

01168
8
2ej.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

DIVISION DE ESTUDIOS DE POSGRADO
DE LA FACULTAD DE INGENIERIA

UN MODELO PARA PRONOSTICO DE LLUVIA DEL EDO. DE MORELOS
ESTUDIO COMPARATIVO DE METODOLOGIAS DE SERIES DE TIEMPO

SAUL F. VALDES CORNEJO

T E S I S

PRESENTADA A LA DIVISION DE ESTUDIOS DE POSGRADO DE LA FACULTAD DE
INGENIERIA DE LA UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO, COMO
REQUISITO PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERIA
(INVESTIGACION DE OPERACIONES)

TESIS CON
FALLA DE ORIGEN

CIUDAD UNIVERSITARIA
MAYO DE 1991



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 está 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.

PARTE I

DESCRIPCION DEL PROBLEMA

1.1 INTRODUCCION.

Es de gran importancia entender la influencia de los fenómenos meteorológicos sobre la variedad de las actividades del hombre; uno de los fenómenos meteorológicos más importantes es la lluvia, tanto para las regiones agrícolas como para sus zonas urbanas. Será posible tener un mayor conocimiento de sus posibilidades de producción o desarrollo, si ayudamos a conocer los alcances de un recurso básico, el agua. De esta forma, un estudio que relacione un análisis estadístico con el problema de pronóstico de lluvia, será una ayuda para poder llevar a cabo una planeación para un uso más eficiente del agua tanto en el campo como en las ciudades resulta vital.

El problema de pronóstico de lluvia podemos plantearlo tanto a corto plazo como a largo plazo. El pronóstico a corto plazo se relaciona con el problema de predecir lluvia para un día determinado basándose en las características de la lluvia o de otros fenómenos meteorológicos de los días que anteceden al día en cuestión. Este tipo de pronóstico está determinado en gran medida por las condiciones atmosféricas que prevalecen en la región en el momento de hacer el pronóstico.

En el pronóstico a largo plazo se pretende determinar la cantidad de lluvia que se espera tener en una región dada para un cierto periodo del año, bajo una medida de probabilidad. La mayoría de las veces éste pronóstico se hace para un mes o una semana

cualesquiera del año.

En el caso particular de México, hasta la fecha, las posibles soluciones que se han ofrecido sobre el pronóstico de lluvia a largo plazo no han resultado satisfactorias. Aunque la Secretaría de Agricultura y Recursos Hidráulicos (S.A.R.H.), la Comisión Federal de Electricidad (C.F.E.), y algunas otras dependencias gubernamentales poseen registros de datos de lluvias para un gran número de regiones del país, no se ha obtenido un gran beneficio de toda esta información, quizá debido a que no se ha aplicado una metodología apropiada que permita resolver el problema.

El propósito fundamental de este trabajo es ofrecer una solución al problema de pronóstico de precipitación pluvial a largo plazo presentando una metodología estadística que se adecue al tipo de Serie de Tiempo que presenta el problema.

Inicialmente debemos definir el periodo de pronóstico, ya que es necesario tener un registro de lluvia para ese periodo, entonces utilizando ese registro y bajo la condición de que un registro (mensual) de lluvia es una realización de un proceso estocástico, el problema se reduce a encontrar un modelo que explique el comportamiento de dicho proceso. Una vez hecho lo anterior es posible hacer inferencia sobre la cantidad de lluvia esperada en un determinado periodo.

1.2 Aspectos Generales Sobre el Estado de Morelos.

1.2.1. Regiones naturales.

El estado de Morelos cuenta con una gran diversidad de condiciones climáticas en su pequeño territorio, predominando el clima cálido, que se manifiesta en las zonas bajas de los ríos Amacuzac y Nexapa.

El Clima de tipo semicálido se presenta, en una franja que va de Este a Oeste situada en la zona de transición entre la sierra y los valles. El templado se distribuye en las partes más altas de los valles de Cuernavaca y Cuautla. Los climas semifrios se localizan en pequeñas áreas de las partes más altas de la sierra en el extremo norte, como son la Cordillera Neovolcánica y la Sierra Nevada.

1.2.2. Hidrología y Balance del Agua.

El estado queda comprendido en la región hidrológica "Río Balsas" (NO. 18) con una superficie de 4 958.22 Km², estando dividido en las siguientes cuencas:

- Cuenca del Río Atoyac, con una superficie dentro del estado de 653.17 km²; es uno de los principales formadores del río Balsas. Se origina con los deshielos que descienden del franco oriental del Volcán Iztaccíhuatl, la aportación al estado de Morelos es mínima, ya que en la entidad sólo son escurrimientos que drenan hacia la corriente principal (Río Atoyac).
- Cuenca del Río Balsas-Mezcala, tiene una superficie dentro del estado de 1.66 km². El río Balsas corriente principal de esta cuenca, recibe en su largo recorrido varios nombres: Zacatula, Atoyac y Mezcala. El aprovechamiento de este Río dentro del estado de Morelos es casi nulo.

- Cuenca del Río Grande Amacuzac, tiene superficie dentro del estado de 4,303.39 km² esta cuenca es la que ocupa la mayor parte del territorio y es uno de los principales afluentes del río Balsas y se origina en las faldas del Nevado de Toluca.

Los principales almacenamientos del estado, su capacidad y uso:

NOMBRE	MUNICIPIO	Cap. Mill. m	USO
Laguna el Rodeo	Miacatlan	28.00	riego.
Emiliano Zapata	Puente de Ixtla	6.00	riego.
La Poza	Tepalcingo	1.45	riego.
Las Teclas	Tepalcingo	0.42	riego.
La Unión	Villa de Ayala	0.30	riego.
Cerro de la Era	Zacualpan	0.30	riego.
El Venado	Jantetelco	0.27	riego.
El Tecolote	Jonacatepec	0.20	riego.
La Palapa	Villa de Ayala	0.13	riego.
Los Lavaderos	Jantetelco	0.12	riego.
Palo Prieto	Tepalcingo	0.10	riego.
El Ciruelo	Jantetelco	0.10	riego.
Sinón Cárdenas	Villa de Ayala	0.10	riego.
Los Tepetates	Axochiapan	0.90	riego.
Coyotamate	Jonacatepec	0.80	riego.
La Esperanza	Jantetelco	0.40	riego.

Las áreas dedicadas a la agricultura de riego se distribuyen al centro y en menor proporción al este y oeste de la entidad. La práctica de la agricultura de temporal es la más usual

en el estado; dentro del sector primario, la agricultura sigue siendo la actividad más dinámica, destacando el cultivo de la caña de azucar. En términos generales el valor de la producción agrícola ascedió en 1986 a 32.7 mil millones de pesos, de los cuales la caña de azucar aportó el 38.1%, maíz 13.8%, jitomate 13.3% y cebolla 11.4%, que en conjunto aportaron el 76.6% del valor total de la producción.

La superficie cosechada en 1985 fué de:

47 990.1 has. de Riego.

93 545.0 has. de Temporal. Total 141 535.1 has.

La superficie total irrigada por sistema de riego:

44 619.4 has. por gravedad.

7 258.7 has. por bombeo de pozo.

2 551.6 has. bombeo por corriente.

El problema de pronóstico de lluvia para el estado de Morelos se estudia de acuerdo a la información recopilada para un total de catorce (14) estaciones climatológicas ubicadas en variados lugares del estado, ver tabla 1.1.

Las estaciones seleccionadas poseen información pluviométrica que comprende el periodo que va desde Enero de 1937 hasta Diciembre de 1989.

El total de las catorce estaciones con información pluviométrica fué clasificado de acuerdo a regiones de igual precipitación (isoyetas) obtenidas del mapa de la Región México Carta de Precipitación Total Anual 1:1.000,000 (segunda impresión,

TABLA 1.1

ESTACIONES CLIMATOLÓGICAS DEL ESTADO
DE MORELOS

MOR1	CUAUTLA
MOR2	CUERNAVACA
MOR3	EL RODEO
MOR4	HUAJINTLAN
MOR5	TEMILPA
MOR6	TEMIXCO
MOR7	TEPALcingo
MOR8	TILZAPOTLA
MOR9	TLACOTEPEC
MOR10	TLAYACAPAN
MOR11	TOTOLAPAN
MOR12	TRES CUMBRES
MOR13	YAUTEPEC
MOR14	ZACATEPEC

1984) que publica la Secretaría de Programación y Presupuesto a través del Instituto Nacional de Estadística Geografía e Informática, ver tabla 1.2.

La isoyeta de 800 a 1000 mm. será denotada ISO810, la isoyeta de 1000 a 1200 mm., ISO1012, y la de 1500 a 2000 mm., ISO1520, se enlistan en las Tablas 1.3, 1.4 y 1.5 respectivamente.

De ésta manera y de acuerdo a información que publica el Instituto Nacional de Estadística Geografía e Informática, en la Serie "Cuadernos para la Planeación Estatal" Edición 1988, se encuentran las siguientes estadísticas: la mayor actividad agrícola se concentra en ISO810, alrededor del 90% de la superficie total de riego. La ISO1012 sólo ocupa el 4% de la superficie irrigada, la ISO1520 ocupa una sólo una pequeña fracción tanto de extensión como de concentración de la población y de la actividad económica de alrededor del 2%. Sin embargo aproximadamente el 33% de la superficie cultivada de temporal se encuentra dentro de ISO1012 y un 65% en ISO810. or lo que a la distribución de la población se refiere, alrededor del 35% se localiza dentro de la ISO810 y un 60% en ISO1012. En cuanto a la actividad industrial en ISO1012 se concentra el 40% de ella mientras que en ISO1012 alrededor del 55%.

TABLA 1.2

ESTACIONES CLIMATOLÓGICAS DEL ESTADO
DE MORELOS

CLASIFICACION POR ISOYETAS

ISOYETA DE 800 A 1000 MM

MOR1	CUAUTLA
MOR3	EL RODEO
MOR4	HUAJINTLAN
MOR5	TEMILPA
MOR6	TE MIXCO
MOR7	TEPALCINGO
MOR8	TILZAPOTLA
MOR9	TLACOTEPEC
MOR13	YAUTEPEC
MOR14	ZACATEPEC

ISOYETA DE 1000 A 1200 MM

MOR2	CUERNAVACA
MOR10	TLAYACAPAN
MOR11	TOTOLAPAN

ISOYETA DE 1500 A 2000 MM.

MOR12	TRES CUMBRES
-------	--------------

Tabla 1.4

PROMEDIO DE PRECIPITACION PLUVIA MENSUAL ACUMULADO
ISEVETA 1000 - 1200 mm.

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	ANO
5.77	5.10	9.43	32.10	88.57	176.03	219.17	184.17	239.67	104.70	2.83	4.87	1937
1.23	12.70	45.37	11.87	45.07	209.43	262.27	175.93	149.27	50.53	25.40	3.73	
.50	8.13	7.37	25.00	60.67	230.17	195.93	196.90	128.33	170.37	8.13	8.70	
5.06	6.63	33.17	.00	43.23	245.37	297.90	219.73	156.63	22.17	17.17	5.23	
3.07	5.50	6.70	5.13	31.97	272.20	131.47	175.23	211.97	149.63	15.60	15.17	
5.70	6.77	4.80	.60	29.17	207.17	187.90	205.73	340.03	33.63	30.50	.00	
5.10	.60	7.27	4.57	39.10	227.63	132.47	219.70	337.93	70.63	5.10	15.23	
9.67	.00	20.63	21.63	34.00	191.73	182.07	253.00	230.37	81.63	10.57	.00	
1.07	.07	.00	15.20	87.93	140.87	115.93	289.67	213.30	86.83	7.03	9.50	
8.27	.00	5.27	74.53	39.53	217.60	227.10	166.40	134.60	93.67	29.10	.00	
12.13	.79	.00	13.43	103.27	193.83	200.53	222.63	207.93	141.90	.00	17.83	
26.62	.07	.57	21.27	119.20	165.70	221.42	178.57	205.23	81.30	76.57	.00	
.00	.00	4.47	4.70	47.37	193.57	99.73	128.53	158.50	70.43	.00	.00	
5.90	6.47	3.93	16.40	90.17	185.57	221.50	214.57	244.83	62.67	13.10	.00	
.00	.60	22.17	5.40	44.27	181.67	236.20	231.03	261.47	73.63	12.23	.00	
15.77	9.40	4.60	35.40	109.27	235.80	208.57	218.30	264.63	59.10	44.87	.00	
1.67	.00	3.20	.80	16.90	169.10	253.97	252.67	212.92	117.37	17.67	3.93	
.07	14.57	4.57	52.43	92.77	226.03	207.53	190.67	321.03	96.93	11.00	.17	
.57	3.63	3.57	7.73	23.63	228.00	273.00	243.87	334.10	106.57	8.00	15.53	
41.00	22.00	11.50	34.77	103.47	258.27	221.10	166.17	276.83	141.80	74.10	26.77	
3.43	5.60	.00	24.93	44.90	222.07	155.07	228.27	213.60	107.03	7.00	1.67	
111.13	20.57	.00	6.53	72.70	342.30	427.60	250.53	339.40	226.43	158.67	15.83	
26.97	4.17	.00	56.97	26.23	295.47	207.70	167.03	147.50	228.67	13.73	.00	
21.70	.00	.00	5.87	17.57	140.30	225.20	175.47	179.17	100.07	3.00	.00	
.00	2.33	3.33	11.23	34.73	219.10	290.33	204.50	250.23	11.27	20.73	.00	
.00	.00	.00	21.53	33.20	242.03	147.83	165.63	296.42	64.60	14.70	6.50	
.00	.00	14.53	.00	57.67	153.67	236.80	117.40	176.50	84.17	2.57	1.73	
15.00	.00	21.23	4.00	171.77	221.10	141.50	101.17	243.50	41.00	.00	10.17	
2.10	24.10	.00	30.83	74.53	173.20	130.73	163.60	161.00	195.70	21.23	30.83	
2.33	1.53	24.60	61.87	22.90	86.40	206.43	309.53	204.30	95.90	.00	21.10	
71.20	1.70	3.47	4.37	45.70	234.20	182.17	247.27	211.60	134.30	10.17	25.93	
5.50	22.50	5.80	42.63	21.90	215.83	215.17	133.67	167.30	94.00	1.00	44.30	
.00	.00	22.00	20.00	15.97	91.50	292.97	323.67	323.70	33.00	1.00	.00	
.00	.00	.00	1.67	17.50	218.47	270.53	174.77	324.60	40.47	17.83	.00	
.00	.00	.00	22.53	41.60	273.40	146.10	247.60	277.67	197.20	5.60	5.67	
.00	.00	26.83	43.83	84.83	285.10	287.63	101.67	218.50	67.27	35.87	.00	
.00	5.67	21.23	40.83	56.67	232.47	220.97	301.50	323.63	158.73	13.20	1.10	
.00	6.00	22.00	21.80	49.17	200.00	214.60	144.00	190.33	61.00	4.00	.00	
.00	.00	.00	19.73	19.73	19.73	19.73	19.73	19.73	19.73	19.73	19.73	
.00	.00	.00	31.13	31.13	31.13	31.13	31.13	31.13	31.13	31.13	31.13	
.00	.00	.00	73.73	73.73	73.73	73.73	73.73	73.73	73.73	73.73	73.73	
.00	21.17	21.17	41.77	41.67	139.27	185.70	230.67	230.63	100.47	23.00	17.17	
.00	18.20	18.20	41.20	41.20	137.20	183.50	228.00	228.00	71.00	18.20	13.20	
.00	.00	.00	73.73	73.73	73.73	73.73	73.73	73.73	73.73	73.73	73.73	
.00	28.10	28.10	41.00	41.00	139.00	185.00	230.00	230.00	61.00	28.10	13.00	
.00	22.00	22.00	38.00	38.00	138.00	184.00	229.00	229.00	60.00	22.00	12.00	
.00	13.00	13.00	31.00	31.00	136.00	182.00	227.00	227.00	59.00	13.00	11.00	
.00	11.00	11.00	28.00	28.00	135.00	181.00	226.00	226.00	58.00	11.00	10.00	
.00	10.00	10.00	27.00	27.00	134.00	180.00	225.00	225.00	57.00	10.00	9.00	
.00	9.00	9.00	26.00	26.00	133.00	179.00	224.00	224.00	56.00	9.00	8.00	
.00	8.00	8.00	25.00	25.00	132.00	178.00	223.00	223.00	55.00	8.00	7.00	
.00	7.00	7.00	24.00	24.00	131.00	177.00	222.00	222.00	54.00	7.00	6.00	
.00	6.00	6.00	23.00	23.00	130.00	176.00	221.00	221.00	53.00	6.00	5.00	
.00	5.00	5.00	22.00	22.00	129.00	175.00	220.00	220.00	52.00	5.00	4.00	
.00	4.00	4.00	21.00	21.00	128.00	174.00	219.00	219.00	51.00	4.00	3.00	
.00	3.00	3.00	20.00	20.00	127.00	173.00	218.00	218.00	50.00	3.00	2.00	
.00	2.00	2.00	19.00	19.00	126.00	172.00	217.00	217.00	49.00	2.00	1.00	
.00	1.00	1.00	18.00	18.00	125.00	171.00	216.00	216.00	48.00	1.00	.00	
.00	.00	.00	17.00	17.00	124.00	170.00	215.00	215.00	47.00	.00	.00	
.00	.00	.00	16.00	16.00	123.00	169.00	214.00	214.00	46.00	.00	.00	
.00	.00	.00	15.00	15.00	122.00	168.00	213.00	213.00	45.00	.00	.00	
.00	.00	.00	14.00	14.00	121.00	167.00	212.00	212.00	44.00	.00	.00	
.00	.00	.00	13.00	13.00	120.00	166.00	211.00	211.00	43.00	.00	.00	
.00	.00	.00	12.00	12.00	119.00	165.00	210.00	210.00	42.00	.00	.00	
.00	.00	.00	11.00	11.00	118.00	164.00	209.00	209.00	41.00	.00	.00	
.00	.00	.00	10.00	10.00	117.00	163.00	208.00	208.00	40.00	.00	.00	
.00	.00	.00	9.00	9.00	116.00	162.00	207.00	207.00	39.00	.00	.00	
.00	.00	.00	8.00	8.00	115.00	161.00	206.00	206.00	38.00	.00	.00	
.00	.00	.00	7.00	7.00	114.00	160.00	205.00	205.00	37.00	.00	.00	
.00	.00	.00	6.00	6.00	113.00	159.00	204.00	204.00	36.00	.00	.00	
.00	.00	.00	5.00	5.00	112.00	158.00	203.00	203.00	35.00	.00	.00	
.00	.00	.00	4.00	4.00	111.00	157.00	202.00	202.00	34.00	.00	.00	
.00	.00	.00	3.00	3.00	110.00	156.00	201.00	201.00	33.00	.00	.00	
.00	.00	.00	2.00	2.00	109.00	155.00	200.00	200.00	32.00	.00	.00	
.00	.00	.00	1.00	1.00	108.00	154.00	199.00	199.00	31.00	.00	.00	
.00	.00	.00	.00	.00	107.00	153.00	198.00	198.00	30.00	.00	.00	
.00	.00	.00	.00	.00	106.00	152.00	197.00	197.00	29.00	.00	.00	
.00	.00	.00	.00	.00	105.00	151.00	196.00	196.00	28.00	.00	.00	
.00	.00	.00	.00	.00	104.00	150.00	195.00	195.00	27.00	.00	.00	
.00	.00	.00	.00	.00	103.00	149.00	194.00	194.00	26.00	.00	.00	
.00	.00	.00	.00	.00	102.00	148.00	193.00	193.00	25.00	.00	.00	
.00	.00	.00	.00	.00	101.00	147.00	192.00	192.00	24.00	.00	.00	
.00	.00	.00	.00	.00	100.00	146.00	191.00	191.00	23.00	.00	.00	
.00	.00	.00	.00	.00	99.00	145.00	190.00	190.00	22.00	.00	.00	
.00	.00	.00	.00	.00	98.00	144.00	189.00	189.00	21.00	.00	.00	
.00	.00	.00	.00	.00	97.00	143.00	188.00	188.00	20.00	.00	.00	
.00	.00	.00	.00	.00	96.00	142.00	187.00	187.00	19.00	.00	.00	
.00	.00	.00	.00	.00	95.00	141.00	186.00	186.00	18.00	.00	.00	
.00	.00	.00	.00	.00	94.00	140.00	185.00	185.00	17.00	.00	.00	
.00	.00	.00	.00	.00	93.00	139.00	184.00	184.00	16.00	.00	.00	
.00	.00	.00	.00	.00	92.00	138.00	183.00	183.00	15.00	.00	.00	
.00	.00	.00	.00	.00	91.00	137.00	182.00	182.00	14.00	.00	.00	
.00	.00	.00	.00	.00	90.00	136.00	181.00	181.00	13.00	.00	.00	
.00	.00	.00	.00	.00	89.00	135.00	180.00	180.00	12.00	.00	.00	
.00	.00	.00	.00	.00	88.00	134.00	179.00	179.00	11.00	.00	.00	
.00	.00	.00	.00	.00	87.00	133.00	178.00	178.00	10.00	.00	.00	
.00	.00	.00	.00	.00	86.00	132.00	177.00	177.00	9.00	.00	.00	
.00	.00	.00	.00	.00	85.00	131.00	176.00	176.00	8.00	.00	.00	
.00	.00	.00	.00	.00	84.00	130.00	175.00	175.00	7.00	.00	.00	
.00	.00	.00	.00	.00	83.00	129.00	174.00	174.00	6.00	.00	.00	
.00	.00	.00</										

PARTE II

DESCRIPCION DE METODOLOGIAS DE SERIES DE TIEMPO.

2.1 PROMEDIOS MOVILES.

Dado un conjunto de N observaciones, se toman las primeras T observaciones como un conjunto inicializador y el resto de $T+1$ hasta N como un conjunto prueba. El método de promedios móviles de orden T , o $MA(T)$ es el siguiente: tomar el promedio de todos los datos en el conjunto inicialización como el pronóstico para el periodo $(T+1)$.

$$\bar{X} = \sum_{i=1}^T X_i / T = F_{T+1}$$

(Entonces como la observación del periodo $(T+1)$ esta disponible, es posible calcular el error:)

$$e_{T+1} = X_{T+1} - F_{T+1}$$

ahora para el periodo $(T+2)$ la situación es: el conjunto inicializador será de X_1, \dots, X_T, X_{T+1} y el conjunto prueba será: X_{T+2}, \dots, X_N . Es decir para el periodo $(T+1)$ se tiene que:

Promedio Móvil

$$\bar{X} = \frac{X_2 + \dots + X_{T+1}}{T}$$

\vdots

$$\bar{X}_1 = \frac{X_{1+1} + \dots + X_{T+1}}{T}$$

Pronóstico

$$F_{T+2} = \bar{X} = \sum_{i=2}^{T+1} X_i / T$$

\vdots

$$F_{T+1} = \bar{X}_1 = \sum_{j=1+1}^{T+1} X_j / T$$

Se observa que el promedio móvil de orden T tiene las siguientes características:

- esta relacionado solamente con los últimos T periodos de datos conocidos.
- el número de observaciones en cada periodo no cambia cuando el tiempo avanza.

Pero tiene las siguientes desventajas:

- todas las últimas T observaciones deben ser almacenadas, no solamente el promedio.
- no puede manejarse tendencia o estacionalidad adecuadamente.

(Por ejemplo, el MA(12), para datos mensuales, suavizaría efectos estacionales de las series y sería útil en la descomposición de la serie, pero no sería efectivo sobre sí mismo como herramienta de pronóstico en una serie que ya muestre tendencia y estacionalidad.

2.2 SUAVIZAMIENTO EXPONENCIAL

2.2.1 Suavizamiento exponencial Simple

El método de suavizamiento exponencial se caracteriza por asignar mayor importancia a las observaciones más recientes, las cuales tendrán mayor influencia en el pronóstico. El metodo puede obtenerse de una variación del método de promedios móviles.

Considere que se cuenta con N observaciones, y se selecciona un MA(N).

Por lo que se tiene:

$$F_t = \frac{X_{t-N} + \dots + X_{t-1}}{N}$$

$$F_{t+1} = \frac{X_{t-N+1} + \dots + X_{t-1}}{N} \quad 2.1$$

Así que el caso más simple de suavizamiento exponencial puede desarrollarse de la ecuación 2.1, o más específicamente de una variación de esa ecuación,

$$F_{t+1} = F_t \left(\frac{X_t}{N} - \frac{X_{t-N}}{N} \right) \quad 2.2$$

Supongamos que no se cuenta con la observación más vieja X_{t-N} , y usamos un valor aproximado en su lugar. Puede ser reemplazada por el pronóstico del periodo previo F_t . Sustituyendo en la ecuación 2.1, llegamos a:

$$F_{t+1} = F_t + \left(\frac{X_t}{N} - \frac{F_t}{N} \right) \quad 2.3$$

o bien

$$F_{t+1} = \frac{1}{N} X_t + \left(1 - \frac{1}{N} \right) F_t \quad 2.4$$

donde observando que N es número positivo, $1/N$ tenderá a una constante entre cero, si N es muy grande, y uno, si $N = 1$.

De la ecuación 2.4 se puede ver que este pronóstico F_{t+1} está basado sobre el peso de las observaciones más recientes con un valor de peso $1/N$ y el pronóstico previo, al más reciente, con un peso de $[1 - (1/N)]$. Observando que N es número positivo, $1/N$ tenderá a una constante entre cero, si $N \rightarrow \infty$, y uno, si $N = 1$. Sustituyendo $1/N$ por α la ecuación 2.4 queda:

$$F_{t+1} = \alpha X_t + (1 - \alpha) F_t \quad 2.5$$

La ecuación 2.5 representa la forma más general para el cálculo de pronósticos, con el método de suavizamiento exponencial. Nótese, de 2.5 que F_{t+1} puede escribirse como:

$$F_{t+1} = \alpha X_t + \alpha(1-\alpha) X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \dots + \alpha(1-\alpha)^{N-1} X_{t-(N-1)} + \alpha(1-\alpha)^N F_{t-N}. \quad 2.6$$

2.2.2 SUAVIZAMIENTO EXPONENCIAL SIMPLE CON APROXIMACION ADAPTIVA LINEAL.

Al igual que el método anterior, este método elimina errores cuando se detecta una tendencia polinomial lineal en la serie observada. La ecuación básica para pronóstico con este método es similar a la ecuación 2.5 excepto que α es reemplazada por α_t :

$$F_{t+1} = \alpha_t X_t + (1 - \alpha_t) F_t \quad 2.7$$

donde

$$\alpha_{t+1} = \left| \frac{E_t}{P_t} \right| \quad 2.8$$

$$\text{con } E_t = \beta e_t + (1 - \beta) E_{t-1} \quad 2.9$$

$$P_t = \beta |e_t| + (1 - \beta) P_{t-1} \quad 2.10$$

$$e_t = X_t - F_t \quad 2.11$$

donde α y β son parámetros entre cero y uno.

El reemplazar α por α_t permite detectar los cambios provocados por la tendencia presente en la serie de tiempo.

La ecuación 2.8 indica que el valor de α a usarse en el periodo de pronóstico ($t+2$) esta definido como el valor absoluto de la razón de un término de error suavizado (E_t) y un término de error absoluto suavizado (P_t). Esos dos términos suavizados se obtienen unando suavizamiento exponencial simple, como se muestra de las ecuaciones 2.9 y 2.10.

2.2.3 DOBLE SUAVIZAMIENTO EXPONENCIAL:

METODO LINEAL DE BROWN.

El método aplica un doble suavizamiento exponencial a la serie observada. De la ecuación 2.5 se tiene que:

$$S'_t = \alpha X_t + (1 - \alpha) S'_{t-1}. \quad 2.12$$

$$S''_t = \alpha S'_t + (1 - \alpha) S''_{t-1}. \quad 2.13$$

donde S'_t es el valor del suavizamiento exponencial simple y S''_t es el valor del doble suavizamiento exponencial. En seguida se tienen los parámetros:

$$a_t = S'_t + (S'_t - S''_t) = 2S'_t - S''_t. \quad 2.14$$

$$b_t = \frac{\alpha}{1 - \alpha} (S'_t - S''_{t-1}). \quad 2.15$$

y el pronóstico viene dado por: $F_{t+m} = a_t + b_t m, \quad 2.16$

donde m es el número de predicciones deseadas.

2.2.4 TRIPLE SUAVIZAMIENTO EXPONENCIAL:

METODO CUADRATICO DE BROWN

Sólo se usa para pronosticar datos cuando en la serie observada se detecta una tendencia polinomial de orden dos. Para ir del suavizamiento lineal al cuadrático, la aproximación básica es incorporar un nivel adicional de suavizamiento y trabajar con una ecuación de pronóstico cuadrático. Las ecuaciones de suavizamiento cuadrático son:

$$S'_t = \alpha X_t + (1 - \alpha) S'_{t-1} \quad (\text{primer suavizamiento}). \quad 2.17$$

$$S''_t = \alpha S'_t + (1 - \alpha) S''_{t-1} \quad (\text{segundo suavizamiento}). \quad 2.18$$

$$S'''_t = \alpha S''_t + (1 - \alpha) S'''_{t-1} \quad (\text{tercer suavizamiento}). \quad 2.19$$

el pronóstico viene dado por:

$$F_{t+m} = a_t + b_t m + 1/2 c_t m^2. \quad 2.20$$

con $a_t = 3S'_t - 3S''_t + S'''_t. \quad 2.21$

$$b_t = \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S'_t - (10-8\alpha)S''_t - (4-3\alpha)S'''_t]. \quad 2.22$$

$$c_t = \frac{\alpha^2}{(1-\alpha)^2} (S'_t - 2S''_t + S'''_t). \quad 2.23$$

y

$$F_{t+m} = a_t + b_t m + 1/2 (c_t m^2). \quad 2.24$$

2.2.5 TRIPLE SUAVIZAMIENTO EXPONENCIAL: METODO DE WINTER

Los métodos de promedios móviles, y de suavizamiento exponencial examinados hasta ahora pueden abordar casi cualquier tipo

de datos estacionarios y no-estacionarios siempre y cuando los datos sean no-estacionales. Cuando la estacionalidad exista, sin embargo, éstos métodos podrían acarrear a un pobre trabajo de pronóstico.

El método de Winter trata de explicar comportamientos del tipo estacional y tendencia lineal, de tal manera que se expliquen dichas componentes que se presentan en la serie observada.

Si los datos son estacionarios, entonces los promedios móviles o los métodos de suavizamiento exponencial simple son apropiados. Si los datos exhiben una tendencia lineal, tanto el modelo lineal de Brown, como de Holt son apropiados. Pero si los datos son estacionales, esos métodos no pueden manejar bien el problema. El método de Winters, sin embargo, puede manejar directamente la estacionalidad.

El método de Winters, esta basado en tres ecuaciones de suavizamiento -una para la estacionariedad, una para la tendencia y una para la estacionalidad. El método esta basado en las siguientes ecuaciones.

sobre suavizamiento:

$$S_t = \alpha \frac{X_t}{I_{t-1}} + (1-\alpha) (S_{t-1} + b_{t-1}). \quad 2.25$$

suavizamiento de tendencia:

$$b_t = \gamma (S_t - S_{t-1}) + (1-\gamma) b_{t-1}. \quad 2.26$$

suavizamiento de estacionalidad:

$$I_t = \beta \frac{X_t}{S_t} + (1 - \beta) I_{t-1}. \quad 2.27$$

pronóstico:

$$F_{t+m} = (S_t + b_t m) I_{t-L+m}$$

2.28

donde L es la longitud de la estacionalidad presente en la serie observada, b es la componente de tendencia, I es el factor de ajuste estacional y F_{t+m} es el pronóstico para m periodos hacia adelante.

2.3 MODELOS DE DESCOMPOSICION DE SERIE

Con la metodología de descomposición de series se pretende identificar y estimar a cada una de las componentes de la serie de tiempo por separado. La representación general de la descomposición de la serie es de la forma:

$$X_t = f(S_t, T_t, I_t)$$

donde

X_t es la serie de tiempo observada al tiempo t .

S_t es la componente estacional al tiempo t .

T_t es la componente de tendencia al tiempo t .

I_t es la componente estocástica al tiempo t .

En los métodos de descomposición se identifica y separa de la serie de tiempo cada una de las componentes presentes, con el objetivo de realizar proyecciones sobre cada una de éstas.

Hoy en día existen diversas metodologías desarrolladas con el propósito de realizar la parte del ajuste estacional de la serie; en los métodos basados en modelos, a cada componente de la serie se le pretende asociar su correspondiente modelo de comportamiento estocástico (metodología Box-Jenkins). De hecho el modelo básico, que muestra explícitamente la relación que guardan las componentes de una serie observada es:

$$X_t = F(S_t, T_t, I_t) = S_t + T_t + I_t.$$

2.29

A partir del modelo básico, el problema para el ajuste estacional se convierte en el de estimación de las componentes T_t , S_t e I_t para cada periodo $t = 1, 2, 3, \dots, N$.

Aunque la forma funcional exacta de 2.29 depende del método de descomposición realmente usado.

Cuando se usan métodos de regresión se aplican generalmente en los supuestos de que la estacionalidad, y en ocasiones la tendencia, pueden representarse de manera determinista mediante funciones de tiempo; para ello es común representar a la tendencia mediante una curva polinomial y a la estacionalidad mediante funciones periódicas (combinaciones de senos y coseno). Cuando se usan métodos de promedios móviles, éstos presuponen que tanto la tendencia como la estacionalidad, siguen comportamientos dinámicos en el tiempo; y en consecuencia la estimación de las componentes se usa localmente, de forma que la tendencia para un cierto tiempo t_0 se estima como un promedio de observaciones previas y futuras al momento de observación. En la práctica se prefiere generalmente trabajar con los métodos de promedios móviles; o bien, con métodos basados en modelos (filtro Kalman).

Para todos los métodos de descomposición el proceso es similar y consiste de los siguientes pasos. tales pasos constituyen el conjunto de reglas empíricas de lo que debe hacerse en el ajuste estacional de una serie (mensual, en este caso).

- 1) De la serie original, calcule un promedio móvil de 12 términos centrado, para estimar inicialmente la componente de tendencia.
- 2) Reste de la serie observada el promedio móvil de 1) para obtener como diferencia la componente de estacionalidad-irregularidad.
- 3) Extrapole seis valores al principio y al final de la serie de estacionalidad-irregularidad, para sustituir los perdidos en 1).

- 4) Calcule un promedio móvil de 3 términos de cada uno de los meses de la componente de estacionalidad-irregularidad, para estimar inicialmente a la componente estacional.
- 5) Ajuste las desviaciones estacionales para que su suma sea cero; para ello reste de las desviaciones estacionales su promedio móvil de 12 términos centrado.
- 6) Reste de la componente de estacionalidad-irregularidad obtenidas en 3) la componente obtenida en 5) para estimar a la componente irregular.
- 7) Calcule la desviación estándar móvil de 5 términos de la componente irregular obtenida en 6) para detectar valores extremos; reemplace valores extremos.
- 8) Repita el paso 4) con la serie obtenida en 7) para obtener otra estimación de la componente estacional.
- 9) Repita 5) con la serie de 8).
- 10) Repita 6) con las series obtenidas en 7) y 9).
- 11) Reste de la serie original la componente estacional obtenida en 9) para obtener una primera estimación de la serie ajustada por estacionalidad.
- 12) Calcule un promedio móvil Henderson (véase Kenny Durvin, 1882) a la serie ajustada por estacionalidad para obtener otra estimación de la componente de tendencia.
- 13) Reste de la serie original la componente de tendencia calculada en 12) para otra estimación de la componente estacionalidad irregularidad.
- 14) Repita 7) con la componente de tendencia calculada en 10).
- 15) Repita el paso 4) con la serie obtenida en 14), pero ahora calcule un promedio móvil de 7 términos.
- 16) Repita el paso 5) con la serie obtenida en 15).
- 17) Repita el paso 11) usando como estimación de la componente estacional la serie 16).
- 18) Reste de la serie ajustada estacionalmente obtenida en 17), la componente de tendencia obtenida en 12), para obtener la estimación de la componente irregular.

Las dificultades computacionales en que se incurre al usar métodos de descomposición son realmente considerables, en consecuencia es necesario un método más adecuado y computacionalmente eficiente.

2.4. EL METODO DE BOX-JENKINS

La metodología propuesta por Box-Jenkins (1976) para determinar los valores de p y q de un modelo ARMA, consiste básicamente en el reconocimiento de patrones de las FAC teóricas, a partir, del comportamiento observado en las FAC muestrales.

2.4.1 FUNCION DE AUTOCORRELACION DE UN MODELO AR.

El término Autorregresivo que se le da al proceso representado por se refiere al hecho de que también puede expresarse como:

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

la cual es básicamente una ecuación de regresión lineal, con la característica especial de que el valor de la variable dependiente Z en el período t no depende de los valores de un cierto conjunto de variables independientes, como sucede en el modelo de regresión, si no de sus propios valores observados en los períodos anteriores a t y ponderados de acuerdo a los coeficientes autorregresivos $\phi_1, \phi_2, \dots, \phi_p$.

Las propiedades generales de este modelo serán indicadas a partir del análisis de los casos particulares AR(1) y AR(2).

i) Modelo AR(1)

En este caso se tiene

$$Z_t - \phi Z_{t-1} = a_t \quad 2.30$$

Como se sabe para que esta serie sea estacionaria se requiere que la raíz de la ecuación

$$1 - \phi x = 0 \quad 2.31.$$

se encuentre fuera del círculo unitario, esto es que, $|\phi| < 1$ o equivalentemente que $U_d(B) = 1$ y por lo tanto 2.31 toma la forma

$$Z_t - \phi Z_{t-1} = a_t \quad 2.32$$

Otra posible representación del proceso autorregresivo estacionario 2.32, en términos de la serie de errores aleatorios exclusivamente, es

$$Z_t = \frac{1}{1 - \phi B} a_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots \quad 2.33$$

de donde se observa que si $|\phi| < 1$

$$E[Z_t] = E[a_t + \phi a_{t-1} + \phi a_{t-2} + \dots] = 0. \quad 2.34$$

$$\gamma_0 = \sigma_a^2 (1 + \phi^2 + \phi^4 + \dots) = \frac{\sigma_a^2}{1 - \phi^2}.$$

y

$$\gamma_m = \frac{\sigma_a^2 \phi^{|m|}}{1 - \phi^2} = \phi \gamma_{m-1} \quad ; \quad m = 1, 2, \dots \quad 2.35$$

y debido a que $\gamma_k = \gamma_{-k}$, se obtiene la fórmula general

$$\gamma_m = \frac{\sigma_a^2 \phi^{|m|}}{1 - \phi^2} \quad ; \quad m \in \mathbb{Z} \quad 2.36$$

de donde se obtiene que la autocorrelación

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \phi^{|m|} \quad 2.37$$

lo cual indica que, conforme m crece, la función de autocorrelación tiende a cero.

ii) Modelo AR(2).

El siguiente esquema autorregresivo a considerar será el autorregresivo de segundo orden, esto es, se pasa ahora a estudiar el proceso AR(2) definido mediante

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = a_t \quad 2.38$$

Para que el proceso 2.38 sea estacionario, debe cumplirse que las raíces del polinomio

$$1 - \phi_1 X - \phi_2 X^2 = 0.$$

se encuentren fuera del círculo unitario o equivalentemente que $U_d(B) = 1$ y por lo tanto 2.38 toma la forma:

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = a_t \quad 2.39$$

De acuerdo con el teorema de Schur (véase Chipman 1951, pag 119 - 120) las condiciones de estacionariedad estan dadas por

$$|\phi_2| < 1, \quad \phi_2 + \phi_1 < 1 \quad \text{Y} \quad \phi_2 - \phi_1 < 1 \quad 2.40$$

Si el proceso AR(2) en estudio resulta ser estacionario, lo único que hace falta para describirlo completamente es obtener su funcion de autocovarianza, lo cual se logra como sigue: si $E[Z_t] = 0$ se tiene que

$$\begin{aligned} \gamma_m &= E[Z_t Z_{t+m}] = \phi_1 E[Z_t Z_{t+m-1}] + \phi_2 E[Z_t Z_{t+m-2}] + E[Z_t a_{t+m}] \\ &= \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2 & \text{si } m = 0. \\ \phi_1 \gamma_{m-1} + \phi_2 \gamma_{m-2} & \text{si } m > 0. \end{cases} \quad 2.41 \end{aligned}$$

luego entonces, si $\gamma_0 < \infty$, se pueden dividir las autocovarianzas γ_1 y γ_2 entre γ_0 y obtener las llamadas ecuaciones de YULE-WALKER

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad 2.42$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

que permiten obtener los valores de ρ_1 y ρ_2 en función de los parámetros ϕ_1 y ϕ_2 .

En general la FAC satisface

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} \quad ; \quad m \geq 3$$

iii) Modelo AR(p)

Como caso general de un proceso autorregresivo, se puede ahora considerar el proceso AR(p) que se describe mediante la ecuación (8), la cual es equivalente a

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad 2.43$$

Un proceso AR(p) será estacionario si y sólo si las raíces de la ecuación característica

$$1 - \phi_1 X - \phi_2 X^2 - \dots - \phi_p X^p = 0.$$

se encuentran fuera del círculo unitario, esto es, $U(B) = 1$ y por lo tanto 2.43 toma la forma

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \dots - \phi_p Z_{t-p} = a_t \quad 2.44$$

De acuerdo con el teorema de Schur las condiciones de estacionariedad surgen del requerimiento de que los p determinantes que se muestran a continuación, sean positivos

$$D_1 = \begin{vmatrix} -1 & \phi_p \\ \phi_p & -1 \end{vmatrix} \quad , \quad D_2 = \begin{vmatrix} -1 & 0 & \phi_p & \phi_{p-1} \\ \phi_1 & -1 & 0 & \phi_p \\ \phi_p & 0 & -1 & \phi_1 \\ \phi_{p-1} & \phi_p & 0 & -1 \end{vmatrix} \quad , \quad \dots$$

$$\dots, D_p = \begin{pmatrix} -1 & 0 & \dots & 0 & \phi_p & \phi_{p-1} & \dots & \phi_1 \\ \phi_1 & 1 & \dots & 0 & 0 & \phi_p & \dots & \phi_2 \\ \vdots & \vdots \\ \phi_{p-1} & \phi_{p-2} & \dots & -1 & 0 & 0 & \dots & \phi_p \\ \phi_p & 0 & \dots & 0 & -1 & \phi_1 & & \phi_{p-1} \\ \phi_{p-1} & \phi_p & & 0 & 0 & -1 & & \phi_{p-2} \\ \vdots & \vdots \\ \phi_1 & \phi_2 & \dots & \phi_p & 0 & 0 & \dots & 1 \end{pmatrix}$$

Ahora se tiene que

$$\begin{aligned} \gamma_m &= E[Z_t Z_{t+m}] = \sum_{i=1}^p \phi_i E[Z_t Z_{t+m-1}] + E[Z_t a_{t+m}] \\ &= \sum_{i=1}^p \phi_i \gamma_{m-1} \quad ; \quad \text{si } m > 0 \end{aligned} \quad 2.45$$

Luego entonces, si $\gamma_0 < \infty$, se pueden dividir las autocovarianzas $\gamma_1, \gamma_2, \dots, \gamma_p$ entre γ_0 y obtener las ecuaciones de Yule-Walker

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \end{aligned} \quad 2.46$$

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

las demas autocorrelaciones se pueden obtener de la relación

$$\rho_m = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} \quad ; \quad m \geq p+1 \quad 2.47$$

que es en realidad otra manera de escribir la ecuación en diferencias

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m \quad ; \quad m \geq p+1$$

cuyas condiciones iniciales están dadas por 2.47.

Nótese que si el proceso AR(p) es estacionario, es posible representarlo como una suma ponderada de choques aleatorios; es decir, debe poder escribirse

$$Z_t = a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \quad 2.48$$

con $\sum_{i=1}^{\infty} |\psi_i| < \infty$. Los coeficientes ψ_i , $i = 1, 2, \dots$ pueden obtenerse

a partir del hecho de que un proceso AR(p) estacionario debe poderse expresar como

$$\phi_p(B) Z_t = a_t \quad \text{y} \quad Z_t = \psi(B) a_t$$

es decir

$$\psi(B) = \frac{1}{\phi(B)} \quad \text{o} \quad 1 = \phi(B) \psi(B) \quad 2.49$$

El argumento anterior conduce a pensar en la existencia de procesos que puedan representarse en términos de choques aleatorios. A este tipo de procesos se les denomina de promedios móviles y se les denota por MA.

Por último es importante notar que existe solamente una FAC asociada con el proceso AR en estudio; para el caso general del proceso AR(p), dicha FAC queda completa y únicamente determinada por 2.49 y las condiciones iniciales que surgen de las ecuaciones de Yule-Walker.

2.4.2 MODELOS DE PROMEDIOS MOVILES.

La idea básica de estos modelos consiste en representar a un modelo estocástico Z_t , cuyos valores pueden ser dependientes unos de otros, como suma finita ponderada de choques aleatorios independientes a_t ; o sea

$$Z_t = \theta_q(B) a_t$$

Las propiedades generales de estos modelos serán indicadas a partir del análisis de los casos particulares MA(1) y MA(2).

iv) Modelo MA(1).

El proceso de promedios móviles de orden uno es el más simple; dicho proceso MA(1) se expresa mediante

$$Z_t = a_t - \theta a_{t-1}$$

para este modelo es inmediato obtener

$$E[Z_t] = 0 \quad ; \quad \gamma_0 = \text{Var}(Z_t) = \sigma_a^2 (1 + \theta^2)$$

además, las autocovarianzas vienen dadas por

$$\gamma_m = E[Z_t Z_{t+m}] = \begin{cases} \theta \sigma_a^2 & ; \text{ si } m = 1 \\ 0 & ; \text{ si } m \geq 2 \end{cases} \quad 2.50$$

Por lo tanto, la FAC es

$$\rho_m = \begin{cases} \frac{-\theta}{1 + \theta^2} & ; \text{ si } m = 1 \\ 0 & ; \text{ si } m \geq 2 \end{cases} \quad 2.51$$

lo anterior indica que el proceso MA(1) no "recuerda" más allá de lo ocurrido en el período anterior, es decir, tiene memoria limitada a un sólo período.

Como se vió anteriormente, los procesos autorregresivos estacionarios pueden representarse también mediante modelos de promedios móviles; en particular, el proceso AR() tiene una representación equivalente en el modelo MA(1), pues

$$(1 - \phi_1 B - \phi_2 B^2 - \dots)^{-1} a_t = (1 - \phi_1 B) a_t = a_t - \phi a_{t-1}$$

una conclusión que surge de esto es que el proceso MA(1) puede representarse también en forma autorregresiva si $|\theta| < 1$.

En general, cuando un proceso puede expresarse mediante un modelo AR, se dirá que dicho proceso es INVERTIBLE, lo cual significa que tiene como posible representación a

$$\Pi(B) = Z_t = a_t$$

en donde

$$\Pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$$

con la condición de que

$$\Pi(X) = 1 - \sum_{i=1} \pi_i X^i$$

converge dentro o sobre el círculo unitario $|X| \leq 1$.

Nótese que en los argumentos recién expuestos aparece la idea de DUALIDAD entre estacionariedad para un proceso AR e invertibilidad para un proceso MA; en particular, todo proceso MA es estacionario, mientras que todo proceso AR es invertible. Además, las condiciones para invertibilidad de un proceso MA pueden obtenerse de manera similar a las condiciones de estacionariedad para un proceso AR; esto es, la condición de invertibilidad para un proceso MA se expresa también en términos del polinomio $\theta(B)$ a través del requerimiento de que las raíces de la ecuación

$$\theta(X) = 0$$

se encuentren fuera del círculo unitario.

La importancia del concepto invertibilidad radica en que todo proceso invertible está determinado de manera única por su FAC,

lo cual no ocurre con un proceso que no sea invertible.

v) Modelos MA(2)

El siguiente proceso a considerar será el MA(2) cuyo modelo es:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

Dicho proceso tiene como media y varianza a

$$E[Z_t] = 0 \quad ; \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2) \sigma_a^2$$

y como función de autocovarianzas a

$$\begin{aligned} \gamma_k &= E[Z_t Z_{t+m}] \\ &= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}) (a_{t+m} - \theta_1 a_{t+m-1} - \theta_2 a_{t+m-2})] \\ &= E[a_t a_{t+m} + \theta_1^2 a_{t-1}^2 + a_{t+m-1}^2 - \theta_1 (a_t a_{t+m-1} + a_{t-1} a_{t+m}) - \\ &\quad - \theta_1 \theta_2 (a_{t-1} a_{t+m-2} - a_{t-2} a_{t+m-1}) - \theta_2 (a_t a_{t+m-2} + a_{t-2} a_{t+m}) + \\ &\quad + \theta_1^2 a_{t-2} a_{t+m-2}] \\ &= \begin{cases} \theta_1 + \theta_1 \theta_2) \sigma_a^2 & ; \quad \text{si } m = 1 \\ \theta_2 \sigma_a^2 & ; \quad \text{si } m = 2 \\ & ; \quad \text{si } m \geq 3 \end{cases} \quad 2.52 \end{aligned}$$

Por lo tanto la FAC correspondiente es:

$$\left\{ \begin{array}{ll} \frac{\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2} & ; \quad \text{si } m = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & ; \quad \text{si } m = 2 \quad 2.53 \\ 0 & ; \quad \text{si } m \geq 3 \end{array} \right.$$

De las fórmulas anteriores es claro que el proceso es estacionario:

para que el proceso sea invertible se requiere que las raíces de

$$1 - \theta_1 X - \theta_2 X^2 = 0$$

se encuentren fuera del círculo unitario o equivalentemente que

$$|\theta_2| < 1, \quad \theta_1 + \theta_2 < 1 \quad \text{y} \quad \theta_2 - \theta_1 < 1$$

vi) Modelo MA(q)

En general, un proceso estocástico se dirá que sigue un esquema de promedios móviles $q \geq 1$ si tiene la representación dada por (9), esto es

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

dicho proceso tiene como media y varianza a

$$E[Z_t] = 0 \quad \text{y} \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

y como función de covarianza

$$\gamma_m = E[Z_t Z_{t+m}] = \begin{cases} (-\theta_m + \theta_1 \theta_{m+1} + \theta_2 \theta_{m+2} + \dots + \theta_{q-m} \theta_q) \sigma_a^2; & \text{sim}=1, 2, \dots, q \\ 0 & ; \text{sim} \geq q + 1 \end{cases} \quad 2.53$$

Por lo tanto, la correspondiente FAC viene a ser

$$\rho_m = \begin{cases} \frac{-\theta_m + \theta_1 \theta_{m+1} + \dots + \theta_{q-m} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & ; \text{ si } m=1, 2, \dots, q \\ 0 & ; \text{ si } m \geq q + 1 \end{cases} \quad 2.54$$

Lo cual muestra que el proceso MA(q) TIENE UNA MEMORIA LIMITADA a q PERIODOS. Se puede ver que en las fórmulas anteriores que ni la media, ni la varianza, ni las covarianzas del proceso MA(q), dependen del tiempo, lo cual indica que el proceso es estacionario. Si el proceso resulta ser invertible, admitirá también la representación

$$\Pi(B) Z_t = 1 - \pi_1 Z_{t-1} - \pi_2 Z_{t-2} - \dots$$

en donde los coeficientes π_1, π_2, \dots podrían obtenerse de la relación

$$\Pi(B) \theta(B) = 1.$$

2.5 MODELOS ARMA.

Una generalización de los modelos AR y MA previamente descritos, consiste en combinar ambas clases de modelos para obtener lo que se conoce como modelos AUTORREGRESIVOS Y DE PROMEDIOS MOVILES (ARMA). Estos modelos tienen la representación dada por 2.11.

Dicha generalización surge del hecho de que las serie de tiempo que se observan en la práctica, muchas veces presentan características tanto de proceso AR como del proceso MA.

vii) Modelos ARMA(1,1)

El proceso autorregresivo y de promedios móviles de orden (1,1), está definido por

$$(1 - \phi B) Z_t = (1 - \theta B) a_t \quad 2.56$$

y, puesto que contiene tanto características autorregresivas como de promedios móviles, no tiene por que ser invertible ni estacionario, pero las condiciones de estacionariedad e invertibilidad se derivan fácilmente de las condiciones respectivas para procesos MA(1) y AR(1); es decir, para que el proceso (33) sea invertible se requiere que la raíz de $1 - \theta X = 0$ esté fuera del círculo unitario y, similarmente para que sea estacionario, es requisito que la raíz de $1 - \phi X = 0$ se encuentre fuera del círculo unitario.

Si el proceso resulta ser estacionario e invertible, entonces las representaciones

$$Z_t = \psi(B) a_t = 1 - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots$$

y

2.57

$$\Pi(B) Z_t = Z_t - \pi_1 Z_{t-1} - \pi_2 Z_{t-2} - \dots = a_t$$

son tales que las sumas

$$\sum_{i=1}^{\infty} \psi_i \quad \text{y} \quad \sum_{j=1}^{\infty} \pi_j$$

son convergentes.

Para obtener las autocovarianzas de 2.56 es conveniente escribirlo como

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1}$$

en donde se sigue considerando que la media de Z_t es igual a cero, entonces la varianza se obtiene

$$\gamma_0 = \phi E[Z_t Z_{t-1}] + E[Z_t a_t] - \theta E[Z_t a_{t-1}] \quad 2.58$$

en donde para todo t

$$E[Z_t a_t] = E[Z_{t-1} a_t] + E[a_t^2] - \theta E[a_{t-1} a_t] = \sigma_a^2$$

puesto que Z_{t-1} depende de los errores aleatorios hasta el tiempo t-1, pero no depende de a_t ; asimismo

$$\begin{aligned} E[Z_t a_{t-1}] &= E[Z_{t-1} a_{t-1}] = E[a_t a_{t-1}] - \theta E[a_{t-1}^2] \\ &= \phi \gamma_a^2 - \theta \sigma_a^2 \end{aligned}$$

Por lo tanto, 2.58 se convierte en

$$\gamma_0 = \phi \gamma_1 + \sigma_a^2 - \theta(\phi \sigma_a^2 - \theta \sigma_a^2) = \phi \gamma_1 + [1 - \theta(\phi - \theta)] \sigma_a^2$$

De manera similar, las autocovarianzas vienen a ser

$$= \begin{cases} \phi \gamma_0 - \theta \sigma_a^2 & ; & \text{si } m = 1 \\ \phi \gamma_{m-1} & ; & \text{si } m \geq 2 \end{cases} \quad 2.59$$

nótese en esta expresión cómo las autocovarianzas para más de un retraso siguen la misma relación que un proceso AR(1).

Notese que γ_0 y γ_1 se obtiene del siguiente sistema de ecuaciones

$$\gamma_0 = \phi\gamma_1 + [1 - \theta(\phi - \theta)]\sigma_a^2$$

$$\gamma_1 = \phi\gamma_0 - \theta\sigma_a^2$$

que tiene por solución

$$\gamma_m = \frac{(1 - 2\phi\theta + \theta^2)\sigma_a^2}{1 - \phi^2}$$

2.60

$$\gamma_1 = \frac{(1 - \phi\theta)(\phi - \theta)\sigma_a^2}{1 - \phi^2}$$

De 2.59 y 2.60 puede concluirse que la función de autocovarianza es

$$\gamma_m = \frac{\phi^{m-1}(1 - \phi\theta)(\phi - \theta)\sigma_a^2}{1 - \phi^2} \quad ; \quad m = 1, 2, \dots$$

pues $\gamma_m = \phi^{m-1}\gamma_1$, lo cual da origen a la FAC

$$\rho_m = \frac{\phi^{m-1}(1 - \phi\theta)(\phi - \theta)}{1 - 2\phi\theta + \theta^2} \quad ; \quad m = 1, 2, \dots \quad 2.61$$

nótese que debido al factor ϕ^{m-1} y a la condición de estacionariedad $|\phi| < 1$, se sigue que ρ_m tiene un decaimiento exponencial a cero a partir de ρ_1 .

viii) Modelos ARMA (p,q).

El caso general de un proceso se representa mediante 2.30, esto es por

$$\phi_p(B) Z_t = C + \theta_q(B) a_t$$

Para que el proceso autorregresivo y de promedios móviles ARMA(p,q) sea estacionario se requiere que todas las raíces de $\phi_p(X)$ estén fuera del círculo unitario, esto es, $U_d(B) = 1$ y por lo tanto tomando $C = 0$, 2.30 toma la forma

$$\phi_p(B) Z_t = \theta_q(B) a_t \quad 2.62$$

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-p}$$

con $\phi_p(B)$, $\theta_q(B)$, a_t , Z_t definidas como en 2.30.

Se sigue que para $m \geq 0$

$$\gamma_m E[Z_t Z_{t-m}] = \sum_{i=1}^p \phi_i E[Z_{t-1} Z_{t-m}] - \sum_{j=1}^q \theta_j E[a_{t-j} Z_{t-m}] \quad 2.63$$

obsérvese que Z_t está afectada por los choques aleatorios a_{t-m} , a_{t-m-1} , ... pero es independiente a a_{t-m+1} , a_{t-m+2} , ...; de donde se sigue que

$$E[a_{t-1} Z_{t-m}] = 0 \quad ; \quad \text{si } m > 1$$

y por lo tanto, 2.63 da por resultado

$$\gamma_m = \phi_1 \gamma_{m-1} + \phi_2 \gamma_{m-2} + \dots + \phi_p \gamma_{m-p} \quad ; \quad m > q \quad 2.64$$

mientras que para $m \leq q$, γ_m involucrará a los parámetros θ_m , θ_{m+1} , ... , θ_q . En particular, la varianza está dada por la ecuación

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2 - \theta_1 [a_t Z_t] - \dots - \theta_q E[a_{t-q} a_t] \quad 2.64'$$

la cual, debido a que involucra a γ_1 , γ_2 , ... , γ_p tiene que ser resuelta simultáneamente para estas p autocovarianzas.

En lo que respecta a las autocorrelaciones para procesos ARMA(p,q), éstas podrán obtenerse para retrasos mayores a q, de la relación

$$\rho_m = \frac{\gamma_m}{\gamma_0} = \phi_1 \rho_{m-1} + \phi_2 \rho_{m-2} + \dots + \phi_p \rho_{m-p} \quad ; \quad \text{para } m > q \quad 2.65$$

y para retrasos menores o iguales a q, la autocorrelación ρ_m involucra a los parámetros $\theta_m, \theta_{m+1}, \dots, \theta_q$. Lo cual significa que si $q < p$, la FAC estará determinada tan solo por la ecuación en diferencias

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_m = 0 \quad ; \quad \text{para } m > q \quad 2.66$$

sujeta a las condiciones iniciales $\rho_1, \rho_2, \dots, \rho_p$.

Se sabe que la representación de los procesos definidos en (2.65), con las condiciones impuestas sobre $\phi_p(B)$ y $\theta_q(B)$, aseguran que a_t es la innovación de Z_t en el tiempo t. Se dirá que esta expresión es MINIMAL, si $\phi_p(B)$ y $\theta_q(B)$ no tienen raíces comunes, se sabe también que dicha representación minimal existe, es única y es obtenida a partir de cualquier representación del tipo (2.65) simplemente al eliminar las posibles raíces comunes de $\phi_p(B)$ y $\theta_q(B)$. Las ideas expuestas anteriormente dan la pauta para una primera caracterización de los procesos ARMA(p,q), la cual se llamará "representación minimal ARMA(p,q)". A continuación se da una definición que posteriormente ayudará para formalizar estas ideas.

Considérese una sucesión de números reales S_m ; $m \in \mathbb{Z}$ la cual satisface una ecuación en diferencias lineal de orden r para el rango R ; esto es, existen $\phi_1, \phi_2, \dots, \phi_r$; $\phi_r \neq 0$ tales que

$$S_m - \phi_1 S_{m-1} - \dots - \phi_y S_{m-y} \begin{cases} = 0 & ; \text{ si } m \geq R \\ \neq 0 & ; \text{ si } m = R - 1 \end{cases}$$

tal sucesión también satisface otras ecuaciones en diferencias para m mayores que algún rango (posiblemente diferente de R). Entre todas las ecuaciones, una tiene un orden minimal r_0 . El rango asociado con este orden se denotará por R_0 y la sucesión S_m ; $m \in \mathbb{Z}$ se dirá que satisface una ecuación en diferencias de orden minimal r_0 para el rango minimo R_0 . Usando la definición anterior, se puede dar la primera caracterización de una representación minimal ARMA(p,q) (ver Beguín y Gourieroux (1980) para más detalles).

Nótese que hasta ahora se ha centrado la exposición en modelos que son estacionarios, pero en la práctica lo más común es que las series que se analizan sean no estacionarias, ya sea porque exhiben algún tipo de tendencia, porque su varianza no sea constante o porque estén influenciadas por algún factor de tipo semideterminista como puede ser la ESTACIONALIDAD. El caso de series influenciadas por elementos estacionales se verá posteriormente; si, en cambio, el problema es que se aprecia una Tendencia en el comportamiento de la serie, es posible que dicha tendencia sea de carácter polinomial adaptivo y que por lo tanto, pueda eliminarse mediante la aplicación del operador diferencia lo cual da origen a los modelos ARIMA. Por otro lado, si la no estacionariedad se debe también a que la varianza no es constante, quizá la causa sea que en cada punto de observación t , la variable Z_t tiene varianza σ_t^2 la cual es función de su media μ_t ; de ocurrir esto, un argumento derivado del trabajo de Bartlett (1947), conduce a determinar una transformación potencia del tipo

$$T(Z_t) = \begin{cases} Z_t^\lambda & ; \text{ si } \lambda \neq 0 \\ \log(Z_t) & ; \text{ si } \lambda = 0 \end{cases}$$

dicha transformación puede ser útil para estabilizar la varianza de la serie, antes de cancelar la posible tendencia polinomial adaptiva.

Del estudio de la función de autocorrelacion se puede deducir un criterio para reconocer el proceso MA e identificar su orden : si $\rho_m = 0$ para $m > q$. Entonces el proceso Z_t tiene una representación MA(q).

Se trata entonces de encontrar otro criterio de este tipo que permita reconocer los procesos AR. Este criterio se obtendrá a partir de las ecuaciones de Yule-Walker 2.46, las cuales pueden escribirse como

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{bmatrix} = B(n,0) \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{bmatrix}$$

en donde $B(n,m)$ es una matriz de orden $n \times n$ de finida por

$$B(n,m) = \begin{bmatrix} \rho_m & \rho_{m-1} & \dots & \rho_{m-n+1} \\ \rho_{m+1} & \rho_m & \dots & \rho_{m-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m+n-1} & \rho_{m+n-2} & \dots & \rho_m \end{bmatrix}$$

Ahora se define la autocorrelación parcial como

$$\phi_{nn} = \frac{|A(n,0)|}{|B(n,0)|} \quad 2.67$$

en donde $A(n,m)$ es una matriz compuesta por las $n-1$ primeras columnas de la matriz $B(n,m)$ y con la n -ésima columna dada por $V = (\rho_{m+1}, \rho_{m+2}, \dots)$

... , ρ_{m+n}). Esto es, si se supone que el proceso es un ARMA(n,0), entonces ϕ_{nn} es la solución de las primeras n ecuaciones de Yule-Walker obtenidas mediante la regla de Cramer, para el último coeficiente autorregresivo.

Si no se conoce n, que es lo que ocurre en la práctica, habrá que hacer el cálculo para $n = 1, 2, \dots$ lo que da la siguiente relación

$$\phi_{nn} = \begin{cases} \rho_1 & ; \quad \text{si } n = 1 \\ \frac{|A(n,0)|}{|B(n,0)|} & ; \quad \text{si } n > 1 \end{cases} \quad 2.68$$

que es conocida como la Función de Autocorrelación Parcial (FACP). obsérvese que si Z_t es un AR(p) entonces

- i) si $n < p$; ϕ_{nn} no será cero en general
- ii) si $n = p$ $\phi_{nn} = \phi_p$ no es cero
- iii) si $n > p$ $\phi_{nn} = 0$

Se puede ver que en el caso de un proceso MA o de un modelo ARMA, la FACP no tiene ninguna razón para ser cero cuando los retrasos sobrepasan un cierto valor. Aunque se observa que en ambos casos, ésta decae en forma exponencial.

De esta manera se ha llegado a un criterio para reconocer los procesos AR e identificar su orden:

si $\phi_{nn} = 0$ para $n > p$ entonces el proceso Z_t es un AR(p).

En el caso en donde ni ρ_m ni ϕ_{nn} se vuelvan cero para m o n suficientemente grande, habrá que buscar otro tipo de modelo, por ejemplo un ARMA.

Bartlett (1946) dedujo, en el caso Gaussiano, la varianza de este estimador. En el caso en que el cálculo se hace para un MA(q) y para $m > q$ esta varianza vale aproximadamente

$$\text{var}(\hat{\rho}_m) = \frac{1}{N} \left\{ 1 + 2 \sum_{j=1}^q \rho_j \right\} \quad 2.69$$

y puede ser estimada por

$$\hat{\text{var}}(\hat{\rho}_m) = \frac{1}{N} \left\{ 1 + 2 \sum_{j=1}^q \hat{\rho}_j \right\} \quad 2.70$$

Igualmente, para un AR(p) y para $m > p$, Guenoville (1942) obtuvo la varianza de $\hat{\phi}_{nn}$ y se encontró que aproximadamente

$$\hat{\text{var}}(\hat{\phi}_{nn}) = \frac{1}{N} \quad 2.71$$

Además, para N suficientemente grande $\hat{\rho}_m$, $\hat{\phi}_{nn}$, tendrán una distribución Gaussiana. Se sabe que el valor absoluto de una variable aleatoria Gaussiana tiene una probabilidad de aproximadamente 5% de sobrepasar dos veces su desviación estándar.

Frente a una serie de tiempo, se calculan $\hat{\rho}_m$, $\hat{\phi}_{nn}$, $\hat{\text{var}}(\hat{\rho}_m)$ y $\hat{\text{var}}(\hat{\phi}_{nn})$, se grafican las $\hat{\rho}_m$, indicando sus ordenadas

$$(2 \sqrt{\hat{\text{var}}(\hat{\rho}_m)}, -2 \sqrt{\hat{\text{var}}(\hat{\rho}_m)})$$

Se podrá considerar que cualquier $\hat{\rho}_m$ comprendida entre esos dos casos límite, corresponde a un valor cero de ρ_m . Análogamente para la FACP se establece que $\hat{\phi}_{nn}$ se encuentra fuera del intervalo definido por $\pm 2 \sqrt{\hat{\text{var}}(\hat{\phi}_{nn})}$ para $n > p$.

En conclusión, para llevar acabo la etapa de identificación es conveniente recordar las características de la FAC y de la FACP; dichas características se resumen en el cuadro (1) para una serie estacionaria Z_t (ver Guerrero, 1983).

COMPORTAMIENTO DE LAS FAC Y FAC'S PARA PROCESOS AR, MA, Y ARMA

PROCESO	FAC	FACP
AR(p)	CONVERGENCIA A CERO, CON COMPORTAMIENTO DICTADO POR LA ECUACION $\phi(B)\rho_m = 0$, PARA $m \geq p$	SOLAMENTE LAS PRIMERAS AUTOCORRELACIONES PARCIALES SON DISTINTAS DE CERO
MA(q)	SOLO LAS PRIMERAS q AUTOCORRELACIONES SON DISTINTAS DE CERO	SUCESION INFINITA CONVERGENTE A CERO
ARMA(p,q)	COMPORTAMIENTO IRREGULAR DE LAS PRIMERAS q AUTOCORRELACIONES Y DESPUES CONVERGENCIA A CERO DE ACUERDO CON $\phi(B)\rho_m = 0$, PARA $m > q$	SUCESION INFINITA CONVERGENTE A CERO

En resumen, la técnica de Box-Jenkins para identificar modelos de series de tiempo, se basa en inferencias construidas a partir de patrones de la FAC Muestral (FACM) y de la FACP Muestral (FACPM) de la serie. Esto es, la técnica de Box-Jenkins usa el hecho de que si Z_t es un ARMA(p,0), entonces $\phi_{nn} \neq 0$ para $n = p$ y $\phi_{nn} = 0$, para $n > p$, mientras que si Z_t es un ARMA(0,q), entonces $\rho_m = 0$ para $m > q$. Dicha técnica emplea la inspección visual de las gráficas de las FACM y FACPM. Sin embargo, cuando p y q son ambos diferentes de cero, la FACP y la FAC no poseen gráficas que determinen de manera única los valores de p y q por una simple inspección, en este caso las correspondientes características de las funciones teóricas para varios tipos de modelos ARMA(p,q) han sido usadas como patrones de referencia (ver por ejemplo Box-Jenkins 1976 y Guerrero 1983).

2.6. SERIES DE TIEMPO ESTACIONALES.

Por una serie de tiempo estacional se entenderá una serie de

tiempo que, aparte de contener tendencia (y/o ciclo) de larga duración, muestra fluctuaciones que se repiten anualmente; quizá con cambios graduales através de los años.

Una característica fundamental de las series estacionales con observaciones mensuales, es la similitud de comportamiento entre observaciones para el mismo mes en años consecutivos. De igual manera, si la serie estacional es de observaciones semestrales, existirán similitudes entre los mismos semestres para años consecutivos, de tal forma que los meses de diciembre fuesen similares a los meses de junio; además de que los diciembres muestran comportamiento similar entre sí y de que los junios se comporten también de manera parecida.

Por lo anterior, es conveniente hablar genéricamente de un período estacional que comprenda E observaciones contiguas.

OPERADORES Y MODELOS PURAMENTE ESTACIONALES.

El operador de diferencia estacional ∇_E^m se define como

$$\nabla_E^m Z_t = (1 - B)^m Z_t = \sum_{j=1}^m \binom{m}{j} (-1)^j Z_{t-jE}$$

para $m = 0, 1, \dots$ y $E = 1, 2, \dots$ 2.72

mientras que un polinomio de retraso estacional de orden K con coeficientes constantes $\alpha_1, \alpha_2, \dots, \alpha_k$ viene dado por

$$G(B^E) = 1 - \sum_{j=1}^k \alpha_j B^{jE} \quad 2.73$$

Con la notación recién introducida es posible obtener representaciones puramente estacionales de tipo ARIMA(P,D,Q) cómo

$$\Phi_p(B^E) \nabla_E^m (Z_t - \mu) = \Theta_Q(B^E) a_t \quad 2.74$$

en donde μ es la media de Z_t , $\Phi(B^E)$ representa a un polinomio autorregresivo estacional de orden p , $\Theta(B^E)$ denota a un polinomio de promedios móviles estacional de orden Q y la sucesión $\{a_t\}$ es ruido blanco. Para observar la analogía del modelo estacional 2.73 con el modelo no estacional (10), es conveniente considerar algunos casos particulares de 2.74.

Considérese el proceso ARIMA $(1,0,0)_E$ descrito por

$$Z_t = \phi Z_{t-E} + a_t \quad 2.75$$

para este proceso se tiene la varianza

$$\gamma_0 = E[(\phi Z_{t-E} + a_t)^2] = \phi^2 E[Z_{t-E}^2] + 2\phi E[Z_{t-E} a_t] + E[a_t^2]$$

en donde, suponiendo estacionariedad para $\{Z_t\}$, se obtiene

$$\gamma_0 = \phi^2 \gamma_0 + \sigma_a^2$$

es decir

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi^2}$$

y, de manera semejante, se obtienen también las autocovarianzas

$$\gamma_E = \frac{\phi \sigma_a^2}{1 - \phi^2}$$

$$\gamma_{2E} = \frac{\phi^2 \sigma_a^2}{1 - \phi^2}$$

$$\gamma_{iE} = \frac{\phi^i \sigma_a^2}{1 - \phi^2}$$

para $i \geq 0$

mientras que $\gamma_m = 0$ para $m \neq iE$.

Por lo tanto, la FAC de 2.75 viene a ser

$$\rho_m = \begin{cases} \phi^1 & \text{para } m = iE \text{ con } i = 0, 1, 2, \dots \\ 0 & \text{para } m \neq iE \end{cases} \quad 2.76$$

Otro caso particular de 2.74 lo constituye el proceso $ARIMA(0,0,1)_E$ representado mediante la expresión

$$Z_t = a_t - \Theta a_{t-E}$$

y para el cual se tiene la función de autocovarianza

$$\gamma_0 = (1 + \Theta^2) \sigma_a^2$$

$$\gamma_E = -\Theta \sigma_a^2$$

$$\gamma = 0$$

para $m \neq E$

que, en términos de la FAC, implica

$$\rho_m = \begin{cases} \frac{\Theta}{1 - \Theta^2} & \text{si } m = E \\ 0 & \text{si } m \neq E \end{cases} \quad 2.77$$

o sea que existirá únicamente una autocorrelación diferente de cero, correspondiente al retraso $m = E$.

De 2.37 y 2.76 se puede apreciar la analogía que existe entre las FAC'S de los procesos $AR(1)$ y $AR(1)_{E_j}$; lo mismo ocurre con las relaciones 2.51 y 2.77. Dicha analogía se generaliza de manera obvia a los modelos más generales $AR(P)_E$; y de igual forma a los modelos mixtos puramente estacionales $ARMA(P,Q)_E$ y $ARIMA(P,D,Q)_E$. De esta manera la metodología de construcción de modelos para el caso no estacional expuesta previamente se aplica sin mayores modificaciones a

los procesos que son puramente estacionales. Sin embargo este hecho, aunque aparentemente importante, es de poca utilidad en la práctica ya que en raras ocasiones la serie en estudio será puramente estacional y, por el contrario, mostrará tanto efectos estacionales como no-estacionales.

1.7. MODELO MULTIPLICATIVO ESTACIONAL.

Para tener en cuenta ambos tipos de efectos, estacionales y no-estacionales, BOX-JENKINS (1970) han propuesto un modelo general de la forma

$$\Phi(B^E) \nabla_D^E (Z_t - \mu) = \Theta(B^E) \alpha_t \quad 2.78$$

en donde las variables $\{\alpha_t\}$ no se suponen ruido blanco, sino generadas por un proceso ARIMA(p,d,q), o sea

$$\phi(B) \nabla^d \alpha_t = \theta(B) a_t \quad 2.79$$

con $\{a_t\}$ un proceso de ruido blanco. De estas últimas expresiones se obtiene el modelo multiplicativo estacional

$$\phi(B) \Phi(B^E) \nabla^d \nabla_E^D (Z_t - \mu) = \theta(B) \Theta(B^E) a_t$$

denotado por ARIMA(p,d,q) × (P,D,Q)_E. La importancia de este modelo radica en que, para series con observaciones mensuales, como las del problema en cuestión, se tienen en cuenta posibles relaciones

- i) entre meses contiguos dentro de los años, debido a 2.79 y
- ii) entre años para los mismos meses, a causa de 2.78.

En particular esto significa que por ejemplo, la elevación

Del nivel de la precipitación pluvial en los meses de Junio, Julio y Agosto en algunos estados de la República Mexicana, debe poder explicarse tanto por el comportamiento de estas observaciones en unos cuantos meses previos a Junio, Julio y Agosto, así como por el nivel alcanzado por la precipitación pluvial en los meses de Junio, Julio y Agosto del o de los años anteriores.

Como es de suponer, a mayor complejidad del modelo deberá corresponder una estructura de autocorrelación más compleja.

2.8 IDENTIFICACION DE MODELOS PARA SERIES DE TIEMPO ESTACIONALES.

El proceso de construcción de modelos para series de tiempo estacionales es el mismo que para series de no-estacionales, pero ahora la etapa de identificación es un tanto más complicada debido a que el número de modelos que podrían postularse para representar a una serie de tiempo dada, aumenta considerablemente.

Cabe notar que en la forma explícita de los modelos que aquí se trataran se hace uso, no de la variable original Z_t , sino de la variable estacionaria W_t definida por

$$W_t = \nabla^d \nabla_E^D T(Z_t) \quad 2.80$$

para algunas variables de d , D y alguna transformación apropiada $T(Z_t)$. Los valores D y d deben determinarse de tal manera que la serie resultante, $\{W_t\}$, sea aproximadamente estacionaria; para esto podrían graficarse las FAC muestrales de $\{T(Z_t)\}$, $\{\nabla T(Z_t)\}$, $\{\nabla^2 T(Z_t)\}$

$\{\nabla_E T(Z_t)\}$, $\{\nabla \nabla T(Z_t)\}$ y $\{\nabla^2 \nabla_E T(Z_t)\}$, la inspección visual de estas gráficas permitirá decidir el grado de diferenciación requerido para estacionarizar la serie. Otra ayuda para decidir los valores de d y D consiste en calcular la desviación estándar de las series anteriores y

quizá de la serie diferenciada una vez más, tanto regular como estacionalmente; la idea es pues calcular los valores

$$s^2(j,i) = \frac{1}{N - E_1 - J - 1} \sum_{t=E_1+J+1}^N \{ \nabla^j \nabla_E^i T(Z_t) - \sum \nabla^j \nabla_E^i T(Z_t) / (N - E_1 - J) \}^2$$

para J, i = 0, 1, 2. 2.81

ya que la serie estacionaria será aquella que satisfaga el criterio

$$S(d,D) = \min \{ S(i,j) ; i,j = 0, 1, 2 \}$$

de manera aproximada. 2.82

PARTE III

ESTUDIO COMPARATIVO Y SELECCION DE METODOLOGIA.

INTRODUCCION.

El problema de pronóstico de lluvia para el estado de Morelos se estudia de acuerdo a la información recopilada para un total de catorce (14) estaciones climatológicas ubicadas en variados lugares del estado.

Las estaciones seleccionadas poseen información pluviométrica que comprende el periodo que va desde Enero de 1937 a Diciembre de 1989.

El total de las catorce estaciones con información pluviométrica fue clasificado de acuerdo a regiones de igual precipitación (isoyetas) obtenidas del mapa de la Región México Carta de Precipitación Total Anual 1:1.000,000 (segunda impresión, 1984) que publica la Secretaría de Programación y Presupuesto a través del Instituto Nacional de Estadística Geografía e Informática.

Se estudia la isoyeta de 1000 mm. a 1200 mm. de precipitación total anual, Tabla 3.1, para elaborar su pronóstico conforme a algunos de los métodos descritos en la Parte II y comparar los resultados que ellos arrojan, a fin de seleccionar un método adecuado para estudiar el tipo de serie de tiempo que se presenta en éste problema y realizar un pronóstico para todo el estado.

3.1 ESTUDIO DE LA ISOYETA 1000 - 1200 mm.

3.1.1 METODOLOGIA BOX-JENKINS.

En esta parte se ejemplifica la metodología para la construcción de modelos de Series de Tiempo desarrollada por Box-Jenkins, haciendo uso de la información correspondiente a la isoyeta 1000 - 1200 mm. del Estado de Morelos.

Todos los modelos fueron obtenidos de manera análoga siguiendo la estrategia de construcción de modelos para series de tiempo desarrollada por Box-Jenkins (1970).

- i) Identificación de un posible modelo dentro de la clase de modelos ARIMA $(p,d,q) \times (P,D,Q)$.

En ésta etapa se tiene como objetivo determinar los ordenes de los polinomios autorregresivo (AR) y de promedios móviles (MA), tanto regulares como estacionales, así como el número de veces que se aplicará el operador diferencia para cancelar la no-estacionariedad.

La tabla 3.1 muestra los datos de la serie de precipitación pluvial mensual acumulada del periodo Enero de 1937 a Diciembre de 1989 de la isoyeta 1000 - 1200 mm., denotada ISO-1012, y en la figura 3.2 se presenta la gráfica de los datos observados.

- i) Identificación del modelo.

Este tipo de serie no requiere estabilizar su varianza, así que se procede a aplicar el operador diferencia para estabilizar el nivel de la serie. Para conocer el grado de diferenciación obtenemos

18010 J 1
 PROMEDIO DE PRECIPITACION PLUVIAL MENSUAL ACUMULADA
 ISOYETA 1000 - 1200 mm.

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	ANO
5.77	5.10	9.43	32.10	88.57	196.03	219.17	184.17	239.67	104.70	2.83	4.87	1937
1.23	12.70	49.37	1.87	45.07	209.43	262.27	179.93	148.27	50.53	25.60	3.73	
.50	8.13	7.37	23.00	50.67	230.17	195.93	196.90	129.33	170.37	8.13	8.90	
5.99	8.03	23.17	.00	43.25	245.37	287.90	219.73	156.63	23.17	17.13	5.23	
3.37	9.50	5.70	5.17	31.97	272.20	131.47	179.23	211.97	149.63	15.00	15.17	
5.70	6.77	4.80	.03	29.17	207.17	189.90	205.73	340.03	33.63	30.50	.00	
5.10	.00	7.27	4.67	39.10	229.33	132.47	229.70	337.93	70.83	5.10	15.23	
9.67	.00	20.80	21.63	34.00	191.73	182.07	253.00	230.37	81.63	10.97	.00	
1.97	.07	.00	15.20	87.93	140.83	119.93	299.97	213.30	86.83	7.03	8.50	
8.57	.00	5.27	74.53	39.53	217.80	223.10	188.40	134.60	93.87	29.10	.00	
12.13	.90	.00	13.43	103.27	198.83	200.50	222.03	207.93	141.00	.33	17.80	
30.63	.07	.57	21.27	119.20	168.33	231.43	178.57	205.23	81.30	76.97	.00	
.00	.00	4.17	4.70	47.37	193.57	99.73	128.63	159.50	70.43	.20	.00	
5.90	4.47	3.93	16.40	90.17	158.57	231.50	214.57	244.83	62.83	13.13	.00	
.00	.00	12.13	5.43	46.27	181.67	236.20	201.03	281.43	73.03	12.23	.07	
15.77	9.40	4.80	35.43	109.27	235.80	208.57	218.50	284.83	59.10	46.60	.00	
1.07	.00	3.30	.80	16.80	169.10	253.97	252.67	212.93	119.37	17.67	3.93	
.07	14.57	4.97	52.43	92.77	226.03	207.53	190.67	321.03	96.93	11.00	.17	
.97	3.00	3.87	7.73	28.03	226.00	270.03	249.57	100.57	8.13	13.93	.00	
41.00	22.00	1.50	34.37	105.47	258.27	221.10	186.17	278.80	91.80	76.10	26.77	
3.43	5.60	.00	24.90	44.90	222.07	158.07	228.27	213.60	107.03	7.03	1.67	
111.13	20.57	.00	9.53	72.70	362.30	427.60	250.53	339.40	226.43	155.63	19.53	
26.93	4.17	.00	58.97	28.23	293.47	207.70	169.03	149.50	228.07	15.73	.00	
21.70	.00	.00	5.87	17.37	140.33	225.20	175.43	179.17	100.07	3.33	.33	
.00	2.33	3.33	1.23	34.73	319.10	202.33	206.53	258.23	16.27	20.77	.00	
.00	.00	.00	31.50	32.20	262.03	147.83	189.63	296.43	66.83	14.70	8.50	
.00	.00	14.97	.33	57.67	183.57	296.80	113.60	136.50	66.17	2.67	1.33	
15.97	.00	1.33	4.00	172.77	221.13	141.50	120.17	263.50	41.00	9.33	13.30	
30.67	24.10	.00	50.83	74.53	175.27	152.77	388.80	106.00	105.70	21.23	3.80	
7.33	1.50	26.60	61.87	22.90	88.43	206.43	309.53	204.30	95.90	.00	2.10	
21.20	1.30	3.07	4.37	43.70	236.20	162.17	263.27	311.60	154.30	10.17	25.93	
5.50	22.50	5.07	43.60	81.90	315.93	218.17	133.67	183.30	94.30	1.77	44.30	
7.30	.67	23.07	1.33	13.97	83.50	292.97	529.07	237.70	55.03	1.00	.10	
.00	.00	.00	1.67	17.90	218.47	273.53	176.77	324.60	40.43	19.87	.00	
.47	.00	3.03	23.93	41.60	273.40	146.10	247.80	377.67	193.20	9.00	8.67	
.00	.00	20.50	45.00	84.83	358.10	253.53	121.67	215.50	67.27	30.67	.67	
.00	8.67	2.33	40.83	58.83	332.47	230.97	376.50	329.63	198.33	10.33	1.03	
.50	6.00	32.13	21.80	69.17	211.33	214.50	166.00	195.33	8.67	6.33	.00	
21.83	1.83	.00	2.00	69.33	230.00	241.33	262.40	232.47	51.67	.00	.00	
2.50	5.67	12.00	38.83	37.83	102.33	369.50	230.83	146.67	212.17	35.33	7.50	
3.17	4.00	.00	7.00	75.33	167.17	118.33	233.40	171.67	47.00	12.17	1.67	
3.00	10.83	28.17	4.33	41.87	227.17	168.73	251.63	314.13	122.47	22.07	7.17	
11.70	15.37	10.77	20.27	48.17	193.03	255.80	242.37	201.77	76.60	5.23	20.73	
91.17	2.23	.00	11.13	70.93	136.37	193.60	441.77	251.23	118.17	17.23	9.00	
37.47	28.23	74.30	31.73	48.30	308.40	371.33	352.03	239.50	60.10	.67	.00	
.00	10.77	2.23	18.57	139.80	235.50	195.30	96.40	162.13	106.07	.67	.67	
20.10	31.90	5.50	.00	18.03	151.93	265.63	163.37	369.40	59.50	16.10	2.37	
23.30	3.53	.53	.00	61.03	224.37	226.23	189.13	244.60	48.60	.00	6.53	
4.57	7.00	2.83	17.17	26.87	209.67	252.60	119.80	190.80	40.13	23.03	.00	
.00	1.50	.50	24.40	85.53	187.63	119.97	197.63	150.37	94.27	20.60	1.00	
1.33	.40	10.93	29.10	30.53	230.37	365.10	162.23	219.60	18.90	26.43	.00	
.00	7.40	14.80	8.37	23.50	312.90	201.37	288.10	168.90	45.60	14.53	.00	
.00	.37	3.73	11.73	30.73	181.17	180.83	223.30	216.87	77.20	22.33	1.57	1989

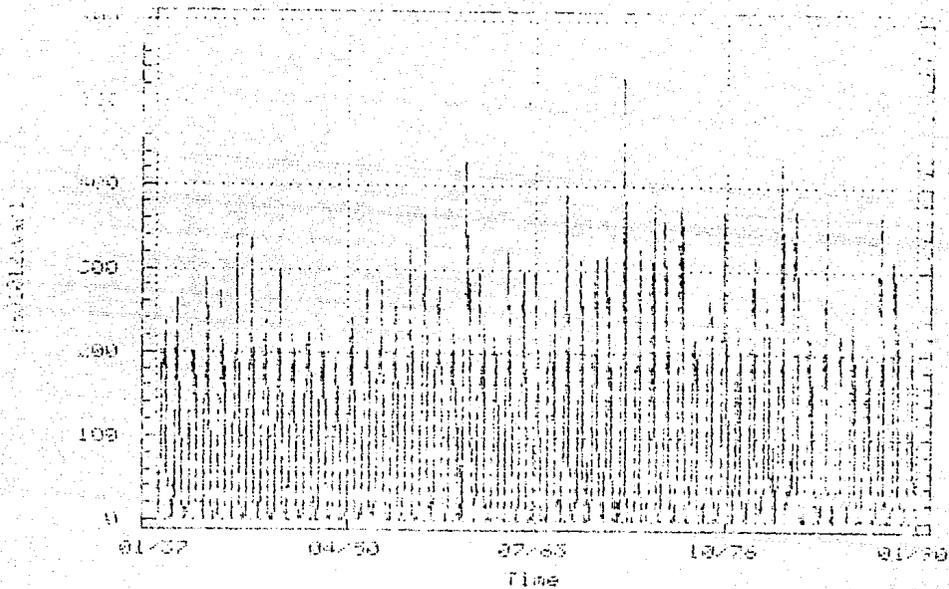


fig. 3.2

la FAC muestral, ya que como sabemos, un decaimiento rápido a cero de las autocorrelaciones indica que la serie es estacionaria en cuanto al nivel se refiere.

En la figura 3.3 se presentan las gráficas de la FAC muestral de la Serie y de algunas de sus diferencias; la figura (3.3d) presenta un decaimiento rápido a cero excepto en el retraso 12, por lo que se empieza a notar que se requiere una diferencia estacional para estacionarizar la serie ; complementariamente se calculan las desviaciones estandar muestrales para la serie diferenciada tanto regular como estacionalmente, ya que al tomar diferencias sucesivas la varianza de la serie se altera de tal manera que decrece hasta que la serie es estacionaria y comienza a crecer con la sobre diferenciación.

Las desviaciones estándar de las diferencias indicadas fueron las siguientes:

S (0,0)	=	105.5816
S (1,1)	=	90.6001
S (2,0)	=	130.5769
S (0,1)	=	66.7506 *
S (1,1)	=	90.0649
S (2,1)	=	153.2154
S (0,2)	=	118.0196
S (1,2)	=	159.1524
S (2,2)	=	271 5380

Se decide, por lo tanto, que las diferencias requeridas son: una estacional de orden 12; es decir, la serie aproximadamente estacionaria que se analizará resulta ser:

$$W_t = \nabla_{12} Z_t .$$

I 20 1012

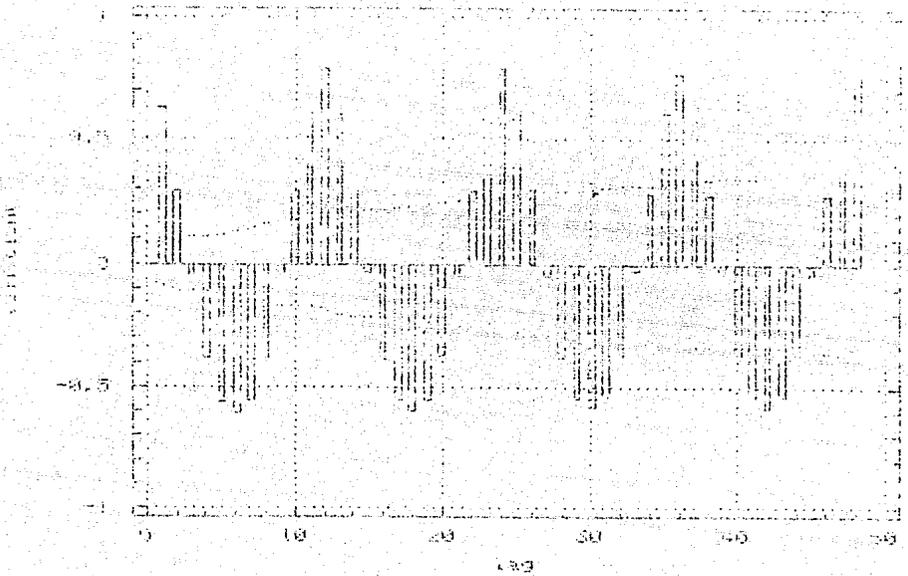


fig. 3.3 FAC de Z_t

Estimated Autocorrelations

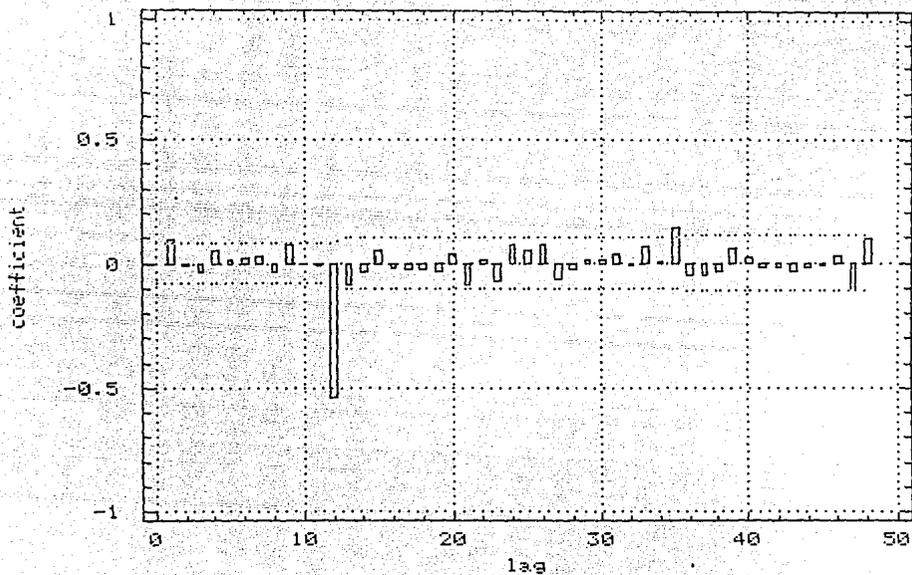


fig. 3.3 d) FAC de $W_t = \nabla_{12} Z_t$; ISO1012

es decir se tiene que:

$$\nabla_{12} Z_t = Z_t - Z_{t-12}$$

al aplicar una diferencia estacional a la serie de precipitación pluvial mensual acumulada, se obtiene la serie de flujos anuales de precipitación. Puesto que la serie Z_t de referencia es una serie mensual que cubre el periodo Enero de 1937 a Diciembre de 1989, por lo tanto consta de 636 observaciones originales, tal número de observaciones se reduce a 624 al aplicar el operador diferencia ∇_{12} .

Nota: para realizar los cálculos estadísticos se hace uso del paquete PC SCA Statistical System. de la Scientific Computing Associates.

El siguiente paso en la etapa de identificación consiste en asociar la FAC muestral con un posible proceso generador del tipo ARIMA (p,d,q) x (P,D,Q).

La gráfica de las autocorrelaciones de la figura 3.3 d) es característica de un modelo multiplicativo estacional:
ARIMA (0,0,1) X (0,1,1)

$$\nabla_{12} Z_t = (1 - \theta_1 B) (1 - \theta_2 B^{12}) a_t. \quad 3.1.1$$

ii) Estimación de los parámetros involucrados en el modelo.

Lo que procede es obtener los mejores valores de los parámetros θ_1 y θ_2 para que dicho modelo represente apropiadamente a la serie en consideración.

Una vez propuesto el modelo 3.1 se realiza la estimación de los parámetros que están involucrados.

El valor obtenido en la estimación de los parámetros resultó ser:

$$\begin{aligned}\hat{\theta}_1 &= -0.0899 \\ \hat{\theta}_2 &= 0.9530\end{aligned}$$

no estando correlacionados los parámetros.

iii) Verificación de que el modelo proporciona un ajuste adecuado y que los supuestos básicos se satisfacen. En caso de que los supuestos no se cumplan, se determinan las modificaciones necesarias al modelo. Habrá que elegir el modelo que presente menos fallas, hasta que la verificación indique resultados aceptables.

El análisis de los residuales nos permite observar las posibles violaciones que presente el modelo a los supuestos. Así, los residuales son aquella parte de las observaciones que no pueden ser explicadas por el modelo y nos miden por lo tanto, la discrepancia entre los valores observados y los valores estimados por el modelo.

Cuando el tamaño de muestra es grande, los errores aleatorios y los residuales son esencialmente iguales; por lo que al analizar los residuales observados $\{ \hat{a}_t \}$ se analiza básicamente lo que debería ser una realización del proceso de ruido blanco $\{ a_t \}$.

En la gráfica de residuales que se presenta en la figura 3.4 para el modelo 3.14 podemos apreciar que ellos son estadísticamente cero por lo tanto el modelo es prácticamente aceptable.

$\{ a_t \}$ tiene media cero.

Una vez teniendo calculada la media y la desviación estándar

coefficient

	-1.0	-.8	-.6	-.4	-.2	.0	.2	.4	.6	.8	1.0
1	.00					I					
2	-.01					+ I +					
3	.02					+ I +					
4	.05					+ IX+					
5	.04					+ IX+					
6	.02					+ I +					
7	.05					+ IX+					
8	.00					+ I +					
9	.07					+ IXX					
10	.00					+ I +					
11	.03					+ IX+					
12	.03					+XI +					
13	.03					+XI +					
14	.00					+ I +					
15	.04					+ IX+					
16	.01					+ I +					
17	.01					+ I +					
18	.02					+ I +					
19	.00					+ I +					
20	.04					+ IX+					
21	.03					+XI +					
22	.03					+ IX+					
23	.02					+ I +					
24	.06					+ IX+					
25	.01					+ I +					
26	.08					+ IXX					
27	.05					+XI +					
28	.03					+XI +					
29	.02					+ I +					
30	.01					+ I +					
31	.03					+ IX+					
32	.01					+ I +					
33	.06					+ IX+					
34	.02					+ I +					
35	.11					+ IX+X					
36	.02					+ I +					
37	.02					+ I +					
38	.00					+ I +					
39	.01					+ I +					
40	.00					+ I +					
41	.00					+ I +					
42	.01					+ I +					
43	.02					+ I +					
44	.01					+ I +					
45	.01					+ I +					
46	.02					+ IX+					
47	.07					XXI +					
48	.03					+ IX+					
49	.03					+ IX+					
50	.00					+ I +					

fig. 3.4

FAC de los Residuales de 3.1.1

coefficient

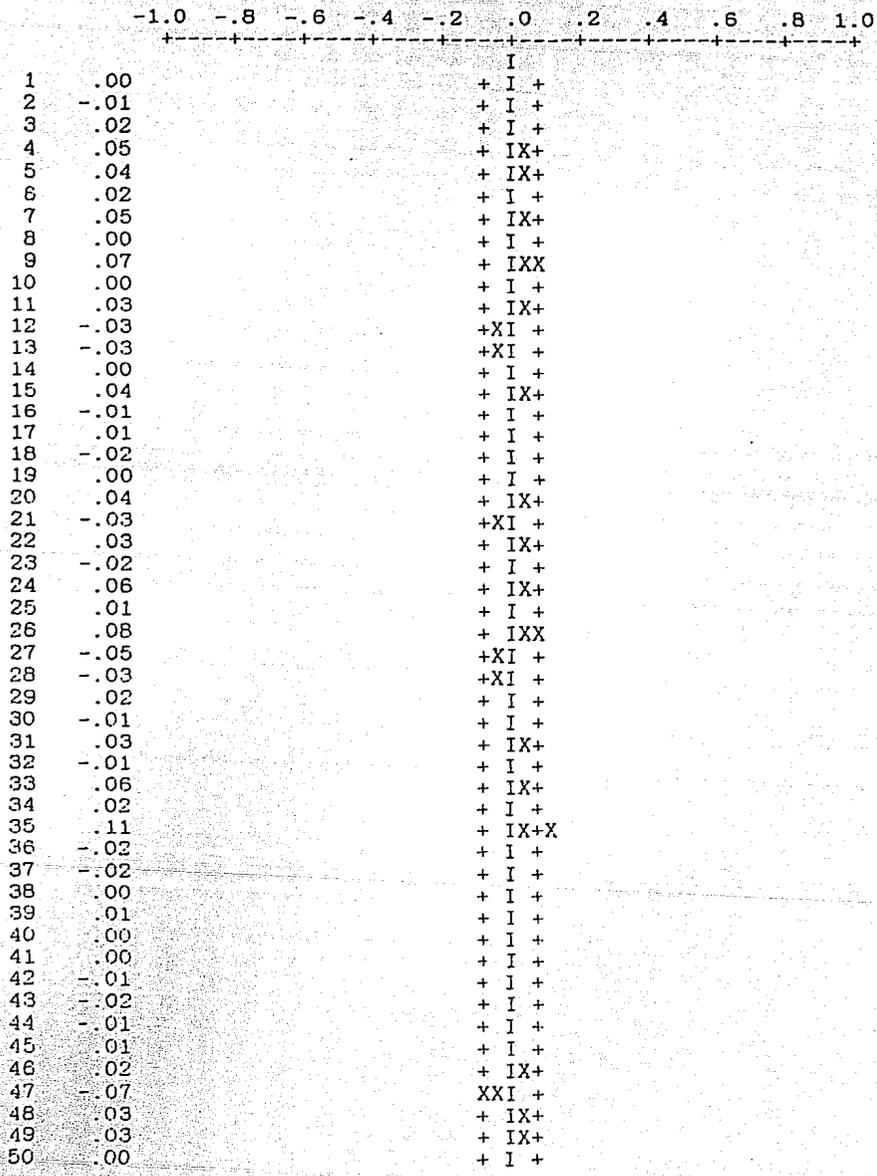


fig. 3.4 FAC de los Residuos de 3.1.1

muestral de los residuales se checa la hipótesis siguiente:

$$\text{Si } | (n)^{1/2} m(a) / \sigma(a) | < 2.$$

se dice que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero y por lo tanto no se rechaza el supuesto.

Así resulta que:

$$m(\hat{a}) = 0.4778 \quad \text{y} \quad \sigma(\hat{a}) = 47.4444.$$

por lo que el cociente

$$\frac{\sqrt{624} (0.4778)}{47.4444} = 0.2516 < 2$$

y en consecuencia se tiene media cero.

Para determinar la significación estadística individual de las autocorrelaciones de los residuales se calcula $| r_k(a) | \geq .08$, se dirá que la autocorrelación k-ésima es significativamente distinta de cero. Sin embargo esta prueba de significación no es suficiente y se realiza una prueba conjunta.

Se calcula el estadístico Q de Box y Pierce para realizar la prueba de significación conjunta. de un total de k autocorrelaciones simultáneamente, en nuestra caso se escoge k=24, por lo que Q sigue aproximadamente una distribución Ji-cuadrada con k-r grados de libertad, donde r es el número de parámetros involucrados. Es decir, el valor de Q calculado será comparado con valores de tablas de Ji-cuadrada con los correspondientes grados de libertad, debiendo ser recomendablemente menor.

Para el modelo 3.11 se tienen 22 grados de libertad y un

valor de: $Q = 15.4$ para la autocorrelación número 24.

El valor de Ji-cuadrada con 22 grados de libertad al 0.05, es de 33.92 .

$\{ a_t \}$ tiene distribución normal, para toda t .

Sí se cumple que la media de los residuales sea cero, se espera que a lo más un total de $(623)/20$ observaciones se localizaran fuera del intervalo $(-2 \hat{\sigma}_a, 2 \hat{\sigma}_a)$.

Debe verificarse también que no existan observaciones aberrantes, o ajenas a la serie de estudio.. Por ejemplo un residual que se encuentre fuera de $(-3 \hat{\sigma}_a, 3 \hat{\sigma}_a)$ implicará que: sucedió un evento cuya probabilidad de ocurrencia era de aproximadamente 0.2% o el residual en cuestión corresponde a una observación que no fué generada por el mismo proceso que generó el resto de la serie.

Existen otros elementos que considerar para verificar el modelo, y que es bueno mencionar:

-El modelo debe ser parsimonioso; es decir, que no podemos reducir el número de parámetros involucrados, puesto que todos son necesarios para explicar el comportamiento del fenómeno. Por lo tanto construimos intervalos de confianza de aproximadamente el 95% del tipo:

$$(\hat{\theta} - 2(\text{s.e.}), \hat{\theta} + 2(\text{s.e.})).$$

donde s.e. es el error estandar calculado al parámetro, por lo que resulta que:

sus intervalos de confianza para θ_1 y θ_2 respectivamente son:

(-0.1697, -0.0101)

(0.9316, 0.9744)

y observamos si los valores cero ó uno se encuentran dentro del intervalo. Pueden existir situaciones en las que el conocimiento del fenómeno sea indicativo de que el parámetro deba aparecer, aun cuando la verificación no rechace la hipótesis de que su valor sea cero.

-El modelo es estable en los parámetros.

Esto indica que debemos cuidar las correlaciones altas (> de 0.5 positivas o negativas) entre parámetros estimados, podrían ser causantes de posible inestabilidad. Debemos entonces calcular las correlaciones entre parejas de parámetros estimados. En caso de existir problema podemos cancelar alguno de los parámetros o bien eliminando factores aproximadamente comunes en los polinomios AR y MA, sin embargo si la inclusión es necesaria para obtener una representación adecuada de la serie y no se encuentran factores aproximadamente comunes, se aceptará el hecho de que existan correlaciones altas.

El modelo que sea aceptado para realizar el pronóstico deberá violar la menor cantidad de supuestos. Por lo que el modelo propuesto 3.1.1 está cumpliendo con los supuestos y se acepta para realizar el pronóstico.

Sin embargo, si analizamos de manera semejante a 3.1.1 el modelo siguiente:

$$(1 - \phi_1 B) \nabla_{12} Z_t = (1 - \theta_2 B^{12}) a_t. \quad 3.1.2$$

observamos que todos los supuestos sobre el modelo se cumplen de igual

manera:

La estimación de los parámetros nos indica que:

$$\hat{\theta}_1 = 0.9525 \quad \text{con un error estandar de: } 0.0110.$$

$$\hat{\phi}_1 = 0.0880 \quad \text{con error estandar de: } 0.0399.$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

$$(0.9305, 0.9745)$$

$$(0.0082, 0.1678)$$

la matriz de correlación no indica correlaciones altas entre los parámetros.

Al realizar el análisis de los residuales para el modelo 3.1.2 se tiene que:

$$m(\hat{a}) = 0.6091 \quad \text{y} \quad \sigma(\hat{a}) = 47.4979.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) |$ resulta ser igual a $0.3201 < 2$.

No se presentan residuales individuales significativamente distintos de cero, como se aprecia en la figura 3.5.

El estadístico $Q = 15.2$, es menor que una $\chi^2_{\alpha}(22) = 33.92$.
con $\alpha = .050$.

Tampoco existen observaciones cuyos residuales sean estadísticamente distintos de cero. Por lo que el modelo 3.1.2 cumple

Coefficient

	-1.0	-.8	-.6	-.4	-.2	.0	.2	.4	.6	.8	1.0
	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										
1	.00					I					
2	-.01					+ I +					
3	.02					+ I +					
4	.05					+ IX+					
5	.04					+ IX+					
6	.02					+ I +					
7	.05					+ IX+					
8	.00					+ I +					
9	.07					+ IXX					
10	.00					+ I +					
11	.03					+ IX+					
12	-.03					+XI +					
13	-.03					+XI +					
14	.00					+ I +					
15	.04					+ IX+					
16	-.01					+ I +					
17	.01					+ I +					
18	-.02					+ I +					
19	.00					+ I +					
20	.04					+ IX+					
21	-.03					+XI +					
22	.03					+ IX+					
23	-.01					+ I +					
24	.06					+ IX+					
25	.01					+ I +					
26	.08					+ IXX					
27	-.05					+XI +					
28	-.03					+XI +					
29	.02					+ I +					
30	-.01					+ I +					
31	.03					+ IX+					
32	-.01					+ I +					
33	.06					+ IX+					
34	.02					+ I +					
35	.11					+ IX+X					
36	.02					+ I +					
37	.02					+ I +					
38	.00					+ I +					
39	.01					+ I +					
40	.00					+ I +					
41	.00					+ I +					
42	-.02					+ I +					
43	-.02					+ I +					
44	-.01					+ I +					
45	.00					+ I +					
46	.02					+ IX+					
47	-.07					XXI +					
48	.03					+ IX+					
49	.03					+ IX+					
50	-.01					+ I +					

Fig. 3.5 F.A.C. de los Residuales de 3.1.2

de igual manera que 3.1.1 con los supuestos, y por lo tanto debe ser considerado como un modelo representativo del proceso generador de la serie.

En general todas las etapas de verificación de modelos podriamos resumirlas en la tabla siguiente:

RESUMEN DE RESULTADOS DE LA ESTIMACIÓN DE MODELOS

SERIE $Z_t = \text{ISO1012}$

PERIODO = ENERO DE 1937 A DICIEMBRE DE 1989.

NUMERO DE OBSERVACIONES = 636 OBSERVACIONES (MENSUALES).

MODELO	PARAMETROS	INTERVALOS DE CONFIANZA	PARAMETROS CORREL.
$\nabla_{12} Z_t =$ $(1-\theta_1 B)(1-\theta_2 B^{12}) a_t$	$\hat{\theta}_1 = -.0899$ $\hat{\theta}_2 = .9530$	$(-.1697, -.0101)$ $(.9316, .9744)$	
$(1-\phi_1 B) \nabla_{12} Z_t =$ $(1-\theta_2 B^{12}) a_t$	$\hat{\theta}_2 = .9525$ $\hat{\phi}_1 = .0880$	$(.9305, .9745)$ $(.0082, .1678)$	

ANÁLISIS DE RESIDUALES				
$\hat{m}(a)$	$\hat{\sigma}(a)$	COCIENTE	$Q =$ G.L. =	$r_k(a) \neq 0$
0.4778	47.4444	0.2516	15.4 22	
0.6091	47.4979	0.3201	15.2 22	

Como podemos observar ambos modelos, 3.1.1 y 3.1.2 cumplen satisfactoriamente los supuestos sobre la estimación de modelos. Sin embargo, veamos nuevamente los modelos que se tienen, para decidir con que modelo resulta más adecuado para realizar el pronóstico:

$$\nabla_{12} Z_t = (1 - \theta_1 B) (1 - \theta_2 B^{12}) a_t. \quad 3.1.1$$

$$(1 - \phi_1 B) \nabla_{12} Z_t = (1 - \theta_2 B^{12}) a_t. \quad 3.1.2$$

El modelo 3.1.1 posee sólo la parte correspondiente a el promedio móvil (regular y estacional), mientras que el modelo 3.1.2 posee tanto la parte de promedio móvil (estacional) como autorregresiva (regular). Esto ayuda a una realizar una interpretación más adecuada del fenómeno, como consecuencia tanto de una situación aleatoria como de otra, "semideterminista" y no solamente de efectos puramente aleatorios como lo indica 3.1.1.

Así se tiene que la lluvia, para la isoyeta 1000 - 1200 mm. del estado de Morelos para un cierto mes del año, es consecuencia de

efectos aleatorios debidos a la lluvia del mes respectivo del año anterior $t-12$, así como a la lluvia del mes anterior al mes en cuestión $t-1$, y no solamente de efectos aleatorios debidos al mes respectivo del año anterior y al mes anterior del mismo año.

De ésta forma el modelo que resulta más adecuado para el realizar el pronóstico es 3.1.2, puesto que la construcción misma del modelo está íntimamente ligada con el pronóstico, como vimos al calcular los residuales se obtienen primero los valores estimados por el modelo, que no son otra cosa más que los valores pronosticados de la variable en el tiempo t , con base en las observaciones hasta el tiempo $t-1$, es decir, los pronósticos un periodo hacia adelante.

Sólo mencionaremos que el criterio para determinar la optimalidad del pronóstico será el del error cuadrático medio mínimo, es decir, $\hat{W}_t(h)$ deberá satisfacer:

$$E_t [W_{t+h} - \hat{W}_t(h)]^2 = \min_{\tilde{W}_t(h)} E_t [W_{t+h} - \tilde{W}_t(h)]^2.$$

donde $\tilde{W}_t(h)$ es un pronóstico cualquiera que se obtenga como combinación lineal de los valores de la serie $\{W_t\}$, y en consecuencia de los errores $\{a_t\}$, el pronóstico óptimo es $\hat{W}_t(h)$, E_t denota a la esperanza condicional, dada toda la información hasta el momento t .

El paquete estadístico S.C.A. arroja los siguientes pronósticos (En el Macro 1 se lleva a cabo la generación del pronóstico y se presenta en el Anexo 1):

Pronóstico de lluvia mensual acumulada para ISO1012, para el periodo Ene 1990 a Dic 1990.

Enero

11.57

Febrero	6.82
Marzo	9.68
Abril	19.60
Mayo	57.15
Junio	217.11
Julio	225.21
Agosto	223.29
Septiembre	230.22
Octubre	85.79
Noviembre	15.37
Diciembre	4.97

en la figura 3.6 se muestra la gráfica de los datos ajustados por el modelo 3.1.2 considerando sólo el periodo: Enero de 1985 a diciembre de 1989. En la figura 3.7 se muestran los datos ajustados por el modelo junto con el pronóstico y los datos históricos de la serie de ISO1012.

3.2. METODO DE WINTERS

Winters propuso dos métodos de suavizamiento exponencial para series de tiempo una componente estacional: un método aditivo y uno multiplicativo. Estos métodos difieren en su suposición de cómo la componente estacional afecta la serie de tiempo.

Así, el método aditivo supone que los datos siguen el modelo:

$$Z_{n+j} = T_{n+j} + S_{n+j} + a_{n+j}, \quad 3.2.1.$$

donde $T_{n+j} = \mu_n + \beta_n j$ es una componente de tendencia, S_{n+j} es un factor aditivo estacional, y μ_n y β_n son el nivel y slope de la serie al tiempo $t=n$. El periodo estacional para el modelo es s y existen s factores estacionales tales que:

- i) $S_1 = S_{1+s} = S_{1+2s} = \dots$ $i = 1, 2, \dots, s, \quad Y$
- ii) $S_1 + S_2 + \dots + S_s = 0$.

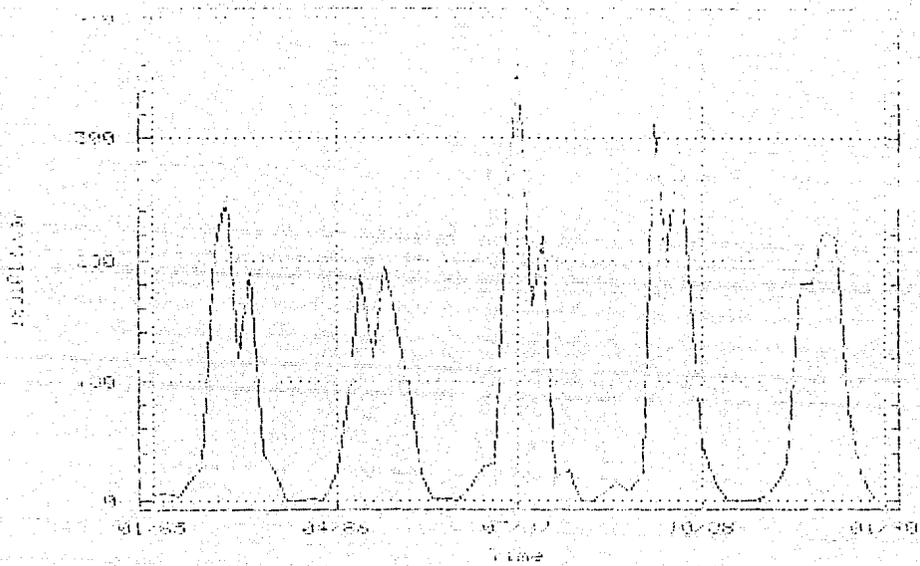


fig. 3.6

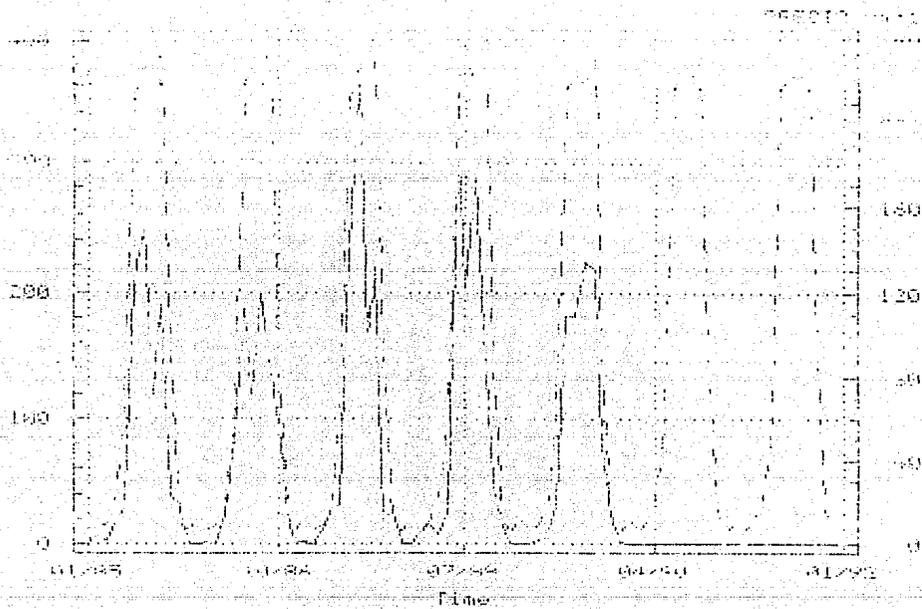


fig. 37

El método resulta apropiado para series de tiempo en las cuales la amplitud del efecto estacional no depende del nivel medio de la serie (Montgomery y Johnson, 1976). Como vimos el método de Winter es una extensión del método de dos parámetros de Holt en el cual en término estacional es incluido. Los pronósticos para el método aditivo involucran pesos actualizados de el nivel, la slope y los factores estacionales. Tres diferentes constantes de suavizamiento podrian ser empleadas para actualizar μ , β y los factores estacionales. Los pronósticos para observaciones futuras son:

$$\hat{Z}_n(l) = \hat{\mu}_n + \hat{\beta}_{n,l} + \hat{S}_{n+1-s} \quad l = 1, 2, \dots, s.$$

$$\hat{Z}_n(l) = \hat{\mu}_n + \hat{\beta}_{n,l} + \hat{S}_{n+1-2s} \quad l = s+1, s+2, \dots, 2s.$$

donde

$$\hat{\mu} = \alpha_1 (Z_n - \hat{S}_{n-s}) + (1-\alpha_1) (\hat{\mu}_{n-1} + \hat{\beta}_{n-1})$$

$$\hat{\beta} = \alpha_2 (\hat{\mu}_n - \hat{\mu}_{n-1}) + (1-\alpha_2) \hat{\beta}_{n-1}.$$

$$\hat{S}_n = \alpha_3 (Z_n - \hat{\mu}_n) + (1-\alpha_3) \hat{S}_{n-s}.$$

Calculo de $\hat{Z}_n(l)$.

Es necesario especificar tres constantes de suavizamiento ($\alpha_1, \alpha_2, \alpha_3$) en el calculo del l-ésimo paso del pronóstico de $\hat{Z}_n(l)$. La estimación de los otros parámetros son calculados internamente por el sistema.

Relación con un modelo ARIMA.

Los pronósticos derivados del método de suavizamiento exponencial aditivo de Winters son equivalentes a aquellos de un modelo ARIMA:

$$(1-B)(1-B^E) Z_t = (1-\alpha_1 B - \alpha_2 B^2 - \dots - \alpha_E B^E - \alpha_{E+1} B^{E+1}) a_t. \quad 3.2.3$$

donde

$$\begin{aligned} \alpha_1 &= 1 - \alpha_1(1+\alpha_2) \\ \alpha_j &= -\alpha_1 \alpha_2 \quad j=2,3,\dots,E-1. \\ \alpha_E &= (1-\alpha_3) - \alpha_1(\alpha_2-\alpha_3), \\ \alpha_{E+1} &= -(1-\alpha_1)(1-\alpha_3) \quad \dots 3.2.4. \end{aligned}$$

con $\alpha_1, \alpha_2, \alpha_3$ las constantes de suavizamiento.

El método multiplicativo de Winters, análogo al método de Winters aditivo, asume que la serie de tiempo sigue un modelo:

$$Z_{n+j} = (\mu_n + \beta_n j) S_{n+j} + a_{n+j}, \quad 3.2.5$$

donde $(\mu_n + \beta_n j)$ es una componente de tendencia, S_{n+j} es un factor multiplicativo estacional, y μ_n y β_n son el nivel y la slope de la serie al tiempo $t=n$. El modelo multiplicativo es usualmente aplicado y se adecua a series de tiempo en las cuales la amplitud del patron estacional es proporcional a el nivel de la serie. Por lo que se usa el modelo aditivo estacional de Winters para realizar el pronóstico de la serie ISO1215 del estado de Morelos.

En el Anexo 1 se presenta el Macro3 usado para realizar el pronóstico mediante Winters aditivo, también através del sistema S.C.A.

Se tienen los siguientes resultados:

Enero	5.80	Julio	211.08
Febrero	4.47	Agosto	207.45
Marzo	3.92	Septiembre	204.51
Abril	7.12	Octubre	69.04
Mayo	34.96	Noviembre	25.00
Junio	210.41	Diciembre	4.05

El pronóstico resulta muy sensible al valor de las constantes de suavizamiento las cuales fueron obtenidas usando las ecuaciones 3.2.4 y tomando la ventaja de la relación con los modelos ARIMA. El pronóstico arroja un error medio estandar $MSE = 3389.3$.

3.3 CONCLUSIONES

El análisis de Box-Jenkins ofrece una metodología mucho más completa de la realización del pronóstico comparada al método de Winters, se tiene por una parte, mayor variedad de posibles interpretaciones para el pronóstico de la lluvia arrojado por el modelo. El método de Winters incurre en errores muy significativos al variar las constantes de suavizamiento del modelo. Es decir, el Método de Winters resulta sumamente sensible a los cambios de valor de las constantes.

Desde el punto de vista computacional; resulta comodo realizar un pronóstico usando el sistema S.C.A. y la metodología Box-Jenkins.

Sin embargo, Box-Jenkins incurre en deficiencias para realizar un pronóstico a periodos estacionales grandes (a más de tres periodos estacionales los pronósticos vienen a ser prácticamente los mismos), así, resulta adecuado para realizar pronóstico de uno a dos periodos estacionales, se recomienda la actualización con datos reales para ejecuciones posteriores. Por éstas razones la metodología seleccionada para efectuar el pronóstico de la lluvia para el estado de Morelos resulta ser Box-Jenkins ya implantada en el sistema de computo estadístico S.C.A. y adecuado a la información que se posee.

PARTE IV

IMPLANTACION DEL MODELO

PRONOSTICO PARA EL ESTADO DE MORELOS

En éste capítulo se implanta un modelo de pronóstico haciendo uso de la metodología Box-Jenkins, que resultó más adecuada para tratar el tipo de Series de Tiempo de lluvia que se presentan en el Estado de Morelos.

Se ejemplifica la metodología para la construcción de modelos de Series de Tiempo desarrollada por Box-Jenkins, haciendo uso de la información correspondiente a todas las isoyetas que se presentan en el Estado de Morelos.

Todos los modelos fueron obtenidos de manera análoga siguiendo la estrategia de construcción de modelos para series de tiempo desarrollada por Box-Jenkins (1970), y presentada en detalle en la parte 3.1.

El propósito es seguir la metodología de manera sistemática y realizar el pronóstico haciendo uso de las ventajas de calculo que provee el paquete de cómputo estadístico S.C.A. de forma tal que, usando Box-Jenkins, realizar el pronóstico sea accesible al usuario.

4.1 Pronóstico de la Isoyeta 800 - 1000 mm.

- i) Identificación de un posible modelo dentro de la clase de modelos ARIMA $(p,d,q) \times (P,D,Q)$.

La figura 4.1 muestra los datos de la serie de precipitación pluvial mensual acumulada del periodo Enero de 1937 a Diciembre de 1989 de la isoyeta 800 - 1000 mm., denotada ISO-810. En la figura 4.2 se presenta la gráfica de los datos observados . En la figura 4.3 se presentan las graficas de la FAC muestral de la Serie y de algunas de sus diferencias; para determinar el grado de diferenciación necesario para volver estacionaria la serie; también se calculan las desviaciones estandar muestrales para la serie diferenciada tanto regular como estacionalmente.

Las desviaciones estándar de las diferencias indicadas fueron las siguientes:

S (0,0)	=	85.7098
S (1,0)	=	74.0941
S (2,0)	=	107.9951
S (0,1)	=	53.0348 *
S (1,1)	=	73.6421
S (2,1)	=	127.1380
S (0,2)	=	93.5158
S (1,2)	=	129.2799
S (2,2)	=	222.8723

Se decide, por lo tanto, que las diferencias requeridas son: una estacional de orden 12; es decir, la serie aproximadamente estacionaria que se analizará resulta ser:

$$W_t = \nabla_{12} Z_t .$$

se tiene:

$$\nabla_{12} Z_t = Z_t - Z_{t-12} .$$

al aplicar una diferencia estacional a la serie de precipitación

ISCIYETA 500 - 1000 aa.

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	ANO
5.17	1.56	7.94	17.90	106.71	216.29	191.87	136.09	222.85	58.17	4.12	3.05	1937
.00	5.04	2.21	2.50	62.76	136.41	136.28	177.22	133.71	54.59	20.84	.10	
.00	.00	2.59	28.78	70.75	175.06	175.86	134.46	100.25	167.75	.00	1.46	
1.96	.62	37.66	.00	56.59	235.15	273.55	181.82	71.78	11.77	4.73	4.71	
.73	6.73	3.57	7.13	30.26	265.7	126.71	115.65	76.35	161.17	16.27	10.50	
5.16	4.61	.73	.00	40.17	225.70	165.78	155.88	206.72	29.24	28.42	.00	
4.75	.00	2.85	4.22	42.71	377.21	85.12	169.24	277.57	55.17	7.93	12.85	
.00	.00	1.24	.10	22.50	167.60	154.05	202.98	196.56	33.02	4.25	.00	
.00	.00	.98	7.11	54.30	71.43	111.31	247.63	157.10	55.47	4.09	.00	
5.90	.00	7.57	52.74	57.53	210.35	168.50	157.20	93.96	79.89	14.61	.00	
7.21	.52	.00	10.15	164.87	197.14	123.60	204.24	158.50	105.59	.15	17.52	
24.08	.00	.00	22.52	100.18	171.59	185.12	145.91	118.91	76.14	70.34	.00	
.00	.00	.00	.00	61.07	217.87	137.12	77.70	115.15	23.47	.67	.00	
4.72	.57	2.16	2.67	120.19	124.29	177.66	112.80	168.79	55.71	.25	.00	
.00	.00	10.89	5.52	75.23	139.56	153.25	155.81	221.38	66.64	4.24	.00	
7.10	.60	.00	21.23	135.41	190.06	219.34	243.02	202.34	2.84	34.32	.00	
.00	.00	.00	.39	19.22	183.07	176.43	162.60	82.94	64.95	11.54	.23	
1.70	9.79	7.38	36.98	115.44	242.16	172.78	142.59	238.96	100.57	11.39	.00	
1.47	.30	3.93	.12	17.42	157.72	256.79	247.62	320.17	95.87	9.68	8.68	
.00	6.21	1.03	22.10	157.61	258.04	178.54	188.75	195.31	19.58	2.56	.43	
.00	3.25	.00	7.83	29.75	201.61	76.95	173.41	147.84	36.53	5.32	.00	
81.37	9.87	.00	17.70	37.63	228.67	267.67	184.76	277.55	136.86	116.45	22.25	
9.14	.00	.00	29.84	40.76	253.61	189.97	232.34	114.12	148.51	2.76	.00	
6.90	.00	.00	.50	58.56	122.66	239.15	161.92	156.88	85.45	.00	.52	
1.48	.23	1.46	1.36	47.69	236.98	232.85	143.60	196.28	34.16	46.05	.99	
.00	.10	.29	18.60	41.65	122.62	131.77	189.79	198.08	58.20	5.79	.33	
.00	.00	16.81	4.75	92.71	202.21	285.23	126.20	139.58	91.53	1.49	.52	
18.49	.00	.50	2.28	142.10	203.16	134.17	86.10	245.86	40.31	22.26	7.40	
21.45	12.62	.00	32.59	56.38	158.92	151.75	278.57	124.97	101.17	8.21	3.88	
1.28	1.43	15.78	56.38	87.88	194.25	175.84	148.28	173.80	93.89	.03	1.24	
57.23	.00	1.51	8.79	54.68	197.09	114.80	241.62	205.49	91.57	10.46	2.11	
3.14	10.36	2.15	32.56	59.12	245.29	99.56	130.84	119.77	65.18	1.43	28.81	
17.75	.02	14.61	.87	10.97	77.68	166.57	365.78	217.51	49.26	.01	.00	
.01	.04	.00	.60	7.98	218.41	154.42	109.24	222.48	44.04	6.65	.00	
4.45	.00	3.63	8.48	71.51	189.31	136.42	178.16	237.16	121.30	13.08	5.57	
.00	.00	11.89	29.91	61.95	241.73	160.85	84.57	173.20	40.75	34.25	.35	
.00	1.25	.00	28.98	53.58	187.86	205.85	245.75	178.60	108.07	3.84	.29	
1.25	.74	8.05	2.76	56.64	237.30	208.08	161.00	161.37	8.83	3.69	.00	
10.14	.46	.00	1.81	131.20	230.64	162.69	199.21	129.94	40.65	1.03	.00	
.04	2.83	4.08	29.06	31.65	139.30	268.80	178.80	136.92	143.16	16.03	7.20	
4.16	5.65	.02	17.07	106.47	247.51	114.44	242.83	122.62	53.67	14.76	1.27	
3.41	1.85	24.58	2.85	31.09	224.66	130.00	143.01	216.40	111.65	13.06	9.17	
.00	5.60	.26	7.41	84.05	87.84	234.73	137.36	161.93	34.73	1.57	12.85	
84.53	.00	.00	13.44	75.94	155.43	122.31	261.00	156.57	61.87	8.35	.00	
17.78	10.73	13.57	23.97	43.76	317.00	278.34	225.29	172.78	81.43	4.88	.10	
.00	10.64	2.17	11.04	72.94	174.78	111.62	77.95	141.52	65.54	3.63	.33	
12.06	18.27	3.58	.28	20.81	154.56	192.50	157.61	205.50	16.69	9.35	.10	
15.72	6.14	.88	.00	79.35	230.74	188.44	165.51	198.85	34.80	.33	.11	
.20	1.80	4.85	15.33	61.18	317.43	213.52	157.37	182.01	55.82	15.34	1.09	
.02	1.20	.03	29.87	94.31	223.87	89.52	147.65	120.14	68.58	7.40	3.08	
.00	2.25	4.41	11.36	28.65	162.84	235.49	134.37	224.73	9.48	23.27	.30	
.00	.02	10.70	1.09	11.97	225.55	176.08	177.57	130.69	36.88	2.43	.08	
.00	.00	.58	1.19	19.27	179.96	145.10	252.47	158.65	85.16	8.18	.96	1989

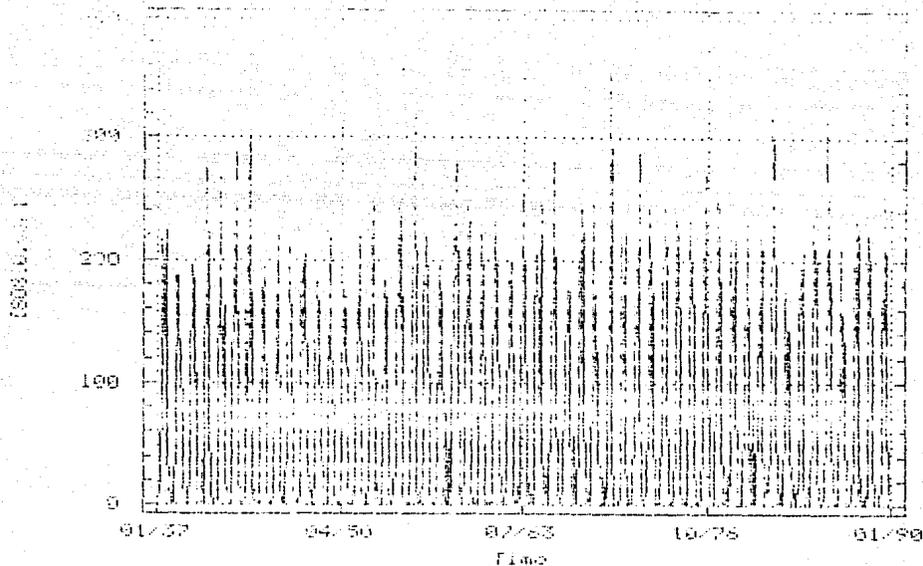


Fig. 4.2

$Z_t = 17$

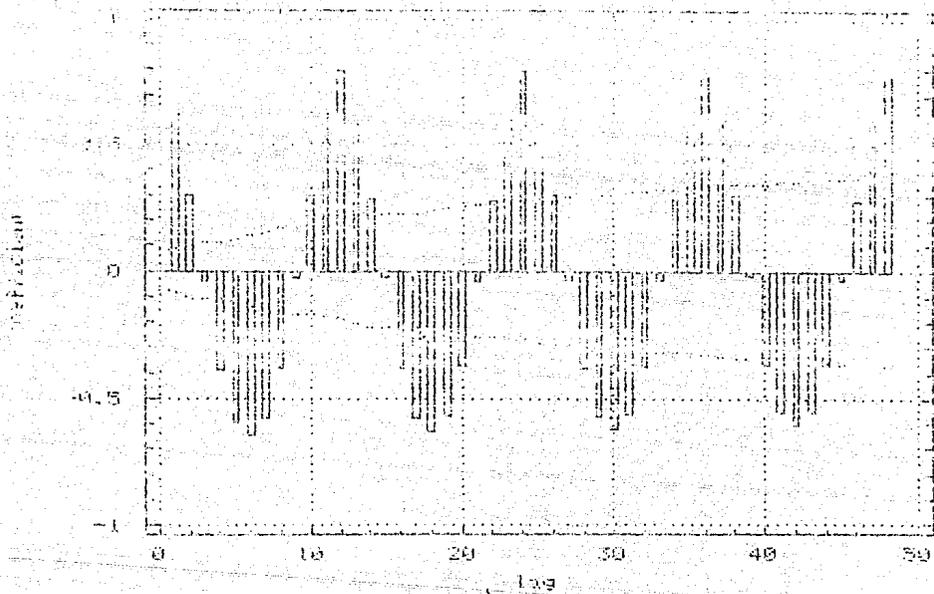


fig. 4.3 $D(0,1) Z_t$

pluvial mensual acumulada, se obtiene la serie de flujos anuales de precipitación. Puesto que la serie Z_t de referencia es una serie mensual que cubre el período Enero de 1937 a Diciembre de 1989, por lo tanto consta de 636 observaciones originales, tal número de observaciones se reduce a 624 al aplicar el operador diferencia ∇_{12} .

La gráfica de las autocorrelaciones de la figura 4.3 d) es característica de un modelo multiplicativo estacional:

ARIMA (0,0,0)X(0,1,1)

4.1.1 Primer modelo propuesto.

$$\nabla_{12} Z_t = (1 - \theta_2 B^{12}) a_t. \quad 4.1.1$$

ii) Estimación de los parámetros involucrados en el modelo.

Una vez propuesto el modelo 4.1.1 se realiza la estimación de los parámetros que están involucrados.

El valor estimado del parámetro resultó ser:

$$\hat{\theta}_2 = 0.9265$$

su intervalo de confianza es:

$$(0.8987, 0.9543)$$

iii) Verificación de que el modelo proporciona un ajuste adecuado y que los supuestos básicos se satisfacen. En caso de que los supuestos no se cumplan, se determinan las modificaciones necesarias al modelo.

En la gráfica de residuales que se presenta en la figura 4.4 para el modelo 4.1.1 podemos apreciar que ellos son estadísticamente cero por lo tanto el modelo es prácticamente aceptable.

Estimated Autocorrelations

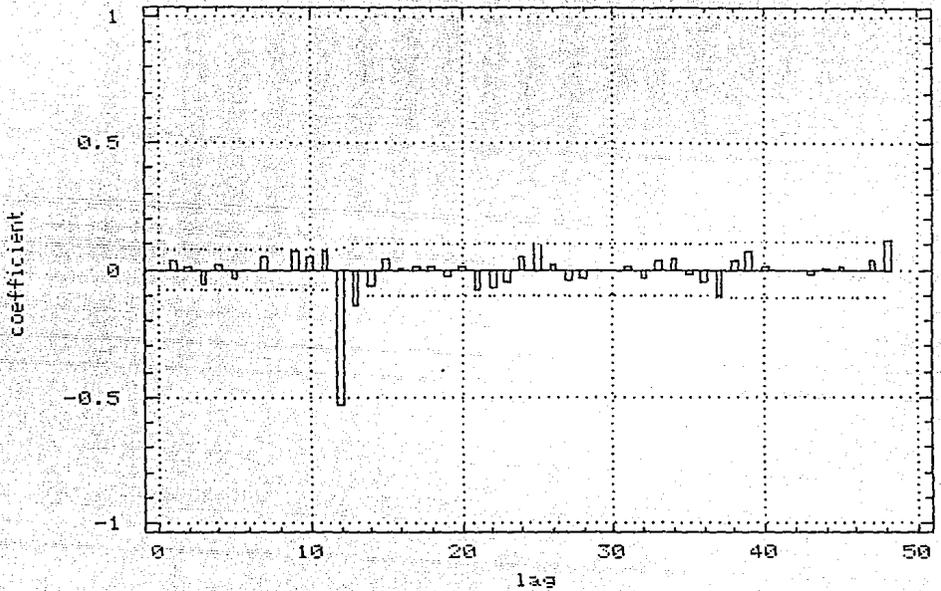


fig. 4.3 d) FAC de W_t ISO810

{ a_t } tiene media cero.

Una vez teniendo calculada la media y la desviación estándar muestral de los residuales se checa la hipótesis:

Si $| (n)^{1/2} m(a) / \sigma(a) | < 2$, se dice que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero y por lo tanto no se rechaza el supuesto.

Así que se tiene:

$$m(\hat{a}) = 1.8863$$

$$\sigma(\hat{a}) = 37.9096.$$

$$\frac{\sqrt{624} (1.8863)}{37.9096} = 1.2429$$

por lo tanto se considera con media cero.

Para determinar la significación estadística individual de las autocorrelaciones de los residuales se calcula $| r_k(a) | \geq .08$. y se tiene $r_{13}(a) = .13$. (Problema No. 1).

Sin embargo esta prueba de significación no es suficiente y se realiza una prueba conjunta.

Se calcula el estadístico Q de Box y Pierce de un total de k autocorrelaciones simultaneamente, en nuestra caso se escoge k=24, por lo que Q sigue aproximadamente una distribución Ji-cuadrada con k-r grados de libertad, donde r es el número de parámetros involucrados. El valor de Q calculado será comparado con valores de tablas de Ji-cuadrada con los correspondientes grados de libertad, debiendo ser menor.

Para el modelo 4.1.1 se tienen 23 grados de libertad y un valor de: $Q = 25.2$ para la autocorrelación número 24.

El valor de Ji-cuadrada con 23 grados de libertad al 0.05, es de 35.17.

$\{ a_t \}$ tiene distribución normal, para toda t .

Sí se cumple que la media de los residuales sea cero, se espera que a lo más un total de $(623)/20$ observaciones se localizaran fuera del intervalo $(-2 \hat{\sigma}_a, 2 \hat{\sigma}_a)$.

No se encuentran residuales fuera de $(-3 \hat{\sigma}_a, 3 \hat{\sigma}_a)$, es decir, no se tienen observaciones aberrantes, observaciones que no fueron generadas por el mismo proceso que generó el resto de la serie.

Los otros elementos a considerar para verificar el modelo, y que siempre es bueno revisar, los podemos enlistar rápidamente:

-El modelo debe ser parsimonioso; es decir, que no podemos reducir el número de parámetros involucrados, puesto que todos son necesarios para explicar el comportamiento del fenómeno.

su intervalo de confianza de θ_2 es:

(0.8987, 0.9543)

y observamos sí los valores cero ó uno se encuentran dentro del intervalo. Pueden existir situaciones en las que el conocimiento del fenómeno sea indicativo de que el parámetro deba aparecer, aun cuando la verificación no rechace la hipótesis de que su valor sea cero.

-El modelo es estable en los parámetros.

Cuidado de las correlaciones altas (> de 0.5 positivas o negativas) entre parámetros estimados, podrían ser causantes de posible inestabilidad. En este caso no se tiene correlación alta entre los parámetros.

4.1.2. Segundo Modelo propuesto.

$$v_{12} z_t = (1 - \theta_1 B^{12}) (1 - \theta_2 B^{13}) a_t. \quad 4.1.2$$

observamos que todos los supuestos sobre el modelo se cumplen de igual manera:

La estimación de los parámetros nos indica que:

$$\hat{\theta}_1 = 0.9273 \quad \text{error estandar : } 0.0146.$$

$$\hat{\theta}_2 = 0.1180 \quad \text{error estandar : } 0.0400.$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

$$(0.8981, 0.9565)$$

$$(0.0380, 0.1580)$$

la matriz de correlación no indica correlaciones altas entre los parámetros.

Análisis de los residuales para el modelo 4.1.2 se tiene que:

$$m(\hat{a}) = -2.1448 \quad \text{y} \quad \sigma(\hat{a}) = 37.6153.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) | = 1.4243 < 2.$

No se presentan residuales individuales significativamente distintos de cero.

El estadístico $Q = 14.8$, es menor que una $\chi^2_{\alpha}(22) = 33.92$.
 con $\alpha = 0.050$.

Tampoco existen observaciones cuyos residuales sean estadísticamente distintos de cero. Por lo que el modelo 4.1.2 cumple con los supuestos, y por lo tanto debe ser considerado como un modelo representativo del proceso generador de la serie.

4.1.3. Tercer Modelo propuesto.

$$(1 - \phi_1 B^{12}) \nabla_{12} Z_t = (1 - \theta_1 B^{12}) (1 - \theta_2 B^{13}) a_t. \quad 4.1.3$$

observamos que todos los supuestos sobre el modelo se cumplen de igual manera:

La estimación de los parámetros nos indica que:

$$\hat{\phi}_1 = -0.1106 \quad \text{error estandar : } 0.0413.$$

$$\hat{\theta}_1 = 0.9259 \quad \text{error estandar : } 0.0149.$$

$$\hat{\theta}_2 = 0.0967 \quad \text{error estandar : } 0.0407.$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

$$(-0.1932, -0.0280)$$

$$(0.8961, 0.9557)$$

$$(0.0153, 0.1781)$$

La matriz de correlación no indica correlaciones altas entre los parámetros.

Análisis de los residuales para el modelo 4.1.3.

se tiene:

$$m(\hat{a}) = -2.1508 \quad \text{y} \quad \sigma(\hat{a}) = 37.3353.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) | = 1.4250 < 2$.

No se presentan residuales individuales significativamente distintos de cero, como se aprecia en la figura (4.5).

El estadístico $Q = 13.9$, es menor que una $\chi^2_{\alpha}(21) = 32.67$.
con $\alpha = 0.050$.

Tampoco existen observaciones cuyos residuales sean estadísticamente distintos de cero. Por lo que el modelo 4.1.3 viola algunos supuestos, y por lo tanto no puede ser considerado como un modelo representativo del proceso generador de la serie.

4.1.4 Cuarto Modelo Propuesto.

$$(1 - \phi_1 B^{13}) \nabla_{12} Z_t = (1 - \theta_1 B^{12}) a_t. \quad 4.1.4.$$

La estimación de los parámetros nos indica que:

$$\hat{\phi}_1 = -0.1157 \quad \text{error estandar} : 0.0399.$$

$$\hat{\theta}_1 = 0.9322 \quad \text{error estandar} : 0.0141.$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

$$\begin{aligned} & (-0.1955, -0.0359) \\ & (0.9040, 0.9604) \end{aligned}$$

La matriz de correlación no indica correlaciones altas entre los parámetros.

Análisis de los residuales para el modelo 4.1.4.

se tiene:

$$m(\hat{a}) = 2.1317 \quad \text{y} \quad \sigma(\hat{a}) = 37.5232.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) | = 1.4043 < 2.$

No se presentan residuales individuales significativamente distintos de cero, como ~~se aprecia en la figura~~ ().

El estadístico $Q = 18.8$, es menor que una $\chi^2_{\alpha}(22) = 33.92$.
con $\alpha = 0.050$.

Tampoco existen observaciones cuyos residuales sean estadísticamente distintos de cero. Por lo que el modelo 4.1.4 no viola supuestos, por lo que puede ser considerado como un modelo representativo del proceso generador de la serie.

En general todas las etapas de verificación de modelos podríamos resumirlas en la tabla siguiente:

RESUMEN DE RESULTADOS DE LA ESTIMACIÓN DE MODELOS

SERIE $Z_t = \text{ISO810}$.

PERIODO = ENERO DE 1937 A DICIEMBRE DE 1989.

NUMERO DE OBSERVACIONES = 636 OBSERVACIONES (MENSUALES).

MODELO	PARAMETROS	INTERVALOS DE CONFIANZA	PARAMETROS CORREL.
$\nabla_{12} Z_t = (1-\theta_1 B^{12}) a_t$	$\hat{\theta}_1 = 0.9265$	(.8987, .9543)	$r_{13} = 0.13$
$\nabla_{12} Z_t = (1-\theta_1 B^{12}) (1-\theta_2 B^{12}) a_t$	$\hat{\theta}_1 = .9273$ $\hat{\theta}_2 = .1180$	(.8981, .9565) (.0380, .1580)	
$(1-\phi_1 B^{12}) \nabla_{12} Z_t = (1-\theta_1 B^{12}) (1-\theta_2 B^{13}) a_t$	$\hat{\phi}_1 = -.1106$ $\hat{\theta}_1 = 0.9259$ $\hat{\theta}_2 = 0.0967$	(-.1932, -.0280) (.8961, .9557) (.0153, .1781)	
$(1-\phi_1 B^{13}) \nabla_{12} Z_t = (1-\theta_1 B^{12}) a_t$	$\hat{\phi}_1 = -0.1157$ $\hat{\theta}_1 = 0.9322$	(-.1955, -.0359) (.9040, .9604)	

ANÁLISIS DE RESIDUALES

$\hat{m}(a)$	$\hat{\sigma}(a)$	COCIENTE	$Q =$ C.L. =	$r_k(a) \neq 0$
-1.8863	37.9096	1.2429	25.2 23	
-2.1448	37.6153	1.4243	14.8 22	
2.1506	37.3353	1.4250	13.9 21	
2.1317	37.5232	1.4043	18.8 22	

Observemos que el modelo 4.1.2 es un modelo parecido al modelo propuesto 3.1 para ISO1012, sólo que introduciendo el factor $(1-\theta_2 B^{13})$ (debido a que la autocorrelación 13 resulta ser estadísticamente distinta de cero), así que por las mismas razones no se considera.

Los siguientes modelos, 4.1.3 y 4.1.4 cumplen satisfactoriamente los supuestos sobre la estimación de modelos. Sin embargo, veámoslos nuevamente, para decidir que modelo resulta más adecuado para realizar el pronóstico:

$$(1-\phi_1 B^{12}) \nabla_{12} Z_t = (1-\theta_1 B^{12}) (1-\theta_2 B^{13}) a_t. \quad 4.1.3$$

$$(1-\phi_1 B^{13}) \nabla_{12} Z_t = (1-\theta_2 B^{12}) a_t. \quad 4.1.4$$

Si observamos las tablas de resultados podemos darnos cuenta

de varios factores por los que el modelo 4.1.3 resulta más conveniente que el modelo 4.1.4 para efectuar el pronóstico:

4.1.3 tiene menor $\hat{\sigma}_a$ y Q que el modelo 4.1.4. Además de que con el fin de proporcionar una explicación para el pronóstico de la lluvia en el modelo 4.1.3 proporciona una explicación más adecuada. Así, la lluvia para la isoyeta 800 - 1000 mm. de un cierto mes t, es debida a la lluvia del mismo mes del año anterior t-12 así como a efectos aleatorios más un efecto aleatorio del mes anterior al mes t-13 correspondiente del año pasado.

De ésta forma el modelo que resulta más adecuado para el realizar el pronóstico es 4.1.3.

Nuevamente se tienen los pronósticos.

Pronóstico de lluvia mensual acumulada para ISO810, para el periodo Ene 1990 - Dic 1990.

Enero	9.57	Julio	181.39
Febrero	5.05	Agosto	170.23
Marzo	5.51	Septiembre	164.98
Abril	13.21	Octubre	58.99
Mayo	62.04	Noviembre	8.42
Junio	209.86	Diciembre	3.08

En la figura 4.4 se presenta la gráfica de los datos ajustados a la serie ISO810 mostrando sólo el periodo Enero de 1985 a Diciembre de 1989.

En la figura 4.5 se presenta la gráfica de los datos ajustados con el pronóstico para 1990 y la serie ISO810 de datos históricos, mostrando sólo el periodo enero de 1985 a diciembre de 1990 (para la serie ajustada y su pronóstico)

4.2 PRONOSTICO DE LA ISOYETA 1500 - 2000 mm .

Figure 4.4: Continuous Plot

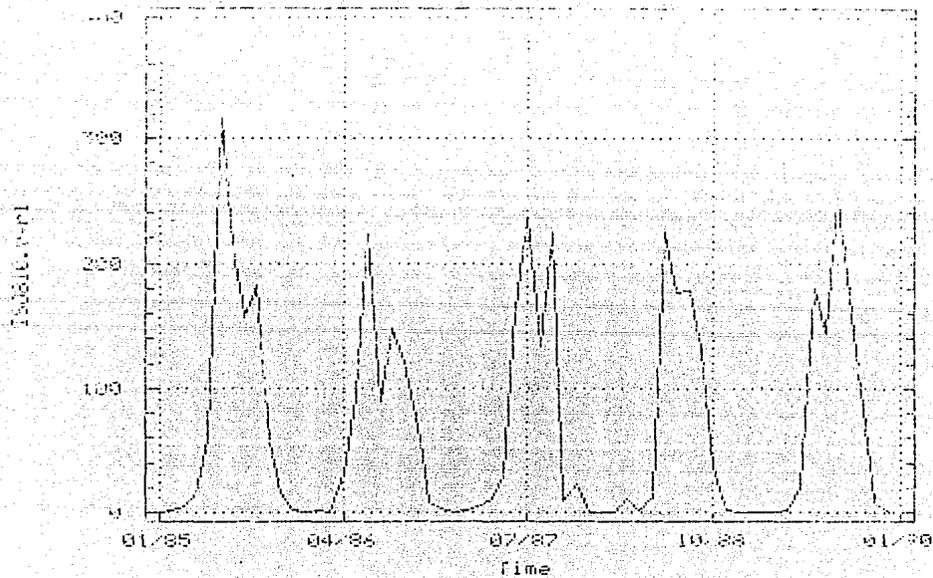


Fig. 4.4

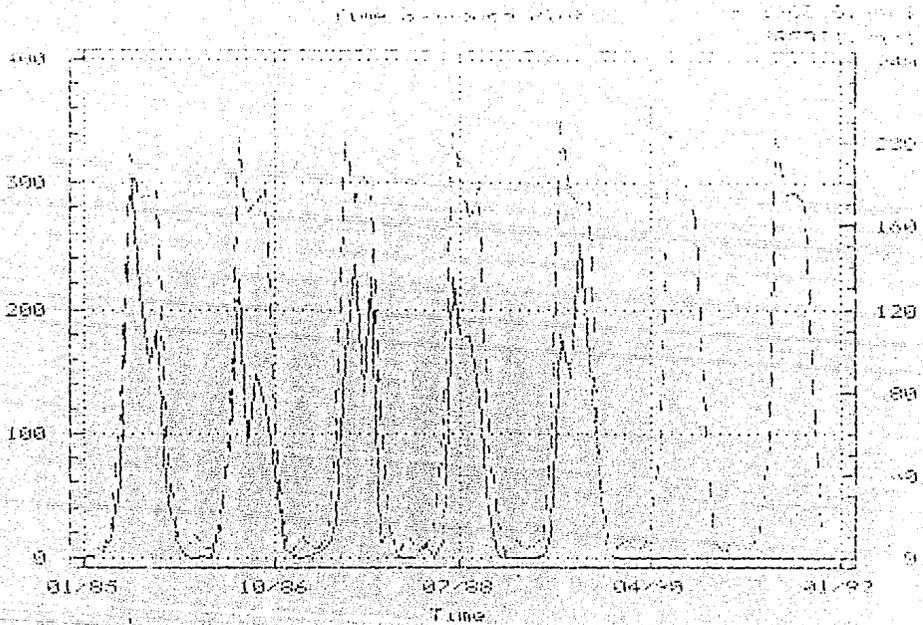


fig. 4.5

- i) Identificación de un posible modelo dentro de la clase de modelos ARIMA (p,d,q) x (P,D,Q).

La tabla 4.8 muestra los datos de la serie de precipitación pluvial mensual acumulada del periodo Enero de 1937 a Diciembre de 1989 de la isoyeta 1500 - 2000 mm., denotada ISO1520. En la figura 4.9 se presenta la gráfica de los datos observados. En la figura 4.10 se presentan las graficas de la FAC muestral de la con UNA DIFERENCIA ESTACIONAL ; para determinar el grado de diferenciación necesario y volver estacionaria la serie; también se calculan las desviaciones estandar muestrales para la serie diferenciada tanto regular como estacionalmente.

Las desviaciones estándar de las diferencias indicadas fueron las siguientes:

S (0,0)	=	168.3066
S (1,0)	=	140.1971
S (2,0)	=	197.1794
S (0,1)	=	111.5636 *
S (1,1)	=	148.0685
S (2,1)	=	257.0472
S (0,2)	=	192.0573
S (1,2)	=	255.1143
S (2,2)	=	443.2477

Se decide, por lo tanto, que las diferencias requeridas son: una estacional de orden 12; es decir, la serie aproximadamente estacionaria que se analizará resulta ser:

$$W_t = \nabla_{12} Z_t .$$

como antes al aplicar una diferencia estacional a la serie de precipitación pluvial mensual acumulada, se obtiene la serie de flujos anuales de precipitación. Puesto que la serie Z_t de referencia es una serie mensual que cubre el periodo Enero de 1937 a Diciembre de 1989,

Tabla 1.6

PROMEDIO DE PRECIPITACION PLUVIAL MENSUAL ACUMULADA
ISOYETA 1500 - 2000 m.

ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	AÑO
17.90	11.90	23.00	37.30	147.30	304.30	418.20	271.50	357.90	96.60	3.40	9.10	1937
.00	29.50	8.50	5.70	77.80	295.30	419.50	266.60	213.70	91.00	23.10	.00	
.00	.00	8.70	62.30	176.60	293.60	271.30	337.90	136.70	175.30	.00	4.90	
3.10	1.60	114.20	.00	77.80	371.20	617.80	382.60	118.70	20.40	4.90	12.70	
1.80	23.10	10.90	8.80	46.50	409.50	790.30	265.40	126.10	258.20	13.00	18.70	
7.40	17.40	1.80	.00	48.30	339.50	357.50	328.00	473.10	50.80	30.90	.00	
10.50	.00	9.60	12.10	59.40	431.10	202.60	334.10	516.10	74.70	4.40	37.00	
1.60	.00	3.00	.20	26.30	264.00	345.00	452.70	334.00	79.10	5.90	.00	
.00	.00	.00	37.50	88.90	117.20	206.10	547.20	292.20	89.40	6.60	.00	
.00	.00	18.40	161.40	58.50	344.10	408.60	329.00	156.00	110.50	8.90	.00	
24.70	2.30	.00	39.70	211.90	288.30	276.90	462.30	247.30	217.80	.00	43.10	
57.20	.00	.00	49.30	163.60	265.00	421.10	322.80	153.00	122.80	79.20	.00	
.00	.00	.00	.00	55.90	340.70	110.60	266.70	153.80	82.10	.00	.00	
.00	2.30	13.80	2.60	121.90	204.10	364.70	226.20	249.80	118.50	.20	.00	
.00	.00	42.50	15.90	103.30	274.40	452.90	308.40	440.90	65.10	5.50	.00	
22.30	.00	.00	47.10	166.90	304.30	369.50	383.60	330.40	1.40	54.60	.00	
.00	.00	1.50	.00	18.50	504.50	291.90	410.00	226.40	200.20	47.00	6.90	
.00	14.50	5.30	6.70	154.70	272.70	471.70	353.00	377.80	118.00	1.00	.00	
4.00	.00	14.90	1.10	39.20	320.40	610.00	428.00	517.70	180.30	7.90	20.80	
.00	38.90	1.70	63.30	248.40	363.10	391.60	225.60	323.90	33.00	1.30	.00	
.00	19.20	.00	5.20	51.00	279.80	200.90	352.20	292.30	64.60	5.50	.00	
240.20	43.30	.00	40.80	61.60	310.90	657.40	382.90	435.30	204.60	106.80	47.10	
25.20	.00	.00	71.00	38.30	382.10	482.10	615.30	198.70	272.70	3.30	.00	
20.90	.00	.00	.90	46.10	198.30	530.10	328.20	327.00	172.80	.00	.00	
6.40	28.50	8.50	.00	50.30	628.30	632.10	199.70	547.60	32.50	31.30	26.90	
.00	.00	.60	36.00	95.80	359.00	301.60	361.40	342.60	216.70	24.00	12.00	
.00	.00	36.00	22.60	83.60	220.20	567.30	552.40	488.20	346.00	.00	.00	
151.60	.00	.30	20.00	194.20	379.20	332.40	173.10	439.40	22.70	.50	2.20	
17.00	45.60	.00	1.20	85.70	157.50	320.40	453.40	294.10	148.80	.60	37.00	
.00	11.40	53.40	199.70	94.80	252.80	368.50	392.40	219.20	83.90	.00	14.00	
42.20	.00	24.20	40.60	67.40	401.70	380.30	294.50	590.10	159.10	.00	1.20	
4.00	42.60	30.00	109.70	96.50	382.00	241.10	258.80	217.00	51.80	.40	12.10	
74.00	.00	34.00	10.00	18.20	198.00	410.40	696.00	213.90	87.50	.60	.50	
.00	.00	.00	.00	74.00	410.60	357.30	210.00	149.10	21.00	10.00	.00	
.00	.00	10.00	11.00	23.80	742.00	261.80	204.00	231.00	97.00	.00	18.00	
.00	.00	14.00	29.20	80.00	122.50	548.90	398.90	301.70	104.20	22.30	.00	
.40	.00	.00	131.80	149.20	501.80	782.40	953.50	408.70	358.40	10.00	.00	
.00	4.00	16.00	22.00	108.00	331.20	512.00	544.00	403.40	10.00	6.00	.00	
42.00	1.50	.00	.00	138.00	382.00	379.50	317.00	252.50	33.50	5.00	.00	
.00	5.00	.00	74.50	108.50	217.00	343.50	425.00	301.00	404.00	46.00	27.00	
1.50	1.00	.00	9.30	110.90	267.50	246.00	285.50	225.00	40.00	15.00	16.50	
6.60	10.50	53.00	6.00	52.00	295.50	233.50	391.00	214.00	118.00	21.00	26.00	
3.50	18.50	5.00	38.00	52.00	180.00	315.50	309.00	166.50	6.50	.00	32.50	
164.50	.00	.00	46.00	88.50	248.50	153.50	515.50	282.50	89.50	20.50	.00	
70.00	37.00	13.00	89.50	105.50	378.00	483.50	357.50	193.50	110.50	4.00	12.50	
.00	7.50	22.00	14.00	137.90	209.50	383.00	294.00	77.00	127.10	2.00	2.50	
49.00	76.30	10.00	.00	33.00	65.50	519.50	273.00	300.00	65.50	40.00	.00	
9.50	41.50	4.50	.00	112.70	331.50	389.00	279.50	226.50	59.60	2.50	.00	
.80	6.70	6.20	54.70	61.70	464.00	433.80	338.50	320.10	106.30	42.30	.00	
8.00	.00	1.00	48.00	58.70	185.10	121.00	237.50	299.20	136.00	31.50	7.00	
.00	2.50	10.00	33.00	51.50	316.30	491.00	474.50	152.50	.00	28.00	.00	
.00	.00	17.00	.00	65.00	328.50	518.50	421.50	162.00	67.00	1.50	.00	
.00	1.20	13.50	17.50	58.20	321.50	504.70	428.00	157.20	33.50	14.70	.00	1989

1000

Factorial experiment 1 (1961)

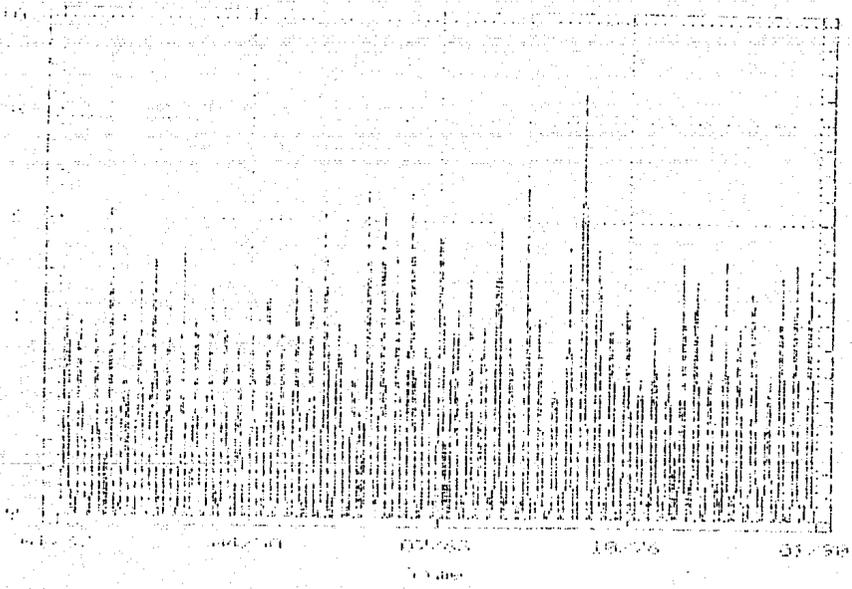


fig 1.9

Estimated Autocorrelations

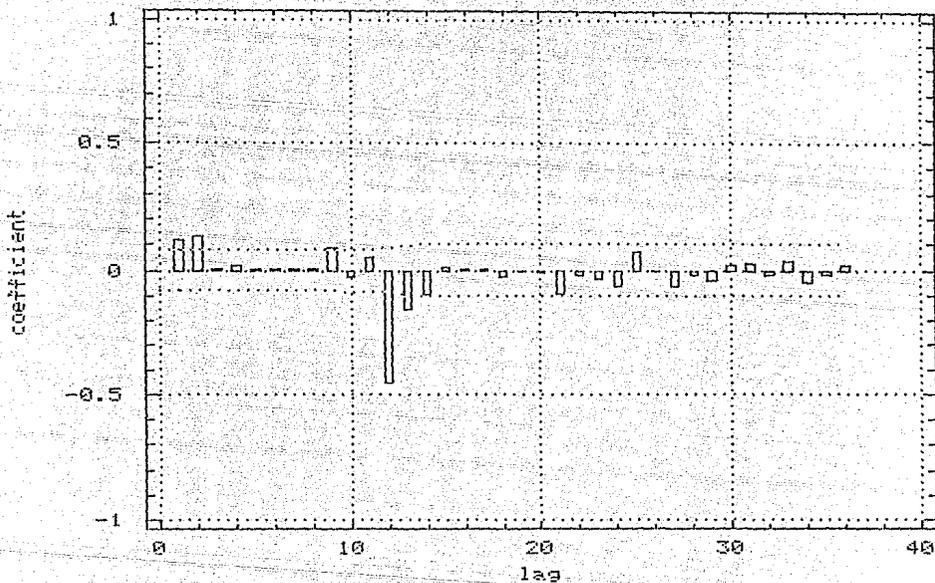


fig. 1.10 FAC of $W_t = \nabla_{12} Z_t$

por lo tanto consta de 636 observaciones originales, tal número de observaciones se reduce a 624 al aplicar el operador ∇_{12} .

La gráfica de las autocorrelaciones de la figura ^{1.6} 4.2.3 es característica de un modelo multiplicativo estacional:

ARIMA (0,0,0)X(0,1,1)

4.2.1 Primer modelo propuesto.

$$\nabla_{12} Z_t = (1-\theta_1 B)(1-\theta_2 B^{12}) a_t. \quad 4.2.1$$

ii) Estimación de los parámetros involucrados en el modelo.

Una vez propuesto el modelo 4.2.1 se realiza la estimación de los parámetros que están involucrados.

El valor estimado a los parámetros resulta ser:

$$\begin{aligned} \hat{\theta}_1 &= -0.1108 \\ \hat{\theta}_2 &= 0.9279 \end{aligned}$$

sus intervalos de confianza son respectivamente:

$$\begin{aligned} &(-0.1505, -0.0711) \\ &(0.9136, 0.9422) \end{aligned}$$

por lo que no se tiene problema con los parámetros.

iii) Verificación de que el modelo proporciona un ajuste adecuado y que los supuestos básicos se satisfacen. En caso de que los supuestos no se cumplan, se determinan las modificaciones necesarias al modelo.

{ a_t } tiene media cero.

Una vez teniendo calculada la media y la desviación estándar muestral de los residuales se checa la hipótesis:

Si $| (n)^{1/2} m(\hat{a}) / \sigma(\hat{a}) | < 2$, se dice que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero y por lo tanto no se rechaza el supuesto.

Así que se tiene:

$$m(\hat{a}) = -2.0353 \quad \sigma(\hat{a}) = 83.2552.$$

$$\frac{\sqrt{624} (2.0353)}{83.2552} = 0.6107 < 2.$$

por lo tanto se considera con media cero.

Para determinar la significación estadística individual de las autocorrelaciones de los residuales se calcula $| r_k(a) | \geq .08$. y se encuentra que $r_1(a) = .12$, $r_2(a) = .13$ y $r_{11}(a) = .11$. (Problema No. 1).

De hecho cuando se presenta un problema de este tipo se recomienda probar otro modelo y abandonar el que se estudia.

En la prueba conjunta se calcula el estadístico Q de un total de k autocorrelaciones simultáneamente, en nuestro caso se escoge $k = 24$, por lo que Q sigue aproximadamente una distribución Ji-cuadrada con $k-r$ grados de libertad, donde r es el número de parámetros involucrados. El valor de Q calculado es comparado con valores de tablas de Ji-cuadrada con los correspondientes grados de libertad, debiendo ser menor.

Para el modelo 4.2.1 se tienen 22 grados de libertad y un valor de: $Q = 35.0$ para la autocorrelación número 24.

El valor de Ji-cuadrada con 22 grados de libertad al 0.05, es de 33.92, por lo que el valor de Q es mayor (Problema No. 2).

Lo mejor en este caso es proponer otro modelo y dejar de revizar los supuestos restantes.

4.2.2. Segundo Modelo propuesto.

$$(1-\phi_1 B) \nabla_{12} Z_t = (1-\theta_1 B)(1 - \theta_2 B^{11})(1 - \theta_3 B^{12}) a_t. \quad 4.2.2$$

observamos que los supuestos sobre el modelo se cumplan de igual manera:

La estimación de los parámetros nos indica que:

$$\hat{\phi}_1 = 0.6539 \quad \text{error estandar : } 0.1521$$

$$\hat{\theta}_1 = 0.5231 \quad \text{error estandar : } 0.1714$$

$$\hat{\theta}_2 = 0.9226 \quad \text{error estandar : } 0.0157$$

$$\hat{\theta}_3 = -0.1260 \quad \text{error estandar : } 0.0401$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

$$(0.3497, 0.9581)$$

$$(0.1803, 0.8659)$$

$$(0.8912, 0.9540)$$

$$(-0.2062, -0.0458)$$

la matriz de correlación indica correlaciones altas entre los parámetros ϕ_1 y θ_1 del 98%.

Análisis de los residuales para el modelo 4.2.2

se tiene que:

$$m(\hat{a}) = -1.0018 \quad \text{y} \quad \sigma(\hat{a}) = 82.1115.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) | = 0.3045 < 2$.

Se presenta ahora el residuale $r_{13}(a) = .10$. (Problema No. 1).

El estadístico $Q = 17.9$, es menor que una $\chi^2_{\alpha}(20) = 31.41$.
con $\alpha = 0.050$.

El modelo 4.2.2 no cumple adecuadamente con los supuestos, y por lo tanto debe ser considerado otro modelo representativo del proceso generador de la serie.

4.2.3. Tercer Modelo propuesto.

$$(1 - \phi_1 B - \phi_2 B^2) \nabla_{12} Z_t = (1 - \theta_1 B^{11}) (1 - \theta_2 B^{12} - \theta_3 B^{13}) a_t. \quad 4.2.3$$

observemos que los supuestos sobre el modelo se cumplen de igual manera:

La estimación de los parámetros nos indica que:

$$\hat{\phi}_1 = 0.1173 \quad \text{error estandar} : 0.0399.$$

$$\hat{\phi}_2 = 0.1144 \quad \text{error estandar} : 0.0400.$$

$$\hat{\theta}_1 = -0.1275 \quad \text{error estandar} : 0.0400.$$

$$\hat{\theta}_2 = 0.9395 \quad \text{error estandar} : 0.0135.$$

$$\hat{\theta}_3 = 0.0364 \quad \text{error estandar} : 0.0145.$$

por lo tanto los intervalos de confianza del 95% son respectivamente:

(0.0375, 0.1971)
(0.0344, 0.1944)
(-0.2075, -0.0475)
(0.9125, 0.9665)
(0.0074, 0.0654)

La matriz de correlación no indica correlaciones altas entre los parámetros.

Análisis de los residuales para el modelo 4.2.3.

se tiene:

$$m(\hat{a}) = -1.8936 \quad \text{y} \quad \sigma(\hat{a}) = 81.7853.$$

por lo que el cociente $| (n)^{1/2} m(a) / \sigma(a) | = 0.5774 < 2.$

No se presentan residuales individuales significativamente distintos de cero.

El estadístico $Q = 14.0$, es menor que una $\chi^2_{\alpha}(19) = 30.14$.
con $\alpha = 0.050$.

Tampoco existen observaciones cuyos residuales sean estadísticamente distintos de cero. Por lo que el modelo 4.2.3 puede ser considerado como un modelo representativo del proceso generador de la serie, por lo tanto es usado para el pronóstico de ISO1520.

En general podemos resumir los resultados de los tres modelos propuestos en las siguiente tablas.

RESUMEN DE RESULTADOS DE LA ESTIMACIÓN DE MODELOS

SERIE $Z_t = \text{ISO1520}$.

PERIODO = ENERO DE 1937 A DICIEMBRE DE 1989.

NUMERO DE OBSERVACIONES = 636 OBSERVACIONES (HENSUALES).

MODELO	PARAMETROS	INTERVALOS DE CONFIANZA	PARAMETROS CORREL.
$\nabla_{12} Z_t =$ $(1-\theta_1 B)(1-\theta_2 B^{12}) a_t$	$\hat{\theta}_1 = -0.1108$ $\hat{\theta}_2 = 0.9279$	$(-.1505, -.0711)$ $(.9136, .9422)$	
$(1-\phi_1 B) \nabla_{12} Z_t =$ $(1-\theta_1 B)(1-\theta_2 B^{11})$ $(1-\theta_3 B^{12}) a_t$	$\hat{\theta}_1 = 0.5231$ $\hat{\theta}_2 = 0.9226$ $\hat{\theta}_3 = -0.1260$ $\hat{\phi}_1 = .6539$	$(.1803, .8659)$ $(.8912, .9540)$ $(-.2062, -.0458)$ $(.3497, .9581)$	
$(1-\phi_1 B - \phi_2 B^2) \nabla_{12} Z_t =$ $(1-\theta_1 B^{11})$ $(1-\theta_2 B^{12} - \theta_3 B^{13}) a_t$	$\hat{\theta}_1 = -0.1275$ $\hat{\theta}_2 = 0.9395$ $\hat{\theta}_3 = 0.0364$ $\hat{\phi}_1 = 0.1173$ $\hat{\phi}_2 = 0.1144$	$(-.2075, -.0475)$ $(.9125, .9665)$ $(.0074, .0654)$ $(.0375, .1971)$ $(.0344, .1944)$	

ANÁLISIS DE RESIDUALES				
$\hat{m}(\hat{a})$	$\hat{\sigma}(\hat{a})$	COCIENTE	$Q =$ G.L. =	$r_k(a) \neq 0$
-2.0353	83.2552	0.6107	35.0 22	$r_1 = .12$ $r_2 = .13$ $r_{11} = .12$
-1.0018	82.1115	0.3045	17.9 20	$r_{13} = .10$
-1.8936	81.7853	0.5774	14.0 19	

El modelo 4.2.3 es el que cumple satisfactoriamente los supuestos sobre la estimación de modelos. por lo tanto es usado para realizar el pronóstico de ISO1520.

Interpretamos la consecuencia de la lluvia, para un cierto mes t ; debida a efectos aleatorios de los meses $t-1$ y $t-2$ (del mismo año) así como a la lluvia de los meses $t-11$, $t-12$ y $t-13$ del año anterior (meses posterior y anterior del mes correspondiente pero del año anterior).

Nuevamente se tienen los pronósticos.

Pronóstico de lluvia mensual acumulada para ISO810, para el periodo Ene 1990 - Dic 1990.

Enero	15.39	Julio	419.68
Febrero	11.98	Agosto	361.19
Marzo	8.15	Septiembre	246.91
Abril	28.83	Octubre	118.79
Mayo	88.08	Noviembre	27.31
Junio	319.91	Diciembre	4.27

En la figura 4.9 se muestra la gráfica de los datos ajustados por el modelo a la serie ISO1520, mostrando sólo el periodo Ene. 1985 a Dic. 1989, y en la figura 4.10 se presenta la gráfica de los datos ajustados y el pronóstico junto con los datos históricos de la serie ISO1520.

4.3 PRONOSTICO DE LA ISOYETA 1000 - 1200 mm.

El pronóstico para ésta isoyeta fué estudiado en la parte 3.1, y nos proporcionó los siguientes datos.

Enero	11.57	Julio	225.21
Febrero	6.82	Agosto	223.29
Marzo	9.68	Septiembre	230.22
Abril	19.60	Octubre	85.79
Mayo	57.15	Noviembre	15.37
Junio	217.11	Diciembre	4.97

4.4 PRONOSTICO GENERAL PARA EL ESTADO DE MORELOS.

Hagamos un breve resumen del pronóstico obtenido mediante la metodología de Box-Jenkins; en las siguientes tablas se presenta el modelo y el pronóstico obtenido, con dicho modelo, además de los intervalos de confianza para el pronóstico arrojado por la metodología Box-Jenkins, para cada región de interés en el estado de Morelos y catalogadas como ISO810, ISO1012 e ISO1520.

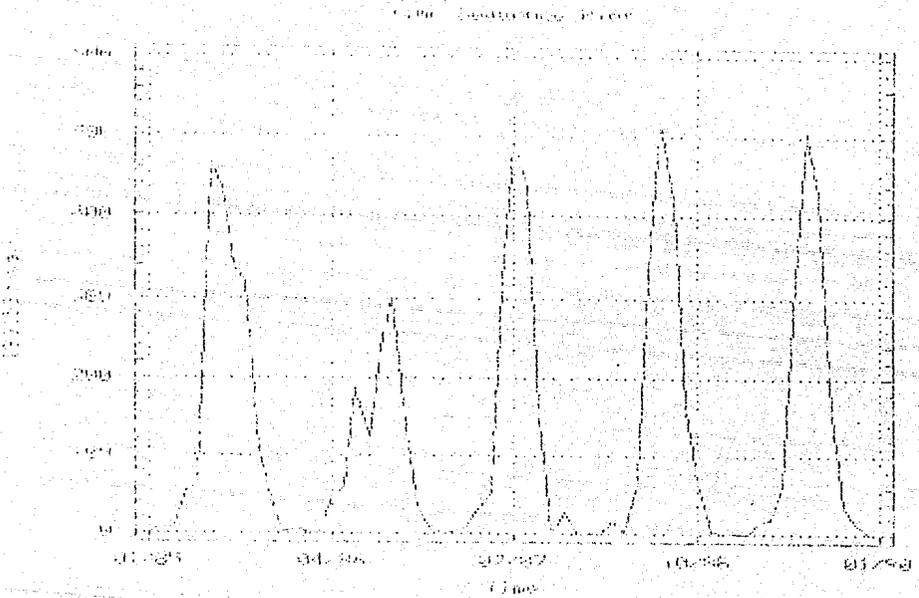


Fig. 1.11

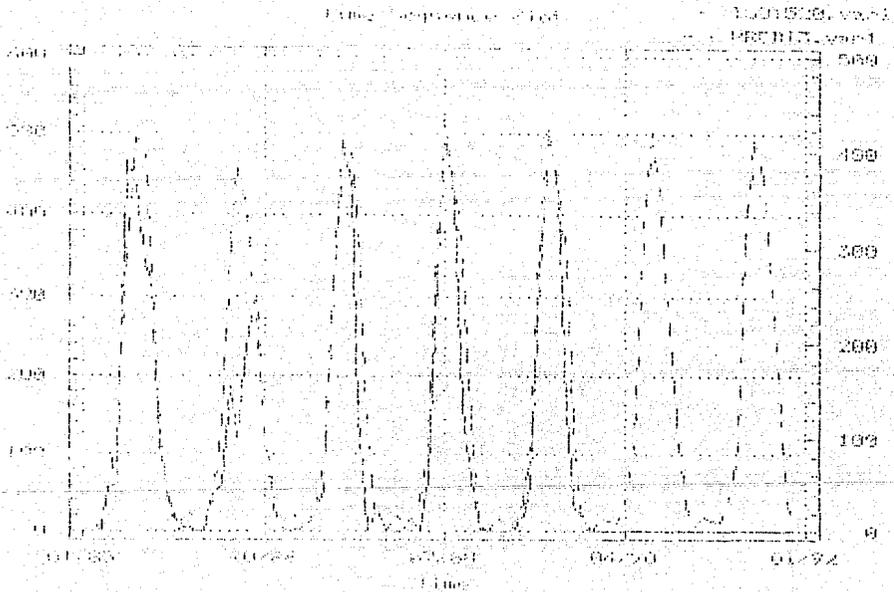


fig. 4 12

MORELOS

REGION: ISO810
 ISOYETA 800 - 1000 mm.

MODELO												
$(1+0.1106 B^{12})V_{12} Z_t = (1-0.9259 B^{12}) (1-0.0967 B^{13}) a_t$												
PRONOSTICO												
AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
LS	81.9	77.2	78.1	85.2	129.9	276.0	248.6	250.2	245.4	133.9	83.9	76.0
91	9.5	5.0	5.5	13.2	62.0	209.8	181.3	170.2	164.9	58.9	8.4	3.0
LI	0.0	0.0	0.0	0.0	0.0	128.6	101.2	102.8	98.0	0.0	0.0	0.0

REGION: ISO1012
 ISOYETA 1000 - 1200 mm.

MODELO												
$(1-0.0880 B^{13})V_{12} Z_t = (1-0.9525 B^{12}) a_t$												
PRONOSTICO												
AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
LS	105.3	100.3	103.1	113.1	150.6	310.6	318.7	316.7	323.7	179.3	108.8	98.4
91	11.5	6.8	9.6	19.6	57.1	217.1	225.2	223.2	230.2	85.7	15.3	4.9
LI	0.0	0.0	0.0	0.0	0.0	123.6	131.7	129.7	136.7	0.0	0.0	0.0

MORELOS

REGION: ISO1520
ISOYETA 1500 - 2000 mm.

MODELO

$$(1-0.1173 B - 0.1144 B^{12}) \nabla^{12} Z_t = (1+.1275 B^{11})(1-.9395B^{12}-0.0364B^{13}) a_t$$

PRONOSTICO

AÑO	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
LS	182.5	175.8	174.2	197.2	252.5	473.4	575.7	537.5	421.9	284.8	192.0	168.8
91	15.3	11.9	8.1	28.8	88.0	319.9	419.6	361.1	246.9	118.7	27.3	4.2
LI	0.0	0.0	0.0	0.0	0.0	144.6	246.9	208.7	93.1	0.0	0.0	0.0

ANEXO 1

- A) MACRO 1 (==DATOS)
- B) MACRO 2 (==MODELO)
- C) MACRO 3 (==MODELO1)

==DATOS

C PROCEDIMIENTO USADO PARA TRANSMISION DE DATOS AL SISTEMA SCA
C OPCIONES DE ARCHIVO PARA MORELOS: ISO810.DAT, ISO1012.DAT Y ISO1520.DAT

INPUT VARIABLE IS ZT. @
FORMAT IS FREE(1,110). @
FILE IS 'A:ISO810.DAT'.

C PROCEDIMIENTO USADO PARA CALCULO DE LAS FAC DE ZT.

ACF VARIABLE IS ZT. @
MAXLAG IS 50. @
HOLD ACF(FAC).

C PROCEDIMIENTO USADO PARA CALCULO DE LAS FAC DE D-1(ZT).

ACF VARIABLE IS ZT. @
DFORDER IS 1. @
MAXLAG IS 50. @
HOLD ACF(FAC10).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC D2-0(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 1,1. @
MAXLAG IS 50. @
HOLD ACF(FAC20).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D0-1(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 12. @
MAXLAG IS 50. @
HOLD ACF(FAC01).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D1-1(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 1,12. @
MAXLAG IS 50. @
HOLD ACF(FAC11).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D2-1(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 1,1,12. @
MAXLAG IS 50. @
HOLD ACF(FAC21).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D0-2(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 12,12. @
MAXLAG IS 50. @
HOLD ACF(FAC02).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D1-2(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 1,12,12. @
MAXLAG IS 50. @
HOLD ACF(FAC12).

C PROCEDIMIENTO USADO PARA CALCULO DE LA FAC DE D2-2(ZT)

ACF VARIABLE IS ZT. @
DFORDER IS 1,1,12,12. @
MAXLAG IS 50. @
HOLD ACF(FAC22).

RETURN

END.

==MODELO

C ESTIMACION DE MODELOS CORRESPONDIENTES A LAS ISOYETAS DEL
C ESTADO DE MORELOS. ISOYETAS: ISO810.DAT,
C ISO1012.DAT Y ISO1520.DAT

C LECTURA DE LA ISOYETA: ISO810.DAT
INPUT VARIABLE IS ZT. @
FORMAT IS FREE(1,110). @
FILE IS 'A:ISO810.DAT'.

C PROCEDIMIENTO USADO PARA ESPECIFICAR EL MODELO

C A PROBAR
UTSMODEL NAME IS MARIMA. @
MODEL IS (1-PHI1*B**12)ZT((1-B**12))= @
(1-TETA1*B**12)(1-TETA2*B**13)NOISE. @
SHOW. @
CHECK.

C PROCEDIMIENTO USADO PARA ESTIMAR PARAMETROS DEL MODELO

UESTIM MODEL IS MARIMA. @
STOP-CRITERIA ARE MAXIT(20). @
OUTPUT IS LEVEL(DETAILED). @
HOLD RESIDUALS(RESI),FITTED(XT),VARIANCE(G). @

C PROCEDIMIENTO USADO PARA PRONOSTICAR VALORES DE LA SERIE ZT.

UFORCAST MODEL IS MARIMA. @
METHOD IS CONDITIONAL. @
HOLD FORECASTS(PREDI), STD_ERRS(DES). @

C PROCEDIMIENTO USADO PARA CALCULAR FAC DE LOS RESIDUALES

ACF VARIABLE IS RESI. @
MAXLAG IS 50.

C PROCEDIMIENTO PARA SALVAR VARIABLES: XT, RESI, PREDI Y DES.

SAVE VARIABLE ARE XT. @
FORMAT IS '6F9.2'. @
FILE IS 'A:PREDI1.DAT'. @
REMARKS ARE 'DATOS AJUSTADOS'. @
REWIND.

SAVE VARIABLE ARE RESI. @
FORMAT IS '6F9.2'. @
FILE IS 'A:PREDI1.DAT'. @
REMARKS ARE 'RESIDUALES'.

SAVE VARIABLE ARE DES. @
FORMAT IS '6F9.2'. @
FILE IS 'A:PREDI1.DAT'. @
REMARKS ARE 'DESVIACION ESTANDAR'.

SAVE VARIABLES ARE PREDI. @
FORMAT IS '6F9.2'. @
FILE IS 'A:PREDI1.DAT'. @
REMARKS ARE 'PRONOSTICOS PARA 1990/1991 DE ISO810'.

RETURN
END.

==MODELO

C PRONOSTICO GENERAL MEDIANTE SUAVIZAMIENTO EXPONENCIAL
C USANDO EL METODO ADITIVO DE WINTER
C CORRESPONDIENTES A LA ISOYETA DEL ESTADO DE MORELOS:
C ISO1012.DAT.
C OPCIONES DE ARCHIVO: ISOB10.DAT, ISO1012.DAT, ISO1520.DAT

C LECTURA DE LA ISOYETA: ISO1012.DAT

INPUT VARIABLE IS ZT. @
FORMAT IS FREE(1,110). @
FILE IS 'E:ISO1012.DAT'.

C ESTE PROCEDIMIENTO ES USADO PARA EFECTUAR EL PRONOSTICO
C CORRESPONDIENTE MEDIANTE EL METODO DE WINTER ADITIVO

GFORECAST VARIABLE IS ZT. @
METHOD IS AWINTERS. @
WEIGHTS ARE .57, .007 , .43 . @
SEASONALITY IS 12. @

RETURN.
END.

BIBLIOGRAFIA

- BRILLINGER, DAVID R.
TIME SERIES DATA ANALYSIS AND THEORY
HOLT, RINEHART AND WINSTON Inc.
- BOX, G.E.P. Y JENKINS, G.
TIME SERIES ANALYSIS, FORECASTING AND CONTROL.
HOLDEN-DAY
- BOX, G.E.P.; HILLMER, S.C. Y TIAO, G.C.
ANALYSIS OF MODELLING OF SEASONAL TIME SERIES
A. ZELLNER ED.
- BURRIDGE, P. Y WALLIS, K.F.
CALCULATING THE VARIANCE OF SEASONALLY ADJUSTED SERIES
JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, 80, 541-552.(1985).
- CONTRERAS PIEDRAGIL CESAR E.
DESESTACIONALIZACION POR EXTRACCIÓN DE SEÑAL
TESIS DE MAESTRIA EN MATEMATICAS
U.A.M.-IZTAPALAPA.
- GRANGER, C.W.J.
SEASONALITY: CAUSATION, INTERPRETATION AND IMPLICATIONS
A, ZELNAR ED., SEASONAL ANALYSIS OF ECONOMIC TIME SERIES
- GUERRERO, V.
ANALISIS ESTADISTICO DE SERIES DE TIEMPO ECONOMICAS
CINVESTAV I.P.N.
- ISLAS CAMARGO ALEJANDRO
METODOS PARA IDENTIFICACION DEL ORDEN DE MODELOS
AUTORREGRESIVOS Y DE PROMEDIOS MOVILES
TESIS DE MAESTRIA EN MATEMATICAS
U.A.M.-IZTAPALAPA.

-LIU LON-MU, HUDAK GREGORY B.
THE SCA STATISTICAL SYSTEM
REFERENCE MANUAL
SCIENTIFIC COMPUTING ASSOCIATES

-MAKRIDAKIS, S.; WHEELWRIGHT S.C. AND MCGEE V.E.
FORECASTING: METHODS AND APPLICATIONS.
JOHN WILLY AND SONS.

-NELSON, CHARLES R.
APPLIED TIME SERIES ANALYSIS
FOR MANAGERIAL FORECASTING
HOLDEN-DAY, Inc.

-SALAS, J.D.; DELLEUR, J.W.; YEVJEVICH, V. AND LANE W. L.
APPLIED MODELING OF HIDROLOGICAL TIME SERIES
W.R.P.