

LICENCIATURA EN ESTADÍSTICA

“Series”

Trabajo Práctico 3

Autores: Tomás Anderson

Docentes:

05/06/2024

Tabla de contenidos

Introducción	1
Análisis descriptivo	1
Identificación del modelo	4
Análisis de residuos	8
Evaluación predictiva del modelo	11

Introducción

Argentina se destaca como uno de los principales exportadores de productos agrícolas a nivel mundial, lo que convierte al sector agrícola en un pilar esencial de su economía. Este sector no solo contribuye significativamente al Producto Interno Bruto (PIB), sino que también genera un gran número de empleos, haciendo que su prosperidad sea fundamental para el desarrollo del país.

Una métrica clave para cuantificar la actividad económica de este sector es el Valor Agregado Bruto (VAB). Esta magnitud económica mide el valor añadido generado por los productores en un área específica, reflejando el valor que se agrega a los bienes y servicios en las diversas etapas del proceso productivo. En términos sencillos, el VAB se calcula restando el total de costos de producción del valor del producto final, proporcionando una visión clara de la contribución económica real del sector.

Este estudio se centrará en el análisis del VAB trimestral del sector de agricultura, ganadería, caza y silvicultura desde 2004 hasta la actualidad ($n = 82$). El objetivo es modelar la evolución de esta métrica a lo largo del tiempo, lo que permitirá predecir el estado futuro del sector agrícola y su potencial desarrollo.

Análisis descriptivo

Como primer medida se grafica el VAB a traves del tiempo para observar la estructura que presenta los datos.

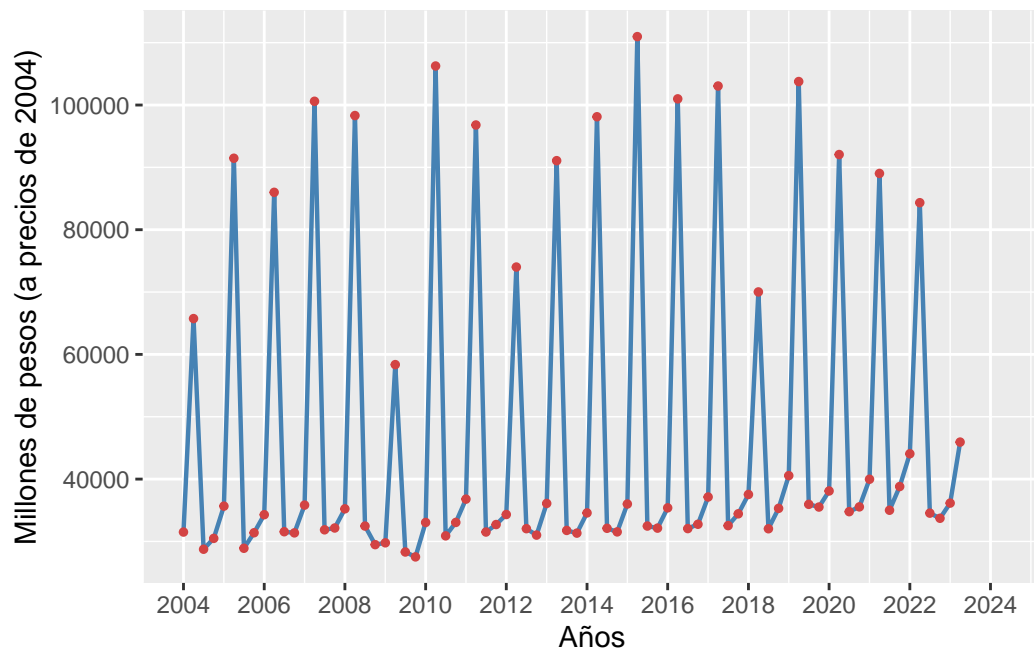


Figura 1: Serie del VAB. Periodo 2004-2024

En la figura 1 se puede observar que el Valor Agregado Bruto (VAB) trimestral no presenta un comportamiento estacionario en términos de la media, ya que muestra un incremento continuo a lo largo del tiempo. Es importante destacar que los segundos trimestres de cada año son los períodos en los que se registran los mayores ingresos en los sectores de agricultura, ganadería, caza y silvicultura. Además, la variabilidad del VAB no parece estar relacionada con el paso del tiempo.

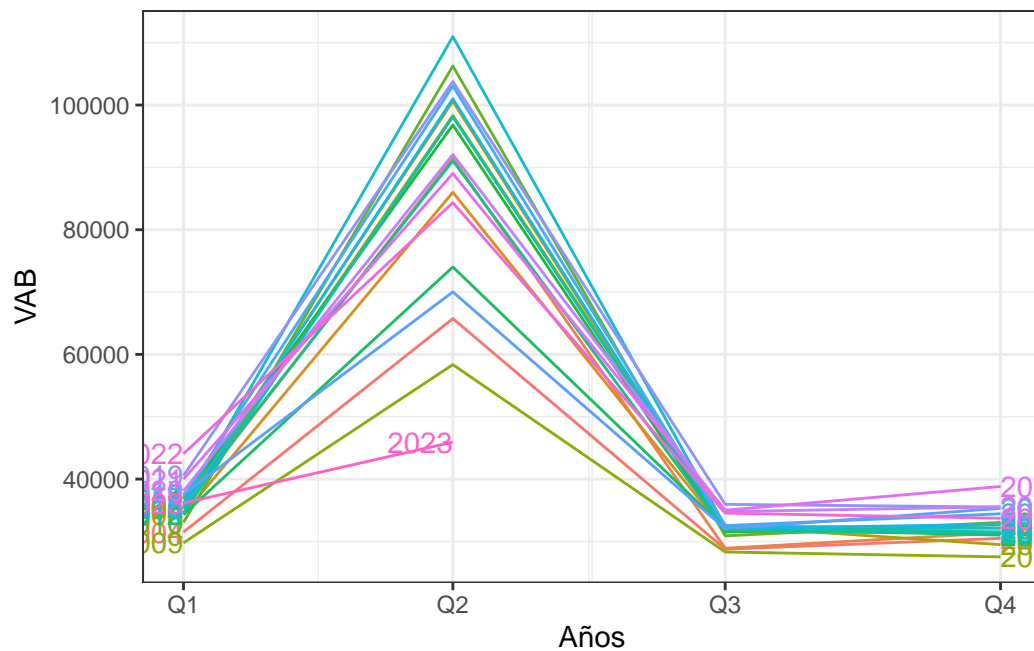


Figura 2: Comportamiento anual del VAB

En la figura 2 se puede apreciar un comportamiento estacional en el VAB, donde los primeros, terceros y cuartos trimestres presentan patrones bastante similares. En contraste, el segundo trimestre destaca por un VAB significativamente superior al del resto del año, tendencia que se repite cada año. Esto indica que el período más favorable para la agricultura se concentra entre abril y junio.

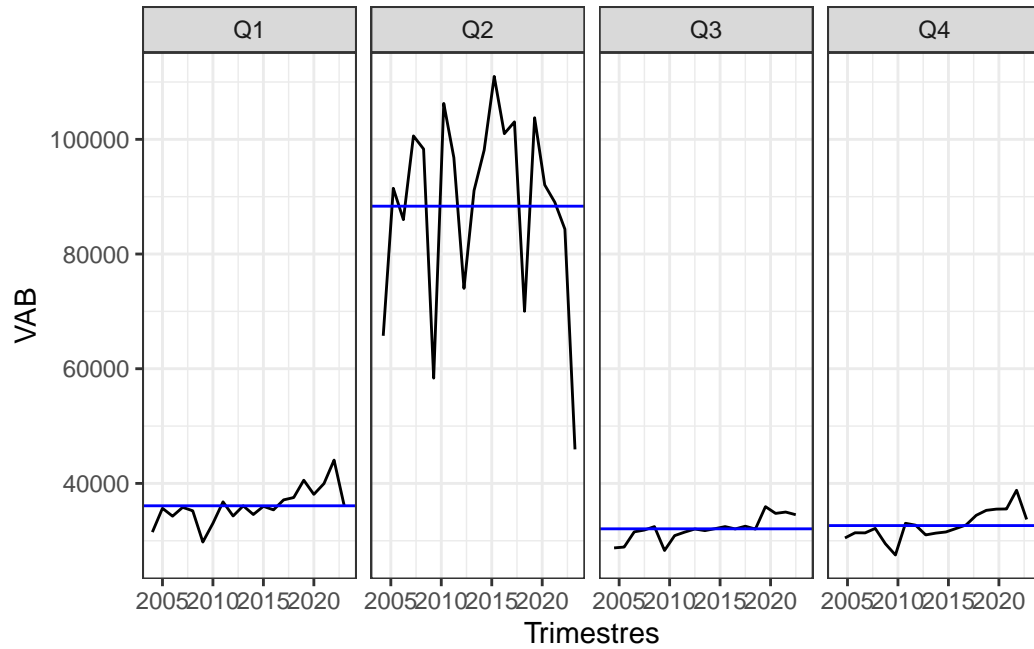


Figura 3: Comportamiento del

Como en la anterior, en la figura 3 se observa claramente la diferencia en el VAB medio entre los trimestres. Además, se puede apreciar de manera más evidente la falta de estacionaridad, ya que el VAB parece mostrar un aumento constante a lo largo de los años.

Identificación del modelo

Antes de proceder con la identificación del modelo, se reservan las últimas 4 observaciones para evaluar su capacidad predictiva. De este modo, la identificación del modelo se lleva a cabo con las primeras 78 observaciones de la serie.

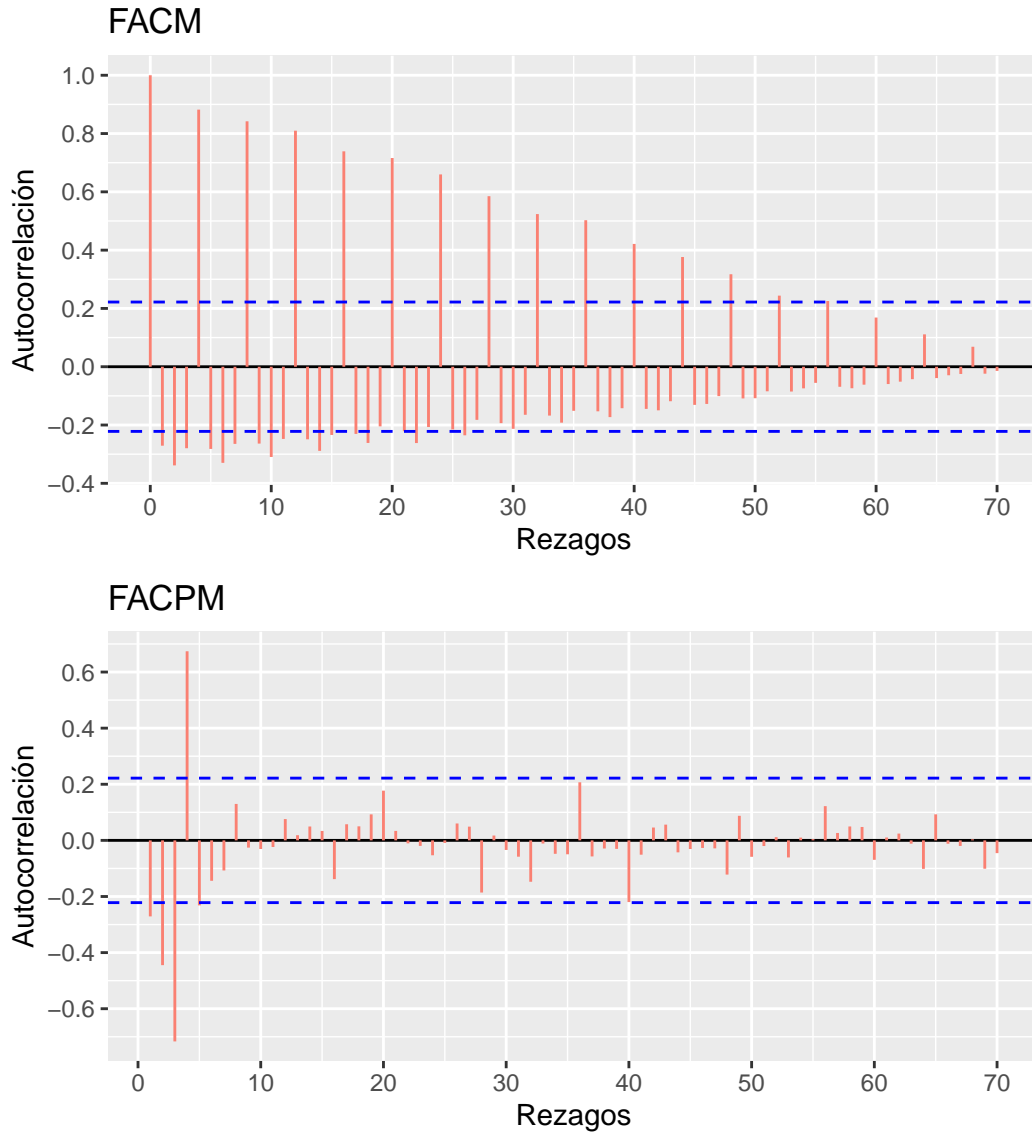
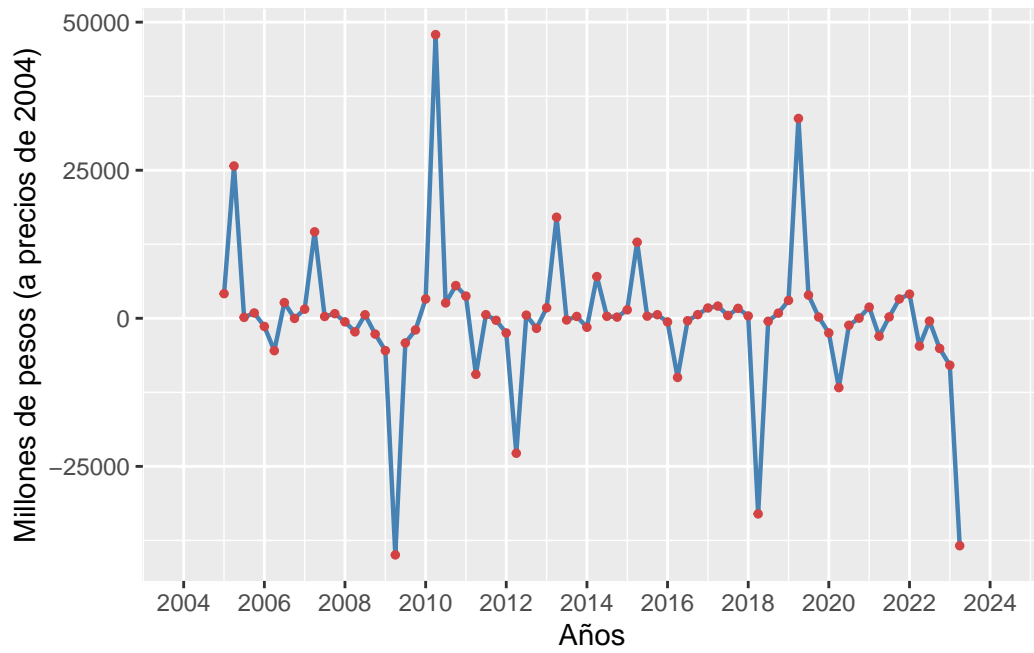


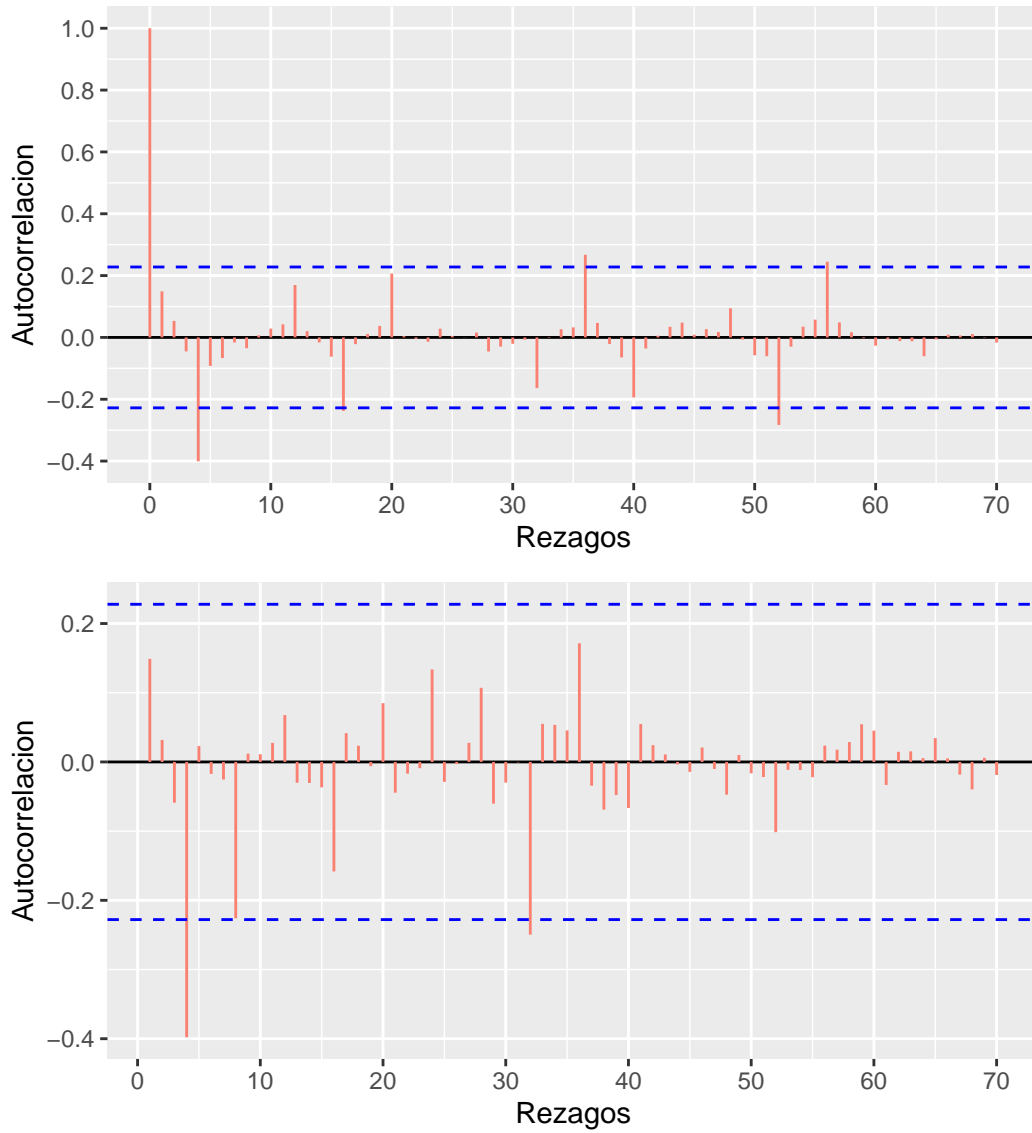
Figura 4: Autocorrelación y autocorrelación parcial muestral

En la figura 4 se evidencia la necesidad de realizar una diferenciación estacional, ya que los valores de los rezagos múltiplos de 4 de la Función de Autocorrelación Muestral (FACM) muestran un decrecimiento lineal. Esto sugiere que, al ajustar un modelo autoregresivo en la parte estacional, la estimación del parámetro estará próxima a 1.

Se aplica una diferenciación estacional, lo que resulta en la siguiente serie:

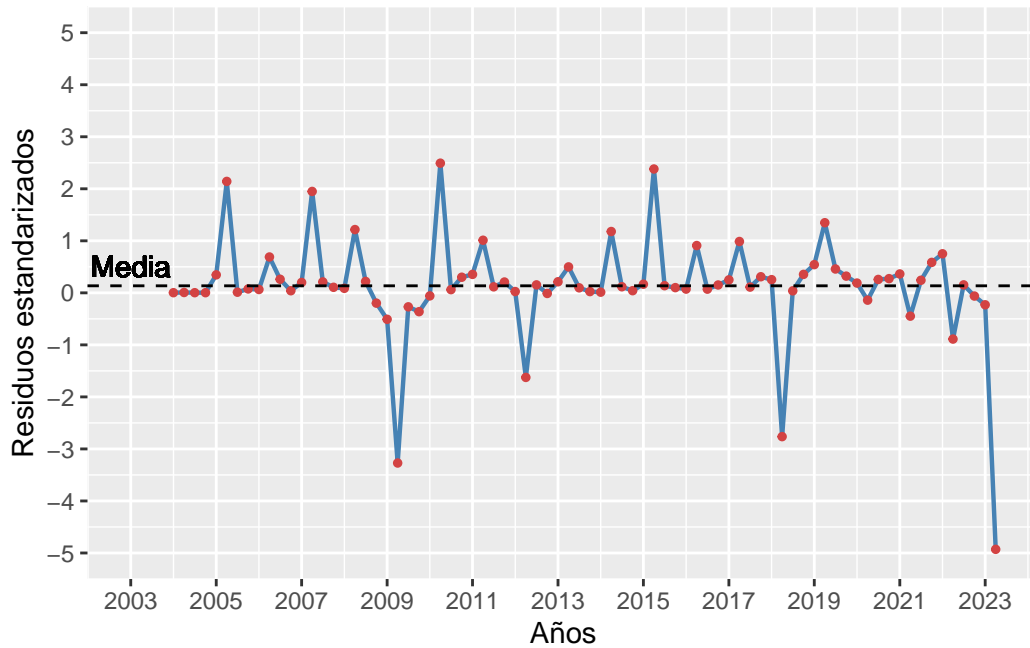


Al aplicar la diferencia, la nueva serie de tiempo resulta estacionaria en media.

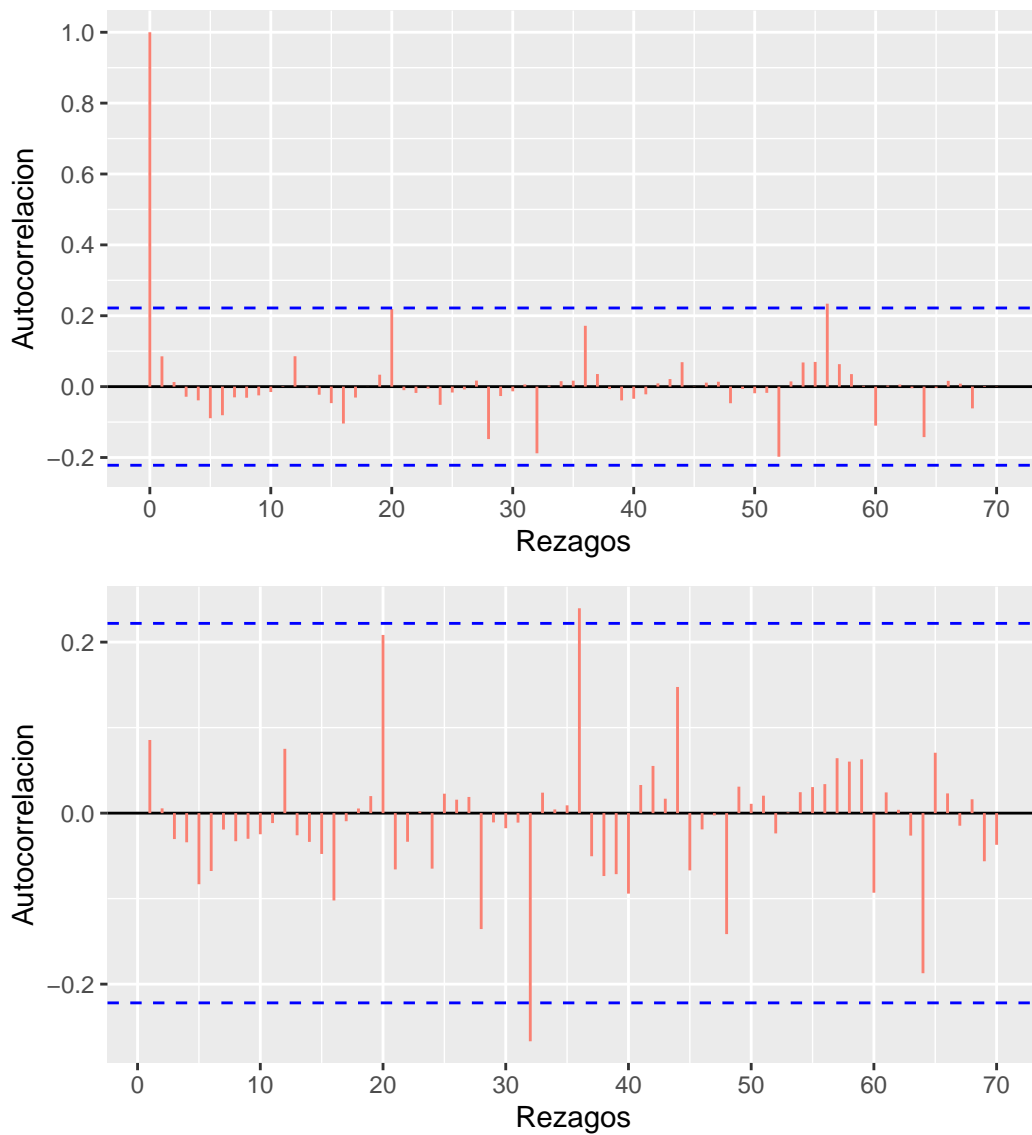


Según la estructura de autocorrelación y autocorrelación parcial muestral presentada en la figura X, se concluye que la componente estacional puede modelarse con un $MA(1)$, dado que se observa un decrecimiento exponencial de los rezagos estacionales en la FACPM, y solo el rezago 4 resulta significativo en la FACM. En cuanto a la componente regular, no se identifican valores significativos, por lo que se modelará como ruido blanco. Así, el modelo propuesto es un $SARIMA(0,0,0)(0,1,1)_4$.

Analisis de residuos



Se puede observar que la media de los residuos es aproximadamente igual a 0. Además, de forma visual, parece corroborarse que la varianza de los residuos se mantiene constante a lo largo del tiempo, lo que sugiere que se cumple la hipótesis de homocedasticidad. Sin embargo, se identifican dos observaciones cuyos residuos se encuentran fuera del intervalo $[-3, 3]$, lo que podría indicar que dichas observaciones son atípicas o outliers.

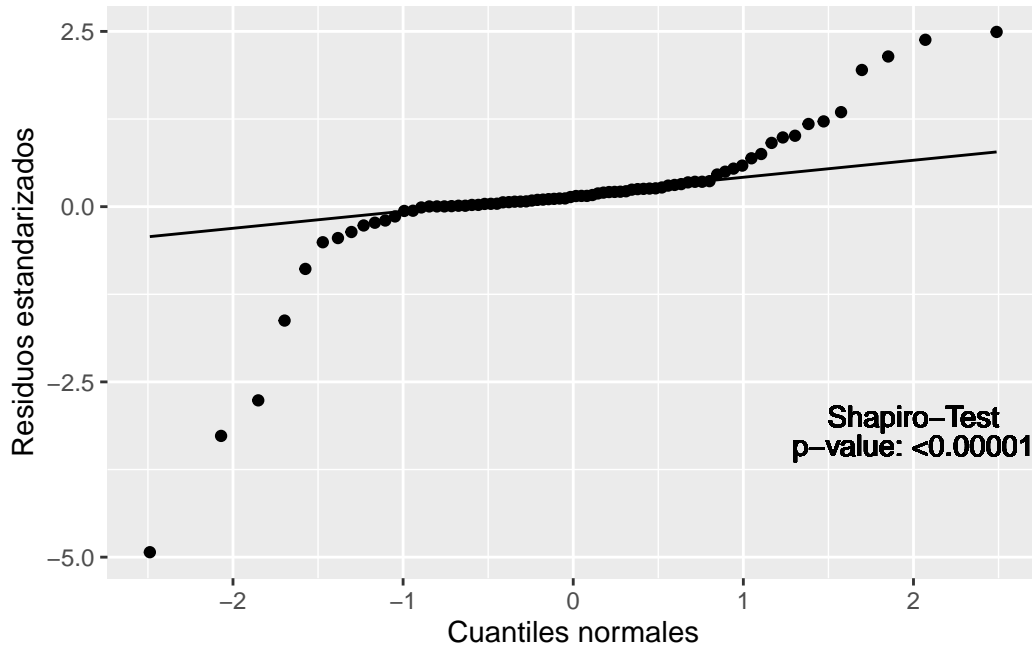


Visualmente, solo dos rezagos en la función de autocorrelación parcial muestral de los residuos caen ligeramente fuera de las bandas de significación. Esto sugiere que no hay evidencia clara de correlación en los residuos. Sin embargo, para confirmar de manera más rigurosa que estos residuos sean ruido blanco, se procede a realizar la prueba de Ljung-Box.

```
[1] 0.9383558 0.9737784 0.9956534 0.9981755 0.9731291 0.9942762 0.9934424
[8] 0.9798928 0.9647485 0.9863599 0.9941211 0.9981332 0.9747396 0.6224440
[15] 0.5526384 0.3752179 0.4263317 0.1011966 0.1564359          NA
```

Introducir tabla.

No se rechaza ningún test, por lo que se cumple el supuesto de incorrelación de los residuos. Esto indica que no hay evidencia suficiente para sugerir que los residuos estén correlacionados, lo que respalda la hipótesis de que los residuos son independientes y, por lo tanto, se puede considerar que el modelo ajustado cumple con este supuesto.



En la figura X se observa que los cuantiles de los residuos no se ajustan a los de una distribución normal. Además, se aplicó el test de Shapiro-Wilk sobre los residuos para evaluar su normalidad, el cual rechaza la hipótesis de que los residuos siguen una distribución normal.

Series: agricultura
Model: ARIMA(0,0,0)(0,1,1)[4]

Coefficients:

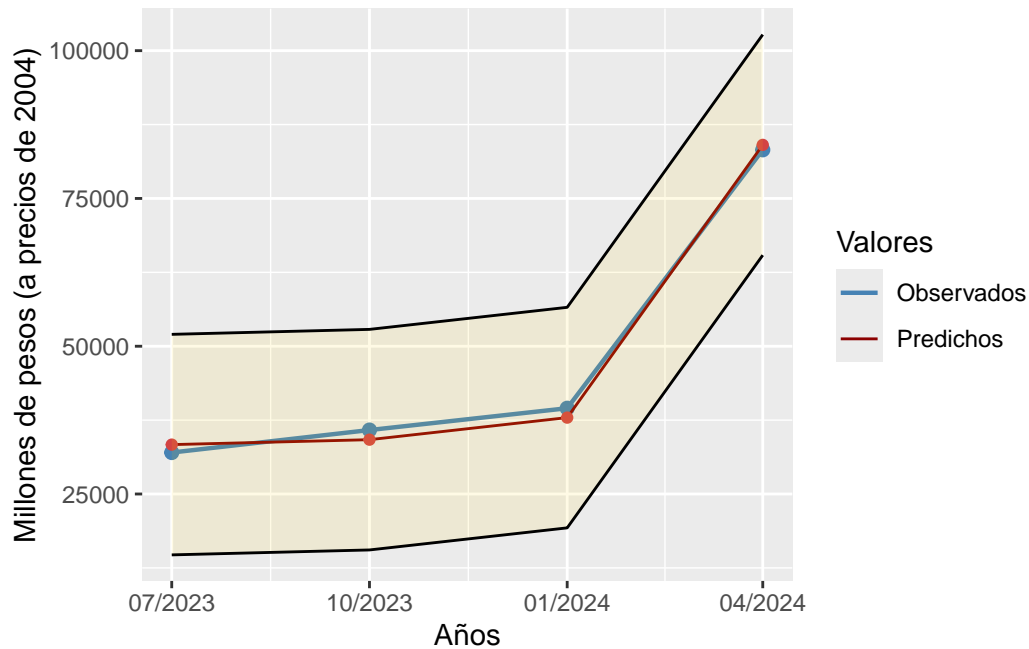
 sma1
 -0.8426
s.e. 0.1833

sigma^2 estimated as 90559331: log likelihood=-784.87
AIC=1573.73 AICc=1573.9 BIC=1578.34

Dado que el modelo planteado tiene un único parámetro, se puede considerar como el modelo más parsimonioso posible.

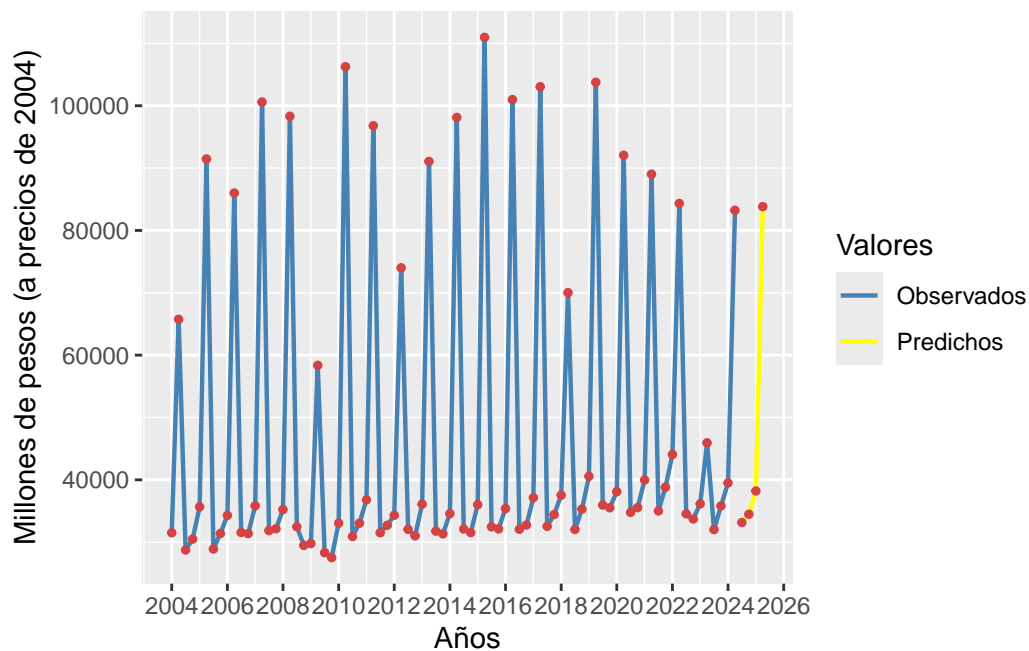
Evaluacion predictiva del modelo

```
colors <- c("Observados" = "steelblue", "Predichos" = "darkred")
prediccion = modelo %>% forecast(h = 4) %>% as.data.frame() %>% mutate(indice_tiempo = datos_
datos_total[79:82,] %>%
  ggplot() +
  geom_line(aes(x = indice_tiempo,
                y = agricultura,
                color = "Observados"),
            size = .78) +
  labs(x = "Años",
        y = "Millones de pesos (a precios de 2004)",
        color = "Valores") +
  scale_x_date(breaks = seq(from = as.Date("2004-01-01"),
                             to = as.Date("2024-04-01"), by = "3 month"),
               date_labels = "%m/%Y") +
  geom_point(aes(x = indice_tiempo, y = agricultura), color = "steelblue", size = 2) +
  geom_line(data = prediccion, aes(x = indice_tiempo, y = .mean, color = "Predichos")) +
  geom_point(data = prediccion, aes(y = .mean, x = indice_tiempo, color = ), color = "#D34343") +
  geom_ribbon(data = prediccion,
             aes(ymin=quantile(agricultura, p = .975),
                 ymax=quantile(agricultura, p = .025),
                 x = indice_tiempo), alpha=0.1, color = "black", fill = "gold") + scale_col
```



```
mape = mean(abs((datos_total[79:82,2] - prediccion$.mean)/datos_total[79:82,2]))[,1])*100
```

El MAPE es igual a 3.4362526%, lo que indica que las estimaciones del modelo son bastante precisas.



Se puede observar que el modelo predice que el final del año 2024 y el comienzo del 2025 serán similares al año anterior. Esto sugiere que las condiciones que están modeladas y observadas en el período actual tienden a mantenerse estables en el futuro cercano, según las tendencias identificadas por el modelo. Esta predicción podría implicar que no se anticipan cambios significativos en las variables de interés a corto plazo.