



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

Control 3: Series de Tiempo

Diplomado en Estadística - Curso Métodos Estadísticos - Pontificia
Universidad Católica de Chile

Profesor: Oscar Ortiz de la Hoz Ayudante: Diego Aravena
Estudiante: Martin Venegas Márquez

16/09/2023

Tabla de contenidos

Instrucciones	3
Desarrollo	4
(a) Descargue la serie mensual histórica de Imacec no minero hasta diciembre de 2022.	4
(b) Utilizando la función <code>ts(...)</code> genere un objeto con atributos de serie de tiempo	4
(c) Utilizando <code>forecast::BoxCox.lambda(...)</code> y <code>forecast::BoxCox(...)</code> evalúe una transformación.	5
(d) Utilizando <code>forecast::ndiffs(...)</code> y <code>forecast::nsdiffs(...)</code> evalúe los grados de integración de la serie.	6
(e) Ajuste un modelo SARIMA usando <code>forecast::auto.arima(...)</code>	6
(f) Chequee la significancia estadística de coeficientes estimados, blancura, normalidad y homocedasticidad de los residuos.	6
(g) Proponga una mejora al modelo propuesto en (e) utilizando el argumento <code>fixed</code> en la función <code>forecast::Arima(...)</code>	8
(h) Con <code>forecast::forecast(...)</code> genere pronósticos hasta diciembre de 2027. . .	11
(i) Compare el desempeño de sus pronostico con respecto a los valores reales observados hasta junio de 2023.	13

Instrucciones

De acuerdo con la información preliminar, el Imacec de junio de 2023 cayó 1,0% en comparación con igual mes del año anterior (Gráfico 1). La serie desestacionalizada aumentó 0,5% respecto del mes precedente y cayó 1,1% en doce meses. El mes registró la misma cantidad de días hábiles que junio de 2022. La variación anual del Imacec fue explicada por la caída del comercio y la industria, compensada en parte por el resto de bienes. En tanto, el crecimiento del Imacec desestacionalizado estuvo determinado por el desempeño de la minería. El Imacec no minero presentó una disminución de 1,3% en doce meses, mientras que en términos desestacionalizados, cayó 0,2% respecto del mes anterior.

<https://www.bcentral.cl>

Ejercicio

- (a) Descargue la serie mensual histórica de Imacec no minero hasta diciembre de 2022.
- (b) Utilizando la función `ts(...)` genere un objeto con atributos de serie de tiempo
- (c) Utilizando `forecast::BoxCox.lambda(...)` y `forecast::BoxCox(...)` evalúe una transformación.
- (d) Utilizando `forecast::ndiffs(...)` y `forecast::nsdiffs(...)` evalúe los grados de integración de la serie.
- (e) Ajuste un modelo SARIMA usando `forecast::auto.arima(...)`.
- (f) Chequee la significancia estadística de coeficientes estimados, blancura, normalidad y homocedasticidad de los residuos.
- (g) Proponga una mejora al modelo propuesto en (e) utilizando el argumento `fixed` en la función `forecast::Arima(...)`.
- (h) Con `forecast::forecast(...)` genere pronósticos hasta diciembre de 2027.
- (i) Compare el desempeño de sus pronóstico con respecto a los valores reales observados hasta junio de 2023.

Desarrollo

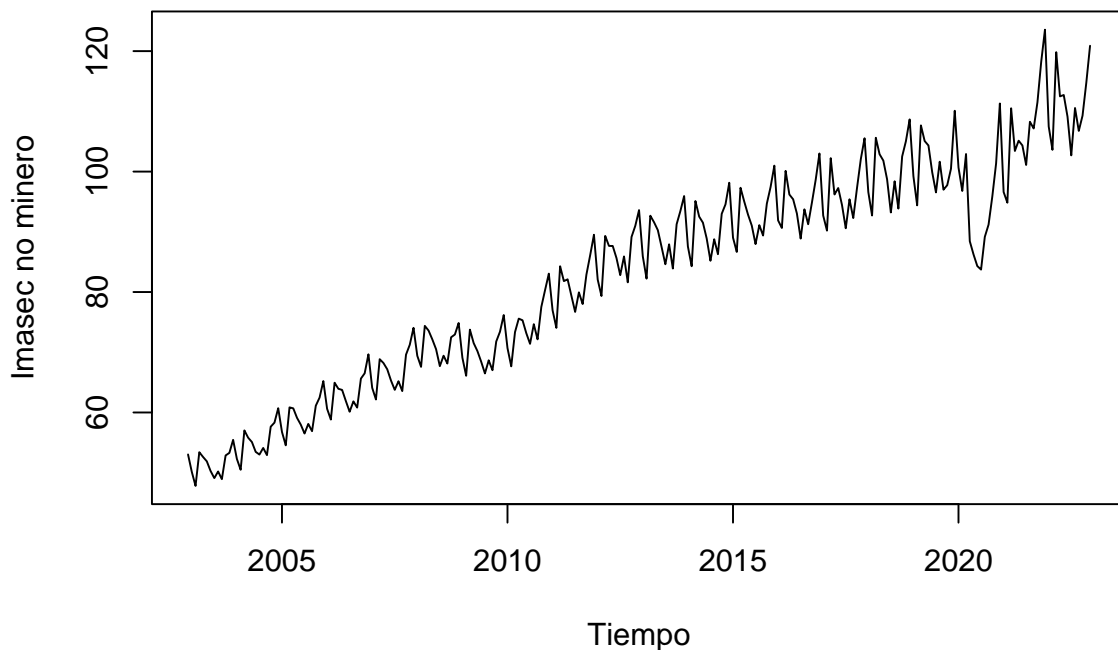
(a) Descargue la serie mensual histórica de Imacec no minero hasta diciembre de 2022.

```
df <- readxl::read_xlsx("../input/data/original/IMACEC_PIB.xlsx") %>%  
  rename(annio = AÑO,  
         mes = MES,  
         imacec = IMACEC_NM  
  ) # Cargar data y renombrar variables
```

(b) Utilizando la función `ts(...)` genere un objeto con atributos de serie de tiempo

```
Y <- ts(  
  df$imacec,  
  start = c(df$annio[1], df$mes[1]),  
  end = c(2022, 12),  
  frequency = 12  
) # Crear objeto TS
```

Figure 1: Serie de Imacec no minero desde 2002 hasta 2022



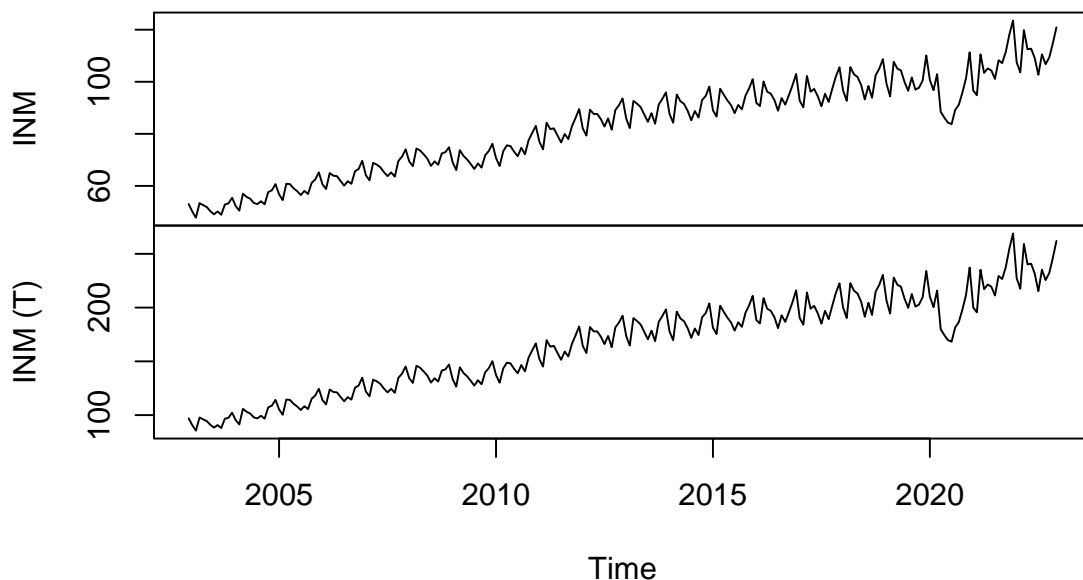
En la Figura 1 se presenta una serie de tiempo del Índice Mensual de Actividad Económica (Imacec) no minero entre los años 2002 y 2022. A grandes rasgos, se observa que la serie no cumple con la condición de estacionaridad para ajustar un modelo de series de tiempo. Esto ya que: a) se observa una tendencia al alza en el Imacec; b) se observa un quiebre por evento conocido en los años cercanos al 2020, probablemente relacionado al efecto de la pandemia por COVID-19; c) se observa una posible efecto estacional en la serie, probablemente explicado por la mayor cantidad de disponibilidad de empleo en época de fiestas y fin de año; d) después del año 2010 se observa un incremento leve en la variabilidad, lo que sugiere la necesidad de transformar los datos para un potencial ajuste de modelo.

En consideración de lo expuesto anteriormente, en los puntos siguientes se procederá a transformar la serie a fin de que cumpla los supuestos de estacionaridad y, de esta manera, poder ajustar un modelo que prediga el comportamiento futuro de este índice.

(c) Utilizando `forecast::BoxCox.lambda(...)` y `forecast::BoxCox(...)` evalúe una transformación.

```
lambda <- forecast::BoxCox.lambda(Y, method = "loglik") # Da un lambda de 1.2
Y_v <- forecast::BoxCox(Y, lambda = 1.2)
```

Figure 2: Comparación entre serie de Imacec NM no transformado y transformado por BoxCox



Al observar la Figura 2 no se identifican grandes diferencias entre la serie transformada y no transformada. Esta afirmación se corrobora al aplicar el test de Breusch-Pagan, en

tanto no existe evidencia suficiente para rechazar la hipótesis nula de homocedasticidad de varianza (*valor p* = 3.156e-07 < .05).

(d) Utilizando `forecast::ndiffs(...)` y `forecast::nsdiffs(...)` evalúe los grados de integración de la serie.

Considerando la existencia de una tendencia y estacionalidad, se consideran la posibilidad de una diferenciación. A fin de conocer el nivel de la diferenciación, se aplican los siguientes tests:

```
# Componente no estacional
forecast::ndiffs(Y, test = "kpss")
forecast::ndiffs(Y, test = "adf")
forecast::ndiffs(Y, test = "pp")

# Componente estacional
forecast::nsdiffs(Y, test = "seas")
forecast::nsdiffs(Y, test = "ocsb")
```

Todos los test sugieren la aplicación de una diferenciación ($d = 1$), tanto para la componente no estacional como para la componente estacional.

(e) Ajuste un modelo SARIMA usando `forecast::auto.arima(...)`.

```
m_auto <- auto.arima(Y)
```

(f) Chequee la significancia estadística de coeficientes estimados, blancura, normalidad y homocedasticidad de los residuos.

En la Tabla 1 se presentan los parámetros obtenidos por el `auto.arima()`. La función sugiere que el mejor modelo es un SARIMA de orden $(p = 2, d = 0, q = 1)$ ($P = 2, D = 0, Q = 1$). Sin embargo, vemos que no todos los parámetros son estadísticamente significativos, sino que el Imacec no minero se puede predecir a partir de un $AR(2)$ (*valor p* < .0001) junto con su componente estacional $SAR(2)$ (*valor p* < .001). Además, el modelo sugiere que el componente de media móvil $MA(1)$ y un *Drift* son parámetros relevantes para la predicción del Imacec (ambos con *valor p* < .001).

En lo que respecta a los supuestos, se aplican los test de Shapiro-Wilk y de Breusch-Pagan para determinar la existencia de normalidad y homocedasticidad de los residuos. Ambos tests sugieren que no se cumple ni con normalidad en los residuos (*valor p* = 1.475e-13 < .05) ni con homocedasticidad (*valor p* = 4.872063e-05 < .05).

Table 1: SARIMA (2, 0, 1)(2, 1, 1) modelo automático

	Model 1
AR(1)	0.11 (0.10)
AR(2)	0.71*** (0.10)
MA(1)	0.91*** (0.07)
SAR(1)	-0.26 (0.14)
SAR(2)	-0.33*** (0.09)
SMA(1)	-0.27 (0.14)
Drift	0.25*** (0.04)
AIC	895.67
AICc	896.32
BIC	923.14
Log Likelihood	-439.83
Num. obs.	229

*** p < 0.001; ** p < 0.01; * p < 0.05

Figure 3: Prueba de Ljung-Box para SARIMA(2,0,1)(2,1,1)

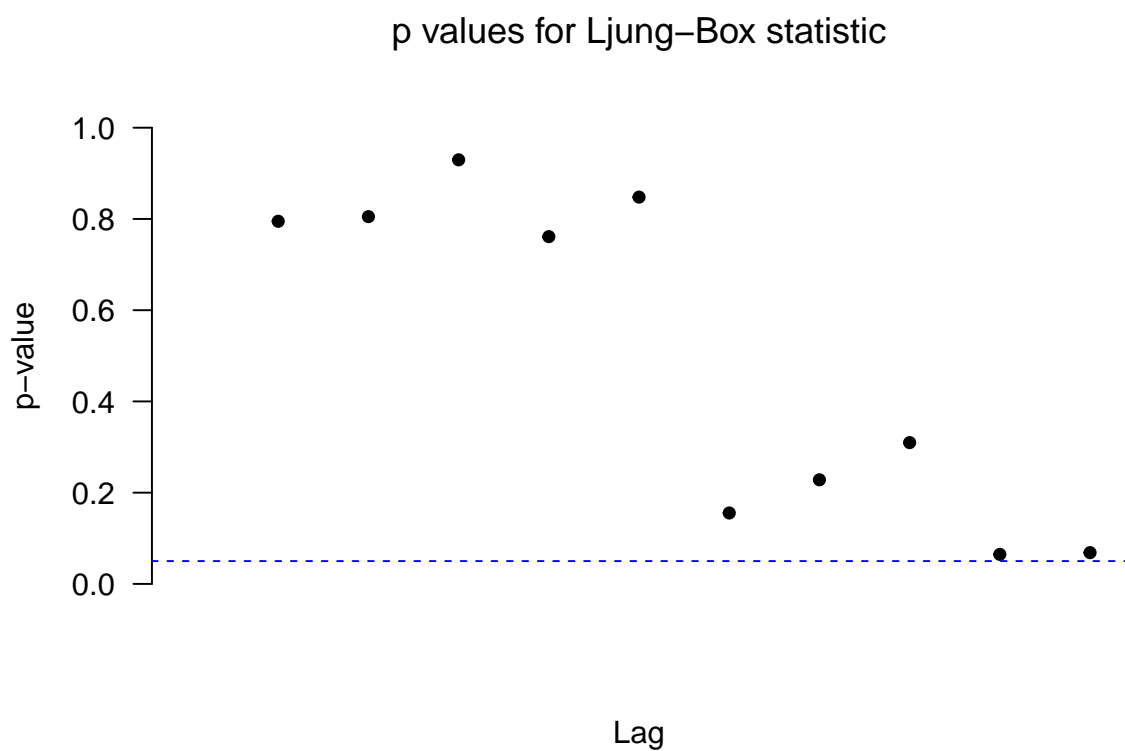
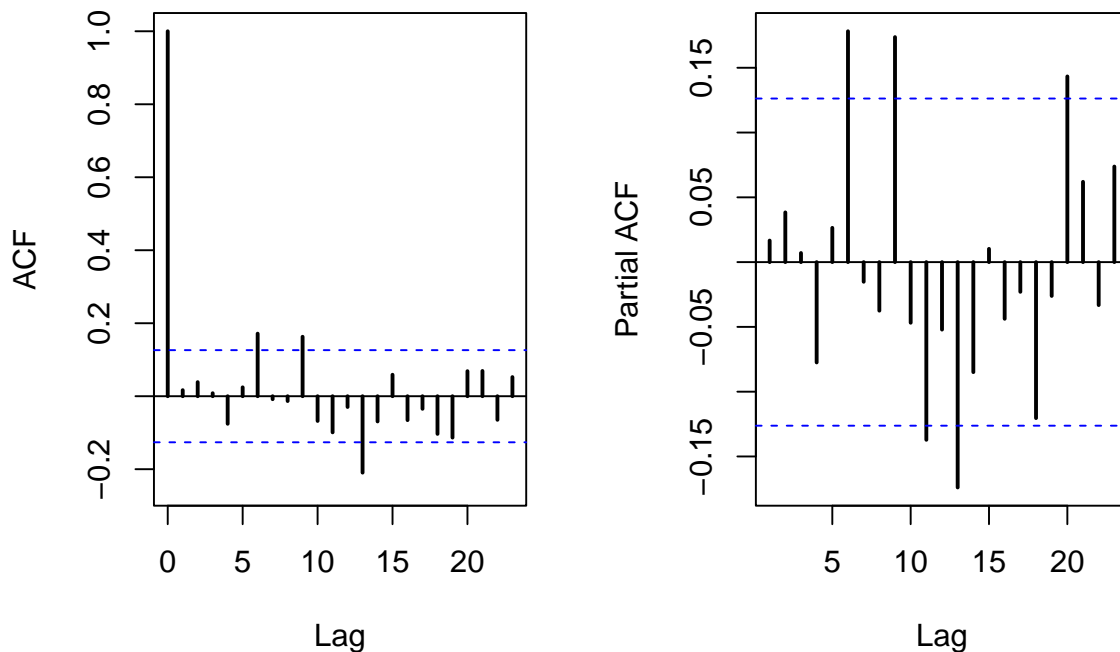


Figure 4: Gráficos de ACF y PCF para evaluación de autocorrelación en los residuos para SARIMA(2, 0, 1)(2, 1, 1)



La prueba de Ljung-Box sugiere que se cumple con la blancura en los residuos (Figura 3). Sin embargo, los gráficos de ACF y PACF (Figura 4) sugieren que aún hay autocorrelación que se puede extraer del modelo. En concreto, se observa que para el componente de autorezago (AR), los lags 6, 9, 11 y 13 están correlacionados significativamente con Y_t . Así también, se observa que para el componente de media móvil (MA), los lags 6, 9 y 13 correlacionan significativamente con Y_t .

(g) Proponga una mejora al modelo propuesto en (e) utilizando el argumento `fixed` en la función `forecast::Arima(...)`.

En consideración de lo anterior, se procederá a probar un modelo distinto al sugerido por el software. En detalle, a fin de no sobre calcular parámetros, se estimará un modelo con el lag significativo más bajo en el componente de AR, fijando a 0 los lags anteriores. Esto corresponde a un $AR(6)$. A fin de mantener parsimonia, se mantendrá un $MA(1)$, sin embargo, si se incorporará una diferenciación en el componente no estacional, a fin de seguir las recomendaciones de los tests empleados anteriormente.

```
# SARIMA de orden 6 en p y P y
# de orden 1 en q y Q y
# diferenciado en ambas componentes.

AR <- c(rep(0, 5), NA) #p # recordar ver PACF
```



```

MA <- NA #q # recordar ver ACF
SAR <- c(rep(0, 5), NA)
SMA <- NA

m_fixed <- forecast::Arima(Y,
  order = c(6, 1, 1),
  seasonal = c(6, 1, 1),
  fixed = c(AR, MA, SAR, SMA)
) ### Ajustar modelo fijado

```

Table 2: SARIMA (6, 1, 1)(6, 1, 1) calculado manual

	Model 1
AR(1)	0.00
AR(2)	0.00
AR(3)	0.00
AR(4)	0.00
AR(5)	0.00
AR(6)	0.15* (0.07)
MA(1)	0.09 (0.07)
SAR(1)	0.00
SAR(2)	0.00
SAR(3)	0.00
SAR(4)	0.00
SAR(5)	0.00
SAR(6)	0.25** (0.09)
SMA(1)	-0.56*** (0.06)
AIC	903.34
AICc	903.61
BIC	920.49
Log Likelihood	-446.67
Num. obs.	228

*** p < 0.001; ** p < 0.01; * p < 0.05

En la Tabla 2 se observa el modelo propuesto. Se observa que efectivamente $AR(6)$ y su componente estacional son estadísticamente significativos (*valor p* < .05). Sin embargo, vemos que el AIC es mayor al del modelo calculado automáticamente ($AIC_{SARIMAfixed} = 903.34 > AIC_{SARIMAauto} = 895.67$), además de tampoco cumplir con los supuestos en los residuos (Shapiro-Wilk: *valor p* = 6.001e-14 < .05; Breusch-Pagan: *valor p* = 2.793176e-06 < .05).

En lo que respecta al test de blancura (Figura 5), se observa que los residuos no presentan una correlación general estadísticamente significativa. A la par, los gráficos de ACF y

PACF (Figura 6) dan cuenta de que el componente de media móvil mejoro un poco en términos de su autocorrelación en comparación al modelo propuesto por el software, sin embargo, el componente de autorezago aun presenta altas cantidades de autocorrelación.

Figure 5: Prueba de Ljung-Box para SARIMA(6,1,1)(6,1,1)

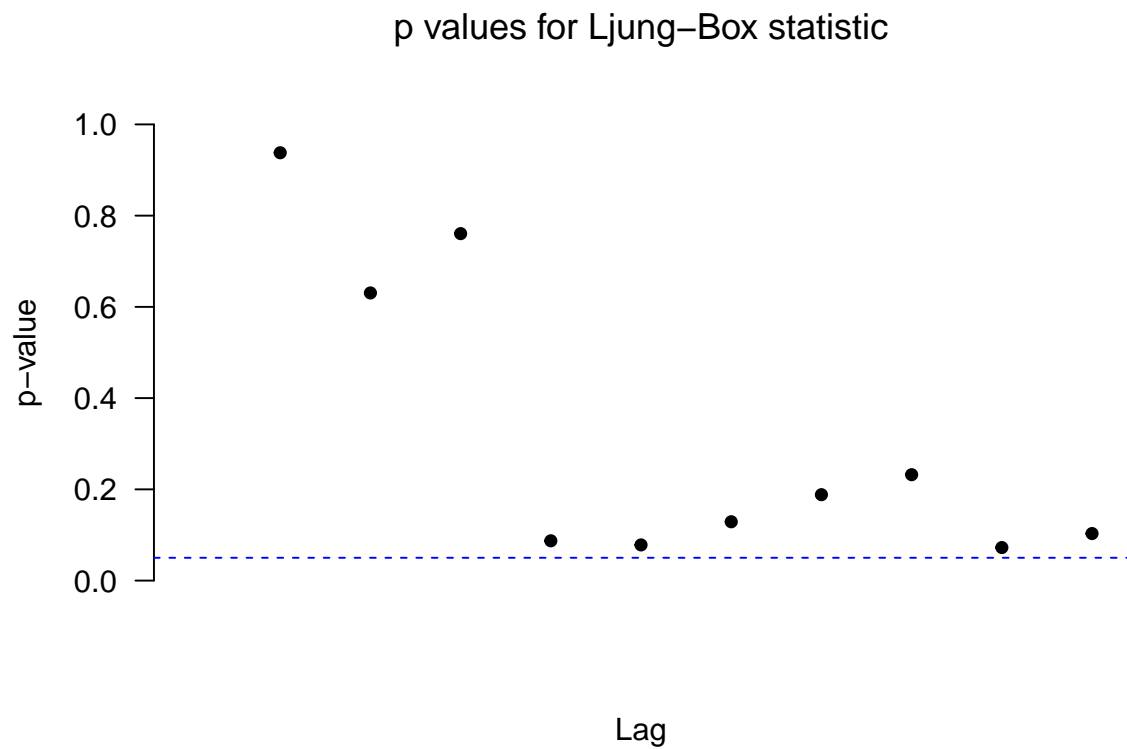
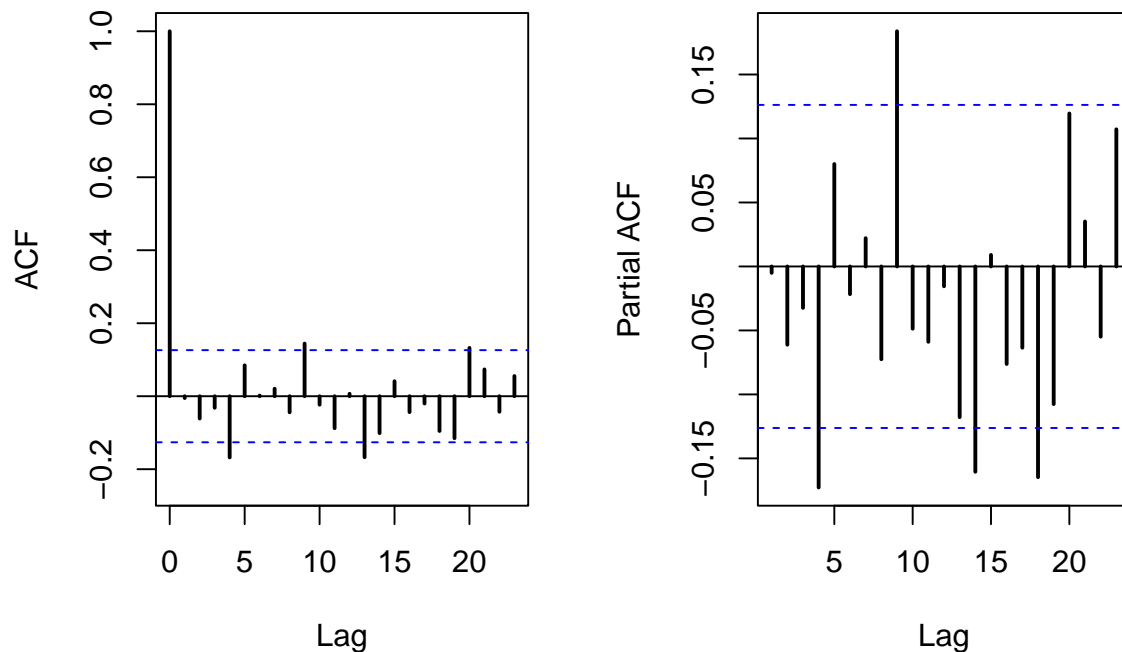


Figure 6: Gráficos de ACF y PCF para evaluación de autocorrelación en los residuos para SARIMA(6, 1, 1)(2, 1, 1)



(h) Con `forecast::forecast(...)` genere pronósticos hasta diciembre de 2027.

Considerando la comparación anterior, se opta por mantener el modelo propuesto por el software. Con este se realizarán las predicciones solicitadas.

```

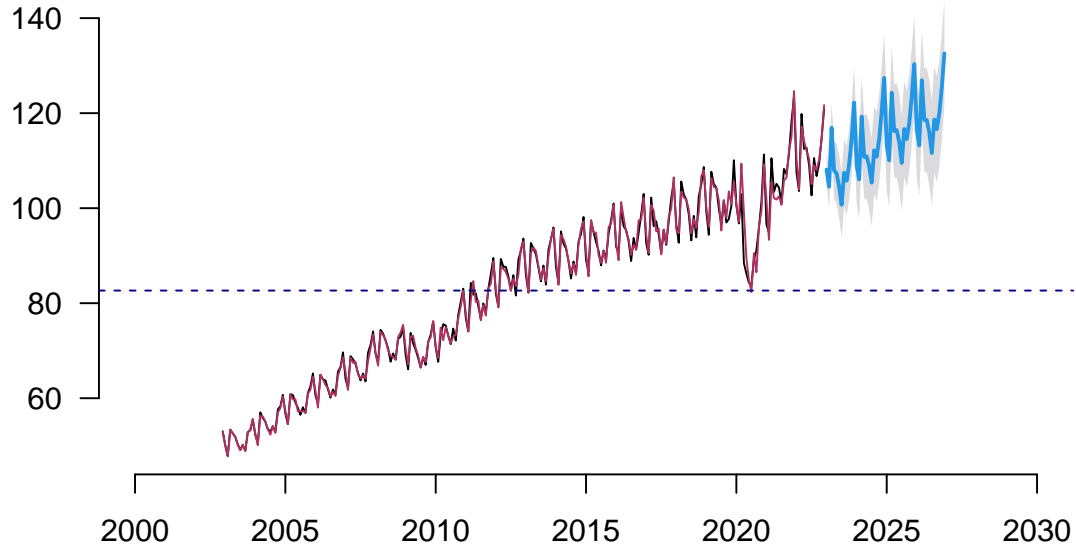
año_p <- 2023:2027
mes_p <- rep(1:12, 4)
tiempo_p <- año_p + (mes_p - 1)/12
tiempo_p <- tiempo_p[1:length(tiempo_p)] # Para empezar de Enero del 2023

df_preds <- forecast::forecast(m_auto,
                               h = length(tiempo_p), level = 0.95)

### Generar tabla de predicciones
P <- cbind(
  df_preds$mean,
  df_preds$lower,
  df_preds$upper)

```

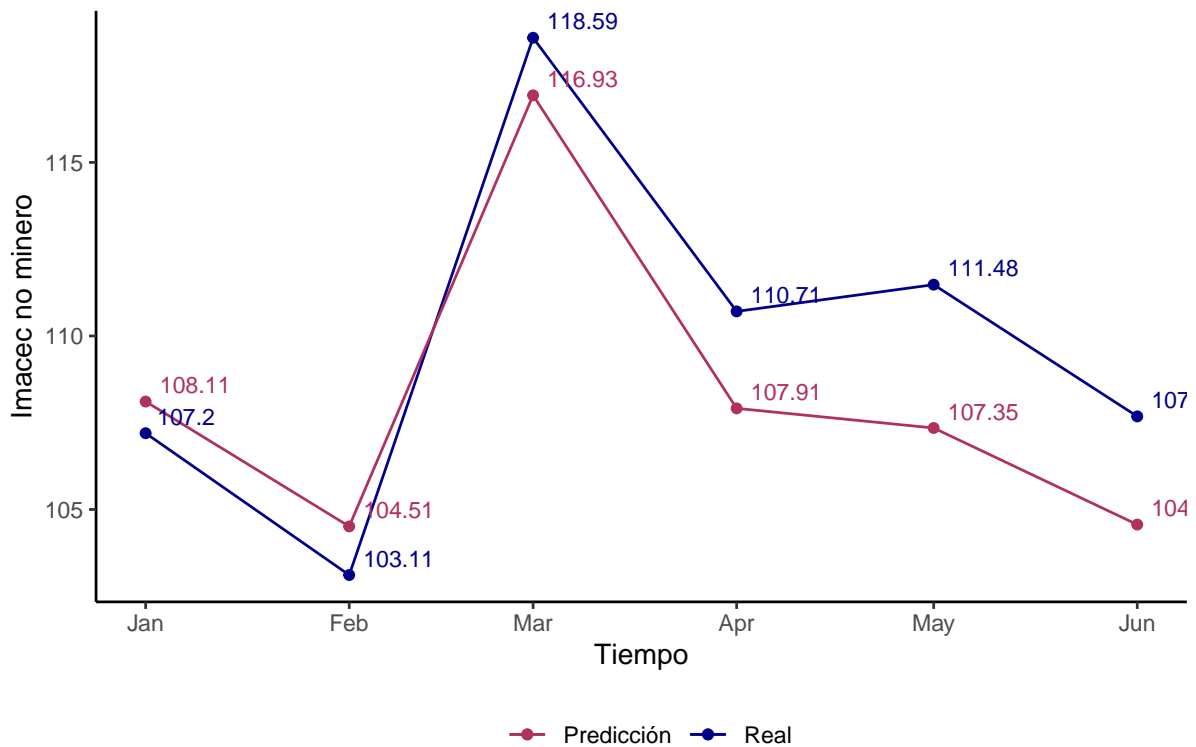
Figure 7: Predicción de Imacec no minero hasta 2027



En la Figura 7 se presenta la predicción del Imacec no minero hasta el 2027 en base al modelo SARIMA de orden $(2, 0, 1)(2, 1, 1)$. La línea azul representa los valores reales, la línea rosada representa los valores ajustados y la línea azul claro con contorno gris representa la predicción. A grandes rasgos, se observa que la tendencia y la estacionalidad de la serie se mantendrán, partiendo en enero del 2023 con una estimación de 108.11, bajando a 104.51 en febrero y manteniendo una fluctuación entre 105 y 116 hasta diciembre, donde la estimación tiene su mayor alza con un 122.21. Comparando los periodos de peak en la actividad económica no minera, se estima que para el 2024 el Imacec de diciembre tenga un alza porcentual del 4%, para el 2025 un alza de 4.4% y para el 2026 un alza del 1.68%, todos con respecto al valor del año anterior.

(i) Compare el desempeño de sus pronostico con respecto a los valores reales observados hasta junio de 2023.

Figure 8: Valores reales y predichos para Imacec no minero para el 2023



En la Figura 8 se comparan los valores predichos por el modelo para la primera mitad del 2023 con los valores reales del Imacec. Se observa que, en los primeros tres meses hay estimaciones bastante cercanas a los valores reales. Por ejemplo, para enero el valor real fue de 107.2, en tanto el valor predicho fue de 108.11 lo cual corresponde a una diferencia del 0.84%. Sin embargo, en los meses de abril, mayo y junio los valores predichos se alejan más de los valores reales. Por ejemplo, para junio el valor real del Imaced no minero es de 107, en tanto el valor predicho es de 104, lo cual corresponde a una diferencia porcentual del 2.8%.