



NOTA TÉCNICA
Nº 05 | 2024

Pronósticos univariados de inflación en Costa Rica: evaluación y selec- ción de modelos

Alberto Vindas Quesada
Carlos Brenes Soto
Susan Jiménez Montero
Andrea Jiménez Morales
Adriana Sandí Esquivel

Fotografía de portada: "Presentes", conjunto escultórico en bronce, año 1983, del artista costarricense Fernando Calvo Sánchez. Colección del Banco Central de Costa Rica.

Pronósticos univariados de inflación en Costa Rica: evaluación y selección de modelos

Alberto Vindas Quesada[†]

Carlos Brenes Soto[‡]

Susan Jiménez Montero[§]

Andrea Jiménez Morales[¶]

Adriana Sandí Esquivel^{||}

Las ideas expresadas en este documento son de los autores y no necesariamente representan las del Banco Central de Costa Rica.

Resumen

Esta nota presenta la metodología que usa el Banco Central de Costa Rica para la evaluación y selección de los modelos univariados de proyección de inflación en el corto plazo. La metodología consiste en la cuantificación de varias propiedades deseables en modelos de pronóstico, la asignación de puntajes y su combinación para obtener un puntaje final. Se evalúa la robustez de la selección de los modelos a cambios en el periodo de evaluación, dados los cambios recientes en la dinámica inflacionaria. La selección del modelo es sensible a este periodo, por lo que se recomienda actualizar la selección con regularidad.

Palabras clave: inflación, pronósticos, volatilidad estocástica, GARCH, evaluación.

Clasificación JEL: E31, E37, C52, C53.

[†]Departamento de Investigación Económica. División Económica, BCCR.VINDASQA@bccr.fi.cr.

[‡]División Económica, BCCR.BRENESSC@bccr.fi.cr.

[§]Departamento de Investigación Económica. División Económica, BCCR.JIMENEZMS@bccr.fi.cr.

[¶]Departamento Gobierno y Gestión de Datos. División de Análisis de datos y Estadísticas, BCCR.JIMENEZMD@bccr.fi.cr.

^{||}Depto. de Integración y Análisis de Datos. División de Análisis de datos y Estadísticas, BCCR.SANDIEA@bccr.fi.cr.

Univariate inflation forecasts in Costa Rica: model evaluation and selection

Alberto Vindas Quesada[†]

Carlos Brenes Soto[‡]

Susan Jiménez Montero[§]

Andrea Jiménez Morales[¶]

Adriana Sandí Esquivel^{||}

The ideas expressed in this paper are those of the authors and not necessarily represent the view of the Central Bank of Costa Rica.

Summary

This document presents the methodology that the Central Bank of Costa Rica uses to evaluate and select the univariate models for short-horizon forecasting purposes. This methodology consists on quantifying several properties that are deemed desirable for forecasting models, assigning scores and combining them to obtain a final score. The robustness of the model selection to the evaluation period is analyzed, given the recent inflation dynamics. The selection is sensitive to this period, leading to the recommendation of regular selection processes.

Key words: inflation, forecasting, stochastic volatility, GARCH, evaluation.

JEL Codes: E31, E37, C52, C53.

[†]Research Department. Economic Division, BCCR.VINDASQA@bccr.fi.cr.

[‡]Economic Division, BCCR.BRENESSC@bccr.fi.cr.

[§]Research Department. Economic Division, BCCR.JIMENEZMS@bccr.fi.cr.

[¶]Data Governance and Management Dept. Data Analysis and Statistics Division, BCCR.JIMENEZMD@bccr.fi.cr.

^{||}Data Integration and Analysis Department. Data Analysis and Statistics Division, BCCR.SANDIEA@bccr.fi.cr.

1. Introducción

El Banco Central de Costa Rica (BCCR) mantiene actualmente un esquema monetario de metas de inflación en el cual la toma de decisiones se centra en el cumplimiento del objetivo inflacionario de la Institución. Esto requiere un esfuerzo continuo por entender los efectos de la política monetaria en la inflación y el comportamiento de ésta variable a lo largo del tiempo.

La política monetaria que mantiene el BCCR se centra principalmente en el uso de la Tasa de Política Monetaria con el fin de influir en el ritmo de variación de los precios, y así cumplir su meta de inflación en el mediano plazo. Esta relación, entre las tasas de interés en una economía y el ritmo de variación del nivel de precios, fue propuesta hace ya varios años, los escritos de Knut Wicksell son las exposiciones más antiguas sobre este tema. Muchos bancos centrales, actualmente, siguen estas ideas.

Los formuladores de política económica y académicos saben que los rezagos con los que la política monetaria actúa son largos y variables. Esta manera de describir los efectos fue acuñada por Friedman (1959), la cuál sigue gozando de vigencia, como expone Dupor (2023).

Los efectos rezagados de la política monetaria tienen implicaciones importantes para la toma de decisiones de un banco central. Cuando los rezagos son largos, las decisiones requieren de un enfoque prospectivo. Por lo tanto, se debe de considerar la evolución esperada de las principales variables económicas. Así, el desarrollo de herramientas de pronóstico es de gran importancia para informar la toma de decisiones.

Que los efectos de la política monetaria muestren rezagos variables también tiene implicaciones para la toma de decisiones de un banco central. Esta variabilidad tiene su origen en la heterogeneidad de choques y condiciones presentes en una economía. Al ir cambiando en el tiempo, es de esperar que las reacciones hacia la política monetaria cambien, por lo que es importante mantener los modelos de pronóstico actualizados.

Estas dos características de la economía y de los efectos de la política monetaria son las que justifican la presente nota técnica. Es preciso mantener herramientas de pronóstico de inflación que reflejen la información más reciente. Unas herramientas que usa el BCCR para pronóstico en horizontes cortos son los modelos de pronóstico univariados, los cuales consideran únicamente la evolución de la inflación y son la clase de modelos en la que se centra este documento.

La metodología se basa en el enfoque desarrollado en Fuentes y Rodríguez (2016), el cual considera varios modelos y los compara entre sí. Este documento difiere principalmente en dos aspectos: considera un periodo más extenso, incluye datos hasta finales del año 2023, y modifica los criterios de selección para la elección final del modelo.

Los criterios de selección se basan en un puntaje ponderado, que cuantifica varias propiedades que son deseables en modelos de pronóstico. Así, se desarrolla un procedimiento que permite la valoración de varios modelos de una manera coherente y que da flexibilidad en la elección de los ponderadores.

Los cambios recientes en la inflación proveen un escenario adecuado para evaluar si el proceso de elección de los modelos es sensible al periodo de evaluación que se usa. Al usar distintos periodos de evaluación, la elección final cambia, lo cual resalta la importancia de actualizar la elección regularmente.

El documento está organizado en siete secciones, que comienzan con esta introducción. La segunda sección describe los datos de inflación, la tercera detalla los modelos a considerar. Las secciones cuatro y cinco describen la creación y evaluación de los pronósticos, junto con la prueba de robustez considerada. La sexta sección analiza los resultados y la séptima concluye.

2. Datos

Esta nota, al centrarse en el uso de modelos de pronóstico univariados, requiere únicamente de una serie de tiempo: el índice de precios al consumi-

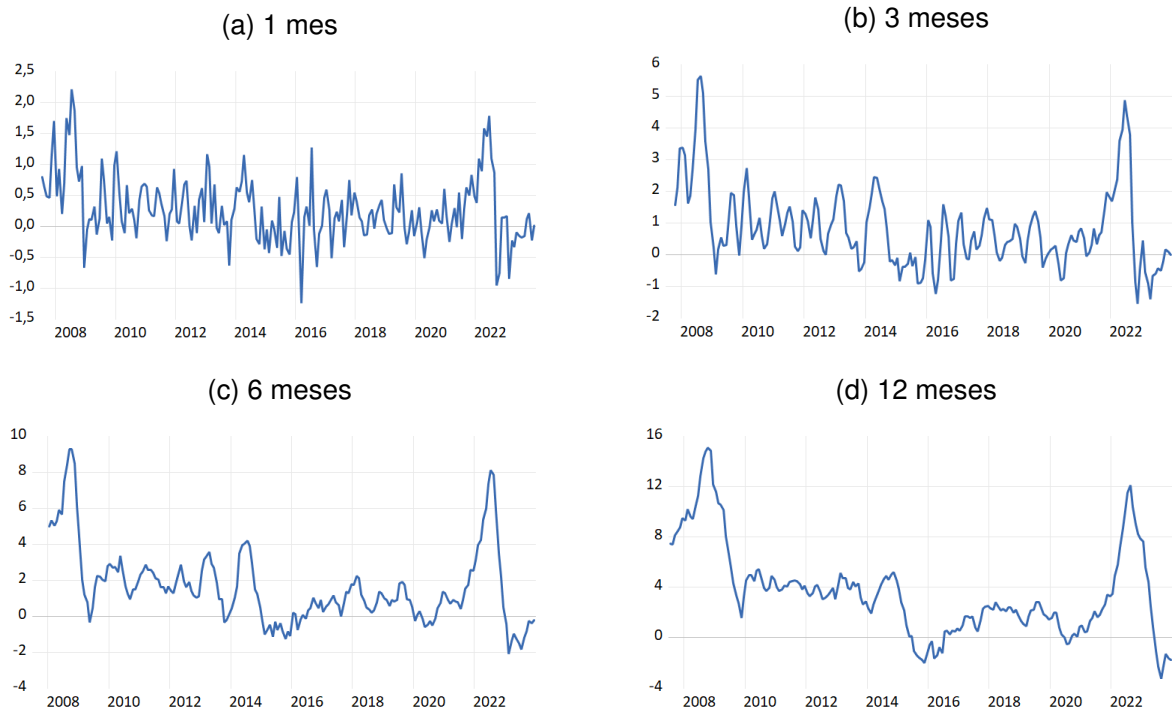
dor (IPC). El periodo de estudio comprende los meses entre julio de 2007 y diciembre de 2023. Durante este periodo, el Instituto Nacional de Estadística y Censos (INEC), a cargo del IPC, ha publicado tres series: la que toma como periodo base julio de 2006, la que toma como periodo base junio de 2015 y la que toma como periodo base diciembre de 2020. Estas tres se enlazan por medio de sus variaciones porcentuales, para obtener la serie que se usa en esta nota.

La variación porcentual del IPC a distintos plazos se presenta en el gráfico 1. La variación a 12 meses plazo es el indicador más frecuentemente usado para medir la inflación en el país. Esta variación, a partir del año 2022, ha sido atípica por varias razones. Por un lado, la relajación de las medidas sanitarias, debido al reciente episodio pandémico, creó tensiones en las cadenas de suministro a nivel mundial. Esto desembocó en un aumento en los costos de los fletes y de los productos en los mercados internacionales. Por otro lado, la invasión de Rusia a Ucrania resultó en aumentos en los precios de los hidrocarburos y granos básicos mundialmente. Así, ambos eventos desembocaron en aumentos de precios locales, llegando a su punto máximo en agosto de 2022. A partir de ese mes, los choques se empezaron a revertir. Esto causó la caída en el IPC y condujo a que la variación interanual llegara al mínimo histórico de -3,23 % en agosto de 2023. Las variaciones a otros plazos muestran un comportamiento similar puesto que, a fin de cuentas, forman parte de esos cambios a 12 meses.

La presencia de cambios de esa magnitud podría determinar un quiebre estructural en la dinámica de la variación de precios, lo cual debería considerarse dentro de la estrategia de modelación. Por esto, se realizaron pruebas econométricas para identificar quiebres estructurales en las variaciones intermensuales del IPC. Los resultados, que se presentan en detalle en el apéndice A, sugieren la presencia de un único quiebre en marzo de 2010. Este periodo coincide con la estabilización de la inflación en Costa Rica posterior a dos eventos importantes: la gran crisis financiera de 2008 y la migración hacia un

Gráfico 1: Variaciones porcentuales del IPC

Junio 2007 - diciembre 2023



esquema monetario de metas de inflación en Costa Rica. Torres (2012), que estudia la dinámica inflacionaria durante el periodo 1997-2011, identifica un quiebre en mayo de 2009. Esta discrepancia se debe al cambio del periodo analizado. Siguiendo la recomendación de Pesaran y Timmermann (2007), se incluyen las observaciones previas al quiebre estructural.

El cuadro 1 muestra algunos estadísticos de resumen para las variaciones del IPC a distintos horizontes. Este confirma que, como sería de esperarse, las muestras antes y después del quiebre estructural identificado son suficientemente distintas.

Cuadro 1: Estadísticos de resumen de variaciones porcentuales del IPC

Horizonte	Muestra	Media	Mediana	Desviación estándar	Coef. de asimetría	Curtosis
1 mes	Jul. 2007 - Mar. 2010	0,660	0,626	0,660	0,401	2,738
	Abr. 2010 - Dic. 2023	0,195	0,160	0,458	0,359	4,229
	Jul. 2007 - Dic. 2023	0,272	0,196	0,525	0,681	4,361
3 meses	Jul. 2007 - Mar. 2010	2,039	1,857	1,642	0,598	2,707
	Abr. 2010 - Dic. 2023	0,592	0,475	1,041	1,110	5,524
	Jul. 2007 - Dic. 2023	0,815	0,553	1,262	1,290	5,348
6 meses	Jul. 2007 - Mar. 2010	4,080	2,908	2,801	0,447	2,107
	Abr. 2010 - Dic. 2023	1,229	0,948	1,709	1,268	5,934
	Jul. 2007 - Dic. 2023	1,630	1,203	2,136	1,398	5,328
12 meses	Jul. 2007 - Mar. 2010	8,681	9,304	3,676	-0,069	2,224
	Abr. 2010 - Dic. 2023	2,601	2,410	2,618	0,718	4,571
	Jul. 2007 - Dic. 2023	3,614	3,076	3,614	1,000	3,958

3. Modelos

3.1. Modelos ARIMA

Los modelos ARIMA se utilizan frecuentemente para el pronóstico de series de tiempo en horizontes cortos. Esto ha resultado en una amplia gama de estudios sobre sus propiedades, que ha afianzado su utilidad en el campo de los pronósticos. Para este trabajo, se consideran 9 modelos ARIMA.

Los modelos ARIMA se describen usando tres parámetros, p , d y q :

- El parámetro p especifica el orden del proceso autorregresivo (AR). En un proceso autorregresivo, la variable a modelar, x_t , sigue un proceso estocástico en donde sus realizaciones pasadas forman parte de su comportamiento:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + u_t \quad (1)$$

donde ϕ_i es el coeficiente AR asociado al rezago i y u_t es un choque aleatorio.

- El parámetro d especifica el orden de integración de la serie. Es decir, indica la cantidad de veces que la serie debe diferenciarse para obtener una que sea estacionaria, la cual puede ser modelada con un proceso ARMA estable.
- El parámetro q especifica el orden del proceso de medias móviles (MA). En un proceso estocástico de medias móviles, la variable a modelar se caracteriza por estar determinada por un choque aleatorio y sus rezagos:

$$x_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

La media de la serie es μ , ε_{t-j} son los choques aleatorios y θ_j son los coeficientes MA asociados al rezago j .

El orden de integración de la variación intermensual del IPC es 0,¹ por lo que se pueden estimar modelos sin tener que hacer más transformaciones en los datos. Cuando $d = 0$, se puede omitir el orden de integración en la descripción del modelo, y se expresa como $ARMA(p,q)$.

Las variaciones analizadas, al ser intermensuales, típicamente muestran un comportamiento estacional. Siendo así, es apropiado usar variables dicotómicas para ajustar por este comportamiento a lo largo del año. Por lo tanto, los modelos a considerar son del tipo:

$$x_t = \sum_{m=1}^{12} \mu_m \mathbb{I}(t, m) + \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (3)$$

donde $\mathbb{I}(t, m)$ es una función indicador que es igual a 1 cuando el periodo t corresponde con el mes m y 0 cuando no, y μ_m es la media durante el mes m .

La selección de los órdenes de los procesos autorregresivos y de medias móviles, p y q , se basa en los criterios de información de Akaike, Schwarz y Hannan-Quinn. Se consideran modelos con p y q variando entre 0 y 12 rezagos, por lo que se evalúan 169 combinaciones. Los primeros 3 modelos elegidos son los que minimizan cada uno de estos criterios de información.

Los siguientes 6 modelos a elegir resultan de dos variaciones de los modelos ya evaluados. Estas variaciones se hacen con el fin de lidiar con la estacionalidad de periodicidad anual (que en este caso corresponden a rezagos en múltiplos de 12 periodos). La primera variación considera la opción de incluir términos AR estacionales y la segunda, términos MA estacionales. Esto implica considerar la inclusión de rezagos de 12 y 24 meses en los términos AR y MA. Esto se denota en la descripción del modelo con SAR(1), SAR(2) y SAR(1,2) para el caso de los términos AR estacionales, y con SMA(1), SMA(2) y SMA(1,2) para el caso de los términos MA estacionales. La selección, igualmente, se hace según los criterios de información de Akaike, Schwarz y Hannan-Quinn: tres especificaciones para los modelos que agre-

¹El resultado de las pruebas analizadas se incluye en el cuadro 6 del apéndice A.

gan términos AR estacionales y tres para los modelos que agregan términos MA estacionales.

Este proceso de elección da como resultado 9 modelos para evaluar. Sin embargo, es posible que algunos criterios den como resultado una misma especificación.

3.2. Modelos estructurales de componentes no observables

En esta sección, se describen los modelos de componentes no observables partiendo de la descomposición clásica de series de tiempo. Esta plantea que una serie de tiempo puede descomponerse de la siguiente manera:

$$x_t = \mu_t + \psi_t + \gamma_t + \varepsilon_t \quad (4)$$

donde μ_t representa el componente tendencial de la serie, ψ_t el componente cíclico de la serie, γ_t el componente estacional y ε_t el componente irregular. Estos no son directamente observables, por lo que deben ser estimados a través de un modelo.

El primer componente, el tendencial, está compuesto de su realización en el periodo anterior (μ_{t-1}), una pendiente que puede cambiar a lo largo del tiempo (β_{t-1}) y un choque aleatorio. Se modela de la siguiente manera:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (5)$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad (6)$$

donde η_t y ζ_t son perturbaciones del tipo ruido blanco, que tienen media cero y variancia constante. Esta estructura implica que la tendencia es estocástica, dado que las perturbaciones de η_t y ζ_t se acumulan a lo largo de los periodos.

El componente cíclico, ψ_t , se modela como una función de tiempo cíclica

con frecuencia λ_ψ :

$$\begin{pmatrix} \psi_t \\ \check{\psi}_t \end{pmatrix} = \rho \begin{pmatrix} \cos \lambda_\psi & \text{sen } \lambda_\psi \\ -\text{sen } \lambda_\psi & \cos \lambda_\psi \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \check{\psi}_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \check{\epsilon}_t \end{pmatrix} \quad (7)$$

En esta especificación, el ciclo es una combinación trigonométrica de senos y cosenos que se representa de manera recursiva. Las perturbaciones, ϵ_t y $\check{\epsilon}_t$, no están correlacionadas y tienen una misma variancia.

Para el componente estacional, se modelan tres especificaciones distintas. Primeramente, se considera un modelo que no incluye estacionalidad. Después, se considera un modelo que incorpora la estacionalidad a través de variables dicotómicas y, finalmente, se emplea un modelo en el cual la estacionalidad se comporta de manera similar al componente cíclico.

Para incorporar la estacionalidad a través de variables dicotómicas, se pueden incluir variables indicadoras. Estas típicamente se restringen de manera que sumen cero a lo largo de su ciclo:

$$\sum_{m=1}^{12} \gamma_{t-m+1} = 0 \quad (8)$$

El modelo considerado muestra un poco más de flexibilidad, pues no impone que la suma de estos términos siempre sea cero. En este caso, se supone que la suma de estos términos estacionales es igual a una perturbación aleatoria, ω_t , que tiene media cero y variancia constante:

$$\sum_{m=1}^{12} \gamma_{t-m+1} = \omega_t \quad (9)$$

La tercera especificación para el término estacional se hace a través de un conjunto de términos trigonométricos, de manera similar al término cíclico. En este caso, el término estacional se obtiene sumando unos términos en las

frecuencias estacionales $\lambda_{\gamma,j} = 2\pi j/12$, con $j = 1, 2, \dots, 6$:

$$\gamma_t = \sum_{j=1}^6 \gamma_{j,t} \quad (10)$$

donde

$$\begin{pmatrix} \gamma_{j,t} \\ \check{\gamma}_{j,t} \end{pmatrix} = \rho \begin{pmatrix} \cos \lambda_{\gamma,j} & \text{sen } \lambda_{\gamma,j} \\ -\text{sen } \lambda_{\gamma,j} & \cos \lambda_{\gamma,j} \end{pmatrix} \begin{pmatrix} \gamma_{j,t-1} \\ \check{\gamma}_{j,t-1} \end{pmatrix} + \begin{pmatrix} e_{j,t} \\ \check{e}_{j,t} \end{pmatrix} \quad (11)$$

Las perturbaciones aleatorias $e_{j,t}$ y $\check{e}_{j,t}$ tienen media cero, no están correlacionadas y tienen una misma variancia.

Los tres modelos de componentes no observables expuestos (que varían en su abordaje del componente estacional) son de tipo recursivo. Siendo así, se pueden escribir en su forma estado-espacio para estimar mediante el filtro de Kalman.

3.3. Modelos con volatilidad no constante

Los modelos ARIMA y de componentes no observables expuestos anteriormente tienen en común el supuesto de la homocedasticidad. Flexibilizar este supuesto permite modelar más apropiadamente episodios donde la volatilidad de la serie de tiempo cambia. Por ejemplo, la variación interanual del IPC se mantuvo alrededor de un 4% entre 2009 y 2014, entre el 2015 y el 2020 fluctuó alrededor de un 1%, a partir de 2021 incrementó hasta llegar al pico de un 12,13% en agosto de 2022, para después llegar al mínimo histórico de -3,23% en agosto de 2023. Esto se puede observar en el panel (d) del gráfico 1. Estos cambios también se ven reflejados en las variaciones a 1, 3 y 6 meses, por lo que es razonable permitir que la volatilidad de las perturbaciones cambie.

Para modelar esta heterocedasticidad se consideran dos enfoques distintos: la estimación de modelos GARCH e EGARCH y la estimación de modelos de volatilidad estocástica.

3.3.1. Modelos GARCH e EGARCH

Estos dos modelos tienen su origen en Engle (1982), donde se introduce el modelo de heterocedasticidad condicional autorregresiva (ARCH, por sus siglas en inglés). En los modelos ARIMA expuestos anteriormente, se trabaja bajo el supuesto de que las perturbaciones ε_t se distribuyen con una variancia constante σ^2 . En un modelo ARCH, se trabaja bajo el supuesto de que la variancia va cambiando a lo largo del tiempo, según un proceso que depende de las q realizaciones más recientes de las perturbaciones:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (12)$$

Los modelos GARCH son una generalización de los modelos ARCH, presentados inicialmente en Bollerslev (1986). En estos, se cambia la estructura que determina la variancia, de manera que también depende de sus últimas p realizaciones:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2 \quad (13)$$

Los modelos EGARCH suponen una distribución exponencial para la variancia, que pasa por la transformación logarítmica para hacer los componentes aditivamente separables. Este modelo permite separar el efecto de los choques estocásticos en su componente de magnitud y de signo, lo que permite incorporar la posibilidad de asimetrías. Al igual que en los modelos GARCH, se denomina con p el orden de rezago de las últimas realizaciones de la volatilidad; ahora q denota el orden de las magnitudes de los últimos choques estocásticos y r el de su asimetría:

$$\log(\sigma_t^2) = \omega + \sum_{j=1}^q \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^p \alpha_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{k=1}^r \gamma_k \frac{\varepsilon_{t-k}}{\sigma_{t-k}} \quad (14)$$

El punto de partida para estas dos alternativas de modelaje son los modelos ARMA seleccionados anteriormente. Para cada uno de ellos, se mo-

de la variancia de los errores según una especificación GARCH(1,1) y una EGARCH(1,1,1).² Esto da como resultado 18 modelos adicionales.

3.3.2. Modelos de volatilidad estocástica

Los modelos de volatilidad estocástica fueron propuestos inicialmente en Taylor (1986). Estos consideran una segunda manera de lidiar con variancias que no son constantes: modelarlas de acuerdo con un proceso estocástico. Esto da más flexibilidad, dado que el componente estocástico permite adaptar la volatilidad a cambios abruptos, pensados en las reacciones a noticias observadas en series financieras. Computacionalmente, estos modelos son más complejos, por lo que su adopción ha sido más lenta.

En este documento se consideran dos modelos, siguiendo el trabajo de Chan (2013). El modelo base descompone la variable observada (y_t) en dos componentes: el que determina la media de la variable (μ_t) y el que determina las perturbaciones (ε_t^y):

$$y_t = \mu_t + \varepsilon_t^y \quad (15)$$

El componente que determina la media se modela de dos maneras. Primeramente, se considera una especificación autorregresiva:

$$\mu_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} \quad (16)$$

y la segunda es una de componentes no observados:

$$\mu_t = \tau_t \quad (17)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim N(0, \sigma_\tau^2) \quad (18)$$

El componente que determina las perturbaciones se modela con volatilidad estocástica. En el caso de Chan (2013), se incorpora un componente de

²Por simplicidad, se decidió no considerar distintas especificaciones GARCH e EGARCH.

medias móviles para mejorar el desempeño del modelo:

$$\varepsilon_t^y = u_t + \psi_1 u_{t-1} + \dots + \psi_q u_{t-q}, \quad u_t \sim N(0, e^{h_t}) \quad (19)$$

$$h_t = \mu_h + \phi_h (h_{t-1} - \mu_h) + \varepsilon_t^h, \quad \varepsilon_t^h \sim N(0, \sigma_h^2) \quad (20)$$

Por lo tanto, la cantidad de modelos de volatilidad estocástica considerados son dos.

4. Creación de pronósticos

Los modelos de pronóstico univariados se usan en el BCCR para tener una idea del trayecto esperado de la inflación, en especial a horizontes cortos. Eso implica crear, con los datos al periodo t , pronósticos a distintos plazos. Son de particular interés los pronósticos a 1, 3, 6 y 12 meses, los cuales son evaluados en esta nota. Para esto, es necesario dividir la muestra en dos periodos: en el primer periodo se lleva a cabo la estimación y diagnóstico inicial de los modelos, en el segundo periodo se evalúan las propiedades de los pronósticos. Al igual que en Fuentes y Rodríguez (2016), se toman los últimos dos años de información para la evaluación. En este caso, el periodo de estimación y diagnóstico cubre las observaciones entre julio de 2007 y diciembre de 2021, y el de evaluación entre enero de 2022 y diciembre de 2023.

Los modelos de componentes no observables y los de volatilidad estocástica tienen una estructura ya definida, como se expuso en la sección anterior. Los modelos ARIMA, GARCH e EGARCH requieren definir los órdenes de los componentes AR y MA. Estos se definen según los criterios de información de Akaike, Schwarz y Hannan-Quinn y dependen de la muestra que se use para estimar el modelo. En este caso, se usa el periodo de estimación y diagnóstico para definir esos órdenes, que se mantienen fijos para la creación de los pronósticos.

Para la creación de los pronósticos a evaluar, es importante recalcar dos

aspectos sobre la metodología: la ventana creciente del periodo de estimación y el ajuste de los horizontes de pronóstico. La primera ventana de estimación del modelo para hacer pronósticos comienza en julio de 2007 y termina en diciembre de 2021. Con esta ventana, se estiman los pronósticos para $t + 1$, que sería enero de 2022, $t + 3$, que sería marzo de 2022, $t + 6$, que sería junio de 2022 y $t + 12$, que sería diciembre de 2022. Con estos datos, se crean las primeras observaciones de las series de pronósticos. Para crear las segundas observaciones, se alarga la ventana de estimación en un periodo, ahora terminando en enero de 2022. Con esta nueva ventana, se reestiman los parámetros de los modelos (manteniendo la misma estructura) y se vuelven a hacer los pronósticos a $t + 1$, $t + 3$, $t + 6$ y $t + 12$. Así se continúa hasta completar las series de pronósticos en diciembre de 2023. La idea de este ejercicio es evaluar los pronósticos considerando la manera en la que se usarían, para pronósticos fuera de muestra. Así, la serie de pronósticos a 12 meses plazo comienza en diciembre de 2022, para un total de 13 observaciones. Las series de pronósticos a 6, 3 y 1 meses plazo comienzan en junio, marzo y enero de 2022, para un total de 19, 22 y 24 observaciones. Estos detalles se presentan en el cuadro 2.

El segundo aspecto es el ajuste de los horizontes de pronóstico. Los modelos están especificados en frecuencia mensual, y el interés es evaluar la capacidad predictiva de la variación del IPC a distintos plazos. Por lo tanto, para $t + 3$, $t + 6$ y $t + 12$, se deben acumular los pronósticos de variación, aplicando el cambio intermensual de los pronósticos.

5. Evaluación de pronósticos

En esta sección, se explican tres aspectos del proceso de evaluación: los criterios a considerar, la combinación de los resultados y la evaluación de la robustez al periodo de evaluación.

Cuadro 2: Distribución de meses para la creación de los pronósticos a evaluar

Periodo para estimación del modelo	Mes del pronóstico a			
	1 mes	3 meses	6 meses	12 meses
Jul. 2007 - Dic. 2021	Ene. 2022	Mar. 2022	Jun. 2022	Dic. 2022
Jul. 2007 - Ene. 2022	Feb. 2022	Abr. 2022	Jul. 2022	Ene. 2023
Jul. 2007 - Feb. 2022	Mar. 2022	May. 2022	Ago. 2022	Feb. 2023
...
Jul. 2007 - Dic. 2022	Ene. 2023	Mar. 2023	Jun. 2023	Dic. 2023
Jul. 2007 - Ene. 2023	Feb. 2023	Abr. 2022	Jul. 2023	-
...				
Jul. 2007 - Jun. 2023	Jul. 2023	Set. 2023	Dic. 2023	-
Jul. 2007 - Jul. 2023	Ago. 2023	Oct. 2023	-	-
...				
Jul. 2007 - Set. 2023	Oct. 2023	Dic. 2023	-	-
Jul. 2007 - Oct. 2023	Nov. 2023	-	-	-
Jul. 2007 - Nov. 2023	Dic. 2023	-	-	-
Total de pronósticos	24	22	19	13

5.1. Criterios de evaluación

Para determinar el desempeño de cada modelo al predecir la inflación futura, esta nota se basa en siete criterios, los cuales se explican a continuación. Estos reflejan varias propiedades que se consideran deseables para un modelo de pronóstico. La evaluación de cada propiedad dará como resultado un indicador, que posteriormente se usa para la asignación de los puntajes finales.

5.1.1. Precisión

La precisión de una serie de pronósticos puede medirse según lo cercano que se encuentran los pronósticos de los valores realizados. Hay varias maneras de medir la precisión. En esta ocasión, se calculan dos medidas ampliamente utilizadas y similares entre sí: la raíz del error cuadrático medio (RECM) y el índice de desigualdad de Theil. Adicionalmente, se considera la proporción de aciertos en la dirección de los cambios.

Para presentar las fórmulas de estas medidas, es preciso definir alguna notación. Se denota con y_t y \hat{y}_t la observación y el pronóstico del periodo t , la cantidad de periodos de pronóstico con h y el periodo a partir del cual se hacen pronósticos con T . Δ^p representa la variación al plazo de p periodos, sgn es una función que indica el signo y \mathbb{I} una función indicadora que es igual a 1 si se cumple la condición a evaluar e igual a 0 en otro caso. Los tres indicadores de precisión que se calculan son:

$$\text{RECM} = \sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h} \quad (21)$$

$$\text{Theil} = \frac{\sqrt{\sum_{t=T+1}^{T+h} (y_t - \hat{y}_t)^2}}{\sqrt{\sum_{t=T+1}^{T+h} y_t^2} + \sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^2}} \quad (22)$$

$$\text{Aciertos} = \sum_{t=T+1}^{T+h} \frac{\mathbb{I}(\text{sgn}(\Delta^p y_t) = \text{sgn}(\Delta^p \hat{y}_t))}{h} \quad (23)$$

Los primeros dos indicadores son similares, ya que usan la suma del cuadrado de los errores de pronóstico. La diferencia aparece en la normalización que se les da: para la RECM se usa la cantidad de pronósticos y para el índice de Theil la suma de los cuadrados de las variables. La proporción de aciertos es un indicador más laxo de la precisión de los pronósticos. En vez de medir lo cercano que están los pronósticos de los valores observados, mide la proporción de veces que aciertan la dirección de los cambios en la variable.

5.1.2. Insesgamiento

Un pronóstico es insesgado si, en promedio, es igual al valor observado. Este es un criterio de alta importancia, ya que permite discernir entre los modelos que crean pronósticos sistemáticamente por encima o por debajo de los valores observados.

Para evaluar esta propiedad, se realiza la prueba de insesgamiento de Mincer y Zarnowitz (1969). En esta, se estima la siguiente regresión que involucra los valores observados (y_t) y los pronósticos (\hat{y}_t):

$$y_t = \beta_0 + \beta_1 \hat{y}_t + \varepsilon_t \quad (24)$$

Se concluye que la serie de pronóstico es insesgada si no se puede rechazar la hipótesis de que $\beta_0 = 0$ y que $\beta_1 = 1$ conjuntamente, a través de una prueba de Wald. Para evaluar esta propiedad, se toman en cuenta dos indicadores. Primeramente, se considera una variable dicotómica que es igual a uno en caso de no rechazar la hipótesis nula con un nivel de significancia del 5%. Adicionalmente, se considera el valor p asociado a la prueba; un valor más alto indica evidencia más fuerte de insesgamiento en comparación con un valor que esté cercano al umbral del 5%.

5.1.3. Capacidad superior de pronóstico

Para evaluar la capacidad de un modelo de producir mejores pronósticos que otro, se evalúan dos hipótesis: si la capacidad predictiva de los modelos es distinta y si los modelos están reflejando la misma información. En ambos casos, la manera de evaluar es haciendo pruebas entre todos los pares de pronósticos disponibles.

La primera hipótesis se evalúa con una prueba de una cola, basada en la prueba de Diebold y Mariano ajustada para muestras pequeñas. La idea de esta prueba, presentada inicialmente en Diebold y Mariano (1995), es comprobar si los errores de pronóstico son suficientemente parecidos. Para evaluar los pronósticos del modelo i (\hat{y}_t^i) contra los del modelo j (\hat{y}_t^j), se toman los errores de pronóstico

$$\varepsilon_t^i = y_t - \hat{y}_t^i \qquad \varepsilon_t^j = y_t - \hat{y}_t^j \qquad (25)$$

para evaluar la siguiente proposición sobre sus cuadrados:

$$(\varepsilon_t^i)^2 - (\varepsilon_t^j)^2 = 0 \qquad (26)$$

La prueba original es de dos colas, por lo que, de rechazarse la hipótesis nula, no permite discernir cuál de los dos pronósticos muestra mayores errores. Así, se ajusta la prueba para que sea de una cola, donde la hipótesis nula es que el modelo i muestra un mejor desempeño que el modelo j , determinado según el cuadrado de sus errores de pronóstico:

$$(\varepsilon_t^i)^2 - (\varepsilon_t^j)^2 < 0 \qquad (27)$$

Adicionalmente, se aplica la corrección de muestra pequeña de Harvey et al. (1997) para obtener resultados más robustos. Como las comparaciones se hacen entre pares de modelos, el indicador para resumir el desempeño de cada uno es la cantidad de modelos para los que muestra un error de pronóstico menor.

La segunda hipótesis se evalúa con la prueba de inclusión de pronósticos de Chong y Hendry (1986). La idea es determinar si un modelo usa toda la información relevante que se refleja en otro. Esto puede indicar si algún modelo es redundante o si pueden haber ganancias en la combinación de los modelos. Para ello, se estima la siguiente regresión:

$$y_t = \beta_0 + \beta_i \hat{y}_t^i + \beta_j \hat{y}_t^j + \epsilon_t \quad (28)$$

Si toda la información del modelo j ya está incorporada en el modelo i , debe cumplirse que $\beta_0 = 0$, $\beta_i = 1$ y $\beta_j = 0$. De la misma manera, si toda la información del modelo i está incorporada en el modelo j , debe cumplirse que $\beta_0 = 0$, $\beta_i = 0$ y $\beta_j = 1$. Si alguno de estos dos casos no se cumple, se concluye que ambos modelos contienen información útil, y suficientemente distinta a la del otro modelo. En esta aplicación, los modelos son univariados, por lo que todos usan la misma información. Lo que una prueba de esta índole revelaría es si los distintos modelos están usando la información de manera suficientemente diferente. Al igual que la prueba anterior, esta se hace comparando pares de modelos, por lo que el indicador es la cantidad de veces que un modelo usa la información de manera suficientemente distinta al otro.

5.1.4. Estructura de autocorrelación

En este apartado, se explica la propiedad de autocorrelación de los errores de pronóstico que se evalúa. Esta se basa en el análisis de Diebold y Lopez (1996), en el que se presentan propiedades que deben cumplir los pronósticos óptimos.

La autocorrelación de un pronóstico tiene que ver con el grado de correlación que tienen sus errores a lo largo del tiempo. Para el caso de un horizonte de pronóstico de un periodo, un modelo debe tener errores que sean ruido blanco. En términos de la estructura de medias móviles, esto significa que es MA(0). La generalización al horizonte de pronóstico h implica que debe de ser, a lo sumo, MA($h - 1$). Para evaluarlo, se usa la prueba de Cumby y Hui-

zinga (1992). Esta supone que los errores siguen un proceso $MA(q)$, donde q está entre 0 y $h - 1$. La hipótesis alternativa es que hay algún rezago mayor a $h - 1$ con autocorrelación. El indicador para esta propiedad se resume por medio de una variable dicotómica, que indica el no rechazo de la hipótesis nula con un nivel de significancia del 5%, también se considera el valor p de la prueba, para así poder discernir entre los modelos que muestren mayor o menor evidencia a favor de la estructura de autocorrelación deseada.

5.2. Combinación de resultados

En la sección anterior, se presentaron nueve propiedades deseables para un modelo de pronóstico, que se resumen cuantitativamente a través de indicadores. Así, se sigue de cerca el trabajo de Fuentes y Rodríguez (2016). Un modelo que sea uniformemente superior al resto mostraría mejores indicadores en todas las propiedades. Encontrar tal modelo sería deseable, pero muy poco probable, en especial al tratarse de varios modelos. El enfoque que se toma, por lo tanto, es el de crear un puntaje ponderado, que requiere de dos pasos. Inicialmente, se deben asignar puntajes a cada propiedad, que van a depender del desempeño relativo. Posteriormente, se deben determinar los ponderadores, para asignar distintas importancias a las propiedades que se están evaluando. El cuadro 3 resume los criterios considerados en este puntaje, los cuales se describen a continuación.

A cada propiedad se le asigna un puntaje entre 0 y 1, que depende de qué tan buen o mal desempeño muestra el modelo en comparación con los otros. Si todos los modelos muestran igual desempeño, reciben un puntaje igual a 0,5. En la medida en que muestren mejor o peor desempeño relativo, estarán recibiendo puntajes mayores o menores a este punto medio. Se consideran dos maneras de asignar los puntajes, dependiendo de si sus indicadores se calculan comparando con los otros modelos o no.

Para describir estas dos maneras de asignación de puntajes, es preciso introducir más notación. Se denota cada modelo con m , que va de 1 hasta M

Cuadro 3: Cuadro de resumen de criterios considerados

Criterio de evaluación	Propiedad	Fórmula de puntaje	Ponderación (%)
Precisión	RECM	(33)	12,5
	Índice de Theil	(33)	12,5
	% de aciertos de dirección de cambios	(33)	5
Insesgamiento	Prueba de Mincer y Zarnowitz (1969)	(33)	25
	Valor p de la prueba	(33)	5
Capacidad superior	Prueba de Harvey et al. (1997)	(34)	20
	Prueba de Chong y Hendry (1986)	(34)	10
Autocorrelación	Prueba de Cumby y Huizinga (1992)	(33)	5
	Valor p de la prueba	(33)	5

y cada indicador con i , que va de 1 hasta I . También, se definen los siguientes estadísticos:

- El valor del indicador i para el modelo m :

$$x^i(m) \quad (29)$$

- El máximo de los valores de cada indicador:

$$\max^i = \text{máx}\{x^i(1), \dots, x^i(M)\} \quad (30)$$

- El mínimo de los valores de cada indicador:

$$\min^i = \text{mín}\{x^i(1), \dots, x^i(M)\} \quad (31)$$

- La diferencia entre el máximo y el mínimo de cada indicador:

$$d^i = \max^i - \min^i \quad (32)$$

Los indicadores que no dependen de la comparación directa entre modelos

(todos, a excepción de los de capacidad superior de pronóstico y de diferencias de variancias), reciben un puntaje de acuerdo a la siguiente fórmula:

$$p^i(m) = \begin{cases} 1/2 + (2x^i(m) - \max^i - \min^i) / 2d^i & \text{si } \max^i \neq \min^i \\ 1/2 & \text{si } \max^i = \min^i \end{cases} \quad (33)$$

Esta fórmula asigna 0 al valor más bajo y 1 al más alto, y distribuye los puntajes linealmente, siempre y cuando haya dispersión en los datos. Si todos los indicadores tienen un mismo valor, obtienen un puntaje igual a 0,5.

Los indicadores que dependen de la comparación directa entre modelos reciben un puntaje de acuerdo a una fórmula distinta:

$$p^i(m) = \frac{1}{2} + \frac{1}{2} \left(\frac{2x^i(m) - \max^i - \min^i}{M - 1} \right) \quad (34)$$

En vez de normalizar por la distancia entre los valores de los indicadores, se normaliza por la cantidad de modelos contra los cuáles se compara. Esto hace que, cuando hay dispersión en los datos, el modelo con el indicador más alto reciba el puntaje más alto, pero no necesariamente uno, y que el modelo con el indicador más bajo reciba el puntaje más bajo, pero no necesariamente cero. Si todos los indicadores tienen un mismo valor, reciben un puntaje de 0,5, igual que con el puntaje anterior.

Esta medida evita penalizar fuertemente algunos casos que se presentan en los datos. Por ejemplo, cuando todos los modelos tienen un mismo valor en su indicador a excepción de uno, el cual tiene un valor mayor. Según la fórmula (33), un modelo obtendría un puntaje de uno y el resto de cero, a pesar de solo ser marginalmente mejor. Según la fórmula (34), la diferencia en los puntajes sería menor, reflejando que la diferencia entre el valor de los indicadores es pequeña.

La determinación de las ponderaciones para cada propiedad permite incorporar diferencias en la importancia relativa que cada una de estas pueda tener en la elección final. Esto permite que, por ejemplo, el insesgamiento ten-

ga un mayor impacto en la elección, o que las medidas de precisión puedan valorarse sin sobrecargar esta dimensión al contarse con varios indicadores. Los ponderadores asignados, que suman un 100 %, son:

- Raíz del error cuadrático medio: 12,5 %.
- Índice de desigualdad de Theil: 12,5 %.
- Porcentaje de aciertos del signo del cambio: 5 %.
- Insesgamiento al 5 % de significancia: 25 %.
- Valor p de la prueba de significancia de insesgamiento: 5 %.
- Capacidad superior de pronóstico HLN: 20 %.
- Capacidad superior de pronóstico de inclusión: 10 %.
- Estructura de autocorrelación deseada: 5 %.
- Valor p de la prueba de estructura de autocorrelación deseada: 5 %.

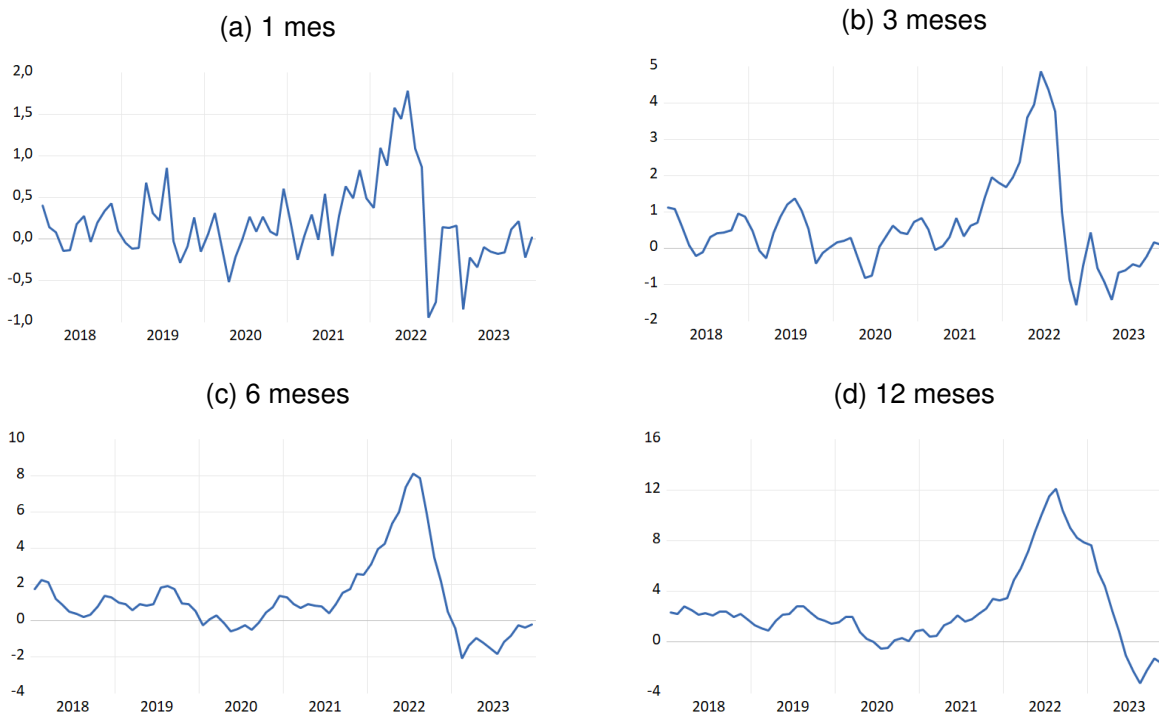
De esta manera, los modelos que muestren un puntaje ponderado más cercano a uno serán preferidos, pues reflejarán un mejor desempeño en los indicadores seleccionados.

5.3. Robustez al periodo de evaluación

El periodo de evaluación de los pronósticos está compuesto por las 24 últimas observaciones de la muestra. Es decir, comienza en enero de 2022 y termina en diciembre de 2023, como se muestra en el cuadro 2. Los años 2022 y 2023 fueron atípicos en términos de la variación del IPC, gracias a los choques de oferta causados por las tensiones en las cadenas de valor global y los conflictos geopolíticos a raíz de la invasión rusa en Ucrania, junto con la reacción de política monetaria del BCCR ante estos eventos. Durante ese periodo, la variación interanual del IPC llegó a un máximo de 12,9 % y

Gráfico 2: Variaciones porcentuales del IPC durante periodos de evaluación

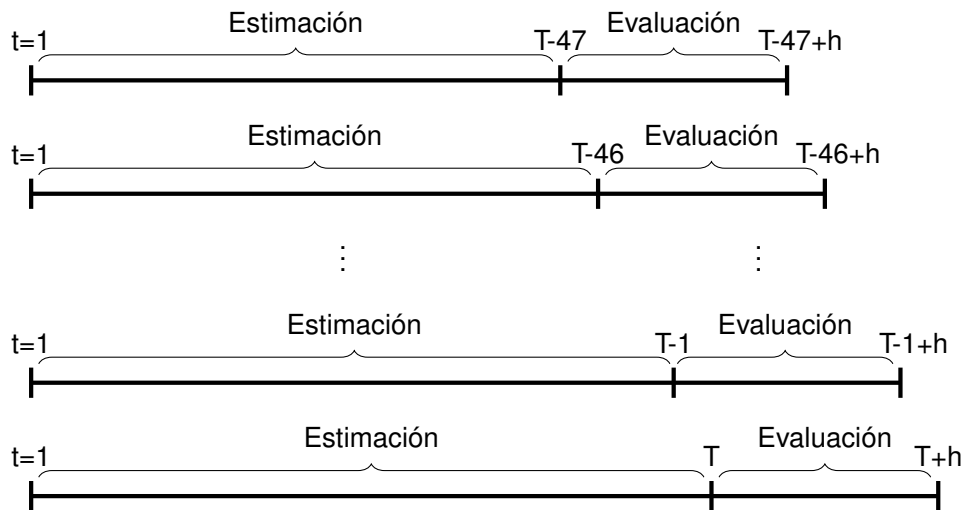
Enero 2018 - diciembre 2023



un mínimo de -3,3%. Este máximo ha sido el más alto después del episodio inflacionario causado por la crisis financiera internacional de 2008 y 2009, y el mínimo el más bajo desde la crisis de los años 80.

Cabe preguntarse, entonces, si esta combinación de eventos tuvo la capacidad de alterar significativamente la elección de los modelos. Por esto, se examina la robustez de la selección del modelo ante cambios en el periodo de evaluación, acortando el periodo de estimación en hasta 4 años. Con esto, se pretende corroborar si al dejar por fuera este episodio, se mantiene el orden del desempeño de los modelos. Es deseable tener información sobre los cambios de orden: si el orden se altera poco, es posible hacer revisiones de los modelos esporádicamente, pero si el orden es sensible al periodo de evaluación, es prudente hacer las revisiones con mayor frecuencia. El gráfico 2 muestra las variaciones intermensuales e interanuales del IPC para la totalidad del periodo que se considera para evaluación de los pronósticos. Como

Gráfico 3: Periodos de estimación para evaluación de robustez



se puede observar, durante 2018 y 2019, la variación interanual del IPC fue estable, que también se puede apreciar en la variación intermensual. En los años posteriores, las variaciones fueron más volátiles, lo cual podría afectar la elección de los modelos de pronóstico.

El gráfico 4 muestra cómo se cambia el periodo de estimación para evaluar esta robustez. El periodo más largo de estimación inicial abarca entre julio de 2007 y diciembre de 2021. Se denotan estos meses con $t = 1$ y $t = T$. Se toman $h = 24$ meses como periodo de evaluación, que abarca el periodo entre $t = T + 1$ hasta $t = T + h$. Esto está representado en la última línea del gráfico. Acortar el periodo de estimación en 4 años implica estimar con datos entre $t = 1$ y $t = T - 47$ para evaluar con datos entre $t = T - 47 + 1$ y $t = T - 47 + h$, donde se mantiene la cantidad de periodos de evaluación. Esto está representado en la primera línea del gráfico. En total, se consideran 48 periodos de estimación con h periodos de evaluación, lo que permite saber si el desempeño de estos modelos varía conforme se va agregando información más reciente.

6. Resultados

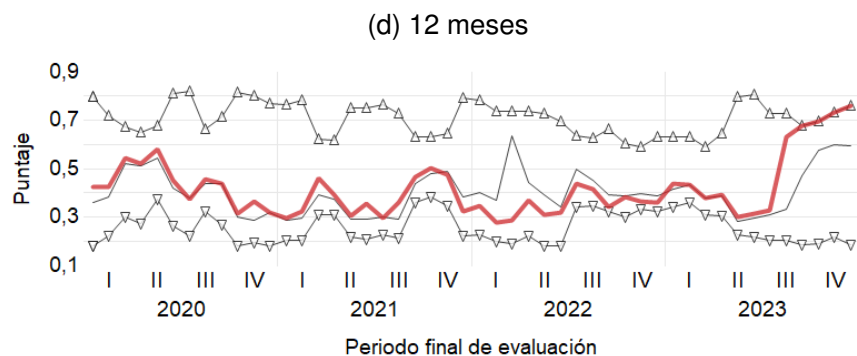
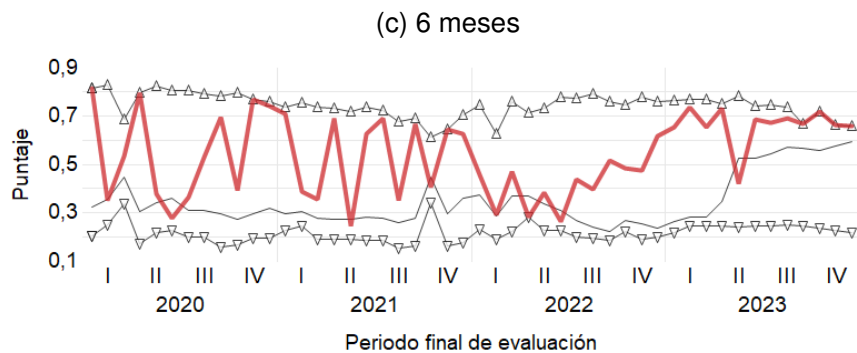
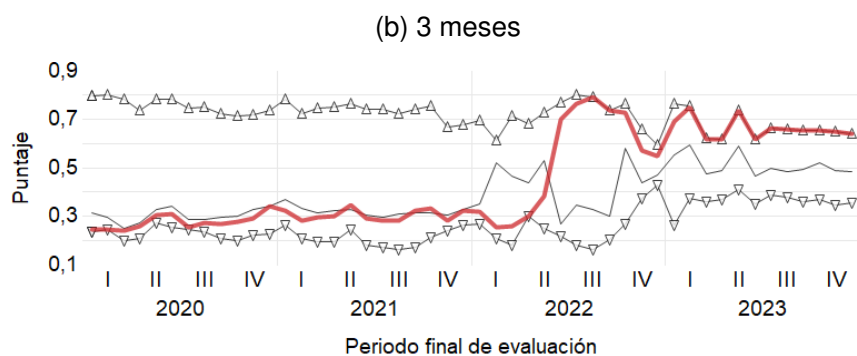
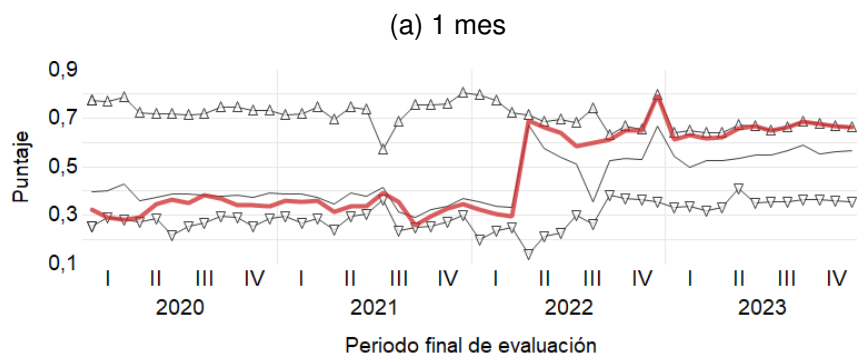
Los modelos con mejor desempeño, evaluados con información entre enero de 2021 y diciembre de 2023 son:

- A un mes plazo: modelo ARMA(2,4)-SMA(2), con un puntaje de 0,637.
- A tres meses plazo: ARMA(2,2)-SMA(2), con un puntaje de 0,644.
- A seis meses plazo: modelo estructural de componentes no observables con términos estacionales trigonométricos, con un puntaje de 0,653.
- A doce meses plazo: ARMA(2,4)-SMA(2)-EGARCH(1,1,1), con un puntaje de 0,746.

Los modelos a un mes plazo mostraron un desempeño similar y en promedio, obtuvieron un puntaje de 0,546. Los puntajes finales mostraron una desviación estándar del 0,070. Esto quiere decir que, en buena medida, los modelos arrojaron pronósticos similares entre sí. El modelo con mejor puntaje, el ARMA(2,4)-SMA(2), mostró un buen desempeño en todas las dimensiones, típicamente quedando entre los primeros diez lugares. Como es de esperarse, no tuvo un desempeño uniformemente mejor, y sus peores desempeños se manifestaron en los valores p de las pruebas de insesgamiento y de autocorrelación. A pesar de tener evidencia estadística del cumplimiento de las propiedades deseadas, no fue tan fuerte como la de otros modelos. Por esto, obtuvo los lugares 25 y 9.

Los modelos a tres meses plazo mostraron una dispersión similar en sus puntajes, con una desviación estándar de 0,065 y el menor promedio en los puntajes, de 0,508. El modelo ARMA(2,2)-SMA(2) mostró el mejor desempeño en cuatro de las propiedades evaluadas. Las propiedades donde mostró el peor desempeño fueron aquellas de capacidad superior de pronóstico, donde no mostró desempeño estadísticamente superior a algún otro modelo (lugar 14) y los valores p de las pruebas de insesgamiento y de autocorrelación (lugares 13 y 11).

Gráfico 4: Percentiles de puntajes y puntajes de modelos seleccionados



Percentil 0 Percentil 50 Percentil 100 Modelo seleccionado

Los modelos a seis meses plazo también mostraron una dispersión similar en sus puntajes, con una desviación estándar de 0,074 y el mayor promedio, de 0,572. El modelo estructural de componentes no observables con términos estacionales trigonométricos mostró buen desempeño en varias categorías. Su peor desempeño, similarmente a los otros horizontes, se dio en los valores p de las pruebas de insesgamiento y de autocorrelación (ambos en el lugar 25).

Los modelos a doce meses plazo mostraron un puntaje promedio de 0,513 y una desviación estándar considerablemente mayor, de 0,163. El modelo con el mayor puntaje, ARMA(2,4)-SMA(2)-EGARCH(1,1,1), mostró un desempeño considerablemente mejor que el resto de los modelos, por lo que obtuvo un puntaje tan alto. Solo en dos propiedades no obtuvo el primer lugar: la prueba de inclusión de pronósticos (lugar 29) y el valor p de la prueba de autocorrelación (lugar 30).

Las pruebas de robustez a cambiar el periodo de evaluación mostraron que estos resultados no son estables a lo largo del tiempo. Por lo tanto, es recomendable llevar a cabo la actualización de los modelos periódicamente. El gráfico 4 ilustra este resultado por medio de cuatro series de tiempo: el puntaje mínimo (línea discontinua inferior), la mediana de los puntajes (línea discontinua intermedia), el puntaje máximo (línea discontinua superior) y el puntaje que obtuvo el modelo seleccionado usando ese periodo distinto de evaluación (línea roja continua). La última observación coincide con el periodo de evaluación que inicia en enero de 2021 y finaliza en diciembre de 2023, por lo que para cada horizonte, la línea roja coincide con la línea del puntaje máximo.

El modelo seleccionado a un horizonte de un mes muestra un buen desempeño a partir de 2022, pero previo a eso, estuvo mostrando desempeños típicamente por debajo de la mediana. El modelo a tres meses se comporta de manera similar. El modelo seleccionado a 6 meses muestra un comportamiento más errático, alcanzando los puntajes más altos de manera estable

hasta el final del periodo de evaluación. Finalmente, el modelo seleccionado a 12 meses mostró una mejoría considerable, pero hasta los últimos periodos de evaluación.

7. Conclusiones

El Banco Central de Costa Rica, al mantener un esquema monetario de metas de inflación, hace un seguimiento constante de sus herramientas de análisis y proyección macroeconómica. En esta nota técnica, se detalla el proceso que actualmente está utilizando para definir qué modelos univariados usa para el pronóstico de la inflación en el corto plazo.

Este proceso implica la medición de varias propiedades deseables en modelos de pronóstico, que se combinan para obtener un puntaje final. La revisión de estos puntajes a lo largo del tiempo, como prueba de robustez, indica que los modelos con mejor desempeño varían en el tiempo. Esto implica que la revisión frecuente de los puntajes es deseable y constituye una de las recomendaciones más importantes de esta nota.

Referencias

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Chan, J. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2), 162–172.
- Chong, Y. Y. y Hendry, D. (1986). Econometric evaluation of linear macroeconomic models. *The Review of Economic Studies*, 53(4), 671–690.
- Cumby, R. y Huizinga, J. (1992). Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. *Econometrica*, 60(1), 185–95.

- Diebold, F. y Lopez, J. (1996). Forecast evaluation and combination. NBER Technical Working Papers 0192, National Bureau of Economic Research, Inc.
- Diebold, F. y Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–63.
- Dupor, B. (2023). Examining Long and Variable Lags in Monetary Policy. *The Regional Economist*.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Friedman, M. (1959). *A Program for Monetary Stability*. Fordham University, The Millar Lectures. Fordham University Press.
- Fuentes, M. y Rodríguez, A. (2016). Pronósticos univariados de inflación para costa rica con volatilidad estocástica y efectos garch. *Documentos de trabajo BCCR 04-2016*.
- Harvey, D., Leybourne, S., y Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), 281–291.
- Mincer, J. y Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (pp. 3–46). National Bureau of Economic Research, Inc.
- Pesaran, M. y Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, 137(1), 134–161.
- Taylor, S. J. (1986). *Modelling Financial Time Series*. John Wiley.

Torres, C. (2012). Costa rica: Determinación de cambios estructurales en el nivel de la tasa de inflación: periodo 1997-2011. *Documentos de trabajo BCCR 02-2012*.

Wicksell, K. (2008). *Interest and Prices*. Josephs Press.

A. Pruebas econométricas y resultados finales detallados

Cuadro 4: Prueba de quiebres estructurales, variación intermensual del IPC

Pruebas Bai-Perron de 1 a M quiebres determinados globalmente				
Muestra:	2007M07 - 2023M12			
Observaciones :	198			
Quiebres secuenciales según estadístico F:	5			
Mayores quiebres significativos según estadístico F:	5			
Quiebres según estadístico UDMax:	1			
Quiebres según estadístico WDMax:	1			
	Estad. F	Estad. F	Valor	
Quiebres	Estad. F	Reajustado	Ponderado	crítico
1	25,726	25,726	25,726	8,580
2	16,175	16,175	19,221	7,220
3	12,517	12,517	18,020	5,960
4	9,970	9,970	17,143	4,990
5	7,472	7,472	16,396	3,910
Prueba	Estadístico	Valor crítico		
UDMax	25,726	8,880		
WDMax	25,726	9,910		
Fechas de quiebre estimadas:				
1:	2010M03			
2:	2010M03, 2020M09			
3:	2010M03, 2014M09, 2021M08			
4:	2010M03, 2014M09, 2017M04, 2021M08			
5:	2010M02, 2012M07, 2014M12, 2017M05, 2021M08			

Cuadro 5: Prueba de quiebres estructurales, variación intermensual del IPC

Pruebas Bai-Perron de L+1 vs. L quiebres determinados secuencialmente			
Muestra:	2007M07 - 2023M12		
Observaciones:	198		
Quiebres según estadístico F, secuencialmente:	1		
	Estad. F	Reajustado	Valor crítico
Prueba de quiebres	Estad. F	Reajustado	Valor crítico
0 vs. 1	25,726	25,726	8,580
1 vs. 2	5,971	5,971	10,130
Fechas de quiebre estimadas:			
	Secuencial	Repartición	
1:	2010M03	2010M03	

Cuadro 6: Pruebas de raíz unitaria y quiebres estructurales, variación intermensual del IPC

Especificación	Quiebre	Tipo de quiebre	Criterio de selección	Fecha de quiebre	Estadístico	Valor <i>p</i>
Intercepto	Intercepto	Inovacional	Min DF	2008M11	-9,233	<0,01
Intercepto	Intercepto	Inovacional	Min t	2010M02	-8,524	<0,01
Intercepto	Intercepto	Inovacional	Max t	2021M07	-7,941	<0,01
Intercepto	Intercepto	Inovacional	Max abs-t	2010M02	-8,524	<0,01
Intercepto	Intercepto	Aditivo	Min DF	2008M11	-9,280	<0,01
Intercepto	Intercepto	Aditivo	Min t	2009M03	-8,703	<0,01
Intercepto	Intercepto	Aditivo	Max t	2021M09	-8,007	<0,01
Intercepto	Intercepto	Aditivo	Max abs-t	2009M03	-8,703	<0,01
Intercepto y tendencia	Tendencia	Inovacional	Min DF	2009M06	-8,797	<0,01
Intercepto y tendencia	Tendencia	Inovacional	Min t	2009M06	-8,797	<0,01
Intercepto y tendencia	Tendencia	Inovacional	Max t	2009M06	-8,797	<0,01
Intercepto y tendencia	Tendencia	Inovacional	Max abs-t	2009M06	-8,797	<0,01
Intercepto y tendencia	Tendencia	Aditivo	Min DF	2009M08	-8,855	<0,01
Intercepto y tendencia	Tendencia	Aditivo	Min t	2009M08	-8,855	<0,01
Intercepto y tendencia	Tendencia	Aditivo	Max t	2009M08	-8,855	<0,01
Intercepto y tendencia	Tendencia	Aditivo	Max abs-t	2009M08	-8,855	<0,01
Intercepto y tendencia	Intercepto	Inovacional	Min DF	2008M07	-9,260	<0,01
Intercepto y tendencia	Intercepto	Inovacional	Min t	2009M08	-8,617	<0,01
Intercepto y tendencia	Intercepto	Inovacional	Max t	2020M06	-8,677	<0,01
Intercepto y tendencia	Intercepto	Inovacional	Max abs-t	2020M06	-8,677	<0,01
Intercepto y tendencia	Intercepto	Aditivo	Min DF	2008M07	-9,330	<0,01
Intercepto y tendencia	Intercepto	Aditivo	Min t	2009M03	-8,775	<0,01
Intercepto y tendencia	Intercepto	Aditivo	Max t	2020M07	-8,729	<0,01
Intercepto y tendencia	Intercepto	Aditivo	Max abs-t	2009M03	-8,775	<0,01
Intercepto y tendencia	Intercepto y tendencia	Inovacional	Min DF	2008M08	-9,458	<0,01
Intercepto y tendencia	Intercepto y tendencia	Inovacional	Min t	2008M08	-9,458	<0,01
Intercepto y tendencia	Intercepto y tendencia	Inovacional	Max t	2008M08	-9,458	<0,01
Intercepto y tendencia	Intercepto y tendencia	Inovacional	Max abs-t	2008M08	-9,458	<0,01
Intercepto y tendencia	Intercepto y tendencia	Aditivo	Min DF	2008M08	-9,595	<0,01
Intercepto y tendencia	Intercepto y tendencia	Aditivo	Min t	2008M08	-9,595	<0,01
Intercepto y tendencia	Intercepto y tendencia	Aditivo	Max t	2008M08	-9,595	<0,01
Intercepto y tendencia	Intercepto y tendencia	Aditivo	Max abs-t	2008M08	-9,595	<0,01

Cuadro 7: Especificación y puntaje final de modelos evaluados

Modelo	Especificación	Puntaje final			
		1m	3m	6m	12m
1	Componentes no observables sin estacionalidad	0,4977	0,5983	0,4765	0,5236
2	Componentes no observables, est dicotómica	0,5376	0,5964	0,4651	0,5126
3	Componentes no observables, est. trigonométrica	0,3382	0,5491	0,6526	0,4300
4	Volatilidad estocástica, autorregresivo	0,5358	0,3896	0,2129	0,2089
5	Volatilidad estocástica, componentes no observados	0,4026	0,3714	0,5625	0,5965
6	ARMA(4,6)	0,5751	0,5210	0,5512	0,2843
7	ARMA(3,1)	0,6095	0,5571	0,5907	0,3250
8	ARMA(3,1)	0,6095	0,5571	0,5907	0,3250
9	ARMA(6,8)-SMA(1)	0,5712	0,5619	0,5786	0,3101
10	ARMA(3,1)	0,6095	0,5571	0,5907	0,3250
11	ARMA(3,1)	0,6095	0,5571	0,5907	0,3250
12	ARMA(2,4)-SMA(2)	0,6372	0,6023	0,6279	0,3125
13	ARMA(2,4)-SMA(2)	0,6372	0,6023	0,6279	0,3125
14	ARMA(2,2)-SMA(2)	0,6214	0,6440	0,6032	0,2742
15	ARMA(4,6)-GARCH(1,1)	0,4939	0,4258	0,6162	0,7087
16	ARMA(3,1)-GARCH(1,1)	0,5356	0,4682	0,5916	0,6169
17	ARMA(3,1)-GARCH(1,1)	0,5356	0,4682	0,5916	0,6169
18	ARMA(6,8)-SAR(1)-GARCH(1,1)	0,4977	0,4868	0,5668	0,6261
19	ARMA(3,1)-GARCH(1,1)	0,5356	0,4682	0,5916	0,6169
20	ARMA(3,1)-GARCH(1,1)	0,5356	0,4682	0,5916	0,6169
21	ARMA(2,4)-SMA(2)-GARCH(1,1)	0,5986	0,4796	0,5948	0,6737
22	ARMA(2,4)-SMA(2)-GARCH(1,1)	0,5986	0,4796	0,5948	0,6737
23	ARMA(2,2)-SMA(2)-GARCH(1,1)	0,5664	0,4936	0,5982	0,6364
24	ARMA(4,6)-EGARCH(1,1,1)	0,4629	0,3946	0,5716	0,6001
25	ARMA(3,1)-EGARCH(1,1,1)	0,5113	0,4757	0,5747	0,6013
26	ARMA(3,1)-EGARCH(1,1,1)	0,5113	0,4757	0,5747	0,6013
27	ARMA(6,8)-SAR(1)-EGARCH(1,1,1)	0,4375	0,4935	0,5508	0,3546
28	ARMA(3,1)-EGARCH(1,1,1)	0,5113	0,4757	0,5747	0,6013
29	ARMA(3,1)-EGARCH(1,1,1)	0,5113	0,4757	0,5747	0,6013
30	ARMA(2,4)-SMA(2)-EGARCH(1,1,1)	0,6270	0,5392	0,6220	0,7460
31	ARMA(2,4)-SMA(2)-EGARCH(1,1,1)	0,6270	0,5392	0,6220	0,7460
32	ARMA(2,2)-SMA(2)-EGARCH(1,1,1)	0,5950	0,4882	0,5912	0,7006

Cuadro 8: Indicadores de propiedades evaluadas, enero 2021 - diciembre 2023

Modelo	RECM				Theil				Aciertos			
	Horizonte				Horizonte				Horizonte			
	1	3	6	12	1	3	6	12	1	3	6	12
1	0,60	1,44	2,37	3,46	0,47	0,41	0,38	0,41	0,26	0,52	0,61	0,67
2	0,60	1,42	2,41	3,59	0,46	0,40	0,38	0,42	0,35	0,57	0,56	0,67
3	0,69	1,74	2,95	4,32	0,53	0,48	0,46	0,48	0,13	0,52	0,50	0,58
4	0,60	2,02	4,38	10,10	0,44	0,51	0,58	0,73	0,35	0,33	0,33	0,33
5	0,58	1,85	3,38	4,54	0,49	0,60	0,63	0,61	0,26	0,33	0,22	0,33
6	0,60	1,82	3,69	6,58	0,44	0,50	0,59	0,67	0,39	0,48	0,22	0,25
7	0,59	1,70	3,35	5,33	0,44	0,49	0,56	0,61	0,43	0,43	0,39	0,25
8	0,59	1,70	3,35	5,33	0,44	0,49	0,56	0,61	0,43	0,43	0,39	0,25
9	0,61	1,73	3,44	6,07	0,43	0,48	0,56	0,65	0,30	0,52	0,44	0,33
10	0,59	1,70	3,35	5,33	0,44	0,49	0,56	0,61	0,43	0,43	0,39	0,25
11	0,59	1,70	3,35	5,33	0,44	0,49	0,56	0,61	0,43	0,43	0,39	0,25
12	0,58	1,52	3,23	6,51	0,40	0,42	0,50	0,64	0,43	0,52	0,61	0,42
13	0,58	1,52	3,23	6,51	0,40	0,42	0,50	0,64	0,43	0,52	0,61	0,42
14	0,58	1,48	3,19	6,97	0,40	0,39	0,47	0,63	0,43	0,52	0,44	0,33
15	0,63	1,68	3,08	3,73	0,51	0,56	0,60	0,53	0,35	0,57	0,56	0,58
16	0,62	1,75	3,34	4,53	0,49	0,55	0,62	0,61	0,48	0,52	0,50	0,42
17	0,62	1,75	3,34	4,53	0,49	0,55	0,62	0,61	0,48	0,52	0,50	0,42
18	0,66	1,86	3,35	4,53	0,50	0,55	0,59	0,61	0,43	0,52	0,39	0,50
19	0,62	1,75	3,34	4,53	0,49	0,55	0,62	0,61	0,48	0,52	0,50	0,42
20	0,62	1,75	3,34	4,53	0,49	0,55	0,62	0,61	0,48	0,52	0,50	0,42
21	0,58	1,62	3,10	3,89	0,45	0,52	0,60	0,58	0,39	0,43	0,50	0,42
22	0,58	1,62	3,10	3,89	0,45	0,52	0,60	0,58	0,39	0,43	0,50	0,42
23	0,59	1,63	3,09	4,16	0,46	0,51	0,58	0,60	0,39	0,48	0,44	0,33
24	0,64	1,99	3,52	4,58	0,51	0,63	0,63	0,61	0,43	0,43	0,44	0,33
25	0,62	1,78	3,43	4,72	0,49	0,55	0,61	0,61	0,43	0,57	0,33	0,33
26	0,62	1,78	3,43	4,72	0,49	0,55	0,61	0,61	0,43	0,57	0,33	0,33
27	0,67	1,86	3,41	4,86	0,50	0,54	0,58	0,61	0,35	0,48	0,33	0,42
28	0,62	1,78	3,43	4,72	0,49	0,55	0,61	0,61	0,43	0,57	0,33	0,33
29	0,62	1,78	3,43	4,72	0,49	0,55	0,61	0,61	0,43	0,57	0,33	0,33
30	0,56	1,56	2,80	2,74	0,43	0,49	0,54	0,40	0,48	0,57	0,61	0,67
31	0,56	1,56	2,80	2,74	0,43	0,49	0,54	0,40	0,48	0,57	0,61	0,67
32	0,57	1,62	3,03	3,14	0,45	0,52	0,58	0,49	0,48	0,48	0,50	0,42

Cuadro 9: Indicadores de propiedades evaluadas, enero 2021 - diciembre 2023 (continuación)

Modelo	Valor p insesgamiento				Pronóstico HLN				Pronóstico inclusión				Valor p autocorrelación			
	Horizonte				Horizonte				Horizonte				Horizonte			
	1	3	6	12	1	3	6	12	1	3	6	12	1	3	6	12
1	0,73	0,15	0,01	0,00	0	0	0	0	29	29	30	26	0,05	0,73	0,41	0,59
2	0,81	0,16	0,01	0,00	0	0	0	0	31	29	30	25	0,12	0,49	0,42	0,58
3	0,59	0,73	0,33	0,00	1	0	0	0	31	29	31	11	0,11	0,93	0,37	0,62
4	0,68	0,18	0,00	0,00	0	0	1	0	31	31	31	26	0,09	0,71	0,29	0,34
5	0,45	0,77	0,90	0,08	0	2	3	0	25	19	27	25	0,04	0,55	0,54	0,32
6	0,68	0,73	0,22	0,00	0	3	7	0	31	31	29	27	0,62	0,75	0,44	0,32
7	0,89	0,93	0,60	0,01	0	2	1	0	31	31	29	26	0,73	0,85	0,46	0,33
8	0,89	0,93	0,60	0,01	0	2	1	0	31	31	29	26	0,73	0,85	0,46	0,33
9	0,56	0,86	0,47	0,00	0	0	1	0	31	31	29	27	0,96	0,89	0,42	0,32
10	0,89	0,93	0,60	0,01	0	2	1	0	31	31	29	26	0,73	0,85	0,46	0,33
11	0,89	0,93	0,60	0,01	0	2	1	0	31	31	29	26	0,73	0,85	0,46	0,33
12	0,57	0,70	0,36	0,00	0	0	2	0	31	31	29	25	0,73	0,58	0,31	0,35
13	0,57	0,70	0,36	0,00	0	0	2	0	31	31	29	25	0,73	0,58	0,31	0,35
14	0,47	0,78	0,25	0,00	0	0	0	0	31	31	30	11	0,60	0,77	0,26	0,48
15	0,82	0,32	0,69	0,51	0	0	0	0	31	20	27	29	0,77	0,37	0,58	0,32
16	1,00	0,75	0,92	0,10	0	0	1	0	31	27	27	27	0,39	0,52	0,50	0,32
17	1,00	0,75	0,92	0,10	0	0	1	0	31	27	27	27	0,39	0,52	0,50	0,32
18	0,63	0,90	0,78	0,10	7	2	0	0	31	27	29	27	0,37	0,78	0,37	0,32
19	1,00	0,75	0,92	0,10	0	0	1	0	31	27	27	27	0,39	0,52	0,50	0,32
20	1,00	0,75	0,92	0,10	0	0	1	0	31	27	27	27	0,39	0,52	0,50	0,32
21	0,95	0,35	0,90	0,53	0	0	0	0	31	29	27	28	0,31	0,59	0,39	0,35
22	0,95	0,35	0,90	0,53	0	0	0	0	31	29	27	28	0,31	0,59	0,39	0,35
23	0,80	0,48	0,95	0,31	0	0	0	0	31	29	29	27	0,33	0,54	0,33	0,36
24	0,95	0,85	0,60	0,09	0	9	2	0	31	19	26	26	0,09	0,51	0,65	0,32
25	0,93	0,81	0,75	0,05	0	2	3	0	31	23	27	27	0,17	0,53	0,52	0,32
26	0,93	0,81	0,75	0,05	0	2	3	0	31	23	27	27	0,17	0,53	0,52	0,32
27	0,52	0,79	0,61	0,03	2	2	0	0	31	29	29	27	0,42	0,87	0,36	0,32
28	0,93	0,81	0,75	0,05	0	2	3	0	31	23	27	27	0,17	0,53	0,52	0,32
29	0,93	0,81	0,75	0,05	0	2	3	0	31	23	27	27	0,17	0,53	0,52	0,32
30	0,40	0,06	0,07	0,83	0	0	0	0	31	29	29	11	0,78	0,77	0,40	0,32
31	0,40	0,06	0,07	0,83	0	0	0	0	31	29	29	11	0,78	0,77	0,40	0,32
32	0,82	0,24	0,62	0,26	0	0	0	0	31	29	29	27	0,17	0,60	0,32	0,36