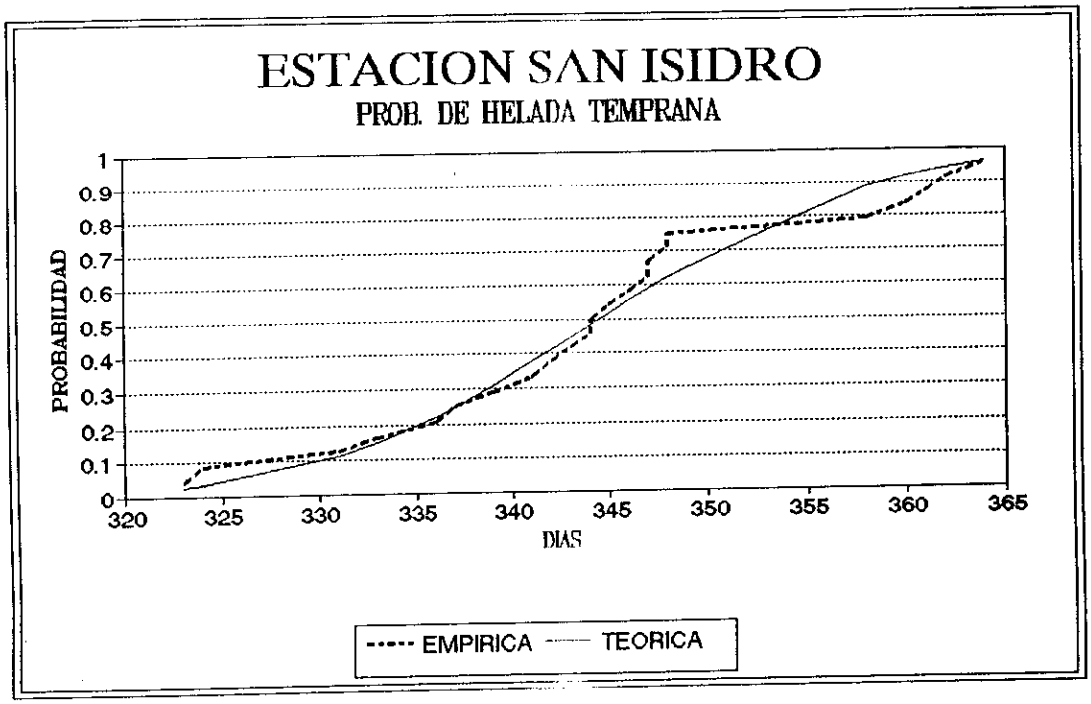
Representación Gráfica I^* e $\hat{I}$

Fig 2.2



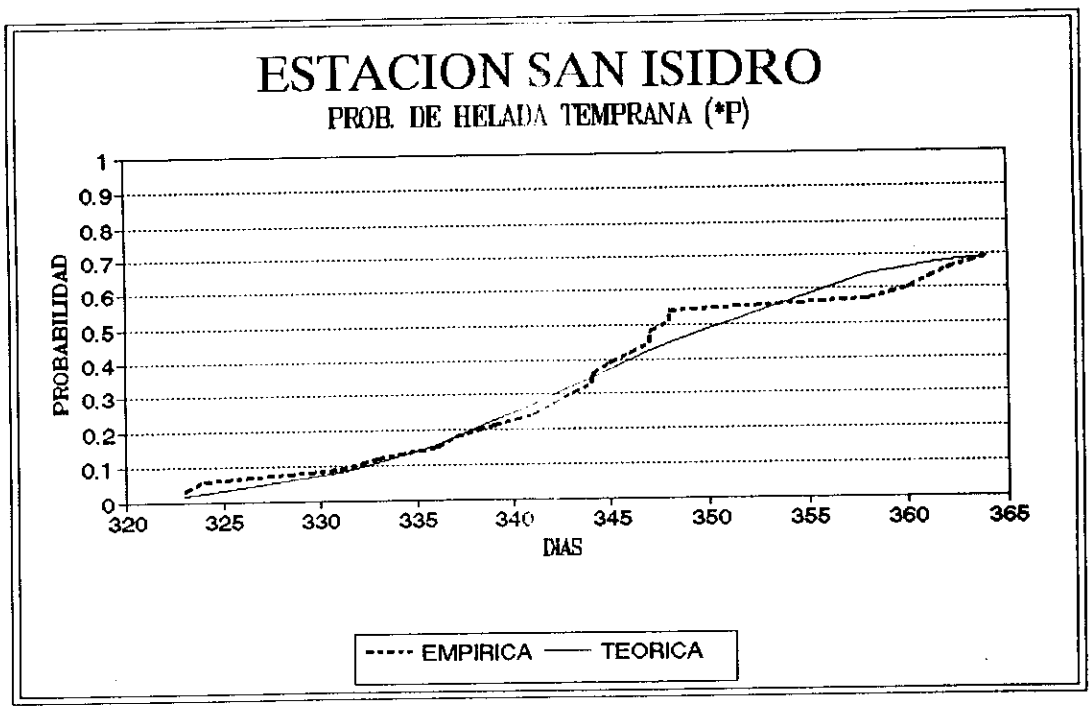
Representación Gráfica de H^* y $\hat{H}$

Fig 2.3



Representación Gráfica de F^* y $\hat{F}$

Fig. 2.4



Representación Gráfica de J^* y $\hat{J}$

Aunque el ajuste de la distribución normal parezca ser satisfactorio (Ver las Figs 2.1, 2.2, 2.3, 2.4), es conveniente realizar una comparación más eficaz sobre la concordancia entre la distribución normal y la distribución no paramétrica. Para ello utilizamos la prueba Kolmogorov-Smirnov [5], determinando la Diferencia Máxima ($|\mathbf{d}|$) entre la distribución normal y la distribución no paramétrica, la cual se define como:

$$P(\text{Max}|\mathbf{d}| \leq c) = 1 - \alpha.$$

donde $\mathbf{d} = |^{*} - |^{\wedge}$ en el primer semestre y $\mathbf{d} = F^{*} - F^{\wedge}$ en el segundo semestre. α es el nivel de significancia.

La aplicación de esta prueba en el ejemplo anterior demostró que el ajuste es satisfactorio. En el Cuadro 2.1, al pie de cada serie se muestra los resultados de esta aplicación.

2.4. FECHA MEDIA DE HELADA.

A partir de que es posible enumerar cada una de las fechas del año de acuerdo a su orden (días Juliano), podemos hablar de número de día del año, de número de día en que se presenta una helada Tardía o Temprana y por lo tanto el número de día esperado en que se puede presentar una helada Tardía o Temprana; a este último se le llama fecha Media de Helada, una para la Tardía y otro para la Temprana.

El cálculo de las fechas Medias de Heladas para una serie incompleta, en el sentido de que no ocurren heladas en todos los años, es más difícil de interpretar que cuando la serie es completa. El desarrollo teórico planteado por Thom [1], conduce a las siguientes expresiones:

$$\bar{u} = P_s \bar{x}_S \qquad (2.4.1)$$

$$\bar{v} = 365 Q_a + P_a \bar{x}_A, \qquad (2.4.2)$$

donde $\bar{x}_S$ y $\bar{x}_A$ son las fechas medias de la serie de helada Tardía y helada Temprana, respectivamente, sobre los años en que ocurrieron heladas; y $\bar{u}$ y $\bar{v}$ son las fechas esperadas para cada distribución, sobre la serie total de años.

2.5. PERIODO LIBRE DE HELADA PROMEDIO

El Período Libre de Helada Promedio es el número esperado de días comprendidos entre la fecha media de helada Tardía y la fecha media de la helada Temprana, queda determinado por la ecuación:

$$\bar{Y} = 365 Q_a + P_a \bar{x}_A - P_s \bar{x}_S \qquad (2.6.1)$$

donde $\bar{Y}$ representa el promedio del número de días Libres de Helada. En el Cuadro 2.2 se presentan los datos ya estimados, que nos serviran para evaluar dicho período.

Cuadro 2.2. Cálculo de la Fecha Media de Helada.

Primer Semestre						
n	m	$\bar{x}$	fecha	$\hat{p}$	$\bar{u}$	fecha
32	31	47.87	16/02	0.97	46.38	15/02

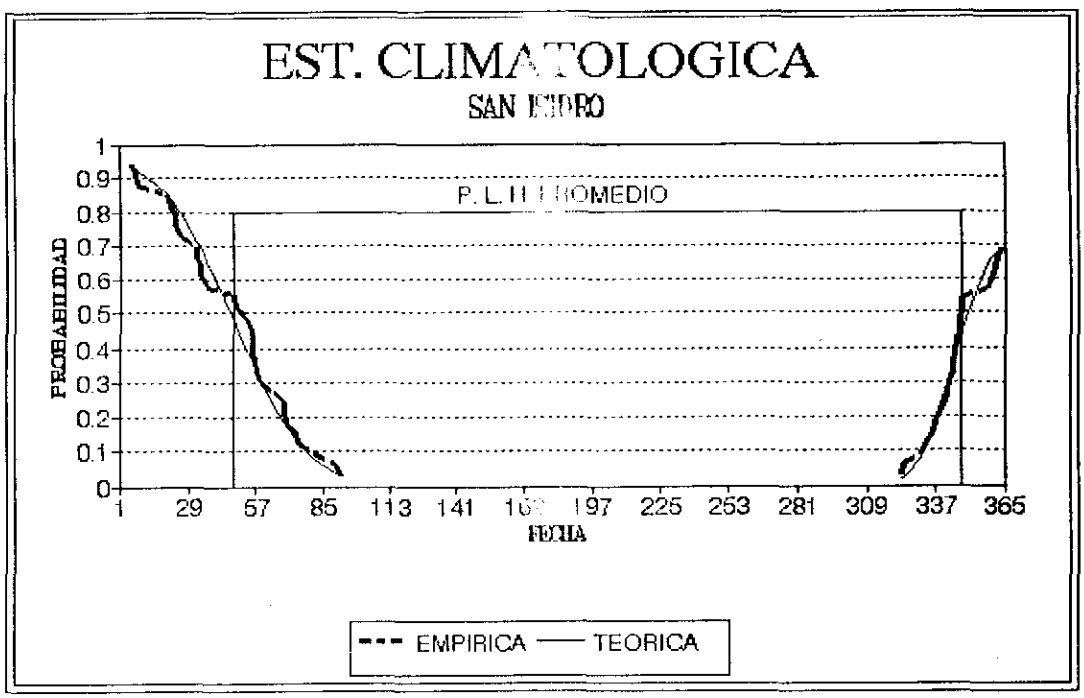
Segundo Semestre						
n	m	$\bar{x}$	fecha	$\hat{p}$	$\bar{v}$	fecha
32	23	344.47	10/12	0.71	350.25	16/12

Tomando los resultados de $\bar{u}$, $\bar{v}$ y sustituyéndolos en la ecuación (2.6.1) se obtiene:

$$\bar{Y} = 303.8 \text{ días},$$

que representan el P. L. H. en promedio (ver la fig 2.5).

Fig 2.5



Período Libre de Heladas Promedio

CAPITULO 3

METODOLOGIA DE BOX-JENKINS

INTRODUCCION

El método para determinar el Período Libre de Helada, expuesto en el capítulo anterior, es útil para calcular en determinado lugar y determinada fecha la probabilidad de que se presente una helada en ese lugar, antes o después de esa fecha.

Sin embargo, el problema de la predicción de la ocurrencia de una helada también puede abordarse desde el punto de vista del estudio completo de las temperaturas que se presentan cada día del año. Sin duda, una de las variables meteorológicas más importantes es la temperatura y más concretamente la temperatura mínima diaria. Aunque este fenómeno se rige por leyes físicas, esa leyes dependen de una gran cantidad de variables muchas de las cuales no se pueden controlar y por lo tanto la ocurrencia o no de este fenómeno se puede considerar aleatorio; de ahí el por qué de la necesidad de buscar métodos estocásticos que nos permitan establecer con mayor claridad su comportamiento.

Los métodos que mejor describen el comportamiento de esa

variables meteorológicas son aquellos que están basados en la Teoría del Análisis de Series de Tiempo. En este capítulo desarrollaremos un modelo para describir el comportamiento de las temperaturas mínima del mes, todo esto, en base a la metodología que Box y Jenkins [1] desarrollaron para la Construcción de Modelos en la descripción del comportamiento de series de tiempo.

Esta metodología consiste en los siguientes pasos:

Paso 1: **Identificación del modelo.** Proponer un modelo tentativo empleando la FunciÓn de Autocorrelación Muestral (FACM) y la Función de Autocorrelación Parcial Muestral (FACPM).

Paso 2: **Estimación de parámetros.** Estimar los parámetros del modelo propuesto.

Paso 3: **Comprobación y Diagnóstico.** Verificar el modelo; si el modelo propuesto resulta inadecuado, el diagnóstico sugiere un nuevo modelo tentativo, que se somete a estos tres pasos y así sucesivamente hasta obtener un modelo adecuado.

Paso 4: **Pronóstico.** Cálculo de valores futuros usando el modelo establecido en los pasos anteriores.

En la sección siguiente describimos el proceso de identificación del modelo para datos no estacionales, en la sección 3.2 presentamos la identificación del modelo para datos estacionales, en la 3.3 se presenta el procedimiento de estimación

de parámetros; la comprobación y diagnóstico del modelo la presentamos en la sección 3.4 y finalizamos el capítulo con la sección 3.5, que contiene el uso del modelo para pronóstico.

3.1. IDENTIFICACION DEL MODELO (Datos no estacionales).

Como el elemento esencial del modelo desarrollado por Box-Jenkins son las series de tiempo, empezaremos por caracterizar a las mismas.

Una serie de tiempo es un conjunto de observaciones generadas secuencialmente en el tiempo y es continua o discreta de acuerdo a que lo sea ese conjunto. En particular, nosotros estamos interesados en las series de tiempo discretas; así, la serie de observaciones en los tiempo $t_1, t_2, \ldots, t_i, \ldots, t_n$ se denotarán por $y_{t_1}, y_{t_2}, \ldots, y_{t_i}, \ldots, y_{t_n}$. Más en particular aún en las series del tipo $y_1, y_2, \ldots, y_i, \ldots, y_n$ que describen observaciones a intervalos de tiempo equidistantes $t_0 + h$, $t_0 + 2h$, $\ldots$, $t_0 + ih$, $\ldots$, $t_0 + nh$, donde t_0 es la observación inicial y h la unidad de tiempo.

Una serie de tiempo puede considerarse como la realización de un proceso estocástico $y_{t_1}, y_{t_2}, \ldots, y_{t_n}$ donde y_{t_i} son variables aleatorias con cierta distribución de probabilidad. En este trabajo tomamos este enfoque de la series de tiempo.

El proceso de identificación del modelo Box-Jenkins comienza por determinar si la serie de tiempo con la cual se trabaja es estacionaria o no estacionaria; de no ser estacionaria se hace una transformación sobre la serie para lograr la estacionariedad. Después, con el uso de la FACM y la FACPM obtenemos una aproximación de los valores p, d, q, del modelo general de Box-Jenkins definido como:

$$\Phi_p(B)\nabla^d Y_t = \delta + \Theta_q(B)a_t. \qquad (3.1.1)$$

Una vez que se obtiene el modelo que representa a la serie, se hace una estimación tentativa de sus parámetros.

3.1.1. Series de tiempo estacionarias y no estacionarias.

Una clase importante de modelos estocásticos que describen a una serie de tiempo son los llamados modelos estacionarios, refiriéndose a aquelos procesos en los cuales sus propiedades no se ven afectadas por un cambio en el tiempo de la observación inicial; es decir, la función de distribución de probabilidad conjunta asociada a las observaciones y_{t_1}, y_{t_2}, ..., y_{t_i}, ..., y_{t_n} hechas en los tiempos t_1, t_2, ...,t_i, ..., t_n es la misma que la asociada a las observaciones y_{t_1+k}, y_{t_2+k}, ..., y_{t_i+k}, ..., y_{t_n+k}. En particular, los modelos de Box-Jenkins pertenecen a estas clases, es decir, basan sus teorías sobre series de tiempo definidas como estacionarias, las cuales tienen

la característica de que la distribución conjunta de las observaciones no debe verse afectada por un cambio en todos los tiempos de observación, adelantados o retrasados en k unidades de tiempo. Así tenemos:

Definición 3.1. Diremos que una serie de tiempo es estacionaria si el proceso estocástico asociado a ella es estacionario, es decir, si la media, la varianza y la autocovarianza satisfacen:

$$\mu = \mathbf{E}\ [Y_t] = \mathbf{E}\ [Y_j]$$

$$\gamma_0 = \sigma^2_{Y_t} = \sigma^2_{Y_j}$$

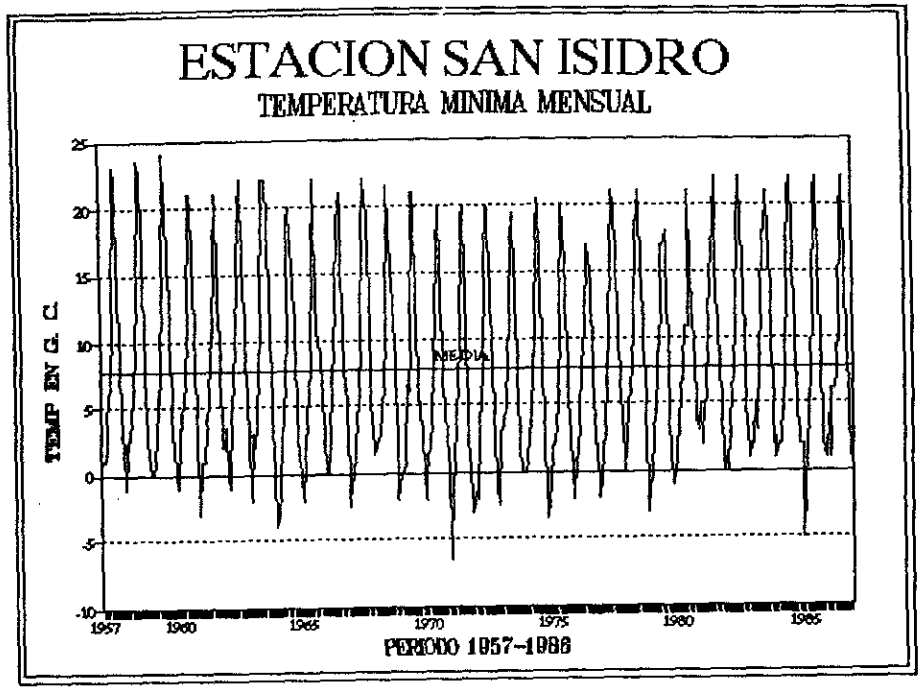
$$\gamma_k = \mathrm{cov}[Y_t, Y_{t+k}] = \mathrm{cov}[Y_j, Y_{j+k}]$$

para toda t, j = 1, 2, ..., n.

La definición anterior nos permite asegurar que una serie de tiempo es estacionaria si los valores de la serie fluctúan alrededor de una media con varianza constante (ver la Fig 3.1). Esa definición se justifica porque en la práctica es común suponer que la distribución asociada con las series de tiempo es la normal, se sigue que es suficiente conocer la μ y la función de autocovarianza para caracterizar a una serie estacionaria. Sin embargo, para evitar la influencia de las unidades de medida, es preferible trabajar con las autocorrelaciones ρ_1, ρ_2, ρ_3, ... definidas a través de

$$\rho_k = \frac{\mathrm{cov}[Y_t, Y_{t+k}]}{\sigma^2_{Y_t}}$$

Fig 3.1



En la práctica, es común que la serie de tiempo con que se trabaja no sea estacionaria, pero en muchos de los casos la no estacionariedad de la serie es homogénea, lo que significa que es únicamente el nivel de la serie el que se ve afectado por la no estacionariedad. En tal caso, una transformación adecuada sobre la serie no estacionaria genera una serie estacionaria.

Una transformación muy usual es la aplicación del operador diferencia ∇^d. Para la definición de este operdor definiremos primeramente al operador de retraso.

Definición 3.2. Sean Y_1, Y_2, Y_3,...,Y_n los n valores de una serie de tiempo representados por Y_t. Se define el operador de retraso B como:

$$BY_t = Y_{t-1} \quad \forall\ t = 2,3, \ldots , n.$$

Este operador retrasa en una unidad de tiempo a la serie Y_t. Además, B^d significa aplicar d veces este operador sobre la serie y B^0 es el operador identidad.

$$B^2Y_t = B(BY_t) = B(Y_{t-1}) = Y_{t-2}$$

Definición 3.3. Sean Y_1, Y_2, Y_3,...,Y_n los n valores de una serie de tiempo representados por Y_t. Se define el operador de diferencia como:

$$Z_t = \nabla^d Y_t = (1 - B)^d Y_t \quad \forall\ t. \qquad (3.1.2)$$

En otras palabras, al aplicar el operador diferencia sobre la serie de tiempo Y_t obtenemos una nueva serie Z_t la cual puede resultar estacionaria.

En la práctica, la experiencia indica que, si la serie de tiempo original Y_t no es estacionaria, la aplicación de la primera diferencia:

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

o la segunda diferencia:

$$Z_t = Y_t - 2Y_{t-1} + Y_{t-2}$$

produce una serie de tiempo estacionaria.

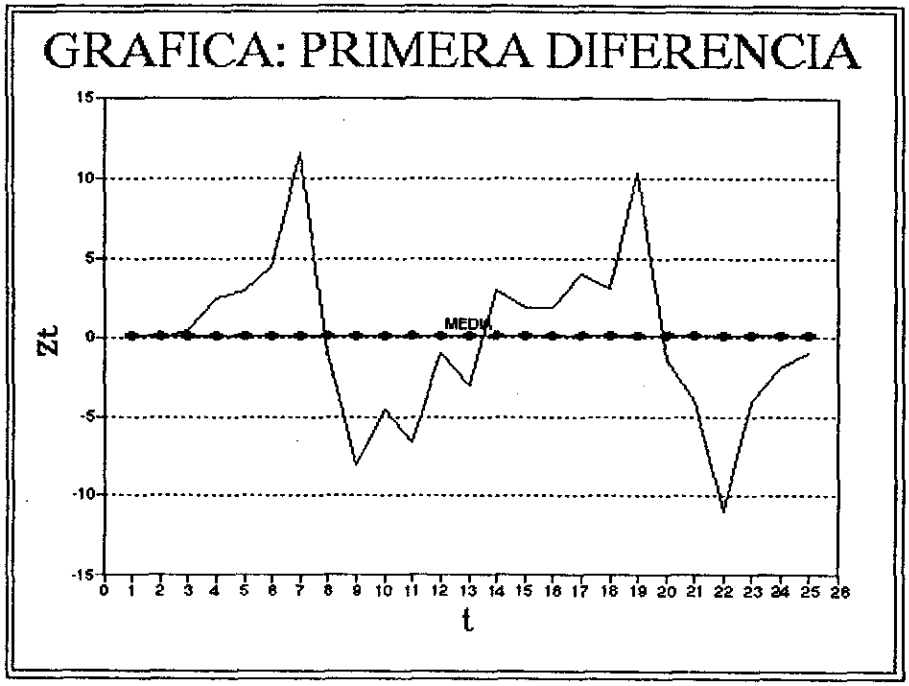
Con el objeto de mostrar como se emplea el operador diferencia, en el Cuadro 3.1 presentamos un ejemplo. Para tal efecto utilizamos los primeros 25 valores de la Tabla 1b del apéndice.

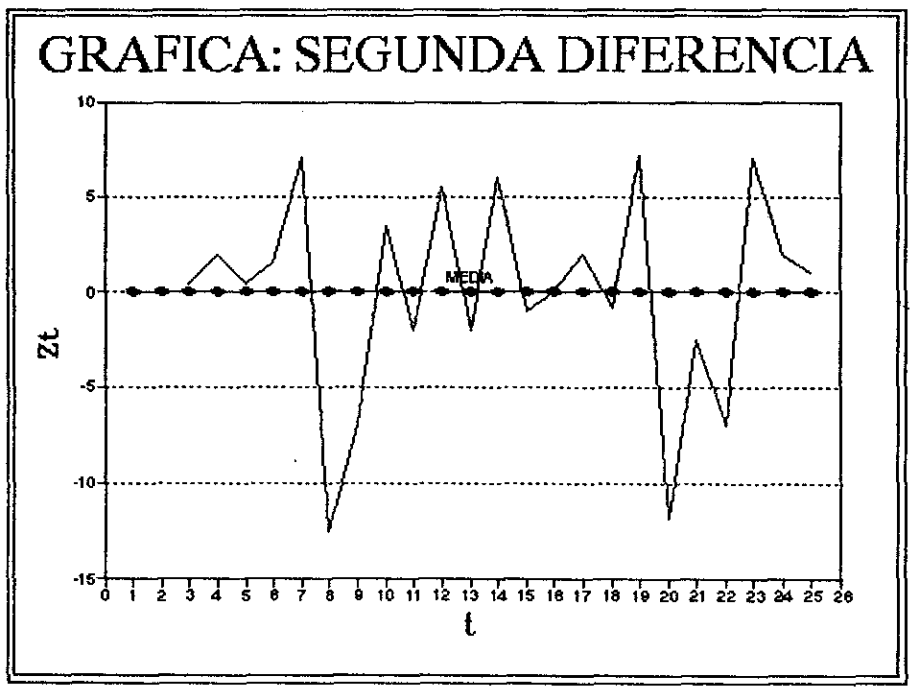
CUADRO 3.1

t	Y_t	$Z_t=\nabla Y_t=Y_t-Y_{t-1}$	$Z_t=\nabla^2 Y_t=Y_t-2Y_{t-1}+Y_{t-2}$
		---	---
1	1.0	0.0	---
2	1.0	0.5	0.5
3	1.5	2.5	2.0
4	4.0	3.0	0.5
5	7.0	4.5	1.5
6	11.5	11.5	7.0
7	23.0	-1.0	-12.5
8	22.0	-8.0	-7.0
9	14.0	-4.5	3.5
10	9.5	-6.5	-2.0
11	3.0	-1.0	5.5
12	2.0	-3.0	-2.0
13	-1.0	3.0	6.0
14	2.0	2.0	-1.0
15	4.0	2.0	0.0
16	6.0	4.0	2.0
17	10.0	3.2	-0.8
18	13.2	10.3	7.1
19	23.5	-1.5	-11.8
20	22.0	-4.0	-2.5
21	18.0	-11.0	-7.0
22	7.0	-4.0	7.0
23	3.0	-2.0	2.0
24	1.0	-1.0	1.0
25	0.0		

En la figura 3.2 se muestran las gráfica de la primera y segunda diferencia, respectivamente. En ellas podemos ver que los valores de cada una de las series fluctúan alrededor de su media, por lo que podríamos pensar que son series estacionarias.

Fig. 3.2


GRAFICA: PRIMERA DIFERENCIA
15
10
5
0
-5
-10
-15
MEDIA
Zt
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
t


GRAFICA: SEGUNDA DIFERENCIA
10
5
0
-5
-10
-15
MEDIA
Zt
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
t

Hemos visto que la función de autocorrelación, junto con la media y varianza define a un proceso estacionario. Usaremos la Función de Autocorrelación Muestral (FACM) para determinar si la serie de tiempo es estacionaria o no.

Definición 3.4. Sea Z_1, Z_2, ..., Z_n los n valores de una serie de tiempo estacionaria Z_t. Definimos la Autocorrelación Muestral (ACM), de Z_t en el retraso k, como:

$$r_k = \frac{C_k}{C_0} \qquad (3.1.3)$$

donde

C_k representa la autocovarianza muestral del proceso Z_t, definida como:

$$C_k = \frac{1}{n}\sum_{t=1}^{n}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z}),$$

C_0 la varianza muestral definida como:

$$C_0 = \frac{1}{n}\sum_{t=1}^{n}(Z_t - \bar{Z})^2$$

considerando que

$$\bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t$$

es la media muestral. La FACMes la representación gráfica de las ACM en los k retraso y en base a ésta diremos que:

1.- Si el comportamiento de la FACM de la serie de tiempo Z_t se corta bastante rápido o tiende a cero rápidamente, entonces la serie de tiempo será considerada estacionaria.

2.- Si la FACM de la serie de tiempo Z_t tiende a cero de forma extremadamente lenta, entonces la serie se considera no estacionaria.

Los términos "cortarse rápidamente", "tender a cero rápidamente" y "tender a cero extremadamente lento" se definen heurísticamente; es la experiencia el criterio básico que se usa para caraterizar esos comportamientos.

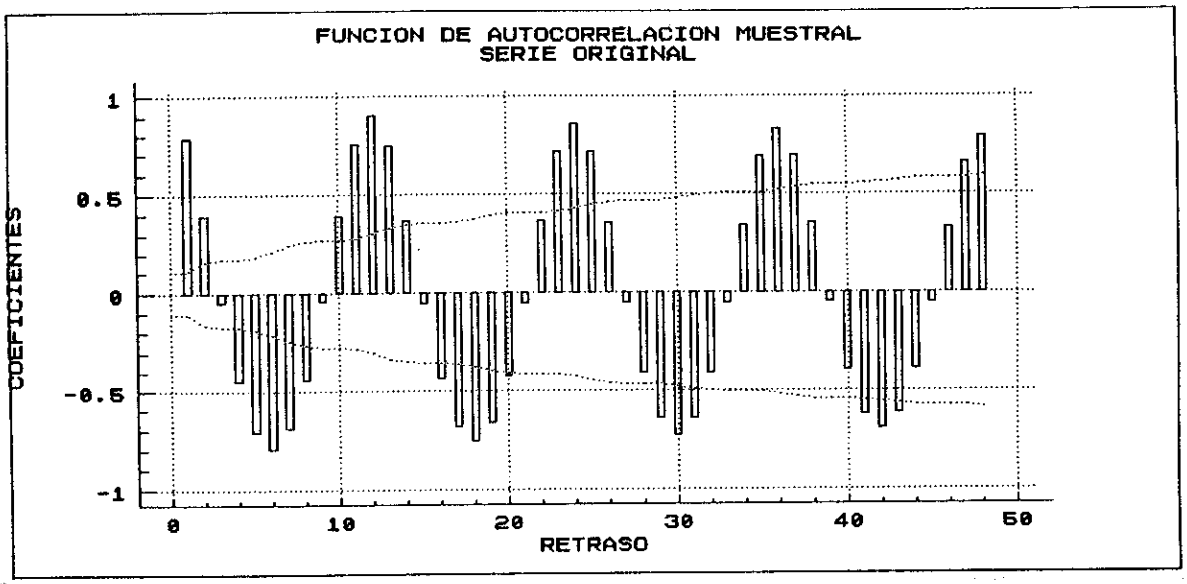
En la figura 3.3, se muestran las gráficas de la FACM para la serie original, primera diferencia y la segunda diferencia. En ellas observamos que tienden a cero lentamente por lo que se consideran no estacionarias. Además, observamos que mantienen un patron repetitivo cada 12 retrasos, a éste comportamiento se le denomina estacionalidad. (ver la definición en el Anexo 2).

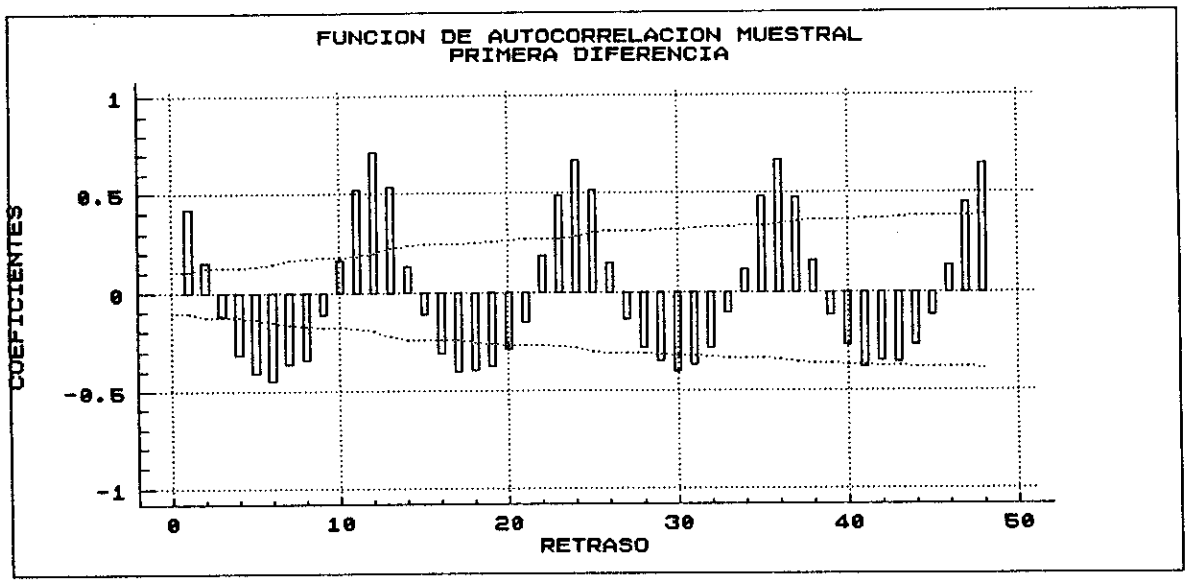
Otro elemento importante en la identificación del modelo es la Función de Autocorrelación Parcial Muestral (FACPM), la cual describe el comportamiento de los coeficientes de la autocorrelación parcial muestral, de la siguientte manera:

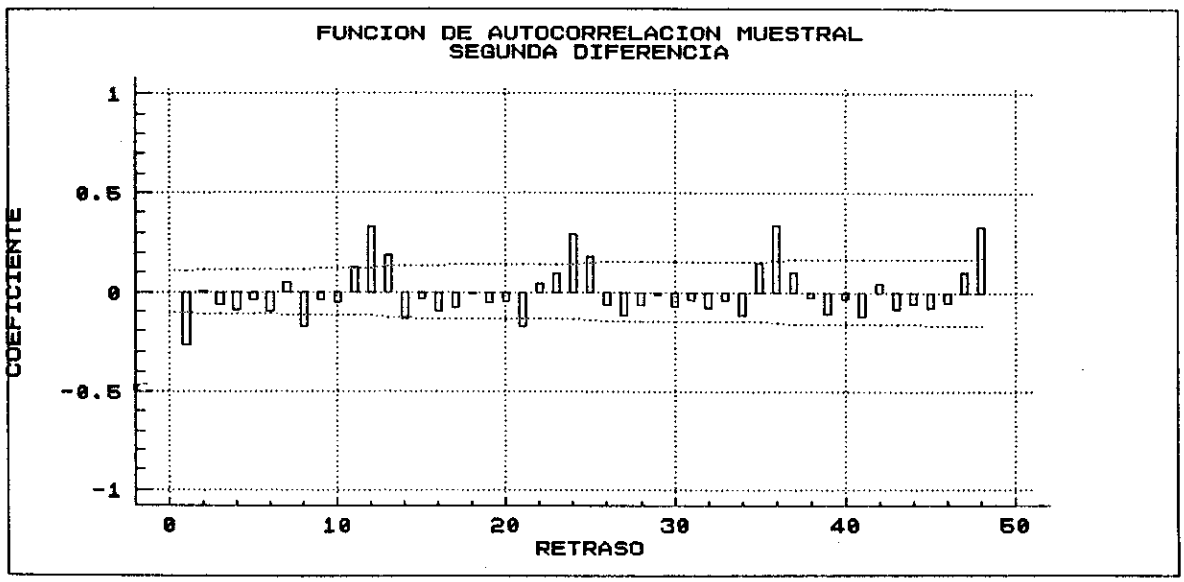
$$r_{kk} = \begin{cases} r_1 & \text{si } k=1 \\ \dfrac{r_k - \sum r_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_j} & \text{si } k=2,\ 3,\ \ldots \end{cases}$$

donde $r_{kj} = r_{k-1,j} - r_{kk}\, r_{k-1,k-j}$ para toda $j=1,\ 2,\ \ldots,\ k-1$.

Fig. 3.3


FUNCION DE AUTOCORRELACION MUESTRAL
SERIE ORIGINAL
COEFICIENTES
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION MUESTRAL
PRIMERA DIFERENCIA
COEFICIENTES
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION MUESTRAL
SEGUNDA DIFERENCIA
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

3.1.2. Uso de la FACM y de la FACPM para la identificación del modelo.

El modelo general de Box-Jankins para datos no estacionales se define como:

$$\Phi_p(B) Z_t = \delta + \Theta_q(B) a_t \qquad (3.1.4)$$

ó

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

$$a_t - \theta_1 a_t - \theta_2 a_{t-2} + \ldots + \theta_q a_{t-q}.$$

Este modelo se llama Autorregresivo-Promedio Móvil de orden (p,q) o simplemente ARMA(p,q) donde:

1) $\Phi_p(B)$ es el operador autorregresivo definido como:

$$\Phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p),$$

2) $\Theta_q(B)$ es el operador de promedio móvil definido como:

$$\Theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q),$$

Z_t es una serie de tiempo estacionaria, es decir, producto de una transformación de la serie original Y_t,

a_t es una variable aleatoria con distribución $\mathbf{N}(0,\sigma^2)$, a la cual se le conoce com proceso de ruido blanco.

3) δ es un término constante determinado por:

$$\delta = \mu(1 - \phi_1 - \phi_2 - \ldots - \phi_p).$$

Donde μ es la media de la serie Z_t. Nótese que si el grado del operador autorregresivo es cero (p=0) entonces $\delta=\mu$.

Con la finalidad de presentar un modelo tentativo a partir del modelo general (3.1.3) necesitamos determinar:

1) Si se incluye la constante δ.
2) El grado de cada operador, es decir, los valores de p y q.

La constante δ se incluye en el modelo si estadísticamente μ es diferente de cero.

Sea $\bar{Z} = \frac{1}{n} \sum_{t=1}^{n} Z_t$

la media muestral de los valores de la serie de tiempo Z_1, Z_2, ..., Z_n, un estimador de μ.

Se dice que μ es estadísticamente distinta de cero si:

$$\left| \frac{\bar{Z}}{S_z / \sqrt{n+1}} \right| > 2$$

donde

$$S_z = \left(\frac{\sum_{t=1}^{n} (Z_t - \bar{Z})^2}{(n-1)} \right)^{1/2} .$$

Si $\bar{Z}$ es estadísticamente diferente de cero y es un estimador de μ, entonces δ es diferente de cero.

Para determinar qué operador y de qué orden se va a utilizar

en el modelo se presenta la siguiente guía [2], que se basa en la FACM y de la FACPM.

GUIA 1: Operadores no estacionales.

1.- Si el comportamiento de la FACM presenta "picos" en los retrasos 1, 2, ..., q y después del retraso q se corta, y la FACPM decae rápidamente, entonces utilice el operador de promedio móvil de orden q.

2.- Si el comportamiento de FACM decae rápidamente, y la FACPM presenta "picos" en los retrasos 1, 2, ..., p y después del retraso p se corta, entonces use el operador autorregresivo de orden p.

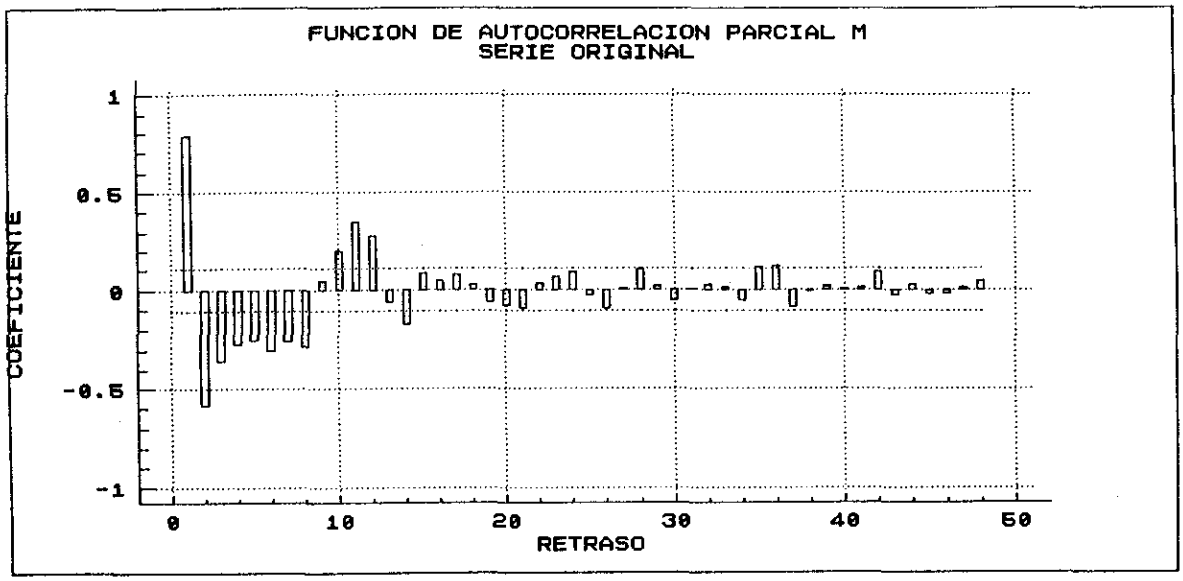
3.- Si el comportamiento de la FACM presenta "picos" en los retrasos 1, 2, ..., q y después del retraso q se corta, y la FACPM presenta "picos" en los retrasos 1, 2, ..., p y después del retraso p se corta, entonces utilice el operador de promedio móvil de orden q si la FACM se corta más abruptamente que FACPM. Utilice el operador autorregresivo si la FACPM se corta más abruptamente que la FACM. En caso de que ambas FACM y FACPM se cortan abruptamente pruebe cual de las dos presenta un mejor modelo.

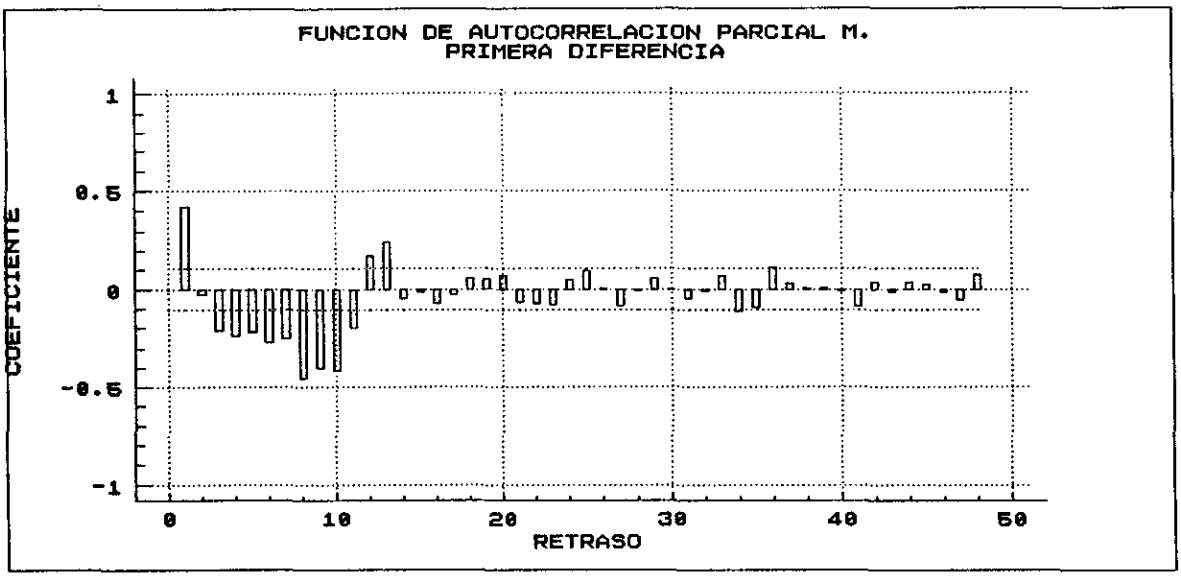
4.- Si el comportamiento de la FACM y la FACPM no presenta "picos", entonces no se utiliza operador no estacional.

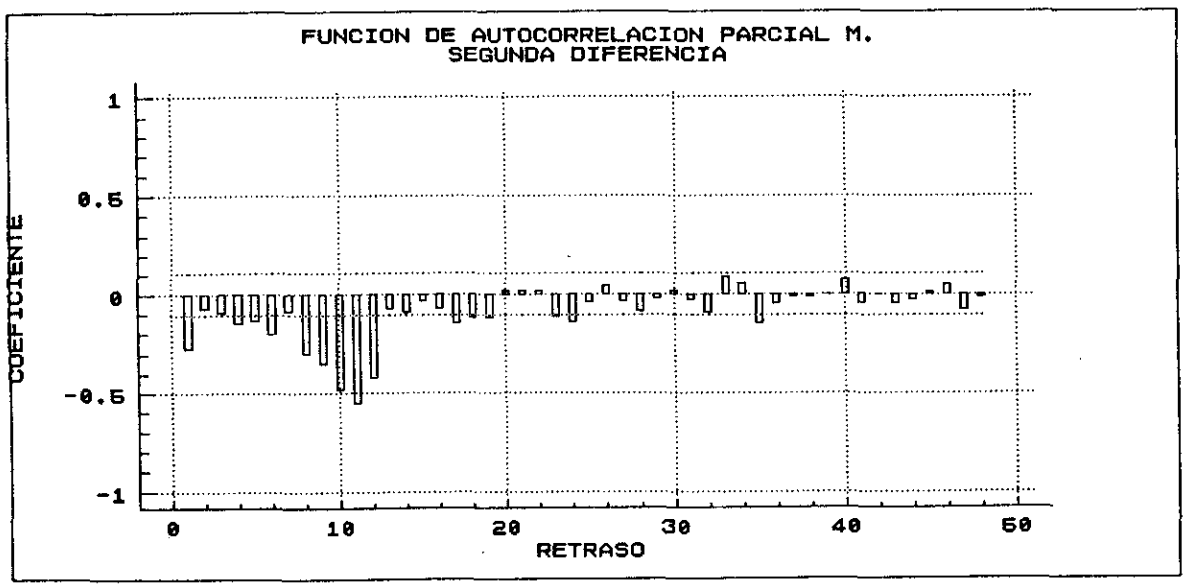
5.- Si el comportamiento de la FACM y la FACPM decaen rápidamente, entonces utilice ambos operadores. Si el comportamiento de éstas sigue un patrón de decaimiento exponencial, utilice ambos operadores con p=1 y q=1.

En las figuras 3.3 y 3.4, presentamos las gráficas de la FACM Y la FACPM, respectivamente; en ellas podemos observar que la FACPM se corta más abruptamente que la FACM; esto sugiere un operador autorregresivo de orden 12; sin embargo la FACM no decae rápidamente debido a la presencia de la componente estacional por lo que se requiere una análisis adicional. Para ello presentamos la siguiente sección.

Fig 3.4


FUNCION DE AUTOCORRELACION PARCIAL M
SERIE ORIGINAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION PARCIAL M.
PRIMERA DIFERENCIA
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION PARCIAL M.
SEGUNDA DIFERENCIA
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

3.2. IDENTIFICACION DEL MODELO (Datos Estacionales)

En el Anexo 2 se definen los elementos que componen a una serie de tiempo. En esta sección analizamos el caso cuando una serie de tiempo contiene la componente estacional y proponemos un modelo tentativo para ella, a partir del modelo general de Box-Jenkins definido como:

$$\Phi_p(B)\ \Phi_P(B^s)\ Z_t = \delta + \Theta_q(B)\ \Theta_Q(B^s)\ a_t. \qquad (3.2.1)$$

En la sección anterior se expuso un ejemplo en el cual se dice que la serie de timpo contiene una componente estacional de longitud 12, identificado con la FACM y la FACPM.

Con el propósito de proponer un modelo tentativo para series de tiempo de este tipo nos basaremos en:

1) Analizar la serie de tiempo en el nivel estacional.
2) Analizar la serie de tiempo en el nivel no estacional.

3.2.1. Análisis de la componente estacional.

Una serie de tiempo que contiene la componente estacional puede ser transformada a una nueva serie de tiempo que no contenga a esta componente. Para esto utilizamos el operador de diferencia estacional ∇_s^D que se define como:

Definición 3.5. Sean Y_1, Y_2, Y_3,...,Y_n los n valores de una serie de tiempo representados por Y_t. Se define el operador de diferencia estacional como:

$$W_t = \nabla_s^D Y_t = (1 - B^s)^D Y_t \quad \forall\ t. \tag{3.2.2}$$

donde

D es el grado de diferenciación y generalmente toma los valores de 0, 1 ó 2 y,

s es la longitud estacional.

En el Cuadro 3.2, se muestran los datos utilizados en el Cuadro 3.1 y la primera diferencia estacional, con una longitud 12 (s = 12).

Una vez que se obtiene W_t, el caso se reduce a la sección anterior, es decir, determinar el grado de diferenciación de W_t para obtener una serie de tiempo estacionaria Z_t. Esto se expresa como:

$$Z_t = \nabla^d W_t = \nabla^d \nabla_s^D Y_t. \tag{3.2.3}$$

3.2.2. Uso de la FACM y la FACPM para la identificación del modelo.

La Ec. 3.2.1 representa el modelo para series de tiempo con efecto estacional. En ella podemos observar que si el grado del operador autorregresivo estacional definido como:

$$\Phi_P(B^s) = (1 - \phi_{1,s}B^s - \phi_{2,s}B^{2s} - \ldots - \phi_{P,s}B^{Ps})$$

y el grado del operador de promedio móvil estacional definido como:

$$\Theta_Q(B^s) = (1 - \theta_{1,s}B^s - \theta_{2,s}B^{2s} - \ldots - \theta_{Q,s}B^{Qs})$$

toman el valor de cero, ésta se transforma en un modelo para series de tiempo no estacionales.

Obsérvese, además, que la constante δ del modelo 3.2.1 queda definida como:

$$\delta = \mu\Phi_p(B)\ \Phi_P(B^s) \qquad (3.2.4)$$

$$\delta = \mu(1 - \phi_1 - \phi_2 - \ldots - \phi_p)(1 - \phi_{1,s} - \phi_{2,s} - \ldots - \phi_{P,s}).$$

Para identificar el modelo particular del modelo general de Box-Jenkins que describe el proceso estacionario Z_t necesitamos determinar.

1) Si se incluye la constante δ
2) El grado (p, P, q, Q) de cada operador.

La constante δ se incluye en el modelo si cumple con lo establecido en la sección anterior.

Para determinar el grado de los operadores no estacionales nos basaremos en la guía de la sección anterior analizando la FACM y la FACPM en el nivel no estacional (del primer retraso al s-3).

El grado de los operadores estacionales se determinan

CUADRO 3.2. **Primera Diferencia Estacional (s = 12)**

t	Y_t	$W_t = \nabla^1_{12} Y_t = Y_t - Y_{t-12}$
1	1.0	-----
2	1.0	-----
3	1.5	-----
4	4.0	-----
5	7.0	-----
6	11.5	-----
7	23.0	-----
8	22.0	-----
9	14.0	-----
10	9.5	-----
11	3.0	-----
12	2.0	-----
13	-1.0	-2.0
14	2.0	1.0
15	4.0	2.5
16	6.0	2.0
17	10.0	3.0
18	13.2	1.7
19	23.5	0.5
20	22.0	0.0
21	18.0	4.0
22	7.0	-2.5
23	3.0	0.0
24	1.0	-1.0
25	0.0	1.0

analizando el comportamiento de la FACM y la FACPM en el nivel estacional, es decir el comportamiento en los retraso s, 2s, 3s, 4s y los retrasos cercanos (s-2, s-1, s+1, s+2, 2s-2, 2s-1, 2s+1,

2s+2, 3s-2, 3s-1, 3s+1, 3s+2, 4s-2, 4s-1, 4s+1 y 4s+2), apoyandonos en la guía 2.

GUIA 2: Operadores estacionales.

1.- Si el comportamiento de la FACM tiene picos en los retrasos s, 2s,..., Qs y se corta después del retraso Qs y la FACPM decae rápidamente, entonces se utiliza el operador de promedio móvil estacional de orden Q.

2.- Si el comportamiento de la FACM decae rápidamente y la FACPM tiene picos en los retrasos s, 2s,..., Ps y se corta después del retraso Ps, entonces utilice el operador autorregresivo estacional de orden P.

3.- Si el comportamiento de la FACM tiene picos en los retrasos s, 2s,..., Qs y se corta después del retraso Qs, y la FACPM tiene picos en los retrasos s, 2s,..., Ps y se corta después del retraso Ps, entonces utilice el operador de promedio móvil estacional de orden Q si la FACM se corta más abruptamente que la FACPM. Utilice el operador autorregresivo estacional de orden P si FACPM se corta más abruptamente que FACM. Si ambas se cortan abruptamente entonces probar cual de los dos presenta un mejor modelo.

4.- Si el comportamiento de la FACM y la FACPM no presentan picos en el nivel estacional, entonces ninguno de los

operadores estacionales se incluye.

5.- Si la FACM decae rápidamente en el nivel estacional y la FACPM decae rápidamente en el nivel estacional, entonces utilice los dos operadores.

Ejemplo. Con el fin de proponer un modelo tentativo para el ejemplo anterior, en la figura 3.5 se muestra la gráfica de la FACM y la FACPM de la primera diferencia estacional. En ella observamos:

1) Estacionariedad.

Dado que la FACM decae rápidamente entonces la serie es estacionaria.

2) Comportamiento en el nivel no estacional.

El comportamiento de la FACM y la FACPM decae aproximadamente en forma exponencial y se corta después del retraso 4. Según el punto cinco de la guía 1 se sugiere el uso de ambos operadores de orden uno.

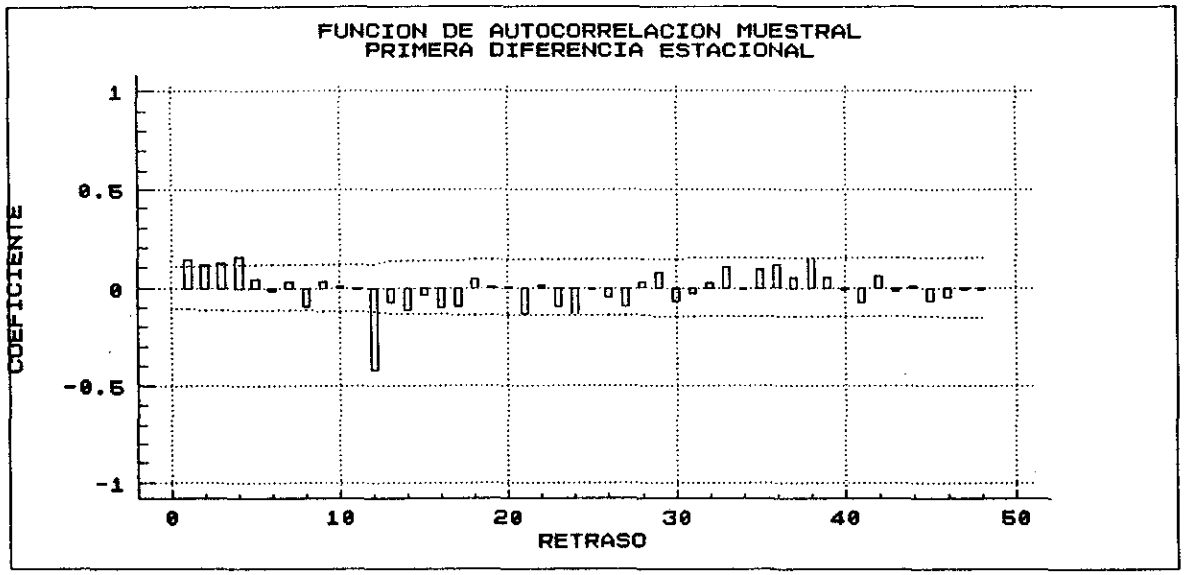
3) Comportamiento en el nivel estacional.

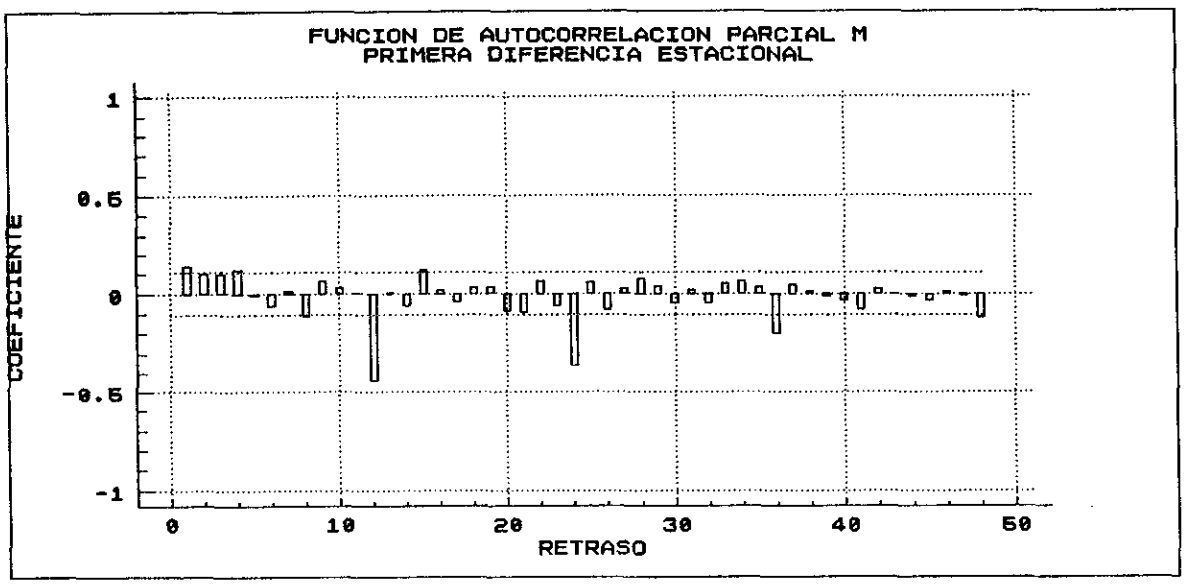
El comportamiento de la FACMP sigue un patrón de decaimiento exponencial mientras que la FACM presenta un pico en el retraso 12 y después se corta. Según el punto uno de la guía 2 se sugiere el uso del operador de promedio móvil de orden uno.

Un modelo tentativo, después de este análisis sería.

$$\Phi_1(B)\, Z_t = \delta + \Theta_1(B)\Theta_1(B^s)\, a_t \qquad (3.2.5)$$

Fig. 3.5


FUNCION DE AUTOCORRELACION MUESTRAL
PRIMERA DIFERENCIA ESTACIONAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION PARCIAL M
PRIMERA DIFERENCIA ESTACIONAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

3.3. ESTIMACION

Del paso de identificación, resultó la Ec. 3.2.5, un modelo tentativo para nuestro ejemplo, que podemos reescribir como:

$$(1 - \phi_1 B) Z_t = \delta + (1 - \theta_1 B)(1 - \theta_{1,12} B^{12}) a_t \qquad (3.3.1)$$

ó

$$Z_t = \delta + \phi_1 Z_{t-1} + a_t - \theta_1 a_{t-1} - \theta_{1,12} a_{t-12} + \theta_1 \theta_{1,12} a_{t-13}.$$

Ahora es preciso someterlo a las condiciones de estacionariedad e invertibilidad para poder determinar si es o no capaz de describir la serie de tiempo.

3.3.1. Condición de Estacionariedad e Invertibilidad.

La metodología de Box-Jenkins requiere que los modelos particulares que resultan del modelo general, representados por la Ec. 3.2.1, y usados para describir y pronosticar una serie de tiempo, sean estacionarios e invertibles; estos significa que los parámetros de los operadores $\Phi(B)$ y $\Theta(B)$, que intervienen en el modelo deben satisfacer estas condiciones.

En el Cuadro 3.3, se presentan las condiciones de estacionariedad e invertibilidad para los parámetros de primero y segundo orden de cada operador.

CUADRO 3.3. Condición de estacionariedad e invertibilidad sobre los parámetros de primer y segundo orden para $\Phi_p(B)$, $\Phi_p(B^s)$, $\Theta_q(B)$, y $\Theta_Q(B^s)$.

OPERADOR	CONDICION	
	ESTACIONARIEDAD	INVERTIBILIDAD
$(1 - \theta_1 B)$	Ninguna	$\lvert\theta_1\rvert < 1$
$(1 - \theta_1 B - \theta_2 B^2)$	Ninguna	$\theta_1 + \theta_2 < 1$; $\theta_2 - \theta_1 < 1$; $\lvert\theta_2\rvert < 1$
$(1 - \theta_{1,s} B^s)$	Ninguna	$\lvert\theta_{1,s}\rvert < 1$
$(1 - \theta_{1,s} B^s - \theta_{1,s} B^{2s})$	Ninguna	$\theta_{1,s} + \theta_{2,s} < 1$; $\theta_{2,s} - \theta_{1,s} < 1$; $\lvert\theta_{2,s}\rvert < 1$
$(1 - \phi_1 B)$	$\lvert\phi_1\rvert < 1$	Ninguna
$(1 - \phi_1 B - \phi_2 B^2)$	$\phi_1 + \phi_2 < 1$; $\phi_2 - \phi_1 < 1$; $\lvert\phi_2\rvert < 1$	Ninguna
$(1 - \phi_{1,s} B^s)$	$\lvert\phi_{1,s}\rvert < 1$	Ninguna
$(1 - \phi_{1,s} B^s - \phi_{1,s} B^{2s})$	$\phi_{1,s} + \phi_{2,s} < 1$; $\phi_{2,s} - \phi_{1,s} < 1$; $\lvert\phi_{2,s}\rvert < 1$	Ninguna

Las condiciones de estacionariedad e invertibilidad sobre los parámetros en la forma general las resumimos como sigue:

Caso 1.- En un modelo autorregresivo puro, es decir, cuando únicamente los operadores $\Phi(B)$ y $\Phi(B^s)$ intervienen en el modelo, se exige la condición de estacionariedad pero no la de invertibilidad. Una condición necesaria pero no suficiente sobre los parámetros de cualquiera de los operadores $\Phi_p(B)$ o $\Phi_p(B^s)$ es que la suma de los valores de los parámetros sea menor que 1.

Caso 2.- En un modelo promedio móvil puro, es decir, cuando únicamente los operadores $\Theta_q(B)$ y $\Theta_Q(B^s)$ intervienen en el modelo, se exige la condición de invertibilidad pero no la de estacionariedad. Una condición necesaria pero no suficiente sobre los parámetros de cualquiera de los operadores es que la suma de los valores de los parámetros sea menor que 1.

Caso 3.- Para los modelo ARMA(p,q) y ARIMA(p,P,q,Q) se exige la condición de estacionariedad sujetándose al Caso 1 y la invertibilidad a el Caso 2.

Caso particular, para un modelo ARMA(1,1) (autorregresivo de orden uno y promedio móvil de orden uno), dado por:

$$(1 - \phi_1 B) Z_t = \delta + (1 - \theta_1 B) a_t$$

la condición para que sea estacionario es que $|\phi_1| < 1$, y la condición de invertibilidad es que $|\theta_1| < 1$.

3.3.2. Estimación preliminar.

Una estimación preliminar consiste en fijar los valores iniciales para los parámetros de los operadores que intervienen en el modelo que se va a utilizar, de modo que cumplan las condiciones anteriores. Después, através de un proceso iterativo se estiman los valores finales de los parámetros, los cuales deberán cumplir las mismas condiciones. Generalmente los valores iniciales de los parámetros de cada uno de los operadores toman el valor de 0.1, de tal manera que el valor inicial para δ (de la Ec 3.2.4) se estima como:

$$\delta = \bar{Z}(1 - .1 - .1 - .1 - \ldots)(1 - .1 - .1 - \ldots).$$

3.3.3. Estimación.

Box-Jenkins prefieren la función de máxima verosimilitud para aproximar los cálculos de los estimadores puntuales; sin embargo, para efecto de cálculos se utilizan mínimos cuadrados para tal aproximación, y puede demostrarse que si los valores que toman los errores a_t están normalmente distribuídos, la estimación por mínimos cuadrados resulta aproximadamente igual si utilizamos la función de máxima verosimilitud.

Asociado con los estimadores puntuales de cada parámetro del modelo de Box-Jenkins se encuentra el error estándar y el

valor t de la distribución de student que nos permite determinar si tal parámetro es o no igual a cero y por lo tanto si se incluye o no en el modelo.

Definición 3.5. Sea $\hat{\psi}$ el estimador de ψ, donde ψ es cualquier parámetro del modelo de Box-Jenkins y $s_{\hat{\psi}}$ el error estándar de $\hat{\psi}$. Se define el valor de t como:

$$t_{\hat{\psi}} = \frac{\hat{\psi}}{s_{\hat{\psi}}} .$$

Si el valor absoluto de $t_{\hat{\psi}}$ es "grande", entonces $\hat{\psi}$ es "grande" esto implica que $\hat{\psi}$ no es igual a cero, así que se rechaza la hipótesis H_0: $\psi = 0$, lo cual implica que ψ se incluye en el modelo.

Resumiendo, diremos que $\hat{\psi}$ se incluye en el modelo si:

$$|t_{\hat{\psi}}| > t_{\alpha/2}^{[n-np]}$$

donde n es el número de observaciones y np número de parámetros y α es el nivel de significancia de la prueba.

Ejemplo: Utilizamos el paquete STATGRAPHICS (Ver Anexo 3), para calcular los valores estimados para el modelo 3.3.1 y se obtienen los siguientes resultados:

Parámetro	Estimación	Error estándar	$t_{\hat{\psi}}$	Prob.
ϕ_1	0.77860	0.12833	6.06732	0.00000
θ_1	0.61570	0.15619	3.94198	0.00010
$\theta_{1,12}$	0.83092	0.03140	26.46471	0.00000

$\bar{Z}$	0.00111	0.03492	0.03186	0.97460
δ	0.00025			

De donde vemos que se cumple:

1). Estacionariedad

$|\phi_1| < 1.$

2). Invertibilidad

$|\theta_1| < 1.$

$|\theta_{1,12}| < 1.$

3). Los parametros ϕ_1, θ_1 y $\theta_{1,12}$ se incluyen en el modelo ya que la probabilidad es menor que $\alpha/2$ con $\alpha = 0.05$; es decir, se cumple

$$|t_{\hat{\psi}}| > t_{\alpha/2}^{[n-np]} \qquad \text{para } \psi = \phi_1,\ \theta_1 \text{ y } \theta_{1,12}.$$

Por otro lado, el valor de $\bar{Z}$ estimado no cumple con la prueba t de student ya que la probabilidad de este valor es muy grande. Esto implica que la constante δ no se incluye en el modelo, de tal manera que el modelo 3.3.1 queda expresado como:

$$(1 - \phi_1 B) Z_t = (1 - \theta_1 B)(1 - \theta_{1,12} B^{12}) a_t. \qquad (3.3.2)$$

3.4. COMPROBACION Y DIAGNOSTICO

Una vez que se obtiene la estimación de los parámetros que van a intervenir en el modelo, necesitamos comprobar si éste es adecuado o no. Un análisis sobre la FACM aplicada a los residuos (FACMR), nos permitirá sugerir un nuevo modelo en caso de que el modelo resultara inadecuado. Enseguida ilustramos el procedimiento.

3.4.1. Comprobación.

Existen dos pruebas estadísticas para determinar si el modelo es o no adecuado. Estas son:

1) Prueba de Box-Pierce

$$Q = n' \sum_{i=1}^{k} r_i^2(\hat{a})$$

2) Prueba de Ljung-Box

$$Q^* = n'(n'-2) \sum_{i=1}^{k} (n'-i) r_i^2(\hat{a})$$

donde

$n' = n - (d + sD)$, n es el número de observaciones de la serie de tiempo original, s la estacionalidad, d y D son el grado de diferenciación en el nivel no estacional y estacional, respectivamente.

$r_i^2(\hat{a})$ es el cuadrado de la autocorrelación residual muestral

en el retraso i, es decir, la autocorrelación muestral del residuo separado por una unidad de tiempo en el retraso i.

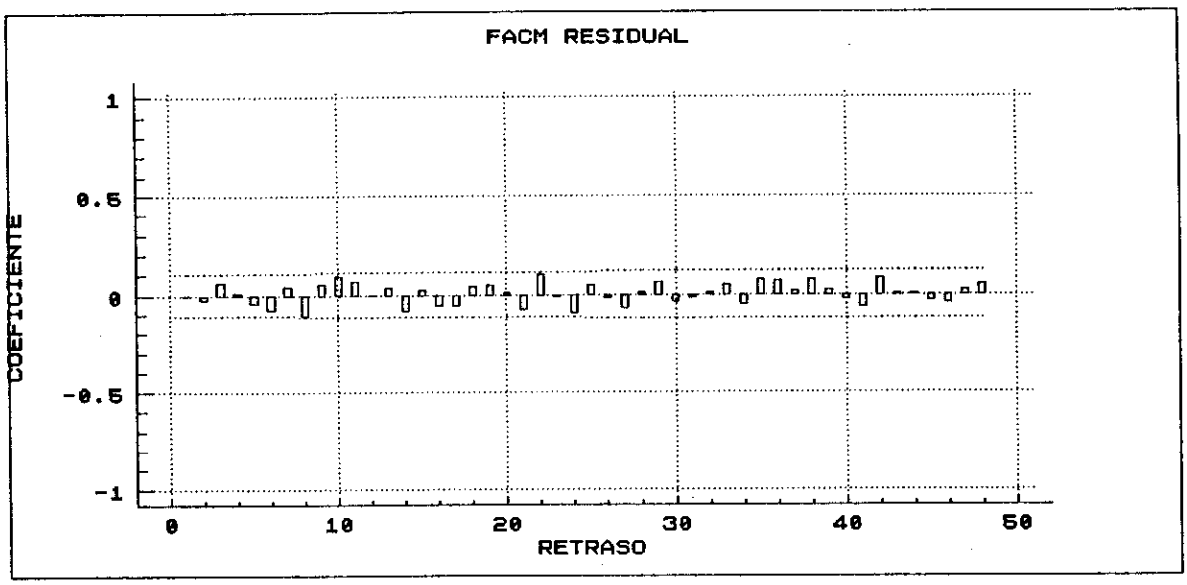
El procedimiento para la comprobación de que el modelo es adecuado o no es el mismo para ambos estadísticos. Presentaremos a continuación como comprobar que el modelo es adecuado utilizando el estadístico de Box-Pirce, Q.

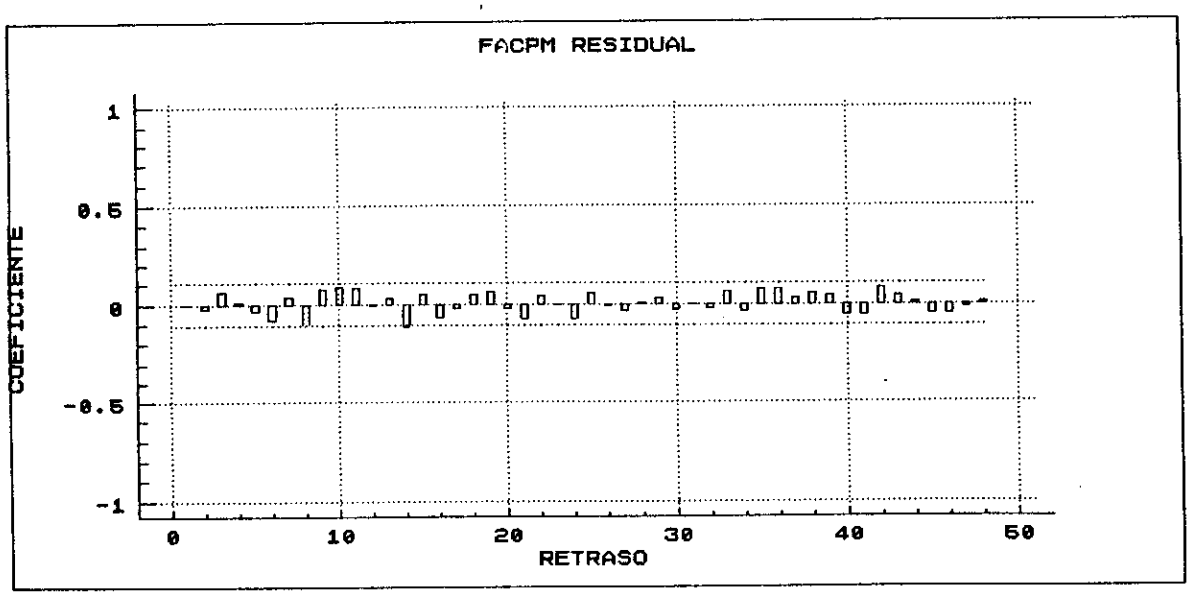
Como podemos darnos cuenta, el valor de Q depende básicamente del factor $r_i^2(\hat{a})$, de tal manera que si Q toma un valor pequeño implica que $r_i^2(\hat{a})$ toma un valor pequeño; ésto significa que los residuos no están relacionados, indicando que el modelo es adecuado. En otras palabras, podemos rechazar lo adecuado del modelo bajo la probabilidad de cometer un error de tipo I igual a α si se cumple la siguiente condición.

$$Q > \chi^2_{[\alpha,k-np]} \text{ con } \alpha = 0.05.$$

Ejemplo. Se hizo una nueva estimación para los parámetros del modelo 3.3.2 y se obtuvo un valor para Q = 21.4378 en los primeros 20 residuos (k=20); este valor es menor que $\chi^2_{[\alpha,k-np]} = 27.6$, con α=0.05 y gl=17 (20 - 3). Esto implica que el modelo es adecuado. En la Fig. 3.6, se muestran las gráficas de la FACM y la FACPM residuales, respectivamente. El comportamiento de cada una nos permite asegurar que los valores de los residuos no estan correlacionados

Fig. 3.6


FACM RESIDUAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FACPM RESIDUAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

3.4.2. Diagnóstico.

Si el modelo es inadecuado significa que los residuos están relacionados, es decir, no son estadísticamente independientes. En otras palabras. Un modelo adecuado es aquel que contiene la estructura de variación de la serie de tiempo y deja fuera, como residuo, únicamente ruido blanco, esto es, errores normales e independientes. Un análisis adicional del comportamiento de la FACM y la FACPM sobre los residuos nos permitirá proponer un nuevo modelo.

Ejemplo. Supóngase que el modelo propuesto anteriormente hubiera sido

$$Z_t = (1 - \theta_{1,12}B^{12})a_t.$$

Para este modelo el valor de Q = 50.835 es mayor que $\chi^2_{[\alpha,k-np]} = 30.1$, con α=0.05 y gl=19 (20 - 1). Esto implica que el modelo es inadecuado, es decir, los residuos (a_t) estan relacionados. Sea η_t los residuos de este modelo, construimos la serie η_t como:

$$\eta_t = \frac{Z_t}{1 - \theta_{1,12}B^{12}} \qquad (3.4.1)$$

La FACM y la FACPM para η_t (denominadas función de autocorrelación residual muestral y función de autocorrelación parcial residual muestral, respectivamente), se muestran en la figura 3.7. En ellas podemos observar que en el nivel no estacional existen valores grandes; de acuerdo al punto 5 de la

Guía 1, se sugiere el uso de ambos operadores de orden uno, con lo que obtenemos:

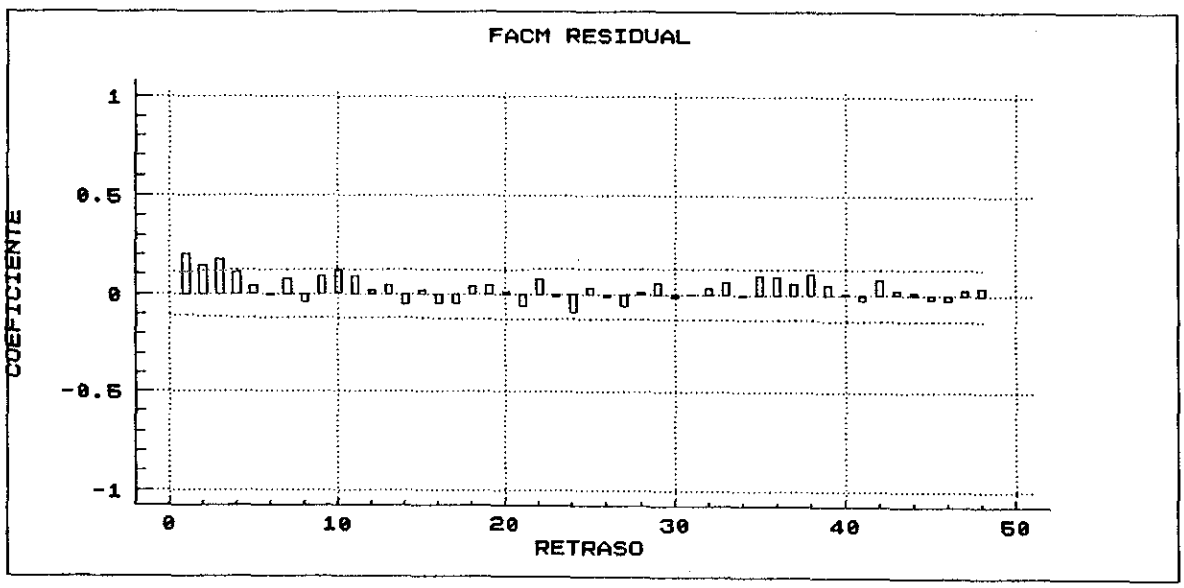
$$(1 - \phi_1 B)\eta_t = (1 - \theta_1 B)a_t. \qquad (3.4.2)$$

De la Ec 3.4.1 y la Ec 3.4.2 se obtiene el modelo

$$(1 - \phi_1 B)Z_t = (1 - \theta_1 B)(1 - \theta_{1,12} B^{12})a_t$$

que es el mismo que se obtuvimos en la sección 3.3, representado por la Ec. 3.3.2.

Fig. 3.7



3.5. PRONOSTICO

El propósito general de la metodología de Box-Jenkins es encontrar un modelo adecuado para pronósticar series de tiempo.

La mejor forma de calcular el pronóstico es por medio de la ecuación en diferencia que representa al modelo. Por ejemplo, de la Ec. 3.3.2, tenemos:

$$Z_{t+1} = \phi_1 Z_{t+1-1} + a_{t+1} - \theta_1 a_{t+1-1} - \theta_{1,12} a_{t+1-12}$$

$$+ \theta_1 \theta_{1,12} a_{t+1-13}.$$

Si $\hat{Z}_t(1)$ representa la estimación de Z_t en el período t desplazado 1 unidades a la derecha de t, entonces, éste la podemos representar por la función

$$\hat{Z}_t(1) = [Z_{t+1}] = \mathbf{E}\ [Z_{t+1} | \phi_1,\ \theta_1,\ \theta_{1,12},\ Z_t\ Z_{t-1},\ \ldots]$$

Por último, falta recordar que el proceso Z_t surgio de una tranformación de la serie Y_t del tipo:

$$Z_t = \nabla_{12} Y_t = Y_t - Y_{t-12}$$

ó

$$Y_t = Z_t + Y_{t-12}.$$

Por lo tanto, para estimar un valor de la serie Y_t en la posición *l*, se utiliza la ecuación definida como:

$$\hat{Y}_t(l) = \hat{Z}_t(l) + Y_{t-12}.$$

En la Tabla 1b se encuentran los datos de temperatura mínima extrema mensual correspondiente a la estación climatológica San Isidro, localizada en latitud 28° 49′ 52", longitud 111° 39′ 17", a una altura de 30 m.s.n.m. en la Costa de Hermosillo.

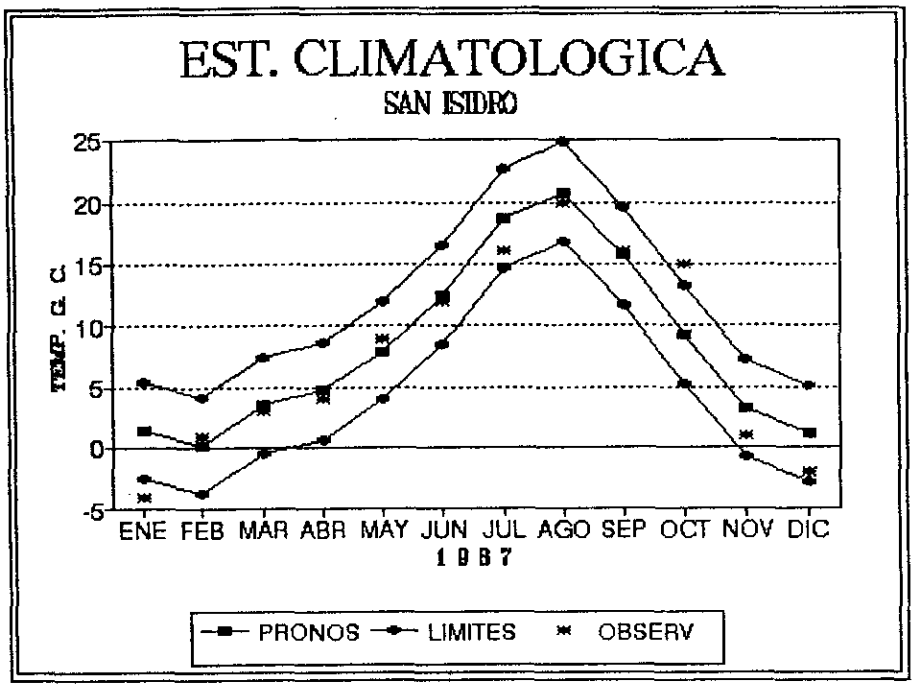
Los análisis realizados en los ejemplos de las secciones anteriores corresponden a ésta, utilizando los datos comprendidos en el período 1957-1986. Esto con el fin de utilizar el modelo de la Ec. 3.3.2, para pronosticar sobre los años 1987 y 1988 y comparar con las observaciones registradas en ese período.

En la Fig. 3.8 se muestran las gráficas de pronóstico para el año de 1987 y 1988. En cada una de ellas obsevamos tres lineas; la del centro representa el pronóstico y las otras dos los límites de confianza a un nivel del 95%. Además se representa con un asterisco las observaciones ocurridas en esas fechas.

Se hizo un segundo análisis utilizando los datos comprendidos en el período 1957-1987 con el fin de preonosticar 1988 y detectar posibles diferencia en el modelo.

En la Fig. 3.9 se muestra la gráfica del pronóstico; en ella podemos observar que el patrón es similar al obtenido anteriormente; esto se debe a que el modelo resultante fué el mismo aunque con pequeña variación en los parámetros (ϕ_1, θ_1, $\theta_{1,12}$)

Fig. 3.8


EST. CLIMATOLOGICA
SAN ISIDRO
TEMP. d. C.
25
20
15
10
5
0
-5
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1987
PRONOS
LIMITES
OBSERV

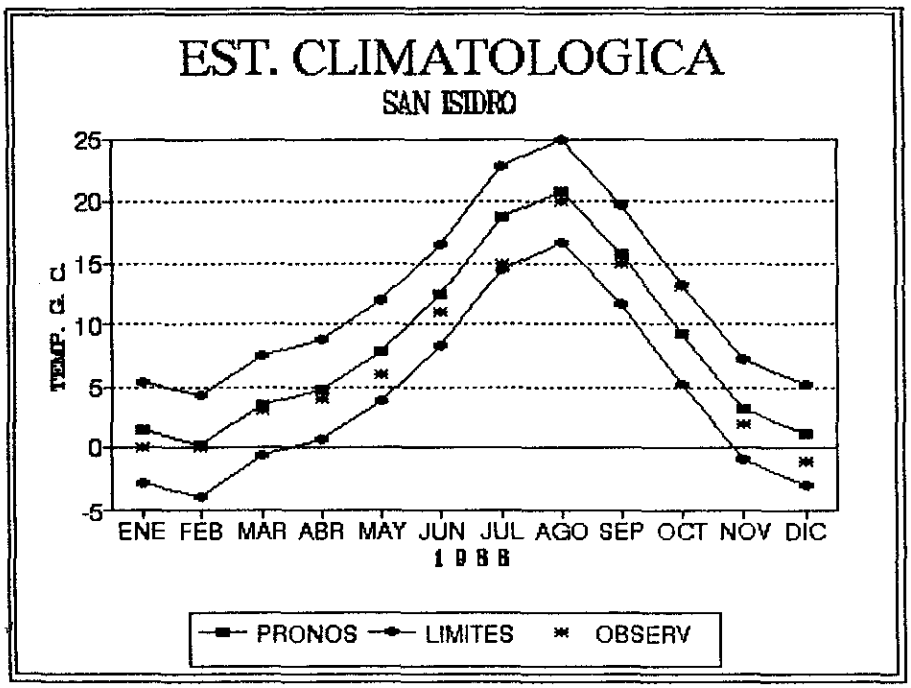
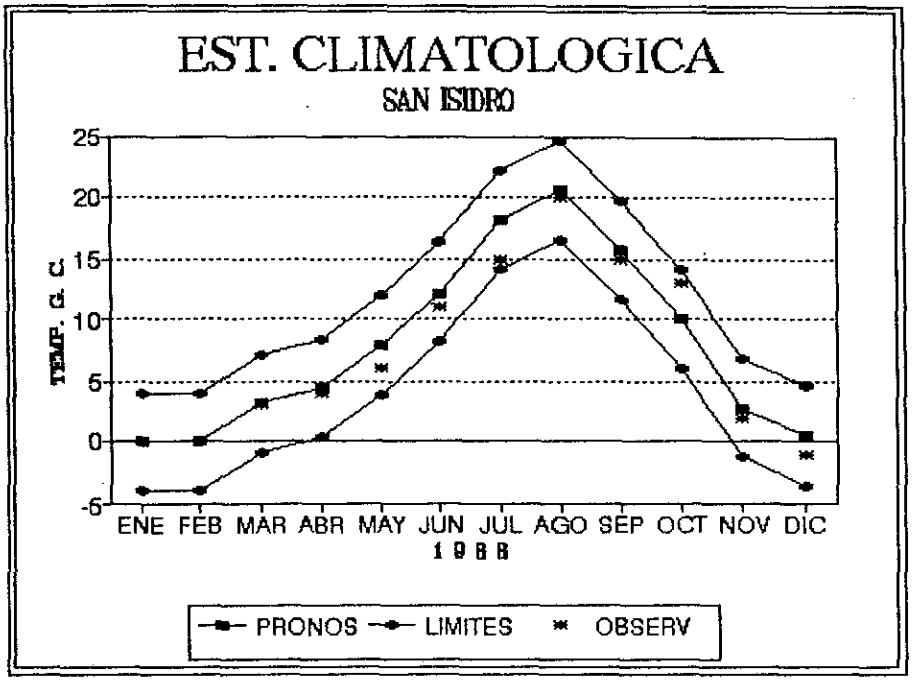

EST. CLIMATOLOGICA
SAN ISIDRO
TEMP. d. C.
25
20
15
10
5
0
-5
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1988
PRONOS
LIMITES
OBSERV

Fig. 3.9.


EST. CLIMATOLOGICA
SAN ISIDRO
TEMP. G. C.
25
20
15
10
5
0
-5
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1988
PRONOS
LIMITES
OBSERV

CAPITULO 4

APLICACIONES Y RESULTADOS

En este capítulo presentamos la aplicación de cada uno de los métodos expuestos anteriormente a las siguientes estaciones climatológicas.

1) Adolfo Ruiz Cortínez
2) Agua Prieta
3) Alvaro Obregon
4) Hermosillo
5) El Oregano
6) Sonoita

La selección de estas estaciones climatológicas se debe a que cada una de ellas pertenece a un clima distinto [1].

La información necesaria para la aplicación de cada uno de los métodos se tomó de los archivos que guarda la Comisión Nacional del Agua. En estos archivos existe un expediente para cada estación climatológica.

4.1. Adolfo Ruiz Cortínez

Esta estación climatológica, más conocida como presa "Mocuzari", se encuentra localizada en:

Latitud 27° 13′ 30"

Longitud 109° 06′ 00"

Altitud 147 M. S. N. M.

Municipio Alamos

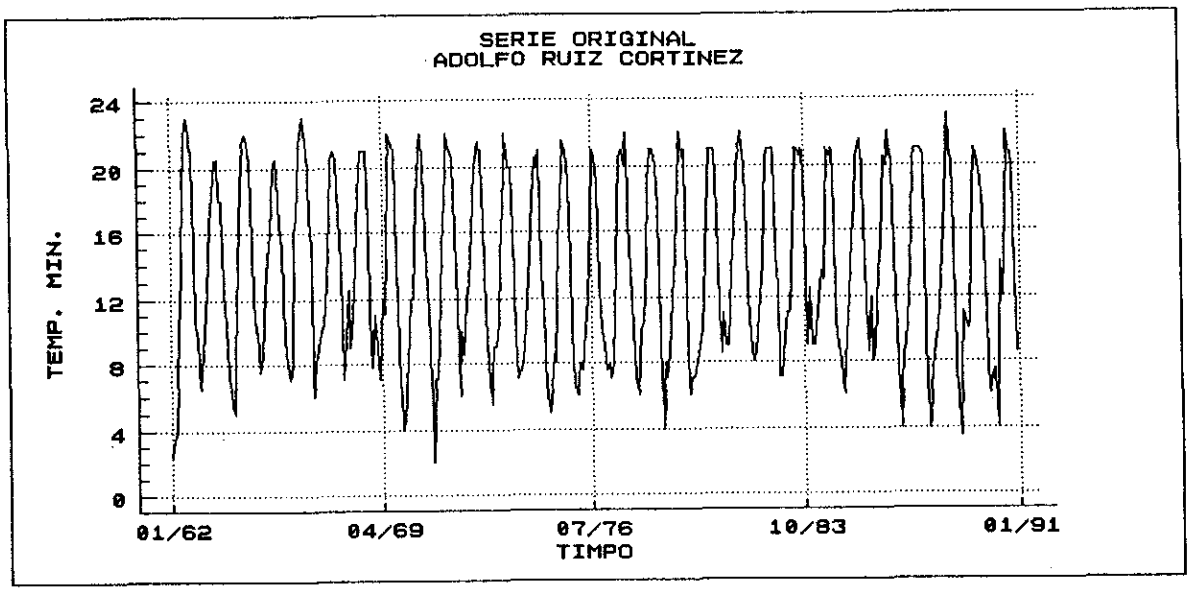
4.1.1. **Metodología de Thom.**

Según la información que se encuentra en el expediente de esta estación climatológica, podemos observar que durante el período comprendido entre Julio de 1955 hasta Junio de 1991, nunca se registró una helada. De esta manera, por el método de Thoms, la probabilidad de ocurra una helada es practicamente cero. Por razones obvias las gráficas de este caso se omiten.

4.1.2. **Metodología de Box-Jenkins.**

En la figura 4.1, se muestra la gráfica de las temperatura mínima mensual correspondiente a esta estación.

Fig. 4.1



Con el propósito de proponer un modelo tentativo para esta serie de tiempo, en la fig 4.2, se presentan las gráficas de la FACM y la FACPM, respectivamente.

Según el comportamiento que presentan estas funciones y basándonos en las guía 1 y 2 del capítulo anterior promponemos el siguiente modelo

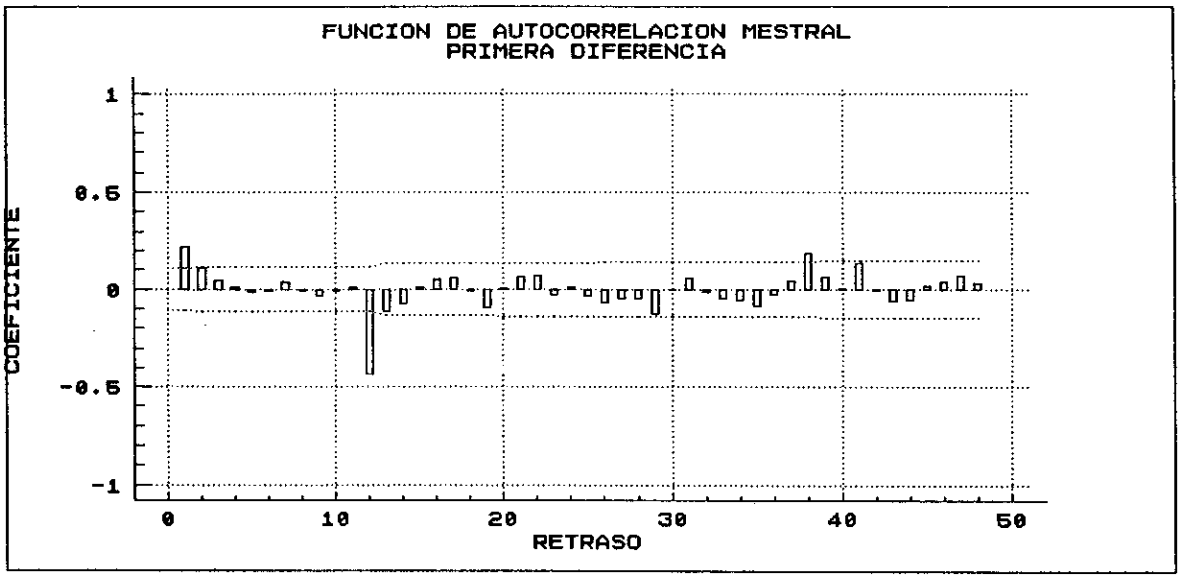
$$\phi_1(B)\,Z_t = \delta + \theta_1(B^{12})\,a_t,$$

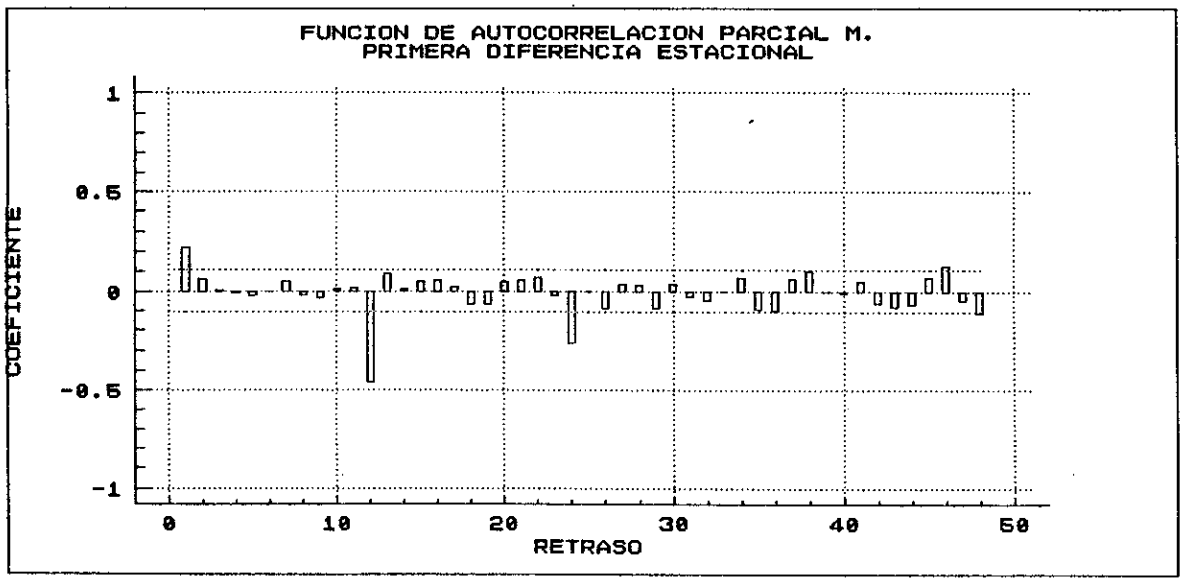
donde $Z_t = \nabla_{12} Y_t$, y Y_t representa a la serie original.

Se llevo a cabo la estimación de los parámetros que intervienen en el modelo, obteniéndose el siguiente resultado:

$$\phi_1 = 0.30235; \qquad \theta_{1,12} = 0.71331.$$

Fig. 4.2


FUNCION DE AUTOCORRELACION MESTRAL
PRIMERA DIFERENCIA
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO


FUNCION DE AUTOCORRELACION PARCIAL M.
PRIMERA DIFERENCIA ESTACIONAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

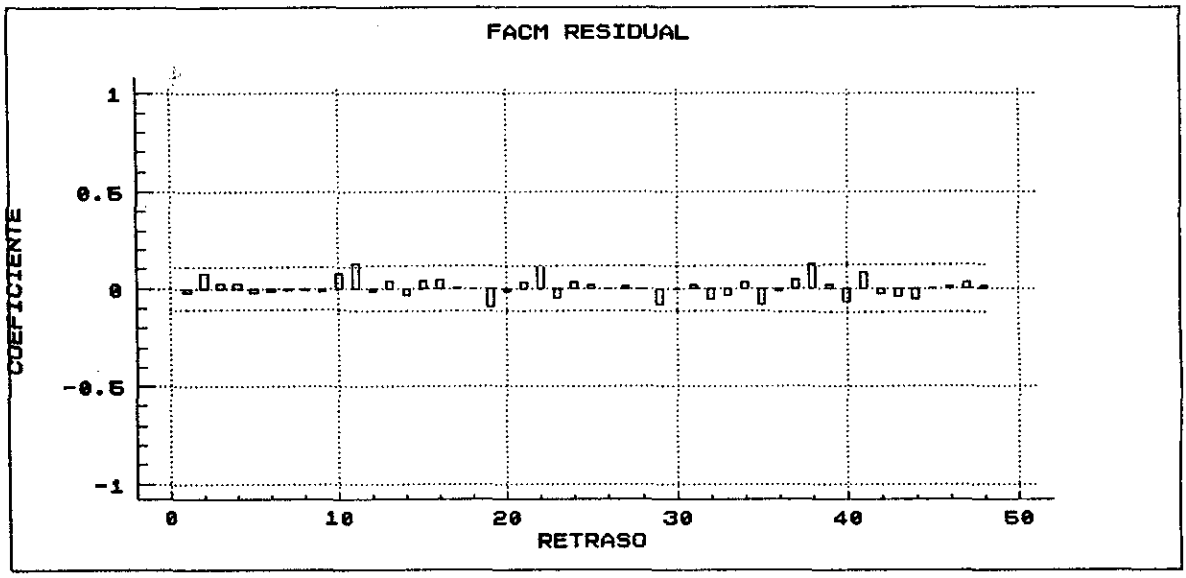
Además la constante δ no se acepta en el modelo. Obsérvese que los parámetros cumplen las condiciones de estacionariedad e invertibilidad, es decir:

$$|\phi_1| < 1 \quad \text{y} \quad |\theta_{1,12}| < 1.$$

En el paso de la comprobación se concluyó que el modelo es adecuado, ya que el valor calculado para Q = 15.2444 es menor que χ^2 = 28.9 con α = 0.05 y 18 gl. En la figura 4.3 se muestra la gráfica de FACM y la FACPM de los residuos, respectivamente. Dado su comportamiento, concluímos que los valores del ruido blanco son estadísticamente independientes.

Por último, en la figura 4.4 se muestra la gráfica de pronóstico para el período 1991 - 1992. (Esto se debe a que en el expediente de esta estación, se cuenta con información hasta el mes de Junio de 1991).

Fig 4.3


FACM RESIDUAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

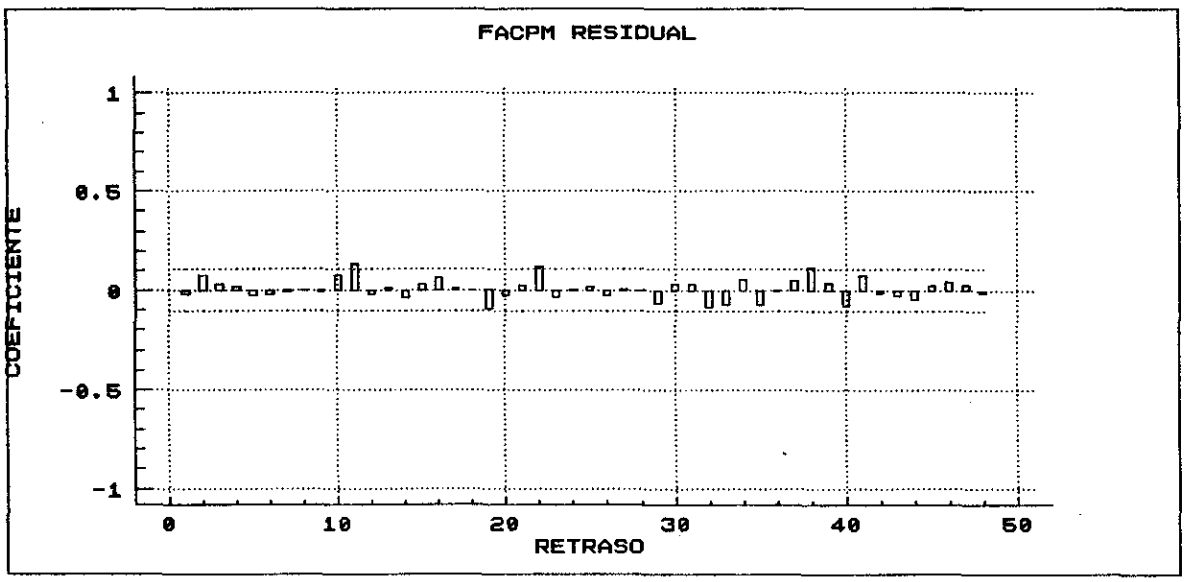
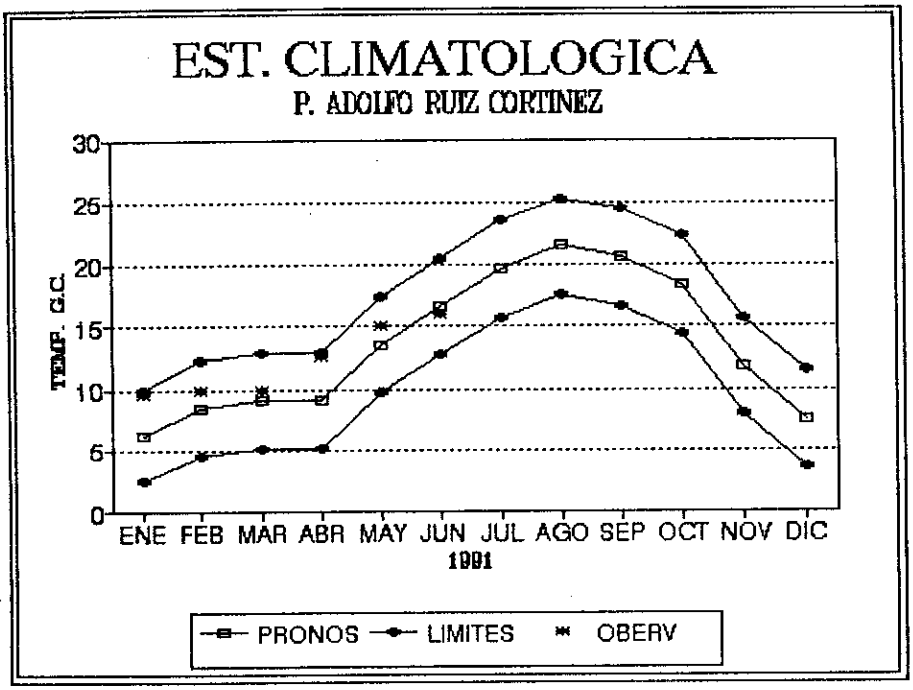
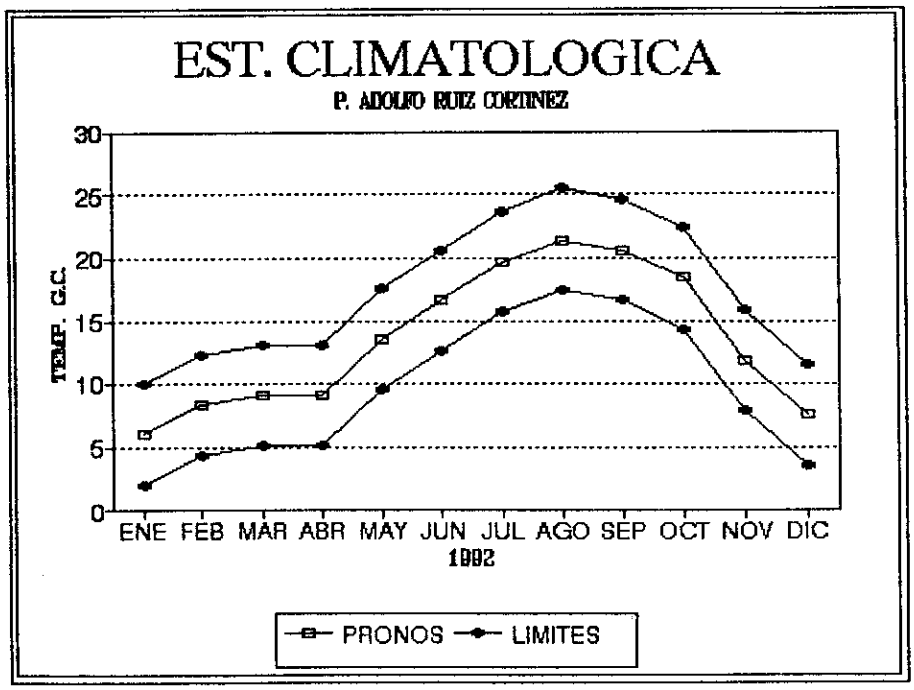

FACPM RESIDUAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

Fig 4.4


EST. CLIMATOLOGICA
P. ADOLFO RUIZ CORTINEZ
TEMP. G.C.
30
25
20
15
10
5
0
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1991
PRONOS
LIMITES
OBERV


EST. CLIMATOLOGICA
P. ADOLFO RUIZ CORTINEZ
TEMP. G.C.
30
25
20
15
10
5
0
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS
LIMITES

4.2. AGUA PRIETA

Esta estación se encuentra localizada en:

Latitud 31° 19′ 37"

Longitud 109° 32′ 28"

Altitud 1220 M. S. N. M.

Municipio Alamos

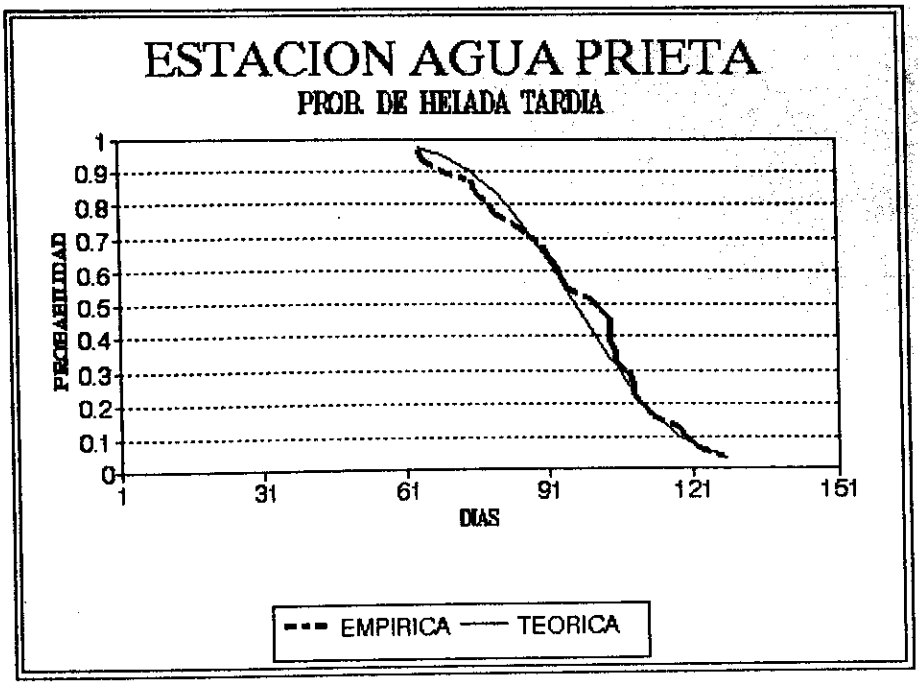
4.2.1 Metodología de Thom

A diferencia de la estación anterior, en esta estación se presentan fechas con helada en todos los años, tanto en el período de helada tardía com en el de helada temprana. De tal manera que la probabilidad de que ocurra una helada resulta ser un evento seguro.

En la figura 4.5 se muestran las gráficas de la distribución continua y la distribución mixta respectivamente, que representan la probabilidad de que ocurra una helada después de una fecha x. En cada una de ellas se muestran la función empírica y la función téorica.

Las gráficas correspondientes a la probabilidad de que ocurra una helada temprana se presentan en la figura 4.6.

Fig. 4.5


ESTACION AGUA PRIETA
PROB. DE HELADA TARDIA
PROBABILIDAD
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
31
61
91
121
151
DIAS
EMPIRICA
TEORICA

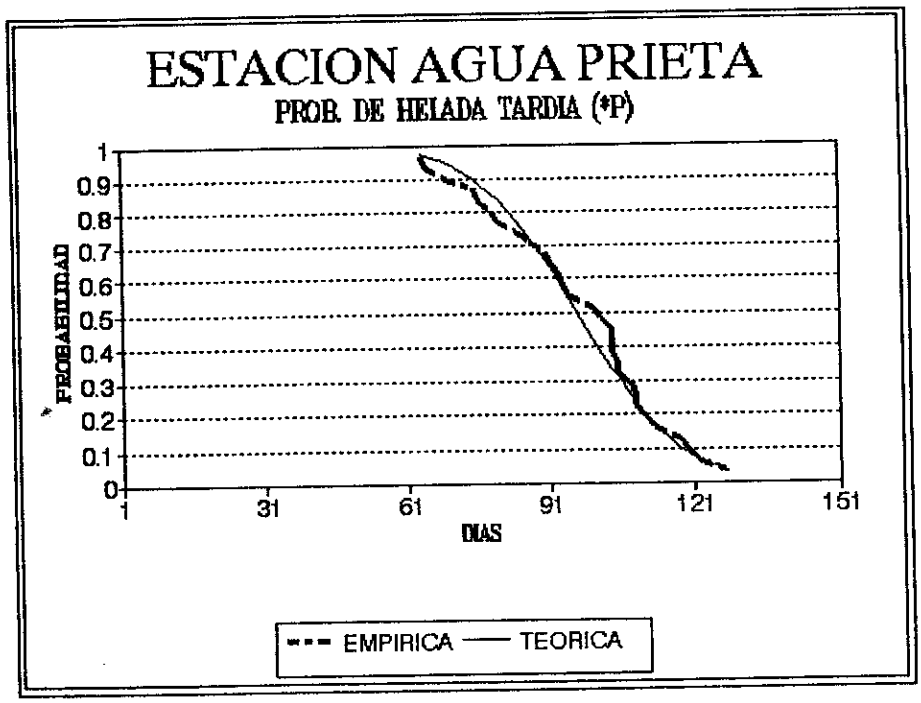
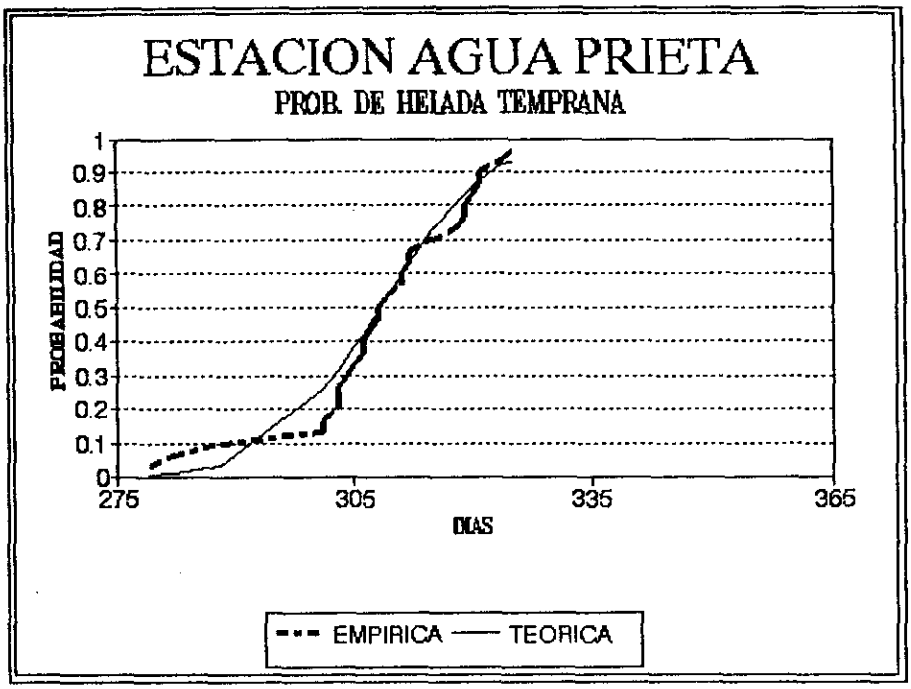
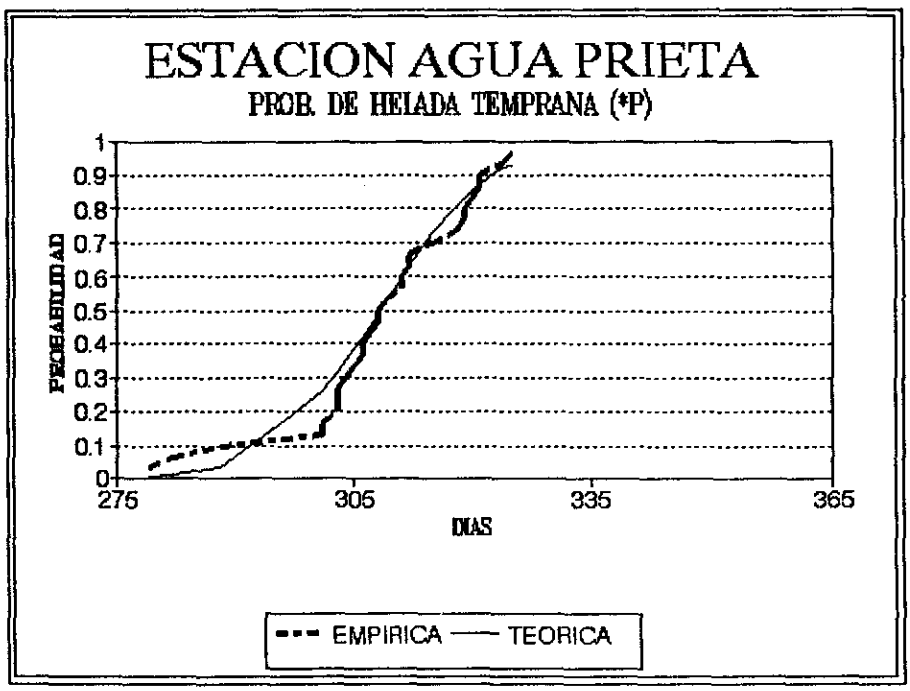

ESTACION AGUA PRIETA
PROB. DE HELADA TARDIA (*P)
PROBABILIDAD
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
31
61
91
121
151
DIAS
EMPIRICA
TEORICA

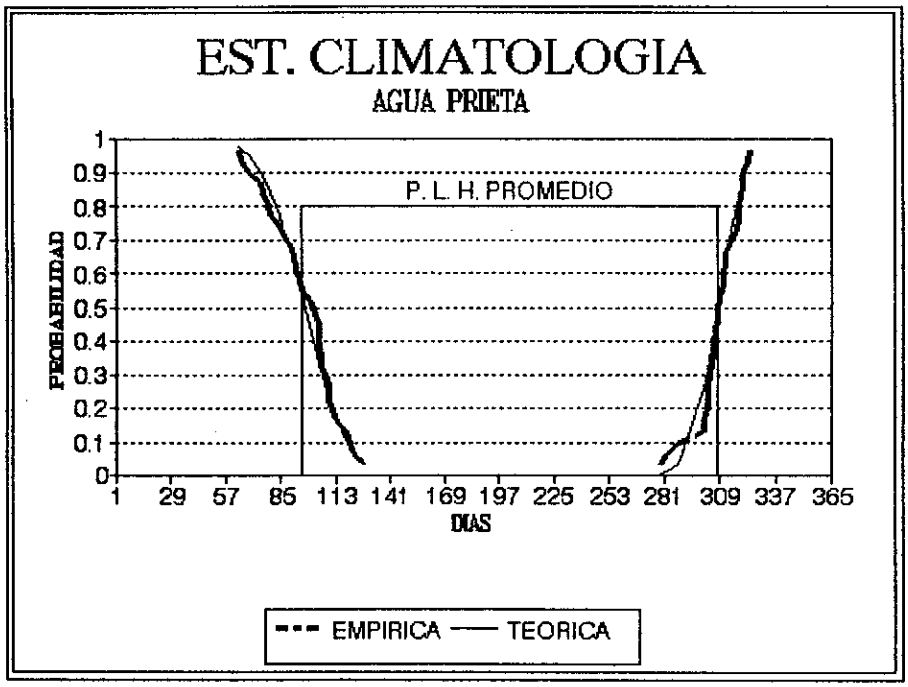
Fig. 4.6


ESTACION AGUA PRIETA
PROB. DE HELADA TEMPRANA
PROBABILIDAD
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
275
305
335
365
DIAS
EMPIRICA
TEORICA


ESTACION AGUA PRIETA
PROB. DE HELADA TEMPRANA (*P)
PROBABILIDAD
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
275
305
335
365
DIAS
EMPIRICA
TEORICA

La gráfica para el período libre de helada se presenta en la siguiente figura.

Fig. 4.7



4.2.2. Metodología de Box-Jenkins.

El modelo tentativo para la serie de tiempo que representa a los datos de esta estación, después de realizar un análisis a la FACM y FACPM (ver la figura 4.8), es de la forma:

$$Z_t = \delta + \theta_1(B^{12})a_t$$

donde $Z_t = \nabla_{12}Y_t$ y Y_t representa a la serie original.

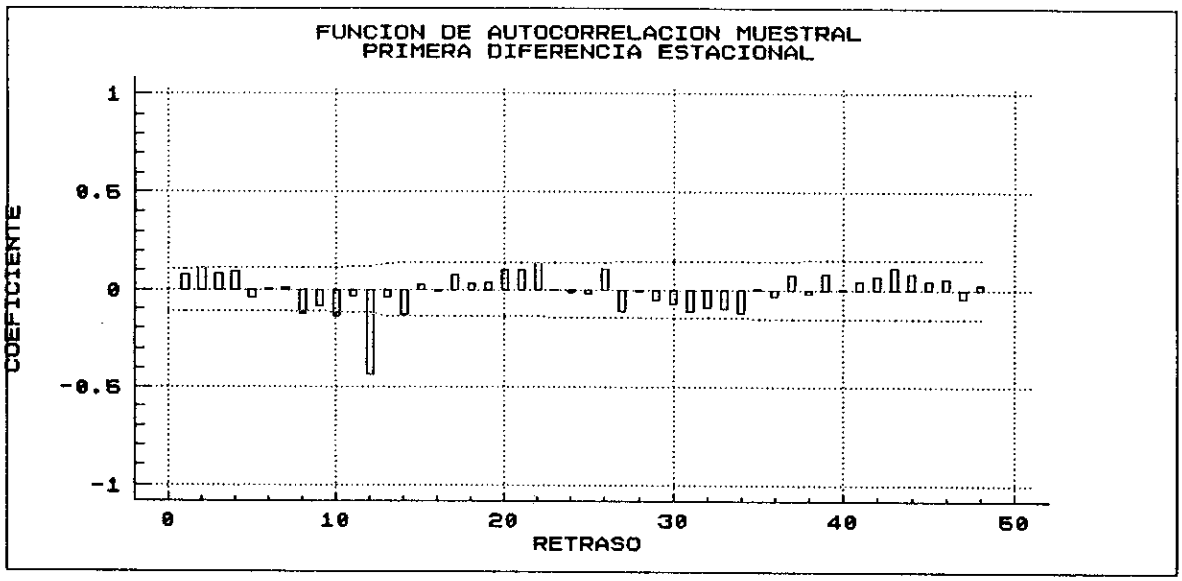
En el procedimiento de estimación de parámetros se determinó que la constante δ no intervenga en el modelo. El valor resultó:

$$\theta_{1,12} = 0.64783.$$

El valor calculado para Q = 17.9778 resultó menor que el valor de la χ^2 = 30.1 con α = 0.05 y 19 gl. Esto implica que el modelo es adecuado.

En la Fig. 4.9 se muestra el pronóstico para el período 1992 - 1993 utilizando el modelo anterior.

Fig. 4.8


FUNCION DE AUTOCORRELACION MUESTRAL
PRIMERA DIFERENCIA ESTACIONAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

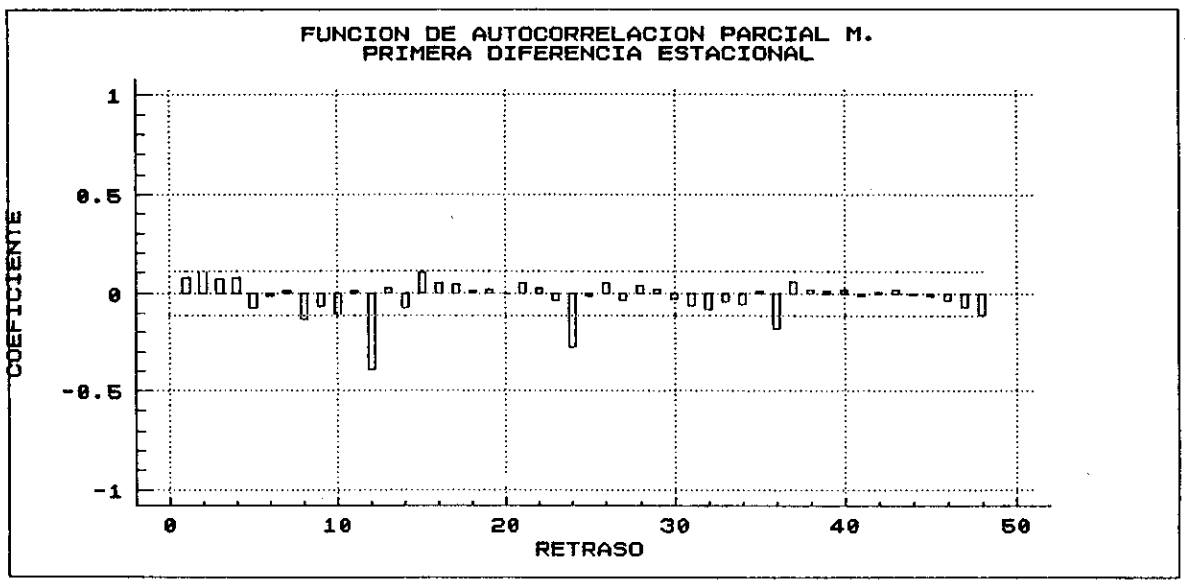
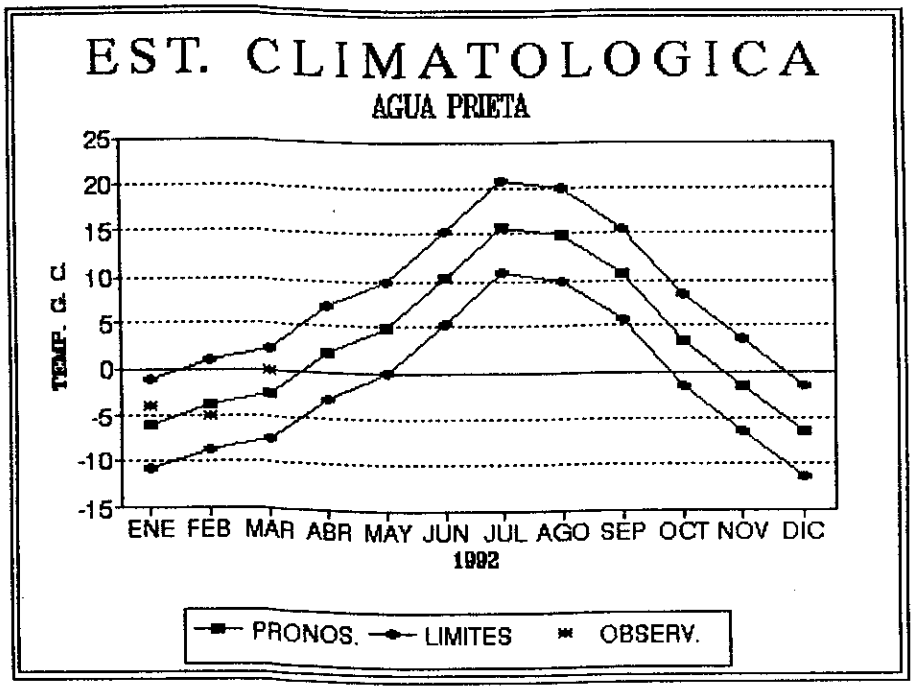
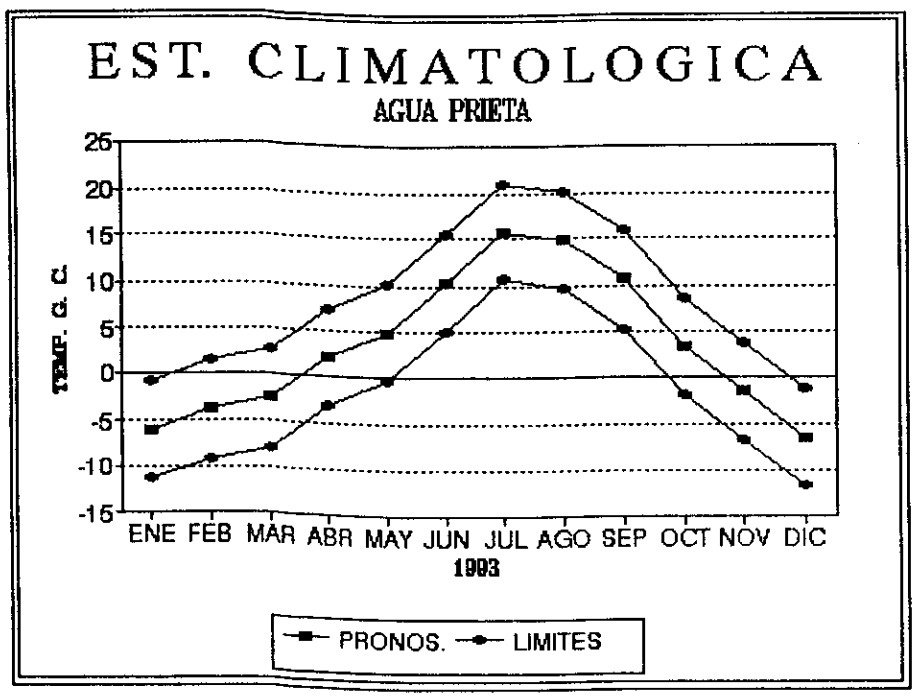

FUNCION DE AUTOCORRELACION PARCIAL M.
PRIMERA DIFERENCIA ESTACIONAL
COEFICIENTE
1
0.5
0
-0.5
-1
0
10
20
30
40
50
RETRASO

Fig 4.9

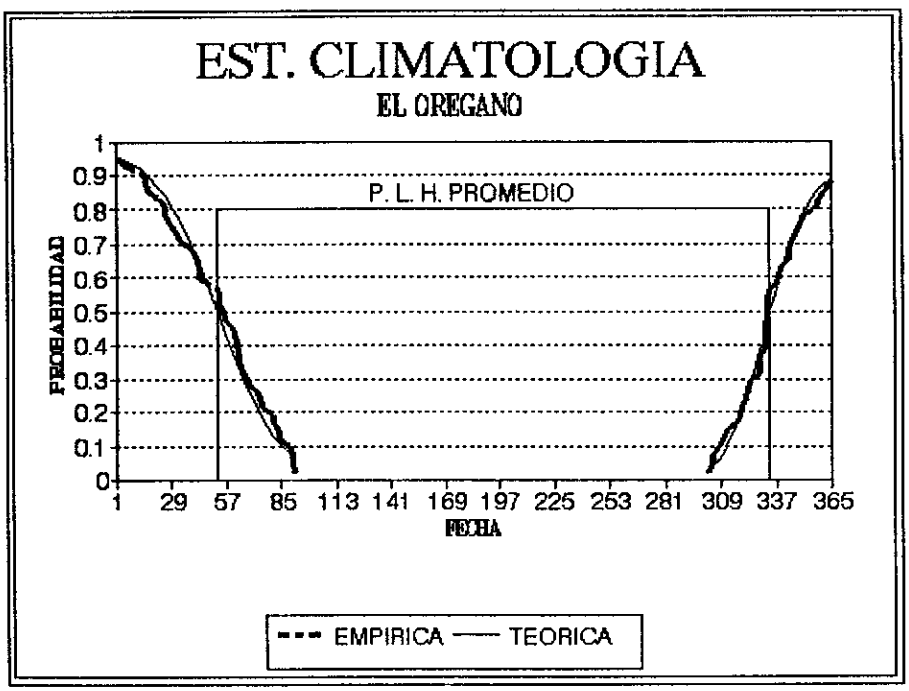

EST. CLIMATOLOGICA
AGUA PRIETA
TEMP. G. C.
25
20
15
10
5
0
-5
-10
-15
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS.
LIMITES
OBSERV.

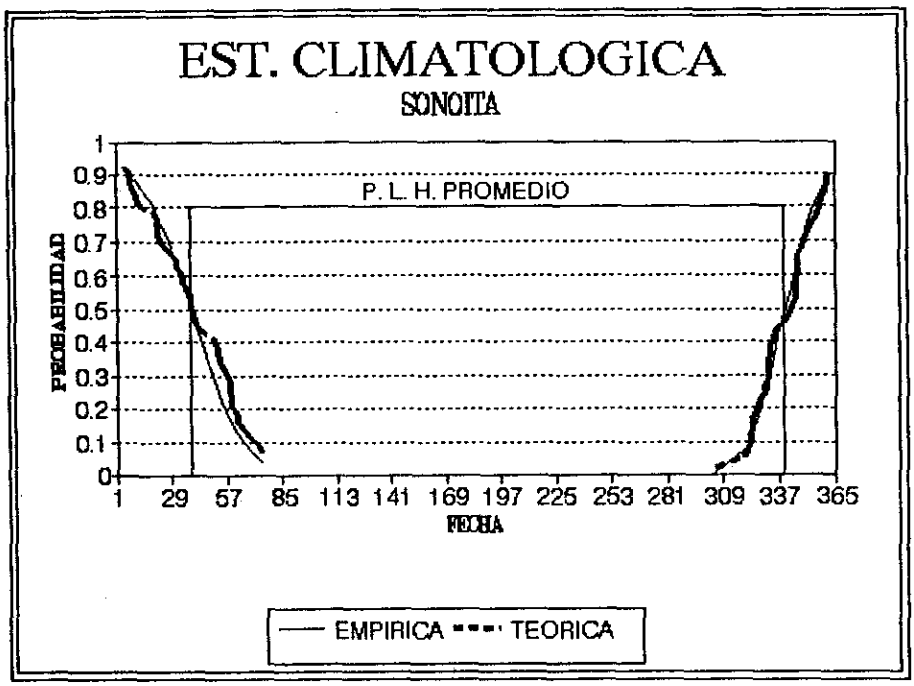

EST. CLIMATOLOGICA
AGUA PRIETA
TEMP. G. C.
25
20
15
10
5
0
-5
-10
-15
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1993
PRONOS.
LIMITES

4.3. METODOLOGIA DE THOM (para las estaciones restantes).

Los expedientes correspondientes a las estaciones climatológicas Presa Alvaro Obregón y Hermosillo indican que en éstas jamás ha ocurrido una helada; esto implica que la probabilidad de helada, estimada por el método de Thoms es practicamente cero.

A continuación presentamos las gráficas del período libre de helada, correspondiente a las estaciones El Oregano y Sonoita.



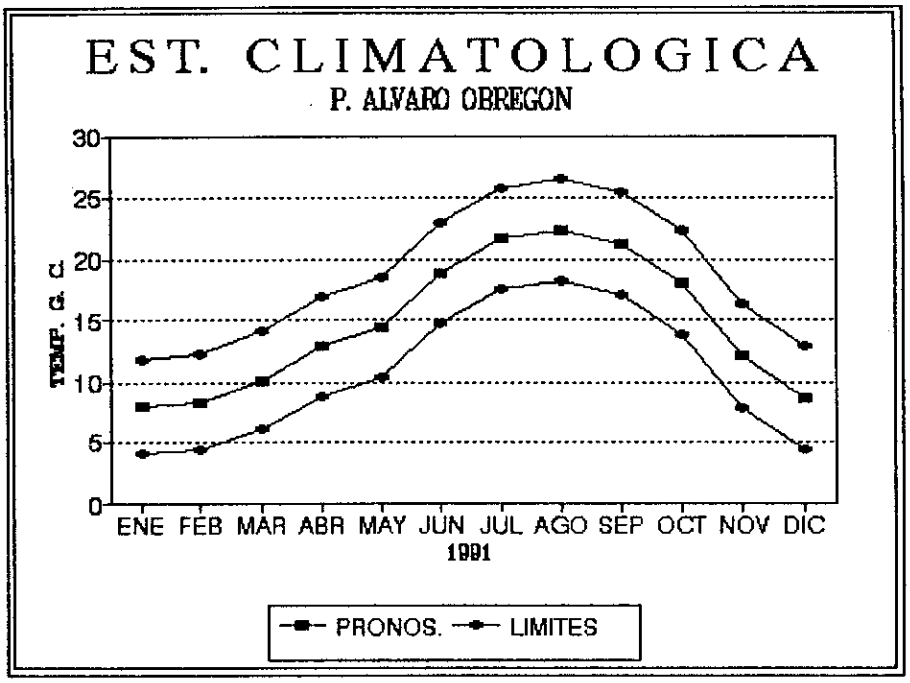


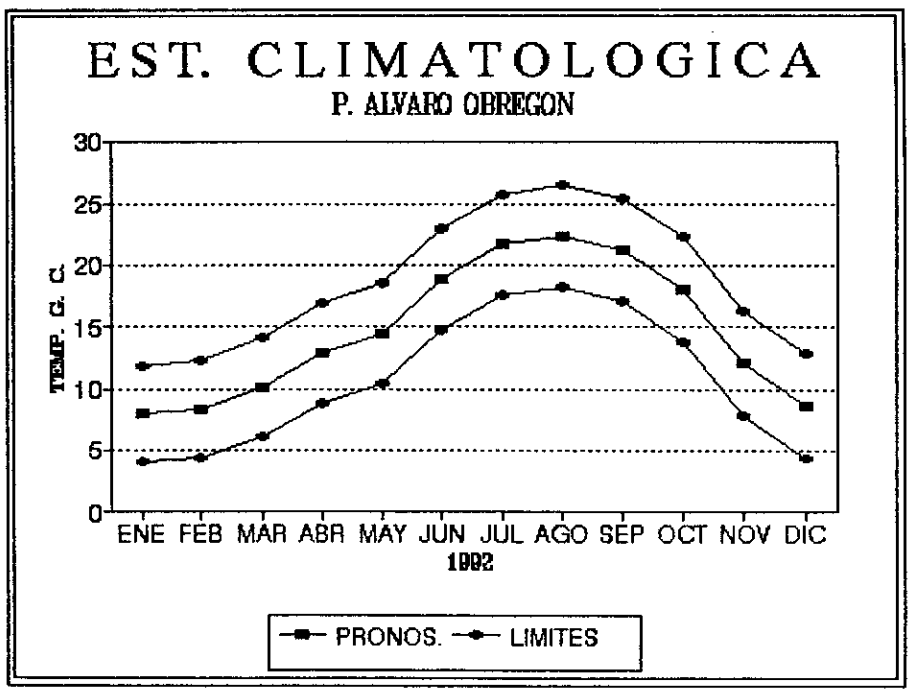
4.4. MET. DE BOX-JENKINS (para las estaciones restantes).

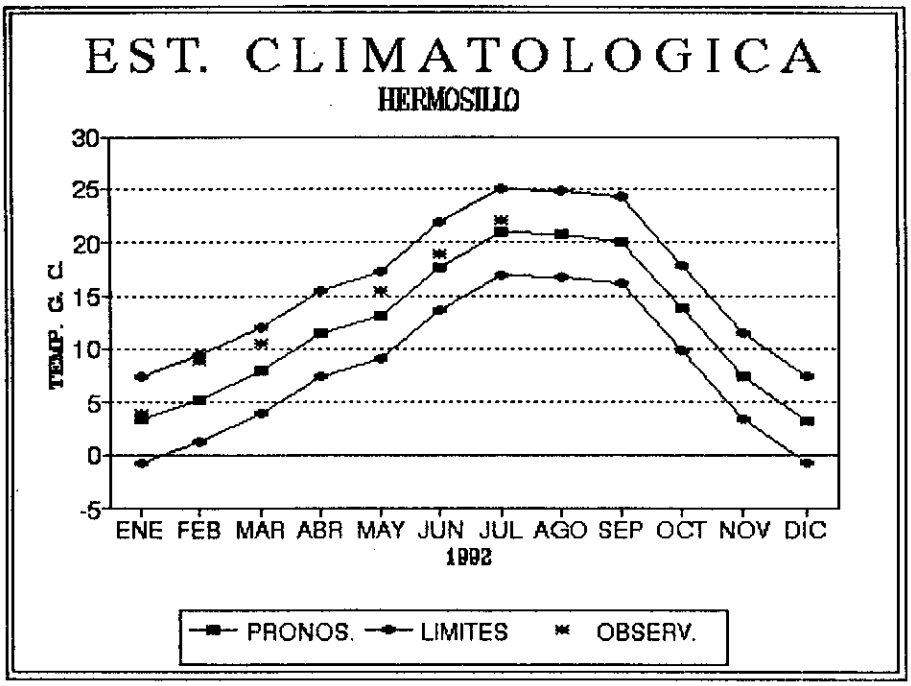
El procedimiento de la aplicación de la metodología de Box-Jenkins sobre las estaciones restantes, se llevó a cabo de manera similar a la expuesta anteriormente. Presentamos a continuación, en el Cuadro 4.1, los resultados obtenidos de esta aplicación. Y posteriormente la gráficas de pronóstico en cada estación. En cada caso el modelo propuesto resultó adecuado.

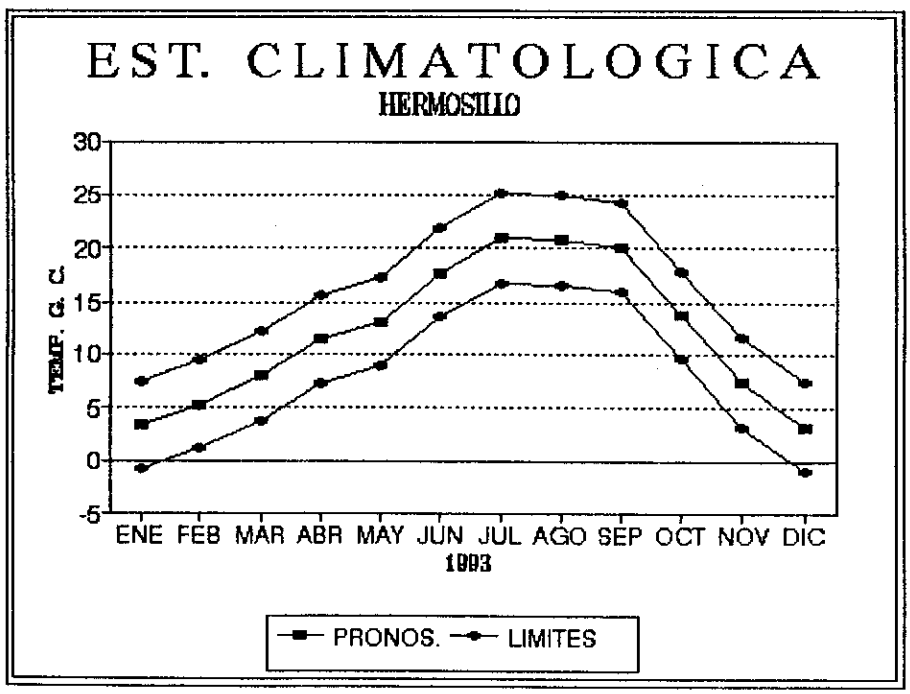
CUADRO 4.1. **Resumen de modelos econtrados para las presas restantes.**

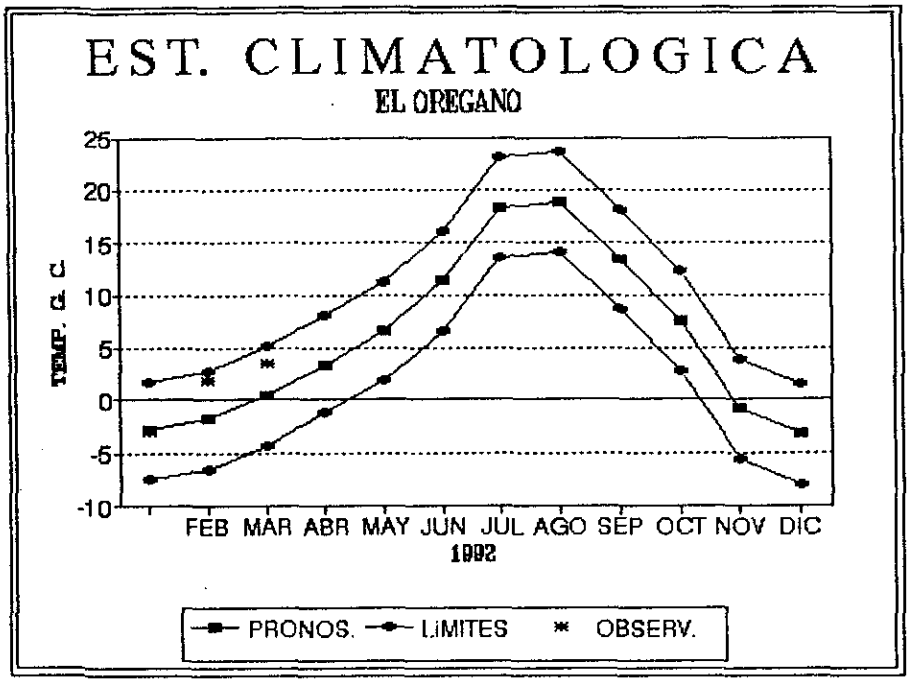
No. Est.	Modelo	Est. de Par.	Est. Q
3	$\phi_1(B)Z_t = \delta + \theta_1(B)\theta_1(B^{12})a_t$	$\phi_1 = 0.95359$ $\theta_1 = 0.79155$ $\theta_{1,12} = 0.7707$	21.9
4	$\phi_1(B)Z_t = \delta + \theta_1(B^{12})a_t$	$\phi_1 = 0.13300$ $\theta_{1,12} = 0.76051$	19.8
5	$\phi_1(B)Z_t = \delta + \theta_1(B)\theta_1(B^{12})a_t$	$\phi_1 = 0.85838$ $\theta_1 = 0.73641$ $\theta_{1,12} = 0.8794$	19.1
6	$Z_t = \delta + \theta_1(B)\theta_1(B^{12})a_t$	$\theta_1 = -.10997$ $\theta_{1,12} = 0.83324$	16.9

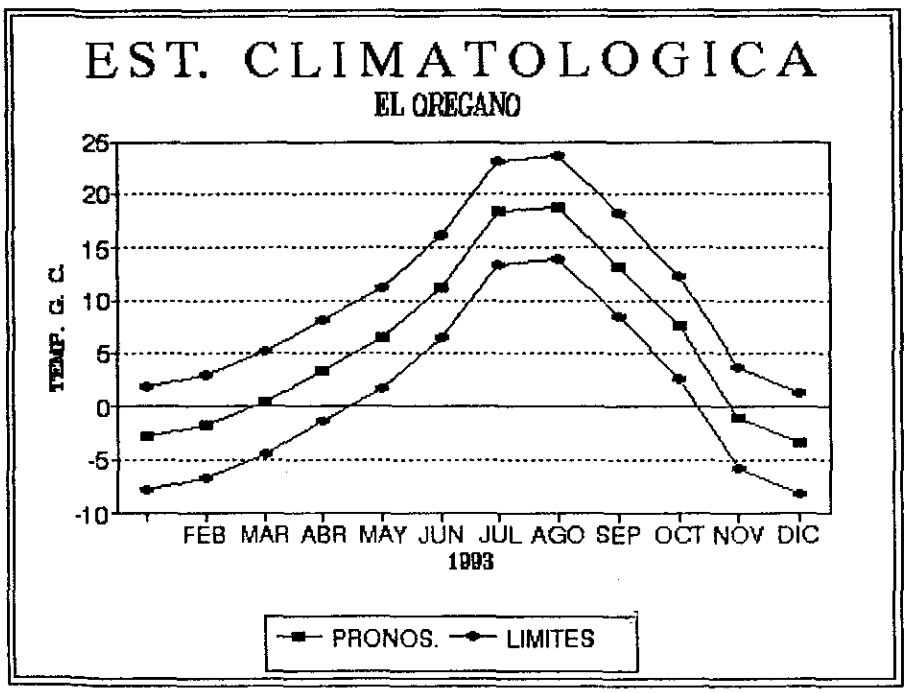

EST. CLIMATOLOGICA
P. ALVARO OBREGON
TEMP. G. C.
30
25
20
15
10
5
0
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1991
PRONOS.
LIMITES

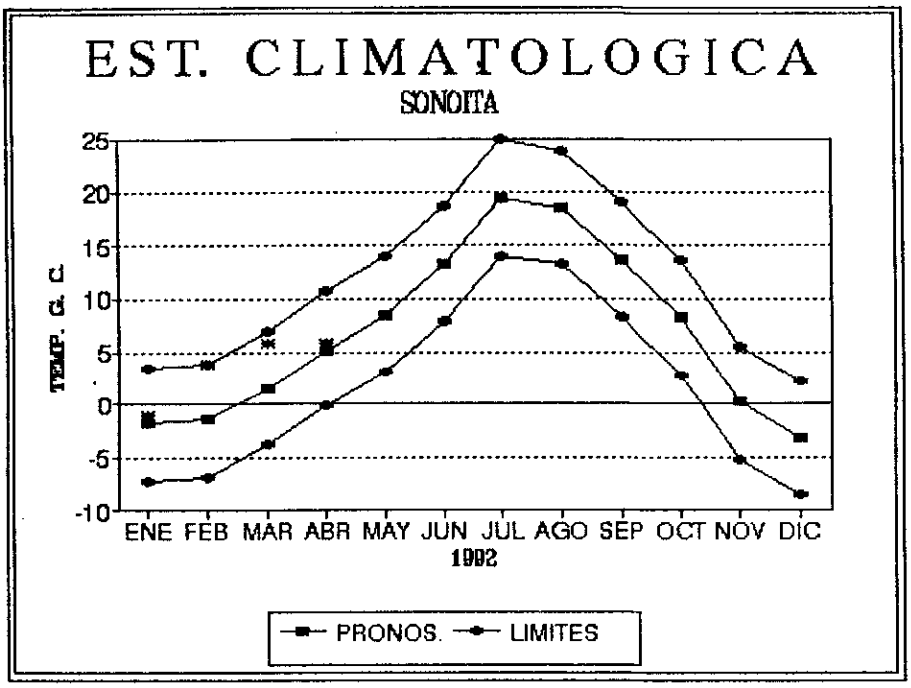

EST. CLIMATOLOGICA
P. ALVARO OBREGON
TEMP. G. C.
30
25
20
15
10
5
0
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS.
LIMITES

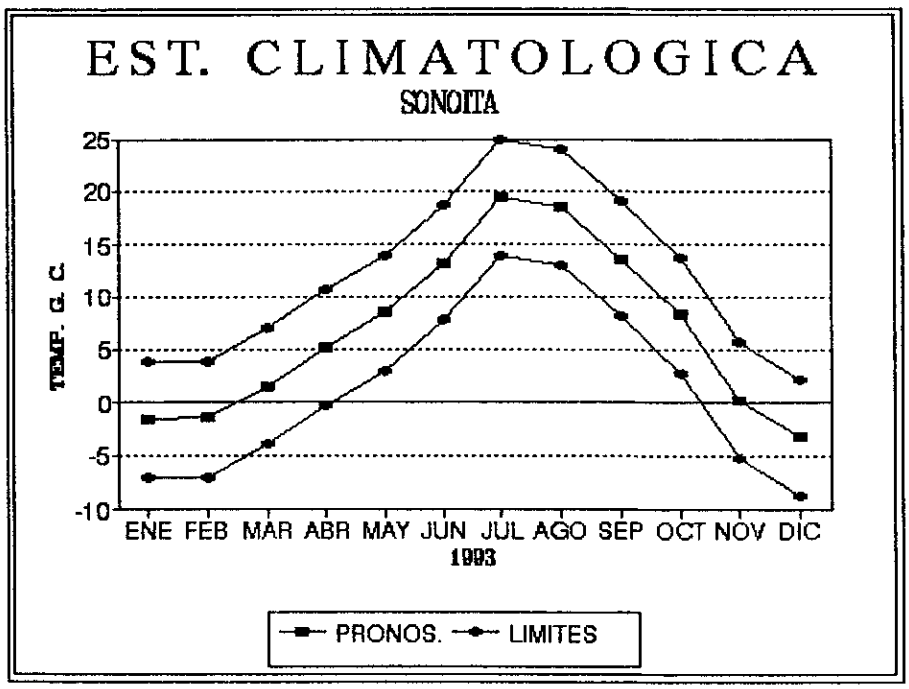

EST. CLIMATOLOGICA
HERMOSILLO
TEMP. G. C.
30
25
20
15
10
5
0
-5
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS.
LIMITES
OBSERV.


EST. CLIMATOLOGICA
HERMOSILLO
TEMP. G. C.
30
25
20
15
10
5
0
-5
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1993
PRONOS.
LIMITES


EST. CLIMATOLOGICA
EL OREGANO
TEMP. G. C.
25
20
15
10
5
0
-5
-10
FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS.
LIMITES
OBSERV.


EST. CLIMATOLOGICA
EL OREGANO
TEMP. G. C.
25
20
15
10
5
0
-5
-10
FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1993
PRONOS.
LIMITES


EST. CLIMATOLOGICA
SONOITA
TEMP. G. C.
25
20
15
10
5
0
-5
-10
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1992
PRONOS.
LIMITES


EST. CLIMATOLOGICA
SONOITA
TEMP. G. C.
25
20
15
10
5
0
-5
-10
ENE FEB MAR ABR MAY JUN JUL AGO SEP OCT NOV DIC
1993
PRONOS.
LIMITES

CAPITULO 5

CONCLUSIONES Y RECOMENDACIONES

En este capítulo presentamos conclusiones y recomendaciones producto de la realización de esta tesis.

Exponemos primero la justificación del uso de la temperatura mínima extrema; después, en la sección 5.2, presentamos algunas recomendaciones para el uso de la metodología de Thom. Por último, en la sección 5.3, presentamos la relación que guarda esta metodología con la de Box-Jenkins, para aquellas estaciones climatológicas en las cuales tenga sentido aplicar ambas.

5.1. JUSTIFICACION DEL USO DE TEMPERATURA MINIMA EXTREMA.

Por el Anexo 1 sabemos que en cada estación climatológica se hacen tres lecturas de temperaturas, cada 24 horas (todos los días a las 8:00 hs.). Estas son la Máxima, la Mínima y la de Ambiente.

La temperatura Máxima es la temperatura más alta

registrada en las últimas 24 hs, anteriores a la hora de tomar la lectura.

La temperatura Ambiente es la que esta ocurriendo a la hora de la lectura.

La temperatura Mínima es la temperatura más baja registrada en las últimas 24 hs, anteriores.

Por otro lado, sabemos que una helada se presenta, en una fecha x, cuando el termómetro registra una temperatura menor o igual a cero grados centígrados.

Partiendo de las definiciones anteriores, podemos concluir que si en una fecha x ocurre una helada, ésta se debió a que se registró una temperatura Mínima menor o igual a cero grados centígrados. Aquí surge la siguiente pregunta; ¿Puede registrarse una temperatura menor o igual a cero grados centígrados y considerarse como Máxima ó Ambiente?.

La respuesta es que sí, existen lugares, Yécora es un ejemplo, en los que en más de 24 hs continuas, la temperatura se mantiene por abajo de cero grados. Sin embargo, la temperatura Mínima será siempre menor a la temperatura Máxima, inclusive a la de Ambiente.

Otro motivo del porqué usamos la temperatura Mínima es

el siguiente: Supóngase que en cierta estación climatológica el último mes del invierno es Marzo (el mes en que el frío empieza a desaparecer). Si en una día x de ese mes, ocurre una temeratura menor o igual a cero grados centígrados, ésta quedará registrada como temperatura Mínima y, además, será tomada como fecha en que ocurre una helada.

Por último, explicamos el motivo del porqué usamos en esta tesis la temperatura Mínima Extrema (temperatura mínima de las Mínimas en el mes). De acuerdo con lo anterior, la menor temperatura registrada de las temperaturas Mínimas ocurridas en un mes es la que nos indica si hubo un helada.

5.2. RECOMENDACIONES PARA EL USO DE LA METODOLOGIA DE THOM.

Como pudimos darnos cuenta en el capítulo anterior, existen lugares en el Estado de Sonora en el que jamás ha ocurrido una helada. La aplicación de la metodología de Thom en esos lugares indicaría que la probabilidad de que ocurra una helada es cero.

Con el fin de obtener mejores resultados con el uso de la metodología de Thom, recomendamos su aplicación en aquellas estaciones en las que por lo menos exista una ocurrencia de fecha con helada para cada serie en por lo menos la mitad de los años en estudio. Esto quiere decir que si en una estación se cuenta con 30

años de estudio, por lo menos han de existir 15 años con al menos una fecha con heladas para cada serie.

5.3. RELACION ENTRE LOS METODOS.

La relación que existe entre ambos métodos la podemos plantear de la siguiente manera:

La metodología de Thom determina la probabilidad de que ocurra una helada después de una fecha x, en el primer semestre del año, y la probabilidad de que ocurra una helada antes de la fecha x, en el segundo semestre.

La metodología de Box-Jenkins proporciona el comportamiento futuro de la temperatura mínima en todo el año.

Por lo tanto, con la primera metodología obtenemos la probabilidad con la que se puede presentar una helada antes o después de una fecha x cualquiera y con la segunda, la temperatura mínima extrema esperada para cada mes; Por ejemplo.

En el cuadro 2.1 podemos observar que la probabilidad de que ocurra una helada después del 2 de Febrero es del .713, esto implica que existe una mayor probabilidad de que ocurra una helada en el mes de Enero. La gráfica 3.9 revela que la temperatura mínima esperada para el mes de Enero es de cero grados centígrados.

ANEXO 1: NOMENCLATURA

Con la finalidad de contar con un vocabulario común, conviene definir una serie de conceptos.

Meteorología.- Ciencia que estudia las propiedades de la atmósfera y los fenómenos físicos y dinámicos que en ella tienen lugar.

Estación Meteorológica.- Sitio donde se hacen observaciones del comportamiento de la atmósfera y del medio ambiente, ya sea con fines climatológicos, agropecuarios, etc.

Estación Meteorológica Sinóptica.- Es la que tiene como finalidad conocer el "Estado del Tiempo" y su previsión a corto plazo.

Estación Meteorológica Climatológica.- Observa el estado del tiempo, acumulando datos para fines estadísticos.

El valor de estos datos es mayor cuando se extienden a un período más largo. Dichos datos permiten caracterizar y clasificar el clima de un lugar e inclusive conocer sus tendencias futuras.

Los fenómenos principales que se miden en estas estaciones

son: Temperatura Ambiente, Temperatura Máxima, Temperatura Mínima, Precipitación, Evaporación y Días con Heladas.

Otros fenómenos, de menor importancia, que también se miden son: Intensidad de los vientos, granizada, nevada, etc.

Temperatura.- Una de las cualidades físicas de la materia consiste en la propiedad que tiene de almacenar calor en mayor o menor grado. Comparando la cantidad de calor almacenada en un cuerpo con la cantidad almacenada en otro, se obtiene el concepto de temperatura.

Para la medición de las temperaturas ambiente, máxima y mínima se utiliza el termómetro llamado "Six" y generalmente se realiza a las 8:00 hs.

El termómetro "Six" esta formado por un tubo capilar en forma de "U"; utiliza mercurio y alcohol y un índice esmaltado con alma de hierro en cada rama de la U, de manera que una rama marca la temperatura máxima (derecha) y la otra la mínima. Se usa un imán para volver los índices a su posición original.

Helada.- Temperatura del aire menor o igual a 0°C, medida por el termómetro "Six" ubicado en una caseta reglamentaria a una altura de 1.5 m.

ANEXO 2: DEFINICIONES

Definición.- Una serie de tiempo es una sucesión de observaciones cronológicas sobre una variable en particular.

Una serie de tiempo generalmente esta formada por cuatro componentes, que son:

1).- Tendencia.

2).- Componente Cíclica.

3).- Componente Estacional.

4).- Componente Aleatoria.

Definición. Se dice que una serie de tiempo contiene la componente tendencia, si mantiene un movimiento ascendente o descendente sobre un período largo de tiempo.

Definición. El efecto cíclico en una serie de tiempo es identificado como movimientos suaves hacia arriba y hacia abajo siguiendo la curva de tendencia.

Definición. El efecto estacional se refiere a movimientos ascendentes y descendentes en un período no mayor de un año; además, este patrón se repite año con año.

Definición. La aleatoriedad son movimientos erráticos en una serie de tiempo, que no siguen un patrón regular.

TABLA 1a.

ESTACION CLIMATOLOGICA SAN ISIDRO

FECHA DE HELADAS TARDIA Y TEMPRANAS

	HELADA	
AÑO	TARDIA	TEMPRANA
1957	Feb 2	Dic 3
1958	Ene 24	Nov 19
1959	Feb 27	No hubo
1960	Feb 25	Dic 7
1961	Feb 28	No hubo
1962	Mar 16	No hubo
1963	Ene 24	Dic 24
1964	Mar 10	No hubo
1965	Mar 6	No hubo
1966	Feb 26	Dic 28
1967	Feb 22	Dic 10
1968	Ene 6	Dic 5
1969	Feb 2	Dic 30
1970	Mar 22	Dic 13
1971	Mar 11	Dic 9
1972	Feb 17	Dic 14
1973	Ene 7	Nov 20
1974	Mar 10	Dic 10
1975	Mar 31	Dic 2
1976	Mar 15	Nov 29
1977	Abr 2	No hubo
1978	Ene 26	Dic 8
1979	Feb 6	Dic 13
1980	Ene 8	No hubo
1981	Feb 24	Dic 26
1982	Feb 3	No hubo
1983	Ene 22	No hubo
1984	Feb 25	Nov 27
1985	Feb 3	Dic 11
1986	No hubo	Dic 12
1987	Ene 20	Dic 14
1988	Feb 18	Dic 27

TABLA 1b.

ESTACION CLIMATOLOGICA SAN ISIDRO

TEMPERATURAS MINIMAS

REGISTRO DE TEMPERATURAS MINIMAS

(expresadas en grados centigrados)

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL
1957	1.0	1.0	1.5	4.0	7.0	11.5	23.0
1958	-1.0	2.0	4.0	6.0	10.0	13.2	23.5
1959	0.0	-0.1	1.0	5.0	8.0	14.0	24.0
1960	0.0	-1.0	3.0	4.0	7.0	11.0	21.0
1961	1.0	1.0	1.0	3.5	6.0	11.0	16.0
1962	3.5	0.0	-1.0	5.0	7.0	10.0	20.0
1963	-2.0	3.0	2.0	4.0	8.0	11.0	22.0
1964	-4.0	-3.0	-1.0	3.0	4.0	11.0	19.0
1965	0.0	-2.0	1.0	5.0	5.0	9.0	22.0
1966	0.0	0.0	2.0	7.0	9.5	13.0	19.0
1967	-0.5	-0.5	2.0	4.0	5.0	11.0	18.0
1968	2.0	2.5	3.0	4.5	7.0	10.0	21.5
1969	0.0	0.0	1.0	6.5	6.0	9.5	21.0
1970	-2.0	0.0	2.0	4.0	8.0	12.0	16.5
1971	-6.5	-0.5	-1.5	3.5	7.5	9.0	19.0
1972	-1.5	-2.0	3.5	3.5	7.5	16.0	20.0
1973	-2.5	3.0	3.5	5.0	8.5	12.0	17.5
1974	0.0	0.0	1.0	4.0	8.0	10.0	18.5
1975	-2.0	-1.0	2.0	2.0	5.0	12.0	20.0
1976	-2.0	1.0	1.0	4.0	8.0	11.0	17.0
1977	-2.0	-0.5	1.5	3.0	7.0	11.0	21.0
1978	0.0	2.0	5.0	6.0	8.0	11.0	18.0
1979	-1.0	-0.5	3.0	4.0	8.0	12.0	17.0
1980	0.0	1.0	4.0	5.0	8.0	11.0	11.0
1981	5.0	2.0	5.0	5.0	11.0	15.0	16.0
1982	1.0	0.0	3.0	6.0	7.0	12.0	18.0
1983	2.0	2.0	5.0	3.0	9.0	12.0	18.0
1984	1.0	2.0	2.0	5.0	7.0	16.0	21.0
1985	2.0	-5.0	2.0	4.0	8.0	12.0	21.0
1986	4.0	1.0	6.0	6.0	7.5	11.7	19.0
1987	-4.0	1.0	3.0	4.0	9.0	12.0	16.0
1988	0.0	0.0	3.0	4.0	6.0	11.0	15.0

Cont. **TABLA 1b.**

ESTACION CLIMATOLOGICA SAN ISIDRO

TEMPERATURAS MINIMAS

REGISTRO DE TEMPERATURAS MINIMAS

(expresadas en grados centigrados)

AÑO	AGO	SEP	OCT	NOV	DIC
1957	22.0	14.0	9.5	3.0	2.0
1958	22.0	18.0	7.0	3.0	1.0
1959	20.0	16.0	12.0	3.0	2.0
1960	21.0	16.0	9.0	3.0	-3.0
1961	21.0	14.0	8.0	2.0	2.0
1962	22.0	16.0	10.0	5.0	3.0
1963	22.0	19.0	12.0	5.0	0.0
1964	20.0	15.0	11.0	3.0	1.0
1965	22.0	11.0	9.0	6.0	3.0
1966	21.0	19.0	8.0	5.0	-2.5
1967	22.0	19.0	10.5	5.5	1.5
1968	18.0	15.0	9.0	1.0	-2.0
1969	21.0	11.5	6.0	3.0	0.0
1970	20.0	15.0	6.5	4.5	1.0
1971	20.0	14.0	2.0	1.5	-3.0
1972	19.5	13.5	8.5	3.0	-1.0
1973	19.5	14.0	6.0	2.0	0.0
1974	20.5	16.0	7.0	4.0	-3.5
1975	18.0	15.0	6.0	3.0	0.0
1976	16.0	15.0	7.0	2.0	2.0
1977	20.0	16.0	11.0	5.5	4.0
1978	21.0	14.0	11.5	4.0	-3.0
1979	17.0	18.0	4.0	1.0	-1.0
1980	21.0	15.0	9.0	4.0	3.0
1981	22.0	14.0	11.0	5.0	0.0
1982	22.0	15.0	8.0	5.0	1.0
1983	18.0	21.0	13.0	2.0	4.0
1984	22.0	18.0	9.0	2.0	2.0
1985	22.0	14.0	9.0	2.0	1.0
1986	22.0	14.0	10.0	4.0	1.0
1987	20.0	16.0	15.0	1.0	-2.0
1988	20.0	15.0	13.0	2.0	-1.0

B I B L I O G R A F I A

CAPITULO 2

1.- H. C. S. Thom, "The Dristributions of Freeze-Data and Freeze Free Period for Climatoological Series with Freezeless Years", Monthly Weather Review, April 1959.

2.- Benjamin Alberto Grassi Cantero, "Riesgo de Primeras y Ultimas Heladas en Puebla y Tlaxcala respecto a los Cultivos Básicos", Tesis de Maestría en Ciencias, 1983.

3.- W. G. Reed, "The Probable Growing Season", Monthly Weather Review, Sept. 1916.

4.- H. C. S. Thom and R. H. Shawn, "Climatological of Freeze Data for Iowa", Monthly Weather Review, Julio 1958.

5.- Erwin Kreyszig, "Introducción a la Estadística Matemática Principios y Métodos, LIMUSA.

6.- Enriqueta García "Modificaciones al Sistema de Clasificación Climática de Köppen", 1981.

CAPITULO 3

1.- George E. P. Box and Gwilym M. Jenkins "Time Series Analysis Forecasting and Control", Holden-Day, 1976.

2.- Bruce L. Bowerman and Richard T. O'Connell, "Unified Cooncepts and Computer Implementation", Second Edition, Duxbury Press, 1978.

3.- Victor Guerrero, "Modelos Estadísticos para Series de Tiempo Univariadas", V Coloquio del Depto. de Matemáticas CINESTAV IPN, 1987.