



Estudio Periódico de Temáticas de Madera N° 5

PREDICCIÓN DE PRECIOS DE PRODUCTOS FORESTALES USANDO MODELOS ARIMA

ÁREA DE INFORMACIÓN Y ECONOMÍA FORESTAL
INSTITUTO FORESTAL
2022



Las fotografías e imágenes incorporadas en tapas o texto de la presente publicación provienen de archivo institucional o fueron obtenidas o elaboradas durante el desarrollo de las actividades del trabajo que origina esta publicación.

Estudio Periódico de Temáticas de Madera N° 5

**PREDICCIÓN DE PRECIOS DE PRODUCTOS FORESTALES
USANDO MODELOS ARIMA**

Wilson Mejías Caballero¹

**Proyecto Fortalecimiento de las Capacidades Tecnológicas del Instituto Forestal para el
Desarrollo de la Industria Secundaria de la Madera, a través de Bienes Públicos
Orientados al Sector de la Construcción**

ÁREA DE INFORMACIÓN Y ECONOMÍA FORESTAL

**INSTITUTO FORESTAL
2022**

¹ Investigador Área de Información y Economía Forestal, Instituto Forestal, Chile. wmejias@infor.cl



INFOR

Instituto Forestal

Sucre 2397, Ñuñoa, Santiago

Chile

F. 223667115

www.infor.cl

ISBN N° 978-956-318-232-3

Registro Propiedad Intelectual N° 2022-A-6911

Se autoriza la reproducción parcial de esta publicación siempre y cuando se efectúe la cita correspondiente:

Mejías Caballero, Wilson, 2022. Modelos ARIMA aplicados a la predicción de precios de productos forestales. Estudio Periódico de Temáticas de Madera N°5. Instituto Forestal, Chile. Documento de Divulgación N°66. 28 p.

ÍNDICE

PRÓLOGO	1
INTRODUCCIÓN	3
METODOLOGÍA	5
RESULTADOS	7
CONCLUSIONES	25
REFERENCIAS	27

PRÓLOGO

La información es un elemento fundamental para la toma de decisiones estratégicas, tanto en el sector privado como en el público. El Instituto Forestal (INFOR) es un referente en esta área, con una larga tradición en la recopilación y análisis de información de la industria forestal chilena, manejando un gran volumen de datos en su plataforma de estadísticas forestales.

En un mundo globalizado y sujeto a diversos factores de incertidumbre, es importante contar con proyecciones confiables de cifras estadísticas que sirvan de apoyo en la planificación de diferentes organizaciones.

La madera es un material que ofrece ventajas sociales, ambientales y económicas a la construcción, por lo que en los últimos años se ha emprendido una serie de iniciativas y proyectos para fomentar el uso de este material, lo cual ha implicado una importante inversión estatal y privada que apunta a cambios a nivel normativo y tecnológico.

Es de importancia conocer el comportamiento futuro de los precios de la madera, para entregar certidumbre al análisis de proyectos actuales y futuros. Por este motivo, el desarrollo de modelos predictivos representa una apropiada herramienta de gestión para potenciar el uso de este material en Chile.

Este estudio fue llevado a cabo por el Área de Información y Economía Forestal de INFOR en el marco del proyecto *"Fortalecimiento de las Capacidades Tecnológicas del Instituto Forestal para el Desarrollo de la Industria Secundaria de la Madera, a través de Bienes Públicos, orientados al Sector de la Construcción"*, financiado por la Corporación de Fomento de la Producción (CORFO).

INFOR, agradece a las empresas que han colaborado en los procesos de levantamiento de datos trimestrales de precios, permitiendo la generación de información estadística y el desarrollo de estudios de análisis sectorial.


Sandra Gacitúa Arias
Directora Ejecutiva
Instituto Forestal



INTRODUCCIÓN

En Chile se producen cerca de 15 millones de m³ de trozas aserrables de pino radiata, las cuales son el insumo para la producción de aproximadamente 7,9 millones de m³ de madera aserrada, cifras que posicionan Chile como uno de los 15 mayores productores de este material en el mundo (INFOR, 2021). A pesar de esto, los datos del Instituto Nacional de Estadísticas (INE, 2021) indican que aproximadamente el 13% de la superficie de viviendas que se construye anualmente en el país utiliza madera como material estructural predominante, lo cual implica un consumo de madera estimado en unos 500.000 m³ (INFOR, 2020).

Según las empresas dedicadas a la construcción, uno de los factores más determinantes a la hora de decidir sobre el uso de la madera es su precio, por lo cual contar con modelos que sean capaces de generar estimaciones robustas ayudaría a tener mayor claridad y proyectar de mejor manera los costos tanto para la industria forestal como para la industria de la construcción, así como también para la planificación de la inversión pública en esta materia.

Series de Tiempo

Una serie de tiempo es un grupo de datos registrados durante un periodo mensual, trimestral o anual (Lind *et al.*, 2015). Su análisis es útil para la toma de decisiones en el presente y generar proyecciones de la información, suponiendo que los patrones de comportamiento pasados, continuarán en el futuro. Las series de tiempo constan de tres componentes:

- a) Tendencia: Se puede definir como un cambio a largo plazo que se produce en la relación al nivel medio, o el cambio a largo plazo de la media. La tendencia se identifica con un movimiento suave de la serie a largo plazo.
- b) Variación estacional: Muchas series temporales presentan cierta periodicidad o, dicho de otro modo, variación de cierto período (mensual, semestral, etc.).
- c) Variación Irregular (Aleatoria): Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada en una serie de tiempo.

Dependiendo de su comportamiento las series temporales se pueden clasificar en:

- a) Estacionarias: Una serie es estacionaria cuando es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.
- b) No estacionarias: Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

En la bibliografía se describen diversos modelos de series de tiempo (Hildebrand, 1998; Chatfield, 2003; Guerrero, 2003; Bowerman, 2009), sin embargo, se destaca Los modelos ARIMA por su simplicidad y practicidad en la modelación y pronóstico (Amaris *et al.*, 2017). Este modelo consiste en la combinación de un término autorregresivo (AR) y un término de media móvil (MA)

con un elemento diferenciador dado por la letra I, basado en un estudio realizado por Yaglom (1955). En general, estos modelos se referencian con la palabra ARIMA (p,d,q)(P,D,Q)[s], en donde el primer paréntesis (p, d, q) se refiere a la parte regular de la serie, el segundo paréntesis (P, D, Q) se refiere a las variaciones estacionales y [s] corresponde al orden de la estacionalidad el cual es 12 en las series anuales, 4 en las trimestrales y 2 en las semestrales. Los modelos ARIMA se pueden generalizar como:

$$Y_t = -(\Delta^d Y_t - Y_t) + \phi_0 + \sum_{i=1}^p \phi_i \Delta^d Y_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Donde:

Y_t corresponde al valor actual de la serie.

Δ^d son las diferenciaciones que necesarias para eliminar la tendencia de la serie.

ϕ_1, \dots, ϕ_p son los parámetros correspondientes a la parte autorregresiva del modelo (AR).

$\theta_1, \dots, \theta_q$ corresponden a los parámetros de la parte de media móviles del modelo (MA).

ϕ_0 es una constante.

ε_t es el término de error.

Existen numerosos estudios donde se han aplicado los modelos ARIMA para estudiar y predecir el comportamiento en el corto y mediano plazo de los precios de productos de la industria forestal (Yin, 1999; Mei *et al.*, 2010; Broz y Viego, 2014), teniendo buenos resultados en dicho objetivo.

Dados estos antecedentes, en el presente estudio se realizó el ajuste de modelos ARIMA para las series de tiempo de precios de productos forestales chilenos, con la finalidad de probar la utilidad de esta herramienta en la predicción de los precios de trozas aserrables y madera aserrada en regiones seleccionadas del país.

METODOLOGÍA

En el presente estudio se analizó el comportamiento de los valores nominales de dos productos forestales; trozas aserrables y madera aserrada (en pesos chilenos/m³) de pino radiata. En base a la información disponible en la Plataforma de Estadísticas Forestales (PEF) de INFOR, se construyeron series de tiempo de los precios de los productos seleccionados, considerando las regiones de Biobío, La Araucanía y Los Lagos, ajustando modelos de series de tiempo locales. Las series de tiempo de precios disponibles en la PEF son de carácter trimestral y abarcan desde el año 1980 hasta hoy.

Para el análisis de las series de tiempo se utilizó el protocolo de Box y Jenkins (Box y Jenkins, 1976) quienes especifican los siguientes pasos para definir el mejor modelo ARIMA:

a) Identificación: Se representa gráficamente la serie de tiempo con lo que se indica si es o no una Serie Estacionaria. En caso de que la serie tenga tendencia, se realizan diferenciaciones hasta hacerla estacionaria, probando esto mediante la prueba ADF de Dickey y Fuller (1979). La diferenciación corresponde a una operación en donde a cada uno de los valores de la serie se resta el anterior de manera de eliminar cualquier tendencia que esta vaya acumulando. Una vez realizadas las transformaciones sobre la serie, se determina su función de autocorrelación simple (ACF) y función de autocorrelación parcial (PACF), con lo que se puede determinar la estructura de la serie.

b) Estimación y verificación: Observando las dos gráficas del ACF y PACF de la serie transformada se plantean los distintos modelos candidatos ARIMA (p,d,q)(P,D,Q)[s]. Se escogieron aquellos modelos que tuvieron los residuales semejantes al de un ruido blanco (residuos no autocorrelacionados), para lo que se aplicó la Prueba de Ljung-Box (Ljung y Box, 1978). El estadístico asociado a esta evaluación estadística se define como:

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k}$$

Donde:

$\hat{\rho}_k$: Autocorrelación de la muestra en el retraso

k, h : Número de retardos que se están probando

n : Tamaño de la muestra

El nivel de significancia asociado a la región crítica para el rechazo de la hipótesis de aleatoriedad es:

$$Q > \chi^2_{1-\alpha, h}$$

Como un segundo criterio de evaluación se seleccionaron aquellos modelos que presentaron mejores métricas de desempeño, para lo cual se calcularon los valores del “Criterio de información de Akaike” (AIC) (Akaike, 1974) y el “Criterio de información Bayesiana” (BIC) (Schwarz, 1978) los cuales se definen como:

$$AIC = 2k - 2 \ln(L)$$

$$BIC = k \ln(n) - 2 \ln(\hat{L})$$

k : Número de parámetros estimados

\hat{L} : Máximo valor de la función de verosimilitud del modelo

Estos criterios de información ayudan a discriminar la eficiencia que ganan los modelos al ir agregando nuevos parámetros, de manera de favorecer la sencillez y flexibilidad de los modelos.

c) Predicción: Una vez seleccionado el modelo, se realizaron predicciones de los valores de la serie. Para evaluar la precisión de las predicciones del modelo, se determinaron los indicadores 'Raíz del Error Cuadrático Medio' (RMSE) y el 'Error absoluto medio' (MAE), los cuales también servirán para discriminar según la capacidad predictiva de los modelos.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}$$

$$MAE = \frac{|\hat{y}_t - y_t|}{n}$$

Donde:

\hat{y}_t : Valor estimado por el modelo

y_t : Valor observado

n : Número de observaciones

Los últimos registros de las series de tiempo serán utilizados para validar las predicciones del modelo y determinar la diferencia porcentual entre los valores registrados y los estimados.

Fuera de la metodología planteada por Box y Jenkins, para cada una de las series temporales se ajustó un modelo automático el cual es sugerido por la función *Autoarima* del programa estadístico R. Para realizar los análisis de las series de tiempo se utilizaron las siguientes librerías del programa R (R Core Team ,2021):

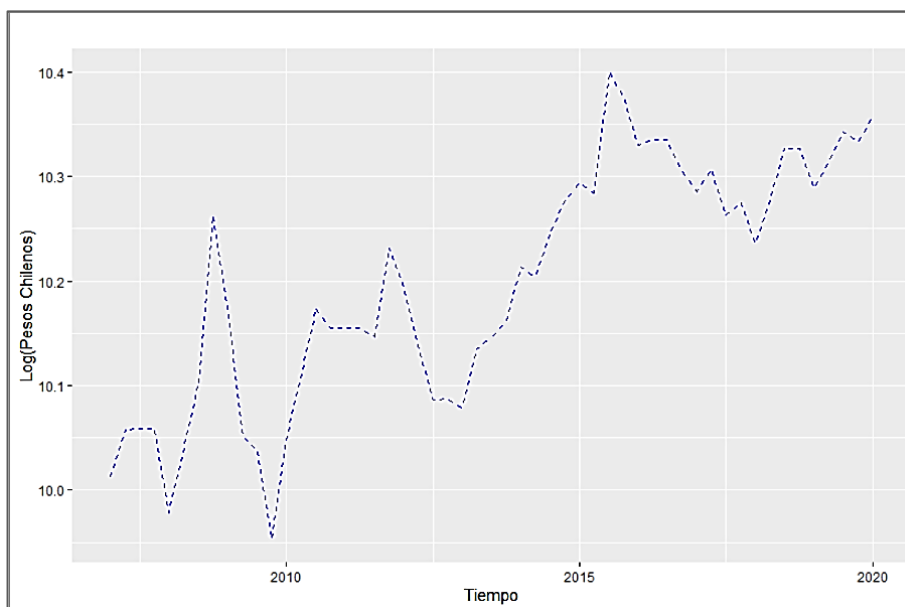
- tseries (Trapletti y Hornik, 2020): Procesamiento de series de tiempo
- forecast (Hyndman y Khandakar, 2008): Ajuste de modelos ARIMA
- ggplot2 (Wickham, 2016) y gridExtra (Aguirre, 2017): Gráficas

RESULTADOS

Trozas Aserrables

a) Región del Biobío

Para el análisis de esta región se consideró la información disponible de precios de trozas aserrables de la Plataforma de Estadísticas Forestales (PEF) desde el año 2007 hasta el primer trimestre del año 2021, debido a que en este periodo presentaba registros más uniformes. Se completó, mediante interpolación lineal, la medición faltante para el tercer trimestre del año 2011. Los datos fueron transformados mediante una función logarítmica para poder mejorar la predicción del modelo. Para ajustar los modelos candidatos se excluyeron los últimos cinco registros de la serie que serán utilizados para el proceso de validación (Figura N°1).

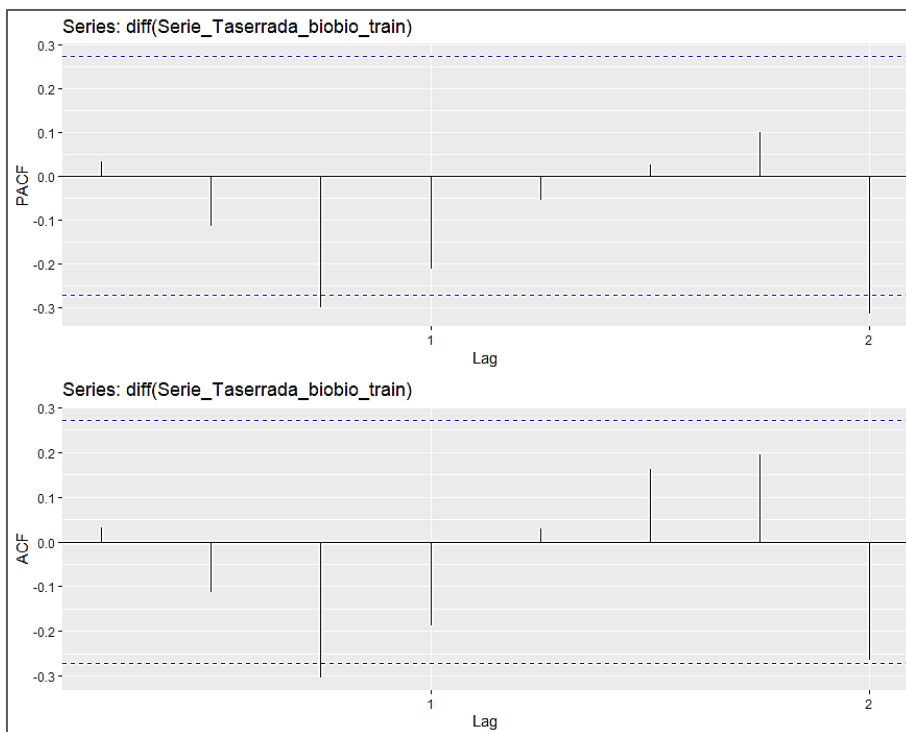


(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°1
SERIE DE TIEMPO LOGARÍTMICA DE PRECIOS DE TROZAS ASERRABLES DE
PINO RADIATA EN LA REGIÓN DEL BIOBÍO

Al aplicar la prueba ADT a la serie de tiempo graficada en la Figura N°1, esta indicó que se necesitaba una diferenciación para eliminar la tendencia de la serie y que esta fuera estacionaria. Las PACF y ACF asociadas a la serie diferenciada de los precios de trozas aserrables en la Región del Biobío presentó valores significativos en el tercer retraso (Figura N°2). Además, el PACF marca un retraso significativo en el segundo año lo cual es similar al

comportamiento en el ACF. Con esta información se ajustaron los modelos candidatos para esta serie.



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°2
ACF Y PACF DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE
TROZAS ASERRABLES EN LA REGIÓN DE BIOBÍO

De los modelos ARIMA ajustados para la serie de precios de trozas aserrables en la Región del Biobío (Cuadro N°1), el Modelo V presentó mejores métricas de desempeño, tanto a nivel de datos de entrenamiento como de validación.

Si bien los valores de AIC y BIC son inferiores a los presentados por otros modelos candidatos, el “Modelo V” tiene valores similares, indicando que sigue siendo eficiente en términos de la utilización de parámetros en función del ajuste que realiza.

Así mismo, el modelo en cuestión cumple con la validación de los residuos según el estadístico de Ljung-Box, por lo que cumple con el supuesto de distribución aleatoria de los residuos.

Cuadro N°1
MODELOS AJUSTADOS Y MÉTRICAS DE RENDIMIENTO DE SERIE DE TIEMPO DE PRECIOS DE
TROZAS ASERRABLES EN LA REGIÓN DE BIOBÍO

Modelo		AIC	BIC	Métricas Entrenamiento		Métricas Validación		Estadístico de Ljung-Box
				RMSE	MAE	RMSE	MAE	
Modelo I	ARIMA(0,1,0)(0,0,2)[4]	-159,42	-158,92	0,0482	0,0362	0,0434	0,0412	Rechazado
Modelo II	ARIMA(0,1,0)(2,0,2)[4]	-165,65	-155,90	0,0393	0,0294	0,0355	0,0333	Aceptado
Modelo III	ARIMA(0,2,0)(2,0,2)[4]	-130,17	-128,84	0,0550	0,0424	0,0231	0,0213	Rechazado
Modelo IV	ARIMA(0,1,1)(2,0,2)[4]	-163,74	-161,87	0,0394	0,0294	0,036	0,0339	Aceptado
Modelo V	ARIMA(1,2,1)(2,0,2)[4]	-154,27	-151,66	0,0388	0,0292	0,0142	0,013	Aceptado

(Fuente: Elaboración propia)

Una vez elegido el modelo se realizó una validación de las predicciones con la fracción de datos que no se utilizaron en el ajuste del modelo, determinando la variación porcentual entre la predicción del modelo y los valores observados.

En el Cuadro N°2 se puede apreciar que el nivel de error del modelo para valores de la serie sin ajuste logarítmico varió entre 0,32% y 1,83%.

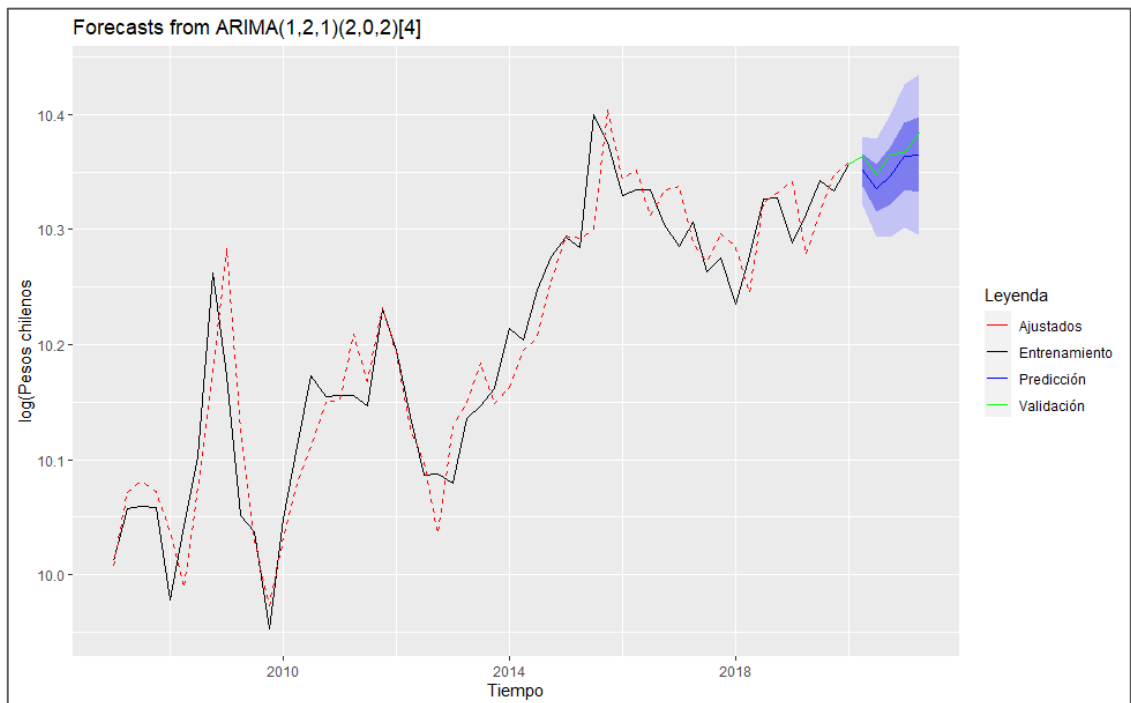
Cuadro N°2
VARIACIÓN ENTRE PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE BIOBÍO
ESTIMADOS CON EL MODELO V Y LOS VALORES OBSERVADOS

Año	Trimestre	Serie logarítmica			Serie sin transformar		
		Observado	Predicción	Variación (%)	Observado (\$/m³)	Predicción (\$/m³)	Variación (%)
2020	T2	10,364	10,352	0,12	31.685	31.304	1,20
2020	T3	10,348	10,336	0,12	31.195	30.823	1,19
2020	T4	10,366	10,347	0,18	31.750	31.151	1,89
2021	T1	10,367	10,364	0,03	31.792	31.689	0,32
2021	T2	10,384	10,365	0,18	32.325	31.735	1,83

(Fuente: Elaboración propia)

En la Figura N°3 se pueden visualizar los datos originales de la serie de tiempo y los valores que ajusta y predice el Modelo V.

Se puede apreciar que los datos de validación se encuentran dentro del rango definido para la predicción generada por el modelo, siguiendo la tendencia que presentan estos datos.



(Fuente: Elaboración propia, en base a INFOR, 2021.)

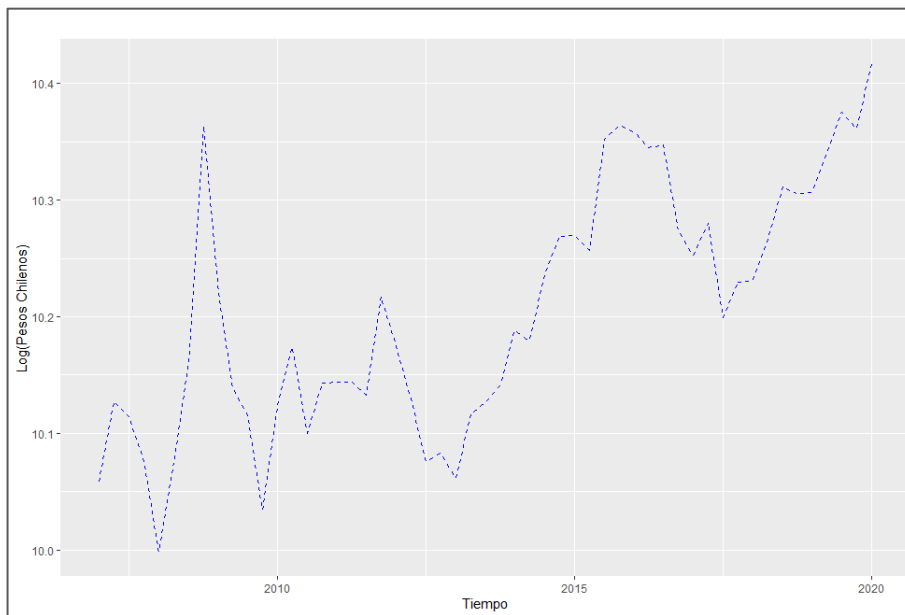
Figura N°3
AJUSTES Y PREDICCIONES REALIZADAS POR EL MODELO V PARA LA SERIE DE TIEMPO DE
PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE BIOBÍO

b) Región de la Araucanía

Para la Región de la Araucanía se utilizaron los precios recopilados en la PEF desde el año 2007 hasta el primer trimestre del año 2021, ya que este periodo tiene información más uniforme.

La medición faltante para el tercer trimestre del año 2011 fue interpolada linealmente. Los datos fueron transformados mediante una función logarítmica de manera de aumentar la eficiencia predictiva de los modelos a ajustar (Figura N°4).

Esta serie presenta una marcada tendencia al alza, aumentando sus valores en el tiempo.



(Fuente: Elaboración propia, en base a INFOR, 2021)

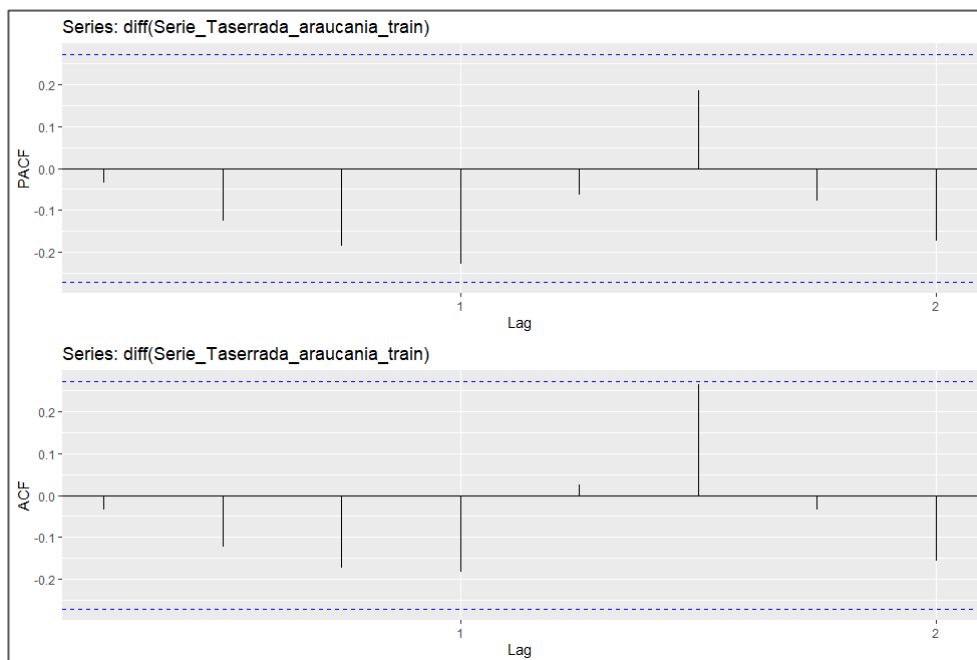
Figura N°4
SERIE DE TIEMPO LOGARÍTMICA DE PRECIOS DE TROZAS ASERRABLES DE PINO RADIATA EN LA REGIÓN DE LA ARAUCANÍA

La prueba de ADT indicó que la serie de precios de trozas aserrables para la Región de la Araucanía necesita una diferenciación para poder eliminar la tendencia de los datos. Por otro lado, las PACF y ACF asociadas a esta serie de tiempo (Figura N°5) muestran que no existen rezagos significativos, lo que indica una carencia o muy baja influencia del término autorregresivo (AR) y el término de la media móvil (MA). Utilizando la información de estos análisis se ajustaron los modelos que se presentan en el Cuadro N°3.

Cuadro N°3
MODELOS AJUSTADOS Y MÉTRICAS DE RENDIMIENTO DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LA ARAUCANÍA

Modelo		AIC	BIC	Métricas Entrenamiento		Métricas Validación		Estadístico de Ljung-Box
				RMSE	MAE	RMSE	MAE	
Modelo I	ARIMA(0,1,0)	-146.39	-144.45	0.0576	0.0425	0.0258	0.0236	Aceptado
Modelo II	ARIMA(1,1,0)	-144.42	-140.53	0.0575	0.0425	0.0249	0.0249	Aceptado
Modelo III	ARIMA(0,1,1)	-144.44	-140.54	0.0575	0.0425	0.0246	0.0225	Aceptado
Modelo IV	ARIMA(1,1,1)	-144.76	-138.91	0.0562	0.0417	0.0340	0.0261	Aceptado

(Fuente: Elaboración propia, en base a INFOR, 2021.)



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°5
ACF Y PACF DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE
TROZAS ASERRABLES EN LA REGIÓN DE LA ARAUCANÍA

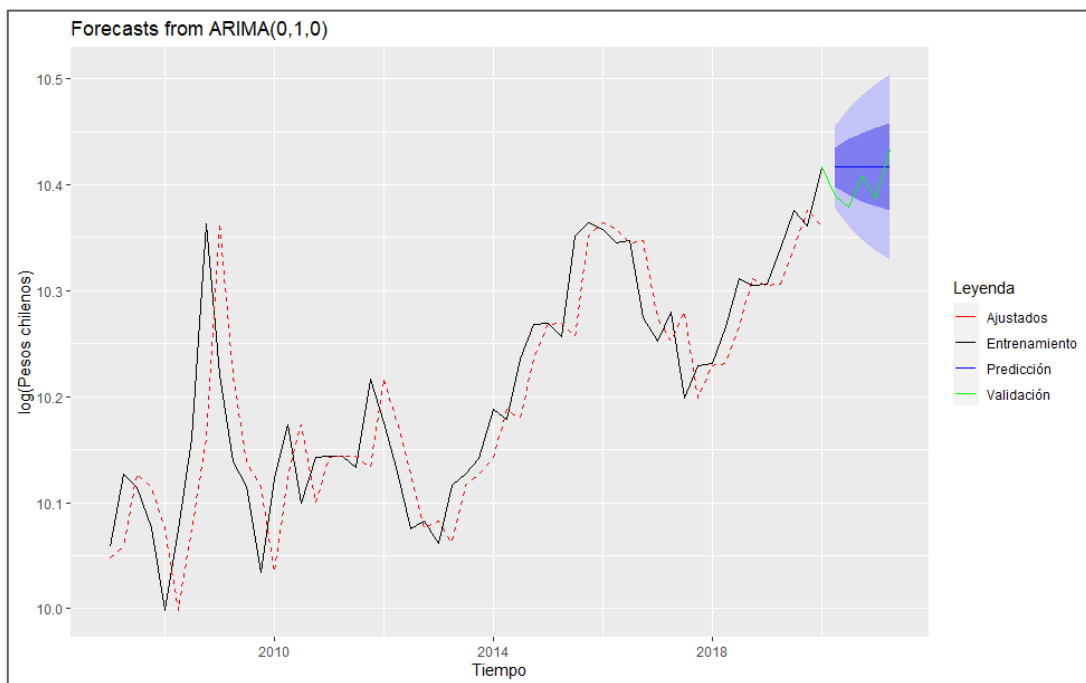
De los modelos ajustados para la serie trozas aserrables en la Región de la Araucanía (Cuadro N°3), se obtuvo que el Modelo I es aquel que optimiza de mejor manera la cantidad de parámetros en función del ajuste final, dado que posee los valores más bajos para el AIC y BIC, aunque el resultado no difiere mucho con respecto a los otros modelos implementados. Por otra parte, el Modelo IV obtuvo el mejor desempeño en las métricas de entrenamiento, pero con un bajo desempeño ajustado a valores desconocidos para el modelo (Validación). Cabe destacar que, si bien el Modelo III muestra las mejores métricas de validación, estas no difieren demasiado con las que presenta el Modelo I.

Al validar las proyecciones del Modelo I con datos que no fueron usados para su ajuste (Cuadro N°4), se aprecia que existe una variación porcentual menor a un 1% cuando se trabaja con datos logarítmicos y de un rango entre 0,79% y 2,29% con los datos sin transformar. Cabe mencionar que este modelo en particular, proyecta que el precio de los próximos períodos es constante, ya que el resto de las posibles alternativas generan un mayor error en la estimación. En la Figura N°6 se aprecia el ajuste que realiza el modelo de los datos ¿entrenados? además de la proyección del modelo, la que se aproxima de buena manera a los valores de validación.

Cuadro N°4
VARIACIÓN ENTRE PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LA ARAUCANÍA
ESTIMADOS CON EL MODELO I Y LOS VALORES OBSERVADOS

Año	Trimestre	Serie logarítmica			Serie sin transformar		
		Observado	Predicción	Variación (%)	Observado (\$/m³)	Predicción (\$/m³)	Variación (%)
2020	T2	10,390	10,417	-0,26	32.523	33.414	-2,74
2020	T3	10,379	10,417	-0,36	32.177	33.414	-3,85
2020	T4	10,409	10,417	-0,08	33.152	33.414	-0,79
2021	T1	10,388	10,417	-0,28	32.463	33.414	-2,93
2021	T2	10,433	10,417	0,16	33.964	33.414	1,62

(Fuente: Elaboración propia)

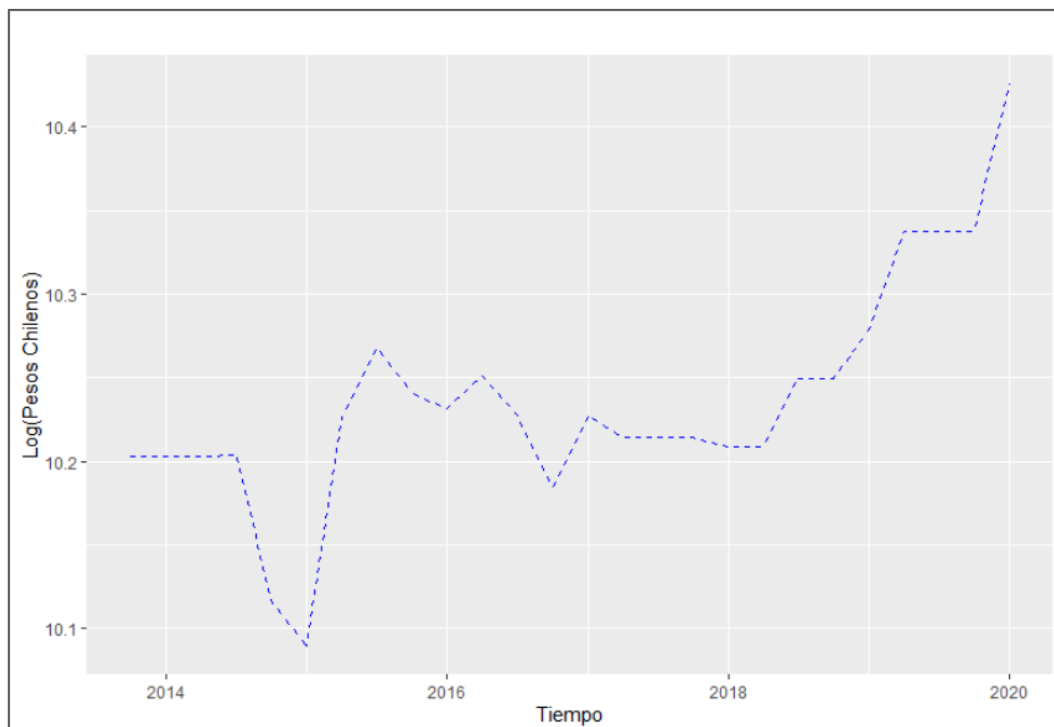


(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°6
AJUSTES Y PREDICCIONES REALIZADAS POR EL MODELO I PARA LA SERIE DE TIEMPO DE
PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LA ARAUCANÍA

c) Región de Los Lagos

Para la región de Los Lagos se utilizaron los precios recopilados en la PEF desde el último trimestre del año 2013 hasta el primer trimestre del año 2021, puesto que este periodo tiene información más uniforme. Los datos fueron transformados mediante una función logarítmica de manera de aumentar la eficiencia predictiva de los modelos a ajustar (Figura N°7).



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°7
SERIE DE TIEMPO LOGARÍTMICA DE PRECIOS DE TROZAS ASERRABLES
DE PINO RADIATA EN LA REGIÓN DE LOS LAGOS

La prueba de ADT muestra que son necesarias dos diferenciaciones para que la serie de precios de trozas aserrables para la Región de Los Lagos sea estacionaria.

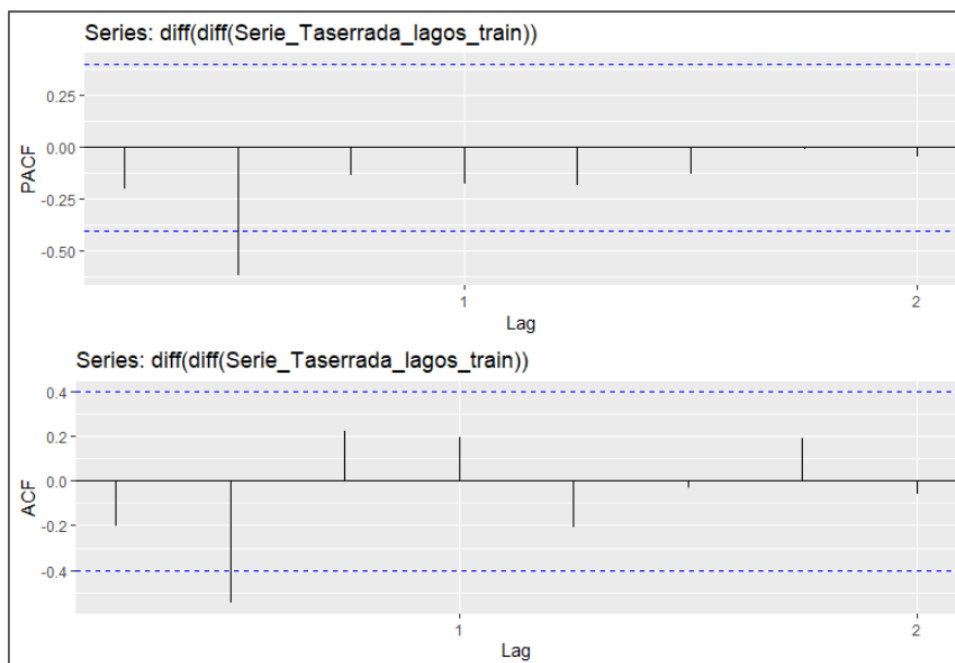
Las PACF y ACF asociadas a esta serie de tiempo (Figura N°8) tienen un rezago significativo, lo que indica que existe influencia de término autorregresivo (AR) y término de la media móvil (MA).

Utilizando la información de estos análisis se ajustaron los modelos que se presentan en el Cuadro N°5.

Cuadro N°5
MODELOS AJUSTADOS Y MÉTRICAS DE RENDIMIENTO DE SERIE DE TIEMPO
DIFERENCIADA DE PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LOS LAGOS

Modelo		AIC	BIC	Métricas Entrenamiento		Métricas Validación		Estadístico de Ljung-Box
				RMSE	MAE	RMSE	MAE	
Modelo I	ARIMA(0,1,0)	-83.327	-82.108	0.0431	0.0271	0.0929	0.0717	Aceptado
Modelo II	ARIMA(1,2,1)	-72.886	-69.352	0.0427	0.0268	0.0583	0.0582	Rechazado
Modelo III	ARIMA(2,2,1)	-76.171	-71.458	0.0387	0.0271	0.0359	0.0335	Aceptado
Modelo IV	ARIMA(2,2,2)	-75.140	-69.250	0.0378	0.0273	0.0419	0.0381	Aceptado

(Fuente: Elaboración propia)



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°8
ACF Y PACF DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE TROZAS
ASERRABLES EN LA REGIÓN DE LOS LAGOS

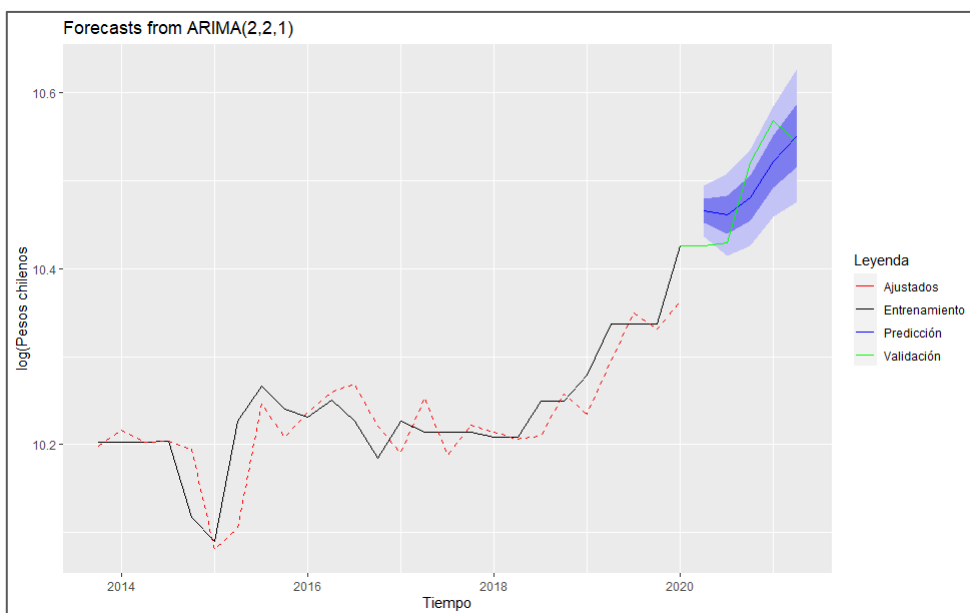
Los modelos ajustados para la serie de precios de trozas aserrables para la Región de Los Lagos se muestran en el Cuadro N°5. El Modelo III presenta las mejores métricas de validación en comparación con el resto de los que se ajustaron, además de ser el segundo mejor modelo en cuanto a las métricas de. Si bien el Modelo I tiene mejores AIC y BIC, este tiene valores de RMSE y MAE de validación de más del doble de los que presenta el Modelo III. Bajo este análisis, el Modelo III es el que presenta mejores predicciones de todos los ajustados.

Cuadro N°6
VARIACIÓN ENTRE DE PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LOS LAGOS
ESTIMADOS CON EL MODELO III Y LOS VALORES OBSERVADOS

Año	Trimestre	Serie logarítmica			Serie sin transformar		
		Observado	Predicción	Variación (%)	Observado (\$/m³)	Predicción (\$/m³)	Variación (%)
2020	T2	10,4256	10,4657	0,38	33.712	35.091	-4,09
2020	T3	10,4292	10,4615	0,31	33.835	34.943	-3,28
2020	T4	10,5205	10,4806	-0,38	37.067	35.618	3,91
2021	T1	10,5682	10,5220	-0,44	38.878	37.124	4,51
2021	T2	10,5430	10,5521	0,09	37.911	38.257	-0,91

(Fuente: Elaboración propia)

Al contrastar las predicciones obtenidas con el “Modelo III” y los valores registrados (Cuadro N°6), se calculó que la variación entre los valores entregados por el modelo con serie logarítmica fue menor al 1%, mientras que en la escala original de la serie este error fue menor al 2%. En la Figura N°9 se aprecia el ajuste del modelo y las variaciones entre la predicción y el conjunto de datos de validación.



(Fuente: Elaboración propia, en base a INFOR, 2021)

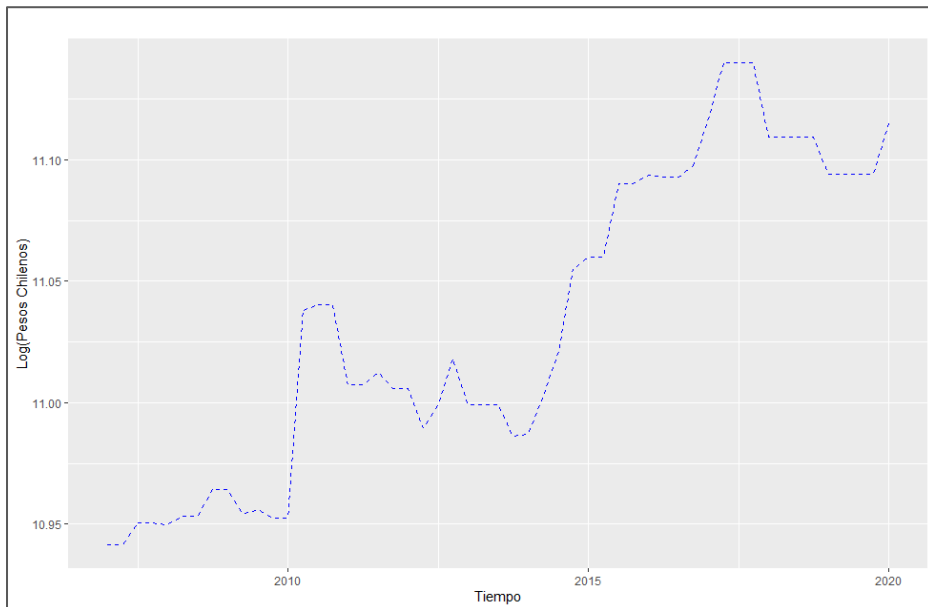
Figura N°9
AJUSTES Y PREDICCIONES REALIZADAS POR EL MODELO III PARA LA SERIE DE TIEMPO DE
PRECIOS DE TROZAS ASERRABLES EN LA REGIÓN DE LOS LAGOS

Madera Aserrada

a) Región del Biobío

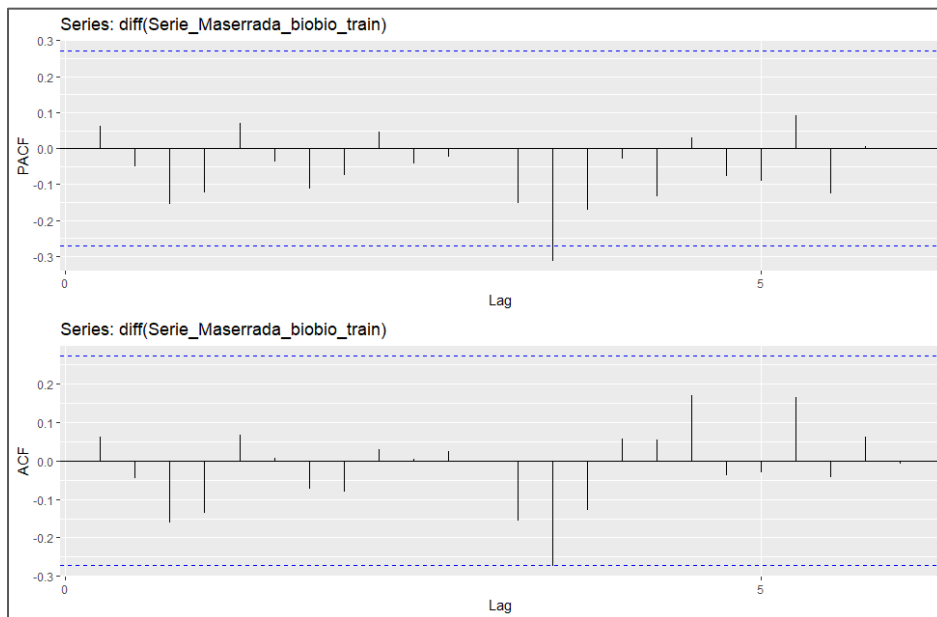
La serie de precios de madera aserrada en la Región del Biobío, tiene información completa desde el primer trimestre del año 2007 hasta el segundo trimestre del año 2021, con la excepción del registro correspondiente al tercer trimestre del año 2011 el cual fue interpolado linealmente.

Se aplicó una transformación logarítmica en la serie de tiempo y se excluyeron los últimos cinco registros de la serie que serán utilizados para el proceso de validación (Figura N°10).



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°10
SERIE DE TIEMPO LOGARÍTMICA DE PRECIOS DE MADERA ASERRADA DE
PINO RADIATA EN LA REGIÓN DE BIOBÍO



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°11
ACF Y PACF DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE
MADERA ASERRADA DE PINO RADIATA EN LA REGIÓN DE BIOBÍO

Mediante la prueba ADT se determinó que es necesaria una diferenciación para eliminar la tendencia de la serie de precios de madera aserrada en la Región del Biobío.

Una vez diferenciada la serie se determinaron sus PACF y ACF, las cuales se muestran en la Figura N°11 e indican que existe un rezago significativo asociado al componente autorregresivo y otro al de la media móvil.

De los modelos ajustados para la serie de precios de madera aserrada en la Región del Biobío (Cuadro N°7), el Modelo II es el que tiene las mejores métricas de validación, prediciendo de mejor manera valores de la serie en periodos más allá del rango con que se entrenó.

Si bien no tiene los mejores valores de AIC y BIC, estos son similares al Modelo I, que es el que tiene los mejores valores en esas métricas.

Cuadro N°7
MODELOS AJUSTADOS Y MÉTRICAS DE RENDIMIENTO DE SERIE DE TIEMPO DE
PRECIOS DE MADERA ASERRADA EN LA REGIÓN DE BIOBÍO

Modelo	AIC	BIC	Métricas Entrenamiento		Métricas Validación		Estadístico de Ljung-Box
			RMSE	MAE	RMSE	MAE	
Modelo I ARIMA(0,1,0)	-273.347	-271.396	0.0170	0.0091	0.2897	0.2162	Aceptado
Modelo II ARIMA(0,1,1)	-271.803	-267.901	0.0170	0.0094	0.2881	0.2148	Aceptado
Modelo III ARIMA(2,2,1)	-271.809	-267.907	0.0170	0.0094	0.2883	0.2150	Aceptado
Modelo IV ARIMA(2,2,2)	-269.810	-263.956	0.0170	0.0094	0.2882	0.2149	Aceptado

(Fuente: Elaboración propia)

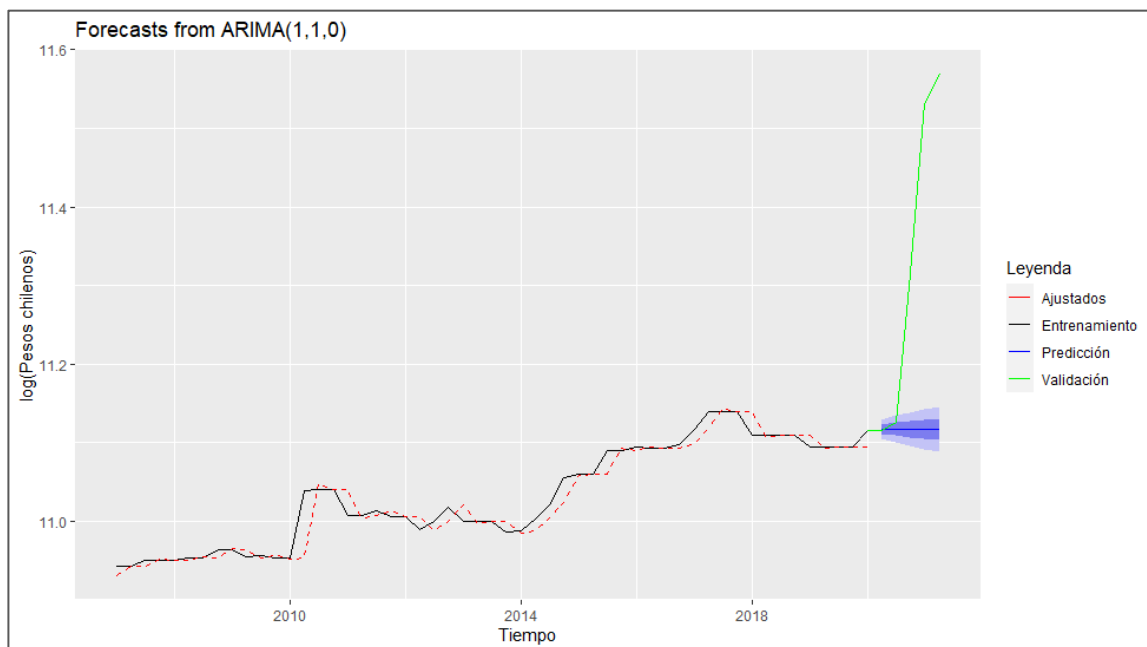
Validando las predicciones que realiza el Modelo II (Cuadro N°8), se aprecia que en los primeros dos trimestres este modelo tiene mucha más precisión ya que presenta variaciones menores al 1% con respecto a los valores registrados.

El resto de los trimestres proyectados van presentando mayores variaciones, las que llegan a ser superiores al 35%. Las grandes diferencias que se registran en ese periodo se deben al explosivo aumento que registró el precio de la madera aserrada durante la pandemia del COVID19, como resultado del aumento de la demanda de este material.

Cuadro N°8
VARIACIÓN ENTRE PRECIOS DE MADERA ASERRADA EN LA REGIÓN DEL BIOBÍO
ESTIMADOS CON EL MODELO II Y LOS VALORES OBSERVADOS

Año	Trimestre	Serie logarítmica			Serie sin transformar		
		Observado	Predicción	Variación (%)	Observado (\$/m³)	Predicción (\$/m³)	Variación (%)
2020	T2	11,1148	11,1167	-0,02	67.155	67.287	-0,20
2020	T3	11,1250	11,1169	0,07	67.848	67.299	0,81
2020	T4	11,3150	11,1169	1,75	82.046	67.300	17,97
2021	T1	11,5313	11,1169	3,59	101.850	67.300	33,92
2021	T2	11,5686	11,1169	3,90	105.730	67.300	36,35

(Fuente: Elaboración propia)



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°12
AJUSTES Y PREDICCIONES REALIZADAS POR EL MODELO IIPARA LA SERIE DE TIEMPO DE
PRECIOS DE MADERA ASERRADA EN LA REGIÓN DEL BIOBÍO

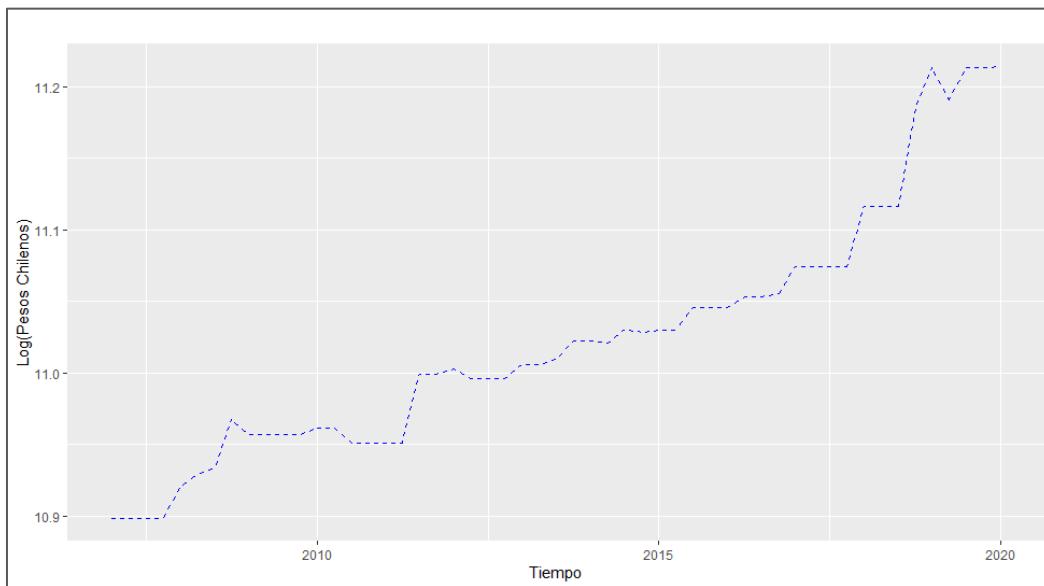
b) Región de la Araucanía

La serie de precios de madera aserrada en la Región de la Araucanía, tiene información que abarca desde el primer trimestre del año 2007 hasta el segundo trimestre del año 2021 periodo que presenta información más completa, con excepción del registro correspondiente al tercer trimestre del año 2011, el cual fue interpolado linealmente.

Se aplicó una transformación logarítmica en la serie de tiempo y se excluyeron los últimos cinco registros de la serie que serán utilizados para el proceso de validación (Figura N°13).

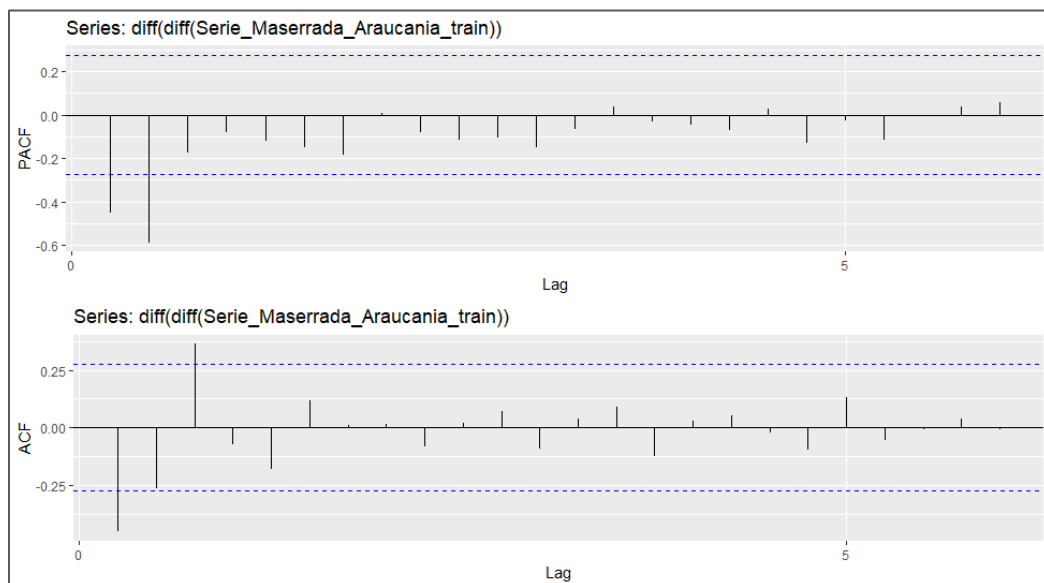
Al aplicar la prueba ADT se determinó que son necesarias dos diferenciaciones para eliminar la tendencia de la serie de precios de madera aserrada en la Región de la Araucanía.

Una vez diferenciada la serie se determinaron sus PACF y ACF, las cuales se muestran en la Figura N°14 e indican que existe un rezago significativo asociado al componente autorregresivo y otro al de la media móvil.



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°13
SERIE DE TIEMPO LOGARÍTMICA DE PRECIOS DE MADERA ASERRADA DE
PINO RADIATA EN LA REGIÓN DE LA ARAUCANÍA



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°14
ACF Y PACF DE SERIE DE TIEMPO DIFERENCIADA DE PRECIOS DE MADERA
ASERRADA DE PINO RADIATA EN LA REGIÓN DE LA ARAUCANÍA

De los modelos propuestos para la serie de precios de madera aserrada de pino radiata en la Región de la Araucanía (Cuadro N°9), el que presenta los mejores resultados en el ajuste en las métricas de validación es el Modelo III, al mismo tiempo que no tiene indicadores significativamente menores en las métricas de entrenamiento.

Cuadro N°9
MODELOS AJUSTADOS Y MÉTRICAS DE RENDIMIENTO DE SERIE DE TIEMPO DE
PRECIOS DE MADERA ASERRADA EN LA REGIÓN DE LA ARAUCANÍA

Modelo		AIC	BIC	Métricas Entrenamiento		Métricas Validación		Estadístico de Ljung-Box
				RMSE	MAE	RMSE	MAE	
Modelo I	ARIMA(0,1,0)	-284.613	-280.710	0.0150	0.0102	0.2864	0.2164	Aceptado
Modelo II	ARIMA(1,2,1)	-272.419	-266.623	0.0152	0.0096	0.2838	0.2143	Aceptado
Modelo III	ARIMA(2,2,1)	-272.949	-265.222	0.0148	0.0096	0.2723	0.2067	Aceptado
Modelo IV	ARIMA(2,2,2)	-272.961	-263.301	0.0146	0.0094	0.2729	0.2076	Aceptado

(Fuente: Elaboración propia)

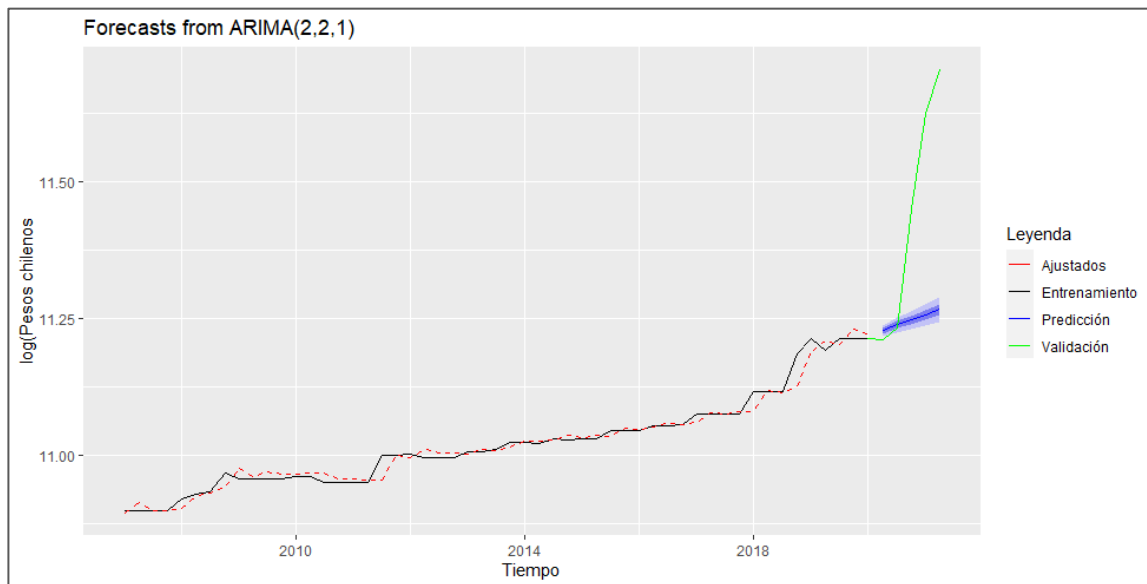
Al evaluar las diferencias porcentuales entre los valores inferidos por el modelo y los usados de validación (Cuadro N°10), se observa que la diferencia con respecto a los dos primeros trimestres es menor a un 2%, mientras que a partir de este punto comienza a elevarse, llegando a un 35,56%.

El incremento en las diferencias comienza en el mismo periodo que en la Región del Biobío, por lo que también es atribuible al aumento del precio de la madera aserrada durante la pandemia del COVID19, derivado del aumento en la demanda de este material.

Cuadro N°10
VARIACIÓN ENTRE DE PRECIOS DE MADERA ASERRADA EN LA REGIÓN DEL BIOBÍO
ESTIMADOS CON EL MODELO III Y LOS VALORES OBSERVADOS

Año	Trimestre	Serie logarítmica			Serie sin transformar		
		Observado	Predicción	Variación (%)	Observado (\$/m³)	Predicción (\$/m³)	Variación (%)
2020	T2	11,2113	11,2273	-0,14	73.963	75.152	-1,61
2020	T3	11,2329	11,2385	-0,05	75.579	75.997	-0,55
2020	T4	11,4514	11,2470	1,79	94.036	76.646	18,49
2021	T1	11,6246	11,2563	3,17	111.819	77.369	30,81
2021	T2	11,7056	11,2663	3,75	121.250	78.139	35,56

(Fuente: Elaboración propia)



(Fuente: Elaboración propia, en base a INFOR, 2021)

Figura N°15
AJUSTES Y PREDICCIONES REALIZADAS POR EL MODELO III PARA LA SERIE DE TIEMPO DE
PRECIOS DE MADERA ASERRADA EN LA REGIÓN DE LA ARAUCANÍA

Debido a la falta de información para diversos periodos, no fue posible construir una serie de tiempo para los precios de la madera aserrada en la Región de los Lagos con la robustez necesaria para aplicar correctamente la metodología planteada en este estudio.

CONCLUSIONES

La información recolectada y administrada por INFOR tiene un importante potencial que va más allá de ser un registro histórico de los precios de los diversos productos forestales. Como se demuestra en el presente trabajo, también puede ser usada para generar nuevas herramientas de gestión que permitan proyectar información útil para el mercado de la madera y con ello fomentar su uso en la construcción.

Los modelos ajustados para trozas aserrables de pino radiata, predijeron de buena manera los precios de este producto. En términos generales, las variaciones obtenidas fueron menores a un 5%, validando con mediciones de 5 trimestres. Los mejores resultados se obtuvieron para la Región del Biobío, en donde las variaciones no superaron el 2%.

En el caso de las series de tiempo de precios de madera aserrada, los modelos presentaron un buen ajuste al inicio del periodo usado para validar la información (Segundo y tercer trimestres del 2020), teniendo variaciones que no superaban el 2%. En los trimestres posteriores se presentaron diferencias superiores al 30%, lo cual es atribuible al inusual aumento de precio de este material experimentado durante la pandemia del COVID 19, como efecto de la inyección de subsidios estatales en la población y el retiro de fondos de pensión

Si se comparan los resultados obtenidos en este estudio con los de otros autores, se observa que son similares a los de Broz y Viego (2014), quienes lograron tener diferencias porcentuales de menos de 10% al estimar los valores de precios de productos forestales en un periodo de nueve meses.

Los modelos ARIMA son una buena herramienta, que permite reducir la incertidumbre respecto de las proyecciones de precios de productos forestales. Sin embargo, una limitación para su uso son las fluctuaciones bruscas de los mercados debido a situaciones extraordinarias, como fue el caso de la pandemia de COVID-19 y la necesaria inyección de recursos monetarios a la población.

REFERENCIAS

- Aguirre, B., 2017.** gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3.
- Akaike, H., 1974.** A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.
- Amaris, G; Ávila, H. y Guerrero, T., 2017.** Aplicación de modelo ARIMA para el análisis de series de volúmenes anuales en el río Magdalena. Revista Tecnura, Tecnura. 2017.2.a07.
- Bowerman, N., 2009.** Pronósticos, series de tiempo y regresión: un enfoque aplicado. 720p. International Thomson Editores De México, DF.
- Box, G. y Jenkins, G., 1976.** Time Series Analysis: Forecasting and Control. Series in Probability and Statistics. Holden-Day, U.S.A.
- Broz, D. y Viego, V., 2014.** Predicción de precios de productos de *Pinnus spp.* con modelos ARIMA. Madera y Bosques 20 (1): 37-46.
- Chatfield, C., 2003.** The Analysis of Time Series, An Introduction. 352p. Chapman and Hall/CRC, United States of America.
- Dickey, D. y Fuller, W., 1979.** Distribution of the Estimators for Autoregressive Time Series with a unit root. Journal of the American Statistical Association 74(366):427-431.
- Guerrero, V., 2003.** Análisis estadístico de series de tiempo económicas. 395p Thomson, México, D.F.
- Hildebrand, D., 1998.** Estadística aplicada a la administración y a la economía. Serie de probabilidad y estadística. 945p. Addison Wesley, México, DF.
- Hyndman, R.J., y Khandakar, Y., 2008.** Automatic time series forecasting: the forecast package for R. _Journal of Statistical Software, *26*(3), 1-22.
- INFOR, 2020.** El Mercado de la Madera Aserrada para uso estructural en Chile. Santiago, Chile. Instituto Forestal (INFOR). 115p
- INFOR, 2021.** La Industria del Aserrío 2021. Instituto Forestal. 124p.
- INE, 2021.** Bases de datos de permisos de edificación, años 2002 al 2019. [En línea]. Disponible en <https://www.ine.cl/estadisticas/economia/edificacion-y-construccion/permisos-de-edificacion> [Consulta: 20 de agosto de 2021]
- Lind, D., Marchal, W. y Wathen, S., 2015.** Estadística aplicada a los negocios y la economía. 736p, McGraw-Hill, México. DF.
- Ljung, G. M., y Box, G. E., 1978.** On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.
- Mei, B., Clutter, M., y Harris, T., 2010.** Modeling and forecasting pine sawtimber stumpage prices in the US South by various time series models. Canadian Journal of Forest Research, 40(8), 1506-1516.

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Schwarz, E., 1978. Estimating the dimension of a model». *Annals of Statistics* 6 (2): 461-464

Trapletti A. y Hornik K., 2020. tseries: Time Series Analysis and Computational Finance. R package version 0.10-48.

Wickham H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York,

Yaglom, A., 1955. Correlation theory of processes with random stationary nth increments. In *Matematicheskii Sbornik* (N. S.), volume 37(79), number 1, pages 141–196

Yin, R., 1999. Forecasting short-term timber prices with univariate ARIMA models. *South Journal of Applied Forestry* 23(1):53-58.



INFOR



www.infor.cl