



**Universitat de les
Illes Balears**

Facultat d'Economia i Empresa

Memòria del Treball de Fi de Grau

Análisis sobre la predicción de las llegadas turísticas a Mallorca utilizando el índice de Google Trends

Daniel Ramos Hoogwout

Grau d'Economia

Any acadèmic 2019-20

DNI de l'alumne: 43202637N

Treball tutelat per Dr. Victor Emilio Troster
Departament d' Economia Aplicada de la Universitat de les Illes Balears

S'autoritza la Universitat a incloure aquest treball en el Repositori Institucional per a la seva consulta en accés obert i difusió en línia, amb finalitats exclusivament acadèmiques i d'investigació	Autor		Tutor	
	Sí	No	Sí	No
	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Paraules clau del treball:
Google Trends; Predicció; Demanda turística; Modelos ARDL; Causalidad de Granger

Índice

Resumen.....	3
Abstract.....	4
Introducción.....	5
Metodología.....	9
Análisis empírico	12
Conclusiones.....	19

Índice de tablas

Tabla 1: Estadísticos descriptivos	14
Tabla 2: Matriz de correlaciones lineales	14
Tabla 3: Modelo ARDL entre las llegadas turísticas y GT para “Mallorca”	15
Tabla 4: Modelo ARDL entre las llegadas turísticas y Δ GT para “Majorca”	17

Índice de figuras

Figura 1: Evolución de las llegadas turísticas en Mallorca y los índices de GT	13
Figura 2: Correlograma de los residuos del modelo ARDL entre las llegadas turísticas y GT para “Mallorca”	16
Figura 3: Correlograma de los residuos del modelo ARDL entre las llegadas turísticas y Δ GT para “Majorca”	18

Resumen

En este trabajo se analiza la posible existencia de una relación entre las llegadas turísticas a Mallorca y los índices de Google Trends para el período 2015-2019. Para poder hacer predicciones sobre la demanda turística en Mallorca el estudio se basará en una amplia variedad de literatura previa que demuestra la eficacia de este índice de Google sobre la predicción de demandas en diferentes ámbitos turísticos. A lo largo de este estudio se aplicarán diferentes metodologías econométricas, estimando un modelo ARDL con la ayuda de correlogramas, aplicando distintos contrastes como Dickey & Fuller (1979) y la prueba de causalidad de Granger (1969) observando que el índice de Google Trends sí es realmente es un buen indicador para poder predecir la demanda turística. Esto tiene varias implicaciones muy interesantes que podrían servir para la toma de decisiones de agentes económicos que dependan directamente de la demanda turística para la obtención de beneficios y que en base a estas predicciones puedan tomar decisiones acertadas.

Palabras clave: Google Trends; Predicción; Demanda turística; Modelos ARDL; Causalidad de Granger

Abstract

This paper analyzes the existence of a relationship between monthly tourist arrivals in Mallorca and searches in the Google Trends index for the period of 2015-2019. To make predictions about tourist demand in Mallorca, the study reviews the literature that demonstrates the effectiveness of the Google Trends on predicting demand at different touristic destinations. This study estimates an autoregressive distributed lag (ARDL) model between tourist arrivals in Mallorca and Google Trends index to verify whether the Google Trends index is a good predictor of tourism demand in Mallorca. The results have interesting implications that can be used for decision-making by economic agents who depend directly on tourism demand to obtain benefits.

Keywords: Google Trends; Forecast; Tourism demand; ARDL models; Granger-causality

1. Introducción

En las últimas décadas la demanda turística global ha crecido de manera exponencial, por este mismo motivo también ha habido un aumento significativo en el número de estudios dedicados a predecir esta demanda tal y como señalan Song & Li (2008). El auge de este tipo de estudios viene dado por la propia naturaleza de los productos turísticos, ya que estos son perecederos y no almacenables (Gunter & Önder, 2015). Por este motivo es tan importante hacer una buena predicción de la demanda turística, por lo que los agentes turísticos deben poder hacer ajustes con anterioridad para poder amoldar su oferta y utilizar sus recursos de la manera más eficiente posible e incluso poder hacer ajustes a corto plazo para maximizar el beneficio.

Tal y como señalan Song & Li (2008), la variable más popular para realizar estudios sobre la demanda turística es el número de llegadas de turistas, por lo tanto, incluiremos esta variable para hacer nuestras predicciones. Por otro lado, el principal punto de información para el turismo es Internet y en concreto el buscador de Google (NetMarketShare, 2020), en el cual miles de personas planean sus viajes a diario por la facilidad y la comodidad que ofrecen tanto los navegadores como las propias páginas web que ofertan productos turísticos (Doblemente, 2015). También se estima que el 75% de las búsquedas de viajes se realizan a través de internet y que el 67% de los viajes se inspiraron de manera digital (Europa Press, 2018). Por lo tanto, es interesante estudiar si se pueden utilizar herramientas como Google Trends (<https://trends.google.es/>) para predecir la demanda turística.

Existe una amplia literatura que ha estudiado el uso de índices de Google para hacer predicciones sobre distintos ámbitos económicos como Choi & Varian (2012) que explican la utilidad de estos índices para hacer predicciones a corto plazo sobre distintas variables económicas como la venta de automóviles, el desempleo, la confianza del consumidor y la planificación de viajes.

Además, existe una literatura que se centra exclusivamente en el uso del índice de Google Trends para hacer predicciones sobre demandas turísticas. Park et al. (2017) modelizan la cantidad de turistas japoneses que viajan a Corea del Sur desde 2004 a 2015 utilizando este índice para construir su modelo, los autores concluyen que el índice de Google puede mejorar la precisión de la demanda turística del número de turistas japoneses que llegan a Corea del Sur. Önder (2017) compara la precisión de estas previsiones para distintas ciudades y países como Barcelona, Viena, Bélgica y Austria desde principios de 2008 hasta finales de 2012 utilizando el índice de Google en un modelo autorregresivo con estacionalidad, concluyendo que el índice de Google es una opción viable para predecir la demanda turística tanto en países como en ciudades.

Gunter & Önder (2015) demuestran que merece la pena aplicar diferentes modelos econométricos dependiendo del país de origen de la demanda turística para obtener datos más precisos, es decir, un solo modelo que utilice Google Trends no es representativo para los distintos mercados turísticos, pero utilizando un conjunto de modelos se podrían obtener datos más precisos que los agente turísticos podrían usar para predecir mejor la demanda, de hecho en este mismo estudios se aplica este criterio y se generan dos modelos distintos para representar mejor las distintas nacionalidades que visitan Mallorca .

Incluso hay autores como Bokelmann & Lessmann (2019) que reconocen la utilidad de estos índices para hacer predicciones turísticas y proponen diferentes métodos para sanear los datos de Google Trends y así evitar patrones espurios que impacten negativamente en la precisión de estas predicciones sobre la demanda turística. Bangwayo-Skeete & Skeete (2015) concluyen la importancia de este índice a la hora de hacer predicciones ya que cada vez más personas utilizan Google para planear sus viajes y estas bases de datos pueden ser utilizadas por diferentes agentes económicos para poder hacer ajustes y proyectar de manera correcta la demanda futura de turísticas en distintos lugares y en concreto en el Caribe.

Li & Law (2019) estudian la posibilidad de mejorar las predicciones sobre la demanda turística descomponiendo las búsquedas en los buscadores de internet, en concreto Google Trends, este método se aplicó a las llegadas turísticas mensuales a Hong Kong desde nueve países distintos concluyendo que el uso de este tipo de índices descompuestos superan el modelo de Choi & Varian (2012) y también demuestran que este tipo de metodología es superior a la hora de pronosticar puntos de inflexión.

Xie et al. (2020) explican que las búsquedas en internet pueden utilizarse para crear índices de búsquedas web y utilizarse para mejorar la predicción de la demanda turística demostrando que los índices de búsquedas web son excelentes predictores para la predicción de la demanda turística. Bi et al. (2020) analizan la posibilidad de mejorar la predicción sobre la demanda turística de dos atracciones turísticas en china utilizando el volumen histórico de turistas, la previsión del tiempo y datos sobre las búsquedas en internet consiguiendo así mejores resultados en la predicción de la demanda turística, evidenciando que las búsquedas en internet son de gran importancia a la hora de predecir la demanda turística.

En definitiva, una larga serie de estudios aparte de los mencionados más arriba han confirmado que el uso de Google Trends puede mejorar de manera significativa la precisión sobre las predicciones de la demanda turística (Pan et al., 2012; Gunter & Önder, 2016; Rivera, 2016).

Para el caso de España ya existen numerosos estudios que modelizan la demanda turística. El autor de Oliveira Santos (2009) desagrega la demanda turística en diferentes componentes de acuerdo a variables como la residencia, el motivo del viaje, el tipo de transporte y el alojamiento concluyendo que desagregando la demanda turística se obtienen predicciones ligeramente más precisas que el método generalmente utilizado de seleccionar la demanda total. Garin-Munoz & Amaral (2000) miden el impacto de los determinantes económicos de la demanda internacional de servicios turísticos en España

utilizando datos de panel para ver los efectos de la renta per cápita, el tipo de cambio y los precios reales sobre la demanda turística entre el periodo de 1985 y 1995.

Rodríguez & Rivadulla (2012) analizan los principales determinantes de la demanda del turismo internacional en España en el periodo 2001-2009 y cuantifican su incidencia. El turismo en España no se distribuye de manera homogénea por su territorio por lo tanto los autores crean 5 modelos, uno para la totalidad del país y los cuatro restantes agregando las Comunidades Autónomas dependiendo de su perfil turístico.

Respecto a la literatura sobre la demanda turística de las Baleares, ya se ha intentado modelizar con anterioridad la demanda turística mediante variables meteorológicas en el cual los autores afirman que incluir este tipo de variables puede incrementar la capacidad predictiva del modelo (Álvarez-Díaz & Rosselló-Nadal, 2010). También se han creado modelos con índices basados en variables macroeconómicas para predecir los puntos de inflexión de las llegadas de distintos países, en este caso Alemania y Gran Bretaña aunque en este caso el autor afirma que persiste incertidumbre (Rossello-Nadal, 2001).

Es curioso que siendo uno de los destinos turísticos referentes en Europa exista una escasez de literatura reciente en materia de predicción de la demanda turística independientemente de que se utilicen índices para la creación de estos.

El objetivo de este estudio es modelizar la demanda turística de Mallorca mediante el índice de Google Trends para demostrar si es posible hacer predicciones sobre dicha demanda. Para lograr este objetivo se hará uso de datos como el número de llegadas y el índice de Google Trends que se incluirán en dos términos diferentes “Mallorca” y “Majorca” con el fin de ajustar mejor el modelo a diferentes nacionalidades tal y como se ha mencionado antes al citar a los autores Gunter & Önder (2015). Se observará su correlación y con ambos términos de Google Trends y las llegadas turísticas a Mallorca se construirán dos

modelos autorregresivos con retardos distribuidos (ARDL) incluyendo el contraste de Granger (1969) para verificar si realmente los índices de Google Trends son útiles para poder predecir las llegadas turísticas a Mallorca. A continuación, se extraerán los multiplicadores de los modelos para que sea más fácil lograr una predicción de la demanda turística futura. Por último, se procederá a la conclusión, en la cual se afirma que el índice de Google Trends es útil a la hora de hacer predicciones sobre la demanda turística, tal y como afirman también otros autores de los estudios mencionados con anterioridad.

El resto del estudio se organiza de la siguiente manera. En la Sección 2 se llevará a cabo la explicación sobre metodología econométrica utilizada en este estudio. En la Sección 3 se ejecutará la metodología explicada en el punto anterior obteniendo los modelos econométricos y sus resultados. Finalmente, en la Sección 4 se mostrarán las conclusiones de este estudio.

2. Metodología

En esta sección, se explicará la metodología utilizada para la elaboración de este estudio. La manera habitual de comenzar un modelo de predicción es utilizando valores pasados de la variable dependiente, por tanto, es adecuado hacer uso de un modelo autorregresivo (AR), en el cual la variable dependiente es predicha por sus propios valores retardados más un error aleatorio no correlacionado con sus valores pasados. El número de estos retardos utilizados para hacer esta predicción se llama orden de la auto regresión. Sea Y_t las llegadas turísticas en el mes t , la expresión matemática de un autorregresivo de orden p , AR(p), sería:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + u_t, \quad (1)$$

donde u_t es un error aleatorio de ruido blanco (con media cero, varianza constante e independiente de sus valores pasados y futuros).

En este estudio se hará uso de modelos econométricos dinámicos uniecuacionales, en concreto, modelos autorregresivos con retardos distribuidos

ARDL (“Autoregressive Distributed Lag”), ya que estos modelos son especialmente útiles para hacer predicciones sobre series temporales al tener en cuenta otras variables y sus respectivos retardos en vez de solo utilizar los retardos de la variable dependiente. Sea X_t el índice de búsqueda en el Google Trends en el mes t . La expresión matemática de un autorregresivo de orden p , ARDL(p,r), es:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \delta_1 X_{t-1} + \dots + \delta_r X_{t-r} + u_t. \quad (2)$$

Al tratar con series temporales se hará uso de correlogramas de los residuos, que son representaciones gráficas de la autocorrelación de los residuos, con sus retardos, $Corr(u_t, u_{t-j})$ para $j \geq 1$, que se emplean para detectar posibles problemas de autocorrelación en los errores del modelo u_t . De no ser así, los modelos no estarán correctamente especificados, por lo que pierden su validez. Estos correlogramas emplean el contraste de autocorrelación de Ljung-Box (Ljung & Box, 1978) que contrasta la hipótesis nula de ausencia de autocorrelación de los errores u_t , $Corr(u_t, u_{t-j}) = 0 \forall j \geq 1$, frente a la hipótesis alternativa de que los errores están autocorrelacionados con al menos un retardo:

$$H_0: Corr(u_t, u_{t-j}) = 0, \text{ para todo } j \geq 1,$$

$$H_A: Corr(u_t, u_{t-j}) \neq 0, \text{ para algún } j \geq 1.$$

En caso de encontrar autocorrelación de los residuos con algún retardo, se procedería al aumento de la orden del AR o a añadir un AR estacional, un modelo autorregresivo en el cual cada valor de la serie viene influenciado por el valor que ha tomado la serie momentos antes, con el fin que el correlograma de los residuos quede como de uno de ruido blanco.

Además, se intentará minimizar los criterios de Akaike (AIC) (Akaike, 1974) y Schwarz (BIC) (Schwarz, 1978) con el fin de seleccionar el mejor modelo ARDL(p,r), ya que un valor más bajo en alguno de estos criterios se traduce en

un mejor ajuste del modelo. Este hecho se consigue mediante un modelo con pocas variables explicativas pero que sean muy relevantes para explicar las llegadas turísticas. Las expresiones matemáticas para estos dos criterios serían:

$$AIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + (p + 1) \frac{2}{T}, \quad (3)$$

$$BIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + (p + 1) \frac{\ln T}{T}, \quad (4)$$

donde T es el número de observaciones de la serie y $SSR(p)$ es la suma de los cuadrados de los residuos.

Antes que nada, empezaremos con el análisis de correlación lineal entre las llegadas turísticas y el índice de Google Trends. Además, verificaremos si los retardos del índice de Google Trends ayudan a predecir las llegadas turísticas. Sin embargo, la existencia de correlación entre las llegadas y el índice de Google Trends no implica causalidad. Por lo tanto, en este estudio incluiremos un contraste para estudiar si los retardos de las variables explicativas son útiles para hacer predicciones sobre la variable dependiente. Se tomará como hipótesis nula el hecho de que ninguno de estos retardos tiene información útil para predecir la variable dependiente, esta hipótesis se denomina contrastes de causalidad en sentido Granger (Granger, 1969):

$$H_0: E(Y_t | Y_{t-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-r}) = E(Y_t | Y_{t-1}, \dots, Y_{t-p}), \quad (5)$$

$$H_A: E(Y_t | Y_{t-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-r}) \neq E(Y_t | Y_{t-1}, \dots, Y_{t-p}), \quad (6)$$

Donde no hay causalidad en sentido Granger bajo la hipótesis nula, es decir, los retardos del índice de Google Trends no ayudan a predecir las llegadas turísticas a Mallorca. El contraste de las Ecuaciones (5)-(6) se hace mediante un contraste F de omisión conjunta de los retardos de X_t en un modelo ARDL(p,r).

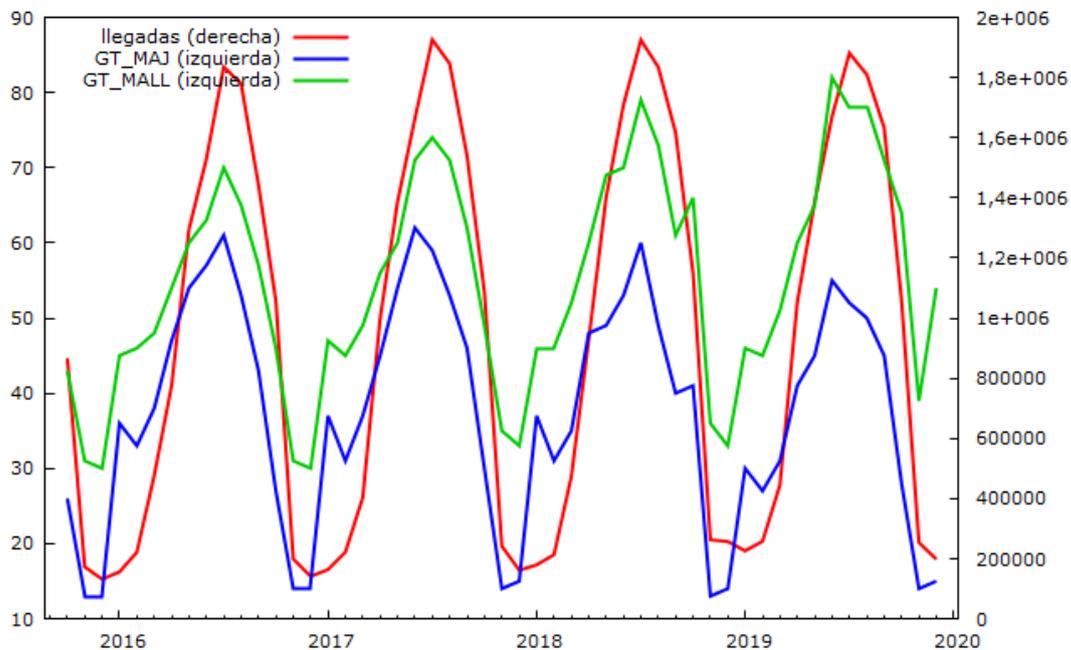
3. Análisis empírico

En esta sección, se hará el análisis empírico de la relación entre búsquedas en el Google Trends y llegadas turísticas a Mallorca. Los datos utilizados para el análisis se han extraído del IBESTAT (<https://ibestat.caib.es/>) y del Google Trends (<https://trends.google.es/>). El IBESTAT es el Instituto de Estadística de las Islas Baleares que recoge el número de llegadas de turistas tanto nacionales como internacionales a Mallorca de manera mensual. Google Trends es una herramienta del buscador de Google que mide la popularidad de un término a lo largo del tiempo mediante un índice porcentual, cuantas más búsquedas se realizan sobre dicho término, más crece el índice (Google, 2020). Estos datos quedan almacenados a diario en los servidores de Google y pueden extraerse en series mensuales que datan de 2004 hasta el momento actual.

Se utilizarán los términos “Mallorca” y “Majorca” para medir su popularidad en Google Trends con el fin de observar su relación con las llegadas turísticas y porque utilizando estos dos términos nos permite centrarnos más en dos mercados diferentes, el alemán y el inglés tal y como proponen Gunter & Önder (2015) tal y como se ha mencionado en la Introducción de este estudio siendo Alemania el país que más búsquedas ha realizado utilizando el término “Mallorca” y Gran Bretaña el término “Majorca” (Google Trends, 2020). El período abarcado en este estudio data del 1/10/2015 hasta 1/12/2019 por la disponibilidad de los datos en IBESTAT.

La Figura 1 abajo presenta la evolución de las llegadas turísticas en Mallorca y los índices de búsquedas de Google Trends (GT). Se puede observar un comportamiento muy parecido en la evolución entre las llegadas y el índice de Google Trends. Así mismo, hay indicios de la existencia de una correlación entre ambas.

Figura 1. Evolución de las llegadas turísticas en Mallorca y los índices de GT



Fuente: Elaboración propia, datos extraídos de IBESTAT y Google Trends.

En la Tabla 1 que se muestra a continuación se pueden observar los estadísticos descriptivos de las variables seleccionadas. En primer lugar, hay que destacar que el cociente de asimetría es positivo lo que indica una concentración de valores a la izquierda de la distribución, pero al ser este cociente muy cercano a cero podemos asumir que la distribución es simétrica. En segundo lugar, la Curtosis de esta distribución es negativa lo que indica una Curtosis platicúrtica, es decir, una distribución más plana que una normal y con colas más alargadas. En tercer lugar, en el contraste de Jarque-Bera indica que la serie de llegadas tiene una distribución normal al 5% de significación, por lo que podemos suponer que la distribución se rige por la de una normal. Por último, la prueba estadística de Dickey-Fuller Aumentada no se rechaza para la hipótesis nula de la existencia de raíces unitarias al 5% de significación para el índice GT_MAJ, por lo que se ha utilizado la diferencia de las búsquedas de Google para el término “Majorca”, con esta modificación se puede suponer que no hay indicios de raíces unitarias.

Tabla 1. Estadísticos descriptivos

	Llegadas _t	GT _{MALL}	GT _{MAJ}	ΔGT _{MAJ}
Mínimo	131580	30	13	-28
Mediana	998590	54	38	1
Máximo	1925700	82	62	23
Media	932080	54,804	37,549	-0,22
Desviación Típica	661840	14,536	15.161	10,758
Asimetría	0,118	-0.31133	-0.26548	-0.004
Exceso de Curtosis	-1,598	-0.988	-1.06	0.061
P-valor JB	0,063	0,353	0,225	0,996
ADF	-7,193	-6,760	-0,536	-5,418

Notas: El *p*-valor JB es el *p*-valor del contraste de normalidad de Jarque y Bera (Jarque & Bera, 1987) que se utiliza para comprobar si los datos tienen la asimetría y la curtosis de una distribución normal. ADF es el estadístico de prueba del contraste de raíces unitarias de Dickey y Fuller (Dickey & Fuller, 1979), que contrasta la existencia de raíces unitarias. Los valores del ADF en negrita indican un rechazo de la hipótesis nula al 5% de significación de que la serie tiene raíces unitarias.

En la Tabla 2 abajo se puede observar la alta correlación que tienen las llegadas turísticas con los índices de Google Trends, para el caso de “Mallorca” se puede observar una correlación del 90% y para “Majorca” un 84% por lo que se cumple la observación antes mencionada sobre la alta correlación que tienen estos términos vistos en la Figura 1.

Tabla 2. Matriz de correlaciones lineales

<i>j</i>	Corr(Llegadas _t , GTMAJ _{t-j})	Corr(Llegadas _t , GTMALL _{t-j})
0	0,831**	0,903**
1	0,893**	0,851**
2	0,734**	0,606**
3	0,416**	0,254**

Nota: GTMAJ se refiere al índice de Google Trends “Majorca” y GTMALL es el índice que se refiere a “Mallorca” *j* indica el número de retardos, siendo 0 el momento actual. Los asteriscos ** indican que la correlación es significativa al 5% de significación.

La alta correlación tanto del índice contemporáneo como de los 2 primeros retardos probablemente se deba al hecho de que hoy en día los individuos buscan información antes y durante su estancia en un destino turístico, ya sea para reservar tanto el transporte como el hotel o simplemente para hacerse una idea de antemano de la situación que se van a encontrar al llegar y poder planificar visitas, actividades o excursiones por la isla.

Es interesante mencionar que en la tabla anterior se pueden observar comportamientos distintos dependiendo de la nacionalidad. El término “Majorca” representa en su mayoría a británicos los cuales hacen búsquedas en Google con mucha más antelación que los individuos que utilizan el término “Mallorca” su mayoría alemanes los cuales utilizan búsquedas en Google ligeramente más en el momento actual.

A continuación, se analizarán los mismos datos mediante modelos dinámicos, los cuales son útiles para modelizar el efecto que tienen los retardos de las búsquedas en Google Trends sobre la variable llegadas, ya que su efecto sobre esta se distribuye a lo largo del tiempo.

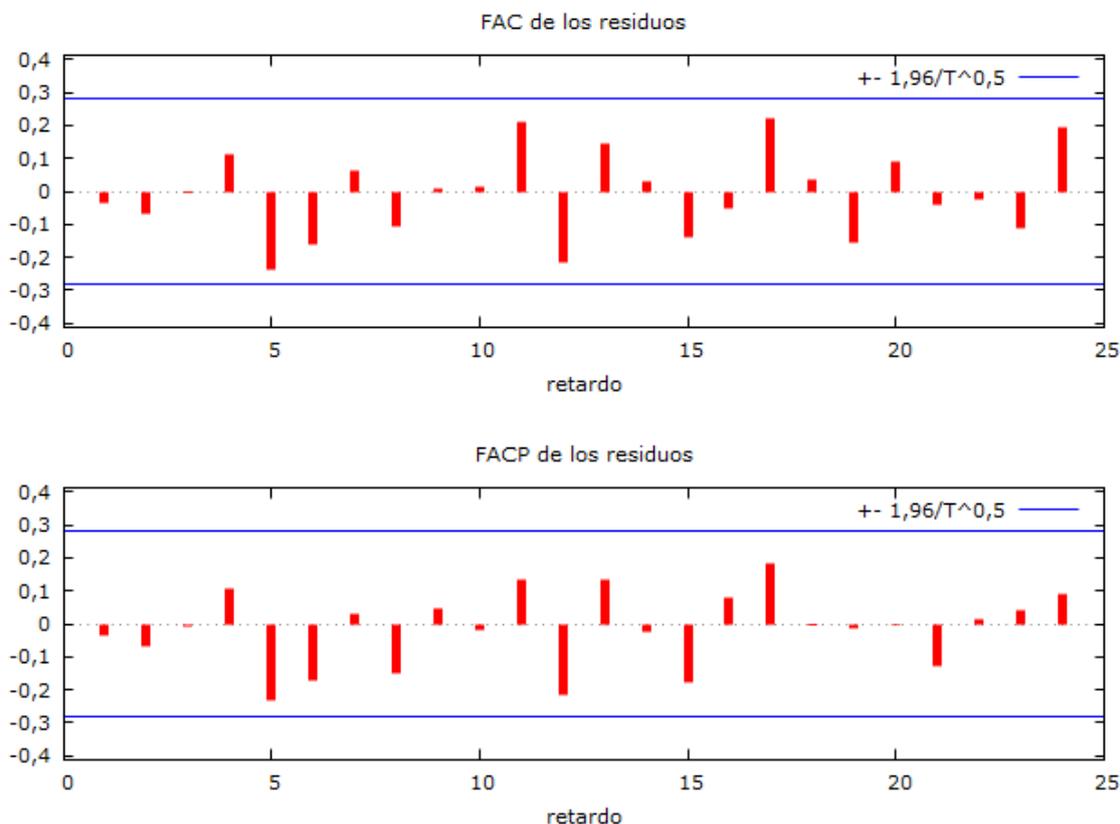
En la Tabla 3 siguiente se puede observar el modelo ARDL seleccionado entre las llegadas turísticas y las búsquedas en el Google Trends para el término Mallorca. Este modelo se ha realizado incluyendo un proceso autorregresivo con retardos específicos uno, dos y cuatro a la vez que un proceso autorregresivo estacional. Éste el mejor modelo según los criterios de AIC y BIC y los residuos de este modelo no están autocorrelacionados.

Tabla 3. Modelo ARDL entre las llegadas turísticas y GT para “Mallorca”

	Coefficiente	Desv. Típica	P-valor
β_0	475579**	197547	0,017
β_1	0,388***	0,093	0,000
β_2	0,363***	0,105	0,001
β_4	-0,519***	0,086	0,000
ϕ_1	0,963***	0,016	0,000
δ_2	5221,870**	2200,150	0,018
δ_3	3504,550	2194,860	0,110
AIC= 1247,054; BIC= 1262,023			

Notas: Las β_i son los coeficientes parciales del modelo, la ϕ_1 es el coeficiente de un AR(1) estacional en el modelo y las δ_i son los retardos del término “Mallorca” y el subíndice i se refiere al número del retardo. Los asteriscos indican que el coeficiente es significativo al 1% de significación con tres asteriscos y al 5% de significación con dos asteriscos.

Figura 2. Correlograma de los residuos del modelo ARDL entre las llegadas turísticas y GT para “Mallorca”



Fuente: Elaboración propia, datos extraídos de IBESTAT y Google Trends.

Una vez obtenido el modelo final y su respectivo correlograma se procede a realizar el contraste de causalidad tipo Granger (1969), donde la hipótesis nula se obtiene igualando los coeficientes retardos de GT_MALL a cero, en concreto, δ_2 y δ_3 :

$$H_0: \delta_2 = \delta_3 = 0,$$

frente a la alternativa de que al menos uno de los coeficientes ($\delta_{2,3}$) es distinto de cero.

Una vez aplicados el conjunto de restricciones nos aparece un p -valor de 0,022, siendo menor que 5% de significación, por lo tanto, se rechaza la hipótesis nula de que no hay causalidad de tipo Granger. Esto significa que los retardos del índice de búsquedas de GT ayudan a predecir las llegadas turísticas a Mallorca. Según el modelo ARDL de la Tabla 3, un cambio unitario en el índice de GT

“Mallorca” de hace 2 meses impactaría en aproximadamente 5221 turistas, puesto que el coeficiente del GT de dos retardos es significativo al 5% de significación e igual a 5221,87.

La siguiente tabla muestra el modelo dinámico final para el término “Majorca”. Esta vez se ha realizado incluyendo un proceso autorregresivo con retardos específicos dos y cuatro a la vez que un proceso autorregresivo estacional para evitar que el correlograma de los residuos mostrase autocorrelación.

Tabla 4. Modelo ARDL entre las llegadas turísticas y Δ GT para “Majorca”

	Coeficiente	Desv. Típica	P-valor
β_0	960977***	190397	0,000
β_1	0,374***	0,097	0,000
β_2	0,465***	0,111	0,000
β_4	-0,502***	0,075	0,000
ϕ_1	0,967***	0,016	0,000
δ_2	2494,080*	1500,350	0,096
δ_3	3226,330*	1658,720	0,052
AIC= 1259,04; BIC= 1272,14			

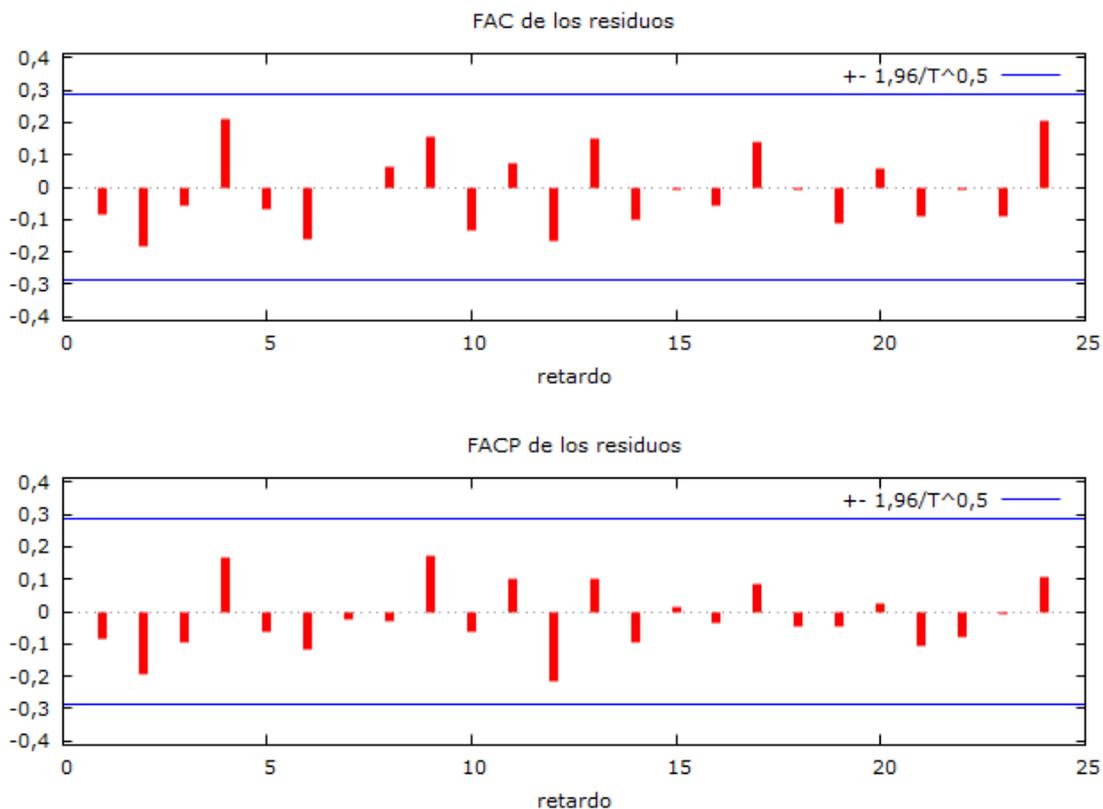
Notas: Las β_i son los coeficientes parciales del modelo, la ϕ_1 es el coeficiente de un AR(1) estacional en el modelo y las δ_i son los retardos de las diferencias del término “Majorca” (Δ GT) y el subíndice i se refiere al número del retardo. Los asteriscos indican que el coeficiente es significativo al 1% de significación con tres asteriscos y al 10% de significación con un asterisco.

Ahora que se ha obtenido el modelo se aplica una vez más el contraste causalidad de Granger (1969), donde la hipótesis nula se obtiene igualando los coeficientes de los retardos de Δ GT_MAJ, en concreto, δ_2 y δ_3 a cero:

$$H_0: \delta_2 = \delta_3 = 0,$$

frente a la alternativa al menos uno de los coeficientes ($\delta_{2,3}$) es distinto de cero. En este caso el resultado del p -valor del contraste es 0,095 por lo que se rechaza la hipótesis nula, la cual indica que no existe una causalidad de tipo Granger al 10% de significación. Esto significa que un cambio unitario en el índice de GT “Majorca” de hace 2 meses impactaría en aproximadamente 2494 turistas, mientras que un cambio unitario en el índice con tres meses de antelación resultaría en 3226 turistas, tal y como indicado en la Tabla 4.

Figura 3. Correlograma de los residuos del modelo ARDL entre las llegadas turísticas y Δ GT para “Majorca”



Fuente: Elaboración propia, datos extraídos de IBESTAT y Google Trends.

Nuestros resultados coinciden con los de Choi & Varian (2012), Pan et al. (2012), Bangwayo-Skeete & Skeete (2015), Gunter & Önder (2016), Rivera (2016), Park et al. (2017), Önder (2017), Bokelmann & Lessmann (2019), a la hora de predecir las llegadas de turistas, por lo que el índice de Google Trends es útil para predecir la demandas turística.

Si comparamos el índice de búsquedas GT con el del año anterior, este ha caído 32 puntos, siendo en abril de 2019 un valor de 70 y este año en el mismo mes 38, una bajada de casi un 50% respecto al año anterior posiblemente a causa de la crisis epidemiológica del coronavirus. Por tanto, siguiendo el modelo para el término “Mallorca” descrito en este estudio, cada cambio unitario en el índice de GT impactaría en las llegadas en 5221 turistas aproximadamente con un retardo

de 2 meses, es decir, que para junio de este año se espera una reducción de 198398 turistas aproximadamente.

4. Conclusiones

El objetivo de este estudio es demostrar la utilidad de Google Trends a la hora de realizar predicciones sobre la demanda turística de Mallorca, ya que siendo uno de los destinos principales de Europa para veranear es útil desarrollar un modelo de predicción turística que no se base exclusivamente en las llegadas turísticas del año anterior.

Por ello en este estudio se han desarrollado dos modelos autorregresivos con retardos distribuidos (ARDL) para dos términos distintos “Mallorca” y “Majorca” para distinguir a los principales consumidores del turismo balear, los alemanes y los británicos. Utilizando diferentes contrastes se ha comprobado el correcto ajuste del modelo, la inexistencia de raíces unitarias evitando así problemas de inferencia y el test de causalidad de Granger confirmando así utilidad de los retardos de los índices de Google Trends a la hora de predecir la demanda turística de Mallorca.

Tal y como mencionan Gunter & Önder (2015), los productos turísticos son perecederos y no almacenables, por este motivo es tan importante disponer de modelos que ayuden a obtener predicción de la demanda turística lo más precisa posible para que así los agentes turísticos podrán hacer ajustes con anterioridad para amoldar su oferta y utilizar sus recursos de la manera más eficiente posible e incluso poder hacer ajustes a corto plazo para maximizar el beneficio.

Bibliografía

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Álvarez-Díaz, M., & Rosselló-Nadal, J. (2010). Forecasting British Tourist Arrivals in the Balearic Islands Using Meteorological Variables. *Tourism Economics*, 16(1), 153–168.
- Bangwayo-Skeete, P. F., & Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, 454–464.
- Bi, J. W., Liu, Y., & Li, H. (2020). Daily tourism volume forecasting for tourist attractions. *Annals of Tourism Research*, 83(September 2019), 102923.
- Bing, P., Doris, C. W., & Haiyan, S. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality and Tourism Technology*, 3(3), 196–210.
- Bokelmann, B., & Lessmann, S. (2019). Spurious patterns in Google Trends data - An analysis of the effects on tourism demand forecasting in Germany. *Tourism Management*, 75(September 2018), 1–12.
- Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record*, 88(SUPPL.1), 2–9.
- de Oliveira Santos, G. E. (2009). Research Note: Forecasting Tourism Demand by Disaggregated Time Series – Empirical Evidence from Spain. *Tourism Economics*, 15(2), 467–472.
- Dickey, D., & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *JASA. Journal of the American Statistical Association*, 74.
- Doblemente. (2015). ¿Por qué los turistas prefieren hacer sus reservas en las agencias de viaje online? <https://www.doblemente.com/por-que-los-turistas-prefieren-hacer-sus-reservas-en-las-agencias-de-viaje-online/>
- Europa Press. (2018). *El turista es digital: el 70% de las reservas de hotel ya se hacen online.*
<https://www.eleconomista.es/tecnologia/noticias/8883498/01/18/EI-70-de->

las-reservas-de-hotel-se-hicieron-online-en-2017-segun-Google-Espana.html

- Garin-Munoz, T., & Amaral, T. P. (2000). An econometric model for international tourism flows to Spain. *Applied Economics Letters*, 7(8), 525–529.
- Google. (2020). *Preguntas frecuentes sobre los datos de Google Trends*.
https://support.google.com/trends/answer/4365533?hl=es&ref_topic=6248052
- Google Trends. (2020). *Comparativa entre “Mallorca” y “Majorca.”*
<https://trends.google.es/trends/explore?cat=67&q=Majorca,Mallorca>
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424–438.
- Gunter, U., & Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tourism Management*, 46, 123–135.
- Gunter, U., & Önder, I. (2016). Forecasting city arrivals with Google Analytics. *Annals of Tourism Research*, 61, 199–212.
- Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2), 163–172.
- Li, X., & Law, R. (2019). Forecasting Tourism Demand with Decomposed Search Cycles. *Journal of Travel Research*, 59(1), 52–68.
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- NetMarketShare. (2020). *Search Engine Market Share*.
<https://netmarketshare.com/search-engine-market-share.aspx?options=%257B%2522filter%2522%253A%257B%2522%2524and%2522%253A%255B%257B%2522deviceType%2522%253A%257B%2522%2524in%2522%253A%255B%2522Desktop%2524laptop%2522%255D%257D%257D%255D%257D%252C%2522date>
- Önder, I. (2017). Forecasting tourism demand with Google trends: Accuracy comparison of countries versus cities. *International Journal of Tourism Research*, 19(6), 648–660.

- Park, S., Lee, J., & Song, W. (2017). Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data. *Journal of Travel and Tourism Marketing*, 34(3), 357–368.
- Rivera, R. (2016). A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. *Tourism Management*, 57, 12–20.
- Rodríguez, X. A., & Rivadulla, R. (2012). Tourism in Spain: disaggregated analysis of the international demand. *Catalonia*, 25, 21–25.
- Rossello-Nadal, J. (2001). Forecasting Turning Points in International Visitor Arrivals in the Balearic Islands. *Tourism Economics*, 7(4), 365–380.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.*, 6(2), 461–464.
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting-A review of recent research. *Tourism Management*, 29(2), 203–220.
- Xie, G., Li, X., Qian, Y., & Wang, S. (2020). Forecasting tourism demand with KPCA-based web search indexes. *Tourism Economics*.