

Modelización de serie de tiempo de precios mayoristas de manzana de Uruguay.

Emanuelle Marsella, Maximiliano Saldaña

Junio 2022

Introducción

Introducción

- ▶ Se modeliza la serie de precios mensuales mayoristas del kilo de manzana en la Unidad Agroalimentaria Metropolitana (ex Mercado Modelo).
- ▶ Se cuentan con los datos desde enero de 2013 a mayo de 2022 y se considerará el promedio mensual de los precios, por lo que se cuenta con 113 observaciones.
- ▶ Datos originales bisemanales, se opta por la frecuencia mensual debido a la dificultad de emplear las herramientas de modelado SARIMA para tales tipos de series.
- ▶ Interés especial en la predicción.

Análisis descriptivo

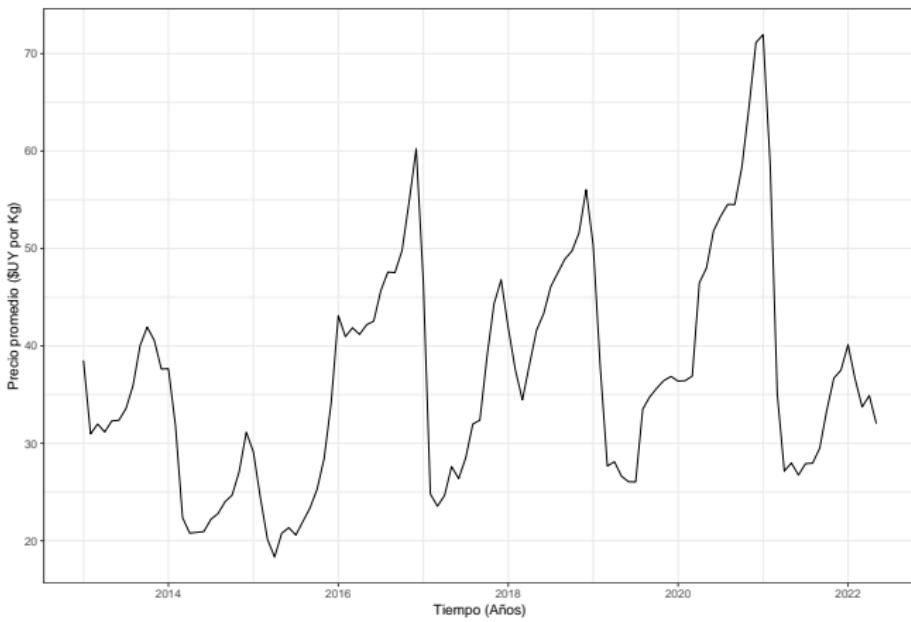


Figura 1: Serie de precios mensuales del Kg de manzana en pesos uruguayos.

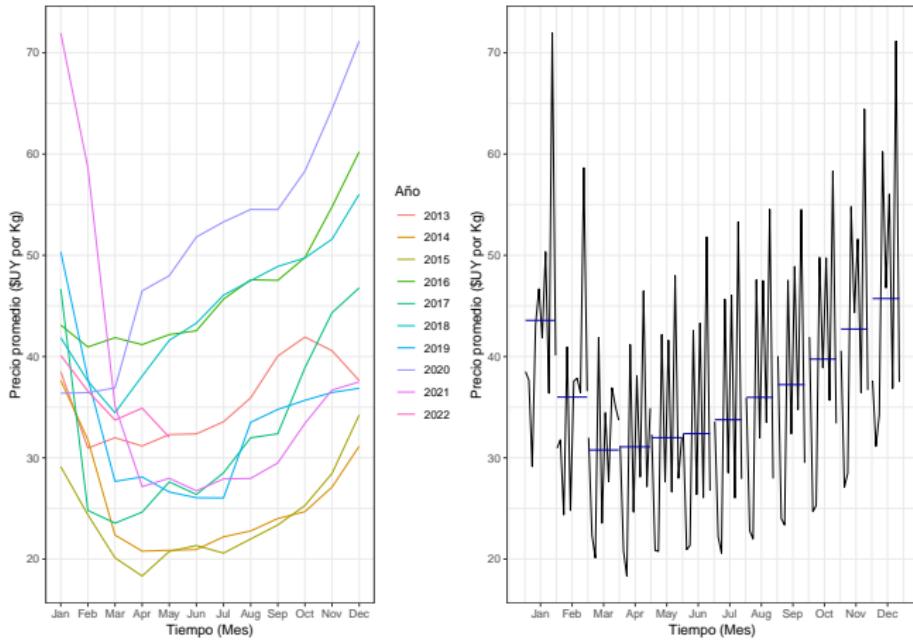


Figura 2: Serie de precios mensuales del Kg de manzana en pesos Uruguayos.

- ▶ Marcada estacionalidad anual
- ▶ En los últimos 5 años, ciclo corto que se repite cada 2 años.
- ▶ Se realizó además la descomposición de la serie en tendencia/ciclo, estacionalidad y componente irregular.
- ▶ Fuerza de la estacionalidad, definida como en (Hyndman and Athanasopoulos 2018):

$$F_s = \max \left(0, 1 - \frac{\text{Var}(R_t)}{\text{Var}(R_t + S_t)} \right) = 0,51$$

Metodología y resultados

Metodología

- ▶ Se divide la serie original en muestras de entrenamiento y test:
 - ▶ Muestra de entrenamiento hasta mayo de 2021, 101 observaciones
 - ▶ Muestra de testeo entre junio de 2021 y mayo de 2022, 12 observaciones.

Identificación manual de un modelo

- ▶ Estructura de autocorrelación
- ▶ Cantidad de parámetros con los que contará la especificación
- ▶ Posible diferenciación de la serie
- ▶ Posible transformación

Estructura de autocorrelaciones

Buscamos identificar la estructura de dependencia temporal de la serie mediante mediante las ACF y PACF muestrales.

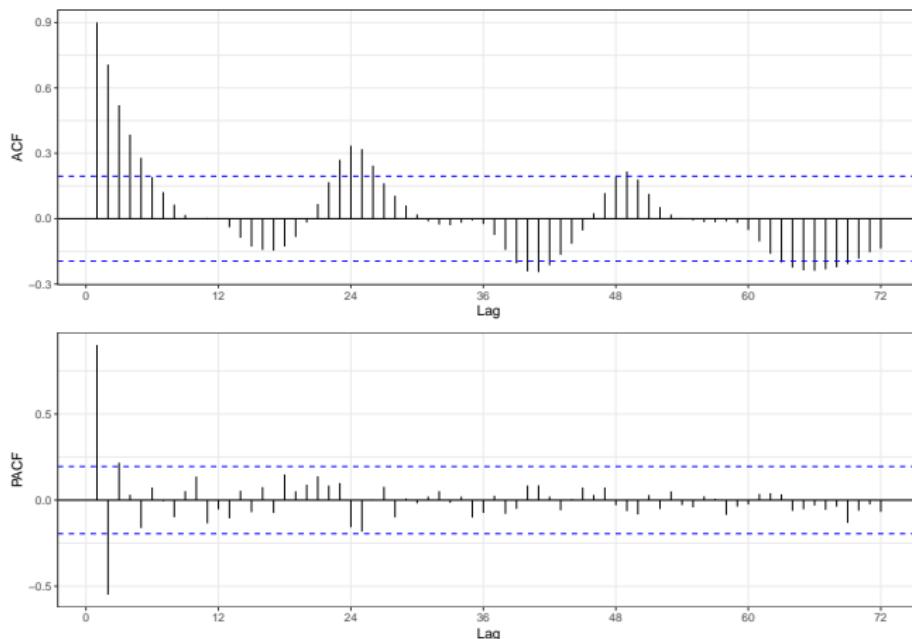


Figura 3: Funciones de autocorrelación y autocorrelación parcial muestrales de la serie de precios de manzana.

Tests de raíces unitarias

- ▶ Dado que las funciones de autocorrelación dieron indicios de que el proceso no es estacionario, resulta de interés poner a prueba si el proceso es $I(1)$, es decir, si cuenta con una raíz unitaria. En dicho caso el proceso sería no estacionario la cual implicaría que no puede ser modelado en el marco de los ARMA.
- ▶ Utilizamos el test de Dickey-Fuller y el de Phillips-Perron.
 - ▶ Hipótesis nula: la serie cuenta con una raíz unitaria
 - ▶ Ambos tests llevan a no rechazar la hipótesis a un nivel de significación de 95 %.

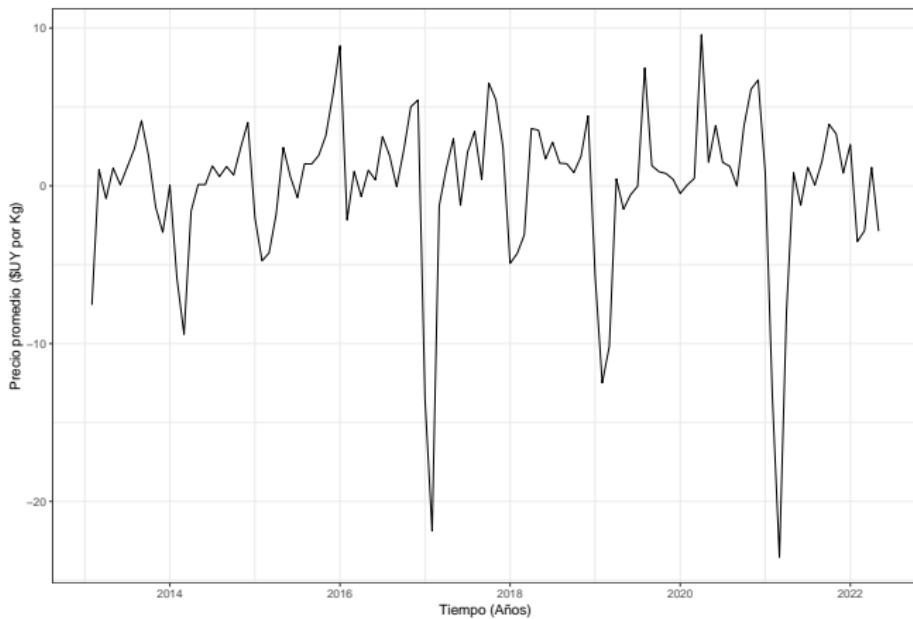


Figura 4: Serie de precios mensuales del Kg de manzana con una diferencia regular, en pesos uruguayos.

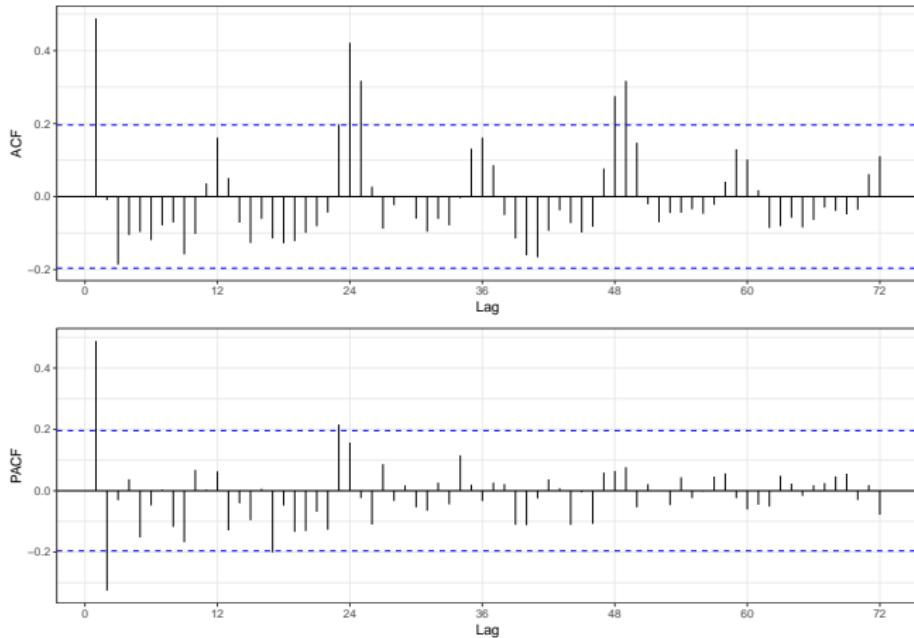


Figura 5: Autocorrelograma y autocorrelograma parcial de la serie de primeras diferencias de precios de manzana.

Especificación manual del modelo

- ▶ La primer autocorrelación resulta significativamente distinta de 0, indicio de un componente MA de primer orden.
- ▶ Los para valores en torno a los lags 24 y 48 también lo son, indicio de dependencia estacional autorregresiva de orden 2.
- ▶ Solo los dos primeros valores de la autocorrelación parcial son significativos, indicio de una estructura AR de orden 2.

Teniendo en cuenta este acercamiento metodológico, inicialmente se plantea un modelo $ARIMA(2, 1, 1)(2, 0, 0)$

Selección mediante criterios de información

Alternativamente especificamos un modelo mediante el Criterio de Información de Akaike corregido (AICc), mediante la función `auto.arima()`. Su fórmula para un determinado modelo es:

$$AICc = T \log \hat{\sigma}_{MV}^2 + T \frac{1 + k/T}{1 - (k + 2)/T}$$

- ▶ T el número de observaciones
- ▶ k el número de parámetros
- ▶ $\hat{\sigma}_{MV}^2$ el estimador MLE de la varianza de los errores.

El modelo seleccionado es aquel con el menor valor del AICc, lo cual a nivel algorítmico se hace empleando la selección *stepwise* a partir de un conjunto inicial de modelos (Hyndman and Athanasopoulos 2018).

El modelo seleccionado mediante esta metodología resulta en un *ARIMA(0, 1, 2)(0, 0, 2)*.

Estimación

- ▶ Luego de haber identificado la especificación de los modelos, estimamos sus parámetros por máxima verosimilitud.
- ▶ Para esto es necesario asumir que los errores son gaussianos.
- ▶ El supuesto anterior es fuerte pero las estimaciones resultantes serán razonables aunque el supuesto no se cumpla (Hamilton 1994).
- ▶ En particular, el método empleado es la estimación máximo verosímil condicional, donde se supone que la primera observación de la serie es determinística y se maximiza la verosimilitud condicionada a dicha observación. Esto simplifica las expresiones de las funciones, y si el tamaño de muestra es razonablemente grande, la primera observación no tendrá gran efecto sobre la verosimilitud estimada.

Pruebas de significación de los parámetros

Para completar la especificación se realizan pruebas de hipótesis sobre la significación de los parámetros estimados. La hipótesis para cada uno de ellos, siendo λ un parámetro cualquiera del modelo:

$$H_0) \lambda = 0$$

$$H_1) \lambda \neq 0$$

Donde el estadístico empleado es:

$$z = \hat{\lambda} / \hat{\sigma}_\lambda$$

Para el que se cumplirá la Normalidad asintótica debido a que la estimación se realizó por máxima verosimilitud (Hamilton 1994, 143).

Estimación de los parámetros para el modelo especificado manualmente

- ▶ Todos resultan significativos al 5 % de confianza con la excepción del parámetro de medias móviles (MA) y el primer parámetro autorregresivo estacional (sAR).
- ▶ Como consecuencia, optamos por eliminar el componente MA de la especificación y restringir el primer parámetro sAR a 0.

La especificación del *modelo manual* final es:

$$\Phi_2(L^{12})\phi_2(L)\Delta^1 Y_t = \varepsilon_t$$

Donde:

$$\Phi_2(L^{12}) = 1 - 0,415L^{24}$$

$$\phi_2(L) = 1 - 0,557L - (-0,277)L^2$$

Diagnóstico

Para evaluar si el modelo es adecuado se ponen a prueba los supuestos realizados sobre los residuos:

- ▶ Media 0 de los residuos: contraste no paramétrico de Wilcoxon, donde se testea $H_0) \mu_\varepsilon = 0$ contra $H_1) \mu_\varepsilon = 0$.
- ▶ Incorrelación de los errores: test de autocorrelación conjunta de Ljung-Box, donde la hipótesis nula es que los residuos son incorrelacionados contra la alternativa de que no lo son.
- ▶ Homoscedasticidad: test de Box-Ljung sobre los residuos al cuadrado del modelo McLeod and Li (1983). La hipótesis nula de esta prueba es la homoscedasticidad entre los k rezagos considerados.
- ▶ Normalidad: tests de Jarque-Bera y Shapiro-Wilk.
 $H_0) \varepsilon_t \sim N(0, \sigma^2)$ versus $H_0) \varepsilon_t \not\sim N(0, \sigma^2)$

Modelo manual: incorrelación de los errores

Modelo manual: homocedasticidad

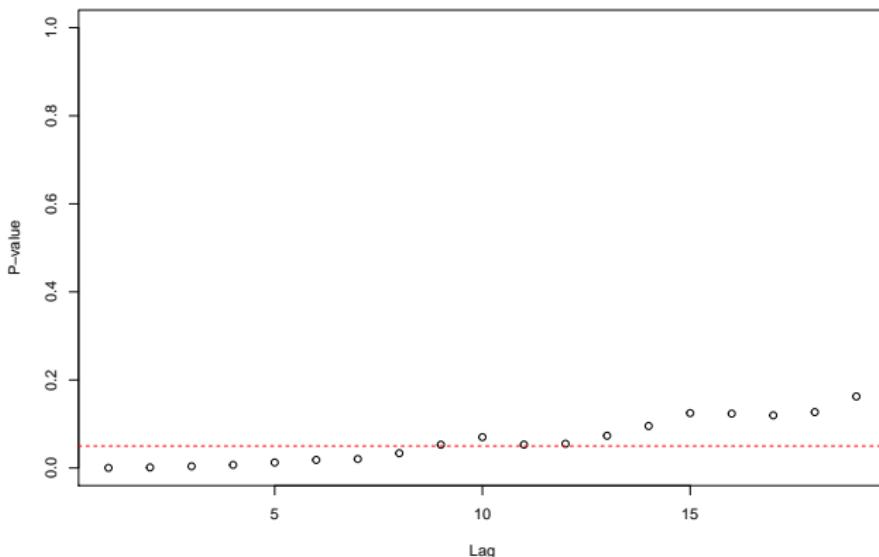


Figura 6: P-valores del test de McLeod-Li para los primeros 19 rezagos

Indicación que los modelos del tipo SARIMA puede no ser los más adecuados y se tendría que recurrir a los del tipo ARCH/GARCH donde se busca modelizar la varianza de los errores.

Modelo manual: normalidad

- ▶ En ambos contrastes se rechaza la hipótesis nula de normalidad al 95 % de confianza, por lo que se recurrió a incorporar el efecto de los atípicos en el modelo manual.
- ▶ Consideramos como outliers a los residuos e_t que no se encuentran en el intervalo $(Q_1 - 3IQR, Q_3 + 3IQR)$ Hyndman (n.d.).
 - ▶ TC: enero de 2017
 - ▶ AO: febrero de 2017 y marzo de 2021
- ▶ Incluimos estas intervenciones como regresores del modelo, y sus parámetros son significativos al 5 %.
- ▶ Para el modelo intervenido, se cumple el supuesto de normalidad pero no el de autocorrelación de los residuos, que sí se cumplía para el anterior. Por lo tanto, deshechamos la intervención.

Modelos alternativos

Una vez se llegó al final de la etapa de diagnóstico del modelo planteado de forma manual, optamos por evaluar otros modelos alternativos seleccionados a partir de métodos automáticos, para tener varias opciones a la hora de evaluar el desempeño de las predicciones y compararlos con el modelo manual.

Modelo seleccionado por AICc

- ▶ Se seleccionó un modelo utilizando el Criterio de Información de Akaike corregido ($AICc$), que tiene la forma $ARIMA(0, 1, 2)(0, 0, 2)$.
- ▶ Los parámetros de medias móviles son significativos, así como el parámetro de medias móviles estacionales de segundo orden, pero no el de primer orden.
- ▶ No se rechaza la hipótesis de no autocorrelación ni tampoco la de media nula, pero la homocedasticidad no se cumple para todos los rezagos, y se rechaza la hipótesis de normalidad.
- ▶ Luego, se realizó una intervención sobre este modelo para corregir el efecto de algunos outliers que se detectaron, y se llegó a un modelo que cumple todos los supuestos sobre los residuos (excepto el de homocedasticidad para algunos rezagos).

Modelo seleccionado por *tso*

- ▶ Se seleccionó otro modelo a partir de las intervenciones aplicadas con el procedimiento automático de detección de outliers de la función *tso* presente en el paquete *tsoutliers* (*modelo tso*).
- ▶ El resultado del mismo es un $ARIMA(0, 1, 1)(2, 0, 0)$, que tiene 7 intervenciones (5 TC y 2 AO), pues se restringió el tipo de outliers posibles a estos dos.
- ▶ Todos los parámetros son significativos, excepto el correspondiente a la estacionalidad autorregresiva de orden 1.
- ▶ Se cumplen todos los supuestos que se suelen pedir sobre los residuos, incluyendo el de la homocedasticidad para todos los rezagos (que en modelos anteriores no se cumplía).

Modelo seleccionado por *tsoutliers*

- ▶ Otra opción considerada es la de un modelo intervenido con por métodos de la función *tsoutliers* del paquete *forecast*, que reemplaza los outliers a través de interpolación lineal (*modelo clean*).
- ▶ En este caso, se selecciona un modelo $ARIMA(1, 1, 0)(2, 0, 0)$, cuyo parámetros son significativos excepto el autorregresivo estacional de orden 1, pero que no cumple los supuestos de normalidad y no autocorrelación de los residuos, por lo cual optamos por no utilizarlo.

Modelo seleccionado por *ets*

Finalmente se selecciona un modelo de una clase distinta a los ARIMA, de la familia de los de suavizado exponencial, que hacen uso de promedios ponderados de las observaciones (*modelo ets*), y tienen como base modelos estructurales o espacio-estado. Este tipo de modelos permiten una estructura de varianzas que cambia en el tiempo, por lo cual no se requiere el supuesto de homocedasticidad de los residuos.

Por su parte, el supuesto de no autocorrelación de los mismos sí se cumple.

Predictión

Una vez se cuenta con un modelo (o una selección de modelos) que se suponen correctos, tanto en especificación como en estimación, se puede proceder a la predicción de los valores futuros de la serie condicionando a los valores con los cuales se cuenta. Para los datos empleados, los precios mayoristas de manzana, resulta de particular interés la predicción a futuro, dado que lo que ocurra con el precio incidirá tanto en la producción como en el consumo del bien y puede ser un insumo para la toma de decisiones a nivel de política económica.

- ▶ Función de costos cuadrática: $C(e) = ae^2$. Minimizar el valor esperado de esta función implica minimizar el error cuadrático medio a un horizonte predictivo de s pasos, condicional a las t observaciones con las que se cuenta: $E(Y_{t+1} - Y_{t+1|t}^*)^2$. Se llega a que la predicción óptima es la esperanza condicional:
$$Y_{t+1|t}^* = E(Y_{t+1}|Y_t).$$

$$\hat{E}[Y_{t+s}|\varepsilon_t, \varepsilon_{t-1}, \dots] = \psi_s \varepsilon_t + \psi_{s+1} \varepsilon_{t-1} + \psi_{s+2} \varepsilon_{t-2} + \dots$$

Medidas de error de predicción

Para cuantificar el error que se comete al comparar predicción contra valor verdadero, se recurre a cuatro medidas de error.

La raíz del error cuadrático medio (RMSE) y el error absoluto medio (MAE) son dos medidas ampliamente usadas, debido a su relevancia en el modelado estadístico (Hyndman and Koehler 2006). Ambas medidas son dependientes de la escala, por lo que no pueden usarse para comparar el error de modelos que no están en la misma escala y además son sensibles a los outliers. Sea $e_s = Y_{t+s} - Y_{t+s|t}^*$:

- ▶ Raíz del error cuadrático medio (RMSE) = $\sum_{i=1}^s e_i^2$
- ▶ Error absoluto medio (MAE) = $\sum_{i=1}^s |e_i|$

Otro tipo de medidas son las basadas en errores porcentuales. Dichos errores están dados por: $p_s = 100e_t / Y_s$. Este tipo de medidas son independientes de la escala, lo que las hace idóneas para comparar el desempeño predictivo de modelos con distintas escalas (como es el caso de nuestra selección).

Predictión dentro de la muestra de entrenamiento

- ▶ Como una forma de medir el ajuste del modelo a la serie trabajada en general se calculan las predicciones a un paso del modelo dentro de la muestra de entrenamiento, comparándose con el valor verdadero de la serie.
- ▶ No resulta una buena referencia para evaluar el modelo en cuanto a sus predicciones fuera de la muestra.

Modelo	RMSE	MAE	MAPE	MASE
modelo_manual	3.9	2.7	8.0	0.20
modelo_aicc	4.0	2.8	8.2	0.20
modelo_aicc_int	3.1	2.4	7.0	0.17
modelo_tso	2.4	1.9	5.2	0.13
modelo_ets	4.1	2.6	7.5	0.19

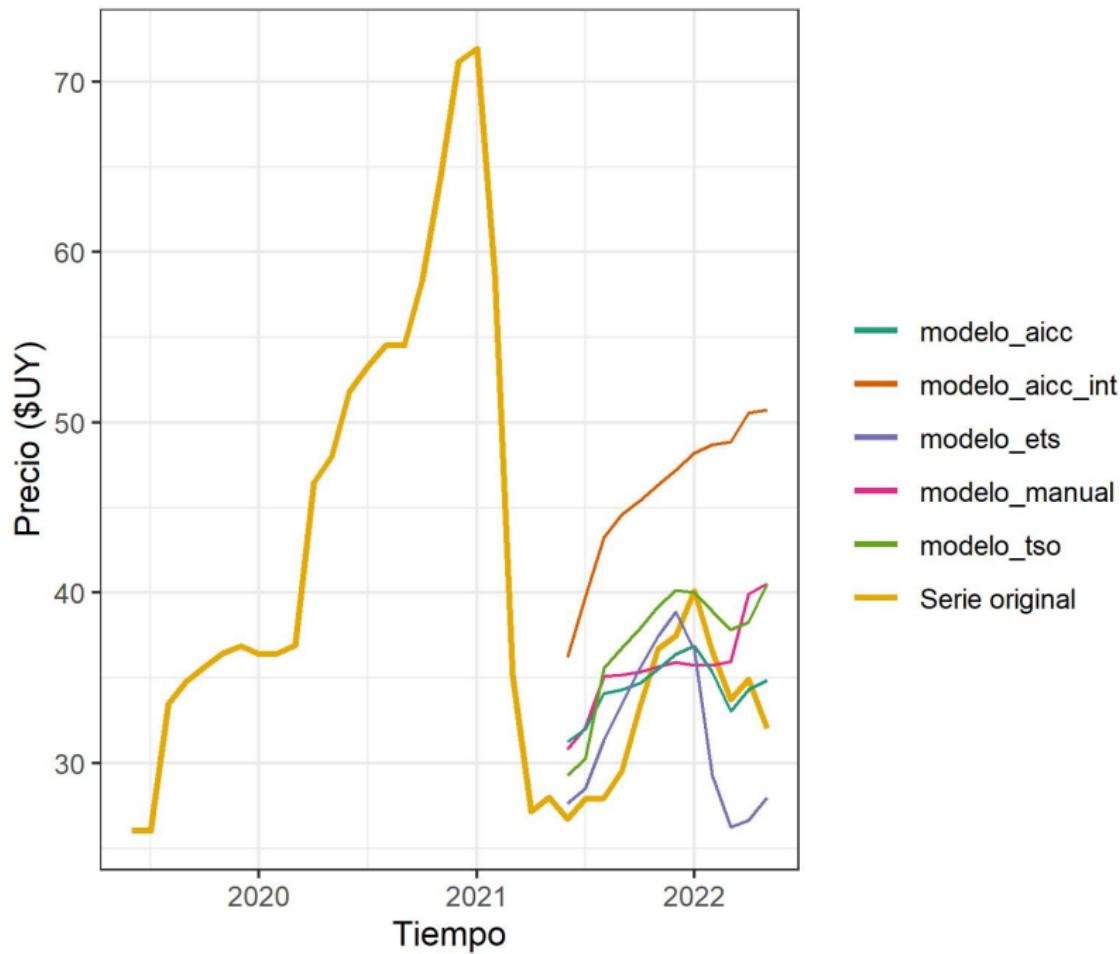
Tabla 1: Errores de los distintos modelos en la muestra de entrenamiento.

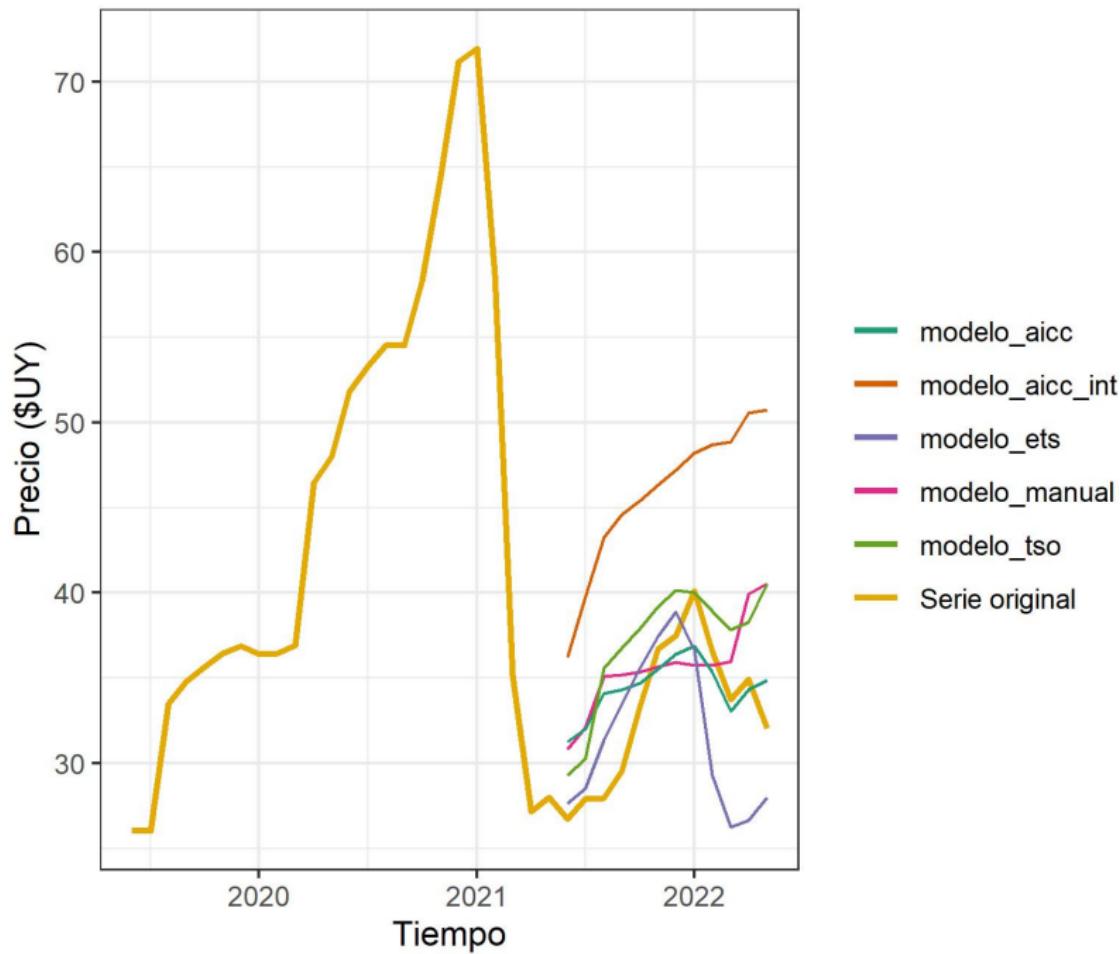
Predictión fuera de la muestra de entrenamiento

- ▶ Una forma de evaluar las predicciones que busca simular el proceso de predecir al futuro pero teniéndose la posibilidad de evaluar la calidad de las predicciones es mediante la predicción fuera de la muestra de entrenamiento.
- ▶ Se predice a 12 pasos (a un año dado que se trabaja con datos mensuales) y se compara con los 12 valores que se apartaron para conformar la muestra de prueba. Luego, se evalúan las medidas de error anteriormente discutidas.

Modelo	RMSE	MAE	MAPE	MASE
modelo_manual	4.5	3.9	12.4	0.28
modelo_aicc	3.2	2.6	8.6	0.19
modelo_aicc_int	13.1	12.7	39.4	0.92
modelo_tso	4.7	4.0	12.7	0.29
modelo_ets	4.5	3.6	10.9	0.26

Tabla 2: Errores de los distintos modelos en la muestra de testeо.





Conclusiones

Conclusiones: metodología de Box y Jenkins

- ▶ Siguiendo la metodología de Box y Jenkins se llegó a un modelo autorregresivo con una diferencia, con un componente estacional también autorregresivo.
- ▶ Este cumplió los supuestos que se hicieron en su especificación con la excepción del de normalidad y homoscedasticidad. El primero afecta los intervalos de predicción y pudo ser subsanado mediante *bootstrap*, mientras que el segundo puede ser una señal de que debe ser considerado un modelo que modele la heteroscedasticidad, como por ejemplo los del tipo ARCH/GARCH.
- ▶ Como alternativa a este modelo trabajado de manera “manual” se planteó un conjunto de modelos seleccionados mediante metodologías de carácter automatizado, como lo son la selección mediante criterios de información.

Conclusiones: modelos con intervenciones

- ▶ La serie cuenta con valores atípicos los cuales se intentaron subsanar incluyendo intervenciones, pero en el caso de los modelos que las incluían se observó mayor error a la hora de la predicción fuera de la muestra.
- ▶ En particular, la intervención de valores atípicos al final de la serie aumentaba el error en el que se incurría, por más que en estos casos se cumplían los supuestos subyacentes a los modelos y el ajuste dentro de la muestra era mejor que para los no intervenidos.
- ▶ Esto lo atribuimos al hecho de que incluir una cantidad alta de intervenciones implica un sobreajuste del modelo, en el caso del *modelo tso* un 6 % de las observaciones fueron intervenidas.
- ▶ La modelización con mejor desempeño predictivo fue la elegida mediante el criterio de Akaike corregido y sin intervenciones, siendo la segunda mejor la especificada manualmente.

Conclusiones: efecto de la pandemia

- ▶ El hecho de que la predicción se haga con el trasfondo de la pandemia de Coronavirus aumenta la incertidumbre de una forma que es difícil cuantificar, porque todavía no se tiene la distancia temporal suficiente para concluir si causó un quiebre en la estructura de los precios.
- ▶ De hecho, las predicciones fuera de la muestra las realizamos para el periodo entre junio de 2021 y mayo de 2022, por lo cual los últimos 3 meses se ubican dos años después del inicio de la pandemia, lo cual puede explicar el pobre desempeño del modelo para dichos meses teniendo en cuenta el orden 2 de los componentes estacionales especificados.

Bibliografía

Bibliografía

Ericsson, Neil. 2001. "Forecast Uncertainty in Economic Modeling." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.266494>.

Hamilton, James. 1994. *Time Series Analysis*. Princeton University Press.

Hyndman, Rob J. n.d. "Detecting Time Series Outliers." Accessed June 29, 2021. <https://robjhyndman.com/hyndts/toutliers/>.

Hyndman, Rob J., and George Athanasopoulos. 2018. *Forecasting: Principles and Practice*. 2nd ed. Australia: OTexts.

Hyndman, Rob J., and Anne B. Koehler. 2006. "Another Look at Measures of Forecast Accuracy." *International Journal of Forecasting* 22 (4): 679–88.

<https://doi.org/10.1016/j.ijforecast.2006.03.001>.

McLeod, A., and W. Li. 1983. "Diagnostic Checking Arma Time Series Models Using Squared-Residual Autocorrelations." *Journal of Time Series Analysis* 4 (4): 269–73.

<https://doi.org/10.1111/j.1467-9892.1983.tb00373.x>