

Comparación de pronósticos para la dinámica del turismo en Medellín, Colombia

Marisol Valencia, Juan Vanegas, Juan Correa y Jorge Restrepo

Marisol Valencia, Juan Vanegas, Juan Correa y Jorge Restrepo

Comparación de pronósticos para la dinámica del turismo en Medellín, Colombia

Resumen: *El turismo es un tema de interés para muchos países, pero en Colombia ha sido poco explorado. Conocer sobre periodos con mayor afluencia de turistas es importante para predecir la cobertura en servicios turísticos. En este trabajo se compara la estimación entre modelos de regresión clásica y bayesianos para elegir la mejor alternativa de predicción de la cantidad de turistas que arribarían a la ciudad de Medellín. Se busca también determinar las variables que más inciden sobre la llegada de turistas y los modelos que permiten ajustar mejor sus dinámicas. Dentro de los resultados obtenidos se observa mejor desempeño de los modelos bayesianos versus los clásicos. Además, la variable mes es significativa para explicar las demandas por parte de turistas colombianos y extranjeros. Los periodos de mayor afluencia de turistas se presentan entre diciembre-enero y junio-julio, comportamiento que se repite anualmente, lo cual es decisivo para la planeación de recursos hoteleros.*

Palabras clave: *demanda turística, evaluación de modelos y selección, métodos de pronóstico y predicción, estadística bayesiana, Medellín.*

Clasificación JEL: *L83, C13, C52, C53, C65, R11.*

Comparing forecasts for tourism dynamics in Medellín, Colombia

Abstract: *Tourism is a topic of interest to many economies around the world, but it has received limited attention in Colombia. Knowing the periods of larger tourist inflows is important for predicting coverage in services for tourists. In this paper, we compare the estimation between classical and Bayesian regression in order to choose the best alternative to predict the number of tourist arrivals to Medellín. We also identify the most significant variables affecting the influx of tourists and the models providing better fit to the associated dynamics. According to our results, the Bayesian approach shows better estimates than the classic one. In addition, the variable month is significant to explain the demands for both Colombians and foreigners. The periods with the highest incidence of visits to the city are December-January and June-July, a pattern that repeats itself every year, which is crucial for planning hotel resources.*

Keywords: *tourism demand, model evaluation and selection, forecasting and prediction methods, Bayesian statistics, Medellín.*

JEL Classification: *L83, C13, C52, C53, C65, R11.*

Comparaison des prévisions pour la dynamique du tourisme dans la ville de Medellín, Colombie

Résumé: *Le tourisme est un sujet de recherche dans des nombreux pays, mais en Colombie il reste peu exploré. Par exemple, le fait de connaître les périodes de l'année avec les plus de touristes permet de mieux prédire la couverture des services touristiques. Cet article établit une comparaison entre les estimations issues à partir des modèles classiques et des modèles de régression bayésienne, afin de choisir la meilleure méthode de prédire le nombre de touristes qui arrivent à la ville de Medellín. L'article vise également une étude de la détermination des variables qui expliquent l'arrivée des touristes, afin d'identifier les modèles qui mieux expliquent sa dynamique. Les résultats obtenus montrent que la performance des modèles bayésiens est beaucoup mieux que celle des modèles classiques. En outre, la variable « mois » est significatif dans l'étude de la demande des services touristiques, aussi dans le cas des touristes colombiens qu'étrangers. Les périodes de l'année avec le plus d'affluence sont Décembre-Janvier et Juin-Juillet. Il s'agit d'une information très utile à la planification des ressources hôteliers.*

Mots-clés: *demande de services touristiques, évaluation et sélection des modèles, méthodes de prédiction et de prévision, statistique bayésienne, Medellín.*

Classification JEL: *L83, C13, C52, C53, C65, R11.*

Comparación de pronósticos para la dinámica del turismo en Medellín, Colombia[†]

Marisol Valencia, Juan Vanegas, Juan Correa y Jorge Restrepo*

–Introducción. – I. Revisión de la literatura. –II. Métodos. –III. Resultados. –IV. Discusión.
–Conclusiones. –Referencias.

doi: 10.17533/udea.le.n86a08

Primera versión recibida el 22 de julio de 2015; versión final aceptada el 6 de febrero de 2016

Introducción

El pronóstico de la demanda del sector turístico es un tema reconocido a nivel mundial. Diversos estudios emplean modelos de pronósticos para su caracterización y análisis de comportamiento, identificando, por ejemplo, covariables con incidencias significativas o patrones estacionales y tendencias

[†] Este artículo es un producto derivado del proyecto de investigación “Flujos turísticos en Colombia: perspectivas y determinantes”, proyecto financiado por la Fundación Universitaria Autónoma de las Américas.

* *Marisol Valencia Cárdenas*: Docente, Institución Universitaria Tecnológico de Antioquia. Dirección electrónica: mvalencia@unal.edu.co.

Juan Gabriel Vanegas López: Docente-Investigador en la Fundación Universitaria Autónoma de las Américas. Facultad de Ciencias Económicas y Administrativas, Grupo de Investigación GICEA, Medellín, Colombia. Dirección electrónica: juan.vanegas@uam.edu.co; jg.tecnologico@gmail.com

Juan Carlos Correa Morales: Profesor Asociado, Universidad Nacional de Colombia, Sede Medellín. Dirección electrónica: jccorrea@unal.edu.co.

Jorge Aníbal Restrepo Morales: Docente, Institución Universitaria Tecnológico de Antioquia. Dirección electrónica: jrestrepo@tdea.edu.co

(Guizzardi & Mazzocchi, 2010). Muchos de estos modelos van desde la teoría clásica, o econométrica, hasta propuestas que comparan modelos clásicos con redes neuronales, como sucede en el trabajo de Claveria y Torra (2014).

Algunos temas relacionados con la teoría de modelos de pronósticos clásicos, como es el caso de la regresión, presentan problemas asociados con la presencia de sesgos al pronosticar, debido a que cuando se prolonga el período de predicción en el modelo con parámetros fijos se puede incurrir en predicciones erróneas. En otros casos hay algunos supuestos de los residuales que no se cumplen con facilidad, por presentar problemas como valores muy extremos, o un cambio repentino en el curso normal de una serie de tiempo.

Cuando la escala de la variable de respuesta inicial es de tipo discreto se pueden usar modelos de pronóstico basados en la distribución Poisson, y si toma valores de respuesta muy altos (más de 100) es posible realizar pronósticos aproximados con modelos basados en datos con distribución Normal. Otros estudios aplicados a datos de conteo muestran algunos modelos simples como distribuciones empíricas o modelos de suavización exponencial. Kolasa (2016) compara pronósticos realizados con distribuciones empíricas, con la Poisson y Binomial Negativa; por otro lado, Wallström y Segerstedt (2010) comparan el método de suavización exponencial con otros como el de Croston.

Se observa, entonces, que la literatura exhibe estudios que muestran resultados de pronósticos elaborados con modelos clásicos basados en la distribución normal o en la Poisson, pero también estudios de modelos con diferentes teorías de estimación como la bayesiana (Zellner, 1996; Valencia, 2016).

La comparación de pronósticos de la demanda del sector turístico en Medellín es un tema que no ha sido abordado en Colombia usando modelos de regresión clásica y de regresión Bayesiana. En este trabajo se propone comparar un modelo de pronósticos econométrico clásico con dos bayesianos: uno basado en la regresión del autor Zellner, que asume una distribución normal para los datos y parámetros, y otro basado en MCMC (Monte Carlo por Cadenas de Markov), que asume la distribución Poisson para los datos y parámetros bajo la Distribución Normal Multivariada. Estos modelos son aplicados al caso de entradas de turismo a la ciudad de Medellín, Colombia,

sin modificar las distribuciones de probabilidad iniciales para este trabajo. Esto puede proponerse en futuras aplicaciones, incluso, con diferentes formas de muestreo.

Con esta investigación se busca responder los siguientes interrogantes: ¿cuáles variables tienen más incidencia sobre la llegada de turistas a la ciudad (fechas, clima, estaciones, meses)?, ¿con cuáles modelos es posible ajustar mejor la dinámica de demanda de turismo?, ¿la teoría Bayesiana es útil para efectuar pronósticos sobre la demanda de turismo?

Este artículo se presenta en el siguiente orden. En la sección I, se muestra una revisión de literatura. En la sección II, se muestran los modelos a estimar e indicadores comparativos. En la sección III, se presentan los resultados, donde se mostrará una exploración de datos y los modelos estimados para ambas series de turistas (colombianos y extranjeros). La sección IV muestra la discusión de resultados. Finalmente se presentan las conclusiones y referencias.

I. Revisión de la literatura

El turismo en el mundo ha sido un tema de interés para la economía de muchos países (Crouch, 1995; Guizzardi & Mazzocchi, 2010; Lim, 1999; Peng, Song & Crouch, 2014; Zhou, Bonham & Gangnes, 2007). La demanda en este sector ha sido analizada desde el punto de vista de pronóstico en numerosos estudios a nivel internacional (Guizzardi & Stacchini, 2015; Song, Gao & Lin, 2013; Liu et al., 2014). Algunas de las revisiones literarias muestran una descripción completa de los métodos usados en la modelación y pronóstico de las demandas turísticas, que van desde las perspectivas econométricas clásicas hasta las nuevas técnicas de pronóstico emergentes, que emplean teoría estadística o matemática (Witt & Witt, 1995; Li, Song & Witt, 2005; Song & Li, 2008).

No obstante, en Colombia algunos estudios realizados se han centrado en el análisis del gasto, o en los servicios que se ofrecen al turista. Dentro de este grupo de trabajos se puede resaltar el efecto de la seguridad y el comercio, abordado por Bonilla y Moreno (2010), quienes, mediante un mo-

delo de datos de panel, encuentran que los arribos de viajeros extranjeros están inversamente relacionados con los secuestros, y el índice de intercambio comercial lo hace de forma positiva. Otros trabajos han analizado las dinámicas locales de los movimientos de viajeros. Cerda y Leguizamón (2005), por ejemplo, encuentran por medio de modelos hedónicos que la demanda interna de agentes nacionales por consumo de un bien turístico depende en gran manera del perfil del jefe de hogar, su poder adquisitivo y la composición familiar. A nivel de municipalidades o locaciones específicas, por ejemplo el caso de Cartagena, los estudios se centran en observar el impacto de las fluctuaciones cambiarias en la demanda turística (Galvis & Aguilera, 1999)¹. Ahora bien, no se ha hecho un comparativo en pronósticos que evalúe los efectos estacionales de las demandas turísticas para la ciudad de Medellín, caracterizando a su vez, algunos factores controlables, que son importantes para proveer herramientas de planeación para el sector turístico y fomentar estrategias de mejoras en los servicios que se ofertan.

El tipo de modelos propuestos para el pronóstico, a nivel internacional, varía acorde con algunas características de las series de demandas turísticas, por ejemplo, hay estudios de su comportamiento lineal o no lineal (Chen, 2011); pero también se estudia la relación de flujos de turistas con las capacidades de alojamientos (Zhou et al., 2007), o la inherencia de esta demanda con su ciclo de negocios (Guizzardi & Mazzocchi, 2010).

Existen modelos en los que se relaciona el pronóstico de demandas, con la optimización de recursos, como el caso de demandas e inventarios; para ello, se han usado muchos modelos estadísticos clásicos (Bes & Sethi, 1988; Gutiérrez & Vidal, 2008; Choi, Li & Yan, 2003; Samaratunga, Sethi & Zhou, 1997; Sarimveis et al., 2008) y luego su salida es usada para modelar los inventarios. Vidal, Londoño y Contreras (2004), por ejemplo, implementan técnicas sencillas de pronósticos y de control de inventarios. Cohen y Dunford (1986) proponen un modelo basado en una distribución empírica para la demanda y afirman que el requerimiento de una distribución normal para los residuales de los modelos ARIMA (Modelos Integrados Autorregresivos de

¹ También, se destaca el trabajo de López y Márquez (2012), en el que formulan un observatorio para el análisis del turismo y presentan su organización; sin embargo, a nivel de pronósticos de demanda, son pocos los trabajos existentes.

Medias Móviles) y de regresión no se cumple en muchos casos. Shoensmith y Pinder (2001) emplean modelos VAR (Vector Autorregresivo) y BVAR (Bayesian Vector Autorregresivo), usando pocos datos de la demanda de los productos estudiados.

Frente a modelos para datos de conteo, existen trabajos comparativos recientes, basados en distribuciones empíricas, Poisson y Binomial negativa, presentados en Kolasa (2016), el cual está basado en Syntetos et al. (2011). A partir de una comparación de indicadores de error con probabilidad, estos trabajos encuentran un mejor desempeño de modelos basados en distribución empírica en comparación con los de la distribución Poisson y Bootstrap. Por su parte, Wallström y Segerstedt (2010) presentan el suavizado exponencial simple (SES-Single Exponential Smoothing) versus Croston y algunas variaciones, aplicados para demandas intermitentes o no², y encuentra que el modelo SES tiene mejor desempeño frente al de Croston cuando utiliza diferentes indicadores de error para elegir el mejor modelo. Para el caso del tema turístico, se encuentra el trabajo de Hellström (2002) que estima modelos de valor entero autorregresivos de media móvil (INARMA) para estudiar la interacción entre la elección del número de viajes de ocio de los hogares y el número de pernoctaciones.

La estadística Bayesiana ha ganado importancia para efectuar pronósticos hace algún tiempo. Por ello, se han elaborado diferentes trabajos y textos basados en inferencia predictiva, valor esperado o modelos dinámicos bayesianos que muestran su utilidad para elaborar pronósticos, de forma univariada o multivariada, basados en información a priori para los parámetros, para los datos y, quizá, información de expertos (Fei, Lu & Liu, 2011; Harrison & Stevens, 1976; De Mol, Giannone & Reichlin, 2008; Neelamegham & Chintagunta, 1999; West & Harrison, 1997; Valencia, 2016). Un modelo de regresión Bayesiana, que es desarrollado en Zellner (1996), se basa en una distribución no informativa para los parámetros y una distribución normal para los datos, lo cual es posible suponer de manera aproximada cuando la serie de datos tienen valores muy grandes, así sean discretos, como son la cantidad de turistas de las series temporales de este trabajo. Este modelo será usado para pronos-

² Modelos que han sido usados por autores como Brown (1959; 1963).

ticar la serie mensual de turistas que ingresan a la ciudad de Medellín, en el periodo temporal de 2007 a 2014, y su ajuste será comparado con un modelo de pronósticos econométrico clásico, y otro bayesiano basado en MCMC (con MCMCpoisson), para determinar así, cual es el mejor modelo que facilite la mejor interpretación del comportamiento de dicha serie de turistas.

II. Métodos

Para iniciar el análisis en series temporales, se hace un corte en la serie de datos, de N valores, T ($T < N$) serán usados para analizar correlación y análisis descriptivo y para estimar modelos; los $N - T$ restantes serán usados para validar la estimación y para determinar si el modelo tiene una capacidad de pronóstico adecuada. Se realiza un análisis estadístico detectando inicialmente valores atípicos y problemas de variabilidad de los valores de demanda turística de la ciudad de Medellín.

A. Modelo econométrico

El modelo econométrico es un modelo de regresión que contempla variables autorregresivas endógenas, así como otras exógenas que pueden estar asociadas a la respuesta z_t (Caridad y Ocerin, 1998; Bowerman, Oconnell & Koheler, 2007), partiendo de la ecuación general dada por (1):

$$z_t = \beta_0 + \dots \beta_k z_{t-k} + \dots \beta_i X_i + \varepsilon. \quad (1)$$

Dentro de las covariables a estudiar se encuentran: rezagos de primer orden y estacional, el tiempo, una variable sinusoidal, variable mes como factor, temperatura, IPC, TRM y el consumo de energía por actividad económica. Se realizará un proceso de selección de variables de tipo paso a paso, validando el modelo final con las premisas de residuos (normalidad, varianza e incorrelación) y obteniendo al final el indicador de ajuste MAPE (error medio absoluto porcentual) y error relativo de pronóstico, con mayor eficiencia para la predicción. Estos indicadores de error serán comparados con el modelo de regresión clásica.

El MAPE suele ser útil para expresar de una forma simple, en términos porcentuales genéricos, el error cometido, incluso para alguien que no tiene idea de lo que constituye un error de “grande” como, por ejemplo, lo relativo a grandes montos de dinero; en especial, este indicador puede ser adecuado para cantidades que son mayores de una unidad. Cuando toma valores cercanos a cero, se recurre más a menudo a utilizar el indicador MSE (Mean Square Error), RMSE (Root of Mean Square Error), sMSE (scale Mean Square Error) (Kolasa, 2016), o incluso la medida sMAPE (symmetric Mean Absolute Percentage Error) (Wallström & Segerstedt, 2010), ya que no incrementa más el porcentaje de error en estos casos cercanos al cero. Autores como Petris, Petrone y Campagnoli (2009), quienes muestran en su libro aplicaciones de los modelos dinámicos bayesianos usando el software R, mencionan que el MAPE es una de las estadísticas más usadas para comparación de pronósticos, y se utiliza junto con otras dos: MAD (Mean Absolute Deviation) y MSE o RMSE, para comparar modelos.

En este trabajo se usan los criterios RMSE, MAPE y sMAPE (ver ecuaciones (2), (3) y (4)) para medir el desempeño en el ajuste y pronóstico sobre estimaciones de modelos clásicos y bayesianos, con respuestas que toman valores muy altos y diferentes de cero, como son las series de tiempo de cantidad de turistas usadas aquí. Además de esto, se realizan procesos de selección y pruebas con relación a la significancia de covariables y, adicionalmente, pruebas de normalidad.

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \frac{|Z_t - \hat{Z}_t|}{Z_t}. \quad (2)$$

$$\text{sMAPE} = \frac{1}{T} \sum_{t=1}^T \frac{|Z_t - \hat{Z}_t|}{(Z_t + \hat{Z}_t)/2}. \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (Z_t - \hat{Z}_t)^2}. \quad (4)$$

B. Modelo de regresión lineal bayesiano basado en la normal para los datos

En la teoría de la toma de decisiones, la probabilidad juega un papel fundamental. Las probabilidades previas de estados de la naturaleza pueden ser consideradas a priori con un carácter subjetivo, pero pueden diferir un poco de las reales (Hillier & Hillier, 2007). Con base en los principios del teorema de Bayes, al combinar el producto entre una probabilidad a priori con una condicionada a un hecho se obtiene lo que se conoce como probabilidad a posteriori. Estos principios son utilizados por la estadística bayesiana para inferir, por ejemplo, usando predicciones (Gill, 2007; Gelman et al., 2004).

En la Estadística Bayesiana, los parámetros de la ecuación de regresión no son fijos, sino que se pueden comportar acorde con una función de distribución, llamada a priori. Al utilizar el teorema de Bayes, es posible usar la premisa acerca de que el producto entre la distribución a priori para el parámetro y la verosimilitud de los datos genera un conocimiento de dicho parámetro, por medio de la distribución a posteriori. Zellner (1996) realiza la construcción del modelo de regresión a partir de las premisas: una distribución a priori no informativa: $1/\sigma$, para los parámetros β, τ del modelo, y una distribución normal para los datos; con ello se construye la verosimilitud de la muestra. El producto entre dicha distribución a priori y la verosimilitud genera la a posteriori de dichos parámetros. Seguido a esto, la integral entre la distribución de los datos, multiplicada por la distribución a posteriori, genera la distribución predictiva bayesiana, función finalmente usada para pronósticos de la variable respuesta.

Zellner (1996) propone el uso de una distribución a priori no informativa: $1/\sigma$ para los parámetros. La función de verosimilitud de los datos, basada en la distribución normal, está dada por (5):

$$L(y_t | y_0, \beta) \propto \tau^T \exp^{-\frac{\tau}{2}(Y-X\beta)'(Y-X\beta)}. \quad (5)$$

Luego del producto entre la distribución a priori por la verosimilitud (5) se obtiene la distribución a posteriori (6):

$$\xi(\beta, \tau \mid \beta_0, \tau_0, y_0, y_t) \propto \tau^{\frac{T+1}{2}} \exp^{-\frac{\tau}{2} \left[A \left((\beta - \hat{\beta})' A^{-1} (X'X)(\beta - \hat{\beta}) + 1 \right) \right]}, \quad (6)$$

donde $A = vs^2$ y $\hat{\beta} = (X'X)^{-1} X'Y$.

La distribución predictiva, obtenida luego de hacer la integral entre la distribución de los datos y la distribución a posteriori (6), es una distribución T de Student (Zellner, 1996), con la cual se generan las estimaciones de la respuesta buscada, en este caso: turistas colombianos y turistas extranjeros, usando X como la matriz de diseño donde se incluyen las covariables.

C. Modelo de regresión Poisson bayesiano

Este modelo parte de las premisas de una distribución Poisson para los datos y una distribución a priori Normal Multivariada para el vector de parámetros β , que permiten obtener una distribución a posteriori de los parámetros del modelo de regresión. La estimación del modelo puede hacerse con el paquete MCMCpack del programa R (Martin, Quinn & Park, 2011), cuya función MCMCpoisson, permite muestrear los parámetros de dicha distribución a posteriori, usando un algoritmo llamado Metropolis de caminata aleatoria. Con esto se obtienen estadísticas para cada uno de los parámetros, como la media o los percentiles, que se reemplazan como parámetros del modelo de regresión y, de ésta forma, encontrar la respuesta media, que es el logaritmo natural de la serie original de los datos.

En resumen, la base teórica parte de una distribución de los datos, $y_i \sim Poisson(mu_i)$, con la función link inversa para la media $mu_i = \exp(X\beta)$, y el vector de parámetros $\beta \sim N(b_0, B_0^{-1})$. Luego de hacer el producto de la distribución a priori por la de los datos, se obtiene la distribución a posteriori, de la cual se establecen condicionales por cada parámetro, y con éstas se hace el muestreo MCMC. De este modo se obtienen estadísticas de los parámetros β para la función inversa: $\hat{Z} = Ln(mu_i) = X\beta$, por ejemplo, la media y la mediana, y con cada una de éstas es posible escribir el modelo ajustado, pero

es necesario transformar nuevamente la respuesta estimada \hat{Z} , con $\exp(Z) = mu_i$ en la escala original.

III. Resultados

Variables respuesta: Serie A: Turistas colombianos, Serie B: Turistas extranjeros. Periodo: años 2007 a 2014. $N = 96$, $T = 72$, periodos de pronóstico: 24 (2 años).

A. Análisis exploratorio

Es común medir la actividad económica por medio del PIB; no obstante, las cifras estadísticas oficiales pueden presentar un sesgo optimista sobre el crecimiento de la economía. Por tanto, se puede interpretar un indicador efectivo como el IPC para que los agentes económicos de la industria del turismo en Colombia puedan establecer unas asociaciones posibles y, con ello, utilizar dichas relaciones en los procesos de toma de decisiones. Similarmente, la actividad económica asociada al consumo de energía puede servir, en general, para diagnosticar si las tendencias del turismo en Colombia se relacionan o no con dicho consumo, sin excluir el sector informal de la economía³, estimando su efecto con los modelos que se presentan aquí.

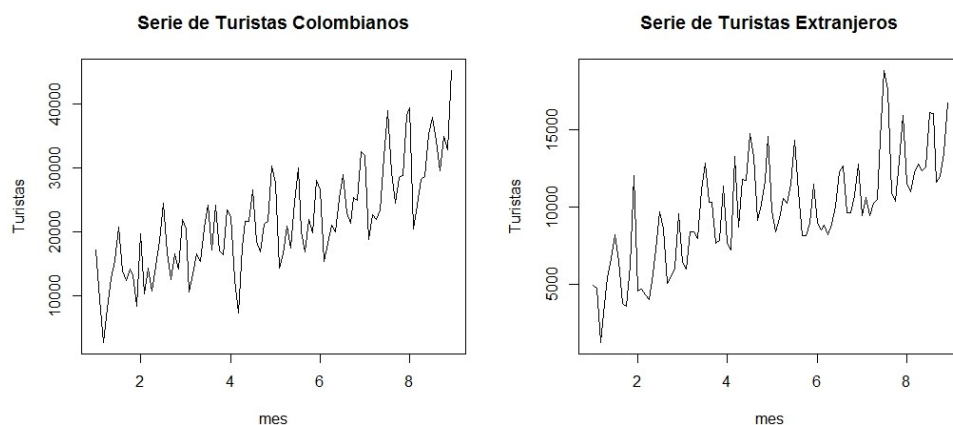
Para alcanzar diagnósticos aceptables, se probaron diferentes variables utilizadas en la literatura, optando por explorar factores e índices que, por su naturaleza, guardan una relación estrecha con el dinamismo de la economía y, por supuesto, con el turismo. Tal es el caso de la cantidad de kilovatios hora de energía gastados en la ciudad en actividades económicas residencial, industrial y comercial para el mes respectivo, covariables que son usadas en la exploración de los modelos clásicos. Sin embargo, como se verá, no todas estas variables muestran asociaciones importantes con el flujo turístico.

Las series tienen tendencias lineales crecientes, con una variación que podría ser cíclica (Gráfico 1). La característica de las variables discretas puede

³ Según el presidente ejecutivo de la Asociación Hotelera y Turística de Colombia (Cotelco), Gustavo Adolfo Toro, la informalidad alcanza un 30 % del gremio (Agencia EFE, 2014).

ocasionar que sea necesaria una transformación en los datos para inferir acerca de su comportamiento; sin embargo, para efectos de comparar la capacidad predictiva del modelo regresivo normal con el bayesiano, no se realizará dicha transformación.

Gráfico 1. *Serie de tiempo: turistas colombianos, extranjeros, porcentaje de ocupación hotelera*

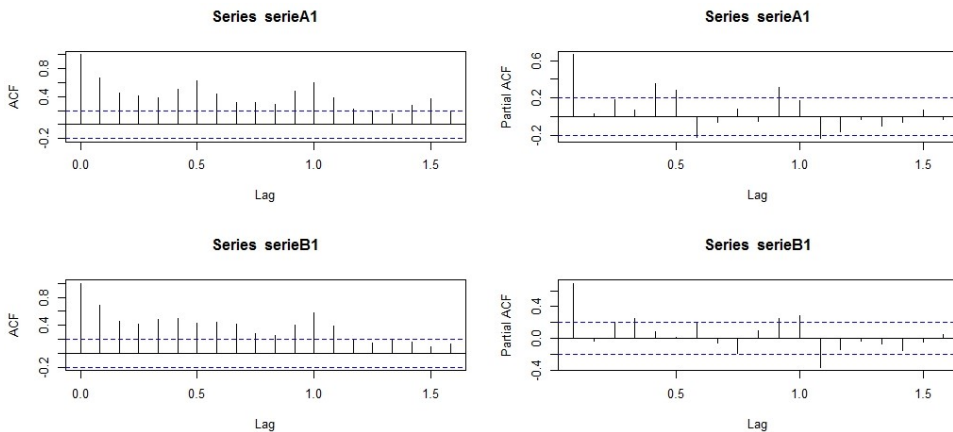


Fuente: elaboración propia.

La serie A (turistas colombianos) presenta autocorrelación estacional, tal como se aprecia en los valores de autocorrelaciones significativas saliendo de las bandas de confianza, en los periodos estacionales semestrales en el Gráfico 2. Similarmente ocurre para la serie B (turistas extranjeros). Las características encontradas en las series temporales muestran la necesidad de incorporar componentes que modelen estacionalidad y correlación (como variables autorregresivas), así como variables indicadoras (que toman valores 1 o 0) (como el mes del año) o trigonométricas (senuidal o cosenoidal), probando hasta encontrar las variables con mayor significancia en el modelo. Si bien las series evidentemente no son estacionarias, dado que los valores de autocorrelación de los periodos entre 1 y L son estadísticamente diferentes de cero (ver Gráfico 2), la estimación del modelo de regresión tiene grandes ventajas por la inclusión de covariables que permiten modelar los componentes estacionales, como variables indicadoras y variables rezagadas que capturan este tipo de

comportamientos y correlación de la serie. De esta manera se pueden generar residuales incorrelacionados, como se demuestra con el análisis presentado.

Gráfico 2. *Autocorrelación y autocorrelación parcial para las series de tiempo: turistas colombianos, extranjeros*

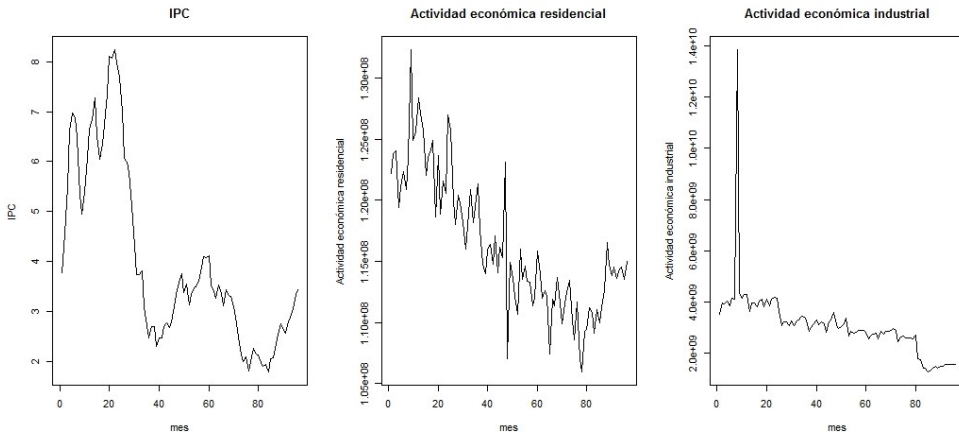


Fuente: elaboración propia.

Al realizar un análisis del comportamiento de variables, como IPC, y el gasto energético de actividades económicas, como residencial e industrial (ver Gráfico 3), se aprecia un aparente grado de disminución en el tiempo, lo que es contrario a las series de turistas que tienen tendencia a incrementarse.

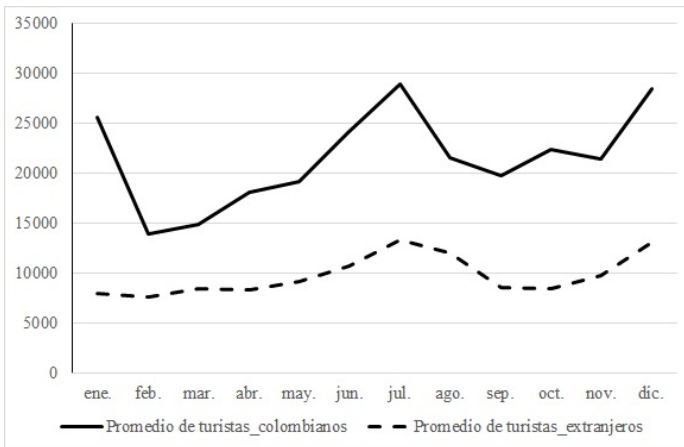
La exploración de los promedios de turistas por mes, que se observan en el Gráfico 4, refuerza el análisis de la incidencia de dicho factor sobre las series, pues se aprecia que en los meses de junio, julio, enero y diciembre es mayor el total de turistas colombianos, mientras que en los meses de julio, agosto y diciembre vienen más turistas extranjeros, ya que estos meses corresponden a las vacaciones de los países respectivos.

Gráfico 3. *Comportamiento de variables explicativas IPC, actividad económica residencial, actividad económica industrial*



Fuente: elaboración propia.

Gráfico 4. *Medias de turistas colombianos y extranjeros por mes*



Fuente: elaboración propia.

B. Modelo lineal econométrico

Los modelos se estiman tomando en consideración la autocorrelación detectada previamente, para ello se incluyen rezagos de la misma serie; para la

estacionalidad detectada, se incluye el mes como variable indicadora. Estos modelos ayudan de manera exploratoria a detectar los efectos más significativos en relación con las covariables mencionadas, y con éstas mismas se pueden ajustar los modelos bayesianos posteriormente.

La Tabla 1 muestra las tablas ANOVA tipo III estimadas para cada una de las series analizadas con los modelos de regresión.

Tabla 1. ANOVA tipo III para los dos modelos estimados

Serie A: Turistas colombianos				
	Sum Sq	Df	F value	$Pr(> F)$
(Intercepto)	90219940	1	7,843	0,00638
Z_{t-1}	1096719834	1	95,34	2,47E-15
Indicadora mes	1944928987	11	15,37	1,18E-15
Activ_económica.res	66144896	1	5,75	0,01879
Residuales	931775044	81		
Serie B: Turistas extranjeros				
	Sum Sq	Df	F value	$Pr(> F)$
(Intercepto)	32025832	1	16,81	9,86E-05
Z_{t-1}	63860544	1	33,52	1,33E-07
Indicadora mes	310503156	11	14,82	3,67E-15
IPC	19151493	1	10,05	0,00216
Activ_económica.ind	9283949	1	4,873	0,03015
Residuales	152420303	80		

Fuente: elaboración propia.

En las tablas ANOVA se aprecia que todas las variables quedan significativas para explicar la respuesta (valor $p < 0,05$) en los respectivos modelos, luego del proceso de selección por pasos. Los coeficientes estimados se aprecian en la segunda columna de la Tabla 2. El rezago de orden 1 (Z_{t-1}) es significativo para explicar la cantidad de turistas, tanto colombianos como extranjeros, lo cual indica la dependencia con respecto al valor del mes anterior.

Tabla 2. *Valores de los coeficientes para los modelos estimados*

Coeficientes del modelo para Serie A: Turistas colombianos				
Coeficientes:	Estimate	Std error	t value	P-value ($P(> t)$)
(Intercepto)	3,18E+04	1,14E+04	2,801	0,006378
Z_{t-1}	7,62E-01	7,80E-02	9,764	2,47E-15
factor(mes)ago.	-7,25E+03	2,02E+03	-3,598	0,000551
factor(mes)dic.	5,56E+03	1,79E+03	3,101	0,002654
factor(mes)ene.	4,90E+02	2,00E+03	0,245	0,80688
factor(mes)feb.	-1,23E+04	1,90E+03	-6,5	6,10E-09
factor(mes)jul.	3,58E+03	1,82E+03	1,966	0,052774
factor(mes)jun.	2,60E+03	1,72E+03	1,509	0,135306
factor(mes)mar.	-2,67E+03	1,70E+03	-1,568	0,120766
factor(mes)may.	-1,44E+03	1,71E+03	-0,843	0,401463
factor(mes)nov.	-2,26E+03	1,81E+03	-1,251	0,214582
factor(mes)oct.	5,48E+02	1,73E+03	0,316	0,752588
factor(mes)sep.	-3,28E+03	1,80E+03	-1,825	0,071751
Act_ecca_res	-2,15E-04	8,98E-05	-2,398	0,018788

(Continúa)

Tabla 2. *Continuación*

Coeficientes del modelo para Serie B: Turistas extranjeros				
Coeficientes:	Estimate	Std error	<i>t</i> value	P-value ($P(> t)$)
(Intercepto)	6,19E+03	1,51E+03	4,1	9,86E-05
Z_{t-1}	5,57E-01	9,63E-02	5,789	1,33E-07
factor(mes)ago.	1,30E+03	8,98E+02	1,45	1,51E-01
factor(mes)dic.	4,01E+03	7,02E+02	5,707	1,87E-07
factor(mes)ene.	-2,30E+03	8,10E+02	-2,84	5,72E-03
factor(mes)feb.	-4,16E+02	6,93E+02	-0,601	5,50E-01
factor(mes)jul.	3,78E+03	7,25E+02	5,209	1,45E-06
factor(mes)jun.	2,10E+03	6,94E+02	3,027	3,33E-03
factor(mes)mar.	5,13E+02	6,95E+02	0,738	4,63E-01
factor(mes)may.	8,91E+02	6,91E+02	1,29	2,01E-01
factor(mes)nov.	1,52E+03	6,90E+02	2,202	3,05E-02
factor(mes)oct.	4,81E+00	6,90E+02	0,007	9,94E-01
factor(mes)sep.	-1,72E+03	7,68E+02	-2,241	2,78E-02
IPC	-4,13E+02	1,30E+02	-3,17	2,16E-03
act_ecca_ind	-3,11E-07	1,41E-07	-2,207	3,02E-02

Fuente: elaboración propia.

La Tabla 3 presenta los indicadores de ajuste para cada modelo estimado.

Tabla 3. *Indicadores de error y validación de cada modelo*

	Serie A	Serie B
Indicadores MAPE (%)	15,029	13,886
Indicadores SMAPE (%)	12,74	11,28
Indicadores RMSE (%)	3115,44	1260,044
Valor P de la prueba de normalidad de Jarque Bera	8,66E-05	0,000175875
Valor P de la prueba de incorrelación de Ljung Box, lag = 12	0,1631454	0,5993553

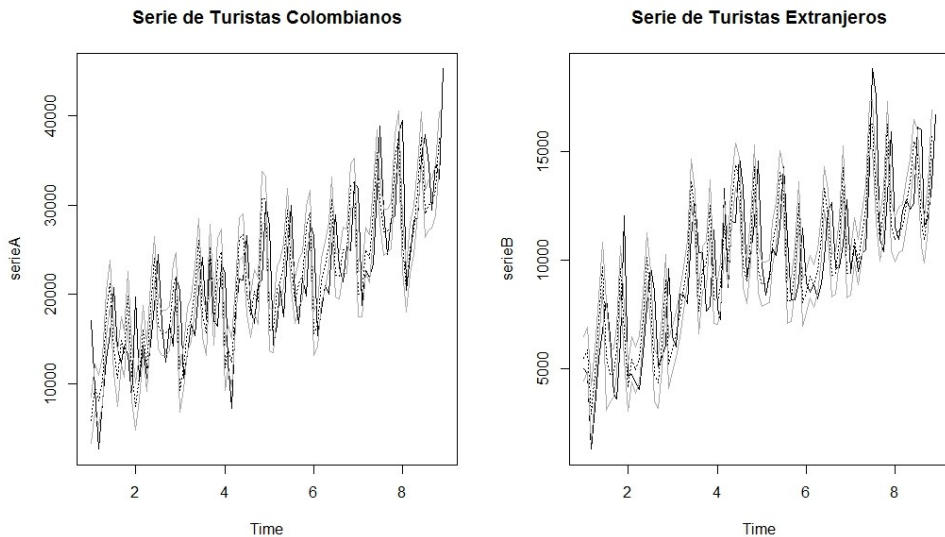
Fuente: elaboración propia.

La tabla 3 muestra, en la primera línea, los valores MAPE de ajuste para el modelo de regresión clásica. El menor valor de error MAPE se da en la serie de los turistas extranjeros, con 14 %, y los colombianos, con 15 %. La incorrelación en todos los residuales se cumple; es decir, los residuales tienen estacionariedad, pero no se comportan bajo la distribución normal, en los dos primeros modelos de las series de turistas (valor $p < 0,05$). Los residuales no se comportan con dicha distribución porque existen puntos atípicos que lo implican, pero no deben eliminarse porque, de hecho, su información es muy valiosa para la estimación de estos modelos. Se hacen transformaciones de las variables respuesta como logaritmo natural, raíz cuadrada y potencia de Box Cox. Sin embargo, ninguna mejora el problema de no normalidad⁴.

En el Gráfico 5 se aprecian los valores observados (en negro), el ajuste (en línea punteada) de los modelos finales econométricos estimados para las series de turistas colombianos y extranjeros, y los respectivos intervalos de confianza al 95 %, bajo la distribución normal (en gris). Aunque se aprecia la mayoría de valores reales (en negro) dentro de los intervalos quedan por fuera algunos un poco extremos.

⁴ Una razón que justifica aún más el uso del modelo bayesiano descrito y cuyo algoritmo en R se desarrolla y presenta en este artículo (Zellner, 1996), ya que los modelos propuestos no se restringen a este supuesto restrictivo sobre sus residuos.

Gráfico 5. *Ajuste de modelos econométricos para las series de turistas colombianos y extranjeros*



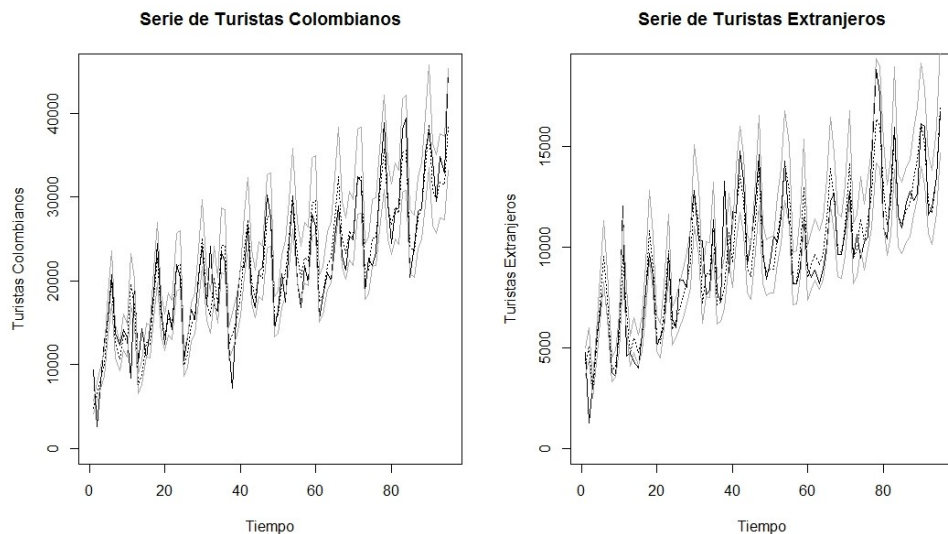
Fuente: elaboración propia.

C. Modelos de regresión bayesiana

1) Regresión bayesiana basada en la distribución normal para los datos

Las variables explicativas que fueron encontradas significativas en el modelo de regresión anterior, que reflejan el comportamiento que se visualizó en el análisis de autocorrelación, serán incluidas en estos modelos de regresión Bayesiana. En el primer modelo Bayesiano ajustado, se utiliza la fórmula con las covariables: rezago de orden 1, tiempo (tendencia), y como variable estacional, el mes. En el Gráfico 6 se aprecia la línea de valores reales (en negro) vs valores ajustados (en línea punteada), para cada una de las series de tiempo de izquierda a derecha: turistas colombianos y extranjeros, con los respectivos intervalos del 95% de probabilidad (en gris) de la distribución predictiva T de Student, con la cual se realizan los pronósticos.

Gráfico 6. Ajuste de los modelos de regresión bayesianos bajo la distribución normal para los datos



Nota: Intervalos del 95 % (límites bilaterales, en los percentiles 2,5 % y 97,5 %).

Fuente: elaboración propia.

El ajuste del modelo bayesiano visto en el Gráfico 6 se basa en el percentil 50 de la distribución predictiva T de Student del modelo con la cuarta ecuación de la Tabla 4. Se observa cómo el ajuste se acerca mucho a la línea de datos reales. Además, los intervalos son un poco más amplios que los del modelo clásico y contienen más valores reales que los intervalos de éste.

Se estimaron diferentes modelos de regresión bayesiana con la distribución no informativa explicada en la sección I.B. Se usaron los percentiles 25 %, 50 % y 75 % de la distribución predictiva para encontrar diferentes valores ajustados acorde con cada ecuación previamente establecida y así determinar el mejor valor posible. Se estiman tres indicadores, MAPE, sMAPE, RMSE, con todos los valores observados vs los estimados. Estos resultados se muestran en la Tabla 4.

Tabla 4. *Indicadores estimados por serie, percentil y modelo*

Covariables incluidas	Serie A			Serie B		
MAPE (%)						
Percentil	25 %	50 %	75 %	25 %	50 %	75 %
y_{t-1} , t, mes	13	12,33	14,4	12,3	12,25	14,7
y_{t-1} , mes	16,2	16	17,6	14,29	14,3	16,5
y_{t-1} , Sin, mes	16,1	15,6	17,3	14,25	14,24	16,36
y_{t-1} , t, Sin, mes	12,6	12,3	14,3	12,2	12	14,6
SMAPE (%)						
Percentil	25 %	50 %	75 %	25 %	50 %	75 %
y_{t-1} , t, mes	12,03	11,12	12,45	10,9	10,27	12,06
y_{t-1} , mes	14,12	13	13,74	12,34	11,61	13
y_{t-1} , Sin, mes	14,02	12,66	13,47	12,3	11,55	12,81
y_{t-1} , t, Sin, mes	12,04	11,0996	12,4	10,9	10,1	11,9
RMSE						
Percentil	25 %	50 %	75 %	25 %	50 %	75 %
y_{t-1} , t, mes	2834,46	2620,74	2947,42	1333,61	1241,59	1382,48
y_{t-1} , mes	3413,87	3241,79	3505,9	1461,5	1379,26	1505,3
y_{t-1} , Sin, mes	3396,82	3221,85	3491,01	1460,71	1377,5	1505,31
y_{t-1} , t, Sin, mes	2835,2	2605,42	2950,2	1329,2	1229,4	1379,15

Fuente: elaboración propia.

Los mínimos indicadores hallados son menores que los encontrados para el modelo de regresión econométrico de la Tabla 3, para los tres indicadores, siendo la cuarta la mejor ecuación de covariables: y_{t-1} , t , **Sin**, **mes** (tabla 4). Los datos señalados en cursiva para esta ecuación en cada serie son los mejores, los cuales corresponden a la mediana (50 %). Los tres indicadores coinciden en el resultado, lo cual muestra la efectividad que tienen para este tipo de valores tan altos en las series de tiempo.

2) Regresión bayesiana basada en la distribución Poisson para los datos

Tomando en cuenta los resultados de significancia determinados con los modelos de regresión para las series de turistas, se usaron las covariables finales de mayores efectos sobre la respuesta para estimar los modelos de regresión Poisson bayesiana, vía MCMC, llegando a los siguientes resultados.

Los coeficientes se simulan de la distribución a posteriori, y la salida de R provee estadísticas de dichas muestras de cada coeficiente, como la media y diferentes percentiles. Con cada percentil es posible ajustar los valores de la regresión Poisson bayesiana. La Tabla 5 muestra la salida de R para la serie de datos de turistas colombianos, una tabla similar sale para la serie de turistas extranjeros, pero con las variables IPC y actividad económica industrial, como las variables encontradas para el modelo clásico. Con los valores de las medias y los percentiles 25 %, 50 % y 75 %, se pueden estimar 4 modelos por cada serie, cuyos indicadores de error MAPE, sMAPE y RMSE se muestran en la Tabla 6.

Tabla 5. Resultados de MCMC poisson para turistas colombianos

Media y desviación estándar empírica por variable (coeficiente)		
	Media	Desv. est.
(Intercepto)	1,10E+01	2,43E-02
Yt1	3,26E-05	1,44E-07
mesN2	-6,38E-01	3,81E-03
mesN3	-1,78E-01	4,51E-03

(Continúa)

Tabla 5. Continuación

mesN4	-1,69E-02	3,97E-03	
mesN5	-6,98E-02	3,92E-03	
mesN6	1,27E-01	3,66E-03	
mesN7	1,39E-01	3,07E-03	
mesN8	-3,03E-01	3,29E-03	
mesN9	-1,44E-01	3,49E-03	
mesN10	3,80E-02	3,57E-03	
mesN11	-9,06E-02	3,56E-03	
mesN12	2,24E-01	3,19E-03	
act_res	-1,41E-08	1,83E-10	
Percentiles por variable			
	25 %	50 %	75 %
(Intercepto)	1,09E+01	1,09E+01	1,10E+01
Yt1	3,25E-05	3,26E-05	3,27E-05
mesN2	-6,42E-01	-6,38E-01	-6,36E-01
mesN3	-1,81E-01	-1,78E-01	-1,75E-01
mesN4	-1,98E-02	-1,68E-02	-1,42E-02
mesN5	-7,28E-02	-6,98E-02	-6,71E-02
mesN6	1,25E-01	1,27E-01	1,30E-01
mesN7	1,37E-01	1,39E-01	1,41E-01
mesN8	-3,05E-01	-3,03E-01	-3,01E-01
mesN9	-1,46E-01	-1,44E-01	-1,42E-01
mesN10	3,57E-02	3,77E-02	4,01E-02
mesN11	-9,33E-02	-9,06E-02	-8,82E-02
mesN12	2,22E-01	2,24E-01	2,26E-01
act_res	-1,42E-08	-1,41E-08	-1,40E-08

Fuente: elaboración propia.

En la Tabla 6 se aprecian indicadores más altos que los encontrados en los modelos anteriores. Sin embargo, no son valores muy malos en comparación con la regresión bayesiana anterior, pero son mejores que los del modelo clásico. Lo anterior podría dejar este modelo como alternativa frente a otros casos de datos discretos que no tengan escalas muy altas para aproximar el modelo de regresión bayesiana bajo la escala normal.

Tabla 6. *Indicadores de error de los modelos de la regresión Poisson bayesiana*

Producto	Serie de turistas colombianos				Serie de turistas extranjeros			
	Media	25 %	50 %	75 %	Media	25 %	50 %	75 %
MAPE (%)	14,33	14,54	14,33	14,96	14,7	14,43	14,7	15,34
sMAPE (%)	11,46	12,17	11,47	11,57	11,65	11,75	11,66	11,96
RMSE	2874,7	2970,6	2874,7	2980,55	1319,76	1337,22	1319,64	1345,4

Fuente: elaboración propia.

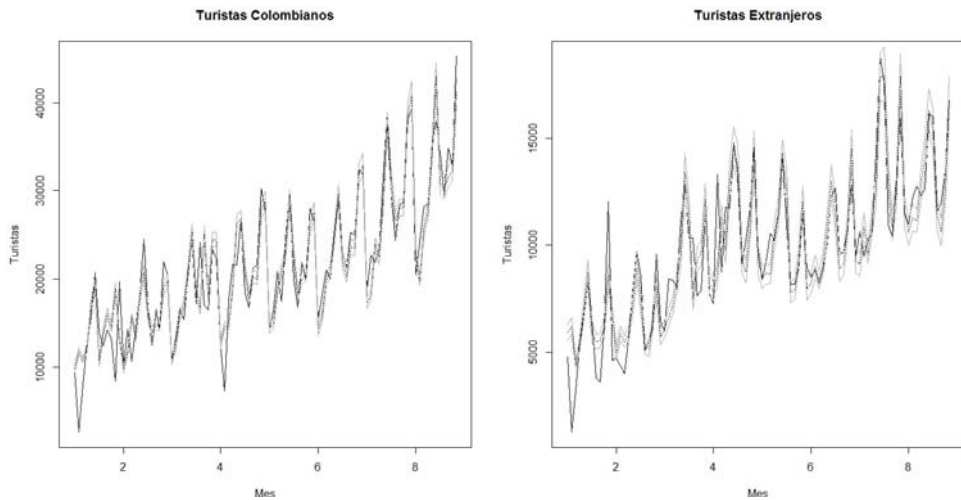
En el Gráfico 7, donde se aprecia el ajuste en línea punteada y límites de confianza en color gris, se observan un poco más datos por fuera de los intervalos estimados para este modelo vía MCMC, situación que en los anteriores modelos puede derivar en mayores errores al pronosticar.

IV. Discusión

Es importante notar que la regresión Bayesiana, tanto el modelo bayesiano basado en la aproximación normal de los datos como el basado en la distribución Poisson con muestreo por MCMC, puede proporcionar un resultado adecuado para predecir demandas turísticas, en modelos con covariables muy similares a las encontradas en la regresión clásica, con mejores indicadores de error de ajuste. Modelos que también pueden ser comparados con otros como los usados con distribución empírica, binomial negativa, de regresión Poisson con bootstrap, en (Kolasa, 2016), o modelos dinámicos lineales, o incluso, con intervención de expertos, aspecto que incrementaría la

información para realizar planeación de recursos y disponibilidades hoteleras por mes, mejorando las posibilidades de ingresos.

Gráfico 7. *Ajuste de los modelos Poisson bayesianos via MCMC: turistas colombianos, extranjeros*



Nota: Intervalos al 95 % de la muestra de MCMC (Límites bilaterales, en los percentiles 2,5 % y 97,5 %).

Fuente: elaboración propia.

Conclusiones

Se encontraron modelos con adecuada representación de la variación temporal de las series de turistas que llegan a Medellín, colombianos o extranjeros, diferentes a los modelos clásicos, basados en Estadística Bayesiana. Así mismo, se hallaron indicios de comportamientos estacionales, reflejados en la presencia de la variable indicadora mes, además, se encontró un impacto significativo del valor de turistas rezagado de orden 1, para explicar el comportamiento de la cantidad de ambos tipos de turistas que llegan a esta ciudad, es decir, tiene una dependencia temporal de primer orden. Esto permitió caracterizar el comportamiento de este sector de la economía colombiana, y con ello, ayudar en la planeación de dicho sector.

Al usar los indicadores de errores sMAPE y RMSE —que son muy útiles en casos generales, en especial robustos cuando las respuestas se acercan a cero— se corrobora que el indicador MAPE puede ser apropiado, pero solo en casos en los que los datos toman valores muy altos, como se ve en este trabajo, alrededor de 5000 o más turistas por mes.

La variable mes también es una variable estadísticamente significativa en las demandas de turismo de colombianos y extranjeros, resultado que permite identificar los periodos de mayor incidencia para la planeación de recursos hoteleros. Por ejemplo, para los colombianos, las vacaciones entre diciembre-enero y junio-julio representan temporadas para visitar la ciudad de Medellín, y anualmente se repite dicho comportamiento.

Las asociaciones negativas, encontradas con los gastos energéticos, podrían deberse a diversas razones, una de ellas puede ser que los periodos de mayor flujo turístico son las vacaciones de los sectores educativos; sin embargo, es necesario usar otro tipo de técnicas investigativas para determinar si hay algún grado de disminución de actividades o alguna otra razón.

Puede considerarse para futuros estudios la planeación de disponibilidad de hoteles en la ciudad, analizando la incidencia del mes, dados los resultados de éste estudio. Con ello, se puede requerir mayores recursos en los meses encontrados con mayor incidencia de los turistas extranjeros, como son: julio, agosto y diciembre, temporadas en que también se programan eventos de ciudad como la feria de las flores, en agosto, y los alumbrados, en diciembre. Por todo esto es importante programar estrategias de atención y cobertura de los servicios requeridos por este tipo de turistas.

Referencias

AGENCIA EFE (18 de septiembre de 2014). Informalidad del turismo en Colombia impide su crecimiento, según expertos. *El Heraldo*. Recuperado de: <http://www.elheraldo.co/turismo/informalidad-del-turismo-en-colombia-impide-su-crecimiento-segun-expertos-167437>

BES, Christian & SETHI, Suresh (1988). “Concepts of forecast and decision horizons: Applications to dynamic stochastic optimization problems”,

Valencia, Vanegas, Correa y Restrepo: Comparación de pronósticos para la dinámica...

Mathematics of Operations Research, Vol. 13, No. 2, pp. 295-310.

- BONILLA, Jorge Mauricio & MORENO, María Margarita (2010). *Determinantes de la demanda de turismo en Colombia 2004-2007: seguridad, comercio y otros factores*. Tesis de Pregrado para optar al título de Economista, Facultad de Economía, Universidad del Rosario, Colombia.
- BOWERMAN, Bruce L.; OCONNELL, Richard T. & KOHELER, Anne B. (2007). *Pronósticos, series de tiempo y regresión: un enfoque aplicado*. México: Cengage Learning Editores S.A.
- BROWN, Robert Goodell (1959). *Statistical Forecasting for Inventory Control*. New York: McGraw-Hill.
- BROWN, Robert Goodell (1963). *Smoothing, Forecasting and Prediction of Discrete Time Series*. New Jersey: Prentice Hall, Englewood Cliffs.
- CARIDAD Y OCERIN, José María (1998). *Econometria: Modelos Econométricos y series temporales*. Barcelona: Editorial Reverté.
- CERDA, Rosa & LEGUIZAMÓN, Manuel (2005). "Análisis del comportamiento de la demanda turística urbana de Colombia", *Turismo y Sociedad*, Vol. 4, pp. 70-98.
- CHEN, Kuan-Yu (2011). "Combining linear and nonlinear model in forecasting tourism demand", *Expert Systems with Applications*, Vol. 38, No. 8, pp. 10368-10376.
- CHOI, T.; LI, D. & YAN, H. (2003). "Optimal two-stage ordering policy with Bayesian information updating", *Journal of the Operational Research Society*, Vol. 54, No. 8, pp. 846-859.
- CLAVERIA, Oscar & TORRA, Salvador (2014). "Forecasting tourism demand to Catalonia: Neural networks vs. time series models", *Economic Modelling*, Vol. 36, pp. 220-228.
- COHEN, Rochelle & DUNFORD, Fraser (1986). "Forecasting for Inventory Control: An Example of When 'Simple' Means 'Better'", *Interfaces*, Vol. 16, No. 6, pp. 95-99.

- CROSTON, J.D. (1972). "Forecasting and stock control for intermittent demands", *Operational Research Quarterly*, Vol. 42, No. 3, pp. 289-303.
- CROUCH, Geoffrey (1995). "A meta-analysis of tourism demand", *Annals of tourism research*, Vol. 22, No. 1, pp. 103-118.
- DE MOL, Christine; GIANNONE, Domenico & REICHLIN, Lucrezia (2008). "Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?", *Journal of Econometrics*, Vol. 146, No. 2, pp. 318-328.
- FEI, Xiang; LU, Chung-Cheng & LIU, Ke (2011). "A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction", *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 6, pp. 1306-1318.
- GALVIS, Luis Armando & AGUILERA, María Modesta (1999). "Determinantes de la demanda por turismo hacia Cartagena, 1987-1998", *Lecturas de Economía*, No. 51, pp. 47-87.
- GELMAN, Andrew; CARLIN, John; STERN, Hal & RUBIN, Donald (2004). *Bayesian Data Analysis* (2nd Ed.). Boca Raton, FL: CRC Press, Taylor & Francis Group.
- GILL, Jeff (2007). *Bayesian methods: A social and behavioral sciences approach* (2nd Ed.). United States of America: Chapman & Hall.
- GUIZZARDI, Andrea & MAZZOCCHI, Mario (2010). "Tourism demand for Italy and the business cycle", *Tourism Management*, Vol. 31, No. 3, pp. 367-377.
- GUIZZARDI, Andrea & STACCHINI, Annalisa (2015). "Real-time forecasting regional tourism with business sentiment surveys", *Tourism Management*, Vol. 47, pp. 213-223.
- GUTIÉRREZ, Valentina & VIDAL, Carlos (2008) "Modelos de gestión de inventarios en cadenas de abastecimiento: Revisión de la literatura", *Revista Facultad de Ingeniería*, Vol. 43, pp. 134-149.

- HARRISON, J. & STEVENS, C. (1976). "Bayesian Forecasting", *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 38, No. 3, pp. 205-247.
- HELLSTRÖM, Jörgen (2002). "Count data modelling and tourism demand", *Umeå Economic Studies*, No. 584. Sweden: Umeå University.
- HILLIER, Frederik & HILLIER, Mark (2007). *Métodos cuantitativos para administración* (3th Ed). México D.F.: McGraw-Hill.
- KOLASA, Stephan (2016). "Evaluating predictive count data distributions in retail sales forecasting", *International Journal of Forecasting*, Vol. 32, No. 3, pp. 788-803.
- LI, Gang; SONG, Haiyan & WITT, Stephen (2005). "Recent developments in econometric modeling and forecasting", *Journal of Travel Research*, Vol. 44, No. 1, pp. 82-99.
- LIM, Christine (1999). "A meta-analytic review of international tourism demand", *Journal of Travel Research*, Vol. 37, No. 3, pp. 273-284.
- LIU, Xiangyun; PENG, Hongqin; BAI, Yun; ZHU, Yujun & LIAO, Lueling (2014). "Tourism Flows Prediction based on an Improved Grey GM(1,1) Model", *Procedia - Social and Behavioral Sciences*, Vol. 138, pp. 767-775.
- LÓPEZ, Ledys & MÁRQUEZ, José Ignacio (2012). "Proyecto de observatorio de turismo para Medellín y Antioquia", *Revista Soluciones de Posgrado EIA*, Vol. 5, No. 9, pp. 79-96.
- MARTIN, Andrew; QUINN, Kevin & PARK, Jong (2011). "MCMCpack: Markov Chain Monte Carlo in R.", *Journal of Statistical Software*, Vol. 42, No. 9, pp. 1-21.
- NEELAMEGHAM, Ramya & CHINTAGUNTA, Pradeep (1999). "A Bayesian model to forecast new product performance in domestic and international markets", *Marketing Science*, Vol. 18, No. 2, pp. 115-136.
- PENG, Bo; SONG, Haiyan & CROUCH, Geoffrey (2014). "A meta-analysis of international tourism demand forecasting and implications for practice", *Tourism Management*, Vol. 45, pp. 181-193.

- PETRIS, Giovanni; PETRONE, Sonia & CAMPAGNOLI, Patrizia (2009). *Dynamic Linear Models with R*. New York: Springer.
- SAMARATUNGA, Chand; SETHI, Suresh & ZHOU, Xun Yu (1997). “Computational evaluation of hierarchical production control policies for stochastic manufacturing systems”, *Operations Research*, Vol. 45, No. 2, pp. 258-274.
- SARIMVEIS, Haralambos; PATRINOS, Panagiotis; TARANTILIS, Chris & KIRANOUDIS, Chris (2008). “Dynamic modeling and control of supply chain systems: A review”, *Computers & Operations Research*, Vol. 35, No. 11, pp. 3530-3561.
- SHOESMITH, G.L. & PINDER, J.P. (2001). “Potential Inventory Cost Reductions Using Advanced Time Series Forecasting Techniques”, *Journal of the Operational Research Society*, Vol. 52, No. 11, pp. 1267-1275.
- SONG, Haiyan & Li, Gang (2008). “Tourism demand modelling and forecasting—A review of recent research”, *Tourism Management*, Vol. 29, No. 2, pp. 203-220.
- SONG, Haiyan; GAO, Bastian Z. & LIN, Vera (2013). “Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system”, *International Journal of Forecasting*, Vol. 29, No. 2, pp. 295-310.
- SYNTETOS, Aris; BABAI, M. Zied; LENGU, David & ALTAY, Nezih (2011). “Distributional assumptions for parametric forecasting of intermittent demand”. In: Altay, Nezih & Litteral, Lewis (Eds.), *Service Parts Management. Demand Forecasting and Inventory Control* (pp. 31-52). London: Springer.
- VALENCIA, Marisol (2016). *Dynamic model for the multiproduct inventory optimization with multivariate demand*. Tesis para optar al título de: PhD en Ingeniería-Industria y organizaciones, Facultad de Minas, Escuela de Ingeniería de la Organización, Universidad Nacional de Colombia.

Valencia, Vanegas, Correa y Restrepo: Comparación de pronósticos para la dinámica...

- VIDAL, Carlos; LONDOÑO, Julio & CONTRERAS, Fernando (2004) “Aplicación de Modelos de Inventarios en una Cadena de Abastecimiento de Productos de Consumo Masivo con una Bodega y N Puntos de Venta”, *Ingeniería y Competitividad*, Vol. 6, No. 1, pp. 35-52.
- WALLSTRÖM, Peter & SEGERSTEDT, Anders (2010). “Evaluation of forecasting error measurements and techniques for intermittent demand”, *International Journal of Production Economics*, Vol. 128, No. 2, pp. 625-636.
- WEST, Mike & HARRISON, Jeff (1997). *Bayesian Forecasting and Dynamic Models* (2nd Ed.). New York. USA. : Springer Series in Statistics.
- WITT, Stephen & WITT, Christine (1995). “Forecasting tourism demand: A review of empirical research”, *International Journal of Forecasting*, Vol. 11, No. 3, pp. 447-475.
- ZELLNER, Arnold (1996). *An introduction to bayesian inference in econometrics* (2nd Ed.). Canadá: Wiley Classics Library.
- ZHOU, Ting; BONHAM, Carl & GANGNES, Byron (2007). “Modelling the supply and demand for tourism: a fully identified VECM approach”, *Working Papers*, No. 2007-17. Department of Economics, University of Hawaii at Manoa.