

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.



ALGORITMO DE COMBINACIÓN DE PRONÓSTICOS DE VARIABLES
MACROECONÓMICAS: APLICACIÓN AL IGAE

TESINA
QUE PARA OBTENER EL GRADO DE
MAESTRO EN ECONOMÍA

PRESENTA
FRANCISCO VINICIO ARREOLA GASPAR

DIRECTOR DE LA TESINA: DR. DANIEL VENTOSA SATAULARIA

LECTOR DE LA TESINA: M.E. EDWIN SALIM TAPIA ALVARADO

A mi familia, por su amor y apoyo incondicional. Este logro es tan suyo como mío.

Agradecimientos

Quiero agradecer: Al Dr. Daniel Ventosa y al M.E. Edwin Tapia, por su valioso apoyo y orientación a lo largo de la elaboración de este trabajo. Sus conocimientos y guía han sido invaluable para mí.

A mi madre Sara, mi tía Nora, y mis hermanos León y Ahmed, junto con Silvia. Gracias por su amor incondicional y por comprender mis ausencias durante mi estancia en el CIDE. Su paciencia y cariño fueron el pilar que me sostuvo en los momentos más difíciles. Isabel, gracias por tu amor, comprensión y apoyo constante. Tu presencia ha sido una fuente de fortaleza y motivación para mí.

A mis amigos y compañeros: Pablo, Josué, Samuel, Luis, Armando, Juan, Celeste, Wolf, Yorlyn y Paola. Gracias por acompañarme en las largas noches de estudio, por su apoyo, sus consejos y, sobre todo, por brindarme el ánimo necesario para seguir adelante en los días más pesados. La experiencia de la maestría ha sido maravillosa gracias a su presencia. Su amistad es uno de los tesoros más valiosos que me llevo de esta etapa de mi vida.

En especial, quiero agradecer a mis amigos Josué y Pablo. Josué, prácticamente comenzamos esto juntos y tu compañerismo y apoyo han sido esenciales. Eres un gran amigo y siempre estaré agradecido por todas las cosas que hemos superado juntos. Pablo, gracias por ese primer año y por siempre estar. La amistad y el apoyo de ambos me han enseñado el significado de la amistad. También tú Celeste, tu amistad en este último semestre fue un gran apoyo, gracias por los jueves de hamburguesitas.

Finalmente, a mis laboratoristas y amigos, Mitch y Paquito, gracias por su ayuda y amistad. Su apoyo fue fundamental en muchos momentos de este camino.

Y, por supuesto, un gran agradecimiento al CIDE. Gracias por otorgarme el título de Maestro en Economía y por ser el escenario de esta increíble aventura académica.

Resumen

Se presenta un algoritmo de Evaluación de Combinaciones de Pronósticos, mismo que se aplica para optimizar la precisión de las predicciones del Indicador Global de la Actividad Económica (IGAE). A través de dicha metodología, el algoritmo integra diversas técnicas de combinación de pronósticos y evalúa su eficacia comparativa en un marco pseudo fuera de muestra. Los resultados muestran que los métodos de combinación de pronósticos no solo mejora significativamente la precisión de los pronósticos en comparación con los modelos individuales, sino que también supera consistentemente las predicciones basadas en la media aritmética, especialmente en ventanas de evaluación y validación más cortas. Este enfoque sistemático y adaptativo proporciona una herramienta robusta para los analistas económicos, ofreciendo un medio para poder seleccionar la mejor estrategia de combinación de pronósticos, reduciendo los riesgos que se asumen al seleccionar un modelo individual o una estrategia de combinación que no sea la adecuada para el tipo de serie que se pretende pronosticar.

Índice general

1. Introducción	1
2. Motivación y Revisión de Literatura	4
3. Metodología	9
3.1. Técnicas de Combinación de Pronósticos	9
3.1.1. Promedio Simple	10
3.1.2. Bates y Granger (1969)	10
3.1.3. Método de Rango Inverso	11
3.1.4. Mínimos Cuadrados con Restricciones (CLS)	11
3.2. Métodos Adicionales	11
3.3. Modelos de Pronósticos	12
3.3.1. ARIMA	12
3.3.2. Suavizado Exponencial Holt-Winters	12
3.3.3. TBATS	13
3.3.4. Theta	14
3.3.5. Redes Neuronales Autorregresiva (NNAR)	15
3.4. Procedimiento de Evaluación de Pronósticos	16
3.5. Criterios de desempeño de pronósticos.	17
3.5.1. Mean Error (ME)	17
3.5.2. Mean Absolute Error (MAE)	18
3.5.3. Mean Percentage Error (MPE)	18
3.5.4. Mean Absolute Percentage Error (MAPE)	18
3.5.5. Root Mean Square Error (RMSE)	19
3.5.6. Comparación entre las Métricas	19

4. Datos	20
5. Resultados	22
5.1. Desempeño de los Modelos individuales en Estabilidad Secuencial	22
5.2. Resultado de la combinación de pronósticos	26
5.3. Evaluación de Métodos de Combinación	29
6. Conclusión	33
Referencias	35
Bibliografía	38
Apéndice	40
Apéndice A. Técnicas de combinación de pronósticos	40
A.1. Técnicas de combinación de pronósticos	40
A.1.1. Promedio Simple	40
A.1.2. Mediana	41
A.1.3. Media Recortada	41
A.1.4. Media Winsorizada	42
A.1.5. Bates y Granger (1969)	42
A.1.6. Newbold y Granger (1974)	42
A.1.7. Método de Rango Inverso	43
A.1.8. Regresión de Mínimos Cuadrados Ordinarios (OLS)	43
A.1.9. Regresión por Desviación Absoluta Mínima (LAD)	44
A.1.10. Regresión con Restricciones de Mínimos Cuadrados (CLS)	45
A.1.11. Método del Vector Propio Estándar	45
A.1.12. Método de Vectores Propios Ajustados por Sesgo	46
A.1.13. Método de Vectores Propios Recortados	46
A.1.14. Enfoque de Vectores Propios con Corrección de Sesgo y Recorte	47
Apéndice B. Desempeño de los Modelos individuales en Estabilidad Secuencial	48
Apéndice C. Combinación de pronosticos	51

Índice de figuras

3.1. Esquema del funcionamiento de una ventana móvil.	17
4.1. Indicador Global de la Actividad Económica, enero de 1993 a enero de 2024 Serie desestacionalizada (índice 2018-100)	20
5.1. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de En- trenamiento (150), Validación (12) y Evaluación (12), Desplazamiento (2). .	23
5.2. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de En- trenamiento (100), Validación (12) y Evaluación (12), Desplazamiento (2). .	24
5.3. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de En- trenamiento (50), Validación (12) y Evaluación (12), Desplazamiento (2). .	25
5.4. Comparativa de Desempeño de Modelos Individuales y Combinados (RM- SE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	27
5.5. Comparativa de Desempeño de Modelos Individuales y Combinados (RM- SE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)	28
5.6. Comparativa de Desempeño de Modelos Individuales y Combinados (RM- SE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)	29
5.7. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Supe- rar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)	30

5.8. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Superar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)	30
5.9. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Superar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	31
B.1. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (150), Validación (9) y Evaluación (9), Desplazamiento (2). . .	48
B.2. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (150), Validación (6) y Evaluación (6), Desplazamiento (2). . .	49
B.3. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (100), Validación (9) y Evaluación (9), Desplazamiento (2). . .	49
B.4. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (100), Validación (9) y Evaluación (9), Desplazamiento (2). . .	50
C.1. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	51
C.2. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)	52
C.3. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	52
C.4. Comparativa de Desempeño de Modelos Individuales y Combinados (MAPE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)	53

C.5. Comparativa de Desempeño de Modelos Individuales y Combinados (MA-PE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)	53
C.6. Comparativa de Desempeño de Modelos Individuales y Combinados (MA-PE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	54
C.7. Comparativa de Desempeño de Modelos Individuales y Combinados (MAE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)	54
C.8. Comparativa de Desempeño de Modelos Individuales y Combinados (MPE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)	55
D.1. Portada de la Aplicación Shiny	59
D.2. Pronóstico de la Inflación en la Aplicación Shiny	59

Índice de cuadros

C.1. MAE de diferentes métodos de combinación de pronósticos en iteraciones sucesivas ventana de validación (9) y evaluación (9)	56
C.2. MAE de diferentes métodos de combinación de pronósticos en iteraciones sucesivas ventana de validación (12) y evaluación (12)	57

Capítulo 1

Introducción

En la actualidad, los pronósticos de variables macroeconómicas cobran gran relevancia tanto en el sector público como en el privado, y parte de este fenómeno se asocia con la creciente disponibilidad y fácil acceso a la información. Con grandes volúmenes de datos disponibles, una gran cantidad de trabajos de investigación se han centrado principalmente en encontrar el modelo que mejor se ajuste a la variable de interés, lo que significa que, en el mejor de los casos, solo se puede tener una aproximación razonable de la realidad. Estos métodos requieren un profundo conocimiento de la estructura subyacente de las variables, la cual rara vez es clara. Además, se requiere realizar numerosas suposiciones, lo que hace crucial contar con un modelo robusto.

Por tanto, el reto radica en cómo gestionar de manera correcta la información de la que se dispone. Con los avances econométricos de las últimas décadas, así como los avances en capacidad de procesamiento de los datos, han surgido muchas formas de poder generar un pronóstico para una variable. No obstante, una serie de decisiones relacionadas con la modelización, como la selección y transformación de los datos, el periodo muestral de estimación, la técnica de estimación empleada, la especificación dinámica, el marco teórico aplicado, entre otros factores, benefician a uno u otro modelo, y la incorrecta inclusión o exclusión de algunos regresores en el proceso de selección, generará una batería de pronósticos los cuales serán distintos del proceso generador de datos de la variable económica, el cual siempre será desconocido para el pronosticador (Timmermann, 2006).

En consecuencia, el pronosticador deberá seleccionar el modelo que más información pueda integrar dentro de esta variedad de pronósticos, la mayoría de las veces optando por aquel que resulta ser el más competitivo, aun cuando este no equilibre de forma óptima los

errores de predicción y sus costos relativos sean altos al concentrar todo el riesgo en una sola opción. Sin embargo, desde el trabajo seminal de Bates y Granger (1969), ha surgido una alternativa para diversificar este riesgo: las combinaciones de pronósticos, mismas que han probado teórica y empíricamente que suelen producir pronósticos más precisos debido a que incorporan toda la información relevante y disponible al momento de realizar la predicción, consiguiendo pronósticos eficientes. Bates y Granger (1969), argumentan que, dado que el verdadero proceso generador de datos suele ser desconocido y el mejor modelo puede variar con el tiempo, combinar modelos permite integrar información relevante de cada uno en distintos momentos. Además, Castle et al. (2019), menciona que las combinaciones ofrecen diversificación, reduciendo el error cuadrático medio en comparación con cualquier modelo individual, al igual que una cartera de inversiones puede reducir el riesgo global en relación con una única acción.

La práctica de combinar pronósticos ha sido ampliamente adoptada, desde el sector privado hasta los bancos centrales. Ejemplos incluyen su uso para pronosticar la volatilidad del tipo de cambio (Benavides & Capistran, 2009) y el tipo de cambio en general (Bilson, 1983), así como para prever la inflación (Dinartes, 2009; Engle et al., 1984; Kapetanios et al., 2008). También se ha aplicado en la estimación del PIB (Reid, 1968), la demanda de dinero (Figlewski & Urich, 1983) y el crecimiento económico (Rapach & Strauss, 2008).

A pesar de las claras ventajas y amplias aplicaciones de la combinación de pronósticos, no existe un consenso sobre el mejor método para combinar pronósticos, es decir, cuál es la forma óptima de hacerlo. Esta interrogante ha dado lugar a diversos estudios que presentan una serie de propuestas sobre cómo deberían combinarse los pronósticos individuales. Estos métodos de combinación van desde los más populares como la media, la mediana y la media recortada, hasta métodos más complejos como la regresión por mínimos cuadrados ordinarios, técnicas bayesianas y los métodos basados en vectores propios. En la práctica, es difícil determinar qué estrategia de combinación elegir, ya que existen resultados mixtos. La evidencia empírica y numerosas simulaciones demuestran consistentemente que el promedio simple con pesos iguales a menudo supera a los esquemas de ponderación más complejos. Este resultado, inesperado para muchos, ha sido un tema central en la literatura, conocido como “el Puzzle de la combinación de pronósticos”, pero también hay literatura que apoya el hecho de que ciertas técnicas de combinación son mejores para cierto tipo de series de tiempo.

Por tanto, el objetivo de este trabajo es contribuir a la literatura sobre la combinación de pronósticos con el desarrollo de un algoritmo de Evaluación de Combinaciones de Pronósticos, el cual consiste en un sistema que toma a los pronósticos individuales, genera las ponderaciones para los pronósticos combinados a partir de distintas técnicas de combinación y los evalúa pseudo fuera de muestra, indicando si existe una mejor técnica de combinación que la media aritmética.

Este trabajo se basa en las siguientes hipótesis: Primero, las combinaciones de pronósticos basadas en ponderaciones óptimas, mínimos cuadrados ordinarios, y métodos basados en vectores propios, pueden superar a la media aritmética. Segundo, la técnica de combinación óptima depende del tipo de serie temporal y su comportamiento. Tercero, el uso de un algoritmo de evaluación de combinaciones mejora la precisión de los pronósticos combinados.

Para alcanzar los objetivos planteados, se pretende responder las siguientes preguntas de investigación: Primero, ¿Las técnicas de combinación sugieren ponderaciones iguales independientemente de la información disponible? Segundo, ¿cómo se comparan las técnicas de combinación en términos de precisión de pronóstico fuera de muestra entre ellas y entre los pronósticos individuales? Tercero, ¿existen métodos óptimos de combinación que superen a la media aritmética en diferentes horizontes de pronóstico? Y cuarto, ¿un algoritmo de evaluación de combinaciones puede identificar la técnica óptima para una serie de tiempo específica como el IGAE?

Este trabajo está organizado de la siguiente manera. El **Capítulo 1** presenta el problema, el objetivo, las hipótesis y las preguntas de investigación. El **Capítulo 2** proporciona una revisión de la literatura relevante y un marco conceptual sobre pronósticos macroeconómicos y técnicas de combinación. El **Capítulo 3** describe la metodología utilizada para desarrollar el algoritmo de evaluación y las técnicas de combinación de pronósticos. En el **Capítulo 4** se presentan los resultados de las pruebas empíricas y comparaciones entre diferentes técnicas de combinación. Finalmente, el **Capítulo 5** ofrece conclusiones, limitaciones del estudio y recomendaciones para futuras investigaciones.

Capítulo 2

Motivación y Revisión de Literatura

Los pronósticos han adquirido una gran importancia en diversos ámbitos de la vida, principalmente en sectores como la economía y las finanzas, donde ejercen una fuerza significativa en las decisiones de política. El origen de los pronósticos, así como su importancia, se remonta a los orígenes de la escritura, desde los primeros intentos humanos por predecir las mejores temporadas de caza por grupos de cazadores-recolectores, hasta el rastreo del cielo y sus constelaciones por parte de los babilonios (Castle et al., 2019). Por tanto, a lo largo de la historia, los pronósticos han permeado todos los aspectos de la vida, facilitando la toma de decisiones al proveer información crucial sobre eventos futuros de interés, desde aspectos muy simples como la duración aproximada de un viaje o el estado del tiempo mañana hasta aspectos más complejos como lo puede ser el tipo de cambio o el estado de la economía en un futuro. Todas estas decisiones están intrínsecamente ligadas al ámbito económico, donde los pronósticos son indispensables para la formulación de estrategias y políticas.

Los pronósticos han evolucionado y se han convertido en herramientas indispensables para los tomadores de decisiones. En este sentido, se ha vuelto fundamental que la elaboración y evaluación de cualquier pronóstico se realice dentro de un marco teórico orientado a la toma de decisiones, destacando que su valor radica exclusivamente en su capacidad para informar decisiones económicas. A pesar de los grandes desarrollos en metodologías estadísticas para la estimación de pronósticos, software especializado, mayor nivel computacional y la mejora de las fuentes de información, los pronósticos económicos se ven continuamente desafiados por los grandes choques económicos ante los cuales reaccionan con fallos, algo que no es propio de la economía, sino que es un fenómeno presente en varias disciplinas (Castle et al., 2019).

Por tanto, el éxito de los métodos de pronóstico dependerá de equilibrar entre incluir regularidades y posibles irregularidades. Estos pueden seguir enfoques muy distintos, desde adivinar o usar reglas prácticas, modelos simples y enfoques formales, cada uno asociado a suposiciones y niveles de riesgo particulares. Además, cualquier pronóstico se enfrenta a un sinfín de retos, entre los que se encuentran las situaciones con datos incompletos, cambios significativos en las relaciones económicas, estimaciones preliminares (muchas series económicas están sujetas a continuas revisiones conforme se acumula más información), errores de medición y problemas de bases cambiantes (Castle et al., 2009). De igual manera, para considerar que un modelo de pronóstico es útil y preciso, hay que asegurar un conjunto de condiciones: aceptar la incertidumbre de los resultados futuros (de no haber incertidumbre, no es necesario un pronóstico), medir correctamente esta incertidumbre, evitar asumir linealidades donde no existen, seleccionar el método estadístico apropiado y permitir que el modelo se pueda ajustar ante cambios en los datos (Castle et al., 2019).

Un pronóstico eficiente es aquel que incorpora toda la información relevante y disponible al momento de realizar la predicción, es decir, no posee un error sistemático (sesgo de error) y no puede ser mejorado al incorporar información del pasado, específicamente errores de autocorrelación. Logrando la mayor precisión posible dada la información disponible, dicho pronóstico no busca superar a otros, sino ser lo más preciso dentro de sus propios términos y limitaciones sin comparar con otros métodos de pronóstico en un entorno competitivo (Nordhaus, 1987). La efectividad del pronóstico será medida por su capacidad para minimizar los costos asociados a los errores de pronósticos, los cuales no solo estarán condicionados a la precisión del modelo, sino también a las preferencias del pronosticador y las decisiones económicas específicas que interactúan con el pronóstico (Elliott & Timmermann, 2008). Por tanto, la falta de teoría específica lleva normalmente a los pronosticadores a basarse en adaptaciones empíricas y a flexibilizar la forma funcional de los modelos de pronóstico para intentar captar la realidad económica, la cual se encuentra en constante cambio.

Por su parte, un pronóstico competitivo, aunque se aproxime a la minimización de pérdidas, puede no equilibrar de forma óptima los errores de predicción y sus costos relativos. La gran diversidad de modelos económicos de pronóstico resultante de la libertad en la construcción de pronósticos plantea la tarea de comparar distintos modelos (Castle et al., 2019). La literatura sobre la comparación de pronósticos muestra una amplia variedad de modelos intentando predecir la misma secuencia de resultados, pero extraer reglas generales de estos

estudios resulta en un desafío debido a la dependencia del tipo de variable, la frecuencia de los datos y el periodo de muestra considerado (Makridakis & Hibon, 2000). Trabajos académicos de este tipo o las competiciones de pronóstico que tienen como objetivo clasificar procedimientos de pronósticos a través de la evaluación, haciendo uso de criterios específicos como la precisión, la capacidad de generalización en conjuntos de evaluación, la interpretabilidad e incluso la eficiencia computacional ha resultado especialmente valiosos para saber qué modelos son más competitivos que otros.

Por lo anterior, los responsables de la toma de decisiones a menudo tienen acceso a múltiples pronósticos y pueden emplear dos distintas estrategias: buscar el mejor modelo individual es decir aquel que a resultado ser el más competitivo o intentar combinar pronósticos de varios modelos. De manera general combinar pronósticos de múltiples modelos suele superar el rendimiento individual el rendimiento de pronósticos de un solo modelo, dicho hallazgo es respaldado por múltiples estudios que evalúan la efectividad de la combinación de pronósticos frente a una amplia batería de pronósticos individuales (Wang et al., 2023). En este sentido la diversificación contra la incertidumbre del modelo es un argumento que se ha vuelto clave a favor de la combinación de pronósticos, ya que las combinaciones de pronósticos se consideran aproximaciones simplificadas de una realidad compleja y en constante evolución que consigue capturar distintas partes de un todo y condensarlas en un solo número, liberando en gran parte al pronóstico resultante del error de especificación y cambios repentinos en el comportamiento de las variables pronosticadas (Bates & Granger, 1969). En parte podemos ver la combinación de pronósticos como el consenso de distintos expertos independientes sobre el valor futuro de una variable.

La primera referencia que se encuentra en la literatura sobre la combinación de pronósticos es el trabajo empírico de Barnard (1963), donde examina el enfoque de pronósticos mediante la metodología de Box-Jenkins y el método de suavización exponencial desarrollado por Brown para pronosticar ventas. Observó que, aunque el método de Box-Jenkins, de manera general, producía mejores pronósticos (minimizaban el error cuadrático medio), la media aritmética entre el pronóstico mediante el método de Box-Jenkins y el método de Brown mejoraba el proceso de predicción.

A partir del trabajo de Barnard (1963), las investigaciones empíricas se han enfocado en explorar si existe alguna combinación de pronósticos que supere el rendimiento del promedio simple. El trabajo considerado pionero en esta área es el de Bates y Granger (1969), quienes

exploraron distintos métodos para combinar pronósticos, encontrando que si los pronósticos individuales no contienen errores de correlación serial, la combinación de ellos necesariamente será igual o más precisa que el pronóstico individual más competitivo. En dicho caso, un pronóstico combinado con una varianza de error significativamente menor que cualquiera de los otros modelos implicaría necesariamente que se puede mejorar algo de los demás modelos individuales y, con ello, obtener indicios de un modelo más verdadero y de mayor valor analítico.

La combinación de pronósticos individuales emerge como una estrategia óptima para minimizar la varianza del error del predictor, superando así las limitaciones de los modelos individuales. Como menciona Elliott y Timmermann (2008), la combinación de pronósticos a menudo resulta ser una estrategia más razonable que seleccionar un solo pronóstico individual, principalmente en ausencia de conocimiento sobre cuál modelo incluye el modelo verdadero. Además, resultan comportarse como un seguro ante quiebres o no estacionariedades futuras. La combinación de pronósticos puede depender del estado de la economía, ya que la rapidez con la que distintos pronósticos incorporan los cambios económicos podría variar y ser mejor capturada por algún pronóstico individual específico (Clemen, 1989). Por otra parte, Wang et al. (2023), mencionan que, bajo el criterio de minimización de la varianza, la combinación de pronósticos individuales, ya sea con errores independientes o correlacionados, siempre es óptima y nunca genera un pronóstico peor que el pronóstico individual más competitivo. Lo que destaca la flexibilidad de la combinación de pronósticos, su adaptabilidad a distintas situaciones económicas y su superioridad sobre los pronósticos individuales.

La técnica de combinar pronósticos individuales ha encontrado aplicaciones extensas y variadas en el campo de la macroeconomía. Tal es el caso para la estimación del PIB Reid (1968), para pronosticar la inflación Engle et al. (1984), la demanda de dinero Figlewski y Urich (1983), para el caso de la tasa de crecimiento Rapach y Strauss (2008), para el tipo de cambio Bilson (1983) y en casi cualquier campo de la macroeconomía en el cual los pronósticos sean de importancia. Algunos trabajos que realizan una gran recopilación de estos trabajos son el de (Clemen, 1989; Elliott & Timmermann, 2008; Wang et al., 2023), que ofrecen una amplia recopilación de referencias sobre la combinación de pronósticos.

A pesar de la aparente simplicidad del método propuesto por Bates y Granger (1969), su adopción fue lenta debido a argumentos que cuestionaban la necesidad de combinar mode-

los para mejorar el pronóstico, asociando que esta necesidad se debía principalmente a una mala especificación de los modelos. Pero los fuertes resultados estadísticos han permitido a esta técnica dejar las críticas atrás. No existe un modelo dominante que subsuma completamente el contenido informativo de los otros; cada modelo aporta un incremento marginal de información que puede ser explotado de manera útil en la combinación de pronósticos (Cooper & Nelson, 1975). La combinación de pronósticos puede interpretarse más allá de un nivel pragmático, llegando a la conclusión de que la combinación de pronósticos conlleva a la optimización de modelos y la integración de conjuntos de información. (Granger & Ramanathan, 1984). Por otra parte, Clemen y Guerard Jr (1989), prueba que la combinación de pronósticos, incluyendo métodos econométricos y de extrapolación, muestra un potencial significativo para mejorar la precisión y provee información incremental valiosa de manera individual que otros pronósticos ignoran.

Capítulo 3

Metodología

Esta sección detalla la metodología que se empleara para seleccionar los modelos que se usarán en las diversas combinaciones de pronósticos. También se describen las técnicas de combinación de pronósticos implementadas, con el fin de determinar una ponderación óptima de distintos modelos individuales de pronóstico para obtener una proyección combinada del Indicador Global de la Actividad Económica (IGAE) con el objetivo que esta logre un pronostico eficiente con respecto a los pronósticos individuales.

3.1. Técnicas de Combinación de Pronósticos

Con el objetivo de establecer un marco notacional claro, definamos la matriz $F_{T \times P}$ que representa los pronósticos, con T indicando la cantidad de periodos temporales y P el número de modelos de pronóstico. Así, disponemos de P pronósticos en cada instante temporal. El pronóstico generado por el modelo i -ésimo se denota como f_i , siendo $i \in \{1, \dots, P\}$. Se prescindirá del subíndice t que designa la dimensión temporal del pronóstico cuando esto no induzca a ambigüedad. En lo que concierne a métodos de combinación que exigen una ordenación de los pronósticos componentes, $f^{(i)}$ simboliza la estadística de orden i -ésima obtenida de la muestra de pronósticos en un punto dado. El peso asignado al pronóstico del modelo i dentro de la combinación final se expresa como w_i , y el pronóstico combinado resultante se designa por f_c .

A lo largo del tiempo, como menciona Wang et al. (2023), han surgido estrategias más complejas para combinar pronósticos. Estas estrategias se clasifican en simples, combinaciones lineales, combinaciones basadas en el rendimiento, combinaciones con pesos bayesianos,

combinaciones mediante el aprendizaje y combinaciones no lineales. Para el objetivo de esta investigación, hemos optado por aquellos esquemas que se utilizan con mayor frecuencia dentro de la literatura (Weiss et al., 2018).

En los resultados se presentan cuatro métodos específicos: BG (Bates/Granger, 1969), InvW (por Inversión de Rango), CLS (Mínimos Cuadrados Restringidos) y SA (por Media Simple). Esta selección se basa en Castle et al. (2019), quien indica que han sido las técnicas de combinación de pronóstico más usuales. En el apéndice se provee el código para otras técnicas de combinación de pronósticos, ofreciendo una guía completa para su implementación y comparación.

3.1.1. Promedio Simple

El método más intuitivo para combinar pronósticos es utilizar el promedio de todos ellos. Con el tiempo, este enfoque tan sencillo se ha establecido como un excelente punto de referencia, a pesar o quizás debido a su simplicidad (Wang et al., 2023).

El pronóstico combinado se da de forma directa por

$$f^c = \frac{1}{P} \sum_{i=1}^P f_i. \quad (3.1)$$

Este método ha sido denominado en la literatura contemporánea como el rompecabezas de la combinación de pronósticos”, debido a que, incluso después de más de cincuenta años, sigue siendo la estrategia predominante en este contexto.

3.1.2. Bates y Granger (1969)

En el trabajo pionero de Bates y Granger (1969), se propuso una metodología para la combinación de pronósticos basada en la teoría de diversificación de carteras. Esta técnica utiliza los elementos diagonales de la matriz de errores cuadráticos medios estimados para calcular los pesos de combinación de los pronósticos. La fórmula para el pronóstico combinado se expresa como:

$$f^c = \sum_{i=1}^P f_i \times \frac{\hat{\sigma}^{-2}(i)}{\sum_{j=1}^P \hat{\sigma}^{-2}(j)} \quad (3.2)$$

donde $\hat{\sigma}^2(i)$ es el error cuadrático medio estimado del modelo i . Este enfoque omite la

correlación entre los pronósticos de los componentes debido a las dificultades para estimar con precisión la matriz de covarianza.

3.1.3. Método de Rango Inverso

Propuesto inicialmente por Aiolfi y Timmermann (2006), este método ordena los modelos de pronóstico de acuerdo con su rendimiento hasta el tiempo N . El modelo con el menor error cuadrático medio recibe el primer rango, el segundo menor el segundo rango, y así sucesivamente. El pronóstico combinado se deriva entonces con la siguiente fórmula:

$$f^c = \sum_{i=1}^P f_i \times \frac{Rank_i^{-1}}{\sum_{j=1}^P Rank_j^{-1}} \quad (3.3)$$

Según Aiolfi y Timmermann (2006), este procedimiento, similar al de Bates y Granger (1969), no toma en cuenta la correlación entre los errores de los diferentes pronósticos. No obstante, su fortaleza radica en su resistencia a los valores atípicos, pues la estructura de los rangos no se ve significativamente afectada por pronósticos extremos.

3.1.4. Mínimos Cuadrados con Restricciones (CLS)

La regresión con restricciones de mínimos cuadrados (CLS) aborda el control de los coeficientes extremos por medio de la minimización de la suma de los errores cuadráticos, imponiendo restricciones específicas en los coeficientes. En CLS, se asegura que todos los coeficientes estimados w_i sean no negativos ($w_i \geq 0$) y que su suma totalice uno $\sum_{i=1}^P w_i = 1$.

Para abordar el problema de la multicolinealidad, que puede influir negativamente en las estimaciones de la regresión CLS, se ha implementado el método desarrollado por Babaie-Kafaki y Roozbeh (2017), el cual es una versión modificada de la descomposición de Cholesky basada en la regresión Ridge. Esta versión es fundamental para mitigar los efectos adversos que la multicolinealidad puede ocasionar en la precisión de las estimaciones del modelo (Weiss et al., 2018).

3.2. Métodos Adicionales

Aunque los métodos mencionados anteriormente son los principales enfoques utilizados en los resultados presentados en este trabajo, existen muchas otras técnicas de combinación

de pronósticos que pueden ser útiles en diferentes contextos. En el apéndice se provee el código para implementar otras técnicas de combinación de pronósticos, ofreciendo una guía completa para su implementación y comparación. Estas técnicas adicionales incluyen métodos basados en regresión, métodos de combinación basados en vectores propios, y otros enfoques avanzados.

3.3. Modelos de Pronósticos

3.3.1. ARIMA

Los modelo ARIMA, que significa AutoRegressive Integrated Moving Average, desarrollados en los años setenta por Box y Jenkins (1970), combina elementos autoregresivos (p), de diferenciación (d) y de media móvil (q) en un único modelo de series temporales. La serie diferenciada y'_t se modela como:

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t, \quad (3.4)$$

donde los ϕ_i y θ_j son los coeficientes para los términos autoregresivos y de media móvil respectivamente, y c es una constante. Este enfoque permite modelar series que no son estacionarias en su forma original, utilizando la diferenciación para estabilizar la media.

La notación de operador de rezago simplifica las fórmulas, permitiendo una representación compacta y manejable de los componentes del modelo:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t, \quad (3.5)$$

Seleccionar los parámetros p , d y q adecuados es crucial y puede automatizarse con herramientas como la función `auto.arima()` del paquete `forecast`, que optimiza estos valores basándose en algún criterio de información.

3.3.2. Suavizado Exponencial Holt-Winters

El método Holt-Winters se fundamenta en el suavizado exponencial para proporcionar pronósticos de series temporales que incorporan tendencias y patrones estacionales. Este modelo es una extensión del método de suavizado exponencial de Holt, diseñado para series que

muestran estacionalidad además de tendencia. Al omitir el componente estacional del modelo Holt-Winters, se simplifica al método de Holt, orientado a series con tendencia pero sin patrones estacionales recurrentes (Hyndman & Athanasopoulos, 2018).

El método de Holt, conocido como el método de tendencia lineal de Holt, se caracteriza por incluir ecuaciones de suavizado para el nivel y la tendencia de la serie:

- **Ecuación de pronóstico:** $\hat{y}_{t+h|t} = \ell_t + hb_t$
- **Ecuación de nivel:** $\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$
- **Ecuación de tendencia:** $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$

Donde:

- ℓ_t es la estimación del nivel de la serie en el tiempo t ,
- b_t es la estimación de la tendencia (pendiente) en el tiempo t ,
- α y β^* son los parámetros de suavizado para el nivel y la tendencia, respectivamente, ambos en el rango de 0 a 1.

Diferente del suavizado exponencial simple, donde el pronóstico es constante, el método de Holt integra un pronóstico que incrementa o decrece a lo largo del tiempo. El pronóstico a h periodos en el futuro se calcula sumando h veces el valor estimado de la tendencia al último nivel estimado.

Este modelo se describe en Hyndman y Athanasopoulos (2018), donde se explica tanto la versión estacional como la no estacional de manera más amplia.

3.3.3. TBATS

El Modelo TBATS, introducido por De De Livera et al. (2011), extiende los modelos de suavizado exponencial para abordar series temporales con múltiples patrones estacionales y tendencias. Este modelo se basa en componentes como transformación Box-Cox, errores tipo ARMA, y componentes tanto de tendencia como estacionales, diseñados para manejar patrones complejos de forma más eficaz.

Dentro del modelo TBATS, se utiliza una la transformación Box-Cox para manejar datos con variaciones de amplitud. Las componentes estacionales se modelan usando sumas de

funciones trigonométricas para capturar patrones estacionales de varios períodos, facilitadas por la teoría de series de Fourier.

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \quad (3.6)$$

donde cada término trigonométrico se actualiza según:

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1 d_t, \quad (3.7)$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2 d_t. \quad (3.8)$$

Lo que permite una representación flexible y precisa de patrones estacionales y de tendencia múltiples. El software de R para pronósticos elige automáticamente los mejores parámetros basándose en el criterio de información de Akaike, con la función automática (`tbats()`) del paquete `forecast`.

3.3.4. Theta

El Modelo Theta desarrollado por Assimakopoulos y Nikolopoulos (2000) para pronosticar series de tiempo univariadas ajustando las tendencias mediante un coeficiente aplicado a las segundas diferencias de la serie. Este método se basa en la idea de modificar la curvatura de las series para reflejar más precisamente las tendencias y patrones subyacentes.

El proceso comienza transformando una serie $\{Y_1, \dots, Y_t\}$ en $\{Y_{1,\theta}, \dots, Y_{t,\theta}\}$, donde cada $Y_{n,\theta}$ se define como:

$$\Delta^2 Y_n = \Delta^2 Y_{n,\theta} \quad (3.9)$$

para $n = 3, \dots, T$. Este ajuste se modela con la relación:

$$Y_{n,\theta} = a_0 + b_0(t-1) + \theta Y_n, \quad (3.10)$$

donde a_0 y b_0 son constantes, y θ se optimiza para minimizar las desviaciones cuadráticas entre Y_n y $Y_{n,\theta}$.

Para proyectar valores futuros, el Modelo Theta emplea un promedio de proyecciones basadas en ajustes de cero y doble diferenciación, representado por:

$$\hat{Y}_{t+h} = \frac{1}{2}(Y_{t+h,0} + Y_{t+h,2}), \quad (3.11)$$

facilitando una estimación que equilibra las tendencias observadas con ajustes estructurales de la serie, esta metodología se valora principalmente por su eficacia en series con comportamientos de tendencia claros.

3.3.5. Redes Neuronales Autorregresiva (NNAR)

Los modelos de redes neuronales son métodos de pronóstico que simulan estructuras neuronales del cerebro, permitiendo analizar patrones complejos y no lineales en los datos. Estas redes están estructuradas en varias capas, desde la capa de entrada que recibe los datos hasta la capa de salida que ofrece los resultados del modelo. Entre estas, pueden existir una o más capas ocultas que ayudan a capturar las no linealidades de los datos.

En la configuración más simple, una red sin capas ocultas es equivalente a una regresión lineal. La salida se calcula como una combinación lineal de entradas, representada por:

$$\hat{y} = \ell + \sum h_i b_i, \quad (3.12)$$

donde ℓ representa el nivel y b_i son los pesos aplicados a cada entrada h_i .

Una red más compleja incluye al menos una capa oculta que permite modelar no linealidades. Cada nodo en esta capa combina las entradas de la capa anterior mediante una suma ponderada, seguida de una transformación no lineal, típicamente una función sigmoide:

$$z_j = b_j + \sum_{i=1}^4 w_{ji} x_i, \quad (3.13)$$

$$s(z) = \frac{1}{1 + e^{-z}}, \quad (3.14)$$

donde w_{ji} son los pesos asignados a las conexiones entre las entradas x_i y el nodo j en la capa oculta, y b_j es el sesgo.

Un ejemplo de esta es la regresión NNAR la cual utiliza valores pasados de la serie temporal como entradas para predecir futuros valores. Este modelo, NNAR(p, k), donde p es el número de retrasos y k el número de neuronas en la capa oculta, integra tanto entradas recientes como estacionales. Es efectivo tanto para datos estacionales como no estacionales,

ajustando los parámetros automáticamente para optimizar dado un criterio de información.

3.4. Procedimiento de Evaluación de Pronósticos

En esta sección se desarrolla la metodología para comparar la capacidad de predicción de los modelos combinados. Un hecho común es que, al evaluar qué tan bueno es un pronóstico para realizar predicciones, parece natural utilizar la mayoría de los datos de las series de tiempo para entrenar los modelos de pronóstico y evaluar con los últimos datos de la serie qué tan buenos son nuestros modelos prediciéndolos. El problema con este procedimiento radica en que se requeriría de una gran cantidad de series temporales con las mismas características para obtener resultados generales, lo que resulta en muchos contextos difícil o prácticamente imposible, dado que cada serie de tiempo tiende a tener características muy particulares.

Para solucionar estos problemas, lo que se hace en la práctica se realiza es diseñar algún algoritmo de validación cruzada de ventana móvil o "rolling window" que divida la serie de tiempo en distintos segmentos más pequeños en los cuales para cada segmento se pueda realizar la evaluación de los métodos de pronóstico a evaluar. De manera formal el procedimiento propuesto es el siguiente:

Supóngase que se dispone de una serie temporal Y_T con T observaciones. Lo primero que se hace es definir el tamaño de la ventana de entrenamiento t , una ventana de validación de tamaño h y una ventana de evaluación, también de tamaño h .

Para determinar el tamaño de cada ventana, se tiene en cuenta la naturaleza de los datos y el objetivo del análisis. El tamaño de la ventana de entrenamiento debe ser suficiente para capturar la dinámica subyacente de la serie, mientras que las ventanas de validación y evaluación deben permitir una adecuada verificación del rendimiento predictivo a corto, mediano o largo plazo.

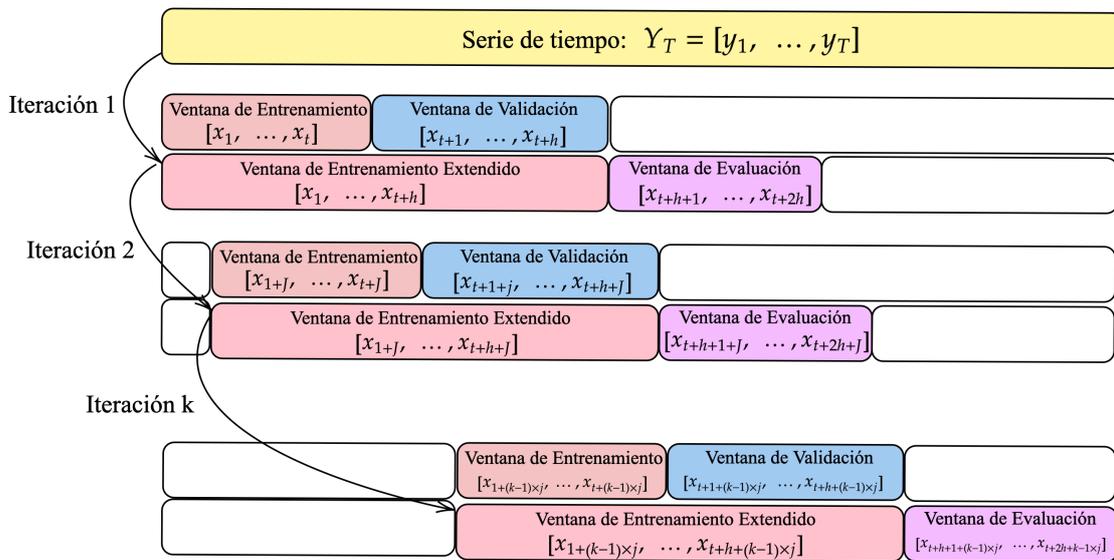
Una vez seleccionados los tamaños de las ventanas, se procede a entrenar los modelos en la ventana de entrenamiento. Las predicciones se generan entonces para el tamaño de la ventana de validación h , y mediante los distintos métodos de combinación de pronósticos presentados en la sección anterior, se calculan los pesos de combinación en esta ventana.

Con los pesos obtenidos, se ajustan nuevamente los modelos utilizando los datos de las ventanas de entrenamiento y validación combinadas. Las predicciones combinadas se generan utilizando los pesos de la ventana de validación. Estas predicciones son evaluadas en la

ventana de evaluación para obtener diversas métricas y funciones de pérdida, proporcionando así una estimación robusta del desempeño de los modelos fuera de muestra.

El desplazamiento de la ventana se realiza avanzando j periodos hacia adelante después de cada iteración. Esto implica que se “olvidan” los primeros j datos de la serie temporal y se añaden j datos más recientes al final de la ventana. Este proceso se repite hasta agotar las observaciones disponibles en la serie temporal.

Figura 3.1. Esquema del funcionamiento de una ventana móvil.



Fuente: elaboración propia.

3.5. Criterios de desempeño de pronósticos.

Para un horizonte de evaluación pseudo fuera de muestra de un pronóstico combinado o individual ($h = 6, h = 9$ y $h = 12$, se definen los siguientes criterios de desempeño de pronóstico como:

3.5.1. Mean Error (ME)

$$ME_j = \frac{1}{H} \sum_{i=1}^H (Y_{t+i} - \hat{Y}_{j,t+i})$$

El ME proporciona una indicación directa del sesgo de un modelo. Un ME cercano a cero sugiere que el modelo no tiene sesgo sistemático en sus predicciones, mientras que un valor positivo o negativo indica una tendencia a sobrestimar o subestimar, respectivamente.

3.5.2. Mean Absolute Error (MAE)

El Error Absoluto Medio (MAE) se calcula como:

$$MAE_j = \frac{1}{H} \sum_{i=1}^H |Y_{t+i} - \hat{Y}_{j,t+i}|$$

El MAE mide la magnitud promedio de los errores de un modelo sin considerar su dirección. Esto lo hace útil para evaluar la precisión general de un modelo en términos de cómo de grande es el error típico en sus predicciones, independientemente de si esas predicciones son consistentemente altas o bajas.

3.5.3. Mean Percentage Error (MPE)

El Error Porcentual Medio (MPE) evalúa el desempeño de las técnicas de pronóstico en términos porcentuales, que es crucial cuando el tamaño del error relativo es importante:

$$MPE_j = \frac{1}{H} 100 \sum_{i=1}^H \left(\frac{Y_{t+i} - \hat{Y}_{j,t+i}}{Y_{t+i}} \right)$$

El MPE, al no tomar valores absolutos, puede mostrar un sesgo por compensación entre errores positivos y negativos, lo que puede ser engañoso si dichos errores son grandes.

3.5.4. Mean Absolute Percentage Error (MAPE)

El Error Porcentual Absoluto Medio (MAPE) se define como:

$$MAPE_j = \frac{1}{H} 100 \sum_{i=1}^H \left| \frac{Y_{t+i} - \hat{Y}_{j,t+i}}{Y_{t+i}} \right|$$

Al igual que el MPE, el MAPE proporciona una medida de la precisión en términos porcentuales sin la influencia de la dirección del error, haciéndolo ideal para comparar la precisión de diferentes técnicas de pronóstico sobre una base porcentual.

3.5.5. Root Mean Square Error (RMSE)

El RMSE mide la magnitud de los errores de predicción, elevando al cuadrado cada error antes de promediarlos y tomando la raíz cuadrada del resultado. Esto da mayor peso a los errores grandes, haciendo que RMSE sea especialmente útil en situaciones donde los errores grandes son más perjudiciales que los pequeños. La fórmula es:

$$RMSE_j = \sqrt{\frac{1}{H} \sum_{i=1}^H (Y_{t+i} - \hat{Y}_{j,t+i})^2}$$

A diferencia del MAE, el RMSE eleva cada error al cuadrado, incrementando el peso de los errores más grandes en el cálculo final. Esto hace que el RMSE sea particularmente útil en situaciones donde los errores grandes son más perjudiciales que muchos errores pequeños (Makridakis & Hibon, 2000).

3.5.6. Comparación entre las Métricas

El ME y el MPE son útiles para detectar sesgos en las predicciones, con el ME indicando si un modelo tiende a sobreestimar o subestimar y el MPE proporcionando esta misma información pero en términos porcentuales. Ambos son cruciales cuando el equilibrio de las predicciones es un factor importante.

Por otro lado, el MAE y el MAPE evalúan la precisión sin considerar la dirección de los errores. El MAE mide la magnitud promedio de los errores, siendo robusto frente a los outliers, mientras que el MAPE ajusta los errores a la escala de los valores reales, facilitando comparaciones entre modelos de distintas escalas.

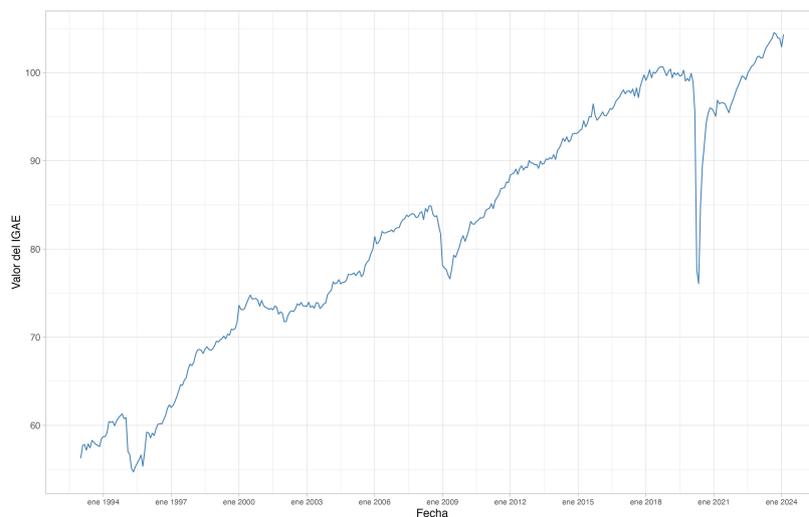
El RMSE es sensible a los errores grandes, lo que lo hace valioso en contextos donde estos errores son especialmente perjudiciales. Su capacidad para penalizar fuertemente los errores grandes lo hace preferido en áreas como el control de calidad y la investigación científica.

Capítulo 4

Datos

En el presente estudio, se empleará la serie del Indicador Global de la Actividad Económica (IGAE) como principal variable de análisis. Este indicador, se construye a partir de índices de volumen físico que reflejan la producción efectiva del país. Con una periodicidad mensual, la serie abarca desde enero de 1993 hasta enero de 2024 (INEGI, 2024). Es importante destacar que los datos han sido ajustados para excluir efectos estacionales¹, permitiendo así una comparación más precisa entre meses consecutivos al eliminar fluctuaciones periódicas anuales provocadas por factores exógenos.

Figura 4.1. Indicador Global de la Actividad Económica, enero de 1993 a enero de 2024 Serie desestacionalizada (índice 2018-100)



Fuente: elaboración propia con datos del INEGI.

¹ Tapia Alvarado (2021) argumenta que los efectos estacionales son la principal fuente de varianza en las series temporales y complican la comprensión de señales económicas por parte de los indicadores.

La selección de la serie temporal del Índice Global de Actividad Económica (IGAE) se fundamenta en su relevancia macroeconómica, ya que se considera una aproximación al Producto Interno Bruto (PIB). Esto proporciona una perspectiva valiosa sobre las tendencias de la economía mexicana a corto plazo, aportando información crucial para la toma de decisiones informadas en el ámbito económico.

Como se describió anteriormente, nuestro principal insumo, además de la serie de datos, incluye los pronósticos generados para cada ventana de validación. En estas ventanas se obtienen los pesos para la combinación de pronósticos y se evalúan diversas métricas para comparar los métodos de combinación. Cada modelo, ajustado automáticamente mediante el criterio de información de Akaike (AIC), cuenta con características particulares. Se ha puesto especial atención en seleccionar el modelo más adecuado de cada familia, asegurando que la predicción sea lo más precisa posible.

Para garantizar una estimación robusta, siempre se mantuvo un mínimo de 36 datos en la ventana de entrenamiento, lo cual es esencial para estimar modelos como Theta, NNAR y TBATS. En particular, utilizamos conjuntos de 50, 100 y 150 datos para observar la evolución de los errores de pronóstico a lo largo del tiempo. Adicionalmente, se empleó un conjunto de 300 datos para obtener las métricas que permiten comparar los modelos combinados. Específicamente para el modelo TBATS, se indicó la ausencia de estacionalidad, dado que la serie ya había sido previamente tratada en este aspecto.

La selección de las ventanas de validación y evaluación tomó en cuenta que los datos son mensuales y que las predicciones se planean a 6, 9 y 12 meses. Para el desplazamiento de las ventanas, se eligió un intervalo de $j = 2$, lo que permite identificar el comportamiento de los diferentes modelos de combinación de pronósticos ante la incorporación de nueva información en un contexto real.

Capítulo 5

Resultados

En esta sección se presentan los principales resultados relacionados con la generación de pronósticos para el Indicador Global de la Actividad Económica (IGAE) y los resultados de implementar diversas técnicas de combinación de pronósticos en los modelos específicos para el IGAE, tal como se discutió en la sección (3.3).

Inicialmente, se emplea una estrategia para evaluar el comportamiento de los errores basada en la metodología introducida por Bates y Granger (1969), la cual se describe en la sección (A.1.5). Este enfoque permite identificar la sensibilidad de los pronósticos frente a variaciones en el volumen de datos utilizados y observar el patrón de los errores en cada modelo de pronóstico. Este análisis se lleva a cabo para diversos tamaños de ventana de validación y evaluación: 6, 9 y 12 meses. Posteriormente, se examina la evolución de varias métricas de precisión para pronósticos fuera de muestra, con el objetivo de evaluar el desempeño de las técnicas de combinación de pronósticos. Finalmente, se comparan estas métricas utilizando como referencia la media aritmética.

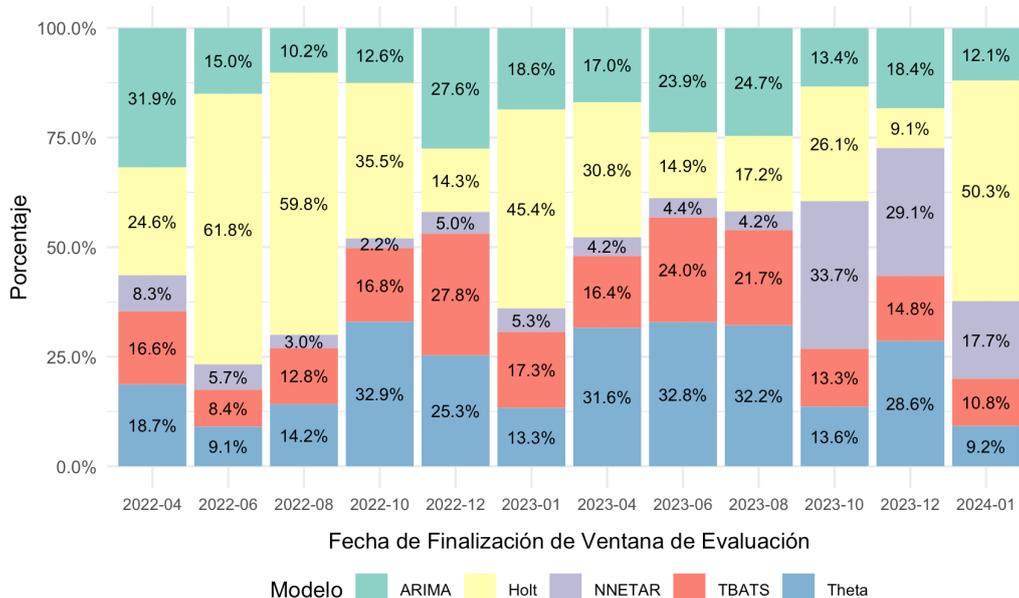
5.1. Desempeño de los Modelos individuales en Estabilidad Secuencial

Con el objetivo de responder a la pregunta ¿Las técnicas de combinación sugieren ponderaciones iguales independientemente de la información disponible?, se utiliza el método de Bates y Granger (1969), el cual se describe detalladamente en la sección (A.1.5). Con este método, se evalúa el desempeño de los pronósticos y la ponderación asignada respectivamente. El motivo por el que usamos esta técnica es porque, al ser una ponderación respecto al

error cuadrático medio, nos permite argumentar, en promedio, cuánto aporta un pronóstico individual y hablar en términos de porcentaje respecto al total de la batería de pronósticos, características que no son generales entre las técnicas de combinación de pronósticos.

Para visualizar los resultados, la información se presenta en un Histograma de Estabilidad Secuencial, donde cada barra representa la importancia relativa, o peso, asignado a un modelo específico de pronóstico para una ventana de evaluación temporal estimados a través de la metodología de Bates y Granger (1969). El histograma ayuda a identificar qué modelos mantienen un peso consistente a lo largo de varias ventanas temporales, además permite comparar visualmente la importancia relativa de cada modelo en la combinación de pronósticos y, al poder visualizar los cambios en los pesos a lo largo del tiempo, es posible evaluar cómo diferentes contextos o cambios en la serie temporal afectan la eficacia de cada modelo.

Figura 5.1. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (150), Validación (12) y Evaluación (12), Desplazamiento (2).

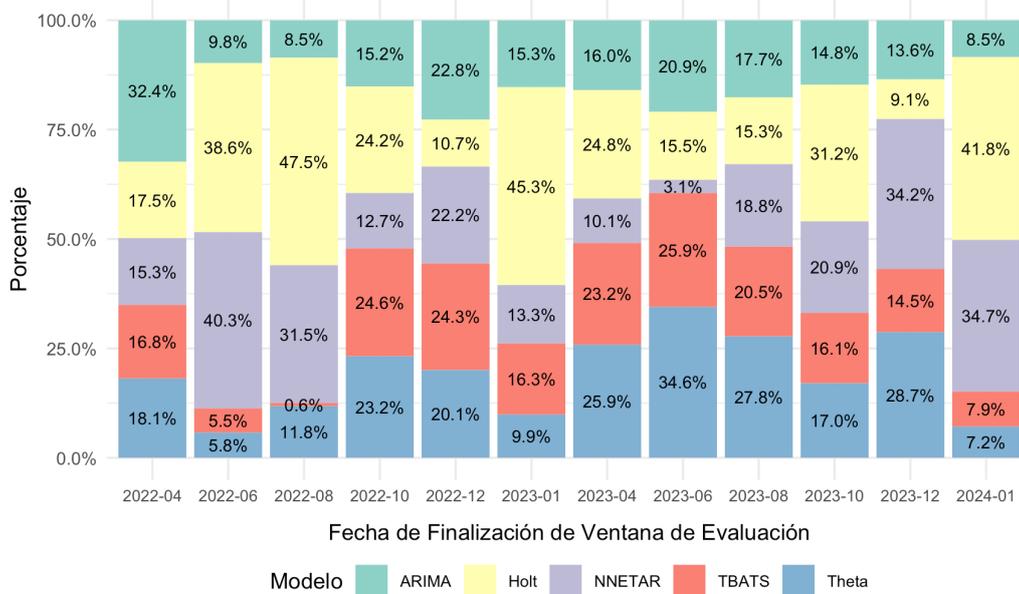


Fuente: elaboración propia con datos del INEGI.

La figura (5.1) ilustra el Histograma de Estabilidad Secuencial para una muestra de 200 observaciones que abarcan de enero de 2007 a enero de 2024. Este histograma revela que las ponderaciones asignadas a los modelos no son consistentes a lo largo del tiempo, indicando que el error de los modelos individuales varía entre las distintas ventanas temporales. Además, se observa que ciertos métodos son más efectivos según la ventana temporal es-

pecífica. También se nota que algunos métodos, dependiendo de la ventana de evaluación, pueden adquirir una importancia relativa significativamente mayor en comparación con otros métodos.

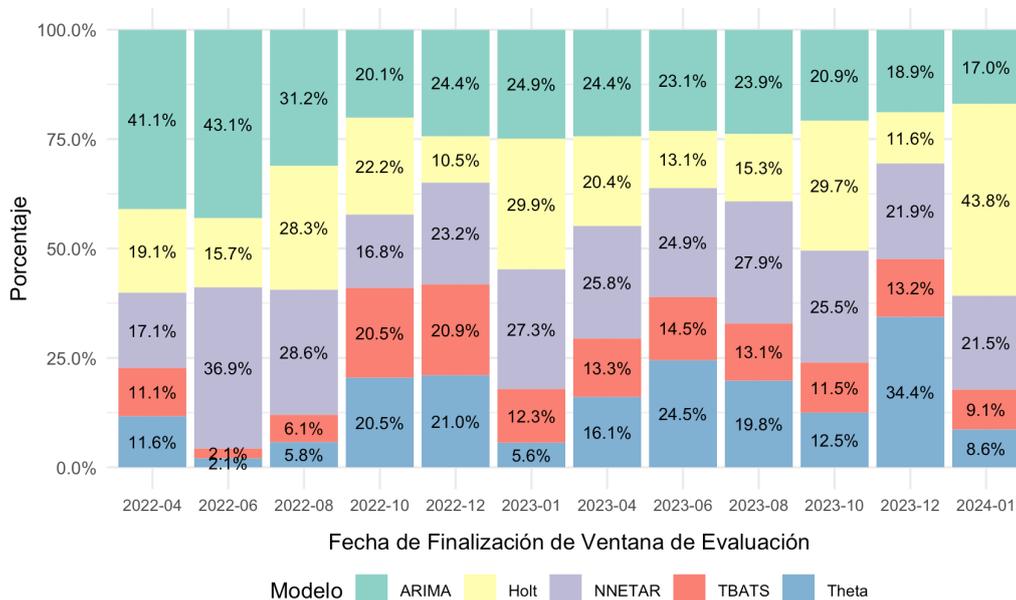
Figura 5.2. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (100), Validación (12) y Evaluación (12), Desplazamiento (2).



Fuente: elaboración propia con datos del INEGI.

De igual manera, la figura (5.2) muestra el Histograma de Estabilidad Secuencial para un conjunto de 150 observaciones, abarcando periodo de septiembre de 2011 a enero de 2024. Al igual que en los resultados previamente discutidos para 200 observaciones en la figura (5.2), las ponderaciones de los modelos continúan mostrando una falta de consistencia temporal. Este patrón subraya que los errores de los modelos individuales fluctúan a través de las distintas ventanas temporales y que la eficacia relativa de los métodos varía significativamente dependiendo de la ventana de evaluación. Estos hallazgos reafirman las observaciones hechas anteriormente, indicando una consistencia en los resultados a pesar de la reducción en el número de observaciones analizadas.

Figura 5.3. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (50), Validación (12) y Evaluación (12), Desplazamiento (2).



Fuente: elaboración propia con datos del INEGI.

Por otra parte, la figura (5.3) muestra el Histograma de Estabilidad Secuencial para un conjunto de 100 observaciones, cubriendo desde diciembre de 2015 hasta enero de 2024. Al igual que en los análisis anteriores para 150 y 200 observaciones en la figura (5.3), las ponderaciones asignadas a los modelos no exhiben consistencia temporal. Sin embargo, se observa que métodos como ARIMA y NNETAR tienden a mostrar una mayor consistencia en comparación con otros métodos.

La mayor consistencia observada en los métodos ARIMA y NNETAR con el uso de conjuntos de entrenamiento más pequeños podría deberse a que el modelo ARIMA es eficaz en la captura de relaciones lineales, y NNETAR es apto para ajustarse a patrones no lineales en series reducidas, lo cual les ayuda a ser más estables respecto a otros modelos en presencia de menos datos. El ejercicio se repitió para los distintos tamaños de muestra pero con distintas ventanas de evaluación y validación (6, 9 y 12 meses) y para los distintos tamaños de muestra, encontrando resultados similares, estos resultados pueden consultarse en el apéndice (B).

5.2. Resultado de la combinación de pronósticos

Tal y como se ha descrito a lo largo del trabajo, se procedió a aplicar algunos de los métodos de combinación de pronósticos discutidos en la Sección 3. En esta sección se presentan los resultados para una selección de cuatro métodos: BG (Bates/Granger, 1969), InvW (por Inversión de Rango), CLS (Mínimos Cuadrados Restringidos) y SA (por Media Simple). Esta selección se basa en Castle et al. (2019), quien indica que han sido las técnicas de combinación de pronóstico más usuales. Además, se evitó tener problemas con algunos métodos sensibles ante la colinealidad entre pronósticos, un problema comúnmente reportado en la literatura (Wang et al., 2023).

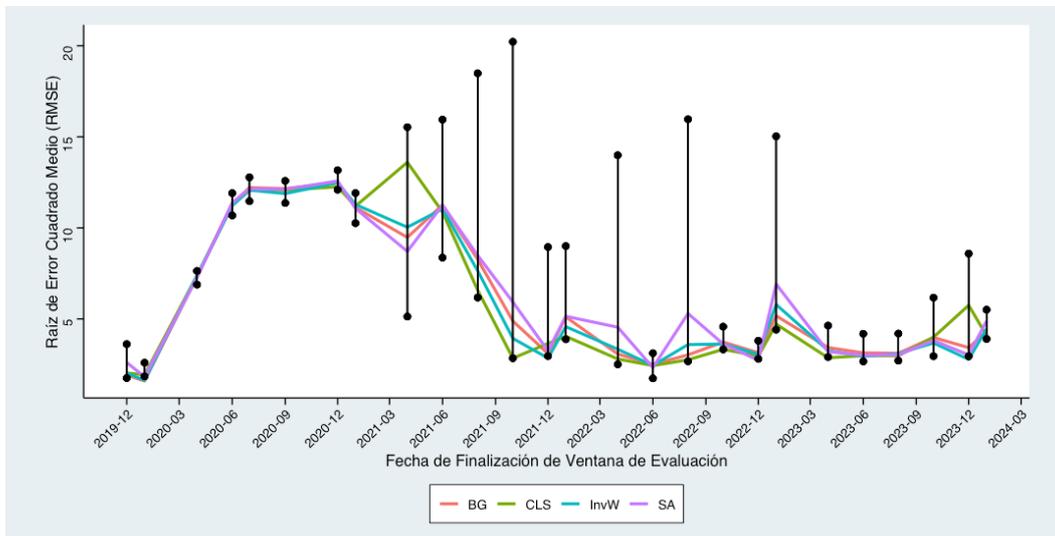
Siguiendo el algoritmo propuesto en la sección (3.4), se utilizó nuevamente la serie del IGAE para obtener los criterios de desempeño de las combinaciones de pronóstico discutidas en la sección (3.5). En particular, en esta sección se ilustra la eficacia relativa de modelos individuales en contraste con modelos combinados, a través del Raíz del Error Cuadrado Medio (RMSE). Los resultados para las otras métricas se encuentran en el apéndice. Para responder a la pregunta, ¿cómo se comparan las técnicas de combinación en términos de precisión de pronóstico fuera de muestra entre ellas y entre los pronósticos individuales?, se presentan los resultados detallados de cada técnica y su comparación con los modelos individuales.

Utilizando las fechas de finalización de ventanas de evaluación como eje temporal, las gráficas revelan cómo evolucionan los errores de estos modelos a lo largo del tiempo. Se destacan mediante líneas diferenciadas por cada método de combinación, identificando los valores extremos que representan los modelos individuales con los máximos y mínimos rendimientos. Las barras de error ilustran la variabilidad del rendimiento, mostrando claramente la tendencia de los modelos individuales a alcanzar los extremos de la escala de error en comparación con los modelos combinados, que generalmente exhiben un comportamiento más estable.

Los resultados indican una variabilidad significativa en el rendimiento de los modelos individuales, lo que podría reflejar una alta especialización o sensibilidad a condiciones específicas del entorno, lo que se hace más visible cuando las ventanas incorporan la información referente a los meses de pandemia a partir de marzo del 2020. Por el contrario, los modelos combinados ofrecen una aproximación más consistente y menos volátil, aún en contextos de inestabilidad económica, lo que sugiere su utilidad en aplicaciones donde se busca reducir el

riesgo de predicciones erráticas y blindarse ante el riesgo de seleccionar un solo modelo.

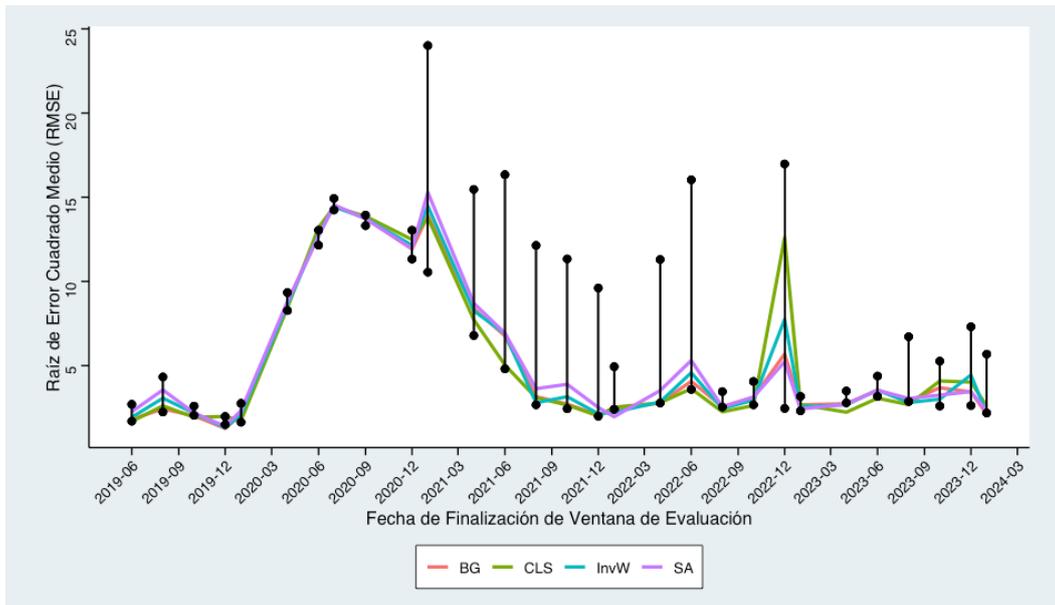
Figura 5.4. Comparativa de Desempeño de Modelos Individuales y Combinados (RMSE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



Fuente: elaboración propia con datos del INEGI.

Los resultados indican una variabilidad significativa en el rendimiento de los modelos individuales, lo que podría reflejar una alta especialización o sensibilidad a condiciones específicas del entorno, lo que se hace más visible cuando las ventanas incorporan la información referente a los meses de pandemia a partir de marzo del 2020. Por el contrario, los modelos combinados ofrecen una aproximación más consistente y menos volátil, aún en contextos de inestabilidad económica, lo que sugiere su utilidad en aplicaciones donde se busca reducir el riesgo de predicciones erráticas y blindarse ante el riesgo de seleccionar un solo modelo.

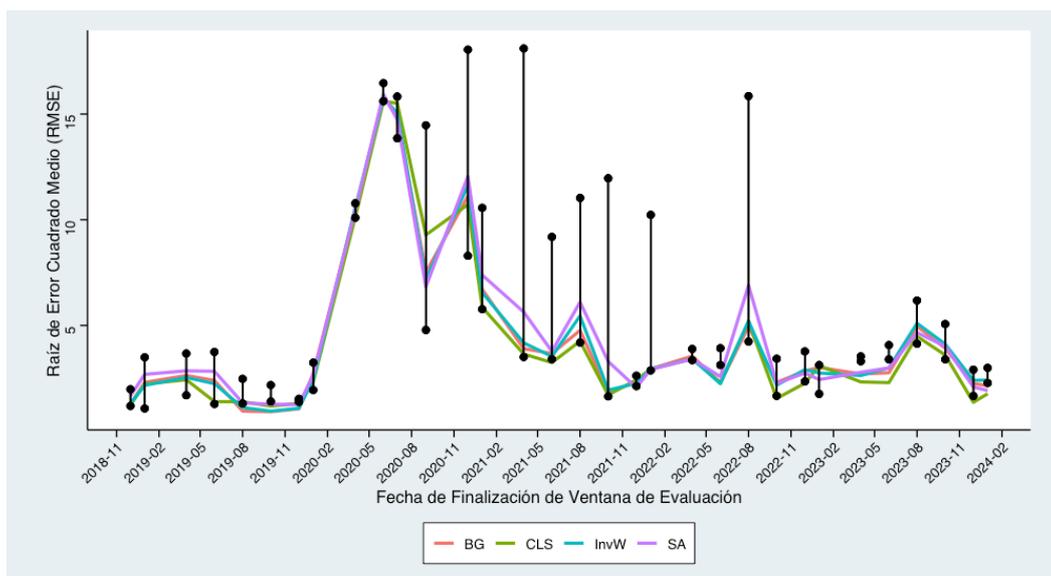
Figura 5.5. Comparativa de Desempeño de Modelos Individuales y Combinados (RMSE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)



Fuente: elaboración propia con datos del INEGI.

Además, se observa que para ventanas más pequeñas de validación y evaluación, el riesgo asociado con los modelos combinados se reduce significativamente. En particular, en el caso donde utilizamos las ventanas de evaluación y validación con un tamaño de 6 periodos como muestra la figura (5.6), hay casos donde el RMSE es excepcionalmente bajo, incluso saliéndose del rango típicamente observado en los modelos individuales. Esto sugiere que los modelos combinados no solo son más estables en general, sino que nos indica que las técnicas de combinación de pronósticos son más efectivas cuando se utiliza información más reciente para obtener los pesos de combinación.

Figura 5.6. Comparativa de Desempeño de Modelos Individuales y Combinados (RMSE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)



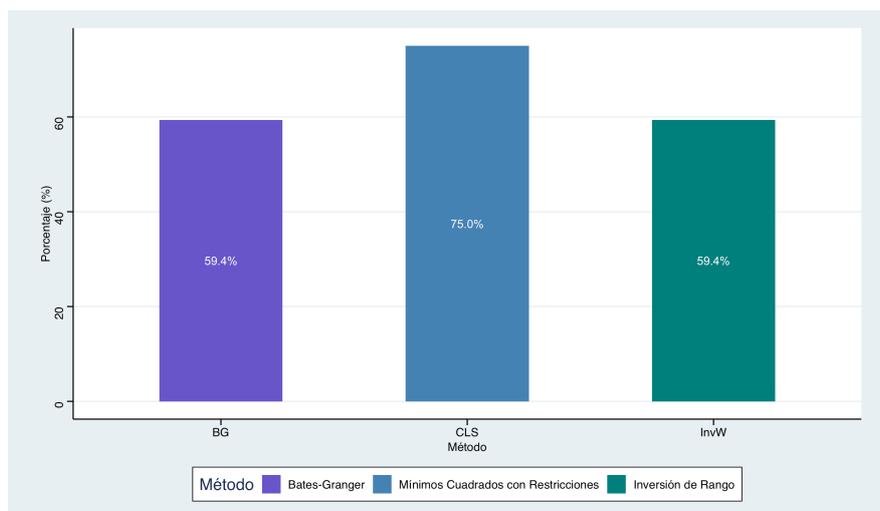
Fuente: elaboración propia con datos del INEGI.

En el apéndice B, específicamente en las tablas 8 y 9, se detallan los resultados obtenidos por los métodos de combinación de pronósticos. Estas tablas ilustran el RMSFE de las distintas técnicas de combinación comparadas con el rendimiento de la combinación de pronósticos de la media. La mejora en la precisión varía según el horizonte temporal y en promedio todas las combinaciones le ganan a la media.

5.3. Evaluación de Métodos de Combinación

En esta sección se presentan los resultados de la evaluación comparativa de diferentes métodos de combinación de pronósticos frente a la media aritmética, utilizando el Indicador Global de la Actividad Económica (IGAE) como la serie temporal de referencia. Las gráficas de barras muestran el rendimiento relativo de cada técnica en términos de la Raíz del Error Cuadrático Medio (RMSE) para diferentes configuraciones de ventanas de validación y evaluación. Se analizan escenarios con ventanas de 9 y 12 meses, proporcionando una visión clara de qué métodos logran superar consistentemente a la media aritmética en precisión de pronóstico fuera de muestra. Con este análisis, respondemos a la pregunta de investigación: ¿existen métodos óptimos de combinación que superen a la media aritmética en diferentes horizontes de pronóstico? A continuación, se detallan los resultados específicos para cada método y configuración.

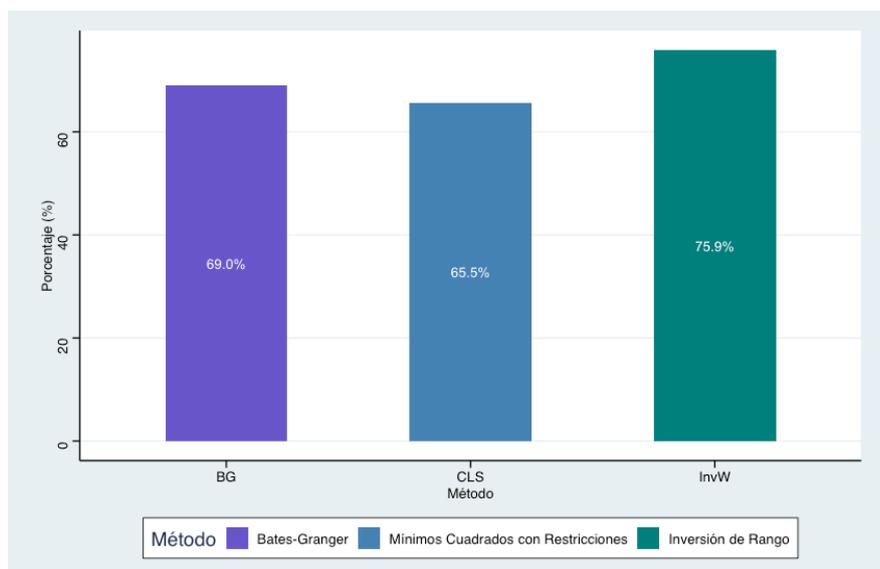
Figura 5.7. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Superar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)



Fuente: elaboración propia con datos del INEGI.

Para el horizonte de donde utilizamos una ventana de de validación y una ventana de evaluación de seis meses el método de Mínimos Cuadrados Restringidos (CLS), como se muestra en la figura (5.8) seguido del método de A.1.5 y el método de Inversión de Rango con el mismo puntaje.

Figura 5.8. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Superar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)

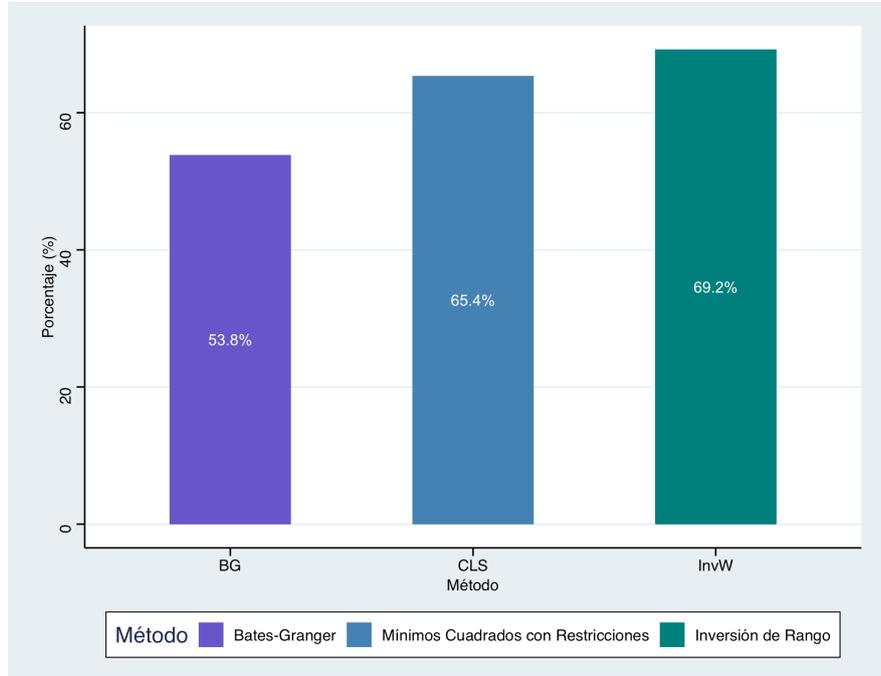


Fuente: elaboración propia con datos del INEGI.

Para el escenario con ventanas de validación y evaluación de 9 meses se obtiene resultados muy distintos en este caso el método de Inversión de Rango produce un RMSE más bajo que

la combinación de la media, con alrededor del 76 % de las veces, seguido del método de A.1.5, y en tercer lugar Mínimos Cuadrados Restringidos (CLS) con 69 % y 65 %, respectivamente.

Figura 5.9. Rendimiento Relativo de Métodos de Combinación de Pronósticos en Superar el Promedio Simple (RMSE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



Fuente: elaboración propia con datos del INEGI.

Para el tercer caso, con ventanas de validación y evaluación de 12 meses los resultados son contrastantes con los otros dos casos, como en el anterior caso el método de Inversión de Rango gana a la media un 69.2 % de las veces, mientras que el método de Mínimos Cuadrados Restringidos tiene un rendimiento del 64.4 % seguido del método de A.1.5 con un 53.8 %.

La evaluación de las técnicas de combinación de pronósticos en la predicción del Indicador Global de la Actividad Económica (IGAE) indica que estas técnicas nunca resultan en un RMSE mayor o igual al del modelo individual con el peor rendimiento. Interesantemente, en las ventanas de evaluación y validación más pequeñas, los pronósticos combinados a menudo logran un RMSE incluso inferior al de los mejores modelos individuales, aunque este no es siempre el caso para todas las ventanas de evaluación ni para todas las técnicas de combinación. En cuanto a la comparación de los métodos usuales de combinación frente a la media aritmética, se identificó que los métodos utilizados superaron en promedio a la media, sin embargo en ninguno de los casos fue de manera consistente es decir el 100 % de las veces lo que podría indicar que si bien las demás técnicas podrían hacerlo, la media es en general el *benchmark* a vencer tal lo son los modelos Arima entre modelos individuales.

Finalmente, la implementación de un algoritmo de evaluación de combinaciones demostró ser una herramienta eficaz para identificar la técnica de combinación óptima para el IGAE. Este algoritmo, al evaluar sistemáticamente el desempeño de cada técnica de combinación bajo diversas configuraciones y criterios, permite discernir cuál de ellas se adapta mejor a las particularidades de la serie temporal en este caso el IGAE. Los resultados indican que la adaptabilidad y la precisión de los métodos de combinación pueden ser significativamente mejoradas mediante la selección cuidadosa de la técnica basada en criterios empíricos de desempeño: selección de muestras, tanto para el ajuste del modelo, como para la selección de ventanas de validación y evaluación.

Capítulo 6

Conclusión

En la presente investigación, se presenta un algoritmo de Evaluación de Combinaciones de Pronósticos, se encuentra que se pueden construir pronósticos eficientes a partir de la combinación de pronósticos individuales. Al implementar este algoritmo, hemos identificado y aplicado técnicas de combinación de pronósticos que se adaptan óptimamente a las variaciones y dinámicas específicas de esta serie de tiempo económica.

Los resultados confirman que las técnicas de combinación de pronósticos reducen significativamente el riesgo de seleccionar modelos individuales, especialmente en escenarios con ventanas de evaluación y validación más cortas. Además, el algoritmo ha demostrado ser capaz de superar consistentemente la precisión de la media aritmética y ofrecer una diversificación de riesgo más efectiva, lo que es crucial en la predicción económica.

A pesar de su eficacia demostrada, el rendimiento del algoritmo sigue siendo altamente dependiente de la calidad y la actualización constante de los datos ingresados. Además, podría explorarse una mejora adicional en la eficacia del algoritmo mediante la exploración de técnicas de pronósticos más sofisticadas.

La implementación práctica de este algoritmo puede beneficiar a instituciones financieras, empresas privadas y organismos gubernamentales, facilitando pronósticos más confiables y robustos para variables de interés económico. Su implementación no solo proporciona una alternativa para reducir el riesgo, sino que también ofrece una herramienta para seleccionar la mejor estrategia para hacerlo.

Las futuras investigaciones podrían centrarse en aplicar el algoritmo a otras variables macroeconómicas para verificar su eficacia en diferentes contextos o integrar técnicas avanzadas de aprendizaje automático para mejorar aún más sus capacidades. También sería beneficio-

so desarrollar aplicaciones web interactivas que hagan accesible este tipo de herramientas a usuarios no especializados, democratizando el uso de técnicas de pronóstico.

En conclusión, el algoritmo de Evaluación de Combinaciones de Pronósticos es una herramienta útil y adaptable para mejorar la precisión de los pronósticos económicos. Su rendimiento depende de la calidad y actualización constante de los datos. Futuras investigaciones pueden aplicar el algoritmo a otras variables macroeconómicas, así como a simulaciones de PGD conocidos y explorar técnicas más avanzadas. La implementación práctica puede beneficiar a instituciones financieras, empresas privadas y organismos gubernamentales. Además, desarrollar aplicaciones web interactivas para democratizar el uso de estas técnicas. En este sentido, y contribuyendo a ello, en el apéndice (D) se presentan resultados preliminares del desarrollo de una plataforma de pronósticos la cual implementa el algoritmo desarrollado en esta tesina.

Referencias

- Aiolfi, M., & Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, 135(1-2), 31-53.
- Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30). Springer.
- Assimakopoulos, V., & Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International journal of forecasting*, 16(4), 521-530.
- Babaie-Kafaki, S., & Roozbeh, M. (2017). A revised Cholesky decomposition to combat multicollinearity in multiple regression models. *Journal of Statistical Computation and Simulation*, 87(12), 2298-2308.
- Barnard, G. A. (1963). New methods of quality control. *Journal of the Royal Statistical Society. Series A (General)*, 126(2), 255-258.
- Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the operational research society*, 20(4), 451-468.
- Benavides, G., & Capistran, C. (2009). Pronosticó de la Volatilidad del Tipo de Cambio: el Desempeño Superior de Combinaciones Condicionales de Pronósticos de Series de Tiempo y Pronósticos Implícitos en Opciones. *Revista Banco de México*, 1(1), 9-28.
- Bilson, J. F. (1983). The evaluation and use of foreign exchange rate forecasting services. En *Managing Foreign Exchange Rate Risk*. Cambridge University Press.
- Box, G. E., & Jenkins, G. M. (1970). *Time series analysis: forecasting and control*. Holden Day.
- Castle, J. L., Fawcett, N. W., & Hendry, D. F. (2009). Nowcasting is not Just Contemporaneous Forecasting. *National Institute Economic Review*, 210, 71-89. <https://doi.org/10.1177/0027950109354412>
- Castle, J. L., Hendry, D. F., & Clements, M. P. (2019). *Forecasting*. Yale University Press.
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International journal of forecasting*, 5(4), 559-583.
- Clemen, R. T., & Guerard Jr, J. B. (1989). Econometric GNP forecasts: Incremental information relative to naive extrapolation. *International Journal of Forecasting*, 5(3), 417-426.

- Cooper, J. P., & Nelson, C. R. (1975). The ex ante prediction performance of the St. Louis and FRB-MIT-PENN econometric models and some results on composite predictors. *Journal of Money, Credit and Banking*, 7(1), 1-32.
- Crane, D. B., & Crotty, J. R. (1967). A two-stage forecasting model: Exponential smoothing and multiple regression. *Management Science*, 13(8), B-501.
- De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association*, 106(496), 1513-1527.
- Dinartes, O. B. (2009). *Combinación de pronósticos de inflación en Nicaragua* (Documento de Trabajo). Foro de Investigadores de Bancos Centrales del Consejo Monetario Centroamericano.
- Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature*, 46(1), 3-56.
- Engle, R. F., Granger, C. W., & Kraft, D. (1984). Combining competing forecasts of inflation using a bivariate ARCH model. *Journal of economic dynamics and control*, 8(2), 151-165.
- Figlewski, S., & Urich, T. (1983). Optimal aggregation of money supply forecasts: Accuracy, profitability and market efficiency. *The Journal of Finance*, 38(3), 695-710.
- Granger, C. W., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of forecasting*, 3(2), 197-204.
- Hsiao, C., & Wan, S. K. (2014). Is there an optimal forecast combination? *Journal of Econometrics*, 178, 294-309.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- INEGI. (2024). Banco de Información Económica (BIE). <https://www.inegi.org.mx/app/indicadores>
- Jose, V. R. R., & Winkler, R. L. (2008). Simple robust averages of forecasts: Some empirical results. *International Journal of Forecasting*, 24(1), 163-169. <https://doi.org/https://doi.org/10.1016/j.ijforecast.2007.06.001>
- Kapetanios, G., Labhard, V., & Price, S. (2008). Forecast combination and the Bank of England's suite of statistical forecasting models. *Economic Modelling*, 25(4), 772-792.
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International journal of forecasting*, 16(4), 451-476.
- Newbold, P., & Granger, C. W. (1974). Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society: Series A (General)*, 137(2), 131-146.
- Nordhaus, W. D. (1987). Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics*, 667-674.

- Nowotarski, J., Raviv, E., Trück, S., & Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*, 46, 395-412.
- Rapach, D. E., & Strauss, J. K. (2008). Forecasting US employment growth using forecast combining methods. *Journal of Forecasting*, 27(1), 75-93.
- Reid, D. J. (1968). Combining three estimates of gross domestic product. *Economica*, 35(140), 431-444.
- Stock, J. H., & Watson, M. W. (2004a). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, 23(6), 405-430.
- Tapia Alvarado, E. S. (2021). Todo lo que se debe saber sobre la estacionalidad. *Lecturas en Lo que Indican los Indicadores: Cómo utilizar la información estadística para entender la realidad económica de México*, 1, 29-48.
- Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, 1, 135-196.
- Wang, X., Hyndman, R. J., Li, F., & Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, 39(4), 1518-1547.
- Weiss, C. E., Raviv, E., & Roetzer, G. (2018). Forecast Combinations in R using the ForecastComb Package. *R Journal*, 10(2).

Bibliografía

- Agnew, C. E. (1985). Bayesian consensus forecasts of macroeconomic variables. *Journal of Forecasting*, 4(4), 363-376.
- Andrawis, R. R., Atiya, A. F., & El-Shishiny, H. (2011). Combination of long term and short term forecasts, with application to tourism demand forecasting. *International Journal of Forecasting*, 27(3), 870-886.
- Fischer, I., & Harvey, N. (1999). Combining forecasts: What information do judges need to outperform the simple average? *International journal of forecasting*, 15(3), 227-246.
- Galton, F. (1907). Vox Populi. *Nature*, 75, 450-451. <https://doi.org/10.1038/075450a0>
- Hadash, G., Kermany, E., Carmeli, B., Lavi, O., Kour, G., & Jacovi, A. (2018). Estimate and Replace: A Novel Approach to Integrating Deep Neural Networks with Existing Applications [preprint]. ArXiv.
- Hendry, D. F., & Clements, M. P. (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling*, 20(2), 301-329.
- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1), 5-10.
- Kang, H. (1986). Unstable weights in the combination of forecasts. *Management Science*, 32(6), 683-695.
- Kour, G., & Saabne, R. (2014, 11-14 de agosto). *Fast classification of handwritten on-line Arabic characters* [ponencia]. 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR), Túnez. <https://doi.org/10.1109/SOCPAR.2014.7008025>.
- Kour, G., & Saabne, R. (2014, 1-4 de septiembre). *Real-Time Segmentation of On-Line Handwritten Arabic Script* [ponencia]. 14th International Conference on Frontiers in Handwriting Recognition (ICFHR), Hersonissos, Grecia. <https://doi.org/10.1109/ICFHR.2014.76>.
- Lichtendahl Jr, K. C., & Winkler, R. L. (2020). Why do some combinations perform better than others? *International Journal of Forecasting*, 36(1), 142-149.

- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, 1(2), 111-153.
- Palm, F. C., & Zellner, A. (1992). To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting*, 11(8), 687-701.
- Stock, J. H., & Watson, M. W. (2004b). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, 23(6), 405-430.
- Utrera, G. (2005, 16-18 de noviembre). *Una aproximación a la combinación de métodos econométricos para pronosticar la inflación en Argentina* [ponencia]. XL Reunión Anual de la Asociación Argentina de Economía Política, La Plata, Argentina.

Apéndice A

Técnicas de combinación de pronósticos

A.1. Técnicas de combinación de pronósticos

Con el objetivo de establecer un marco notacional claro, definamos la matriz $F_{T \times P}$ que representa los pronósticos, con T indicando la cantidad de periodos temporales y P el número de modelos de pronóstico. Así, disponemos de P pronósticos en cada instante temporal. El pronóstico generado por el modelo i -ésimo se denota como f_i , siendo $i \in \{1, \dots, P\}$. Se prescindirá del subíndice t que designa la dimensión temporal del pronóstico cuando esto no induzca a ambigüedad. En lo que concierne a métodos de combinación que exigen una ordenación de los pronósticos componentes, $f^{(i)}$ simboliza la estadística de orden i -ésima obtenida de la muestra de pronósticos en un punto dado. El peso asignado al pronóstico del modelo i dentro de la combinación final se expresa como w_i , y el pronóstico combinado resultante se designa por f_c .

A.1.1. Promedio Simple

El método más intuitivo para combinar pronósticos es utilizar el promedio de todos ellos. Con el tiempo, este enfoque tan sencillo se ha establecido como un excelente punto de referencia, a pesar o quizás debido a su simplicidad Wang et al. (2023).

El pronóstico combinado se da de forma directa por

$$f^c = \frac{1}{P} \sum_{i=1}^P f_i. \quad (\text{A.1})$$

Este método ha sido denominado en la literatura contemporánea como el rompecabezas

de la combinación de pronósticos”, debido a que, incluso después de más de cincuenta años, sigue siendo la estrategia predominante en este contexto.

A.1.2. Mediana

El método de la mediana al igual que el método anterior resulta ser un método simple y ampliamente utilizado, tiene la cualidad de ser resistente a valores atípicos, lo cual es crucial en ciertas aplicaciones. Stock y Watson (2004a) Sugiere que la justificación de utilizar un método como la media basado en el rango se justifica por su resistencia a valores atípicos a diferencia del pronóstico.

Para el método de la mediana, el pronóstico combinado se calcula de la siguiente manera:

- Si P es impar:

$$f^c = f_{(\frac{P}{2}+0,5)} \quad (A.2)$$

- Si P es par:

$$f^c = \frac{1}{2} \left(f_{(\frac{P}{2})} + f_{(\frac{P}{2}+1)} \right) \quad (A.3)$$

A.1.3. Media Recortada

La media recortada es una técnica reconocida por su resistencia a los valores atípicos, que se emplea comúnmente como medida de localización en situaciones donde los valores extremos pueden distorsionar el resultado. Armstrong (2001) Este método utiliza un factor de recorte λ , lo que significa que se descartan los valores extremos superiores e inferiores, representando el $100 \times \lambda \%$ de los datos, del cálculo del pronóstico combinado.

La fórmula para calcular el pronóstico combinado utilizando la media recortada es la siguiente:

$$f^c = \frac{1}{P(1-2\lambda)} \sum_{i=\lambda P+1}^{(1-\lambda)P} f^{(i)} \quad (A.4)$$

Normalmente, se selecciona $\lambda = 0,1$, indicando que eliminamos el 10% de los pronósticos más extremos de cada extremo. La media recortada sirve como un intermediario entre el promedio simple ($\lambda = 0$) y la mediana ($\lambda = 0,5$).

A.1.4. Media Winsorizada

Al igual que la media recortada, el promedio winsorizado es menos sensible a los valores atípicos en comparación con el promedio simple. Este método adopta un enfoque más suave con los valores extremos: en vez de eliminarlos por completo como se hace en el método de la media recortada, este los limita a un cierto nivel máximo. Al restringir los valores atípicos en lugar de descartarlos, estos conservan alguna influencia en el resultado final. Debido a esto, este método es a menudo preferido por su capacidad de mitigar el impacto de los valores extremos, como sugieren Jose y Winkler (2008).

Definimos λ como el factor de recorte, donde los valores extremos superior e inferior que constituyen el $100 \times \lambda\%$ son modificados. Con un número total $K = \lambda P$ de pronósticos modificados a cada extremo, el pronóstico combinado se determina usando la fórmula:

$$f^c = \frac{1}{P} \left[K f_{(K+1)} + \sum_{i=K+1}^{P-K} f_{(i)} + K f_{(P-K)} \right] \quad (\text{A.5})$$

A.1.5. Bates y Granger (1969)

En el trabajo pionero de Bates y Granger (1969) propusieron una metodología para la combinación de pronósticos basada en la teoría de diversificación de carteras. Esta técnica utiliza los elementos diagonales de la matriz de errores cuadráticos medios estimados para calcular los pesos de combinación de los pronósticos. La fórmula para el pronóstico combinado se expresa como:

$$f^c = \sum_{i=1}^P f'_i \times \frac{\hat{\sigma}^{-2}(i)}{\sum_{j=1}^P \hat{\sigma}^{-2}(j)} \quad (\text{A.6})$$

donde $\hat{\sigma}^2(i)$ es el error cuadrático medio estimado del modelo i . Este enfoque omite la correlación entre los pronósticos de los componentes debido a las dificultades para estimar con precisión la matriz de covarianza.

A.1.6. Newbold y Granger (1974)

Ampliando la investigación inicial de Bates y Granger (1969), Newbold y Granger (1974) desarrollaron una metodología que también extrae los pesos de combinación de la matriz de error de predicción cuadrado medio estimado. Sea Σ la matriz de error de predicción cuadrado medio de $F_{N \times P}$ y e un vector $P \times 1$ de unos $(1, 1, \dots, 1')$. La metodología de

Newbold y Granger (1974) se basa en la minimización restringida del error cuadrático medio de predicción, sujetando a la condición de normalización $e'w = 1$. Esto produce el siguiente pronóstico combinado:

$$f^c = F_{N \times P} \times \frac{\Sigma^{-1}e}{e'\Sigma^{-1}e} \quad (\text{A.7})$$

Aunque el método se originó con Newbold y Granger (1974), la variante utilizada en la presente investigación no impone la restricción previa de que Σ sea diagonal. Esta aproximación, utilizada por Hsiao y Wan (2014), representa una generalización del método original.

A.1.7. Método de Rango Inverso

Propuesto inicialmente por Aiolfi y Timmermann (2006), este método ordena los modelos de pronóstico de acuerdo con su rendimiento hasta el tiempo N . El modelo con el menor error cuadrático medio recibe el primer rango, el segundo menor el segundo rango, y así sucesivamente. El pronóstico combinado se deriva entonces con la siguiente fórmula:

$$f^c = \sum_{i=1}^P f_i \times \frac{\text{Rank}_i^{-1}}{\sum_{j=1}^P \text{Rank}_j^{-1}} \quad (\text{A.8})$$

Según Aiolfi y Timmermann (2006), este procedimiento, similar al de Bates y Granger (1969), no toma en cuenta la correlación entre los errores de los diferentes pronósticos. No obstante, su fortaleza radica en su resistencia a los valores atípicos, pues la estructura de los rangos no se ve significativamente afectada por pronósticos extremos.

A.1.8. Regresión de Mínimos Cuadrados Ordinarios (OLS)

La aplicación de técnicas de regresión para sintetizar pronósticos fue inicialmente propuesta por Crane y Crotty (1967), y refinada más adelante por Granger y Ramanathan (1984). Este enfoque establece que el pronóstico compuesto es una combinación lineal de los pronósticos individuales, donde los pesos se asignan a través de un modelo de regresión sobre la variable objetivo:

$$y = \alpha + \sum_{i=1}^P w_i f_i + \epsilon, \quad (\text{A.9})$$

Parte de los datos de pronóstico se utilizan para calibrar el modelo de regresión, opti-

mizando los coeficientes de OLS mediante la minimización de la suma de cuadrados de los errores. Este proceso resulta en el siguiente pronóstico combinado:

$$f^c = \hat{\alpha} + \sum_{i=1}^P \hat{w}_i f_i, \quad (\text{A.10})$$

Uno de los beneficios destacados de este método es que el pronóstico final es insesgado debido al intercepto en la fórmula, incluso cuando algunos de los pronósticos individuales estén sesgados. Sin embargo, un inconveniente relevante es que no hay restricciones en los pesos de los pronósticos combinados (esto es, no suman necesariamente uno y pueden ser negativos), lo cual puede complicar la interpretación de los resultados, sobre todo cuando los coeficientes derivados no son convexos.

A.1.9. Regresión por Desviación Absoluta Mínima (LAD)

En contraste con la regresión por mínimos cuadrados ordinarios (OLS), que calcula los coeficientes minimizando la suma de los cuadrados de los errores ecuación (A.10), la regresión LAD opta por una función de pérdida alternativa, minimizando la suma absoluta de los errores. Podemos ejemplificar esto considerando un modelo que típicamente tiene un rendimiento aceptable, pero ocasionalmente presenta errores significativos. En LAD, estos errores de mayor magnitud reciben una penalización más severa que en OLS, ya que LAD asigna un peso más elevado a estos desvíos.

Este enfoque de penalización más rigurosa de los errores grandes y en ocasiones resulta y se encuentra más conveniente según las preferencias del analista o el impacto económico de errores pronósticos significativos. Adicionalmente, la regresión LAD ofrece una ventaja de estabilidad en situaciones donde los predictores están fuertemente correlacionados, una situación frecuente en la combinación de pronósticos. Las pequeñas variaciones en los datos no provocan ajustes drásticos en los coeficientes bajo el esquema LAD, lo que puede traducirse en una performance más consistente en aplicaciones prácticas. Por estas razones, la adopción del método LAD se recomienda en escenarios donde los componentes del pronóstico mantienen alta correlación, como sugieren Nowotarski et al. (2014).

A.1.10. Regresión con Restricciones de Mínimos Cuadrados (CLS)

Similar a la aproximación LAD, la regresión CLS aborda el control de los coeficientes extremos por medio de la minimización de la suma de los errores cuadráticos, imponiendo restricciones específicas en los coeficientes. En CLS, se asegura que todos los coeficientes estimados w_i sean no negativos ($w_i \geq 0$) y que su suma totalice uno $\sum_{i=1}^P w_i = 1$.

Para abordar el problema de la multicolinealidad, que puede influir negativamente en las estimaciones de la regresión CLS, se ha implementado el método desarrollado por Babaie-Kafaki y Roozbeh (2017), el cual es una versión modificada de la descomposición de Cholesky basada en la regresión Ridge. Esta versión es fundamental para mitigar los efectos adversos que la multicolinealidad puede ocasionar en la precisión de las estimaciones del modelo.

A.1.11. Método del Vector Propio Estándar

Los métodos que emplean vectores propios para la combinación de pronósticos, ideados por Hsiao y Wan (2014), optimizan el error cuadrático medio de predicción (MSPE) bajo una condición de normalización específica. Tradicionalmente, se requiere que los pesos de los modelos sumen exactamente uno, $\sum_{i=1}^P w_i = 1$, según lo han establecido estudios anteriores como los de Newbold y Granger (1974) y Timmermann (2006). Hsiao y Wan (2014) innovaron al demostrar que adaptando esta condición de normalización se puede obtener un mínimo no restringido del MSPE, formulado como:

$$\sum_{i=1}^P w_i^2 = 1 \quad (\text{A.11})$$

Este enfoque sin restricciones del MSPE se ha adoptado como uno de los principales métodos de combinación en la literatura.

En particular el método de Vector Propio Estándar selecciona los pesos de combinación partiendo de la matriz MSPE estimada: se ordenan los P autovalores positivos ($\Phi_1 = \Phi_{\min}, \Phi_2, \dots, \Phi_P$), y κ_i representa el vector propio asociado a Φ_i . Definimos $d_i = e' \kappa_i$, con e siendo un vector de unos ($P \times 1$). Los pesos se determinan entonces a partir del mínimo de $\left(\frac{\Phi_1}{d_1^2}, \frac{\Phi_2}{d_2^2}, \dots, \frac{\Phi_P}{d_P^2} \right)$, identificado como κ_1 , y se calculan como:

$$w = \frac{1}{d_1} \kappa_1 \quad (\text{A.12})$$

Finalmente, el pronóstico combinado resulta de la suma ponderada:

$$f^c = \sum_{i=1}^P f_i w_i \quad (\text{A.13})$$

A.1.12. Método de Vectores Propios Ajustados por Sesgo

Este método mejora las estimaciones de pronósticos combinados ajustando los sesgos inherentes en los modelos componentes. Modifica el método convencional de vectores propios segmentando los errores de pronóstico en tres categorías: sesgos específicos de cada modelo, factores comunes no considerados por los modelos y elementos idiosincráticos que no se correlacionan entre los modelos.

Para la asignación de pesos en este método mejorado, parte de la matriz MSPE ajustada, que se obtiene tras sustraer las medias de cada conjunto de pronósticos de un modelo particular. Los pesos se determinan con la siguiente expresión:

$$w = \frac{1}{\tilde{d}_1} \tilde{\kappa}_1 \quad (\text{A.14})$$

En esta fórmula, \tilde{d}_i y $\tilde{\kappa}_i$ representan los vectores derivados de la descomposición espectral de la matriz MSPE ajustada, a diferencia de la matriz original. El pronóstico final combinado se calcula como:

$$f^c = \alpha + \sum_{i=1}^P f_i w_i \quad (\text{A.15})$$

Donde α sirve para corregir cualquier sesgo residual en los pronósticos combinados.

A.1.13. Método de Vectores Propios Recortados

El método convencional de vectores propios puede enfrentar desafíos significativos debido a la variabilidad en el rendimiento de los modelos componentes. En particular, el rendimiento de este método puede degradarse considerablemente por los modelos que arrojan resultados deficientes, ya que trata de forma simétrica las incertidumbres en las series de tiempo reales, $F_{N \times P}$. Una discusión detallada sobre este principio de ortogonalidad puede encontrarse en la sección correspondiente del trabajo de Hsiao y Wan (2014). El enfoque Recortados de vectores propios se propone como una solución a esta vulnerabilidad.

Este ajuste, originado en las ideas de Timmermann (2006) y extendido por Hsiao y Wan (2014), emplea una técnica de poda en la selección de los modelos de pronóstico. Al aplicar este método, los pesos se derivan igual que en el método estándar pero utilizando la matriz MSPE ajustada, que excluye los modelos con rendimiento inferior.

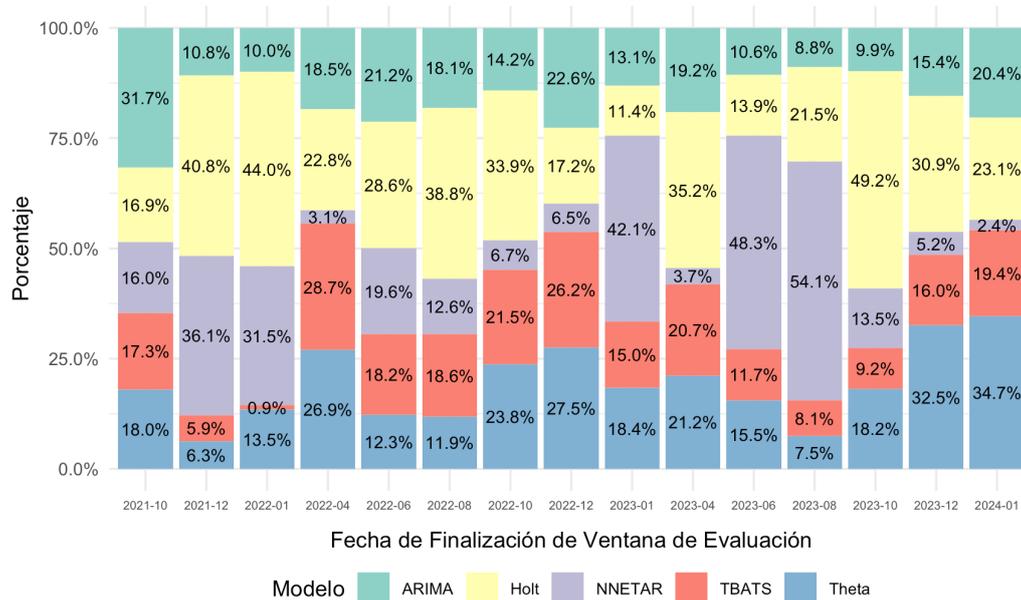
A.1.14. Enfoque de Vectores Propios con Corrección de Sesgo y Recorte

La metodología subyacente del enfoque de vectores propios recortado y corregido por sesgo sigue los principios del enfoque corregido por sesgo. En este método, los pesos se derivan a través de la descomposición espectral de la matriz MSPE centrada. La principal distinción de este método respecto al enfoque solo corregido por sesgo radica en su estrategia de selección previa de modelos: se elige un subconjunto de modelos de pronóstico basándose en su rendimiento, descartando aquellos con los resultados más desfavorables. Esto permite que el método combine los aspectos beneficiosos de las correcciones de sesgo y el recorte de modelos, optimizando así la precisión y la robustez de los pronósticos combinados.

Apéndice B

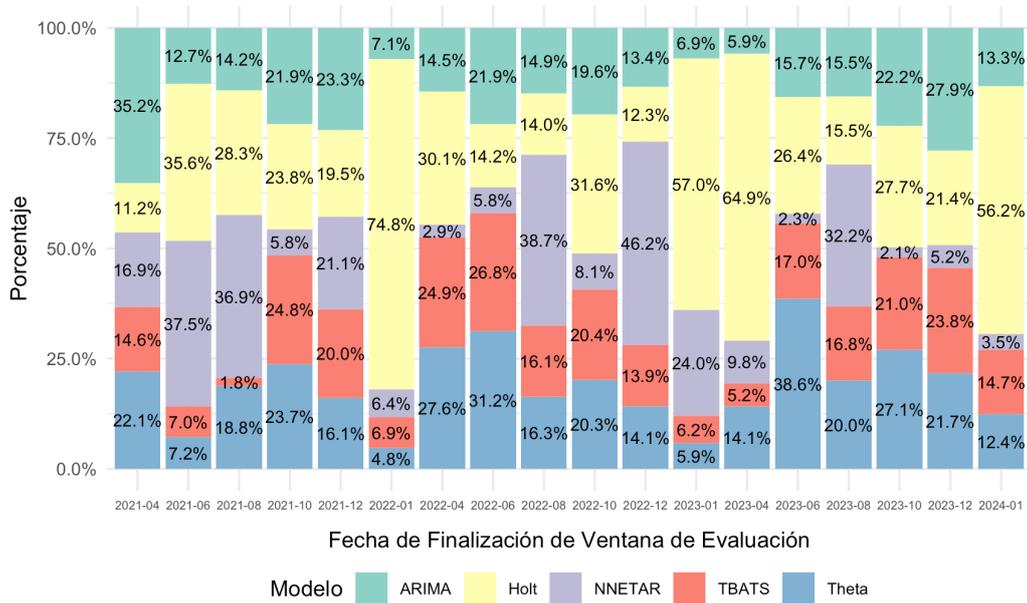
Desempeño de los Modelos individuales en Estabilidad Secuencial

Figura B.1. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (150), Validación (9) y Evaluación (9), Desplazamiento (2).



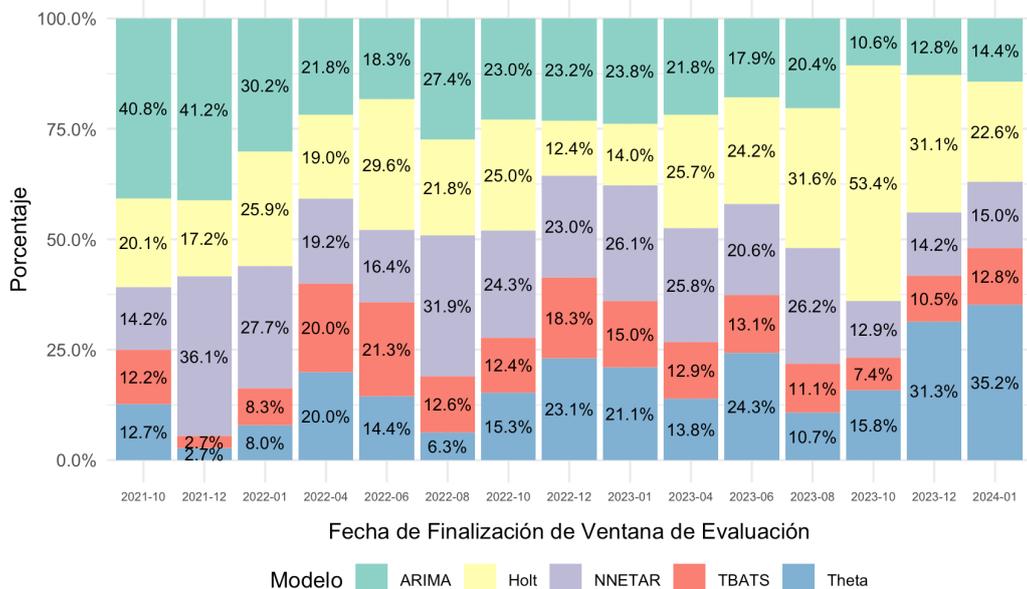
Fuente: elaboración propia con datos del INEGI.

Figura B.2. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (150), Validación (6) y Evaluación (6), Desplazamiento (2).



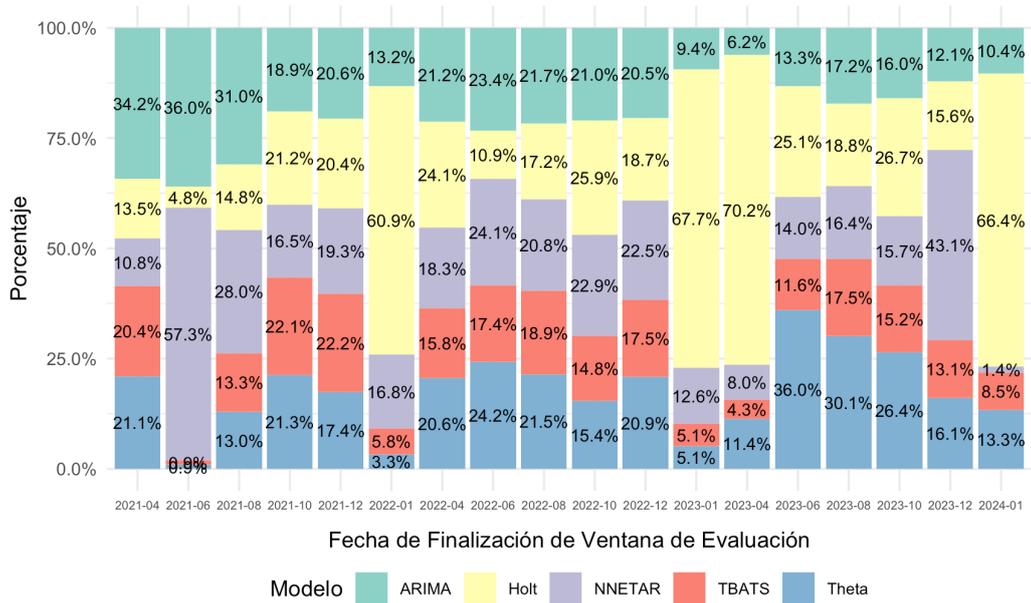
Fuente: elaboración propia con datos del INEGI.

Figura B.3. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (100), Validación (9) y Evaluación (9), Desplazamiento (2).



Fuente: elaboración propia con datos del INEGI.

Figura B.4. Histograma de Estabilidad Secuencial para el IGAE: Pesos de los Modelos en la Combinación Bates/Granger, para 12 Iteraciones con Ventanas de Entrenamiento (100), Validación (9) y Evaluación (9), Desplazamiento (2).

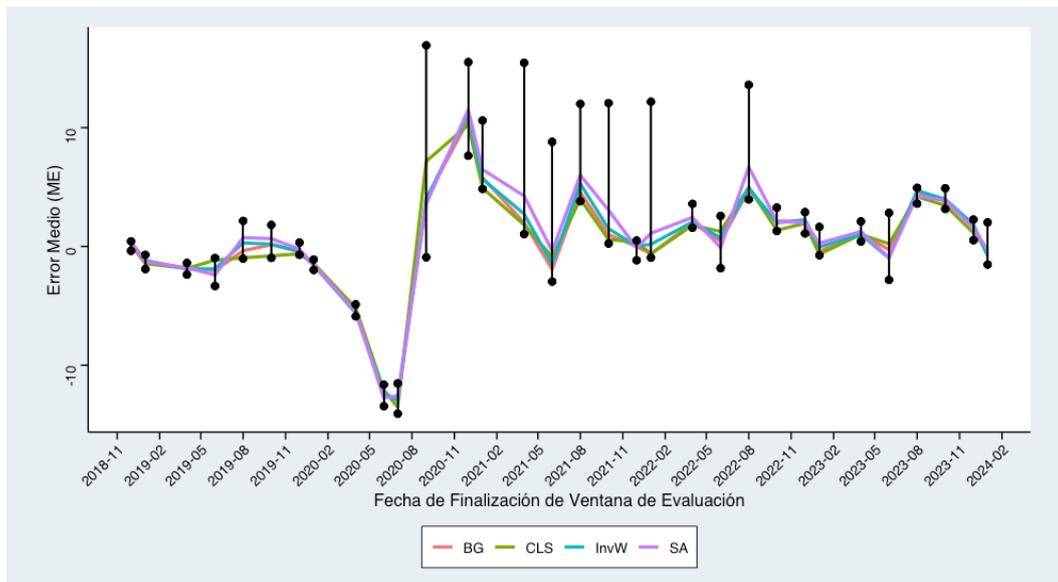


Fuente: elaboración propia con datos del INEGI.

Apéndice C

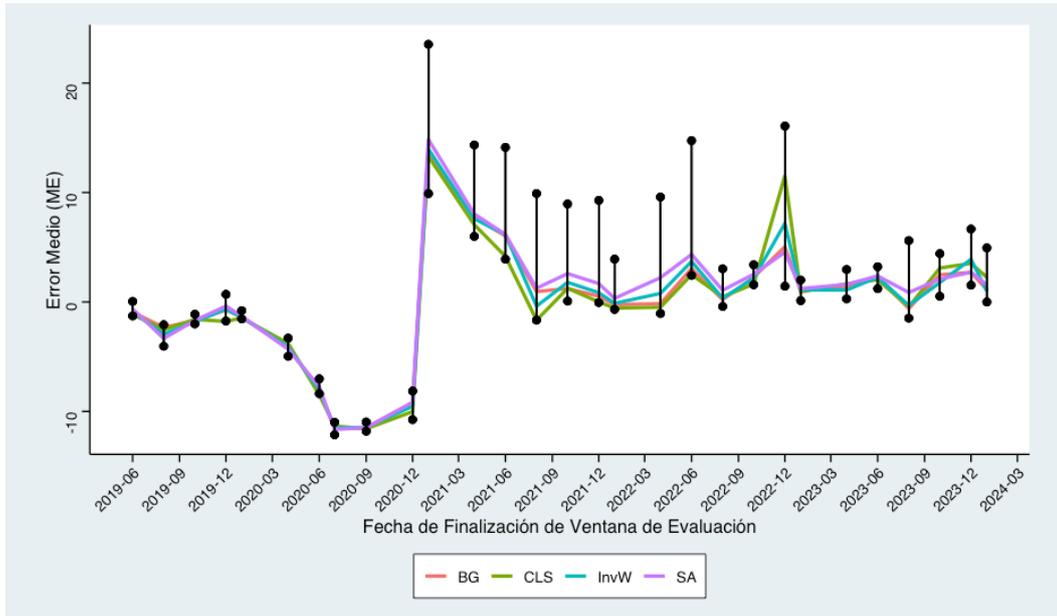
Combinación de pronósticos

Figura C.1. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



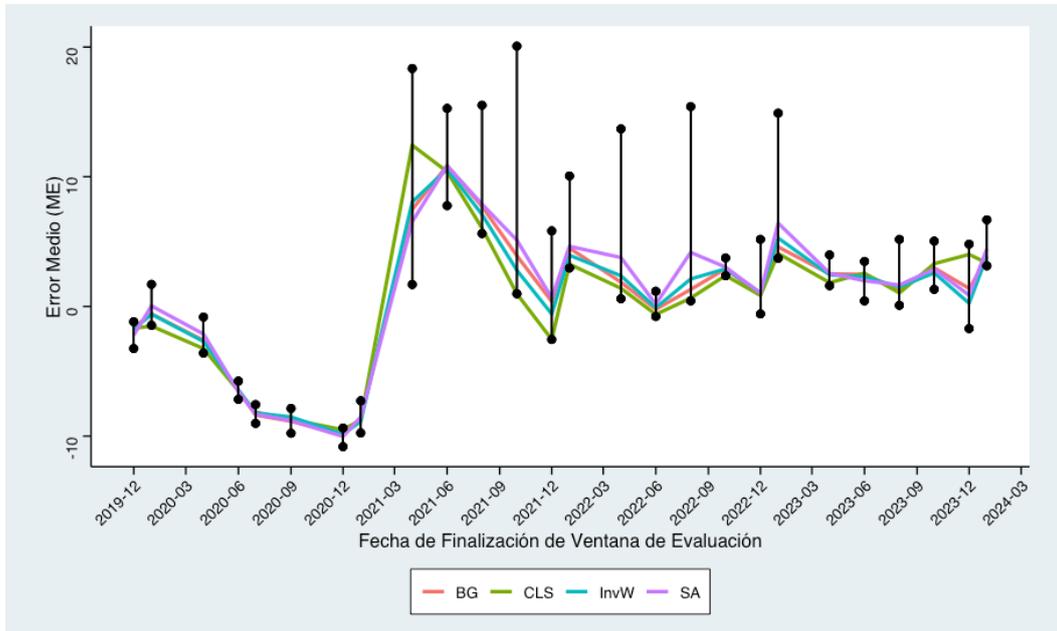
Fuente: elaboración propia con datos del INEGI.

Figura C.2. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)



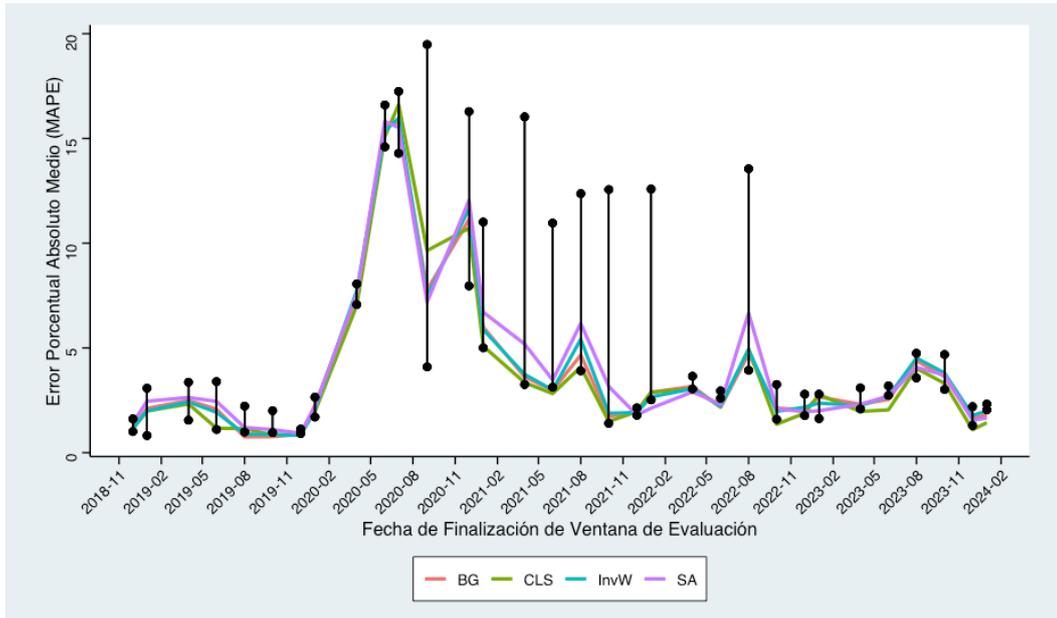
Fuente: elaboración propia con datos del INEGI.

Figura C.3. Comparativa de Desempeño de Modelos Individuales y Combinados (ME); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



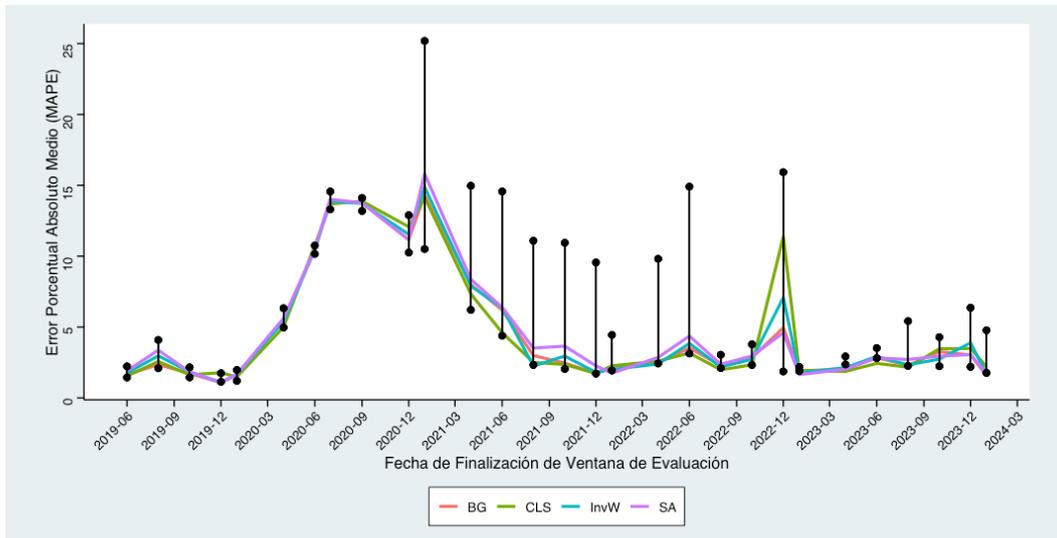
Fuente: elaboración propia con datos del INEGI.

Figura C.4. Comparativa de Desempeño de Modelos Individuales y Combinados (MAPE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)



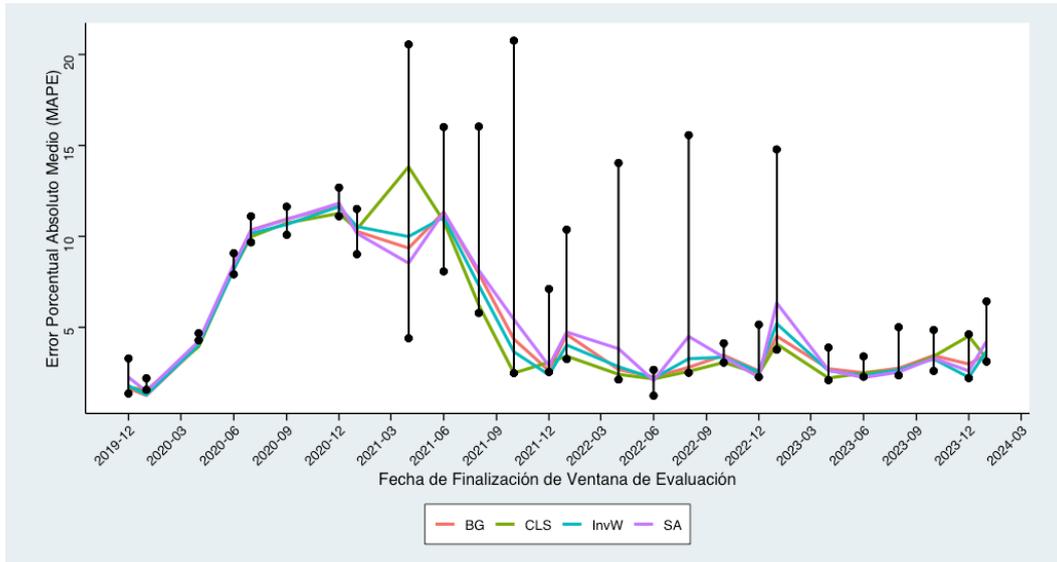
Fuente: elaboración propia con datos del INEGI.

Figura C.5. Comparativa de Desempeño de Modelos Individuales y Combinados (MAPE); ventana de entrenamiento (300), ventana de validación (9) y ventana de evaluación (9)



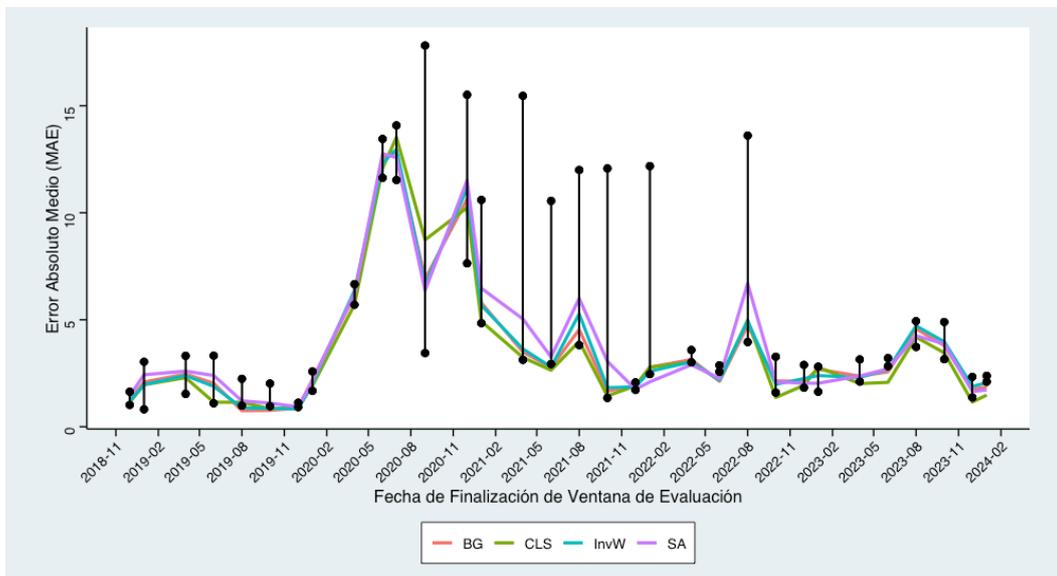
Fuente: elaboración propia con datos del INEGI.

Figura C.6. Comparativa de Desempeño de Modelos Individuales y Combinados (MAPE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



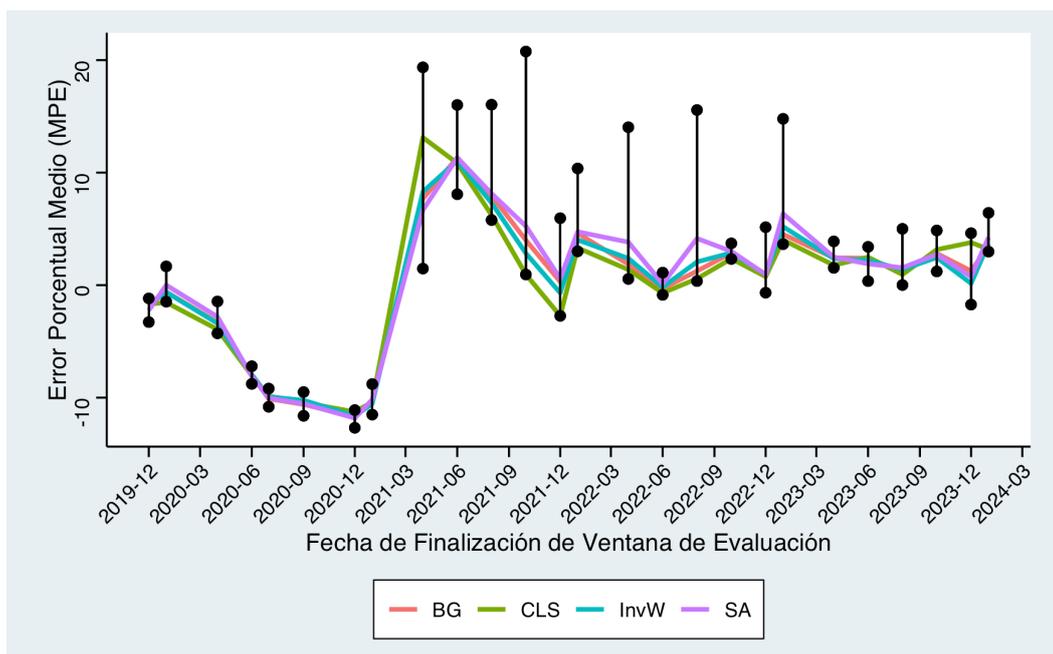
Fuente: elaboración propia con datos del INEGI.

Figura C.7. Comparativa de Desempeño de Modelos Individuales y Combinados (MAE); ventana de entrenamiento (300), ventana de validación (6) y ventana de evaluación (6)



Fuente: elaboración propia con datos del INEGI.

Figura C.8. Comparativa de Desempeño de Modelos Individuales y Combinados (MPE); ventana de entrenamiento (300), ventana de validación (12) y ventana de evaluación (12)



Fuente: elaboración propia con datos del INEGI.

Cuadro C.1. MAE de diferentes métodos de combinación de pronósticos en iteraciones sucesivas ventana de validación (9) y evaluación (9)

Iteration	BG	CLS	InvW	SA
1	1.633478	1.539024	1.748252	1.898869
2	2.334455	2.549412	2.944024	3.336312
3	1.697253	1.618168	1.808655	1.800465
4	1.080826	1.777150	1.105545	1.069177
5	1.583557	1.462636	1.615342	1.637483
6	4.558396	4.095999	4.460935	4.679935
7	8.485777	8.702651	8.448232	8.513058
8	11.582326	11.336502	11.532973	11.649922
9	11.487588	11.590996	11.449654	11.492539
10	9.186404	10.018473	9.534502	9.179803
11	13.440536	13.296877	13.971742	14.837316
12	7.689177	7.060043	7.596478	8.039670
13	6.045204	4.492110	6.201551	6.234150
14	2.856753	2.349797	2.126345	3.359193
15	2.360919	2.276256	2.839721	3.520587
16	1.706237	1.672828	1.765072	2.215846
17	2.034191	2.185809	1.942286	1.697186
18	2.431084	2.515008	2.353034	2.828324
19	3.521494	3.155738	3.854093	4.370358
20	2.234900	1.949119	2.138955	2.350066
21	2.849974	2.330381	2.728313	2.954387
22	5.041654	11.516686	7.175368	4.683412
23	1.936540	1.976243	1.814320	1.689842
24	2.121543	1.917029	2.187496	2.080654
25	2.938505	2.523898	3.025620	2.927847
26	2.315688	2.184146	2.353466	2.774043
27	3.370687	3.596095	2.826698	3.047295
28	3.180296	3.660137	4.083666	3.207110
29	1.664530	2.296943	1.740827	1.787587

Fuente: elaboración propia con datos del INEGI.

Cuadro C.2. MAE de diferentes métodos de combinación de pronósticos en iteraciones sucesivas ventana de validación (12) y evaluación (12)

Iteration	BG	CLS	InvW	SA
1	1.869345	1.805653	2.025978	2.205119
2	1.254061	1.544677	1.228293	1.451966
3	3.322811	3.407898	3.259523	3.553125
4	6.701522	6.539660	6.770626	6.877708
5	8.411554	8.595915	8.520458	8.625327
6	9.148683	9.184053	9.471216	9.282293
7	9.728173	9.704017	9.744105	9.991338
8	8.624577	8.838089	8.859278	8.512683
9	8.434024	14.308491	8.619529	8.136161
10	11.731316	10.391332	10.978562	11.152869
11	8.058896	6.041049	7.206629	7.813110
12	4.237741	2.391288	3.501548	5.316330
13	2.618752	2.897060	2.284206	2.589529
14	4.425904	3.352679	4.183217	4.985671
15	2.568513	2.374380	2.838386	3.478934
16	2.196681	1.926441	2.151622	2.061268
17	2.727214	2.533376	3.292056	4.224118
18	3.455382	3.063016	3.329722	3.474513
19	2.603041	1.982611	2.578417	2.364772
20	4.605594	4.151116	5.141066	6.579033
21	2.782633	2.099559	2.694807	2.687588
22	2.481264	2.810803	2.562906	2.307506
23	2.819282	2.702833	2.794366	2.676444
24	3.393080	3.635691	3.258443	3.322351
25	3.006521	3.695257	2.343586	2.817543
26	3.754586	3.431773	3.973526	4.853707

Fuente: elaboración propia con datos del INEGI.

Apéndice D

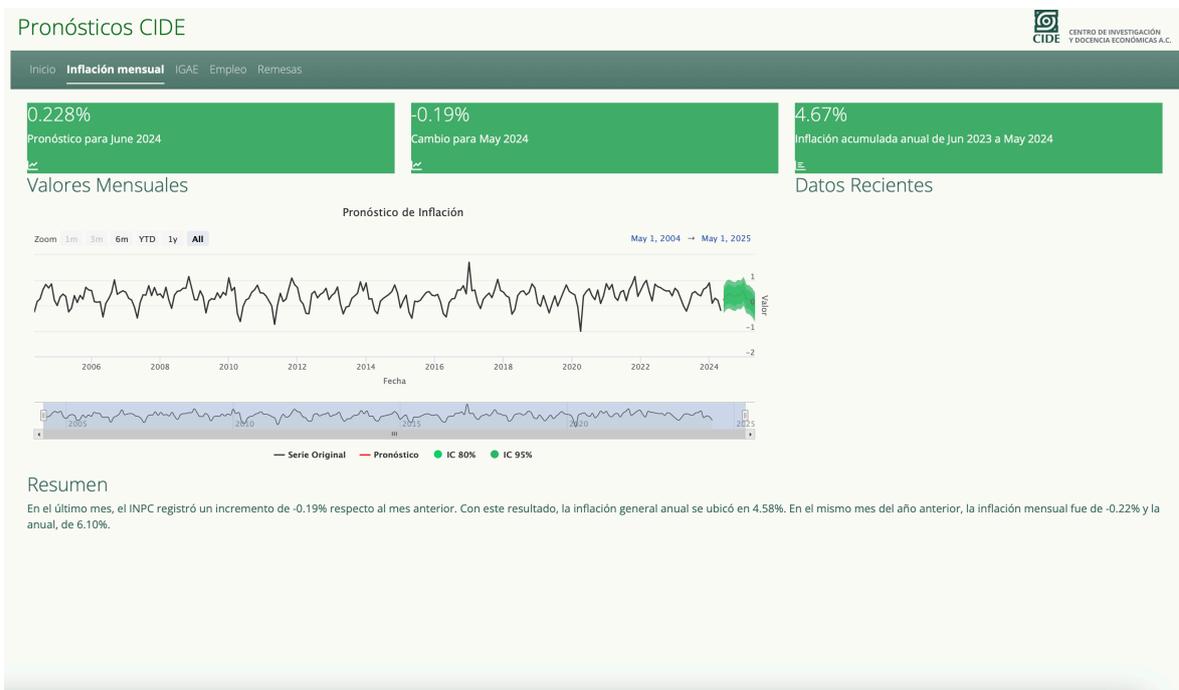
Desarrollo de plataforma Shiny

Figura D.1. Portada de la Aplicación Shiny



Fuente: elaboración propia con datos del INEGI.

Figura D.2. Pronóstico de la Inflación en la Aplicación Shiny



Fuente: elaboración propia con datos del INEGI.