Revista

El Periplo Sustentable
Universidad Autónoma del
Estado de México
http://rperiplo.uaemex.mx/
ISSN: 1870-9036
Publicación Semestral
Número: 41
Julio / Diciembre 2021

Artículo

Título
Forecasting tourism demand in Chile: Regional analysis using the Seasonal Autoregressive Model

Autores:
Cristian Mauricio Mondaca-Marino
Ailin Arriagada Millaman
Pedro V. Piffaut

Fecha Recepción:
23/08/2019
Fecha Reenvío:
05/08/2020
Fecha Aceptación:
10/06/2021

Páginas:
234 - 254

Forecasting tourism demand in Chile: Regional analysis using the Seasonal Autoregressive Model

Pronóstico de la demanda turística en Chile: Análisis regional utilizando un Modelo Autorregresivo Estacional

Abstract

This paper presents Chilean tourism demand describing its behavior for both the country and each of its regions, the analyzed period comprises 2014:01 to 2019:02. The seasonal autoregressive model (SARIMA) process was used to model the series growing dynamics. Results show that best-fitting models capture nonlinear growth, seasonal patterns, and series volatility, and make it possible to describe not so obvious behaviors, such as the seasonal process order or long-term growth trends. From a public policy point of view, this provides relevant information for decision-makers to manage touristic services and infrastructure in a better way. Regional and countries' forecasted demand presents a low error percentage, less than 2%, though in some regions this value is underestimated overestimated in others.

Keywords:
Box-Jenkins methodology, demand forecast, short-term estimate, tourism demand, VAR analysis, Granger causality test.

Resumen

Conocer el comportamiento de la demanda turística permite a los planificadores de política pública tomar mejores decisiones respecto a cómo administrar los servicios turísticos y priorizar las diferentes inversiones e intervenciones en los territorios. El presente trabajo aporta a la comprensión de este comportamiento al describir la demanda turística de Chile, tanto a nivel país como para cada una de sus regiones durante el período 2014:01 a 2019:02, aplicando la metodología SARIMA (Modelo autorregresivo estacional de media móvil) para modelar la dinámica de crecimiento de la demanda en cada caso. Los resultados permiten identificar que aquellos modelos mejor ajustados para cada región y el país capturan los crecimientos no lineales, patrones estacionales y volatilidades de cada serie, permitiendo describir conductas no tan evidentes como el orden del proceso estacional, o tendencias de crecimiento de largo plazo. Las proyecciones de demanda regionales y del país presentan un bajo porcentaje de error, menor al 2%, el cual se encuentra subestimado en ciertos casos y sobreestimado en otros.

Keywords:
Metodología Box-Jenkins, pronóstico de la demanda, estimación de corto plazo, demanda turística, análisis VAR, test de causalidad de Granger.

The
AUTHORS

Cristian Mauricio Mondaca-Marino

Doctor en Economía.
Universidad Austral de Chile, Chile.
ORCID:
https://orcid.org/0000 -0003 -0606 -6249.

cristianmondaca@uach.cl

Ailin Arriagada Millaman

Universidad Austral de Chile, Chile.

Pedro V. Piffaut

Ph. D. Economics.
Columbia University,
New York, USA.

Introduction

Tourism has become a very important activity for achieving sustainable development in many countries around the world, showing a very important growth in tourist arrivals in recent years. Chile has not been an exception to this phenomena and tourism has become a matter of interest in every region, especially in consolidated destinations in the north, center, and south of the country. An issue of great relevance in the study of tourism is the ability to forecast demand behavior especially knowing the number of people that may arrive at a destination. It is crucial for touristic planning being able to forecast the demand, especially when its behavior presents a seasonal component, during which the visitation falls in the low season and presents an important increase in visitation during high season.

In consequence, forecasting can help policymakers define the necessary arrangements in terms of the capacity of tourism load, and the possibility of increasing a destiny's infrastructure its services and accommodations in order to respond to future increases and decreases in touristic visitation. Given the previews arguments, it is manifest that the study of the dynamics of growth and the many determinants that might influence tourism demand has captured the interest of researchers, who have had the interest of modeling the phenomena and forecasting demand behavior in the short, medium and long term in different countries in the world (Song & Li, 2008; Song, Qiu & Park, 2019; Song & Witt, 2011).

This research will contribute to literature by giving new insights about touristic demand behavior in Chile at a regional level, as well as bringing forward a methodology that might be used for forecasting its behavior in the short term. The main contribution of the article is modeling and forecasting tourism demand for the country and its regions using a SARIMA autoregressive seasonal model.

The paper is organized as follows: It starts with a brief introduction and literature review about the subject, the next section describes methodology and data used for assessing the best forecasting model, followed by the estimation and report of the main results and a discussion section. The final section presents a summary of the main findings and concludes with recommendations for public policies for the industry.

Previous studies in tourism demand

Using models to forecast tourism demand is a topic that attracts great interest (Goh & Law, 2011) due to the rising in importance of the tourism activity and because it is useful for public and private planning decisions both in terms of resources allocation and infrastructure investments (Chatziantoniou, Degiannakis, Eeckels & Filis, 2016).

Research in forecasting tourism demand can be arranged into two groups, the first one uses methodologies that study the number of tourists (visitors to a country or territory) in relation to macroeconomic factors such as GDP, inflation, currency value, etc. (Peng, Song & Crouch, 2014; Song, Qiu & Park, 2019), this researchers looks for a causal association able to explain the variations that tourism demand might present (Dogru, Sirakaya-Turk & Crouch, 2017; Martins, Gan & Ferreira-Lopes, 2017; Morley, Rosselló & Santana-Gallego, 2014). In this case, chosen methodologies are the aggregate production function, structural, gravitational and vector autoregression (VAR) models and the biggest dilemma arises from defining which factors are endogenous or exogenous to the demand or causality degree between variables (Song & Li, 2008).

The second group uses modeling tools to shape the intrinsic behavior of the series, this approach assumes that touristic demand strongly responds to its past behavior, specially in the short term analysis (Hassani, Silva, Antonakakis, Filis & Gupta, 2017; Kulendran & Witt, 2003; Lim & McAleer, 2001, 2002). There are a variety of tools to choose in this methodology (Song & Witt, 2011); nevertheless, one of the most popular methods (Goh & Law, 2011) is the one proposed by Box, Jenkins & MacGregor (1974), their approach considers both, the estimated present values and linear combinations of past values, and the presence of random stochastic and seasonal components, this is called Seasonal Autoregressive Integrated Moving Average (SARIMA) process.

The greatest advantage of SARIMA processes is its capacity to model time series with trends, seasonal patterns and short-time correlation using a small amount of parameters and demanding little assumptions about the analyzed variable conduct. Some examples may be found using this methodology to forecast tourism demand behavior such as Brida & Garrido (2011); Chu (2008); Du Preez & Witt (2003); Goh & Law (2002); Gustavsson & Nordström (2001); Hassani *et al.* (2017); Kim *et al*. (2011), you could also find an interesting review of this subject in Goh & Law (2011).

Methodology

This research main objective is to provide a short-term forecast for tourism demand in Chile and its regions using the seasonal autoregressive model (SARIMA), followed by a vector autoregressive analysis (henceforth, VAR) and the Granger test to determine plausible causality relationships on the demand for tourism between different regions throughout Chile. The arrivals monthly data was provided by the Subsecretaria de Turismo de Chile (2020) for the period 2014:01 to 2019:02, this range corresponds to the available data for the territory.

The forecasting time series methodology was developed by Box, Jenkins y Macgregor (1974). It is based on a procedure that relates present values of a series, with its past values, known as Autoregressive Integrated Moving Average process (ARIMA), a method designed to model time series with stochastic correlated components, with constant mean and variance, despite its seasonal patterns. This is a well-used technique for forecasting short term future values. The model is defined as:

$$\left(1-\phi_1 B-\ldots-\phi_p B^p\right)\left(1-B\right)^d y_t=\left(1-\theta_1 B-\ldots-\theta_q B^q\right)e_t \tag{1}$$

Equation (1) is usually referred as to ARIMA(p,d,q) model, where p, d, and q are natural numbers and refer to the autoregressive process order, differences applied to the series to become stationary, and the moving average process order considered for modeling (Box, Jenkins, Reinsel & Ljung, 2015). In general, ARIMA processes are good for short term forecasting in time series, independent of its data periodicity (annual, quarterly, monthly, hourly, etc.).

Seasonality occurs when a regular behavior pattern is repeated every S interval, S is the number of periods that must go forward until the pattern is re-iterated. In this study, S is valued at 12, since the series present peaks in the same month every year. The study should, therefore, incorporate to the time series analysis, a seasonal pattern (S), short term correlations (AR) and trends (MA). Consequently, the methodology applied for the analysis corresponds to the SARIMA process, that incorporates seasonal and non-seasonal determinants and may be articulated as: SARIMA (p, d, q) (P, D, Q), where P is the autoregressive order, D the number of differences applied to the series, and Q the moving average order for the seasonal elements, and can be represented in the following:

$$\phi_p(L)\Phi_P(1-L)^d(1-L^s)^D y_t = \delta + \theta_q(L)\Theta_Q(L^s)e_t \tag{2}$$

Where:

$\phi_P(L) = (1-\phi_1 L - \ldots - \phi_P L^P)$, is the non-seasonal autoregressive process order

$\Phi_P(L^s) = (1-\Phi_1 L - \ldots - \Phi_P L^s)$, is the seasonal autoregressive process order

$\theta_q(L) = (1-\theta_1 L - \ldots - \theta_q L^q)$, is the non-seasonal moving average process order

$\Theta_Q(L^s) = (1-\Theta_1 L - \ldots - \Theta_Q L^s)$, is the seasonal moving average process order

$(1-L)^d$, is the non-stationary difference order *d*

$(1-L^S)^D$ is the seasonal difference order D

e_t, is the disturbance or error term $\sim \text{iid}(0, \sigma^2)$

L is the autoregressive operator and ? the constant. This model brings together the elements linking the seasonal component to its autoregressive or moving average side.

The estimation process consists of an iterative multi-stage procedure: the augmented Dickey-Fuller test is applied to determine the stationarity of the series variance and mean, or if it is stationary on its differences (*d*); secondly the model is described, p and q orders are identified through the study of its autocorrelations, defining a range of possible model(s) for which the actual parameters will be estimated; the model(s) is(are) validated by means of statistic diagnosis, such as the Akaike Information Criterion (AIC), and the best fitting model is chosen; the model is then used to forecast short term tourism demand behavior.

This study also analyzes the models forecasting performance, defining an error function using actual data from the series last six months and minimizing the projected error. The statistics criterion, in this case, are (Goh & Law, 2002): Theil´s U, Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Root Mean Squared Percentage Error (RMSPE), Root Mean Square Error, RMSE) and Mean Absolute Deviation (MAD).

Once the forecast estimation has been concluded, a VAR model is applied to estimate the possible interrelationships between the forecasted demands for tourism series above-described. The rationale behind this is that some regions may influence or drives tourism demand to other regions. The idea is to explore the existence of this type of relationship between one or more regions by the means of a VAR model following by a Granger causality test.

Consider a model with *p* lags:

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t \qquad (3)$$

Where y_t is a vector of variables Kx1, v is a vector of parameters Kx1, A_1-A_p are matrices of parameters KxK and ε_t is a vector of perturbations or errors with mean 0 and covariance matrix Σ, being also an independent and identically normally distributed random variable. Thus, a VAR is a model in which K variables are specified as linear functions of p of their own lags, p lags of the other K-1 variables, and possibly additional exogenous variables. More rigorous treatments can be found in Hamilton (1994), Lütkepohl (2005) and Amisano and Giannini (1997). Stock and Watson (2001) provide an excellent nonmathematical treatment of vector autoregressions. Becketti (2013) introduces VAR analysis with an emphasis on how it is done in practice.

Results

The first interesting result comes from the observation of tourists' arrivals to Chile and its regions for the 2014-2019 period (figure 1). From the graphical analysis, one may notice that most of the country's regions show a seasonal pattern in its tourism demand, with an exception in the XV, I, II and RM regions where no seasonality can be identified. The aggregate country's tourism demand shows a seasonal behavior as well.

In table 1, a summary of basic statistics is presented. This data gives a better understanding of the tourism demand for each region and the country. The statistical analysis shows that tourism demand tends to be heterogeneous in terms of its size, volatility and the strength of its previously identified seasonal pattern.

Table 1. Tourism demand statistic summary for Chile and its regions (2014:01-2019:02)

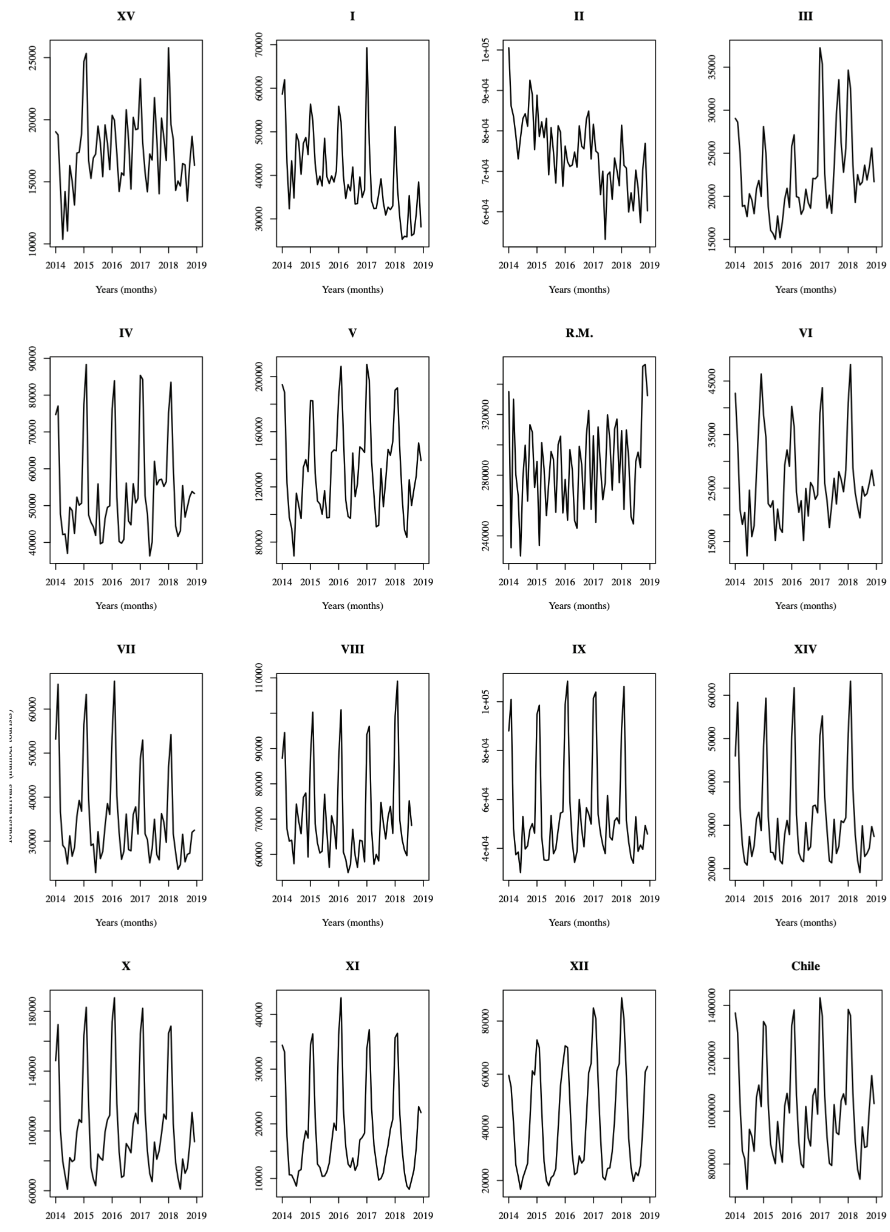
N	Region	Share	Mean	S.D.	Median	Min	Max	Skw	Kur	ADF
XV	Arica y Parinacota	1.7	17,503.2	3,117.1	17,264.7	10,374.4	25,792.0	0.48	0.54	-4.1192***
I	Tarapacá	4.0	40,220.7	9,289.5	38,301.6	25,332.3	69,288.2	0.87	0.56	-4.8443***
II	Antofagasta	7.5	75,399.8	8,637.7	74,950.2	53,194.5	100,542.0	0.14	0.34	-3.2459* .
III	Atacama	2.2	22,378.2	5,222.4	20,909.0	15,009.1	37,255.7	1.09	0.56	-3.9508***
IV	Coquimbo	5.3	53,942.8	14,021.3	50,029.9	36,342.0	88,323.9	1.08	0.04	-5.1949***
V	Valparaíso	13.0	132,545.5	35,051.5	127,471.6	69,949.3	208,770.9	0.55	-0.66	-5.0790***
R.M.	Metropolitana	29.0	282,460.8	26,433.2	285,071.3	226,786.8	335,077.0	-0.14	-0.93	-3.6296** .
VI	O'Higgins	2.6	26,416.0	8435.9	24,282.8	12,323.6	48,097.1	0.84	-0.08	-4.4988***
VII	Maule	3.4	35,426.2	11,014.1	31,697.6	22,864.6	66,333.4	1.33	0.88	-4.7716***
VIII	Bio-bío	6.4	70,400.4	13,427.9	66,595.6	54.843.0	109,081.4	1.21	0.48	-5.0683***
IX	La Araucanía	5.3	55,060.7	22,068.8	47,976.4	30,085.9	108,396.3	1.30	0.33	-5.0282***
XIV	Los Ríos	3.1	32,162.7	11,668.1	29,350.4	19,070.0	63,254.7	1.26	0.54	-4.8644***
X	Los Lagos	10.0	103,158.8	35,595.3	92,055.8	60,995.4	189,037.8	1.07	-0.01	-5.0679***
XI	Aysén	1.8	18,240.4	9,280.2	15,976.3	8,041.6	43,046.8	1.13	0.04	-5.2107***
XII	Magallanes	4.1	42,102.2	20,917.1	36,411.1	16,685.5	88,747.6	0.54	-1.07	-5.8702***
	Chile	100.0	1,006,128.7	192,597.1	991,362.3	704,542.4	1,429,187.0	0.75	-0.47	-5.3514***

***, **, * indicates significance at 1%, 5% and 10% respectively.

Source: Own elaboration based on Subsecretaria de Turismo de Chile (2020).

In terms of its share, the most relevant regions are RM, V, X that together, sum up to the 52% tourists' arrivals to the country; these regions may be defined as consolidated destinations. The non-seasonal pattern is strong in XV, I, II, RM, these regions also show clear volatility in its tourism demand.

Figure 1. Tourist arrivals to Chile and its regions (2014:01-2019:02) (Thousand)



Source: Own elaboration based on Subsecretaria de Turismo de Chile (2020).

A seasonal component is present in the regions III, IV, V, VI, VII, VIII, IX, XIV, X, XI and XII indicating a pattern in its significant oscillation for tourism demand over the year. Tourists flow shows a clear cyclic behavior especially in the southern part of the country, and in the northern central area. Because of its great increase or decrease in the number of tourist arrivals, the IV, V, VII, IX, X and XII stand out from within this group. In these regions, tourism demand shows a great variance depending on the analyzed period.

Tourism demand models, for each region, are expected to present different processes due to the important differences in the series behavior, in terms of its seasonal, and volatility/fluctuation patterns. Moreover, because of the growing trend and seasonal pattern, short-term forecasting models are expected to be non-stationary in its mean and variance.

In order to identify the parameters of each region best fitting forecasting model, mean and variance stationarity was tested, the Augmented Dickey-Fuller (ADF) approach was applied to each region data series. The data tested correspond to logarithm (LOG) returns of the series and first difference of LOG. It is found that, for series, its first difference's mean and variance are stationary (table 1).

To continue with the best SARIMA model identification, the Akaike and Bayesian information criterion were applied (AIC and BCI), the Mean Square Error (MSE) was also evaluated. In all cases, the coefficients are significant with a p-value <0.01. The error's independent and identically distributed (i.i.d) assumption was also tested for the identified models; there was no evidence found against the hypothesis of autocorrelation absence in the error term, neither against the normal distribution of disturbances at a 5% significance. Box-Pierce (Ljung-Box) Modified Chi-Square and the residuals autocorrelation functions were analyzed to get to this conclusion.

The models were also tested to assess its forecasting accuracy using data from the last six months (2018:09-2019:02) and defining a loss function to study its mean percentage error (MPE). Different statistics criteria were applied to the forecasted errors: Theil´s U, Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Root Mean Squared Percentage Error (RMSPE), Root Mean Square Error, RMSE), and Mean Absolute Deviation (MAD), its results are shown in table 2.

Table 2. SARIMA models forecasting accuracy for Chile and its regions tourism demand (2018:09-2019:02)

Tourism Demand	SARIMA Model	EPM	U Theil	MAPE	MSE	RMSPE	RMSE	MAD
XV	SARIMA(2,0,0)(1,1,0)$_{12}$	-0.354	0.468	0.009	0.009	0.030	0.095	0.088
I	SARIMA(0,1,1)(0,1,0)$_{12}$	1.988	1.026	0.020	0.077	0.087	0.278	0.212
II	SARIMA(1,1,0)(1,1,0)$_{12}$	-0.256	0.323	0.006	0.006	0.023	0.077	0.068
III	SARIMA(2,1,1)(0,1,1)$_{12}$	-1.367	1.115	0.015	0.027	0.051	0.164	0.149
IV	SARIMA(1,0,0)(0,1,1)$_{12}$	-0.505	0.561	0.006	0.006	0.023	0.078	0.066
V	SARIMA(0,0,1)(0,1,1)$_{12}$	-0.775	0.964	0.008	0.014	0.034	0.119	0.094
R.M.	SARIMA(0,0,0)(1,1,0)$_{12}$	1.063	1.263	0.011	0.020	0.040	0.143	0.135
VI	SARIMA(0,0,0)(0,1,0)$_{12}$	-0.128	0.353	0.008	0.010	0.032	0.102	0.084
VII	SARIMA(2,0,0)(0,1,0)$_{12}$	-0.263	0.723	0.010	0.015	0.038	0.124	0.099
VIII	SARIMA(2,0,1)(0,1,1)$_{12}$	-0.004	0.182	0.002	0.001	0.010	0.032	0.026
IX	SARIMA(1,0,1)(1,1,0)$_{12}$	-0.927	0.878	0.012	0.030	0.052	0.173	0.134
XIV	SARIMA(0,0,1)(0,1,1)$_{12}$	-1.249	0.622	0.012	0.022	0.046	0.148	0.129
X	SARIMA(2,0,0)(0,1,0)$_{12}$	-0.368	0.385	0.006	0.007	0.025	0.086	0.071
XI	SARIMA(1,0,0)(0,1,0)$_{12}$	0.537	0.333	0.009	0.013	0.037	0.115	0.094
XII	SARIMA(1,0,0)(0,1,1)$_{12}$	-0.557	0.195	0.006	0.004	0.019	0.062	0.060
Chile	SARIMA(1,0,0)(0,1,1)$_{12}$	-0.007	0.265	0.002	0.001	0.009	0.035	0.029

Source: Own elaboration based on Subsecretaria de Turismo de Chile (2020).

All models show a low mean error, generally under ±2% ranging, in most cases, below ±0.8% (see MPE indicator in table 2). It should be noticed, however, that all estimates use transformed LOG variables, so its error values are underestimated when studying the actual series. Forecasted series for the 2018:09-2019:02 period is overestimated in 18.7% of the cases (VX, I, R.M, VII, VIII, XI regions) and underestimated in 81.3% of the regions (II, III, IV, V, VI, IX, XIV, X, XII, and the country). The best-fitting model for each region is presented in table 3, additionally, the actual and simulated series may be observed in figure 2.

In every region´s SARIMA model, a non-stationary time-related variance was identified; this component required a logarithmic and seasonal difference transformation in order to have constant variance over time. The mean, however, did not present a similar behavior in every region, the XV, I, III, and R.M did not require a transformation in order to have a single value when shifted in time. As presented in figure 2, adjusted SARIMA models for each region and the country were capable of satisfactory replicate tourism demand behavior. However, regions that show high volatility and a strong seasonal component, XV, I, II, III, and RM present a greater difference between the real and forecasted values for the series in the analyzed period (2014:01-2019:02).

Table 3. SARIMA models for Chile and its regions tourism demand (2014:01-2019:02)

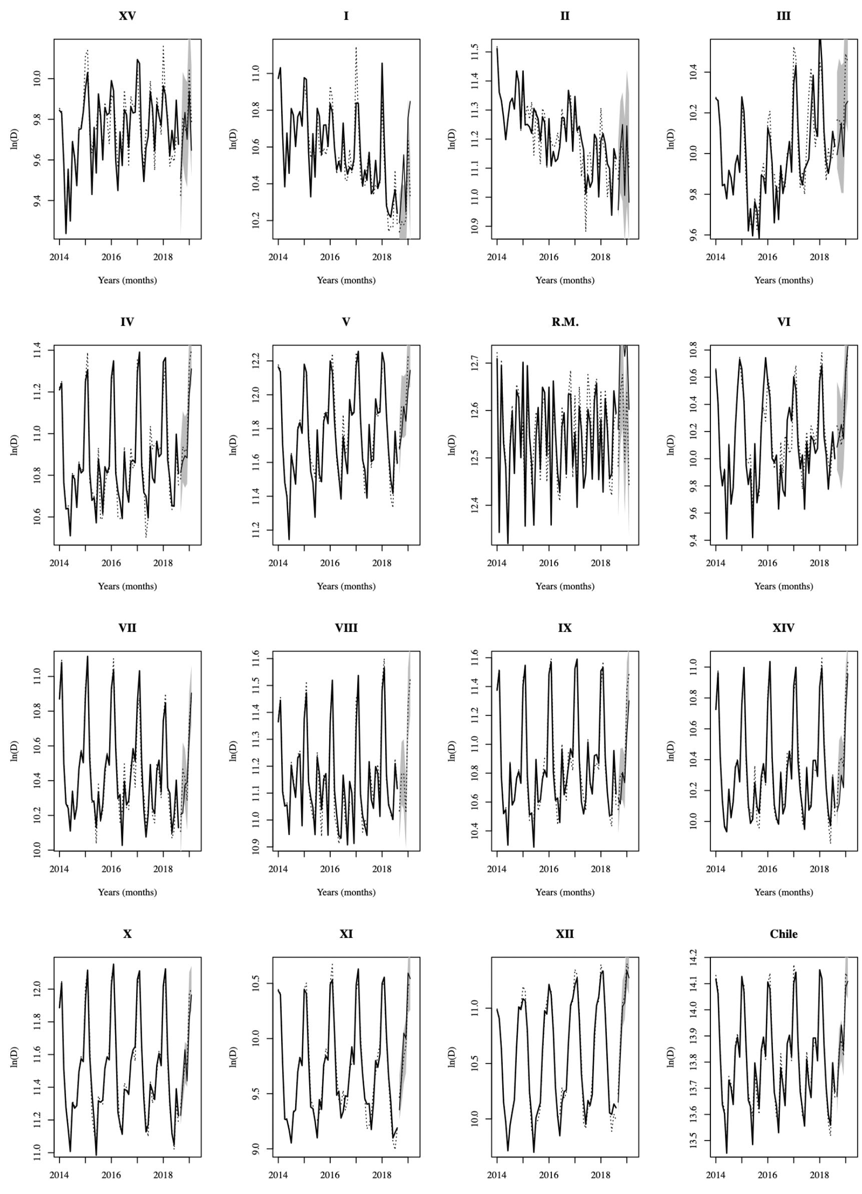
Tourism Demand	SARIMA model	Coefficient					AIC	BIC
		AR	**MA**	**AR_s**	**MA_s**	**δ**		
XV	SARIMA$(2,0,0)(1,1,0)_{12}$	0.4450 (0.1409) 0.3348 (0.1494)		-0.5263 (0.1458)			-65.34	-58.20
I	SARIMA$(0,1,1)(0,1,0)_{12}$		-0.7933 (0.10)				-68.76	-65.24
II	SARIMA$(1,1,0)(1,1,0)_{12}$	-0.3302 (0.1472)		-0.6309 (0.1231)			-100.82	-95.53
III	SARIMA$(2,1,1)(0,1,1)_{12}$	0.4451 (0.2203) -0.2872 (0.1628)	-0.6635 (0.1956)		-0.6222 (0.2919)		-48.83	-40.03
IV	SARIMA$(1,0,0)(0,1,1)_{12}$	0.2599 (0.1451)			-0.8175 (0.5186)	0.0022 (0.0008)	-80.49	-73.35
V	SARIMA$(0,0,1)(0,1,1)_{12}$		0.4792 (0.1427)		-0.5944 (0.3105)	0.0019 (0.0010)	-83.6	-76.47
R.M.	SARIMA$(0,0,0)(1,1,0)_{12}$			-0.5062 (0.1509)			-122.9	-119.4
VI	SARIMA$(0,0,0)(0,1,0)_{1}$						-31.5	-29.7
VII	SARIMA$(2,0,0)(0,1,0)_{12}$	0.2978 (0.1422) 0.3004 (0.1440)					-97.74	-92.39
VIII	SARIMA$(2,0,1)(0,1,1)_{12}$	-0.1345 (0.1664) 0.6467 (0.1265)	0.6465 (0.1265)		-0.8000 (0.6759)		-120.74	-111.82

Tourism Demand	SARIMA model	Coefficient					AIC	BIC
		AR	MA	AR_s	MA_s	δ		
IX	$SARIMA(1,0,1)(1,1,0)_{12}$	0.9302 (0.0746)	-0.6312 (0.1392)	-0.4406 (0.1376)			-83.06	-75.92
XIV	$SARIMA(0,0,1)(0,1,1)_{12}$		0.5495 (0.1401)		-0.7749 (0.4196)	0.0012 (0.0007)	-103.59	-96.45
X	$SARIMA(2,0,0)(0,1,0)_{12}$	0.3657 (0.1397) 0.3885 (0.1480)					-128.2	-122.85
XI	$SARIMA(1,0,0)(0,1,0)_{12}$	0.5534 (0.1237)					-70.77	-67.2
XII	$SARIMA(1,0,0)(0,1,1)_{12}$	0.6118 (0.1289)			-0.4305 (0.2382)	0.0040 (0.0019)	-82.79	-75.65
Chile	$SARIMA(1,0,0)(0,1,1)_{12}$	0.4248 (0.1379)			-0.4848 (0.2830)		-161.6	-156.25

Obs: Significant model parameters present p-value <0.01. Source: Author's computations.

Source: Own elaboration based on Subsecretaria de Turismo de Chile (2020).

Figure 2. Original series, selected SARIMA model series (2014:01-2018:08) and SARIMA forecasts (2018:09-2019: 02) (Thousand).



Obs: (__) original series, (...) SARIMA model series, gray area: 95% confidence interval for simulated data.

Source: Own elaboration based on Subsecretaria de Turismo de Chile (2020).

Results show that diverse models are a consequence of the series of different identified processes and are accountable for the dynamics and seasonal patterns in every region. III, IV, VII, VIII, X, XI, XII, and the country are strongly dependent on recent past values therefore, its tourism demand is mostly influenced by the AR process. The seasonal pattern in the I and V regions is captured by the previous period volatility (MA component), while in the XIV the seasonal pattern relies on its previous period parameters (MA seasonal).

There is another group constituted by XV, II, and IX regions, these demands present high volatility regarding its recent past values and are, almost all, attenuated by its previous season variance, with the exception of the R.M. region where the previous season tourism levels it is unique.

There are some demands with a singular behavior; in region I, tourism demand depends only on its recent volatility and presents a systematic non-stationary seasonal pattern which may be eliminated by a differencing transformation process; R.M´s demand strongly relies on its previous season values; and VI tourism demand may be described by a random walk, difference-stationary model. There are only four models that present a constant drift, that is to say, they follow a systematic trend, these are the IV, V, XIV and XII demand, show an increasing trend. Overall, considering demand parameters and trends, each region tourism behavior may be classified into three relevant categories: growing, stable and declining visitation. III, IV, VI, VIII, XIV, XI, XII and national are blooming demands, R.M, X and IX are stable and XV, I, II, V and VII experience a decline in its demand.

Results are especially important for decision-makers since they may define touristic strategies and policies based on each region´s demand dynamics. Long term measures (from one season to the next) will be successful when a destination demand is associated with its previous season arrivals, this gives a certain time for intervention planning. In other cases, actions must be ready to be implemented in the short term, as a response to increases and decreases in tourist arrivals to a certain destination.

The IV, V, XIV and XII demand are special because of its particular behavior, these show a seasonal pattern moderated by a systematic long-term growth along with a trend. In these cases, intervention planning may account for long-run future results due to its growing/declining dynamics.

The results presented above are similar to those obtained by Brida & Garrido (2011) who describe the behavior of tourism demand in Chilean regions for the period 2004-2009; they determined that there were two identifiable behaviors in the time series of monthly tourist arrivals to the regions; for example, some presented a growth trend (Antofagasta), and others a very marked seasonal component (O'Higgins) in the summer season (January and February). However, when contrasting the behavior of the regions at the beginning of 2000 in comparison with the late 2010, there are some differences; the region of Antofagasta while presenting an increasing trend in the first study, shows now a decreasing trend; Magallanes that had an incipient growing trend, also shows a stable growth trend in the past years; and the Metropolitan region that presented a stationary behavior, currently shows greater volatility and a growing trend.

Reviewing the methodology there is quite a similarity between both studies, especially regarding to the necessary differentiation used to model the series, coinciding the normal and seasonal differentiations required in the former and present analysis. On the other hand, the models present disparities regarding to the orders in its autoregressive and moving average components, which might be explained because of differences in the degrees of volatility and stochastic trends present in the analyzed series. The present period shows a lower seasonality in comparison to the beginning of 2000, and in most regions a growing trend, a sign of a more stable tourism demand over time.

Finally, it is interesting to determine the possible effect of some specific regions on other region's tourism demand. The approach to this question is an Autoregressive Vector (VAR) model. Prior to the analysis, it was determined that the optimum number of lags for the proposed VAR models is one lag based on the usual information criteria (AIC, BIC and HQIC), which is highly expected because the region effect on other region demand for tourism should have low persistence, in other words, the demand effect will not last beyond the year.

Based on the results of VAR models, it is concluded that there is Granger causality relation between the demand for tourism from some specific regions to the others, and these relationships are statistically significant at the 5% level. For instance, regions V, VIII and XIV are the most influential regions meaning that the affluence of tourists to these three specific regions has spillover effects on other regions, increasing the demand for tourism in these, perhaps, less attractive locations.

Even though the great tourist's affluence comes from the Metropolitan Region (RM), this is due to the location of the international airport in the capital, Santiago de Chile, which is practically the only entrance from tourists coming from all over the world. In this regard and according to the Granger test, region V Granger causes tourist demand in regions RM, VIII, IX, X, XI and XIV. By the same token, region VIII Granger causes tourist demand in regions V, IX, X, XI and XIV while region XIV Granger causes tourist demand in regions V, VIII, IX, X and XI, and so on. Table 4 summarizes these Granger causality relationships.

Table 4. Granger causality relationships from VAR model

Region	Granger Causality	χ^2 Statistic	Significance Level
V	RM	7.3255	0.007
	VIII	4.3209	0.038
	IX	10.464	0.001
	X	13.318	0.000
	XI	15.653	0.000
	XIV	7.3851	0.007
VIII	V	6.6053	0.010
	IX	7.056	0.008
	X	6.7463	0.009
	XI	5.3161	0.021
	XIV	8.2151	0.004
XIV	V	10.402	0.001
	VIII	6.4542	0.011
	IX	11.168	0.001
	X	13.507	0.000
	XI	16.242	0.000

Region	Granger Causality	χ^2 Statistic	Significance Level
RM	V	6.3384	0.012
	VIII	5.7518	0.016
	IX	6.0987	0.014
XI	V	4.1427	0.042
	IX	5.6208	0.018
	XIV	10.376	0.001

Source: Author's computations from VAR model.

Regarding the identified dependence between regional tourist flows, determined by the VAR models, where regions V, VIII and XIV are the most influential regions in other regions arrivals, this interdependence might be because there are tourist attractions of international interest in these.

Valparaíso (V region) for example was declared by United Nations Educational, Scientific and Cultural Organization (UNESCO) in 2003 as a World Heritage Site, for being "exceptional testimony of the early phase of globalization of the late nineteenth century, when (Valparaíso) became the leading commercial port of the shipping routes of the Pacific coast of South America" (UNESCO, 2020); or the Bio-Bío region (VIII) characterized for its cultural and artistic activities and mixing urban and rural destinations, this region is also a close to Santiago gateway to various leisure and nature destinations such as beaches, snow mountains, hot springs, and forests; and the Los Ríos region (XIV) which stands out for its nature, this region characteristic is its many rivers and lakes, and an extensive network of public and private nature reserves. These are the reasons why the identified regions can be considered as destinations that possibly generate tourist flows to other areas of the country; these serve as natural way stations to visit the diverse sights the country offers, and facilitate mobility in the extensive geography of Chile.

Conclusions

In this paper, Chilean tourism demand was modeled both for the whole country and independently for its regions, available data allowed for the study to comprise the period 2014:01-2019:02, seasonal autoregressive modeling was used to understand each demand´s behavior. Results show that this methodology captures nonlinear growth, seasonal patterns, and series volatility. It also makes it possible to describe not so obvious behaviors, such as the seasonal process order or long-term growth trends. Forecasted values show a low error percentage, under ±2%, though some region's error terms are underestimated and for others, these values are overestimated.

Knowing tourism demand is beneficial for public policies planning and strategies because provides relevant information, allowing the decision-makers, tourism related groups, and authorities to manage touristic services and infrastructure in a better way.

Another interesting finding comes from the fact that the most demanded regions have spillover effects on tourist demand in other regions. Regions V, known as the region of Valparaiso-Viña del Mar, region VIII, known as region de La Araucanía, and region XIV, also known as region de Los Ríos, are the most influential regions of the country boosting demand for tourism in other regions of the country.

For future research, this study may be extended to regional demand dependency on tourist's distribution centers such as the V, R.M, X, these regions are given this role due to its weight in national tourism demand (over 50% of tourist arrivals). VAR process might be a good choice for analyzing regional demands relations, and its mutual influence.

It should be noticed that, although its robust results, any destination structural change (infrastructure expansion, new communication nodes, better transport supply) may change the series growing trend, which means that model parameters might change significatively, then actual models would no longer be valid.

Acknowledgment: Authors thank reviewers for their comments and suggestions.
JEL Classifications: C32, C53, R22.

References

Amisano, G. and C. Giannini (1997). *Topics in Structural VAR Econometrics*. 2nd ed. Heidelberg: Springer.

Becketti, S. (2013). *Introduction to Time Series Using Stata*. College Station, TX: Stata Press.

Box, G. E., Jenkins, G., Reinsel, G. C. & Ljung, G. M. (2016). *Time series analysis: forecasting and control* (5th ed.). New Jersey: John Wiley & Sons.

Box, G. E. P., Jenkins, G. M. & MacGregor, J. F. (1974). Some Recent Advances in Forecasting and Control. *Applied Statistics*, *23*(2), 158. https://doi.org/10.2307/2346997.

Brida, J. G. & Garrido, N. (2011). Tourism forecasting using SARIMA models in Chilean regions. *International Journal of Leisure and Tourism Marketing*, *2*(2), 176. https://doi.org/10.1504/ijltm.2011.038888.

Chatziantoniou, I., Degiannakis, S., Eeckels, B. & Filis, G. (2016). Forecasting tourist arrivals using origin country macroeconomics. *Applied Economics*, *48*(27), 2571-2585. https://doi.org/10.1080/00036846.2015.1125434.

Chu, F.-L. (2008). A fractionally integrated autoregressive moving average approach to forecasting tourism demand. *Tourism Management*, *29*(1), 79-88. https://doi.org/10.1016/J.TOURMAN.2007.04.003.

Dogru, T., Sirakaya-Turk, E. & Crouch, G. I. (2017). Remodeling international tourism demand: Old theory and new evidence. *Tourism Management*, 60, 47-55. https://doi.org/10.1016/J.TOURMAN.2016.11.010.

Du Preez, J. & Witt, S. F. (2003). Univariate versus multivariate time series forecasting: an application to international tourism demand. *International Journal of Forecasting*, *19*(3), 435-451. https://doi.org/10.1016/S0169-2070(02)00057-2.

Goh, C. & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, *23*(5), 499-510. https://doi.org/10.1016/S0261-5177(02)00009-2.

Goh, C. & Law, R. (2011). The Methodological Progress of Tourism Demand Forecasting: A Review of Related Literature. *Journal of Travel & Tourism Marketing*, *28*(3), 296-317. https://doi.org/10.1080/10548408.2011.562856.

Gustavsson, P. & Nordström, J. (2001). The Impact of Seasonal Unit Roots and Vector ARMA Modelling on Forecasting Monthly Tourism Flows. *Tourism Economics*, *7*(2), 117-133. https://doi.org/10.5367/000000001101297766.

Hamilton, J. D. (1994). *Time series analysis*. United States of American: Princeton University Press.

Hassani, H., Silva, E. S., Antonakakis, N., Filis, G. & Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, *63*, 112-127. https://doi.org/10.1016/J.ANNALS.2017.01.008.

Kim, J. H., Wong, K., Athanasopoulos, G. & Liu, S. (2011). Beyond point forecasting: Evaluation of alternative prediction intervals for tourist arrivals. *International Journal of Forecasting*, *27*(3), 887-901. https://doi.org/10.1016/J.IJFORECAST.2010.02.014.

Kulendran, N. & Witt, S. F. (2003). Forecasting the Demand for International Business Tourism. *Journal of Travel Research*, *41*(3), 265-271. https://doi.org/10.1177/0047287502239034.

Lim, C. & McAleer, M. (2001). Forecasting tourist arrivals. *Annals of Tourism Research*, *28*(4), 965-977. https://doi.org/10.1016/S0160-7383(01)00006-8.

Lim, C. & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management*, *23*(4), 389-396. https://doi.org/10.1016/S0261-5177(01)00098-X.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Martins, L. F., Gan, Y. & Ferreira-Lopes, A. (2017). An empirical analysis of the influence of macroeconomic determinants on World tourism demand. *Tourism management*, 61, 248-260.

Morley, C., Rosselló, J. & Santana-Gallego, M. (2014). Gravity models for tourism demand: theory and use. *Annals of Tourism Research*, 48, 1-10. https://doi.org/10.1016/j.annals.2014.05.008.

Peng, B., Song, H. & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism Management*, *45*, 181-193. https://doi.org/10.1016/J.TOURMAN.2014.04.005.

Song, H. & Li, G. (2008). Tourism demand modelling and forecasting-A review of recent research. *Tourism Management*, 29(2), 203-220. https://doi.org/10.1016/J.TOURMAN.2007.07.016.

Song, H., Qiu, R. T. R. & Park, J. (2019). A review of research on tourism demand forecasting. *Annals of Tourism Research*, *75*, 338-362. https://doi.org/10.1016/J.ANNALS.2018.12.001.

Song, H. & Witt, S. F. (2011). *Tourism demand modelling and forecasting : modern econometric approaches*. Routledge: Taylor & Francis.

Stock, J. H. & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 15(4), 101-115.

Subsecretaria de Turismo de Chile (2020). Estadísticas. Estadisticas website: http://www.subturismo.gob.cl/documentos/estadisticas/, [June 10, 2020].

UNESCO (United Nations Educational, Scientific and Cultural Organization) (2020). Historic Quarter of the Seaport City of Valparaíso. World Heritage List website: https://whc.unesco.org/en/list/959/, [August 15, 2020].