African Journal of Hospitality, Tourism and Leisure. ISSN: 2223-814X

August 2022, Vol 11, No 4, pp. 1381-1394

Modelling Tourist Arrivals in South Africa To Assess The Impact of the COVID-19 Pandemic on the Tourism Sector

Musara Chipumuro*
Department of Mathematical Statistics and Actuarial Science, University of the Free State, South Africa, Email, chipumuromusara@gmail.com, https://orcid.org/0000-0002-4438-4443

Delson Chikobvu
Department of Mathematical Statistics and Actuarial Science, University of the Free State, South Africa, Email, chikobvu@ufs.ac.za

Corresponding Author

How to cite this article: Chipumuro, M. & Chikobvu, D. (2022). Modelling Tourist Arrivals in South Africa To Assess The Impact of the COVID-19 Pandemic on the Tourism Sector. African Journal of Hospitality, Tourism and Leisure, 11(4):1381-1394. DOI: **https://doi.org/10.46222/ajhtl.19770720.297**

Abstract

The paper examines tourism flows from all foreign countries to South Africa (SA) from 2009 to February 2020 using time series models. The resultant model is used to forecast and assess the impact of COVID-19 on tourist arrivals in South Africa by comparing with actual tourist arrivals after February 2020. Monthly data on tourist arrivals to South Africa from the Overseas, the Southern African Development Commission (SADC), other African countries and those who did not specify area of origin were considered. The Box and Jenkins methodology identified the $\text{ARIMA}(1,1,0)(1,1,0)_{12}$ as the best model as confirmed by the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The model has good forecasting power as evidenced by its Mean Absolute Percentage Error (MAPE). Therefore, the fitted model's forecasts may be used in the absence of COVID-19.The COVID-19 adversely affected forecasting of future values as forecasted values differ greatly with actual observations recorded since COVID-19 induced lockdown measures in March of 2020. This study helps to have an overview of the total contribution being realized from tourism receipts that can be inferred through tourist arrivals before and after the COVID-19 pandemic and how best the tourism sector may be rejuvenated. The study highlights are that the number of tourist arrivals to South Africa follows an $\text{ARIMA}(1,1,0)(1,1,0)_{12}$ model. Prior to the COVID-19 pandemic, there was a persistent upward trend and seasonality in tourist arrivals to South Africa since the 1994 democratic elections. South Africa is currently losing more than 90% of its monthly tourist arrivals because of the pandemic, and shows little sign of an imminent recovery, hence a devastating impact on the tourism industry.

Keywords: International tourist arrivals, South Africa (SA), Box-Jenkins methodology, Forecasting

Introduction

The tourism sector has become a key driver of social and economic development as exemplified through foreign currency earnings, infrastructure and technological development. Analysing the flow of tourist arrivals plays a critical role in driving the need for forecasting and implementation of various appropriate strategies and policies through informative forecasting. Modelling tourist arrivals help ensure sustainable expansion of tourism activities of the host country. Accurate and dependable forecasts of international tourists are therefore of great importance to the planning and or marketing of various tourism products. Industries supplying tourism facilities and or services like accommodation providers and airlines as examples, find tourism flow forecasts as a tool for effective planning and budgetary control. Li and Wu (2019) highlighted on the importance of forecasting tourism demand and pointed that forecasting is a vital tool to tourism business planning and policy formulation, thus, has received increasing attention from industry practitioners.

The tourism industry has proved to be a vital source of economic growth and development for many developing and emerging countries, South Africa included. This is evidenced by the benefits the tourism industry has on the host country as exemplified by creation of jobs and foreign currency earnings. Research on tourism demand is therefore of great importance as it is a key determinant of business profitability (Song & Turner, 2006). Foreign Direct Investment (FDI), foreign currency generation, employment creation, infrastructural development, improved human job security policies and overall economic growth are some of the main benefits the tourism industry has to offer to the host country. Oh (2005) postulate that tourism help create job opportunities, promote notable improvements in a country's infrastructure, technological and foreign currency earnings. Increased globalization has also made the tourism industry one of the largest and fastest growing industries worldwide due to its trickle down effects in addressing, political, cultural, social and economic development. Hence, good planning and administration of tourism activities prove vital in this day and age when the tourism sector is taking center stage in political, cultural, economic and social development. Richardson (2010) postulated that tourism development is really an important tool that may be used in promoting economic growth, poverty alleviation and food security advancement. With this in mind, accurate tourism forecasts help build the good foundation for proper planning and administration, entrepreneurship, investment and policy making.

The South Africa peaceful presidential elections from the apartheid era saw a rise in the number of tourist arrivals who arrived from Overseas, SADC and other African countries and it contributed positively to tourist arrivals and the economy across all divides. The World Tourism and Travel Council (WTTC) report (2010) shows that travel and tourism contributed R189.402 billion (in nominal terms) to the GDP of South Africa in 2009. The support of the South Africa's tourism sector to the economy was projected to be R124.4 billion in the year 2020 from a R91.2 billion in the year 2015 in the South African Tourism's strategic plan and annual performance plan of 2015. From this plan, it was highlighted that a strong growth in tourist arrivals of 15.4% increase was realized in the first 6 months of 2016. The tourism sector has proved to be vast enough to be able to create a massive volume of job vacancies and can contribute to growing the other macroeconomic indicators like the GDP. In 2013, tourism contributed 2.9% (R103 557 million) of South Africa's GDP. The tourism sector employs approximately 1 in 25 of workers in South Africa (4.4% of the total workforce).

South Africa recorded GDP growth due to direct travel in the year 2016 to the extent that the growth outperformed the financial and business sectors, transport, manufacturing, public, retail and distribution sectors (Bhorat et al., 2016). According to Statistics South Africa, tourist exports from overseas increased by14% to 245074 tourists in January 2017 compared with January 2016. The highest increase, 38.1% was for tourists from China (from 84 691 in the year 2015 to 116 946 in the year 2016), followed by India, 21.7% (from 78 385 in 2015 to 95 377 in the year 2016) and Germany, 21.5% (from 256 646 in the year 2015 to 311 832 in the year 2016). South Africa's tourism platform encompasses among others, the rich cultural diversity and natural beauty, and tourists come from Overseas and Africa at large. Tourism destinations in South Africa are complex areas of interests that saw many tourists frequenting them, and these include the Kruger and Kalahari national parks, Table Mountain in Cape Town, Victoria and Alfred waterfront among others. In the year 2002, an International Tourism growth strategy was formulated and it was based on intensive market segmentation. This strategy saw a sharp rise in South Africa tourism with total foreign tourist arrivals growing by an unprecedented 11.1% to 6.4 million, Overseas tourist arrivals increased by 20.3% and arrivals from Africa by 7.7%, thus affirming that the strategy had a positive impact on its target audience. This led to the global competitiveness study being commissioned by Department of

Environmental Affairs and Tourism (DEAT), Department of Trade and Industry (DTI) and South Africa tourism in the year 2004 as a way of identifying products, services and infrastructure gaps that could be used in developing interventions that bring lasting solutions to the tourism problems. This competitiveness study was aimed at finding products, services and infrastructure that accommodate all. Singh (2013) highlighted that tourism earnings were one of the five leading sources of foreign currency earning export revenue (for 69 developing countries, South Africa included) during the period 1995 to 1998.

Apart from its natural beauty and rich cultural diversity, South Africa boasts of having hosted many regional and international events since its peaceful elections held in 1994. It hosted the 1995 Rugby World Cup, the African Cup of Nations as well as the World Cup of Golf in 1996, the 1998 World Cup of Athletics and the 2003 Cricket World Cup, leading up to the biggest of them all, the 2010 soccer FIFA World Cup. These events marketed the South Africa tourism industry and consequently fuelled growth of various sectors as exemplified by the growth in the transport, food and accommodation sectors as well as creation of more jobs. Therefore, improving tourism facilities help improve visibility from the international community and subsequently an increase in tourist arrivals, which may in turn, lead to the hosting of other regional and or international major events. Lim (1997) cited the need to improve on the numbers of foreign tourist arrivals as it creates more income, job and investment opportunities, and this is something the South African government has started doing.

The statistical information available on the nature, progress as well as consequences of tourism in South Africa is largely based on arrivals and over-night stay statistics, balance of payments (BoP) data and South African Tourism (SAT) surveys, and this information fail to give a clear picture of the whole economic phenomenon of tourism (Statistics South Africa, 2012). This therefore calls for the need to develop models that will help in making sure that the data collected gives a clearer picture on the dynamics of tourist arrivals from Overseas and other Regions. The monthly data on tourist arrivals from Overseas, Southern African Development Commission (SADC), other African countries and the unspecified group is used in this research. This data was obtained from the South African Tourism and migration statistical release P0351 and these data sets are released monthly with a 2 months collection lag. The data gives an insight on all tourism activities within a given month. The data is routinely gathered by the Department of Home Affairs'(DHA) immigration officers at the air, land and seaports of entry/exit. In this research, total arrivals data is used and this is motivated by the need to understand the flow of tourists to South Africa.

This study aims to model and forecast the flow of total tourist arrivals to South Africa from Overseas, Southern African Development Commission (SADC), other African countries and the unspecified groupfor better planning, budgetary control and policy formulation that attract more Overseas and Regional travel exports in light of the impact of the COVID-19 pandemic. This helps to have an overview of the inferred total contribution being realized from tourism receipts before and after the COVID-19 pandemic and how best the tourism sector may be rejuvenated. Expected forecasts are compared to actual arrivals to assess the impact of the pandemic on the tourism sector up to date, for effective planning, budgetary control and crafting of policies that help improve sustainability in this new normal. The rest of the paper is structured as follows: Section 2 gives the literature review and Section 3 is on the methodology, Section 4 gives the results and discussions, Section 5 concludes.

Literature review

The contribution tourism has on economic growth has attracted a lot of attention from researchers and policy makers on how best, the tourism industry may be improved for better service delivery and overall improved economic growth for most African countries. Many studies have been done on investigation of the association that exists between tourism development and economic growth for various developing countries (Sharma, 2018; Govdeli & Direkci, 2017; Suresh & Senthilnathan, 2014). Most of these studies used panel data, cointegration and Granger casualty approaches and a positive association was observed in all the studies above. A notable number of studies have been conducted in determining the relationship between tourism and economic growth among other variables. Naude & Saayman (2005) used regression techniques in determining the relationship between various explanatory variables and tourist arrivals to Africa. The Holt's additive and multiplicative smoothing models, Autoregressive Integrated Moving Average (ARIMA), Vector autoregressive (VAR) models are some of the widely used forecasting models. Purna (2011) used univariate time series for tourist arrivals in India from December 1990 to January 2010. The SARIMA model was considered the best for forecasting in this dataset. Jere et al. (2020) modelled international tourist arrivals to Zambia from 1995 to 2014 using Holt-Winters Exponential Smoothing (HWES) and ARIMA models. In this case, the HWES outperformed the ARIMA models. Despite the use of other models in analysing time series data, the ARIMA model has been proved to be reliable in modelling and forecasting tourism demand using monthly or quarterly time series data, and they are the most popular linear model for time series forecasting (Saayman & Saayman, 2010; Louw, 2011; Peiris, 2016).

Saayman and Saayman (2010) developed a SARIMA model for forecasting tourist arrivals for South Africa from the Overseas market, making it the first research on forecasting of tourist arrivals in the new South Africa using SARIMA models. The authors used data from 1994 to 2006 in developing their ARIMA models for each and every 5 countries that formed their overseas market, which are United States of America, Germany, Great Britain, France and Netherlands. In their study, results showed that the seasonal ARIMA models delivered better accurate predictions of arrivals over different time horizons. The authors used error measures such as RMSE and MAPE in determining the best SARIMA models. From the year 2010 to date, there has been a positive development on research done on tourism forecasting, though there are still few quantitative models for the tourism data in planning and decision making processes when considering the role of tourism in South Africa.

Other notable researches are also those done by (Saayman & Saayman, 2008; Louw, 2011). Saayman and Saayman (2008) noted the rise in tourist arrivals to South Africa and tried to determine factors that affect inbound tourism to South Africa for sustainable economic growth. These authors used a multivariate framework technique for cointegration analysis using quarterly time series data spanning from 1993 to 2004. From their analysis, the authors found out that the main contributors to tourist arrivals are income, relative prices and travel cost. Climate and capacity factors were also noted as important contributors to tourist arrivals. Louw (2011) studied data for different markets namely: Asia, Australasia, Europe, North America, South America and the United Kingdom as they visited South Africa. The author used a SARIMA model, Autoregressive Distributed Lag Model (ADLM) and the Vector Autoregressive (VAR) model and Vector Error Correction Model (VECM) for each of the identified six markets. The SARIMA model was fitted for each of the origin market as a baseline model in determining the forecasting accuracy of this approach. Results obtained indicated that the SARIMA model achieved more accurate forecasts than the VAR model, and the VAR forecasts were more accurate than the ADLM forecasts. Louw (2011) also highlighted the need for policy-makers to use both the SARIMA and VAR models. This suggests that research on forecasting tourism demand in South Africa is still limited, hence the need for more empirical evidence in this area of tourism forecasting.

Chu (2008) modelled tourism demand for Singapore using a fractionally integrated autoregressive moving average (ARFIMA) model. However, the author, noted that the

SARIMA model performs best in the middle-run whereas the ARFIMA gives highest forecasting accuracy both in the short-run and long-run. Results obtained using the ARFIMA methodology by Chu (2008) and Gil-Alana (2005) show that there is an improvement on the accuracy of forecasts obtained by the ARFIMA approach when compared to the general ARIMA and SARIMA models.

Considering the popularity of the univariate time series models in forecasting tourist arrivals, the SARIMA approach is used in this study to model the total tourist arrivals to South Africa.

Materials and methods

A time series model helps explain a variable of interest with respect to its past behavior together with a random disturbance term. These models have been widely used for tourism demand due to their efficacy. In this section, ARIMA and SARIMA models developed by Box and Jenkins are discussed.

ARIMA (p, d, q) model

ARIMA models are considered the general class of models that can be used in modeling and forecasting time series data. These models are an extension of the Autoregressive Moving Average (ARMA) process. Given an ARIMA model that is not stationary, the data may be made stationary in the mean through various data transformation techniques such as differencing, if the data is not trend stationary, and log transformation if the data is not variance stationary. ARIMA models were pioneered by Box and Jenkins (Box et al., 1994) and they include the Autoregressive (AR) and the Moving Average (MA) models. A time series $\mathrm{Z_t}$ is said to follow an ARIMA model if the d^{th} difference $\Delta^{\mathrm{d}}\mathrm{Z_t} = W_t$ is a stationary ARMA process. The distribution of tourists in any given area or country tend to be influenced by major events and the calendar season.
The ARIMA $(p, d, q)(P, D, Q)_s$ or SARIMA is of the form:

$$\varphi_p(B)\ \Phi_{\mathrm{P}}(\mathrm{B})\Delta^{\mathrm{d}}\Delta_{\mathrm{S}}^{\mathrm{D}}\mathrm{Z_t} = \theta_{\mathrm{q}}(\mathrm{B})\Theta_{\mathrm{Q}}(\mathrm{B})\mathrm{a_{t,}} \quad (1)$$

where

- p represent the order of the non-seasonal autoregressive part,
- d represent the order of integration of the process,
- P represent the order of the seasonal autoregressive seasonal part,
- D represent the seasonal differencing,
- S represent the maximum number of seasons (for example S can be 12 for monthly data),
- q is the order of the non-seasonal moving average part,
- Q is the order of the seasonal moving average part,
- Δ^{d} represent a non-seasonal difference with $\Delta^{\mathrm{d}} = (1-\mathrm{B})^{\mathrm{d}}$,
- $\Delta_{\mathrm{S}}^{\mathrm{D}}$ represent the seasonal difference with $\Delta_{\mathrm{S}}^{\mathrm{D}} = (1-\mathrm{B^S})^{\mathrm{D}}$.

Data transformation

Most time series data is not stationary as they may follow other distributions other than the normal distribution depending on the nature of the variable being studied. As many statistical analyses assume normality, there is a need to make the data approximate a normal distribution before embarking on any statistical analysis. Transformations are used in normalizing non normal data to become normal and in making data that is not stationary become stationary. Detrending is a method normally used in dealing with trend in a given dataset.

The indicated transformations in equation 1, are used to make the data stationary. The data is first made variance stationary by taking a log transformation and then made trend stationary by differencing.

Model adequacy

Model diagnostic checking is done through an examination of residuals for autocorrelation, constant variance, normality and stationarity. Analysis of the residuals is done to test the adequacy of the various models developed including the Auto-Correlation Function ACF, Partial Auto-Correlation Function PACF, Extended Auto-Correlation Function (EACF) and other plots. Measures such as the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and the Mean Absolute Scaled Error (MASE) were also used in checking the model's goodness of fit for in-sample forecasting. The MAPE is to be used in checking the model's forecasting power as a value for MAPE below 10% indicates that the model has a good forecasting accuracy. Measures such as MAPE and MASE are the commonly used (Saayman, 2010). The best model to be considered in this research is the one with smaller values of AIC and BIC.

Box-Ljung test

The Box-Ljung test is a portmanteau test which is a diagnostic tool used in testing for autocorrelation. This test is applied to the residuals of a fitted time series model.
H_0: The residuals are independently distributed.
H_1: The residuals are dependently distributed.
The test statistic is

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

where *n* represent the sample size, $\hat{\rho}_k$ is the sample autocorrelation estimate at lag *k* and *h* represent the number of lags being tested. A good model should have residuals that are independently distributed.

Bayesian Information Criterion (BIC)

This is a criterion for model selection among a number of finite models, and the model with the lowest BIC is preferred. Mathematically, it can be represented as:

$$BIC = ln(n)k - 2ln(\hat{L}), \qquad (3)$$

where $(\hat{L})$represent the maximum value of likelihood function of the model with *n* being the number of data points and *k* being the number of free parameters to be estimated.

Akaike Information Criterion (AIC)

This is an estimator used in obtaining prediction error as well as the relative quality of statistical models for a given dataset. The procedure estimates the relative amount of information lost by a given model, hence the need for a lower AIC value as this implies that the model is of higher quality when it loses less information. It is obtained as:

$$AIC = 2k - 2ln(\hat{L}), \qquad (4)$$

where *k* is the number of estimated parameters in the model and $(\hat{L})$ is the maximum value of the likelihood function for the model.

Data analysis

Monthly tourism data were obtained from the Statistics South Africa's (Stats SA) Tourism and Migration P0531 reports from January 2009 to February 2020. Monthly data from January

2009 to August 2019 were used in developing a forecasting model and the data from September 2019 to February 2020 was then used for comparing the sample forecasted values with the actual to check for the goodness of the model in fitting a given dataset. Later the model was fitted to the whole data set, and the model was then used for out of sample forecasts. The R software package was used in data analysis.

Data analysis on total tourist arrivals in South Africa

Model identification

Model identification in this case started with the computation of the descriptive statistics. From the descriptive statistics in Table 1, the minimum number of tourists who visited South Africa per month during the period under review is 505 431 with its corresponding maximum value of 1 103 940. The range of this dataset indicates that the values are not too far apart and this helps when forecasting future observations especially in the absence of pandemics or natural disasters.

Table 1: Descriptive statistics

Min.	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max.**
505431	696811	767195	770590	839047	1103940

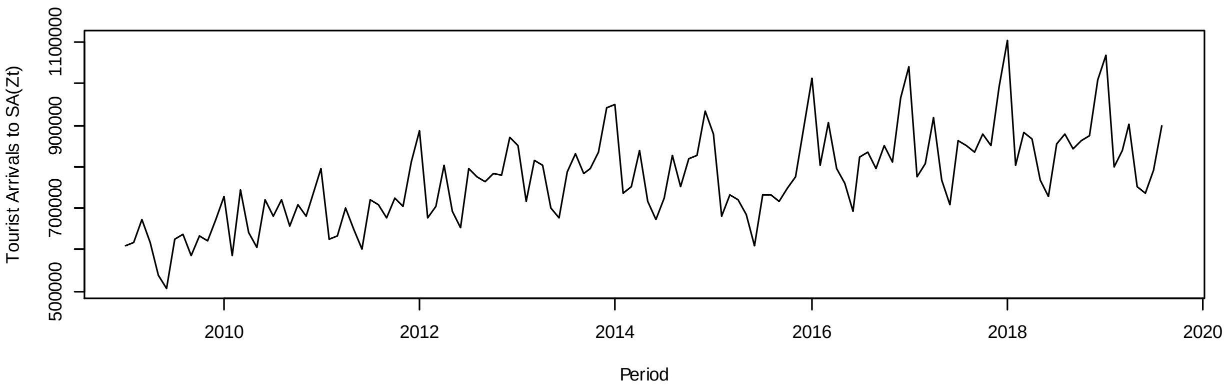

Figure 1: Time series plot on total tourist arrivals to SA

The time plot of total tourist arrivals to South Africa from January 2009 to August 2019 in Figure 1 indicates that the data follow an increasing pattern with an increasing variance and seasonality. Highest figures are recorded in December and January each year. Lowest figures are in February, May and June for each year.

African Journal of Hospitality, Tourism and Leisure, Volume 11 (4) - (2022) ISSN: 2223-814X

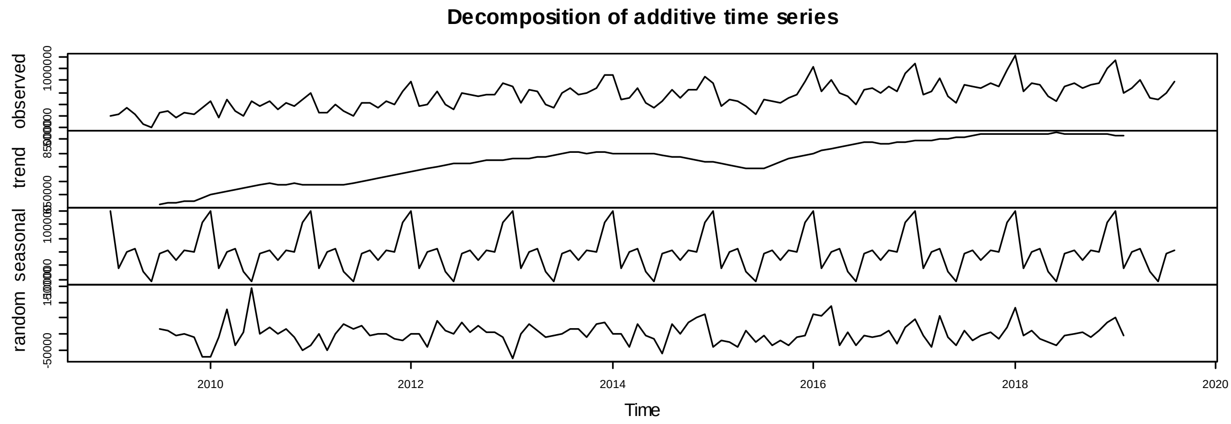

Figure 2: Decomposition of additive time series

Observing the four graphs in Figure 2 above of the decomposition of the additive time series, one can see the need to tame the increasing variance. Seasonality is one significant component of South African tourism data series hence there is also need to deseasonalise the data. The deseasonalised and detrended logged series data was tested for stationarity using the Augmented Dickey- Fuller (ADF) test and a p-value of 0.01 confirmed the data to be stationary. Graphs of the ACF, PACF and the EACF were also drawn and ACF plot indicated a sinusoidal pattern which clearly indicates a seasonal pattern and non-stationarity and the PACF plot confirming the presence of a seasonality component in the series. From the time series plot, there is need to log the data in order to cater for the increasing variance. The ACF and PACF plots suggest the need to log transform and/or difference the series as well as deseasonalising it.

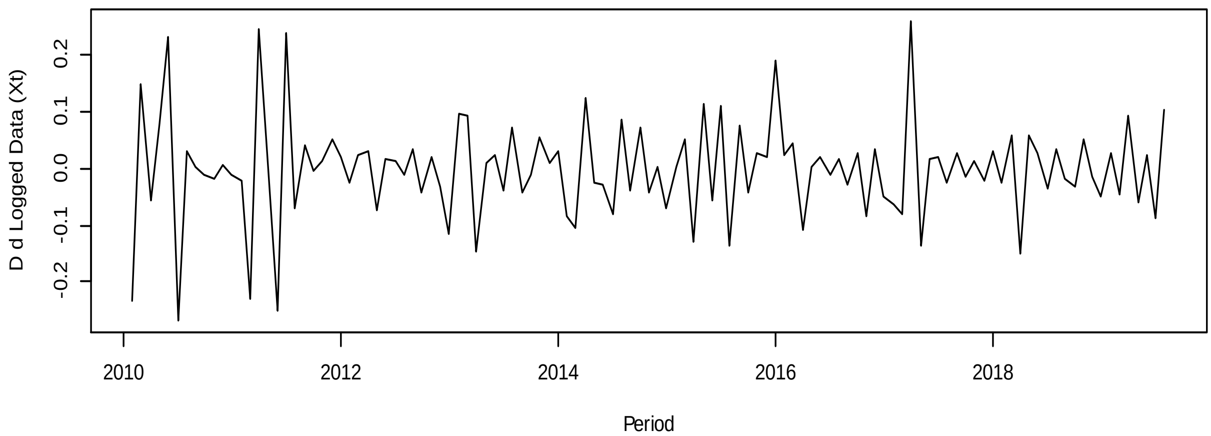

Figure 3: Time series plot of the log, first and seasonal differencedtourism data

The plot of the log, first and seasonal differenced data in Figure 3 indicates that the plot no longer shows any trend or seasonal pattern. D and d logged data stands for data that has been logged, first and seasonal differenced. The ADF test was performed on the deseasonalised and detrended logged series and the p-value obtained indicate stationarity as highlighted in Table 2.

Table 2: ADF test for thelog, first and seasonal differenced tourism data

Augumented Dickey-Fuller test for the**log, first and seasonal differenced tourism data**		
Test statistic = -6.6276	lag order = 4	p-value = 0.01

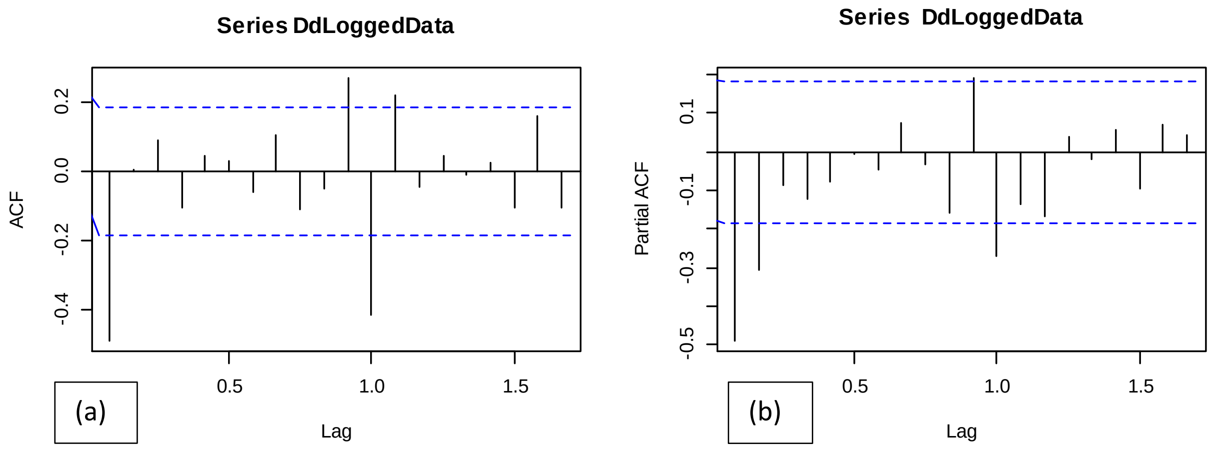

Figure 4: ACF and PACF plot of the log, first and seasonal differenced tourism data

The ACF plot 4(a) indicates that the model cuts of at lag 1. The PACF plot 4(b) indicate a cut off at lag 2 thus suggesting an ARMA(2,1)as the initial model. The ACF and PACF plot indicate strong seasonality due as shown by the presence of some spurious lags.

Table 3: The EACF plot of the log, first and seasonal differenced tourism data

```
AR/MA
 0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o oooooooo x  xx  o
1 x o ooooooooo  x  o  o
2 x x o oooooooo  x  o  o
3 x o ooooooooo  x  o  o
4 x o ooooooooo  x  o  o
5 x x o oooooooo  x  x  o
6 x x o oooooooo  x  o  o
7 x x o oooooooo  x  o  o
```

The EACF plot of the log, first and seasonal differenced tourism data in Table 3 indicate the existence of models such as ARMA(0,1),ARMA(1,1) and ARMA(2,2). The possible models are fitted and the results are tabulated in Table 4.

Table 4: Model fitting and adequacy checking

Model	**AIC**	**BIC**	**RMSE**	**MAPE**	**MAE**	**MASE**
ARIMA $(0,1,1)(0,1,1)_{12}$	-310.26	-302.03	0,0557	0.2891	0.0391	0.5620
ARIMA $(2,1,1)(0,1,1)_{12}$	-306.81	-293.09	0.0556	0.2913	0.0394	0.5662
ARIMA $(1,1,0)(1,1,1)_{12}$	-292.97	-281.99	0.0595	0.3079	0.0417	0.5946
ARIMA $\mathbf{(1,1,0)(1,1,0)_{12}}$	**-279.82**	**-271.61**	**0.0655**	**0.3263**	**0.0441**	**0.6341**
ARIMA $(1,1,2)(0,1,1)_{12}$	-307.7	-293.98	0.0552	0.2915	0.0395	0.5667
ARIMA $(1,1,0)(0,1,1)_{12}$	-294.9	-286.66	0.0596	0.3090	0.0418	0.6006
ARIMA $(1,1,1)(0,1,1)_{12}$	-308.27	-297.29	0.0557	0.2887	0.0391	0.5613
ARIMA $(0,1,1)(1,1,0)_{12}$	-296.19	-287.96	0.0609	0.2994	0.0405	0.5821

Note: Best model is in bold.

The ARIMA $(1,1,0)(1,1,0)_{12}$ was chosen as the best model due to its low values of AIC and BIC as shown in Table 4. This model also has a high accuracy level in forecasting since its MAPE value of 0.3263% is less than 10%. MAE (0.0441) measure the average absolute deviation of forecasted values from original values.

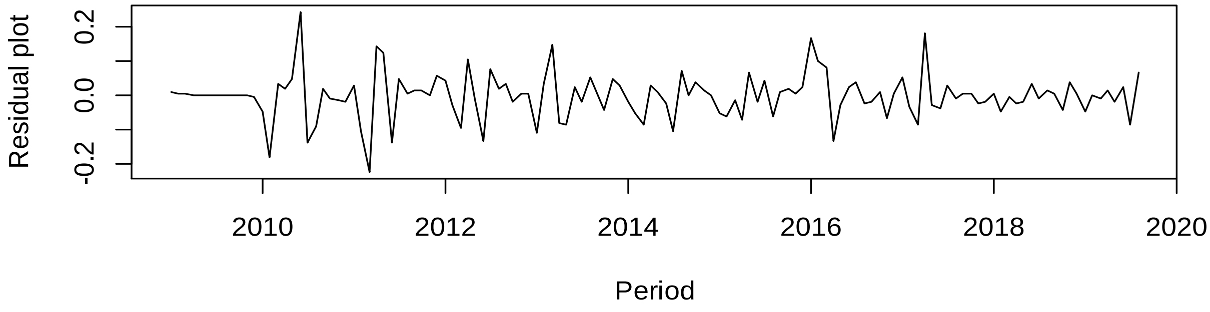

Figure 5: Time plot of the residual values from model ARIMA $(1,1,0)(1,1,0)_{12}$

The time plot of residuals obtained from the ARIMA $(1,1,0)(1,1,0)_{12}$ in Figure 5 indicate that the model developed is a good fit for the data as the residual observations are fluctuating around the zero mean and behave like white noise.

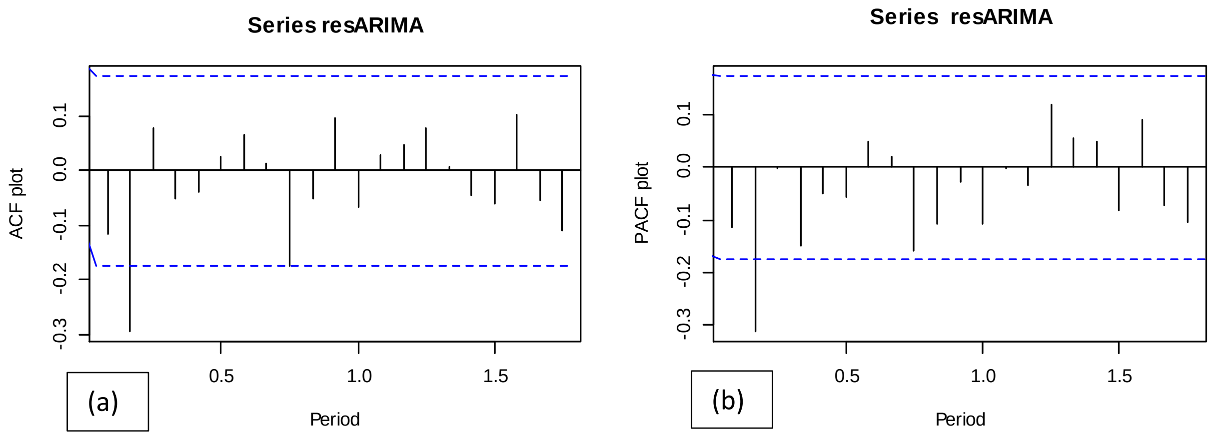

Figure 6: ACF and PACF plot of residuals of the ARIMA$(1,1,0)(1,1,0)_{12}$

The ACF and PACF plot of the residuals as shown in Figure 6 (a) and (b) do indicate a white noise process despite the presence of one lag outside the dotted blue line hence this model is a good fit. The Box-Ljung test has a p-value of 0.1621 which confirms that the model residuals are independently and normally distributed.

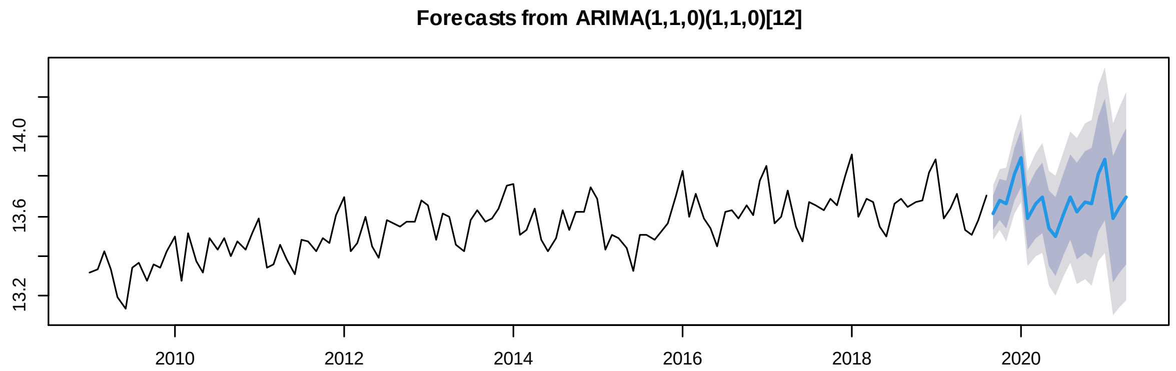

Figure 7: Plot of the forecasted values

The plot of the predicted values in Figure 7 indicates that the forecasted values can be used for decision making and policy review and or formulation. The forecasted values from the model are well behaved.

Table 5: Forecasted values against actual

Month	**Forecast**	**Lo 95**	**Hi 95**	**Actual**
Sept 2019	819336.4	714722.8	939262.2	780381.00
Oct 2019	875123.2	750774	1020068	834275.00
Nov 2019	856277.7	712111.7	1029620	850537.00
Dec 2019	998800.2	814858.7	1224264	981038.00
Jan 2020	1078886	862120.2	1350155	1093268.00
Feb 2020	796976.1	625802.4	1014981	800815.00

Lo 95 is the lower 95% confidence limit and Hi 95 is the higher 95% confidence limit. Data in Table 5 above show that the forecasted values are very close to the actual values, suggesting the ARIMA $(1,1,0)(1,1,0)_{12}$ as a good fit for this dataset.
The model is then fitted to the whole data set and forecasts are now made outside of the sampling period.
The model can be written as follows

$$Z_t = \phi Z_{t-1} + \varphi Z_{t-12} - \phi\varphi Z_{t-13} + a_{t,} \tag{5}$$

Where ϕ and φ are the coefficients of the model and a_t is a random error term. With the parameter estimates, the model becomes:

$$Z_t = -0.4912 Z_{t-1} - 0.4468 Z_{t-12} - 0.2195 Z_{t-13} + a_t. \tag{6}$$

The coefficients of the estimated model indicate that the model follow a stationary process as all the parameter estimates of the fitted model are less than one. The fitted model was used in obtaining forecasts from March 2020 to December 2021. Actual values observed from March 2020 to March 2021 were also obtained as we determine the accuracy of forecasts with relation to actuals figures especially now that we are in the COVID-19 era. Differences between forecasted and actual values from March 2020 to March 2021 were computed and tabulated in Table 6.

Table 6: Forecasts, Actual and Differences

Month	Forecast	Actual	Differences
March 2020	862456.5	535094.00	327362.50
April 2020	889744.80	29341.00	860403.80
May 2020	762875.40	49481.00	713394.40
June 2020	735039.90	62841.00	672198.90
July 2020	822200.90	68914.00	753286.90
August 2020	891454.80	67051.00	824403.80
September 2020	810705.40	75273.00	735432.40
October 2020	849922.10	73988.00	775934.10
November 2020	864874.80	101096.00	763778.80
December 2020	996555.40	198059.00	798496.40
January 2021	1086280.00	139134.00	947146.00
February 2021	802357.90	90165.00	712192.90
March 2021	855139.60	157638.00	697501.60
April 2021	899217.30		
May 2021	761686.20		
June 2021	737735.00		
July 2021	810940.50		
August 2021	897465.50		
September 2021	800170.40		
October 2021	846233.00		
November 2021	861844.40		
December 2021	993520.50		

African Journal of Hospitality, Tourism and Leisure, Volume 11 (4) - (2022) ISSN: 2223-814X

With information on forecasts, actuals and differences (between forecasts and actuals) from March 2020 to date, a time plot on the differences was plotted. The plot can help visualise the effect made on South Africa's tourism industry by COVID-19.

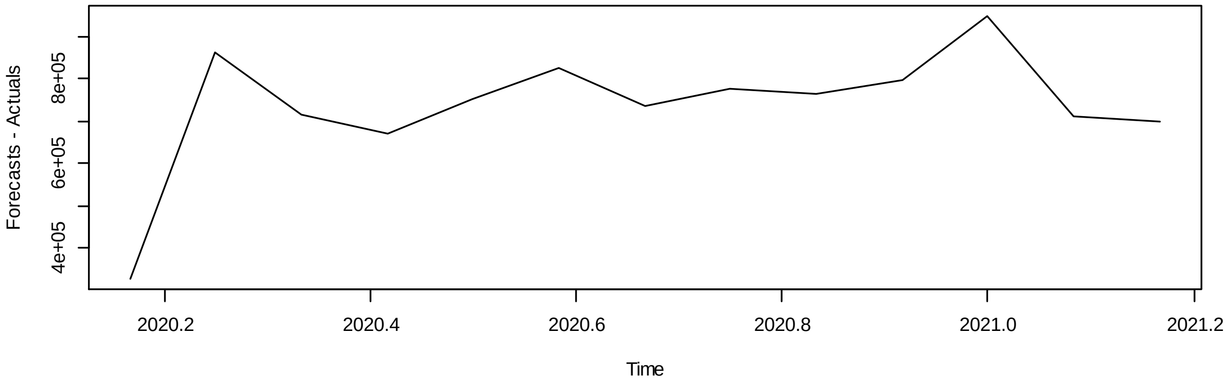

Figure 8: Plot of the difference values

The plot of the differences values obtained by subtracting actual observations from forecasted values from March 2020 to March 2021, do indicate that the tourism industry has suffered a lot due to the COVID-19 pandemic which saw a series of lockdown measures being effected. Most of the lockdowns effected saw the closure of different ports of entry to non- South Africa citizens thus having a negative impact on the tourism sector earnings and opportunities from foreign visitors. From March 2020, tourist arrivals are less than a tenth of what they used to be before the COVID-19 pandemic. Therefore, South Africa is losing more than 90% of its tourists currently because of the pandemic, and shows little sign of an imminent recovery. This has a devastating impact on the tourism industry.

Conclusions

Foreign tourist arrivals are the main driver of revenue for many countries, South Africa included. It is this reason why predicting, forecasting and management of its trend is of great importance and a necessity for tourism growth. This research aims at modeling tourist arrivals to South Africa from January 2009 to Feb 2020. Data used in this research exhibited an increasing variance coupled with trend and seasonal variations. December has been noted as the time the tourism sector revenue and activities are at its peak. A logarithm transformation was used in taming the increasing variance as well as a regular and seasonal differencing. An ARIMA $(1,1,0)(1,1,0)_{12}$ model was fitted and considered the best among other models according to their low AIC and BIC values and other model accuracy measures namely MAPE, MAE, MASE an RSME. The South Africa tourism department need to continue marketing and advertising their products so as to continue increasing their tourism revenue and employment for the locals, but the COVID-19 pandemic needs to be brought under control first with a speedy vaccination programme. Issues of human job security need be prioritized as this has a negative impact on the tourism market. SARIMA models do play an important role in modeling and forecasting tourism data for useful marketing and planning purposes. COVID-19 has had a negative impact on the tourism industry, South Africa included, and the negative effects are trickling down to the lowest levels. The evidence is the decline in the actual number of tourists who visited South Africa from March 2020 to date. Though there seem to be anon-significant increase in between March 2020 to date, there is a lot that need to be done to resuscitate the industry. With the pandemic in mind, the SARIMA model is useful in providing information on the potential number of tourists that South Africa could be receiving, if it were able to tame the COVID-19 pandemic through vaccination together with the rest of the world. South Africa is currently loosing over 90% of potential tourists per month because of the pandemic.

With the South Africa tourism industry contributing heavily to the country's GDP, good forecasts are therefore a necessity for effective planning, restructuring and growth. There is also need to consider data on gender, age and source region distribution over the years as this may help come up with tailor made strategies for the different categories of tourists who visit South Africa. Though SARIMA models have proved to model tourism data well, there is need to consider other models that accommodate time series data such as the ARCH and GARCH, which factor in the component of a non-constant variance. Hybrid models also need to be considered for the tourism dataset. Models which also factor-in components that deal with risk or uncertainity need be looked into as this will help in coming up with models that withstand the various shocks or jumps that may affect the tourism industry in future, such as the outbreak of other pandemics as exemplified by the COVID-19 pandemic. With COVID-19 on our doorsteps, there is need for government, communities and other stakeholders to support tourism related projects for quicker resuscitation of the tourism industry.

References

Akuno, A. O., Oteieno, M. O., Mwangi, C.W. & Bichanga, L. A. (2015). Statistical Models for Forecasting Tourists' Arrival in Kenya. *Open Journal of Statistics*, 5, 60-65

Bhorat, H., Steenkamp, F., Rooney, C., Kachingwe, N. & Lees, A. (2016). Understanding and characterizing the services sector in South Africa: An overview. *WIDER Working Paper, No. 2016/157*

Box, G., Jenkins, G. M. & Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control.* 3rd Edn., Prentice Hall

Chu, F. L. (2008). A fractionally integrated autoregressive moving average approach to forecasting tourism demand. *Tourism Management*, 29, 79-88

Gil-Alana, L. A. (2005). Modelling International Monthly Arrivals Using Seasonal Univariate Long Memory Processes. *Tourism Management*, 26(2), 887-878

Govdeli, T. & Direcki, T. B. (2017). The Relationship Between Tourism and Economic Growth: OECD Countries. *International Journal of Academic Research in Economics and Management Sciences* 6(4), 104-113

Jere, S., Banda, A., Kasense, B., Siluyele, I. & Moyo, E. (2019). Forecasting Annual International Tourist Arrivals in Zambia Using Holt-Winters Exponential Smoothing. *Open Journal of Statistics*, 9, 258-267

Li, G. & Wu, D. C. (2019). Introduction to the special issue: Tourism forecasting – New trends and issues. *Tourism Economics,* 25(3), 305–308

Lim, C. (1997). An Econometric Classification and Review of International Tourism Demand Models. *Sage Journals*, 3(1), 69-81

Louw, R. (2011). Forecasting tourism demand for South Africa. Unpublished master's thesis, North-West University, Potchefstroom.

Naudé, W. A. & Saayman, A. (2005). "Determinants of tourist arrivals in Africa: a panel data regression analysis." *Tourism Economics*, 11(3), 365-391

Oh, C. (2005). The Contribution of Tourism Development to Economic Growth in the Korean Economy. *Tourism Management*, 26, 39–44

RangikaIroshaniPeiris, H. (2016). A Seasonal ARIMA Model of Tourism Forecasting: The Case of Sri Lanka. *Journal of Tourism, Hospitality and Sports,* 22(2312-517), 98-109

Richardson, R.B. (2010). The contribution of tourism to economic growth and food security. *USAID Mali Off. Econ. Growth,* doi: 10.22004/ag.econ.97140

Saayman, A. & Saayman, M. (2008). Determinants of inbound tourism to South Africa. *Tourism Economics,* 14(1), 81-96

Saayman, A. and Saayman, M. (2010). Forecasting Tourist Arrivals in South Africa.' *ActaCommercii*, 10 (1): 281–293

Sharma, N. (2018). Tourism Led Growth Hypothesis: Empirical Evidence from India. *African Journal of Hospitality, Tourism and Leisure*, 7 (2), 1-11

Singh, E. H. (2013). Forecasting Tourist Inflow in Bhutan using Seasonal ARIMA. *International Journal of Science and Research (IJSR),* 2, 242-245

Song H. & Turner, L. (2006). Tourism demand forecasting. In Dwyer, L. & Forsth, P. (Eds.). *International Handbook on the Economics of Tourism,* Edward Elgar Publishing Ltd

Suresh, J. & Senthilnathan, S. (2014). Relationship between tourism and economic growth in Sri Lanka, Published as the 7th chapter of a book entitled *"Economic Issues in Sri Lanka*" compiled by Dr. S. Vijayakumar, 115–132

World Travel and Tourism Council (2010), *The Economic Impact of Travel and Tourism: Mali. London*