

PRONÓSTICO DE INGRESOS POR VISITANTES INTERNACIONALES A MÉXICO: UN ENFOQUE DE SERIES JERÁRQUICAS

REVENUE FORECAST INTERNATIONAL VISITORS TO MEXICO: A HIERARCHICAL APPROACH SERIES

Rómulo Sánchez*
Nora Gavira-Durón**

(Recibido: mayo, 2016/Aceptado: agosto, 2016)

Resumen

El objetivo de la presente investigación es determinar el modelo más eficiente para pronosticar el ingreso monetario por visitantes internacionales a México, tomando en consideración la naturaleza jerárquica de la serie. Desde esta perspectiva, las series de interés contienen todos los elementos necesarios no observados (v.g. aleatoriedad, ciclicidad y tendencia), para realizar un pronóstico razonable de corto plazo y contribuir al desarrollo de la teoría de pronóstico de series jerárquicas. Se utilizó el error promedio absoluto escalado (MASE) como criterio para seleccionar el método más eficiente. En las conclusiones se indica que, para el caso de los ingresos monetarios por visitantes internacionales a México, el método de pronóstico más eficiente es un ARIMA de arriba hacia abajo.

Clasificación JEL: C13, C19, C22, C43, C52.

Palabras clave: pronóstico, modelos jerárquicos, series de tiempo, ingresos, visitantes.

* Profesor-investigador de la Escuela de Negocios, UDLAP. Correo electrónico: romulo.sanchez@udlap.mx.

** Profesora-investigadora de la Escuela de Negocios, UDLAP. Correo electrónico: nora.gavira@udlap.mx.

Abstract

The aim of this paper is to identify the most efficient forecasting model of income from visitors to Mexico. The received theory of forecasting claims that hierarchical methods are systematically the best forecasting models of hierarchical time series, a claim that has not been evaluated in a large number of cases. Consequently, in this research several methods are evaluated, following the conventional approach of selecting two samples (i.e. training and test sets). Based on the mean absolute scaled error (MASE), our results indicate that, in the case of the variable of interest, the bottom-up ARIMA is the most efficient forecasting method.

Keywords: forecast, hierarchical models, time series, income, visitors.

JEL Classification: C13, C19, C22, C43, C52.

1. Introducción

En este trabajo se hace una evaluación de métodos para determinar el más eficiente para pronosticar el ingreso monetario por visitantes internacionales a México, una categoría muy importante para el país en términos de su proporción del producto interno bruto (PIB); debido a que el ingreso de las actividades económicas, como proporción del PIB, es un indicador indirecto de la importancia de las actividades en términos de empleo, inversiones, remuneraciones, etcétera.

La evaluación de los métodos se hace con base en la perspectiva de la teoría de pronósticos. Desde esta perspectiva, las series de interés contienen todos los elementos (no observados) necesarios (v.g. aleatoriedad, ciclicidad y tendencia) para hacer un pronóstico razonable de corto plazo. Los desarrollos relativamente recientes de dicha teoría incorporan otros elementos como, por ejemplo, la naturaleza de la serie: jerárquica o no jerárquica, entre otros. En particular, se afirma que las características propias de la serie objetivo deben considerarse como una restricción para efectos de la evaluación de los métodos.

El propósito de este trabajo es contribuir al desarrollo de la teoría de pronóstico de series jerárquicas. Asimismo, pretende identificar un método útil, práctico y eficiente que atienda las necesidades de pronóstico

de ingreso por visitantes internacionales de organizaciones públicas y privadas. Es necesario señalar que la investigación sobre la eficiencia relativa de los enfoques jerárquicos para evaluar métodos de pronósticos de series jerárquicas es, a pesar de su potencial teórico y práctico, todavía escasa y los resultados distan de ser concluyentes.

El trabajo está dividido en tres secciones. En la primera, se contrasta el enfoque utilizado con otros enfoques; tal como se indicará en su momento, la teoría de pronósticos renuncia a investigar las causas del fenómeno y privilegia la búsqueda del método más eficiente *per se*. Esto es, desde luego, controversial, pero ha probado ser útil, práctico y relevante para realizar pronósticos de corto plazo. En la segunda, se indica el diseño de la investigación, es necesario indicar que se ha puesto cuidado en evaluar métodos de reciente desarrollo para la evaluación de series jerárquicas; en particular, se evalúa el método combinación óptima (*optimal combination*) desarrollada por (Hyndman, Ahmed, Athanasopoulos y Lin S., 2011). Los autores del presente estudio no tienen referencia de un trabajo aplicado previo en este campo que no sea el de Hyndman y sus colegas. En la tercera sección se presenta un comparativo de los resultados, utilizando dos estrategias diferentes para definir el conjunto de entrenamiento (*training set*), mismas que se explicarán más adelante. La conclusión del trabajo indica que el enfoque jerárquico sí permite identificar el método del pronóstico de corto plazo más eficiente para una serie de naturaleza jerárquica como el ingreso por visitantes a México.

2. El modelo de pronóstico de series jerárquicas

El repertorio de métodos avanzados de pronósticos disponible es amplio y diverso. No obstante, todos los métodos de pronóstico, independientemente de la complejidad de su desarrollo algebraico, intentan dar respuesta a tres preguntas fundamentales: ¿de dónde venimos?, ¿dónde estamos?, y, ¿hacia donde vamos?, un modelo en sí mismo no responde estas preguntas; el analista o investigador es quien debe responderlas a partir de los resultados que arroja el modelo. Por lo tanto, la naturaleza propia del fenómeno de interés obliga a reflexionar sobre la viabilidad o pertinencia de modelos particulares independientemente de su nivel de complejidad.

El pronóstico del ingreso monetario por visitantes internacionales puede ser planteado, por ejemplo, en términos de un modelo de regresión múltiple:

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \dots + \beta_k x_{k,t} + e_t \quad (1)$$

donde Y es la variable a pronosticar y, los coeficientes β , indican el impacto de cada predictor k ; en este modelo se asume que e_t es un término de error no autocorrelacionado. La tarea del analista o pronosticador consiste en identificar y eventualmente manipular los predictores k , para que se comporten conforme a lo esperado.

Otra opción sería utilizar un modelo de vectores autoregresivos ARIMA. En este tipo de modelo se pronostica la variable a partir de una combinación lineal de valores pasados de la variable. El modelo puede ser escrito de la siguiente forma:

$$y_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t \quad (2)$$

ARIMA (p, d, q) donde:

- p = orden de la parte autoregresiva;
- d = grado de la primera diferencia incluida;
- q = orden del promedio móvil.

donde y'_t es la serie diferenciada, la cual puede haber sido diferenciada más de una vez. Un modelo ARIMA estacional se forma incluyendo la parte estacional en ARIMA (p, q, d). Para simplificar: ARIMA (p, d, q) (P, D, Q_m), donde m es el número de periodos. Más aún, los modelos anteriores se pueden combinar para definir un modelo de regresión dinámica. Asumiendo que n_t sigue un ARIMA (1,1,1), el modelo queda especificado con la siguiente ecuación:

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \dots + \beta_k x_{k,t} + n_t \quad (3)$$

$$(1 - \phi_1 \beta)(1 - \beta)n_t = (1 + \theta_1 \beta)e_t \quad (4)$$

donde e_t es una serie con ruido blanco. Note que esta ecuación tiene dos términos de error (n_t y e_t). El modelo de regresión en este caso se torna estadísticamente más complejo y el proceso de pronóstico se traduce en pronosticar por separado y combinar la parte de regresión del modelo y la

parte correspondiente al modelo ARIMA. Una particularidad de este tipo de modelos está relacionada con el comportamiento futuros de los predictores, dado que se debe asumir o pronosticar los valores posteriores de las variables predictoras (Stock, 2006). Se debe tener presente que los intervalos de predicción de la regresión no consideran la incertidumbre en el pronóstico de los predictores.

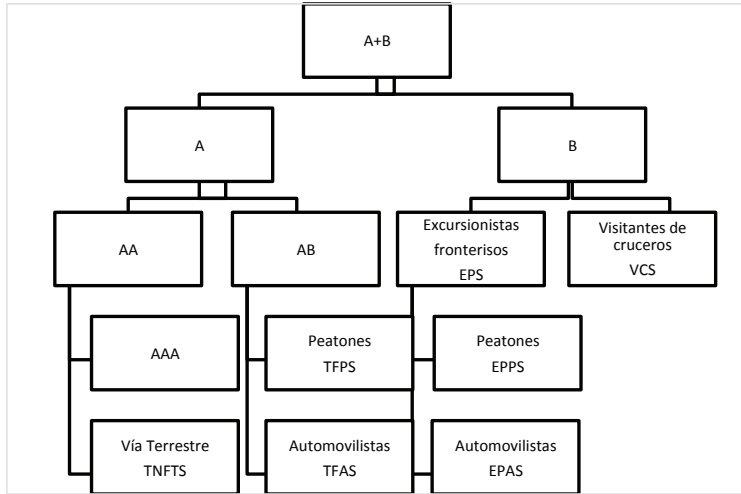
En este trabajo se adoptó el modelo de pronósticos de series jerárquicas, debido a que la variable objetivo es el resultado de agregar información correspondiente a diferentes categorías y subcategorías. Al pronosticar el ingreso por turistas fronterizos, se pronostica el ingreso derivado de los visitantes que pernoctarán y demandarán servicios de alojamiento, alimentos, bebidas y entretenimiento al menos una noche en las fronteras. Este argumento es válido para de igual forma los visitantes de cruceros, los turistas no fronterizos o excursionistas fronterizos.

El modelo jerárquico (Hyndman R. J., 2014) se muestra en la figura 1, el cual consta de $K = 4$ niveles, donde el nivel 1 (representa el agregado total) se define como una serie jerárquica. Cada nivel (de 1 a K) muestra una mayor desagregación hasta llegar al nivel inferior de K (mayor desagregación) de la serie. Las observaciones se registran en los instantes $t = 1, 2, \dots, n$ y el objetivo es realizar pronósticos para cada serie en cada nivel en el tiempo $t = n + 1, n + 2, \dots, n + h$. Sea Y_t la suma de todas las series en el tiempo t . En un modelo jerárquico, las observaciones de niveles más altos se pueden obtener mediante la suma de la serie de abajo, así $y_t = \sum_i y_{it}$, donde los índices i , indican una serie genérica en el nivel 1 de la jerarquía, y $y_{it} = \sum_j y_{ij,t}$, etc. En términos de matrices, Y_{it} denota el vector de todas las observaciones del nivel i y tiempo t , entonces:

$$Y_t = [Y_t, Y_{1,t}, \dots, Y_{K,t}]' \quad (5)$$

Utilizando la “sumadora” de matrices (Hyndman, Ahmed, Athanassopoulos y Lin S., 2011) S , que almacena la estructura de la jerarquía, Y_t se puede encontrar a partir de la serie del nivel inferior.

$$Y_t = SY_{K,t} \quad (6)$$



Fuente: elaboración propia.

Figura 1
Serie jerárquica de siete niveles

La “sumadora” de matrices S , es una matriz de orden $m \times m_K$. Donde m es el número total de series en la jerarquía y el número de series en cada nivel es m_i (para $i = 1, 2, \dots, K$). El modelo de pronóstico general asume que primero se calcula el pronóstico para cada serie en cada nivel base m para cada periodo $n+1, n+2, \dots, n+h$, basado en la información disponible hasta el tiempo n . Bajo el supuesto de (Hyndman, Ahmed, Athanasopoulos y Lin S., 2011) que la combinación óptima del pronóstico base $\hat{Y}_{X,n}(h)$ (donde X denota la serie pronosticada) es independiente de los datos, tenemos que para una serie jerarquía simple con sólo un nivel de desagregación $K=1$, y con nodos m_1 en el nivel 1, los pesos están dados por:

$$s = \begin{pmatrix} 1111111 \\ 1111000 \\ 0000111 \\ 1100000 \\ 0011000 \\ 0000110 \\ 0000001 \\ 1000000 \\ 0100000 \\ 0010000 \\ 0001000 \\ 0000100 \\ 0000010 \\ 0000001 \end{pmatrix} \quad (7)$$

A continuación se presentan de manera resumida los distintos enfoques (Hyndman, Ahmed, Athanasopoulos y Lin S., 2011) y (Hyndman R. J., 2014).

De abajo hacia arriba: en este caso se formulan pronósticos para las series base o subcategorías y luego se agregan en categorías y niveles o nodos, por ejemplo:

$$\begin{aligned} \hat{Y}_{A,h} &= \hat{Y}_{AA,h} + \hat{Y}_{AB,h} + \hat{Y}_{AC,h} \text{ (nodo A)} \\ \hat{Y}_{B,h} &= \hat{Y}_{BA,h} + \hat{Y}_{BB,h} + \hat{Y}_{BC,h} + \hat{Y}_{BD,h} \text{ (nodo B)} \\ \hat{Y}_{T,h} &= \hat{Y}_{A,h} + \hat{Y}_{B,h} \text{ (nodo agregado)} \end{aligned} \quad (8)$$

De arriba hacia abajo: la ventaja de este método radica en que únicamente requiere un sólo pronóstico. Hay varios métodos para desagregar dicho pronóstico:

Promedio de proporciones históricas: las proporciones se calculan para todas las series base. Este método refleja el promedio de las proporciones históricas que representan estas series del agregado total.

$$pj = \frac{1}{n} \sum_{t=1}^n Y_{j,t} / Y_t \quad (9)$$

series $j = 1, \dots, m_K$; Y_t = agregado total

Proporción de promedios históricos: tal como lo indica el nombre del método, en este caso implica estimar el valor histórico promedio de las series base, relativo al valor promedio del agregado a través del tiempo:

$$pj = \sum_{t=1}^n \frac{Y_{j,t}}{n} / \sum_{t=1}^n \frac{Y_t}{n} \quad (10)$$

Pronóstico de proporciones: en este caso, para cada nivel, de arriba hacia abajo, se calcula la proporción de cada pronóstico por serie base respecto al agregado de todos los pronósticos de las series base con K nodos o niveles:

$$\begin{aligned} pj &= \prod_{l=0}^{K-1} \frac{Y_{t,k}^{(l)}}{S_{j,t}^{(l+1)}} \\ pj &= \left(\hat{Y}_{AA,t} \middle| S_{AA,t}^{(1)} \right) \left(\hat{Y}_{AA,t}^{(1)} \middle| S_{AA,t}^{(2)} \right) = \left(\hat{Y}_{AA,t} \middle| S'_{A,t} \right) \left(\hat{Y}_{A,t} \middle| S'_{Total,t} \right) \end{aligned} \quad (11)$$

donde $S_{AA,t}^{(2)} = S'_{Total,t} = \hat{Y}_{A,t} + \hat{Y}_{B,t} + \hat{Y}_{C,t}$ y $S_{AA,t}^{(1)} = \hat{Y}_{AA,t} + \hat{Y}_{AB,t}$

De un nivel intermedio hacia arriba y hacia abajo: este caso combina los dos tipos de métodos anteriores. Se comienza por hacer un pronóstico de nivel intermedio y los subsiguientes. Para obtener los pronósticos de los niveles superiores al nivel de pronóstico se utiliza el método de abajo hacia arriba; y algún método de arriba hacia abajo para desagregar la información de las series base.

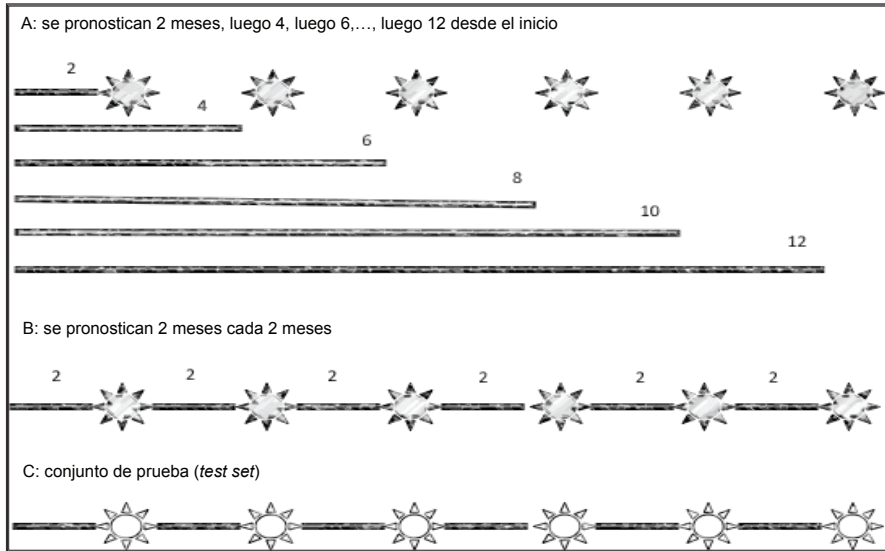
Combinación óptima: este método se basa en la aplicación del método de regresión para reconciliar y combinar óptimamente los pronósticos. La idea general es derivada de los pronósticos base de la jerarquía varios pasos o periodos adelantados o en el futuro:

$$\hat{y}_h = S\beta_h + \varepsilon_h \quad (12)$$

donde: \hat{y}_h es un vector de h pasos o periodos de pronósticos las series base, β_h es la media desconocida de los valores futuro s de los valores de las series base del nivel K , y ε_h representa el error de la regresión.

3. Diseño de la investigación

La evaluación del modelo de pronósticos de series jerárquicas se realizó siguiendo una variación de la metodología estándar de evaluación de modelos de pronósticos. Usualmente, se define un conjunto de entrenamiento (*training set*) y un conjunto de prueba (*test set*). En contraste, en este trabajo se definieron dos estrategias, como se muestra en la figura 2:



Fuente: elaboración propia.

Figura 2
Estrategias de evaluación del pronóstico

En la estrategia A se toma como base la información obtenida hasta finales de 2014 con lo que se pronostica el primer bimestre de 2015, el segundo bimestre, el tercer bimestre, y así, hasta llegar al sexto bimestre de 2015. Mientras que en la estrategia B se pronostica el próximo bimestre a partir del último bimestre pronosticado.

A medida que transcurre el tiempo, los errores, debido a la información disponible en los distintos momentos en que se realiza pronóstico, según la estrategia, difieren, lo cual a su vez, permite determinar si contar con más información en el corto plazo mejoraría la calidad del pronóstico. Nótese que los pronósticos del primer bimestre de las dos estrategias son o deben ser iguales por definición. Se utilizan la media, la desviación estándar y el coeficiente de variación para determinar el método de pronóstico más eficiente por tipo de estrategia.

Asimismo, se seleccionó una medida de error para evaluar los métodos. El error de pronóstico es simplemente la diferencia entre el dato original y el dato pronosticado:

$$e_i = y_i - \hat{y}_i \tag{13}$$

Existe una extensa bibliografía sobre medidas de error en esta área que trasciende los propósitos de este trabajo. No obstante, aquí se adopta el criterio de (Hyndman, R. J., 2014) quien sugiere utilizar la media del error promedio escalado (MASE, por sus siglas en inglés) para evaluar comparativamente los pronósticos obtenidos con métodos diferentes, debido a que dicha medida es neutral respecto a la escala en que se miden las distintas series de tiempo. Esta medida de error se basa en la media del error absoluto (e_j) del método de pronóstico ingenuo (indicado anteriormente) para el *training set*. Tal método genera un pronóstico de un paso para todos los datos de la serie. El modelo es el siguiente:

$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \tag{14}$$

El MASE es menor que uno si resulta de un pronóstico superior al calculado sobre el conjunto de datos de entrenamiento (*training set*). Es mayor que uno en caso contrario. El MASE para datos estacionales se estima con el siguiente ajuste:

$$q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|} \tag{15}$$

Adicionalmente, se comparan los errores para métodos utilizando enfoques jerárquicos y no jerárquicos. A continuación se presentan de manera simplificada los modelos no jerárquicos (una sola serie) más utilizados (Hyndman, R. J., 2014) y (Montemayor, G., 2012):

- Media (*mean*)

Según este método, el pronóstico es igual a la media de los datos históricos $y + \dots + y_T$, h es el periodo a pronosticar:

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + \dots + y_T) / T \tag{16}$$

- Ingenuo (*naive*)

En este caso, cada pronóstico es igual al último valor observado:

$$\hat{y}_{T+h} = y_T \tag{17}$$

- Ingenuo estacional (*seasonal naïve*)

De manera similar al método anterior, pero teniendo presente que el pronóstico debe ser igual al último valor observado correspondiente al valor observado de la misma estación del año (v.g. el mismo mes o trimestre o bimestre del año previo).

$$y_{T+h-km} \text{ donde } m = \text{periodo estacional} = \left\lfloor \frac{(h-1)}{m} \right\rfloor + 1, \quad (18)$$

Siendo $\lfloor(u)\rfloor$ la parte entera de (u) .

- Camino aleatorio con derrape (*random walk with drift*)

Una variación de método ingenuo consiste en permitir que el pronóstico aumente o disminuya en el tiempo, con una cantidad de cambio llamada *drift* igual al promedio de cambio observado en la en la serie histórica. El pronóstico para el tiempo $T+h$ está dado por:

$$y_T + \frac{h}{T-1} \sum_{t=2}^T (y_t - y_{T-1}) = y_T + h \left(\frac{y_T - y_1}{T-1} \right) \quad (19)$$

- Suavizado exponencial y variaciones de *holt* y *holt-winters*

El pronóstico es calculado usando un promedio ponderado a partir de pesos o ponderaciones que decrecen exponencialmente de tal forma que las ponderaciones menores están asociadas con las observaciones iniciales:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 \alpha y_{T-2} + \dots, \quad (20)$$

donde $0 \leq \alpha \leq 1$

Una representación alternativa es la siguiente:

$$\hat{y}_{t+1|t} = \ell_t, \text{ ecuación de pronóstico} \quad (21)$$

$$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}, \text{ ecuación de suavizado} \quad (22)$$

El modelo exponencial fue extendido para hacer un pronóstico de datos con tendencia lineal y es conocido como suavizado lineal de *holt*:

$$\hat{y}_{t+1|t} = \ell_t + hb_t, \text{ ecuación de pronóstico} \quad (23)$$

$$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}), \text{ ecuación de nivel} \quad (24)$$

$$b_t = \beta (\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}, \text{ ecuación de tendencia} \quad (25)$$

Donde ℓ_t denota un estimado en el tiempo t y b_t denota un estimado de la tendencia de la serie en el tiempo t y α y β son parámetros mayores o iguales a 0 y menores o iguales a 1.

Una variación de *holt* al método lineal anterior se permite que el nivel la pendiente sea multiplicativa en vez de aditiva:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t b_t^h \\ \ell_t &= \alpha y_t + (1 - \alpha) (\ell_{t-1} b_{t-1}) \\ b_t &= \beta \frac{\ell_t}{\ell_{t-1}} + (1 - \beta) b_{t-1}\end{aligned}\tag{26}$$

Asimismo, se han realizado variaciones, conocidas en inglés como *damped trend methods*, para incluir un parámetro \emptyset , mayor que 0 y menor a 1, para que los pronósticos de corto plazo se hagan con tendencia pero los de largo plazo sean constantes. Una variación es de tipo aditiva:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t b_t^{(\emptyset + \emptyset^2 + \dots + \emptyset^h)} \\ \ell_t &= \alpha y_t + (1 - \alpha) (\ell_{t-1} + \emptyset b_{t-1}) \\ b_t &= \beta \frac{\ell_t}{\ell_{t-1}} + (1 - \beta) \emptyset b_{t-1}\end{aligned}\tag{27}$$

Otra variación del método es multiplicativa:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t b_t^{(\emptyset + \emptyset^2 + \dots + \emptyset^h)} \\ \ell_t &= \alpha y_t + (1 - \alpha) (\ell_{t-1} b_{t-1}^\emptyset) \\ b_t &= \beta \frac{\ell_t}{\ell_{t-1}} + (1 - \beta) b_{t-1}^\emptyset\end{aligned}\tag{28}$$

Otras variaciones, conocidas como de *holt-winters* para datos con estacionariedad S_t , cuya constante correspondiente γ varía es mayor que 0 y menor que 1, se presentan a continuación:

holt-winters aditivo:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t + h b_t + S_{t-m+h_m} \\ \ell_t &= \alpha (y_t - S_{t-m}) + (1 - \alpha) (\ell_{t-1} + b_{t-1}) \\ b_t &= \beta (\ell_t - \ell_{t-1}) + (1 - \beta) b_{t-1} \\ S_t &= \gamma (y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma) S_{t-m}\end{aligned}\tag{29}$$

holt-winters multiplicativo:

$$\begin{aligned}\hat{y}_{t+h|t} &= (l_t + hb_t)S_{t-m+h_m^*} \\ \ell_t &= \alpha \frac{y_t}{S_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta(\ell_{t-1} - \ell_{t-1}) + (1-\beta)b_{t-1} \\ S_t &= \gamma \frac{y_t}{(\ell_t + b_{t-1})} + (1-\gamma)S_{t-m}\end{aligned}\quad (30)$$

Otra variación importante de mencionar son los *damped trend methods*. Para resumir, se indica el *holt-winters* multiplicativo *damped trend* por su utilidad práctica:

$$\begin{aligned}\hat{y}_{t+h|t} &= (l_t + (\emptyset + \emptyset^2 + \dots + \emptyset^h)b_t)S_{t-m+h_m^*} \\ \ell_t &= \alpha \frac{y_t}{S_{t-m}} + (1-\alpha)(\ell_{t-1} + \emptyset b_{t-1}) \\ b_t &= \beta(\ell_{t-1} - \ell_{t-1} + (1-\beta)\emptyset b_{t-1}) \\ S_t &= \gamma \frac{y_t}{(\ell_t + \emptyset b_{t-1})} + (1-\gamma)S_{t-m}\end{aligned}\quad (31)$$

- Modelos de suavizado exponencial simple con errores (ETS)

En este modelo hay dos ecuaciones: una que describe la serie (2.16) y una de transición (2.17) que captura los componentes no observados (nivel, estacionalidad, tendencia). Note que se trata de una familia de modelos: aditivos y multiplicativos.

$$y_t = \ell_{t-1} + \varepsilon_t \quad (32)$$

$$\ell_t = \ell_{t-1} + \alpha\varepsilon_t \quad (33)$$

- ARIMA (p, d, q) donde:

p = orden de la parte autoregresiva;

d = grado de la primera diferencia incluida;

q = orden del promedio móvil.

El modelo puede ser escrito de la siguiente forma:

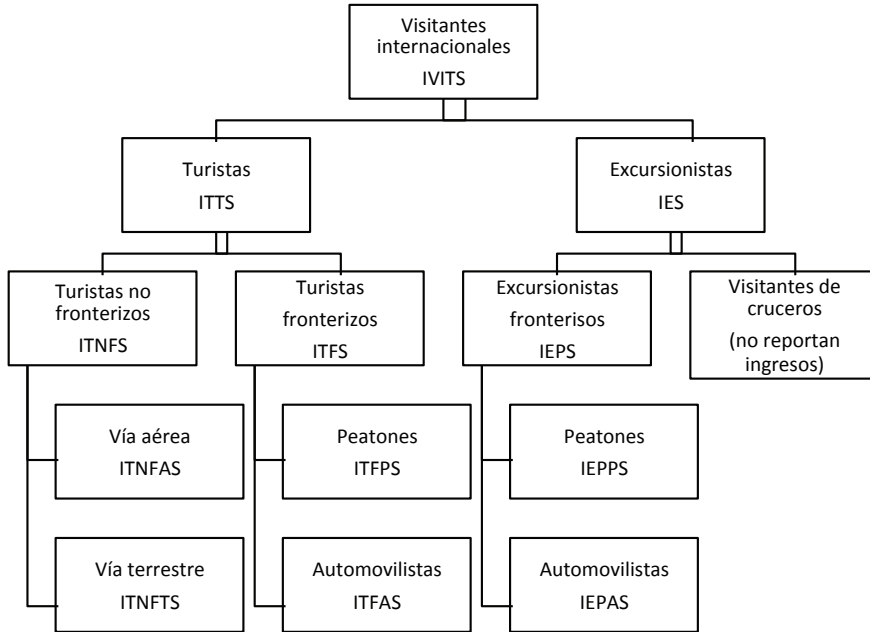
$$y_t = c + \emptyset_1 y_{t-1} + \dots + \emptyset_p y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t \quad (34)$$

Donde y'_t es la serie diferenciada, la cual puede haber sido diferenciada más de una vez. Un modelo ARIMA estacional se forma incluyendo la parte estacional en ARIMA (p, q, d). Para simplificar: ARIMA (p, d, q) (P, D, Q) m , donde m es el número de periodos.

En muchos casos, los modelos no jerárquicos arrojan pronósticos aceptables; más aún, métodos poco sofisticados como la media y los ingenuos pronostican con menos error que los modelos más sofisticados; por esta razón, los analistas deben considerarlos como referencia para comparar los resultados obtenidos con modelos matemáticamente más elaborados. La naturaleza de la serie es un factor que debe tomarse en cuenta a la hora de utilizar un método de pronóstico, existen series que son el resultado de una agregación; en algunos casos, la agregación se realiza con base en criterios *ad hoc* y, en otros casos, se asume o considera un “lógica natural”. El uso de enfoques jerárquicos se justifica cuando se pronostican series agregadas o agrupadas. Tal enfoque debería arrojar pronósticos más exactos, pero como no hay garantía de que esto ocurra, siempre se comparan sus resultados con los resultados de enfoques no jerárquicos.

4. Pronóstico del ingreso monetario por visitantes internacionales a México

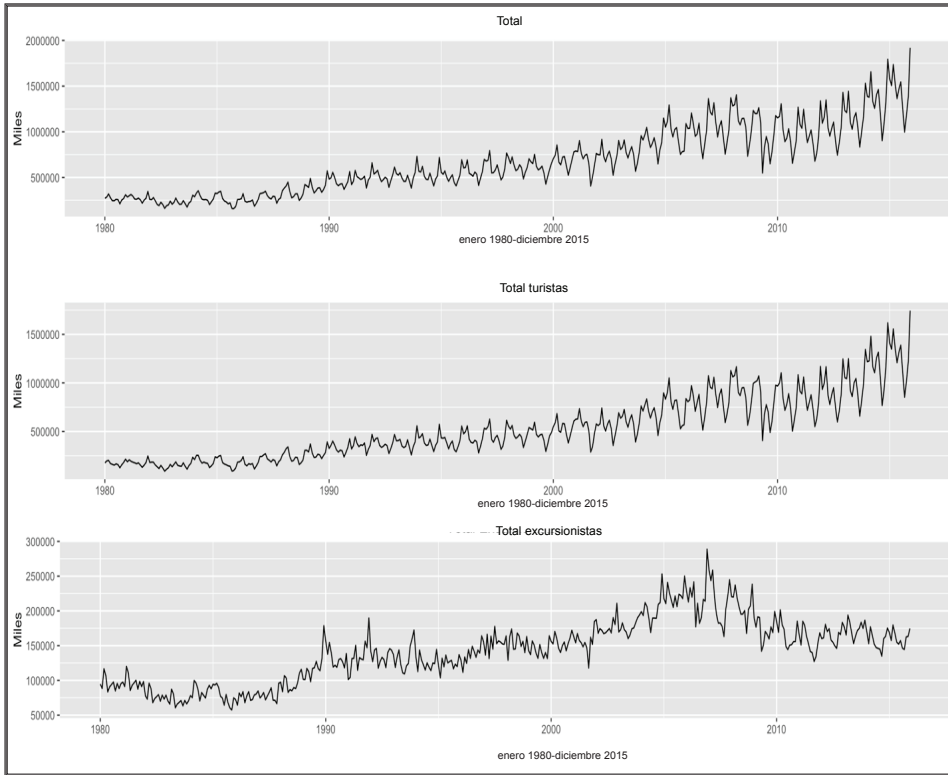
La perspectiva adoptada en este trabajo consiste en que el ingreso monetario asociado con los visitantes internacionales a México debe ser tratado una serie jerárquica, debido a que es una agregación del ingreso derivado de varias subcategorías y categorías de visitantes. En la figura 3 se presenta de forma esquematizada la agrupación de las distintas subcuentas.



Fuente: elaboración propia.

Figura 3
Agrupamiento jerárquico de los ingresos por visitantes internacionales a México

Tal como sugiere la figura 3, la cuenta de ingresos de interés es el resultado de agregar seis subcuentas diferentes en tres nodos que posteriormente se convierten en dos y finalmente en uno, mismo que es el total de ingresos. Este agrupamiento es lógico y natural debido a la naturaleza misma de la variedad de visitantes. Nótese que un agrupamiento distinto complicaría innecesariamente el análisis de las series. Por ejemplo, se podrían agrupar los turistas por modalidad de ingreso al país (vía aérea o terrestre) y, consecuentemente, se obtendrían dos nodos equivalentes que incluirían turistas y excursionistas dentro de las mismas subcategorías; ciertamente, este arreglo resultaría ilógico. En la figura 4, se presenta la evolución del ingreso mensual para el periodo 1980-2015.

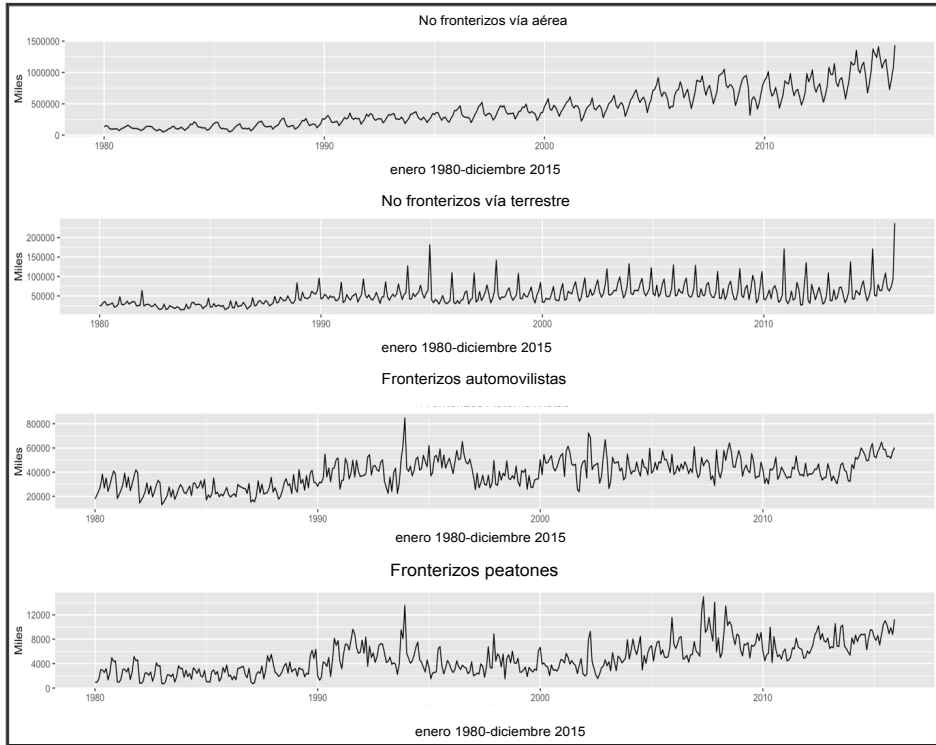


Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016).

Figura 4

Ingresos totales por visitantes internacionales a México por categoría viajeros internacionales, gasto y número de viaje

Es claro que el peso relativo del ingreso de los turistas es mucho más determinante que el peso relativo del ingreso de los excursionistas en la trayectoria del ingreso total. En la figura 5, se observa que la subcategoría de mayor peso en el ingreso de turistas es la formada por los turistas no fronterizos que entran al país por vía aérea. La importancia relativa de las otras subcategorías es considerablemente menor. Es claro también que, salvo en algunos años de marcada variabilidad, la tendencia del ingreso de estas subcategorías es horizontal.



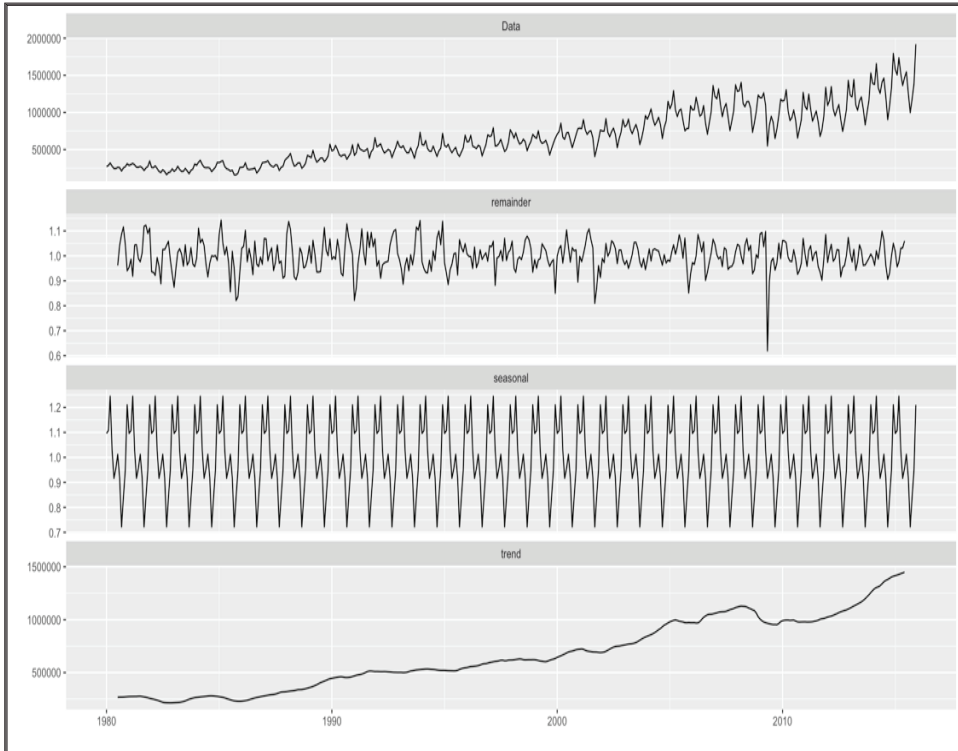
Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

Figura 5
Ingresos por categoría de turistas.

La información indica que la evolución del ingreso total está determinada por el ingreso de los turistas no fronterizos que entran por vía aérea. Esto es esperable debido a que estos turistas pasan más tiempo en el país y por lo tanto consumen más y, presumiblemente, mayor diversidad de bienes y servicios. Por esta razón, no es casual que México, como otros países, se concentre en atraer la mayor cantidad posible de turistas no fronterizos.

En la figura 6 se muestran cada uno de los componentes, aleatoriedad (*random*), estacionalidad (*seasonal*), ciclicidad y secularidad (*trend*) de la serie agregada del total de ingresos monetarios por visitantes internacionales a México, considerando las subseries del total de ingresos por turistas y del total de ingresos por excursionistas (nodo principal y nodos de primer nivel,

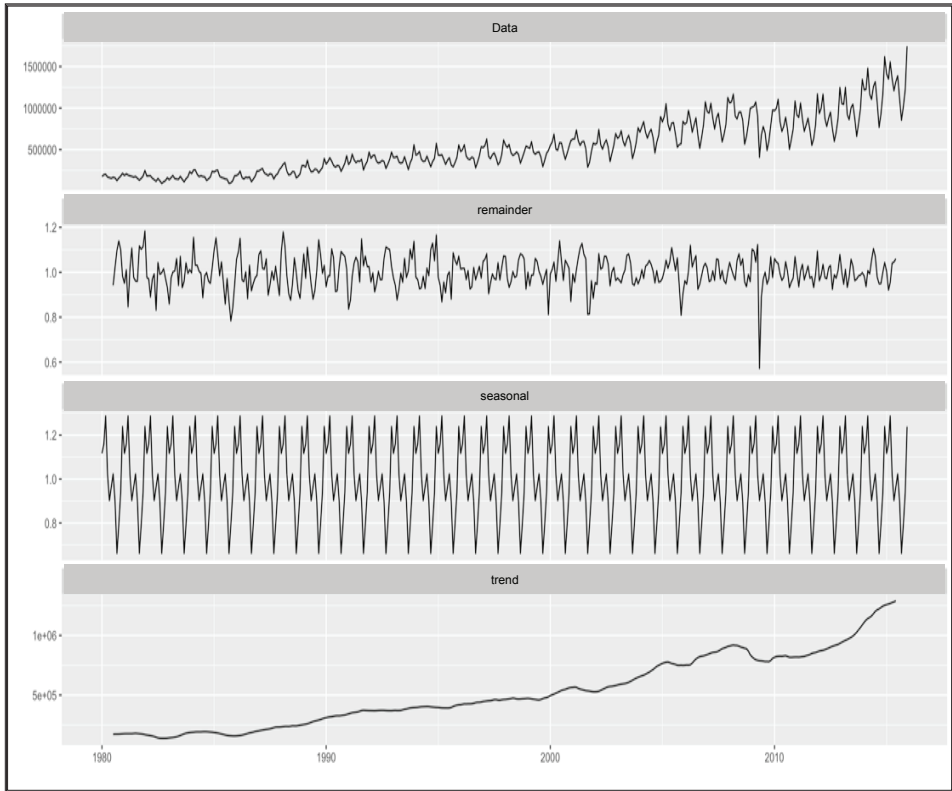
respectivamente). Esta información ilustra el hecho de que la agregación de las categorías incorpora los componentes propios de las otras dos series, mismas que, a su vez, son el resultado de agregar otras subcategorías.



Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

Figura 6
Componentes de la serie total agregada de ingresos monetarios por visitantes (turistas y excursionistas)

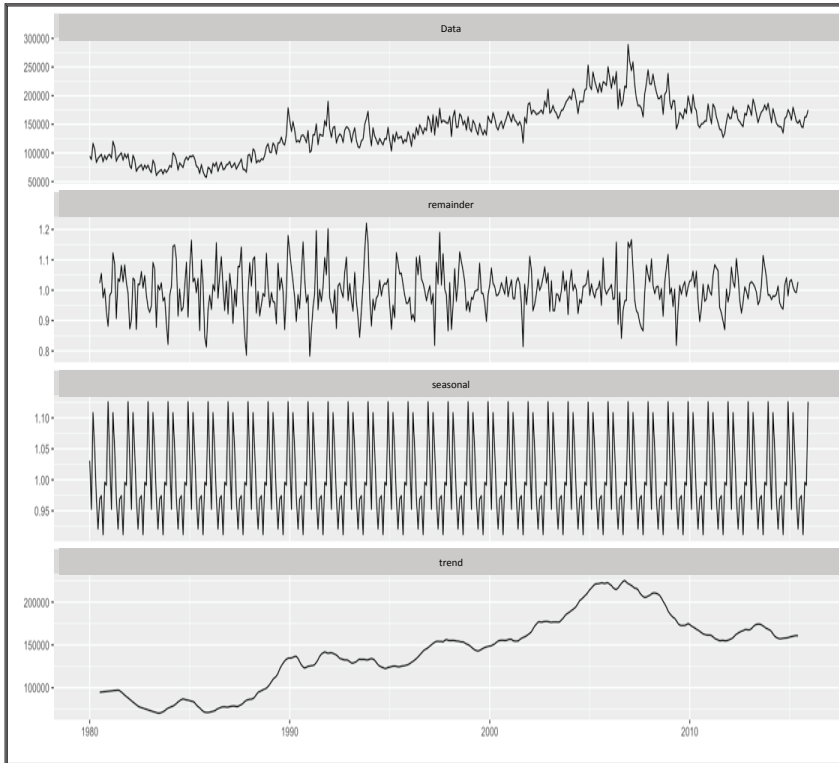
Es evidente que los componentes de la serie total de ingresos son una combinación de los componentes de las series base, correspondientes a las subcategorías primarias o fundamentales. Es indudable que la tendencia secular de los ingresos totales es influida de forma determinante por la tendencia secular de los ingresos de turistas, así como los restantes componentes, como se muestra en la figura 7.



Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

Figura 7
Componentes de la serie total agregada de ingresos
monetarios por turistas

La influencia del ingreso de los excursionistas, misma que contrasta con la del ingreso de turistas, es menos relevante debido a su menor peso relativo, aunque están presentes en la serie total. Como se muestra en la figura 8.



Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

Figura 8
Componentes de la serie total agregada de ingresos
monetarios por excursionistas

Nótese que al realizar la agregación de agregaciones de las series bases, aumenta implícitamente el nivel de ruido en la serie agregada total. Por lo tanto, la razón señal-ruido tiende a ser menor en esta serie que en las series primarias, lo cual es un obstáculo relevante para identificar un modelo eficiente de pronóstico.

4.1 Pronóstico de ingresos por visitantes internacionales con métodos jerárquicos y no jerárquicos

En la tabla 1 se presentan los resultados de la evaluación los métodos de pronóstico no jerárquicos, considerando la estrategia de prueba A.

Tabla 1
Resultados del pronóstico de la estrategia de prueba A aplicando métodos NO jerárquicos.

Panel 1: estrategia de prueba A									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
ARIMA	0.46	0.62	0.57	0.70	0.70	0.93	0.67	0.16	0.24
<i>Holt-winters</i> aditivo	0.53	0.72	0.72	0.71	0.71	0.71	0.69	0.08	0.11
<i>Holt-winters</i> aditivo amortiguado	0.52	0.72	0.77	0.80	0.80	0.74	0.73	0.11	0.15
<i>Holt-winters</i> multiplicativo	0.64	0.89	1.04	0.94	0.94	0.81	0.88	0.14	0.15
<i>Holt-winters</i> multiplicativo amortiguado	0.63	0.89	1.06	0.98	0.98	0.90	0.91	0.15	0.16
ETS	0.59	0.83	1.07	1.03	1.03	1.01	0.93	0.19	0.20
Ingenuo estacional	2.59	2.39	2.05	1.80	1.80	1.69	2.05	0.36	0.18
Ingenuo	4.08	3.38	4.30	4.82	4.82	5.85	4.54	0.84	0.18
Camino aleatorio derrape	4.17	3.53	4.51	5.09	5.09	6.12	4.75	0.89	0.19
Media	14.52	15.21	14.29	13.77	13.77	13.08	14.11	0.73	0.05
Panel 2: diferencia entre ARIMA y otros pronósticos									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
ARIMA	0.46	0.62	0.57	0.70	0.70	0.93	0.67	0.16	0.24
<i>HOLT-WINTERS</i> ADITIVO	-0.07	-0.10	-0.15	-0.01	-0.01	0.22	-0.02	0.08	0.11
<i>Holt-winters</i> aditivo amortiguado	-0.06	-0.10	-0.20	-0.10	-0.10	0.19	-0.06	0.11	0.15
<i>Holt-winters</i> multiplicativo	-0.18	-0.28	-0.46	-0.23	-0.23	0.12	-0.21	0.14	0.15
<i>Holt-winters</i> multiplicativo amortiguado	-0.17	-0.27	-0.49	-0.27	-0.27	0.03	-0.24	0.15	0.16
ETS	-0.13	-0.21	-0.50	-0.33	-0.33	-0.07	-0.26	0.19	0.20
Ingenuo estacional	-2.13	-1.77	-1.47	-1.09	-1.09	-0.76	-1.39	0.36	0.18
Ingenuo	-3.62	-2.76	-3.73	-4.12	-4.12	-4.92	-3.88	0.84	0.18
Camino aleatorio derrape	-3.71	-2.91	-3.93	-4.39	-4.39	-5.19	-4.09	0.89	0.19
Media	-14.06	-14.59	-13.72	-13.07	-13.07	-12.15	-13.44	0.73	0.05

Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

El panel 1 indica de manera ordenada los métodos no jerárquicos según la estrategia A y el panel 2 presenta la diferencia entre el método más eficiente y los otros métodos. En la tabla 2 se presentan los resultados de la evaluación los métodos de pronóstico no jerárquicos, considerando la estrategia de prueba B.

Tabla 2
Resultados del pronóstico de la estrategia de prueba B aplicando métodos NO jerárquicos

Panel 1: estrategia de prueba B									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
<i>Holt-winters</i> aditivo	0.53	0.91	0.14	0.62	0.62	1.08	0.65	0.33	0.50
<i>Holt-winters</i> aditivo amortiguado	0.52	0.92	0.18	0.62	0.62	1.09	0.66	0.32	0.48
ARIMA	0.46	0.72	0.62	0.73	0.73	0.97	0.70	0.17	0.24
ETS	0.59	1.01	0.73	0.58	0.58	0.95	0.74	0.19	0.26
<i>Holt-winters</i> multiplicativo amortiguado	0.63	1.08	0.90	0.56	0.56	0.85	0.76	0.21	0.28
<i>Holt-winters</i> multiplicativo	0.64	1.09	0.91	0.55	0.55	0.85	0.77	0.22	0.29
Ingenuo estacional	2.59	2.17	1.35	1.02	1.02	1.44	1.60	0.64	0.40
Ingenuo	4.08	1.98	1.68	2.32	2.32	7.61	3.33	2.26	0.68
Camino aleatorio derrape	4.17	1.91	1.75	2.34	2.34	7.56	3.35	2.24	0.67
Media	14.52	15.72	12.15	11.83	11.83	15.74	13.63	1.91	0.14
Panel 2: diferencia entre Holt-Winters additive y otros pronósticos									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
<i>Holt-winters</i> aditivo	0.53	0.91	0.14	0.62	0.62	1.08	0.65	0.33	0.50
<i>Holt-winters</i> aditivo amortiguado	0.01	-0.01	-0.04	0.00	0.00	-0.02	-0.01	0.32	0.48
ARIMA	0.07	0.19	-0.48	-0.11	-0.11	0.12	-0.06	0.17	0.24
ETS	-0.07	-0.10	-0.59	0.03	0.03	0.02	-0.09	0.19	0.26
<i>Holt-winters</i> multiplicativo amortiguado	-0.11	-0.16	-0.76	0.06	0.06	0.10	-0.11	0.21	0.28
<i>Holt-winters</i> multiplicativo	-0.12	-0.18	-0.77	0.07	0.07	0.00	-0.12	0.22	0.29
Ingenuo estacional	-2.06	-1.26	-1.21	-0.41	-0.41	-0.59	-0.95	0.64	0.40
Ingenuo	-3.55	-1.07	-1.54	-1.70	-1.70	-6.17	-2.68	2.26	0.68
Camino aleatorio derrape	-3.64	-1.00	-1.61	-1.73	-1.73	0.05	-2.70	2.24	0.67
Media	-13.99	-14.81	-12.01	-11.21	-11.21	-8.18	-12.98	1.91	0.14

Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

En la tabla 3 se presentan los resultados de la evaluación los métodos de pronóstico jerárquicos, considerando la estrategia de prueba A.

Tabla3
Resultados del pronóstico de la estrategia de prueba A aplicando métodos jerárquicos. Panel 1: estrategia de prueba A

Panel 1: estrategia de prueba A									
Modelo	2	4	6	8	10	12	Media	Desición estándar	Coefficiente de variación
Abajo arriba-arima	0.51	0.67	0.60	0.61	0.56	0.65	0.60	0.06	0.10
Comb-arima	0.49	0.64	0.54	0.63	0.81	0.85	0.66	0.14	0.22
Arriba abajo-arima	0.46	0.62	0.57	0.70	0.92	0.93	0.70	0.19	0.27
Medio-arima	0.54	0.67	0.61	0.72	0.92	0.94	0.73	0.16	0.22
Arriba abajo-ets	0.59	0.83	1.07	1.03	0.91	1.01	0.91	0.18	0.20
Medio-ets	1.15	1.27	1.04	0.92	0.82	0.87	1.01	0.17	0.17
Comb-ets	1.22	1.25	1.20	1.11	0.95	1.03	1.13	0.12	0.10
Abajo arriba-ets	3.00	2.34	1.92	1.75	1.50	1.50	2.00	0.58	0.29
Abajo arriba-camino aleatorio	4.08	3.38	4.30	4.82	6.16	5.85	4.77	1.07	0.22
Panel 2: diferencia entre BU-ARIMA y otros pronósticos									
Modelo	2	4	6	8	10	12	Media	Desición estándar	Coefficiente de variación
Abajo arriba-arima	0.51	0.67	0.60	0.61	0.56	0.65	0.60	0.06	0.10
Comb-arima	0.02	0.03	0.06	-0.03	-0.25	-0.20	-0.06	0.14	0.22
Arriba abajo-arima	0.05	0.05	0.03	-0.09	-0.36	-0.28	-0.10	0.19	0.27
Medio-arima	-0.03	0.00	-0.01	-0.11	-0.36	-0.29	-0.13	0.16	0.22
Arriba abajo-ets	-0.09	-0.16	-0.47	-0.43	-0.35	-0.35	-0.31	0.18	0.20
Medio-ets	-0.65	-0.60	-0.44	-0.31	-0.26	-0.22	-0.41	0.17	0.17
Comb-ets	-0.71	-0.58	-0.60	-0.50	-0.39	-0.38	-0.53	0.12	0.10
Abajo arriba-ets	-2.50	-1.67	-1.32	-1.14	-0.94	-0.85	-1.40	0.58	0.29
Abajo arriba-camino aleatorio	-3.57	-2.72	-3.70	-4.22	-5.60	-5.20	-4.17	1.07	0.22

Fuente: elaboración propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

El panel 1 indica de manera ordenada los métodos jerárquicos según la estrategia A y el panel 2 presenta la diferencia entre el método más eficiente y los otros métodos.

En la tabla 4 se presentan los resultados de la evaluación los métodos de pronóstico jerárquicos, considerando la estrategia de prueba B.

Tabla 4
Resultados del pronóstico de la estrategia de prueba B aplicando métodos jerárquicos

Panel 1: estrategia B									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
Abajo arriba-arima	0.51	0.81	0.49	0.64	0.26	1.38	0.68	0.39	0.57
Arriba abajo-ets	0.59	1.01	0.73	0.58	0.24	0.95	0.68	0.28	0.41
Comb-arima	0.49	0.75	0.62	0.71	0.51	1.06	0.69	0.21	0.30
Arriba abajo-arima	0.46	0.72	0.62	0.73	0.64	0.97	0.69	0.17	0.24
Medio-arima	0.54	0.77	0.59	0.73	0.34	1.16	0.69	0.28	0.40
Comb-ets	1.22	1.12	1.02	0.67	0.20	0.90	0.86	0.37	0.44
Medio-ets	1.15	1.20	1.07	0.44	0.42	0.90	0.86	0.35	0.41
Abajo arriba-ets	3.00	1.63	1.50	1.19	0.58	0.79	1.45	0.86	0.59
Abajo arriba-camino aleatorio	4.08	1.98	1.68	2.32	2.71	7.61	3.40	2.23	0.66
Panel 2: diferencia entre BU-ARIMA y otros pronósticos									
Modelo	2	4	6	8	10	12	Media	Desviación estándar	Coefficiente de variación
Abajo arriba-arima	0.51	0.81	0.49	0.64	0.26	1.38	0.68	0.39	0.57
Arriba abajo-ets	-0.09	-0.20	-0.25	0.05	0.03	0.43	0.00	0.28	0.41
Comb-arima	0.02	0.07	-0.13	-0.08	-0.25	0.32	-0.01	0.21	0.30
Arriba abajo-arima	0.05	0.09	-0.13	-0.09	-0.38	0.41	-0.01	0.17	0.24
Medio-arima	-0.03	0.04	-0.11	-0.10	-0.08	0.22	-0.01	0.28	0.40
Comb-ets	-0.71	-0.31	-0.54	-0.03	0.06	0.48	-0.18	0.37	0.44
Medio-ets	-0.65	-0.39	-0.58	0.20	-0.15	0.48	-0.18	0.35	0.41
Abajo arriba-ets	-2.50	-0.82	-1.01	-0.55	-0.32	0.59	-0.77	0.86	0.59
Abajo arriba-camino aleatorio	-3.57	-1.17	-1.19	-1.68	-2.45	-6.23	-2.72	2.23	0.66

Fuente: realización propia en R con datos de la balanza de pagos (Banco de México, 2016). Viajeros internacionales, gasto y número de viaje.

El panel 1 indica de manera ordenada los métodos jerárquicos según la estrategia B y el panel 2 presenta la diferencia entre el método más eficiente y los otros métodos.

Las cuatro tablas permiten apreciar los siguientes resultados:

- i) la eficiencia de los métodos por tipo de estrategia y
- ii) el desempeño de cada método respecto al método más eficiente por estrategia.

Para la estrategia A, podemos observar en las tablas 3 y 4 que en los métodos jerárquicos, el ARIMA resulta ser el más eficiente desde la perspectiva de la media, aunque ésta no es marginalmente muy superior a cualesquiera de los métodos exponenciales. Mientras que los métodos no jerárquicos mostrados en las tablas 1 y 2, muestran que el *holt-winters additive* arroja una desviación estándar menor y coeficiente de variación más pequeño aún más que en los métodos jerárquicos. Para la estrategia B, sucede justamente lo contrario. Desde el punto de vista de la media, el método ARIMA es marginalmente inferior pero arroja una desviación estándar menor.

Con base en los resultados obtenidos, se puede argumentar que la estrategia A es marginalmente superior a la estrategia B, dado que todos los métodos más eficientes arrojan un MASE inferior a 1 y que el método ARIMA de abajo hacia arriba (BU-ARIMA) resulta el más eficiente independientemente del tipo de estrategia; no obstante, este método tiene importantes contendientes en la estrategia B. Por ejemplo, el TD-ETS y el COMB-ARIMA arrojan medias similares, pero desviaciones estándar menores, que el BU-ARIMA. Note además que éste es particularmente ineficiente para el último bimestre.

5. Conclusiones

En principio, un enfoque jerárquico debería identificar el método de pronóstico más eficiente para una serie jerárquica que un enfoque no jerárquico. La evaluación realizada comprueba este supuesto: el método BU-ARIMA es el método más eficiente para pronosticar los ingresos por visitantes a México, independientemente de la estrategia de selección del conjunto de prueba (*training set*). Desde el punto de vista de cualquier indicador (la media, la desviación estándar o el coeficiente de variación), dicho método es

ciertamente superior o más eficiente. Es importante indicar, no obstante, que varios métodos (TD-ETS, COMB-ARIMA, TD-ARIMA y MO-ARIMA) resultan más eficientes que el BU-ARIMA cuando se utiliza la estrategia B. En esta estrategia, el enfoque de arriba hacia abajo, independientemente del método, es ineficiente para pronosticar el último bimestre.

Referencias

- Banco de México (7 de marzo de 2016). *Balanza de Pagos*. Obtenido de CE36 - Cuenta de Viajeros Internacionales. Periodo: enero 198-diciembre 2015, mensual: <http://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?accion=consultarCuadro&idCuadro=CE36§or=1&locale=es>.
- Hyndman, R. J. (2014). *Forecasting: Principles & Practice*. NY. USA.: Prentice Hall.
- Hyndman, R. J.; R. A. Ahmed; G. Athanasopoulos y S. H. Lin (2011). "Optimal combination forecasts for hierarchical time series". *Computational Statistics & Data Analysis*, vol. 55, Issue 9., pp. 2579-2589.
- Montemayor, G., J. E. (2012). *Métodos de Pronósticos para Negocios*. Ciudad de México: Editorial Digital, Tecnológico de Monterrey.
- Stock, J. H. (2006). *Handbook of Economic Forecasting*, vol. 1. Harvard: Graham Elliott, Clive W.J. Granger and Allan Timmermann, pp. 120-175.