World Conference on
Timber Engineering
Oslo 2023

USING ARIMA MODELS TO PROJECT SAWLOGS, AND SAWN WOOD PRICES IN THE CHILEAN CONSTRUCTION MATERIALS MARKET

Wilson Mejías[1], Janina Gysling[2], Carlos Kahler [3], Daniel Soto[4] Evaristo Pardo[5]

ABSTRACT:

One of the most remarkable factors when defining the materiality of a construction is the material cost. In this context, by having the tools to help reduce the uncertainty of wood cost will contribute to carry out projects with this material. The ARIMA models are management tools that have the potential to support projections in the short, and medium term. In the present research, the implementation of local ARIMA models were carried out in Biobío, La Araucanía, and Los Lagos regions, using as supplies the time series of radiata pine (*Pinus radiata*) sawn wood, and sawlogs prices, which have been collected periodically by the Forestry Institute (INFOR). The estimations on the implemented models had positive results in respect of validation data, with the exception of sawn wood prices from the four quarter of 2020, and part of 2021, which presented an unusual increase during the COVID-19 pandemic.

KEYWORDS: Sawn wood, sawlogs, construction with wood, ARIMA models, time series.

1 INTRODUCTION

The Chilean forestry industry produces near to 15 million m³ of radiata pine sawlogs, which are production supplies of approximately 7.9 million m³ of sawn wood, these figures make Chile one of the 15 largest producers of this material in the world [1]. Despite this, the national statistical institute (INE) indicates that approximately 13% of the area which is built annually in the country uses wood as a predominant structural material, which implicates a low consumption of wood estimated in approximately 500.000 m³ [2].

According to INFOR's figures, in 2021, the Biobío, La Araucanía, and Los Lagos regions produced approximately 4.5 million m³ of sawn wood, equivalent to 52% of what Chile produced that year [3]. On the other side, from 2020, a strong increase in sawn wood price has been observed, due to the increase of demand of this material during the COVID-19 pandemic. This increase in consumption was the result of several government policies that allowed a partial withdraw of workers' pension funds, which generated an increase in repairs, and extensions of housing.

According to construction companies, one the most determining factors when deciding on the use of wood is the price [2], therefore, by having models with the ability to generate robust estimations, it would help to have more clarity, and to project costs in a better way for the forestry industry, as well as for the construction industry, and for planification of public investment in this matter.

A time series is a set of data recorded in a monthly, quarterly, or annual period [4]. Its analysis is useful for decision making in the present time, and to generate projection on information, on the assumption that the past behavior patterns will continue in the future. The time series consist in three components:

a) Trend: It can be defined as a change in the long term that it is produced in relation to the average, or the long-term change in the mean. The trend is identified as a smooth movement of series in the long-term period.

b) Seasonal variation: Several time series have certain periodicity, or, in other words, variation of a certain period (monthly, quarterly, etc.).

c) Irregular variation (random): This component does not respond to any behavior pattern, since it is the result of fortuitus, or random factors having a remote impact in a time series.

[1] Wilson Mejías, (INFOR) Instituto Forestal, wmejias@infor.cl
[2] Janina Gysling, (INFOR) Instituto Forestal, jgysling@infor.cl
[3] Carlos Kahler, (INFOR) Instituto Forestal, ckahler@infor.cl
[4] Daniel Soto, (INFOR) Instituto Forestal, dsoto@infor.cl
[5] Evaristo Pardo, (INFOR) Instituto Forestal, epardo@infor.cl

https://doi.org/10.52202/069179-0583

Depending on time-series behavior, they can be classified as:

i) Seasonal: the series are seasonal when it is stable in time, that is to say, when the mean, and variation are constant over time. This is graphically reflected on the fact that the values of series tend to range around a constant mean, and that the variance in respect of that mean also remain constant over time.

ii) Non-seasonal: These are series in which the trend, and/or variability change over time. The changes in the mean determine a long-term increasing, or decreasing trend, so the series do not vary around a constant value

In the biography, several time series models are described [5-7], however, the ARIMA model is noted due to its simplicity, and practicality in modeling, and forecasting [8]. This model consists in a combination of an autoregressive term (AR), and a moving average term (MA) with a differencing element given by the letter I, based on research conducted by Yaglom [9]. In overall, these models are represented with the word ARIMA (p,d,q)(P,D,Q)[s], where the first parenthesis –(p,d,q)-, refers to the regular part of the series, the second parenthesis –(P,D,Q)- refers to the seasonal variations, and [s] corresponds to the seasonality order, which is 12 in annual series, 4 in quarterly series, and 2 in monthly series.

There are many studies where the ARIMA models have been used to analyze, and predict the behavior in short, and medium-term of prices of forestry industry products [10-12], obtaining good results in this regard.

ARIMA models can be generalized as:

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

Where Y_t corresponds to the current value of a series; Δ^d are the differences needed to remove the trend from the series; $\phi_1, \ldots, \phi_p$ are the corresponding parameters of the autoregressive part of the model (AR); $\theta_1, \ldots, \theta_q$ corresponds to the parameters of the moving-average part of the model (MA); ϕ_0 is a constant, and ε_t is the error term.

With this background, in the present research the ARIMA models were fitted to the time series of Chilean forestry product prices to prove the utility of this tool in terms of sawlogs, and sawn wood prices within the selected regions of the country.

2 MATERIALS AND METHODS

In the present research it was analyzed the behavior of nominal values of forestry products: sawlogs, and sawn wood of radiata pine (in Chilean pesos/m³). Based on the information available in INFOR's Forestry Statistical Platform (PEF), time series of selected product prices were created, considering the regions of Biobío, La Araucanía, and Los Lagos, fitting local time-series models. The time series of prices available in the PEF are of quarterly nature, and comprehends year 1980 until now. The prices used to create the series consider the average of quarterly measures recorded in a region for each product, including different provinces, and delivery locations (sawmills, sales plant, or purchase plant).

In the analysis of time series, the Box, and Jenkins protocol [13] was used, which indicates a divided process in the following stages:

a) Identification: the time series are presented graphically indicating if the series are seasonal or not. If the series have a trend, the differencing method is applied until making it seasonal, proving this through the ADF Dickey, and Fuller test [14]. The difference consists in an operation where each of the series values are subtracted from the previous one in order to remove any trend accumulated. Once the transformations on series are carried out, a simple autocorrelation function (ACF), and partial correlation function (PACF) are determined, which allows to determine the series structure, that is to say, if it has mobile average, or autoregressive components.

b) Estimation, and verification: by observing the two ACF, and PACF graphics of the transformed series, the ARIMA (p,d,q)(P,D,Q)[s] candidate models are raised. The models randomly distributed, or in a white noise pattern were selected by using the Ljung-Box test [15]. The statistic associated to this statistical evaluation is defined as:

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

Where $\hat{\rho}_k$: autocorrelation of the sample in the delay, $k\ y\ h$ are lag numbers being proved, and n is the size of the sample. The significance level associated to the critical region to reject the hypothesis of randomness is:

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

As a second criterion of the evaluation, those models presenting better performance metrics were selected by calculating the values of "Akaike Information Criterion" (AIC) [16], and the "Bayesian Information Criterion" (BIC) [17], which are defined as:

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

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

Where k is the number of estimated parameters, and $\hat{L}$ is the maximum value of the model's likelihood function.

These information criteria help to differentiate the efficiency that the models gain when adding new parameters, therefore benefiting the simplicity, and flexibility of models.

c) Prediction: Once the model has been selected, predictions were made for the values of the series. To evaluate the accuracy of the model predictions, the "Root Mean Square of Error" (RMSE), and "Mean Absolut Error" (MAE) indicators were determined, which would also help in differencing according to the predictive capacity of the models. The formulas of these indicators are:

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

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

Where: $\hat{y_t}$ is the value estimated by the model, y_t is the observed value, and n is the number of observations.

The last 5 records of time series were used to validate the model forecasts, and to determine the percentage difference between the recorded, and estimated values.

In addition to what it is stated in the methodology suggested by Box, and Jenkins, for each one of the time series, an automatic model was fitted which is suggested by "auto-ARIMA" function of the statistical software R [18].

Beside of the Box, and Jenkins methodology, for each one of the time series, an automatic model was fitted, which is suggested by the auto-ARIMA function of the statistical software R: tseries [19], forecast [20], ggplot2 [21], and gridExtra [22].

3 RESULTS

3.1 SAWLOGS

a) Biobío region

For the analysis of this region, it was considered the available information of sawlog prices in the platform of forestry industry (PEF) since 2007, to the last quarter of 2022, since in that period it showed more uniform records. The missing measure for the third quarter of 2011 was completed through linear interpolation. The data was transformed through a logarithmic function in order to improve the model forecasting.

When applying the ADT test to the time series of log prices in Biobío region, it indicated that differencing was needed to remove the trend from the series, for this to be a seasonal type of series. However, models with two differences were fitted, since these presented better metrics.

The PACF, and ACF associated to the differenced series of sawlogs prices in Biobío region presented significant values in the third lag (Figure 1). Besides, the PACF records a significant lag in the second year, which is similar to the ACF behavior.

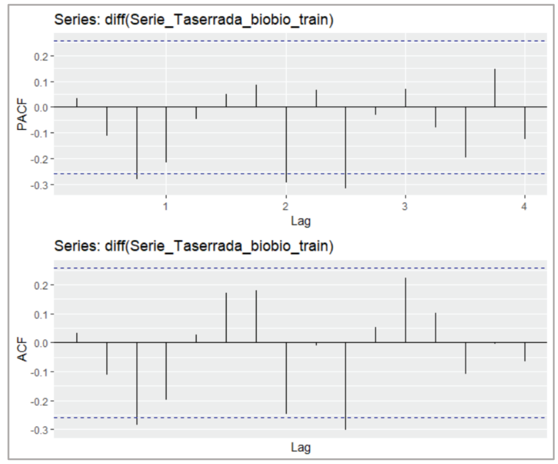


Figure 1: *The ACF, and PACF of differenced time series of sawlogs prices in Biobío region*

From the ARIMA models fitted to the series of sawlog prices in Biobío region (Table 1), the M_4 model showed better performance metrics at validation data level. This model also had no problems with the distribution of its residuals, since the Ljung-Box test did not deliver significant values.

Table 1: *Fitted models, and performance metrics of time-series prices of sawlogs in Biobío region.*

ARIMA Models		AIC	BIC	Train		Test	
				RMSE	MAE	RMSE	MAE
M_1	(0,1,0)(0,0,2)[4]	**-181.69**	**-173.45**	0.0464	0.0354	0.1839	0.1761
M_2	(0,2,0)(0,0,2)[4]	-148.93	-141.79	0.0590	0.0427	0.1103	0.0927
M_3	(1,2,0)(0,0,2)[4]	-152.03	-143.86	0.0567	0.0431	0.0663	0.0515
M_4	(2,2,0)(0,0,2)[4]	-153.40	-143.19	0.0548	0.0410	**0.0459**	**0.0403**
M_5	(3,2,1)(0,0,2)[4]	-173.37	-161.11	**0.0386**	**0.0295**	0.2194	0.2116

Once the model is chosen, a validation of the forecast was carried out with the data slice that was not used in the model fitting, determining the percentage variation between the model prediction, and the observed values.

In Table 2, it can be observed that the error level of the model for the series values without logarithmic fitting varied between 1.58%, and 7.82%.

 https://doi.org/10.52202/069179-0583

Table 2: *Variation of sawlog prices in Biobío region estimated with the M_4 model, and the values observed.*

Año	Trimestre	Observado ($/m3)	Predicción ($/m3)	Variación (%)
2021	Q4	40.271	37.122	7,82%
2022	Q1	39.759	39.129	1,58%
2022	Q2	43.057	41.323	4,03%
2022	Q3	45.501	44.043	3,20%
2022	Q4	45.251	46.653	-3,10%

In Figure 2 it can be seen the original data of time series, and the values that the m4 model fits, and predicts. It also can be observed that the validation data is included, almost entirely, within the defined range of the prediction generated by the model, following the trend that this data presents.

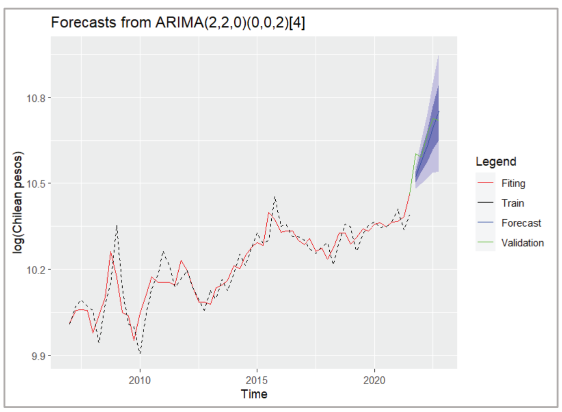


Figure 2: *Logarithmic time series of radiata pine sawlogs prices in Biobío region, forecast, and fitting made with the M_4 model.*

b) La Araucanía Region

For La Araucanía region, the prices recorded in the PEF since 2007, until the last quarter of 2022 were used, since this period has more uniform records. The missing measure in the third quarter of 2011 was linearly interpolated. The data was transformed through a logarithmic function in order to increase the forecast efficiency of the models to be fitted. The ADT test indicated that the prices series of sawlogs in La Araucanía region needed differencing to remove the trend from the data. With this information, models that consider a maximum of two differences were raised.

On the other hand, The PACF, and ACF associated to this time series (Figure 3) show that there is no significant time lag, which indicates a lack, or very low impact of the autoregressive term (AR), and the mobile average term (MA). The models presented in Table 3 were fitted by using the information of this analysis.

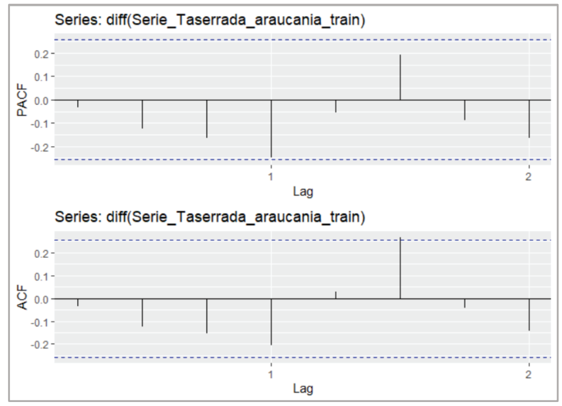


Figure 3: *ACF Y PACF de serie de tiempo diferenciada de precios de trozas aserrables en la Región de la Araucanía.*

From the model fitted for the sawlog series in La Araucanía region (Table 3), it was observed that the M_2 model shows the best validation metrics, which indicates that it generates better projections in relation to the last period data.

Table 3: *Fitted models, and performance metrics of time series of sawlog prices in Biobío region*

ARIMA Models		AIC	BIC	Train		Test	
				RMSE	MAE	RMSE	MAE
M_1	(0,1,0)(0,0,1)[4]	**-167.63**	**-173.45**	**0.0464**	**0.0354**	0.1839	0.1761
M_2	(1,2,0)(0,0,1)[4]	-137.30	-131.17	0.0672	0.0497	**0.0328**	**0.0273**
M_3	(0,2,1)(0,0,1)[4]	-159.01	-152.88	0.0535	0.0394	0.1742	0.1679
M_4	(1,2,1)(0,0,2)[4]	-157.25	-149.08	0.0533	0.0394	0.1796	0.1733
M_5	(0,2,0)(0,0,1)[4]	-123.67	-119.58	0.0774	0.0555	0.0533	0.0437

The residuals of this model were reviewed through the Ljung-Box test, which obtained non-significant values, concluding that the residual distribution corresponds to white noise. By validating the M_2 Model projections with data that was not used in the fitting (Table 4), it is observed that there is a percentage variation which range between 0.26%, to 5.71%.

Table 4: *Variations of sawlog prices in La Araucanía region estimated with the M_2 model, and the observed values.*

Year	Quarter	Observed ($/m3)	Forecast ($/m3)	Variation (%)
2021	Q4	41.032	38.689	5,71%
2022	Q1	42.587	41.431	2,71%
2022	Q2	43.582	43.468	0,26%
2022	Q3	46.179	45.202	2,12%
2022	Q4	45.976	47.216	-2,70%

In Figure 4 it can be seen the original data of time series, and the values that model fits, and forecast for the validation period.

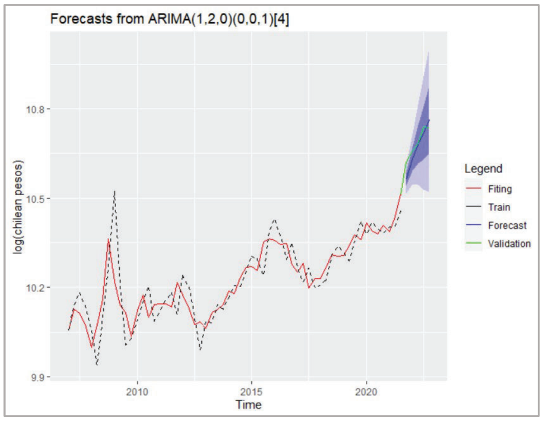


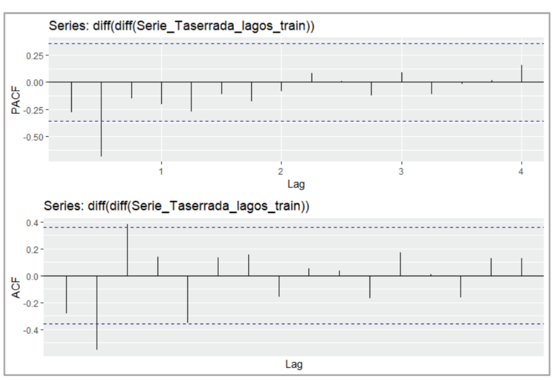
Figure 4: *Logarithmic time series of radiata pine sawlog prices in La Araucanía region, fitting, and forecast carried out with the* M_2 *model.*

c) Los Lagos region

For Los Lagos region, prices recorded in the PEF since the last quarter of 2012 until the first quarter of 2020 were used, since this period has more uniform records.

The ADT test shows that two differences are needed for the series of Los Lagos region's sawlog prices to be seasonal.

Figure 5: *The ACF, and PACF of differenced time series of sawlog prices in Los Lagos region*



The M_4 Model presents the best validation, and train metrics in comparison with the rest that were fitted. While M_1 Model has better AIC, and BIC, this one has RMSE, and MAE validation values more than double that the ones presented in M_4 Model.

Under this analysis, the M_4 Model is the one with the best forecast from all the fitted ones. The M_4 residuals were reviewed through the Ljung-Box, which indicated that the residual distribution corresponds to white noise.

Table 5: *Fitted models, and performance metrics of time series of sawlog prices in Los Lagos region.*

ARIMA Models		AIC	BIC	Train		Test	
				RMSE	MAE	RMSE	MAE
M_1	(0,1,0)	**-101.86**	**-98.99**	0.0432	0.0331	0.1081	0.1046
M_2	(0,2,0)	-79.60	-78.19	0.0601	0.0455	0.1046	0.0836
M_3	(1,2,1)	-92.60	-88.39	0.0441	0.0300	0.0708	0.0698
M_4	(2,2,1)	-98.80	-93.20	**0.0383**	**0.0283**	**0.0277**	**0.0206**
M_5	(2,2,2)	-98.51	-91.51	0.0371	0.0281	0.0357	0.0296

When comparing the forecasts obtained with the M_4 Model, and the recorded values (Table 6), it was calculated that the variation between the values delivered by the model with respect to the real data was lower than 7%. Moreover, in Figure 6 it is observed the model fitting, and variations between the forecast, and the validation data set.

Table 6: *Variation of sawlog prices in Los Lagos region estimated with the* M_4 *model, and the observed values*

Year	Quarter	Observed ($/m3)	Forecast ($/m3)	Variation (%)
2021	Q4	43.911	43.782	0,29%
2022	Q1	46.333	44.343	4,30%
2022	Q2	47.444	44.478	6,25%
2022	Q3	48.778	46.772	4,11%
2022	Q4	50.223	48.476	3,48%

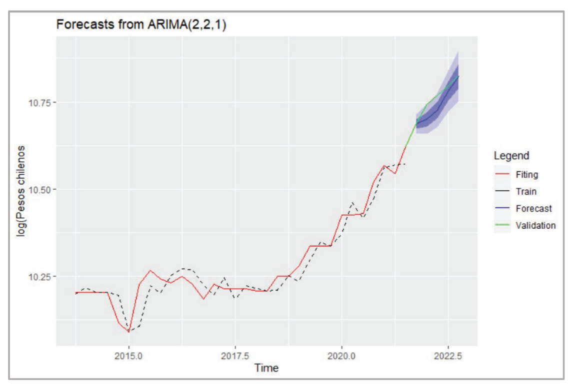


Figure 6: *Logarithmic time series of radiata pine sawlog prices in Los Lagos region, forecast, and fitting carried out with the* M_2 *model*

3.2 SAW WOOD

a) Biobío region

The series of sawn wood prices in Biobío region has complete information since the first quarter of 2007 until the last quarter of 2022, with the exception of the corresponding record in the third quarter of 2011, which was linearly interpolated.

https://doi.org/10.52202/069179-0583

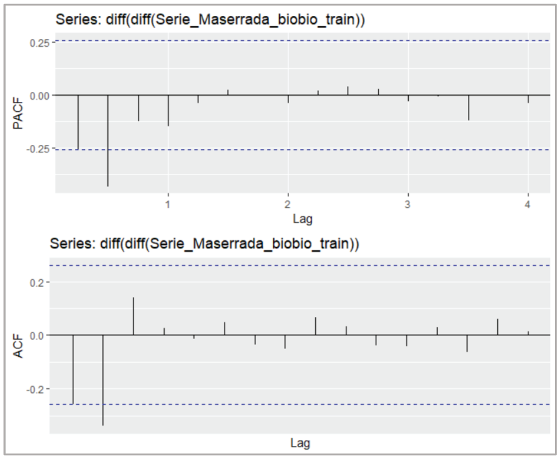


Figure 7: *The ACF, and PACF of differenced time series of sawn wood prices in Biobío region.*

Through the ADT test it was determined that differencing was needed to remove the trend from the series. Once the series are differenced, their PACF, and ACF were determined, which are shown in Figure 7, and they indicate that there is a significant lag associated to the autoregressive component, and another in the mobile average.

Table 7: *Fitted models, and performance metrics of the sawn wood prices time series in Biobío region.*

ARIMA Models		AIC	BIC	Train		Test	
				RMSE	MAE	RMSE	MAE
M_1	(0,2,3)	-209.90	-201.73	0.0350	0.0197	0.3144	0.2680
M_2	(1,1,1)	**-210.73**	**-204.55**	0.0369	0.0192	**0.0348**	**0.0291**
M_3	(1,2,2)	-208.42	-200.25	0.0350	0.0198	0.2115	0.1738
M_4	(2,2,2)	-208.75	-198.54	**0.0347**	**0.0203**	0.3202	0.2792
M_5	(2,2,3)	-206.79	-194.53	0.0347	0.0204	0.3495	0.3067

From the models fitted to the sawn wood prices series in Biobío region (Table 7), the M2 Model is the one with the best AIC, BIC, and validation metrics, forecasting in a better way the values of series beyond the range it was fitted with. The residuals associated to the M2 model were validated through the Ljung-Box test, indicating that it is distributed in a random pattern or as white noise.

Table 8: *Variation of sawn wood prices in Biobío region estimated with M_2 model, and the observed values.*

Year	Quarter	Observed ($/m3)	Forecast ($/m3)	Variation (%)
2021	Q4	134.830	128.604	4,62%
2022	Q1	136.771	136.425	0,25%
2022	Q2	140.650	142.917	-1,61%
2022	Q3	144.530	148.245	-2,57%
2022	Q4	144.530	152.579	-5,57%

When comparing the extrapolations that the M_2 Model produces (Table 8), it is observed that the variation between what is observed, and what is projected by this model is lower than 6%, where the minimum difference was 0.25% in the first quarter of 2022.

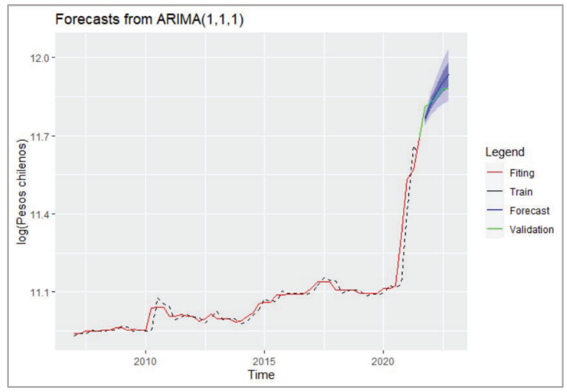


Figure 8: *Logarithmic time series of radiata pine sawn wood prices in Biobío Region, forecasts, and fitting carried out with the M_2 model.*

b) La Araucanía region

The series of sawn wood prices in La Araucanía region has information from the first quarter of 2007, until the last quarter of 2022, which is a period with more complete information, except for the records from the third quarter of 2011, which were linearly interpolated.

When applying the ADT test, it was determined that two differences were needed to remove the trend in the series of sawn wood price in La Araucanía region. Once the differencing was carried out, its PACF, and ACF were determined, which are shown in Figure 8, and they indicate that there is a significant lag associated to the autoregressive component, and another in the mobile average.

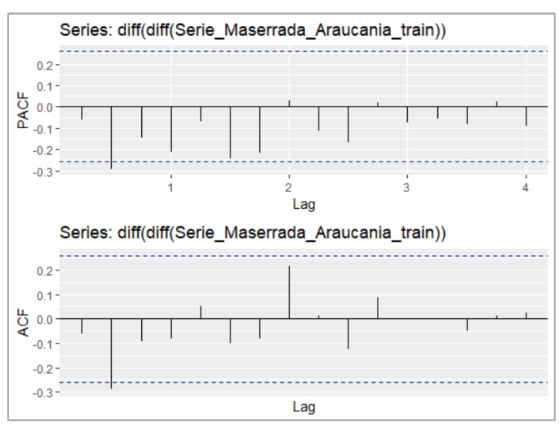


Figure 8: *The ACF, and PACF of differenced time series of sawn wood prices in La Araucanía region.*

From the models raised for the series of prices of radiata pine sawn wood in La Araucanía region (Table 9), the one with the best results in validation metrics fitting is the M_5 Model, at the same it has no significant lower indicators in train metrics, and its residuals are associated to white noise according to the Ljung-Box test.

Table 9: *Fitted models, and performance metrics of time series of sawn wood prices in La Araucanía region.*

ARIMA Models		AIC	BIC	Train		Test	
				RMSE	MAE	RMSE	MAE
M_1	(2,2,1)(2,0,0)[4]	**-224.73**	**-212.47**	**0.0289**	0.0181	0.1522	0.1374
M_2	(1,2,2)	-219.92	-211.75	0.0319	0.0170	0.0756	0.0636
M_3	(2,2,1)	-220.40	-212.23	0.0318	**0.0166**	**0.0679**	0.0562
M_4	(2,2,3)	-216.60	-204.35	0.0314	0.0177	0.0757	0.0642
M_5	(3,2,2)	-217.31	-205.05	0.0312	0.0177	**0.0679**	**0.0557**

When assessing the percentage differences between values provided by the model, and the ones used in validation (Table 10), it is observed that the difference in respect of the first two quarters is higher than 7%, while from this point it starts to decrease, reaching even 0.22%.

Table 10: *Variation of sawn wood prices in La Araucanía region estimated with the M_2 model, and the observed values.*

Year	Quarter	Observed ($/m3)	Forecast ($/m3)	Variation (%)
2021	Q4	133.645	124.421	6,90%
2022	Q1	146.308	129.809	11,28%
2022	Q2	142.267	136.781	3,86%
2022	Q3	144.423	144.110	0,22%
2022	Q4	144.423	151.212	-4,70%

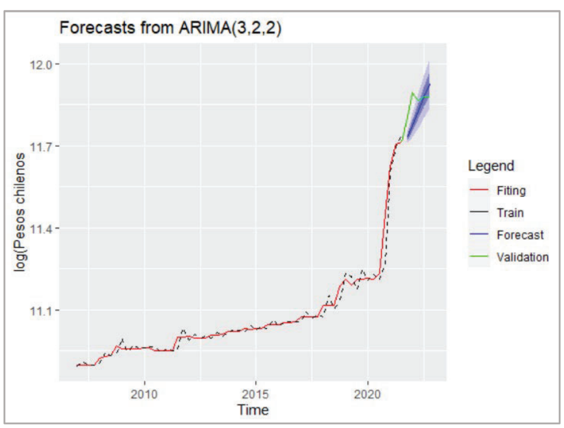


Figure 9: *Logarithmic time series of radiata pine sawn wood prices in La Araucanía region, forecasts, and fitting carried out with the M_2 model, and the observed values.*

Due to the lack of information on different periods, it was not possible to build a time series for the sawn wood prices in Los Lagos region with the necessary consistency to properly apply the methodology presented in this study.

4 CONCLUSIONS

The information collected, and managed by INFOR has an important potential that goes beyond of being a historical record of prices of several forestry products; it also can be used to generate new management tools allowing to promote construction with wood in Chile.

The models fitted for radiata pine sawlogs forecasted in a good way the prices of this product. In overall terms, the variations obtained were lower than 8%, which was validated with five quarter measures. The best results were obtained in La Araucanía region, where variations did not exceed 6%.

In case of time series of sawn wood prices, the models presented differences of up to 11.28%. In La Araucanía region, there were major differences between the last quarter of 2021, and the first quarter of 2022, which is due to unusual price increase of this material during the COVID 19 pandemic.

If the result obtained in this study are compared with other authors, it is observed that results are similar to Broz, and Viego [10], who obtained percentage differences lower than 10% by estimating the values of prices of forestry products in a 9-months period.

The ARIMA models are a good tool to carry out short, and medium-term projections of the forestry product prices, with a low level of uncertainty. Nevertheless, they cannot work efficiently with sudden market fluctuations, as the ones derived from the government policies for economic support in time of a pandemic.

ACKNOWLEDGEMENT

The present research have been carry out under the framework of the project entitled: *"Fortalecimiento de las Capacidades Tecnológicas del Instituto Forestal (INFOR), para el desarrollo de la Industria Secundaria de la Madera, a través de bienes públicos, orientados al sector de la construcción"*. The project was funded by *la Corporación de Fomento a la Producción* (CORFO).

REFERENCES

[1] Alvarez González, V., Poblete Hernández, P., Soto Aguirre, D., et al. La Industria del Aserrío 2021. Instituto Forestal, Santiago, Chile, 2021

[2] Gysling Caselli, J., Kahler González, C., Soto Aguirre, D., et al. El Mercado de la Madera Aserrada para uso estructural en Chile. Instituto Forestal, Santiago, Chile, 2020.

[3] Alvarez González, V., Poblete Hernández, P., Soto Aguirre, D., et al. Anuario forestal 2022. Santiago, Chile. Instituto Forrestal (INFOR). 276p. 2022.

[4] Lind, D., Marchal, W. y Wathen, S. Estadística aplicada a los negocios y la economía. McGraw-Hill, DF, México, 2015.

https://doi.org/10.52202/069179-0583

[5] Chatfield, C. The Analysis of Time Series, An Introduction. 352p. Chapman and Hall/CRC, United States of America, 2003.

[6] Guerrero, V. Análisis estadístico de series de tiempo económicas. 395p Thomson, D.F, México, 2003.

[7] Bowerman, N. Pronósticos, series de tiempo y regresión: un enfoque aplicado. 720p. International Thomson Editores D.F., México, 2009.

[8] Amarís, G.; Ávila, H. & Guerrero, T. Aplicación de modelo ARIMA para el análisis de series de volúmenes anuales en el río Magdalena. *Revista Tecnura*.2017.2. a07, 2017

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

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

[11] Broz, D. & Viego, V. Predicción de precios de productos de *Pinnus spp.* con modelos ARIMA. *Madera y Bosques* 20 (1): 37-46, 2014.

[12] Mei, B., Clutter, M. & Harris, T. 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, 2010.

[13] Box, G. & Jenkins, G. Time Series Analysis: Forecasting and Control. Series in Probability and Statistics. Holden-Day, U.S.A, 1970

[14] Dickey, D. & Fuller, W. Distribution of the Estimators for Autoregressive Time Series with a unit root. *Journal of the American Statistical Association* 74(366):427-431, 1979

[15] Ljung, G. M., & Box, G. E. On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303, 1978

[16] Akaike, H. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723, 1974

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

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

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

[20] Hyndman, R.J., y Khandakar, Y. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1-22, 2008

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

[22] Aguirre, B. gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3, 2017