

JUAN ANTONIO MONDÉJAR JIMÉNEZ
Área de Comercialización e Investigación de Mercados
Facultad de Ciencias Sociales
Universidad de Castilla-La Mancha
Avda. de los Alfares, 44 — 16.071-Cuenca (España)
Teléfono +34 902 204 100 Ext. 4242 — Fax +34 902 204 130
JuanAntonio.Mondejar@uclm.es

JOSÉ MONDÉJAR JIMÉNEZ
Área de Estadística
Facultad de Ciencias Sociales
Universidad de Castilla-La Mancha
Avda. de los Alfares, 44 — 16.071-Cuenca (España)
Teléfono +34 902 204 100 Ext. 4246 — Fax +34 902 204 130
Jose.Mondejar@uclm.es

CARLOTA LORENZO ROMERO
Área de Comercialización e Investigación de Mercados
Facultad de Ciencias Económicas y Empresariales
Universidad de Castilla-La Mancha
Plaza de la Universidad, 1 — 02.071-Albacete (España)
Teléfono +34 902 204 100 Ext. 2310 — Fax +34 902 204 130
Carlota.Lorenzo@uclm.es

ALTERNATIVES OF PREDICTION IN TOURISM SERIES

ABSTRACT

Last years, tourist sector has been seen strongly affected by exogenous factors fruit of the economic globalization. In this scene the handling of the time series available is from special difficulty. The extraction of excellent signal in the series, as well as the accomplishment of predictions, receives a special difficulty. In this work it is tried to make a comparison between the traditional and bayesian methods of extraction of signal, with which we very obtain smoother profiles of special importance in times money changers. For one better understanding such finalizes the work with the application to the tourist series of “pernoctaciones” corresponding to all provinces of the Spanish community of Castilla-La Mancha.

KEYWORDS: Tourism, time series, signal extraction.

RESUMEN

En los últimos años, el sector turístico se ha visto fuertemente afectado por factores exógenos fruto de la globalización económica. En este escenario cada vez más complejo resulta de especial dificultad el manejo de las series temporales disponibles. La extracción de señal relevante en las series, así como la realización de predicciones, cobran una especial dificultad. En este trabajo se pretende realizar una comparación entre los métodos tradicionales y bayesianos de extracción de señal, con la que obtenemos perfiles más suaves de especial importancia en tiempos muy cambiantes. Para una mejor comprensión de los mismos se finaliza el trabajo con la aplicación a las series turísticas de pernoctaciones correspondientes a todas las provincias de la comunidad autónoma española de Castilla-La Mancha.

PALABRAS CLAVE: Turismo, series temporales, extracción de señal

1.- INTRODUCCIÓN.

Los investigadores han mostrado en los últimos años su reiterado interés por investigar el sector turístico. Concretamente, la relación entre la investigación sobre turismo en economía de la empresa, especialmente desde el punto de vista de Marketing, ha ido creciendo como muestra del creciente interés de los investigadores y permitiendo, a su vez, mejorar su gestión (Sánchez y Marín, 2003). La principal razón de esta proliferación de estudios está directamente relacionada con la introducción de los estudios de turismo en la Universidad española (Esteban, 2000).

La propia naturaleza del turismo como conjunto de actividades relativas a alojamiento, transporte, animación, promoción, comercialización, entre otras, así como la amplia gama de dimensiones que confluyen en la configuración del producto argumentan sobradamente su interpretación como hecho multiforme. De esta forma, su estudio se aborda desde distintos ámbitos disciplinares que tratan de analizar el turismo con sus correspondientes metodologías y herramientas y desde sus propios fundamentos epistemológicos (Vera e Ivars, 2001).

A partir de este creciente interés por analizar el sector, surge la necesidad de establecer herramientas que permitan canalizar la gran cantidad de información que, a nivel institucional, se deriva de las estadísticas oficiales.

Así, el Instituto Nacional de Estadística (INE) elabora numerosas estadísticas oficiales relacionadas con el sector, a saber: encuesta de ocupación hotelera, encuesta de ocupación en acampamentos turísticos, encuesta de ocupación en apartamentos turísticos, encuesta de ocupación en alojamientos de turismo rural, índice de precios hoteleros e índice de ingresos hoteleros, índice de precios de acampamentos y de apartamentos turísticos, encuesta sobre la estructura de las empresas hoteleras, encuesta sobre la estructura de empresas de agencias de viajes, etc... Además, el instituto de Estudios Turísticos analiza turismo receptor y turismo nacional en base a sus encuestas FRONTUR y FAMILITUR junto a otras muchas estadísticas oficiales, de gran interés para los investigadores de la materia.

2.- ESTADÍSTICAS TURÍSTICAS: LA ENCUESTA DE OCUPACIÓN HOTELERA

Como se ha mencionado, entre las numerosas estadísticas oficiales disponibles en España a nivel turístico, la Encuesta de Ocupación Hotelera (EOH), elaborada mensualmente por el INE con un nivel de desagregación provincial, que representa una radiografía de los

turistas en España, sobre la base de unidades de análisis formada por todos los establecimientos hoteleros inscritos como tales en el correspondiente registro de las Consejerías de Turismo de cada Comunidad Autónoma. Es decir, aquellos establecimientos que prestan servicios de alojamiento colectivo mediante precio con o sin otros servicios complementarios (hotel, hotel-apartamento o apartahotel, motel, hostel, pensión,...).

En los datos facilitados se refleja la doble vertiente que se considera en el estudio del turismo: del lado de la demanda, se ofrece información sobre viajeros, pernoctaciones y estancia media distribuidos por país de residencia de los viajeros y categoría de los establecimientos que ocupan o por Comunidad Autónoma de procedencia en el caso de los viajeros españoles; del lado de la oferta se proporciona el número de establecimientos abiertos estimados, plazas estimadas, grados de ocupación e información sobre el empleo en el sector, según la categoría del establecimiento (INE, 2006).

La definición de variables de la EOH implica las siguientes consideraciones:

- Establecimientos hoteleros abiertos estimados: se entiende por establecimiento hotelero abierto de temporada, aquel en el que el mes de referencia está comprendido dentro de su periodo de apertura.
- Plazas estimadas: el número de plazas estimadas por la encuesta de los establecimientos abiertos de temporada. Este número de plazas equivale al número de camas fijas del establecimiento.
- Viajeros entrados: todas aquellas personas que realizan una o más pernoctaciones seguidas en el mismo alojamiento. Los viajeros se clasifican por su país de residencia, y para los residentes en España, por la comunidad autónoma donde residen habitualmente.
- Pernoctaciones o plazas ocupadas: se entiende por pernoctación cada noche que un viajero se aloja en el establecimiento. Al igual que en la entrada de viajeros, las plazas ocupadas se desglosan según el lugar de residencia.
- Estancia media: esta variable es una aproximación al número de días que, por término medio, los viajeros permanecen en los establecimientos y se calcula como cociente entre las pernoctaciones y el número de viajeros.

- Grado de ocupación por habitaciones: relación, en porcentaje, entre la media diaria de habitaciones ocupadas en el mes y el total de habitaciones disponibles.
- Grado de ocupación por plazas: relación, en porcentaje, entre el total de las pernoctaciones y el producto de las plazas, incluyendo las camas supletorias, por los días a que se refieren las pernoctaciones.
- Grado de ocupación por plazas en fin de semana: relación, en porcentaje, entre las pernoctaciones del viernes y sábado que caen dentro de la semana de referencia y el producto de las plazas, incluyendo las camas supletorias de esos dos días, por los días a que se refieren las pernoctaciones, en este caso dos.
- Grado de ocupación por habitaciones esperado: se trata de conocer el porcentaje de habitaciones del establecimiento que se considera que van a estar ocupadas durante el segundo y tercer mes siguiente al que incluye la semana de referencia. Por ejemplo, si la encuesta corresponde al mes de enero se pide información de los meses de marzo y abril.
- Personal ocupado: se define como el conjunto de personas, remuneradas y no remuneradas, que contribuyen mediante la aportación de su trabajo, a la producción de bienes y servicios en el establecimiento durante el mes que incluye el período de referencia de la encuesta, aunque trabajen fuera de los locales del mismo.
- Punto turístico: municipio donde la concentración de la oferta turística es significativa.
- Zona turística: conjunto de municipios en los que se localiza de forma específica la afluencia turística. Se ofrece información de las principales zonas de interés turístico.

La información se presenta en diferentes niveles de desagregación geográfica: nacional, de comunidad autónoma, provincial, de zonas y de puntos turísticos. Se han considerado las zonas (conjunto de municipios), así como los municipios en los que se localiza de forma específica la afluencia turística.

2.- EXTRACCIÓN DE SEÑAL EN SERIES TEMPORALES

El concepto de filtrado incluye una serie de procedimientos derivados del hecho de que la señal que transportan los datos está contaminada con componentes que no son deseables y que se pueden considerar ruido en este contexto. Se trata, por tanto, de limpiar esta señal, es decir, de depurar la serie de los componentes irregulares y estacionales.

Lo que se entiende por ruido, sin embargo, no es concebido del mismo modo por todos los autores. Así, por ejemplo, Gómez y Maravall (1998) o Espasa y Cancelo (1993) comienzan el planteamiento depurando la estacionalidad y los componentes irregulares para obtener el componente ciclo tendencia. Sin embargo, otros autores como Baxter y King (1995) no tratan el problema de la estacionalidad y se preocupan de obtener el componente cíclico a través de un filtro, imponiendo restricciones de estacionaridad y eliminando componentes irregulares. En un tercer enfoque, Canova (1998) plantea el problema de una forma diferente: se trata de extraer el componente tendencial de la propia serie ciclo-tendencia, obteniendo como residuo el componente puramente cíclico.

Estos distintos enfoques pueden parecer complementarios y no excluyentes desde el momento en que la extracción de ambos componentes, irregular y estacional, proporciona una serie de ciclo-tendencia sobre la que aplicar una segunda fase de extracción de ciclo y tendencia. Sin embargo, considerar la aplicación de los distintos filtros de forma independiente puede llevar a dos tipos de problemas: el primero es el tratamiento de los componentes estacionales que, sistemáticamente, han recibido un tratamiento aislado y diferente del filtrado y la extracción del resto de los componentes. Algunos de los filtros de desestacionalización, típicamente el procedimiento X-11, suponen aplicar tratamientos previos para depurar la serie original de los componentes estacionales. Esto conduce al segundo problema y es el derivado de las consecuencias de aplicar filtros en cascada.

Podemos, por tanto, caracterizar los diferentes procedimientos de filtrado, describiendo sus efectos desde una perspectiva de análisis espectral, es decir, caracterizándolos como filtros de frecuencias. En la propia noción de filtro subyace la idea de dejar pasar determinadas frecuencias y eliminar otras. Un filtro ideal es aquel que está diseñado para permitir el paso de información en una banda de frecuencias y eliminar o inhibir la información contenida en otras bandas de frecuencia consideradas no deseables. Se

trata de analizar qué tipo de frecuencia deja pasar cada filtro y cuáles de ellas son eliminadas, así como la eficacia relativa de cada filtro respecto a las frecuencias filtradas.

Cualquier filtro tiene un coste informativo que se traduce en las observaciones que se pierden en el tramo final e inicial que hacen necesario sustituir los datos por predicciones provisionales sujetas a revisión. No es nada nuevo en el trabajo de los analistas de coyuntura tener que recalcular las series filtradas cuando llega un nuevo dato que, a veces, llega a alterar de forma dramática los resultados previamente obtenidos. Este desfase temporal de las series filtradas será analizado mediante el análisis de fase. Se trata, pues, de clarificar las consecuencias de la aplicación de diferentes filtros ya que como señala Melis (1991): *“El desarrollo de las técnicas ARIMA ha contribuido indirectamente, al facilitar la predicción a corto, a oscurecer la frontera entre señal y predicción”*.

Pasaremos, a continuación, a exponer brevemente los métodos más importantes de filtrado y sus características, clasificando los filtros en distintos tipos (tan solo se hará un desarrollo más profundo del filtro bayesiano):

- ✓ **Filtros autoregresivos (AR):** El filtro univariante más sencillo es el filtro AR, es decir, el operador diferencia. Dentro de los filtros AR el más simple es el AR(1), es decir, la primera diferencia ordinaria. Este tipo de filtros tiene unas propiedades características: cuando se aplica sobre series que son estacionarias es un filtro de paso alto, es decir, elimina las frecuencias bajas pero acentúa las frecuencias altas. En series suaves permite obtener estimaciones de los componentes estacional e irregular y aproximar la tasa de crecimiento de la serie original cuando se filtra el logaritmo y, por lo tanto, ha sido utilizado como indicador del ciclo cuando las series tienen comportamientos suaves. Sin embargo, la vasta utilización que se ha hecho de este filtro se debe a que convierte en estacionarias series I(1), es decir, series integradas de orden 1, permitiendo la utilización del análisis ARMA para la predicción y estimación. La clave está en analizar correctamente el comportamiento de la serie original. Para cualquier comportamiento de la serie original, el componente cíclico debe ser en todo caso estacionario. Sin embargo, el componente tendencial puede ser estacionario o no. Según cuál sea el comportamiento de la tendencia estaremos ante un ciclo clásico o de crecimiento.

- ✓ **Filtros aditivos o de medias móviles (MA):** Son uno de los filtros más usados para extraer componentes en análisis del ciclo. En términos frecuenciales son filtros pasabanda, es decir, filtros que dejan pasar intacta la información contenida en determinada banda de frecuencias mientras que eliminan o acentúan las restantes. Se trata de eliminar movimientos tendenciales correspondientes a las frecuencias más bajas y los movimientos irregulares o vinculados a las frecuencias más altas. En cualquier proceso MA(m), la función de ganancia cuando los pesos son iguales es la siguiente:

$$G(w) = \frac{\text{sen}(mw / 2)}{m \text{sen}(w / 2)}$$

Esta función presenta ceros en $\frac{2\pi k}{m}$, $k=1,2,\dots$ y máximos locales (lóbulos laterales) en $\frac{\pi(2k + 1)}{m}$, $k= 1,2,\dots$. Precisamente estas bandas típicas de las medias móviles simples suponen el principal inconveniente para su uso como filtros, ya que parte de la varianza de la serie original pasa a la serie filtrada contaminándola de oscilaciones cortas de alta frecuencia. Una de las medias más usadas es la media móvil de periodo estacional. Se trata de uno de los filtros más sencillos para desestacionalizar extrayendo el componente ciclo-tendencia. La función de transferencia tiene ceros en los armónicos estacionales, aunque los lóbulos laterales que induce hacen que esté muy alejado del filtro ideal.

- ✓ **Filtro bayesiano en espacio de estados:** El planteamiento bayesiano del filtrado de indicadores está íntimamente ligado a los modelos dinámicos en espacio de estados. El aprendizaje bayesiano combina, mediante el teorema de Bayes, la información de las observaciones, expresada por la función de verosimilitud, con el estado de conocimiento del analista antes de disponer de las observaciones; se dispone así de un mecanismo de actualización del conocimiento sobre la serie que se puede esquematizar como sigue:

1.- Dado el indicador Y_t y su señal de ciclo-tendencia, θ_t , se considera que su relación está perturbada por un ruido, y que la señal presenta una evolución markoviana como se recoge en la siguiente ecuación:

$$Y_t = F\theta_t + V_t$$

$$\theta_{t-1} = G\theta_t + W_t$$

2.- Denotemos por D_{t-1} el conocimiento existente en el instante (t-1). Podemos expresar la información a priori sobre la señal en el instante t mediante la distribución de θ_t/D_{t-1} .

3.- La predicción a partir de la información a priori se generará a través de la ecuación de observación

$$Y_t/D_{t-1} = (F_t \theta_t/D_{t-1}) + (V_t/D_{t-1}) = F_t(\theta_t/D_{t-1}) + (V_t/D_{t-1})$$

4.- Ahora, considerando $D_t = D_{t-1} \cup Y_t$, podemos obtener la distribución a posteriori, que será proporcional al producto de la distribución a priori por la verosimilitud.

$$(\theta_t/D_t) \propto (Y_t/\{\theta_t, V_t\}) \cdot (\theta_t/D_{t-1})$$

5.- Por último, utilizando la ecuación de transición, podemos obtener la distribución a priori para la señal en el instante (t+1) como

$$(\theta_{t+1}/D_t) = G_t(\theta_t/D_t) + (W_t/D_t)$$

cerrándose así el ciclo «a priori - verosimilitud - a posteriori - a priori», fundamento del aprendizaje bayesiano.

Esta formulación permite una gran flexibilidad, tanto de análisis como de intervención, propiedad deseable dentro de un contexto dinámico como son los indicadores de coyuntura.

Asumiendo la normalidad para la distribución a priori y mediante el algoritmo expuesto, podemos actualizar el conocimiento sobre la serie de ciclo-tendencia, quedándonos las distribuciones a posteriori

$$(\theta_t/D_t) \rightarrow t_{n_t}[m_t, C_t]$$

$$(\varphi_t/D_t) \rightarrow G\left[\frac{n_t}{2}, \frac{d_t}{2}\right]$$

donde las actualizaciones son

$$m_t = a_t + A_t e_t \quad C_t = \left(\frac{S_t}{S_{t-1}}\right) [R_t - A_t A_t' Q_t]$$

$$e_t = Y_t - f_t \quad A_t = \frac{R_t F_t}{Q_t}$$

$$n_t = \delta_t n_{t-1} + 1 \quad d_t = \delta_t d_t + \frac{S_{t-1} e_t^2}{Q_t}$$

completándose así el proceso de aprendizaje de los datos.

Hemos introducido toda la información disponible en la serie dentro de nuestro modelo de una forma coherente y bastante versátil. En particular, la formulación de un modelo lineal dinámico permite la introducción de información extramuestral en cualquier instante del período con sólo alterar la distribución a priori en ese período.

Otra ventaja interesante es la detección y tratamiento de anómalos. Supongamos que llamamos M al modelo propuesto por el algoritmo anterior y M_A a un modelo alternativo. Si definimos ahora

$$H_t = \frac{p(Y_t/D_{t-1}, M)}{p(Y_t/D_{t-1}, M_A)}$$

llamado factor de Bayes, este cociente nos permite cuantificar la evidencia a favor o en contra del modelo propuesto frente al alternativo. Con este mecanismo, Jeffreys (1961) sugiere que existe evidencia a favor del modelo alternativo.

$$H_{t \dots}(k) = \frac{p(Y_t, Y_{t-1}, \dots, Y_{t-k+1}/D_{t-k}, M)}{p(Y_t, Y_{t-1}, \dots, Y_{t-k+1}/D_{t-k}, M_A)}$$

De esta forma, podemos detectar la presencia de outlier cuando el factor de Bayes indique evidencia contraria al modelo propuesto, siendo el modelo alternativo el resultante de la eliminación de la información contenida en el dato presuntamente anómalo. El hecho de tomar k períodos permite la detección de sólo aquellos anómalos que presentan influencia relativamente importante.

En el caso de estar interesados en un horizonte de predicción más amplio, $k \geq 1$, las distribuciones predictivas proporcionadas por el modelo lineal dinámico serían

$$\begin{aligned}(\theta_{t+k}/D_t) &\rightarrow t_{\delta n_t}[a_t(k), R_t(k)] \\(Y_{t+k}/D_t) &\rightarrow t_{\delta n_t}[f_t(k), Q_t(k)]\end{aligned}$$

donde las actualizaciones futuras se hacen en función de

$$\begin{aligned}a_t(k) &= G_{t+k} a_t(k-1) & a_t(0) &= m_t \\R_t(k) &= G_{t+k} R_t(k-1) G'_{t+k} + W_{t+k} & R_t(0) &= C_t \\f_t(k) &= F'_{t+k} a_t(k) & Q_t(k) &= F'_{t+k} R_t(k) F_{t+k} + S_t\end{aligned}$$

Por último, para obtener la distribución del ciclo-tendencia histórica, basta con retrotraer la información desde el último periodo disponible, utilizando las distribuciones:

$$\begin{aligned}(\theta_{t-k}/D_t) &\rightarrow t_{n_t(-k)}[a_t(-k), R_t(-k)] \\(\varphi_{t-k}/D_t) &\rightarrow G \left[\frac{n_t(-k)}{2}, \frac{d_t(-k)}{2} \right]\end{aligned}$$

donde se tiene que

$$\begin{aligned}a_t(-k) &= m_{t-k} - B_{t-k} [a_{t-k+1}(-k) - a_t(-k+1)] \\R_t(-k) &= C_{t-k} - B_{t-k} [R_{t-k+1} - R_t(-k+1)] B'_{t-k} \\B_t &= C_t G'_{t+1} + R_{t+1}^{-1} \\n_t(-k) &= n_{t-k} + \delta_{t-k+1} [n_t(-k+1) - \delta_{t-k+1} n_{t-k}] \\S_t^{-1}(-k) &= S_{t+1}^{-1} + \delta_{t-k+1} (S_t^{-1}(-k+1) - R_{t-k}^{-1}) \\d_t(-k) &= n_t(-k) S_t(-k)\end{aligned}$$

Así, obtenemos una estimación suavizada del ciclo-tendencia dado el conjunto de información disponible¹.

- ✓ **Extracción de la señal relevante con Seats:** Por todo lo comentado antes, nos basaremos en el enfoque basado en modelos ARIMA. Para la extracción de señal utilizaremos el programa SEATS (Gómez y Maravall, 1997). El nombre de este programa se corresponde con las letras iniciales de la traducción al inglés de ‘extracción de señal en series de tiempo ARIMA’ (“**S**ignal **E**xtraction in **A**rima **T**ime **S**eries”). Vamos a describir de manera breve las utilidades que proporciona este

¹ Algunas aplicaciones de esta metodología pueden encontrarse en Mondéjar (2006), Alfaro, Mondéjar y Vargas (2004) y Mondéjar (2003).

programa con el objeto de explicar qué hace el programa. Para ello, aportamos unos extractos de la descripción del programa que Gómez y Maravall (1997) hacen²:

“SEATS es un programa para la identificación de componentes no observados en series temporales siguiendo el enfoque llamado ‘basado en modelos ARIMA’. Los componentes tendenciales, estacionales, irregulares, y cíclicos son estimados y predichos con técnicas de extracción de señal aplicadas a los modelos ARIMA. Son obtenidos los errores estándar de las estimaciones y predicciones y la estructura basada en modelo se explota para contestar a cuestiones de interés en el análisis en el corto plazo de los datos. (...) Cuando se usan [los programas TRAMO y SEATS] para ajuste estacional, TRAMO previamente adapta la serie que va a ser ajustada por SEATS.” (Maravall y Gómez 1997).

- ✓ **Otros filtros:** En el proceso de filtrado es posible que estemos interesados en aplicar una combinación de varios métodos de filtrado. Un filtro compuesto es la aplicación de varios filtros simples sucesivamente sobre la serie original. El módulo de la función de transferencia resultado de la combinación de filtros tendrá un módulo igual al producto de los módulos de los filtros componentes y una fase igual a la suma de las fases de los mismos. La aplicación sucesiva de un filtro MA y un filtro AR tendrá como consecuencia respectiva la atenuación de altas frecuencias y la atenuación de bajas frecuencias, resultando un pico en la función de transferencia de la serie filtrada y una atenuación de determinadas frecuencias intermedias. Los filtros más ampliamente usados, sin embargo, son filtros ad-hoc, de expresión variable. Concretamente, filtros de la familia Butterworth a la cual pertenece el filtro de Hodrick y Prescott (1980), Prescott (1986) y los de tipo Henderson entre los que destaca el utilizado en el método de desestacionalización X-11.

El filtro de Hodrick y Prescott (HP) adopta la expresión:

$$y^c = \left(\frac{\lambda(1-L)^2(1-L^{-1})^2}{1 + \lambda(1-L)^2(1-L^{-1})^2} \right) y$$

Donde λ es el parámetro que penaliza la variación en el crecimiento y para el cual los autores recomiendan un valor arbitrario de $\lambda=1600$ cuando se utilizan datos

² Una exposición adecuada requiere la lectura íntegra de dicha referencia y otras que allí se indican.

trimestrales. Este filtro elimina el componente no estacionario, es simétrico y no causa desfase en la serie filtrada. Este filtro es parecido a un filtro de paso alto, es decir, asigna valor cero a la frecuencia 0 y valores cercanos a uno para valores de frecuencias altas. Precisamente para este valor de $\lambda=1600$, se aproxima a un filtro lineal tipo Butterworth que permite el paso de la información para frecuencias de corte $w = \frac{\pi}{16}$, es decir, permite el paso de información cíclica por encima de 32 trimestres.

El filtro X-11, ampliamente usado para desestacionalizar las series económicas, es un filtro tipo Henderson cuya expresión inicial en su primera iteración en datos mensuales sería del tipo 1-MA(12). Se trata de otro filtro de paso alto centrado en el procedimiento, además, se combina con el análisis ARIMA utilizando estimaciones para los extremos de la media móvil que se pierden en el centrado.

3.- EL SECTOR TURÍSTICO EN CASTILLA-LA MANCHA

Nuestro trabajo empírico se centrará en la Comunidad Autónoma española de Castilla-La Mancha, formada por las provincias de Albacete, Ciudad Real, Cuenca, Guadalajara y Toledo.

En este sentido, es preciso destacar que la única fuente de información sobre la demanda turística extrapolable al conjunto de provincias de la Comunidad es, a priori, la proporcionada por el Instituto Nacional de Estadística y por los estudios del Instituto de Estudios Turísticos. En el ámbito provincial existen estadísticas de ciertas tipologías de alojamiento, pero ninguna de ellas es común para todas las provincias y por lo tanto no se pueden estimar datos para la Comunidad.

3.1. La demanda en Castilla-La Mancha

Castilla-La Mancha ocupa el 9º lugar como comunidad autónoma de destino de viajeros nacionales alojados en establecimientos hoteleros. Representa el 4,08% de los viajeros nacionales alojados en este tipo de establecimientos en España durante el año 2004.

Además, la región ocupa el 10º lugar como comunidad autónoma de destino de viajeros extranjeros alojados en establecimientos hoteleros en el año 2004. Con ello, nuestra

región representa el 1,1% de los viajeros extranjeros alojados en establecimientos hoteleros entrados en España.

Castilla-La Mancha recibe según las cifras oficiales del INE alrededor de 2.000.000 de viajeros que pernoctan en algún establecimiento de la Comunidad. Desde el año 2001 el número de viajeros ha ido descendiendo ligeramente (2,3%), tanto si se trata de viajeros extranjeros como nacionales. En el año 2004 según los últimos datos se experimentó otra vez un crecimiento del 6% respecto al año anterior.

La proporción entre viajeros nacionales y extranjeros se mantiene más o menos constante en los últimos años. El mercado nacional es el mayoritario oscilando entre el 82,4% en el año 2002 y el 84,3% en el año 2004, el mercado extranjero oscila entre el 15,7% del año 2004 y el 17,8% del año 2002.

En la curva de estacionalidad se pueden observar dos picos, uno menor en Abril, coincidiendo con el periodo vacacional de Semana Santa, y otro más marcado en los meses de Julio a Septiembre. En los últimos años no ha habido grandes variaciones en la distribución de los viajeros por meses. En el análisis según el tipo de alojamiento utilizado, se observa claramente un retroceso en Hoteles y Pensiones, en cambio aumentan las otras modalidades de alojamiento.

El número de viajeros extranjeros representa aproximadamente un 17% sobre el total de los turistas. Los países emisores más importantes para Castilla-La Mancha son Francia, que representó en el año 2004 el 19,3% sobre el total del mercado extranjero; el Reino Unido representa el 11,5%; Estados Unidos, 9,7%; Alemania, 9%; y Portugal e Italia, 8,7% y 6,9%, respectivamente.

Más de la mitad, el 54%, de los viajeros extranjeros de Castilla-La Mancha, que representaron en el año 2004 el 15,7% del total, se alojó en la provincia de Toledo, el 14% en la provincia de Ciudad Real, el 13,1% en Guadalajara, el 10,2% en Cuenca y el 8% en la provincia de Albacete.

La estacionalidad no es idéntica para todas las provincias aunque es similar para Ciudad Real, Cuenca y Toledo, con dos picos marcados en abril-mayo y agosto. Guadalajara

presenta una curva de estacionalidad más regular y Albacete registra mayor entrada de viajeros en el mes de septiembre.

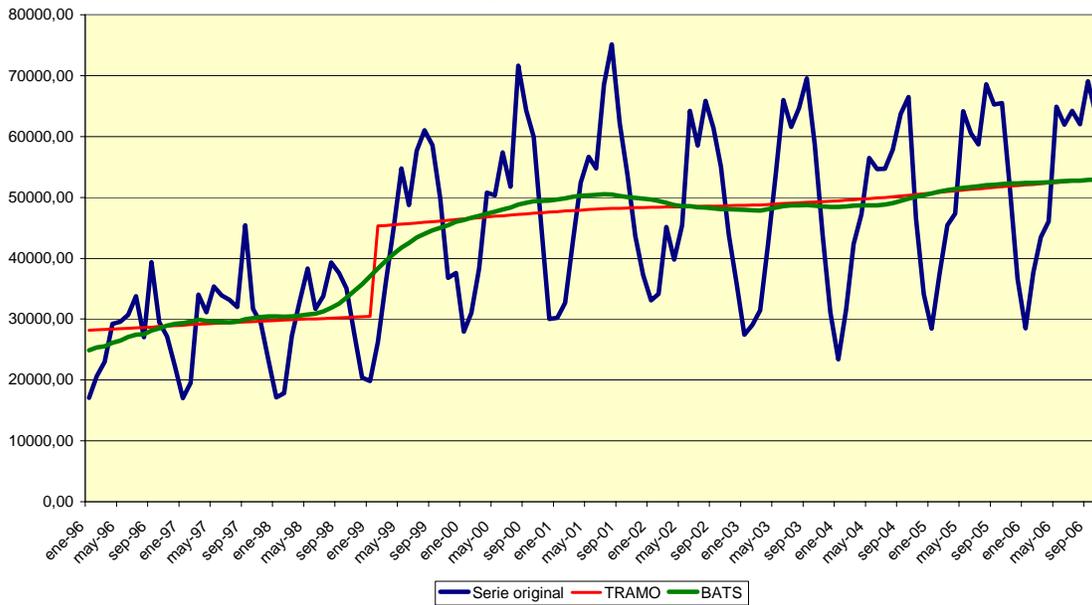
Los datos de la presente temporada consolidados hasta el mes de septiembre muestran tendencias positivas para el turismo en la región. Castilla-La Mancha es la octava región de España que más crece, un 6,9%, 2,3 puntos por encima de la media nacional. Castilla-La Mancha está entre las regiones que más aumenta la recepción de turistas internacionales, un 21,7% (cuando la media nacional es del 6,3%). Es la tercera región de España que más crece en pernoctaciones, cerca del 10%, tras Cantabria y Madrid. Es la Comunidad Autónoma que más crece en afiliados a la Seguridad Social en el sector turístico, con un incremento medio del 11,5%, más del doble que la media nacional. Por último, Castilla-La Mancha está entre las regiones que más aumenta el gasto por turista, un 36,2% (cuando la media nacional es del 4,3%), además, el gasto medio diario se sitúa en 80,9 euros, el octavo más elevado de cuantos se producen en el resto de Comunidades Autónomas (Consejería de Industria y Tecnología, 2006).

4.- PREDICCIÓN DE SERIES TURÍSTICAS EN CASTILLA-LA MANCHA

En esta fase se pretende llevar a cabo una aplicación práctica con el fin de eliminar de cada serie de turismo el ruido inherente propio de cada serie; dejando, por tanto, la componente desestacionalizada y corregida de efecto calendario. Para realizar dicha extracción se han considerado la posibilidad de utilizar el software *TRAMO-SEATS*, basado en un modelo ARIMA univariante para la serie, a partir del cual se obtienen modelos univariantes para los componentes ciclo-tendencia, estacional y ruido (metodología UCARIMA).

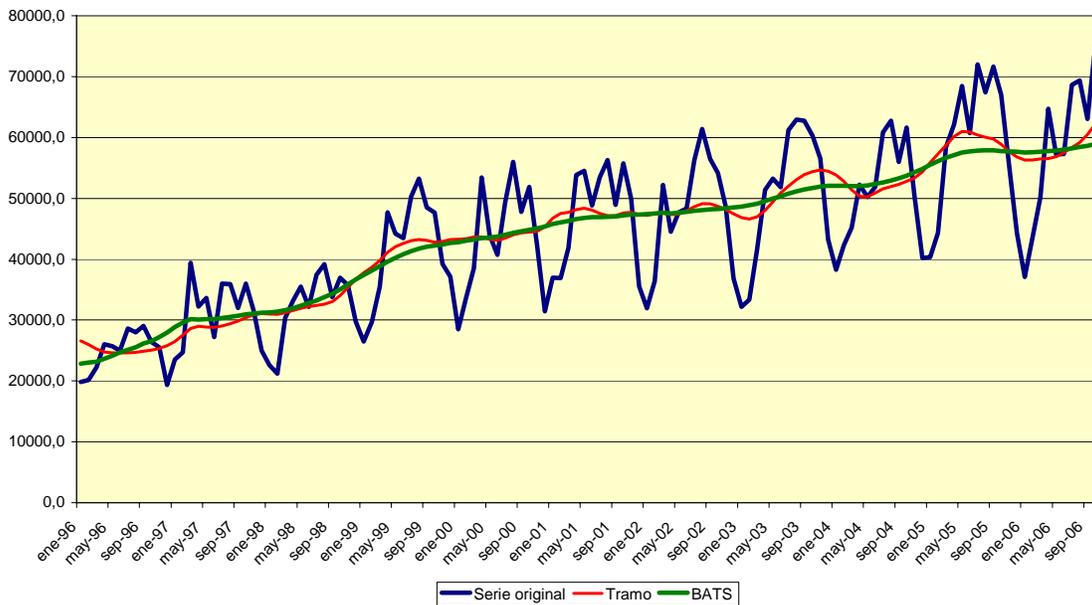
Debido a las particularidades de las series del turismo, estas se ven fuertemente afectadas por la estacionalidad (vacaciones de verano, puentes, festivos, etc...) y por un importante “efecto calendario” (de especial importancia para Castilla-La Mancha resulta el efecto de Pascua móvil). Es por lo que, para realizar una análisis más en profundidad, deberán extraerse las series desestacionalizadas y corregidas de efecto calendario, tal y como podemos observar en los siguientes gráficos:

Figura 1. Evolución mensual de las pernoctaciones en la provincia de Albacete
(serie original y ciclo-tendencia)



Fuente: Elaboración propia a partir de Datos INE (2006).

Figura 2. Evolución mensual de las pernoctaciones en la provincia de Ciudad Real
(serie original y ciclo-tendencia)

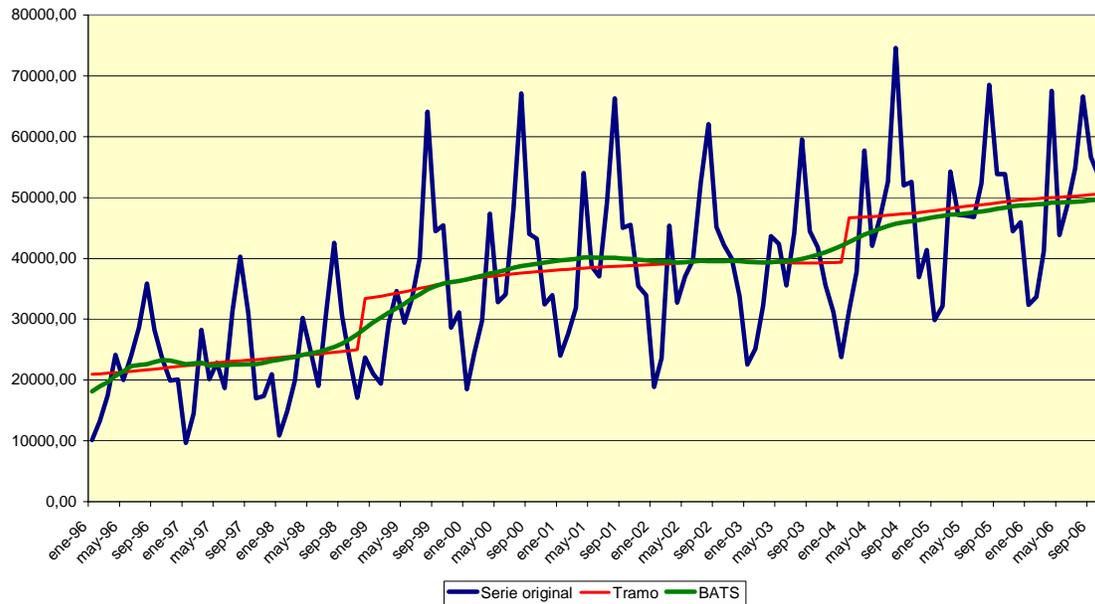


Fuente: Elaboración propia a partir de Datos INE (2006).

Como se observa en los gráficos, la serie original muestra un elevado componente estacional, derivado de las características del sector turístico que antes hemos descrito. En este

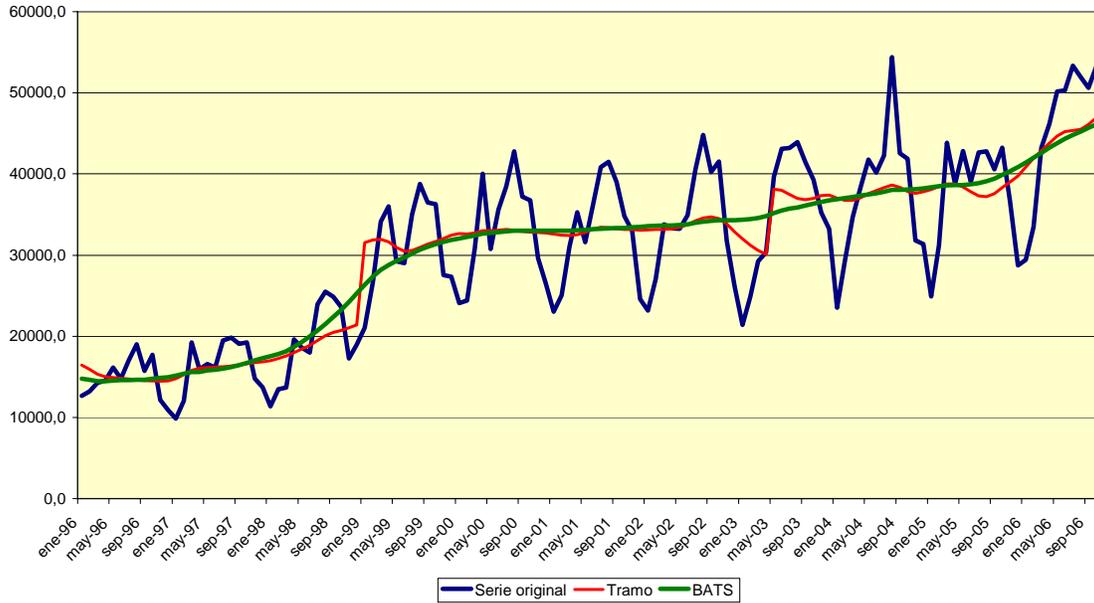
sentido, el efecto de la desestacionalización de las series en base a nuestra metodología supone obtener unas líneas más adecuadas a una evolución que pueda ser comparable desde el punto de vista analítico. Así, la figura 1 muestra la evolución de la provincia de Albacete, donde se observa un crecimiento sostenido y equilibrado, mientras que en las figuras 2 y 3, las provincias de Ciudad Real y Guadalajara, respectivamente, presentan mayor incremento en la evolución mensual de las cifras de turistas alojados en sus establecimientos hoteleros.

Figura 3. Evolución mensual de las pernoctaciones en la provincia de Cuenca
(serie original y ciclo-tendencia)



Fuente: Elaboración propia a partir de Datos INE (2006).

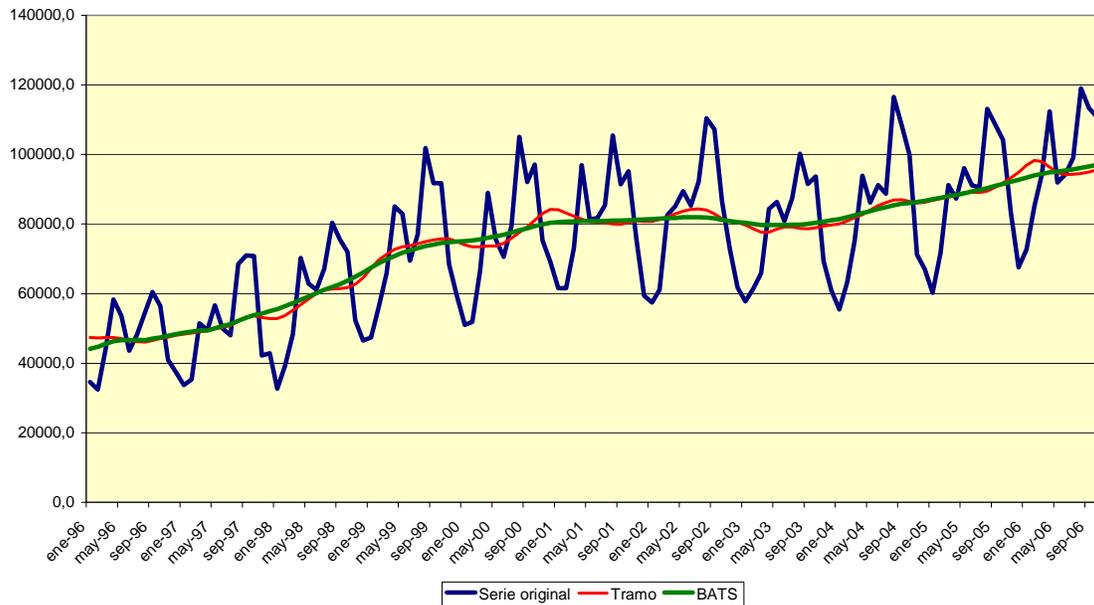
Figura 4. Evolución mensual de las pernoctaciones en la provincia de Guadalajara
(serie original y ciclo-tendencia)



Fuente: Elaboración propia a partir de Datos INE (2006).

Asimismo, Toledo presenta (Fig. 5) una evolución constante al alza, donde observamos la importante corrección que reciben los datos tras la desestacionalización, evitando las grandes fluctuaciones derivadas del efecto calendario.

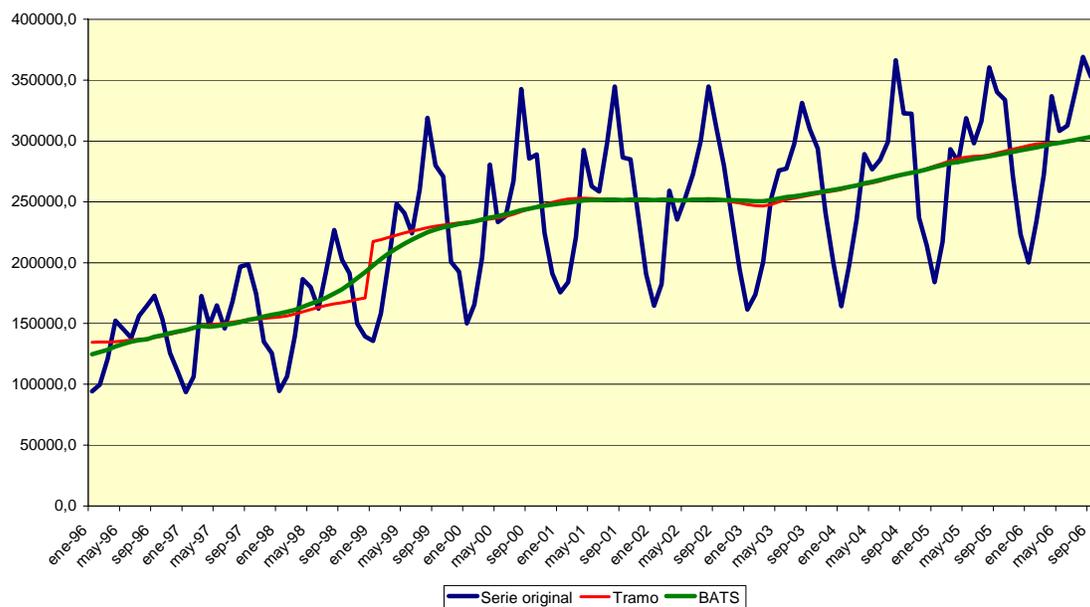
Figura 5. Evolución mensual de las pernотaciones en la provincia de Toledo
(series original y ciclo-tendencia)



Fuente: Elaboración propia a partir de Datos INE (2006).

Por último, la figura 6 muestra los datos globales de la región. Evidentemente, el importante efecto producido por la componente estacional de la serie destaca también para los datos agregados.

Figura 6. Evolución mensual de las pernoctaciones en Castilla-La Mancha
(serie original y ciclo-tendencia)



Fuente: Elaboración propia a partir de Datos INE (2006).

Si, además del análisis anterior, realizamos una estimación de las previsiones de evolución de los datos correspondientes a las cinco provincias y al total agregado regional, el resultado aparece en las tablas 1 y 2.

Tabla 1. Previsión de pernoctaciones por provincias (TRAMO-SEATS)

Pernoctaciones	AB-Tramo	CR-Tramo	CU-Tramo	GU-Tramo	TO-Tramo	CLM-Tramo
nov-06	49505,82	61034	43927	42097	84066	278129
dic-06	36886,10	48729	46530	36715	71673	234277
ene-07	28985,66	43999	31674	32004	70238	204980
feb-07	36943,94	47966	34146	37538	79401	239760
mar-07	43343,01	57573	42368	46335	89415	279410
abr-07	50039,62	72139	67608	51043	112919	351979
may-07	63377,31	67857	45857	51344	101439	328758

jun-07	63673,78	64636	50502	51099	99261	328237
jul-07	63373,58	76313	55260	55668	103265	353116
ago-07	66699,41	78256	70067	58691	127191	393261
sep-07	70573,27	73615	58794	53594	118457	371602
oct-07	65478,93	75444	54145	54872	115908	369882
nov-07	50657,06	66039	44996	44322	87962	291346
dic-07	37680,65	52725	47662	38655	74994	245411
ene-08	29593,55	47608	32445	33695	73492	214721
feb-08	37635,64	51803	34811	39558	83230	251070
mar-08	49796,74	72073	61739	54006	107176	345111
abr-08	45377,05	67465	48680	48544	103139	312699
may-08	65147,64	73904	47757	53886	105471	344785
jun-08	64991,17	69937	51730	53799	103860	343836
jul-08	64229,03	82033	55674	58797	108735	369464
ago-08	69048,73	85789	74189	61401	131414	412915
sep-08	71022,01	78617	58262	56786	125521	388353
oct-08	66833,59	81632	55462	57772	121279	387460

Fuente: Elaboración propia.

Tabla 2. Previsión de pernотaciones por provincias (BATS)

Pernотaciones	AB-BATS	CR-BATS	CU-BATS	GU-BATS	TO-BATS	CLM-BATS
nov-06	50505	57902	44485	42449	86932	282273
dic-06	37461	46558	45290	37656	75786	242753
ene-07	30356	42087	33312	34641	75174	215573
feb-07	38226	46939	36035	39411	85314	245927
mar-07	45328	55353	47466	48695	97662	294507
abr-07	48651	64830	61574	49622	109977	334655
may-07	64017	62849	48020	53443	101527	329857
jun-07	63076	60888	50421	52847	101323	328556
jul-07	63427	71413	56978	55873	104005	351698
ago-07	66195	71359	71654	57222	125197	391628
sep-07	69246	68080	57726	53993	119132	368178
oct-07	65851	73099	55996	55617	115259	365822
nov-07	51645	60048	46425	46803	91991	296913
dic-07	38601	48704	47229	42011	80846	257393
ene-08	31496	44233	35252	38995	80233	230213
feb-08	39366	49085	37974	43765	90373	260567
mar-08	46468	57499	49406	53049	102721	309147
abr-08	49791	66976	63513	53976	115036	349295
may-08	65157	64995	49959	57797	106586	344497
jun-08	64216	63034	52361	57201	106382	343196

jul-08	64567	73559	58917	60227	109065	366338
ago-08	67335	73505	73594	61576	130257	406268
sep-08	70386	70226	59666	58347	124192	382818
oct-08	66991	75245	57936	59971	120318	380462

Fuente: Elaboración propia.

5.- CONCLUSIONES, LIMITACIONES Y OPORTUNIDADES DE INVESTIGACIÓN

El filtrado de los datos proporciona al investigador unas series más adecuadas para el análisis y estudio del sector turístico debido a la fuerte estacionalidad del mismo. El uso de un determinado modelo de filtrado condiciona los resultados obtenidos: mientras los modelos tradicionales (metodología ARIMA) conforman unos mejores resultados a corto plazo, los modelos bayesianos son más recomendables para el estudio de la evolución a largo plazo.

Este filtrado de series supone el punto inicial de un estudio más ambicioso donde se realizará un análisis de cointegración de las variables estudiadas, para establecer las relaciones de cointegración que pueden presentar las cinco provincias de la comunidad autónoma de Castilla-La Mancha y su comportamiento a nivel agregado para el conjunto de la región.

Desde el punto de vista de las limitaciones de nuestro trabajo, debemos destacar el hecho de haber considerado exclusivamente datos de una Comunidad Autónoma, lo que deriva inmediatamente la primera línea de investigación futura en que trabajamos actualmente: ampliar el estudio al resto de comunidades autónomas, de tal forma que podamos efectuar una comparación a nivel nacional.

A este mismo nivel se refiere el hecho de trabajar exclusivamente con un reducido número de variables de las que forman parte de la EOH. Es posible ampliar el análisis a nuevos ámbitos de estudio que permiten obtener una mayor riqueza informativa desde el punto de vista de la investigación turística.

6.- BIBLIOGRAFÍA

ALFARO, J.L., MONDÉJAR, J. Y VARGAS, M. (2004): Time structure of agricultural macromagnitudes in Castilla-La Mancha. *57th International Atlantic Economic Conference*, Lisboa (Portugal), 10 a 14 de marzo.

BAXTER, M. Y KING, M. (1995): "Measuring business cycles approximate band-pass filters for economic time series". *NBER Working Paper n. 5022*, Cambridge (Massachusetts).

CANOVA, F. (1998): "Detrending and business cycle facts". *Journal of Monetary Economics*, vol. 41, pp. 475-512.

CHÍAS, J. (2005): *El negocio de la felicidad. Desarrollo y marketing turístico de países, regiones, ciudades y lugares*. Financial Times – Prentice Hall, Madrid.

CONSEJERÍA DE INDUSTRIA Y TECNOLOGÍA (2006): *Plan de Ordenación y Promoción del Turismo en Castilla-La Mancha*. Dirección General de Turismo y Artesanía, Junta de Comunidades de Castilla-La Mancha, Toledo.

ESPASA, A. Y CANCELO, J. R. (1993): *Métodos cuantitativos para el análisis de la coyuntura económica*. Alianza Economía, Madrid.

ESTEBAN, A. (2000): "La investigación turística en la Universidad Española", *Estudios Turísticos*, nº 144-145, págs. 155-180.

GÓMEZ, V. Y MARAVALL, A. (1997): "Programs TRAMO (Times Series Regression with ARIMA Noise, Missing Observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series). Instructions for the User.

GÓMEZ, V. Y MARAVALL (1998): "Seasonal adjustment and signal extraction in economic time series". *Documento de trabajo 9809*. Servicio de Estudios del Banco de España. Madrid.

GÓMEZ, M. A., MONDÉJAR, J.A. Y SEVILLA, C. (coords.) (2005): *Gestión del turismo cultural y de ciudad*. Ediciones de la Universidad de Castilla-La Mancha, Cuenca.

HODRICK, R. Y PRESCOTT E. C. (1980): "Post-war U.S. Business cycle: an empirical investigation", mimeo, Carnegie-Mellon University, Pittsburgh P.A.

INSTITUTO NACIONAL DE ESTADÍSTICA (2006): Encuesta de Ocupación Hotelera (Datos 2006 y anteriores). Disponible en www.ine.es.

JEFFREYS, H. (1961): *Theory of Probability*. Oxford University Press, London.

JIMÉNEZ, J. F.; GÁZQUEZ, J. C. Y SÁNCHEZ, R. (2006): "La capacidad predictiva en los métodos Box-Jenkins y Holt-Winters: una aplicación al sector turístico". *Revista Europea de Dirección y Economía de la Empresa*, vol. 15, nº. 3, pp. 185-198.

MELIS, F. (1991): "La estimación del ritmo de variación en series económicas". *Estadística Española*, vol. 33, nº. 126, pp. 7-56.

MONDÉJAR, J. (2003): Evaluation methods for complex indicators. *56th International Atlantic Economic Conference*. Québec City (Canadá), 16 a 19 de octubre.

MONDÉJAR, J. (2006): *Análisis cuantitativo de la coyuntura económica. Una aplicación de la representación en espacio de estados de series temporales múltiples*. Tesis Doctoral, Ediciones de la Universidad de Castilla-La Mancha, Cuenca.

PRESCOTT, E. (1986): "Theory ahead of business cycle measurement". *Carnegie-Rochester Conference series on Public Policy*, nº. 25 , pp. 11-66.

SÁNCHEZ, M. Y MARÍN, M.B. (2003): "La investigación en turismo y economía de la empresa publicada en revistas especializadas españolas: 1996-2001". *Papers de turisme*, nº 33, págs. 6-39.

VERA, J.F. E IVARS, J.A. (2001): La formación y la investigación turística en España: una visión de síntesis. *Papers de turisme*, nº 29, págs. 6-27.