

# PROYECTO DE GRADO

Modelo predictivo para el precio  
nacional del cacao.

Karen Dayana Culma Ramírez

Asesor: Santiago Pérez Angarita

Co-Asesor: Adriana Abrego Pérez

2022-1

# CONTENIDOS

01

## INTRODUCCIÓN

Contexto y motivación

02

## OBJETIVOS

03

## DATOS

Adquisición, preprocesamiento y base de datos.

04

## METODOLOGÍA

Exploración de modelos predictivos.

05

## RESULTADOS

Selección de mejor modelo.

06

## CONCLUSIONES Y RECOMENDACIONES

07

## BIBLIOGRAFIA

08

## ANEXOS

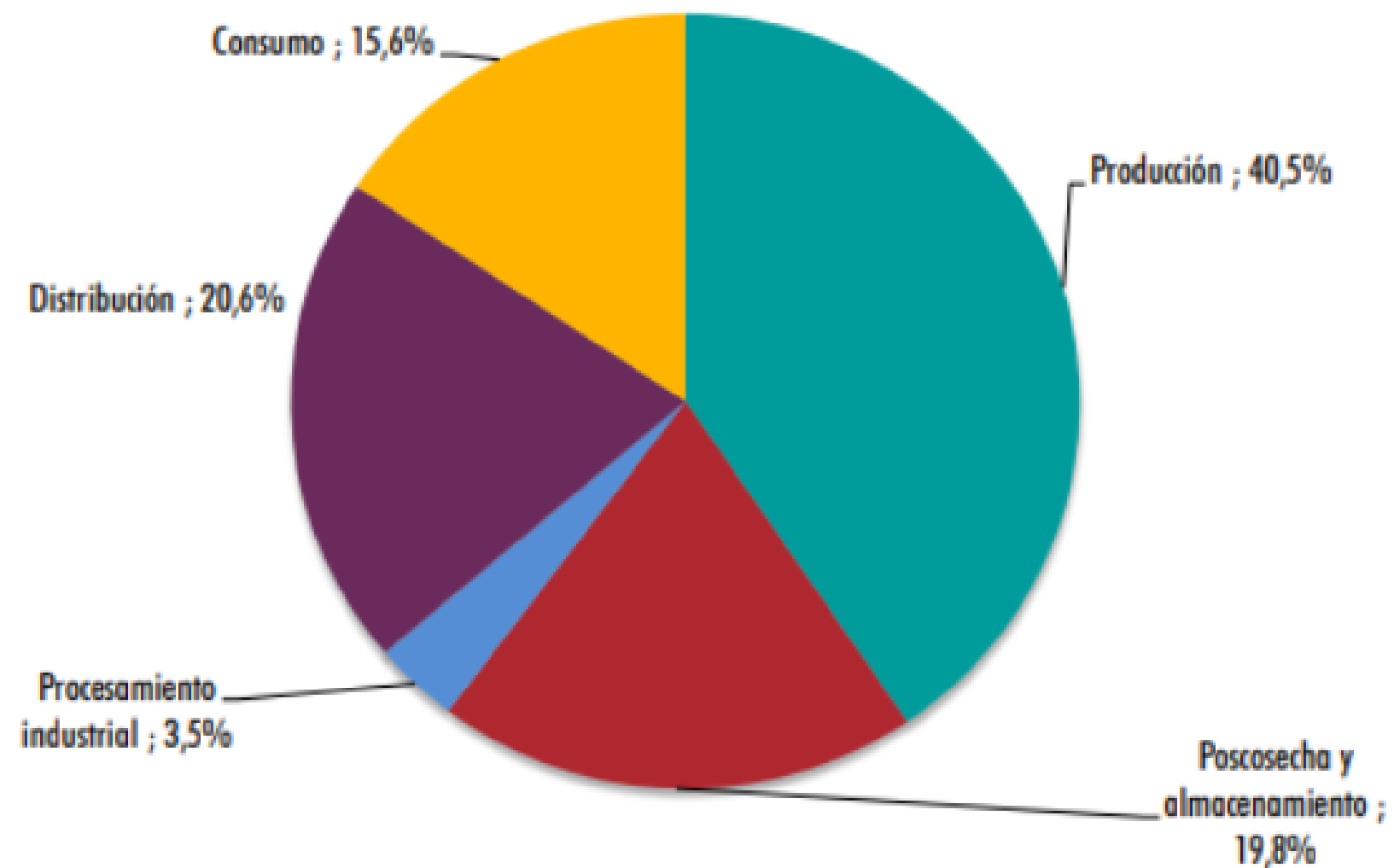
Modelo predictivo univariado para el precio Internacional de cacao.  
Aplicativo

# INTRODUCCIÓN

**Contribución al PIB (7% en 2017)**

**Demanda de productos agrícolas crece.**

**Se pierde y desperdicia 9,76M de t.**



Fuente: cálculos DNP.

## COMERCIALIZACIÓN DE CACAO

- Mercado oligopsónico.
- Precio Internacional como referente.
- Casa Luker y Grupo Nutresa tienen 85% de participación.



Tomado de <https://sac.org.co/wp-content/uploads/2019/01/CACAO.jpg>

# PREGUNTA DE INVESTIGACIÓN

¿Cómo pueden los modelos predictivos ser de ayuda para conocer el comportamiento futuro del precio nacional del cacao y así apoyar en las decisiones sobre la producción?

## OBJETIVOS



**Analizar y consolidar una base de datos.**



**Construir diferentes modelos predictivos.**



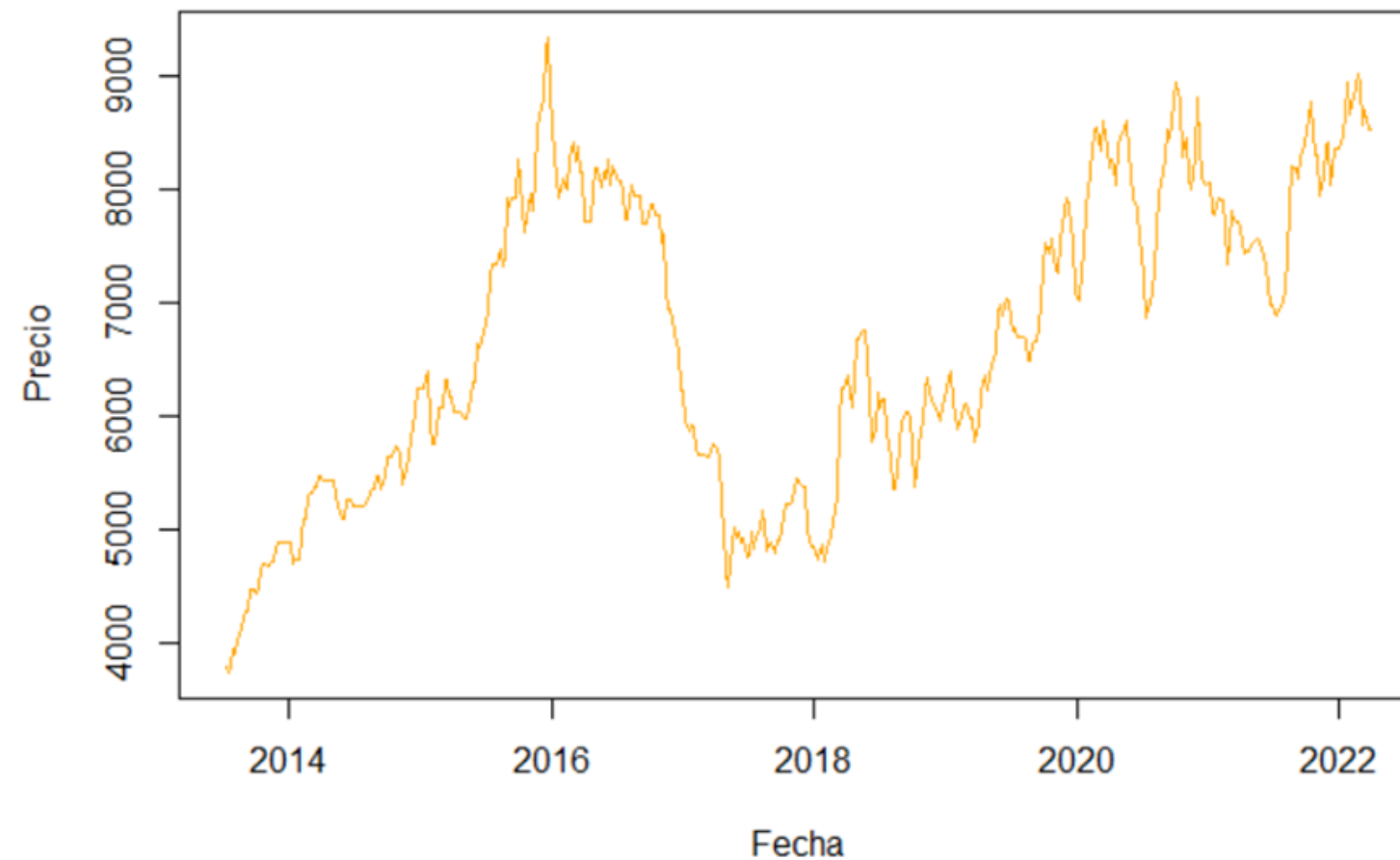
**Seleccionar el mejor modelo.**

# DATOS

01

## PRECIO NACIONAL

Suministrada por el Fondo Nacional del Cacao a través de la red de información y comunicación del sector Agropecuario Colombiano.

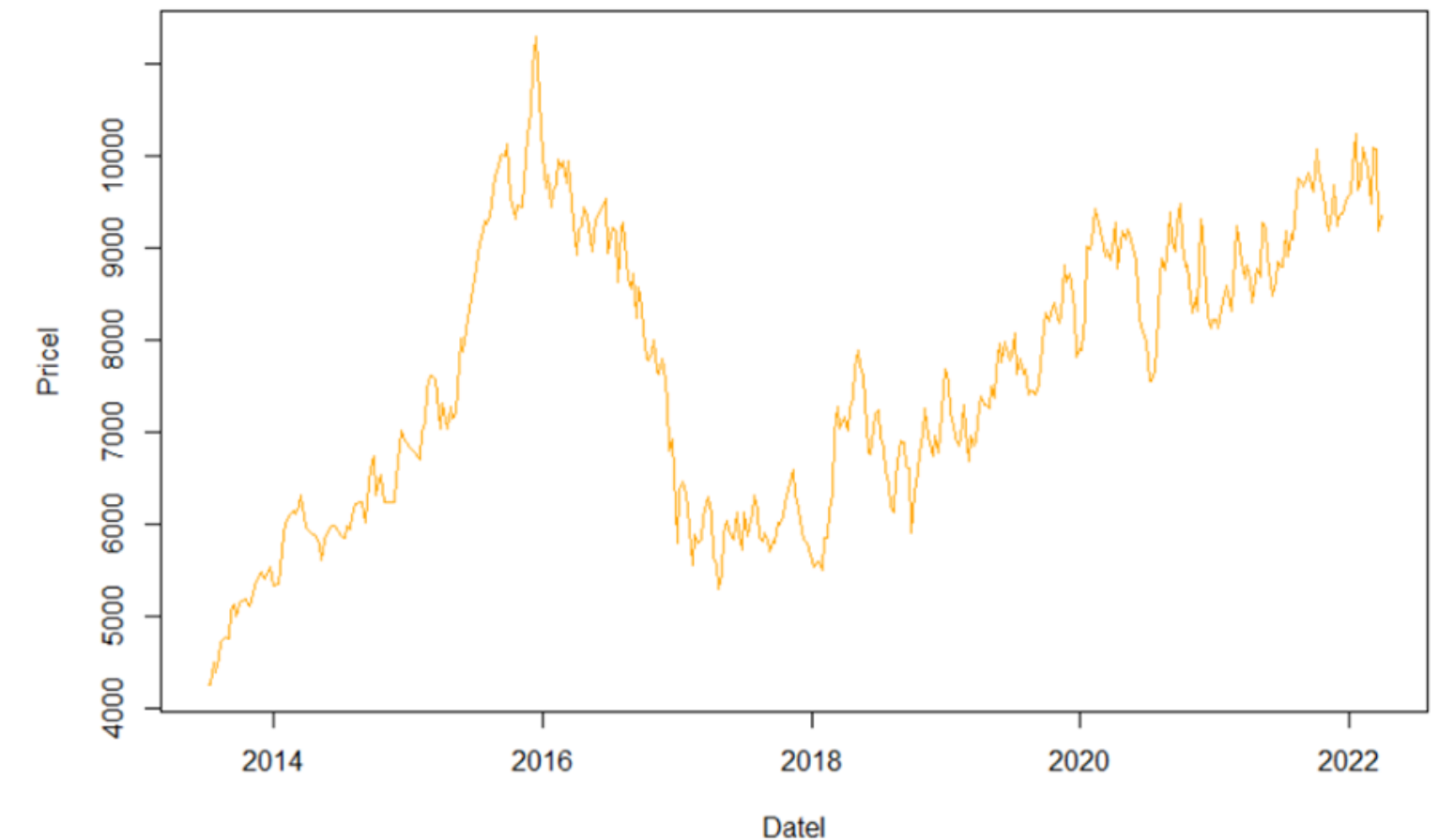


Media	Varianza	Desviación Estándar	Curtosis	Sesgo	Máximo	Mínimo	Cuenta
6.669,74	1775981	1.332,66	-1,2167	-642,2	9.342,5	3.737,5	455

02

## PRECIO INTERNACIONAL

Se cotiza en la bolsa de valores de Nueva York, suministrado por IFC Markets e IndexMundi



Media	Varianza	Desviación Estándar	Curtosis	Sesgo	Máximo	Mínimo	Cuenta
7.644,5	2266295	1.505,42	-1,056	-254,24	11.283,75	4.253,62	407

# DATOS

03

## OTRAS VARIABLES

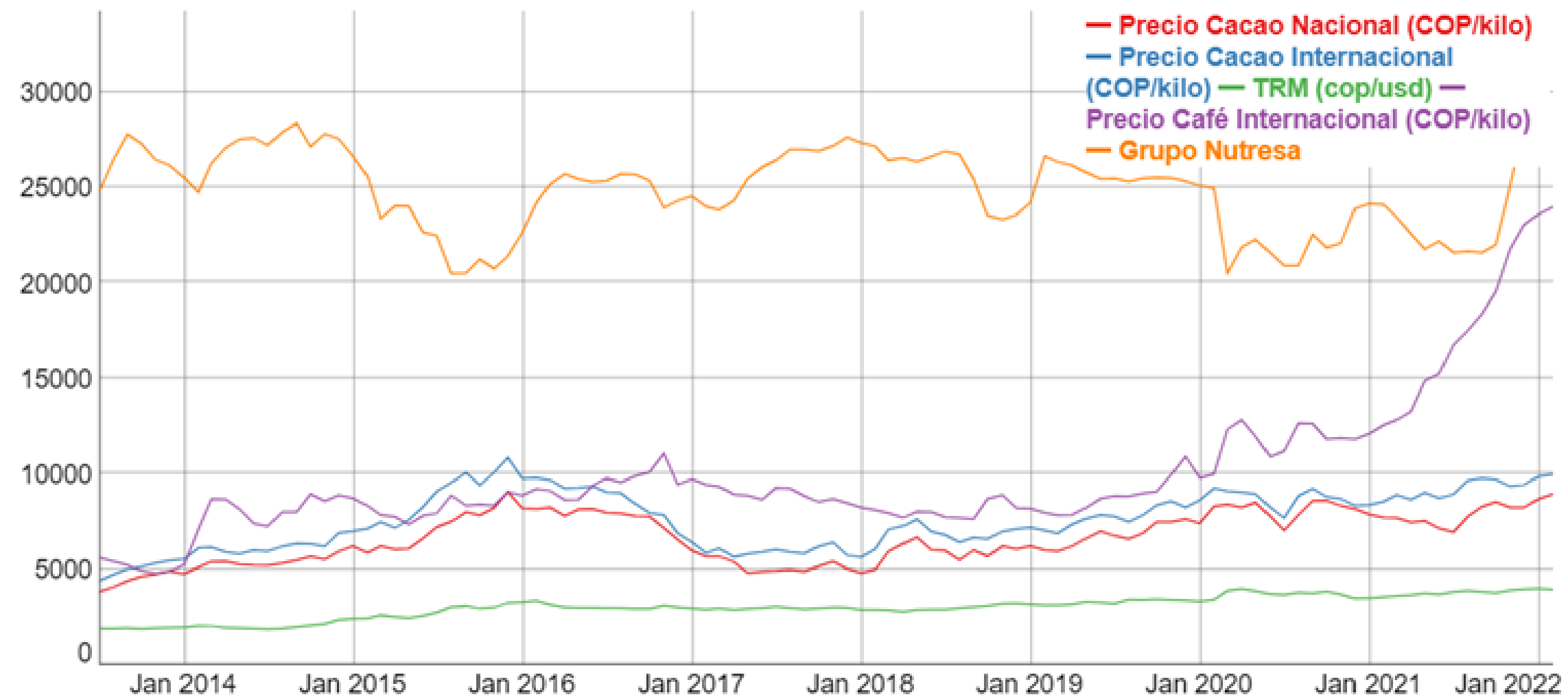
Precio internacional del cacao.

TRM

Precio del café,

Valor de importación y exportación

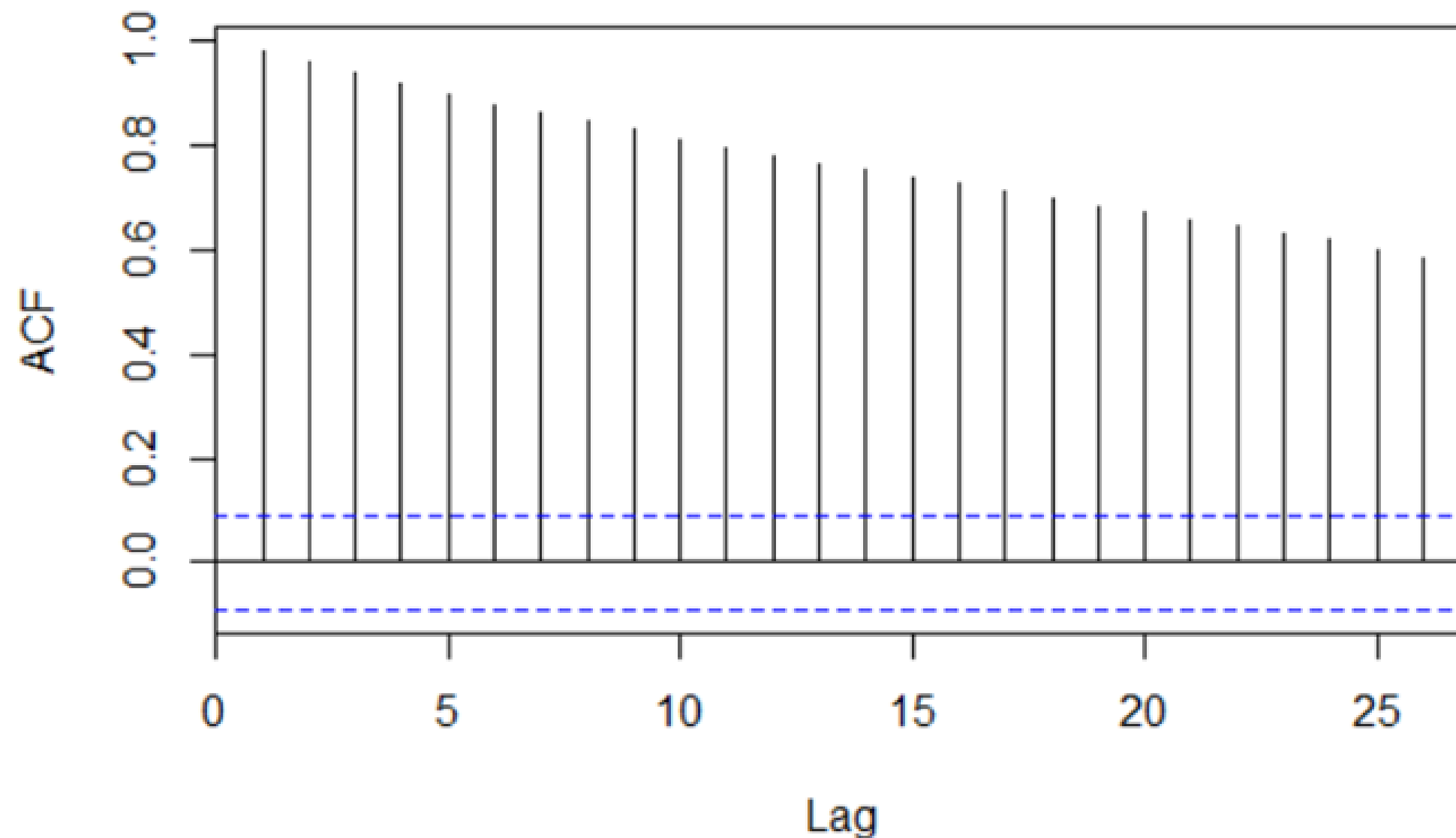
Cotización de Grupo Nutresa.



# METODOLOGÍA

## ARIMA UNIVARIADO

Metodología de Box-Jenkins



## Prueba Dickey-Fuller

H0: La serie no es estacionaria.

H1: La serie es estacionaria

Augmented Dickey-Fuller Test

data: Price

Dickey-Fuller = -2.0782, Lag order = 7, p-value = 0.5448

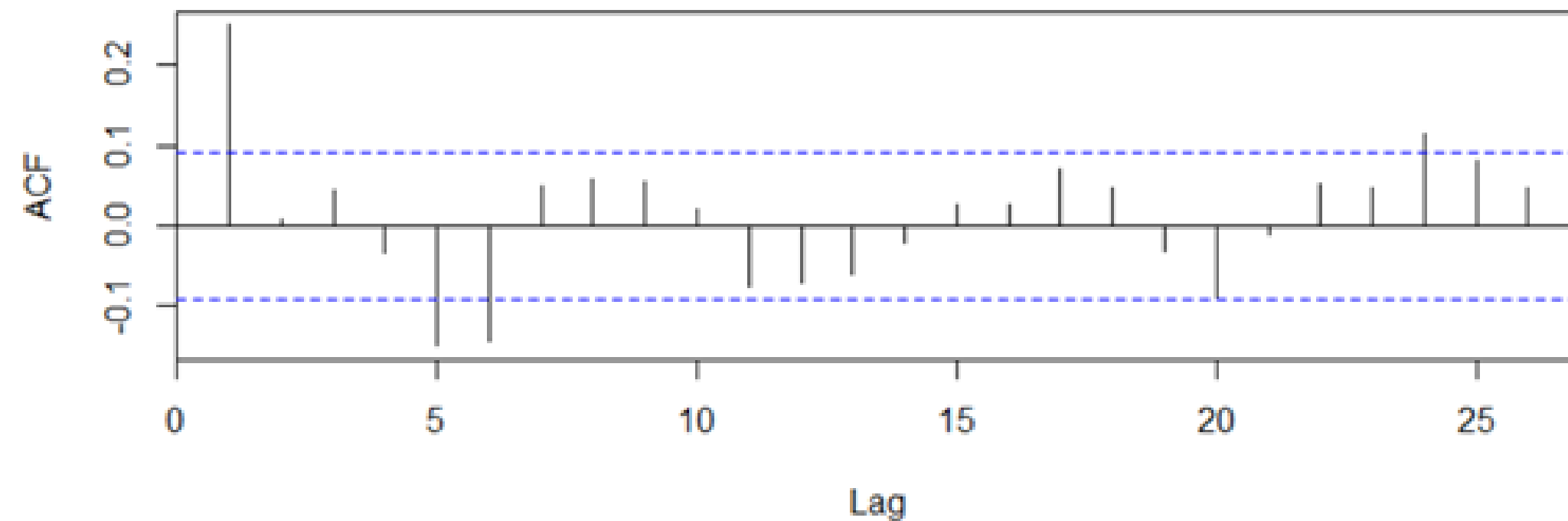
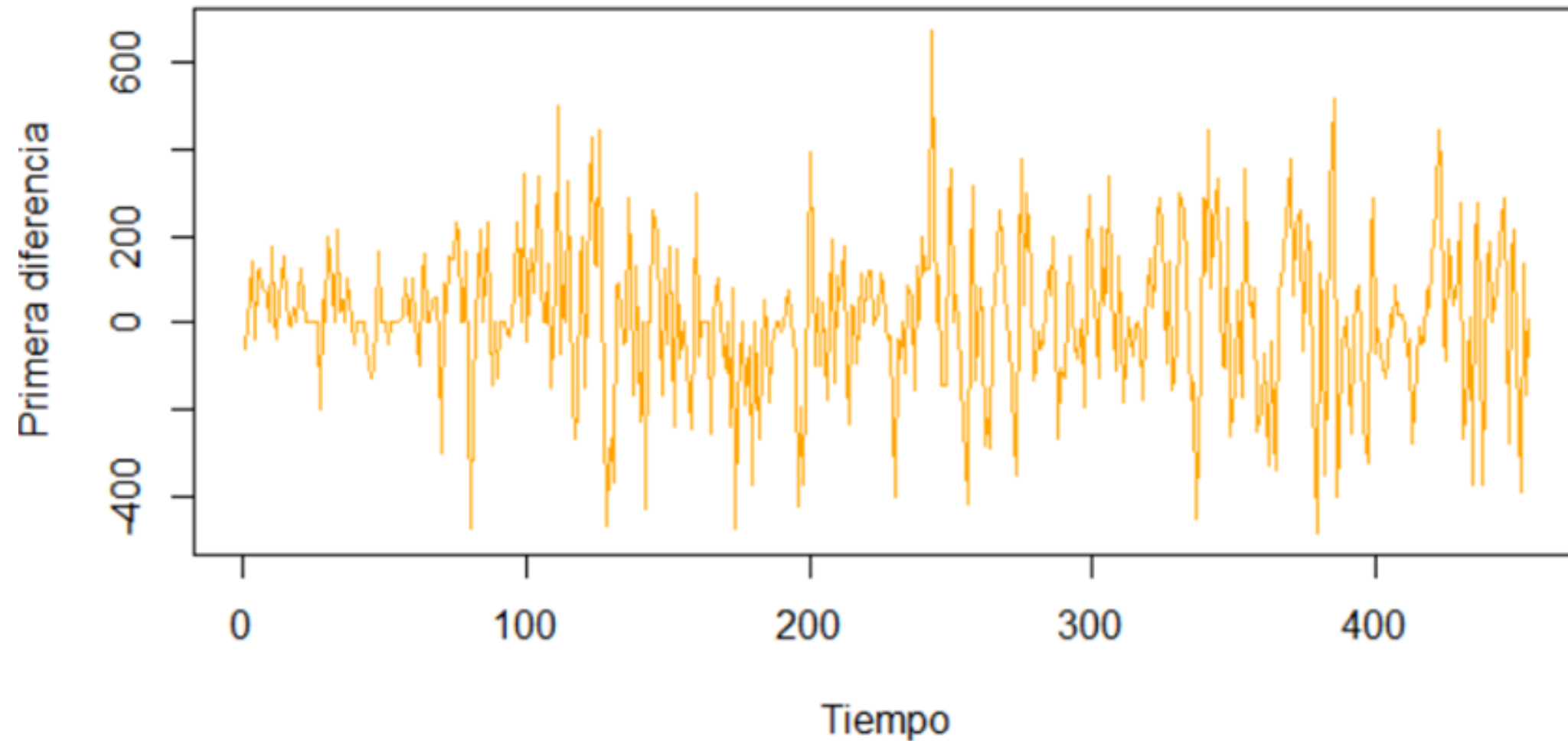
alternative hypothesis: stationary

La serie no es estacionaria



# ARIMA UNIVARIADO

Metodología de Box-Jenkins



## Prueba Dickey-Fuller

H0: La serie no es estacionaria.

H1: La serie es estacionaria

Augmented Dickey-Fuller Test

```
data: difserie
Dickey-Fuller = -7.4171, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

**La serie es estacionaria**



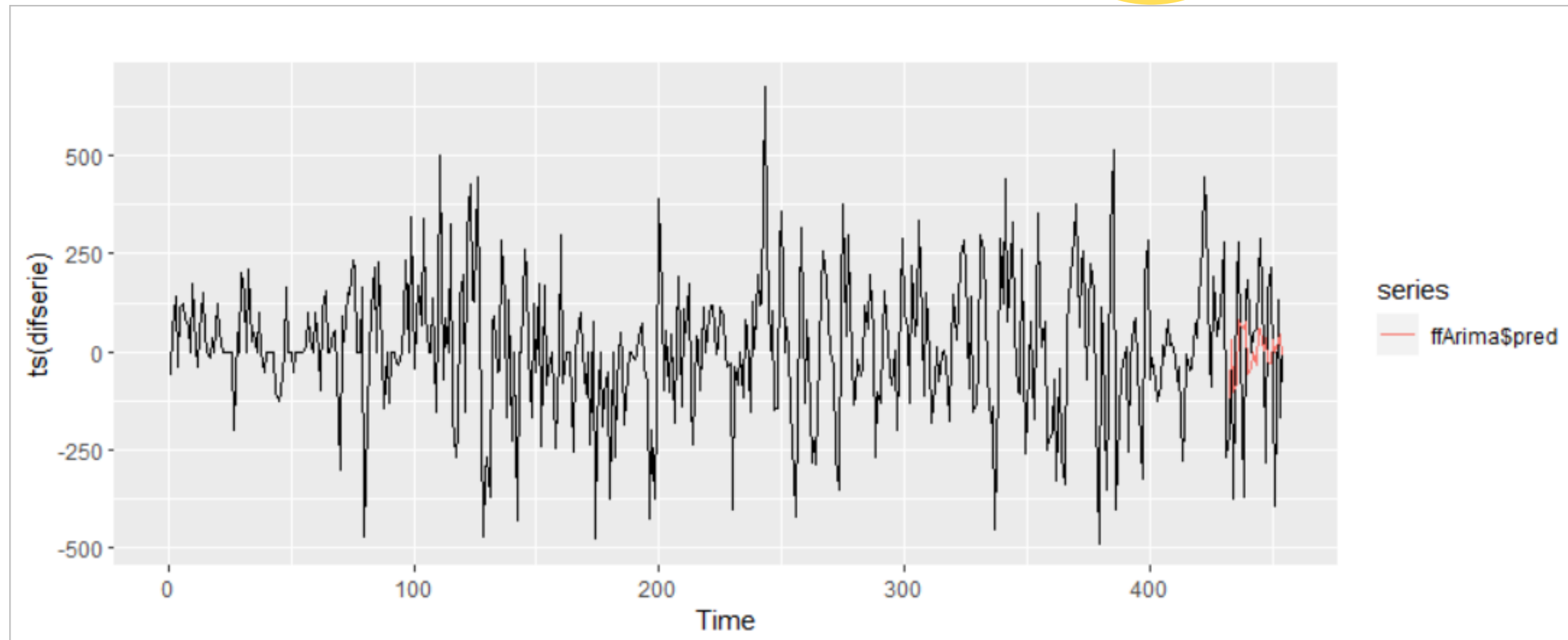
# ARIMA UNIVARIADO

Metodología de Box-Jenkins

Entrenamiento-95%

Prueba-5%.

Modelo	ARIMA(0,0,1) (auto.arima)	ARIMA(1,0,1)	ARIMA(1,0,5)	ARIMA(5,0,5)	ARIMA(5,0,6)	ARIMA(6,0,7)
AIC	5649.98	5650,9	5647,84	5636,88	5635,84	5638,09



# METODOLOGÍA

## Modelo Autorregresivo Vectorial (VAR)

### Prueba Dickey-Fuller en diferenciación

Serie	Valor p
Precio Nacional del Cacao	0,01492
Precio Internacional del Cacao	0,02011
TRM	0,01
Precio del Café	0,09974
Importación	0,01
Exportación	0,01
Grupo Nutresa	0,04566

Las series son estacionarias

```
Estimation results for equation cacaoNacional:
=====
cacaoNacional = cacaoNacional.l1 + cacaoInternacional.l1 + trm.l1 +
cafe.l1 + importación.l1 + exportacion.l1 + nutresa.l1 + const

              Estimate Std. Error t value Pr(>|t|)
cacaoNacional.l1  -4.560e-01  1.596e-01  -2.858  0.00527 **
cacaoInternacional.l1  6.873e-01  1.524e-01   4.511  1.88e-05 ***
trm.l1             -2.743e-01  4.361e-01  -0.629  0.53091
cafe.l1             4.984e-02  5.977e-02   0.834  0.40652
importación.l1     -1.692e-05  1.411e-05  -1.199  0.23370
exportacion.l1     -1.292e-06  7.505e-06  -0.172  0.86367
nutresa.l1         -8.841e-03  2.893e-02  -0.306  0.76060
const              2.749e+01  3.605e+01   0.763  0.44765
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 347.6 on 93 degrees of freedom
Multiple R-Squared: 0.2151,    Adjusted R-squared: 0.1561
F-statistic: 3.642 on 7 and 93 DF,  p-value: 0.001626
```

# METODOLOGÍA

## Modelo Autorregresivo Vectorial (VAR)

Entrenamiento-95%

Prueba-5%.

Estimation results for equation precioNacional1d:

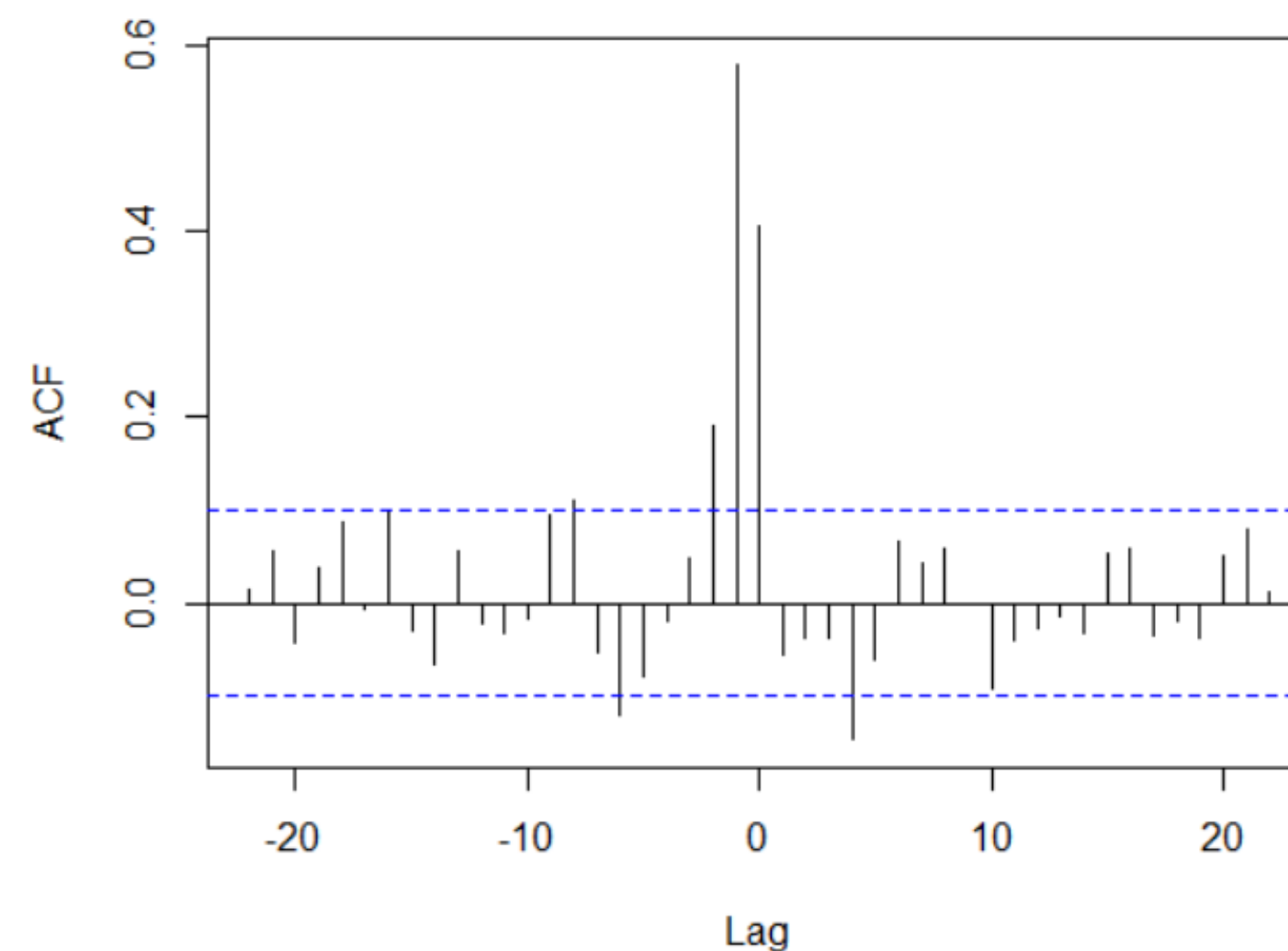
```
=====
precioNacional1d = precioNacional1d.l1 + precioInternacionald.l1 + p
recioNacional1d.l2 + precioInternacionald.l2 + precioNacional1d.l3 +
precioInternacionald.l3 + const
```

	Estimate	Std. Error	t value	Pr(> t )	
precioNacional1d.l1	-0.18638	0.05880	-3.170	0.00165	**
precioInternacionald.l1	0.45367	0.03278	13.840	< 2e-16	***
precioNacional1d.l2	-0.12750	0.05718	-2.230	0.02635	*
precioInternacionald.l2	0.24134	0.04123	5.854	1.05e-08	***
precioNacional1d.l3	0.03571	0.04461	0.800	0.42396	
precioInternacionald.l3	0.10361	0.03923	2.641	0.00860	**
const	3.71695	7.72676	0.481	0.63076	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 150.3 on 376 degrees of freedom  
Multiple R-Squared: 0.3925, Adjusted R-squared: 0.3828  
F-statistic: 40.49 on 6 and 376 DF, p-value: < 2.2e-16

trainVAR\$precioInternacionald & trainVAR\$precioNacional1



# METODOLOGÍA

## Modelo Autorregresivo Vectorial (VAR)

Entrenamiento-95%

Prueba-5%.

### Prueba de causalidad de Granger:

H0: La serie x no es Granger-Cause de la serie y.

H1: La serie x es Granger-Cause de la serie y.

Granger causality test

```
Model 1: trainVAR$precioNacional1d ~ Lags(trainVAR$precioNacional1d, 1:3)
+ Lags(trainVAR$precioInternacional1d, 1:3)
```

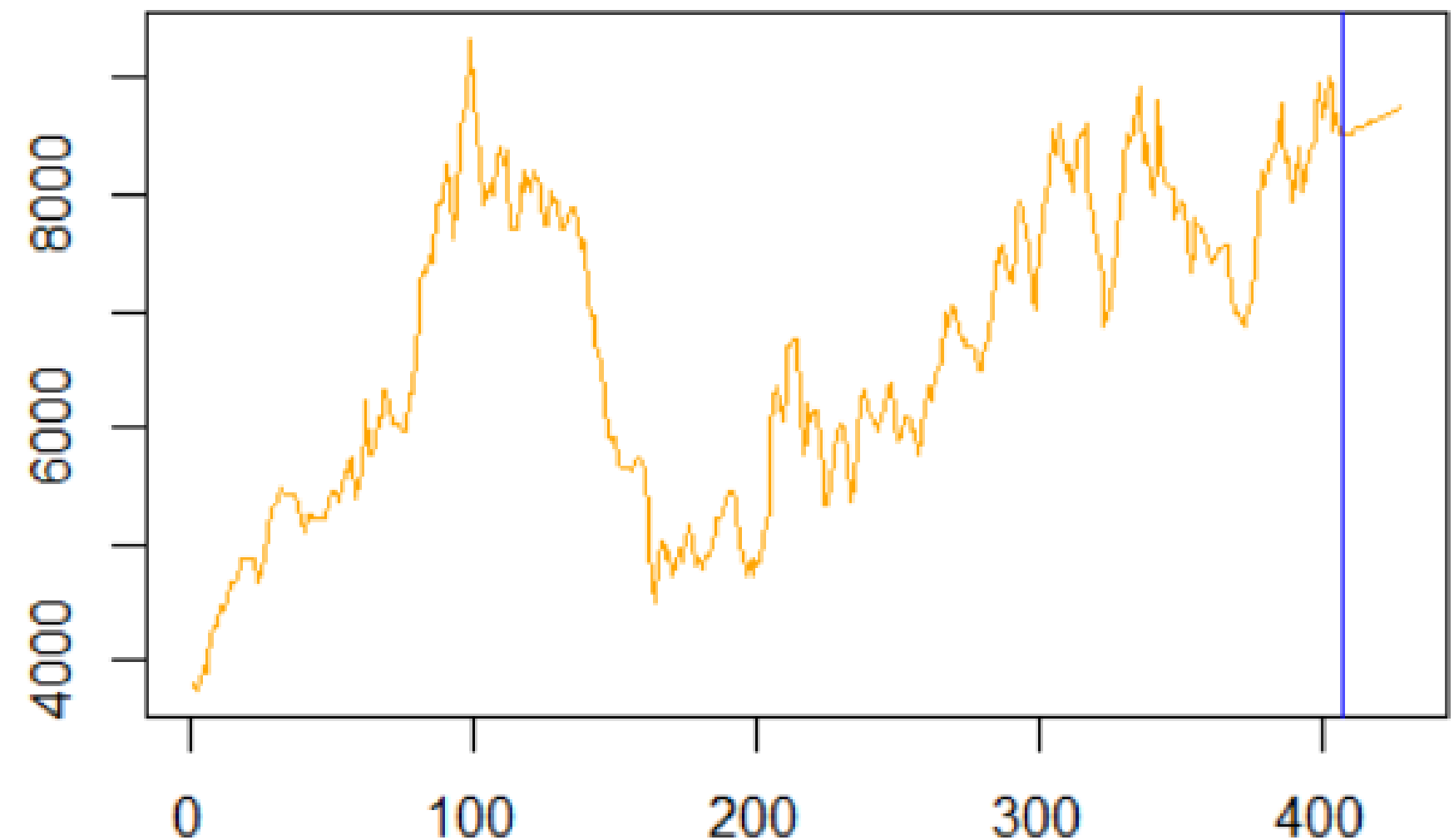
```
Model 2: trainVAR$precioNacional1d ~ Lags(trainVAR$precioNacional1d, 1:3)
```

	Res.Df	Df	F	Pr(>F)
1	376			
2	379	-3	65.537	< 2.2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

La serie de precios internacionales es Granger-Cause de la serie de precios nacionales.

## Precio Nacional



# METODOLOGÍA

## Modelo de Vector de Corrección de Errores (VECM)

Entrenamiento-95%

Prueba-5%.

### Prueba de cointegración de Phillips-Ouliaris

H0: No existe cointegración entre las series.

H1: Existe cointegración entre las series.

#### Phillips-Ouliaris Cointegration Test

```
data: cbind(precioNacional1, precioInternacional)
Phillips-Ouliaris demeaned = -61.367, Truncation lag
parameter = 4, p-value = 0.01
```

Las series estan cointegradas

```
###Model VECM
#####
Full sample size: 387   End sample size: 384
Number of variables: 2   Number of estimated slope parameters 12
AIC 8049.626   BIC 8100.984   SSR 36076525
Cointegrating vector (estimated by ML):
      precioNacional1 precioInternacional
r1              1              -0.8552125
```

```
Equation precioNacional1      ECT
Equation precioInternacional -0.0945(0.0224)***
                                -0.0748(0.0409).
                                Intercept
Equation precioNacional1      19.4831(8.3768)*
Equation precioInternacional 25.6967(15.2964).
                                precioNacional1 -1
Equation precioNacional1      -0.1057(0.0565).
Equation precioInternacional -0.1182(0.1031)
                                precioInternacional -1
Equation precioNacional1      0.3676(0.0368)***
Equation precioInternacional -0.0079(0.0673)
                                precioNacional1 -2
Equation precioNacional1      -0.0221(0.0438)
Equation precioInternacional -0.0472(0.0800)
                                precioInternacional -2
Equation precioNacional1      0.1502(0.0386)***
Equation precioInternacional 0.0281(0.0704)
```

# METODOLOGÍA

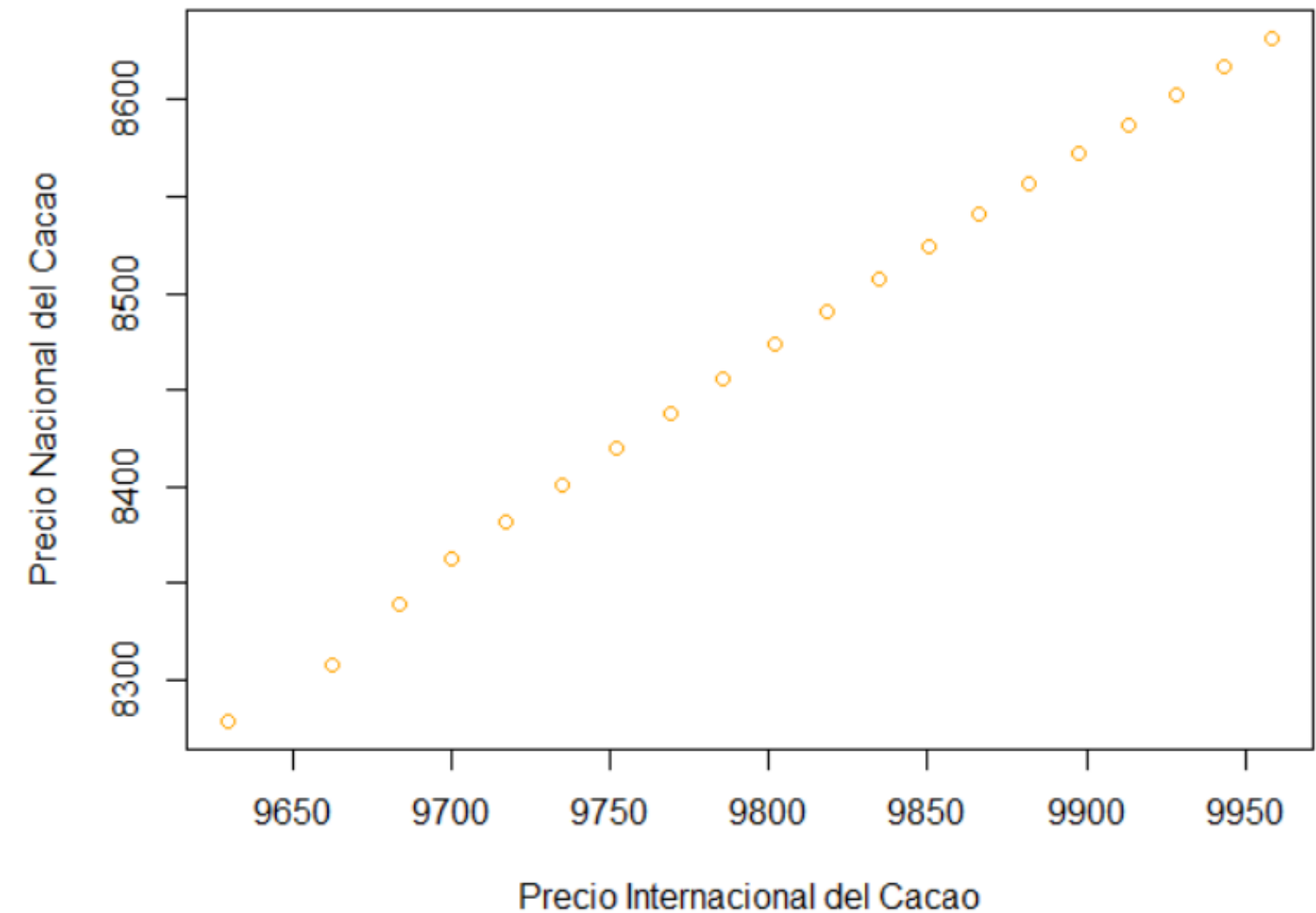
## Modelo de Vector de Corrección de Errores (VECM)

Entrenamiento-95%

Prueba-5%.

```
###Model VECM
#####
Full sample size: 387   End sample size: 384
Number of variables: 2   Number of estimated slope parameters 12
AIC 8049.626   BIC 8100.984   SSR 36076525
Cointegrating vector (estimated by ML):
  precioNacional1 precioInternacional
r1          1          -0.8552125
```

	ECT
Equation precioNacional1	-0.0945(0.0224)***
Equation precioInternacional	-0.0748(0.0409).
	Intercept
Equation precioNacional1	19.4831(8.3768)*
Equation precioInternacional	25.6967(15.2964).
	precioNacional1 -1
Equation precioNacional1	-0.1057(0.0565).
Equation precioInternacional	-0.1182(0.1031)
	precioInternacional -1
Equation precioNacional1	0.3676(0.0368)***
Equation precioInternacional	-0.0079(0.0673)
	precioNacional1 -2
Equation precioNacional1	-0.0221(0.0438)
Equation precioInternacional	-0.0472(0.0800)
	precioInternacional -2
Equation precioNacional1	0.1502(0.0386)***
Equation precioInternacional	0.0281(0.0704)



# METODOLOGÍA

## ARIMA con regresores

Entrenamiento-95%

Prueba-5%.

```
Series: trainArimaE$precioNacional1  
Regression with ARIMA(0,1,0) errors
```

```
Coefficients:
```

```
      xreg  
      0.2874  
s.e.  0.0328
```

```
sigma^2 = 30465:  log likelihood = -2539.81  
AIC=5083.62    AICc=5083.65    BIC=5091.53
```

## ARIMA(0,1,0)

$$Y_{i+1} - Y_i = \epsilon_{i+1}$$

$$Y_{i+1} = Y_i + \epsilon_{i+1}$$

**Caminata aleatoria, un caso  
especial de los modelos ARIMA**



# METODOLOGÍA

## ARIMA con regresores

### Se revisan los residuales:

#### Prueba de Ljung-Box

H0: Los residuales se distribuyen de forma independiente.

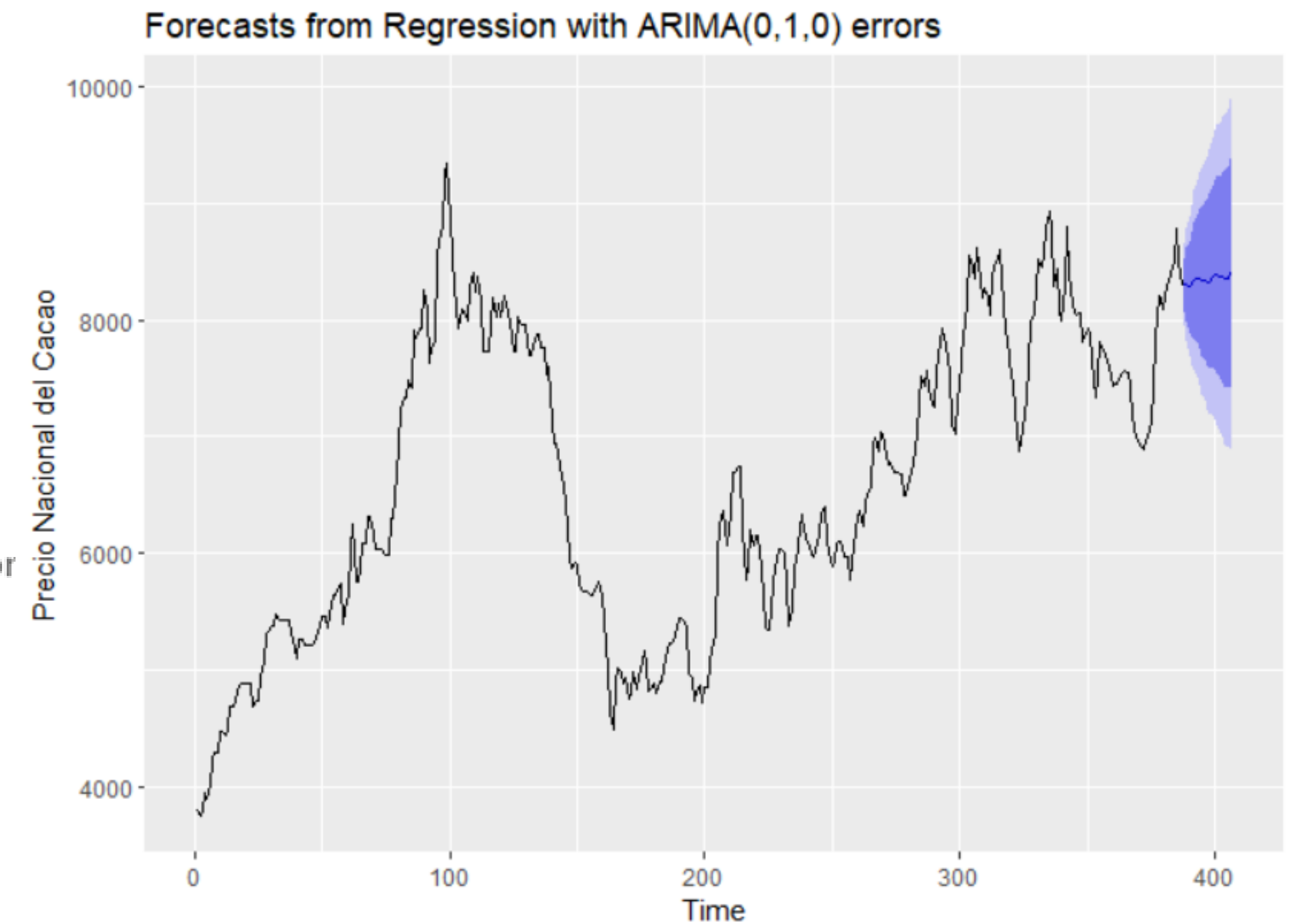
H1: Los residuales exhiben correlación serial.

#### Ljung-Box test

```
data: Residuals from Regression with ARIMA(0,1,0) error  
Q* = 13.552, df = 9, p-value = 0.1392
```

```
Model df: 1. Total lags used: 10
```

Las residuales se distribuyen de forma independiente.



# METODOLOGÍA

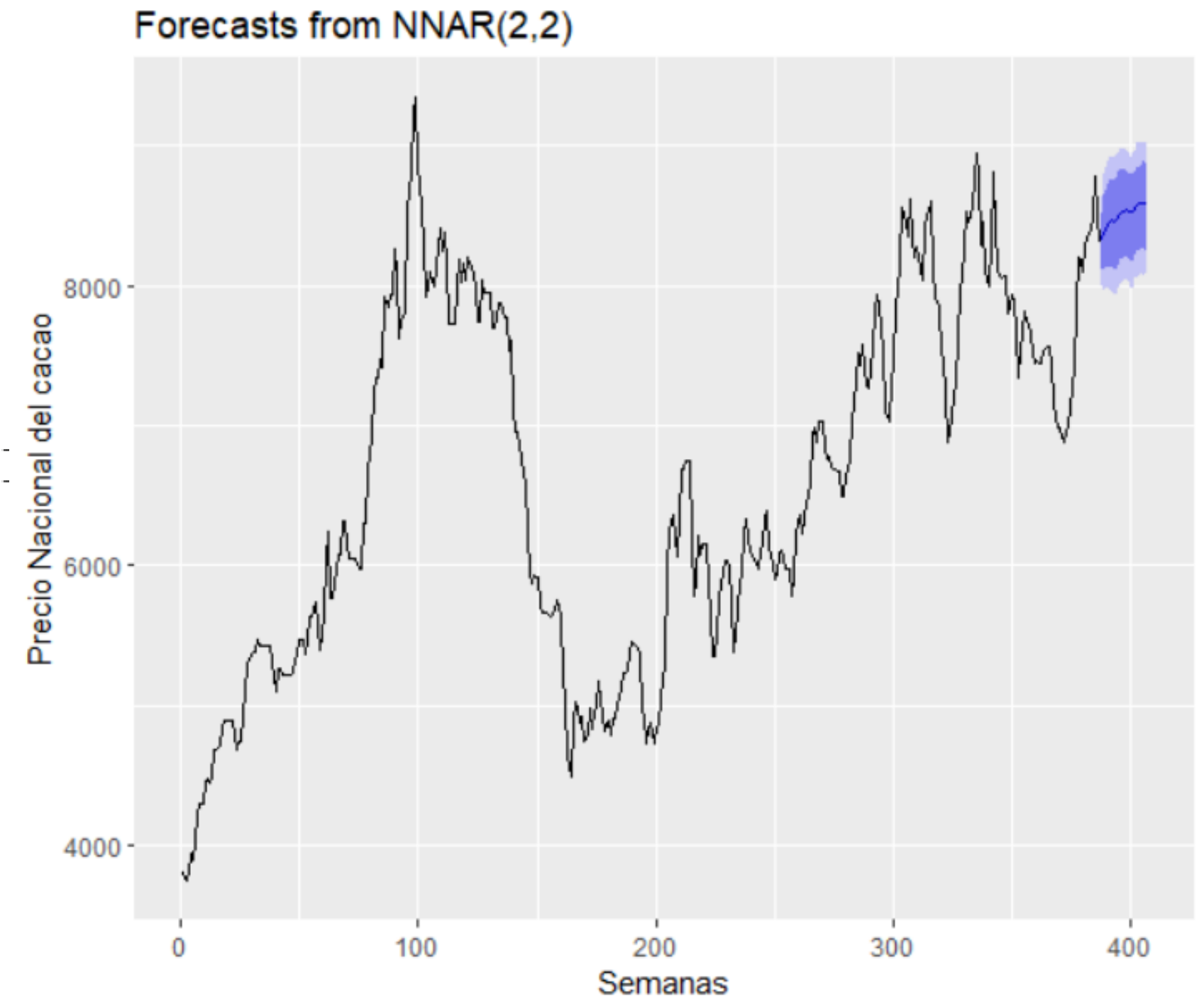
## Modelo de redes neuronales

Librería forecast de R se utiliza la función nnetar que estima un modelo NNAR

```
Series: trainRN[, 1]  
Model: NNAR(2,2)  
Call: nnetar(y = trainRN[, 1], xreg = trainRN[, 2],  
  lambda = "auto")
```

Average of 20 networks, each of which is  
a 3-2-1 network with 11 weights  
options were - linear output units

sigma^2 estimated as 5.96

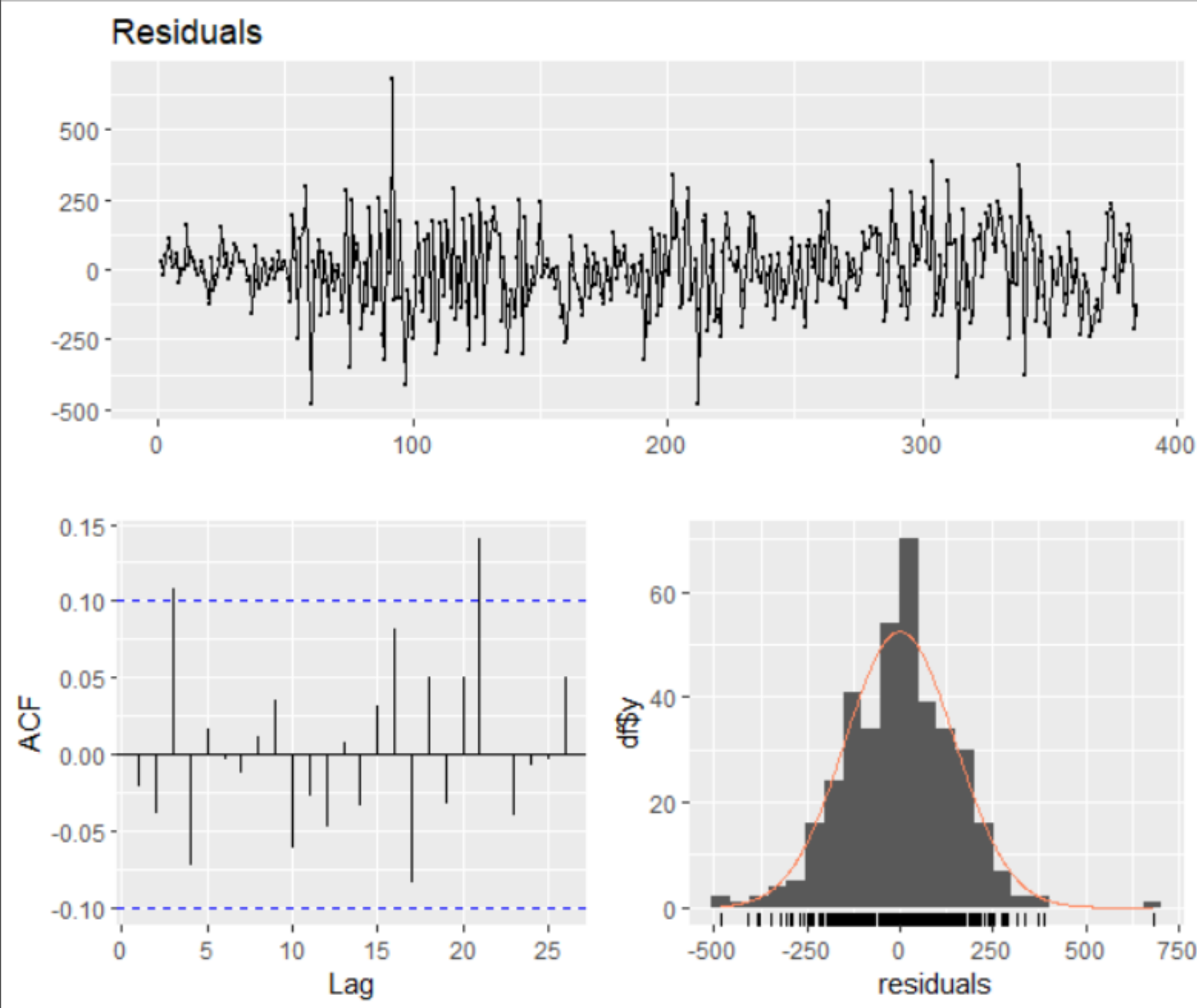


# RESULTADOS

Modelo	VAR	VECM	ARIMA con regresores
AIC	10212,12	8049,63	5083,62

Modelo	VAR	VECM	ARIMA con regresores	NNAR
RMSE	281,12	233,87	315,62	252,40
MAPE	0,0280	0,0210	0,0294	0,0231

# VALIDACIÓN



# RESULTADOS

## Prueba de Breusch-Godfrey

H0: Los residuales se distribuyen de forma independiente.

H1: Los residuales exhiben correlación serial.

Breusch-Godfrey LM test

```
data: Residuals of VAR object varmod1  
Chi-squared = 22.145, df = 20, p-value = 0.3327
```

## Prueba de ARCH

H0: No existe efecto ARCH.

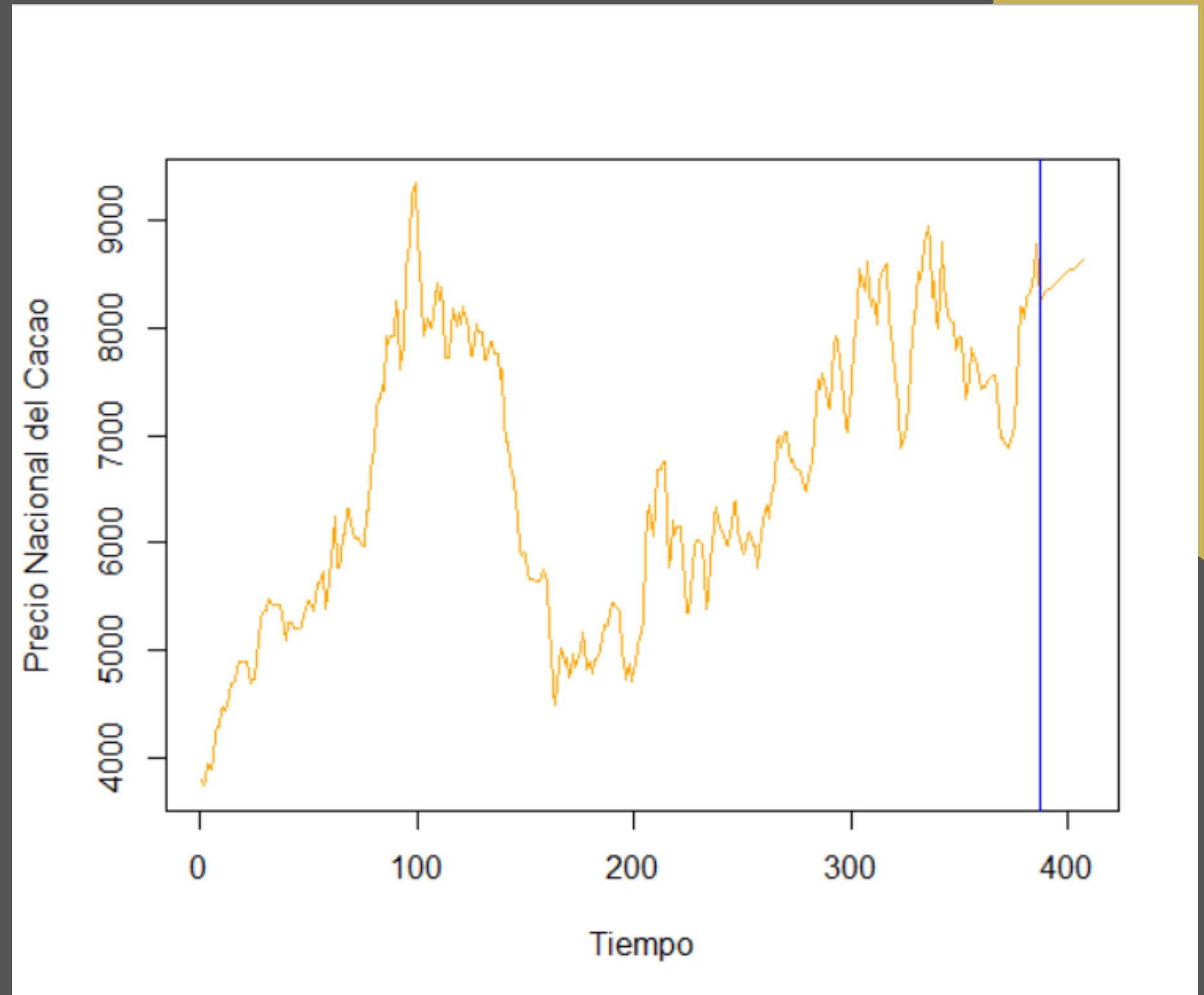
H1: Existe efecto ARCH.

ARCH (multivariate)

```
data: Residuals of VAR object varmod1  
Chi-squared = 147.92, df = 135, p-value = 0.211
```

**Los residuales son independientes entre sí y no hay presencia de efecto ARCH.**

## PRONÓSTICO DE MODELO VEC



# CONCLUSIONES Y RECOMENDACIONES

**Se selecciona el modelo predictivo VEC, el cual tiene un MAPE de 2,1%**

**Mejorar la periodicidad de las series**

**Identificar otras variables explicativas**

En particular, aquellas que describen factores del clima. y variación de producción

**Construir modelos econométricos y de aprendizaje automático**

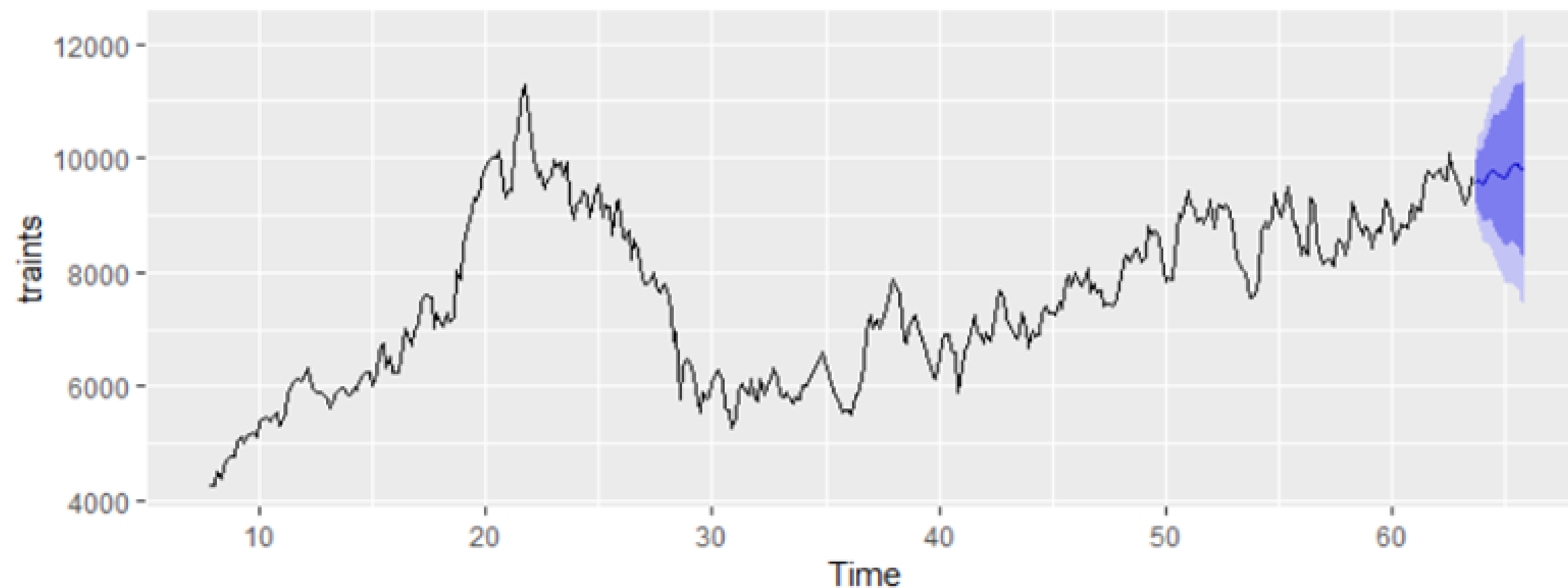
# BIBLIOGRAFÍA

- Ahumada, O., & Villalobos, J. (2009). Application of planning models in the agri-food supply chain: A review. Elsevier, 196 (1), 1-20.
- Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723
- Alonso Rodríguez, A. (2019). Un modelo de red neuronal para el Índice de producción de la Construcción Total en España. Anuario Jurídico y Económico Escurialense.
- Calo, D., Mena, L. (2020). Modelo arima para el pronóstico de la producción de cacao en el Ecuador para el periodo 2020-2024 [Proyecto de investigación, Universidad Central del Ecuador]
- Delgado, Y. (2021). Desarrollo de un modelo predictivo de precio de mora de castilla en Bogotá implementando técnicas de aprendizaje automático [Trabajo de maestría, Universidad Jorge Tadeo Lozano]. Recuperado de <https://expeditiorepositorio.utadeo.edu.co/bitstream/handle/20.500.12010/25006/PROYECTO%20DE%20GRADO%20V6.pdf?sequence=1&isAllowed=y>
- Geisendorf, S., & Pietrulla, F. (2017). The circular economy and circular economic concepts. Thunderbird International Business Review, 60, 771-782.
- Luna, O. (2020). Modelo de predicción de precios de productos agropecuarios como base para la priorización de Proyectos Integrales de Desarrollo Agropecuario y Rural con enfoque territorial en Colombia en el marco de los impactos producidos por la COVID-19 [Trabajo de maestría, Universidad Jorge Tadeo Lozano]. Recuperado de: <https://expeditiorepositorio.utadeo.edu.co/bitstream/handle/20.500.12010/17247/II-2020%20Luna.%20Modelo%20predicci%C3%B3n%20precios%20para%20priorizaci%C3%B3n%20de%20proyectos%20agropecuarios.pdf?sequence=1&isAllowed=y>
- Martínez Covaleta, H. J., Ortiz Hermida, L., & Espinal G, C. F. (2005). La cadena del cacao Una mirada global de su estructura y dinámica 1991-2005. Bogotá: Ministerio de Agricultura y Desarrollo Rural.
- Oliveros, D., y Pérez, S. (2013). Medición de la competitividad de los productores de cacao en una región de Santander, Colombia. Revista Le Bret (5). Bucaramanga, Colombia: Universidad Santo Tomás, pp 243-267
- Riveros, O. (2019). Desarrollo de un modelo predictivo basado en Machine Learning para observar el comportamiento de los precios de los tubérculos en la ciudad de Bogotá. Recuperado de: <https://repository.udistrital.edu.co/bitstream/handle/11349/16286/RiverosReyOscarArley2019.pdf?sequence=1&isAllowed=y>
- Sodhi, M., & Tang, C. (2014). Supply-Chain Research Opportunities with the Poor as. Production and Operations Management, 1483-1494.
- Vásquez-Barajas, E. F., García-Torres, N. E., Bastos-Osorio, L. M., & Lázaro-Pacheco, J. M. (2018). Análisis económico del sector cacaotero en Norte de Santander, Colombia y a nivel internacional. Rev.investig. desarro.innov, 8(2), 237-250.
- Zhu, Z., Chu, F., Dolgui, A., Chu, C., Zhou, W., & Piramuthu, S. (2018). Recent advances and opportunities in sustainable. International Journal of Production Research, 5700-5722.

# ANEXO

## Modelo predictivo univariado del precio internacional del cacao

Modelo	RMSE	MAPE
Simple Exponential Smoothing	326,2273	0,02959976
Metodos con Tendencia. holt ()	326,1477	0,02959970
Método HoltWinters ()	308,0072	0,02643114
Suaviamiento exponencial. Ets ()	333.1557	0,02923134
Método Hw ()	682,1321	0,06032489





# ANEXO

## Herramienta de apoyo a la decisión



Ir a <https://karen0c.shinyapps.io/Aplicativo/>