



Asepelt
España

Comunicaciones XIV Reunión

**ANÁLISIS UNIVARIANTE Y DE INTERVENCIÓN EN
SERIES TEMPORALES DEL MERCADO DE
TRABAJO**

Joan Baró Llinàs

M. Àngels Cabasés Piqué

M. Jesús Gómez Adillón -

M.Jesus.Gomez@econap.UdL.es

Universitat de Lleida

Anales de Economía Aplicada

Oviedo 2³
Junio 2000 4



Reservados todos los derechos.

Este documento ha sido extraído del CD Rom “Anales de Economía Aplicada. XIV Reunión ASEPELT-España. Oviedo, 22 y 23 de Junio de 2000”.

ISBN: 84-699-2357-9

TÍTULO: Análisis univariante y de intervención en series temporales del mercado de trabajo

AUTORES:

BARÒ LLINÀS, JOAN; CABASÀ PIQUÈ, M. ÀNGELS y GÀMEZ ADILLÀN, M. JESÚS. E-mail: M.Jesus.Gomez@econap.UdL.es

Universitat de Lleida

Departamento de Economía Aplicada

RESUMEN:

En este estudio se efectúa el análisis univariante y de intervención de las series del mercado de trabajo en la provincia de Lleida, a partir de los trabajos de Box y Tiao (1975), que nos ha de permitir realizar predicciones en un horizonte temporal de dos años.

Se efectúa una descripción de los modelos teóricos y en la parte empírica, se analizan los valores de las variables para los años de 1977 a 1997, con detalle trimestral y se efectúa la construcción de modelos univariantes para las series históricas. Las variables población desocupada y población ocupada en el sector industria han sido intervenidas, detectándose anomalías de tipo aditivo y de cambio de nivel, efectuando la correspondiente corrección.

PALABRAS CLAVE: análisis de intervención, mercado de trabajo, predicción dinámica.

ÀREA TEMÀTICA: G2 MÈTODOS DE ESTADÍSTICA ECONÓMICA

Análisis univariante y de intervención en series temporales del mercado de trabajo

Introducción

El estudio de las variables del mercado de trabajo empieza, utilizando la metodología de análisis de series temporales, con la elaboración de modelos univariantes estocásticos (Box-Jenkins 1970) y con el análisis de intervención (Box-Tiao 1975).

Dejaremos a un lado otros procedimientos como el de medias móviles, técnicas de regresión, alisados, análisis armónica y espectral, aunque a veces, la referencia a estos modelos para hacer comparaciones será inevitable.

Antes de proceder a la exposición de estos tipos de modelos, se introducen algunos conceptos teóricos necesarios para el análisis empírico.

1. Definiciones

Consideraremos la definición de un proceso estocástico como una familia de variables aleatorias $\{Y(t, \omega) ; \omega \in \Omega, t \in T\}$ que evoluciona a lo largo del tiempo, asociada a cualquier variable indicadora del mercado de trabajo: población ocupada, población activa, etc... de acuerdo con unas determinadas leyes de probabilidad, donde T es el dominio temporal o el conjunto de puntos de tiempo para los cuales el proceso está definido, ω es el resultado de un acontecimiento aleatorio producido en el momento t y Ω es el espacio muestral sobre el cual está definida la variable aleatoria ω .

Para simplificar la notación, omitiremos la variable ω y nos referiremos al proceso estocástico $Y(t, \omega)$ simplemente por $Y(t)$ si

se trata de un proceso continuo o por Y_t si es un proceso en tiempo discreto.

Un proceso estocástico en tiempo continuo $\{Y(t), t \in T\}$ es aquel proceso en que el dominio temporal es $T = \mathbb{R}$, y por tanto,

$$-\infty < t < \infty.$$

En cambio un proceso estocástico en tiempo discreto $\{Y_t, t \in T\}$ es aquel proceso en que el dominio temporal es $T = \mathbb{Z}$, y por tanto $t = 0, \pm 1, \pm 2, \dots$

Hay una manera más sencilla, aunque incompleta, de describir un proceso estocástico a partir de sus momentos, en concreto a partir de los modelos de primero y segundo orden.

Podemos decir que una serie temporal aplicada al mercado laboral se define como:

$\{Y_t, t = 1, 2, 3, \dots, n\}$ un conjunto de observaciones de una variable discreta medida secuencialmente a lo largo del tiempo.

Las series del mercado de trabajo que se utilizaran son discretas, y los datos se miden en intervalos fijos de tiempo, en concreto, trimestralmente.

2. Análisis de intervención

Las series temporales del mercado de trabajo se encuentran influenciadas por sucesos externos, políticos, normativos, etc. En la provincia de Lleida, la pérdida de representación de la población ocupada en el sector agrícola sobre el total de ocupados, etc. son un ejemplo de los sucesos que pueden afectar las series temporales. Estos sucesos externos los denominaron Box y Tiao (1975) como *intervenciones*. La técnica del análisis de intervención consiste en evaluar el efecto de las intervenciones en el proceso de

comportamiento de una serie temporal. Para poder hacer intervenciones en un modelo, hemos de identificar dos características del modelo:

- 1) El período en el que empiezan los sucesos externos
- 2) La forma del impacto de las intervenciones.

El análisis de intervención se caracteriza porque con antelación se sospecha que algún suceso externo ha influido considerablemente en la serie temporal objeto de predicción.

Pero hay ocasiones en que no conocemos la existencia de estos sucesos o cuando empiezan, y si es posible que se pongan de manifiesto porque hay ciertos datos atípicos: anomalías. Así, si la causa y el tiempo de ocurrencia de los sucesos se desconocen nos referiremos a anomalías en las series temporales.

Si en la serie se detectan anomalías, durante el proceso anterior, es decir, observaciones no uniformes al resto de observaciones del grupo, se han de tener en cuenta. Una anomalía también es posible que sea producida por un error de medida o por un hecho puntual que normalmente no sucede.

Es importante disponer de procedimientos para detectar las anomalías y eliminar sus efectos, porque podrían influir en el análisis de las series temporales, haciendo que los resultados de la inferencia sean inválidos.

3. Modelos de intervención

La forma más simple de intervenir un modelo consiste en utilizar una variable *dummy*, que tome el valor 1 en el período de intervención y cero en el resto. Pero esta no es la única posibilidad existente, según el período en que empieza la intervención y la forma del impacto, la especificación de los modelos de intervención variará.

Podemos estudiar tres tipos de intervenciones con diferentes variables:

1) Variable escalón

Pueden existir sucesos extraordinarios que afectan la serie temporal cambiando el nivel de forma permanente. En la serie que estamos analizando algún suceso hace que los valores de esta serie, a partir del mismo, disminuyan o aumenten de forma continuada.

Una variable que tomará el valor cero antes del suceso extraordinario y 1 a partir de este momento.

$$I_t \begin{cases} 0 & \text{antes del suceso} \\ 1 & \text{después del suceso} \end{cases}$$

En este caso, es posible especificar el modelo con un cambio de nivel o variable escalón (*step variable*), se representa por $S_t^{t_0}$, donde t_0 es el período temporal en el que empieza el suceso.

2) Variable impulso

Pueden existir efectos transitorios sobre la serie, es decir, sucesos extraordinarios que afectan la serie durante un período de tiempo. Si hay un efecto de este tipo, hemos de observar un valor de gran magnitud en el residuo de la serie. La forma de intervención será:

$$I_t \begin{cases} 0 & \text{período de ocurrencia} \\ 1 & \text{resto de períodos} \end{cases}$$

La variable impulso (*pulse variable*) la representaremos por $P_t^{t_0}$ donde t_0 es el período de ocurrencia del suceso.

Las variables *escalón* y *impulso* se pueden relacionar con la expresión:

$$P_t^{t_0} = (1-L) S_t^{t_0}$$

3) Finalmente las series temporales pueden presentar cambios estructurales en que la tasa de crecimiento de la serie varia a partir de un determinado momento. En este caso es posible introducir una variable escalón que tome los valores 1,2,... a partir del período de este suceso y cero posteriormente.

Alternativamente, podemos introducir una variable impulso en la ecuación de la tasa de crecimiento de la serie.

Para hacer este análisis es necesario conocer el período en que se inician las intervenciones, que puede ser gradual o brusco, y la forma general del impacto y duración del efecto, es decir, como afecta la serie, que puede ser permanente o temporal.

Los modelos de intervención se pueden especificar con la siguiente expresión general:

$$X_t = v(L) I_t^{t_0},$$

donde $I_t^{t_0}$ es la variable impulso o escalón,

$$v(L) = \frac{\omega(L)L^b}{\delta(L)}$$

donde:

$$\omega(L) = \omega_0 - \omega_1L - \omega_2L^2 - \dots - \omega_sL^s$$

$$\delta(L) = 1 - \delta_1L - \delta_2L^2 - \dots - \delta_rL^r$$

Los valores de b , r y s representan los diferentes modelos de intervención. En todos los casos $s = 0$, por tanto,

$$\omega(L) = \omega_0 = \omega.$$

Si generalizamos el modelo simple a un modelo con k intervenciones, escribiremos:

$$X_t = \sum_{j=1}^k v_j(L) I_{jt}^{t_0j}$$

donde:

$$v_j = \frac{\omega_j(L)L^{b_j}}{\delta_j(L)}$$

4. Identificación de los modelos de intervención

A partir de la serie temporal $\{Y_t\}$, en la que algún suceso externo o intervención producen modificaciones en el comportamiento de la serie, es necesario tener en cuenta en la modelización la inclusión de alguna intervención o más según los anteriores modelos.

A partir de la forma general del modelo multiplicativo *SARIMA* (p,d,q) $(P,D,Q)_S$

$$Y_t = \frac{\Theta_Q(L^S)\theta_q(L)}{\Phi_P(L^S)\phi_p(L)(1-L^S)^D(1-L)^d} u_t$$

si incorporamos el modelo de intervención, se transforma y se obtiene la siguiente ecuación:

$$Y_t = v(L)I_t^{t_0} + \frac{\Theta_Q(L^S)\theta_q(L)}{\Phi_P(L^S)\phi_p(L)(1-L^S)^D(1-L)^d} u_t$$

Otra forma de obtener la expresión anterior es:

$$Y_t = v(L)I_t^{t_0} + Z_t,$$

donde:

$$Z_t = \frac{\Theta_Q(L^s)\theta_q(L)}{\Phi_p(L^s)\phi_p(L)(1-L^s)^D(1-L)^d} u_t$$

Generalizando para el modelo con k intervenciones, se obtiene:

$$Y_t = \sum_{j=1}^k v(L)I_{jt}^{t_0j} + Z_t$$

5. Tipos y efectos de las anomalías

Cuando el tiempo y las causas de los factores externos creadores de observaciones atípicas son desconocidas, necesitamos analizar los *outliers* o anomalías de la serie.

Ajustando la presencia de las anomalías podemos, en consecuencia, mejorar la calidad de las predicciones que se obtienen según Hillmer (1984), Ledolter (1989) y Trivez (1992).

Podemos describir diferentes tipos de anomalías¹:

1) Anomalía aditiva (AO) o *additive outlier*: es un suceso que afecta a la serie en un instante temporal.

$$Y_t^\lambda = Z_t + \omega I_t^{t_0}$$

$$I_t^{t_0} \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$$

¹ Hay cuatro tipos de anomalías que se han considerado en la bibliografía por: Hillmer, BellyTiao (1983), Muirhead (1986) Tsay (1988) ChenyLiu (1990), Chen, LiyuHudak (1990), ChenyTiao (1990)y ChenyLiu (1993).

2) Anomalía de cambio de nivel (LS) o *level shift*: es un suceso que afecta una serie con un período dado y donde el efecto es permanente.

$$Y_t^\lambda = Z_t + \frac{1}{1-L} \omega I_t^{t_0} \quad \text{donde } S_t^{t_0} = \omega I_t^{t_0}$$

$$S_t^{t_0} \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases}$$

3) Anomalía innovacional (IO) o *innovational outlier*: es un efecto que se propaga de conformidad con el modelo *SARIMA* del proceso, que afecta todos los valores observados después de su ocurrencia.

$$Y_t^\lambda = Z_t + Z_t \omega I_t^{t_0}$$

4) Anomalía de cambio temporal (TC) o *temporary change*: es un suceso que tiene un impacto inicial, y su efecto decae exponencialmente de conformidad con un factor de amortiguación.

$$Y_t^\lambda = Z_t + \frac{1}{1-\delta L} \omega I_t^{t_0} \quad 0 < \delta < 1$$

Como podemos apreciar, las anomalías AO y LS son casos particulares de TC.

Si $\delta = 0$ anomalía aditiva

Si $\delta = 1$ anomalía de cambio de nivel

En términos genéricos, si denominamos Y_t a la serie observada y Z_t a la serie libre de anomalías suponiendo que Z_t sigue un proceso estacional multiplicativo *SARIMA* $(p,d,q)(P,D,Q)_S$:

$$\Phi_P(L^S) \phi_p(L) (1-L^S)^D (1-L)^d Z_t = \Theta_Q(L^S) \theta_q(L) u_t$$

Se considera que son operadores estacionarios e invertibles, sin factores comunes, y donde u_t es una secuencia de ruidos blancos.

Otra forma de escribir la ecuación es:

$$Z_t = \frac{\Theta_Q(L^s)\theta_q(L)}{\Phi_p(L^s)\phi_p(L)(1-L^s)^D(1-L)^d} u_t$$

Si tenemos en cuenta que una serie temporal puede contener diferentes tipos de anomalías, por ejemplo k , el modelo general se puede escribir así:

$$Y_t = \sum_{j=1}^k \omega_j v_j(L) I_{jt}^{t_{0j}} + Z_t$$

Donde:

Z_t : es la serie libre de anomalías

$v_j(L)$:

1 per una AO en $t = t_{0j}$

Z_t per una IO en $t = t_{0j}$

$1/(1-L)$ per una LS en $t = t_{0j}$

$1/(1-\delta L)$ per una TC en $t = t_{0j}$

donde:

δ es un factor de amortiguación que está entre 0 y 1.

Para detectar las anomalías se proponen diferentes procedimientos iterativos, como el de Chang y Tiao (1983) y Tsay (1986), con el cual la detección de las anomalías se hace a partir de los residuos del modelo univariante previamente especificado y estimado.

Más adelante, Chen y Liu (1990) y (1993) han hecho propuestas desarrollando un método según el cual la detección de los *outliers*, la estimación conjunta de los parámetros del modelo y los efectos de las anomalías, se lleva a cabo en la fase de estimación del modelo univariante.

La opción de intervención permite tratar las anomalías sin que sean eliminadas. Si las anomalías se intervienen, se corrige la anomalía, por diferentes métodos según el tipo de anomalía, con el resto de datos. Finalmente se obtienen las estimaciones de las anomalías, que son los valores corregidos.

Con el tratamiento informático, la detección y la corrección de los efectos atípicos se puede hacer de forma automática o puede no hacerse. Si se opta por la opción automática, es posible escoger el tipo de atípicos que se desea identificar y el rango de la muestra en que se realiza la búsqueda. Si esta búsqueda es automática se recomienda hacerla sobre toda la muestra.

Se llevaron a cabo simulaciones con programas de identificación automática con el FORENET y el FORCE IV, realizados en el marco de un proyecto de investigación que tiene como objetivo integrar diferentes paquetes estadísticos existentes, como: SCA, SEATS y TRAMO. Con este programa hay desarrollada una herramienta de previsión que integra la metodología Box-Jenkins, la desestacionalización de series temporales, el análisis de coyuntura, etc. Contiene los métodos más eficientes en referencia al tiempo de respuesta del Dr. Lewandowski, de manera que permiten tratar un gran número de series en un espacio corto de tiempo.

En un segundo paso, y después de ver las opciones de cada uno de los programas anteriores, el software escogido para llevar a cabo el

tratamiento informatizado de los datos se ha utilizado el programa FORCE IV² ya que posee las ventajas de SCA, SEATS y TRAMO, y que además permite la identificación automática del modelo a partir de unas coordenadas fijadas, una mejor presentación gráfica de series, correlogrames y periodograma de las previsiones. Además de facilitar, las medidas para evaluar la bondad del ajuste de forma más exhaustiva.

La identificación de las series es automática con la función IDAUT del programa FORCE IV, siguiendo la aplicación informática del paquete estadístico SCA, aunque TRAMO y SEATS. Este programa genera previsiones a partir de los datos del pasado de la serie temporal per poder extrapolar después los datos al futuro, con la siguiente secuencia:

- Busca el modelo que mejor se ajusta a las observaciones.
- Identifica el modelo comparando los correlogrames FAS (Función de autocorrelación simple) y FAP (Función de autocorrelación parcial) en modelos teóricos de series temporales.

Se realizan las transformaciones necesarias para conseguir la estacionariedad, y después escoger la estructura estocástica con el orden de p , P , q y Q , para formular el modelo o posibles modelos para elaborar las estimaciones y después evaluar la bondad.

² El programa informático "FORCE IV Advanced Methods for Forecasting" está desarrollado por la sección de Tècniques Quantitatives per a la Gestió de la UPC.

En algunos casos se ha detectado una cierta tendencia a sobrediferenciar, aunque los efectos de una sobrediferenciación son menos graves que los que tendría una infradiferenciación.

La primera opción es escoger el método de estimación de los parámetros entre:

- Estimación con máxima verosimilitud exacta
- Estimación con mínimos cuadrados no condicionales (incondicionales)

En el primer caso, queremos maximizar la función de la verosimilitud, mientras que en el segundo, minimizamos la suma de los cuadrados de los residuos que se obtienen modelando.

La siguiente opción es el filtro a utilizar, en este caso se puede escoger entre:

a) Filtro de las raíces cuadradas: es una versión del filtro de Kalman, que se utiliza para series no diferenciadas.

b) Algoritmo de Morf, SidhuyKailath, mejorado por Mélard. Es adecuado cuando la serie ha estado diferenciada.

c) Filtro de Kalman. Se utiliza para series no diferenciadas.

d) Filtro de mínimos cuadrados condicionales. Es adecuado cuando nos podemos encontrar con dificultades numéricas, pero es mucho más lento.

Estos filtros se utilizan para calcular la máxima verosimilitud o los mínimos cuadrados con más precisión si se quiere efectuar desestacionalización.

Seguidamente podemos escoger que a partir de los polinomios que obtenemos modelando, nos calcule el valor de las raíces.

El sistema experto evalúa los parámetros y valora el modelo con sus reglas, que el usuario puede cambiar. En esta segunda parte se necesita la ayuda de herramientas de cálculo potentes, con un tratamiento informatizado de los datos. Hay la opción de definir el t-valor por debajo del cual los parámetros se han de reestimar.

Las dos primeras fases de la tercera etapa son automáticas con el programa informático, mientras que la tercera se ha de efectuar de manera manual.

También tiene la opción de detectar y corregir diversos tipos de anomalías y de realizar las intervenciones de las posibles anomalías del modelo, así como encontrar interpolaciones óptimas de las observaciones perdidas.

El tratamiento de las anomalías permite modificar cuatro parámetros. El primer parámetro que deja escoger es el correspondiente a la detección o no de anomalías que presenta las opciones:

a) Detección automática de los cuatro tipos de anomalías IO, AO, LS y TC; y una vez corregida la serie con las anomalías encontradas en la primera iteración el programa vuelve a identificar automáticamente el modelo, y busca nuevas anomalías en caso que el modelo haya cambiado.

b) No realizar la detección de las anomalías.

El segundo parámetro hace referencia a como se hacen las correcciones de las anomalías. En el caso de escoger la detección y la corrección de anomalías aparecen dos opciones más, en las cuales se ha de escribir el valor que se desea:

1) El valor crítico para la detección de anomalías.

2) La reducción del valor crítico para la detección de las anomalías. Es a decir, el porcentaje en que el valor de la función del número de observaciones se reduce en cada iteración.

Finalmente, se puede escoger el método para realizar la estimación del modelo y la detección de las anomalías:

a) Utilizar el método de Hannan-Rissanen para la estimación de los parámetros en la detección y la corrección automática de anomalías.

b) Utilizar el método de la máxima verosimilitud.

c) Utilizar primero el método de mínimos cuadrados no condicionales, y después el de la máxima verosimilitud en la estimación de las raíces del modelo automático. La estimación de las anomalías se hace utilizando el procedimiento de Hannan-Rissanen.

d) Igual que en el caso c) excepto para la estimación de las anomalías, ya que se hace con el método exacto de la máxima verosimilitud.

Una vez evaluados los modelos, y según el criterio AIC, se decide el modelo adecuado. Según las medidas de evaluación de la capacidad predictiva estas influyen de manera determinante en la bondad de las previsiones que se calculan posteriormente.

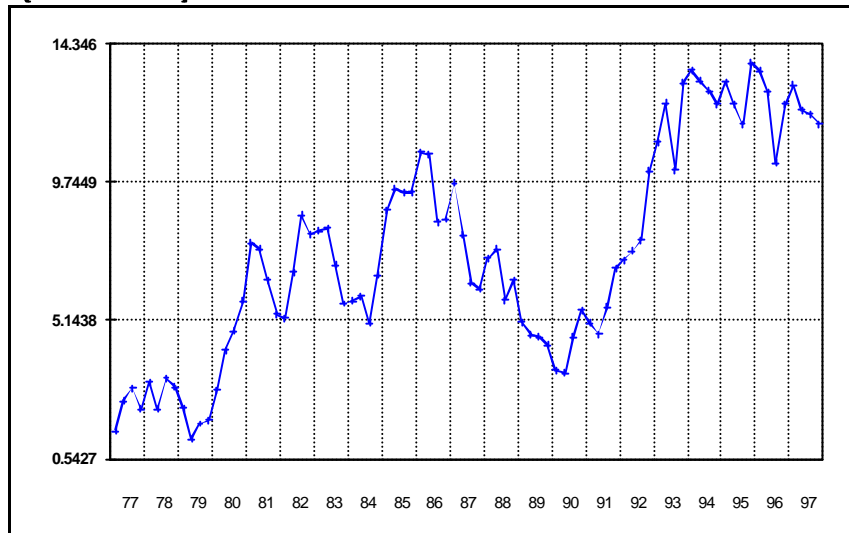
Las predicciones correspondientes se determinan automáticamente, a partir de la modelización efectuada que serán óptimas en caso que el modelo calculado sea óptimo.

A partir de los datos de las variables se podrá comprobar cual es el escenario futuro de la provincia de Lleida, en relación con el mercado laboral. Con esta metodología utilizada se presenta un estudio empírico con las diferentes fases hasta conseguir la obtención de las previsiones para dos variables.

6. An.lisis empirica

Para efectuar una aplicación práctica presentamos la evolución de la serie de población parada (U_t) y ocupada en el sector industria (E_t), trimestralmente, en la provincia de Lleida y con observaciones desde 1977 hasta 1997.

Gr.fico 1. Serie de poblaciøn parada en la provincia de Lleida (1977-1997)



En primer lugar, si se analiza el gráfico de la serie de población parada en la provincia de Lleida, se observa su evolución y en particular, el diferente comportamiento antes de 1990 y a partir de esta fecha hasta 1993.

A continuación, en la tabla 1, se presenta el modelo escogido junto con la estimación de los parámetros.

Tabla 1. Estimación de la serie U_t

SARIMA (0,1,0) (1,0,0)₄

$$(1-L)(1-0,2357 L^4) U_t = u_t$$

(2,17)

AIC : 93,121 $R^2 = 0,906$

Desviación residual $\sigma_u = 1,07974$

En el proceso de estimación se detectan anomalías, y por tanto es necesario conocer su naturaleza, además de identificar el periodo en que empiezan y la forma general del impacto.

Se detectan dos anomalías de tipo AO en el tercer trimestre de 1993 y de 1996, y que corresponden a los períodos 67 y 79 respectivamente. Así pues, son intervenidas y la estimación con anomalías de la serie se presenta en la tabla 2.

Tabla 2. Estimación con anomalías de la serie U_t

SARIMA (0,1,0) (1,0,0)₄

$$(1-L)(1-0,1718 L^4)U_t = (1+0,3353L) u_t + (-2,5854 AO_{67}(1-L) -1,7817 AO_{79}(1-L))$$

(1,51)

(-3,17)

(4,75)

(3,25)

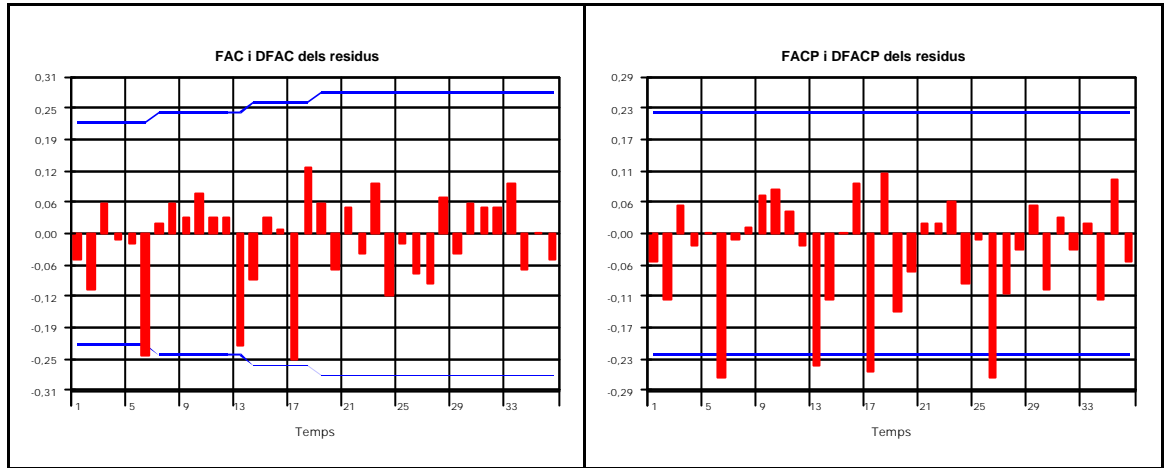
AIC : 70,353 $R^2 = 0,931$

Desviación residual $\sigma_u = 0,923077$

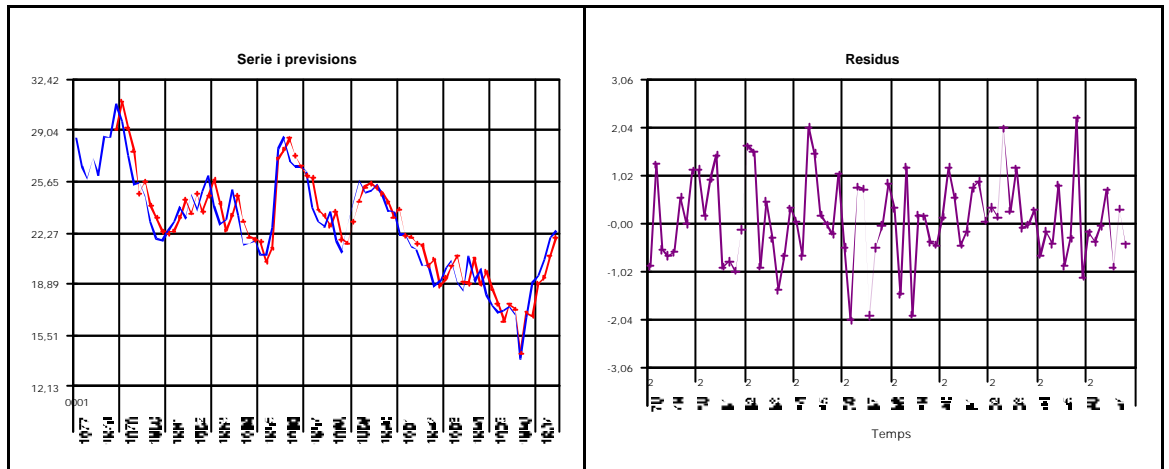
Si comparamos los dos modelos comprobamos que con el tratamiento de las anomalías se reduce la estimación de la desviación residual σ_u , que pasa de un 1,07974 a un 0,923077, es decir, en el segundo caso se obtiene un valor inferior al modelo anterior sin intervención. También aumenta la capacidad predictiva

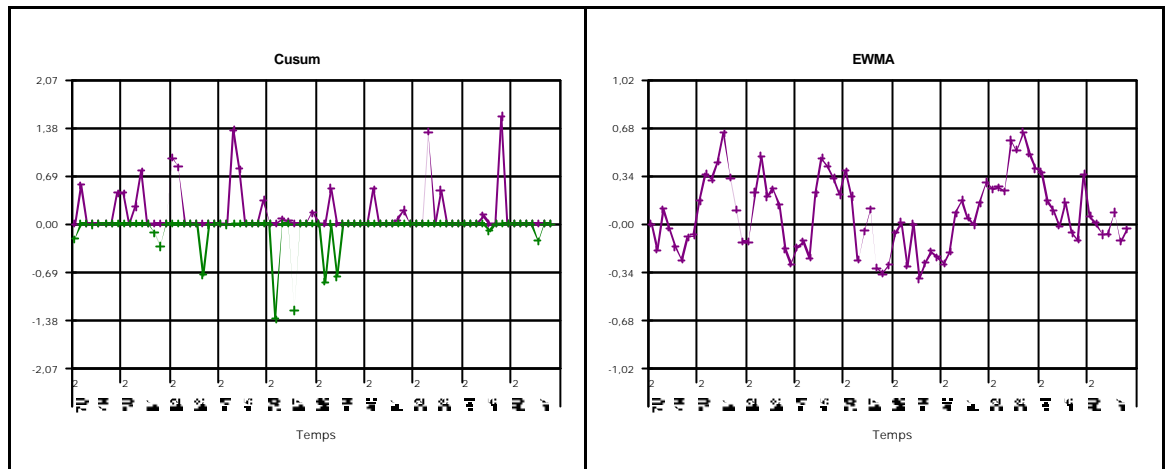
del modelo y la R^2 que proporciona información sobre el ajuste, aumenta hasta un 93,1%.

Gr.fico 2. Correlogramas de los residuos de la serie U_t



Gr.fico 3. Representaciçn de la modelizaciçn de la serie U_t





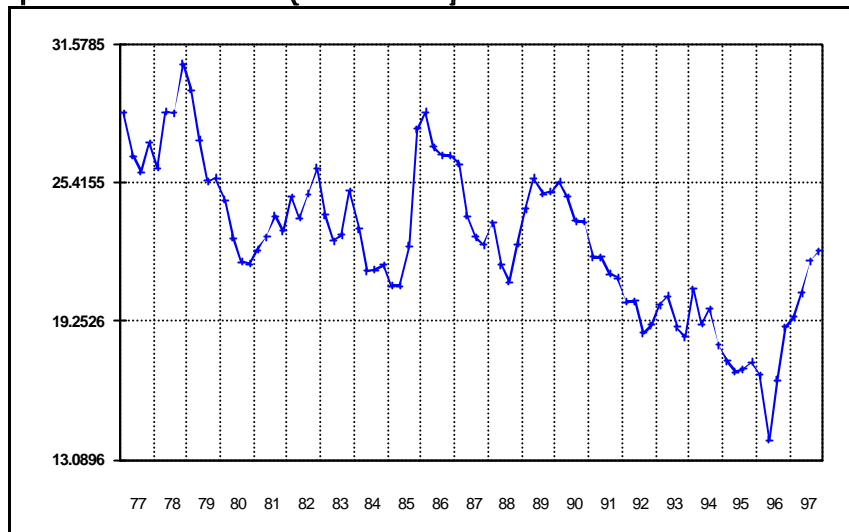
En la siguiente tabla 3, se indican las previsiones para los próximos 12 trimestres, de la serie población desempleada en la provincia de Lleida, con una clara tendencia a la estabilidad.

Tabla 3. Previsiones de la serie U_t

PREVISIONES (en miles de personas)				PERIODOGRAMA
Año	Tri.	Previsión	Intervalo	
1998	1	12,51	(11,59-13,43)	
1998	2	12,62	(11,08-14,16)	
1998	3	12,48	(10,50-14,45)	
1998	4	12,46	(10,13-14,78)	
1999	1	12,54	(9,83 -15,25)	
1999	2	12,56	(9,49 -15,63)	
1999	3	12,54	(9,14 -15,93)	
1999	4	12,53	(8,84 -16,22)	
2000	1	12,55	(8,58 -16,52)	
2000	2	12,55	(8,31- 16,79)	
2000	3	12,55	(8,06- 17,04)	
2000	4	12,55	(7,82 -17,27)	

En segundo lugar la serie que se analiza corresponde a la población ocupada en la industria en la provincia de Lleida.

Gráfico 4. Serie de población ocupada en la industria en la provincia de Lleida (1977-1997)



El gráfico presenta la evolución del total de ocupados en el sector industria entre 1977 i 1997. Se observa una tendencia descendente hasta el segundo trimestre de 1996, a partir del cual hay un aumento en la ocupación industrial. También, presenta oscilaciones, que podrían relacionarse con los ciclos económicos.

En una primera fase de identificación después de estudiar la evolución gráfica y los correlogramas de la serie E_t , continuamos analizando por un lado, la parte regular y por otro lado la estacional de orden 4.

En la parte regular es necesaria una diferenciación que convierte en estable la serie con tendencia. En la parte estacional no es necesario aplicar diferencias para convertirla en estacionaria.

Tabla 4. Estimación de la serie E_t

SARIMA (0,1,0) (1,0,0)₄

$$(1-L)(1+0,28344L^4) \ln E_t = u_t$$

(-2,51)

AIC : -377,248 $R^2 = 0,831$

Desviación residual $\sigma_u = 0,065007$

La serie presenta 2 anomalías, una anomalía de cambio de nivel, LS en el cuarto trimestre de 1985, que corresponde al período 36 y otra anomalía aditiva, AO en el segundo trimestre de 1996, que corresponde al período 78. En la tabla 5 se muestra la estimación del modelo con anomalías.

Tabla 5. Estimación con anomalías de la serie E_t

SARIMA (0,1,0) (1,0,0)₄

$$(1-L)(1+0,2922 L^4) \ln E_t = u_t + (-0,1852 AO_{78} (1-L)+0,1599/(1-L)*LS_{36} (1-L))$$

(-2,58)

(-5,47)

(3,37)

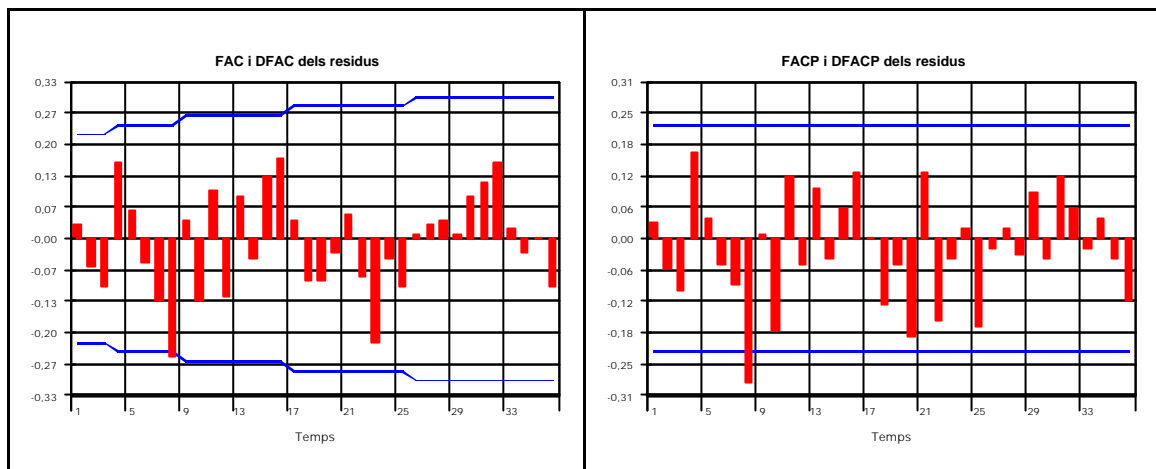
AIC : -386,146 $R^2 = 0,89$

Desviación residual = 0,0487806

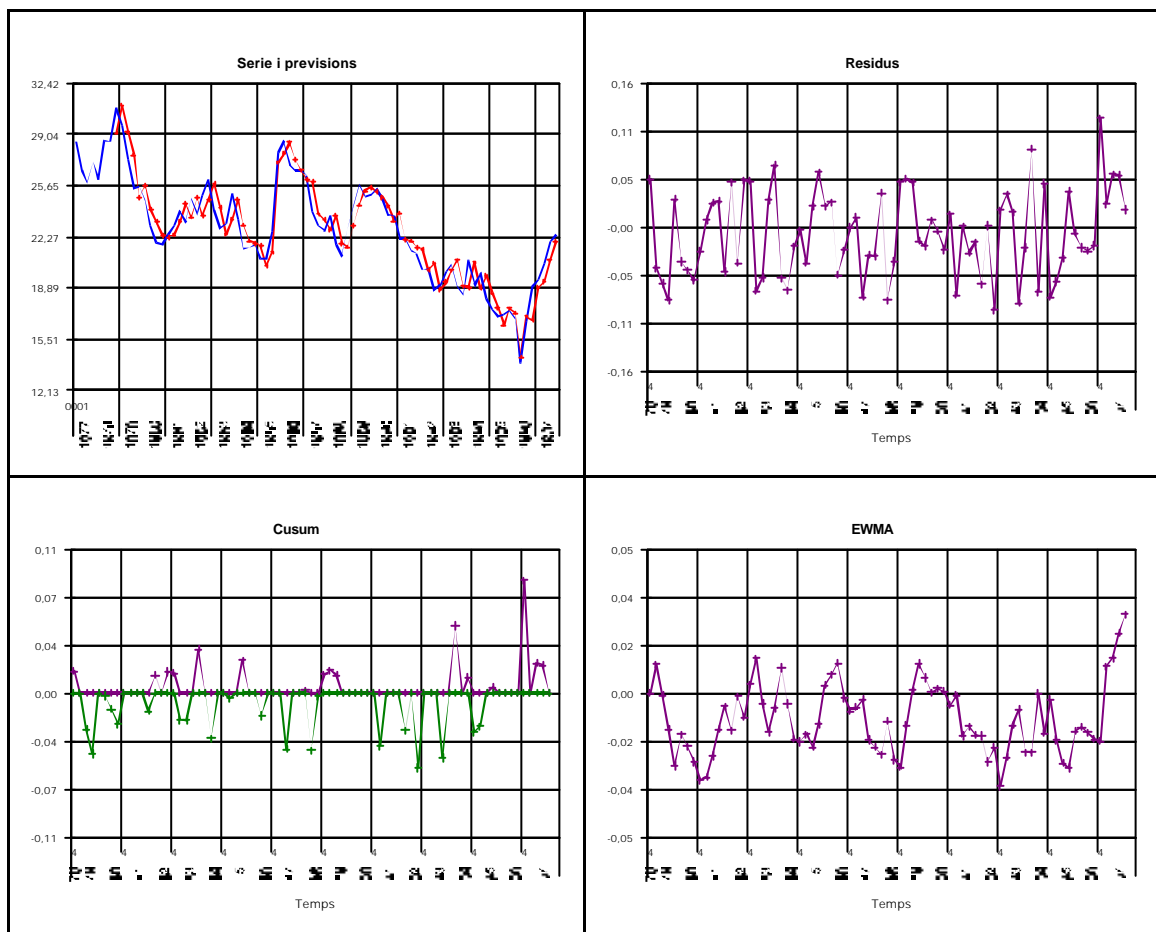
La serie de ocupación industrial no sigue en la parte regular ninguna estructura, en cambio en la parte estacional podemos apreciar un comportamiento AR(1), con un coeficiente significativo (Estadístico t-value >2).

La desviación residual ha disminuido respecto al primer modelo, y el coeficiente R^2 ha aumentado.

Gr.fico 5. Correlogramas de los residuos de la serie EI_t



Gr.fico 6. Representaçiçn de la modelizaciçn de la serie EI_t



En último lugar, en la tabla 6, presentamos el conjunto de previsiones, para los próximos tres años, con detalle trimestral para la población ocupada en el sector industrial en Lleida, observándose en primer lugar un cierto descenso en la ocupación y una posterior recuperación en los últimos periodos.

Tabla 6. Previsiones de la serie E_{it}

PREVISIONES (en miles de personas)				PERIODOGRAMA
Año	Tri.	Previsión	Intervalo	
1998	1	21,30	(20,26-22,34)	
1998	2	20,52	(19,11-21,94)	
1998	3	20,41	(18,69-22,13)	
1998	4	20,11	(18,16-22,07)	
1999	1	19,74	(17,60-21,89)	
1999	2	19,65	(17,31-21,99)	
1999	3	19,96	(17,49-22,43)	
1999	4	20,19	(17,60-22,79)	
2000	1	20,24	(17,55-22,93)	
2000	2	20,35	(17,56-23,14)	
2000	3	20,48	(17,59-23,38)	
2000	4	20,53	(17,54-23,51)	

7. Conclusiones

En el presente trabajo se han efectuado en primer lugar unas indicaciones teóricas sobre el análisis de series temporales aplicadas al mercado de trabajo, realizando un estudio de las definiciones, del análisis y de los modelos de intervención.

Destacamos la aplicación informática utilizada, el FORCE IV, que permite la identificación automática de la serie y facilita este primer paso para la posterior predicción de las series del mercado de trabajo.

El estudio se ha centrado en la provincia de Lleida, durante el periodo 1977-1997 para efectuar predicciones sobre los próximos tres años. En concreto, en este trabajo se presentan dos series para

realizar el análisis empírico, correspondientes a la población desocupada y a la población ocupada en el sector industrial, presentado la modelización con estimación de las anomalías y la posterior predicción.

En ambos casos se han detectado anomalías, de tipo aditivo y de cambio de nivel, que incorporadas a la estimación del modelo, nos facilitan una predicción más eficiente para los próximos tres años.

Para la serie de población desocupada, se predice una tendencia estable y para la serie de población ocupada industrial, un retroceso inicial, y una recuperación en los últimos trimestres.

8. Bibliografía

- BOX, G.E.P. i JENKINS, G.M. (1970) *Time Series Analysis: Forecasting and control*. Holden-Day.
- BOX, G.E.P., JENKINS, G.M. i REINSEL, G.C. (1994) *Time Series Forecasting and control*. Prentice-Hall, Englewood Cliffs, New Jersey.
- CAJA DE MADRID (6/1995-12/1997) *"Previsión y seguimiento de la Economía Española"*. Madrid.
- INSTITUTO NACIONAL DE ESTADÍSTICA. *"Encuesta de Población Activa. Resultados detallados. Primer trimestre 1977-Cuarto trimestre 1997"*. Madrid.
- Dirección General de Previsión y Coyuntura (1988) "Nota sobre el enlace provisional de las series laborales tras la reforma de la EPA en el segundo trimestre de 1987". Ministerio de Economía y Hacienda.
- AZNAR, A. i TRIVEZ, F.G. (1993) *Métodos de predicción en Economía*, Ariel Económica.
- BOX, G.E.P. i TIAO, G.C. (1975) *"Intervention Analysis with applications to Economic and Environmental problems"* Journal of the American Statistica Association, 70, pàgs. 70-79.
- CHANG, I. i TIAO, G.C. (1983) *Estimation of time series parameters in the presence of outliers*, Technical Report 8, University of Chicago, Statistics Research Center.

- ESPASA, A. (1977) "El problema de la desestacionalización de las series económicas: métodos utilizados y su interpretación" Boletín de Estudios Económicos, vol XXXII, 101, pàgs. 461-478.
- ESPASA, A. (1984-a) *El ajuste estacional en series económicas*, Banco de España - Documento de Trabajo, 8410.
- GÓMEZ, V. i MARAVALL, A. (1994) "Estimation prediction, and interpolation for non stationaty series with the Kalman filter" Journal of the American Statistical Association, Vol. 89, pàgs. 611-624.
- HAMILTON, J.D. (1994) *Time Series Analysis*, Pricenton University Press.
- HARVEY, A.C. (1989) Forecasting, structural time series models and the Kalman filter, *Cambridge University Press, Cambridge*.
- MARAVALL, A. (1984-a) "Notas sobre la extracción de una señal en un modelo ARIMA". Banco España - Documento de Trabajo nùm 8404.
- MARAVALL, A. (1985) "On structural Time Series Models and the Characterization of Components". *Journal of Business and Economic Statistics*. Vol. 3, nùm 4, pàg. 350-355.
- MARAVALL, A. (1987) "Descompsición de series temporales: especificación, estimación e inferencia". -*Estadística Española*, nùm 114, pàg. 11-69.
- MARAVALL, A. (1988-a) "Two papers on ARIMA signal extraction". Banco de España - Documento de Trabajo nùm 8801.
- MARAVALL, A. (1988-b) "A note on revisions in ARIMA-based signal extraction". Banco de España - Documento de Trabajo nùm 8421.
- MARAVALL, A. (1989) "La extracción de señales y el análisis de coyuntura". *Revista Española de Economía*. Vol 6, nùm 1-2, pàg. 109-132.
- STOCK, J.H. i WATSON, M.W. (1988) "Variable trends in economic time series". *Journal of Economic Perspectives*. Vol. 2, nùm 3, pàg. 147-174.
- THEIL, H. i WAGE, S. (1964) "Some observations on adaptive forecasting". *Management Sciencie*, nùm 10, pàg. 198-206.
- WEI, W.W.S. (1994) *Time Series Analysis*. Addison-Wesley Pub. Comp. Inc.
- WIENER, N. (1949) *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. -John Wiley.
- WILSON, G.J. (1969) "Factorization of the generating function of a pure moving average process" *S.I.A.M.- Journal of Numerical Analysis*, nùm 6, pàg. 1-7.
- YULE, G.U. (1897) "On the theory of correlation". *Journal of the Royal Statistical Society*, nùm 60, pàg. 812-854.