



Nowcasting y Proyecciones de Corto y Mediano Plazo de la Dinámica del Sector Turístico de la República Dominicana: Un enfoque basado en Probabilidades de *Google Trends* y un Modelo Bayesiano Estructural de Series de Tiempo

**Lisette Josefina Santana Jiménez**

**INTEC**

**Instituto Tecnológico de Santo Domingo**

Febrero 2017

# ***Nowcasting y Proyecciones de Corto y Mediano Plazo de la Dinámica del Sector Turístico de la República Dominicana: Un enfoque basado en Probabilidades de Google Trends y un Modelo Bayesiano Estructural de Series de Tiempo***

Lisette Josefina Santana Jiménez<sup>1</sup>

## **Resumen**

El turismo es considerado como la espina dorsal de la economía dominicana, así como también la actividad que genera los vínculos más relevantes con el resto de los sectores productivos. El objetivo primario de esta investigación consiste en pronosticar la dinámica futura del sector turístico de la República Dominicana, medida en términos de la potencial llegada de pasajeros no residentes desde los países que históricamente han exhibido el mayor influjo de visitantes: Estados Unidos de América, Canadá y Alemania, a fin de poder contar con información que pueda ser incorporada en los procesos de toma de decisiones, tanto de corto como de mediano plazo. Se llevan a cabo procesos de *nowcasting*, proyecciones de corto y mediano plazo utilizando datos procedentes de la plataforma *Google Trends* y usando un Modelo Bayesiano Estructural de Series de Tiempo con un tipo especial de prior denominado "*Spike & Slab*". El procedimiento empleado exhibe una significativa minimización del error de pronóstico, medido a través del error cuadrático medio (ECM), así como también una mejora en la bondad de ajuste, en comparación con otros modelos como estructuras Estado-Espacio y Vectores Autoregresivos (VAR) de Frecuencia Mixta, los cuales son usados como puntos de comparación para evaluar el desempeño del modelo seleccionado. Los resultados obtenidos llevan a la conclusión de que el manejo de grandes conjuntos de datos, como la información procedente de *Google Trends*, a través de modelos especializados como el BSTS ayuda a mejorar la precisión de los resultados. La importancia y el valor de estas herramientas continúan en ascenso y su desempeño refleja su potencial para coadyuvar los procesos de toma de decisiones.

***Palabras Claves:*** Turismo, Google Trends, Nowcasting, Turismo Sostenible.

***Clasificación JEL:*** L8, C02, C11, C14, C6.

---

<sup>1</sup> Departamento de Ciencias Básicas, Docente Maestría de Matemática Aplicada.

## **I. Introducción**

La actividad turística en la República Dominicana representa aproximadamente un 8% del Producto Interno Bruto (PIB) del país. Este sector de la economía es considerado como una de las principales fuentes de generación de empleo; su aporte, en términos monetarios, ascendió a US\$6,500 millones en el año 2016<sup>2</sup>. Durante los últimos tres años, la dinámica del sector turístico ha experimentado un crecimiento promedio del orden de 8.7% y la llegada de no residentes durante el año 2016 alcanzó la cantidad de seis millones de turistas, siendo su contribución al PIB de un 6%, en el último trimestre del 2016.

Dadas estas consideraciones, el turismo se ha posicionado como la espina dorsal de la economía dominicana y también como la actividad que genera los más importantes vínculos con el resto de los sectores productivos. El efecto multiplicador de la demanda de turismo en la evolución de las actividades relacionadas, como el mercado laboral, transporte, acomodamiento, bebidas y alimentos, y la construcción de nueva infraestructura (*e.g.* hoteles, vías de comunicación, entre otros) es también un elemento que sustenta esta aseveración.

Considerando la trascendencia de la actividad turística en esta economía, la evaluación de su trayectoria futura se plantea como una de las dimensiones esenciales en el marco de la toma de decisiones, lográndose con esto una minimización de costos y una locación de recursos más eficiente, manejándolos de manera más apropiada de acuerdo a los picos o caídas que puedan predecirse en la cantidad de turistas que arriban al país.

A fin de determinar el impacto de la actividad turística en la economía, resulta necesario aproximar el influjo prospectivo de turistas al país a nivel de los horizontes temporales que podrían resultar relevantes para estos fines. No obstante, uno de los problemas que se

---

<sup>2</sup> Informe Trimestral de la Economía, Banco Central de la Republica Dominicana, Diciembre 2016.

visualizan al pensar en llevar a cabo esta tarea es que varias de las variables macroeconómicas claves de interés son publicadas con rezagos o con baja frecuencia.

El Banco Central de la República Dominicana, bajo su estricta política de transparencia, hace un continuo y arduo esfuerzo por reforzar y mejorar la disponibilidad de las estadísticas de acceso público; sin embargo, la posibilidad de manejar un espectro de datos más amplio, disponible en una frecuencia alta y eficaz, resulta más que atractivo, pues enriquecería significativamente el acervo de información relevante para la elaboración de los informes que servirían de base para la toma de decisiones.

La literatura económica moderna presenta una alternativa al problema de falta de información periódica y de alta frecuencia, la cual está basada en datos que provienen de grandes compañías de internet que han invertido en el procesamiento y almacenamiento de información; este es el caso de Google con su plataforma *Google Trends*. El fenómeno creciente de *big data* ha revolucionado la teoría económica en los últimos años, explotando las ventajas de los “conjuntos modernos de datos”, que capturan el *rastreo* de los usuarios, poniendo esta información a disposición del público.

El potencial de *Google Trends* para mejorar la precisión de los modelos económicos es explorado en la literatura reciente (Varian & Choi, 2012; Bortoli & Combes, 2015), conjuntamente con el surgimiento de nuevas técnicas para analizar y hacer inferencia en base a esta metadata. Dadas estas consideraciones, así como también el hecho de que el turismo es una de las actividades de mayor búsqueda intensiva en el internet<sup>3</sup>, tanto para obtener información sobre los paquetes de viaje como para adquirir dichos paquetes, es

---

<sup>3</sup> De acuerdo al portal de estadísticas StatisticBrain.com, aproximadamente un 60% de las reservaciones de viajes son realizadas vía internet.

lógico pensar que esta información sería útil para pronosticar la demanda de servicios turísticos<sup>4</sup>.

Este estudio pretende verificar si el uso de la información procedente de *Google Trends* conjuntamente con un modelo económico basado en métodos de *machine learning*, en este caso el Modelo Bayesiano Estructural de Series de Tiempo (BSTS), resultan útiles para derivar señales sobre el comportamiento futuro del sector turístico, para una economía emergente como la de República Dominicana, lo cual es relevante para hacer inferencia sobre la proporción del ingreso generado por este segmento particular.

Las proyecciones se realizan siguiendo la metodología propuesta por Scott & Varian (2014), la cual consiste de un modelo Bayesiano Estructural de Series de Tiempo (BSTS, por sus siglas en inglés). Este modelo emplea un tipo particular de *prior* denominado “*Spike & Slab*”, el cual castiga regresores irrelevantes y espurios. El enfoque de Scott y Varian (2014) es extendido para lograr proyecciones de mediano plazo, a través de métodos iterativos en los cuales se calculan las predicciones para el horizonte  $h$  y se retroalimenta el modelo con los resultados de dichas proyecciones para generar la proyección de  $h + 1, \dots, h + n$ .

Las simulaciones realizadas en base al modelo BSTS reflejan una significativa minimización del error de pronóstico, medido en términos del error cuadrático medio (ECM), en ejercicios realizados fuera de muestra, con relación al desempeño de los modelos alternativos, *i.g.* VAR de Frecuencia Mixta y Modelo Estado-Espacio, así como también un incremento en la bondad de ajuste del modelo, evaluado en función del  $R^2$ .

---

<sup>4</sup> Se ha reportado que un número significativo de la cantidad total de búsquedas en la web están relacionadas a viajes (Jansen et al, 2008), y esta cantidad continúa en ascenso.

El resto de este trabajo está estructurado de la manera siguiente: en la sección subsiguiente se presenta una revisión de la literatura teórica y empírica concerniente al desenvolvimiento del sector turístico; en la Sección III se esbozan los aspectos asociados al proceso de compilación de datos utilizados en esta investigación; en la Sección IV se discuten las técnicas metodológicas empleadas para llevar a cabo el *nowcasting* y las proyecciones de corto y mediano plazo. En la Sección V se presentan los resultados esta investigación y, finalmente, las conclusiones y recomendaciones de política.

## **II. Revisión de Literatura**

El turismo se ha constituido como el sector líder en el segmento de los servicios en los pequeños estados insulares en desarrollo (SIDS, por sus siglas en inglés), con alto potencial para crear un valor agregado significativo no sólo para el establecimiento de estructuras directamente vinculadas a dicho sector, sino también para la sociedad como un todo y para el desarrollo y sostenibilidad económica y social en el largo plazo (Craigwell, 2007; Tisdell & Bandara, 2004).

La literatura relacionada al estudio del sector turístico se ha inclinado esencialmente al análisis de las relaciones de dicho segmento de la economía con el crecimiento económico y, en los últimos años, a la evaluación de políticas y medidas encauzadas a garantizar la sostenibilidad de esta industria, por medio de campañas de respeto al medio ambiente, satisfacción de la demanda y de las preferencias heterogéneas de los consumidores, así como también publicidad orientada a enaltecer los aspectos idiosincráticos de cada región.

Eugenio-Martin *et al* (2004) llevan a cabo un estudio usando datos de panel, a fin de estudiar la relación entre el desempeño del sector turístico y el crecimiento económico para un conjunto de países latinoamericanos, partiendo de la premisa de que en la mayoría de los países en desarrollo los turistas usualmente demandan cuatro tipos de bienes y servicios: alojamiento, alimentos, facilidades de transporte y entretenimiento, de manera que, para satisfacer esta demanda, es necesario aumentar la producción, lo que tiene dos

efectos positivos en la economía: en primer lugar, un incremento en el ingreso y, en segundo lugar, un aumento en el nivel de empleo, dado que el turismo es trabajo-intensivo. Los autores llegan a la conclusión de que el crecimiento per-cápita en los países considerados guarda una alta correlación con el desarrollo del sector turismo y que factores tales como la infraestructura y seguridad del país anfitrión juegan un rol trascendente en la decisión de los visitantes al momento de viajar.

En lo concerniente a la sostenibilidad del turismo, desde la óptica medio-ambiental (Ogonowska & Torre, 2012), es importante señalar que, a partir de la década de los 90's, el tema de la protección al medio ambiente ha adquirido cada vez mayor preponderancia y, tanto los consumidores como los países anfitriones, se han vuelto más conscientes de los problemas que se pueden derivar del turismo masivo sin las regulaciones y normas pertinentes. Consecuentemente, ha surgido un nuevo segmento de turistas con preferencias matizadas por el deseo de proteger el medio ambiente y con predilección por entornos tranquilos en que puedan entrar en contacto directo con la naturaleza. Esta modificación de preferencias ha llevado a los proveedores de servicios turísticos a evolucionar y a adaptar su oferta en congruencia con los estándares requeridos.

En adición a los temas relacionados a la sostenibilidad del turismo y a las preferencias, otro de los aspectos que se encuentran en la palestra en este momento está vinculado al turismo accesible y a su impacto económico, entendiéndose como turismo accesible el conjunto de facilidades e infraestructuras que atiendan las necesidades de las personas con discapacidad física, pues de acuerdo a estadísticas de la Organización Mundial del Turismo (OMT), un destino que no atienda apropiadamente las necesidades de este segmento de la población excluye un prometedor mercado (Figura 1). En palabras de Luis Felipe Aquino, Presidente de la Bolsa Turística del Caribe (BTC), “la República Dominicana tiene una gran oportunidad en el turismo accesible y por tanto es un nicho que se debe explotar”.

Bajo el esquema de una economía sustentada primordialmente en la prestación de servicios turísticos, uno de los objetivos y preocupaciones de los tomadores de decisiones consiste en maximizar la cantidad de turistas que llegan al país y su nivel de gasto,

propiciando las condiciones e instalaciones necesarias para materializar dicho propósito. La sostenibilidad de la demanda de turismo, entendida como un tipo particular de demanda, en la cual “el producto turismo” se concibe como un vector de bienes complementarios y servicios (Song *et al*, 2010) está supeditada esencialmente a la satisfacción de las necesidades reales de los visitantes, en términos de preferencias, comodidades, servicios y facilidades provistas.

Dadas las particularidades y divergencias en las predilecciones de servicios turísticos entre los diferentes agentes económicos, así como también la naturaleza perecedera de los productos ofertados en este sector, tales como asientos de líneas aéreas, habitaciones de hoteles, alimentos, apartamentos, servicios de automóviles que no son utilizados, entre otros, la proyección de la demanda de turismo es el tema al cual los economistas prestan mayor atención (Witt & Witt, 1995), de manera que la generación de predicciones exactas sobre la llegada de turistas es reconocida como primordial para asistir la toma de decisiones.



*\*Fuente: Google Trends y elaboración propia.*

La literatura empírica tradicional para llevar a cabo proyecciones de demanda de turismo es amplia y está fundamentalmente basada en el análisis de series de tiempo (Balaguer & Cantavella-Jordá, 2010), siendo las principales variables consideradas los ingresos, costos de viaje, tipo de cambio, costos en el destino de viaje y un componente de persistencia (capturado por un proceso autoregresivo).

No obstante, la falta oportuna de información para llevar a cabo dichas proyecciones es uno de los problemas que se plantean para inferir decisiones a partir de estos modelos. La literatura económica moderna presenta una alternativa al problema de falta de información periódica y de alta frecuencia, la cual está basada en datos que provienen de grandes compañías de internet que han invertido en el procesamiento y almacenamiento de información; este es el caso de Google con su plataforma *Google Trends*.

La necesidad de generar proyecciones cada vez más acertadas para la demanda de turismo y asistir los procesos de toma de decisiones, busca lidiar con la naturaleza perecedera de muchos de los bienes de esta industria (Archer, 1987), *e.g.* la caducidad de reservaciones de asientos de aerolíneas, apartamentos o habitaciones de hoteles que no se ocupen, entre otros. Para el caso del sector turístico, diversos horizontes temporales y niveles de agregación resultan relevantes para la planificación (Witt & Witt, 1995). En este sentido, el *nowcasting* y las proyecciones de corto plazo resultan apropiados para la organización del staff, marketing y fijación de precios, mientras que las proyecciones de mediano y largo plazo resultan importantes desde la óptica de la planeación de inversión en facilidades viales, acomodación, reacondicionamiento de ambientes de acuerdo a nuevas preferencias, entre otros.

En este contexto, surge la siguiente pregunta: ¿Pueden los datos de Google Trends coadyuvar a predecir el comportamiento de variables macroeconómicas para una economía emergente, en este caso la llegada de turistas a la República Dominicana, empleando a la vez un modelo especializado para lidiar con grandes conjuntos de datos que permita depurar regresores espurios y no significativos?

Para responder esta pregunta, resulta pertinente remontarse a dos tipos de limitaciones: en primer lugar, las restricciones de la modelación tradicional, que no necesariamente contemplan el tratamiento de grandes conjuntos de datos, los cuales requieren modelos y algoritmos de la teoría de *machine learning*; por otra parte, la falta de información oportuna en muchos casos constituye un problema, ya que las entidades encargadas de divulgar las variables necesarias para llevar a cabo estudios económicos hacen publicaciones con rezagos y no necesariamente son series completas.

La toma de decisiones oportunas para trazar las políticas de inversión deben basarse en información precisa, por lo que es pertinente contar con metodologías que permitan orientar de manera efectiva las inversiones en el sector turístico, en función de la demanda que se generaría dadas las variaciones esperadas en el número de visitantes al territorio dominicano.

Las técnicas cimentadas en *machine learning* para el manejo de *Big Data* se plantean como herramientas con gran potencial para este propósito, permitiendo evaluar la trayectoria esperada de las variables macroeconómicas, empleando información disponible de manera oportuna y que, indudablemente, ayudarían a fortalecer los procesos de toma de decisiones.

Para los académicos, el término *Big Data* se refiere a una “colección de datos con complejidad, diversidad, heterogeneidad y alto potencial” (Hassani & Silva, 2015). El fenómeno creciente de *big data* ha revolucionado la teoría económica en los últimos años, explotando las ventajas de los “conjuntos modernos de datos”, que capturan el *rastreo* de los usuarios, poniendo esta información a disposición del público. La literatura empírica y teórica en este ámbito es amplia y muestra cómo los datos de búsqueda de internet constituyen una fuente valiosa para descubrir nuevas tendencias así como también patrones de comportamiento.

En un reciente documento (Bholat, 2016) analiza la trascendencia del fenómeno de *Big Data* desde la óptica de los bancos centrales, explicando cómo los nuevos métodos para

obtener y analizar datos no sólo suponen un incremento cuantitativo en los datos, sino que también representarían un nuevo enfoque para analizar el sistema económico y financiero. La clave para los bancos centrales es entender que aceptar el fenómeno de *big data* como parte de sus sistemas de modelación no necesariamente requiere que los bancos amplíen su colección de datos, sino que puedan hacer más con los datos “nanoeconómicos” disponibles.

Las investigaciones pioneras en esta línea de trabajo (Varian & Choi, 2008; Carriere-Swallow & Labbe, 2010) han llevado al surgimiento de nuevas tendencias que emplean el enfoque de Google Trends como una fuente primaria para obtener información de búsquedas de internet, demostrando cómo esta plataforma se ha posicionado como una valiosa herramienta con alto potencial para mejorar las proyecciones de variables.

Guzmán (2011) emplea datos de Google Trends para pronosticar expectativas de inflación, Artola & Galán (2012) para generar predicciones del turismo en España, a nivel de provincias. Choi & Varian (2011) presentan uno de los primeros trabajo que hacen *nowcasting* basado en Google Trends para variables tales como compra de motores de vehículos, partes de autos y beneficios de desempleo en Estados Unidos. Se encuentra que un simple modelo AR que incluye información exógena de Google Trends tiende a superar el desempeño de otros modelos que no contemplan la inclusión de dichos regresores en un 5% a un 10%.

Kulkarni *et al* (2009) demuestran cómo los datos de búsqueda de Google proveen una manera certera de predecir el comportamiento de actividades financieras; estos autores aplican una metodología para proyectar la tendencia de los precios de mercado de las viviendas. El uso de probabilidades de Google Trends genera proyecciones fuera de muestra con un error cuadrático medio menor que el modelo base que emplea datos convencionales sin considerar los datos de Google Trends. Se encuentra que el modelo utilizado genera mejores resultados que los de la *National Association of Realtors* en un 23.6% para futuras ventas de casas en Estados Unidos. Del mismo modo, se demuestra

cómo esta dinámica puede ser trasladada a otros mercados como la venta de accesorios del hogar.

Schmidt & Vosen (2009) analizan el comportamiento del consumo privado basándose en series de tiempo provistas por Google Trends; construyen un indicador fundado en factores extraídos de búsquedas relacionadas al consumo. El desempeño del nuevo indicador en términos de su poder predictivo es examinado con respecto a los indicadores basados en encuestas más comúnmente usados: el Indicador de Sentimiento del Consumidor de la Universidad de Michigan y el Índice de Confianza del Consumidor del *Conference Board*. Los resultados muestran que prácticamente en todos los experimentos conducidos tanto fuera como dentro de muestra el indicador de Google Trends tiene mejor desempeño que los índices contruidos en base a encuestas. Estas conclusiones sugieren que la incorporación de la información de Google Trends ofrece beneficios significativos para la proyección del consumo privado.

Askitas & Zimmermann (2009) demuestran cómo la información procedente de Google Trends resulta particularmente útil en tiempos de crisis económica, donde el flujo tradicional de información suele ralentizarse y por tanto no es tan simple proveer una base propia para generar decisiones económicas; los autores analizan el potencial de esta información en la predicción del comportamiento económico bajo circunstancias cambiantes y complejas, observándose una fuerte correlación entre el desempleo y el indicador construido a partir de *Google Trends*.

En lo concerniente a las técnicas de modelación empleadas para el manejo de esta data, el enfoque Bayesiano se presenta como una alternativa natural y apropiada, considerando su eficacia para la selección de las variables explicativas. Brodersen *et al* (2015) proponen un modelo Bayesiano estado-espacio para obtener predicciones contrafactuales de la actividad de mercado; los resultados muestran que el modelo seleccionado ayuda a sobrellevar algunos de los problemas que presentan los modelos estadísticos tradicionales, los cuales adolecen de ciertas características esenciales y por tanto no resultan apropiados para predecir procesos estocásticos.

Ludwig et al (2015) analizan el desempeño de un Modelo Bayesiano Estructural de Series de Tiempo, aplicado a la predicción de los precios de la electricidad, demostrando que este enfoque es apropiado para la selección de variables en el contexto de grandes conjuntos de regresores.

Scott & Varian (2014) presentan un sistema para hacer nowcasting de series económicas, basándose en datos de Google Trends y Google Correlates, combinando modelos estructurales de serie de tiempo con técnicas bayesianas con priors “Spike & Slab”. Los autores hacen aplicaciones para el caso del desempleo, turismo y ventas minoristas, llegando a la conclusión de que los predictores de Google Trends no solamente son apropiados para hacer *nowcasting*, sino también para predecir puntos de giro en la serie de tiempo.

Para las proyecciones de mediano plazo y largo plazo, la propuesta de *machine learning* se centra en tres técnicas esenciales (Huijskens, 2016; Taieb, 2014) : Método de Proyección Iterativo, Método de Proyección Directo de H-Pasos, Modelo de Múltiples Insumos y Múltiples Productos. La selección de un modelo de proyección para un horizonte de mediano o largo plazo va a depender de la naturaleza del problema, así como también considerando si la meta final es explicativa o predictiva.

### **III. Datos**

Los datos procedentes de la plataforma *Google Trends* fueron la fuente primaria de información empleada para esta investigación. Esta plataforma provee series de tiempo dadas como índices que miden el volumen de búsquedas llevadas a cabo por los usuarios en determinada área geográfica. Los índices son normalizados y escalados de manera tal que el valor máximo sea 100 y el valor mínimo 0, de acuerdo a la popularidad relativa de determinado tópico en un intervalo de tiempo. El período de tiempo considerado para esta investigación es de Diciembre 2012 a Octubre 2017, con datos de *Google Trends* capturados con frecuencia semanal; estos datos son publicados con un rezago de dos días. Las áreas geográficas, en este caso, están representadas por los países que históricamente han

exhibido el mayor flujo de turistas hacia República Dominicana: Estados Unidos, Canadá y Alemania (Figura 2, Anexo I).<sup>5</sup>

El proceso de selección de variables para cada país se efectuó en base a dos criterios: en primer lugar, se construyó un conjunto de palabras claves relacionadas a las posibles búsquedas que podrían realizar los turistas. En este sentido, Jansen *et al* (2008) apuntan que las búsquedas vinculadas a locaciones geográficas, tales como nombres de ciudades, países y regiones representan aproximadamente un 60% de las búsquedas relacionadas a viajes; otras variables consideradas para cada país están asociadas con las búsquedas de alojamiento, alquiler de autos, paquetes de viaje, compra de tickets, actividades culturales y entretenimiento.

En segundo lugar, la herramienta *Google Correlates* fue empleada para realizar búsquedas más meticulosas que contemplasen aspectos idiosincráticos inherentes a cada país considerado. *Google Correlates* busca millones de candidatos, a fin de identificar aquellas búsquedas que presenten un patrón de comportamiento similar al de la serie objetivo.

Con el objetivo de establecer estas similitudes entre vectores de información, se emplea el coeficiente de correlación de Pearson, el cual tiene la particularidad de ser invariable bajo transformaciones lineales. El método usado para computar la distancia aproximada entre vectores es el algoritmo de *Asymmetric Hashing*.<sup>6</sup> El primer paso del algoritmo es mapear la función de correlación de Pearson en una distancia Euclidiana, a través de una normalización apropiada de los vectores; el segundo paso consiste en un fase de “entrenamiento e indexación”, en la cual se asume una colección de vectores  $V = v^{(1)}, v^{(2)}, \dots, v^{(M)}$  en  $\mathbb{R}^N$ . Estos vectores son divididos en particiones k-dimensionales compuestas de dimensiones consecutivas en cada vector. Un conjunto de proyecciones  $\pi_1, \dots, \pi_{N/k} \in \mathbb{R}^N \mapsto \mathbb{R}^k$  desde el vector completo a las diferentes particiones; para cada una de estas proyecciones se emplea el algoritmo de k-vecindad (Duda *et al*, 2000) para encontrar los centroides. Finalmente, se completa el proceso de búsqueda de palabras

---

<sup>5</sup> Información procedente del Informe Trimestral de la Economía, Banco Central de la República Dominicana.

<sup>6</sup> Véase Vanderkam *et al* (2013).

claves en función de la distancia euclidiana entre el vector objetivo y la base de datos reconstruida.

Los datos mensuales de llegada de turistas procedentes de Estados Unidos, Canadá y Alemania se obtuvieron del portal del Banco Central de la República Dominicana; estas series fueron trabajadas con una transformación logarítmica, a fin de suavizar su comportamiento y posteriormente emplearlas como insumos en la plataforma *Google Correlates* para encontrar términos adicionales que mantuviesen un alto grado de correlación con estos datos.

#### IV. Metodología

Para llevar a cabo el ejercicio de *nowcasting* se emplea el modelo propuesto por Scott & Varian (2014), el cual está compuesto de un módulo que se encarga de capturar los patrones estacionales y la tendencia de los datos y otro módulo consistente de un componente de regresión a través del cual se incorpora la información exógena procedente de *Google Trends*.

##### 4.1. Módulo Estructural

Denótese  $y_t$  como la observación en  $t$  de una serie de tiempo observada en tiempo real. Un modelo estructural de series de tiempo puede ser descrito como un sistema consistente de un par de ecuaciones que relacionan  $y_t$  a un vector de variables latentes de estado  $\alpha_t$ :

$$y_t = Z_t^T \alpha_t + \varepsilon_t \quad \varepsilon_t \sim N(0, H_t) \quad (1)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim N(0, Q_t). \quad t = 1, \dots, n, \quad (2)$$

La ecuación (1) se denomina ecuación de observación, ya que permite establecer un vínculo entre los datos observados contenidos en el vector  $y_t$  con el estado latente no observado  $\alpha_t$ . La ecuación (2) se denomina ecuación de transición, ya que define cómo el estado latente evoluciona a través del tiempo. La variable latente  $\alpha_t$  no puede ser observada directamente, y por tanto el análisis se enfoca en las observaciones  $y_t$ . Se asume que los estados se rigen por un modelo Markoviano de transición.

Se supone que inicialmente las matrices  $Z_t, T_t, R_t, H_t$  son conocidas y que los términos de error, representados por  $\varepsilon_t, \eta_t$ , son serialmente independientes, así como también independientes uno de otro en todos los puntos del tiempo. Las matrices  $Z_t$  y  $T_{t-1}$  pueden depender de  $y_1, \dots, y_{t-1}$ . Se espera que el vector inicial de estado  $\alpha_1$  sea  $N(\alpha_1, P_1)$  e independiente de  $\varepsilon_1, \dots, \varepsilon_n$  y de  $\eta_1, \dots, \eta_n$ , asumiéndose que se conocen  $\alpha_1$  y  $P_1$ . La ecuación (1) tiene la estructura de un modelo de regresión lineal donde el coeficiente de vectores  $\alpha_t$  es dinámico. La expresión (2) representa un Vector Autoregresivo (VAR) de primer orden.

Para los fines de esta investigación, se considera la especificación estado espacio sugerida por Scott & Varian (2014):

$$y_t = \mu_t + \tau_t + \beta^T x_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \delta_{t-1} + u_t \quad (3)$$

$$\delta_t = \delta_{t-1} + v_t$$

$$\tau_t = -\sum_{s=1}^{S-1} \tau_{t-s} + \omega_t$$

Donde  $\eta_t = (u_t, v_t, \omega_t)$  contiene componentes independientes de ruido Gaussiano. En este caso,  $Q_t$  es una matriz diagonal con los elementos  $\sigma_\varepsilon^2$ ,  $\sigma_v^2$  y  $\sigma_\omega^2$  en su diagonal principal;  $H_t$  es una constante  $\sigma_\varepsilon^2$ .

El nivel de la tendencia es  $\mu_t$  y la pendiente es  $\delta_t$ . El componente estacional  $\tau_t$  puede ser representado como un conjunto de variables con coeficientes dinámicos que se ajustan a la restricción de tener un valor esperado equivalente a lo largo de un ciclo de  $S$  estaciones. Los parámetros en el sistema (3) son las varianzas  $\sigma_\varepsilon^2, \sigma_u^2, \sigma_v^2, \sigma_\omega^2$  y el vector de coeficientes de regresión  $\beta$ , el cual está representado por el conjunto de búsquedas procedentes de *Google Trends*.

El filtro de Kalman (Kalman, 1960) y el Kalman *smoother* (Fahrmeir & Kaufmann, 1991; Yu et al, 2004, Briers *et al*, 2010 ) son las técnicas más usadas para el tratamiento de modelos estado espacio como el sistema (3). No obstante, cuando no es plausible asumir normalidad y linealidad, no es posible generar computaciones exactas basadas en estos enfoques; estos métodos resultan ineficientes en problemas de alta dimensión (Pnevmatikakis et al, 2012). En este caso, resulta necesario apelar a otras metodologías que contemplan supuestos no Gaussianos. En este sentido, se han propuesto diversos enfoques (Fahrmeir and Wagenpfeil, 1997; Fahrmeir and Tutz, 1994, Jungbacker & Koopman, 2007), que incorporan métodos de suavizamiento bayesiano y simulaciones basadas en Cadenas de Markov con simulaciones Monte Carlo.

La computación fundamentada en métodos de aumento bayesiano de datos (Neal & Kypraios, 2015; Taylor *et al*, 2015) produce simulaciones a partir de  $p(\alpha_t | \mathbf{y})$ , donde  $\mathbf{y} = y_{1:n}$  y  $\alpha = \alpha_{1:n}$

denotan el conjunto completo de variables observadas y latentes; esta alternativa surge dado que no es posible hacer *draws* de cada  $\alpha_t$ , a partir de  $p(\alpha_t|\mathbf{y})$ , ya que la correlación serial entre  $\alpha_t$  y  $\alpha_{t+1}$  debe ser respetada. El enfoque bayesiano es inherente a una teoría en la que los parámetros de estado son variables aleatorias y las observaciones son fijas. Para el tratamiento bayesiano se requieren métodos basados en simulaciones para llevar a cabo la estimación de parámetros adicionales.

Uno de estos métodos basados en simulaciones es el algoritmo de Durbin & Koopman (2002), el cual es usado en esta investigación (de acuerdo a las sugerencias de Scott & Varian, 2014). La ventaja de este enfoque radica en el hecho de que permite lidiar con una de las limitaciones del suavizador de Kalman, extendiendo el alcance de esta herramienta para dar respuesta a cuestiones concernientes a la covarianza de variables no observables a través del tiempo. El algoritmo de Durbin y Koopman se encuentra entre las técnicas más rápidas y convenientes de implementar en términos computacionales y se resume (Jarocinski, 2015) de la manera siguiente:

**Paso 1:** Generar los *draws*  $\alpha^+$  y  $y^+$  a través de medios recursivos, a partir de (3), donde el proceso de recursión inicia con el *draw*  $\alpha_1^+ \sim N(0, P_1)$ .

**Paso 2:** Construir la serie artificial  $y^* = y - y^+$  y computar  $\hat{\alpha}^* = E(\alpha|y^*)$  pasando  $y^*$  a través del filtro de Kalman y del Kalman *smoother*.

**Paso 3:** Tomar  $\hat{\alpha} = \hat{\alpha}^* + \alpha^+$ ;  $\hat{\alpha}$  es un *draw* derivado de la distribución de  $\alpha$  condicional en  $y$ .

## 4.2. Componente de Regresión

La representación estado espacio (3) considerada para las estimaciones realizadas en esta investigación contiene un componente de regresión que permite la incorporación de un vector de factores exógenos, con potencial de mejorar la exactitud de los pronósticos.

Existen diversas formas de organizar las matrices del modelo, a fin de incluir el componente de regresión en la especificación estado-espacio. Un método apropiado (como sugieren Scott & Varian, 2014) es adjuntar una constante 1 a cada  $\alpha_t$ , y adjuntar  $\beta^T x_t$  a  $Z_t$  en la ecuación de observación; esta especificación tiene la ventaja de que solamente incrementa la dimensión del vector de estado en una unidad, independientemente de la cantidad de predictores.

Dado que la cantidad de variables procedentes de *Google Trends* y el hecho de generalmente las series económicas no son extensas, es posible que la cantidad de regresores sea mucho mayor que la cantidad de observaciones. En este sentido, el enfoque de regresión basado en *priors Spike y Slab* se plantea como una opción lógica para la estimación de los coeficientes de regresión, ya que reduce el grado de dispersión, disminuyendo el tamaño del problema de regresión y evitando caer en la *maldición de la dimensionalidad*<sup>7</sup>. Esta propiedad de los *priors Spike & Slab* para la selección de variables hace que este enfoque sea atractivo y popular en el paradigma bayesiano (Ishwaran et al, 2010).

El término *Spike & Slab* fue introducido inicialmente por Mitchel & Beauchamp (1988), haciendo referencia a un tipo particular de prior que asume que los coeficientes de regresión son mutuamente excluyentes con una distribución bimodal, siendo la idea subyacente hacer

---

<sup>7</sup> En el contexto de *machine learning*, la maldición de la dimensionalidad se refiere a cómo puede afectarse el desempeño de algoritmos de aprendizaje al lidiar con conjuntos de datos de gran dimensión.

cero los coeficientes  $\beta_K$  que no son relevantes en la regresión, asignándole valores suficientemente pequeños a la media del posterior.

Sea  $\gamma_k = 1$  si  $\beta_k \neq 0$ , y  $\gamma_k = 0$  si  $\beta_k = 0$ . Sea  $\beta_\gamma$  el subconjunto de elementos de  $\beta$ , donde  $\beta_K \neq 0$ . Un prior *Spike & Slab* puede ser expresado de la manera siguiente:

$$p(\beta, \gamma, \sigma_\varepsilon^2) = p(\beta_\gamma | \gamma, \sigma_\varepsilon^2) p(\sigma_\varepsilon^2 | \gamma) p(\gamma) \quad (4)$$

Es conveniente simplificar la ecuación (4) usando un prior independiente de Bernoulli (Scott & Varian, 2014):

$$\gamma \sim \prod_{k=1}^K \pi_k^{\gamma_k} (1 - \pi_k)^{1-\gamma_k} \quad (5)$$

La ecuación (5) puede reducirse asumiendo que todos los  $\pi_k$  tienen el mismo valor  $\pi$ ; el valor de  $\pi$  puede ser establecido en base al criterio del investigador sobre la cantidad esperada de regresores en el modelo, de manera que si se espera una cantidad  $p$  de predictores distintos de cero, entonces  $\pi = p/K$ , donde  $K$  es la dimensión de  $\mathbf{x}_t$ . Una estrategia alternativa es segmentar los predictores en grupos, asignando a todos los elementos de cada grupo la misma probabilidad de inclusión. Para el caso de una matriz simétrica  $\Omega^{-1}$ , denótese  $\Omega_\gamma^{-1}$  como las filas y columnas de  $\Omega^{-1}$  correspondientes a  $\gamma_k = 1$ . Los *priors* condicionales  $p(1/\sigma_\varepsilon^2 | \gamma)$  y  $p(\beta_\gamma | \sigma_\varepsilon^2, \gamma)$  pueden ser expresados como el par conjugado:

$$\beta_\gamma \left| \sigma_\varepsilon^2, \gamma \sim N \left( b_\gamma, \sigma_\varepsilon^2 (\Omega_\gamma^{-1})^{-1} \right) \quad \frac{1}{\sigma_\varepsilon^2} \left| \gamma \sim Ga \left( \frac{v}{2}, \frac{SS}{2} \right), \quad (6)$$

Donde  $Ga(r, s)$  denota la distribución gamma con media  $r/s$  y varianza  $r/s^2$ .

Con respecto a la posterior de  $\beta$  y  $\sigma_\varepsilon^2$ , dada  $\gamma$ , sea  $y_t^* = y_t - Z_t^{*T} \alpha_t$ , donde  $Z_t^*$  es la matriz de observación de la ecuación (1) con  $\beta^T x_t$  equivalente a cero. Sea  $\mathbf{y}^* = y_{1:n}^*$ , de manera que  $\mathbf{y}^*$  es  $\mathbf{y}$  sin considerar el componente de series de tiempo.

La densidad conjunta posterior para  $\beta$  y  $\sigma_\varepsilon^2$  condicional en  $\gamma$  está disponible a partir de las fórmulas:

$$\beta_\gamma \mid \sigma_\varepsilon, \gamma, \mathbf{y}^* \sim N(\widetilde{\beta}_\gamma, \sigma_\varepsilon^2 (V_\gamma^{-1})^{-1}) \quad \frac{1}{\sigma_\varepsilon^2} \mid \gamma, \mathbf{y}^* \sim Ga\left(\frac{N}{2}, \frac{SS_\gamma}{2}\right) \quad (7)$$

Donde las estadísticas relevantes son:

$$\begin{aligned} V_\gamma^{-1} &= (\mathbf{X}^T \mathbf{X})_\gamma + \Omega_\gamma^{-1} & \widetilde{\beta}_\gamma &= (V_\gamma^{-1})^{-1} (\mathbf{X}_\gamma^T \mathbf{y}^* + \Omega_\gamma^{-1} b_\gamma) \\ N &= v + n & SS_\gamma &= ss + \mathbf{y}^{*T} \mathbf{y}^* + b_\gamma^T + b_\gamma^T \Omega_\gamma^{-1} b_\gamma - \widetilde{\beta}_\gamma^T V_\gamma^{-1} \widetilde{\beta}_\gamma. \end{aligned} \quad (8)$$

#### 4.3. Estimación de la distribución posterior: Algoritmo de Cadena de Markov Monte Carlo

La estimación de la media de la distribución posterior en los modelos estructurales de series de tiempo puede simularse usando el Algoritmo de Cadena de Markov Monte Carlo (MCMC, por sus siglas en inglés). La idea subyacente es evaluar la media posterior de  $x(\alpha)$  del vector de parámetros  $\boldsymbol{\phi} = (\theta, \beta, \sigma_\varepsilon^2, \alpha)$ , por medio de simulaciones, seleccionando muestras a partir de

una densidad conjunta aumentada  $p(\phi, \alpha | \mathbf{Y})$ . Tras la inicialización para  $\phi$ , es decir,  $\phi = \phi^{(0)}$ , el ciclo de los siguientes pasos se repite:

1. Simular el estado latente  $\alpha$  a partir de  $p(\alpha | y, \theta, \beta, \sigma_\varepsilon^2)$ .
2. Simular  $\theta \sim p(\theta | y, \alpha, \beta, \sigma_\varepsilon^2)$ .
3. Simular  $\beta$  y  $\sigma_\varepsilon^2$  a partir de una cadena de Markov con distribución estacionaria  $p(\beta, \sigma_\varepsilon^2 | y, \alpha, \theta)$ .

Los *draws* procedentes del paso 3 pueden ser generados por medio de un algoritmo de búsqueda estocástica de selección de variables (George y McCulloch, 1997).

## V. Proyecciones

Bajo el paradigma bayesiano, las proyecciones derivadas del modelo BSTS están basadas en la distribución posterior (Scott & Varian, 2014). Sea  $\tilde{y}$  el vector de variables a ser proyectadas; la distribución posterior predictiva de  $\tilde{y}$  es:

$$p(\tilde{y} | \mathbf{y}) = \int p(\tilde{y} | \phi) p(\phi | \mathbf{y}) d\phi \quad (9)$$

Es posible realizar iteraciones en el sistema de ecuaciones (1)-(2) de  $\alpha_n^{(g)}$ , con parámetros  $\theta^{(g)}, \beta^{(g)}$  y  $\sigma_\varepsilon^{2(g)}$ , a fin de hacer un muestro a partir de  $p(\tilde{y} | \mathbf{y})$ , a través de  $p(\tilde{y}^{(g)} | \phi^{(g)})$ , dado un conjunto de *draws* aleatorios  $\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(n)}$  de  $p(\phi | \mathbf{y})$ ; considerando que diferentes elementos del vector  $\beta$  serán equivalentes a cero en diferentes *draws* de la simulación Monte Carlo, automáticamente los *draws* de la distribución posterior ayudan a corregir la dispersión; esta técnica de proyección genera una muestra de *draws* derivados de la distribución posterior

predictiva, que puede ser resumida por sus medias, que es la estimación Monte Carlo de  $E(\tilde{y}|\mathbf{y})$ .

### 5.1. Modelo Satélite

Un elemento adicional que debe ser considerado en el proceso de proyección es la construcción de un modelo o sistema de modelos que funjan como modelos satélites para generar los supuestos de pronóstico, en este caso, el comportamiento de los índices procedentes de *Google Trends*. Para fines de esta investigación se emplea un conjunto de Modelos Arima Estacionales (ESARIMA) para generar los supuestos del modelo Bayesiano Estructural, a fin de generar pronósticos en horizontes superiores a un período ( $h > 1$ ). La intención de contar con un modelo generador de supuestos es observar el desempeño del modelo central no solamente en los ejercicios de *nowcasting*, sino también en las proyecciones de corto plazo. Los modelos ESARIMA son una extensión de los modelos ARMA, que permiten incluir una dinámica específica, en este caso poder lidiar con el comportamiento estacional de la serie. La representación viene dada por la siguiente ecuación (Box & Jenkins, 1970):

$$\Phi_p(B^s)\phi(B)\nabla_s^D\nabla^d Z_t = \alpha + \Theta_Q(B^s)\theta(B)a_t \quad (10)$$

Y se denota como  $\text{ARIMA}(p, d, q) \times (P, D, Q)_s$ , donde  $\phi(B)\nabla^d$  es un operador autoregresivo generalizado y  $\nabla^d X_t$  es una cantidad que se hace estacionaria por medio de la diferenciación. La especificación del modelo ESARIMA para el caso particular de este estudio es la siguiente:

$$x_t = \alpha + \beta_1 x_{t-1} + \beta_2 x_{t-52} + \varepsilon_t \quad (11)$$

Donde  $x_t$  es un vector con el Índice de Búsquedas de *Google Trends* en el período contemporáneo  $t$ ;  $x_{t-1}$  es el primer rezago del índice y  $x_{t-52}$  representa un rezago de 52 semanas, a fin de capturar el componente estacional durante el año completo.

## **VI. Resultados**

### **6.1. Nowcasting y Corto Plazo**

Los ejercicios de *nowcasting* y las proyecciones de corto plazo realizados en el marco de esta investigación fueron generados usando la técnica propuesta por Scott & Varian (2014), descrita en la Sección IV, considerando un intervalo de tiempo de Diciembre 2012 a Diciembre 2017, para datos semanales procedentes de *Google Trends* y un intervalo de tiempo análogo con frecuencia mensual para la llegada de turistas de Estados Unidos, Canadá y Alemania.

Los resultados se presentan de manera independiente para cada uno de los países considerados y evaluados en términos del Error Cuadrático Medio (ECM), la bondad de ajuste y en comparación con otros modelos usados como puntos de control: Vector Autoregresivo (VAR) de Frecuencia Mixta y Modelo Estado-Espacio. Por otra parte, se obtiene un indicador adelantado de la llegada de turistas al país, a través de una estimación retrospectiva para Agosto 2016- Julio 2017 (Figuras 4a-4c), considerándose un intervalo de tiempo equivalente a uno ( $h=1$ ), con el propósito de verificar si las proyecciones derivadas del modelo, para cada punto del período contemplado, convergen a la llegada efectiva de turistas para el mismo intervalo de tiempo.

Se consideran conjuntos individuales de 95, 89 y 120 variables (Tablas 1,2,3), respectivamente, para Estados Unidos, Canadá y Alemania, considerando los factores que podrían determinar las

preferencias de viaje de los agentes económicos. Por otra parte, usando la herramienta *Google Correlates* (descrita en la Sección III) , la información correspondiente a la llegada de turistas para cada uno de los países considerados fue utilizada para identificar los índices de *Google Trends* que mantuviesen un mayor coeficiente de correlación de Pearson en comparación con el flujo de turistas procedente de cada país.

En el proceso de compilación de variables se identificaron algunos patrones individuales; por ejemplo, en el caso de Estados Unidos, un aspecto que debe ser enfatizado es que la mayoría de las búsquedas asociadas a viajes incorporan variables relacionadas con actividades aptas para el disfrute de toda la familia (Tabla 1), siendo el foco geográfico de estas búsquedas el polo Sureste de la República Dominicana (específicamente Punta Cana). En el caso de Canadá, las búsquedas relacionadas a hoteles de lujo y spa fueron las más frecuentes (Tabla 2). Finalmente, para Alemania las variables contempladas están relacionadas básicamente con el ecoturismo, “mochileo” (*backpacking*) y visita a diferentes ciudades y puntos turísticos del país (Tabla 3).

Los resultados generados a través del sistema (3) fueron computados usando el software R, específicamente el paquete BSTS (Scott, 2016), mientras que los supuestos procedentes del modelo satélite ESARIMA y los resultados de los modelos comparativos fueron generados usando el software Eviews.

Una vez finalizado el proceso de selección de variables, la primera fase de la estimación consistió en mezclar la frecuencia de los datos, ya que la información procedente de *Google Trends* está disponible con periodicidad semanal, mientras que los datos de flujo de turistas procedentes de cada país están en frecuencia mensual.

En una segunda etapa, se llevó a cabo la especificación y estimación del componente de series de tiempo del modelo, dado por una Tendencia Lineal Local y un Componente Estacional, así como también la inclusión del componente de regresión, donde los datos de *Google Trends* son incorporados de manera exógena, como se ha señalado previamente. Las Figuras 4a, 4b y 4c muestran la contribución de cada componente particular, para los países considerados.

Siguiendo el enfoque planteado por Scott & Varian (2014), se emplearon *priors Spike & Slab*, siendo posible inferir información para estos *priors* a partir del tamaño del modelo, el  $R^2$  esperado, un tamaño de muestra  $v$  y usando shrinkage diagonal, que corresponde al prior  $g$  de Zellner<sup>8</sup>.

Se especificaron cuatro predictores para la estimación del modelo para cada país; esta cantidad de predictores implica establecer un umbral de inclusión<sup>9</sup> de 0.031, 0.04 y 0.026, para Estados Unidos, Canadá y Alemania, respectivamente (número esperado de predictores/ número total de variables). El valor de los otros *priors*, *i.g.* bondad de ajuste y tamaño de muestra, fue establecido de acuerdo al valor que proporciona por *default* el paquete BSTS (Scott, 2016), siendo  $R^2 = 0.5$  y  $v = 0.01$ . Para hacer un muestreo a partir de la distribución posterior se emplea el algoritmo MCMC, con 1,500 iteraciones.

Tras la estimación del modelo, se obtuvieron los predictores más probables para cada país, con sus respectivos coeficientes. En el caso de Estados Unidos (Figura 5a), los predictores con una probabilidad de inclusión que excede el umbral de inclusión son: Lugares de Vacaciones para la

---

<sup>8</sup> Véase Scott (2016) para verificar detalles técnicos asociados al establecimiento del *shrinkage* diagonal.

<sup>9</sup> La probabilidad de inclusión permite determinar los regresores que serían incluidos en el modelo, siempre y cuando excedan determinada probabilidad.

Familia (75%), Playas de Samaná (27%), Lugares para rentar en República Dominicana (26%) y Vacaciones Express (25%). La brecha entre el último predictor por encima del umbral de inclusión es significativa, considerando que la probabilidad de inclusión de la siguiente variable en línea, que es “Hoteles con piscina”, equivale a 12 por ciento.

Para el caso de Canadá (Figure 5b), se consideraron los mismos criterios de selección de *priors* y número de iteraciones; los predictores seleccionados en función de su probabilidad de inclusión fueron “Bávaro” (62%), “Resorts y Spa” (48%), “Vacaciones en República Dominicana” (22%) y “Vuelos Económicos a Punta Cana” (18%); las variables vinculadas a la búsqueda de hoteles específicos no resultaron ser relevantes para la especificación del modelo.

Los factores identificados como determinantes en la afluencia de turistas desde Alemania (Figura 5c) fueron “Airbnb” (89%), “Skyscanner” (36%), “Dominikanische” (16%), “Oberstdorf.de”<sup>10</sup> (7%); estas variables indican que existe una alta probabilidad de que los turistas alemanes obtengan soporte de agencias especializadas al momento de planificar sus vacaciones.

La Tabla 6 muestra los coeficientes estimados para los predictores por encima del umbral de inclusión. Todos los coeficientes, a excepción de la variable “oberstdorf.de” (en el caso de Alemania), exhiben una correlación positiva con la llegada de turistas procedentes de cada país.

---

<sup>10</sup> Oberstdorf es un poblado localizado en la región de los Alpes; su correlación negativa con la llegada de turistas alemanes hacia República Dominicana puede estar vinculada al hecho de que, al momento de planificar vacaciones, los alemanes también consideran este destino como una posibilidad.

Es importante enfatizar que los coeficientes de estas variables, así como también las probabilidades de inclusión de las mismas pueden variar en función del número de iteraciones seleccionado y de los datos. Del mismo modo, es pertinente señalar que cada solución del modelo puede generar distintas variables que se sitúen por encima del umbral de inclusión, en la medida que se tenga nueva información.

Una vez que fueron identificados los regresores relevantes para cada país, se generaron las proyecciones derivadas del Modelo Bayesiano Estructural de Series de Tiempo, usando las búsquedas de *Google Trends* como supuestos para la fase de *nowcasting*. Por otra parte, los supuestos para llevar a cabo las proyecciones de mediano plazo ( $h > 1$ ) fueron computados a través de los modelos ESARIMA (explicados en la Sección 5.1), para cada indicador y para un horizonte de tiempo de dos períodos. Se generaron proyecciones fuera de muestra para el año 2016 y 2017 y los resultados fueron evaluados en términos del error de pronóstico (ECM) y la bondad de ajuste, representada por el  $R^2$  (Tabla 5 y Tablas 6a-6c).

Con el propósito de verificar el valor agregado de las variables exógenas, representadas por la información procedente de *Google Trends*, se estima una representación estado-espacio<sup>11</sup> en la cual se captura la tendencia local y el componente estacional. Del mismo modo, se estima un VAR de Frecuencia Mixta<sup>12</sup>, que incorpora la misma información de *Google Trends*, salvo que en

---

<sup>11</sup> La idea de incorporar una representación Estado-Espacio es verificar la influencia del componente de regresión, representado por las variables procedentes de *Google Trends*, sobre la exactitud de las proyecciones.

<sup>12</sup> El enfoque de frecuencias mixtas se plantea como una herramienta simple para llevar a cabo *nowcasting* de series de tiempo; no obstante, para el tratamiento de grandes conjuntos de datos se presenta una gran desventaja: la falta de una técnica de *shrinkage* (Para detalles técnicos sobre estos modelos, véase, por ejemplo, Arnesto *et al.* 2010).

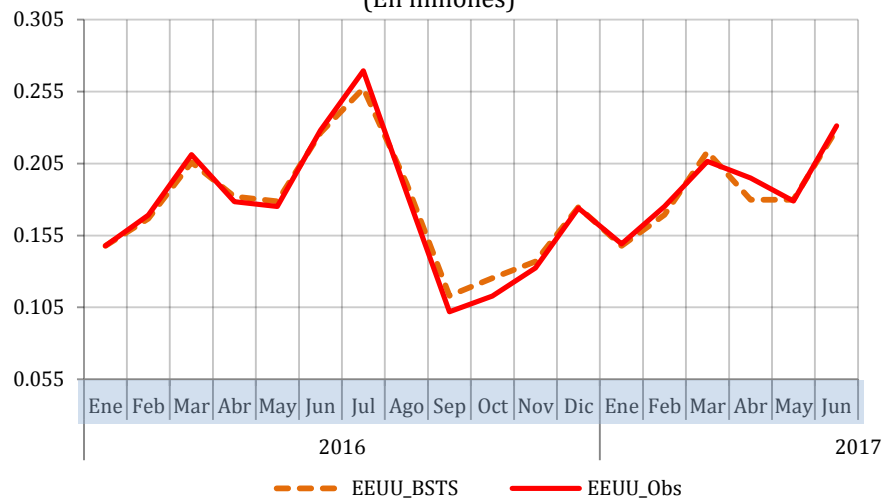
este caso la intención es emplear una técnica de estimación distinta, considerando el mismo conjunto informativo.

Los resultados correspondientes al modelo base (Tabla 5) con respecto a los modelos alternativos reflejan un decrecimiento significativo en el error de pronósticos tanto con respecto a la representación estado-espacio como al VAR de Frecuencia Mixta. Se puede constatar que el Modelo Bayesiano Estructural supera el desempeño de los modelos alternativos para la mayoría de los meses considerados, de acuerdo a los criterios seleccionados, así como también prediciendo con mayor exactitud los puntos de giro de las series, gracias a las señales extraídas de los indicadores de *Google Trends*, especialmente útiles en el *nowcasting*.

El ECM global, para cada caso, es también un indicador de cómo el modelo base supera el desempeño de los modelos alternativos, minimizando el error de pronóstico en aproximadamente un 90% para los tres países considerados (Tabla 4). Los resultados obtenidos justifican la selección del Modelo Bayesiano Estructural de Series de Tiempo, para el tratamiento de grandes conjuntos de datos, caracterizados por tener una cantidad de potenciales regresores que superan el número de observaciones.

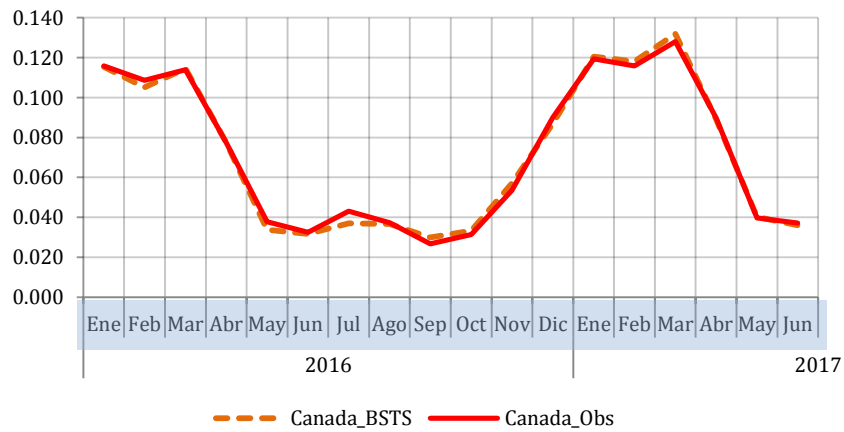
A través de los resultados obtenidos en esta fase, se elaboró un indicador adelantado con proyecciones dentro de muestra para el período Enero 2016-Junio 2017 de la llegada de turistas (por nacionalidad) hacia República Dominicana (Figura 6a-6c), en las cuales se observan los resultados del ejercicio de *nowcasting* y las proyecciones de corto plazo.

**Figura 6a.** Indicador Adelantado Llegada de Turistas: EEUU  
(En millones)

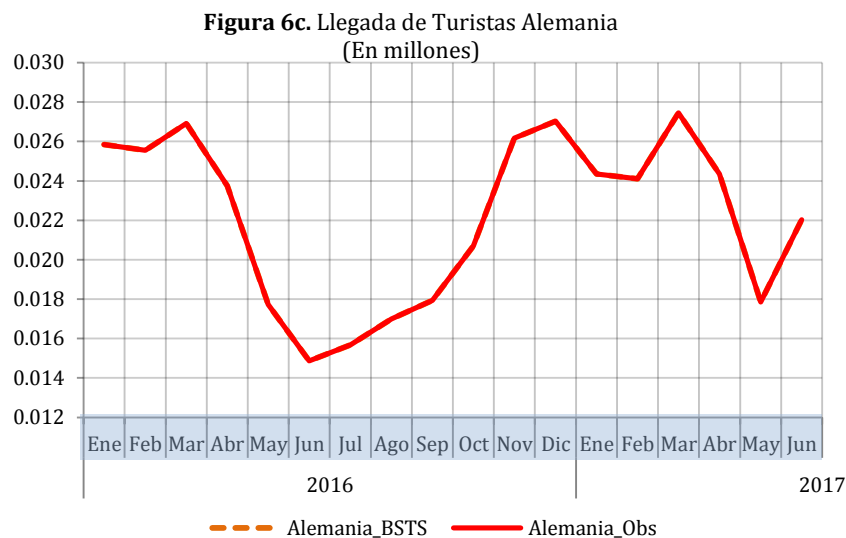


\*Fuente: Modelo Bayesiano Estructural (BSTS).

**Figura 6b .** Indicador Adelantado Llegada de Turistas :Canadá  
(En millones)



\*Fuente: Modelo Bayesiano Estructural (BSTS).



\*Fuente: Modelo Bayesiano Estructural (BSTS).

Estos indicadores muestran resultados prometedores, ya que se observa que las señales provistas por *Google Trends* son apropiadas para predecir la dinámica turística en República Dominicana y, más aún, los puntos de giro en las series, lo cual se visualiza por medio de la convergencia entre dichas señales y los valores de las variables efectivamente observados.

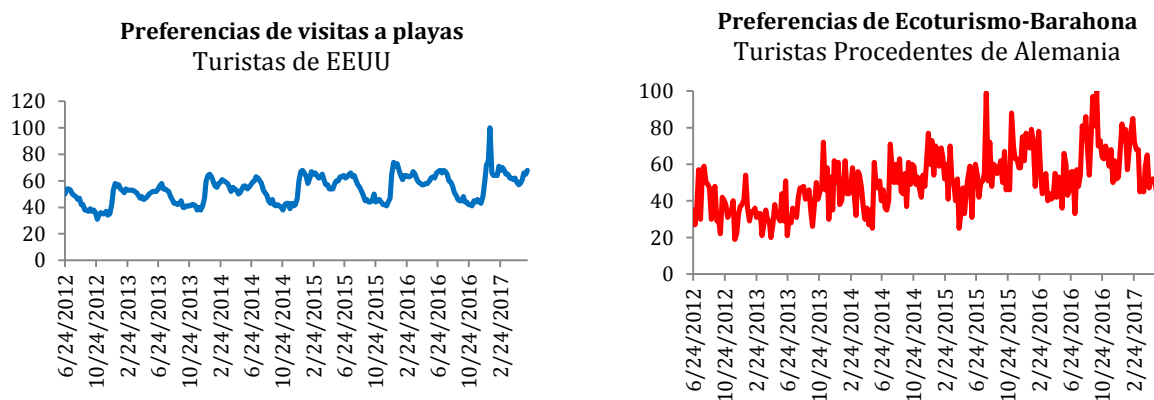
## 6.2. Mediano Plazo

Las proyecciones de corto plazo y el *nowcasting* permiten dirigir el enfoque hacia la coyuntura prevaleciente, así como también a actividades relacionadas al marketing y a la promoción de paquetes turísticos y a la fijación de precios. Este tipo de información resulta sumamente útil y el hecho de contar con la misma de manera oportuna y en tiempo real se presenta como un punto a favor de la técnica planteada, dada la política de expansión del Estado Dominicano en el ámbito turístico, la cual contempla la conquista de nuevos mercados, incluyendo el

Suramericano (Informe del Ministerio de Turismo, 2016) y la continuación de una agresiva publicidad hacia los principales mercados tradicionales (Estados Unidos de América y Europa).

Adicionalmente, una de las bondades de la herramienta *Google Trends*, aún en el corto plazo, es que coadyuva a identificar la popularidad de las actividades que desean realizar los turistas, considerando el volumen de búsquedas de las mismas; *i.g.* si los intereses de los visitantes al país están concentrados en realizar ecoturismo, senderismo, visitas a playas, etc. (Figura 7 ), e inferir una planificación basada en estos distintos tipos de actividades turísticas.

**Figura 7 .** Ejemplos de Preferencias de las Actividades de los Turistas.



**\*Fuente:** Google Trends y elaboración propia.

Por otra parte, dichas herramientas también permiten identificar las provincias en las cuales sería pertinente activar o re-activar la dinámica turística, instalar nueva infraestructura y llevar a cabo obras orientadas a satisfacer a los turistas, ya que permiten establecer concretamente cuáles son las zonas del país en las cuales se esperaría un mayor flujo de visitantes.

No obstante, pese a la gran cantidad de información que pueden proporcionar tanto el *nowcasting* como las proyecciones de corto plazo, otros aspectos tales como la planificación y

organización de políticas orientadas a realizar cambios estructurales en el sector turístico del país, tales como la construcción de obras de infraestructura viales, nuevas habitaciones<sup>13</sup>, acondicionamiento de áreas recreativas, servicios, implementación de nuevas legislaciones, modificaciones en el marco competitivo y adopción de medidas de estabilidad orientadas a garantizar la seguridad del turista están amparados en un contexto de proyecciones con un horizonte temporal superior al corto plazo.

En este sentido, la etapa final de esta investigación consiste en extender el enfoque utilizado al mediano plazo ( $h > 1$ ), empleando los mismos regresores resultantes de los ejercicios de nowcasting y corto plazo, a fin de alinear los resultados obtenidos con el horizonte de operación de la política monetaria, y manteniendo las ventajas inherentes a las técnicas de *machine learning*, a través del uso de proyecciones basadas en métodos iterativos (Chevillon, 2005; Huijskens, 2016). Las predicciones se computan para el horizonte  $h$  y se retroalimenta el modelo con los resultados de dichas proyecciones para generar  $h + 1$ . Posteriormente se emplea  $h + 1$  para generar  $h + 2$  y así sucesivamente hasta  $h + n$ , con el objetivo de aproximar:

$$E[y_{(t+1):(t+H)} | y_{(t-n+1):t}] \quad (12)$$

Donde  $n$  es el orden del proceso auto-regresivo. Dichas aproximaciones fueron generadas en el contexto del Modelo Bayesiano Estructural, propuesto en la Sección VII, excepto que en este caso la dinámica en términos de los insumos o supuestos del modelo es distinta al proceso de *nowcasting* y corto plazo, ya que el horizonte de proyección es distinto y se debe velar por la continua minimización de los errores de pronósticos en el corto plazo, que es una de las

---

<sup>13</sup> Es importante puntualizar que los avances en materia de ingeniería permiten llevar a cabo construcciones de infraestructura vial, habitacional, recreacional, entre otras, con una mayor celeridad y en el orden de prioridad que se requiera.

bondades de la técnica seleccionada). Las proyecciones de mediano fueron realizadas dentro y fuera de muestra para el período Junio 2017- Junio2018, a fin de verificar el desempeño del modelo en horizontes temporales superiores al *nowcasting* y corto plazo. El marco de referencia para validar los resultados de las proyecciones lo constituyen la información reportada por la Organización Mundial del Turismo (OMT) con respecto a las perspectivas regionales del sector en el cuatrienio 2016-2020, así como también la información consignada en el plan estratégico del Ministerio de Turismo de la República Dominicana en lo concerniente a las expectativas de desarrollo del turismo.

La importancia de las proyecciones de mediano plazo radica en el hecho de que ofrecen una visión que se extiende a un horizonte temporal en el cual resulta factible realizar cambios estructurales en el sector turismo de acuerdo a las tendencias subyacentes a nivel económico, social, demográfico y político, *e.g.* desarrollo de nuevos productos, competencias y marcos legislativos, realización de obras de infraestructura habitacional y vial en beneficio del turismo, formación de capital humano, entre otros aspectos.

El efecto multiplicador del turismo se visualiza a nivel de diferentes aristas macroeconómicas (Tabla 7), afectando las siguientes dimensiones de la economía: el desenvolvimiento del mercado laboral, tanto de manera directa (como resultado de los gastos turísticos directos) como de manera indirecta (gastos de apoyo al sector turismo); disminución del déficit de la balanza de pagos, a través de la generación de divisas vinculadas al flujo de turistas; aumento de la recaudación tributaria y por consiguiente mayores ingresos públicos; demanda de bienes y servicios, incluyendo un aumento en la demanda de inversión, que se traduce en una

dinamización general de la actividad económica, medida en términos del Producto Interno Bruto (PIB).

**Tabla 7.** Principales Indicadores del Sector Turismo de la República Dominicana (2010-2017).

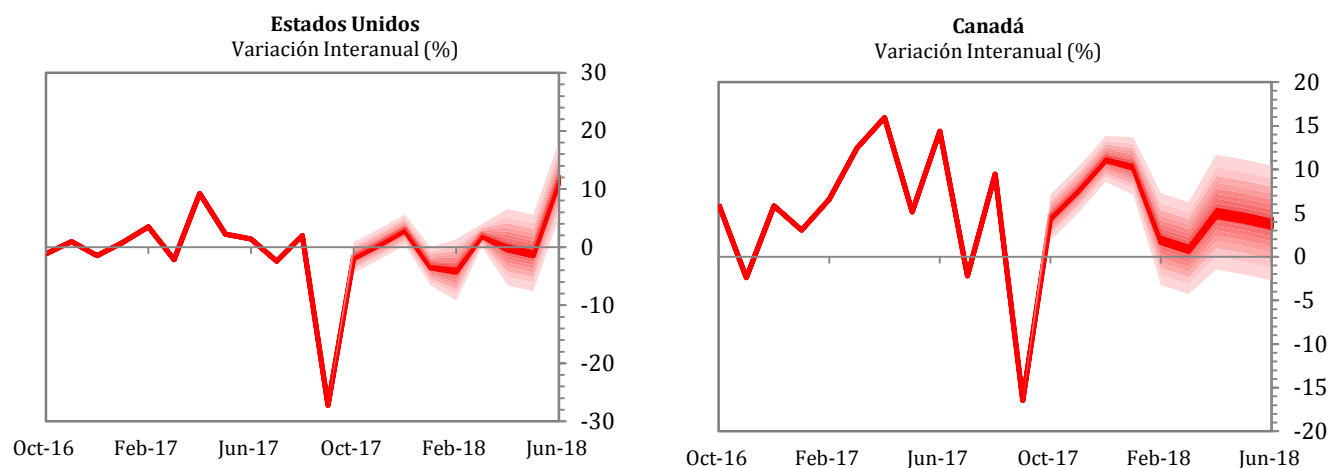
| Período | No. Habitaciones<br>(Unidades) | Tasa de<br>Ocupación<br>Hotelería (%) | Cantidad de<br>Turistas | Ingresos p/ | Balance             | Inversión<br>Extranjera | Empleos Generados |          |            |
|---------|--------------------------------|---------------------------------------|-------------------------|-------------|---------------------|-------------------------|-------------------|----------|------------|
|         |                                |                                       |                         | Turismo     | Cuenta<br>Corriente |                         | Total             | Directos | Indirectos |
|         |                                |                                       |                         |             | Millones US\$       |                         |                   |          |            |
| 2010    | 67,220.00                      | 66.60                                 | 4,124,543               | 4,161.70    | (4,023.50)          | 180.00                  | 195,371           | 55,820   | 139,551    |
| 2011    | 66,450.33                      | 69.30                                 | 4,306,431               | 4,385.70    | (4,334.60)          | 107.80                  | 201,597           | 57,599   | 143,998    |
| 2012    | 67,049.00                      | 70.30                                 | 4,562,606               | 4,680.10    | (3,880.90)          | 162.00                  | 201,235           | 57,496   | 143,739    |
| 2013    | 67,838.67                      | 71.70                                 | 4,689,770               | 5,054.70    | (2,567.90)          | 256.50                  | 216,543           | 61,869   | 154,674    |
| 2014    | 69,868.33                      | 74.80                                 | 5,141,377               | 5,629.80    | (2,170.20)          | 300.70                  | 247,025           | 70,578   | 176,447    |
| 2015    | 71,154.33                      | 75.50                                 | 5,599,859               | 6,115.90    | (1,280.30)          | 671.90                  | 254,146           | 71,686   | 182,460    |
| 2016*   | 74,005.33                      | 78.00                                 | 5,959,347               | 6,723.30    | (977.60)            | 790.00                  | 315,153           | 88,777   | 226,376    |
| 2017*   | 74,078.00                      | 77.10                                 | 6,187,542               | 7,177.50    | (165.10)            | 525.30                  | -                 | -        | -          |

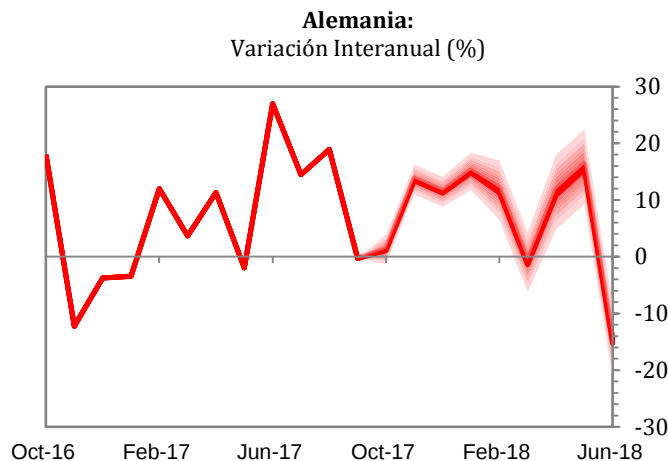
Fuente: Banco Central de la República Dominicana y Asociación de Hoteles y Turismo de la República Dominicana (ASONAHORES)

\*Cifras Preliminares

Los resultados de las proyecciones del flujo de turistas para el mediano plazo procedentes de los países considerados se muestran a continuación:

**Figura 8.** Gráficos de Abanicos: Proyecciones de Mediano Plazo para la llegada de Turistas procedentes de Estados Unidos, Canadá y Alemania.





**\*Fuente:** Modelo BSTS y elaboración propia.

Es evidente que la llegada de turistas a República Dominicana se vio significativamente afectada por el paso de fenómenos atmosféricos (tormentas Irma y María), que mermaron considerablemente la llegada de turistas al país. No obstante, las proyecciones generadas sugieren una paulatina recuperación de estas cifras, con tasas de crecimiento promedio de la llegada de turistas no residentes de 0.78%, 5.60% y 7.17% para Estados Unidos, Canadá y Alemania<sup>14</sup>, respectivamente hacia junio del 2018. Estos valores se muestran coherentes con las perspectivas de la OMT y del Ministerio de Turismo de la República Dominicana (MITUR), en cuanto al crecimiento turístico del país, al tiempo que reafirman el compromiso de las autoridades gubernamentales de alcanzar los 10 millones de turistas para el año 2020. El balance de riesgo de las proyecciones en el mediano plazo estaría sesgado a la baja, siendo visualizado como un indicador de las perspectivas favorables que se plantean para el sector turístico y que incidirían positivamente en la confianza y en la percepción de calidad de los consumidores de este servicio en República Dominicana. Sin embargo, es importante acotar que este balance de riesgos puede verse afectado por fenómenos atmosféricos y cambios ambientales, *e.g. ciclones, tormentas, caso de invasión de algas, entre otros*, que se van haciendo

<sup>14</sup> Esta trayectoria se vio afectada significativamente por el paso de las tormentas Irma y María, que fueron fenómenos atmosféricos que mermaron la llegada de turistas procedentes de estos destinos a la República Dominicana. No obstante, se espera un repunte del sector en un horizonte temporal de mediano plazo.

cada vez más agresivos como consecuencia del cambio climático y que constituyen un factor de incertidumbre. En este sentido, la inversión destinada a los proyectos turísticos debe contar con planes de contingencia que permitan mitigar los efectos de estos fenómenos y así minimizar o atenuar sus consecuencias sobre el flujo de turistas.

Dado el desenvolvimiento actual del sector turístico en la República Dominicana, matizado por elementos favorables tales como una tasa de crecimiento promedio de 6.6% en el período 2014-2016 y reafirmando su posición de liderazgo a nivel regional como captador de visitantes extranjeros (Figura 2) y generador de divisas, es de esperarse que las perspectivas de inversión que se vislumbran para el mediano y el largo plazo resulten optimistas, y que a la vez se perciban como señales del interés y confianza por parte de los inversionistas en el sector y en la estabilidad macroeconómica del país, así como en el marco regulatorio existente para la inversión turística, el cual promueve mediante incentivos fiscales el desarrollo y expansión de esa actividad en todo el país. En palabras del ministro de turismo, Francisco Javier García, “se puede afirmar categóricamente que el turismo se ha convertido en la verdadera locomotora de la economía dominicana”, desplazando los sectores de telecomunicaciones, minería y zonas francas, en términos de recepción de inversión extranjera directa.

De manera concreta, de acuerdo a los últimos datos del Consejo de Fomento Turístico (CONFOTUR), entre los proyectos de inversión predominan los del ámbito “inmobiliario-turístico”, con un 46.5% del total; “inmobiliario-hoteler” (42.3%) y de turismo de salud (5.2%), destacándose que la mayor proporción de las inversiones se localizan en la provincia El Seibo (72%), seguida de La Altagracia (11.9%), María Trinidad Sánchez (7.3%), Santo Domingo (4.4%), San Pedro de Macorís (3.5%), entre otras, con inversiones esencialmente de capital vernáculo, dominico-estadounidense, francés, así como también ítalo-dominicano (Figura 10).

En adición a la inversión inmobiliaria- hotelera, una oportunidad de mercado que recientemente ha sido explorada por pequeños comerciantes, distintos de grandes cadenas de hoteles, es la modalidad de alojamiento provista por portales de internet como “*Airbnb*” y “*Booking.com*”. En el caso de Republica Dominicana, estas ofertas de estadía son relativamente recientes y no son tan populares como los paquetes tradicionales ofrecidos por las grandes cadenas hoteleras. Sin embargo, es de esperarse que, si bien esta tendencia es incipiente en el país, adquiera preponderancia a través del tiempo (Documento del *Bankwest Curtin Economics Centre*, 2017) , principalmente en el segmento más joven de la población (Glusac, 2016) ; ya en otros países el interés en este tipo de oferta hotelera es cada vez más creciente y, de hecho, la mayor proporción de turistas que buscan realizar actividades distintas a las modalidades tradicionales de “paquetes todo incluido” tienen predilección por esta clase de alojamiento, donde encuentran casas y apartamentos a precios más accesibles. Este es el caso de los turistas procedentes de Alemania, uno de los países seleccionados en este estudio, donde la variable “*Airbnb*” se presenta como uno de los regresores posicionados por encima del umbral de inclusión del modelo BSTS<sup>15</sup>.

El crecimiento económico de la República Dominicana durante los últimos años se considera como uno de los más robustos de la región de América Latina y el Caribe (6.1% en el período 2015-2017) y se atribuye esencialmente a un cambio estructural en el patrón productivo, en el cual participa el turismo con un empleo creciente de recursos a tasas creciente de productividad. La entrada de inversión extranjera directa en turismo ha sido parte de una terciarización de la economía, en términos de producto y empleo, que a su vez se ha acompañado de un crecimiento del PIB *per-cápita* (Pérez, 2011).

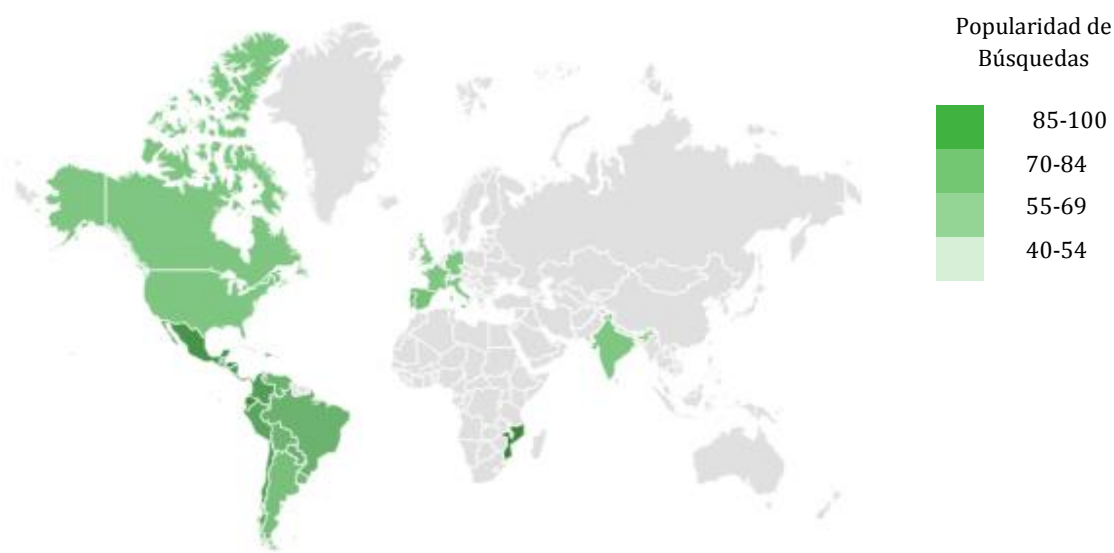
---

<sup>15</sup> Las señales de predilección de este tipo de alojamiento pueden ser capturadas a través de *Google Trends*.

Es importante destacar que un mayor impulso en el crecimiento turístico de la República Dominicana, tanto en el mediano como en el largo plazo, estará supeditado a la consistencia en la implementación del plan estratégico del país para asumir los retos inherentes al sector; estos desafíos incluyen la diversificación de la oferta turística del país, movilizándose a un paradigma bajo el cual no solamente se ofrezcan los populares paquetes de “todo incluido” o el turismo de “sol y playa”.

En este sentido, las perspectivas de crecimiento de mediano plazo del sector turístico requieren modificaciones estructurales que contemplen una evolución concomitante con los cambios en la demanda y la heterogeneidad de las preferencias de los consumidores, reforzando el ecoturismo (Figura 9), el turismo accesible (Figura 1) y de salud, lo que a su vez otorga mayor participación a diferentes comunidades del país con potencial para beneficiarse de los ingresos derivados de estas actividades, siendo esta una manera de promover el turismo sostenible y comunitario como una nueva tendencia. En este sentido, el modelo BSTS, utilizado para llevar a cabo las proyecciones presentadas en esta investigación, constituye una herramienta con potencial para monitorear de manera más cercana y capturar cambios en las preferencias de los consumidores, a través de la extracción de patrones de comportamiento a nivel microeconómico, empleando datos nano-económicos. Estas señales ayudarían a estructurar las decisiones inherentes a dicho sector en base a las tendencias prevalecientes y al potencial de demanda, lo que resulta necesario para conocer las tendencias turísticas de los próximos años, que a su vez resultan relevantes en el marco de las metas consignadas en el plan estratégico del MITUR.

**Figura 9.** Popularidad de Búsqueda del Término “Ecoturismo”



*\*Fuente: Google Trends y elaboración propia.*

El aumento del acervo de inversión extranjera en los proyectos turísticos del país, las perspectivas favorables que se plantean a raíz de esta afluencia de capital y los esfuerzos por diversificar la oferta turística, así como también la magnitud e importancia del turismo en la economía dominicana, existe una alta probabilidad de que se produzcan efectos de derrame (*spillover effects*) tanto positivos como negativos a distintos niveles. Un efecto de derrame positivo se observaría en el mercado laboral, con una mayor demanda de mano de obra, así como también un incentivo a la capacitación de capital humano en las áreas de interés para el sector turismo, de manera que se continúe propiciando la mejora en la competitividad del sector turístico local.

De manera que, el fortalecimiento de la educación a todos los niveles que intervienen en el sector, tanto del personal que proporciona servicios a los visitantes y que entra en contacto directo con el turista, como del personal que labora en las instituciones de envergadura

vinculada a políticas de inversión, planificación y sostenimiento del sector, así como también en la promoción y evaluación de la competencia.

Es importante destacar que, como se ha apuntado previamente, una de las tendencias del mercado turístico la constituye la búsqueda de destinos orientados a “turismo de salud”<sup>16</sup> (Figura 9) , lo que plantearía nuevos retos de formación en esa área<sup>17</sup>, ya que el personal de los lugares turísticos desempeña un papel importante en la reducción de las posibles deficiencias de acceso o la solución de las dificultades imprevistas, y su formación en la conciencia de la discapacidad y la atención al cliente puede minimizar las barreras que encuentran las personas con discapacidad. Este personal debe recibir una formación adecuada respecto a los derechos de las personas con discapacidad, a fin de prestar mejor asistencia y los servicios garantizados por esos derechos, supervisar y prestar los servicios necesarios y explicar el funcionamiento de las instalaciones diseñadas para clientes con discapacidad.

Por otra parte, un *spillover* negativo del crecimiento del turismo se asocia al deterioro y a la explotación de los recursos naturales, principalmente bajo un esquema de turismo orientado a “sol y playa” y a paquetes de “todo incluido”. De manera que, se debe propugnar por un modelo de turismo que tome en cuenta las repercusiones actuales y futuras, económicas y sociales y medioambientales de los visitantes, de la industria, del entorno y de las comunidades; en otras palabras, modelos de turismo sostenible.

---

<sup>16</sup> El turismo de salud o turismo médico es un fenómeno global que consiste en el viaje a otra ciudad o país para recibir algún tipo de tratamiento o atención médica (como cirugías, tratamientos, rehabilitación) o de bienestar. En el turismo de salud entra el turismo accesible, pues crea las condiciones para que las personas puedan tener un cómodo acceso a las instalaciones existentes y a los cuidados requeridos.

<sup>17</sup> De acuerdo a datos de *Google Trends*, República Dominicana se posiciona en el segundo lugar de la región, en cuanto a la intensidad de las búsquedas a nivel mundial vinculadas a “turismo de salud”.

En respuesta las externalidades negativas generadas por el turismo masivo, en los últimos años ha tenido lugar la proliferación de sub-grupos de consumidores, cada vez más conscientes de las consecuencias dañinas al medio ambiente provocadas por este fenómeno (Ogonowska y Torre, 2012). Estas cohortes de la población demandan productos “eco-amigables”, que defiendan la sostenibilidad del medio ambiente, creando un nicho de mercado para los proveedores que ofertan bienes ecológicos y sostenibles; se debe tener un cuidado especial al momento de lidiar con estos subconjuntos de agentes económicos, ya que su comportamiento y crecimiento resulta muy sensibles a choques positivos, aún si estos son de pequeña magnitud, e influyen en la convergencia de la dinámica de comportamiento hacia un equilibrio estable.

Es importante señalar que la cuantificación exacta de los costos- beneficios derivados de estos *spillovers*, tanto positivos como negativos, escapa de los límites de este estudio, pero puede ser considerada para futuras extensiones de la investigación.

Finalmente, en el marco de las discusiones de mediano plazo, es importante hacer referencia al carácter contra-cíclico de los programas de estímulo al sector turismo, especialmente en periodos de recesión económica. La literatura empírica (Vellas, 2011) enfatiza el rol del turismo para incentivar el crecimiento económico, dado que los flujos comerciales generados por una economía que depende esencialmente de este sector tienen un mayor efecto en el índice de confianza de los consumidores y en la incertidumbre financiera. De manera más específica, el estudio referido demuestra que en los países del continente americano, el turismo representa un elemento clave para lograr un sustancial repunte del crecimiento económico. De esta manera, se infiere que el sector turístico puede jugar un importante rol en planes de recuperación económica, considerando que este sector es un catalizador de la dinámica económica, a través de la reacción en cadena que crea su impulso, incluyendo el mercado

laboral, como fuente generadora de empleos para complementar otros sectores, en especial el de manufactura.

## **VII. Consideraciones Finales y Recomendaciones de Política**

La revolución generada por el creciente fenómeno de *Big Data* hace plausible la evaluación de la dinámica futura de diferentes variables económicas, incorporando información de alta frecuencia y disponible en tiempo real que, indiscutiblemente, ha propiciado un aumento significativo en el acervo de datos nano-económicos, que permiten verificar los patrones de comportamiento de los agentes económicos. Consecuentemente, estos avances han contribuido a reforzar los procesos de toma de decisiones, tanto a nivel macroeconómico como microeconómico, empleando técnicas de análisis inherentes a la teoría de *machine learning* (aprendizaje computacional), las cuales se perfilan como herramientas de gran valor para estos propósitos y con un alto potencial de desarrollo.

La investigación presentada cumple un objetivo dual: en primer lugar, la introducción de una nueva técnica de modelación, específicamente un Modelo Bayesiano Estructural de Series de Tiempo (BSTS), que emplea datos de alta frecuencia procedentes de la plataforma *Google Trends* para generar proyecciones de corto y mediano plazo sobre el flujo de turistas procedentes de Estados Unidos, Canadá y Alemania; en segundo lugar, esbozar las perspectivas del sector turismo en el corto y mediano plazo, de acuerdo a los resultados obtenidos y a las tendencias observadas, y examinar como esa trayectoria esperada y la consiguiente movilización de la cadena de producción por servicios turísticos se concilia con el desenvolvimiento de otras dimensiones de la economía como la inversión, el consumo y el mercado laboral. En adición a la concatenación de las perspectivas del sector turismo con el

desempeño de otras aristas de la economía, se examinan las nuevas tendencias que deben contemplarse en la oferta de turismo en la República Dominicana, tales como el ecoturismo y el turismo medicinal, a fin de cumplir con la demanda de los visitantes al país, bajo una esfera de sostenibilidad del sector.

El verdadero reto para las entidades públicas y privadas consistirá en implementar las estrategias apropiadas para explotar este amplio espectro de datos, con la intención de escudriñar y monitorear de manera más cercana dicha información, a fin de optimizar los beneficios derivados de esta información y, al mismo tiempo, tener una mejor comprensión de los mecanismos que permiten la generación de los vínculos existentes entre el sector turismo y otras áreas de la economía.

Ahora bien, es prudente cuestionarse cómo la estructura y desarrollo de este trabajo permiten responder estas preguntas de investigación y, más aún, cómo estos resultados se concilian con las expectativas de desempeño de estas nuevas herramientas de análisis que se plantean. En este sentido, esta investigación no sólo explica meticulosamente cómo es posible extraer señales del comportamiento de determinadas variables relacionadas al sector turístico de la Republica Dominicana (en adición a las estadísticas publicadas por organismos oficiales), que históricamente se ha posicionado como uno de los pilares fundamentales de la economía, sino que también explica cómo utilizar esta información para la proyección y construcción de indicadores adelantados basados en la misma. Esta ventana de posibilidades resulta bastante atractiva en un momento en que el país está invirtiendo esfuerzos y recursos para guiar al sector turismo al cumplimiento de metas más ambiciosas, que incluyen elevar a diez millones el número de visitantes no residentes en el país para el año 2020, por lo que la concatenación de este objetivo con el uso de métodos que producen información oportuna y una evaluación

más meticulosa del comportamiento de este sector contribuye a reforzar los procesos de toma de decisiones. Por otra parte, esta información no es usada exclusivamente como insumo para realizar proyecciones en base a técnicas que incluyen variables procedentes de fuentes novedosas y confiables, sino que también resulta posible capturar señales relativas al comportamiento de indicadores vinculados a este sector. Al mismo tiempo, la metodología considerada para el manejo y realización de las proyecciones realiza un proceso de selección por medio del cual se obtienen los indicadores más relevantes en términos de su probabilidad de inclusión en el modelo de pronósticos, gracias a la técnica consistente de un Modelo Bayesiano Estructural de Series de Tiempo (BSTS), que emplea *priors* “*Spike & Slab*”, a fin de descartar regresores irrelevantes dentro de un gran conjunto de variables, simplificando el proceso de selección de variables, que se presenta al manejar grandes conjuntos de datos.

Los resultados generados demuestran que el modelo BSTS tiene un desempeño superior a los modelos usados como alternativas de comparación, evaluados en función del error de pronóstico (Error Cuadrático Medio, ECM), bondad de ajuste y ECM general, observándose una reducción de aproximadamente un 40% en el ECM general con respecto a las representaciones estado espacio sin componente de regresión y en más de un 100% en comparación con el VAR de Frecuencia Mixta. Los indicadores construidos para cada país evidencian que las señales capturadas a través de la plataforma *Google Trends* resultan útiles no sólo a lo largo del proceso de predecir la trayectoria futura de una variable, sino también para pronosticar los puntos de giro en el comportamiento de las series de tiempo.

En términos del aporte a la literatura económica dominicana, cabe señalar que no existe evidencia empírica de trabajos desarrollados en el sector turismo que utilicen las herramientas planteadas en esta investigación; siendo éste uno de los trabajos pioneros en términos de

aplicaciones de *machine learning* en el sector considerado. En el marco del plan estratégico del Ministerio de Turismo de la República Dominicana (MITUR), las señales, proyecciones y patrones de comportamiento de los agentes económicos, determinados por medio de este análisis son idóneos para ayudar a estructurar las decisiones inherentes a este segmento de la economía, en base a las tendencias prevalecientes y al potencial de demanda en el sector , siendo éstos factores que la política turística dominicana deberá tener en cuenta para sus estrategias futuras de incremento de la competitividad.

Si bien el valor agregado de los ejercicios de *nowcasting* y las proyecciones de corto plazo radica en la gestación rápida de operaciones y tácticas de promoción de actividades, establecimiento de precios así como también en la evaluación de la coyuntura sectorial prevaleciente, son las proyecciones generadas a horizontes de mediano plazo las que permiten estructurar las medidas de mayor nivel de profundidad. De acuerdo a las metas consignadas en el plan estratégico del MITUR, la esquematización de la política del sector turismo, en el mediano y largo plazo, está encauzada a potenciar el crecimiento de este segmento de la economía, contemplándose una afluencia de diez millones de turistas para el año 2020. Para la concretización de este propósito, se apelaría a políticas de inversión en infraestructura vial y habitacional<sup>18</sup>, así como también a la diversificación de la oferta turística del país, atendiendo a tendencias como el ecoturismo, el turismo medicinal y creando las condiciones para el turismo accesible, sin perder de vista el tema de la sostenibilidad turística que abogue por la protección del medio ambiente y por regular el desgaste de los recursos naturales, que puede generar la masificación del turismo en una comunidad.

---

<sup>18</sup> Como se mencionó en la Sección VI del documento, estas inversiones ya han empezado a gestarse, por medio de capital procedente de diferentes países.

En adición al ambiente de sostenibilidad turística en que debe ampararse la materialización de estos objetivos, se requiere un fortalecimiento de la educación a todos los niveles que intervienen en el sector, tanto del personal que proporciona servicios a los visitantes, como del personal que labora en las instituciones de envergadura vinculada a políticas de inversión, planificación y sostenimiento del sector, así como también en la promoción y evaluación de la competencia. Del mismo modo, se considera como prioridad el fortalecimiento de la seguridad ciudadana, a fin de generar mayor confianza en el país como destino turístico y de manera especial en las zonas de mayor flujo de visitantes.

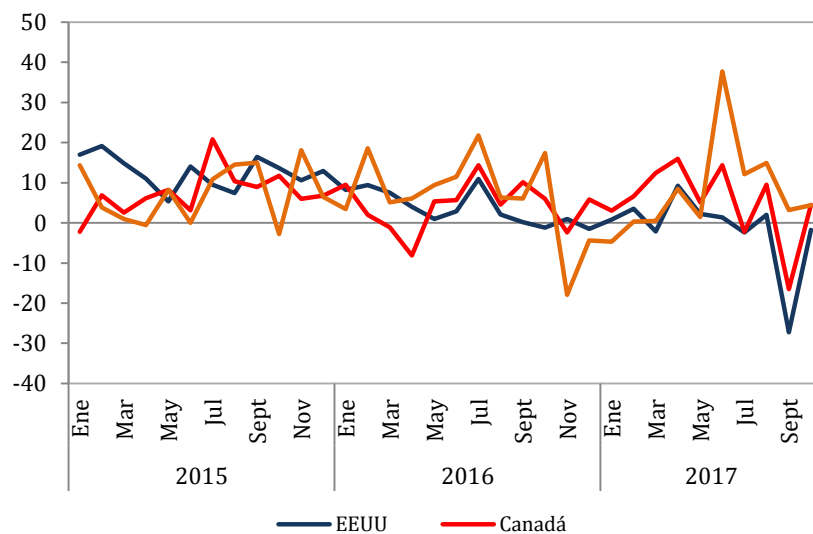
La inversión en turismo genera la movilización de una cadena de producción por los servicios demandados, produciendo efectos tanto directos como indirectos sobre el comportamiento de diversas variables macroeconómicas, principalmente el mercado laboral y la balanza de pagos. Del mismo modo, los impuestos, tanto nacionales y provinciales, son percibidos por el sector público en todos sus niveles sin discriminación de la persona que los consume; por lo tanto, los turistas en su calidad de consumidores de bienes y servicios nacionales contribuyen a aumentar los ingresos del Estado. El rol de modelos y técnicas de análisis suficientemente sofisticadas, así como también un seguimiento cercano de datos nano-económicos que permitan potenciar los beneficios de estas decisiones de política resulta inminente para garantizar un mayor dinamismo de la actividad económica, a través de los efectos del multiplicador del gasto turístico y de su rol como variable contra-cíclica para estimular la actividad económica.

Finalmente, es importante señalar que las futuras extensiones de este trabajo podrían incorporar la construcción de un índice de confianza del turista, así como también una cuantificación exacta de los costos- beneficios derivados de los *spillovers*, tanto positivos como

negativos, de las nuevas tendencias de oferta de turismo que están emergiendo en República Dominicana, a implementarse en el marco del plan estratégico de las autoridades competentes. En términos econométricos, la realización de ejercicios de impulso-respuesta se plantea como una posibilidad en el contexto del Modelo BSTS, por medio de la aplicación de choques a las variables seleccionadas y estableciendo comparaciones con un escenario base.

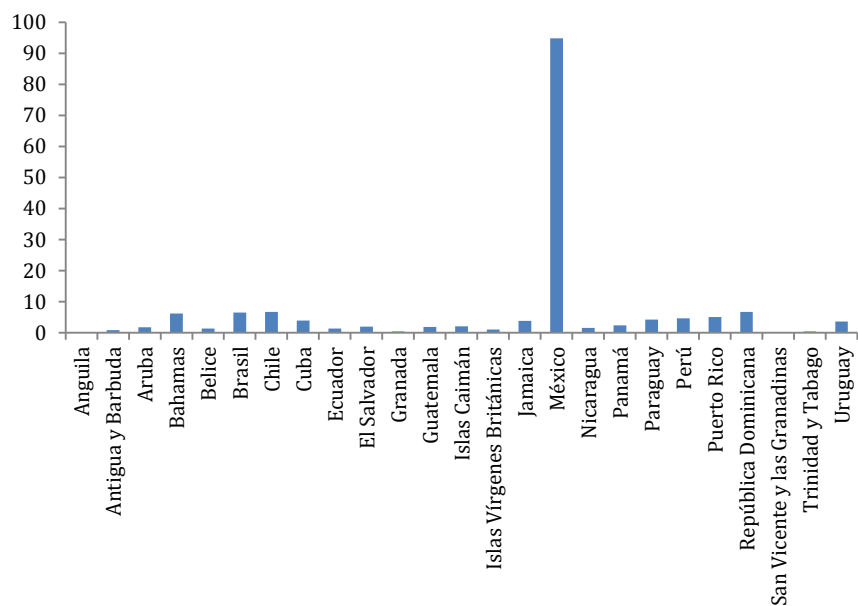
## ANEXO I

**Figura 1.** Llegada de Turistas a República Dominicana de Estados Unidos, Canadá y Alemania (Variación Interanual).



\*Fuente: Banco Central de la República Dominicana.

**Figura 2 .** Llegada de Turistas América Latina y el Caribe (En millones) . 1/



\*Cifras Preliminares

\*\*Fuente: Organización Mundial del Turismo (OMT).

1/ Datos revisados a Diciembre 2017.

**Tabla 1. Variables incluidas en el Modelo BSTS, Estados Unidos.**

| <b>Categorías</b>    | <b>Búsquedas Google Trends EEUU</b>  |
|----------------------|--------------------------------------|
| Búsqueda de Ciudades | Punta Cana                           |
|                      | Las Terrenas                         |
|                      | Puerto Plata                         |
|                      | Bavaro                               |
|                      | Cabarete                             |
|                      | Excursions                           |
|                      | Cities to visit, dominican republic  |
|                      | Samana Beaches                       |
|                      | Best beaches, Dominican Republic     |
|                      | Vacations, Las Terrenas              |
|                      | Best Beaches, Dominican Republic     |
|                      | Ecotourism, Samana                   |
|                      | Ecotourism, Barahona                 |
| Alojamiento          | All Inclusive Resorts, Punta Cana    |
|                      | Family Resorts, Punta Cana           |
|                      | Beach Houses, Punta Cana             |
|                      | Vacation Houses, Santo Domingo       |
|                      | 3 Bedroom Rental, Dominican          |
|                      | Hotel Las Terrenas                   |
|                      | Best All Inclusive Hotels            |
|                      | All Inclusive Hotel, Adults          |
|                      | Kids Friendly Resorts, Santo Domingo |
|                      | Junior Suite, Punta Cana             |
|                      | 4 Bedroom House Rental, Punta Cana   |
|                      | Bedroom Apartment, Punta Cana        |
|                      | Places to rent, Dominican Republic   |
|                      | Alsol Luxury Village                 |
|                      | Spa Resorts, Santo Domingo           |
|                      | Cheap Hotels, Santo Domingo          |
|                      | Family Friendly Resorts              |
|                      | Hotels with pool, Punta Cana         |
|                      | Amanera Resort                       |
|                      | Grand bahia principe resort          |
|                      | Book Hotel, Dominican Republic       |
|                      | Hard rock hotel                      |
|                      | Cheap Resorts, Dominican Republic    |
|                      | Hotels Near Beach                    |
|                      | Hotels Waterpark                     |
|                      | Best Resorts Caribbean               |
|                      | Paradisus Punta Cana                 |
|                      | Bayhipe Hotels                       |
|                      | Juan Dolio Resorts                   |

| <b>Categorías</b>           | <b>Búsquedas Google Trends EEUU</b>  |
|-----------------------------|--------------------------------------|
| Reserva de Vuelos           | Book Flight, Santo Domingo           |
|                             | Book Flight, Punta Cana              |
|                             | Book Flight, Dominican Republic      |
|                             | Cheapest all Inclusive Travel        |
|                             | Trip Advisor                         |
|                             | Flight Packages, Santo Domingo       |
| Entretenimiento             | Nightlife, Santo Domingo             |
|                             | What to do in Dominican Republic     |
|                             | Zona Colonial Tours                  |
|                             | Family Activities, Santo Domingo     |
|                             | Fun Spots, Santo Domingo             |
|                             | Family things to do, Punta Cana      |
|                             | Kid Things to do                     |
|                             | Kids Attractions                     |
|                             | Kids Activities                      |
|                             | Best Restaurants, Punta Cana         |
|                             | Best Snorkeling                      |
|                             | Best Restaurants, Santo Domingo      |
|                             | Healthy Places to eat, Santo Domingo |
|                             | Family Activities, Santo Domingo     |
| Variables Google Correlates | Dominican Republic, Vacation         |
|                             | Rent Car                             |
|                             | Wedding, Punta Cana                  |
|                             | Bride dress: (United States)         |
|                             | Key Rental, Punta Cana               |
|                             | Lose Weight, fast results            |
|                             | Punta Cana Airport                   |
|                             | Get Visa Tourist                     |
|                             | Flights to Orlando                   |
|                             | Dominican Republic                   |
|                             | Flight Miami                         |
|                             | Chicken breast nutrition             |
|                             | Walgreens Passport                   |
|                             | Airport, Santo Domingo               |
|                             | Back workout                         |
|                             | Airport , Santo Domingo              |
|                             | Back Workout                         |
|                             | How to book, Punta Cana              |

| <b>Categorías</b>           | <b>Búsquedas Google Trends EEUU</b> |
|-----------------------------|-------------------------------------|
| Variables Google Correlates | MLB Prospects                       |
|                             | Trip Packing List                   |
|                             | What to pack for a trip             |
|                             | Trip packing list, punta cana       |
|                             | Driving From Santo Domingo          |
|                             | Best Vacation destination           |
|                             | Scuba Diving                        |
|                             | All Inclusive Deals                 |
|                             | Vacation Spots                      |
|                             | All Inclusive Deals, Santo Domingo  |
|                             | Best Beach Vacation                 |
|                             | Travel Package,Dominican Republic   |
|                             | Workout, Weight Loss                |
|                             | Flights From                        |
|                             | Vacations Express                   |
|                             | Driving from                        |
|                             | On vacation                         |
|                             | Going on Vacations                  |

**Tabla 2. Variables incluidas en el Modelo BSTS, Canadá.**

| <b>Categorías</b> | <b>Búsquedas Google Trends, Canadá</b> |
|-------------------|----------------------------------------|
| Búsqueda Ciudades | Punta Cana                             |
|                   | La Romana                              |
|                   | Punta Cana, Dominican Republic         |
|                   | Cabarete                               |
|                   | Playa Dorada                           |
|                   | Bavaro                                 |
|                   | Sosua                                  |
|                   | Cayacoa                                |
|                   | Puerto Plata                           |
|                   | Cayo Levantado                         |
|                   | Samana                                 |
|                   | Bavaro-Punta Cana                      |
|                   | Santo Domingo                          |
|                   | Las Terrenas                           |
| Alojamiento       | Riu                                    |
|                   | Melia Punta Cana                       |
|                   | Catalonia                              |
|                   | Barcelo                                |
|                   | Resorts & Spa, Punta Cana              |
|                   | Crown Paradise Club                    |
|                   | Paradisus Punta Cana                   |
|                   | Resorts, Puerto Plata                  |
|                   | Bahia Principe                         |
|                   | Rooms for rent                         |
|                   | Apartaments for rent                   |
|                   | Best resorts, Punta Cana               |
|                   | Iberostar                              |
|                   | Resort & Spa, Bavaro                   |
|                   | Where to stay                          |
|                   | Occidental Grand Punta Cana            |
|                   | Puerto Plata Resorts                   |
|                   | Book Apartment                         |
|                   | Booking.com                            |
|                   | Best all inclusive resorts             |
|                   | Airbnb                                 |
|                   | Best all inclusive                     |
|                   | Bahia Principe, Cayacoa                |
|                   | Junior Suites                          |

| <b>Categorías</b>           | <b>Búsquedas Google Trends, Canadá</b> |
|-----------------------------|----------------------------------------|
| Entretenimiento             | What to do in Dominican Republic       |
|                             | Restaurants                            |
|                             | Nightlife                              |
|                             | Excursions                             |
| Reservas Vuelos             | Dominican Republic Travel              |
|                             | Trivago                                |
|                             | Tripadvisor                            |
|                             | Flight Punta Cana                      |
|                             | Skyscanner                             |
| Variables Google Correlates | Vacations                              |
|                             | Cheap Vacations                        |
|                             | Orlando Airport                        |
|                             | Sell-off Vacations                     |
|                             | Punta Cana Weather                     |
|                             | Punta Cana Reviews                     |
|                             | Punta Cana Airport                     |
|                             | Novice                                 |
|                             | Punta Cana Resort Map                  |
|                             | Elliptical                             |
|                             | Graduate Application                   |
|                             | Cheap Trip, Punta Cana                 |
|                             | Destination Voyage                     |
|                             | Dominican Republic travel              |
|                             | Dominican Airport                      |
|                             | Dominican Republic Vacation            |
|                             | Best Beaches                           |
|                             | Voyage Sunwing                         |
|                             | Voyage Aquarelle                       |
|                             | Playa                                  |
|                             | Tout Inclus                            |
|                             | Toronto to Punta Cana                  |

**Tabla 3. Variables incluidas en el Modelo BSTS, Alemania.**

| <b>Categorías</b>  | <b>Búsquedas Google Trends, Alemania</b> |
|--------------------|------------------------------------------|
| Búsquedas Ciudades | Punta Cana                               |
|                    | Bavaro                                   |
|                    | Bayahibe                                 |
|                    | Las Terrenas                             |
|                    | Puerto Plata                             |
|                    | Samana                                   |
|                    | Cabarete                                 |
|                    | Cayo Arena                               |
|                    | La Romana                                |
| Alojamiento        | Bahia Principe                           |
|                    | Hotels Punta Cana                        |
|                    | Barcelo                                  |
|                    | Melia                                    |
|                    | Palm Resort                              |
|                    | Kinderhotel                              |
|                    | Costa Calma Palace                       |
|                    | Sporthotel                               |
|                    | Wellness Hotel                           |
|                    | Resorts Punta Cana                       |
|                    | Resorts Samana                           |
|                    | Iberostar Bavaro                         |
|                    | Apparments Rent                          |
|                    | Barcelo Bavaro                           |
|                    | Club Hotel                               |
|                    | Paradisus                                |
|                    | Riu Bambu                                |
|                    | Apartments                               |
|                    | Iberostar                                |
|                    | Hotel Costa                              |
|                    | Casa Marina                              |
|                    | Princess Hotel                           |
|                    | Hotel Viva                               |
|                    | Iberostar Paraiso                        |
|                    | Nickelodeon hotel, Punta Cana            |
|                    | Houses rent, Dominican Republic          |
|                    | Beach Houses Rent, Dominican Republic    |
|                    | Bedroom Apartment Punta Cana             |
|                    | Hotel Bahia                              |
|                    | Spa Resorts, Dominican Republic          |
|                    | Airbnb                                   |

---



---

Bed and breakfast, Dominican Republic

---



---

| Categorías                 | Búsquedas Google Trends, Alemania                                                                                                                                                                                                                                                                                                                                                                                              |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Reserva Vuelos             | Skyscanner<br>Flights punta cana<br>Booking.com<br>Flights, Dominican Republic<br>Cheap flights, Dominican Republic<br>Trivago<br>Tripadvisor                                                                                                                                                                                                                                                                                  |
| Entretenimiento            | What to do, Dominican Republic<br>Restaurants<br>Nightlife, Dominican Republic<br>Fun Spots, Dominican Republic                                                                                                                                                                                                                                                                                                                |
| Variables Google Correlate | Dominican Republic: (Alemania)<br>Climatic Table<br>Dom Rep<br>Dominican<br>Cheap Travel<br>Akumal<br>maspalomas<br>Shelves<br>stopover: (Alemania)<br>Esquinzo<br>Cabinets<br>Schauinsland<br>Lodge<br>Tulum<br>Kanarem<br>Holiday House<br>Game Reserve<br>Travel Climate<br>Number of rooms<br>Updates<br>Fitness Studio<br>Search<br>Maxxus<br>Mar y Montaña<br>Hotels Inn<br>Flexa<br>Pine Furniture<br>John Travel<br>cz |

---

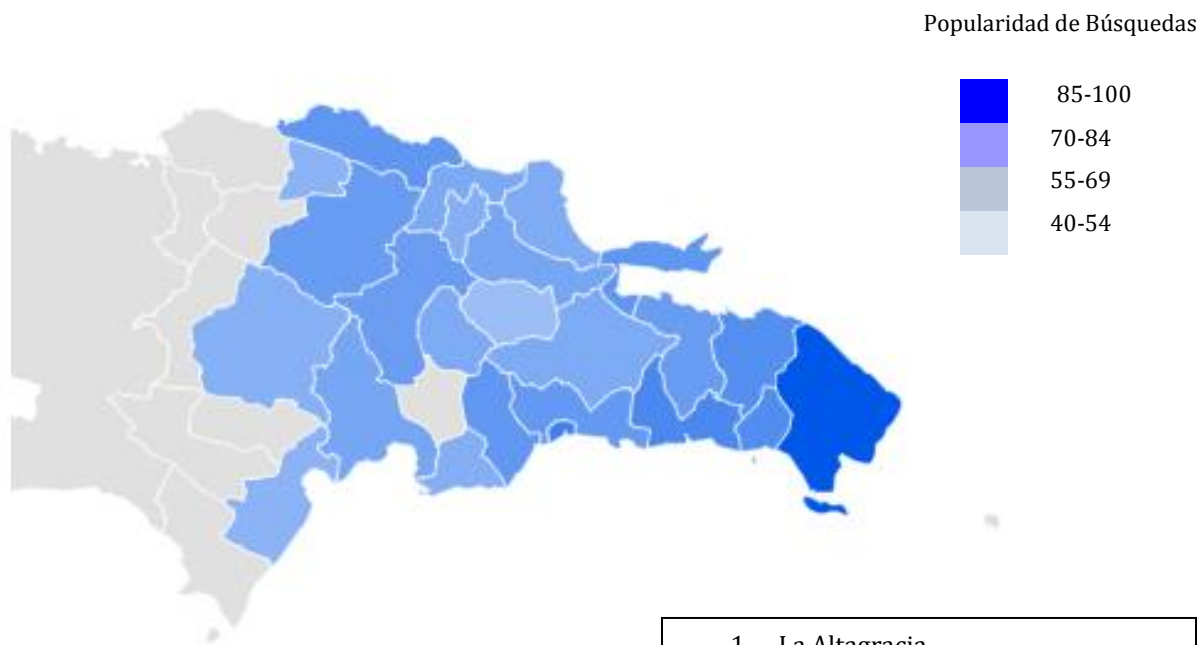


---

| Categorías                  | Búsquedas Google Trends, Alemania |
|-----------------------------|-----------------------------------|
| Variables Google Correlates | Aeg                               |
|                             | Lazne                             |
|                             | Best Travels                      |
|                             | Gouna Hotel                       |
|                             | DSL Offers                        |
|                             | Study Travel                      |
|                             | Austria Travels                   |
|                             | Mountain huts                     |
|                             | Cheap Health Insurance            |
|                             | Travels                           |
|                             | Ram                               |
|                             | Contribution rates                |
|                             | Computer Cabinet                  |
|                             | Walsertal                         |
|                             | Furniture House                   |
|                             | Contribution                      |
|                             | Baan                              |
|                             | Of travel                         |
|                             | Bretz                             |
|                             | Villeroy Boch                     |
|                             | Rassekatzen                       |
|                             | Required to keep                  |
|                             | Natural Furniture                 |
|                             | Wellness Trips                    |
|                             | Interhome                         |
|                             | Mountaineer Resorts               |
|                             | Dansommer                         |
|                             | Flights                           |
|                             | Network Cable                     |
|                             | Fitness Training                  |
|                             | Fitness Equipment                 |
|                             | Transorient                       |
|                             | Alltours                          |
|                             | Fertility                         |
|                             | Walsertal                         |

| <b>Categorías</b>          | <b>Búsquedas Google Trends, Alemania</b> |
|----------------------------|------------------------------------------|
| Google Correlate Variables | Holiday Packages                         |
|                            | Furnitures, Rooms                        |
|                            | Gradner                                  |
|                            | www.oberstdorf.de                        |
|                            | Food, Restaurants                        |
|                            | Dominican                                |
|                            | Bavaro Beach                             |
|                            | Golf Trips                               |
|                            | Holiday Time                             |
|                            | Diving                                   |
|                            | Islands                                  |
|                            | Fit                                      |
|                            | Travel Packages                          |
|                            | Club Ship                                |
|                            | Playa Esmeralda                          |
|                            | Weekend Away                             |
|                            | Dolomitas prices                         |
|                            | Dominican Republic                       |
|                            | Boatable                                 |
|                            | Makadi Oasis                             |
|                            | Hotel Deals                              |
|                            | Hut                                      |
|                            | Coral Bay                                |
|                            | Dive Center                              |
|                            | Online Booking                           |
|                            | Punta Cana Airport                       |
|                            | Backpacker                               |
|                            | Special Offers, Dominican Republic       |
|                            | Special Offers, Punta Cana               |
|                            | Las Americas Airport                     |
|                            | Ecotourism, Dominican Republic           |
|                            | Dominican Republic Vacation              |
|                            | Beach Vacation, Dominican Republic       |

**Figura 10.** Búsquedas de *Google Trends* y noticias asociadas a la distribución de inversión en turismo en República Dominicana,

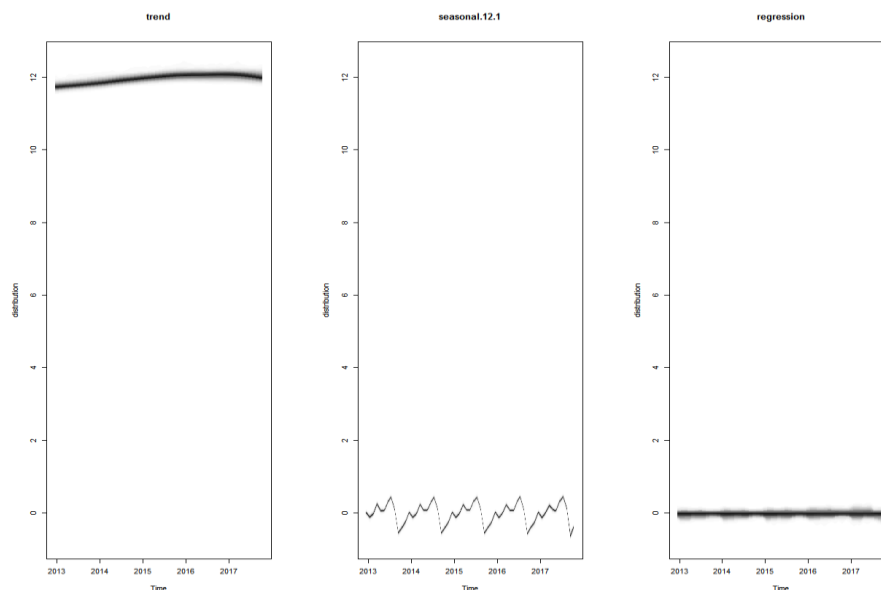


\*Fuente: *Google Trends* y elaboración propia.

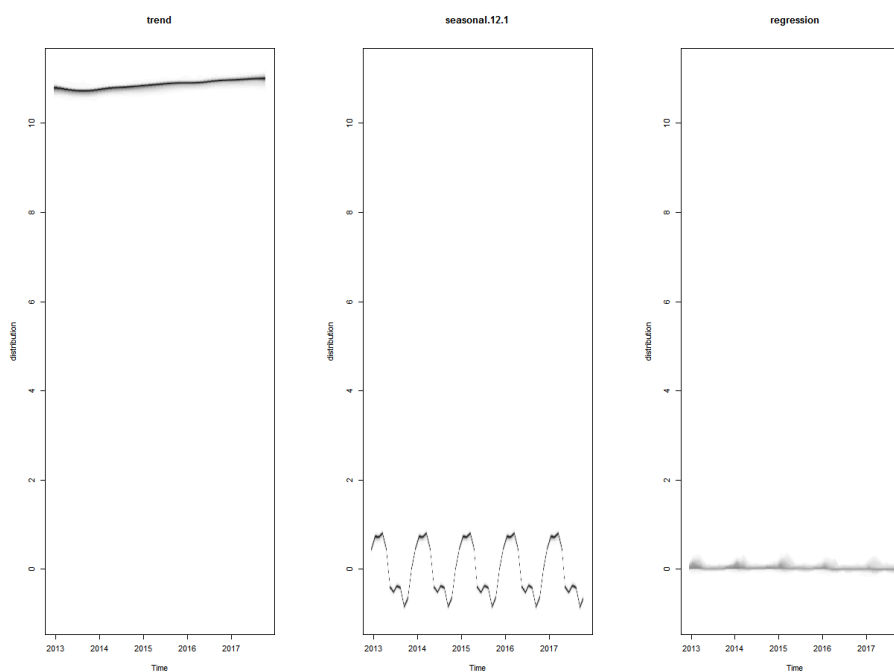
1. La Altagracia
2. Distrito Nacional
3. San Pedro de Macorís
4. La Romana
5. El Seibo

**Figuras 4a-4c.** Contribución por componentes (tendencia, estacional, regresión) de la llegada de turistas a República Dominicana 1/

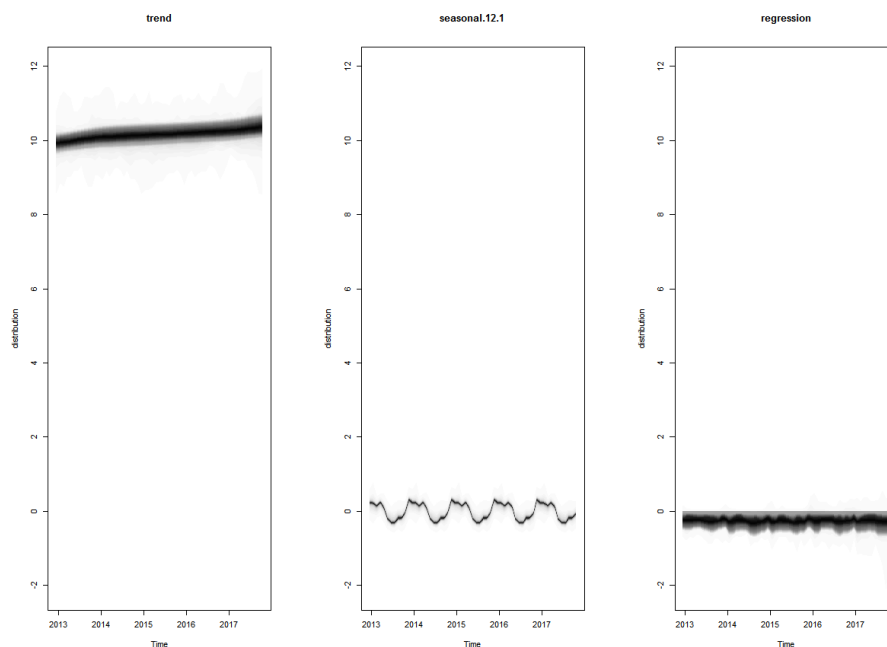
**Figura 4a.** Proyecciones de la llegada de Turistas de Estados Unidos por componente.



**Figura 4b.** Proyecciones de la llegada de Turistas de Canadá por componente.

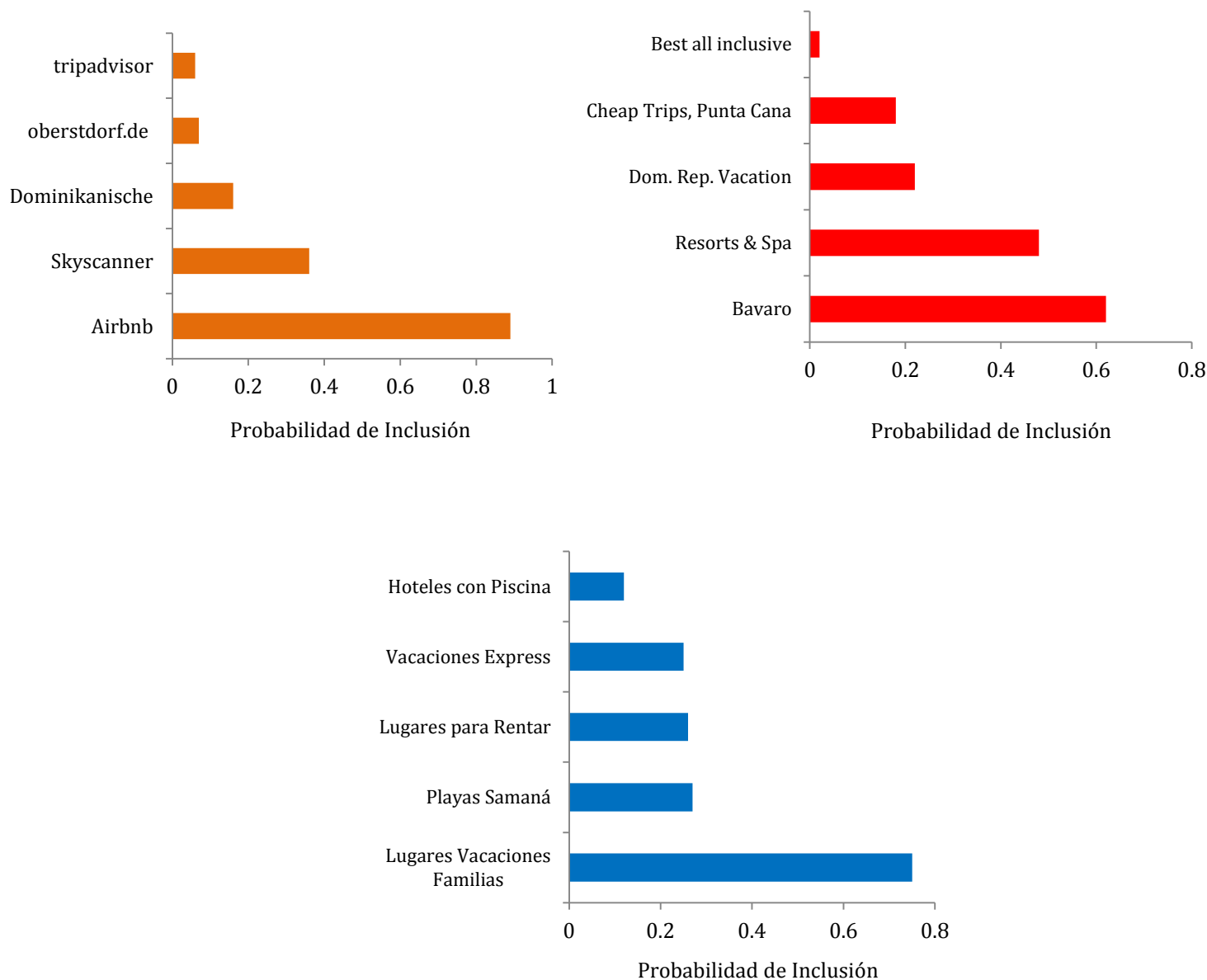


**Figura 4c.** Proyecciones de la llegada de Turistas de Alemania por componente.



**Notas:** 1/ El panel (a) de cada figura representa el comportamiento de tendencia, el panel (b) muestra el componente estacional y el panel (c) el aporte del componente de regresión a las proyecciones correspondientes a cada país.

**Figuras 5a-5c.** Predictores Modelo Bayesiano Estructural de Series de Tiempo para Estados Unidos, Canadá y Alemania 1/ .



**Notas:** 1/ Los paneles (a-c) representan el conjunto de predictores más probables, generados por el Modelos Bayesiano Estructural (BSTS) para Estados Unidos, Canadá y Alemania.

**Tabla 4 .Resumen de Error Cuadrático Medio (ECM), h=2**

| Estados Unidos  | Ago-16 | Sept-16 | Oct-16 | Nov-16 | Dic-16 | Ene-17 | Feb-17 | Mar-17 | Abr-17 | May-17 | Jun-17 | Jul-17 | ECM General |
|-----------------|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------------|
| <b>BSTS</b>     | 0.000  | 0.015   | 0.026  | 0.019  | 0.021  | 0.008  | 0.046  | 0.034  | 0.104  | 0.104  | 0.061  | 0.004  | 0.037       |
| Estado Espacio  | 0.032  | 0.003   | 0.031  | 0.003  | 0.049  | 0.024  | 0.113  | 0.028  | 0.228  | 0.188  | 0.110  | 0.007  | 0.068       |
| VAR Freq. Mixta | 0.158  | 0.061   | 0.231  | 0.025  | 0.101  | 0.349  | 0.366  | 0.053  | 0.353  | 0.036  | 0.008  | 0.158  | 0.158       |

| Canadá          | Ago-16 | Sept-16 | Oct-16 | Nov-16 | Dic-16 | Ene-17 | Feb-17 | Mar-17 | Abr-17 | May-17 | Jun-17 | Jul-17 | ECM General |
|-----------------|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------------|
| <b>BSTS</b>     | 0.003  | 0.065   | 0.019  | 0.005  | 0.176  | 0.079  | 0.195  | 0.033  | 0.095  | 0.054  | 0.071  | 0.049  | 0.070       |
| Estado Espacio  | 0.034  | 0.030   | 0.081  | 0.088  | 0.106  | 0.090  | 0.241  | 0.066  | 0.081  | 0.045  | 0.075  | 0.060  | 0.083       |
| VAR Freq. Mixta | 0.285  | 0.080   | 0.131  | 0.528  | 1.232  | 1.058  | 1.380  | 1.966  | 1.971  | 0.080  | 0.170  | 0.305  | 0.826       |

| Alemania        | Ago-16 | Sept-16 | Oct-16 | Nov-16 | Dic-16 | Ene-17 | Feb-17 | Mar-17 | Abr-17 | May-17 | Jun-17 | Jul-17 | ECM General |
|-----------------|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------------|
| <b>BSTS</b>     | 0.013  | 0.111   | 0.033  | 0.011  | 0.062  | 0.118  | 0.046  | 0.035  | 0.058  | 0.010  | 0.097  | 0.010  | 0.050       |
| Estado Espacio  | 0.019  | 0.090   | 0.014  | 0.053  | 0.119  | 0.165  | 0.057  | 0.053  | 0.043  | 0.070  | 0.107  | 0.010  | 0.061       |
| VAR Freq. Mixta | 0.041  | 0.073   | 0.089  | 0.032  | 0.236  | 0.215  | 0.037  | 0.018  | 0.030  | 0.100  | 0.229  | 0.157  | 0.101       |

\*BSTS=Modelo Bayesiano Estructural de Series de Tiempo.

\*\*El Error Cuadrático Medio (ECM) fue computado para un horizonte temporal de dos períodos, h=2

**Tabla 5 .** Bondad de Ajuste del Modelo BSTS vs Modelo Estado-Espacio.

| Bondad de Ajuste Modelos: EEUU |      |                |
|--------------------------------|------|----------------|
| Periodo                        | BSTS | Estado-Espacio |
| Ago-16                         | 0.90 | 0.87           |
| Sep-16                         | 0.80 | 0.85           |
| Oct-16                         | 0.95 | 0.9            |
| Nov-16                         | 0.87 | 0.75           |
| Dic-16                         | 0.80 | 0.84           |
| Ene-17                         | 0.92 | 0.85           |
| Feb-17                         | 0.98 | 0.82           |
| Mar-17                         | 0.95 | 0.91           |
| Abr-17                         | 0.97 | 0.84           |
| May-17                         | 0.98 | 0.80           |
| Jun-17                         | 0.96 | 0.71           |
| Jul-17                         | 0.83 | 0.88           |
| R-Cuadrado General             | 0.91 | 0.87           |

| Bondad de Ajuste Modelos: Canadá |      |                |
|----------------------------------|------|----------------|
| Período                          | BSTS | Estado-Espacio |
| Ago-16                           | 0.98 | 0.96           |
| Sep-16                           | 0.97 | 0.93           |
| Oct-16                           | 0.97 | 0.94           |
| Nov-16                           | 0.88 | 0.96           |
| Dic-16                           | 0.99 | 0.95           |
| Ene-17                           | 0.96 | 0.97           |
| Feb-17                           | 0.96 | 0.98           |
| Mar-17                           | 0.98 | 0.98           |
| Abr-17                           | 0.98 | 0.98           |
| May-17                           | 0.98 | 0.92           |
| Jun-17                           | 0.99 | 0.98           |
| Jul-17                           | 0.91 | 0.91           |
| R-Cuadrado General               | 0.96 | 0.94           |

| Bondad de Ajuste Modelos: Alemania |      |                |
|------------------------------------|------|----------------|
| Período                            | BSTS | Estado-Espacio |
| Ago-16                             | 0.83 | 0.87           |
| Sep-16                             | 0.85 | 0.83           |
| Oct-16                             | 0.95 | 0.90           |
| Nov-16                             | 0.90 | 0.75           |
| Dic-16                             | 0.87 | 0.84           |
| Ene-17                             | 0.97 | 0.85           |
| Feb-17                             | 0.92 | 0.82           |
| Mar-17                             | 0.98 | 0.91           |
| Abr-17                             | 0.97 | 0.80           |
| May-17                             | 0.65 | 0.80           |
| Jun-17                             | 0.85 | 0.71           |
| Jul-17                             | 0.80 | 0.88           |
| R-Cuadrado General                 | 0.88 | 0.85           |

**Notas:** 1/ Para el caso del Modelo de Frecuencia Mixta, no se computa el R-cuadrado, ya que no se estima una ecuación particular para cada indicador.

**Tabla 6.** Coeficientes de las variables con probabilidades por encima del umbral de inclusión 1/.

| País           | Variables                   | Coeficiente | Probabilidad de Inclusión |
|----------------|-----------------------------|-------------|---------------------------|
| Estados Unidos | Lugares Vacaciones Familias | 0.0048      | [0.75]                    |
|                | Playas Samaná               | 0.0008      | [0.27]                    |
|                | Lugares para Rentar         | 0.0006      | [0.26]                    |
|                | Vacaciones Express          | 0.0004      | [0.25]                    |
| Canada         | Bavaro                      | 0.0087      | [0.62]                    |
|                | Resorts & Spa               | 0.0052      | [0.48]                    |
|                | Vacaciones en Rep. Dom      | 0.0036      | [0.22]                    |
|                | Vuelos Baratos a Rep.Dom.   | 0.0006      | [0.18]                    |
| Alemania       | Airbnb                      | 0.0036      | [0.89]                    |
|                | Skyscanner                  | 0.0019      | [0.36]                    |
|                | Dominikanische              | 0.0042      | [0.16]                    |
|                | oberstdorf.de               | -0.0006     | [0.07]                    |

**Notas:** 1/ Los coeficientes de estas variables, así como también la probabilidad de inclusión, puede variar de acuerdo al número de iteraciones seleccionados, así como también y a los datos; del mismo modo, es importante acotar que cada vez que se resuelve el modelo puede ser distinto. Una posible solución a este problema es fijar las variables en el componente de regresión; no obstante, fijar estos coeficientes no permite la incorporación de nueva información.

2/ Las probabilidades de inclusión fueron 0.031, 0.04 y 0.026 para EEUU, Canadá y Alemania, respectivamente.

\*El número de iteraciones fue 1,500.

## REFERENCIAS

- [1] Archer, B. (1987). "Demand forecasting and estimation". Travel, tourism and hospitality research, Published by: John Wiley & Sons, Inc.
- [2] Armesto, M. Engemann, K. & Owyang, M. (2010) "Forecasting with Mixed Frequencies". Federal Reserve Bank of St. Louis Review, November- December 2010, 92(6), pp. 521-36.
- [3] Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N.; Csaki, F., 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akademiai Kiado, pp. 267-281.
- [4] Artola, C. & Galan, E. (2012). "Tracking the future on the Web: Construction of Leading Indicators using Internet Searches". Bank of Spain, Working Paper No. 1203.
- [5] Askitas N. & Zimmermann K. (2009). "Google econometrics and unemployment forecasting". Applied Economics Quarterly: Vol. 55, No. 2, pp. 107-120.
- [6] Bankwest Curtin Economics Centre (2017). "The Impact of Airbnb on WA's Tourism Industry".
- [7] Bholat, D. (2016). "Big Data and Central Banks". Bank of England, Advanced Analytics Division, Quarterly Bulletin.
- [8] Brodersen, K., Gallusser, F., Koehler, J., Remy, N. & Scott, S. (2015). "Inferring causal impact using Bayesian structural time-series model". Annals of Applied Statistics, Vol. 9, No. 1, pp. 247-274.
- [9] Bortoli, C. & Combes, S. (2015). "Contribution from Google Trends for forecasting the short-term economic outlook in France: limited avenues". National Institute of Statistics and Economic Studies (INSEE).
- [10] Box, G.; Jenkins, G. (1970). "Time Series Analysis: Forecasting and Control". San Francisco: Holden-Day.
- [11] Briers, M. Doucet, A. & Maskell, S. (2010). "Smoothing algorithms for state-space models". Annals of the Institute of Statistical Mathematics (2010) 62:61-89.
- [12] Carriere-Swallow, Y. & Labbe, F. (2010). "Nowcasting with Google Trends in an Emerging Market". Working Paper, Central Bank of Chile, No. 588.

- [13] Chevillon, G., (2005), "Direct multi-step estimation and forecasting", OFCE and Economics Department, University of Oxford. Working Paper No. 2005-10
- [14] Choi, H. and Varian, H. (2008). "Predicting the present with Google Trends". Technical Report, Google Inc.
- [15] Choi, H. & Varian, H. (2012). "Predicting the present with Google Trends". Economic Record 88, pp. 2- 9.
- [16] Craigwell, R. (2007). "Tourism Competitiveness in Small Island Developing States" . United Nations University –World Institute for Development Economic Research. Research Paper No. 2007/19
- [17] Balaguer Jacint & Manuel Cantavella-Jordá (2000). "Tourism as a long run economic growth factor: The Spanish Case". Instituto de Economía Internacional. Departament d'Economia. Universitat Jaume I., WP-EC 2000-10.
- [18] Duda, R., Hart, P. & Stork, D. (2000). "Pattern Classification". Wiley-Interscience; 2nd Edition.
- [19] Durbin, J. & Koopman, S. J. (2001). Time Series Analysis by State Space Methods. Oxford University Press.
- [20] Durbin, J. & Koopman, S. J. (2002). "A simple and efficient simulation smoother for state space time series analysis". Biometrika 89, pp. 603- 616.
- [21] Eugenio-Martin, J., Martin Morales, N. & Scarpa, R. (2004). "Tourism and Economic Growth in Latin American Countries: A Panel Data Approach," FEEM Working Paper No. 26.2004.
- [22] Fahrmeir, L. & Kaufmann, H. (1991). "On kalman filtering, posterior mode estimation and fisher scoring in dynamic exponential family regression". Metrika: International Journal for Theoretical and Applied Statistics, 1991, Vol. 38, issue 1, pp. 37-60.
- [23] Fahrmeir, L. & Tutz, G. (1994). "Multivariate Statistical Modelling based on Generalized Linear Models". Springer Series in Statistics, New York.
- [24] Fahrmeir, L. & Wagenpfeil, S. (1997). "Penalized likelihood estimation and iterative Kalman smoothing for non-Gaussian dynamic regression models. Comp. Stat. Data Analysis 24, 295- 320.
- [25] Fahrmeir, L. (1999). "State space models a brief history and some recent developments". Proceedings of the 52nd ISI Session.
- [26] George, E. & McCulloch, R. (1997). "Approaches for Bayesian Variable Selection". Statistica Sinica, 7, pp. 339-373.

- [27] Gerolimetto, M.(2010). "ARIMA and SARIMA" . Lecture Notes 6, [www.dst.unive.it](http://www.dst.unive.it).
- [28] Geweke, J. & Whiteman,C. (2006). "Bayesian Forecasting". Chapter 1 in Handbook of Economic Forecasting, 2006, vol. 1, pp 3-80 from Elsevier.
- [29] Glusac, E. (2016) "Hotels vs. Airbnb: Let the Battle Begin - The New York Times".
- [30] Guzman, G.(2011). "Internet search behavior as an economic forecasting tool: The case of inflation expectations". The Journal of Economic and Social Measurement, Vol. 36, No. 3.
- [31] Hassani, H. & Silva, E. (2015). "Forecasting with Big Data: A Review". Annals of Data Science, Vol. 2, Issue 1, pp 5-19.
- [32] Huijskens, T., (2016), "Long term forecasting with machine learning models", Disponible en: <https://thuijskens.github.io/2016/08/03/time-series-forecasting/>
- [33] Informe Trimestral del Banco Central de la Republica Dominicana, Banco Central, Diciembre 2016. [www.bancentral.gov.do](http://www.bancentral.gov.do).
- [34] Informe Ministerio de Turismo (2016), República Dominicana. (MITUR), <http://mitur.gob.do/>
- [35] Internet Travel & Hotel Booking Statistics, <http://www.statisticbrain.com/internet-travel-hotel-booking-statistics/>
- [36] Ishwaran, H. & Rao, J. (2005). "Spike and slab variable selection: frequentist and Bayesian strategies". Annals of Statistics, 33, pp. 730-773.
- [37] Ishwaran, H., Kogalur, U., & Rao, J. (2010). "Spikeslab: Prediction and variable selection using spike and slab regression"..R Journal, 2(2), 68-73.
- [38] Jansen,B., Ciamacca, C. & Spink, A. (2008). "An Analysis of Travel Information Searching on the Web". Information Technology & Tourism, Vol. 10.
- [39] Jarocinski, M. (2015). "A note on implementing the Durbin and Koopman simulation smoother". Working Paper, European Central Bank, No.
- [40] 1867.
- [41] Jungbacker, B. & Koopman, S. (2014). "Likelihood-based Dynamic Factor Analysis for Measurement and Forecasting".Econometrics Journal, Vol.
- [42] 17, pp. 1-21.
- [43] Kalman, R. (1960). "A new approach to linear filtering and prediction problems". Journal of Basic Engineering 82, pp. 35-45.

- [44] Kulkarni, R., Haynes, H, Stough, R. , Paelinck, J. (2009) "Forecasting Housing Prices with Google Econometrics". GMU School of Public Policy Research Paper No. 2009-10
- [45] Ludwig, N. Feuerriegel, S. & Neumann, D. (2016). "Time Series Analysis for Big Data: Evaluating Bayesian Structural Time Series Using Electricity Prices". Multikonferenz Wirtschaftsinformatik (MKWI) 2016, volume III, pp. 1569-1580.
- [46] Mitchell, T. & Beauchamp, J. (1988). "Bayesian variable selection in linear regression". J. Amer. Stat. Assoc., 83, pp.1023-1036.
- [47] Organización Mundial del Turismo (OMT) (2014). "Recomendaciones de la OMT por un turismo accesible para todos". Resolución A/RES/637(XX) de la Asamblea General de la OMT en agosto de 2013.
- [48] Ogonowska, M. & Torre, D. "Sustainable Tourism and the emergence of new Environmental Norms". 29èmes Journées de Microéconomie Appliquée (JMA), Jun 2012, Brest, France.
- [49] Pérez, A. (2011). "Inversión Turística y desarrollo en República Dominicana
- [50] Pnevmatikakis, E. Rad, K.,Huggins, J. & Paninski, L.(2014) "Fast Kalman Filtering and Forward-Backward Smoothing via a Low-Rank Perturbative Approach". Journal of Computational and Graphical Statistics, Vol. 23,- Issue 2. pp. 316-339.
- [51] Reyes, P. Artículo Periodístico, Listín Diario 2016  
<https://listindiario.com/economia/2017/01/18/450830/print>
- [52] Schmidt, T. & Vosen, S. (2009). "Forecasting private consumption: Survey-based indicators vs. Google Trends". Ruhr Economic Papers 0155.
- [53] Scott, L., & Varian, H. (2014). "Predicting the Present with Bayesian Structural Time Series." International Journal of Mathematical Modelling and Numerical Optimisation, Vol. 5 (1/2), pp. 4-23.
- [54] Scott, S. (2016). "Package BSTS for R: Time series regression using dynamic linear models fit using MCMC". Cran Project.
- [55] Song, H., Li, G. & Witt,S. & Fei, B. (2010). "Tourism demand modelling and forecasting:how should demand be measured? Journal of Tourism Economics, 2010, 16 (1), 63-81.
- [56] Taieb, S. (2014). "Machine learning strategies for multi-step-ahead time series forecasting" Université Libre de Bruxelles. Belgium. September 2014.

- [57] Taylor, B., Davies, T. Rowlingson, B. & Diggle, P. (2015). "Bayesian Inference and Data Augmentation Schemes for Spatial, Spatiotemporal and Multivariate Log-Gaussian Cox Processes in R". *Journal of Statistical Software*, Vol.63, Issue 7.
- [58] Tisdell, C. and Bandara, R. (2004) "Tourism as a Contributor to Development in Sri Lanka: An Overview and a Case Study", *Working Paper on Economics, Ecology and the Environment*, 94, pp. 1-29.
- [59] Vanderkam,D., Schonberger, R.,Rowley, H. & Kumar, S. (2013). "Nearest Neighbor Search in Google Correlate". Google Research.
- [60] Vellas, F. (2011). "The indirect impact of tourism: An economic analysis". Third Meeting of T20 Tourism Ministers Paris, France, October 2011
- [61] Witt, S. & Witt,C. (1995). "Forecasting tourism demand: A review of empirical research". *International Journal of Forecasting* 11 (1995) 447-475.
- [62] Yu, B., Shenoy, K. & Shahani, M. (2004). "Derivation of Kalman filtering and smoothing equations".<http://www.academia.edu>.