

PREVISIÓN DE SERIES TEMPORALES DE TURISMO EN LAS ISLAS BALEARES MEDIANTE MODELOS ESTADÍSTICOS AVANZADOS

INFORME FINAL

Proyectos de investigación en el ámbito del turismo en el
marco del Proyecto Escaparate Turístico Inteligente.

Miembros del equipo de investigación

Juan José MONTAÑO MORENO

(Departamento de Psicología. Universitat de les Illes Balears)

Andreu SANSÓ ROSSELLÓ

Jaume ROSSELLÓ NADAL

(Departamento de Economía Aplicada. Universitat de les Illes Balears)

Resumen: La continua reducción del tiempo transcurrido entre la reserva del viaje y el check-in que ha caracterizado el mercado turístico durante los últimos años ha supuesto un incremento del interés por reducir los errores de predicción de la demanda turística que se obtiene a través de los modelos tradicionales de previsión. La incorporación de nuevas técnicas con unos instrumentales matemáticos más complejos ha supuesto también la necesidad de disponer de grandes muestras para su correcta estimación. En el caso de la demanda turística, ante la limitación de datos históricos disponibles, no parece estar claro que las nuevas metodologías hayan conseguido sustituir a las metodologías más tradicionales a la hora de obtener mejores resultados de predicción. En este proyecto se pretenden obtener las previsiones de las principales series turísticas referidas a Baleares enfrentando para ello los modelos de redes neuronales artificiales con los modelos ARIMA. Los resultados no permiten descartar ninguna de las dos metodologías a la hora de obtener predicciones ajustadas mostrando incluso como la combinación de predicciones de ambas metodologías ofrece unos resultados muy ajustados.

Palabras clave: Predicción turística; ARIMA; Redes Neuronales Artificiales; Combinación de predicciones; Islas Baleares.

1. INTRODUCCIÓN

La naturaleza perecedera de los productos y servicios turísticos, el creciente peso de la industria turística en el producto interior bruto de la mayoría de los países y la implantación de nuevas herramientas de gestión de precios como el *Revenue Management* por parte de las empresas del sector explican el creciente interés en el sector por tratar de mejorar las predicciones de la demanda turística. Así, durante los últimos años, la incorporación de técnicas más complejas a la predicción de series temporales de naturaleza turística ha sido constante (Peng et al., 2014; Song y Li, 2008). Sin embargo, la mejora del ajuste de la predicción por parte de estas técnicas suele venir condicionada a la disponibilidad de un número elevado de observaciones.

En este trabajo se pretenden obtener previsiones ajustadas de las principales series temporales de turismo en las Islas Baleares –principalmente referidas al número de turistas–, mediante la aplicación de modelos estadísticos clásicos ARIMA (Autoregressive Integrated Moving Average) y modelos basados en redes neuronales artificiales (RNA), a fin de proporcionar información útil a las instituciones públicas y privadas del sector turístico para la correcta planificación de las infraestructuras turísticas necesarias y para el conocimiento de las contribuciones relativas del turismo en la producción, ingresos y empleo en las Islas Baleares. Tal y como han mostrado estudios anteriores realizados por el equipo de este proyecto (Álvarez-Díaz et al., 2009; Álvarez-Díaz y Rosselló-Nadal, 2010 y 2011; Haldrup et al., 2007 y 2011; Molinet et al., 2015; Montañó et al., 2004, 2011 y 2013; Palmer et al., 2005, 2006 y 2008; Pons y Sansó, 2005; Rosselló et al., 2004) pero también otros trabajos descritos en el apartado siguiente, la literatura no ha sido concluyente a la hora de determinar la mejor estrategia para la predicción de series turísticas, por lo que a través de este trabajo se pretende además aportar nueva evidencia al debate.

En base a la experiencia del equipo investigador en la aplicación de estas metodologías en la previsión de series temporales sobre turismo, se pretenden diseñar modelos capaces de realizar previsiones con un horizonte de previsión de un año de anticipación con errores

inferiores al 10%, especialmente en temporada alta. Estos resultados proporcionarían una información valiosa con el objeto de poder anticipar una correcta planificación de nuestra oferta y conocer el futuro impacto de la demanda turística en nuestra economía. Por su parte, estas previsiones se podrían ir actualizando mes a mes y podrían estar implementadas en una aplicación informática accesible para todos los agentes relacionados con la industria turística.

El trabajo procede como sigue. En el apartado 2, se revisa la literatura sobre turismo y previsión de demanda. En el apartado 3 se describen los fundamentos de las metodologías que se van a utilizar. A continuación, en el apartado 4 se presentan y analizan las series temporales sobre las que desarrollan los modelos de predicción. En el apartado 5 se describe el ejercicio para la comparación de las metodologías que conduce a la selección de la metodología en función del horizonte y la serie a predecir. En el apartado 6 se presentan los resultados de las predicciones para el último cuatrimestre de 2015 y los dos primeros del 2016. Finalmente en el apartado 7 se presentan las principales conclusiones.

2. ANTECEDENTES Y REVISIÓN DE LA BIBLIOGRAFÍA

Debido a la naturaleza dinámica de la industria turística, la necesidad de diseñar previsiones precisas se ha convertido en una cuestión crucial (Chandra y Menezes, 2001). Así, tanto los investigadores como los profesionales y los responsables políticos han reconocido abiertamente la necesidad de disponer de previsiones ajustadas en el ámbito del turismo (Sheldon y Var, 1985). En el caso de la demanda turística, unas previsiones precisas ayudarían a directores e inversores a tomar decisiones operativas, tácticas y estratégicas, ejemplos de los cuales son la programación de la oferta y del volumen de personal, la preparación de folletos turísticos y de las inversiones hoteleras, respectivamente. Del mismo modo, los órganos de gobierno necesitan previsiones de la demanda del turismo para planificar las infraestructuras turísticas necesarias, tales como la planificación de espacios en el destino y el desarrollo del transporte, entre otros. Por otra parte, la previsión del gasto turístico es de valor en la determinación de las contribuciones relativas del turismo en la producción, los ingresos y el

empleo en los destinos turísticos (Bull, 1995). Por lo tanto, la gran cantidad de literatura académica que se ha generado en este área no es sorprendente (Morley, 2000).

A pesar del consenso sobre la necesidad de elaborar previsiones precisas y el reconocimiento de sus correspondientes beneficios, no hay un modelo que se destaque en términos de precisión en las previsiones (Law y Au, 1999). En este sentido, uno de los procedimientos más utilizados en la previsión de series temporales es la metodología de Box-Jenkins (Box & Jenkins, 1976), que se basa en un tipo especial de modelo estadístico lineal conocido como ARIMA (Autoregressive Integrated Moving Average). Un inconveniente que padece este procedimiento estadístico de series temporales es que se debe especificar formalmente el modelo y se debe asumir una distribución de probabilidad de los datos (Hansen, McDonald y Nelson, 1999).

En los últimos años, el estudio de las redes neuronales artificiales (RNA) ha despertado un gran interés en campos tan diversos como biología, psicología, medicina, economía, matemáticas, estadísticas e informática. La razón de este interés es que las RNA son aproximadores universales de funciones capaces de ajustar cualquier función lineal o no lineal (Cybenko, 1989). Debido a su flexibilidad como aproximadores de funciones, las RNA son métodos de gran alcance en tareas que implican la clasificación de patrones, la estimación de las variables y la previsión de series temporales (Kaastra y Boyd, 1996). En este último caso, las redes neuronales ofrecen varias ventajas potenciales sobre los métodos alternativos –principalmente los modelos de series temporales ARIMA-- al ser capaces de manejar problemas con datos no lineales y que no siguen una distribución normal (Hansen, McDonald y Nelson, 1999). La primera ventaja es que las RNA son muy versátiles y no requieren una especificación formal del modelo ni la aceptación de una distribución de probabilidad determinada para los datos. En cuanto a la segunda ventaja, Masters (1995) ha demostrado que las RNA son capaces de tolerar la presencia de componentes caóticos mejor que la mayoría de los métodos alternativos. Esta capacidad es particularmente importante, ya que muchas series temporales relevantes poseen componentes caóticos significativos.

Las RNA se han aplicado en los diversos campos mencionados anteriormente y también, de forma pionera, en el campo del análisis de los datos turísticos (Palmer, Montañó y Sesé, 2006). En este sentido, los modelos de redes neuronales se han aplicado recientemente como una técnica estadística en los principales campos de investigación del turismo, tales como la demanda y la previsión del comportamiento del consumidor (Wang, 2004), la segmentación del mercado y análisis de posicionamiento (Bloom, 2005).

Estos estudios indican un creciente interés en el uso de las RNA para representar las actividades cognitivas complejas y no lineales que se estudian a menudo en el campo de la gestión del turismo (Morley, 2000). Sin embargo, debido a la reciente introducción de las redes neuronales en el campo del turismo, el número de aplicaciones de las RNA en la previsión de series temporales en el campo del turismo sigue siendo muy limitado en comparación a la aplicación de otros métodos estadísticos (Uysal y El Roubier, 1999).

El presente proyecto pretende comprobar si los buenos resultados obtenidos por las RNA en estudios anteriores son extensibles a los que obtendremos en su aplicación a los datos turísticos de las Islas Baleares, ya sea directamente o bien a través de la combinación de predicciones con otros métodos.

3. DESCRIPCIÓN DE LOS MODELOS DE ANÁLISIS DE SERIES TEMPORALES

Para la previsión de series temporales sobre turismo se aplicarán modelos estadísticos ARIMA (Autoregressive Integrated Moving Average), modelos basados en redes neuronales artificiales (RNA) así como la posibilidad de combinar las predicciones de ambas metodologías. Estos modelos parten del supuesto que una serie temporal es la realización de un proceso estocástico. A efectos prácticos, esto significa que las observaciones de la serie temporal son fruto de un modelo estadístico subyacente que, aunque desconocido, puede ser adecuadamente aproximado y estimado y, posteriormente, usado para realizar predicciones.

3.1. Modelos ARIMA

Los Modelos ARIMA fueron desarrollados por Box y Jenkins (1976) y constituyen la técnica más conocida hoy en día para el análisis de series temporales. El punto de partida es el Teorema de Representación de Wold, que demuestra que cualquier proceso estocástico estacionario admite una representación en forma de proceso de Medias Móviles (MA) de orden infinito. A su vez, dicho modelo MA de orden infinito puede ser aproximado por un modelo lineal Autorregresivo (AR) y de Medias Móviles, ambos de orden finito. En definitiva, todo proceso estocástico estacionario puede ser aproximado por un modelo lineal ARMA de órdenes finitos.

No obstante, muchas series temporales económicas, entre las que se encuentran las ligadas a la actividad turística, se caracterizan por presentar tendencias a largo plazo así como una amplificación a largo plazo de los patrones estacionales. Véase, por ejemplo, la Figura 3 unas páginas más abajo. Esta inestabilidad tanto de la media como de las oscilaciones estacionales a largo plazo es incompatible con un proceso estocástico estacionario, el cual exige que tanto la media como la varianza sean constantes a lo largo del tiempo.

Por tanto, antes de poder aplicar la metodología de los modelos ARMA debe procederse a transformar las series temporales de modo que las series transformadas muestren un patrón de media y de varianza estable a largo plazo. Este proceso de transformación de las series, conocido como homogeneización, se consigue mediante: 1) la transformación de Box-Cox, tendente a estabilizar la varianza; y 2) diferenciación de las series, esto es, transformar las variables a partir de trabajar con cambios sobre períodos anteriores.

La transformación Box-Cox más usual para variables económicas es la logarítmica, puesto que ésta, por un lado, transforma la interacción multiplicativa entre tendencia y estacionalidad en una relación aditiva en la cual las oscilaciones estacionales tienden a ser de amplitud estable a lo largo del tiempo; y, por otro lado, la posterior aplicación de diferenciaciones, da lugar a una

serie transformada que es fácilmente interpretable como tasa de variación.¹ El número de diferenciaciones que se requieren para transformar una variable en estacionaria se conoce como orden de integración.

Puesto que las series temporales turísticas muestran un importante patrón estacional, los modelos ARMA a utilizar deberán capturar también dicho comportamiento. Para ello, los modelos ARMA usuales se extienden ampliándose con estructuras ARMA adicionales que capturan la estacionalidad. En síntesis, los modelos que se van a utilizar son los ARIMA(p,d,p) (P,D,Q)_s y pueden formularse mediante la siguiente expresión:

$$\Phi_p(L^s)\phi_p(L)\Delta_s^D\Delta^d y_t = \delta + \Theta_Q(L^s)\theta_q(L)\varepsilon_t, \quad [1]$$

donde y_t es la variable estabilizada mediante la transformación Box-Cox, en nuestro caso, transformada en logaritmos; L es el operador retardo, esto es, $Ly_t = y_{t-1}$ y $L^j y_t = y_{t-j}$; $\Delta = (1-L)$ es el operador diferencia regular; d es el orden de integración regular; s es el número de estaciones por año, de modo que para datos mensuales $s=12$; $\Delta_s = (1-L^s)$ es el operador diferencia estacional; D es el orden de integración estacional; $\phi_p(L) = 1 - \phi_1 L - \dots - \phi_p L^p$ es el polinomio autorregresivo regular de orden p ; $\Phi_p(L^s) = 1 - \Phi_1 L^s - \dots - \Phi_p L^{sP}$ es el polinomio autorregresivo estacional de orden P ; $\theta_q(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ es el polinomio media móvil regular de orden q ; $\Theta_Q(L^s) = 1 - \Theta_1 L^s - \dots - \Theta_Q L^{sQ}$ es el polinomio media móvil estacional de orden Q ; δ es la constante del modelo; y ε_t es un proceso de ruido blanco, esto es, de media nula, varianza constante e incorrelacionado.²

Un modelo muy usual en el modelado de series temporales mensuales, el cual ya aparece en el propio libro de Box-Jenkins de 1976, al aplicar su metodología a una serie sobre número de

¹ Por ejemplo, para la serie temporal y_t , para $t=1, \dots, T$, la aplicación de la transformación logarítmica junto con una diferenciación da lugar a una nueva serie temporal de tasas de variación naturales, las cuales son aproximadamente iguales a las tasas de variación, $g_t = y_t/y_{t-1} - 1$. En efecto, para valores bajos de g_t se verifica:

$$\ln y_t - \ln y_{t-1} = \ln(y_t/y_{t-1}) = \ln(1+g_t) \approx g_t$$

² Los polinomios autorregresivos y de medias móviles deben cumplir ciertos requisitos de tipo técnico. En primer lugar, deben de tener todas la raíces fuera del círculo de radio unidad. Ello garantiza que la variable diferenciada sea estacionaria. En segundo lugar, las medias móviles no pueden tener ninguna raíz común con los autorregresivos, puesto que en este caso el modelo puede ser simplificado y se viola el principio de parsimonia.

pasajeros, es el conocido como modelo de líneas aéreas y viene dado por la especificación ARIMA(0,1,1) (0,1,1)₁₂:

$$(1-L^2)(1-L)y_t = \delta + (1-\Theta_1 L^2)(1-\theta_1 L)\varepsilon_t. \quad [2]$$

Nótese que si y_t está en logaritmos, entonces $(1-L^2)y_t$ es aproximadamente la tasa interanual de crecimiento y $(1-L)(1-L^2)y_t$ es el cambio de dicha tasa sobre el período anterior.

Las cuatro etapas que intervienen en la construcción de un modelo ARIMA son: 1) identificación, 2) estimación, 3) validación y 4) predicción o uso del modelo. La fase de identificación trata de determinar los órdenes de integración, autorregresivos y de medias móviles tanto de la parte regular como de la estacional, así como de la transformación Box-Cox más adecuada. Una vez determinados los órdenes de integración, por ejemplo mediante contrastes de raíz unitaria, los órdenes AR y MA se pueden determinar a partir de las funciones de autocorrelación observadas (ACF) y las funciones de autocorrelación parciales (PACF). Puesto que dichas funciones son estimadas y están sujetas a error de estimación, es posible que en esta fase se identifique más de un posible modelo. La fase de estimación calcula los parámetros que finalmente se han especificado en el (los) modelo(s). La fase de validación comprueba si el (los) modelo(s) seleccionado(s) cumple(n) todos los requisitos, de modo que pueda ser utilizado correctamente en la fase de predicción, y selecciona, a partir de diversos criterios, el modelo más adecuado si hay más de un posible candidato. Finalmente, el modelo puede ser usado para realizar predicciones.

3.2. Modelos RNA

Las RNA han sido aplicadas en los últimos años de forma satisfactoria en la clasificación de patrones, en la estimación de variables continuas y en la previsión de series temporales, obteniendo en muchos casos resultados superiores a los modelos estadísticos clásicos. Las RNA son modelos derivados de la Inteligencia Artificial y se pueden definir como sistemas de procesamiento de la información cuya estructura y funcionamiento se inspira en las redes

neuronales biológicas. Tienen tres características fundamentales: procesamiento en paralelo, memoria distribuida y capacidad de adaptación. Estas propiedades les proporcionan una serie de ventajas en comparación con otros sistemas de procesamiento, tales como la robustez y la tolerancia a errores y ruido.

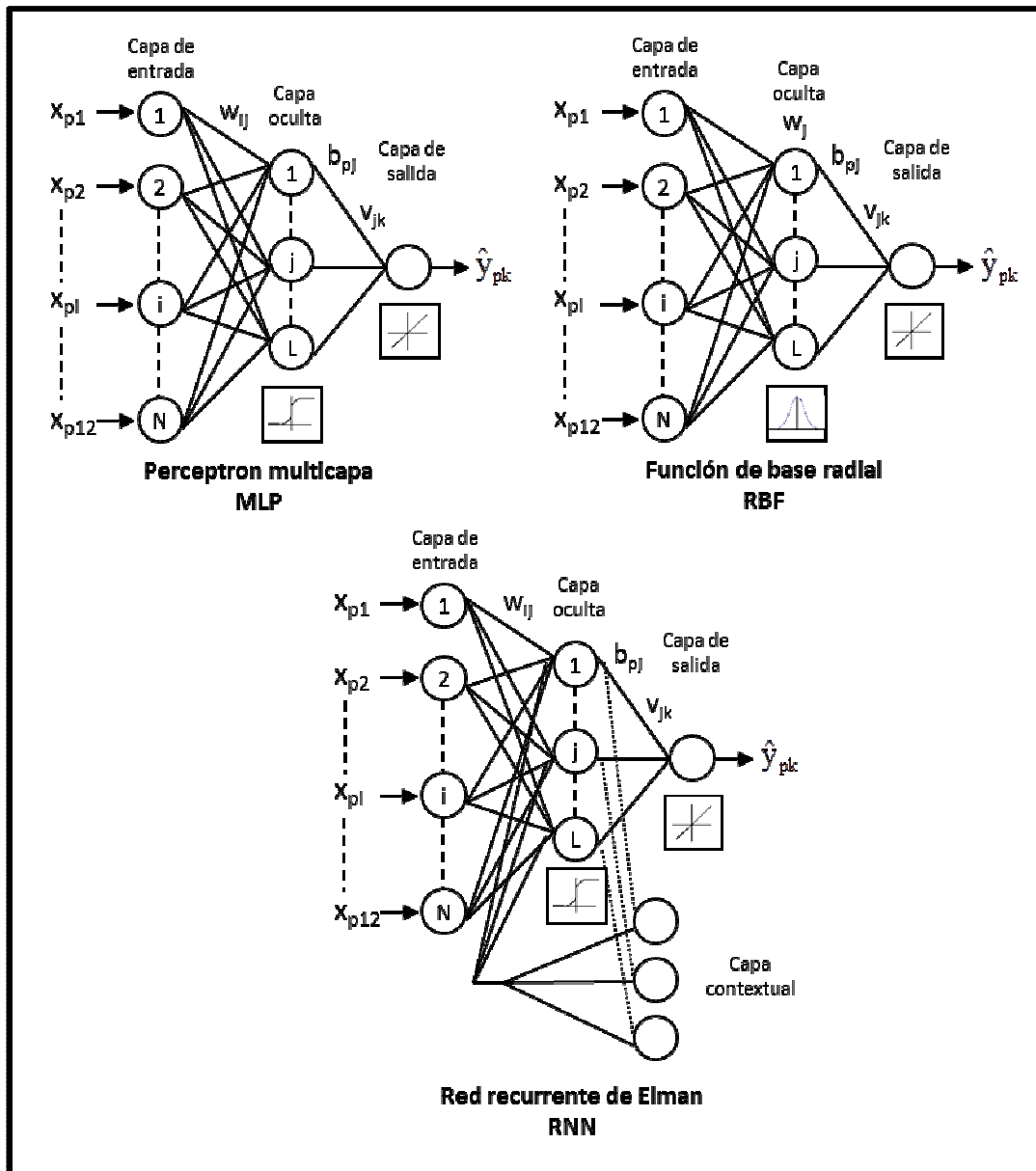


Figura 1. Modelos de red neuronal artificial analizados

En general, una RNA se compone de un gran número de elementos de procesamiento simples conocidos como nodos o neuronas, organizados en capas. Cada neurona está conectada a otras

neuronas por enlaces de comunicación, cada uno de los cuales tiene un valor numérico asociado conocido como peso. Los pesos contienen el conocimiento o la información que la red neuronal posee acerca de un problema específico.

El modelado de una serie temporal univariada mediante RNA se realiza generalmente utilizando un número determinado de términos atrasados de la serie como entrada y las previsiones como salida (Bishop, 1995). Cuando la serie registrada es de tipo mensual, el modelo habitualmente utilizado consiste en predecir el valor de cada mes a partir de los 12 o 24 meses anteriores. No obstante, este tipo de modelos es muy limitado ya que, entre otras cuestiones, determina un horizonte de previsión de un mes. En el presente trabajo se pretende crear modelos de red neuronal con un horizonte de previsión a un año vista. Para ello, cada punto temporal de tipo mensual será predicho como valor de salida a partir de la siguiente información de entrada al modelo de red, de acuerdo con la descripción realizada en Molinet et al. (2015):

- Mes a predecir: representado mediante 12 variables ficticias binarias.
- Temporada a la que pertenece el mes a predecir: tres temporadas (alta, media y baja) representado mediante 3 variables ficticias binarias.
- Valor del mes a predecir tres años atrás y los valores de los dos meses consecutivos.
- Valor del mes a predecir dos años atrás y los valores de los dos meses consecutivos.
- Valor del mes a predecir un año atrás y los valores de los dos meses consecutivos.

Esta estructura de presentación de los datos de entrada y salida, 24 variables o neuronas de entrada y 1 variable o neurona de salida, ha demostrado tener un buen rendimiento en la predicción del número de pasajeros tanto en destinos consolidados como en destinos no consolidados con un horizonte de previsión de un año (Molinet et al., 2015).

Existen diversas arquitecturas y algoritmos de aprendizaje para los modelos de red neuronal. Los modelos más ampliamente utilizados para su aplicación a la previsión de series temporales y que serán empleados en el presente proyecto con la estructura de datos de entrada y salida descritos, son (ver Figura 1): Perceptrón Multicapa, Funciones de Base Radial y Redes Recurrentes.

3.2.1. Perceptrón Multicapa (MLP)

Un modelo MLP está compuesto por una capa de N neuronas de entrada, una capa de M neuronas de salida y una o más capas ocultas; aunque se ha demostrado que para la mayoría de problemas bastará con una sola capa de L neuronas ocultas (Hornik et al. , 1989). En este tipo de arquitectura, las conexiones entre neuronas son siempre hacia delante, es decir, las conexiones van desde las neuronas de una determinada capa hacia las neuronas de la siguiente capa.

La representación matemática de la función que aplican las neuronas ocultas para obtener un valor de salida b_{pj} , ante la presentación de un vector o patrón de entrada $X_p: x_{p1}, \dots, x_{pi}, \dots, x_{pN}$, viene definida por:

$$b_{pj} = f_L \left(\theta_j + \sum_{i=1}^N w_{ij} \cdot x_{pi} \right) \quad [3]$$

donde f_L es la función de activación de las L neuronas ocultas, ϑ_j es el umbral de la neurona oculta j , w_{ij} es el peso de conexión entre la neurona de entrada i y la neurona oculta j y, finalmente, x_{pi} es la señal de entrada que recibe la neurona de entrada i para el patrón p .

Por su parte, la salida de las neuronas de salida se obtiene de forma similar con respecto a las neuronas de la capa oculta, mediante:

$$\hat{y}_{pk} = f_M \left(\theta_k + \sum_{j=1}^L v_{jk} \cdot b_{pj} \right) \quad [4]$$

donde \hat{y}_{pk} es la señal de salida que proporciona la neurona de salida k para el patrón p , f_M es la función de activación de las M neuronas de salida, ϑ_k es el umbral de la neurona de salida k y, finalmente, v_{jk} es el peso de conexión entre la neurona oculta j y la neurona de salida k .

De forma general, se utiliza una función sigmoïdal en las neuronas de la capa oculta para dotar a la red neuronal de capacidad para aprender posibles funciones no lineales, mientras que la función lineal es utilizada en la neurona de salida en caso de la estimación de una variable continua.

El entrenamiento de las redes MLP es de tipo supervisado y se puede realizar mediante la aplicación del clásico algoritmo de gradiente decreciente (Rumelhart et al., 1986) o mediante algún algoritmo de optimización no lineal que, como en el caso del algoritmo de gradientes conjugados (Battiti, 1992), permite acelerar considerablemente la velocidad de convergencia de los pesos con respecto al algoritmo de gradiente decreciente.

3.2.2. Funciones de Base Radial

Las RBF (Broomhead y Lowe, 1988) están compuestas de tres capas al igual que la red MLP. La particularidad de las RBF reside en que las neuronas ocultas operan en base a la distancia euclídea que separa el vector de entrada X_p respecto al vector de pesos W_j que cada una almacena (denominado centroide), cantidad a la que aplican una función radial con forma gaussiana, de forma similar a las funciones kernel en el modelo de regresión kernel (Bishop, 1995).

La representación matemática de la función radial que aplican las neuronas ocultas sobre el vector de entrada X_p para obtener un valor de salida b_{pj} viene definida por:

$$b_{pj} = \exp \left[\frac{-\sum_{i=1}^N (x_{pi} - w_{ij})^2}{2\sigma^2} \right] \quad [5]$$

Si el vector de entrada X_p coincide con el centroide W_j de la neurona j , ésta responde con máxima salida (la unidad). Es decir, cuando el vector de entrada está situado en una región

próxima al centroide de una neurona, ésta se activa, indicando que reconoce el patrón de entrada; si el patrón de entrada es muy diferente del centroide, la respuesta tenderá a cero.

El parámetro de normalización σ (o factor de escala) mide la anchura de la gaussiana, y equivaldría al radio de influencia de la neurona en el espacio de las entradas; a mayor σ la región que la neurona domina en torno al centroide es más amplia.

La salida de las neuronas de salida se obtiene como una combinación lineal de los valores de activación de las neuronas ocultas ponderados por los pesos que conectan ambas capas de igual forma que la expresión matemática asociada a una red ADALINE (Widrow y Hoff, 1960):

$$\hat{y}_{pk} = \theta_k + \sum_{j=1}^L v_{jk} \cdot b_{pj} \quad [6]$$

Como la red MLP, las RBF permiten realizar con relativa facilidad modelados de sistemas no lineales arbitrarios y también constituyen aproximadores universales de funciones (Hartman et al., 1990), con la particularidad de que el tiempo requerido para su entrenamiento suele ser mucho más reducido. Esto es debido en gran medida a que las redes RBF constituyen un modelo de red híbrido, por incorporar aprendizaje supervisado o no supervisado en dos fases diferentes. En una primera fase, los vectores de pesos o centroides asociados a las neuronas ocultas se obtiene mediante un aprendizaje no supervisado a través del algoritmo k-medias. En una segunda fase, los pesos de conexión entre las neuronas ocultas y las de salida se obtienen mediante un aprendizaje supervisado a través de la regla delta de Widrow-Hoff (1960).

3.2.3. Red Recurrente

Las redes recurrentes son de especial utilidad en aquellas situaciones en las que se desea representar las relaciones temporales que se puedan establecer entre las entradas y salidas de la red neuronal (Elman, 1990). En este tipo de redes una capa de neuronas posee conexiones recurrentes, es decir, las salidas de las neuronas son almacenadas temporalmente y enviadas

posteriormente como señales de entrada a esas mismas neuronas o a otras neuronas de la red neuronal. Este proceso se representa en la Figura 2.

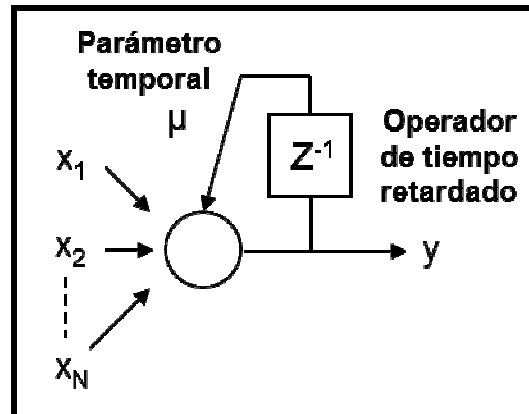


Figura 2. Neurona artificial con conexión recurrente

El operador de tiempo retardado Z^{-1} permite almacenar las señales de salida y obtenidas en los n momentos anteriores de la neurona j . De este modo, se genera una memoria a corto plazo de los valores de activación anteriores generados por la neurona. Por su parte, el parámetro temporal μ determina el peso o traza de memoria de las señales de salida y anteriores de la neurona j . Así, la salida de la neurona j es función de las señales de entrada x_i de las neuronas N de la capa inmediatamente anterior y de las señales de salida y anteriores de esa misma neurona j ponderadas por el parámetro temporal. La huella o traza de memoria de las señales de salida anteriores decrece de forma exponencial:

$$y(n) = \sum_{i=1}^N x_i + \mu y(n-1) + \mu^2 y(n-2) + \mu^3 y(n-3) + \dots + \mu^z y(n-z) \quad [7]$$

donde $0 < \mu < 1$

Las redes recurrentes poseen al igual que los anteriores modelos, una arquitectura similar a un modelo MLP donde una de las capas de neuronas tienen conexiones recurrentes. En el presente estudio se utilizó la red de Elman (Elman, 1990) en la cual, las neuronas de la capa oculta presentan conexiones recurrentes. De este modo, las neuronas de esta capa reciben señales de sí mismas procedentes de momentos anteriores y señales de las neuronas de la capa

de entrada. Tal como pusieron de manifiesto Jehee y Lee (1996), este tipo de recurrencia es de especial utilidad para su aplicación en la previsión de series temporales. Las conexiones recurrentes que almacenan las señales atrasadas de las neuronas se suelen representar mediante una capa de neuronas especial denominada capa de neuronas contextuales.

Respecto al entrenamiento de este tipo de redes neuronales, las conexiones que aparecen en la figura con línea continua son modificadas siguiendo un aprendizaje supervisado al igual que en caso del modelo MLP, mientras que las conexiones que aparecen con línea discontinua (conexiones hacia la capa de neuronas contextuales) son fijadas a un valor constante igual a 1 y no son susceptibles de modificación.

3.3. Combinación de predicciones

Como complemento a las dos metodologías anteriores este trabajo considerará la posibilidad de combinación de predicciones (COMB) obtenidas a partir de los dos métodos anteriores. El principio básico de la combinación de predicciones se sustenta en la aceptación de que es muy difícil encontrar una supremacía entre diferentes métodos (en este caso ARIMA y RNA) y que la combinación de los mismos puede dar lugar a una predicción incluso más ajustada que la obtenida individualmente por cada uno de los métodos por separado (Shen et al, 2011; Wong et al, 2007).

En este trabajo se considera que el valor de la predicción es una media ponderada de las predicciones obtenidas a través de métodos particulares. En concreto, la ponderación para establecer el peso de cada una de los métodos en la predicción final toma como referencia la inversa de la varianza de los errores en términos absolutos obtenidos en el periodo de evaluación de los métodos. De esta manera se consigue dar más peso al método que presenta un menor error de predicción. Los detalles de la aplicación del método se describen en el apartado 5.

4. DATOS

La selección de las series de demanda turística referida a las Islas Baleares se basa en los datos publicados por la Agència de Turisme de les Illes Balears (ATB, 2015). Así, en primer lugar, La ATB recoge periódicamente la llegada mensual de turistas a las Islas Baleares con desagregación por mercado de origen y por islas desde enero de 1999, enero de 2009 o junio de 2010 (según el caso) y estas constituyen las variables de demanda turística principales a analizar. Sin embargo, dado que la llegada de turistas puede tener un efecto diferenciado según la capacidad y comportamiento económico de los turistas junto con las variables de llegadas de turistas, en segundo lugar se analizará la variable de gasto turístico total de los turistas extranjeros que visitan Baleares, serie procedente de Egatur. Esta variable, aunque puede resultar incluso más interesante desde el punto de vista económico en comparación al número de turistas, presenta tres limitaciones prácticas: se dispone de una serie menos larga (desde enero de 2004), considera únicamente los turistas internacionales y estaría distorsionada por el efecto precio (dado que se ofrece a precios corrientes). En tercer lugar y en un intento, precisamente, de recoger las variaciones de precios del producto turístico, el presente estudio analizará finalmente el Índice de Precios Hoteleros (IPH). Es importante remarcar cómo el IPH, a diferencia de un índice de precios al consumo, trata de que recoger los precios percibidos por los empresarios hoteleros (no por los consumidores - turistas), considerando todos sus clientes (tarifa normal, fin de semana, touroperador, grupos y empresas). En este caso, el IPH se encuentra elaborado originariamente por el INE a través de la Encuesta de Ocupación hotelera.

En cuanto a las series de llegadas de turistas y su posible desagregación por mercado de origen cabe tener en cuenta que para el período de referencia 2011-2014, el 83% de las mismas procedieron de mercados extranjeros mientras que el 17% se refirieron al mercado nacional (o doméstico). Además, cabe destacar como, sobre las llegadas totales de turistas, un 30% procedieron de Alemania mientras que un 25% del Reino Unido, evidenciándose una importante concentración de mercado en estas dos nacionalidades. De esta manera, Italia que se sitúa como cuarta nacionalidad en importancia relativa, concentra un 5% de las llegadas a las Baleares, emplazándose así a una distancia significativa del resto de mercados (Alemán,

Británico, Nacional). En consecuencia, junto a la llegada total de turistas, en este estudio se ha optado por considerar los cuatro siguientes subgrupos de turistas: Internacionales, Nacionales, Británicos y Alemanes.

Por lo que se refiere a la desagregación por Islas, se ha optado por seguir la misma clasificación tradicional distinguiendo Mallorca, Menorca e Ibiza-Formentera. Aunque sería posible plantear una sub-desagregación por nacionalidades en el caso de Mallorca (ya que concentra el 72% de las llegadas turísticas), para el caso de Menorca (9%) y de Ibiza-Formentera (19%) existe un problema de representatividad. Además dicha desagregación sólo tendría sentido si hubieran razones de peso para pensar que algunas nacionalidades tienen un comportamiento diferenciado en alguna isla en particular, lo que inicialmente ha sido descartado.

En cuanto a la longitud de las series temporales, cabe señalar que la información disponible de llegadas de turistas para el caso del turismo nacional y la desagregación por islas imposibilita la correcta estimación de cualquier tipo de modelo estadístico tipo ARIMA o Redes Neuronales. La solución adoptado ha sido la de alargar las series temporales seleccionadas a partir del movimiento de pasajeros por vía aérea a los aeropuertos de las Islas Baleares. En este sentido cabe señalar que durante el periodo en el que coinciden ambas series, el coeficiente de correlación entre los pares de series seleccionadas,³ supera en todos los casos el 0,97, por lo que la posibilidad de ajustar el comportamiento de la serie de pasajeros por vía aérea con la de llegada de turistas parece adecuada.

Analíticamente, la incorporación del comportamiento de las series de llegada de pasajeros a las series de demanda turística de referencia se realiza de la siguiente manera. Sea n_i el número de turistas del segmento i (i =Total, España, Mallorca, Menorca, Ibiza-Formentera) durante el mes t (desde enero 2009 o junio 2010 en adelante, según disponibilidad) y sea p_i el número de pasajeros llegados por vía aérea del segmento i durante el mes t (desde enero 1999). Si se define:

³ Esto es, por ejemplo, Número de pasajeros llegados por vía aérea a Mallorca y Número de turistas llegados a Mallorca.

$$x_t^i = \frac{X_t^i}{X_{t-1}^i} - 1 \quad [8]$$

entonces es posible proyectar hacia adelante la variable mediante la relación:

$$Y_t^i = (1 + x_t^i)Y_{t-1}^i \quad [9]$$

Desde enero 2009 o junio 2010 en adelante, según disponibilidad. Para la retroproyección hacia atrás, se puede utilizar la relación:

$$Y_{t-1}^i = \frac{Y_t^i}{1 + x_t^i} \quad [10]$$

Para el resto de observaciones (hasta enero 1999). De esta manera representa el número de turistas del segmento i reconstruidos a partir del crecimiento intermensual en la llegada de pasajeros a los aeropuertos a las Islas Baleares.

Para el caso del gasto turístico y el IPH, al no tener posibilidad de extender la serie, se utilizan los datos disponibles; esto es desde enero de 2004 en el caso del gasto y desde enero de 2001 para el IPH. La Tabla 1 resume las series sobre las que se desarrollan los modelos para la obtención de previsiones periódicas de demanda.

Tabla 1. Selección de las series para la construcción de modelos de predicción

| Series | Símbolo utilizado |
|---|-------------------|
| Llegadas de turistas a las Islas Baleares** | TOT |
| Llegadas de turistas Internacionales a las Islas Baleares | INT |
| Llegadas de turistas procedentes de UK a las Islas Baleares | UK |
| Llegadas de turistas procedentes de Alemania a las Islas Baleares | ALE |
| Llegadas de turistas procedentes de España a las Islas Baleares** | DOM |
| Llegadas de turistas Internacionales a Mallorca*** | MALL |
| Llegadas de turistas Internacionales a Menorca*** | MEN |
| Llegadas de turistas Internacionales a Ibiza-Formentera*** | EF |
| Gasto turístico | DESPESA |
| Índice de Precios Hoteleros | IPH |

Notas: ** Extensión de la serie con la serie de AENA de llegada de pasajeros (Totales / nacionales) desde enero de 1999 hasta enero de 2009. *** Extensión de la serie con la serie de AENA de llegada de pasajeros internacionales desde enero de 1999 hasta mayo de 2010.

La Figura 3 muestra la representación gráfica de las variables de demanda turística donde, tal y como era de esperar, destaca un fuerte comportamiento estacional. Una de las facultades de la metodología ARIMA es la capacidad de capturar el componente ciclo tendencia de las series temporales. En este sentido, en cierta medida, la predicción en el contexto de la metodología ARIMA lo que pretende es alargar este componente ciclo tendencia y basar su predicción en esa prolongación. Es por ello que en el siguiente apartado se centra la atención en la definición y el estudio del comportamiento cíclico tendencial de las series seleccionadas para llevar a cabo el ejercicio de predicción.

La utilización de datos mensuales y la presencia de una elevada estacionalidad (Figura 3) sugieren una primera posibilidad de trabajar con los datos de crecimiento interanuales. Sin embargo, a pesar de conseguir eliminar el componente estacional, el elevado valor del componente irregular dificulta el reconocimiento del componente ciclo-tendencia (Figura 4).

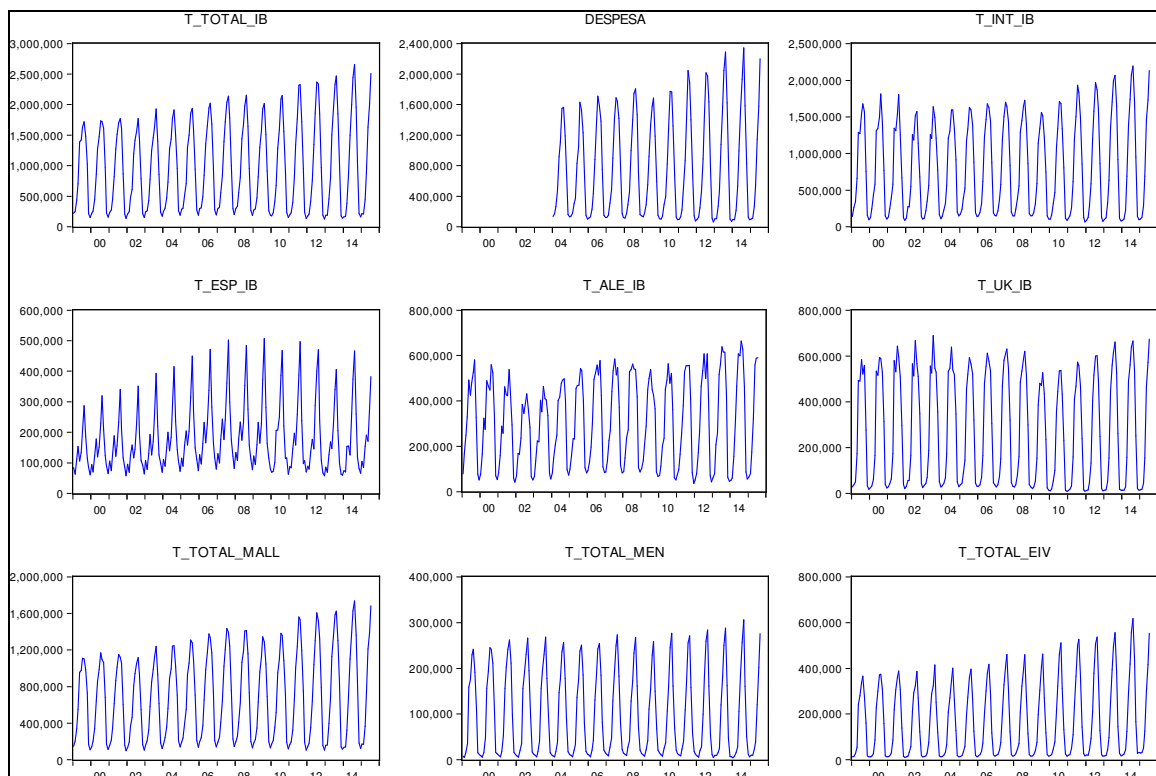


Figura 3. Series de demanda turística de referencia

Fuente: Elaboración propia a partir de los datos de la ATB

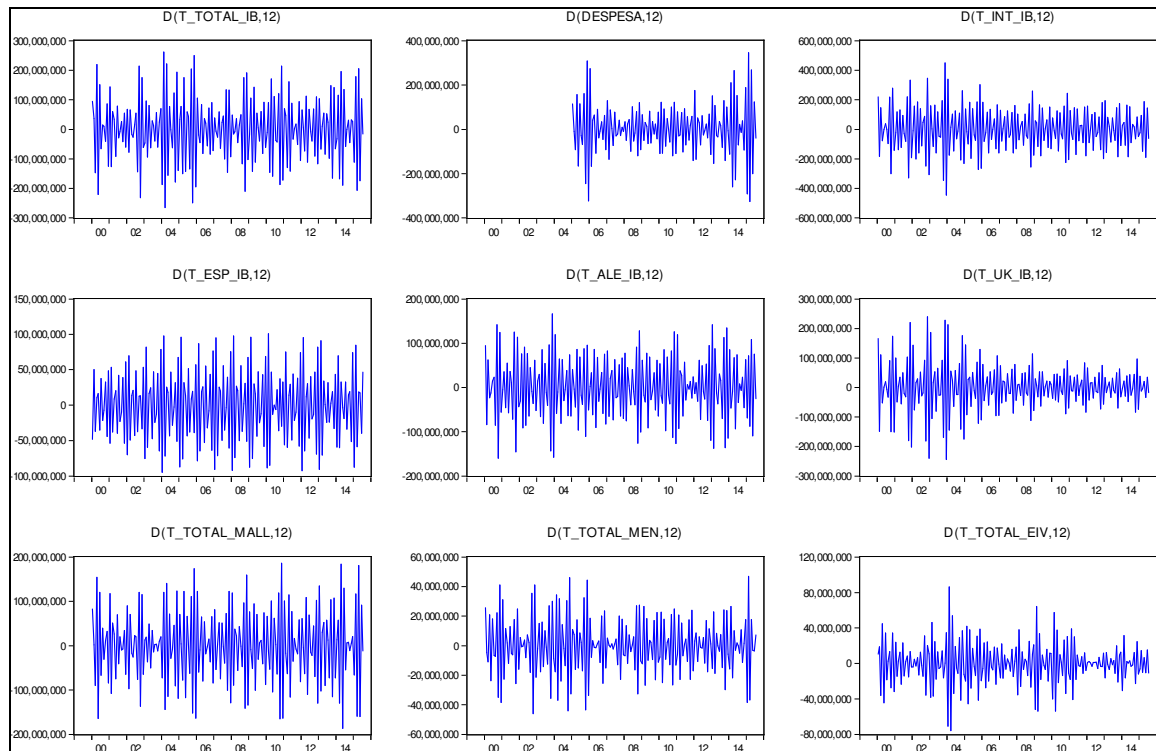


Figura 4. Incrementos interanuales de las variables de referencia

Fuente: Elaboración propia a partir de los datos de la ATB

Es por ello que se opta por el cálculo de la evolución subyacente (TUT), conocida también por $T_{12,12}$, propuesta por Espasa y Cancelo (1993) y utilizado en el marco de la evolución cíclica de las series de demanda turística de las Islas Baleares en Rosselló (2001). Analíticamente:

$$TUT_t^i = \frac{\sum_{k=0}^{11} Y_{t+k}^i}{\sum_{k=1}^{12} Y_{t-k}^i} - 1 \quad [11]$$

El cálculo de la evolución subyacente de las series de turistas totales (TOT), Internacionales (INT) y Españoles (ESP) y del gasto total de los turistas internacionales (DESPEA) permite mostrar (Figura 5) como el número total de turistas ha presentado un crecimiento generalizado con tasas positivas a excepción de dos periodos, 2001 - 2002 y desde mediados del 2007 hasta finales de 2009, coincidiendo así pues con los periodos de crisis económica internacional, y totalmente en sintonía con lo que acontecía con los movimientos de turistas internacionales a nivel mundial (UNWTO, 2015).

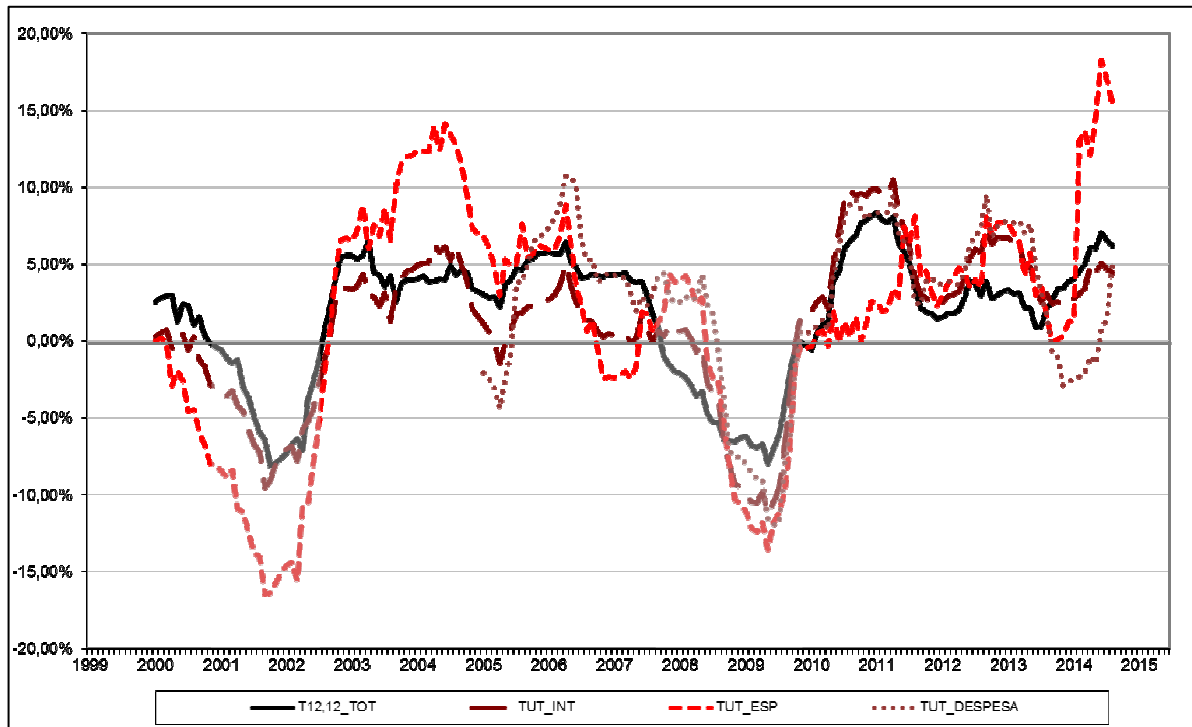


Figura 5. Evolución subyacente de los turistas totales, internacionales y nacionales llegados a las Islas Baleares y del gasto turístico total de los turistas internacionales en Baleares.

Cabe hacer notar cómo el ciclo tendencia de los turistas totales y el de los turistas internacionales coinciden en mostrar unos crecimientos moderados de la demanda turística, de entre el 0% y el 5%, durante el periodo entre las crisis y como, a partir del 2010, después de unos crecimientos notable con tasas cercanas al 10%, la llegada de turistas se ha mantenido en tasas positivas, aunque más moderadas.

El caso de los turistas nacionales (ESP) es especialmente significativo. Si bien se muestra en total sintonía a la hora de datar la crisis del 2001-2002, la recuperación y el crecimiento del mercado nacional destacan por su vigor durante los años posteriores a la primera crisis y hasta la llegada de la última crisis (2007). Es entonces cuando el mercado nacional, si bien no parece sufrir con tanta intensidad los inicios de la crisis, acaba sucumbiendo y alargando la duración de la misma hasta el año 2013 donde parece haber recuperado definitivamente las tasas de crecimiento positivo.

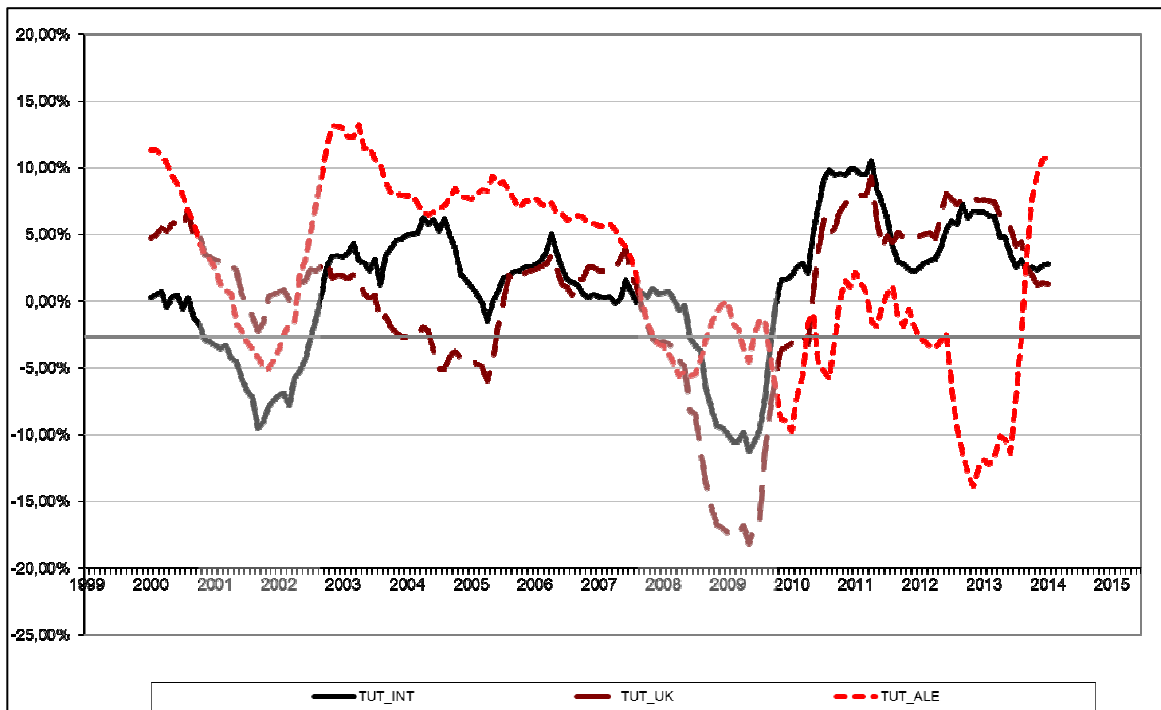


Figura 6. Evolución subyacente de los turistas internacionales, alemanes y británicos llegados a las Islas Baleares

Al analizar conjuntamente la evolución de los turistas internacionales (INT), alemanes (ALE) y británicos (UK) (Figura 6), se pone de manifiesto como, si bien ambas nacionalidades tienen un comportamiento muy parecido en los dos periodos de crisis descritos anteriormente, existen divergencias importantes en la evolución mostrada por alemanes y británicos durante el período 2003-2005.

De esta manera, mientras que el mercado alemán destaca por sus importantes crecimientos, el mercado británico se caracteriza por crecimientos negativos durante este mismo periodo. Por otra parte es remarcable que el inicio de la crisis que afecta a los mercados internacionales es especialmente más importante para el caso del Reino Unido, situación que podría explicarse por un tipo de cambio especialmente desfavorable entre la libra esterlina y el euro que caracterizó el inicio de la crisis económica en Europa. En cualquier caso, la evolución de la evolución subyacente de ambos mercados muestra como la recuperación y el retorno a la senda de crecimiento está caracterizando a ambos mercados durante los últimos años.

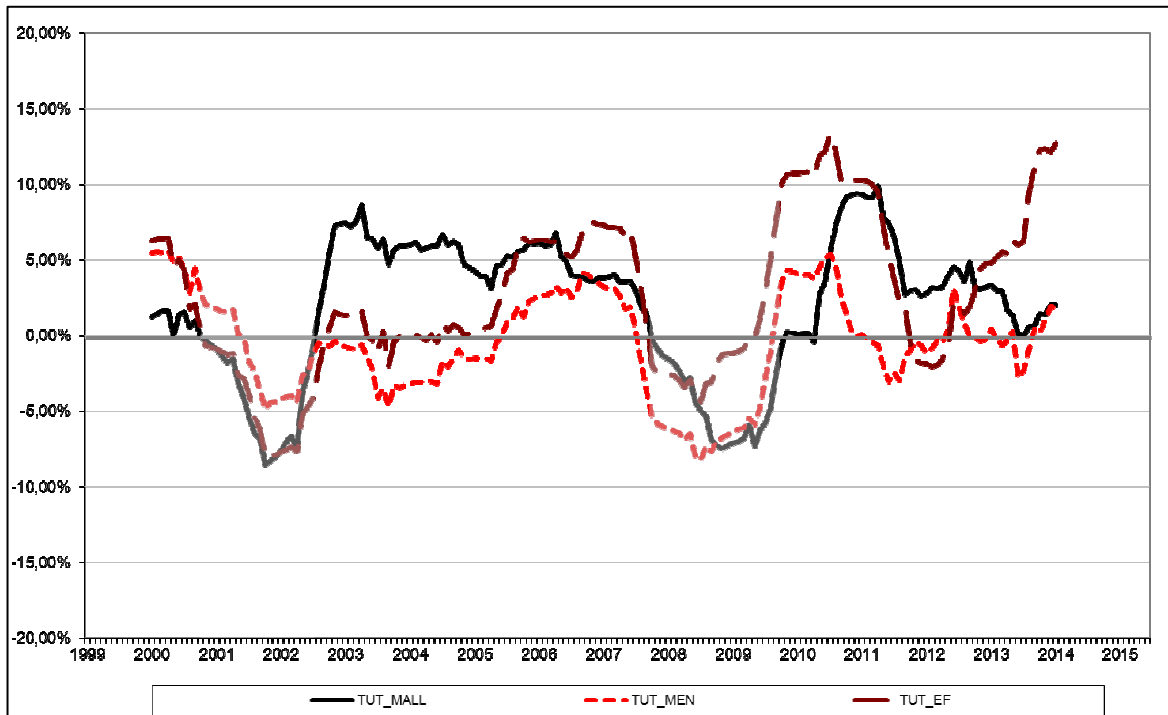


Figura 7. Evolución subyacente de los turistas llegados a Mallorca, Menorca e Ibiza-Formentera

La Figura 7 muestra como la evolución por islas se puede explicar en sintonía con la composición de mercados que suele caracterizar a cada una de ellas. De esta manera, Mallorca, más frecuentada por el mercado alemán, muestra un perfil similar al conjunto de los turistas internacionales de Baleares. Por su parte, Menorca, con un elevado número de turistas Británicos y Nacionales, ha sido la isla más perjudicada, en primer lugar por la caída del turismo británico a principios de siglo y posteriormente por la caída del turismo español durante los últimos años. Por su parte, Ibiza-Formentera, menos focalizada en el turismo español, si bien sufrió las consecuencias de las caídas de turistas británicos de principios de siglo, durante los últimos años se está beneficiando del mayor dinamismo del mercado internacional.

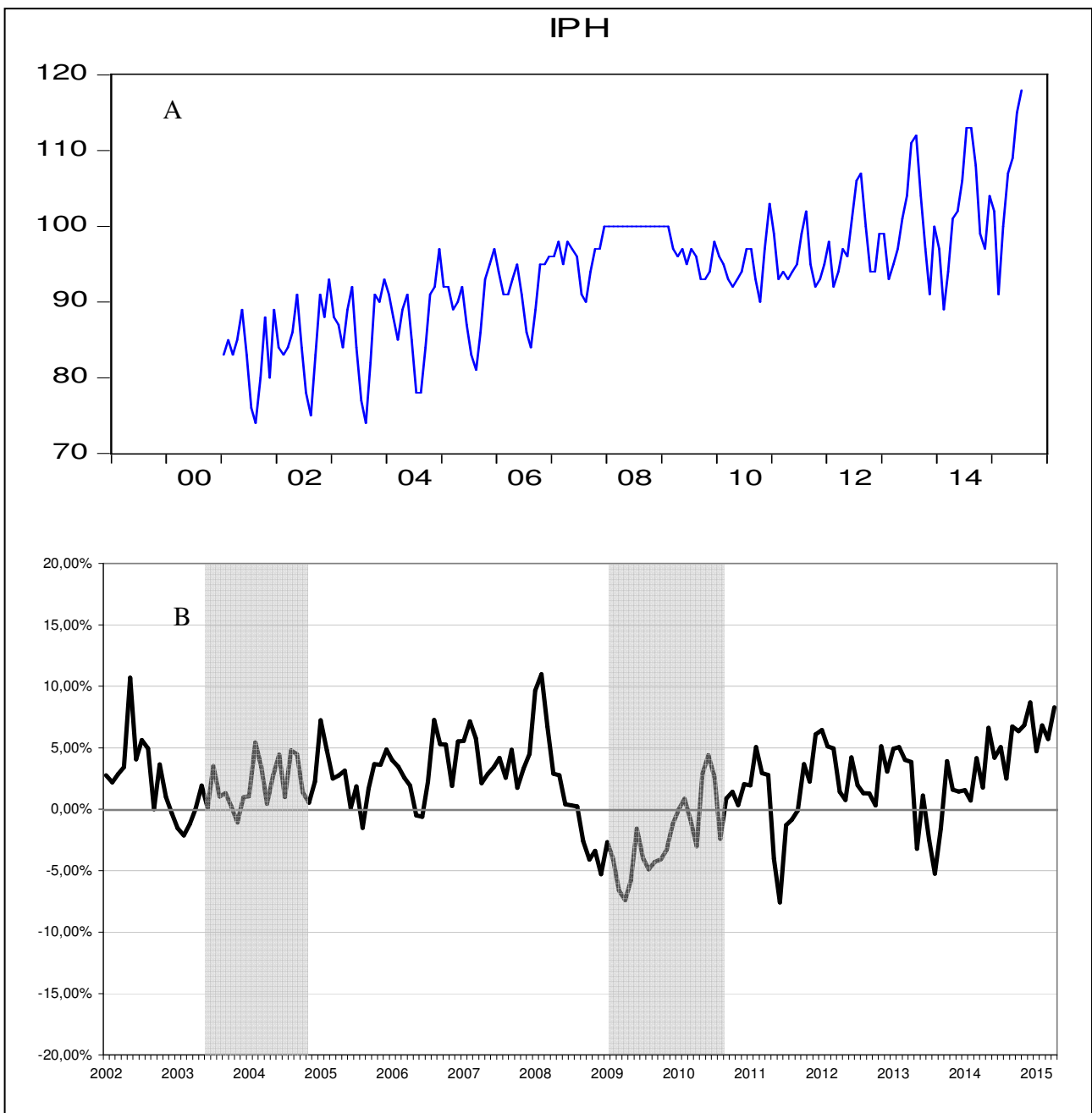


Figura 8. Evoluci3n del IPH (A) y su crecimiento interanual (B)

Por lo que se refiere al IPH, es destacable observar nuevamente una marcada estacionalidad (Figura 8). Sin embargo, en este caso, al calcular la inflaci3n interanual (diferencia en t3rminos porcentuales sobre el mes anterior del IPH) es posible ya distinguir cierto movimiento c3clico donde destaca la clara bajada de precios acontecida en el a3o 2009, circunstancia que no aconteci3 en el a3o 2004 .

5. PROCEDIMIENTO DE DISEÑO, EVALUACIÓN Y SELECCIÓN DE LOS MODELOS DE SERIES TEMPORALES

Teniendo en cuenta que el propósito del proyecto es obtener previsiones de turismo en las Islas Baleares a fin de proporcionar información útil a las instituciones públicas y privadas del sector para la correcta planificación de oferta de servicios e infraestructuras turísticas, la información que se toma por objetivo para proporcionar se resume en las Tablas 2 y 3.

Tabla 2. Ejemplo de tabla a elaborar en Diciembre 2015

| Series | Enero-Abril 2016 | | Mayo-Setiembre 2016 | | Octubre-Diciembre 2016 | | Total 2016 | |
|----------|------------------|-------------|---------------------|-------------|------------------------|-------------|------------|-------------|
| | Previsión | Incremento% | Previsión | Incremento% | Previsión | Incremento% | Previsión | Incremento% |
| TOT | | | | | | | | |
| INT | | | | | | | | |
| UK | | | | | | | | |
| ALE | | | | | | | | |
| DOM | | | | | | | | |
| MALL | | | | | | | | |
| MEN | | | | | | | | |
| EF | | | | | | | | |
| DESPESEA | | | | | | | | |
| IPH | | | | | | | | |

Tabla 3. Ejemplo de tabla a elaborar en Mayo 2016

| Series | Mayo-Setiembre 2016 | | Octubre-Diciembre 2016 | | Total 2016 | | Enero-Abril 2017 | |
|----------|---------------------|-------------|------------------------|-------------|------------|-------------|------------------|-------------|
| | Previsión | Incremento% | Previsión | Incremento% | Previsión | Incremento% | Previsión | Incremento% |
| TOT | | | | | | | | |
| INT | | | | | | | | |
| UK | | | | | | | | |
| ALE | | | | | | | | |
| DOM | | | | | | | | |
| MALL | | | | | | | | |
| MEN | | | | | | | | |
| EF | | | | | | | | |
| DESPESEA | | | | | | | | |

De esta manera, a pesar de que la estimación de los modelos de predicción toma como unidad temporal el mes, la información proporcionada a los agentes sería en términos cuatrimestrales y el total anual. En cualquier caso se prevé la posibilidad de ofrecer actualizaciones de las previsiones que cabe esperar mejoren a medida que se obtengan las últimas observaciones. Para llegar a completar dichas tablas el proceso de estimación y selección de modelos que se ha seguido se ha regido por el siguiente procedimiento:

1ª Etapa: Definición de los periodos de estimación y evaluación. Es importante distinguir el periodo de estimación, que comprende los datos que se utilizan para la estimación del modelo, y el periodo de evaluación de las predicciones, que se utiliza para analizar el ajuste de las predicciones efectuadas con el modelo y las observaciones reales. En este sentido, y ante la posibilidad de que diferentes modelos tengan comportamientos diferenciados según el horizonte de predicción, se han tomado como referencia dos ejercicios. En el primero de ellos se han considerado datos desde enero de 1999 hasta diciembre de 2013 para la estimación de los modelos, y los meses del año 2014 para la evaluación de las predicciones. En el segundo ejercicio se han considerado datos desde enero de 1999 hasta diciembre de 2014 para la estimación de los modelos, y todos los disponibles del 2015 para la evaluación de las predicciones.

2ª Etapa: Competición. Obtención de medidas de ajuste de los diferentes modelos ante los datos de prueba y selección del mejor modelo para la realización de las predicciones. Para ello, se utilizan los índices de ajuste más ampliamente utilizados en el análisis de series temporales: RMSPE (Raíz de la Media Cuadrática del Error en términos porcentuales) y MAPE (Error Porcentual Medio Absoluto).⁴ Se selecciona el mejor modelo en función del error cometido en la previsión y en función del alcance en su horizonte de previsión. En este sentido es importante señalar que las comparativas tienen en cuenta el mejor ajuste no sólo del horizonte

⁴ $MAPE = \frac{1}{n} \sum_{t=k}^{k+n} \left(\frac{|e_t|}{A_t} \right) \cdot 100$ $RMSPE = \sqrt{\frac{\sum_{t=k}^{k+n} \left(\frac{e_t}{A_t} \right)^2}{n}} \cdot 100$ donde e representa el error de la predicción, A el valor real de la variable y n el número de observaciones predichas.

completo de predicción sino que también se evaluarán diferentes subgrupos temporales para detectar si existen diferencias entre la bondad de ajuste en función del horizonte temporal de la predicción (y detectar por ejemplo si un determinado modelo se comporta mejor a corto plazo). Finalmente, en esta segunda etapa, también se considera la combinación de predicciones de los dos modelos. Para cada serie, se obtiene una predicción como media ponderada de las predicciones realizadas con los métodos ARIMA y RNA. El peso de cada una de las predicciones toma como referencia la inversa de la varianza obtenida para cada modelo, otorgando, de esta manera, un mayor peso a aquel modelo que haya obtenido una menor varianza en la Etapa 1.

3ª Etapa: Reestimación de los modelos seleccionados con toda la información disponible y realización de predicciones. Presentación de los resultados en cuatrimestres, total anual, en términos absolutos y en términos de crecimiento porcentual respecto al periodo anterior (Ver Tablas 2 y 3).

A través del diseño y prueba de los diferentes modelos de series temporales se prevé poder dar información sobre la evolución esperada de la cada una de las ocho series temporales con un horizonte de previsión de un año de antelación. En cualquier caso, se propone ofrecer la predicción por temporadas con la posibilidad de ir actualizando y revisando las predicciones con el transcurso de los meses y en la medida de que se disponga de más información.

6. RESULTADOS

De acuerdo con la metodología y el protocolo descritos anteriormente los modelos ARIMA estimados en la Etapa 1 aparecen en la Tabla 4. No existen diferencias en los componentes de los modelos en los dos ejercicios planteados en la Etapa 1, o dicho de otra manera, la incorporación del año 2014 a la estimación no altera finalmente la composición de los modelos ni requiere de ningún tipo de intervención adicional.

Tabla 4. Estructura de los modelos ARIMA utilizados en la Etapa 1.

| Serie | MODELO | Intervenciones |
|---------|------------------------------------|----------------|
| TOT | AR(1) MA(11) SMA(12) | 4 |
| INT | AR(1) MA(11) SMA(12) | 5 |
| UK | AR(1) MA(11) SMA(12) | 5 |
| ALE | AR(2) AR(3) AR(11) MA(12) MA(1) | 4 |
| DOM | AR(12) MA(3) | 3 |
| MALL | AR(1) AR(2) SAR(12) SAR(24) MA(11) | 3 |
| MEN | AR(1) AR(2) MA(11) SMA(12) | 6 |
| EF | AR(1) AR(2) MA(12) | 7 |
| DESPESA | AR(1) SAR(12) | 7 |
| IPH | AR(1) | 4 |

Asimismo, tal y como se detalla en el ANEXO I, todas las series se han tomado en logaritmos naturales y han precisado de una diferencia de orden 12, lo que suele ser habitual en la literatura y que precisamente se conoce como “modelo de líneas aéreas” (Lorenzo y Revuelta, 1996). Por otra parte, cabe señalar que para el caso de la serie del turismo doméstico, se ha rechazado para la estimación ARIMA su alargamiento con la serie de llegadas de AENA tal y como inicialmente se proponía en la sección 4. De esta manera se han utilizado únicamente los datos de llegadas de turistas domésticos desde enero de 2009.

La Tabla 5 muestra los modelos de red neuronal seleccionados que han mostrado el mejor ajuste a los datos, según la serie analizada. Como se puede observar, de los tres tipos de red neuronal analizados (Perceptrón Multicapa, Funciones de Base Radial y Redes Recurrentes) con los que se han obtenido diferentes arquitecturas y configuraciones, el Perceptrón Multicapa muestra un mejor ajuste en 8 series y las Funciones de Base Radial muestra un mejor ajuste en las 2 series restantes.

Los modelos Perceptrón Multicapa (MLP) están compuestos por 24 neuronas de entrada, 1 neurona de salida y un número comprendido entre 2 y 15 neuronas ocultas. Las neuronas ocultas han utilizado la función de transferencia sigmooidal tangente hiperbólica y la neurona de salida ha utilizado la función lineal. Los algoritmos de aprendizaje utilizados han sido el backpropagation y el de gradientes conjugados.

Los modelos de Función de Base Radial (RBF) están compuestos por 24 neuronas de entrada, 1 neurona de salida y un número comprendido entre 5 y 35 neuronas ocultas. Las neuronas ocultas han utilizado la función de transferencia gaussiana de tipo radial y la neurona de salida ha utilizado la función lineal. El algoritmo de aprendizaje utilizado para determinar los centroides de las neuronas ocultas ha sido el de k vecinos más cercanos, mientras que el algoritmo de aprendizaje utilizado para determinar los pesos de la neurona de salida ha sido el de gradiente decreciente.

Tabla 5. Modelos RNA utilizados en la Etapa 1.

| Serie | MODELO |
|---------|--------|
| TOT | MLP |
| INT | MLP |
| UK | MLP |
| ALE | MLP |
| DOM | RBF |
| MALL | MLP |
| MEN | MLP |
| EF | RBF |
| DESPESA | MLP |
| IPH | MLP |

Por lo que se refiere a la combinación de predicciones, de acuerdo con la metodología propuesta de asignación de pesos en función de la inversa de la varianza, la proporción de predicción de RNA y ARIMA que configura la combinación de predicción aparece en la Tabla 6.

Tabla 6. Peso de las predicciones ARIMA y RNA sobre la predicción por combinación.

| Serie | ARIMA | RNA |
|---------|-------|-------|
| TOT | 56,7% | 43,3% |
| INT | 80,4% | 19,6% |
| UK | 66,6% | 33,4% |
| ALE | 80,6% | 19,4% |
| DOM | 32,5% | 67,5% |
| MALL | 60,7% | 39,3% |
| MEN | 82,6% | 17,4% |
| EF | 52,6% | 47,4% |
| DESPESA | 74,8% | 25,2% |
| IPH | 80,3% | 19,7% |

Los resultados de las predicciones en términos de MAPE y RMSPE de acuerdo con la metodología descrita para el ejercicio de evaluación de la primera etapa aparece en las Tabla 7 y 8. En general ARIMA se presenta como la metodología mejor para la mayoría de series temporales a excepción del turismo procedente de Alemania (ALE) el doméstico (DOM), donde la RNA parece superior según el MAPE, y para las llegadas a Ibiza-Formentera (EF) y el gasto (DESP), donde la combinación de predicciones es superior. En cualquier caso, es importante señalar como la combinación de predicciones ofrece el menor error en términos promedios (Prom) a pesar de que no es la alternativa más frecuente.

Tabla 7. Comparativa capacidad predictiva ARIMA vs. RNA a través del MAPE. Ejercicio 1

| MAPE | | TOT | INT | UK | ALE | DOM | MALL | MEN | EF | DESP. | IPH | Prom |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|-------|
| TOTAL Anual-2014 | ARIMA | 10,6% | 13,6% | 13,7% | 15,2% | 13,5% | 11,5% | 22,0% | 10,3% | 7,8% | 2,2% | 12,0% |
| | RNA | 10,7% | 21,1% | 22,0% | 9,9% | 11,6% | 12,3% | 23,7% | 9,3% | 9,1% | 5,0% | 21,3% |
| | Comb. | 10,6% | 15,1% | 16,3% | 13,9% | 12,2% | 11,6% | 22,2% | 9,2% | 7,7% | 2,6% | 11,9% |
| 1Q-2014 | ARIMA | 15,6% | 23,8% | 30,7% | 26,9% | 9,6% | 17,4% | 42,3% | 14,4% | 12,4% | 4,1% | 19,7% |
| | RNA | 18,9% | 39,1% | 50,6% | 6,2% | 17,9% | 22,9% | 57,5% | 8,8% | 15,9% | 5,8% | 29,3% |
| | Comb | 17,0% | 26,8% | 37,4% | 22,4% | 15,1% | 19,6% | 45,0% | 9,9% | 13,3% | 4,4% | 20,4% |
| 2Q-2014 | ARIMA | 4,6% | 5,2% | 4,3% | 8,3% | 11,1% | 4,9% | 6,2% | 8,2% | 4,4% | 0,3% | 5,8% |
| | RNA | 4,6% | 12,4% | 6,4% | 7,5% | 8,3% | 4,5% | 6,2% | 4,5% | 6,0% | 5,0% | 15,9% |
| | Comb | 4,6% | 6,6% | 5,0% | 8,1% | 9,2% | 4,1% | 6,2% | 6,5% | 4,8% | 0,7% | 5,6% |
| 3Q-2014 | ARIMA | 12,4% | 13,6% | 7,0% | 13,0% | 20,1% | 11,6% | 15,0% | 17,8% | 10,2% | 2,4% | 12,3% |
| | RNA | 11,5% | 16,4% | 6,6% | 15,8% | 7,7% | 8,5% | 6,9% | 25,5% | 4,7% | 3,7% | 20,1% |
| | Comb | 12,0% | 10,3% | 6,3% | 12,4% | 11,7% | 10,4% | 11,4% | 21,5% | 7,7% | 2,6% | 10,7% |

Las diferencias de la capacidad predictiva según el MAPE o el RMSPE son escasas. En términos totales sólo las llegadas del resto de España (DOM) y a Menorca (MEN) presentan pequeñas divergencias en el criterio de mejor metodología decantándose, en ambos casos, por la combinación de predicciones si nos centramos en el RMSPE. Esta circunstancia podría explicarse por la propensión de la combinación de predicciones a cometer un menor número de grandes errores.⁵

⁵ De esta manera, cabe recordar que el criterio del RMSPE penaliza los grandes errores en mayor medida que el MAPE.

Tabla 8. Comparativa capacidad predictiva ARIMA vs. RNA a través del RMSPE. Ejercicio 1

| MAPE | | TOT | INT | UK | ALE | DOM | MALL | MEN | EF | DESP. | IPH | Prom |
|----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|-------|
| TOTAL Anual-2014 | ARIMA | 13,6% | 20,4% | 28,6% | 21,3% | 15,1% | 13,8% | 44,5% | 18,0% | 11,2% | 3,1% | 19,0% |
| | RNA | 15,3% | 28,4% | 47,5% | 13,0% | 13,4% | 17,8% | 47,3% | 18,9% | 13,7% | 5,6% | 27,3% |
| | Comb | 14,1% | 20,4% | 34,8% | 18,1% | 13,0% | 14,8% | 44,4% | 17,7% | 11,1% | 3,3% | 18,5% |
| 1Q-2014 | ARIMA | 18,9% | 31,4% | 48,4% | 32,1% | 12,0% | 19,5% | 73,8% | 15,9% | 13,7% | 4,4% | 27,0% |
| | RNA | 22,1% | 42,2% | 81,6% | 8,1% | 19,2% | 29,1% | 80,8% | 10,2% | 21,9% | 6,6% | 34,0% |
| | Comb | 20,2% | 33,0% | 59,5% | 26,7% | 16,3% | 22,8% | 74,7% | 11,1% | 15,4% | 4,8% | 27,3% |
| 2Q-2014 | ARIMA | 5,4% | 5,9% | 5,1% | 8,8% | 11,2% | 5,4% | 6,7% | 8,8% | 4,9% | 0,5% | 6,3% |
| | RNA | 5,6% | 12,7% | 6,7% | 7,8% | 8,8% | 5,6% | 6,7% | 5,4% | 6,8% | 5,3% | 16,5% |
| | Comb | 5,4% | 7,1% | 5,6% | 8,5% | 9,5% | 4,4% | 6,5% | 7,1% | 5,3% | 0,8% | 6,0% |
| 3Q-2014 | ARIMA | 12,9% | 14,9% | 9,3% | 15,8% | 20,4% | 12,8% | 21,0% | 25,4% | 12,9% | 2,9% | 14,8% |
| | RNA | 13,4% | 21,6% | 7,5% | 19,4% | 9,8% | 8,7% | 11,1% | 30,6% | 5,8% | 4,8% | 22,5% |
| | Comb | 12,5% | 10,8% | 8,5% | 14,2% | 12,4% | 10,8% | 16,8% | 27,7% | 10,3% | 2,9% | 12,8% |

Los resultados del desarrollo del segundo ejercicio (predicción para los meses disponibles del 2015) aparecen en las Tabla 9 y 10. De nuevo ARIMA se presenta como la mejor metodología para un mayor número de series temporales a pesar de que la combinación de predicciones consigue, de nuevo, reducir en mayor medida el error promedio de todas las predicciones en su conjunto. Por su parte la RNA aparece como la metodología más adecuada para algunos casos concretos. De esta manera ARIMA es la mejora alternativa para todos los casos citados en la primera etapa, con la excepción del turismo procedente del Reino Unido (UK), donde la combinación de predicciones es ahora superior, y además hay que añadir el turismo procedente de Alemania (ALE) (en este caso, en el primer ejercicio la RNA conseguía superar a ARIMA). Por su parte, la RNA sigue siendo la mejor alternativa para el turismo doméstico (DOM), mientras que la combinación de predicciones sigue también ofreciendo mejores resultados para las llegadas a Ibiza-Formentera (EF) y el gasto (DESP).

Tabla 9. Comparativa capacidad predictiva ARIMA vs. RNA a través del MAPE. Ejercicio 2

| MAPE | | TOT | INT | UK | ALE | DOM | MALL | MEN | EF | DESP. | IPH | Prom |
|-----------------------------|-------|-------|-------|-------|-------|-------|------|-------|-------|-------|-------|-------|
| TOTAL Anual-2015* | ARIMA | 4,7% | 4,9% | 8,7% | 6,5% | 9,8% | 9,0% | 10,8% | 9,0% | 4,2% | 3,4% | 7,1% |
| | RNA | 8,1% | 11,2% | 17,3% | 25,7% | 6,4% | 7,2% | 15,2% | 13,0% | 5,2% | 8,0% | 20,0% |
| | Comb | 5,4% | 5,4% | 7,5% | 6,8% | 7,2% | 4,9% | 10,9% | 7,6% | 3,9% | 3,8% | 6,4% |
| 1Q-2015 | ARIMA | 7,8% | 7,2% | 14,4% | 4,5% | 13,3% | 8,9% | 22,7% | 9,4% | 4,3% | 2,4% | 9,7% |
| | RNA | 12,7% | 9,1% | 29,0% | 39,7% | 7,5% | 9,5% | 20,5% | 19,3% | 6,1% | 5,0% | 22,9% |
| | Comb | 9,9% | 7,4% | 12,5% | 9,7% | 9,4% | 8,6% | 22,1% | 9,9% | 4,2% | 2,9% | 9,9% |
| 2Q-2015 | ARIMA | 1,5% | 2,8% | 3,5% | 6,9% | 8,2% | 5,9% | 1,4% | 10,0% | 4,1% | 4,4% | 4,8% |
| | RNA | 3,4% | 13,4% | 5,6% | 11,7% | 5,3% | 5,0% | 9,9% | 6,7% | 4,2% | 10,9% | 17,1% |
| | Comb | 2,0% | 4,2% | 3,6% | 5,6% | 6,2% | 2,3% | 1,8% | 5,6% | 3,5% | 4,8% | 3,9% |

*Total-Anual 2015 incluye sólo los dos primeros cuatrimestres

De nuevo encontramos algunas divergencias de criterio entre el MAPE y el RMSPE en las series de llegadas de turistas procedentes del Extranjero (INT) y a Menorca (MEN), si bien las diferencias en términos de los errores son escasas.

Tabla 10. Comparativa capacidad predictiva ARIMA vs. RNA a través del RMSPE. Ejercicio 2

| RMSPE | | TOT | INT | UK | ALE | DOM | MALL | MEN | EF | DESP. | IPH | Prom |
|---------------------------------|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| TOTAL Anual- 2015* | ARIMA | 6,5% | 7,1% | 12,0% | 7,5% | 13,4% | 10,9% | 15,8% | 9,5% | 4,6% | 3,8% | 9,1% |
| | RNA | 11,8% | 12,1% | 23,6% | 31,4% | 7,5% | 10,8% | 18,0% | 16,6% | 7,0% | 8,7% | 22,4% |
| | Comb | 8,5% | 6,6% | 11,3% | 9,3% | 8,9% | 7,7% | 15,5% | 9,5% | 4,4% | 4,2% | 8,7% |
| 1Q-2015 | ARIMA | 9,2% | 9,9% | 17,2% | 4,9% | 15,9% | 10,4% | 23,7% | 9,8% | 4,8% | 2,8% | 10,9% |
| | RNA | 17,1% | 10,7% | 34,5% | 44,1% | 7,8% | 15,2% | 24,3% | 23,4% | 8,1% | 5,6% | 25,6% |
| | Comb | 12,5% | 8,8% | 16,2% | 11,7% | 9,6% | 11,2% | 23,2% | 12,2% | 4,4% | 3,1% | 11,4% |
| 2Q-2015 | ARIMA | 2,2% | 3,2% | 4,3% | 7,9% | 12,2% | 6,2% | 1,5% | 10,2% | 4,5% | 4,6% | 5,7% |
| | RNA | 4,3% | 13,6% | 6,3% | 16,4% | 7,7% | 5,6% | 10,0% | 6,7% | 5,8% | 11,0% | 18,1% |
| | Comb | 2,8% | 4,5% | 4,5% | 7,4% | 9,1% | 2,8% | 2,0% | 6,6% | 4,4% | 5,0% | 4,9% |

*Total-Anual 2015 incluye sólo los dos primeros cuatrimestres

De esta manera, dada la coherencia obtenida en los dos ejercicios de la primera etapa, la Tabla 11 presenta la selección de modelos efectuada para cada una de las series objeto de estudio con el fin de efectuar la predicción a 12 meses que aparece, finalmente, en la Tabla 12.

Tabla 11. Modelos seleccionados para la realización de las predicciones.

| Serie | MODELO |
|---------|--------|
| TOT | ARIMA |
| INT | ARIMA |
| UK | ARIMA |
| ALE | ARIMA |
| DOM | RNA |
| MALL | ARIMA |
| MEN | COMB |
| EF | COMB |
| DESPESA | COMB |
| IPH | ARIMA |

Las predicciones para el turismo total muestran un incremento del 1% esperado para el último cuadrimestre del año, un incremento del 0,8% para el primer cuadrimestre y una caída del 1,9% para el segundo cuadrimestre del 2016. Cabe destacar que la caída esperada no se traslada en términos de gasto turístico que consigue un aumento del 2,1% para este mismo último periodo. Por nacionalidades, el turismo británico y, en general, el internacional se espera que presenten

descensos durante el 2016 después de presentar un comportamiento positivo durante el último cuadrimestre del 2015. Por su parte, el turismo alemán se prevé que se comporte totalmente a la inversa, con caídas durante el último cuadrimestre del 2015 y recuperación para el 2016.

Tabla 12. Predicciones Setiembre 2015-Agosto 2016

| Series | Setiembre-Diciembre 2015 | | Enero-Abril 2016 | | Mayo-Agosto 2016 | |
|---------|--------------------------|---------|------------------|---------|------------------|---------|
| | Previsión | Incr. % | Previsión | Incr. % | Previsión | Incr. % |
| TOT | 3.450.199 | 1,1% | 1.753.625 | 0,8% | 8.652.577 | -1,9% |
| INT | 2.887.523 | 0,1% | 1.201.330 | -0,3% | 7.370.059 | -2,0% |
| UK | 851.281 | 4,0% | 224.335 | -2,5% | 2.358.453 | -1,3% |
| ALE | 1.176.607 | -4,6% | 644.691 | 1,2% | 2.365.262 | 1,5% |
| DOM | 526.168 | -0,1% | 585.793 | 9,7% | 1.293.527 | -0,1% |
| MALL | 2.564.486 | 1,7% | 1.446.856 | 1,8% | 5.871.195 | -2,8% |
| MEN | 252.969 | 16,6% | 76.629 | 17,2% | 886.832 | -0,6% |
| EF | 642.917 | -2,6% | 234.368 | -0,5% | 1.931.910 | 3,6% |
| DESPESA | 3.113.707 | 2,0% | 1.345.394 | 3,1% | 8.209.642 | 2,1% |
| IPH | 106,69 | 4,6% | 102,75 | 2,9% | 118,41 | 2,3% |

Notas: Incrementos referidos sobre el mismo periodo del año anterior. Estimaciones realizadas con datos hasta setiembre de 2015 con la excepción del IPH donde se han utilizado datos hasta Octubre 2015.

Por islas, los modelos apuntan a que Mallorca y Menorca presentarán un buen comportamiento en temporada baja mientras que se esperarían caídas en el número de turistas durante la temporada alta del 2016, justamente al contrario de Ibiza y Formentera, que presentaría caídas durante la temporada baja y crecimiento en la temporada alta de 2016. Finalmente por lo que se refiere a los precios de los establecimientos hoteleros, se espera que tras presentar un crecimiento interanual del 4,6% durante el último cuadrimestre del año moderen su crecimiento hasta tasas entre el 2% y el 3% hasta la temporada alta de 2016.

7. CONCLUSIONES

La obtención de previsiones ajustadas sobre el número de visitantes tanto en términos globales en las Islas Baleares como por zonas más específicas (por islas, por zonas estratégicas, etc.) así

como de otras variables turísticas relevantes como los precios hoteleros o el gasto turístico se ha constituido en un instrumento imprescindible para la gestión eficiente de los destinos tanto a nivel público como privado. De esta manera, la toma de decisiones operativas, tácticas y estratégicas, en relación a las infraestructuras turísticas necesarias, la programación de la oferta (espacios, alojamiento, comida, transporte, ocio, etc.) y del volumen de personal, así como de las inversiones hoteleras; a fin de poder responder de forma más adecuada a la demanda turística requiere de unas previsiones actividad turística adecuadas.

La actual coexistencia de diferentes métodos para la realización de predicciones de demanda turística en la literatura pone de manifiesto la dificultad para encontrar una metodología que se adapte perfectamente a las necesidades del sector y la controversia entre la utilización de modelos complejos como las redes neuronales artificiales como alternativa a los modelos más clásicos como el ARIMA. En la misma línea de la que parece desprenderse de los últimos trabajos aparecidos en la literatura especializada, los resultados no han sido concluyentes en favor de una u otra metodología, por lo que no sorprende que la combinación de predicciones aparezca como una buena opción para la obtención de predicciones. De esta manera, si bien el modelo ARIMA presenta, en general, mejores resultados, también es verdad que para ciertas series temporales y para determinados horizontes temporales la RNA consigue superar a la metodología ARIMA, siendo el comportamiento de la combinación de predicciones una alternativa que logra mejorar ambas alternativas en determinadas ocasiones.

Sobre la base de la selección de modelos realizada y de la competición llevada a cabo, las previsiones para la temporada alta del 2016 apuntan a que la llegada de turistas a las Islas Baleares presentará un ligero descenso en términos de turistas después de seguir creciendo hasta entonces. El descenso para la temporada alta vendría determinado por la caída del mercado internacional, en general, y del británico, en particular. Por su parte las previsiones sobre el mercado alemán y doméstico apuntan a crecimientos positivos o estancamiento de cara a la temporada alta de 2016. Por islas sólo Ibiza y Formentera consiguen resultados positivos de cara a temporada alta. En cualquier caso, en términos de gasto, las previsiones son



positivas para los próximos 12 meses, circunstancia que vendrá ligada a unos precios hoteleros que se crecerán de manera moderada.

Referencias

- Álvarez-díaz, M., Mateu-Sbert, J., y Rosselló-Nadal, J. (2009). Forecasting tourist arrivals to Balearic Islands using genetic programming. *International Journal of Computational Economics and Econometrics*, 1(1), 64–75.
- Álvarez-Díaz, M., y Rosselló-Nadal, J. (2010). Forecasting British tourist arrivals in the Balearic Islands using meteorological variables. *Tourism Economics*, 16(1), 153–168.
- Álvarez-Díaz, M., y Rosselló-Nadal, J. (2011). An overview of computational methods in forecasting tourism demand: Recommendations and future research. In *Globalization Policies and Issues* (pp. 179–200). New York: Nova Science Publishers, Inc.
- ATB (2015). Indicadors de demanda turística - Illes Balears Agència de Turisme de les Illes Balears. (Disponible en: <http://www.caib.es/sacmicrofront/home.do?mkey=M10072911244127834137&lang=ca>. Acceso 20 marzo de 2015)
- Battiti, R. (1992). First and second order methods for learning: between steepest descent and Newton's method. *Neural Computation*, 4, 141-166.
- Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Bloom, J.Z. (2005). Market segmentation: a neural network application. *Annals of Tourism Research*, 32(1), 93-111.
- Box, G.E.P. y Jenkins, G.M. (1976). *Time series analysis: forecasting and control*. San Francisco: Holden Day.
- Broomhead, D.S. y Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2, 321-355.
- Bull, A. (1995). *The economics of travel and tourism*. Melbourne: Longman.
- Chandra, S. y Menezes, D. (2001). Applications of multivariate analysis in international tourism research: the marketing strategy perspective of NTOs. *Journal of Economic and Social Research*, 3(1), 77-98.
- Clements, M.P y D.F. Hendry (1998). *Forecasting Economic Time Series*. Cambridge University Press.

- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematical Control, Signal and Systems*, 2, 303-314.
- Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Espasa, A. y Cancelo, J.R. (1993). *Métodos cuantitativos para el análisis de la coyuntura económica*, Ed. Alianza: Madrid.
- Haldrup, N. Montañés, M. y Sansó, A. (2011). Detection of Additive Outliers in Seasonal Time Series. *Journal of Time Series Econometrics*, 3(2).
- Haldrup, N., Hylleberg, S., Pons, G. y Sansó, A. (2007). Common Periodic Correlation Features and the Interaction of Stocks and Flows in Airport Data. *Journal of Business and Economic Statistics*, 25(1), 21-32.
- Hansen, J.V., McDonald, J.B. y Nelson, R.D. (1999). Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models. *Computational Intelligence*, 15(3), 171-184.
- Hartman, E., Keeler, J.D. y Kowalski, J.M. (1990). Layered neural networks with Gaussian hidden units as universal approximators. *Neural Computation*, 2(2), 210-215.
- Hornik, K., Stinchcombe, M. y White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2, 359-366.
- Jhee, W.C. y Lee, J.K. (1996). Performance of neural networks in managerial forecasting. En R.R. Trippi y E. Turban (Eds.), *Neural networks in finance and investing* (pp. 703-733). New York: John Wiley.
- Kaastra, I. y Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10, 215-236.
- Law, R. y Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, 20, 89-97.
- Lorenzo, F y Revuelta, J.M. (1996) Tramo y Seats: Un Marco completo para el análisis Univariante y Extracción de Señales de Series Temporales. *Revista de Economía Aplicada*, 11: 201-214.
- Masters, T. (1995). *Advanced algorithms for neural networks: a C++ sourcebook*. New York: John Wiley and Sons.

- Molinet, T., Molinet, J.A., Betancourt, M.E., Palmer, A. y Montaña, J.J. (2015). Models of Artificial Neural Networks Applied to Demand Forecasting in Non-Consolidated Tourist Destinations. *Methodology – European Journal of Research Methods for the Behavioral and Social Sciences*, 11, 35-44.
- Montaña, J.J., Palmer, A. y Muñoz, P. (2011). Artificial neural networks applied to forecasting time series. *Psicothema*, 23, 322-329.
- Montaña, J.J., Palmer, A., Sesé, A. y Beltrán, M. (2004). Redes neuronales artificiales aplicadas al análisis de datos turísticos. En M. Beltrán (coord.), *Técnicas y métodos estadísticos aplicados al sector turístico* (pp. 278-295).
- Montaña, J.J., Palmer, A., Sesé, A. y Cajal, B. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema*, 25, 500-506.
- Morley, C. (2000). Demand modelling methodologies: integration and other issues. *Tourism Economics*, 6(1), 5-19.
- Palmer, A., Montaña, J.J. y Franconetti, F.J. (2008). Sensitivity analysis applied to artificial neural networks for forecasting time series. *Methodology – European Journal of Research Methods for the Behavioral and Social Sciences*, 4, 80-86.
- Palmer, A., Montaña, J.J. y Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism Management*, 27, 781-790.
- Palmer, A., Sesé, A. y Montaña, J.J. (2005). Tourism and statistics: bibliometric study 1998-2002. *Annals of Tourism Research*, 32, 167-178.
- Peng, B., Song, H., & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism Management*, 45, 181–193. doi:10.1016/j.tourman.2014.04.005
- Pons, G. y Sansó, A. (2005). Estimating Cointegrating Vectors with Time Series Measured with Different Periodicities. *Econometric Theory*, 21(4), 735-756.
- Roselló, J. (2001). Forecasting turning points in International visitor arrivals in the Balearic Islands. *Tourism Economics*, 7(4), 365-380.
- Roselló, J., Riera, A. y Sansó, A. (2004). The Economic Determinants of Seasonal Patterns. Seasonality in Monthly International Tourist Arrivals to the Balearic Islands. *Annals of Tourism Research*, 31(3), 697–711.

- Rumelhart, D.E., Hinton, G.E. y Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart, and J.L. McClelland (Eds.), *Parallel distributed processing* (pp. 318-362). Cambridge, MA: MIT Press.
- Sheldon, P.J. y Var, T. (1985). Tourism forecasting: a review of empirical research. *Journal of Forecasting*, 4(2), 183-195.
- Shen, S., Li, G. y Song, H. (2011) Combination forecasts of International tourism demand. *Annals of Tourism Research*, 38(1), 72–89
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29(2), 203–220. doi:10.1016/j.tourman.2007.07.016
- UNWTO (2015) Historical perspective of world tourism. United Nations World Tourism Organization: Madrid. (Disponible en: <http://www.unwto.org/facts/eng/historical.htm>. Acceso 25 mayo de 2015)
- Uysal, M. y El Roubi, M.S. (1999). Artificial neural networks versus multiple regression in tourism demand analysis. *Journal of Travel Research*, 38, 111-118.
- Wang, C.H. (2004). Predicting tourism demand using fuzzy time series and hybrid grey theory. *Tourism Management*, 25, 367-374.
- Widrow, B. y Hoff, M. (1960). Adaptive switching circuits. En J. Anderson y E. Rosenfeld (Eds.), *Neurocomputing* (pp. 126-134). Cambridge, Mass.: The MIT Press.
- Wong, K., Song, H., Witt, S.F. and Wu, D.C. (2007) Tourism forecasting: To combine or not to combine? *Tourism Management*, 28(4), 1068–1078

ANEXO I

Modelos ARIMA estimados en la Etapa1. Detalle.

LOG(TOT)-LOG(TOT(-12)) = 0.00884275838525 - 0.304061856895*D2010M01 - 0.198695106271*D2002M03(-1) +
0.135130563209*D2003M04 - 0.266644085147*D2012M11 + 0.223680649811*D2011M01 -
0.0929253849126*D2011M12 + 0.0918547544637*D2013M03 +
[AR(1)=0.628538713009,MA(11)=0.555869173781,SMA(12)=-
0.519859078063,BACKCAST=2000M02,ESTSMPL="2000M02 2014M12"]

LOG(INT)-LOG(INT(-12)) = -0.00181008995361 - 0.322680352403*D2002M03 + 0.173597469638*D2011M04 -
0.205406066569*D2011M12 + 0.327979758314*D2003M04(1) + 0.13700753864*D2013M03 +
[AR(1)=0.558619525085,MA(11)=0.446654428253,SMA(12)=-
0.431238355287,BACKCAST=2000M02,ESTSMPL="2000M02 2014M12"]

LOG(UK)-LOG(UK(-12)) = -0.00848212041366 - 0.476269876891*D2009M11 - 0.633026430512*D2009M12 +
0.259370408438*D2011M04 - 0.293721592555*D2010M01 - 0.293926486802*D2010M11 +
[AR(1)=0.383545407292,MA(11)=0.318129377278,SMA(12)=-
0.35637879626,BACKCAST=2000M02,ESTSMPL="2000M02 2014M12"]

LOG(ALE)-LOG(ALE(-12)) = 0.0147364182372 - 0.306182851454*D2002M03 - 0.172325492663*D2006M02 -
0.368436723166*D2011M12 - 0.22684677994*D2010M11 +
[AR(2)=0.270805744174,AR(3)=0.180660218848,AR(11)=0.214568305634,MA(12)=-
0.488290276552,MA(1)=0.398968928977,BACKCAST=2000M12,ESTSMPL="2000M12 2014M12"]

LOG(DOM)-LOG(DOM(-12)) = 0.017729417346 - 0.537433756447*D2010M01 - 0.439746236582*D2010M03 -
0.476988387052*D2012M11 + [AR(12)=-
0.164768367629,MA(3)=0.154006815374,BACKCAST=2001M01,ESTSMPL="2001M01 2014M12"]

LOG(MALL)-LOG(MALL(-12)) = 0.0255188960581 - 0.302759043001*D2012M11 - 0.1880908616*D2013M02 -
0.108836169453*D2011M12 + [AR(1)=0.555348040839,AR(2)=0.225581722437,SAR(12)=-
0.71721039075,SAR(24)=-
0.326011524352,MA(11)=0.887626101281,BACKCAST=2002M03,ESTSMPL="2002M03 2014M12"]

LOG(MEN)-LOG(MEN(-12)) = -0.00398512242122 - 0.414036858005*D2012M11 - 0.525474533322*D2012M12 +
0.086208818226*D2013M02(1) + 0.79169127*D2013M02 - 0.22351168683*D2002M04 -
0.130527484681*D2013M11 +
[AR(1)=0.342428208556,AR(2)=0.292301329163,MA(11)=0.479320339177,SMA(12)=-
0.377070413855,BACKCAST=2000M03,ESTSMPL="2000M03 2014M12"]

LOG(EF)-LOG(EF(-12)) = 0.0335173416462 - 0.130386103925*D2013M02(1) - 0.38065846969*D2002M04 +
0.419851728835*D2000M4 + 0.222549686518*D2011M04 + 0.266032319922*D2003M04 -
0.218209027685*D2004M4 + 0.144077823484*D2002M11 +
[AR(1)=0.339704030738,AR(2)=0.329402419395,MA(12)=-
0.263395322013,BACKCAST=2000M03,ESTSMPL="2000M03 2014M12"]

LOG(DESPEsa_NEW)-LOG(DESPEsa_NEW(-12)) = 0.0317740831714 - 0.265117142555*D2011M11(-1) -
0.190929599069*D2013M02 - 0.315683574948*D2012M11 + 0.102472437708*D2011M04 +
0.237926358387*D2009M01 + 0.189118372655*D2009M11 + 0.169633116092*D2009M02 +
[AR(1)=0.459623549957,SAR(12)=-0.493814477232]

LOG(IPH)-LOG(IPH(-12)) = 0.0172361476202 + 0.0658279615735*D2002M11 - 0.0556759515119*D2011M11 -
0.0804558546056*D2011M12 - 0.0565987916559*D2013M11 + [AR(1)=0.729315019462]

Nota: Las intervenciones tiene el formato D(año)(mes)