AFRICAN JOURNAL OF HOSPITALITY, TOURISM AND LEISURE
2026, Vol. 15, No. 1, pp. 262-271
https://doi.org/10.46222/ajhtl.19770720.736

AJHTL
OPEN ACCESS

Research article

Statistical analysis of tourist arrivals to South Africa from non- Southern African Development Community African countries and the unspecified group, an under-tapped market

Musara Chipumuro* [1], Delson Chikobvu [1] and Tendai Makoni [1]
[1] Department of Mathematical Statistics and Actuarial Science, University of the Free State, South Africa

ABSTRACT
Tourism is a sector that has gained significant attention from policy formulators and decision-makers around the globe, given its contribution to the overall development of the host country and beyond. The paper suggests a suitable Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast non-Southern African Development Community (non-SADC) African countries and the Unspecified tourist arrivals to South Africa (SA). Tourist arrivals to SA from non-SADC African countries and those who did not specify country of origin (the Unspecified group) from July 2010 to February 2024 are examined. The non-SADC African countries comprise 46 countries, and the Unspecified group comprises those tourists who failed to specify their country of origin at their various ports of entry. A time series modelling approach is considered to support better-informed decision-making, forecasting, and assessment of the impact of the COVID-19 pandemic in the tourism sector. A SARIMA(2,1,0)(1,0,1)$_{12}$ was considered the best model based on its lowest Akaike Information Criterion (AIC) value. Residual analysis is done to validate model accuracy and to support the fitted model. A graph of the forecasts and actuals highlights the negative impact of COVID-19 on tourist arrivals, the tourism sector and the country at large. There was a major step change in this category of tourist arrivals to SA due to the COVID-19 containment measures.

KEYWORDS
Tourist arrivals; Box-Jenkins methodology; non-SADC African countries; Unspecified group; SARIMA

Introduction

Sustainable tourism management plays a crucial role in a host country's development, as it concurrently addresses environmental, infrastructural, cultural, social, and economic development. This is realised through foreign currency earnings, foreign direct investment (FDI), employment creation, stimulation of the local economy, infrastructural development, ecosystem management and development, contribution to Gross Domestic Product (GDP), support for small and medium-sized enterprises (SMEs), improved human security, and market positioning, among others. Boniface et al. (2016) postulate that tourism contribution is widespread and boosts consumption, infrastructure development and GDP. UNWTO (2019) and WTTC (2018) highlight that, despite tourism activities being the main drivers of foreign currency generation and improving the balance of payments, international tourism creates employment, income, savings, investment, and economic growth. The increasing interconnectivity of the tourism sector and the need to adopt sustainable practices have drawn more attention to it, given its contribution to the overall economic development of any host country, including SA. According to WTTC (2020), SA accounted for 9.5% of GDP from tourism activities. The WTTC (2022) report points out that SA's tourism industry accounts for almost 8.6% of the country's GDP and employs over 1.5 million people. Though tourism's contribution to SA's GDP dipped during the COVID-19 era, it bounced back in 2023, accounting for 8.2% of SA's GDP,

CORRESPONDING AUTHOR'S CONTACT: Musara Chipumuro chipumuromusara@gmail.com
HOW TO CITE: Chipumuro, M., Chikobvu, D. & Makoni, T. (2026). Statistical analysis of tourist arrivals to South Africa from non- Southern African Development Community African countries and the unspecified group, an under-tapped market. African Journal of Hospitality, Tourism and Leisure, 15(1), 262-271. https://doi.org/10.46222/ajhtl.19770720.736
ISSN: 2223-814X (Online) | © 2026 AJHTL

according to WTTC (2024). The WTTC (2024) report further postulated that SA's tourism sector also employed 1.46 million people in 2023. With these facts and figures, tourism is undeniably a vital contributor to SA's GDP, warranting ongoing research into how SA's tourism industry can best continue to grow, shine, and attract existing and new tourists, especially from other non-SADC African countries and beyond. These may be the forgotten tourists to whom SA can still benefit.

SA classifies tourists by region of origin, thereby enabling segmentation of travel patterns and the planning of appropriate policy interventions. International tourists have been broadly categorised into four groups: SADC tourists, non-SADC African tourists, unspecified tourists, and Overseas tourists. SADC tourists come from the 15 member states of the Southern African Development Community (SADC), such as Zimbabwe and Mozambique, and typically cross overland on short visits usually for trade, shopping, or visiting relatives. Unlike the non-SADC African tourists, they come from the rest of the African continent, including Nigeria, Kenya and Egypt. This segment travels by air and is more business-, medical-, or education-oriented, with longer stays and higher per capita expenditure. Overseas tourists are those coming from outside of the African continent, from Europe, North America, Asia, and Oceania and are leading contributors to foreign exchange through leisure and cultural tourism. From this categorisation, non-SADC African tourists are obtained by subtracting the SADC subset from the overall African category. In contrast, the Unspecified category comprises arrivals for which country-of-origin information is incomplete or ambiguous. The non-SADC African countries are those from East, Central, West, and North Africa. Countries in East and Central Africa include Burundi, Cameroon, Central African Republic, Chad, Comoros, Congo, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Kenya, Réunion, Rwanda, São Tomé and Príncipe, Somalia, and Uganda. Countries in West Africa include Benin, Burkina Faso, Cape Verde, Côte d'Ivoire, The Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Saint Helena, Senegal, Sierra Leone, and Togo. North Africa comprises countries such as Algeria, Egypt, Libya, Morocco, South Sudan, Sudan, Tunisia and Western Sahara. The countries that make up the non-SADC African group are many, with other countries, such as Egypt, Ghana, and Kenya, also advanced in their tourism industries. Yet, this group, coupled with the Unspecified, comprise just a small portion of the total tourist arrivals to SA. Although these groups, including the Unspecified category, account for only a small fraction of total tourist arrivals to SA, they represent untapped potential to boost the country's tourism sector.

Highlighting the combined interests and preferences of the non-SADC African and Unspecified group could enhance SA's appeal and increase visitor numbers. There is a need to get in-depth knowledge about the dynamics of tourists from these countries for informed decision-making on stimulating local and regional development through providing information using time series modelling, making it the first study on tourist arrivals to SA focusing on the marginalised non-SADC African and those who did not specify country of origin (Unspecified group). Tourist arrivals from non-SADC African countries have increasingly become a critical component of SA's tourism sector, despite being overshadowed by SADC tourism and tourist arrivals from Western markets. The non-SADC African countries offer notable economic and strategic advantages to SA, even though many are still developing their tourism appetite, business, and markets. Increased inclusion of tourists from the non-SADC African countries will help mitigate the risk of overdependence on SADC and Western markets. This is supported by UNWTO (2023), which highlighted the importance of mitigating risk by diversifying the tourism portfolio and reducing overdependence on any single region. According to SAT (2022), tourists from non-SADC African countries tend to exhibit higher per capita spending, especially those travelling for business, medical treatment, and/or education. This is so because visitors from non-SADC African countries tend to stay longer than tourists from SADC and Western markets, thereby contributing more significantly to tourism receipts. This increased expenditure benefits accommodation, transport services, restaurants, and the retail sector, for example. Moreover, tourists from non-SADC African countries can contribute to the growth of niche tourism segments such as medical, educational, and meetings, incentives, conferences, and exhibitions (MICE) tourism. This is so because SA is considered a more developed country in the healthcare and tertiary education sectors and attracts a significant number of visitors from Nigeria, Kenya, Ghana, and Ethiopia. Naude and Saayman (2005) pointed out that these arrivals often result in extended stays and wider engagement with domestic economic sectors beyond hospitality and leisure, hence the need for a niche-based tourism market.

Considering the African Union (AU)'s Agenda 2063, South Africa's improvement of tourism business with non-SADC African countries will ensure deeper Pan-African cooperation and integration through inter-African travel, a key pillar for achieving economic unity and shared prosperity. The AU (2020) Agenda 2063 highlights that increased tourism within the African region will help support its agenda by strengthening people-to-people ties, promoting cross-border cultural exchange, and facilitating investment networks. Moreover, SA's engagement in good tourism business will ensure that tourists from non-SADC African countries serve as informal ambassadors upon returning to their countries, thereby influencing travel behaviour to SA in their communities and ensuring a positive impact on SA's tourism receipts. As many non-SADC African economies are experiencing economic growth and rising middle classes, SA can benefit from this under-tapped, high-potential tourism market. Improved air connectivity, digital marketing, and simplified visa regimes could also help unlock this potential and support sustainable growth in tourist arrivals to SA from the non-SADC African countries. The tourism industry worldwide is characterised by fluctuations influenced by seasonal trends, global events and marketing conditions. These fluctuations fall into two categories, the internal and external environment, which greatly impact tourist arrivals, hence the need for proper management of tourism activities. Du Plessis and Cronjea (2021) also postulated on the various fluctuations that influence the tourism industry, such as significant seasonal variations, global economic conditions, political instability and environmental changes. With COVID-19 at our doorstep, the volatility of tourist arrivals became even more evident as the pandemic negatively affected them worldwide. Tourist arrivals from other non-SADC African countries and the Unspecified group play a critical role in South Africa's tourism sector, particularly in driving expansion and growth. The sheer number of tourists arriving in SA from non-SADC African and Unspecified groups, combined, compared with those from Overseas and SADC, indicates a dire need for SA to improve its tourism image in these countries. A look at the dynamics of tourist arrivals before, during, and after the COVID-19 pandemic will help clarify tourism dynamics and influence policy formulation, control, and decision-making at various levels of the decision-making process.

Many researchers have used tourist arrival data in quantitative tourism demand modelling, while others have used tourism expenditure or overnight stays. Researchers who considered tourist arrivals in tourism research include but are not limited to Hopken et al. (2021), Apergis et al. (2023), Zhu et al. (2018), Khairudin et al. (2018), Sun et al. (2019), Gunter and Onder (2015) and Vergori (2025). In this paper, tourist arrivals to SA from non-SADC African countries and the Unspecified group are combined. The reason for combining is that the tourist arrival numbers for both are relatively small. A time series model for tourist arrivals from non-SADC African countries and the Unspecified group is then considered. The paper aims to identify trends, seasonal patterns, and forecasting capabilities of the developed time series model, and to evaluate the major step changes caused by the COVID-19 pandemic in non-SADC African and Unspecified tourist arrivals. This study is structured as follows: Section 2 provides the literature review. Section 3, the methodology, Section 4, the results and discussions and Section 5 concludes and recommends.

Literature review

Forecasting plays a pivotal role in data science by enabling planning, goal setting, and anomaly detection. Kumar et al. (2021) highlight the importance of accurate forecasting of tourist arrivals for effective resource allocation, policy formulation and strategic planning in the tourism industry. Modelling tourist arrivals provides a clearer picture of the relevant information. The derived information will inform the view of the impact of set policies, infrastructure development, cultural diversity, and ecosystems, for example, on tourist arrivals. For many African countries, tourism is one of the biggest foreign currency earners, hence the need for informed decision-making about tourist arrivals. Human security policies, political stability, an enabling business environment, and good governance play a crucial role in attracting tourists to any host country, thereby providing a good foundation for tourism growth. Many studies have examined how tourist arrivals are influenced by weather, economic conditions, seasonality, and holidays, using various methodologies. A systematic approach to time series analysis is presented in the classic work by Box and Jenkins (1970). Methods such as the Autoregressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) have been widely used across various sectors due to their ability to model linear relationships and seasonality, as postulated by Box et al. (2015). A series of time series approaches, such as the Naïve I, Naïve II, Linear Trend, Sine Wave, Holt-Winters and ARIMA were used by Chu (1998) to forecast tourism demand in ten countries, which are Japan, South Korea, Taiwan, Hong Kong, Philippines, Indonesia, Singapore, Thailand, Australia and New Zealand. Nine of the ten countries studied by Chu had the ARIMA model as the most suitable for prediction when Mean Absolute Percentage Error (MAPE) was used to measure model adequacy. Song et al. (2009) demonstrated the effectiveness of SARIMA in capturing seasonal fluctuations in tourist arrivals to Hong Kong, outperforming several naïve and exponential smoothing models.

Tularam et al. (2010) adopted the logistic, ARIMA and Vector Autoregressive (VAR) models for the Australian dataset. The ARIMA (2,2,2) model performed better on the dataset than the logistic and VAR models. Goh and Law (2011) highlighted the effectiveness of SARIMA models in short-term forecasting of inbound tourism to Asia-Pacific countries, mostly for data with consistent seasonal peaks. Wang et al. (2013) used a SARIMA model and compared it with a seasonal moving average (SMA) and a Holt-Winters model. Their findings showed that the SARIMA model outperformed the other models under different forecast intervals. Makoni et al. (2022) demonstrated the use of SARIMA models in modelling and forecasting international tourist arrivals to Zimbabwe. Their findings show that a SARIMA$(1,0,0)(1,0,1)_{12}$ fitted their dataset well. Wikasanti et al. (2025) applied the SARIMA and Triple Exponential Smoothing (Holt-Winters) models to predict international tourist arrivals to West Sumatra in the post-COVID-19 pandemic era. Their findings point out the effectiveness of the SARIMA $(2,0,1)(1,0,1)_{12}$ in providing the most accurate predictions after considering the model's MAPE, MAE and RMSE over the five months. In this paper, a SARIMA model is considered as it allows for trend and seasonal components suspected to be present in the non-SADC African countries and the Unspecified group of tourist arrivals to South Africa.

Methodology

Time series modelling has gained momentum in understanding the behaviour of various real-life phenomena over time due to its ability to account for trend, seasonal, and/or cyclical components that may exist in each real-time data system. The Box-Jenkins methodology (1994) is adopted for the non-SADC African and Unspecified dataset. Model estimates are obtained using the maximum likelihood estimate (MLE) technique.

ARIMA (p, d, q) and SARIMA $(p,d,q)(P,D,Q)_s$ model

The ARIMA model is an extension of the Autoregressive Moving Average (ARMA(p,q)) model that includes an extra differencing step. The differencing component d is discrete, which helps detrend the time series dataset and make it stationary. The p and q are the non-seasonal components of the autoregressive (AR) and moving average (MA) parts, respectively. An ARIMA (p, d, q) model can be expressed as:

$$\phi_p(L)\,\nabla^{\mathrm{d}}\,\mathrm{Y_t} = \mu + \theta_{\mathrm{q}}(L)\,\mathrm{a_t}, \tag{1}$$

where $\phi_p(L) = [1 - \phi_1 L - \phi_2 L^2 \ldots \phi_{\mathrm{p}} L^p]$ and $\theta_{\mathrm{q}}(L) = [1 + \theta_1 L + \theta_2 L^2 \ldots . \theta_q L^q]$
where p, d, and q represent the non-seasonal AR component, non-seasonal difference and non-seasonal MA component, respectively. $\phi_p = [\,\phi_1, \phi_2 \ldots \phi_{\mathrm{p}}\,]$ and $\theta_{\mathrm{q}} = [\,\theta_1, \theta_2 \ldots . \theta_q]$ are vector non-seasonal coefficients of the AR and MA components, respectively. The white noise term is denoted by $\mathrm{a_t}$ and L is the backward shift operator, which can be written as $L\mathrm{Y_t} = \mathrm{Y_{t-1}}$. A SARIMA $(p, d, q)(P, D, Q)_s$ model can be expressed as:

$$\phi_p(L)\,\Phi_{\mathrm{P}}(L_s)\,\nabla^{\mathrm{d}}\,\nabla_s^{\mathrm{D}}\,\mathrm{Y_t} = \mu + \theta_{\mathrm{q}}(L)\Theta_{\mathrm{Q}}\,(L_s)\,\mathrm{a_t}, \tag{2}$$

where P, D, and Q represent seasonal AR, seasonal difference and seasonal MA orders, respectively, with s as the seasonality index. Φ_{P} and Θ_{Q} are vector seasonal AR and MA model parameters, respectively, with $L_s\mathrm{Y_t} = \mathrm{Y_{t-s}}$ and $\nabla_s^{\mathrm{D}}\mathrm{Y_t} = (1 - L^s)^D\,\mathrm{Y_t}$.

Data transformation and stationarity test

The Box-Cox transformation is used to determine the optimal transformation for the combined dataset of non-SADC African and Unspecified. The transformation is done to achieve normality. The Augmented Dickey-Fuller (ADF) test is used to determine the dataset's stationarity.

Model adequacy

The AIC is used for model selection. The mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE), and mean absolute scaled error (MASE) are used to assess model adequacy. The Box-Ljung and the Jarque-Bera tests are used to validate model accuracy through residual diagnostics.

Ethical considerations

The current study is grounded in publicly accessible aggregate-level secondary data on international tourist arrivals, gathered from the official Stats SA website via the Migration Statistical Release P0351. The data is initially collected by the Department of Home Affairs (DHA) at different ports of entry or exit. No personally identifiable data or individual-level information was accessed or used. Consequently, no human subjects were employed in this research, and institutional ethical approval was therefore not required. Still, research ethics of cautious use of data, transparency, and attribution were followed throughout the research. Due citation of the data source has been provided, and the analysis was conducted strictly in accordance with the conditions for the use of publicly available data as laid down by Stats SA.

Results

This section presents the analysis of monthly tourist arrivals in SA between July 2010 and February 2024 by tourists from other non-SADC African countries and the Unspecified group combined, focusing on long-term trends, seasonality and the predictive performance of the developed model. In this study, the data is divided into three sections: training (July 2010 to February 2019), validation (March 2019 to February 2020), and test (March 2020 to February 2024). The datasets were obtained from monthly reports published by the South African Tourism and Migration statistical release P0351 (http://www.statssa.gov.za/publications/P0351). The data is collected by the Department of Home Affairs (DHA) immigration officers at different ports of entry/exit. Descriptive statistics for the tourist arrivals of this targeted group were computed using data from July 2010 to February 2024, and the results are shown in Table 1.

Table 1: Descriptive Statistics: Non-SADC African countries and the Unspecified group combined.

Min	1st Qu.	Median	Mean	3rd Qu.	Max
5	11932	15588	14154	17334	23751

On average, SA received 14,154 tourists from other non-SADC African and unspecified sources per month. The minimum number of tourists received from this group is 5, and the maximum is 23751. The high variability in the range is evidence of the COVID-19 containment measures, which led to the closure of borders and airports, thereby limiting travel. There are 46 countries comprising the non-SADC African countries and the Unspecified combined tourist arrivals group. A time plot of the tourist arrivals is shown in Figure 1.

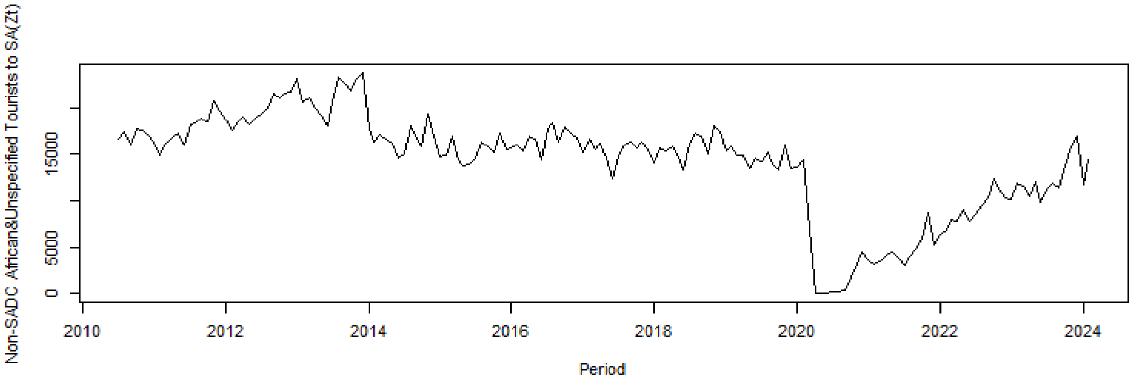

Figure 1: The time plot for non-SADC African and Unspecified tourist arrivals to SA from July 2010 to February 2024

The pattern in Figure 1 shows that the data for combined non-SADC African and Unspecified tourist arrivals to SA were relatively stable up to 2013. The increase in tourists to SA is commensurate with the notable events SA hosted during the period, such as the 2010 World Cup and the 2013 Africa Cup of Nations. The data became even more stable from 2013 till late 2019. The plot shows that the data decreased drastically in 2020 in the aftermath of the COVID-19 pandemic. Since August 2020, a rising, fluctuating pattern has been observed. A decomposition of the training data was performed and shown in Figure 2 to ascertain the components present in the combined non-SADC African and Unspecified tourist arrivals data.

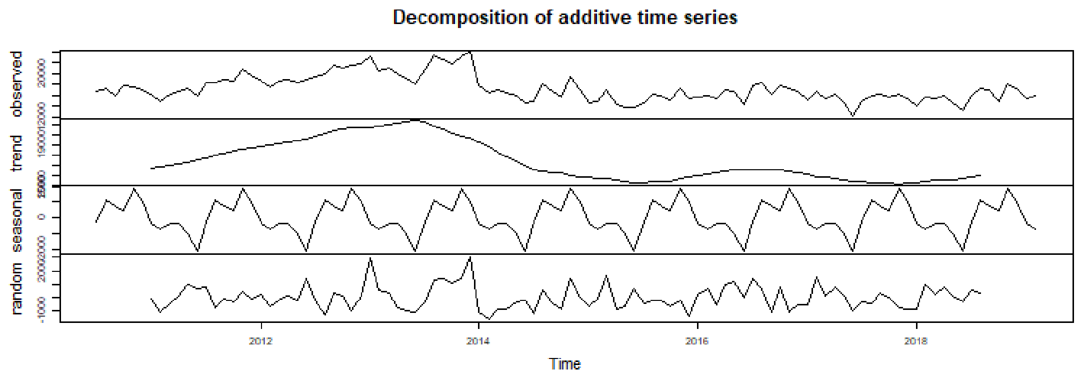

Figure 2: Decomposition of the additive time series for non-SADC African and Unspecified tourist arrivals data

Figure 2 shows the additive decomposition of monthly tourist arrivals from non-SADC African nations and the Unspecified combined for the training data, with clear trend, seasonal, and random components in the time series. The trend component shows a consistent rise in arrivals from 2011 to the end of December 2013, followed by a slow decline and stabilising afterwards. The seasonal part shows regular yearly cycles. The random part also represents irregular short-term variations that the trend or the seasonality cannot explain. This decomposition reveals the presence of trend and seasonality in the data, confirming that the data is not stationary and indicating the need to determine the best transform using the Box-Cox transformation, as shown in Figure 3.

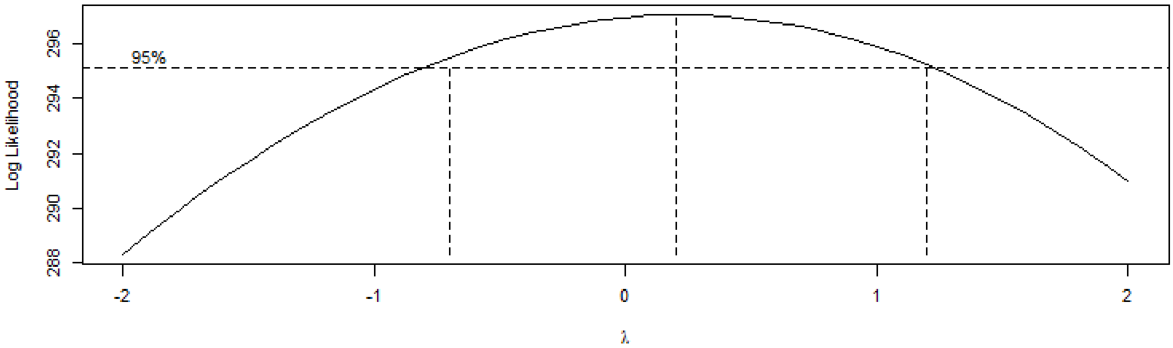

Figure 3: Box-Cox transform of the non-SADC African and Unspecified tourist arrivals to SA

The Box-Cox transformation plot in Figure 3 suggests a log transformation is necessary. The ADF for the log-transformed data is shown in Table 2.

Table 2: Augmented Dickey-Fuller test

Test Statistic	Log Transformed Data	First Difference of Log-Transformed Data
Dickey-Fuller	-2.8442	-4.8597
Lag Order	4	4
p-value	0.2266	0.01

The results in Table 2 show that the first difference of the log-transformed data is now stationary, as indicated by its corresponding p-value, which is less than 0.05. Therefore, a plot of the first difference of the log-transformed data is shown in Figure 4.

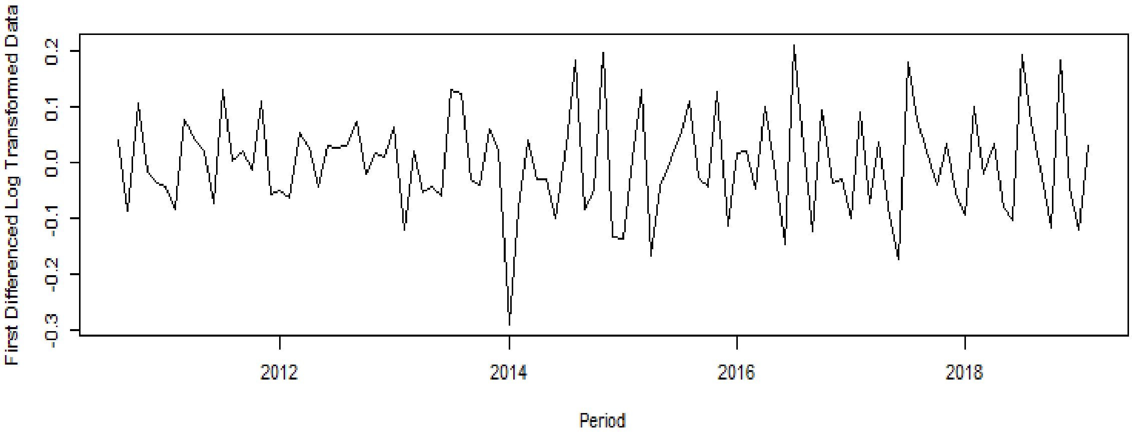

Figure 4: Time plot of the first difference of the log-transformed tourist arrivals data

The plot of the first difference of the log-transformed data in Figure 5 indicates a stationary series, as observations fluctuate around a zero mean, and is ready for model identification. An autocorrelation function (ACF) plot and a partial autocorrelation function (PACF) plot are plotted to ascertain the tentative model, as shown in Figure 5.

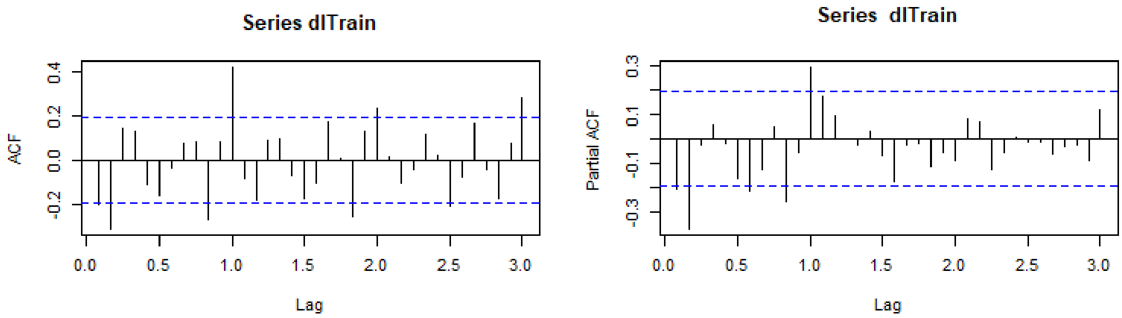

Figure 5: ACF and PACF plot of the first ordinary difference of the log-transformed data

The ACF and PACF plots in Figure 6 suggest an ARIMA(2,1,2) model, considering that the data has been differenced once and the ACF and PACF both cut off after lag 2. A seasonality component at lag 12 is also noted in the ACF plot and once in the PACF plot. An EACF is presented in Table 3.

Table 3: The EACF plot of the first differenced log-transformed data

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

An EACF in Table 3 cements the claims of the presence of the model's AR and MA coefficients (seasonal or non-seasonal). The EACF in Table 3 indicates the existence of models such as ARIMA $(0,1,2)(1,0,1)_{12}$, ARIMA $(1,1,2)(1,0,1)_{12}$, ARIMA $(2,1,1)(1,0,1)_{12}$ and ARIMA $(2,1,0)(1,0,1)_{12}$. The tentative and other potential models are fitted, and their results are tabulated in Table 4.

Table 4: Model fitting and adequacy checking for final model consideration

Model	AIC	RMSE	MAE	MAPE	MASE
ARIMA $(0,1,2)(1,0,1)_{12}$	-242.29	0.0676	0.0539	0.5531	0.4891
ARIMA $(1,1,2)(1,0,1)_{12}$	-244.26	0.0675	0.0539	0.5530	0.4891
ARIMA $(1,1,2)(1,0,1,)_{12}$	-243.85	0.0664	0.0527	0.5411	0.4788
ARIMA $(2,1,1)(1,0,1)_{12}$	-244.59	0.0660	0.0521	0.5349	0.4732
ARIMA $(2,1,0)(1,0,1)_{12}$	-246.56	0.0661	0.0528	0.5346	0.4729
ARIMA $(3,1,1)(1,0,1)_{12}$	-243.99	0.0655	0.0518	0.5313	0.4701
ARIMA $(1,1,0)(1,0,1)_{12}$	-236.71	0.0702	0.0562	0.5774	0.5107

The results in Table 4 highlight the values of various model adequacy measures. The better model considered in this paper is the one with the lowest AIC value. The model also gave the lowest RMSE, MAE, MAPE and MASE. The SARIMA$(2,1,0)(1,0,1)_{12}$ model and its corresponding error metric validate its use for baseline forecasting. There are no shocks in the training period. The model coefficients are computed as shown in Table 5.

Table 5: Model Parameters of the ARIMA $(\mathbf{2,1,0})(\mathbf{1,0,1})_{\mathbf{12}}$

	ϕ_1	ϕ_2	Φ_1	Θ_1
Coefficients	-0.4688	-0.3294	0.9522	-0.7161
Standard error (SE.)	0.0951	0.0929	0.0538	0.1577

The ARIMA $(2,1,0)(1,0,1)_{12}$ model coefficients in Table 5 are all significant, thus cementing the adoption of this model. Model adequacy checking was done on the ARIMA $(2,1,0)(1,0,1)_{12}$ residuals and shown in Figure 6.

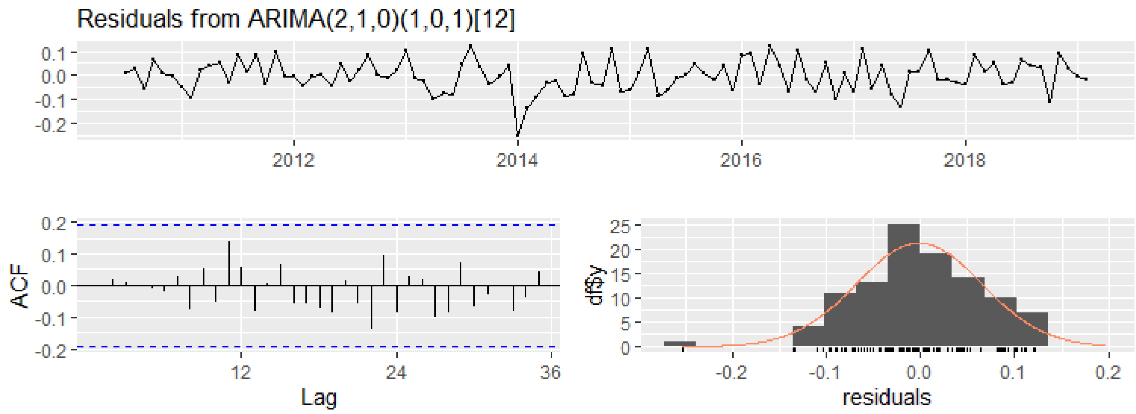

Figure 6: Residual plot of the ARIMA $(2,1,0)(1,0,1)_{12}$.

The residuals in Figure 6 behave like white noise, and the histogram of residuals indicates normality of residuals, thus proving the adequacy of the ARIMA $(2,1,0)(1,0,1)_{12}$ model for forecasting purposes. To confirm the arguments based on the visual picture in Figure 6, more formal tests were conducted to assess residual autocorrelation and normality using the Box-Ljung test and the Jarque-Bera test, respectively. The results of the Box-Ljung and the Jarque-Bera tests are shown in Table 6.

Table 6: Residual test results

Test Statistic	Box-Ljung Test	Jarque-Bera Test
X-squared	0.031944	4.9315
Df	2	2
p-value	0.9842	0.08495

The formal residual tests in Table 6 on the independence and normality of residuals further confirm the claim that the ARIMA $(2,1,0)(1,0,1)_{12}$ is a good fit for the data under study. A Box-Ljung test's p-value of 0.9842 > 0.05 indicates that there is no significant evidence to suggest the presence of autocorrelation in the residuals, and a p-value of 0.08495 > 0.05 also satisfies the normality assumption, thus confirming the adequacy of the model for use.

The expanded ARIMA $(2,1,0)(1,0,1)_{12}$ model for the non-SADC African countries and the Unspecified group is given as:

$$\hat{Y}_t = (1+\phi_1)\hat{Y}_{t-1} - (\phi_1 - \phi_2)\hat{Y}_{t-2} - \phi_2\hat{Y}_{t-3} + \Phi_1\hat{Y}_{t-12} - (\Phi_1 + \phi_1\Phi_1)\hat{Y}_{t-13} + (\phi_1\Phi_1 - \phi_2\Phi_1)\hat{Y}_{t-14} + \phi_2\Phi_1\hat{Y}_{t-15} + \hat{a}_t - \Theta_1\hat{a}_{t-12}. \quad (3)$$

$$\hat{Y}_t = 0.5312\hat{Y}_{t-1} + 0.1394\hat{Y}_{t-2} + 0.3294\hat{Y}_{t-3} + 0.9522\hat{Y}_{t-12} + 0.9152\hat{Y}_{t-13} - 0.1327\hat{Y}_{t-14} - 0.3137\hat{Y}_{t-15} + \hat{a}_t + 0.7161\hat{a}_{t-12}. \quad (4)$$

The train data from July 2010 to February 2019, forecasts made from the ARIMA $(2,1,0)(1,0,1)_{12}$ from March 2019 to February 2026, and the actual observations recorded on tourist arrivals from March 2020 to February 2024 are plotted as shown in Figure 7.

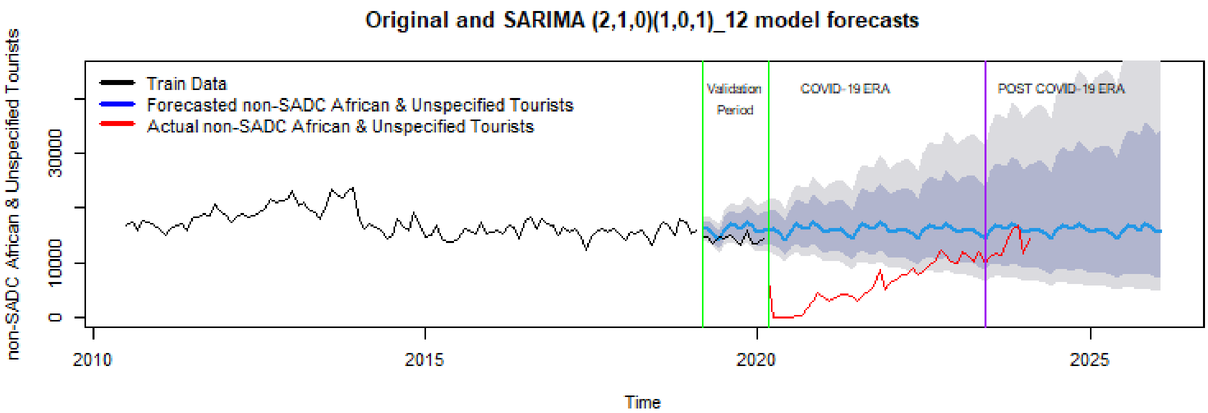

Figure 7: Plot of train, forecasts and validation data on non-SADC African and Unspecified tourist arrivals data

A plot of the train data, forecasts, validation and actual data is shown in Figure 7. Forecasts have been made till February 2026, making it an 84-month-ahead forecast. The COVID-19 pandemic era refers to the period in SA when the pandemic was considered a public health emergency, from March 2020 to 5 May 2023. The COVID-19 pandemic era saw various strict measures implemented across geographic divides to control human movement and contain the virus. The forecasts are used to assess the impact of the COVID-19 pandemic on combined tourist arrivals for the non-SADC African and Unspecified groups. The gap between the forecasts (blue line) and the actuals (red line) shows the effect of the containment measures after the COVID-19 era was declared in SA on combined non-SADC African and Unspecified tourist arrivals. The plot of the forecasts and actuals shows the negative impact of the COVID-19 pandemic containment measures. The negative impact of the COVID-19 pandemic on tourist arrivals is highlighted in the United Nations report (2021). There was an 84% decrease in international tourism between March and December 2020, and an overall 88% decline in global tourism. Forecasts play a pivotal role in future planning, budgetary control, and effective resource allocation. The significant increase in tourist arrivals from non-SADC African countries and the Unspecified combined group after the strict lockdown was lifted in SA around October 2020 indicates an upward trend. The trend offered hope for a full recovery in anticipated arrivals. This recovery is supported by Statistics SA (2023), which recorded 1.4 million same-day tourist visitors in 2022, when strict COVID-19 pandemic measures were relaxed, given the substantial drop in total tourist arrivals in SA from February 2020.

Discussion

This study investigated tourist arrivals to SA from the non-SADC African countries and the Unspecified group, to generate actionable insights to inform policy and strategic decision-making. Strengthening diplomatic and tourism relations with this region/s offers substantial potential to increase visitor inflows and enhance tourism revenues, as noted by UNWTO (2023). Tourism plays a vital role in SA’s economic development, contributing through foreign currency earnings, FDI, employment creation and global visibility (SAT, 2022). Accurate forecasting of tourist arrivals is indispensable for economic planning and policy formulation. This is particularly important because tourism demand is highly sensitive to external shocks such as political instability, natural disasters and pandemics like COVID-19, as highlighted by Gossling et al. (2021). The analysis of arrivals from 46 non-SADC African countries and the Unspecified group underscores the need for intensified research and marketing efforts. These efforts can facilitate more decision-making, support the development of new tourism products, and enable better-targeted marketing campaigns. This aligns with Yangailo (2025), who advocates multi-faceted governmental strategies to overcome barriers to tourism growth across Eastern and Southern Africa. SA could prioritise strengthening ties with non-SADC African countries and diversifying its tourism pull market. Such cooperation will help promote economic integration, shared infrastructure (such as airlines and routes), market access, and harmonised tourism policies, which are critical components for expanding SA’s tourism footprint on the continent, as postulated by the African Union (AU) (2020). Methodologically, this study used a SARIMA model on the COVID-19 pandemic data, divided into training, validation and test datasets. The discrepancies between the actual and forecast values during the testing phase are evidence of the pandemic’s impact. Beyond data modelling, technological advancement is crucial to modernise SA’s tourism infrastructure. With the introduction of biometric identification systems at ports of entry to improve security, additional innovations, such as real-time mobile applications that help with navigation, deliver site-specific content, and provide emergency alerts, can significantly enhance tourist experiences. Features such as location-based danger sensors and emergency response instructions, as exemplified by pressing a designated panic button in a specific threat scenario, can help ensure safety and accessibility, even for tourists without guides (Buhalis & Amaranggana, 2015). Such innovations not only foster a safer and more enriching tourist experience but also boost SA’s competitive edge in the global tourism market. By investing in smart tourism technologies and adopting data-driven policies, SA may attract even more tourist arrivals from underrepresented regions, such as non-SADC African countries, and increase resilience to disruptions, ensuring the development of a long-term sustainable tourism industry.

Conclusions

Based on study results and findings, it is recommended that SA increase its tourism engagement with the non-SADC African countries and the Unspecified group through targeted promotion and investigating high-growth markets, including medical, educational, and business travel, which are associated with longer stays and higher per capita spend (SAT, 2022; Naudé & Saayman, 2005). Enhancing air connectivity and simplifying visa requirements can significantly reduce travel barriers and encourage greater regional mobility, aligning with the African Union’s Agenda 2063 aspirations for intra-African integration and tourism-driven development (AU, 2020). The recovery of COVID-19 among arrivals focuses on building resilience through data-driven forecasting techniques such as SARIMA, which could identify seasonal patterns and account for pandemic disruptions (Goh & Law, 2011; Makoni et al., 2022). Also, the application of smart tourism technologies, such as biometric systems and mobile navigation and safety applications, can help maximise tourist satisfaction and aid crisis management, thereby making SA more competitive in the African tourism market (Buhalis & Amaranggana, 2015).

References

African Union. (2020). *The digital transformation strategy for Africa (2020–2030).* https://au.int-en-documents

Apergis, N., Gavriilidis, K., & Gupta, R. (2023). Does climate policy uncertainty affect tourism demand? Evidence from time-varying causality tests. *Tourism Economics, 29*(6), 1484–1498. https://doi.org/10.1177/13548166221110540

Boniface, B., Cooper, R., & Cooper, C. (2016). *Worldwide destinations: The geography of travel and tourism* (7th ed.). Routledge.

Box, G., Jenkins, G. M., & Reinsel, G. (1994). *Time series analysis: Forecasting and control* (3rd ed.). Prentice Hall.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control.* Holden Day.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control.* Wiley.

Buhalis, D., & Amaranggana, A. (2015). Smart tourism destinations are enhancing the tourism experience through personalisation of services. In I. Tussyadiah & A. Inversini (Eds.), *ENTER 2015 proceedings* (pp. 377–390). Springer.

Chu, F. (1998). Forecasting tourism: A combined approach. *Tourism Management, 19*(6), 515–520. https://doi.org/10.1016/S0261-5177(98)00053-3

Du Plessis, E., & Cronje, D. (2021). What makes South Africa competitive from a tourist's point of view? *Development Southern Africa, 38*(6), 919–937. https://doi.org/10.1080/0376835X.2020.1834355

Goh, C., & Law, C. H. R. (2011). The methodological progress of tourism demand forecasting: A review of related literature. *Journal of Travel and Tourism Marketing, 28*(3), 296–317. https://doi.org/10.1080/10548408.2011.562856

Gössling, S., Scott, D., & Hall, C. M. (2021). Pandemics, tourism and global change: A rapid assessment of COVID-19. *Journal of Sustainable Tourism, 29*, 1–20. https://doi.org/10.1080/09669582.2020.1758708

Gunter, U., & Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tourism Management, 46*, 123–135. https://doi.org/10.1016/j.tourman.2014.06.017

Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2021). Improving tourist arrival prediction: A big data and artificial neural network approach. *Journal of Travel Research, 60*, 998–1017. https://doi.org/10.1177/0047287520921244

Khairudin, S., Ahmad, N., Razali, A., Japeri, A. Z. U.-S. M., & Azmi, A. B. (2018). Forecasting international tourist arrivals in Penang using a time series model. *International Journal of Academic Research in Business and Social Sciences, 8*(16), 38–59. http://dx.doi.org/10.6007/IJARBSS/v8-i16/5117

Makoni, T., Mazuruse, G., & Nyagadza, B. (2022). International tourist arrivals modelling and forecasting: A case of Zimbabwe. *Sustainable Technology and Entrepreneurship.* https://doi.org/10.1016/j.stae.2022.100027

Naudé, W. A., & Saayman, A. (2005). Determinants of tourist arrivals in Africa: A panel data regression analysis. *Tourism Economics, 11*(3), 365–392. https://doi.org/10.5367/000000005774352962

South African Tourism. (2022). *Annual tourism report.* South African Tourism.

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

Sun, S., Wei, Y., Tsui, K. L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management, 70*, 1–10. https://doi.org/10.1016/j.tourman.2018.07.010

Tularam, G. A., Wong, V. S. H., & Nejad, S. A. H. S. (2010). Modeling tourist arrivals using time series analysis: Evidence from Australia. *Journal of Mathematics and Statistics, 8*(3), 348–360. https://doi.org/10.3844/jmssp.2012.348.360

Vergori, A. S., Colacchio, G., & Suppa, D. (2025). Accommodation statistics and unobserved inbound tourism: The Italian case study. *Italian Economic Journal.* https://doi.org/10.1007/s40797-025-00328-3

UNWTO. (2019). *World tourism barometer.* https://www.e-unwto.org/loi/wtobarometereng

UNWTO. (2021). *World tourism barometer.* https://www.e-unwto.org/loi/wtobarometereng

UNWTO. (2023). *Global and regional tourism performance.* https://www.unwto.org/tourism-data/global-and-regional-tourism-performance

Wang, S., Feng, J., & Liu, G. (2013). Application of seasonal time series model in precipitation forecast. *Mathematical and Computer Modelling, 58*(3–4), 677–683. https://doi.org/10.1016/j.mcm.2011.10.034

Wikasanti, D. R., Uqwatul, A. W., & Aidina, F. (2025). Forecasting international tourist arrivals in West Sumatra with SARIMA and triple exponential smoothing for post-pandemic tourism recovery. *Nuansa Informatika, 19*(1). https://doi.org/10.25134/nuansa

World Travel and Tourism Council. (2018). *Travel and tourism: Economic impact 2018.* WTTC.

World Travel and Tourism Council. (2020). *Global economic impact and trends 2020.* WTTC.

World Travel and Tourism Council. (2024). *Global economic impact and trends 2024.* WTTC.

Yangailo, T. (2025). The challenges of tourism growth in Eastern and Southern African countries. *Journal of Tourism and Heritage Research, 8*(1), 45–58.

Zhu, L., Lim, C., Xie, W., & Wu, Y. (2018). Modelling tourist flow association for tourism demand forecasting. *Current Issues in Tourism, 21*, 902–916. https://doi.org/10.1080/13683500.2016.1218827