

APLICACION DE LA METODOLOGIA DE BOX-JENKINS  
A LA PREVISION DE LA PUNTA MENSUAL DE CARGA  
DE UNA EMPRESA ELECTRICA

M. MARTI    A. PRAT    C. HERNANDEZ

En este trabajo se detallan con intención ilustrativa los pasos del análisis y modelado de la serie "Puntas Mensuales de Carga" de la empresa Gas y Electricidad, S.A. (G.E.S.A.) por el método de Box-Jenkins. Una vez modelada la serie, se utiliza el modelo para hacer previsiones de las puntas mensuales de 1978. Un análisis comparativo de los residuos obtenidos con el modelo SARIMA<sup>1</sup> y los resultantes de utilizar un modelo de regresión múltiple muestran la superioridad de aquel. Los resultados ilustran bien el potencial uso de esta metodología y sus extensiones en la industria eléctrica, en particular en el modelado, previsión y explotación de la Curva de Carga.

LISTA DE SIMBOLOS

- $A_t$  = Serie de residuos asociados a un modelo.
- $Z_t$  = Serie temporal estudiada.
- $B$  = Operador retardo:  $BZ_t = Z_{t-1}$
- $E$  = Esperanza matemática.
- F.A.C. = Función de autocorrelación.
- F.A.C.P. = Función de autocorrelación parcial.
- $K$  = Retardos para los que se calculan los F.A.C.P. y F.A.C.
- $\nabla$  = Operador diferencia pura.
- $\nabla_\tau$  = Operador diferencia de orden  $\tau$ .
- $r_k$  = Función de autocorrelación muestral en el retardo  $K$ .
- $\hat{\rho}_{kk}$  = Función de autocorrelación parcial estimada para el retardo  $K$ .
- $\omega$  = Serie transformada previamente -- por algún tipo de transformación tal vez no lineal, log. exp., etc. Por ejemplo  $\nabla \nabla_{12} \log Z_t = w_t$
- $\bar{w}$  = Media de la serie transformada.
- $\sigma_{\bar{w}}$  = Variancia de la media de la serie transformada.
- S-E = "Standard Error" del correspondiente parámetro estimado.
- $\hat{\sigma}_{A_t}$  = Estimación de la desviación tipo residual.
- $\emptyset$  = Parámetro de los polinomios en  $B$ , de la parte autoregresiva.
- $\theta$  = Parámetros de los polinomios en  $B$  de la parte en medios móviles.

1. INTRODUCCION

El conocimiento de la curva de carga es esencial para la asignación eficiente de recursos de una compañía eléctrica: despacho horario, previsiones de caja, manipulación de la curva de carga y su relación con tarificación, diseño de redes e investigación de mercados.

La metodología de Box-Jenkins parece muy apropiada para la previsión y modelado de la curva de carga. La referencia /4/ contiene una apreciación comparativa de su eficiencia en relación con otras técnicas utilizadas -- hasta ahora. No es sorprendente que aunque estas técnicas incluyan modelos con variables correctoras como población, temperatura, etc., la precisión de sus previsiones, sea inferior a la obtenida por métodos del tipo Box-Jenkins. La razón es que frecuentemente estas variables correctoras están correlacionadas con lo que la estimación es inconsistente.

El método de Box-Jenkins permite hacer previsiones proyectivas eficientes, a partir de la única información contenida en una serie temporal, pero además este modelado univariante, es una herramienta imprescindible para poder construir modelos bien especificados, mas complejos, que incluyan otras variables; análisis multivariante y de intervención.

La idea fundamental del método es la de in--

- M. Martí, A. Prat i G. Hernández de la Càtedra d'Estadística de l'Escola Tècnica Superior d'Enginyers Industrials de Barcelona. Diagonal, 647. Barcelona - 28.  
- Article rebut el Novembre del 78.

Tabla 1.1

634	637	611	635	652	688	752	787	788	715	631	721
712	701	705	745	735	759	880	909	917	854	738	808
810	826	802	845	865	913	993	1030	994	901	857	969
951	956	973	969	961	1080	1175	1230	1236	1040	956	1050
1060	1072	1133	1094	1128	1216	1351	1440	1339	1171	1168	1258
1283	1236	1250	1245	1249	1284	1496	1515	1375	1345	1274	1384
1444	1440	1447	1454	1391	1487	1680	1730	1672	1473	1457	1518
1470	1503	1396	1457	1384	1481	1641	1741	1688	1546	1469	1549
1512	1509	1547	1497	1441	1493	1808	1800	1620	1616	1502	1606
1649	1620	1555	1521	1465	1604	1761	1808	1625	1532	1476	1610
1642	1574	1577	1546	1452	1558	1760	1802	1719	1651	1566	1675

PUNTA MENSUAL DE GESA

tentar ajustar alguno(s) de los modelos generales formulados por Box-Jenkins /1/, /2/, - /3/ a nuestra serie temporal mediante un procedimiento de aprendizaje secuencial cuyo mecanismo se resume en el esquema de la fig. 1.1.

En este artículo ilustramos el análisis con datos de una sola serie, modelo univariante. El análisis se ha hecho con 132 observaciones que corresponden a las puntas mensuales de G.E.S.A. de enero de 1967 a diciembre de 1977 ambos inclusive. Los datos de partida están tabulados en la Tabla 1.1.

Es sabido que el estudio de una sola serie temporal por la metodología que nos ocupa, es un proceso en el que interviene en forma interactiva dos elementos básicos:

- a) El razonamiento del analista.
- b) El ordenador.

En el presente artículo exponemos los razonamientos efectuados para modelar la serie de puntas mensuales de G.E.S.A. creyendo que pueden ser de interés para otros analistas de series temporales. Todos los cálculos necesarios han sido efectuados con programas de ordenador especialmente diseñados para el análisis de series temporales según la metodología de Box-Jenkins.

2. IDENTIFICACION DEL MODELO

2.1 Análisis del gráfico de la serie original

Los valores de las puntas y su gráfico aparecen en el listado correspondiente (fig. 2.1).

El gráfico de la serie nos muestra en primer lugar que no es estacionaria en la media ya que existe una clara tendencia al crecimiento con cambio brusco a partir del dato 84 (Diciembre de 1973), cambio debido seguramente a la crisis del petróleo de dicho año. La no estacionariedad en la media deberemos eliminarla diferenciando la serie convenientemente, es decir aplicando el operador  $V=1-B$  una o varias veces a nuestra serie original.

Se observa también en el gráfico una clara -

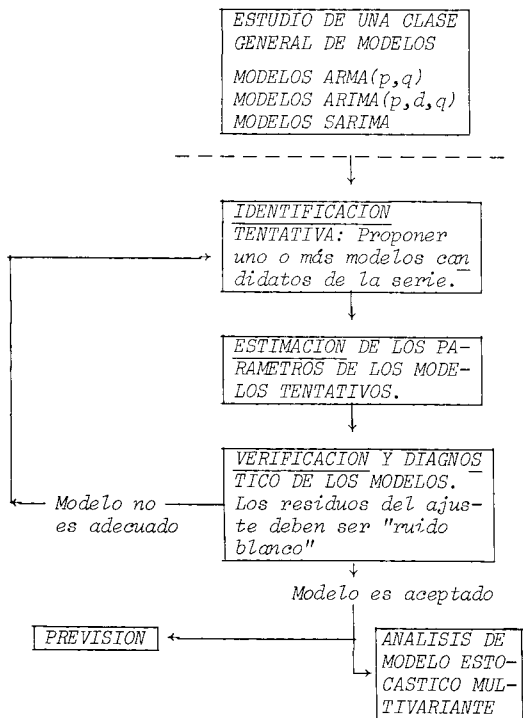
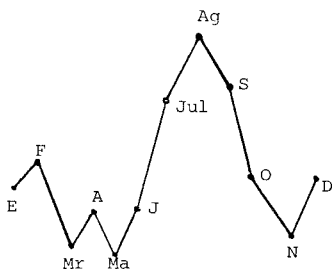


Fig. 1.1

estacionalidad de período  $S=12$  que viene reflejada por la repetición más o menos sistemática del segmento



Otro hecho importante es que la variancia de la serie aumenta al aumentar el nivel de la misma, lo que nos sugiere la conveniencia de una transformación logarítmica de los datos. La conveniencia de esta transformación viene reforzada por el hecho de que el gráfico de la serie  $\log. Z_t$  (Fig. 2.2) presenta una igualdad de variancia mucho mayor que la de la serie original. El gráfico medias-amplitudes nos confirmará la bondad de la transformación logarítmica.

Finalmente no se observa en el gráfico la existencia de datos anormales ("outliers").

## 2.2 Análisis de las F.A.C. y F.A.C.P. de diversas transformaciones de los datos

### 2.2.1 F.A.C. de la serie $\log. Z_t$ .

Los coeficientes de autocorrelación con sus correspondientes desviaciones tipo aparecen en el listado de ordenador correspondiente, listado que contiene además el gráfico de la

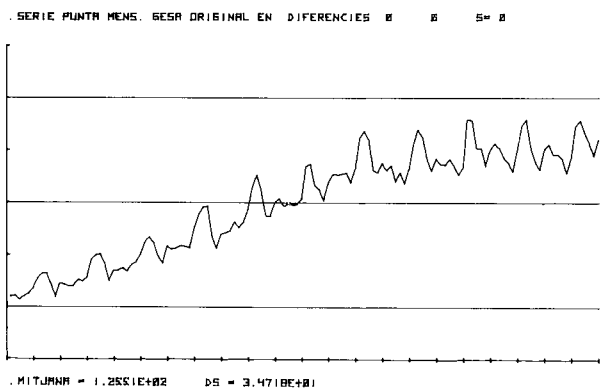


Fig. 2.1

función de autocorrelación estimada para  $K=1,2,\dots,42$  (Fig. 2.3).

Este gráfico nos confirma la no estacionariedad en la media, ya mencionada al observar el gráfico de la serie original y a la vez la estacionalidad de período  $S=12$  aunque dicha estacionalidad esté enmascarada por el hecho dominante de la tendencia.

Para eliminar la no estacionariedad en la media diferenciamos la serie original una vez y pasamos a analizar la serie  $\nabla \log. Z_t$ . Dado que la serie  $\log. Z_t$  no es estacionaria no entramos en el análisis de la F.A.C.P.

### 2.2.2 F.A.C. de $\nabla \log. Z_t$ .

La estimación de la F.A.C. ( $r_k$ ) para la serie  $\nabla \log. Z_t$  y retardos  $K=1,2,\dots,42$  la desviación tipo de cada coeficiente de autocorrelación y el gráfico de dicha función aparecen en el listado de ordenador (Fig. 2.4).

El análisis de la F.A.C. muestra ahora claramente la estacionalidad de período  $S=12$ . Obsérvese el valor de  $r_{12}$ ,  $r_{24}$ ,  $r_{36}$ . El decrecimiento lento de dichos coeficientes sugiere además la existencia de una no estacionariedad estacional y por lo tanto la conveniencia de aplicar el operador  $\nabla_{12} = (1-B^{12})$  a la serie.

El valor relativamente elevado de  $\bar{\omega}/\sigma_{\omega} = 1,29$  es otro dato que confirma la existencia de no estacionariedad. Sería pues conveniente analizar la serie  $\nabla \nabla_{12} \log. Z_t$  pero vamos antes a estudiar la serie  $\nabla_{12} \log. Z_t$ .

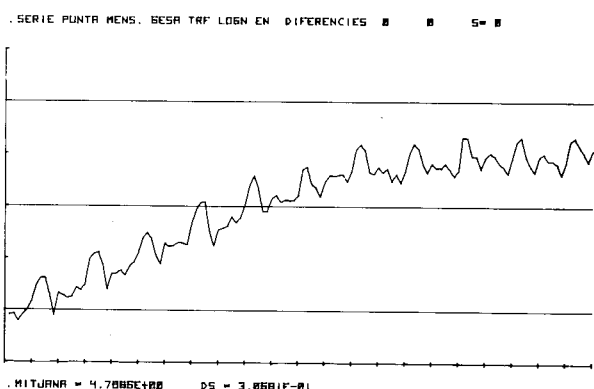


Fig. 2.2

2.2.3 F.A.C. de  $\nabla_{12} \log. Z_t$ .

El gráfico (Fig. 2.5) muestra claramente la no estacionariedad en la media confirmando la necesidad del operador  $\nabla$ .

En cambio parece evidente que hemos conseguido una estacionariedad estacional.

Otro dato favorable a la necesidad de analizar la serie es  $\nabla_{12} \log. Z_t$  en que  $\bar{\omega}/\sigma_{\omega} = 13,74$ .

En consecuencia proseguimos el análisis con la serie transformada  $\nabla_{12} \log. Z_t$ , transformación que además induce la reducción máxima en la variancia inicial.

2.2.4 F.A.C. y F.A.C.P. de la serie  $\nabla_{12} \log Z_t$ .

Los gráficos de las estimaciones de la F.A.C. y de la F.A.C.P. así como los valores de los coeficientes de autocorrelación parcial con sus desviaciones tipo aparecen en el listado correspondiente (Fig. 2.6).

Los coeficientes de autocorrelación significativamente no nulos son

$r_1 = -.35$	$r_{12} = -.43$
S.E. .09	S.E. .11

El coeficiente  $r_{20} = .22$  es anormalmente alto aunque no significativo. Su existencia se justificará más tarde en el análisis de los residuos.

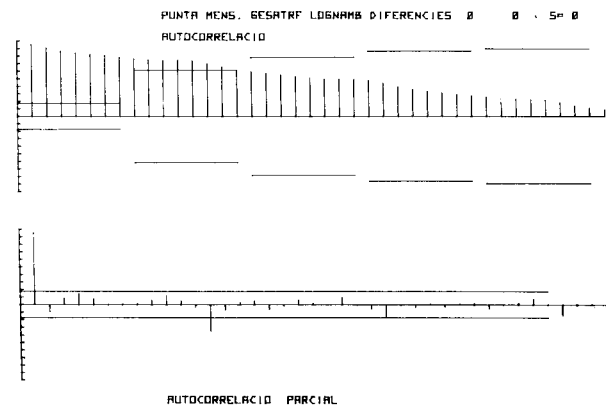


Fig. 2.3

Del análisis conjunto de las dos funciones se deduce la siguiente identificación tentativa:

- PARA LA COMPONENTE ESTACIONAL

Parece lógico que para la componente estacional proponamos un modelo  $ARIMA(0,1,1)_{12}$  -- (ver nota 1), es decir  $(\nabla_{12}) \log. Z_t = (1 - \theta B^{12}) A_t$  ya que la F.A.C. tiene un solo valor no nulo,  $r_{12}$  y la F.A.C.P. decrece exponencialmente un  $K=12, 24$  y  $36$ . La estimación inicial para  $\theta$  resultó ser  $\theta = .60$ .

- PARA LA COMPONENTE NO ESTACIONAL

Para la componente no estacional es posible, a nuestro juicio, proponer dos modelos,

a) Un modelo  $ARIMA(0,1,1)$  dado que la F.A.C. tiene un solo pico  $r_1$ , y la F.A.C.P. presenta un decrecimiento exponencial. La estimación inicial de  $\theta$  resultó ser  $\theta = .50$ .

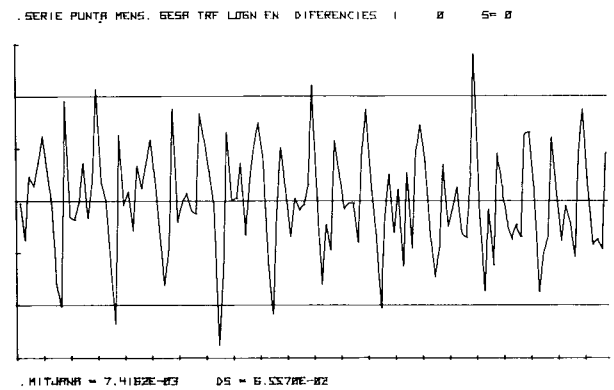


Fig. 2.4a

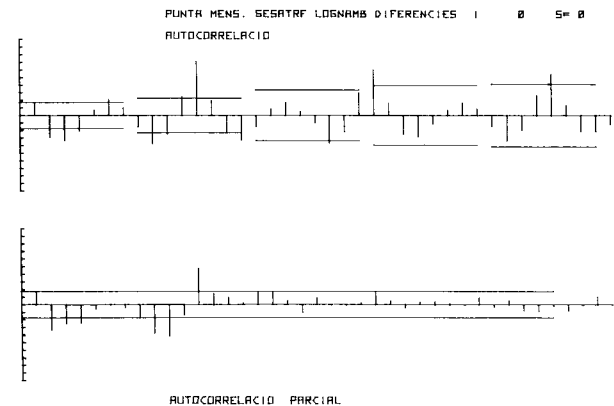


Fig. 2.4b

b) La existencia de dos coeficientes de auto correlación parcial no nulos  $\psi_{11}$  y  $\psi_{22}$  y la "satelización" observada alrededor de  $r_{12}$  en la F.A.C. permite considerar también un modelo ARIMA (2,1,0) con una estimación inicial para los parámetros  $\phi_1 = -.9$  y  $\phi_2 = -.5$ .

Así pues retendremos dos modelos SARIMA para la fase de estimación eficiente.

$$\begin{aligned} \text{M.1. } \nabla \nabla_{12} \log. Z_t &= (1-\theta B)(1-\theta' B^{12}) A_t \\ \theta &= .50 \\ \theta' &= .60 \end{aligned}$$

$$\begin{aligned} \text{M.2. } (1-\phi_1 B - \phi_2 B^2) \nabla \nabla_{12} \log. Z_t &= (1-\theta B^{12}) A_t \\ \phi_1 &= -.9 \\ \phi_2 &= -.5 \\ \theta &= .60 \end{aligned}$$

No introducimos término constante dado que  $\bar{\omega}/\sigma_{\omega} = 0,16$ .

### 3. ESTIMACION EFICIENTE

#### 3.1 Estimación del modelo M.1.

La estimación completa del modelo M.1. se encuentra en el listado correspondiente. Vamos a resumir aquí los resultados más importantes.

El modelo estimado con el intervalo de confianza de sus parámetros es:

$$\begin{aligned} \nabla \nabla_{12} \log. Z_t &= (1-.505B)(1-.596B^{12}) A_t \\ &\quad .343 \quad .439 \\ &\quad .666 \quad .754 \end{aligned}$$

Observamos en primer lugar que los dos coeficientes son significativos.

La matriz de correlaciones entre los coeficientes es:

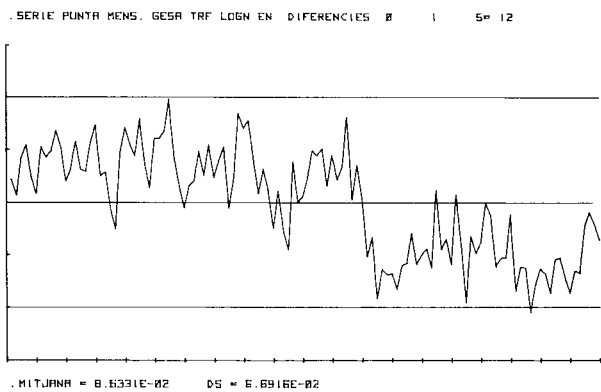


Fig. 2.5a

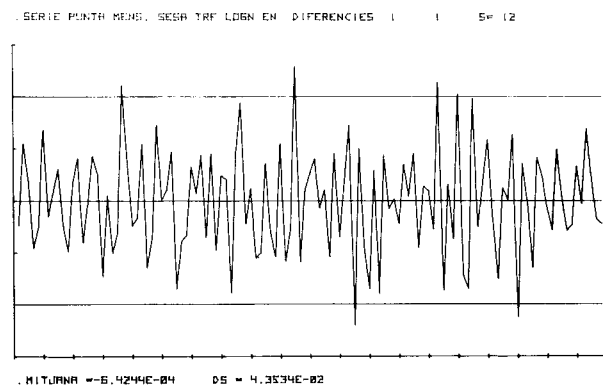


Fig. 2.6a

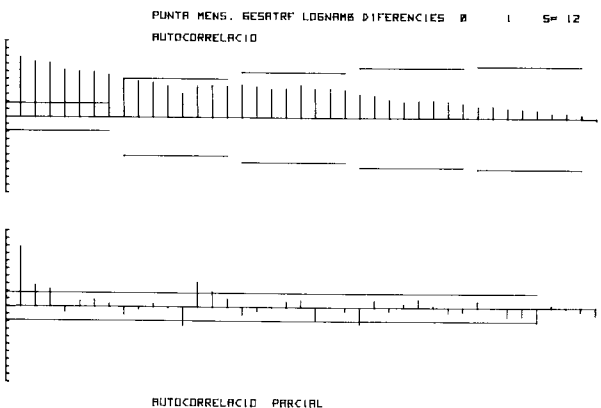


Fig. 2.5b

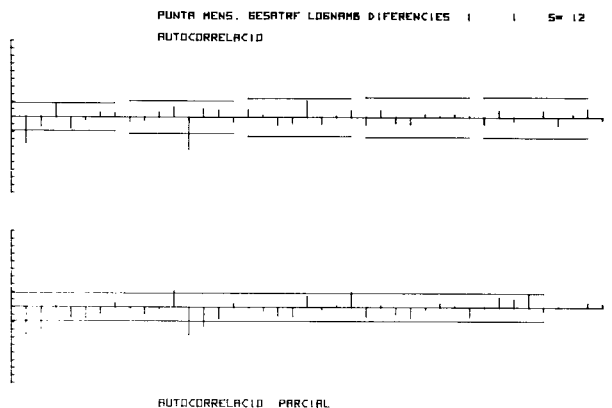


Fig. 2.6b

$\theta$	$\theta'$
0	1.0000
$\theta'$	-0.1464
	1.0000

Dado el valor pequeño del coeficiente -0.1464 vemos que las estimaciones de los parámetros son independientes.

La desviación tipo residual  $\hat{\sigma}_{A_t} = 0.034$  es bastante buena ya que supone una reducción en - varianza sobre la original logarítmica del 99%.

Para el segundo modelo los resultados básicos de la estimación fueron:

$(1 + .44275 B + .287 B^2) \nabla \nabla_{12} \log Z_t = (1 - .625 B^{12}) A_t$	
-0.263	-0.109
	0.467
-0.662	-0.464
	0.782

siendo pues todos los coeficientes significativos.

La independencia entre coeficientes es aceptable tal como se observa en la matriz de correlaciones

	$\phi_1$	$\phi_2$	$\theta$
$\phi_1$	1		
$\phi_2$	.3436	1	
$\theta$	.1492	.0072	1

y la variancia residual en  $\hat{\sigma}_{A_t} = 0.0332$  muy parecida a la del modelo M.1.

Los ceros del polinomio  $\phi(B)$  son complejo-conjugados y cumplen las condiciones de estacionariedad.

El que los dos modelos sean prácticamente -- equivalentes en cuanto a la reducción en varianza está justificado por el hecho de que invirtiendo el polinomio  $(1 - \theta B)$  en el modelo M.1. se obtendría un desarrollo autoregresivo muy parecido a  $\phi(B)$  en el modelo M.2.

#### 4. VERIFICACION DEL MODELO

##### 4.1 Análisis de los residuos de M.1.

No se observan residuos muy anormales. Los - residuos que caen fuera del intervalo  $0 \pm 1,96\sigma_A$  han sido los siguientes:

NOV 69	$A_{35} = +2 \sigma_A$
OCT 70	$A_{46} = -2,02 \sigma_A$
MAR 71	$A_{51} = +2,02 \sigma_A$
NOV 71	$A_{59} = +2,03 \sigma_A$
SEP 72	$A_{69} = -2,03 \sigma_A$
OCT 72	$A_{70} = +2,01 \sigma_A$
EN 74	$A_{85} = -2,01 \sigma_A$
MAR 74	$A_{87} = -2,66 \sigma_A$
SEP 75	$A_{105} = -2,1 \sigma_A$

Dos de los residuos extremos  $A_{85}$  y  $A_{105}$  están separados por 20 meses lo cual podría -- ser la causa que explicase la existencia en la F.A.C. de los residuos del único coeficiente de autocorrelación significativamente no nulo  $r_{20} = .25$  así como del coeficiente de S.E. = .10

autocorrelación parcial  $\hat{\rho}_{20,20} = .28$   
S.E. = .09

En una breve conversación con miembros de -- G.E.S.A. llegamos a la conclusión de que el residuo  $A_{105}$  de septiembre de 1975 podría -- ser debido a la fuga masiva de turistas consecuencia de las ejecuciones del FRAP y ETA de dicho mes. El residuo  $A_{85}$  de enero de -- 1974 podría ser debido a la crisis del petróleo de diciembre de 1973 al igual que  $A_{87}$  correspondiente a la Semana Santa.

Los residuos restantes podrían ser objeto de un estudio más detallado para ver la posibilidad o no de ser explicados por causas externas y ser eliminados en un posterior estudio de análisis de intervención y/o función de transferencia.

De todos modos la configuración de los residuos es francamente aceptable, como se puede comprobar en la Fig. 4.1.

##### 4.3 Interpretación intuitiva de los modelos obtenidos

Es útil interpretar los modelos estimados -- en términos intuitivos. Sea por ejemplo el modelo M.1:

$$(1-B)(1-B^{12})X_t = (1-0.505B)(1-0.596B^{12})a_t$$

$$X_t = \log Z_t$$

podría escribirse

$$(1-B)(X_t - \bar{X}_{t-12}) = (1-0.505B)a_t$$

con,

$$(X_t - \bar{X}_{t-12}) = \frac{1-B^{12}}{1-0.596B^{12}} X_t =$$

$$= \{1-0.404(B^{12}+0.596B^{24}+\dots)\}X_t$$

de donde

$$(1): \bar{X}_{t-12} = 0.404X_{t-12} + 0.24X_{t-24} + 0.097X_{t-36} + \dots$$

es decir si por ejemplo t representa el mes de abril  $\bar{X}_{t-12}$  sería una media ponderada, con pesos decrecientes de los abrils anteriores.

Además

$$\frac{1-B}{1-0.505B} (X_t - \bar{X}_{t-12}) = \{1-0.495(B+0.505B^2+\dots)\} \cdot (X_t - \bar{X}_{t-12}) = (X_t - \bar{X}_{t-12}) -$$

$$- 0.495(X_t - \bar{X}_{t-13}) - 0.25(X_{t-2} - \bar{X}_{t-14}) \dots$$

con lo que finalmente tendríamos:

$$(2): X_t = \bar{X}_{t-12} + 0.495(X_{t-1} - \bar{X}_{t-13}) +$$

$$+ 0.25(X_{t-2} - \bar{X}_{t-14}) + \dots + a_t$$

Es decir, la previsión para, sea, el mes de abril de 1978 sería la media ponderada (1) - de los abrils anteriores corregida con una media también ponderada de las desviaciones de marzo de 78 respecto al "marzo típico", - febrero 1978 respecto al "febrero típico" y poco más ya que los pesos decrecen rápidamente.

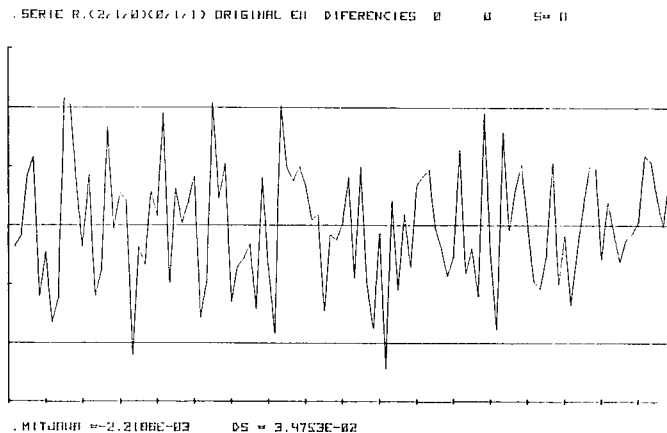


Fig. 4.1a

Esta interpretación explica por qué métodos simples de extrapolación pueden dar previsiones aceptables. Per la previsión de Box-Jenkins que puede traducirse en estas reglas de "interpolación", permite estimar los coeficientes, en este caso 1; 0.495; 0.25... eficientemente, además de desvelar la estructura de la serie.

#### 5. PREVISION DE LAS PUNTAS DE 1978

En la Tabla 5.1 se encuentran las previsiones de las puntas correspondientes a los doce meses del año 1978 hechas a partir del modelo M.1., así como el intervalo de confianza del 95% de dichas previsiones. Las previsiones hechas con el modelo M.2. se recogen en la tabla 5.2.

Siendo conservadores nos concretamos al análisis de las previsiones hechas con el modelo M.1., que es ligeramente inferior al M.2.

En el momento en que realizamos el estudio - disponíamos solo de los datos de la serie -- hasta diciembre de 1977. Actualmente G.E.S.A nos ha suministrado el valor observado de -- las puntas de Enero-Junio de 1978 y que presentamos en la Tabla 5.3 junto a las previsiones hechas con el modelo M.1.

Aunque los resultados son bastante elocuentes por si mismos conviene tener en cuenta - una serie de consideraciones.

- a) Todos los valores observados están dentro del intervalo de confianza del 95% de las

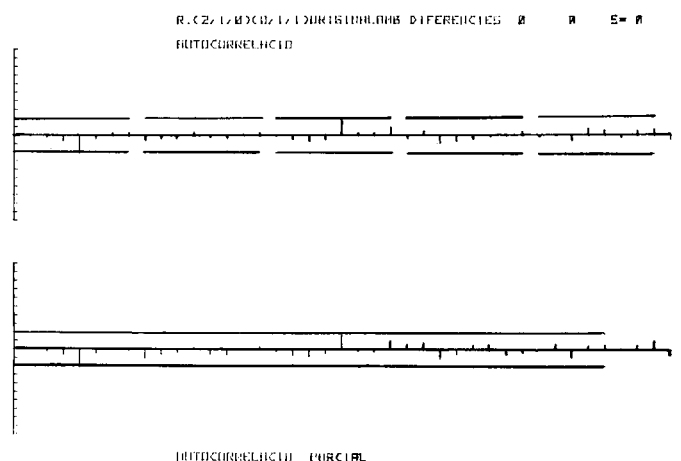


Fig. 4.1b

Tabla 5.1

MODEL 1 FORECASTS AT BASE PERIOD 132 WITH 95 PER CENT CONFIDENCE LIMITS

PERIODS AHEAD	LO. CONF. LIMIT	FORECAST	UP. CONF. LIMIT
1	.1586143+03	.1693010+03	.1807079+03
2	.1542027+03	.1658418+03	.1783594+03
3	.1515567+03	.1641160+03	.1777162+03
4	.1484850+03	.1618065+03	.1763232+03
5	.1410265+03	.1545839+03	.1694446+03
6	.1502395+03	.1655939+03	.1825174+03
7	.1690653+03	.1873192+03	.2075439+03
8	.1723903+03	.1919545+03	.2137390+03
9	.1601919+03	.1792207+03	.2005099+03
10	.1514942+03	.1702637+03	.1913587+03
11	.1434840+03	.1619697+03	.1828370+03
12	.1534080+03	.1739066+03	.1971442+03

Tabla 5.2

MODEL 2 FORECASTS AT BASE PERIOD 132 WITH 95 PER CENT CONFIDENCE LIMITS

PERIODS AHEAD	LO. CONF. LIMIT	FORECAST	UP. CONF. LIMIT
1	.1592070+03	.1698831+03	.1812752+03
2	.1546756+03	.1666061+03	.1794569+03
3	.1520470+03	.1647497+03	.1785136+03
4	.1486387+03	.1626634+03	.1780113+03
5	.1411454+03	.1556548+03	.1716557+03
6	.1501245+03	.1666383+03	.1849687+03
7	.1687116+03	.1885259+03	.2106672+03
8	.1718924+03	.1932800+03	.2173286+03
9	.1596305+03	.1805364+03	.2041802+03
10	.1505553+03	.1712298+03	.1947434+03
11	.1424301+03	.1628615+03	.1862238+03
12	.1521733+03	.1749003+03	.2010215+03

Tabla 5.3

MES	PREVISION	VALOR OBSERVADO	ERROR
Enero	169,30	179,0	-5,4%
Febrero	165,84	170,1	-2,5%
Marzo	164,12	168,5	-2,6%
Abril	161,81	156,6	+3,3%
Mayo	154,58	154,8	-0,1%
Junio	165,59	166,9	-0,7%

previsiones obtenidas.

b) Las previsiones para los seis meses de -- 1978 no están actualizadas. Como es sabido, la metodología utilizada permitiría modificar las previsiones para febrero, marzo, etc., en función de las discrepancias observadas entre el valor de enero de 1978 y el valor previsto para dicho mes. Asimismo, la nueva previsión para marzo, -- abril, etc. se actualizaría a su vez en --

función de la discrepancia observada entre el valor real y el previsto para febrero. La Tabla 5.4 da estas previsiones actuali- zadas para febrero.

c) El mes de enero de 1978 fue anormalmente frío en las Baleares, lo cual explicaría en parte el error de previsión para dicho mes, y la conveniencia de un modelo de -- transferencia, Carga-Temperatura.



Tabla 5.4

MES	PREVISION	VALOR OBTENIDO	ERROR
Febrero	170,30	170,1	+0,01%
Marzo	168,67	168,5	+0,01%
Abril	166,23	156,6	-5,5%
Mayo	159,138	154,8	-2,7%

d) La previsión con los modelos univariantes dinámicos resultó superior a la suministrada por un modelo de regresión múltiple con cinco variables, una de ellas ficticia. Ello indica que como es frecuente en econometría, también en problemas técnicos la presencia de errores de especificación de modelo mina sustancialmente el valor predictivo de modelos de regresión múltiple.

6. CONCLUSIONES

Hemos ilustrado con intención pedagógica una aplicación de la metodología de Box-Jenkins a su nivel inferior: análisis y modelado de una sola serie. Las previsiones obtenidas dan fe de la utilidad potencial de estos métodos en la industria eléctrica en particular y en general en aquellas empresas en que una buena previsión es un parámetro importante en su actividad. Las previsiones obtenidas dan idea de la utilidad del método, previsiones que al menos en este caso concreto fueron superiores a la suministrada por un modelo de regresión múltiple y no inferiores a la obtenida con modelos en el espacio de fases incluyendo variables con temperatura y procedimientos alternativos.

Es conveniente además, percatarse de que el análisis de una serie temporal según Box-Jenkins en su nivel más elemental es un paso inevitable en el modelo multivariante, si se quiere evitar errores de especificación debidos a colinealidad entre los regresores (realimentación), situación muy frecuente.

La previsión suministrada por el modelo univariante podría ser mejorada, permitiendo el cálculo en tiempo real mediante una modificación del método de Box-Jenkins, e incorporando la información que aportan otras variables relacionadas con la curva de carga (exceso de frío, calor, población activa, etc). Si se quiere mejorar la previsión univariante, simular el efecto de intervenciones o ma-

nipulaciones de la curva de carga  $Z_t$ , hemos de relacionarla con otras variables exógenas (explicativas)  $X_{1t}, X_{2t}, X_{3t}, \dots$  (alguna tal vez controlable) identificando y estimando el correspondiente modelo multivariante de transferencia de una salida.



En una próxima contribución abordaremos algunos aspectos de la metodología de Box-Jenkins a este nivel superior que aplicaremos al modelado de la curva de carga en función de otras variables y utilizaremos análisis de intervención para estudiar el efecto de la crisis del petróleo de 1973 y otras intervenciones.

Durante la revisión de este artículo hemos tenido conocimiento de un importante trabajo, Hagan /4/, sobre el valor comparativo de la metodología de Box-Jenkins en la previsión de la carga horaria y diaria y la modificación de dicha metodología para permitir cambios suaves en los parámetros del modelo y obtener previsiones en tiempo real. Las conclusiones del mismo confirman las nuestras.

7. AGRADECIMIENTOS

Agradecemos al Comité de Curvas de Carga de U.N.E.S.A. la asistencia recibida en relación con este trabajo. Ha sido realizado con datos de Gas y Electricidad, S.A. (G.E.S.A.) y forma parte de un estudio hecho para Unidad Eléctrica, S.A. (U.N.E.S.A.) dentro de un proyecto de investigación sobre Curvas de Carga que dicha entidad realiza bajo el régimen de acción concertada de investigación aprobada por la Comisión Asesora de Investigación Científica y Técnica.

## 8. REFERENCIAS

- /1/ ANDERSON, O.D. "Time Series Analysis and Forecasting. The Box-Jenkins Approach". Butterworth. 1976.
- /2/ BOX, G.E.P. & JENKINS, G.M. "Time Series Analysis Forecasting and Control". Holden-Day. 1976.
- /3/ GRANGER, C.W.J. & NEWBOLD, P. "Forecasting Economic Time Series". Academic-Press. 1977.
- /4/ HAGAN, M.T. & KLEIN, R. "On the Maximum Likelihood Estimation for Load Forecasting". Trans. Systems Man and Cybernetics. IEEE. vol. SMC-8 nº 9, Sept. 1978.

que se denomina model SARIMA (p,d,q) (P,D,Q) (seasonal autoregressive integrated moving - average). La parte que opera sobre  $B^s$  recoge la dinámica estacional de la serie.

## 9. NOTAS

1. En la terminología de Box-Jenkins /2/ se denomina modelo mixto ARIMA (p,d,q) (autoregresivo-integrado-de medias móviles) a la representación

$$\begin{aligned} [1 - \phi_1 B - \phi_2 B^2 \dots - \phi_p B^p] (1-B)^d \{Z_t - E(Z_t)\} &= \\ = \Phi_p(B) \nabla^d (Z_t - \bar{Z}_t) &= [1 - \theta_1 B - \theta_2 B^2 \dots - \\ - \theta_q B^q] A_t &= \Theta_q(B) A_t \end{aligned}$$

$A_t \sim \text{Normal}(0, \sigma_A^2)$

$\Phi_p$  parte autoregresiva de orden p.

$\Theta_q$  parte medias móviles de orden q.

d orden de las diferencias puras

$$\nabla = (1-B)$$

Si el modelo contiene además factores polinómicos con coeficientes potencias de  $B^s$  del tipo

$$[1 - \phi_{1,s} B^s - \phi_{2,s} B^{2s} - \dots - \phi_{p,s} B^{ps}] = \Phi_p(B^s)$$

en la parte autoregresiva, y

$$[1 - \theta_{1,s} B^s - \theta_{2,s} B^{2s} - \dots - \theta_{q,s} B^{qs}] = \Theta_q(B^s)$$

en la de medias móviles y  $\nabla_s^d$  es el factor de diferencias puras de tipo s y orden D, el modelo resultante sería

$$\Phi_p(B) \Phi_p(B^s) \nabla_s^d \nabla^D (Z_t - \bar{Z}_t) = \Theta_q(B) \Theta_q(B^s) A_t$$