

# Modelos Predictivos en Series de Tiempo

Informe Practica Profesional I

Estudiante:	Jorge Novoa Contreras
Empresa:	RialStat SPA
Supervisor:	Ricardo Pérez Sáez
Cargo Supervisor:	Director Ejecutivo
Periodo Practica:	Inicio: 9 de Marzo 2023 Termino: 9 de Febrero 2023

# Resumen

En este informe se presenta lo realizado durante la práctica profesional en la empresa RialStat, en el sector de Data Science.

Los objetivos principales de esta practica fueron:

- Revisar y Validar Bases de Datos
- Desarrollar Modelos Estadísticos y de proyección Macroeconómica

En particular, se crearon e incorporaron múltiples modelos predictivos para la predicción a corto plazo de agentes importantes dentro de la economía del país.

La Metodología para revisar y validar las Bases de Datos se ocupo tanto fuentes oficiales (como el Banco Central, Cochilco, INE, etc) y para manipular los datos se ocupo Excel.

La Metodología para el desarrollo de los modelos Estadísticos y Predictivos consistió en ocupar el lenguaje de programación R, en particular se ocupo la interfaz de RStudio y Google Colab, junto con un estudio previo sobre análisis descriptivo de series de tiempo, modelos predictivos clasicos y modelos de Machine Learning en el contexto de las series de tiempo.

Como resultados generales, se encontró que hay modelos (como el ARIMA y el PROPHET) de fácil implementación y de buena calidad de predicción (en el sentido de comparar predicción vs datos reales mediante backtesting), junto con métodos de filtración de datos para la mejora en la calidad de predicción del modelo.

Se concluyo que esta experiencia de practica fue una buena introduccion al mundo de la ciencia de datos, al aprender y profundizar sobre series de tiempo, modelos predictivos, entre otros, los cuales son elementos importantes dentro de la ciencia de datos guiada a la predicción, además, se señalaron limitaciones dentro de la practica relacionadas con la formación teórica/practica dentro de la carrera.

# Índice de Contenidos

<b>1. Contexto</b>	<b>1</b>
<b>2. Objetivos y Metodología</b>	<b>2</b>
<b>3. Desarrollo</b>	<b>4</b>
<b>4. Conclusiones</b>	<b>7</b>
<b>Referencias</b>	<b>8</b>
<b>Anexo A. Ecuaciones de los Modelos Predictivos</b>	<b>9</b>
A.1. ARMA . . . . .	9
A.2. ARIMA . . . . .	9
A.3. ARMAX . . . . .	9
A.4. ARIMAX . . . . .	9
A.5. PROPHET . . . . .	10
<b>Anexo B. Imágenes</b>	<b>11</b>

# Índice de Figuras

B.1. Parte de la Base de Datos utilizada durante la práctica. . . . .	11
B.2. ECM del Modelo para la serie del Cobre. . . . .	11
B.3. ECM del Modelo para la serie del Petroleo BRENT. . . . .	12
B.4. ECM del IMACEC. . . . .	13
B.5. ECM del Índice de Desocupación. . . . .	14
B.6. ECM del IRR. . . . .	15
B.7. Puntos de cambio estructural para la Serie del IMACEC. . . . .	16
B.8. ECM al considerar el cambio estructural vs no considerar el cambio estructural. . . . .	17
B.9. Predicción del Índice de Desocupación vs los valores reales según el mes. . . . .	18
B.10. ECM de múltiples modelos para la serie del Euro. . . . .	19
B.11. ECM de múltiples modelos para la serie de la Libra. . . . .	20
B.12. ECM de múltiples modelos para la serie del Real. . . . .	21
B.13. ECM de múltiples modelos para la serie del Yen. . . . .	22
B.14. ECM de múltiples modelos para la serie del Yuan. . . . .	23
B.15. ECM de múltiples modelos para la serie del Dolar . . . . .	24
B.16. Modelo MARS junto con su predicción. . . . .	25
B.17. ECM de modelo LSTM para el Dolar. . . . .	26
B.18. ECM de modelo Prophet + covariables para el Dolar. . . . .	27
B.19. ECM de modelo ARIMAX para el Dolar. . . . .	28
B.20. ECM de modelo Prophet + covariables y considerando el cambio estructural para el Dolar. . . . .	29

# 1. Contexto

RIALSTAT SPA esta integrado por profesionales con más de 30 años de experiencia en el análisis y desarrollo de Modelos Matemáticos Estadísticos.

Su misión de empresa es brindar un servicio altamente especializado, en particular, en el ámbito de riesgo financiero y predictivo.

En particular, se desarrollan y construyen Modelos Matemáticos Estadísticos de tipo:

- Modelos para el apoyo a la Gestión de Negocios
- Modelos Predictivos.
- Modelos de Provisiones
- Construcción de Score (Comportamiento, Pago, Admisión, entre otros)

Muchos de los clientes que acuden a RIALSTAT SPA buscan asesoría en el ámbito de la predicción o expectativas futuras para anticipar variables claves relacionadas con la gestión comercial de su empresa, como lo son sus ventas, riesgo, precios, entre otros, con el objetivo de diseñar estrategias comerciales conociendo el comportamiento que podrían tener estas variables bajo distintos escenarios futuros.

Para modelar y predecir de buena forma, es necesario considerar el contexto de la empresa y los factores que influyen en esta, así, como es conocido, la macroeconomía del país es uno de estos factores, esto ya que las variables macroeconómicas pueden llegar a influir, por ejemplo, en si al cliente le es asequible algún producto, o si la empresa puede seguir con los mismos costos para generar algún producto y mantener la ganancia, etc.

En resumen, se busca conocer cómo se debe actuar para maximizar la ganancia y minimizar el riesgo, y para esto, se ocupan pronósticos de diferentes variables, como pueden ser variables propias de la empresa, indicadores más especializados sobre el comportamiento de sus clientes, indicadores de gestión (variables internas), indicadores que entreguen información sobre la competencia (entorno competitivo), variables macroeconómicas (entorno macroeconómico), las cuales mencionamos anteriormente y, en particular, se trabajaran a lo largo de la práctica.

Cuando las variables que se buscan “anticipar” (o predecir) son datos cuantitativos, podemos formalizar el problema a un contexto matemático.

En específico, como los datos (las observaciones) son dependientes del tiempo, el problema se transforma en uno de Series de Tiempo, y para la predicción de esto se tienen los conocidos Modelos Predictivos para Series Temporales.

Así, con todo lo explicado anteriormente (bajo que contexto surge y como se evaluara) tenemos nuestro problema a tratar durante la práctica.

## 2. Objetivos y Metodología

Antes de empezar la práctica, se acordaron los siguientes objetivos a nivel general:

- Revisar y Validar Bases de Datos
- Desarrollar Modelos Estadísticos y de proyección Macroeconómica

Considerando que la empresa es una consultoría, lo solicitado por los clientes puede variar, pero en general todo se mantuvo dentro de esas directrices.

Como objetivos particulares para la revisión y validación de bases de datos, se solicito que:

- Ocupar fuentes confiables para la creación y actualización de la Base de Datos.
- Implementación correcta de los datos en Excel.

Como objetivos particulares para los modelos predictivos (los cuales fueron predominantes dentro de la práctica) fueron:

- Investigar sobre series de tiempo, métodos de análisis para estas y como implementarlas en R.
- Crear modelos consistentes y de buen ajuste (en el sentido de modelos de regresión)
- Que cada modelo tenga un buen nivel de predicción

La metodología con respecto a la Base de Datos consistió en, según lo solicitado, ocupar fuentes como la base de datos del Banco Central [1], INE [2], Cochilco [3], entre otros, he incorporarlos en Excel.

Para los modelos predictivos, la metodología para investigar sobre la teoría de las series de tiempo fue el recurrir al libro “ Introduction to Time Series and Forecasting’ ’ [4] de Peter J. Brockwell y Richard A. Davis, y para la implementación de estas en R se recurrió a múltiples artículos de la plataforma RPubls , según el tipo de modelo que se quisiera implementar para la serie.

Para cada serie de tiempo, el primer procedimiento consistía en descomponer la serie según su Tendencia, Estacionalidad y Aleatoriedad. Pese a si se observa alguna marcada tendencia o no, se hizo uso de los test de raíces unitarias para verificar si la serie es estacionaria o no, y dependiendo de esto se realizaba un numero de diferenciaciones hasta que la serie fuese estacionaria (buscando siempre no sobrediferenciar).

Posterior a esto, se aplicaba el método de Box - Jenkins [5] para decidir un modelo ARMA o ARIMA (dependiendo del caso), este consiste en:

1. **Selección** : Se gráfica la autocorrelacion (ACF) y la autocorrelacion parcial (PACF), y dependiendo de la cantidad y el comportamiento de aquellos lag que estaban sobre el intervalo de significancia, se seleccionaba una cantidad de medias móviles y autoregresivos.
2. **Estimación de Parámetros** : Ocupar algoritmos o funciones integradas para seleccionar los coeficientes que mejor ajustan el modelo a la serie (siempre se ocupo funciones integradas para esto)

3. **Verificación** : Si el modelo se ajusto bien, debería cumplir que sus residuales sean constantes en media y en varianza. Además, se realiza el test de Ljung - Box, el cual sirve para ver la autocorrelacion serial.

Tanto el modelo ARMA como ARIMA consisten en un modelo de regresión que tiene dos partes:

- AR : la parte autoegresiva, refiriéndose a como influyen las observaciones pasadas en la variable de interés.
- MA : la parte de media móvil, refiriéndose a como el variable de interés es linealmente dependiente del error pasado y actual.

Posterior a eso, se agrego una segunda instancia para la verificación de la calidad del modelo, esta consistió en realizar backtesting. Para esto se uso una serie con menos datos (quitando una cantidad definida de datos mas recientes) y a partir de ahí ocupar el modelo creado anteriormente para predecir los valores que se eliminaron, y, dependiendo de como es el ECM según los datos predichos vs los datos reales, se acepta el modelo o se rechaza.

Si es que el modelo se rechaza por un alto ECM, entonces se intenta incorporar otros tipos de modelos, en particular, los siguientes:

1. AUTO ARIMA
2. ARIMAX
3. PROPHET

El modelo AUTO ARIMA [6], mas que un modelo, corresponde a un algoritmo integrado en la librería forecast, el cual permite generar un modelo ARIMA de forma automatizada, pese a esto, se encontró que, en casi todos los modelos predictivos que se realizaron durante la práctica, el ocupar Auto Arima empeoró el ECM.

El modelo ARIMAX [7] corresponde a un modelo ARIMA en el cual se le agrega alguna covariable dependiente del tiempo (como puede ser otra serie), la idea de la prediccion de este modelo es que se asume que se conoce el "futuro" de esta covariable, por lo que previamente se deben predecir los valores de esta covariable.

El modelo PROPHET [8] es un modelo de fuente libre de Facebook, el cual consiste en un modelo de regresión aditivo con cuatro componentes principales:

- Una curva de crecimiento de la tendencia linear o logística por partes.
- Un componente estacional anual
- Un componente estacional semanal (dependiendo del caso)
- Una lista de días festivos (esto debido al impacto que generan en ciertas áreas de la economía)

Y en los casos en los que no se logra reducir el ECM, se recurre a filtrar cambios estructurales mas recientes, ocupar covariables (las cuales para identificarlas se ocupo test de Correlación de Pearson y Random Forest, ambos ya incorporados en librerías de R) y recurrir a modelos mas complejos.

### 3. Desarrollo

Antes de empezar con esta sección, es necesario señalar que por cada modelo predictivo se realizó un reporte/informe, por lo que estos se adjuntaron en la bibliografía.

Inicialmente, se crea una base de datos de frecuencia mensual partiendo desde Enero 2010 hasta Diciembre 2022, en la cual, se agregaron múltiples variables, como lo son el dolar, paridades de monedas de alta relevancia, precio del Dolar, IPC (General, Transable y No Transable), Cobre, IMACEC, Índice de Desocupación, Stock del Cobre, BRENT, WTI, entre otros. Además, de vez en cuando se solicitaba actualizar con nuevas variables esta base de datos, llegando a sí a un total de 28 variables, en la Figura B.1 se puede observar algunas de estas variables.

Luego, se creó un modelo predictivo simple para el cobre y se realizó un reporte del mismo [9], en este caso el procedimiento fue igual al señalado en la metodología, terminando con un Modelo ARIMA, pese a esto, el ECM del modelo no era lo suficientemente bueno (como se puede observar en la Figura B.2), además, se observó que al aumentar la cantidad de datos el ECM (predicción vs dato real) tenía una tendencia decreciente, aunque en alguno de estos test (en el contexto de predecir con datos limitados) se mostró un leve aumento en el ECM y luego una bajada, lo que sugirió que quizás el modelo no estaba lo suficientemente bien ajustado (pese a los resultados observados por el Ljung - Box test) o que los datos que se agregaron en los test donde aumentó el ECM estaban alterados y/o faltó limpiarlos.

En esa misma semana se prosiguió con la creación de un modelo predictivo simple para la serie del Petróleo BRENT (junto con su respectivo reporte [10]), la metodología fue exactamente la misma que la del modelo predictivo del cobre. Pese a que, como se muestra en la Figura B.3, el ECM en general disminuyó en comparación al modelo del cobre, no se observó una tendencia decreciente del ECM al aumentar la cantidad de datos, la única explicación que se encontró para eso (aparte del poco ajuste del modelo, lo cual nuevamente contradice a los resultados del Ljung - Box test) fue la falta de significancia de los datos agregados en comparación a los datos de base.

La siguiente tarea consistió en crear 3 modelos predictivos, para el IMACEC, el Índice de Desocupación y el IRR, junto con un único reporte para los 3 modelos [11]. Aplicando la misma metodología de antes, se obtuvieron los ECM mostrados en las Figuras B.4, B.5 y B.6.

Adicionalmente se encontraron cambios estructurales para la serie del IMACEC (Figura B.7), y se eliminaron los datos previos a ese cambio estructural, esto con el motivo de mejorar el ECM, se obtuvo que el ECM aumentó (Figura B.8), por lo que se concluyó que los datos eliminados sí tenían importancia, adicionalmente, se encontró que una de las soluciones sería el ocupar alguna función de peso para disminuir la importancia de los datos previos al cambio estructural, generando así que aporten a la predicción (cabe señalar que esta idea no se llevó a cabo).

Posterior a esto, se solicitó crear un modelo predictivo para el Índice de Desocupación (Desempleo) ocupando como covariable al IMACEC (Índice Mensual de Actividad Económica) junto con un reporte de lo realizado [12], con esto, se procedió directamente con un modelo ARIMAX, ocupando nuevamente la metodología de Box - Jenkins para definir la parte ARIMA del modelo.

Uno de los resultados interesantes al realizar backtesting (Figura B.9) para este modelo es que los valores eran bastante cercanos excepto cuando se agregaron los datos de Julio, Agosto y Septiembre. Y la razón a la que fue atribuido este resultado fue la misma que para los 3 modelos anteriores, la importancia de los datos referentes a esos meses.

Una de las tareas finales mas importantes y complicadas a nivel tanto técnico como teórico fue la creación de modelos predictivos para las paridades de divisas y el dolar.

Cabe señalar que toda la experiencia se encuentra a detalle en el reporte correspondiente [13].

Al realizar los modelos y el backtesting, se llego, en el caso de las series de paridades de monedas, hasta el modelo PROPHET (el cual es el ultimo en la metodología a seguir) debido a que los ECM al probar todos los modelos (Figuras B.10, B.11, B.12, B.13 y B.14) no disminuían como se quería, cabe señalar que, en términos de ECM relativo a los otros modelos, el ECM del modelo PROPHET resulto ser mucho menor que el ECM de los otros modelos ( $\approx 10\%$  del ECM de los otros), y por lo mismo no se muestra en los gráficos. Producto de todo lo explicado anteriormente se eligió el modelo PROPHET para cada una de las paridades de divisas.

En el caso del modelo predictivo para el dolar se consideraron mas modelos y posibles factores, pese a que con el modelo PROPHET se encontro el menor ECM en comparacion al modelo ARIMA, AUTO ARIMA , ARIMAX , etc, (como se observa en la Figura B.15, donde nuevamente se cumple que el modelo PROPHET  $\approx 10\%$  del ECM de los otros modelos) el ECM seguía siendo muy grande en comparación a la magnitud de los valores observados.

Uno de los modelos “nuevos” con los que se intento mejorar el ECM fue el MARS (*Multivariate adaptive regression spline*) [14], el cual, pese a tener un excelente ajuste (esto bajo los test como Ljung - Box y funciones en R que miden esto) su predicción fue contradictoria (como se muestra en la Figura B.16)

Luego , se ocupo un modelo ETS y SARIMA, pero aun así estos empeoraban el ECM en comparación al modelo ARIMA, por lo que se decidió ocupar el modelo PROPHET.

Como el modelo predictivo del dolar fue malo (en el sentido de calidad de predicción bajo backtesting) se realizo nuevamente (junto con un nuevo reporte [15]), considerando ahora los siguientes otros modelos:

- **LSTM** (*Long Short Term Memory*)
- **PROPHET + variables**
- **ARIMAX + nuevas variables**
- **PROPHET + variables + cambio estructural**

Nuevamente todos estos modelos aseguraban un buen ajuste, pero el ECM al realizar backtesting seguía siendo muy alto (como se muestran en las Figuras B.17, B.18, B.19 y B.20).

El modelo que obtuvo una mejor calidad de predicción fue el modelo PROPHET al considerar variables externas y cambio estructural, la diferencia frente al modelo PROPHET solo con variables externas no fue tanta, pero mejora considerablemente en comparación al ECM del modelo PROPHET simple para el dolar, aunque, el ECM seguía siendo alto.



Posteriormente, se intento encontrar (teóricamente) cuales serias las cosas a cambiar para mejorar el modelo, refiriéndonos a su capacidad predictiva, estas fueron las siguientes:

- Encontrar y filtrar de mejor manera los cambios estructurales: esto es indicado gracias a la tendencia creciente del ECM en algunos modelos al agregar datos.
- Agregar mas variables: las variables que se ocuparon fueron resultado de los test para ver correlación de series, el problema es que teníamos una cantidad limitada de variables con las que trabajar, por lo que no se sabe con certeza si es que el existía alguna otra variable la cual pudiese mejorar el modelo.
- Ocupar modelos mas complejos: durante la realización del ultimo modelo predictivo para el dolar, se intento implementar un modelo ANN y un modelo GARCH, pero la falta de conocimiento sobre redes neuronales y como implementar modelos así de complejos en series de tiempo termino con la no incorporación de un modelo ANN y GARCH.

Una de las hipótesis que se mantuvieron al momento de observar que los ECM de los modelos para la serie del Dolar seguían siendo altos fue la poca/nula capacidad de los modelos ocupados para capturar la volatilidad de la serie [16], y, como es conocido, durante los últimos 3 años el comportamiento del dolar ha sido bastante volátil.

Por lo explicado en el párrafo anterior fue que se intento enfocar en aplicar un Modelo GARCH, ya que este tipo de modelo son conocidos por su capacidad de capturar la volatilidad [16].

## 4. Conclusiones

Mientras se realizó la práctica Profesional en la empresa RialStat, se desarrollaron varias habilidades, destacando las de trabajo en equipo, y las habilidades analíticas, siendo las de trabajo en equipo las utilizadas para coordinar y distribuir las labores, y las habilidades analíticas las relacionadas al realizar la selección de los modelos a ocupar según el contexto.

Uno de los logros principales de la práctica fue el entender más a fondo los modelos de regresión, los tipos de modelos y en general los criterios que se ocupan para la selección del modelo y como verificar la calidad del modelo, y en particular, aprender sobre modelos ARIMA o PROPHET, ambos los cuales destacaron dentro de los modelos predictivos que se realizaron durante la práctica debido a su facilidad de implementación junto con la calidad de predicción.

Sobre las series de tiempo, uno de los logros más importantes fue entender parte de la teoría de esta, como verificar cuando hay que eliminar o dar menos importancia a algunos datos según múltiples factores y, en general, la relación de las series de tiempo con los modelos de proyección económica.

Relacionado al último punto, se destaca el llegar a entender las variables dentro de economía del país (desde un punto de vista macroeconómico) y como estas se relacionan unas con otras, además de saber que herramientas matemáticas existen para ver esta relación.

Por último, se destaca el haber aprendido a programar en R, habilidad la cual es una de las 3 más solicitadas dentro de los lenguajes a aprender para la ciencia de datos.

Como limitaciones principales, fueron el poco conocimiento de todo lo relacionado a modelos de regresión y series de tiempo, ya que, pese a que en el ramo MA3402 - Estadística se ve al final un poco de modelos de regresión, no se acercan a lo visto en la actualidad, y no se ve nada de series de tiempo en el mismo ramo.

La otra limitación que se encontró fue la falta de conocimiento de programación en R, lo que hizo el comienzo de la práctica un poco lento, por lo que dentro de la malla de la carrera (o dentro del mismo ramo de Estadística nombrado anteriormente) podrían enseñar sobre este lenguaje.

Por último, me gustaría destacar la excelente introducción que se hizo al mundo de la ciencia de datos por parte de Ricardo y Wilfredo (respectivamente: el supervisor y la persona que me guió a RialStat), además de generar una motivación sobre la teoría que hay detrás de los modelos predictivos y su utilidad tanto en la actualidad como a futuro.

## Referencias

- [1] B.Central, “Base de datos estadísticos del banco central de chile,” 2023, <https://si3.bcentral.cl/siete/>.
- [2] INE, “Repositorio de datos estadísticos del instituto nacional de estadísticas,” 2023, <https://stat.ine.cl/?lang=es>.
- [3] Cochilco, “Base de datos de la comision chilena del cobre,” 2023, <https://www.cochilco.cl/Paginas/Estadisticas/Bases\%20de\%20Datos/Bases-de-Datos.aspx>.
- [4] Brockwell, P. J., “Introduction to time series and forecasting,” <https://link.springer.com/book/10.1007/978-3-319-29854-2>.
- [5] NCSS, “The box-jenkins method,” [https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/The\\_Box-Jenkins\\_Method.pdf](https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/The_Box-Jenkins_Method.pdf).
- [6] Hyndman, R. J., “Automatic time series forecasting: The forecast package for r,” Journal of Statistical Software, vol. 27, no. 3, 2008, doi:<https://dx.doi.org/10.18637/jss.v027.i03>.
- [7] Brockwell, P. J., “Introduction to time series and forecasting,” pp. 327 – 328, 2016.
- [8] Taylor, S. J., “Automatic time series forecasting: The forecast package for r,” The American Statistician, vol. 72, no. 1, pp. 37–45, 2018.
- [9] Novoa, J., “Reporte modelo cobre,” 2023, [https://drive.google.com/file/d/1yIVIPRP\\_gmDcj\\_bZq784YjBP0uwbi4VeF/view?usp=sharing](https://drive.google.com/file/d/1yIVIPRP_gmDcj_bZq784YjBP0uwbi4VeF/view?usp=sharing).
- [10] Novoa, J., “Reporte modelo brent,” 2023, <https://drive.google.com/file/d/1Dt1kKCNr5Hq6x-VngWevx8-cLogspWD7/view?usp=sharing>.
- [11] Novoa, J., “Reporte modelo imacec, desocup e irr,” 2023, [https://drive.google.com/file/d/1ZXAqLpgYErUIY-nQgF-hckui-mL\\_EsMI/view?usp=sharing](https://drive.google.com/file/d/1ZXAqLpgYErUIY-nQgF-hckui-mL_EsMI/view?usp=sharing).
- [12] Novoa, J., “Reporte modelo desocup con covariable imacec,” 2023, <https://drive.google.com/file/d/1SRUtn-sF4UGqn6vDtB-Vk-dNdV1JN68j/view?usp=sharing>.
- [13] Novoa, J., “Reporte modelos multiples divisas,” 2023, <https://drive.google.com/file/d/1XFnXi42fCt1vaNkMNWKhn1mnofD7oWFz/view?usp=sharing>.
- [14] U.N.C., “Nonlinear time series forecasting using mars,” <http://www.scielo.org.co/pdf/dyna/v81n184/v81n184a01.pdf>.
- [15] Novoa, J., “Reporte modelo dolar,” 2023, <https://drive.google.com/file/d/1AnvYtNbOueY3A2Fn2FVgpSGb6oWXR7wb/view?usp=sharing>.
- [16] Brockwell, P. J., “Introduction to time series and forecasting,” pp. 196 – 204, 2016.

## Anexo A. Ecuaciones de los Modelos Predictivos

### A.1. ARMA

Un modelo **ARMA**( $p,q$ ) consiste en lo siguiente:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

donde:

- $p$  : el orden del polinomio autoregresivo
- $q$  : el orden del polinomio de media móvil
- $\phi$  : los coeficientes de regresión (estas definen como influyen los valores de anteriores a  $X_t$ )
- $\theta$  : los coeficientes de medias móviles
- $\varepsilon$  : el error
- $c$  : una constante

### A.2. ARIMA

Si  $d$  es un entero no negativo, entonces  $\{X_t\}$  es un proceso **ARIMA**( $p,d,q$ ) si el siguiente proceso:

$$Y_t := (1 - B)^d X_t$$

es un **ARMA**( $p,q$ ).

### A.3. ARMAX

Si tenemos un modelo **ARMA**( $p,q$ ), junto con una variable  $Z$  que se puede correlacionar con la variable principal del modelo **ARMA**, entonces el siguiente proceso:

$$X_t = \beta Z_t + c + \varepsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

es un modelo **ARMAX**( $p,d,q$ ).

### A.4. ARIMAX

Si  $d$  es un entero no negativo, entonces  $\{X_t\}$  es un proceso **ARIMAX**( $p,d,q$ ) con covariable  $\{Z_t\}$  si el siguiente proceso:

$$Y_t := (1 - B)^d X_t$$

es un **ARMAX**( $p,q$ ) con covariable  $\{Z_t\}$ .

## A.5. PROPHET

El modelo PROPHET se puede considerar como un modelo de regresión no lineal, de la forma:

$$Y_t = g(t) + s(t) + h(t) + \varepsilon_t$$

donde:

- $g(t)$  : corresponde a la tendencia linear por partes
- $s(t)$  : corresponde a la parte estacional principal
- $h(t)$  : captura el efecto de calendario
- $\varepsilon_t$  : corresponde al ruido blanco

El modelo es estimado usando un enfoque Bayesiano para permitir la selección automática de los puntos de cambio (para la tendencia linear por partes) y otras características del modelo.

## Anexo B. Imágenes

A	B	C	D	E	F	G	H	I	Z	AA	AB	AC	AD
year	mes	us	euro	libra	real	yen	yuan	libo180	euro_clp	libra_clp	real_clp	yen_clp	yuan_clp
2010	1	500.66	0.70	0.62	1.77	91.17	6.83	0.39913	715.66	808.99	282.15	5.49	73.34
2010	2	532.56	0.73	0.64	1.85	90.20	6.83	0.38780	729.27	834.06	288.60	5.90	77.99
2010	3	523.16	0.74	0.66	1.79	90.51	6.83	0.40943	710.22	787.79	292.87	5.78	76.64
2022	12	875.66	0.95	0.82	5.24	135.08	6.99	5.16087	925.89	1065.09	166.98	6.48	125.29
2023	1	826.34	0.9287	0.8183	5.198	130.4152	6.795	5.1389	889.73	1009.76	158.98	6.34	121.6

Figura B.1: Parte de la Base de Datos utilizada durante la práctica.

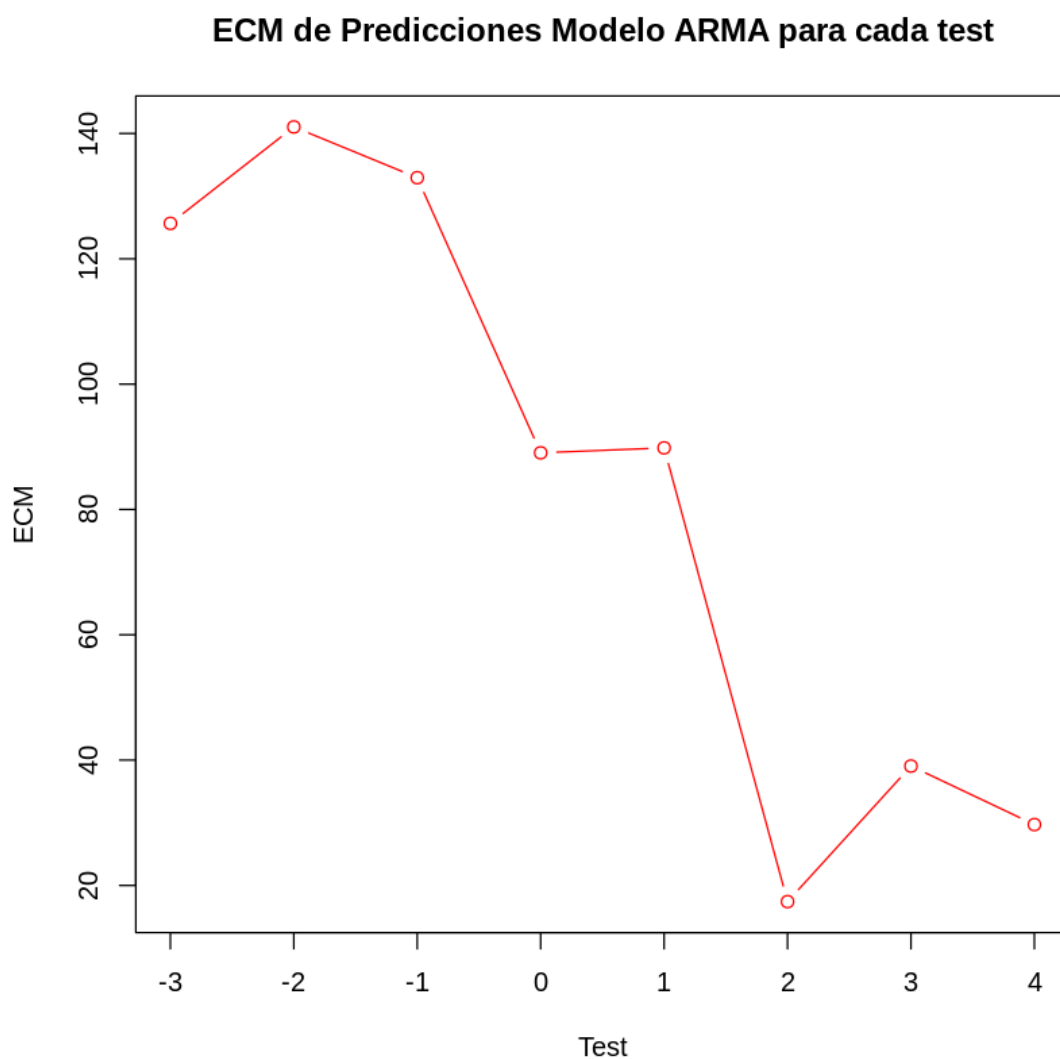


Figura B.2: ECM del Modelo para la serie del Cobre.

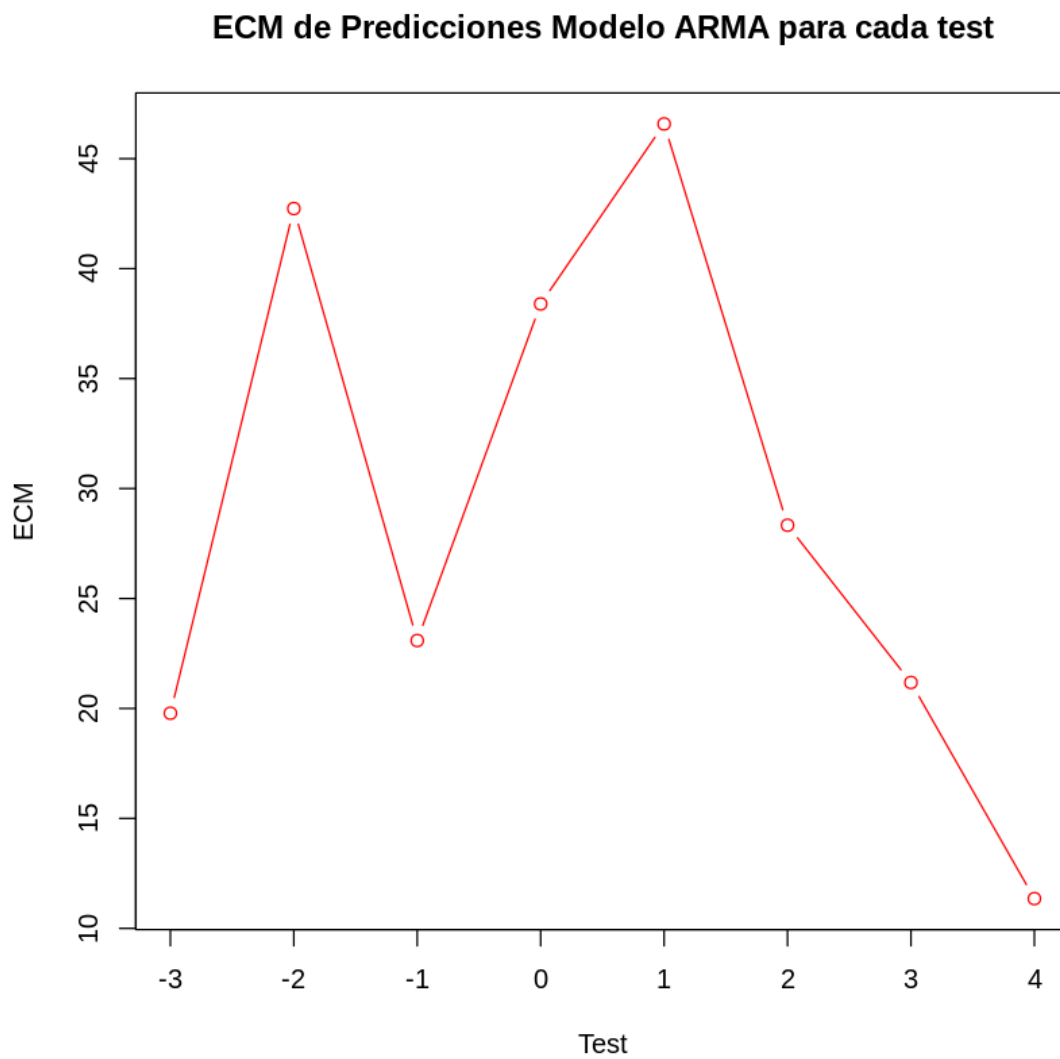


Figura B.3: ECM del Modelo para la serie del Petroleo BRENT.

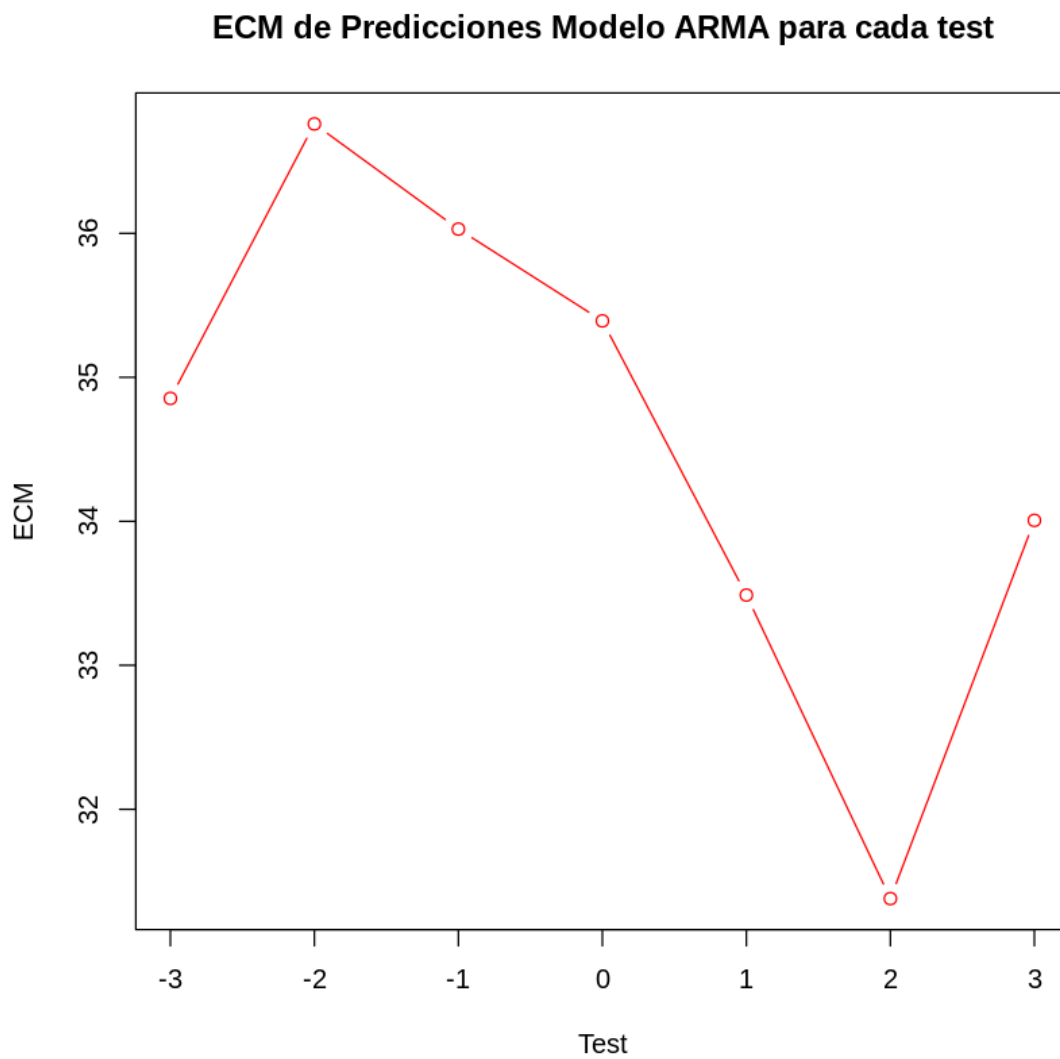


Figura B.4: ECM del IMACEC.



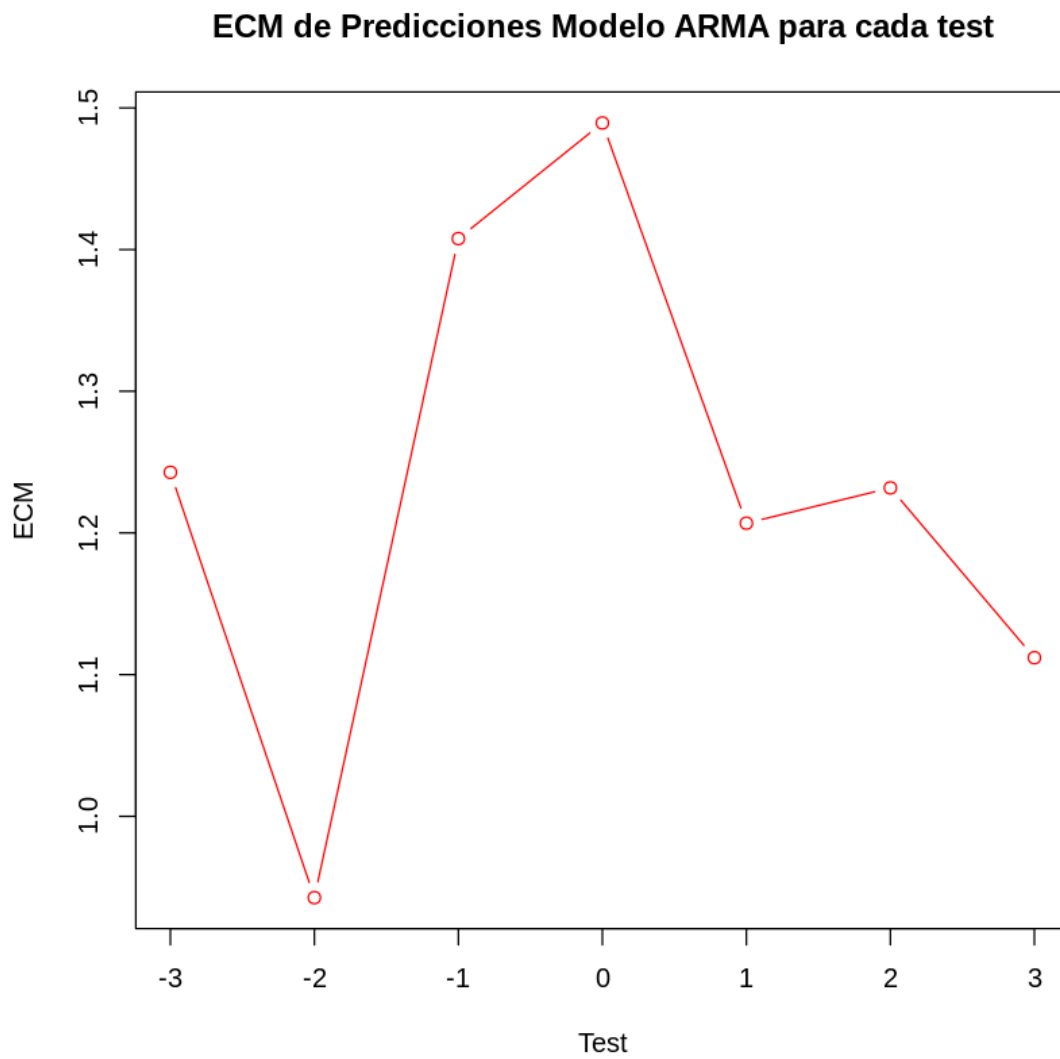


Figura B.5: ECM del Índice de Desocupación.

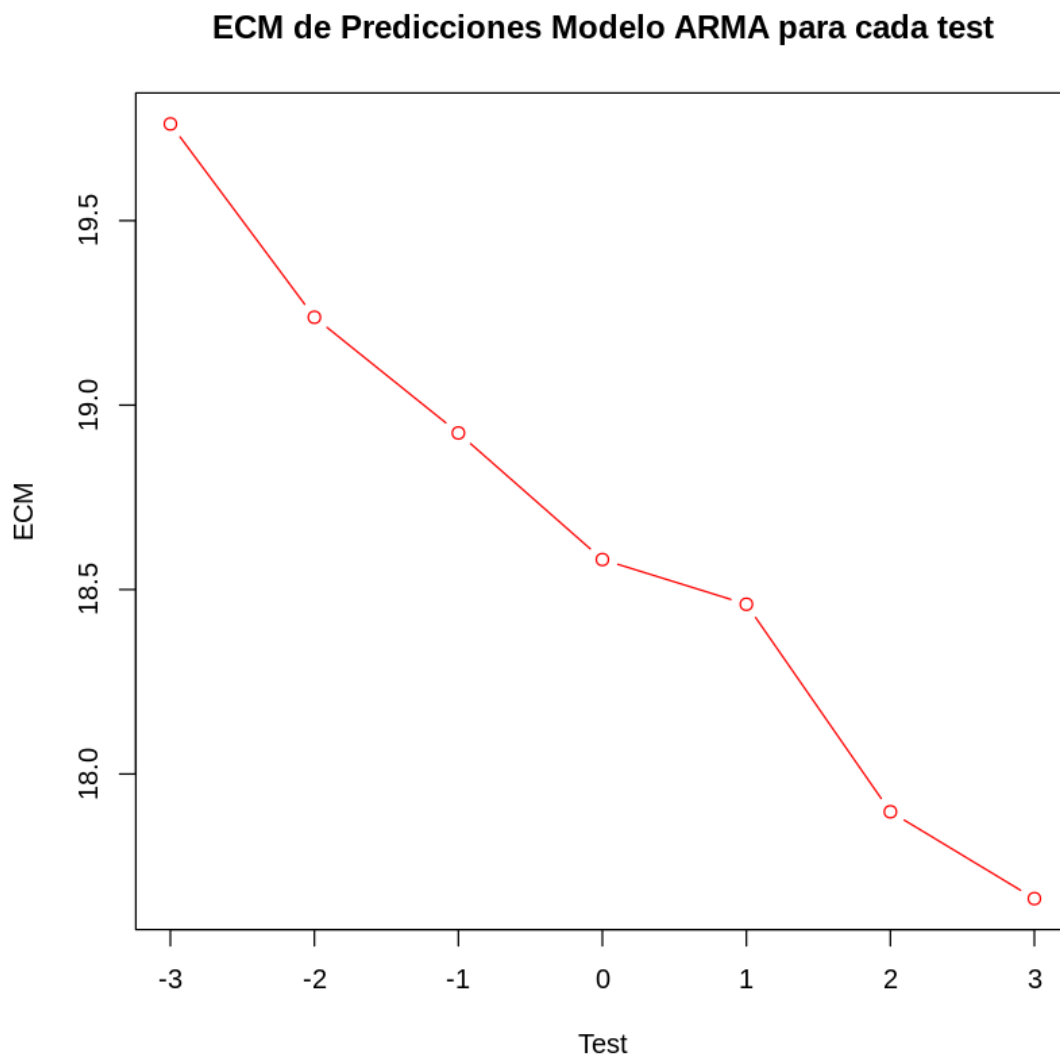


Figura B.6: ECM del IRR.

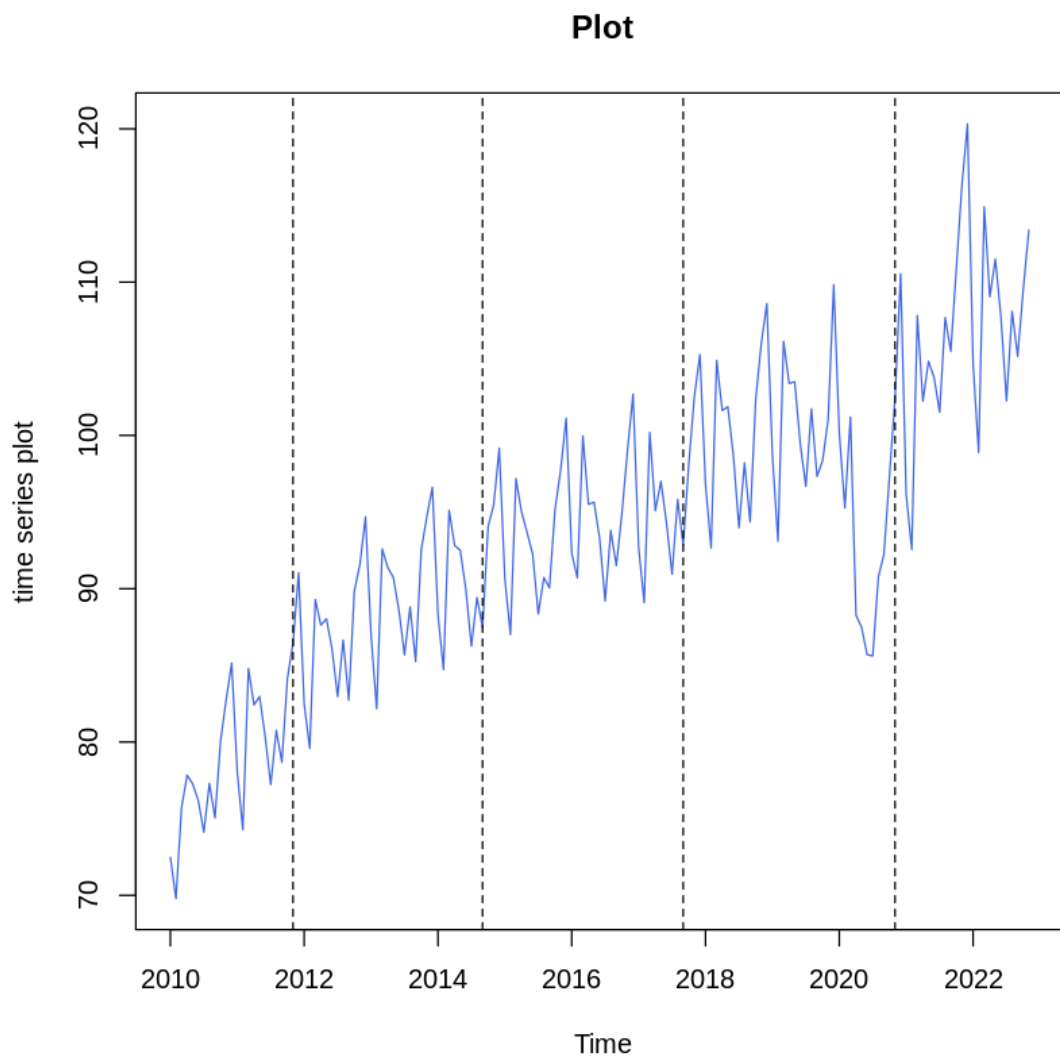


Figura B.7: Puntos de cambio estructural para la Serie del IMACEC.

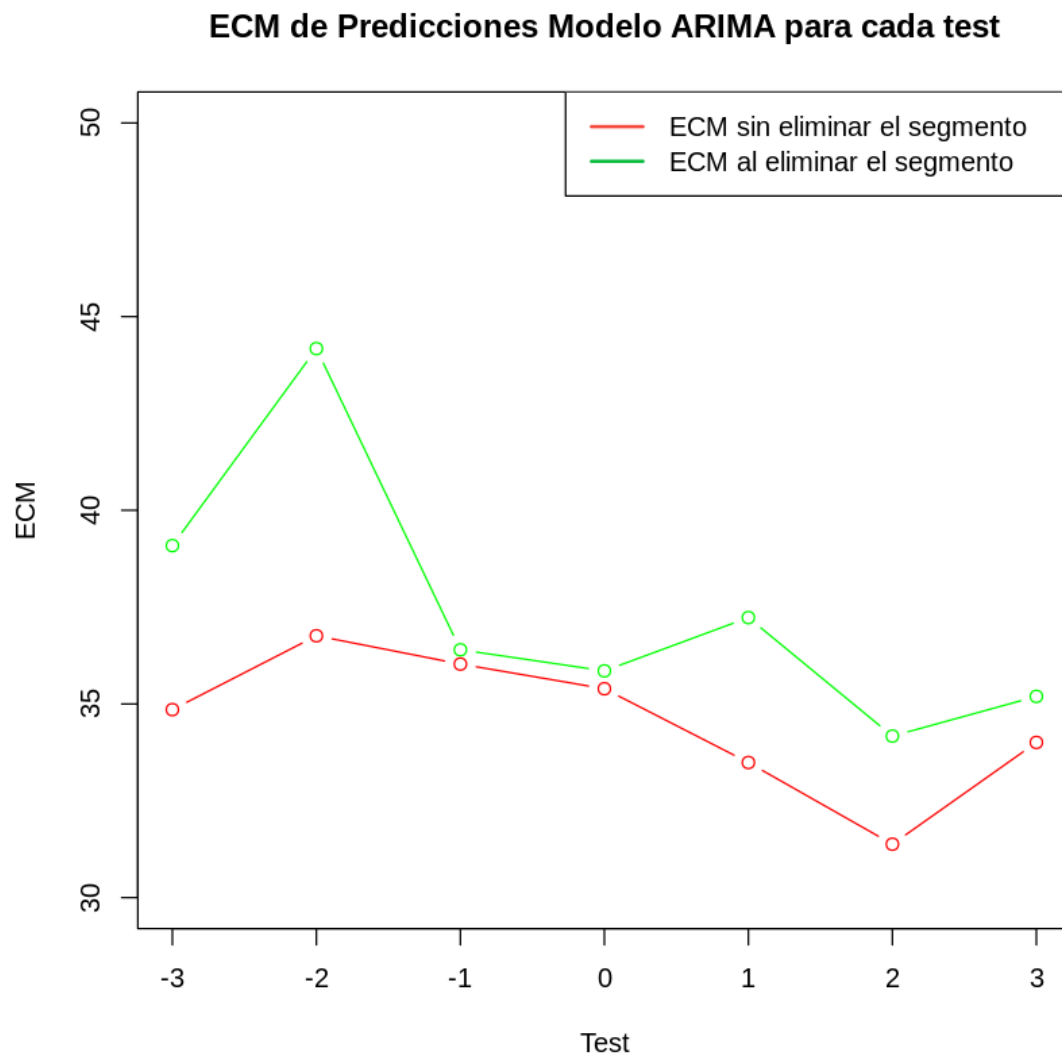


Figura B.8: ECM al considerar el cambio estructural vs no considerar el cambio estructural.

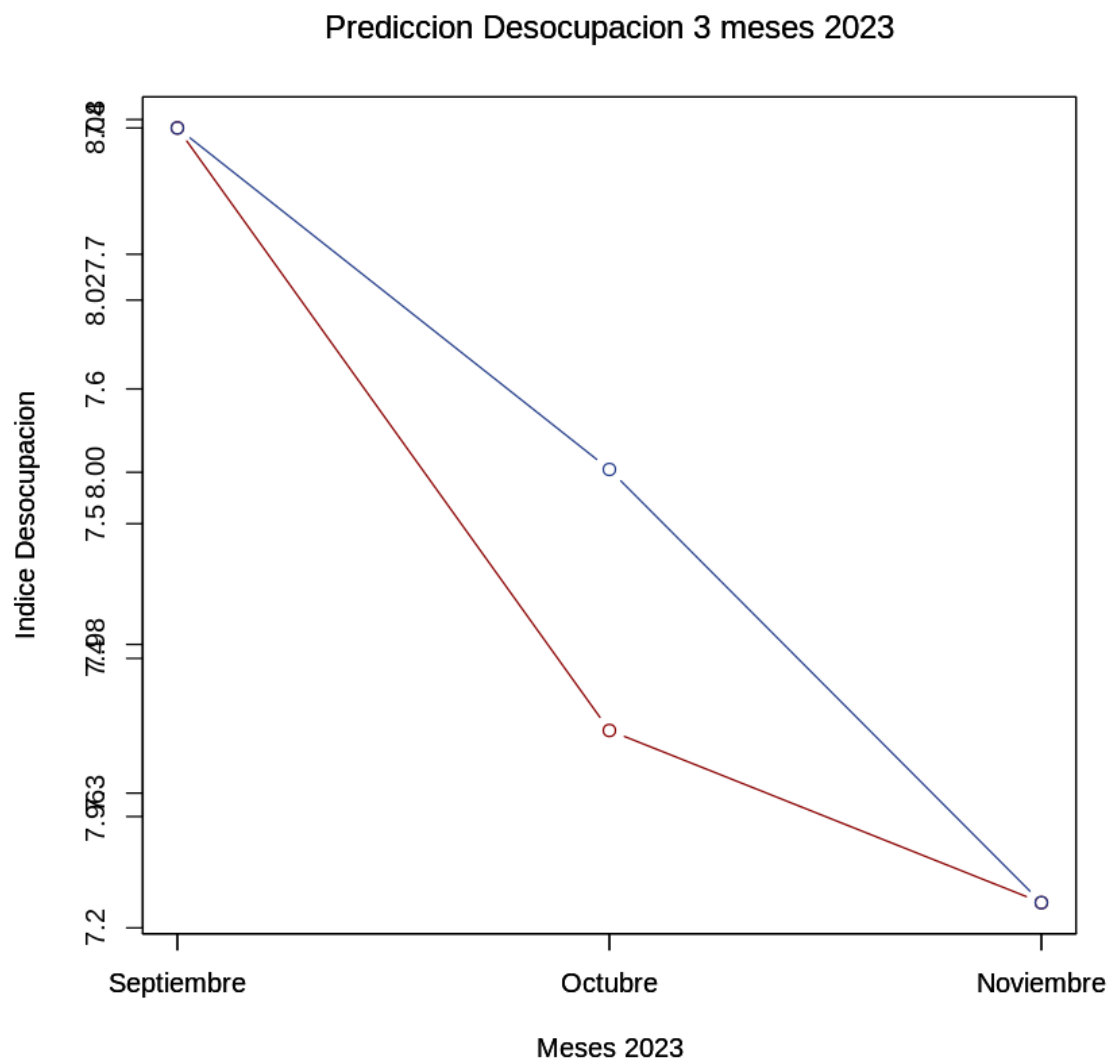


Figura B.9: Predicción del Índice de Desocupación vs los valores reales según el mes.

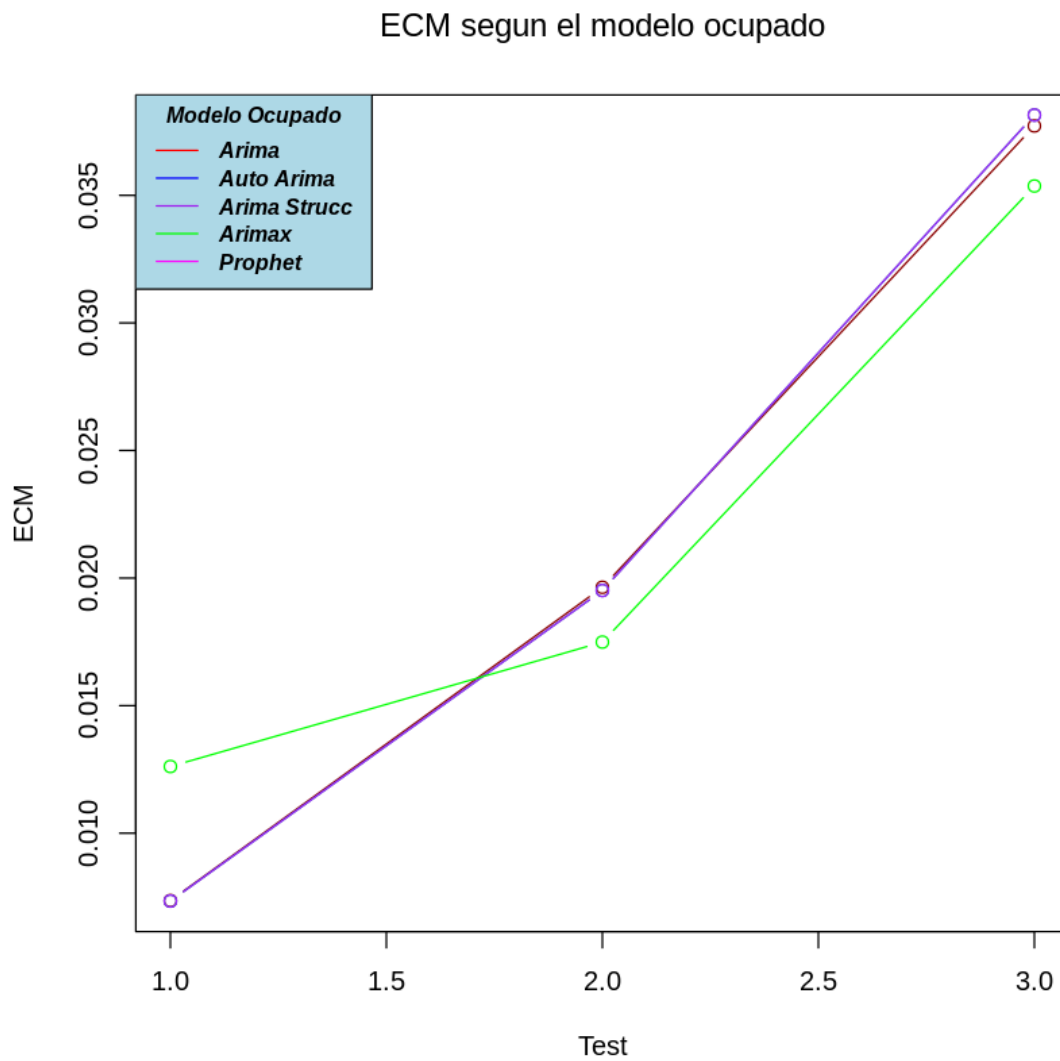


Figura B.10: ECM de múltiples modelos para la serie del Euro.

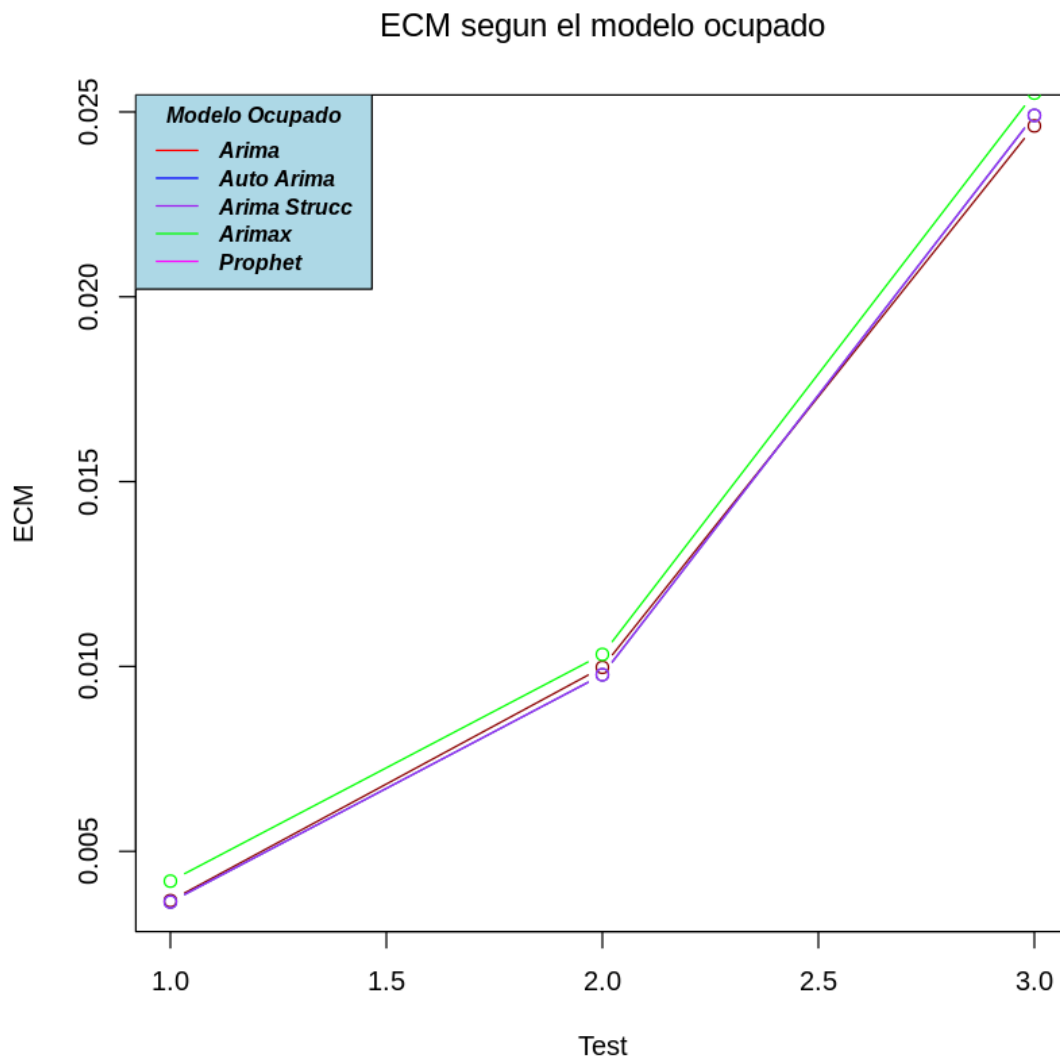


Figura B.11: ECM de múltiples modelos para la serie de la Libra.

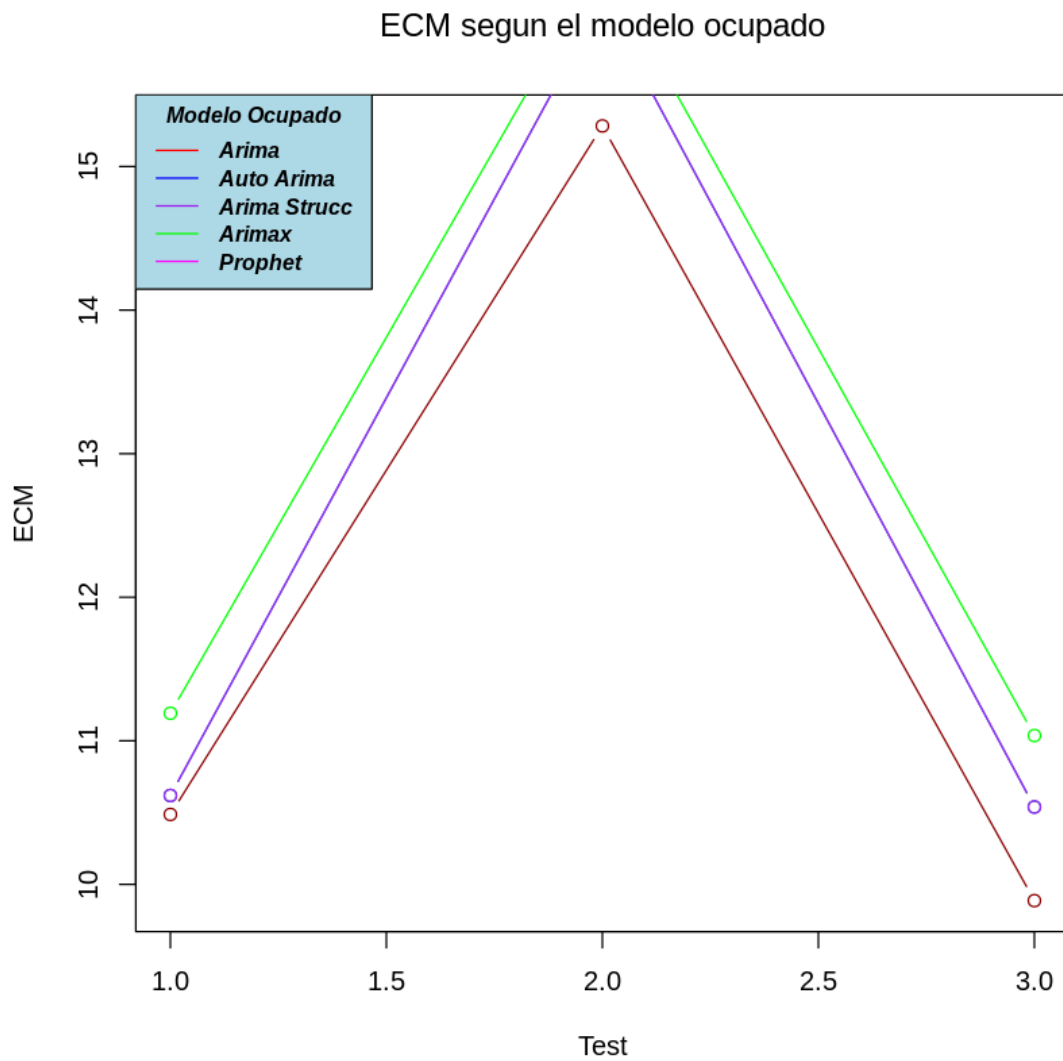


Figura B.12: ECM de múltiples modelos para la serie del Real.



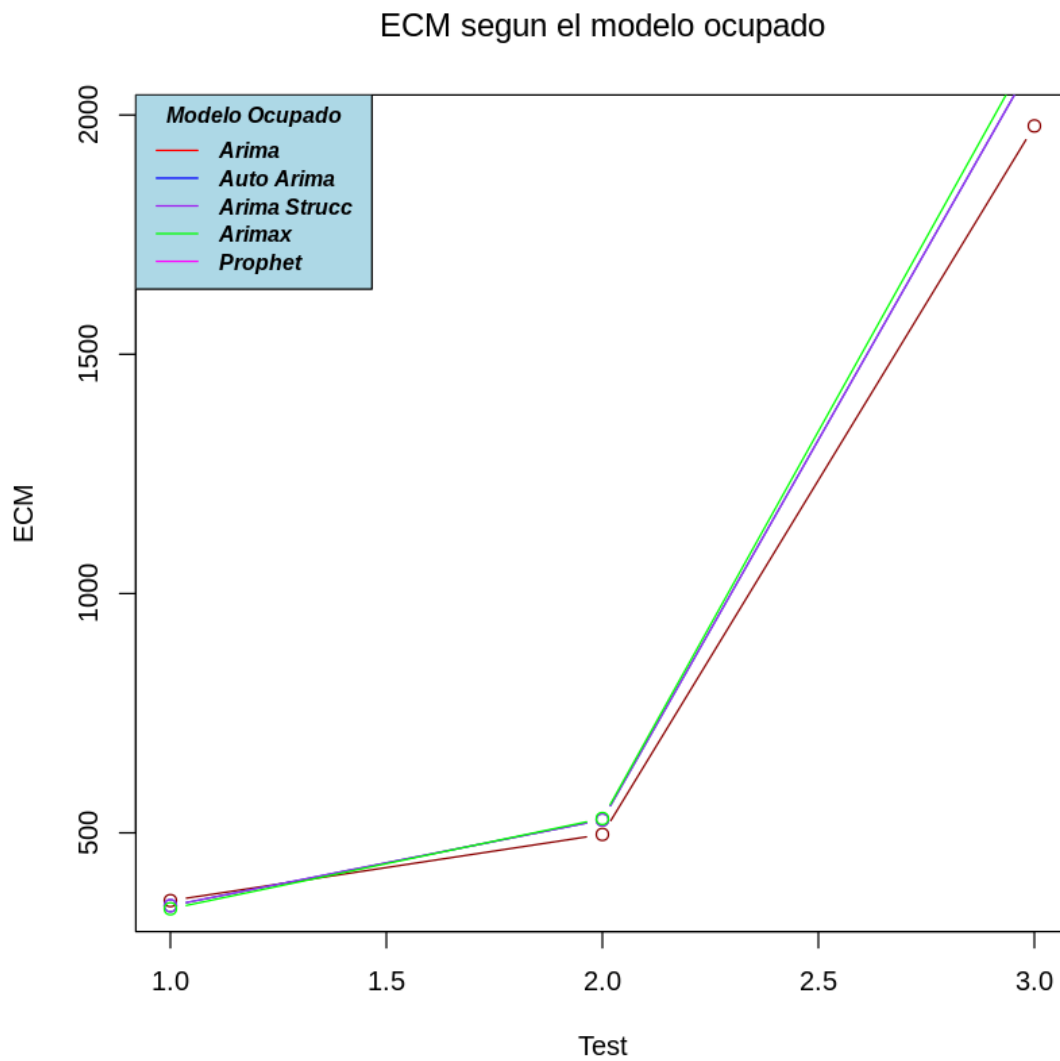


Figura B.13: ECM de múltiples modelos para la serie del Yen.

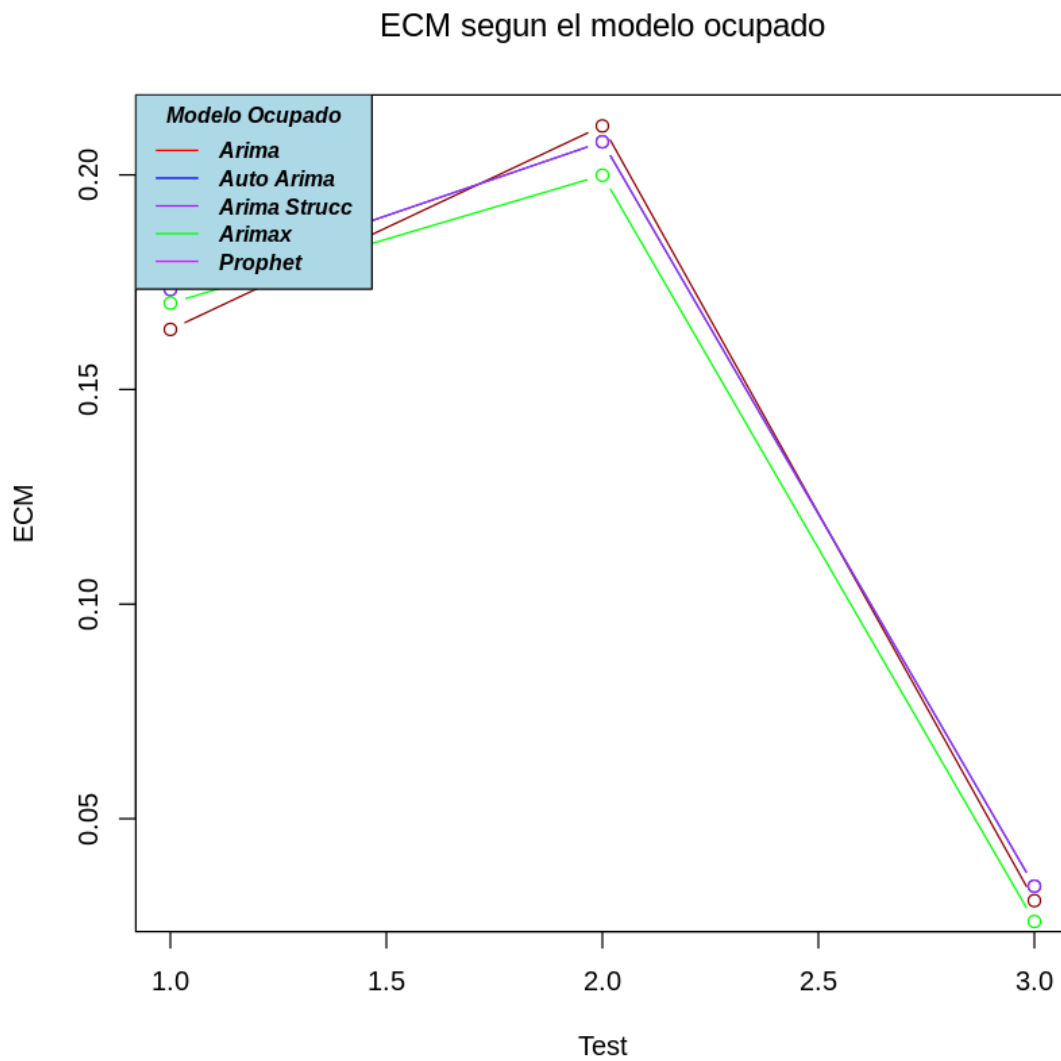


Figura B.14: ECM de múltiples modelos para la serie del Yuan.

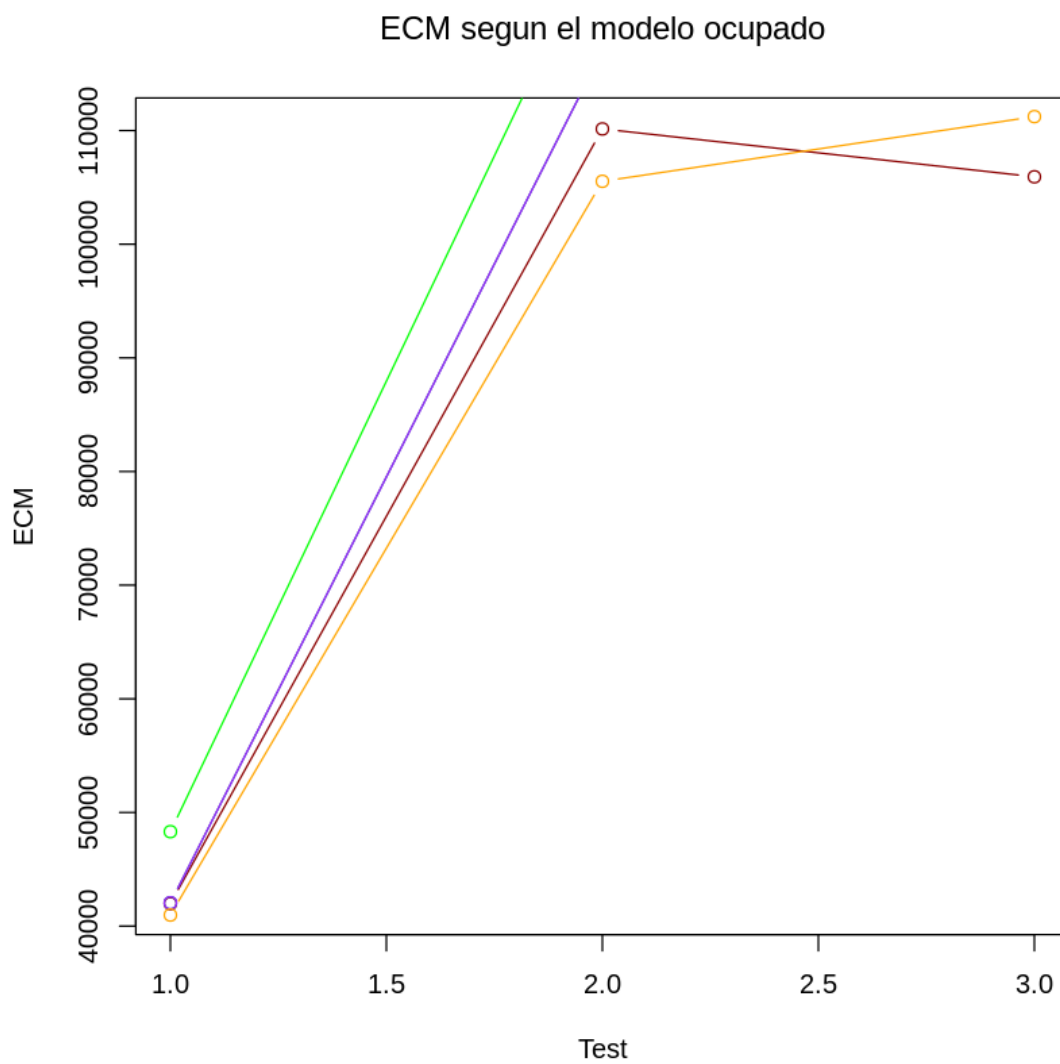


Figura B.15: ECM de múltiples modelos para la serie del Dolar

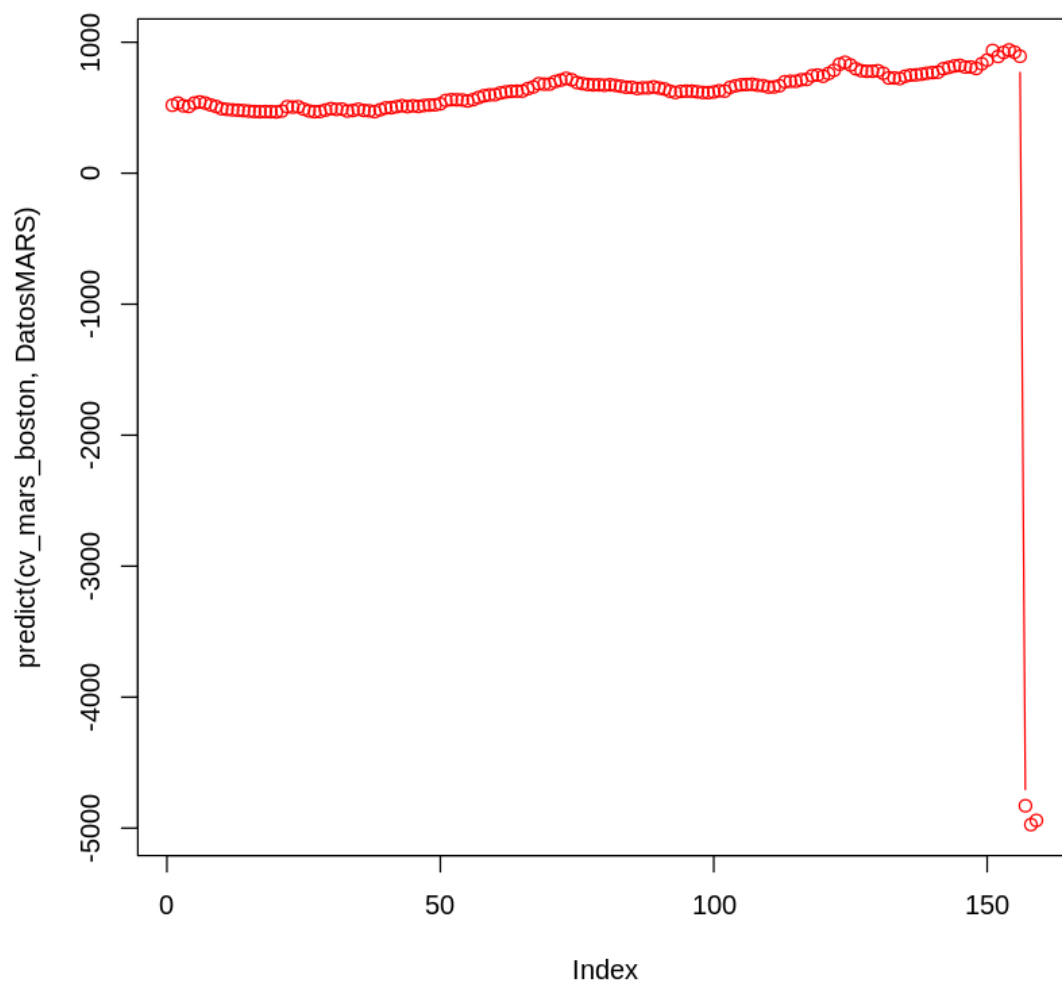


Figura B.16: Modelo MARS junto con su predicción.

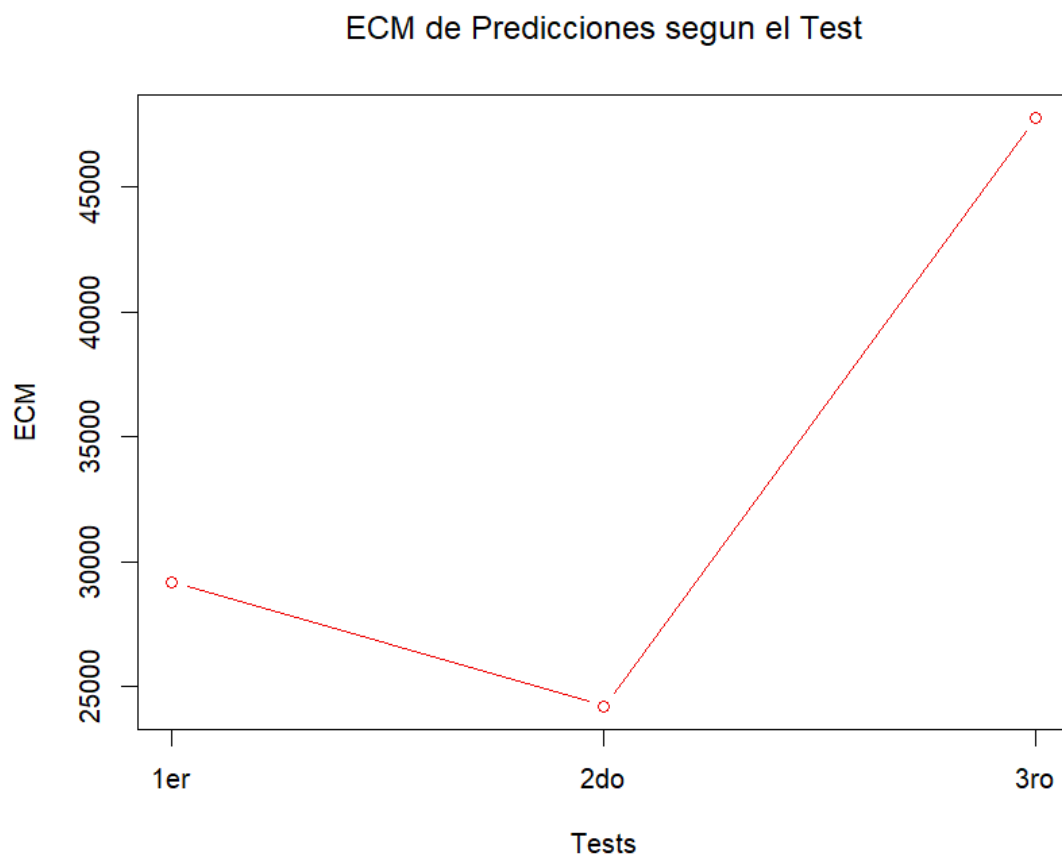


Figura B.17: ECM de modelo LSTM para el Dolar.

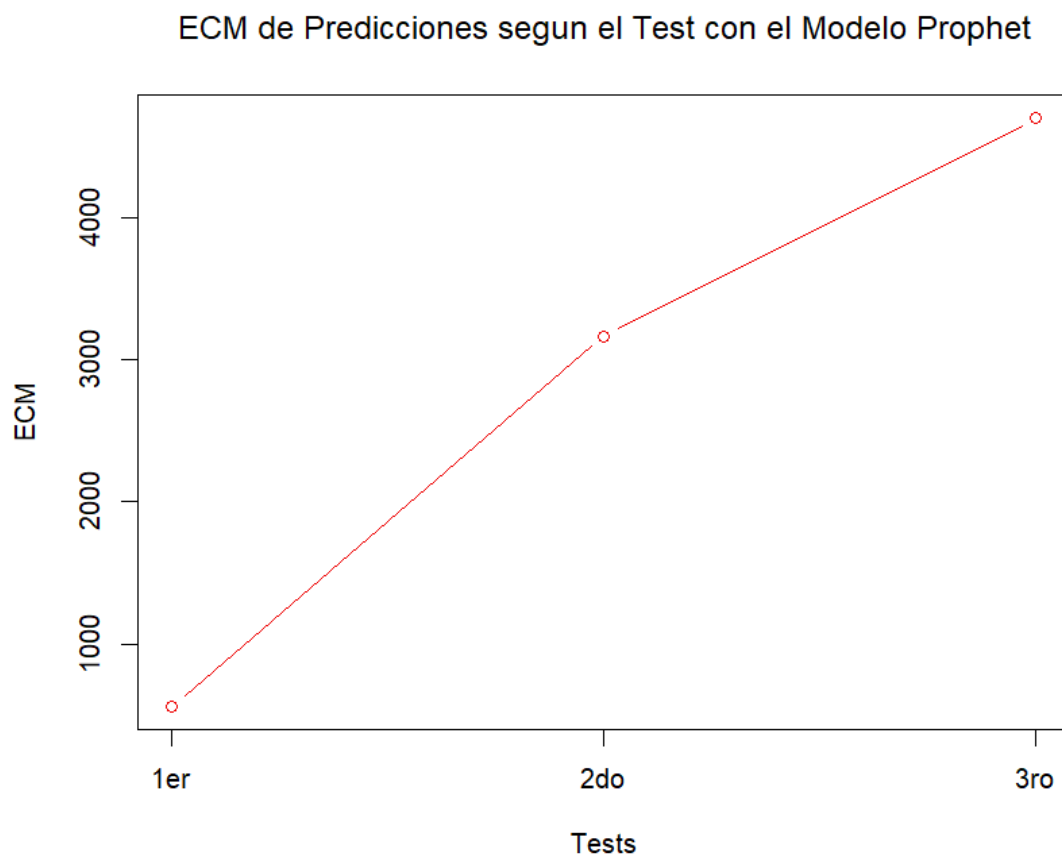


Figura B.18: ECM de modelo Prophet + covariables para el Dolar.

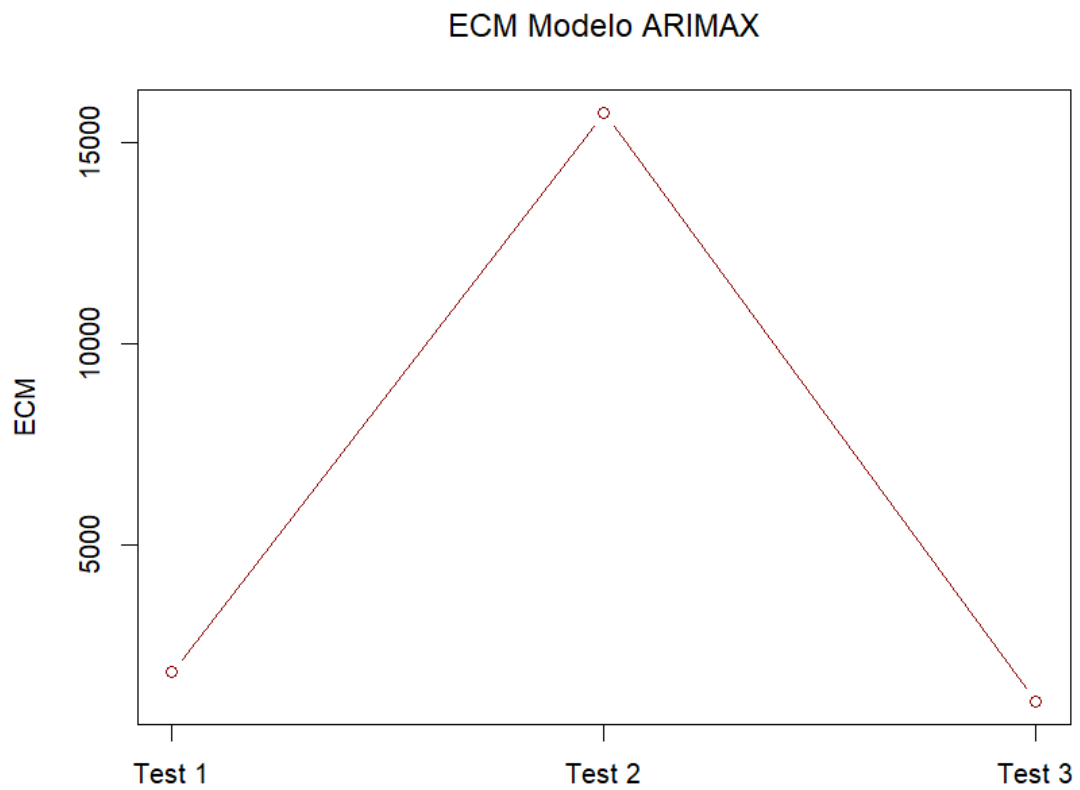


Figura B.19: ECM de modelo ARIMAX para el Dolar.

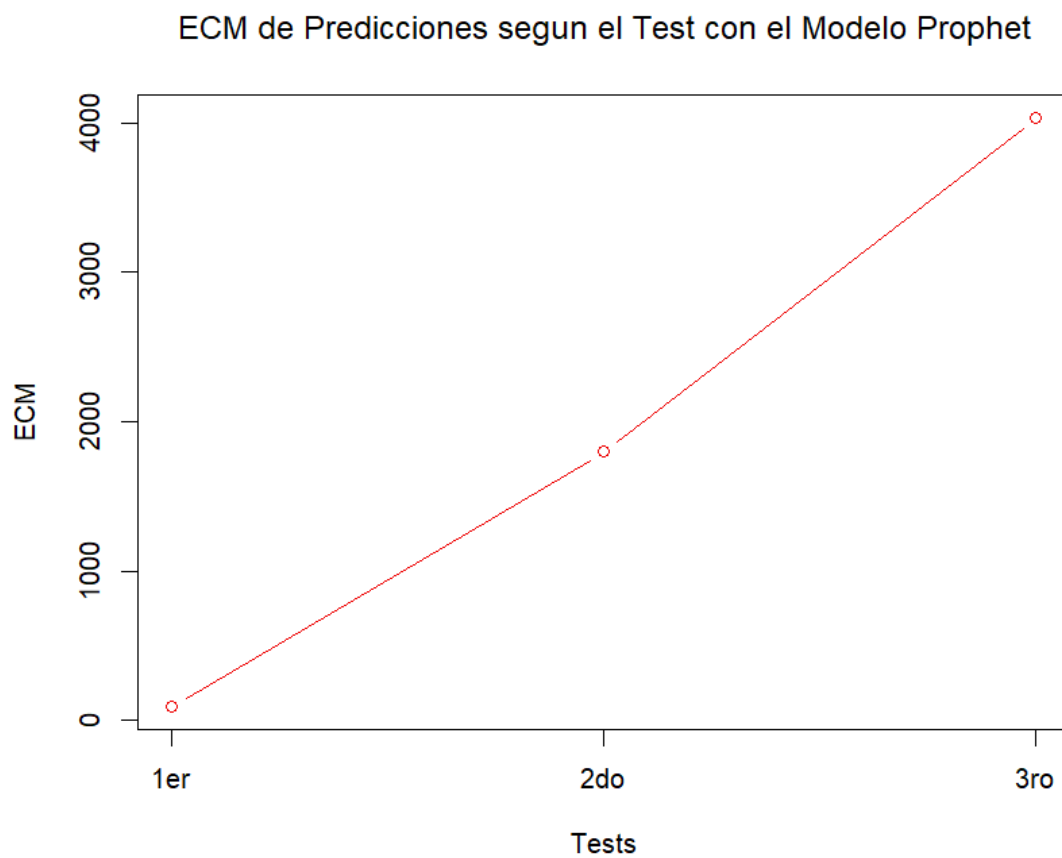


Figura B.20: ECM de modelo Prophet + covariables y considerando el cambio estructural para el Dolar.