

METODOS DE PREDICCION EN SERIES TEMPORALES TURISTICAS

José García Pérez
Departamento de Economía Aplicada
Universidad de Almería
e-mail: nikita97@larural.es

José Felipe Jiménez Guerrero
Departamento de Dirección y Gestión de Empresas
Universidad de Almería
e-mail: jfjimene@ual.es

Yolanda Rosado López
Universidad de Almería
e-mail: yolandarosado@terra.es

Palabras Clave: Predicción, Series Turísticas, Estacionalidad
Área Temática: Métodos Cuantitativos. 7

RESUMEN:

En la presente comunicación se ha pretendido realizar un análisis de la capacidad predictiva de dos métodos ampliamente utilizados, la metodología Box-Jenkins y el método de Holt-Winters basado en los métodos clásicos de descomposición, aplicándolo en ambos casos a una serie con estacionalidad procedente del sector turístico. Tras una breve introducción al trabajo, en un segundo apartado se realiza una revisión de algunos métodos usados para la modelización de series turísticas. A continuación se dedica un tercer bloque a exponer las dos metodologías usadas para el desarrollo de nuestro caso práctico, para concluir con un cuarto y último apartado en el que se realiza la comparación de las predicciones obtenidos por ambos métodos con los valores reales de la serie para ese periodo, usando los estadísticos de la RECM y el EAM.

La serie utilizada para el desarrollo del caso práctico es “Viajeros Alojados en establecimientos hoteleros de la provincia de Almería” en el período 1986-1999, serie que presentaba una estacionalidad muy acentuada y que se prestaba a un análisis de este tipo.

METODOS DE PREDICCION EN SERIES TEMPORALES TURISTICAS

*José García Pérez
José Felipe Jiménez Guerrero
Yolanda Rosado López*

Universidad de Almería

1. INTRODUCCIÓN

La permanente disyuntiva presente en el ámbito económico empresarial respecto a la toma de decisiones dentro de un entorno incierto, es uno de los principales argumentos que avalan la necesidad de analizar los fenómenos económicos. Los decisores esperan obtener ciertas evidencias que les ayuden a reducir la incertidumbre sobre sucesos futuros y así tomar las decisiones más acertadas.

En el presente trabajo pretendemos contrastar la utilidad de los métodos de previsión como herramienta de predicción de fenómenos económicos. Para ello utilizaremos el análisis univariante de series temporales y a partir de la serie de viajeros alojados en establecimientos hoteleros en la provincia de Almería evaluaremos la capacidad predictiva de distintos modelos.

En un primer apartado se hace una breve revisión de la literatura referente a las aplicaciones econométricas realizadas a series estadísticas económicas, y más concretamente turísticas, al objeto de compilar los estudios más recientes sobre este tema. En un segundo apartado y partiendo de la serie de viajeros alojados en establecimientos hoteleros aplicamos a la misma las dos metodologías básicas del análisis univariante

de series temporales, esto es, los métodos clásicos de descomposición y los modelos ARIMA, al objeto de preparar las series para realizar predicciones. Finalizamos nuestro trabajo con la valoración de los resultados obtenidos, comparando las predicciones resultantes de la aplicación de ambos métodos con los datos reales para ese año.

2. ANALISIS ECONOMETRICO EN INDICADORES TURISTICOS

El sector turístico es objeto de estudio en economía cada vez con más insistencia debido al papel tan importante que tiene en la estructura productiva de un país. Los ingresos por turismo “constituyen uno de los soportes básicos del equilibrio de la balanza por cuenta corriente”¹ -Busian, A. - y, por tanto, la obtención de predicciones sobre la futura evolución de estos indicadores puede suponer un argumento de gran interés para decidir políticas institucionales a nivel nacional, autonómico o local.

La metodología usada para el análisis del sector turístico ha sido muy variada, y con ello se ha pretendido llevar a cabo básicamente dos tipos de estudio: por un lado la modelización de los cambios del sector turístico, y por otro la obtención de previsiones de la evolución futura de la demanda turística.

Haciendo una breve revisión a la bibliografía vemos como uno de los primeros trabajos que se llevaron a cabo en esta línea, fue el del profesor Pulido San Román, A. (1966). Su trabajo pretendía encontrar una modelización válida para el estudio del sector turístico contemplando los flujos internacionales de turistas. Para ello, diferenciaba entre 17 países emisores y 14 zonas de destino, aunque uno de sus grandes handicap fue la información estadística con la que trabajar por lo que tuvo que llevar a cabo filtrados de series y elaboración de información estadística inédita.

¹ Busian, A. “Evolución de los ingresos por turismo desde una perspectiva reciente”

Posteriormente Almagro, J. (1979) lleva a cabo una modelización ARIMA estacional de las series “entrada de extranjeros y no residentes” como proxy de los ingresos por turismo. En este caso la serie recogía 156 observaciones que iban desde enero de 1965 a enero de 1978. La mayor aportación de este trabajo es la obtención de un modelo univariante que modelizaba el comportamiento de series turísticas, aunque no puso especial énfasis en las capacidades predictivas.

El Cercle d'estudis economics de les Illes Balears (1981) realiza un trabajo que estudia dos tipos de modelos distintos: por un lado, un análisis multivariante en el que introduce tanto variables turísticas como variables económicas y sociales que se consideran relevantes para el estudio de los ingresos por turismo, realizando distintas regresiones en función del país de procedencia del turista; en una segunda parte del trabajo se realiza una modelización univariante de la serie total de entradas en el aeropuerto de Palma, diferenciando por nacionalidades, de forma que se llevan a cabo previsiones de entrada de turistas de distintos países.

La metodología de Box-Jenkins volvió a ser usada por Bru Parra, S. y Usach Domingo, J (1987), llevando a cabo el análisis de la serie mensual “Pernoctaciones totales en la Comunidad Valenciana” durante el periodo 1975-1985. En este segundo caso el modelo estimado es usado con fines predictivos, aunque según los autores las estimaciones realizadas perdían capacidad previsional a partir del cuarto periodo.

Ya en la década de los noventa Espasa, Gómez-Churruca, y Morales (1993) realizan un análisis econométrico del turismo en España, en el llevan a cabo la estimación de dos modelos relacionados. En uno de ellos la variable endógena elegida son los “Ingresos totales reales” y en el segundo la variable dependiente pasa a ser “Número total de turistas entrados en España”. Para realizar la modelización en ambos casos hacen uso de una serie de variables explicativas, que mantienen una estructura de retardos, y

que en ambos casos son las mismas. La interrelación de los dos modelos podría asemejarse a un modelo VAR, sin embargo las variables endógenas no intervienen como exógena en la modelización de la otra variable dependiente, siendo la intención de los autores realizar un análisis de causalidad dinámica.

Martínez Sánchez (1995), realiza un trabajo titulado “La demanda turística alemana en la costa del sol”. Haciendo uso de la misma metodología en 1999 el autor realiza el análisis para el turismo procedente de Gran Bretaña². En ambos casos el autor analiza el turismo por país de procedencia y elabora un modelo basado en una sola ecuación en la cual introduce como variable dependiente “la entrada de turistas alemanes/ingleses por el aeropuerto de Málaga”, y como variables explicativas la relación entre el coste de la vida en España y en el conjunto de países competidores, los valores del PIB en Alemania/Inglaterra y la relación entre la variación del coste de la vida en España y en Alemania/Inglaterra. El autor realiza la modelización trabajando con la transformación logarítmica de las variables explicativas antes citadas, he introduciendo variables ficticias para modelizar la estacionalidad. Su intención es explicar en que medida influye cada uno de estos factores en la llegada de visitantes a la provincia de Málaga.

Destacamos también dos trabajos de Otero, J.M. (1996) sobre modelizaciones turísticas. En el primer, lleva a cabo un análisis multivariante sobre los factores determinantes de la entrada de pasajeros a través del aeropuerto de Málaga³. Para la modelización estima cuatro modelos alternativos, y el que considera mas adecuado contiene componentes dinámicos para modelizar la estacionalidad. Las variables explicativas⁴ son retardadas cuatro periodos para poner de manifiesto la correlación existente entre un año y el mismo periodo del año anterior. El modelo estimado es

² “La demanda turística inglesa en la Costa del Sol” Papers de Turisme.

³ “Principales determinantes del flujo de pasajeros extranjeros desembarcados en el aeropuerto de Málaga” J.M. otero. Estudios de economía aplicada .nº5, 1996

⁴ Las variables explicativas recogen los efectos de renta y precios, al estar fuertemente relacionadas con la demanda turística. Los datos son cuatrimestrales.

usado para realizar predicciones con un horizonte de cuatro periodos y un análisis casual en función de alteraciones de las variables explicativas.

En un segundo trabajo analiza “Las fluctuaciones cíclicas del turismo en Andalucía”⁵. En este caso el estudio lo realiza partiendo de un análisis univariante sobre la serie de pernoctaciones hoteleras, realizando una modelización ARIMA para la obtención del componente tendencia-ciclo. Con este trabajo se realiza un análisis de ciclos para tratar de prever la tendencia del sector turístico andaluz, además de realizar un análisis de la evolución entre los años 79-98. Como conclusiones el autor extrae las siguientes: en primer lugar, el sector turístico se comporta de forma procíclica, aunque sufre un desfase temporal de dos o tres trimestres respecto de PIB, lo cual supone que el PIB se puede considerar un indicador adelantado del sector turístico. Una segunda conclusión que extrae el autor es que las recesiones son mas acusadas en el sector turístico que en el PIB.

Por ultimo hacer un reseña al documento de trabajo realizado por. Ledesma F.J. y otros⁶ (2000), titulado “Datos de Panel y Demanda de Turismo: El caso de Tenerife” en el que llevan a cabo el estudio de la demanda de servicios turísticos en la isla de Tenerife. Para ello hacen uso de distintos modelos tanto estáticos como dinámicos, introduciendo modelos de ecuaciones simultáneas estimadas tanto por mínimos cuadrados bietápicas como trietápicas, realizando por último un proceso de simulación para los distintos países de procedencia del turista.

Resumiendo, podríamos realizar por tanto dos clasificaciones en las investigaciones llevadas a acabo: de un lado, y en función de la perspectiva tratada dentro del sector turístico, podemos observar como son muchos los trabajos que se han centrado en los flujos de turistas extranjeros en territorio español, existiendo trabajos regionales de gran

⁵ Estudios de Economía Aplicada, nº13, 1999. Págs. 129-142.

⁶ Navarro .M y Pérez J.

calado; por otra parte, los que se han centrado en el análisis de los ingresos por turismo de la región en cuestión o a nivel nacional.

En cuanto a los modelos econométricos considerados, a lo largo de estas páginas hemos podido observar como por un lado se encuentran las modelizaciones multivariantes que toman como base una única ecuación de comportamiento, mientras que toman fuerza las modelizaciones univariantes, gracias a las facilidades que en términos de información estadística presentan. Ello se debe a que la disponibilidad de datos dificulta en muchos casos el afrontar modelos que enlacen distintas ecuaciones, como podrían ser los modelos de ecuaciones simultáneas, o modelos dinámicos de múltiples ecuaciones como los modelos VAR. Espasa, A. (1996) expone en su trabajo “Características de la demanda en los estudios econométricos sobre el turismo e implicaciones de política económica y estrategia empresarial”⁷ la necesidad de realizar estudios econométricos cada vez más desagregados para poder captar todas las particularidades de los grupos homogéneos de visitantes, considerando, no obstante, este trabajo difícil de llevar a cabo debido a las deficiencias de datos que existen.

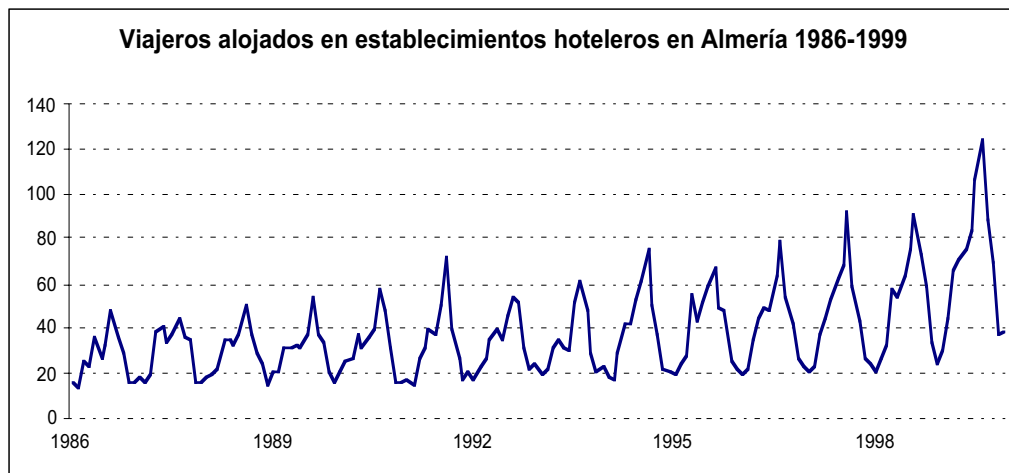
3. ANÁLISIS Y PREDICCIONES EN LA SERIE “VIAJEROS ALOJADOS”

En el presente apartado se pretende estudiar la serie mensual de viajeros alojados en establecimientos hoteleros en la Provincia de Almería durante el periodo 1986-1999 a través de las dos principales modelizaciones del análisis univariante de series temporales, esto es, los métodos clásicos de descomposición de series y los modelos ARIMA.

Los datos proceden de la Encuesta de ocupación hotelera (hasta 1999 Encuesta de movimientos de viajeros en establecimientos hoteleros) que mensualmente realiza el Instituto Nacional de Estadística. Como podemos apreciar en el gráfico, la serie presenta grandes oscilaciones debido a la

⁷ Boletín Económico del ICE, Enero de 1996, nº 749.

fuerte estacionalidad que presenta este indicador con picos en los meses estivales y mínimos en la temporada invernal.

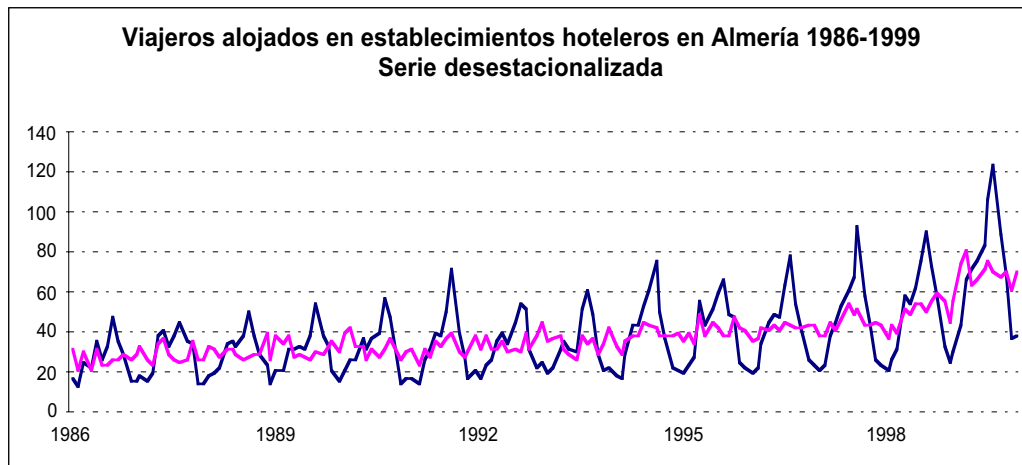


El hecho de analizar una serie en la que el factor estacional está muy acentuado requiere de un tratamiento previo si pretendemos realizar predicciones. Este paso consiste en la desestacionalización de la serie. Para ello utilizamos el método más utilizado de desestacionalización: la razón a la media móvil.

Como resultado de la aplicación del procedimiento anterior obtenemos una nueva serie ,más suavizada, y libre del componente estacional.

Pretendemos a partir de la modelización de esta serie realizar predicciones para el año 2000 evaluando la capacidad predictiva del modelo a través de diferentes criterios de jerarquización, bajo la premisa de conocer los dato reales del año de predicción.

En el siguiente gráfico podemos ver el contraste entre la serie original, muy fluctuante y la serie desestacionalizada, con una línea más suavizada.



3.1. LOS MODELOS ARIMA

La metodología de Box-Jenkins, ampliamente utilizada a lo largo de los últimos años, en nuestro caso se ha llevado a cabo con la ayuda del programa informático TRAMO-SEATS (Gómez. V. y Maravall, A. 1994,1996). Este programa, que realiza la estimación a través del método de máxima verosimilitud, mejorando la precisión que ofrece la estimación tradicional, obtiene el modelo ARIMA del cual procede la serie objeto de estudio.

El output del programa, resultado de analizar la serie, estima un modelo ARIMA estacional $(0,1,1) \times (0,1,1)_{12}$ que responde a la siguiente especificación:

$$(1 - B)(1 - B^{12})L Aloi_t = (1 - 0.81061B)(1 - 0.7336B^{12})\epsilon_t$$

Como paso previo a la obtención del proceso generador de los datos se realiza un análisis de estacionariedad de la serie, decidiendo trabajar con la serie transformada, a la que se le han aplicado dos transformaciones sucesivas: una primera, consistente en tomar logaritmos, con el objeto de suavizar el componente estacional,

$$L Aloi_t = \ln(Aloi_t)$$

y además, una segunda en la que se han llevado a cabo dos diferenciaciones de esta serie transformada, una primera diferencia de la serie y una diferencia estacional.

$$TLal_{oj} = (1 - B)(1 - B^{12})LAlo_{j,t}$$

El programa lleva a cabo la descomposición de la serie en sus componentes tendencia, ciclo, componente estacional e irregular, lo cual permitirá en sucesivos trabajos realizar un estudio detallado de cual es el comportamiento de cada componente a lo largo del tiempo.

Una vez llevada a cabo la estimación, el programa realiza la predicción de la serie original de alojados para 24 periodos, los cuales son objeto de discusión en la presente comunicación y se analizarán más detenidamente en el capítulo de conclusiones. Sin embargo comentar que para llegar a obtener resultados de la capacidad predictiva sólo se han hecho uso de las predicciones correspondientes al primer año, puesto que se contaba con las realizaciones para ese periodo.

3.2. LOS METODOS DE DESCOMPOSICION

Una segunda alternativa a la hora de analizar series temporales son los llamados métodos clásicos de descomposición. En este caso se suele considerar que la serie se puede descomponer en todos o algunos de los siguientes componentes: a) un componente que refleja la evolución a largo plazo de la serie histórica y al que llamaremos tendencia, b) un segundo componente que recoge oscilaciones de carácter periódico, pero no regular, y a medio plazo, al que llamaremos factor cíclico; c) un tercer componente que recoge oscilaciones a corto plazo, es decir, iguales o inferiores al año, y que llamamos estacionalidad

y d) un cuarto y último componente residual que recoge tanto aspectos accidentales como acontecimientos no siempre previsibles, y que llamamos componente irregular.

De estos cuatro componentes es la tendencia junto al componente estacional los que plantean un mayor interés. El primero de ellos refleja la línea maestra de la evolución de una serie económica; por su parte el componente estacional es un elemento en el que el factor tiempo juega un papel más trascendental que en los demás ya que, de un lado la climatología de las estaciones afecta de una forma importante a las series y, además, el propio calendario está permanentemente condicionado por determinados hechos o costumbres.

Según los métodos de descomposición, las series son el resultado de la integración de esos cuatro componentes. Dicha integración puede plantearse a través de dos esquemas básicos: uno aditivo y otro multiplicativo. En el primero de ellos las fluctuaciones a lo largo de la serie son más o menos regulares sin verse afectadas por la tendencia, mientras que en el segundo tipo de esquema las fluctuaciones varían con la tendencia, como ocurre con la serie de viajeros alojados.

Coincidiendo con el tratamiento previo realizado, cuando una serie sigue un esquema multiplicativo y presenta estacionalidad, es el método de la razón a la media móvil el más apropiado, por su consistencia y uso, para eliminar el factor estacional.

Una vez desestacionalizada la serie podremos realizar predicciones para periodos futuros. En las series temporales que siguen una tendencia aproximadamente lineal, y además están sometidas a la incidencia del factor estacional, el método de predicción más adecuado resulta ser el método de Holt-Winter.

La aplicación de este método parte de un modelo teórico que va a servir de base para la predicción y que podemos expresar de la siguiente forma:

$$Y_t = (b_0 + b_1)E_t + u_t$$

donde b_0 representa el componente permanente, b_1 la pendiente de la recta y E_t el factor estacional multiplicativo.

El método plantea tres ecuaciones de alisado para estimar estos componentes. Para el caso de b_0 se estima por la variable de alisado S_t y viene dado por la siguiente ecuación:

$$S_t = \alpha \frac{Y_t}{C_{t-L}} + (1 - \alpha)(S_{t-1} + b_{1t-1}) \quad 0 < \alpha < 1$$

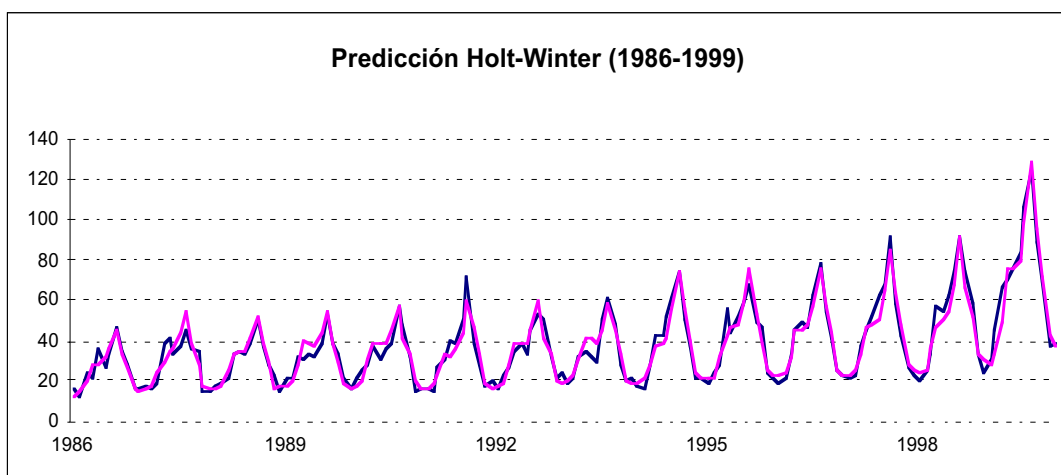
La pendiente (b_1) se estima por la variables de alisado b_{1t} y viene por la ecuación:

$$b_{1t} = \beta(S_t - S_{t-1}) + (1 - \beta)b_{1t-1} \quad 0 < \beta < 1$$

Por último el factor estacional E_t se estima por la variable de alisado C_t cuya expresión es la siguiente:

$$C_t = \gamma \frac{Y_t}{S_t} + (1 - \gamma)C_{t-L} \quad 0 < \gamma < 1$$

Para un valor de $\alpha=0,29$, $\beta=0,04$ y $\gamma=0,13$ la raíz del error cuadrático medio es mínima, por lo que estos valores son los que consiguen unas mejores predicciones. Esto se aprecia con más claridad en el siguiente gráfico en el que casi se superponen los valores originales de la serie con las predicciones obtenidas por Holt-Winter.



4. CAPACIDAD PREDICTIVA DEL MODELO.

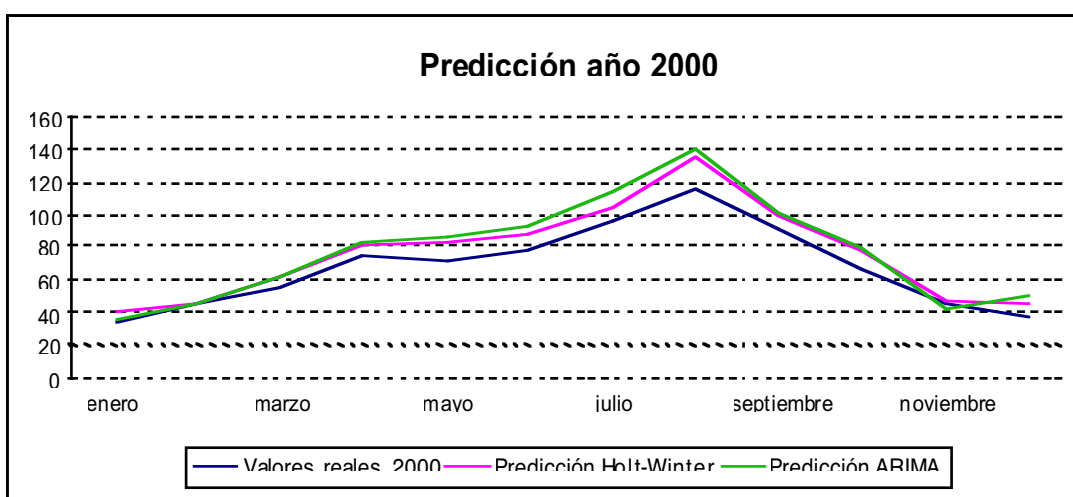
Para concluir con el objetivo planteado en este trabajo, exponemos a continuación los resultados de las predicciones mensuales para el año 2000 de cada uno de los modelos.

Al disponer de los datos reales proporcionados por la Encuesta de ocupación hotelera para ese año, podemos evaluar la capacidad predictiva de ambos métodos. Para ello se utilizan dos estadísticos: la raíz del error cuadrático medio (RECM) y el error absoluto medio (EAM)

	VALOR REAL AÑO 2000	PREDICCIONES AÑO 2000	
		HOLT-WINTER	ARIMA
enero	33,6	39,2	35,5
febrero	45,2	44,7	43,6
marzo	54,7	61,2	62,0
abril	73,5	80,7	82,6
mayo	71,8	83,4	85,0
junio	77,4	86,9	92,7
julio	95,3	105,7	113,9
agosto	115,7	135,1	140,1
septiembre	91,4	100,2	101,7
octubre	67,6	76,8	78,4
noviembre	44,5	46,0	40,2
diciembre	36,6	43,3	49,4
RECM		9,32	12,57
EAM		8,1	10,8

En ambos casos las predicciones por Holt-Winter presentan un mejor comportamiento de la RECM y del EAM que las de la modelización ARIMA. Por ello si quisiéramos realizar predicciones para el año 2001 utilizaríamos los metodología clásica.

Analizando en detalle las predicciones mensuales, podemos observar en el siguiente gráfico como los mayores errores tienen lugar en ambos métodos en los meses de mayor estacionalidad, siendo muy coincidentes en los meses invernales. Además, de forma generalizada, los valores pronosticados son superiores a los reales.



Como podemos deducir del análisis realizado los resultados obtenidos dan prioridad a los modelos clásicos sobre los modelos ARIMA, hecho que aunque parezca sorprendente también se repite al analizar la serie de “Pernoctaciones en establecimientos hoteleros” para el mismo periodo y ámbito espacial. El hecho de trabajar con series en las que la estacionalidad es un componente muy importante pensamos que pueda ser el elemento que justifique estos resultados.

BIBLIOGRAFIA

ALMAGRO, J. (1979) “Aplicaciones del enfoque Box-Jenkins a series del turismo español” *Cuadernos Económicos del ICE* pp: 55-107

BRU PARRRA, SEGUNDO y USACH DOMINGO, JUAN (1987) "Análisis de series temporales del turismo de la Comunidad Valenciana" *Estadística Española* nº114 pp111-132.

BUISAN, ANA (1996) "Evolución de los ingresos por turismo desde una perspectiva reciente" *Economistas* nº69 pp:162-167

CERCLE D'ESTUDIS ECONOMICS DE LES ILLES (1981) "Modelos multivariantes y univariantes de comportamiento y previsión de la demanda turística para las Islas Baleares"

PULIDO SAN ROMAN, A. (1966) "Introducción a un análisis econométrico del Turismo" *Instituto De Estudios Turísticos*.

ESPASA, ANTONI Y CANCEDO, J.R. (1993) "Un análisis econométrico del turismo en España: Implicaciones para el turismo sectorial de las exportaciones y algunas consideraciones de Política Económica" *Alianza.Ed.* Cap 10

ESTEBAN TALAYA, AGUEDA Y REINARES LARA, EVA (1996) "La investigación de la demanda turística en España: Recopilación y análisis" *Estudios Turísticos* nº129 pp: 81-104

LEDESMA RODRIGUEZ, F.J. Y NAVARRO IBAÑEZ, M. (2000) "Datos de panel y demanda de turismo: El caso de Tenerife" Universidad de La Laguna Tenerife. Documento de Trabajo

MARTINEZ SÁNCHEZ, JOSE MANUEL (1995)"La demanda turística alemana en la Costa del Sol" *Estudios Turísticos* nº127 pp:5-17

MARTINEZ SÁNCHEZ, JOSE MANUEL (1999)"La demanda turística inglesa en la Costa del Sol" *Papers De Turisme* nº26 pp:63 y ss.

OTERO J.M. (1996) "Principales determinantes del flujo de pasajeros extranjeros desembarcados en el aeropuerto de Málaga" *Revista De Estudios De Economía Aplicada*.

OTERO J.M.(1999) "Las Fluctuaciones cíclicas del turismo en Andalucía" *Revista de Estudios de Economía Aplicada* nº13 pp:129-142.

URIEL, E. Y MUÑIZ, M. (1993) "Estadística económica y empresarial. Teoría y ejercicios". Editorial AC