The Economics and Finance Letters
2025 Vol. 12, No. 3, pp. 534-544
ISSN(e): 2312-430X
ISSN(p): 2312-6310
DOI: 10.18488/29.v12i3.4416

Forecasting the South African tax-to-GDP ratio series utilizing seasonal autoregressive integrated moving average and artificial neural networks models

Martin Chanza[1+]
Nkosinathi Emmanuel Monamodi[2]
Modisane Seitshiro[3]

(+ Corresponding author)

[1]*Department of Statistics and Operations Research, North-West University, Mmabatho, 2735, South Africa.*
[1]*Email: Martin.Chanza@nwu.ac.za*
[1,3]*National Institute for Theoretical and Computational Sciences, Potchefstroom, 2531, South Africa.*
[2]*Department of Economics and Finance, University of the Free State, Phuthaditjaba, 9869, South Africa.*
[2]*Email: MonamodiNE@ufs.ac.za*
[3]*Centre for Business Mathematics and Informatics, North-West University, Potchefstroom, 2531, South Africa.*
[3]*Email: Modisane.Seitshiro@nwu.ac.za*

Article History
Received: 10 March 2025
Revised: 28 July 2025
Accepted: 18 August 2025
Published: 26 September 2025

Keywords
ANN
SARIMA
Tax to GDP ratio.

JEL Classification:
C22; C45; C58.

ABSTRACT

South Africa has experienced successful tax collections but has not achieved the governance outcomes desired to establish an efficient fiscal contract compared to most developed countries. The objective of this study is to compare the performance of the conventional Seasonal Autoregressive Integrated Moving Average (SARIMA) model with that of the recently developed machine learning approach, the Artificial Neural Network (ANN) model in forecasting the South African Tax-To-GDP ratio. The focus is on accurately forecasting the South African tax-to-GDP ratio using the historical series from January 2008 to November 2024. The sampled period is characterized by random, irregular, and seasonal fluctuations, which are critical for accurate forecasting in the context of macroeconomic policy. The lower mean absolute percentage error indicates that the machine learning model outperformed the conventional time series model in terms of accuracy and reliability when forecasting South Africa's tax-to-GDP ratio. The findings further show that there will be slight growth in the tax-to-GDP ratio during the financial year of 2025, with a sharp decline forecasted between the end of 2025 and the beginning of 2026. These results add to the growing literature on the application of machine learning methods to economic forecasting. For policy considerations, this study suggests that South Africa's policy to expand its tax base, enhance tax administration efficiency, diversify revenue sources, and promote sustainable economic growth to minimize tax distortions and maintain macroeconomic stability during economic downturns.

Contribution/Originality: This study contributes to the existing literature by applying time series and neural network models to forecast South African tax revenue using the tax-to-GDP ratio within the South African context. The results demonstrate the effectiveness of neural networks, offering valuable insights for academic understanding and practical policy development.

1. INTRODUCTION

Tax revenue serves as the primary source of government funding by enabling the execution of various government functions and public expenditures (Tsaurai, 2021). That is, tax revenue is an integral tool that the government utilizes to promote economic growth and employment opportunities and further bridge the inequality gap between the rich and poor. Notably, South Africa's high levels of income and wealth inequality and contracting critical economic sectors force the government to rely heavily on direct and progressive taxes (Monamodi & Choga, 2021). Since 2020, tax rates have remained consistent despite economic shocks such as the COVID-19 epidemic and associated lockdowns, public unrest, rising prices for fuel and food, and power disruptions. South African Revenue Service (SARS) responded to these shocks with targeted temporary solutions. For instance, SARS (2022) indicates that overall tax collection has risen from 1,216.50 billion rand in 2017/18 to 1,563.80 billion rand in 2021/22, representing a 6.5% compound annual growth rate. The personal income tax, value-added tax, and corporate income tax continue to be the primary sources of tax revenue for the South African government, accounting for 81.2% of total collections (SARS, 2022). The tax-to-GDP ratio fell from 23.8% in 2019/20 to 22.3% in 2020/21 before rising to 24.9% in the 2021/22 fiscal year (SARS, 2022). The tax-to-GDP ratio increased further due to a 15 billion rand increase in tax for the year 2024/25 to alleviate immediate budgetary pressures and enable speedier debt stabilization (National-Treasury, 2024). The National-Treasury (2024) further indicates that nearly all the reported increments are due to direct taxes, with no tax rate increases. This means that the personal income tax is not increased by not adjusting tax brackets, rebates, and medical tax credits for inflation. The remaining increments are due to increased indirect taxes. As far as the fuel taxes are concerned, the National-Treasury (2024) report indicates that no increments are expected, leading to a potential revenue loss of 4 billion rand. This is largely offset by the inflation increases in excise duty on alcohol and some tobacco goods proposed in the 2024/25 budgetary framework (OECD, 2023).

The problem of forecasting errors in tax revenue for governments has attracted a lot of attention in the literature. In forecasting government tax revenue, Auerbach (1995) made a distinction between two categories of discrepancies, namely the economic and technical, based on behavioral. Fiscal policy inefficiencies are the major cause of tax policy faults. Williams and Kavanagh (2016) add that technical faults can happen because of unanticipated behavioral reactions or model misspecification, whereas economic inaccuracies are brought on by incorrect macroeconomic parameter estimates used in budget projections. According to Xu, Kayser, and Holland (2017) and Claudio, Heinisch, and Holtemöller (2020), inefficiencies can result from data inaccuracies and economic shocks that couldn't have been anticipated when the forecasts were prepared. Moreover, forecasting tax revenue is difficult in a volatile macroeconomic environment, yet it is crucial. Looking at the dire aftermath effects of the pandemic and the ambiguity in the forecasts of key economic factors like economic growth rate, investment, and unemployment, some uncertainty in forecasts is unavoidable. It is for this reason that this paper attempts to forecast the South African tax-to-GDP ratio series to inform policymakers on taxation and expenditure prospects.

This study contributes to the literature by articulating a comprehensive analysis that includes a conventional and machine learning technique for addressing fiscal policy fallacy that makes forecasting tax revenue difficult. This study investigates and forecasts the tax-to-GDP ratio, a topic that has not yet attracted enough attention in the national context. Furthermore, the results contribute to both the academic discourse and practical policy design by offering strategies for quantitatively forecasting the tax-to-GDP ratio in an emerging economy using the artificial neural network technique.

The aim of the research study is to assess the effect of seasonal autoregressive integrated moving average (SARIMA) and artificial neural networks (ANN) in accurately forecasting the South African tax-to-GDP ratio from January 2008 to November 2024. The remainder of the paper is organized as follows: Section 2 presents a literature review on government tax revenue forecasting. Section 3 outlines the methodology for forecasting the South African tax-to-GDP ratio series. Section 4 presents empirical results and provides an analysis. Finally, Sections 5 and 6 conclude the paper and discuss the limitations of the study and opportunities for further research.

2. LITERATURE REVIEW

The literature outlines several techniques used to forecast tax revenues. These techniques are based on macroeconomic approaches such as tax elasticity, tax buoyancy, Gross Domestic Product (GDP)-based models, Vector Auto Regressions (VARs), and macro simulation methods. For example, Favero and Marcellino (2005) applied

Autoregressive Moving Average (ARMA), Vector Autoregression (VAR), and Small Scale Structural Pooling (SSP) to forecast fiscal variables. The scholars found that simple time series or pooled models performed better than multivariate time series and small semi-structural models. In a similar comparison, Keene and Thomson (2007) found the same results using simple benchmark models where tax revenue was expressed as a product of a tax ratio. Using the two-step regression method (ECM), Koester and Priesmeier (2012) improved the German tax revenue forecasts. ECMs were also used by Corvalão, Samoryl, and Brasil (2010) to forecast VAT revenue in the state of Santa Catarina in Brazil, while Rudzkis and Maciulaityte (2007) forecasted the profit tax revenue series of Lithuania and Zhang states in China. The scholars proved the ECMs as better forecasting models compared to multivariate time series and small-structural models. Krol (2010) used the Bayesian method, which outperformed random walk forecasts and simple VARs. Botrić and Vizek (2012) used trend, random walk, ARIMA, and ECM models to forecast each component of tax revenue. The ARIMA model outperformed the random walk model and ECM.

The study by Koirala (2012) demonstrated that SARIMA models perform better in forecasting Nepal's tax revenue series. Galinski (2013) conducted a comprehensive analysis to compare forecasted values with actual values of local government revenue and expenditures in Poland from 2001 to 2011. The author distinguished between two subperiods: 2001–2008, characterized by underestimation due to incorrect forecasts of general subsidies, and 2009–2011, marked by overestimation resulting from overly optimistic revenue estimates. The study reported forecast accuracy with a mean absolute percentage error (MAPE) of 3.90% for total revenue, 4.53% for total expenditures, and 20.59% for capital expenditures.

In the other research study, seven forecasting techniques were used to forecast the yields of thirteen sources and five aggregates of St. Petersburg's tax revenue in Florida, including the Box-Jenkins technique, single exponential smoothing, the Holt technique, the moving average, linear regression, general adaptive filtering, and the Winters technique (Gianakis & Frank, 1993). The authors have demonstrated that the techniques of moving average and adaptive filtering achieve the highest accuracy scores. Additionally, the Box and Jenkins MAPE of 10.30% is smaller compared to other techniques, while the Winters MAPE is the highest at 42.36%.

Guo and Chen (2021) used a non-linear method known as artificial neural network (ANN), which provided small forecast errors for forecasting Iran's tax revenue time series. This study is among those that employed a different family of forecasting techniques.

Other research studies suggest using sensible expectations, analysis of variance, the Kruskal-Wallis test, root mean squared error (RMSE), mean absolute error (MAE), and mean absolute scaled error (MASE) to evaluate the efficiency of forecasting techniques (Chanza, De Koker, Boucher, Munapo, & Mabuza, 2023; Mbuli, Mathonsi, Seitshiro, & Pretorius, 2020; Rapoo & Xaba, 2017; Singh & Balasundaram, 2007).

The above literature has provided a vast insight into forecasting techniques and their assessment performance error metrics. This research study uses the MAE and MAPE performance error metrics to evaluate the performance of SARIMA and ANN when forecasting the South African tax-to-GDP ratio. Thus, it is done specifically taking into account the severe effects of the 2007 to 2009 financial crisis and the 2020 COVID-19 pandemic, which had an impact on South African public finances. Other factors contributing to the impact may include the implementation of significant expansionary policy actions on government current/operational expenditures and the loss of an important portion of revenue due to job cuts, economic restrictions, and weak economic performance. Furthermore, this research study has discovered that the literature and scientific research on the South African tax-to-GDP ratio series are limited, despite its importance in preventing underfunding and excessive funding, which can lead to unsustainable budget deficits.

3. METHODOLOGY

In this section, the methods of Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Artificial Neural Network (ANN) are presented. These forecasting techniques are used to forecast the South African tax-to-GDP ratio monthly series covering the period from January 2008 to December 2024.

The source of the data series is the South African Reserve Bank (SARB) database. The sampled time series period was chosen mainly because of the volatile behavior of South African tax revenue relative to the country's GDP (Fall, 2022). It is crucial to forecast the South African tax-to-GDP ratio because of several macroeconomic factors, such as economic growth, trade openness, inflation, and unemployment, that influence it.

3.1. The ARIMA Model

The ARIMA model is a linear combination of past observations, past errors, and their differences for homogeneous nonstationary time series. It is known to be flexible when analyzing and forecasting low-frequency types of time series data (Seitshiro, 2006).

The ARIMA model is significant for identifying current trends and predicting future patterns, especially when the data shows repeating cycles and is influenced by its past values. The ARIMA model combines three components given by Autoregressive (AR), Integrated (I), and Moving Average (MA). This model is statistically represented as ARIMA (p, d, q), where the AR component with order p captures the relationship between an observation and its previous values. The Integrated component with order d involves differencing the time series data to achieve stationarity. The MA component with order q models the relationship between an observation and the residual errors from a moving average of past observations (Tsay, 2005). Notably, an ARIMA model is a theoretical model, not based on any economic theory, and can be estimated using the Box-Jenkins approach for calculating p, d, and q.

The Box-Jenkins approach for the ARIMA model estimation process involves four phases. The first phase is identification, which aims to find acceptable values for p, d, and q, using the autocorrelation function (ACF) and partial autocorrelation function (PACF) while maintaining the intermediate autocorrelation between lags (Brooks, 2008). The second phase involves estimating the model's parameters. The third phase assesses the goodness of fit of the chosen ARIMA model, usually assessing whether the residuals follow a white noise process. For residuals that do not follow a white noise process, one should repeat steps one, two, and three with new values for p, d, and q. If the residuals are white noise, the model can be accepted and proceed to phase four, forecasting desirable periods for the time series (Chanza et al., 2023; Gujarati, 2009).

Time series that show seasonal and irregular patterns can be modeled using a conventional ARIMA model, which is modified to include a seasonality component. This is statistically referred to as the SARIMA model and denoted by ARIMA(p, d, q)×(P, D, Q)s, where P and Q respectively represent seasonal autoregressive and moving average orders, D represents the order of seasonal difference, and s represents the number of seasonal cycles (Ebhuoma, Gebreslasie, & Magubane, 2018).

The ARIMA or SARIMA model must be diagnosed and selected among other time series models by determining the smallest Akaike's Information Criterion (AIC). According to Khumalo, Mashele, and Seitshiro (2023), an AIC is used to determine the best model for forecasting among other models and evaluate its effectiveness for both in- and out-of-sample forecasting. The AIC is defined as

$$AIC = -2 \times \ln(L) + 2P, \quad (1)$$

Where L represents the likelihood function of the series, and P represents the sum of p and q.

3.2. The ANN Model

The Artificial Neural Network (ANN) models have become essential tools in various scientific fields because of their ability to model complex relationships and reduce computational time (Bungane, Botha, & Van Der Vyver, 2024).

Khashei and Bijari (2010) allude that the ANN model is a model that learns and becomes wise to execute different applications like optimization, prediction, modeling, clustering, pattern recognition, and simulation. It is made up of an input layer for data collection, hidden layers for processing the information and connecting the input with the output, and an output layer for computed information. According to Chen, Leung, and Daouk (2003) and Singh and Balasundaram (2007), a neuron is a basic processing unit responsible for gathering inputs and constructing output, multiplied by connection weights, products, and biases, and then passed through an activation function. A multilayer feedforward network is a type of ANN that can pass in one direction without feedback.

The backpropagation (BP) process is a supervised learning technique that is commonly employed for training feedforward artificial neural network (ANN) models (Al-Gahtani, Alsugair, Alsanabani, Alabduljabbar, & Almohsen, 2025). Al-Gahtani et al. (2025) indicate that the BP process runs iteratively to learn samples and uses a parameter estimation method known as the gradient descent optimization technique to estimate the parameters of the model. To minimize the mean squared error (MSE), weights are adjusted for each training model, comparing predicted and actual values.

In this research study, the statistical measurements based on relative forecasting errors include the mean percentage error (MPE) and mean absolute percentage error (MAPE). Montgomery, Keats, Runger, and Messina (1994) discovered that the employment of MPE and MAPE are easily compared since they are assessed in percentage, contrary to certain other metrics that may be scaled differently due to varying characteristics. The values of both measures must be at a minimum compared to the others for the forecasting model to be reliable in-sample and out-of-sample forecasts (Rapoo & Xaba, 2017). The following equations present the mathematical representations of MPE and MAPE, respectively, as

$$MPE = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{Y_t - \hat{Y}_t}{Y_t}\right) \times 100\% \qquad (2)$$

and

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \times 100\%, \qquad (3)$$

where *n* denotes sample size, Y_t is the actual variable at time *t*, and $\hat{Y}_t$ is the forecasted variable at time *t*.

Although ANN models have advantages, such as allowing complex non-linear relationships between forecasts, the ANN models present challenges, particularly in terms of interpretability when the dimensions of variables increase to more than two. While the model delivers superior accuracy, the nature of artificial neural networks means that it is harder for non-mathematical scientists to explain the underlying reasons behind the forecasted values. This lack of transparency could be a concern for policymakers who need to understand the rationale behind decisions based on model forecasts. This is concurred by Fall (2022) and further recommends the development of techniques and competent quantitative experts to explain the rationale for their outputs. In contrast, the SARIMA model, although less accurate, provides clearer insights into the relationships between the observed variables, which could be useful for understanding fiscal dynamics in a more interpretable way. Thus, the ANN model is the simplest among the other groups of neural network models. This model is chosen because of its ability to model complex relationships, reduce computational time, and require no underlying assumptions about the time series data.

4. RESULTS

The findings of this research study concerning the SARIMA and ANN models are presented and interpreted below. Estimating the desired SARIMA and ANN models begins with visualizing the characteristics of the data. For this reason, the following section analyzes the descriptive statistics for the tax-to-GDP ratio series.

Table 1. Descriptive statistics.

Variable	Min	Max	Mean	Std. Dev
Tax-to-GDP ratio	21.89	26.01	24.09	1.07

4.1. Descriptive Statistics Analysis

Table 1 presents the summary statistics of the Tax-to-GDP ratio, which has a minimum value of 21.89, a maximum of 26.01, a mean of 24.09, and a standard deviation (Std. Dev.) of 1.07. The standard deviation represents the spread of the data around the mean. The standard deviation is small, indicating very little spread around the mean.

4.2. Modelling Results

The first step is to plot the series to observe its stationarity and primary components of the time series. Hence, Figure 1 presents the graphical plot of the tax-to-GDP ratio series.

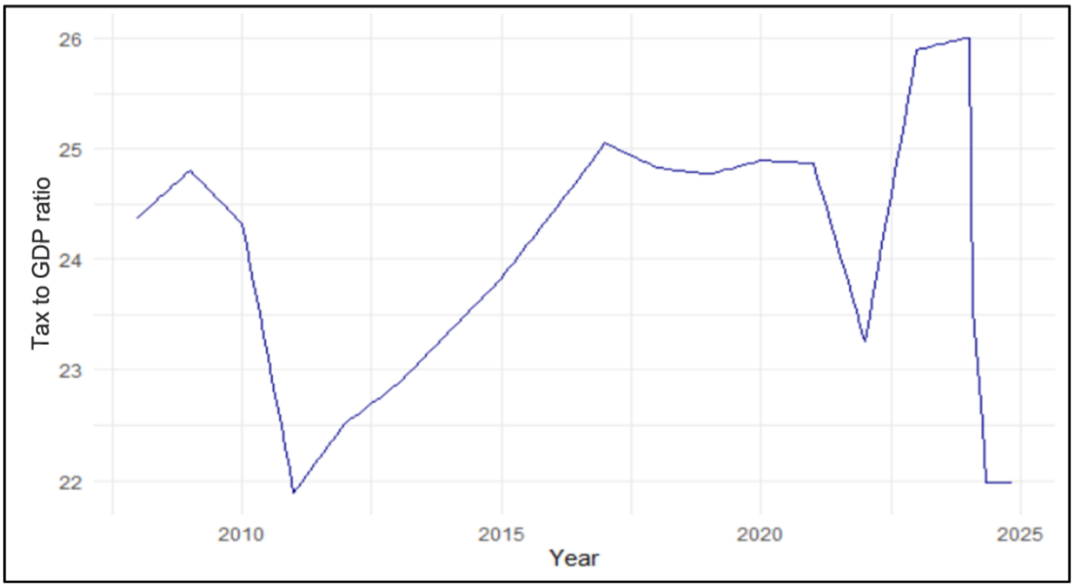

Figure 1. Tax-to-GDP ratio series (January 2008 – November 2024).

The series plotted in Figure 1 shows random, irregular, and seasonal fluctuations, leading to an estimation of the SARIMA model. However, the series needs to be detrended or differenced first using a unit root test. Hence, Table 2 presents the results for the Augmented Dickey Fuller (ADF) unit root test performed at 'intercept and trend' specification.

Table 2. ADF unit root test results.

Variable	Specification	*t* Stat	Critical Value	P (*t* Stat)	Order of integration
Tax-to GDP ratio	Intercept and trend	-6.3321	-3.4823*** -1.4600** -0.9428*	0.0000*** 0.0000** 0.0000*	*I (1)*

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively.

The results presented in Table 2 indicate that the tax-to-GDP ratio series is stationary after being differenced once. This is because the probability value for the corresponding t-statistic is zero, which is less than the 1% significance level.

Therefore, the tax-to-GDP ratio is an I(1) variable. Having determined the I component of the SARIMA model, due to visible random irregular and seasonal fluctuations, Table 3 presents the estimated coefficients for AR and MA components (for the ARIMA part), and seasonal AR, I, and MA.

Table 3. Estimated coefficients for the study's SARIMA model.

Series: Train					
SARIMA components	**AR (1)**	**MA (1)**	**MA (2)**	**MA (3)**	**SMA (1)**
Estimate	0.5502	-0.2070	0.1516	0.1499	0.7096
SE	0.1998	0.2081	0.0915	0.0979	0.1360
σ^2 = 0.03116 Log likelihood = 61.83					
AIC = -111.65 BIC = -91.8					

Note: AR and MA denote autoregressive and moving average, respectively. SMA and SE denote seasonal moving average and standard error, respectively. σ^2 is the estimated variance of the model.

Using the estimated coefficients provided in Table 3, and the SARIMA model standard form of ARIMA (p, d, q)×(P, D, Q)s, the SARIMA model of the study fitted on the series is denoted as ARIMA (1, 1, 3)×(0, 0, 1)12. Similarly, Table 4 presents the estimated results for the study's ANN model.

Table 4. Results for the study's ANN model.

Series: Train
Model: NNAR (12,1,6)[12]
Call: nnetar (y = Train)
Average of 20 networks, each of which is a 12-6-1 network with 65 weights
Options: Linear output units
σ^2 = 0.01668

4.3. Model Accuracy

This section assesses the performance of the estimated SARIMA and ANN models by comparing the forecasting accuracy. Thus, comparing the Mean Percentage Error (MPE) and Mean Absolute Percentage Error (MAPE) of the estimated models. The smaller the MAPE, the better the model's performance (Galeshchuk, 2016). Hence, Table 5 presents the values of the MPE and MAPE matrices for the estimated models.

Table 5. MPE and MAPE values for the estimated SARIMA and ANN models.

Model	**MPE**	**MAPE**
SARIMA	-0.0161	0.2497
ANN	-0.0026	0.1174*

Note: * Indicates the model selected by the MAPE matrix.

The results shown in Table 5 indicate that the ANN model with a lower MAPE of 0.1174 outperformed the SARIMA model with a MAPE of 0.2497 in terms of accuracy. Notably, negative values of MPE suggest underprediction (Montgomery et al., 1994).

4.4. Forecasting with the Best Model

Having selected the study's ANN model as the best forecasting model, Figure 2 presents the training data and forecast series for the South African tax-to-GDP ratio over time using the ANN model.

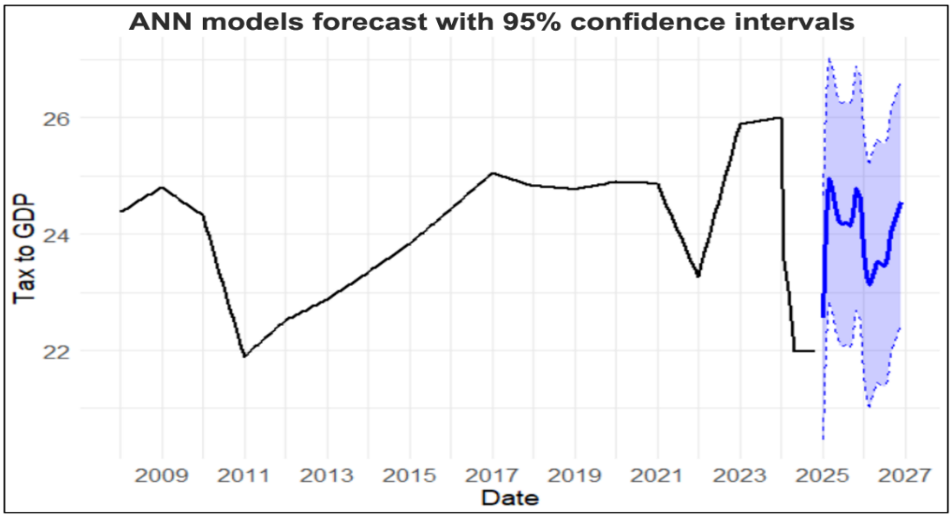

Figure 2. Training data and forecast series for the ANN model.

Figure 3 presents the forecast of the South African tax-to-GDP ratio series for the next two years (2025 to 2026), at a 5% confidence interval.

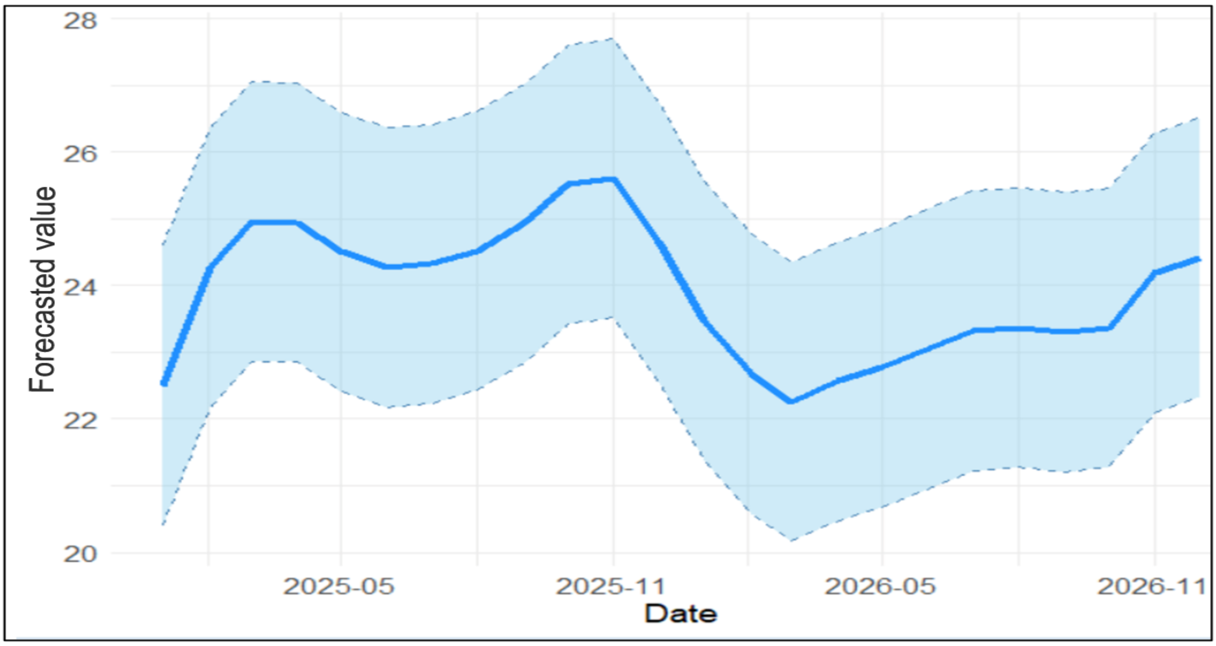

Figure 3. The South African tax-to-GDP ratio forecast for 2025 – 2026.

At a 5% confidence interval, Figure 3 shows a slight forecasted growth in the South African tax-to-GDP ratio in the year 2025. However, a sharp forecasted decline is expected at the end of 2025. Thereafter, a steady increase is forecast during the period 2026, which suggests possible economic instability.

5. CONCLUSIONS, DISCUSSION, AND POLICY SUGGESTIONS

The purpose of this study was to forecast the South African tax-to-GDP ratio series, a critical indicator of government fiscal health and policy effectiveness, particularly in light of the ongoing challenges posed by economic volatility and shocks. The South African government's reliance on tax revenue for redistributive functions, public goods, and economic stability highlights the importance of accurate forecasting. Tax revenue plays a central role in promoting economic growth, addressing inequality, and stabilizing the fiscal landscape, especially during periods of economic downturns such as the 2008 to 2009 financial crisis, the 2018 technical recession, and the 2020 COVID-19 pandemic. The study also illustrates the broader challenge of fiscal policy forecasting in emerging economies like South Africa, where issues of inequality, economic stagnation, and reliance on progressive taxes create a complex fiscal landscape. Policymakers must contend with the unpredictability of these factors, which makes accurate forecasting not just difficult but essential for informed decision-making. By using advanced forecasting techniques such as SARIMA and ANN, the South African government can better anticipate future tax revenues, enabling more effective budget planning and debt management.

The study's comparative analysis of the SARIMA and ANN models reveals significant insights. Both models were able to capture the seasonal fluctuations inherent in the tax-to-GDP ratio series, but the ANN model demonstrated a clear advantage in terms of forecasting accuracy. The ANN model's lower MAPE suggests that it is better equipped to handle the complexities and nonlinear relationships in the data compared to the SARIMA model, which assumes linearity. This is consistent with some of the presented previous studies that highlight the advantage of machine learning techniques, such as artificial neural networks, in forecasting economic indicators with inherent nonlinearities.

The out-of-sample (2025 to 2026) forecast shows that there will be growth in the tax-to-GDP ratio in South Africa in the year 2025, and a sharp decline between 2025 and 2026. For this reason, the South African government may reduce volatility in the tax-to-GDP ratio by expanding its tax base, improving tax administration efficiency, diversifying its revenue sources, and promoting sustainable economic growth, to minimize tax distortions and maintain macroeconomic stability during economic downturn events such as the 2008/9 financial crisis, the 2018 technical recession, and Covid-19.

6. LIMITATIONS AND OPPORTUNITIES FOR FURTHER RESEARCH

The interpretability of machine learning models remains a key challenge. Policymakers may face difficulties in understanding the rationale behind the ANN model's forecasts, which could limit its practical application in fiscal decision-making. Therefore, future research may explore ways to enhance the transparency of machine learning models, potentially integrating them with traditional models to leverage their strengths in both accuracy and interpretability.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Al-Gahtani, K. S., Alsugair, A. M., Alsanabani, N. M., Alabduljabbar, A. A., & Almohsen, A. S. (2025). ANN prediction model of final construction cost at an early stage. *Journal of Asian Architecture and Building Engineering*, *24*(2), 775-799. https://doi.org/10.1080/13467581.2023.2294883

Auerbach, A. J. (1995). Tax projections and the budget: Lessons from the 1980's. *American Economic Review*, *85*(2), 165–169.

Botrić, V., & Vizek, M. (2012). Forecasting fiscal revenues in a transition country: The case of Croatia. *Zagreb International Review of Economics & Business*, *15*(1), 23-36.

Brooks, C. (2008). *RATS Handbook to accompany introductory econometrics for finance*. Cambridge: Cambridge University Press.

Bungane, T., Botha, C., & Van Der Vyver, C. (2024). Developing a conceptual framework for addressing the accuracy of forecasting models in predicting platinum group metals prices. *Kurdish Studies*, *12*(5), 865–876.

Chanza, M., De Koker, L., Boucher, S., Munapo, E., & Mabuza, G. (2023). *Applying machine learning techniques to forecast demand in a South African fast-moving consumer goods company*. Paper presented at the International Conference on Intelligent Computing & Optimization.

Chen, A.-S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers & Operations Research*, *30*(6), 901-923. https://doi.org/10.1016/S0305-0548(02)00037-0

Claudio, J. C., Heinisch, K., & Holtemöller, O. (2020). Nowcasting East German GDP growth: A MIDAS approach. *Empirical Economics*, *58*(1), 29-54. https://doi.org/10.1007/s00181-019-01810-5

Corvalão, E. D., Samoryl, R. W., & Brasil, G. H. (2010). Forecasting the collection of the state value added tax (ICMS) in Santa Catarina: The general to specific approach in regression analysis. *Brazilian Journal of Operations & Production Management*, *7*(1), 105-121.

Ebhuoma, O., Gebreslasie, M., & Magubane, L. (2018). A seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. *South African Medical Journal*, *108*(7), 573–578.

Fall, F. (2022). *Strengthening the tax system to reduce inequalities and increase revenues in South Africa*. OECD Economic Department Working Papers, No. (1745).

Favero, C. A., & Marcellino, M. (2005). Modelling and forecasting fiscal variables for the Euro area. *Oxford Bulletin of Economics and Statistics*, *67*(s1), 755-783. https://doi.org/10.1111/j.1468-0084.2005.00140.x

Galeshchuk, S. (2016). Neural networks performance in exchange rate prediction. *Neurocomputing*, *172*, 446-452. https://doi.org/10.1016/j.neucom.2015.03.100

Galinski, P. (2013). The accuracy of the budget forecasting in local governments in Poland. *Economics and Management*, *18*(2), 218-225.

Gianakis, G. A., & Frank, H. A. (1993). Implementing time series forecasting models: Considerations for local governments. *State & Local Government Review*, *25*(2), 130-144.

Gujarati, D. N. (2009). *Basic econometrics* (5th ed.). New York: McGraw-Hill.

Guo, H., & Chen, C. (2021). Forecasting Revenue Impacts from COVID-19: The Case of Florida Municipalities. *State and Local Government Review*, *53*(1), 78-88.

Keene, M., & Thomson, P. (2007). *An analysis of tax revenue forecast errors*. Treasury Working Paper No. 07/02. New Zealand Treasury.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications*, *37*(1), 479-489. https://doi.org/10.1016/j.eswa.2009.05.044

Khumalo, M., Mashele, H., & Seitshiro, M. (2023). Quantification of the stock market value at risk by using FIAPARCH, HYGARCH and FIGARCH models. *Data Science in Finance and Economics*, *3*(4), 380-400. https://doi.org/10.3934/DSFE.2023022

Koester, G., & Priesmeier, C. (2012). *Estimating dynamic tax revenue elasticities for Germany*. Deutsche Bundesbank Discussion Paper No. 23/2012. Deutsche Bundesbank.

Koirala, T. (2012). Government revenue forecasting in Nepal. *NRB Economic Review*, *24*(2), 47-60.

Krol, R. (2010). *Forecasting state tax revenue: A Bayesian vector autoregression approach*. Working Paper No. 2010-04. Department of Economics, California State University, Northridge.

Mbuli, N., Mathonsi, M., Seitshiro, M., & Pretorius, J.-H. C. (2020). Decomposition forecasting methods: A review of applications in power systems. *Energy Reports*, *6*(Supplement 9), 298-306. https://doi.org/10.1016/j.egyr.2020.11.238

Monamodi, N., & Choga, I. (2021). An estimation of the effect of tax revenue collection on the public debt in South Africa. *Review Pub Administration Manag*, *9*, 292-306.

Montgomery, D. C., Keats, J. B., Runger, G. C., & Messina, W. S. (1994). Integrating statistical process control and engineering process control. *Journal of Quality Technology*, *26*(2), 79-87.

National-Treasury. (2024). *Budget review: Trends and tax proposals. National Treasury, Republic of South Africa.* Retrieved from https://www.treasury.gov.za/documents/National%20Budget/2024/review/Chapter%204.pdf

OECD. (2023). *Revenue statistics in Africa: Key findings for South Africa.* Paris, France: Organization for Economic Cooperation and Development.

Rapoo, M. I., & Xaba, D. (2017). A comparative analysis of artificial neural network and autoregressive integrated moving average model on modeling and forecasting exchange rate. *World Academy of Science, Engineering and Technology International Journal of Economics and Management Engineering, 11*(11), 2504-2507.

Rudzkis, R., & Maciulaityte, E. (2007). Econometrical modelling of profit tax revenue. *Nonlinear Analysis, 12*(1), 95-112. https://doi.org/10.15388/NA.2007.12.1.14724

SARS. (2022). *Tax statistics. South African Revenue Service.* Retrieved from https://www.sars.gov.za/about/sas-tax-and-customs-system/tax-statistics/

Seitshiro, M. B. (2006). Two-sample comparisons for serially correlated data. Master's Thesis. North-West University, South Africa.

Singh, R., & Balasundaram, S. (2007). Application of extreme learning machine method for time series analysis. *International Journal of Intelligent Technology, 2*(4), 256-262.

Tsaurai, K. (2021). Determinants of tax revenue in upper middle-income group of countries. *The Journal of Accounting and Management, 11*(2), 148-160.

Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). Hoboken, NJ, United States: John Wiley & Sons.

Williams, D. W., & Kavanagh, S. C. (2016). Local government revenue forecasting methods: Competition and comparison. *Journal of Public Budgeting, Accounting & Financial Management, 28*(4), 488-526. https://doi.org/10.1108/JPBAFM-28-04-2016-B004

Xu, Q., Kayser, H., & Holland, L. (2017). Forecast errors: Balancing the risks and costs of being wrong. In K. G. Willoughby (Ed.), Government budget forecasting: Theory and practice. In (pp. 479-496). New York, United States: Routledge.