

# ESTACIONALIDAD Y CICLOS EN LAS SERIES PROVINCIALES DE PERNOCTACIONES<sup>1</sup>

Enrique M. Quilis\*

Rafael Frutos Vivar\*\*

\* Instituto Nacional de Estadística

\*\* Agencia Tributaria y Universidad de Alcalá

Junio 2000

## RESUMEN

En este trabajo se analizan las pautas de aglomeración que se forman a nivel provincial en los factores estacionales de las series de pernoctaciones hoteleras, mediante técnicas de análisis de conglomerados. Dichas pautas sirven para la elección de series representativas cuyas fluctuaciones cíclicas son examinadas con el fin de verificar su grado de comunalidad. De esta forma se comprueba, mediante una aplicación real, en qué medida la información estacional es relevante para el estudio del ciclo.

**Palabras clave:** estacionalidad, ciclos, modelos ARIMA, extracción de señales, análisis de conglomerados.

---

<sup>1</sup> Se agradecen los comentarios de Ana M<sup>a</sup> Abad, Salvador Ortíz y de participantes en seminarios del Instituto Gallego de Estadística y de la Universidad Autónoma de Madrid. Las opiniones expresadas corresponden a los autores y no reflejan necesariamente ni las de la Agencia Tributaria ni las del Instituto Nacional de Estadística.

## 1. INTRODUCCION

El papel de la estacionalidad en el análisis aplicado de series temporales es crucial, ya que su presencia influye de forma decisiva y permanente en las propiedades estadísticas de las series, de manera que su tratamiento incorrecto conduce a serios errores de especificación en los modelos estadístico-econométricos, con efectos adversos tanto en su capacidad descriptiva como en su rendimiento predictivo.

Esta relevancia en cuanto a la modelización contrasta agudamente con la escasa importancia que suele tener la estacionalidad en el análisis económico. En efecto, el tipo de análisis más habitual considera que la estacionalidad representa elementos exógenos a los procesos de decisión de los agentes económicos. De esta forma, dichos agentes resuelven sus problemas de optimización incorporando restricciones dependientes de la estación, véase Todd (1990) para una exposición de este enfoque. Así, la estacionalidad es un elemento añadido cuya presencia dificulta una adecuada comprensión del fenómeno económico objeto de estudio y, por lo tanto, el interés se centra en las series desestacionalizadas o de tendencia-ciclo.

Por el contrario, algunos analistas consideran que la diferencia entre estacionalidad y tendencia-ciclo antes expuesta es artificial. Esta línea de desarrollo extiende la conocida crítica a la distinción entre tendencia y ciclo basada en los modelos estocásticos de crecimiento óptimo, véase Prescott (1986). Desde este punto de vista, los agentes económicos resuelven sus programas de optimización de una forma indiferenciada, de manera que los elementos estacional y no estacional de sus decisiones son aspectos distintos de un único proceso de decisión, siendo meramente facetas diferentes de una misma respuesta a impulsos comunes. En Barsky y Miron (1989) se defiende persuasivamente este enfoque.

Por otra parte, desde un punto de vista aplicado, la estacionalidad puede servir como guía en la elección del nivel de desagregación más apropiado. Esta aplicación es especialmente pertinente en el caso espacial, ya que se puede decidir qué agrupaciones territoriales seleccionar siguiendo un criterio objetivo y con significado estadístico. En particular, en este trabajo se examina si la agrupación espacial sugerida por el análisis de conglomerados de los factores estacionales posee o no un comportamiento cíclico común. De esta forma, se somete a comprobación si una determinada “segmentación del mercado” asociada a factores permanentes (estacionalidad) obedece a fluctuaciones comunes o, por el contrario, predominan los factores de tipo idiosincrásico.

La estructura del trabajo es la siguiente. En la segunda sección se exponen los métodos de estimación de los componentes estacional y cíclico. A continuación, se presenta la metodología de formación de conglomerados seleccionada. En la cuarta sección se describen los datos provinciales de pernoctaciones y, en la quinta, se presentan los resultados empíricos. El trabajo termina con un apartado de conclusiones.

## 2. ESTIMACION DE LOS COMPONENTES ESTACIONAL Y CICLICO

El procedimiento utilizado para estimar los componentes estacional y cíclico de las series temporales analizadas consta de tres etapas: corrección de los efectos asociados a las observaciones atípicas y efectos de calendario, extracción basada en modelos ARIMA de las señales estacional y de tendencia-ciclo y, por último, estimación del ciclo por medio de un filtro de paso en banda aplicado a la serie de tendencia-ciclo obtenida en la etapa anterior. A continuación se describe brevemente cada fase.

Se considera que la serie temporal observada, en logaritmos, ( $Z_t$ ) puede ser expresada de acuerdo con la siguiente expresión:

$$[1] \quad Z_t = O_t + N_t$$

donde  $O_t$  representa una combinación de modelos de intervención asociados a factores de tipo extraordinario que afectan a la serie de manera no recurrente junto con los efectos de calendario vinculados con el ciclo semanal y la Pascua móvil. Por su parte,  $N_t$  caracteriza el comportamiento estocástico de la serie. La expresión formal es:

$$[2] \quad O_t = \sum_{h=1}^k V_h(B) I_t^{T_h} + \sum_{j=1}^7 \beta_j D_{j,t} + \gamma P(\tau)_t$$

donde  $I_t^{T_h}$  es una variable binaria de tipo impulso que adopta un valor unitario en la observación  $T_h$  y nulo en los restantes, siendo  $T_h$  la observación en que tiene lugar el acontecimiento atípico o extraordinario. El filtro  $V_h(B)$  recoge los efectos dinámicos asociados a la observación anómala. En este trabajo se consideran tres posibles tipos de atípicos: aditivos, transitorios y cambios de nivel. Análisis detallados de las características de estos tipos de *outliers* y de los procedimientos secuenciales de detección de los mismos se encuentran en Abraham y Box (1979), Chang *et al.* (1988), Chen *et al.* (1990), Chen y Liu (1991), Tsay (1986, 1988) y Gómez y Maravall (1998b).

Asimismo,  $D_{j,t} = [(\text{número de días de tipo } j \text{ en el mes } t) - (\text{número de domingos en el mes } t)]$ , con  $j = \text{lunes, ..., sábados}$  y  $D_{7,t}$  es la diferencia entre la duración del mes  $t$  y la duración media de todos los meses. Acerca del tratamiento y modelización del ciclo semanal véase Hillmer (1982), Bell y Hillmer (1983) y Hillmer *et al.* (1983), entre otros. Finalmente,  $P(\tau)_t$  expresa la proporción que representa la semana de Pascua en el mes  $t$ , habiéndose considerado que su efecto se percibe en los  $t$  días anteriores al Domingo de Resurrección. En este trabajo se asume que  $\tau=8$ . Un análisis detallado de esta clase de efectos se encuentra en Liu (1980, 1983) y en Hillmer *et al.* (1983).

La especificación del componente estocástico sigue una representación autorregresiva, integrada y de medias móviles (ARIMA) de tipo multiplicativo (Box y Jenkins, 1976):

$$[3] \quad N_t = \frac{\theta_q(B)\theta_Q(B^{12})}{\phi_p(B)\phi_P(B^{12})(1-B)^d(1-B^{12})^D} a_t$$

donde  $\phi_p(B)$  y  $\theta_q(B)$  son, respectivamente, polinomios de orden  $p$  y  $q$  en el operador de desfases  $B$ , y  $\phi_P(B^{12})$  y  $\theta_Q(B^{12})$  son polinomios de orden  $P$  y  $Q$  en  $B^{12}$ . Las expresiones  $(1-B)^d$  y  $(1-B^{12})^D$  son operadores de diferenciación regular y estacional controlados por los parámetros enteros  $d$  y  $D$ , respectivamente. Por último,  $a_t$  es una secuencia de ruido blanco gaussiano con esperanza nula y desviación típica constante  $\sigma_a$ .

A su vez, el término estocástico  $N_t$  admite una descomposición según la hipótesis de los componentes subyacentes en tendencia-ciclo ( $P_t$ ), estacionalidad ( $S_t$ ) e irregularidad ( $I_t$ ):

$$[4] \quad N_t = P_t + S_t + I_t$$

Una vez estimado el modelo ARIMA con análisis de intervención (AI) descrito en [1]-[3], es posible extraer tanto una señal estacional como una de tendencia-ciclo aplicando filtros de error cuadrático medio mínimo compatibles con dicho modelo ARIMA, siguiendo la propuesta de Burman (1980), Hillmer y Tiao (1982), Hillmer *et al.* (1983) y Maravall (1987, 1993a, 1993b, 1994). De esta manera se obtiene una estimación de los componentes de interés adaptada a las propiedades de la serie tal y como vienen recogidas en el modelo ARIMA y, merced al principio de descomposición canónica, lo más libre posible de elementos irregulares de tipo ruido blanco.

La expresión general de este proceso de filtrado es, para la estacionalidad:

$$[5a] \quad \hat{S}_t = V_s(B, F) \hat{N}_t = k_s \Pi(B) \Pi(F) \Psi_s(B) \Psi_s(F) \hat{N}_t$$

Y, para la tendencia-ciclo:

$$[5b] \quad \hat{P}_t = V_p(B, F) \hat{N}_t = k_p \Pi(B) \Pi(F) \Psi_p(B) \Psi_p(F) \hat{N}_t$$

donde  $k_i$  ( $i=S,P$ ) son parámetros que normalizan las funciones de ganancia de los filtros,  $\Pi(B)$  es la expansión autorregresiva del modelo ARIMA de  $N_t$ ,  $\Psi_i(B)$  ( $i=S,P$ ) es la expresión de medias móviles del modelo teórico de los componentes y  $\hat{N}_t$  es la estimación del componente estocástico obtenida al eliminar de la serie observada  $Z_t$  sus elementos deterministas  $O_t$ . Una exposición detallada de estos filtros se encuentra en Maravall (1987).

La señal de tendencia-ciclo así obtenida permite la estimación de un componente cíclico al aplicar a aquélla un filtro fijo de paso en banda diseñado desde el dominio de la frecuencia. Dicho filtro es la adaptación simetrizada de uno de la familia Butterworth, especificado para aproximar con una precisión dada *a priori* a uno cíclico de tipo ideal. Detalles de este procedimiento se encuentran en Gómez (1998a).

De esta forma, el componente cíclico se obtiene según la siguiente expresión:

$$[6] \quad \hat{C}_t = H_c(B, F) \hat{P}_t = H_c(B, F) k_p \Pi(B) \Pi(F) \Psi_p(B) \Psi_p(F) \hat{N}_t$$

donde  $H_c(B, F)$  es el filtro cíclico antes mencionado y  $\hat{C}_t$  es la señal cíclica.

Este método bietápico puede ser interpretado de forma bayesiana, ya que combina información *a priori* (un filtro fijo basado en uno tipo Butterworth) con la contenida en la tendencia estimada a partir de la muestra (por medio de un filtro adaptable de tipo Wiener-Kolmogorov). De esta manera, se obtiene la información *a posteriori*: una serie de fluctuaciones de periodicidad comprendida entre dos y ocho años.

Las principales ventajas de este enfoque son:

- a- La estimación de los componentes es compatible con las propiedades agregadas de las series, de forma que se evita la inducción de fenómenos espurios como, por ejemplo, la estimación de un componente estacional en una serie que no tiene estacionalidad.
- b- El filtro usado en la estimación se adapta a las características de la serie observada, de forma que series con características estacionales distintas tendrán asimismo filtros distintos.
- c- El preprocesamiento de las series mediante el modelo ARIMA-AI permite estimar los componentes estocásticos sin la influencia distorsionadora asociada a las observaciones atípicas y a los efectos de calendario, lo que redundará en una mejor estimación de los mismos.

### 3. FORMACION DE CONGLOMERADOS

Una vez estimados los componentes estacionales de las series que se desea analizar, se procede a formar grupos de series con patrones estacionales similares. La formación de estas agrupaciones se realiza mediante el análisis de conglomerados, véase Everitt (1993) para una excelente exposición.

El análisis de conglomerados es una técnica estadística de tipo exploratorio y no paramétrica que agrupa los objetos atendiendo a su semejanza. Trata, por lo tanto, de identificar pautas en los datos utilizando muy poca información *a priori* y sin basarse en modelo alguno sobre el proceso generador de las observaciones. Estas dos características dotan al análisis de grupos de una gran flexibilidad, ya que se adapta muy bien a situaciones difusas y a aquellas que requieren regularidades empíricas como paso previo a la elaboración de modelos estadísticos. Asimismo, el análisis de conglomerados resulta especialmente útil cuando se desea reducir drásticamente la dimensión de grandes masas de información, de forma que se selecciona un representante de cada grupo en lugar de la población completa. En este sentido, este tipo de análisis se asemeja a un muestreo endógeno en el que los propios objetos determinan su representante.

El planteamiento formal del problema es el siguiente. Sea  $x_{i,j}$  el valor observado de la variable  $j$  (con  $j=1..p$ ) en el objeto  $i$  (con  $i=1..n$ ). En este trabajo, las  $p$  variables son los doce factores estacionales correspondientes al año central (1991) y los  $n$  objetos son las cincuenta provincias del Estado español. Asumiendo que se han formado  $g$  grupos (con  $1 \leq g \leq n$ ) con  $n_i$  elementos cada uno, la variabilidad total del conjunto puede ser expresada por medio de la siguiente matriz de dimensión  $p \times p$ :

$$[7] \quad T = n^{-1} \sum_{i=1}^g \sum_{j=1}^p (x_{i,j} - \bar{x})(x_{i,j} - \bar{x})'$$

donde  $\bar{x}$  es el vector de medias. En consecuencia, la matriz  $T$  tiene una interpretación similar a una matriz de varianzas y covarianzas. Dicha matriz se puede descomponer de la siguiente forma:

$$[8] \quad T = B + W$$

siendo:

$$[9] \quad B = \sum_{i=1}^g n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$$

y

$$[10] \quad W = (n - g)^{-1} \sum_{i=1}^g \sum_{h=1}^{n_i} (x_{i,h} - \bar{x}_i)(x_{i,h} - \bar{x}_i)'$$

donde  $\bar{x}_i$  es el vector de medias correspondiente al grupo  $i$ . Naturalmente, tanto  $B$  como  $W$  son matrices de dimension  $p \times p$  y representan medidas de variabilidad entre grupos (*between*) y dentro de los grupos (*within*), respectivamente.

Existen diversos métodos de formación de conglomerados basados en criterios de optimización. Uno de ellos trata de buscar grupos con la máxima heterogeneidad entre ellos y la mínima dentro de cada uno de ellos. Formalmente, se trata de maximizar la traza de  $W$ .

Naturalmente, el mayor problema de estos métodos radica en la explosión combinatoria a que da lugar una búsqueda exhaustiva. Con el fin de dotar al procedimiento de optimización de un contenido operativo en situaciones reales, se han propuesto diversos algoritmos de formación de un número dado de conglomerados. En este trabajo se utiliza el de las  $k$ -medias por su relativa eficiencia y por la solidez de su planteamiento teórico, véase Faber (1994) para una exposición detallada. La secuencia básica es:

- Comienzo: calcular  $g$  centroides  $z_i$   $i=1..g$  elegidos, por ejemplo, mediante una selección aleatoria. El número  $g$  de grupos se supone dado.
- Asignación inicial: se considera que el objeto  $h$  pertenece al grupo  $i$  si la distancia que lo separa del correspondiente centroide es la menor posible:

$$h \in i \Leftrightarrow i = \arg \min D(x_h, z_i)$$

- Cálculo de nuevos centroides como promedio de cada uno de los grupos formados:

$$z_i = n_i^{-1} \sum_{j=1}^p \sum_{h=1}^{n_i} X_{h,j}$$

- Asignación: de la misma forma que antes:

$$h \in i \Leftrightarrow i = \arg \min D(x_h, z_i)$$

Si el objeto  $h$  cambia de grupo, se vuelven a calcular los centroides de los dos grupos afectados por el cambio. En caso contrario, no se recalculan los centroides en este paso.

- El proceso de cálculo de centroides y asignación continúa hasta que se satisface algún criterio de convergencia como, por ejemplo, que la variación en valor absoluto de todos los centroides sea inferior a un umbral predeterminado.

El algoritmo de las  $k$ -medias requiere que el número de conglomerados a formar sea conocido. En algunas aplicaciones, como la presente, esta información preliminar no está disponible y ha de ser obtenida mediante alguna investigación previa. Dado el moderado número de objetos (50), en este trabajo se ha realizado una aglomeración jerárquica mediante el método de Ward (1963) con el fin de disponer de una estimación tentativa del número de grupos. Se ha seleccionado este método porque incorpora explícitamente una función objetivo compatible con los criterios de optimización antes expuestos. El esquema básico del método de Ward es:

- Comienzo: formar  $n$  grupos  $C_1 .. C_n$  conteniendo un solo objeto.
- Paso 1:
  - Buscar el par de grupos más próximos entre sí:  $C_i$  y  $C_j$ .
  - Fusionar  $C_i$  y  $C_j$ .
  - Eliminar  $C_j$ .
- Si el número de conglomerados es uno, terminar. En caso contrario, volver al paso 1.

El criterio de proximidad empleado en el paso 1 del algoritmo de Ward consiste en formar un nuevo grupo en cada etapa de forma que el incremento de varianza intragrupal sea el más pequeño posible. Debe señalarse que la aglomeración final no se ha efectuado con este método por las siguientes razones:

- a- El algoritmo de las k-medias es computacionalmente más eficiente que el de Ward y, en general, que los de tipo jerárquico.
- b- La agrupación formada por métodos jerárquicos puede ser excesivamente rígida en determinadas ocasiones, al no poder reconsiderar las decisiones de fusión tomadas en alguna etapa anterior.
- c- La representatividad de los centroides está garantizada en el caso del algoritmo k-medias pero no lo está, en general, en los demás casos.

En resumen, el proceso de formación de conglomerados consta de dos etapas: (1) determinación preliminar del número de grupos  $g$  mediante el examen del dendrograma generado por el método (jerárquico) de Ward y (2) aplicación del algoritmo de las k-medias tomando  $g$  como número apropiado de conglomerados.



#### 4. DATOS

Los datos utilizados en este trabajo son las pernoctaciones en establecimientos hoteleros procedentes de la *Encuesta sobre Movimientos de Viajeros en Establecimientos Hoteleros* que bajo ese nombre publicó hasta 1999 el Instituto Nacional de Estadística.

La *Encuesta sobre Movimientos de Viajeros en Establecimientos Hoteleros* estaba dirigida a establecimientos hoteleros y campings, sobre la base del marco que proporcionaba el directorio de la Secretaría General de Turismo. La muestra estaba diseñada mediante un muestreo estratificado, siendo los estratos la provincia y la categoría hotelera. La investigación era exhaustiva excepto en aquellos estratos con un número grande de establecimientos para los cuales se seleccionaba una muestra.

La recogida de la información era mensual, contestando cada establecimiento durante un período de siete días seguidos elegidos de manera que entre todos los establecimientos de cada estrato se cubrieran todos los días del mes.

Dentro del marco conceptual de la *Encuesta*, se entiende por pernoctación la ocupación por una persona de una plaza o una cama supletoria dentro de una jornada hotelera y en un mismo establecimiento. La ocupación por una persona en el mismo día de dos o más plazas en establecimientos distintos da lugar a más de una pernoctación.

Se dispone, además, de los datos de pernoctaciones clasificados por nacionalidad del viajero, si bien a partir de 1990 se produce una ruptura al establecerse como criterio de procedencia el país de residencia y no el de nacionalidad.

El período muestral utilizado es 1984-1998. La elección del período está condicionada por la sustitución, a partir de 1999, de la *Encuesta sobre Movimientos de Viajeros en Establecimientos Hoteleros* por la *Encuesta de Ocupación Hotelera*, que, aunque responde a los mismos objetivos y respeta en gran parte los conceptos de la anterior, amplía el ámbito de la investigación y cambia el marco poblacional, perdiéndose así la homogeneidad temporal de la serie. Esta circunstancia resulta además especialmente grave si se tiene en cuenta que en alguna provincia este cambio metodológico coincide con la celebración de eventos cuya repercusión sobre los datos que se tratan de analizar puede confundirse con los efectos inducidos por la modificación de la encuesta (es el caso de la provincia de A Coruña y el Xacobeo).

Una ruptura similar, aunque de signo contrario, se observa en el año 1993 en el que se dejó de encuestar a los establecimientos hoteleros de menor categoría. Esta alteración provocó un cambio de nivel en las series que se salva estimando su valor mediante análisis de intervención y eliminando el escalón correspondiente de las series de ciclo-tendencia.

Por lo que se refiere a las características de las series elegidas que más interesan en este trabajo, destacan la periodicidad mensual y su elaboración a nivel provincial. Ambas características conjuntamente permiten estimar la pauta estacional de cada una de las provincias y su posterior clasificación de acuerdo a la misma. Además, se puede suponer que estas series deben ser un buen indicador del ciclo de la actividad turística, a pesar de que sólo contemplen una parte de dicha actividad.

## **5. RESULTADOS EMPIRICOS**

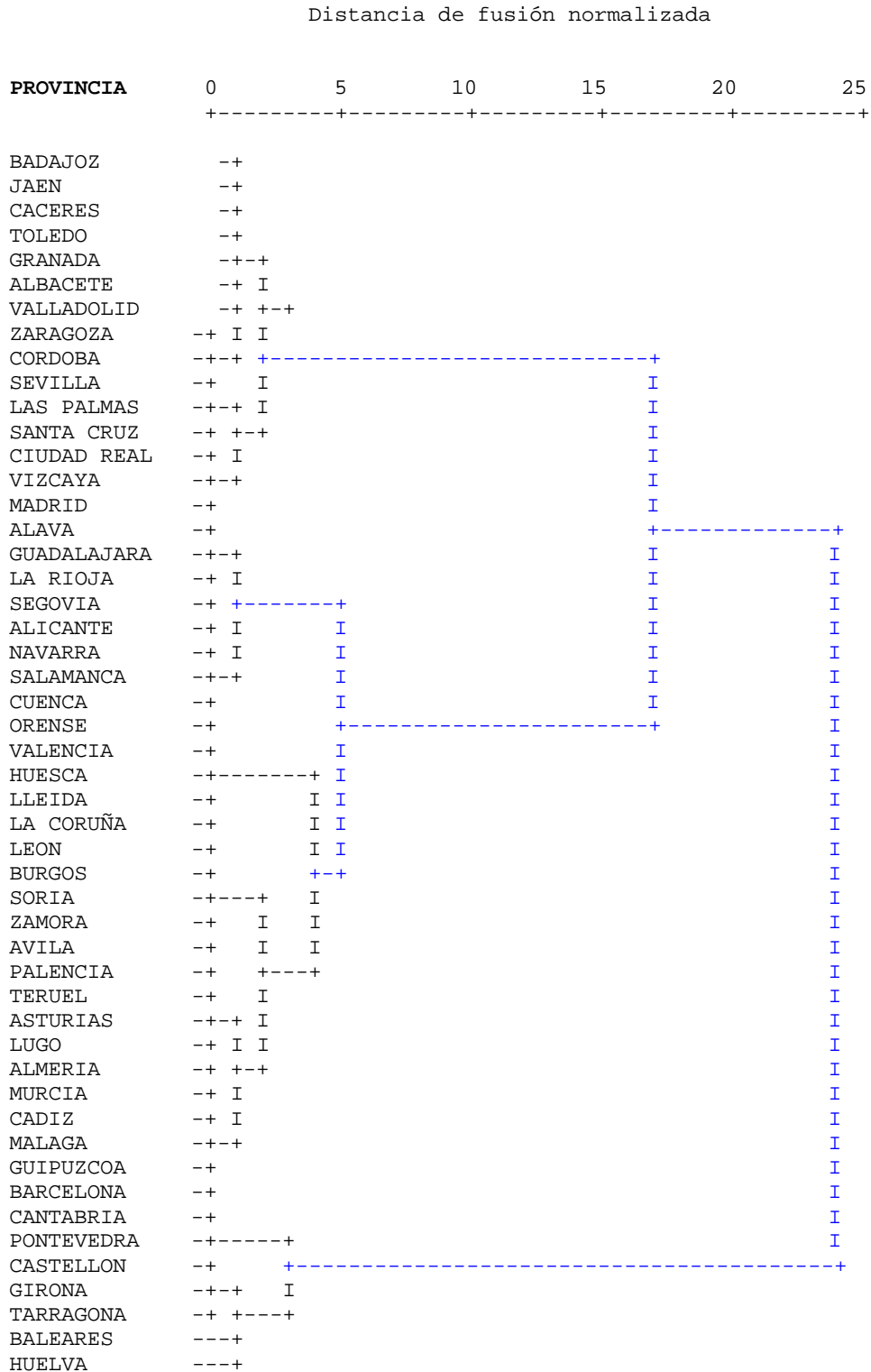
En esta sección se ofrecen los resultados obtenidos de la aplicación de los métodos expuestos en las secciones anteriores a los datos disponibles para este estudio. Se expone, en primer lugar, la aglomeración de los factores estacionales obtenida para, a continuación, examinar la conformidad entre las fluctuaciones cíclicas de un representante de cada uno de los grupos formados.

La selección y estimación de los modelos ARIMA con AI se ha realizado con el programa TRAMO y la estimación de la señal de ciclo-tendencia con el SEATS (Gómez y Maravall, 1996, 1998a). La descomposición ulterior de esta señal se ha efectuado con el programa TRACE (Gómez, 1998b) y la formación de conglomerados con el SAS (SAS, 1991).

### **5.1. Aglomeración estacional**

La aplicación del método de Ward a los factores estacionales del año central (1991) de las series de pernoctaciones de las cincuenta provincias españolas genera el siguiente dendrograma:

Figura 1: Dendrograma de los factores estacionales  
Método aglomerativo jerárquico de Ward



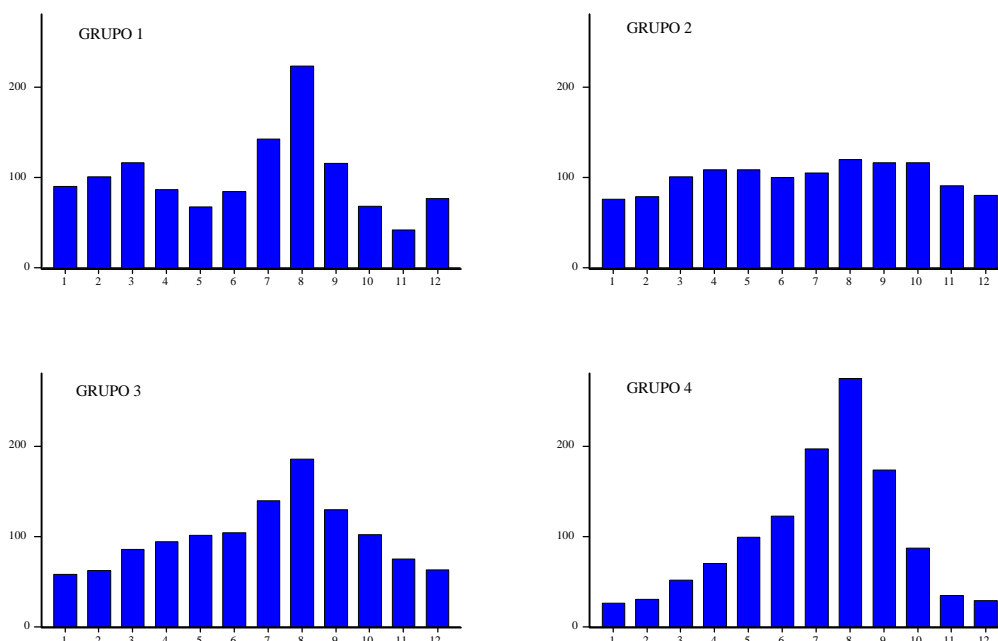
Se aprecia cómo se forman cuatro grupos con una distancia de fusión similar por lo que, con la debida cautela asociada a esta clase de decisiones, se considera  $g=4$  como un valor tentativo para el número de grupos. Considerando  $g=4$ , el algoritmo de las  $k$ -medias descrito en la sección anterior da lugar a la siguiente estructura de conglomerados:

Tabla 1: Composición de los conglomerados

Grupo	Provincia	Distancia	Media	Desviación tip.
1	Huesca	27,03	27,36	0,34
	Lleida	27,70		
2	Alava	30,48	31,32	12,86
	Albacete	24,60		
	Badajoz	12,12		
	Cáceres	25,04		
	Ciudad Real	17,52		
	Córdoba	46,53		
	Granada	35,63		
	Guadalajara	34,74		
	Jaén	16,76		
	Las Palmas	53,63		
	Madrid	42,89		
	Santa Cruz	48,42		
	Segovia	40,01		
	Sevilla	50,68		
	Toledo	18,06		
	Valladolid	18,80		
	Vizcaya	28,52		
Zaragoza	19,30			
3	Alicante	26,37	31,98	12,28
	Almería	43,32		
	Asturias	45,98		
	Avila	19,72		
	Barcelona	63,24		
	Burgos	14,24		
	Cádiz	31,02		
	Cuenca	34,99		
	Guipuzcoa	30,10		
	La Coruña	17,21		
	La Rioja	40,12		
	León	14,44		
	Lugo	47,22		
	Málaga	24,84		
	Murcia	34,65		
	Navarra	27,09		
	Orense	39,87		
	Palencia	13,28		
	Salamanca	38,45		
	Soria	28,24		
Teruel	31,21			
Valencia	46,51			
Zamora	23,42			
4	Baleares	82,65	58,59	15,32
	Cantabria	46,57		
	Castellón	65,97		
	Girona	33,06		
	Huelva	52,60		
	Pontevedra	57,42		
	Tarragona	71,88		

Los factores estacionales de los cuatro centroides aparecen representados en la siguiente figura:

Figura 2: Factores estacionales de los centroides



El primer grupo es el más pequeño y homogéneo. Está formado por las dos provincias pirenaicas cuya estacionalidad es bimodal (abril y agosto), resultado de la combinación de dos temporadas turísticas diferenciadas.

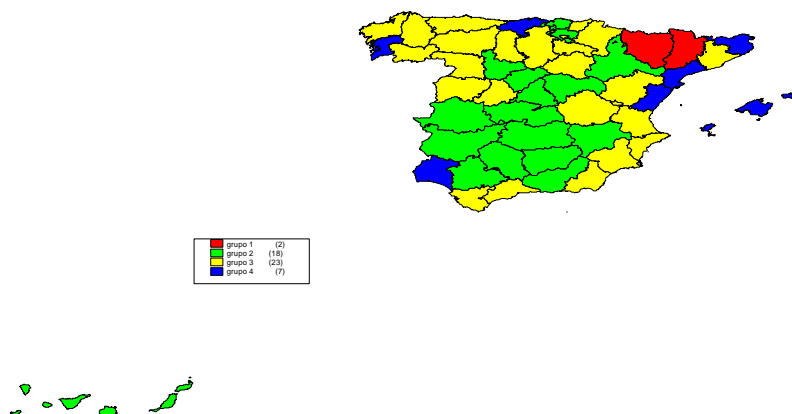
Los principales rasgos que caracterizan al segundo grupo (formado por 18 provincias) son su reducido rango de variación estacional y su localización geográfica predominantemente interior. Este grupo es bastante compacto territorialmente, destacando como excepción la presencia en el mismo de las dos provincias canarias.

El tercer grupo es el más numeroso (23 provincias) y es más homogéneo que el anterior, tanto en lo que se refiere a distancia media al centroide como en cuanto a la variabilidad de dichas distancias. Este grupo posee un claro patrón estacional unimodal concentrado en el trimestre estival y, sobre todo, en el mes de agosto. Territorialmente, abarca la mitad de las provincias costeras de España pero no se identifica exclusivamente con esa característica, ya que la mitad del grupo está formado por provincias interiores, fundamentalmente del cuadrante noroccidental de la península.

Finalmente, el cuarto grupo está compuesto por 7 provincias costeras cuyo patrón estacional es similar al del grupo anterior pero con todos sus rasgos cualitativos sensiblemente acentuados, de forma que la concentración en los meses de verano es muy acusada. Por otra parte, este grupo es bastante heterogéneo, tanto desde la perspectiva de la localización geográfica (sus miembros se distribuyen por todas las costas españolas) como atendiendo a las medidas promediadas de distancias al centroide correspondiente.

En el siguiente mapa se muestra la distribución territorial de los grupos formados.

Figura 3: Aglomeración estacional



Como se aprecia en la tabla 2, el tercer grupo es el más representativo del conjunto, al poseer una configuración estacional intermedia entre las de los grupos dos (extremadamente limitada) y cuatro (extremadamente acusada).

Tabla 2: Distancia final entre grupos

Grupo	1	2	3	4
1	-	142,43	95,00	166,14
2	142,43	-	87,13	227,68
3	95,00	87,13	-	142,56
4	166,14	227,68	142,56	-

## 5.2. Conformidad cíclica

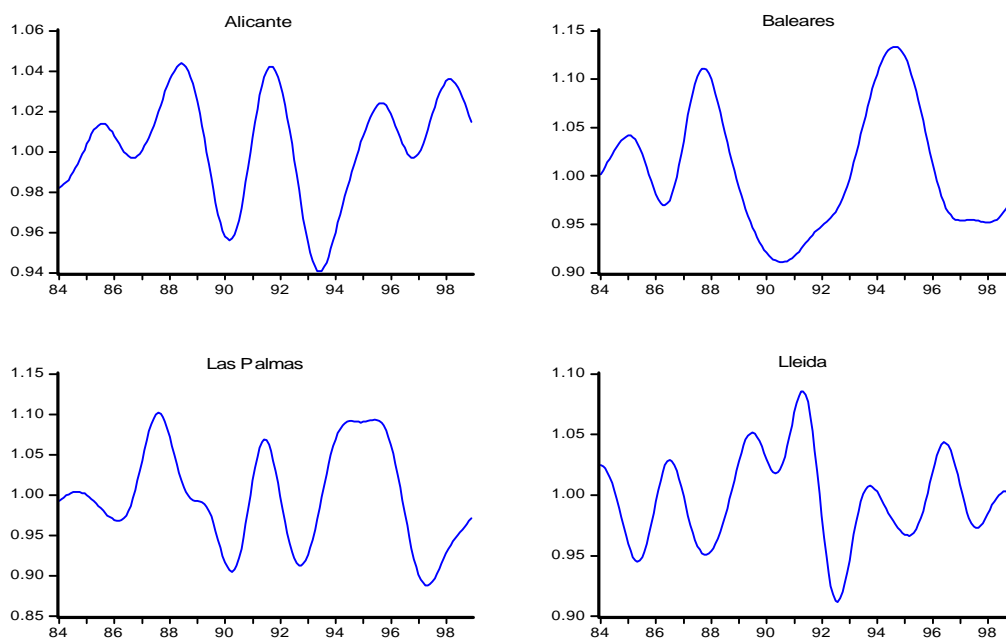
Una vez realizada la agrupación de las provincias de acuerdo a su pauta estacional, se está en disposición de responder a la pregunta planteada inicialmente acerca de la conformidad cíclica de series que evidencian un comportamiento estacional diferente.

La idea en este apartado consiste en seleccionar un representante de cada uno de los cuatro grupos formados anteriormente, estimar sus respectivos componentes cíclicos y evaluar la conformidad de los mismos.

A partir de los cuatro grupos que se han delimitado en el epígrafe anterior se han elegido cuatro representantes: Lleida, Las Palmas, Alicante y Baleares, respectivamente, de forma que cada una de ellas es la que mayor peso tiene en el total dentro de cada grupo. A las series de pernoctaciones de estas cuatro

provincias se les ha extraído el componente puramente cíclico. El resultado aparece en el siguiente conjunto de gráficos:

Figura 4: Perfiles cíclicos



En un primer análisis visual del perfil cíclico de las series ya se obtiene como primera conclusión la disparidad del representante del primer grupo, Lleida, con respecto a las otras tres provincias. Este grupo es el primero que se forma en la jerarquización de las provincias de acuerdo a sus factores estacionales. Es, además, el más localizado geográficamente y el que parece identificarse más claramente con un tipo determinado de turismo como sería el invernal.

El hecho de que este grupo muestre una pauta estacional peculiar, con valores elevados en los primeros meses del año, podría ser un primer argumento en favor de una identificación entre pautas estacionales y perfiles cíclicos. Sin embargo, como se verá más adelante, el segundo grupo (Las Palmas), que también presenta un comportamiento estacional muy característico (en este caso por la virtual ausencia de picos estacionales), ofrece una notable conformidad cíclica con el resto de representantes.

Una forma de establecer la medida en la cual las cuatro provincias elegidas siguen un perfil cíclico que pudiera considerarse común, es intentar resumir las características de las series en un solo factor. Si éste resultara suficientemente explicativo de la varianza de estas cuatro series, se podría pensar que todas ellas obedecen a unas causas explicativas comunes y, en consecuencia, formarían parte de un mismo fenómeno. Si, por el contrario, el factor no fuera capaz de explicar buena parte de la variabilidad total, habría que concluir que hay en las cuatro series perfiles cíclicos distintos.

El resultado de este análisis factorial es claro en este sentido: un primer factor sólo explicaría el 47,7 por ciento de varianza total. Este factor estaría, además, compuesto básicamente por las provincias de Las Palmas y Baleares. Alicante y Lleida entrarían con un peso muy inferior y esta última, además, con signo



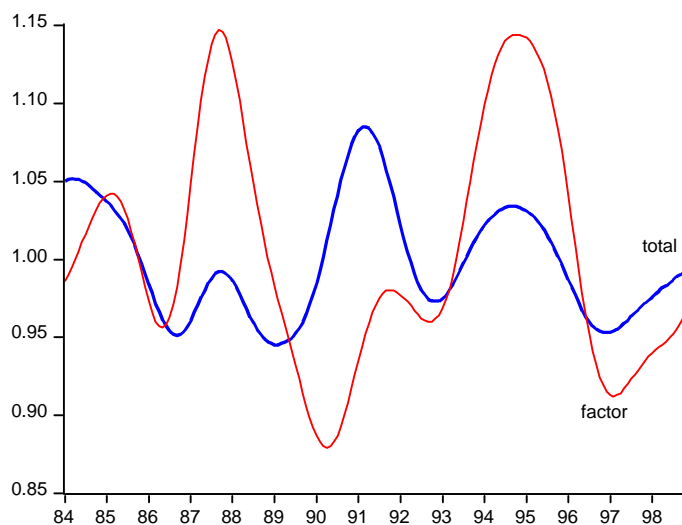
negativo, lo que confirma la primera impresión comentada anteriormente. La siguiente tabla expone los resultados de este análisis.

Tabla 3: Análisis factorial

Provincia	Carga	Comunalidad
Alicante	0,22	0,05
Baleares	0,95	0,90
Las Palmas	0,83	0,69
Lleida	-0,52	0,27
Autovalor	1,91	
Varianza explicada	47,72	

En el gráfico siguiente se representa el factor estimado y se compara con el perfil cíclico de las pernoctaciones totales:

Figura 5: Perfiles cíclicos

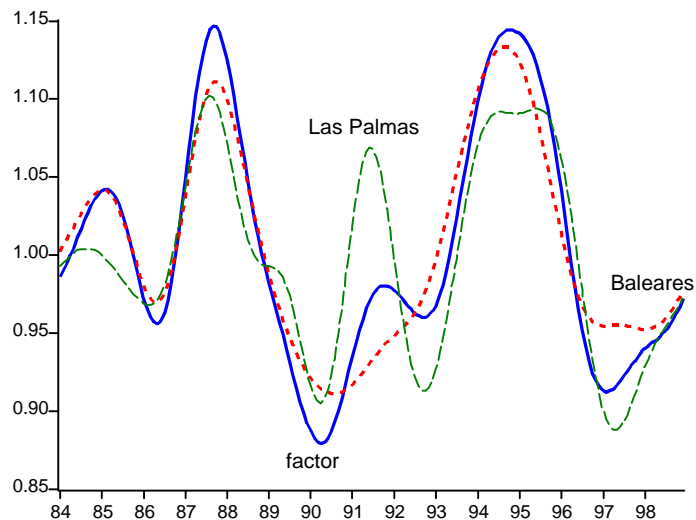


Conviene destacar la elevada coherencia entre ambas series, con la excepción del período 89-92. Es interesante esta relación por cuanto el factor prácticamente sólo utiliza la información de dos provincias (Las Palmas y Baleares), aunque éstas representen en media del período algo más del 34 por ciento de las pernoctaciones totales.

La mayor variabilidad que se observa en el factor también está causada por el hecho de que sean estas dos provincias sus componentes. En ellas el porcentaje de viajeros extranjeros es mayor (el 86 por ciento frente al 63 del total) y están sometidas, por lo tanto, a fluctuaciones ligadas a los factores explicativos de la demanda internacional de turismo (especialmente al tipo de cambio). Un ejemplo de estos efectos sobre el factor se puede encontrar en la recuperación de 1993 tras las sucesivas devaluaciones de la peseta.

Si se compara el factor con los ciclos de Las Palmas y Baleares se observa con más claridad lo comentado en párrafos anteriores:

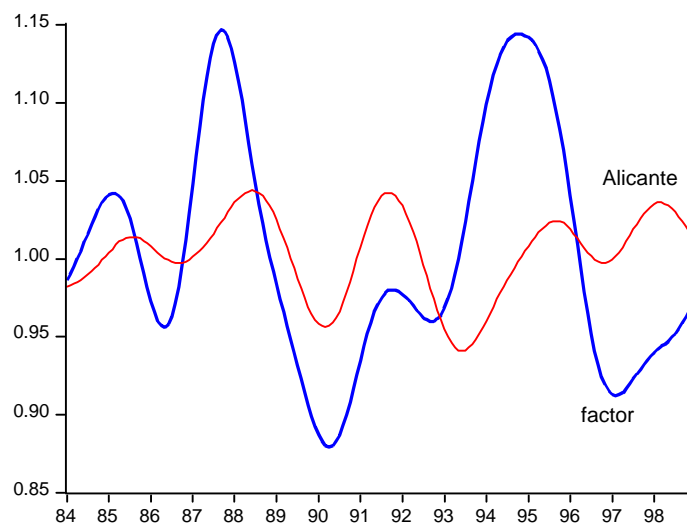
Figura 6: Perfiles cíclicos



Es interesante en este punto hacer notar que estas provincias que muestran una elevada coherencia cíclica tienen comportamientos estacionales muy distintos. Las Palmas representa al grupo caracterizado por la ausencia de picos estacionales ya que sus pernoctaciones se reparten de una manera bastante homogénea a lo largo del año. Por su parte, Baleares destaca precisamente por tener una pauta estacional muy acusada y centrada en los meses de verano.

Más parecido a dicho patrón es el de la provincia de Alicante que, sin embargo, no forma parte nada más que minoritariamente del factor y que parece manifestarse con cierto retraso respecto al mismo. Este hecho lo confirman tanto la función de correlación cruzada como el análisis realizado con los programas <F> y <G> (Abad y Quilis, 1997) que detectan un retraso de aproximadamente 7 meses.

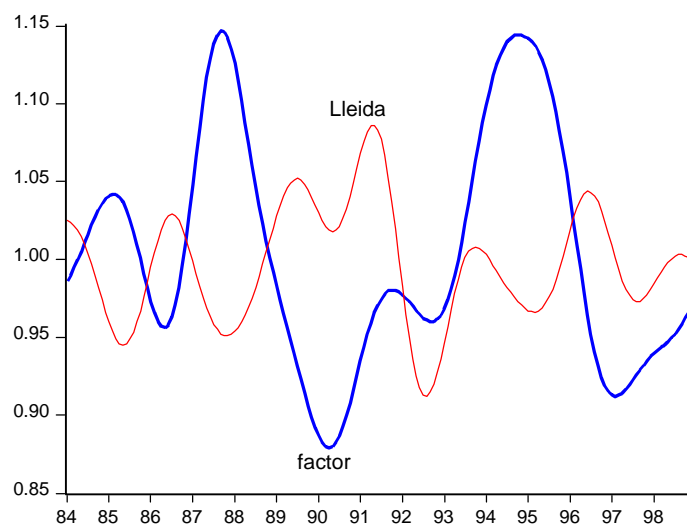
Figura 7: Perfiles cíclicos



Por otra parte, hay que resaltar que el perfil cíclico de Alicante parece guardar un mayor parecido con el de las pernoctaciones totales, aunque siempre produciéndose con retraso. De nuevo puede estar detrás el hecho de que la proporción de pernoctaciones de viajeros extranjeros y españoles es muy distinta entre las provincias elegidas. Así, Alicante tiene unas pernoctaciones divididas casi al 50 por ciento entre unos y otros, lo que hace que los factores que influyen en su evolución difieran o tengan distinto peso que los que explican el ciclo en Las Palmas o Baleares.

Finalmente, la falta de conformidad cíclica de las pernoctaciones causadas en la provincia de Lleida con el factor común se aprecia claramente en el gráfico siguiente.

Figura 8: Perfiles cíclicos



## 6. CONCLUSIONES

La propuesta de otorgar un papel más destacado para la estacionalidad en el análisis económico, en general, y del ciclo, en particular, ha cosechado un éxito parcial, a tenor de los resultados obtenidos en este trabajo.

Por una parte, se ha podido apreciar cómo la información contenida en el componente estacional de las cincuenta series consideradas permite establecer una segmentación del mercado bien fundamentada, de forma que se dispone de una base objetiva sobre la que asentar estudios de carácter cíclico y coyuntural.

Por otra, la evidencia empírica no avala la noción de que series con comportamientos estacionales diferentes posean un patrón cíclico igualmente diferente. En términos espectrales, la información contenida en los picos estacionales no actúa como discriminante de la contenida en la banda cíclica. En este sentido, la práctica habitual de analistas del ciclo y coyunturistas, consistente en ignorar la pauta estacional de las series que analizan, aparece como esencialmente correcta.

Este estudio puede ser ampliado en dos direcciones. La primera consiste en considerar de forma desagregada las pernoctaciones efectuadas por residentes y no residentes. De esta manera se podría comprobar en qué medida sus pautas de comportamiento son similares, tanto en lo que se refiere a la estacionalidad (un rasgo permanente) como al ciclo (un elemento transitorio). La segunda línea de desarrollo aplicaría la metodología aquí expuesta a otros fenómenos de interés en el comportamiento de estas series: ciclo semanal, fiestas móviles e incidencia de observaciones anómalas.

## REFERENCIAS

- Abad, A. y Quilis, E.M. (1997) "Programas de análisis cíclico: <F> , <G> y <FDESC>. Manual del usuario", Instituto Nacional de Estadística, Documento Interno.
- Abraham, B. y Box, G.E.P. (1979) "Bayesian analysis of some outlier problems in time series", *Biometrika*, vol. 66, n. 2, p. 229-236.
- Barsky, R.B. y Miron, J.A. (1989) "The seasonal cycle and the business cycle", *Journal of Political Economy*, vol. 97, n. 3, p. 503-534.
- Bell, W.R. y Hillmer, S.C. (1983) "Modeling time series with calendar variation", *Journal of the American Statistical Association*, vol. 78, n. 383, p. 526-534.
- Box, G.E.P. y Jenkins, G.M. (1976) *Time series analysis, forecasting and control*, Holden Day, San Francisco, U.S.A.
- Burman, J.P. (1980) "Seasonal adjustment by signal extraction", *Journal of the Royal Statistical Society, series A*, n. 143, p. 321-337.
- Chang, I., Tiao, G.C. y Chen, C. (1988) "Estimation of time series parameters in the presence of outliers", *Technometrics*, vol. 30, n. 2, p. 193-204.
- Chen, C. y Liu, L.M. (1991) "Joint estimation of model parameters and outlier effects in time series", SCA Corporation, Working Paper n. 126.
- Chen, C., Liu, L.M. y Hudak, G.B. (1990) "Outlier detection and adjustment in time series modeling and forecasting", SCA Corporation., Documento Interno.
- Everitt, B.S. (1993) *Cluster analysis*, Arnold, London, U.K.
- Faber, V. (1994) "Clustering and the continuous k-means algorithm", *Los Alamos Science*, n. 22, p. 138-144.
- Gómez, V. (1998a) "Butterworth filters: a new perspective", Ministerio de Economía y Hacienda, Documento de Trabajo n. D-98008.
- Gómez, V. (1998b) "Program TRACE. Instructions for the user", Ministerio de Economía y Hacienda, Documento Interno.
- Gómez, V. y Maravall, A. (1996) "Programs TRAMO and SEATS", Documento de Trabajo n. 9628, Banco de España.
- Gómez, V. y Maravall, A. (1998a) "Guide for using the programs TRAMO and SEATS", Documento de Trabajo n. 9805, Banco de España.
- Gómez, V. y Maravall, A. (1998b) "Automatic modeling methods for univariate series", Documento de Trabajo n. 9808, Banco de España.
- Hillmer, S.C. (1982) "Forecasting time series with trading day variation", *Journal of Forecasting*, vol. 1, p. 385-395.
- Hillmer, S.C. y Tiao, G.C. (1982) "An ARIMA model-based approach to seasonal adjustment", *Journal of the American Statistical Society*, vol. 77, n. 377, p. 63-70.
- Hillmer, S.C., Bell, W. y Tiao, G.C. (1983) "Modeling considerations in the seasonal adjustment of economic time series", en Zellner, A. (Ed.) *Applied time series analysis of economic data*, U.S. Department of Commerce, Bureau of the Census, Washington, U.S.A.
- Liu, L.M. (1980) "Analysis of time series with calendar effects", *Management Science*, vol. 26, n. 1, p. 106-112.
- Liu, L.M. (1983) "Identification of time series models in the presence of calendar variation", SCA Corporation, Working Paper n. 102.
- Maravall, A. (1987) "Descomposición de series temporales. Especificación, estimación e inferencia", *Estadística Española*, vol. 29, n. 114, p. 11-69.
- Maravall, A. (1993a) "Stochastic linear trends. Models and estimators", *Journal of Econometrics*, n. 56, p. 5-37.
- Maravall, A. (1993b) "Short-term analysis of macroeconomic time series", European University Institute, Working Paper ECO n. 95/9.
- Maravall, A. (1994) "Unobserved components in economic time series", en Pesaran, H., Schmidt, P. y Wickens, M. (Eds.) *The handbook of applied econometrics*, vol. 1, Basil Blackwell, Oxford, U.K.
- Prescott, E.C. (1986) "Theory ahead of business cycle measurement", Federal Reserve Bank of Minneapolis, *Quarterly Review*, vol. 10, n. 4, p. 9-22.
- SAS (1991) *SAS/STAT. User's guide*, SAS Institute, Cary, North Carolina, U.S.A.
- Todd, R.M. (1990) "Periodic linear-quadratic methods for modeling seasonality", *Journal of Economic Dynamics and Control*, vol. 14, n. 3-4, p. 763-795.
- Tsay, R.S. (1986) "Time series model specification in the presence of outliers", *Journal of the American Statistical Association*, vol. 81, n. 393, p. 132-141.

- Tsay, R.S. (1988) "Outliers, level shifts, and variance changes in time series", *Journal of Forecasting*, vol. 7, p. 1-20.
- Ward, J.H. (1963) "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, n. 53, p. 236-244.