

Modeling and Forecasting Somalia's Economic Growth: A Comparison of Traditional Time Series and Modern Neural Network Approaches

Abdimajid Ali Abdi

cmajiid09@gmail.com

NATIONAL ECONOMIC COUNCIL

Abdihakim Ahmed Hassan

Zamzam University of Science and Technology

Shukri Yusuf Salad

Somalia National Bureau of Statistics

Research Article

Keywords:

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11209588/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Modeling and Forecasting Somalia's Economic Growth: A Comparison of Traditional Time Series and Modern Neural Network Approaches.

Abdimajid Ali Abdi 1 (*Corresponding Author), Abdihakim Ahmed Hassan 2 , Shukri Yusuf Salad 3

1National Economic Council, Federal Republic of Somalia, Mogadishu, Somalia.

2Faculty of Business & Management Sciences, Zamzam University of Science and Technology, Mogadishu, Somalia.

3National Bureau of Statistics, Federal Republic of Somalia, Mogadishu, Somalia.

*Corresponding Author and Email: cmajiid09@gmail.com

Abstract

Reliable forecasts of economic growth can inform fiscal planning and resource allocation, particularly in economies exposed to substantial year-to-year fluctuations. This study compares the out-of-sample forecasting performance of autoregressive integrated moving average (ARIMA), exponential smoothing (ETS), and neural network Autoregression (NNAR) models for Somalia's annual GDP growth rate. The analysis uses 64 annual observations from the World Bank's World Development Indicators for 1961–2024. Each model was fitted using data from 1961–2014, and forecasts for 2015–2024 were generated from a single forecast origin at the end of 2014. Forecast accuracy was assessed using mean absolute error (MAE), root mean squared error (RMSE), and mean absolute scaled error (MASE). ETS(A,N,N) achieved the lowest errors across all three measures (MAE = 3.8308; RMSE = 4.6641; MASE = 0.5771), followed by ARIMA(0,1,1). NNAR (2,2) produced the highest errors. These findings indicate that the level-only ETS model performed best among the three approaches under the evaluation design used in this study. The comparison provides evidence to guide the selection of a forecasting method for Somalia's annual GDP growth, while the small test sample and single forecast origin limit conclusions about performance in other periods. More dependable growth forecasts may support economic planning, although forecasting accuracy alone does not establish development outcomes.

Introduction

Economic growth forecasts help governments anticipate changes in revenue, employment, and demand when preparing budgets and development plans. Macroeconomic forecasting research therefore examines not only how well a model explains past observations, but also how useful its predictions are for decisions made before future outcomes are known.

Producing reliable forecasts is difficult because economic series may contain irregular fluctuations and may change after events that have little precedent in the historical data. A model fitted closely to earlier observations can consequently perform poorly when asked to predict later ones. Forecast

evaluation must preserve the order of time, keep test observations outside model estimation, and assess errors over a clearly defined forecast horizon(Hewamalage, Ackermann et al. 2023)

Forecasting methods differ in how they use historical information. ARIMA represents dependence on earlier observations and forecast errors, while Exponential Smoothing updates estimates as new observations arrive; neural network methods can learn more flexible relationships between past and future values. Yet a flexible model has more patterns to learn, which can be difficult when observations are limited. Research on small time-series datasets identifies data scarcity and overfitting as concerns when applying machine learning methods(Alexandrino, Pacheco et al. 2026)

The choice between simple and complex methods is therefore an empirical question. In a comparison covering ARIMA, ETS, and several machine learning approaches, Beck et al. (2025) found that model performance varied with the predictability of the series and emphasized the value of assessing forecasts against a simple benchmark. Their findings support evaluating competing methods on the same observations instead of assuming that a newer method will be more accurate(Beck, Dovert et al. 2025).

Somalia provides a pertinent setting for this assessment. According to the World Bank (2026), growth moderated to an estimated 3% in 2025, from about 4% in 2023–2024, as drought, reduced foreign aid, and rising living costs weighed on economic activity. These pressures increase the need for forecasts that can inform planning while making it harder to rely on stable historical patterns (World Bank, 2026)

Against this background, the present study compares ARIMA, Exponential Smoothing (ETS), and Neural Network Autoregression (NNAR) in forecasting Somalia’s annual GDP growth. By applying the three methods to the same series and evaluating their predictions over the same test period with consistent accuracy measures, it seeks to identify which approach is most useful for this particular forecasting task(Beck, Dovert et al. 2025).

1.2 problem statement

Reliable forecasts of economic growth support budgeting, resource allocation, and investment decisions. Producing such forecasts is challenging in Somalia, where economic activity is exposed to recurrent shocks and changing conditions. These difficulties make it important to establish which forecasting method predicts annual GDP growth most accurately when applied to the available historical data.

ARIMA, Exponential Smoothing (ETS), and Neural Network Autoregression (NNAR) use past observations in different ways. ARIMA models temporal dependence, ETS updates forecasts as new observations become available, and NNAR can learn more flexible relationships. None can be assumed to perform best for Somalia’s GDP-growth series: a model that fits historical observations closely may still make large errors when forecasting later years.

The specific problem addressed by this study is the limited direct evidence on the relative out-of-sample accuracy of ARIMA, ETS, and NNAR for Somalia’s annual GDP growth under identical

evaluation conditions. The study therefore fits the three methods to the same training series and compares their predictions over the same test period using consistent accuracy measures. This comparison provides an empirical basis for selecting a method for the forecasting task examined here.

2.2 Theoretical literature

2.2.1 Box–Jenkins Framework and ARIMA

The Box–Jenkins framework examines how a variable's current value relates to its past values and past random disturbances (Box and Jenkins 1968). ARIMA applies this framework through autoregressive terms, differencing when needed, and moving-average terms (Hyndman and Khandakar 2008). The autoregressive terms capture relationships with earlier observations, while the moving-average terms use information from earlier errors; differencing can address non-stationarity where appropriate. For this study, ARIMA provides a way to assess whether linear relationships in Somalia's past annual GDP-growth rates can generate useful forecasts of later growth rates.

The Box–Jenkins approach also involves identifying a suitable model, estimating its parameters, checking its residuals, and producing forecasts. These checks help determine whether the specified model has adequately captured the historical relationships it is intended to represent. Nevertheless, a satisfactory fit to the estimation data does not, by itself, establish how accurately the model will predict later observations (Tashman 2000).

2.2.2 Exponential Smoothing Framework.

Exponential smoothing forecasts a time series by updating estimates as new observations become available, while the influence of older observations declines over time (Gardner Jr 1985). The framework includes specifications that represent different features of a series, such as its level, trend, and seasonality; Holt's extension, for example, accommodates a trend (Holt 2004). The appropriate specification depends on the characteristics of the data, making model selection an important part of applying exponential smoothing (Billah, King et al. 2006).

For this study, exponential smoothing offers an alternative to ARIMA for forecasting Somalia's annual GDP growth. Whereas ARIMA explicitly models relationships with past values and errors, exponential smoothing updates its estimates as successive GDP-growth observations become available. Comparing their forecasts can show which approach makes more effective use of the historical series over the study's evaluation period.

2.2.3 Neural Network Autoregression

NNAR uses previous observations of a time series as inputs to a neural network. During training, the network adjusts the weights connecting its computational units to reduce the difference between its predictions and the observed values. Neural networks can represent nonlinear relationships without requiring the researcher to specify their exact form beforehand (Hornik, Stinchcombe et al. 1989). This capacity provides a reason to examine whether NNAR can identify

predictive patterns in Somalia's GDP-growth series that are difficult to represent with the statistical approaches.

Greater flexibility also requires care in selecting and evaluating the network. A network that learns the estimation observations too closely may perform less well when forecasting observations it has not seen, particularly when the historical series is short (Qi and Zhang 2001). NNAR's potential advantage must therefore be assessed through its performance on the test data rather than inferred from its ability to fit the training data.

2.2.4 Integration of the Forecasting Approaches

ARIMA, exponential smoothing, and NNAR differ in how they turn Somalia's historical GDP-growth observations into forecasts. ARIMA represents linear temporal dependence, exponential smoothing updates estimates as observations arrive, and NNAR can learn more flexible relationships. Forecasting research shows that methodological complexity does not guarantee greater accuracy, so no approach should be presumed best before its predictions are evaluated (Makridakis, Spiliotis et al. 2018).

These approaches provide the basis for the study's direct comparison. Each model is applied to the same annual GDP-growth series, and its forecasts are assessed over the same test period using consistent accuracy measures. Out-of-sample accuracy provides the basis for deciding which method performs best for the forecasting task examined in this study (Tashman 2000).

2.3 Empirical literature

One model

To examine whether past growth patterns could predict Somalia's economic performance, quarterly GDP growth data from 1960 to 2022 were analyzed using the Box–Jenkins ARIMA method. ARIMA (5,1,2) was selected, and the model projected average quarterly growth of approximately 4% over the next 14 quarters (Mohamed 2022).

Somalia's unemployment trends were examined through a Box–Jenkins time series analysis covering 1991 Q1 to 2020 Q1. After differencing the data to achieve stationarity, ARIMA (1,1,3) was selected using AIC and BIC; residual diagnostics supported its use for short-term forecasting (Mohamed, Abdirahman et al. 2026).

To understand how income inequality changed over time in Malaysia, researchers used 29 annual observations from 1987 to 2015 to forecast the Gini coefficient. Their optimized NAR–MLP model produced a very low mean squared error of 1.14×10^{-7} , and its residuals showed no apparent correlation (Amin Megat Ali, Zabidi et al. 2024).

Growing pressure on the Juba River Basin's water resources prompted an assessment of future water needs using the WEAP model and a 2002–2022 reference period. The model showed satisfactory performance (NSE = 0.64; $R^2 = 0.87$) and projected domestic water shortfalls of 5.79%

by 2030 and 24.1% by 2050. Raising irrigation efficiency from 45% to 60% reduced the projected shortages (Dhaqane, Murshed et al. 2026).

Across the Horn of Africa, changes in vegetation and land surface temperature provided a way to track emerging drought. Drawing on Landsat observations from 2000 to 2022, the Eco-Phase approach identified worsening drought in southern Somalia, central Ethiopia, and Eritrea. An LSTM model also predicted the drought score one year ahead ($r = 0.689$; $RMSE = 0.0935$) (Abdelrahim and Jin 2026).

Mohamed (2022) and Mohamed, Abdirahman et al. (2026) used ARIMA to forecast Somalia's GDP growth and unemployment, respectively, showing its application to different economic indicators. In contrast, Amin Megat Ali et al. (2024) used a neural network to estimate income inequality, while the Juba River Basin and Horn of Africa studies applied models to water allocation and drought. Together, these studies show how model choice varies with the outcome, data, and forecasting purpose.

Two models

Climate conditions were linked to Somalia's economic growth in an analysis of 31 annual observations from 1993 to 2023. After interpolating missing values and normalizing data from the World Bank, IMF, and climate databases, the researchers compared Random Forest and Bayesian Regression. Higher temperature and CO₂ emissions were associated with lower GDP, while rainfall was associated with higher GDP; Random Forest performed better (Osman, Mohamud et al. 2025).

Across eight East African countries, including Somalia, economic growth, agricultural activity, capital formation, and energy consumption were found to be significantly associated with environmental change. These findings place the relationship between economic growth and the environment within a wider regional context (Mwanga 2025).

Forecasting Bahrain's GDP with approximately 60 quarterly observations from 2010 to 2024, the researchers compared ARIMA and Support Vector Regression (SVR) using MSE, MAE, and MAPE. SVR produced the more accurate forecasts on all three measures, including a MAPE of 4.30% compared with ARIMA's 5.19% (Moussa, Mustafa et al. 2025).

Vietnamese pepper exports to the United States provided a test of two forecasting approaches using 288 monthly observations from January 2002 to December 2025. The researchers compared SARIMA–GARCH with a gated recurrent unit (GRU) model using error measures and statistical tests. GRU performed better, although the reported summary gives no numerical accuracy values (Quang, Thinh et al. 2026).

Kenya's public debt was forecast using 168 monthly observations from January 2010 to December 2023. ARIMA gave the more accurate forecasts for domestic debt ($RMSE = 0.0265$) and total debt ($RMSE = 0.0571$), while Holt's model held a slight advantage for external debt (Obwoye 2025).

Colombia's unemployment rate provided a setting for testing whether complex economic patterns could be made easier to interpret and predict. Drawing on data from 2005 to 2023, the researchers

combined UMAP based local biplots with Gaussian process regression. The approach identified influential economic variables and improved prediction and interpretability compared with conventional methods (Pérez-Rosero, Manrique-Cabezas et al. 2025).

Forecasts of shark and skipjack tuna catches were developed from historical catch records and remote sensing data. When the two approaches were compared, functional networks produced smaller prediction errors than artificial neural networks for both species. The available summary does not specify the observation years (Iglesias, Arcay et al. 2004).

The studies reach different conclusions about which forecasting approach works best. For Somalia's GDP, Random Forest outperformed Bayesian Regression; SVR produced more accurate GDP forecasts than ARIMA in Bahrain, and GRU surpassed SARIMA–GARCH when forecasting Vietnamese pepper exports. Kenya's debt forecasts offer a contrast: ARIMA performed best for domestic and total debt, while Holt's model did slightly better for external debt. Functional networks also outperformed artificial neural networks for both fish species studied. Taken together, these results show that model performance varies with the outcome and dataset, making a direct comparison valuable for forecasting Somalia's economic growth (Osman, Mohamud et al. 2025; Moussa, Mustafa et al. 2025; Quang, Thinh et al. 2026; Obwoye 2025; Iglesias, Arcay et al. 2004).

Three models

Somalia's GDP forecasts improved when three approaches were tested against 19 annual observations from 2005 to 2023. Among Random Forest, XGBoost, and Prophet, Random Forest produced the smallest forecast errors and the highest reported R^2 (0.89). SHAP analysis highlighted current account balance, GDP per capita, and population as influential predictors (Osman and Muse 2024).

Forecasting Somalia's GDP growth with just 32 annual observations poses a challenge for data intensive models. Comparing ANN, CNN, SVR, Random Forest, and XGBoost over 1991–2022, the researchers found that ANN and CNN gave the most accurate forecasts, while Random Forest and XGBoost performed less well with the small sample (Ibrahim, Mohamed et al. 2026).

Somalia's unemployment rate was forecast from 31 annual observations spanning 1991–2021, using economic and sectoral indicators. Of the four models compared—Random Forest, XGBoost, SVR, and ARIMA—Random Forest performed best ($R^2 = 0.92$), while ARIMA performed least well. SHAP analysis identified labor force participation, foreign direct investment, and inflation as the most influential predictors (Mohamed and Ibrahim 2026).

Somalia's cereal production was forecast from 63 annual observations covering 1961–2023. Across eight ARIMA, ETS, and NNETAR specifications, NNETAR achieved the lowest reported SMAPE (40.42%), while ARIMA and ETS showed greater forecasting bias. The comparison suggests that the neural network approach handled fluctuations in this agricultural series more effectively (Ahmed, Abdisamad et al. 2026).

Using 64 annual observations from 1961–2024, the study compared 19 models for forecasting Somalia's meat production. ARFIMA performed best among seven individual models, while

ARIMA–NNAR led the twelve hybrid models. Forecasts for 2025–2036 suggest that production may remain stable or decline slightly (Abdi 2026).

Somalia's deforestation was examined using 34 annual observations from 1990–2023. Gradient Boosting performed best among three models. GDP per capita, urbanization, and population growth were associated with greater deforestation, while renewable energy was associated with less (Osman, Hassan et al. 2026).

Drought forecasting across 55 Ethiopian districts drew on 474 months of satellite data. The researchers compared stochastic, conventional machine learning, and deep learning approaches across 11 locations. Deep learning performed best throughout, with reported R^2 values ranging from 0.42 to 0.95, suggesting potential for earlier warnings and better humanitarian planning (Woldmariam 2025).

Somalia's unemployment outlook was examined using 34 annual observations from 1991 to 2024. Of six standalone and ten hybrid forecasting models, ARFIMA led the standalone models, while ETS–NNAR achieved the best overall accuracy. The forecasts place unemployment at roughly 18.83–18.85% during 2025–2030 (Abdullahi Hassan, Farah Darod et al. 2026).

Somalia's urban population growth was forecast using 62 annual observations from 1961–2022. The researchers compared three standalone models with three hybrids. TBATS led the standalone models, while ARIMA–TBATS–NNAR achieved the best overall accuracy and projected continued growth through 2033 (Dahir, Korse et al. 2025).

Somalia's agricultural land per person was forecast from 63 annual observations covering 1961–2023. After stationarity testing, six models were compared using 27 observations for validation; ARIMA (0,2,3) fitted the full sample best, while Theta gave the most accurate validation forecasts (RMSE = 0.274; MAPE = 5.96%). The projections suggest a continuing decline in agricultural land per person through 2050 (Abdi, Ali et al. 2026).

Forecasting the decline in Somalia's agricultural land per person, Ibrahim (2026) compared six models using annual data from 1961–2023. ARIMA (0,2,3) fitted the historical data best, but Theta produced the most accurate forecasts on the validation data (RMSE = 0.274; MAPE = 5.96%). The forecasts suggest a decline to 2.30 hectares in 2024, with land per person approaching critically low levels by 2050 (Ibrahim 2026).

To examine the future of maternal health in Somalia, researchers compared 16 forecasting models using 39 annual observations from 1985–2023. ARIMA (0,1,1) with drift delivered the most accurate forecasts (MAPE = 4.38%) and projected a gradual decline in the maternal mortality ratio from 2024 to 2030 (Yousuf and Kahie 2026).

Cancer mortality in Somalia was forecast from 42 annual observations spanning 1980–2021. After comparing six standalone and six hybrid models, the researchers found that ARIMA (2,2,1) gave the most accurate predictions (MAPE = 1.127%; RMSE = 100.16), outperforming a naïve

benchmark. Its forecasts suggest that annual cancer deaths could rise to approximately 9,652 by 2030 (Seiman, Abdi et al. 2026)

Using quarterly data, the study applied Box–Jenkins ARIMA models to forecast urbanization and electricity consumption in Somalia. ARIMA (4,1,2) was best for urbanization, while ARIMA (5,1,2) was best for electricity consumption. Both variables were projected to increase through 2030, although the abstract does not report the sample period or number of observations (Mohamed, Farah et al. 2026).

Forecasting neonatal mortality in Somalia, researchers tested six standalone and ten hybrid models against 41 annual observations from 1983–2023. NNAR produced the most accurate forecasts (MAPE = 4.84%), suggesting that neonatal mortality may remain near 35 deaths per 1,000 live births during 2024–2030, although the forecasts become less certain further into the future (Mahamoud Yousuf, Mohamed Kahie Seiman et al. 2026).

To forecast dengue in Somalia, researchers evaluated an SIR model alongside eight standalone and twelve hybrid forecasting models, using annual observations from 1990–2021 and reserving 2017–2021 for testing. The ARIMA–TBATS hybrid performed best (MAPE = 6.49%; RMSE = 66.81), outperforming TBATS, the strongest standalone model. The SIR model also suggested that dengue transmission could persist ($R_0 = 1.015$) (Abdi, Nadarajah et al. 2026).

Road fatalities in Somalia rose from 3,770 in 1997 to 56,247 in 2017–2026, according to an analysis of 30 annual observations. Comparing six standalone and nine hybrid models, the researchers found that ARIMA–ETS–Theta gave the most balanced forecasts (sMAPE = 0.3697; RMSE = 2,551.36). Its projections indicate that fatalities could remain high through 2036 (Abdullahi, Muse et al. 2026).

Women’s unemployment in Somalia was examined through 32,310 observations from the 2020 Somalia Health and Demographic Survey. Among seven models, Random Forest performed best (AUC = 98.6%), followed by LightGBM (94.1%) and XGBoost (84.1%). Age, region, household size, and wealth status emerged as the strongest predictors (Osman, Esse et al. 2026).

Across 39 African countries, researchers examined population growth forecasts using a recurrent neural network (RNN) and three approaches to handling missing data. The available abstract does not identify the observation years or report forecast errors, so it does not establish which missing-data approach performed best (Appiah Takyi 2024).

Somalia’s growth was examined over 1989–2023 using ARDL and DOLS, with the analysis finding a long-run relationship among the variables. Public spending and trade openness were associated with higher growth, while external debt was associated with lower growth; foreign direct investment showed no significant effect (Mohamed, Dirie et al. 2026).

Somalia’s rainfall patterns showed marked variation across space and time in an analysis of CRU TS data from 1901–2021. The distance-weighted spatial autoregressive (SAR) model captured

these patterns better than OLS ($R^2 = 0.45$ versus 0.40), supporting the value of accounting for geographic clustering in rainfall analysis(Alasow, Hersi et al. 2025).

Monthly data from 2009–2024 offered a test of five forecasting approaches: VARIMA, Prophet, RNN, LSTM, and GRU. VARIMA(2,0) performed best, while the RNN and LSTM models showed signs of overfitting. The available summary does not report the number of monthly observations or numerical forecast errors (Karagöz, Güler et al. 2026).

What shapes residential property prices? An analysis of 118 properties compared hedonic regression with an artificial neural network using five-fold cross-validation. Size, bedrooms, proximity to the central business district, safety, age, and air quality all influenced prices. The neural network captured nonlinear patterns more effectively, reducing mean squared error by about 20%, while hedonic regression explained 74.2% of price variation (Nor and Hussein 2025).(Nor and Hussein 2025).

Farm data collected in real time in Balcad District were used to test three models for recommending crops. Decision Tree achieved the highest accuracy (99.2%), ahead of Random Forest (99.0%) and KNN (97.2%), and also performed well on precision, recall, and F1 score. The available summary does not report the number of observations (Abdullahi, Jimale et al. 2024).

Conflict's human toll was examined across seven countries, including Somalia, using 54 variables spanning 32 years. Five statistical methods were applied to forecast battle deaths, displacement, genocide, and undernourishment. The available abstract does not identify a best-performing method or provide numerical accuracy results (Tannehill 2008).

School attendance among Somali children varied sharply by age, region, and place of residence. Drawing on 10,511 children aged 6–18 in the 2022 SHBS, researchers compared ten predictive models; Random Forest performed best (AUC = 0.86). Adolescents and children in Middle Shabelle and nomadic communities faced particularly high risks of being out of school (Ali, Cumar et al. 2025).

Forecasting vegetation biomass in Kenya, researchers compared six models across four forecast horizons using five series measured every 15 days over 20 years. SARIMA held a slight overall advantage over CNN and CNN–LSTM, but no model performed best for every series or horizon (Noa-Yarasca, Osorio Leyton et al. 2024).

Forecasting six international stock indices, researchers tested LSTM, XGBoost, and a hybrid LSTM–XGBoost model using 1,259 daily observations from 1 March 2018 to 1 March 2023. After tuning the models with grid search, the hybrid delivered the most accurate forecasts across the markets. The available abstract does not report numerical forecast errors(Oukhouya 2024).

GDP growth in Albania and Serbia was forecast with an ANN–MLP model trained on 70% of each country's 31 annual observations from 1993–2023. The model recorded an error below 0.3% and identified exports as important in Albania, while consumption and foreign direct investment stood out in Serbia(Caushi 2025).

India's summer monsoon rainfall was forecast from 46 years of data spanning 1950–1995. Comparing an artificial neural network (ANN), persistence forecasting, and multiple linear regression, the researchers found that the three-layer ANN performed best. The available summary does not give numerical accuracy results (Chattopadhyay 2007).

To improve quarterly GDP forecasts, researchers built an EBA–BPNN hybrid using World Bank and OECD data. After selecting 32 economic variables through dynamic time warping, Granger causality, and mutual information, the hybrid reduced mean absolute error by 29.3%, 17.8%, and 15.6% relative to BPNN, PSO–BPNN, and BA–BPNN, respectively. The available abstract does not specify the study years or observation count (Mou, Gan et al. 2025).

Across six Brazilian cities, researchers used 2007–2022 data to forecast leishmaniasis outbreaks with SFNN, DFNN, LSTM, and SVR models. They assessed forecast errors using RMSE and found that weather conditions and the timing of model inputs affected predictive accuracy. The available abstract does not identify an overall best model or report the observation count and error values (Donizette, Rocco et al. 2025).

Can complex models predict conflict more accurately? Using 51 country-year observations from 2009–2015 and seven indicators of conflict intensity, researchers compared LDA, CART, KNN, Random Forest, and neural networks including LSTM. The neural networks did not consistently outperform simpler methods on this dataset, and the available abstract gives neither a single best model nor numerical accuracy results (Ettensperger 2020).

A comparison of 12 forecasting models shows how much the choice of method mattered across roughly 204 monthly observations from 2008 to 2024. Before the pandemic, exponential smoothing gave the most accurate forecasts, with a MAPE of 6.80%. After the pandemic, rolling SARIMA performed best, though its MAPE rose to 18.50%. Rolling forecasts also cut mean absolute error by more than 80% compared with direct forecasts (Luo 2026).

GDP growth proved difficult to capture with a single approach across Germany and Poland. Comparing linear and machine learning models with data from 1991 to 2023, the researchers found that nonlinear methods performed better. XGBoost led in Germany, while a decision tree performed best in Poland; the strongest predictors also differed between the two countries (Valiyev, Batrancea et al. 2025).

Forecasting food security across Ethiopia's regions, Demssie, Assabi et al. (2026) combined Random Forest, XGBoost, LSTM, and CNN in a stacking model. Using monthly regional data, the model achieved a mean absolute error of 0.00201, an RMSE of 0.038, and an R^2 of 0.96. These results suggest strong predictive performance, although the abstract does not specify the years covered by the data (Demissie, Assabie et al. 2026).

The reviewed studies show that forecasting performance varies across outcomes, sample sizes, and evaluation designs. In studies of Somalia's GDP, Random Forest performed best in one comparison, whereas ANN and CNN performed best in another (Osman and Muse, 2024; Ibrahim, Mohamed et al., 2026). Studies forecasting other Somali indicators also reported different leading approaches, including traditional, neural, and hybrid models. These results provide no basis for

assuming that a more complex model will consistently produce more accurate forecasts. They also cannot be ranked directly across studies that use different series, test periods, and accuracy measures.

Although previous studies have compared forecasting models for Somalia's GDP and other indicators, the literature reviewed does not establish how ARIMA, ETS, and neural network Autoregression (NNAR) perform against one another on the same annual Somali GDP-growth series under a common evaluation procedure. This study addresses that specific gap by applying the three methods to the same training data and comparing their forecasts over the same test period using MAE, RMSE, and MASE.

3.0 Methodology

3.1 Research Design

This study employs a quantitative time-series research design to model and forecast Somalia's economic growth. It uses a comparative forecasting approach to assess the performance of two traditional time-series techniques the Autoregressive Integrated Moving Average (ARIMA) model and exponential smoothing—against a modern Artificial Neural Network (ANN) model. The quantitative design is appropriate because the study analyzes numerical economic-growth observations collected over time and applies statistical and computational methods to identify historical patterns and predict future movements. The study is primarily predictive rather than causal, as it aims to identify the model that produces the most accurate forecasts instead of examining the factors responsible for changes in economic growth.

A univariate forecasting framework is adopted, meaning that future economic growth is predicted solely from its previous values. To ensure a fair comparison, ARIMA, exponential smoothing, and ANN will be applied to the same dataset using identical training and testing periods and the same forecasting horizon. ARIMA will capture linear patterns and dependence across time, exponential smoothing will emphasize recent observations and changes in the series' level and trend, while ANN will learn possible nonlinear and more complex patterns. The models will be compared mainly through their performance on unseen observations using suitable forecast-accuracy measures. This consistent evaluation process will establish whether traditional time-series methods or the modern neural-network model provides the most reliable forecasts of Somalia's economic growth.

3.2 Data Source, Measurement

Variable	Notation	Measurement	Source
Economic growth	gt	annual percentage growth rate of real GDP."	World bank

The study uses Somalia's annual real GDP growth rate to measure economic growth. The data come from the World Bank and are expressed as a percentage: each observation shows how much inflation-adjusted GDP changed compared with the previous year. A positive value indicates that

real economic output increased, while a negative value indicates that it declined. This series is the variable forecast by the ARIMA, ETS, and NNAR models.

3.4 ARIMA Model

The study employs the Autoregressive Integrated Moving Average, ARIMA (p, d, q), model to identify and forecast the linear dynamics of Somalia's economic-growth series. The autoregressive order p measures the influence of previous growth observations on the current value, while d represents the number of differences required to make the series stationary. The moving-average order q captures the effects of previous random shocks or forecasting errors. The general ARIMA model is expressed as:

$$\varphi(B)(1 - B)^d Y_t = c + \theta(B)\varepsilon_t$$

where Y_t denotes Somalia's economic-growth rate at time t, B is the lag operator, and $(1 - B)^d$ represents the differencing process. Furthermore, $\varphi(B)$ denotes the autoregressive polynomial, $\theta(B)$ represents the moving-average polynomial, c is the constant term, and ε_t is a random error term assumed to have a zero mean and constant variance.

Several alternative ARIMA specifications will be estimated and compared to identify the model that best represents the historical behaviour of the series. The final specification will be selected primarily using the corrected Akaike Information Criterion (AICc), while also considering parameter significance and model parsimony. Residual diagnostic tests, particularly the Ljung–Box test, will be applied to determine whether the model has adequately captured the temporal structure of the data. An ARIMA specification will be considered suitable when its residuals behave like white noise and show no significant serial correlation, after which it will be used to produce out-of-sample forecasts.

3.5 Exponential Smoothing Model

The study applies exponential smoothing to forecast Somalia's economic growth by assigning greater importance to recent observations while progressively reducing the influence of older values. Because the series is annual, only nonseasonal specifications will be considered, including simple exponential smoothing, Holt's linear-trend model, and the damped-trend model. Where a trend is supported by the data, the general Holt specification is expressed as follows:

$$\ell_t = \alpha Y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1}$$

$$\hat{Y}_{t+h|t} = \ell_t + hb_t$$

where Y_t is Somalia's observed economic-growth rate, ℓ_t denotes the estimated level, b_t represents the estimated trend, and α and β are the smoothing parameters for the level and trend, respectively. The term h is the forecast horizon, while $\hat{Y}_{t+h|t}$ represents the forecast produced at time t for period t + h.

The smoothing parameters and initial states will be estimated from the training sample by minimizing the model's forecast errors. Alternative exponential smoothing forms will be compared using the corrected Akaike Information Criterion (AICc), together with model parsimony and residual adequacy. The selected specification must leave no meaningful temporal pattern in its residuals before it is used to generate out-of-sample forecasts. This procedure avoids imposing a trend in advance and allows the observed behaviour of Somalia's economic-growth series to determine the most appropriate exponential smoothing structure.

3.6 Artificial Neural Network Model

The study used a feedforward artificial neural network in the form of neural network Autoregression, NNAR(p, k), to forecast Somalia's annual GDP growth. This univariate model uses p previous growth observations as inputs, passes them through one hidden layer with k neurons, and produces a forecast of the next value. Its ability to represent nonlinear patterns makes it a useful comparison with ARIMA and ETS. Because the annual series is relatively short, model complexity requires particular care. The general specification is:

$$\hat{Y}_t = \beta_0 + \sum_{j=1}^k \beta_j g(\alpha_{0j} + \sum_{i=1}^p \alpha_{ij} Y_{t-i})$$

where $\hat{Y}_t$ is the forecast GDP growth rate in year t ; Y_{t-i} is the observed growth rate i years earlier; p is the number of input lags; and k is the number of hidden neurons. The function $g(\cdot)$ is the hidden-layer activation function, α_{ij} and β_j are network weights, and α_{0j} and β_0 are bias terms.

The NNAR structure was selected automatically using the 1961–2014 training series, resulting in NNAR(2,2): two input lags and two hidden neurons. The model was fitted using the training data and then used to generate forecasts for 2015–2024 from the end of 2014. Observations from the test period were used to assess forecast accuracy, not to select the NNAR structure.

3.7 Training and Testing Procedure

The study used the original annual GDP-growth series comprising 64 observations from 1961–2024, with no conversion to quarterly frequency. To preserve the chronological structure of the time series, the first 54 observations (1961–2014) were assigned to the training sample, while the remaining 10 observations (2015–2024) were reserved for out-of-sample testing. No random splitting was applied, as this would disrupt the temporal order of the data. The ARIMA, ETS, and NNAR models were therefore estimated using the same training period and evaluated consistently on the unseen test sample using MAE, RMSE, and MASE. Using the models fitted through 2014, forecasts for 2015–2024 were generated from a single forecast origin, without incorporating observed test-period values.

3.8 Forecast-Accuracy Measures and Model Selection

The out-of-sample forecasting performance of the ARIMA, ETS, and NNAR models was evaluated using three standard accuracy measures: Mean Absolute Error (MAE), Root Mean

Squared Error (RMSE), and Mean Absolute Scaled Error (MASE). These measures assess forecast performance from complementary perspectives and are defined as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$$

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

$$MASE = \frac{\frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|}{\frac{1}{T-1} \sum_{t=2}^T |Y_t - Y_{t-1}|}$$

where Y_t denotes the observed GDP-growth value, $\hat{Y}_t$ represents the corresponding forecast value, n is the number of observations in the test sample, and T is the size of the training sample. MAE measures the average absolute forecast error, RMSE gives greater weight to larger forecast errors, and MASE evaluates forecast accuracy relative to a naïve benchmark. Lower values of all three measures indicate better forecasting performance; therefore, the preferred model was identified as the one that consistently produced the lowest MAE, RMSE, and MASE values on the out-of-sample test period.

4.0 Results and Discussion

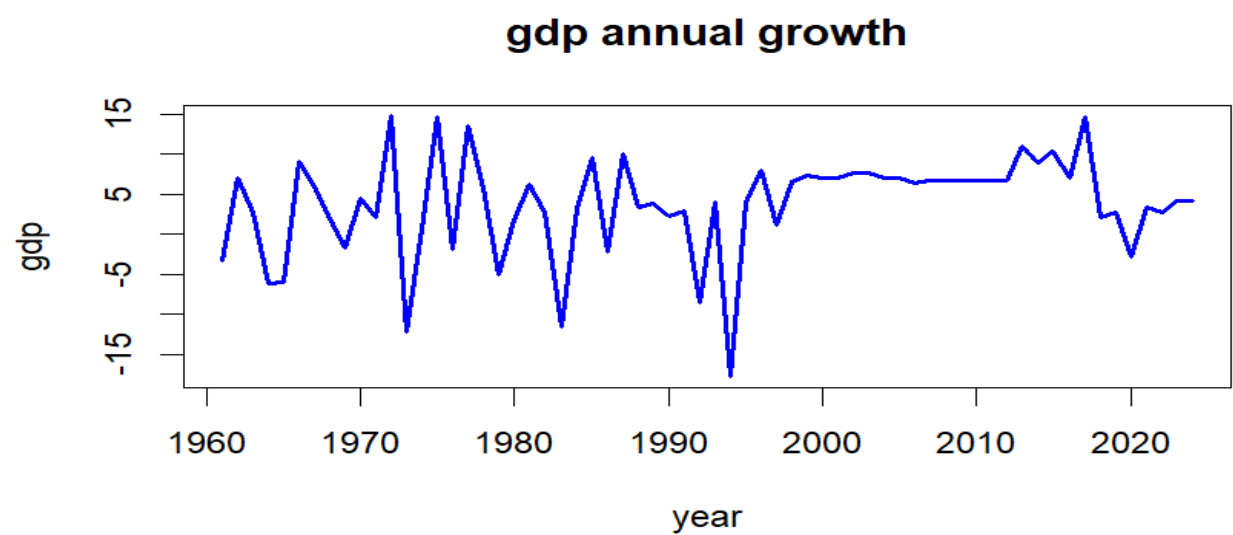
This section presents the empirical findings from the comparative forecasting analysis of Somalia's economic growth using Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing (ETS), and Neural Network Autoregression (NNAR) models. Consistent with the methodological framework, the 64 annual observations covering 1961–2024 were divided chronologically into training and testing samples to preserve the temporal ordering of the series and prevent information leakage. The first 54 observations, spanning 1961–2014, were used for model identification, estimation, and training, while the remaining 10 observations covering 2015–2024 were reserved exclusively for out-of-sample forecast evaluation.

The empirical analysis begins with the ARIMA framework, including an assessment of the statistical properties of the training series, stationarity testing, model identification, parameter estimation, model selection, and residual diagnostic evaluation, followed by out-of-sample forecasting. The same training and testing structure is subsequently applied to the ETS and NNAR models to ensure a consistent basis for comparison. Forecasting performance across the three competing approaches is evaluated primarily using the Mean Absolute Error (MAE), with the Root Mean Square Error (RMSE) and Mean Absolute Scaled Error (MASE) employed as complementary accuracy measures. The preferred forecasting model is ultimately determined by its performance on the unseen test sample, thereby emphasizing genuine out-of-sample predictive ability rather than relying solely on in-sample goodness of fit.

4.1 Autoregressive Integrated Moving Average (ARIMA) model

The Autoregressive Integrated Moving Average (ARIMA) model was applied as the first forecasting approach to capture the linear dynamics embedded in Somalia’s economic-growth series. The model combines autoregressive terms, differencing, and moving-average components to represent persistence, remove non-stationarity when necessary, and account for the influence of past forecast errors. Model identification and selection were guided by the statistical properties of the training series, with the final specification chosen on the basis of parsimony and information criteria before being subjected to residual diagnostic testing and out-of-sample forecast evaluation.

4.1.1 Historical Dynamics and Volatility of Somalia’s Annual GDP Growth, 1961–2024



The GDP growth series exhibits pronounced instability from the early 1960s through the mid-1990s, with periods of strong economic expansion frequently followed by sharp contractions. Several observations fall into negative growth territory, while others record substantial positive increases, highlighting the considerable year-to-year volatility that characterized Somalia’s earlier economic performance. The magnitude and frequency of these fluctuations suggest repeated macroeconomic disturbances rather than a smooth and persistent growth trajectory, reflecting substantial instability in the historical evolution of economic activity.

From the late 1990s onward, GDP growth became comparatively more stable and remained predominantly positive, with noticeably narrower fluctuations throughout much of the 2000s. Nevertheless, episodes of renewed volatility emerged during the late 2010s and early 2020s, indicating that economic growth remained susceptible to short-term disturbances. Across the full sample, the series does not exhibit a clear and persistent upward or downward deterministic trend; rather, it is characterized by time-varying volatility and episodic fluctuations. These observed dynamics provide a strong basis for formally examining the stationarity properties of the series before proceeding with ARIMA model identification and estimation.

4.1.2 Descriptive Statistics of Somalia’s Annual GDP Growth

Table 2 Descriptive Statistics of Somalia’s Annual GDP Growth

Observations(N)	Mean	Median	S.D	Minimum	Maximum	Skewness	Kurtosis
64	3.81	4.38	6.29	-17.85	14.88	-1.04	1.58

The descriptive statistics for Somalia’s annual GDP growth over the 64 observations show an average growth rate of 3.81%, while the median of 4.38% indicates that the typical annual growth performance was slightly higher than the arithmetic mean. The relatively large standard deviation of 6.29 percentage points reveals substantial year-to-year volatility in economic growth. This variability is further reflected in the wide range between the minimum growth rate of −17.85% and the maximum of 14.88%, indicating episodes of severe contraction as well as strong expansion. The negative skewness of −1.04 suggests that the distribution is left-skewed, with relatively extreme negative growth episodes exerting greater influence on the lower tail. In addition, the kurtosis value of 1.58 indicates a more peaked and heavy-tailed distribution than the normal benchmark, implying a comparatively high occurrence of extreme growth outcomes. Overall, the descriptive evidence highlights the considerable instability and asymmetric behavior of Somalia’s historical GDP growth series, providing further motivation for rigorous time-series modeling and forecasting.

4.1.3 Model Training Procedure

Table 3. Distribution of observations across the full, training, and test samples

Total Sample	Training Set	Test Set
64	54	10

The dataset comprises 64 annual observations of Somalia’s GDP growth covering 1961–2024. To preserve the chronological structure of the series and prevent information leakage, the first 54 observations covering 1961–2014 were allocated to the training sample, while the remaining 10 observations covering 2015–2024 were reserved exclusively as an independent test set for evaluating out-of-sample forecasting performance.

4.1.4 Stationarity Assessment and Determination of the ARIMA Differencing Order.

Table 4. Stationarity tests for Somalia’s annual GDP-growth training series, 1961–2014.

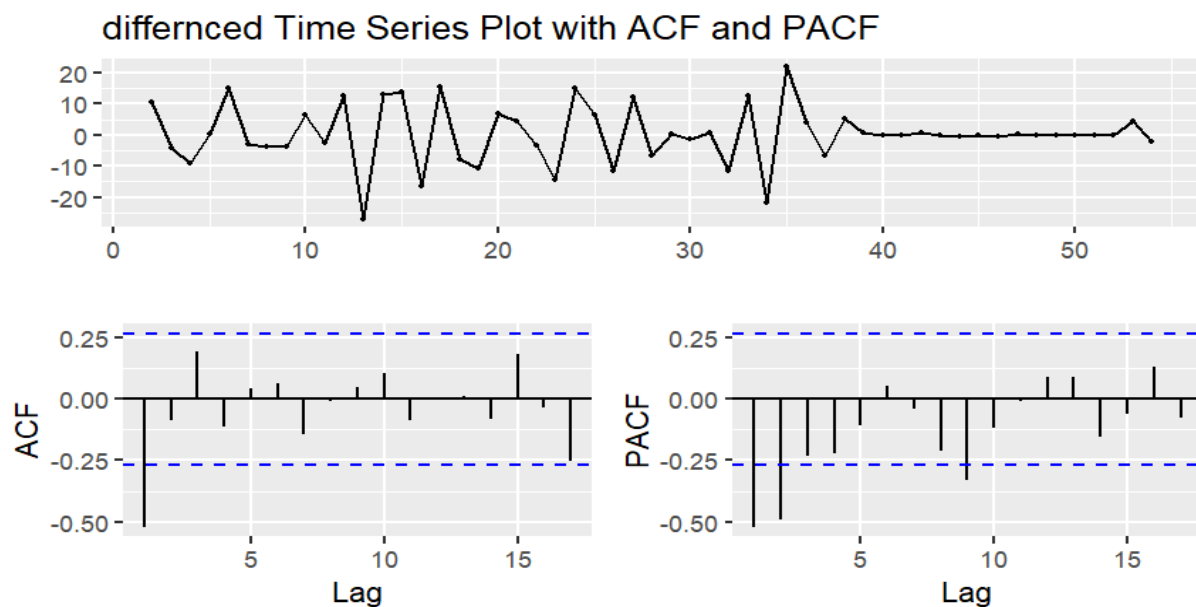
Test	Test Statistic	p-value	Decision at 5%	Conclusion
ADF at level	-3.0644	0.1458	Fail to reject H_0	Non-stationary
PP at level	-60.296	0.0100	Reject H_0	Stationary
KPSS at level	0.4691	0.04863	Reject H_0	Non-stationary
ADF at first difference	-6.0384	< 0.010	Reject H_0	Stationary

The stationarity tests applied to the 1961–2014 training series produced mixed results. The ADF test failed to reject a unit root ($p = 0.1458$), and the KPSS test rejected stationarity ($p = 0.0486$), whereas the Phillips–Perron test rejected a unit root ($p = 0.0100$). Thus, the level results do not establish stationarity consistently across tests. First differencing was examined to address the possibility of persistent non-stationarity; the ADF test on the differenced series rejected a unit root (statistic = −6.0384, $p < 0.01$). On this basis, $\Delta(d=1)$ was used when estimating the reported ARIMA candidates. Because the original variable is already an annual *growth rate* and the level

tests disagree, the choice of $(d=1)$ should be treated as a modeling decision rather than proof that differencing was essential.

4.1.5 First-Differenced Series, ACF–PACF Diagnostics, and Candidate ARIMA Models

Figure 2. First-differenced Somalia GDP-growth training series and its ACF–PACF diagnostics.



The first-differenced training series fluctuates around a relatively stable mean with no persistent trend, indicating that the differencing procedure has effectively reduced the non-stationary behavior present in the original series. The ACF exhibits a pronounced negative spike at lag 1, while most subsequent autocorrelations remain within the approximate 95% confidence bounds, suggesting a strong short-run moving-average component. In contrast, the PACF shows several significant negative spikes at the early lags before gradually diminishing, indicating the possible presence of low-order autoregressive dynamics. Taken together, these patterns support the estimation of parsimonious $\text{ARIMA}(p, 1, q)$ specifications, particularly $\text{ARIMA}(0, 1, 1)$, $\text{ARIMA}(1, 1, 1)$, $\text{ARIMA}(2, 1, 0)$, $\text{ARIMA}(3, 1, 0)$, and $\text{ARIMA}(2, 1, 1)$ as plausible candidate models. These

competing specifications were subsequently compared using AIC, AICc, and BIC, with the preferred model further evaluated through residual diagnostic testing and out-of-sample forecast accuracy.

4.1.6 Comparison and Selection of Candidate ARIMA Models.

Table 5. Comparison of candidate ARIMA models for Somalia's annual GDP growth

Candidate models	σ^2	log likelihood	AIC	AICc	BIC
Auto-ARIMA	44.03	-175.86	355.72	355.96	359.66
ARIMA (0,1,1)	44.03	-175.86	355.72	355.96	359.66
ARIMA (1,1,1)	44.14	-175.4	356.81	357.3	362.72
ARIMA (2,1,1)	44.37	-174.99	357.97	358.8	365.85
ARIMA (3,1,0)	48.22	-176.88	361.76	362.59	369.64
ARIMA (2,1,0)	49.86	-178.21	362.42	362.91	368.34

Best ARIMA Model is ARIMA (0,1,1)

The comparison of the candidate ARIMA specifications indicates that ARIMA (0,1,1) provides the most favorable balance between goodness of fit and model parsimony, recording the lowest AIC (355.72), AICc (355.96), and BIC (359.66) among the competing models. The identical information criteria reported by Auto-ARIMA further confirm that the automatic selection procedure also favored the ARIMA (0,1,1) structure. ARIMA (1,1,1) emerges as the second-best specification, followed by ARIMA (2,1,1), as both models exhibit progressively higher information-criterion values. Although ARIMA (2,1,1) achieves a slightly higher log-likelihood, the gain in fit is insufficient to compensate for its additional complexity. Accordingly, ARIMA (0,1,1) is selected as the preferred candidate model, with ARIMA (1,1,1) and ARIMA (2,1,1) retained as the principal alternative specifications for subsequent diagnostic and forecasting evaluation.

4.1.7 Parameter Estimates of the Selected ARIMA (0,1,1) Model.

Table 6. Parameter estimate of the selected ARIMA (0,1,1) model

Parameter	Estimate	Standard Error	z-statistic	p-value
MA(1)	-0.9066	0.0661	-13.7174	< 0.001

The estimated MA (1) coefficient of -0.9066 is large in magnitude and highly statistically significant, as indicated by the z-statistic of -13.7174 and p-value < 0.001. This confirms that the moving-average component plays a dominant role in explaining the short-run dynamics of the first-differenced GDP growth series. The negative sign implies that a positive forecast error in the preceding period tends to be followed by a downward adjustment in the current period, while a negative error tends to induce an upward correction, reflecting a strong short-run error-reversal

mechanism. The relatively small standard error of 0.0661 further indicates that the coefficient is estimated with high precision. Overall, the parameter estimate provides strong statistical support for the selected ARIMA (0,1,1) specification and suggests that recent forecast innovations exert a substantial influence on subsequent movements in the differenced series.

4.1.8 Estimated Equation of the Selected ARIMA (0,1,1) Model

$$\Delta g_t = \varepsilon_t - 0.9066 \varepsilon_{t-1}$$

where g_t denotes Somalia's annual GDP growth rate, Δg_t represents its first difference, ε_t is the current-period innovation, and ε_{t-1} is the one-period-lagged innovation. The estimated MA (1) coefficient of -0.9066 therefore implies that the current change in the GDP-growth series is strongly influenced by the previous period's forecast innovation in the opposite direction. The coefficient is highly statistically significant ($z = -13.7174$, $p < 0.001$), supporting the inclusion of the moving-average term in the selected specification.

In operator form, the fitted model can be written as $(1 - B)g_t = (1 - 0.9066B) \varepsilon_t$, where B is the backshift operator. This formulation is consistent with an ARIMA (0,1,1) process without drift and makes the estimated structure of the selected model explicit for transparency and reproducibility.

4.1.9 Inevitability Test of the Selected ARIMA (0,1,1) Model.

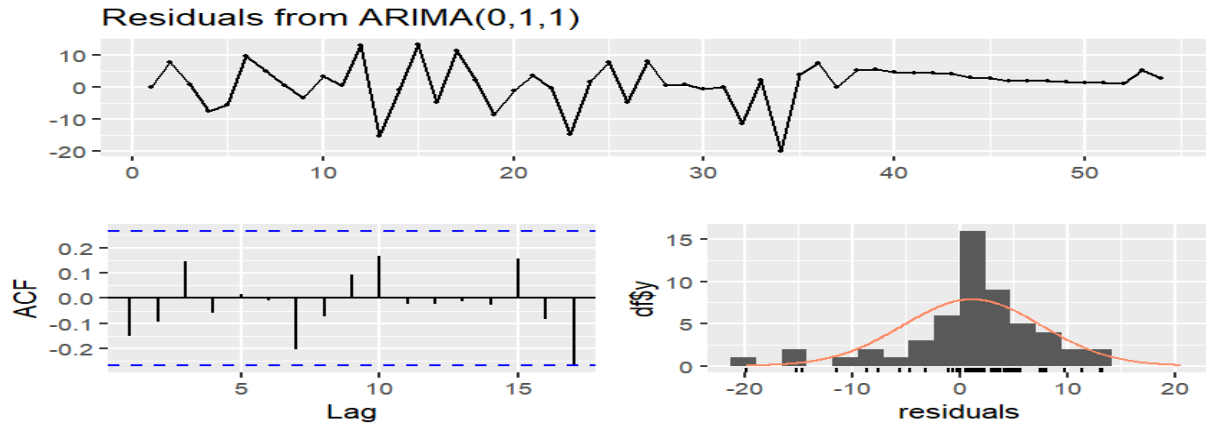
Table 7. Invertibility assessment of the selected ARIMA(0,1,1) model.

Model	MA (1) Coefficient	MA Root	Root Modulus	Decision	Conclusion
ARIMA (0,1,1)	-0.9066	1.1030	1.1030	Root modulus > 1	Invertibility condition satisfied

The invertibility condition of the selected ARIMA (0,1,1) specification was evaluated using the root of the moving-average polynomial. The estimated MA (1) coefficient of -0.9066 produced a corresponding root modulus of 1.1030. Since this value exceeds unity, the root lies outside the unit circle, satisfying the standard invertibility requirement for an MA process. This result confirms that the moving-average component of the model is dynamically well defined and that the ARIMA (0,1,1) representation is statistically admissible. Accordingly, the invertibility test provides additional support for the technical validity and stability of the selected forecasting specification.

4.1.10 Residual Diagnostic Assessment of the Selected ARIMA (0,1,1) Model

Figure 3. Residual diagnostic plots for the selected ARIMA (0,1,1) model



The Ljung–Box test yielded a statistic of 8.8371 with 9 degrees of freedom and a p-value of 0.4524. Since the p-value exceeds the 5% significance level, the null hypothesis of no residual autocorrelation cannot be rejected. This indicates that the residuals do not exhibit statistically significant serial dependence up to lag 10, suggesting that the ARIMA (0,1,1) model has adequately captured the linear temporal dependence in the training series. The residuals are therefore consistent with white-noise behavior in terms of serial correlation, providing further support for the diagnostic adequacy of the selected specification.

4.1.11 Ljung–Box Residual Diagnostic Test

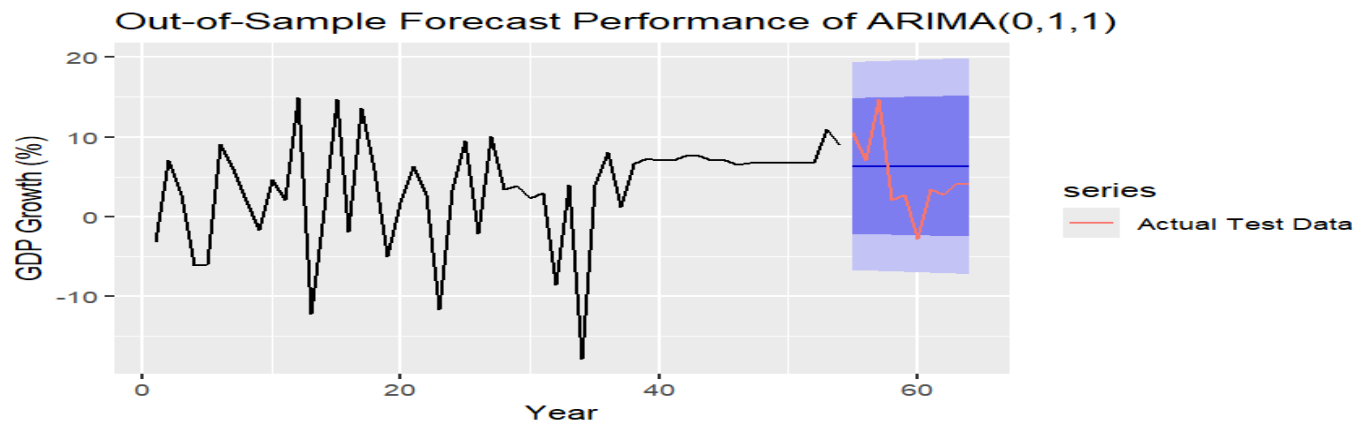
Table 8. Ljung–Box residual autocorrelation test for the selected ARIMA(0,1,1) model.

Model	Lag	Ljung–Box χ^2	df	p-value	Decision	Conclusion
ARIMA (0,1,1)	10	8.8371	9	0.4524	Fail to reject H_0	No significant residual autocorrelation

The residual diagnostics for the selected ARIMA (0,1,1) model indicate that the residuals are generally centered around zero and display no persistent deterministic pattern, suggesting that the model has captured a substantial portion of the systematic temporal structure in the series. The residual ACF is largely contained within the approximate 95% confidence limits, although a noticeable spike appears at one of the higher lags, indicating that residual serial dependence should be assessed formally. Accordingly, the Ljung–Box test is employed separately to determine whether statistically significant autocorrelation remains in the residuals. In contrast, the residual histogram is used to evaluate the distributional shape of the errors and suggests mild asymmetry and the presence of a few relatively extreme observations. Any departure from normality should therefore be assessed independently using an appropriate normality test, such as the Jarque–Bera test, rather than through the Ljung–Box procedure.

4.1.12 Out-of-Sample Forecast Performance of the Selected ARIMA (0,1,1) Model.

Table 9. Out-of-sample forecast accuracy of ARIMA (0,1,1), 2015–2024.



The out-of-sample forecast plot shows that the selected ARIMA (0,1,1) model produces a relatively stable forecast for Somalia’s annual GDP growth over the 2015–2024 test period. The observed growth rates fluctuate more sharply than the point forecast, indicating that the model does not anticipate the timing or size of individual annual changes. The prediction intervals are much wider than the point-forecast path, reflecting uncertainty about future growth. Overall, the figure suggests that ARIMA provides a forecast of the general growth level, while its ability to predict abrupt year-to-year movements is limited.

4.1.13 Out-of-Sample Forecast Accuracy of the Selected ARIMA (0,1,1) Model

Table 9. Out-of-sample forecast accuracy of ARIMA(0,1,1), 2015–2024.

Model	MAE	RMSE	MASE
ARIMA(0,1,1)	4.109335	4.829134	0.6190897

The out-of-sample accuracy results indicate that the selected ARIMA (0,1,1) model achieved an MAE of 4.1093, an RMSE of 4.8291, and a MASE of 0.6191. The relatively close values of MAE and RMSE suggest that, although some larger forecast errors are present, they do not dominate the overall error structure excessively. More importantly, the MASE value is well below one, indicating that the ARIMA model performs better than the corresponding naïve benchmark and reduces the average absolute forecast error by approximately 38.1% relative to that benchmark. Overall, these results provide evidence that ARIMA (0,1,1) delivers a meaningful level of out-of-sample predictive accuracy and can serve as a strong benchmark for subsequent comparison with the ETS and NNAR forecasting models.

4.2 Exponential smoothing

Exponential smoothing is a widely used time-series forecasting method that produces future estimates by giving more weight to recent observations while gradually reducing the influence of older values. Unlike ARIMA, which models autoregressive and moving-average relationships,

exponential smoothing updates the underlying components of a series, particularly its level and, where appropriate, its trend, through smoothing parameters that control how quickly the model responds to new information. In this study, non-seasonal exponential smoothing specifications are considered because the GDP-growth series is observed annually; therefore, a within-year seasonal component is not applicable. The analysis consequently focuses on alternative level and trend structures to identify the most appropriate specification. Model selection is guided by AICc, parsimony, residual adequacy, and out-of-sample forecasting performance, allowing the characteristics of Somalia's GDP-growth series to determine the final smoothing structure rather than imposing one in advance.

All ETS analyses were conducted in R using the forecast package. The GDP-growth series was retained at its original annual frequency (*frequency* = 1), and the same chronological training and test samples defined earlier in the analysis were used throughout. Alternative non-seasonal ETS structures were estimated using the training sample, with AICc and model parsimony guiding structural selection. Parameters of the retained ETS(A,N,N) specification were estimated directly from the training data using the *ets()* function. Forecasts were generated over a horizon equal to the length of the test sample, with 80% and 95% prediction intervals reported to reflect forecast uncertainty. Out-of-sample performance was evaluated using MAE, RMSE, and MASE.

4.2.1 Selection of the Appropriate ETS Structure.

Table 10. Comparison of Candidate Exponential Smoothing Models.

Model	AIC	AICc	BIC	Model Selection Interpretation
ETS(A,N,N)	423.6126	424.0926	429.5795	Parsimonious model; lowest BIC
ETS(A,A,N)	422.7873	424.0373	432.7322	Slightly favored by AIC/AICc
ETS(A,Ad,N)	427.1872	428.9745	439.1211	Less supported by the information criteria

Because Somalia's annual GDP-growth series contains both positive and negative values, the analysis focused on additive-error, non-seasonal ETS specifications. Multiplicative-error models were not considered suitable because they assume that forecast errors vary proportionally with the level of the series, an assumption that becomes problematic when observations include zero or negative values. The model comparison therefore concentrated on alternative trend structures within the additive ETS framework, allowing the data to determine whether a trend component meaningfully improved the representation of GDP growth.

The results show that ETS(A,A,N) achieved the lowest AIC and AICc values of 422.7873 and 424.0373, respectively, while ETS(A,N,N) produced a very similar AICc of 424.0926. The difference of only 0.0553 indicates that the two models receive nearly equivalent support under AICc. However, the lower BIC of 429.5795 favors the simpler ETS(A,N,N) specification, whereas the damped-trend model performs less well across the information criteria. Since adding a trend provides only a negligible improvement in fit while increasing model complexity, ETS(A,N,N) is retained as the preferred parsimonious model for residual diagnostics and out-of-sample forecasting.

Following the structural comparison, the retained exponential-smoothing specification is ETS(A,N,N), which contains an additive error component, a time-varying level, no trend, and no seasonal component. The model can be expressed in innovations form as follows:

$$y_t = \ell_t - 1 + \varepsilon_t \quad (1)$$

$$\ell_t = \ell_{t-1} + \alpha \varepsilon_t \quad (2)$$

$$\ell_t = \alpha y_t + (1 - \alpha) \ell_{t-1} \quad (3)$$

$$\hat{y}_{t+h|t} = \ell_t, \quad h \geq 1 \quad (4)$$

where y_t denotes observed annual GDP growth, ℓ_t is the estimated level at time t , α is the level-smoothing parameter with $0 \leq \alpha \leq 1$, and ε_t is the one-step-ahead innovation. Because ETS(A,N,N) contains neither a trend nor a seasonal component, forecasts beyond the estimation period remain equal to the most recently estimated level, while forecast uncertainty widens with the horizon.

4.2.2 Sensitivity Analysis of Alternative Smoothing Parameters for ETS(A,N,N)

Table 11. Sensitivity Analysis of Alternative Smoothing Parameters for ETS(A,N,N)

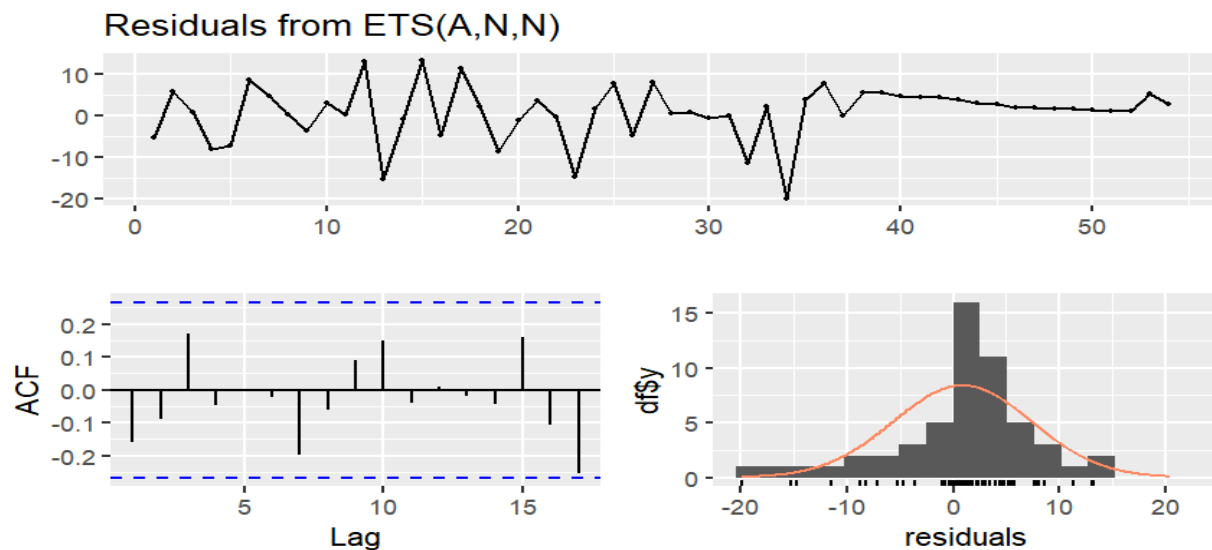
Models	alpha	AIC	AICc	BIC
Model 1 ets	0.0678	423.6126	424.0926	429.5795
Model 2	0.1	421.8263	422.0616	425.8043
Model 3	0.2	424.1424	424.3777	428.1204
Model 4	0.3	427.5002	427.7355	431.4782
Model 5	0.4	431.3336	431.5689	435.3116
Model 6	0.5	435.5226	435.7579	439.5006
Model 7	0.6	440.0213	440.2566	443.9993
Model 8	0.7	444.7900	445.0253	448.7680
Model 9	0.8	449.7989	450.0342	453.7768
Model 10	0.9	455.0386	455.2739	459.0166

The sensitivity analysis reveals systematic differences in model fit across alternative smoothing parameters within the ETS(A,N,N) specification. Model 1 represents the data-estimated ETS(A,N,N) specification, for which the level-smoothing parameter was estimated from the training sample as $\hat{\alpha} = 0.0678$, yielding an AICc of 424.0926. Models 2–10, in contrast, impose fixed values of α ranging from 0.1 to 0.9 to examine the sensitivity of model fit to alternative smoothing intensities. Among these manually fixed specifications, $\alpha = 0.1$ produced the lowest AICc of 422.0616, together with the lowest AIC and BIC values, indicating the strongest fit within the fixed-alpha sensitivity set. The information criteria increased progressively as α rose beyond 0.1, suggesting that greater responsiveness to recent GDP-growth observations does not improve

the level-only model. Overall, the sensitivity results favor relatively low smoothing, indicating that historical observations retain substantial influence on the estimated GDP-growth level. Nevertheless, the fixed-alpha specifications are treated as sensitivity checks, while Model 1, with the data-estimated $\hat{\alpha} = 0.0678$, is retained as the final ETS(A,N,N) model for residual diagnostics and out-of-sample forecasting.

4.2.3 Residual Diagnostics of the Selected ETS(A,N,N) Model.

Figure 5. Residual diagnostics for the selected ETS(A,N,N) model.



The residual diagnostic plots indicate that the fitted ETS(A,N,N) model captures most of the systematic temporal dependence in the annual GDP growth series. Residuals fluctuate around zero without a persistent pattern, while the residual ACF coefficients are generally contained within the 95% confidence bounds, suggesting limited remaining serial correlation. Nevertheless, the magnitude of the residuals is greater in the earlier observations and becomes more stable toward the end of the sample, indicating some variation in residual dispersion over time. The residual distribution is centered near zero, although it is somewhat peaked and departs modestly from normality. Overall, the diagnostic evidence supports the adequacy of the ETS model, subject to confirmation from the Ljung–Box test.

4.2.4 Residual Autocorrelation Diagnostic for the Selected ETS Model.

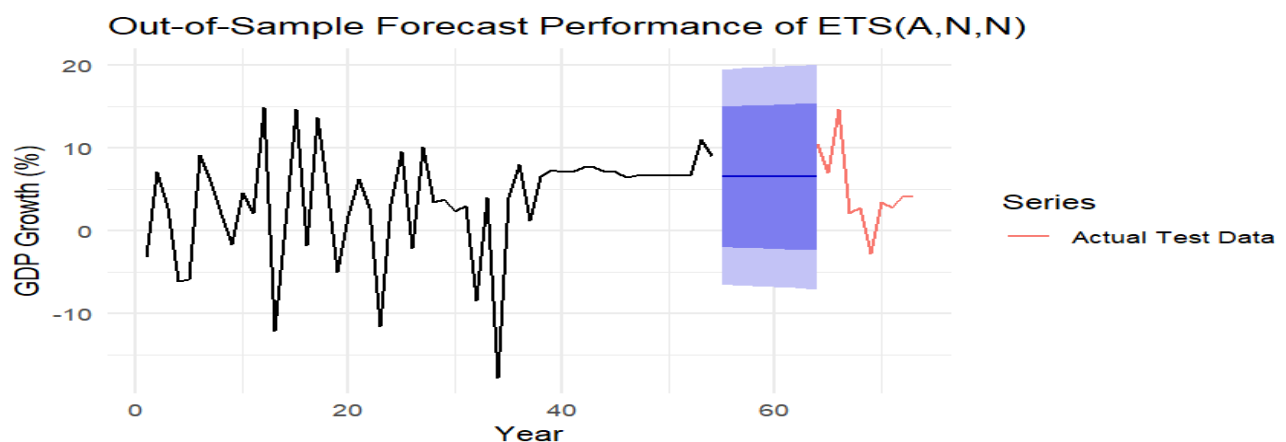
Table 12. Ljung–Box Test for Residual Autocorrelation in the ETS(A,N,N) Model

Diagnostic Test	Test Statistic	df	p-value	Decision at 5% Level	Conclusion
Ljung–Box Test	Q 7.8372	10	0.6447	Fail to reject H0	No significant residual autocorrelation

The Ljung–Box residual autocorrelation test provides further support for the adequacy of the selected ETS(A,N,N) model. The test produced a Q-statistic of 7.8372 with 10 degrees of freedom and a p-value of 0.6447, which is well above the 5% significance level. Accordingly, the null hypothesis of no serial correlation cannot be rejected, indicating that the residuals do not exhibit statistically significant autocorrelation across the examined lags. This result is consistent with the residual ACF evidence and suggests that the model has captured the main temporal dependence in the annual GDP-growth series, leaving residuals that behave approximately as a white-noise process. Overall, the diagnostic findings strengthen the statistical reliability of the selected ETS specification for subsequent forecasting analysis.

4.2.5 Out-of-Sample Forecast Performance of the Selected ETS(A, N, N) Model.

Figure 6. Out-of-sample ETS(A,N,N) forecasts and observed GDP growth, 2015–2024.



The selected ETS(A,N,N) model produces a relatively stable point forecast around the estimated level of annual GDP growth, consistent with a level-only exponential-smoothing specification containing neither trend nor seasonality. The actual test observations exhibit considerably greater year-to-year variation than the forecast path, indicating that the model captures the underlying average growth level more effectively than abrupt short-run fluctuations. Nevertheless, the observed test values generally remain within the prediction intervals, reflecting the considerable uncertainty associated with Somalia’s historically volatile GDP-growth series. Overall, the forecast provides a reasonable baseline representation of the central growth level, while formal out-of-sample performance is assessed using MAE, RMSE, and MASE.

4.2.6 Out-of-Sample Forecast Accuracy of the Selected ETS (A, N, N) Model.

Table 13. Out-of-Sample Forecast Accuracy of the Selected ETS(A,N,N) Model, 2015–2024

Model	MAE	RMSE	MASE
ETS(A,N,N), $\alpha = 0.0678$	3.830792	4.664092	0.577126

The out-of-sample accuracy results demonstrate that the final ETS(A,N,N) specification provides a credible forecasting performance for Somalia’s annual GDP growth. The model records an MAE

of 3.8308, indicating that its forecasts deviate from the observed growth rates by approximately 3.83 percentage points on average. The RMSE of 4.6641 is somewhat larger than the MAE, implying that a few relatively large forecast errors increased the overall squared-error loss. More importantly, the MASE value of 0.5771 is below one, showing that the ETS model outperforms the naïve benchmark over the test period. Taken together, these results indicate that the selected level-only ETS model captures the underlying growth level reasonably well and provides useful out-of-sample predictive performance despite the substantial volatility and abrupt movements evident in Somalia's GDP-growth series.

4.3 Nonlinear Time-Series Forecasting Using Artificial Neural Networks.

Artificial Neural Networks (ANNs) are flexible forecasting models that capture complex and nonlinear patterns from historical observations. In this study, an NNAR model is used to forecast Somalia's annual GDP growth from its own lagged values and to provide a nonlinear comparison with ARIMA and exponential smoothing. Since the data are annual, no seasonal component is included.

Training and Test Periods. The annual GDP-growth series covering 1961–2024 was divided chronologically into a training sample comprising the first 54 observations (1961–2014) and a test sample comprising the remaining 10 observations (2015–2024). The NNAR model was estimated using only the training sample, while the test sample was reserved for evaluating out-of-sample forecasting performance.

Hyperparameter Specification. The NNAR model was specified using the training sample with two autoregressive lagged inputs ($p = 2$) and two hidden neurons ($k = 2$). This produced an NNAR(2,2) model with a 2–2–1 architecture consisting of two input nodes, two neurons in a single hidden layer, and one linear output unit. To improve model stability, 20 networks with different random initial weights were estimated and their forecasts were averaged.

Scaling and Preprocessing. Before NNAR estimation, the input observations were standardized using the mean and standard deviation of the training sample. This scaling places the inputs on a comparable numerical scale and improves the stability of neural-network estimation. The test data were not used during model training or scaling, thereby avoiding information leakage.

Training Procedure. The NNAR model was trained using only the training sample. During estimation, the network weights were iteratively adjusted to minimize the difference between the observed and predicted GDP-growth values. To improve model stability, 20 networks were estimated using different random starting weights, and their forecasts were averaged to produce the final NNAR predictions.

Reproducibility. A fixed random seed (`set.seed(123)`) was applied before NNAR estimation to ensure reproducibility of the network initialization and forecasting results.

4.3.1 Neural Network Autoregression Model Specification

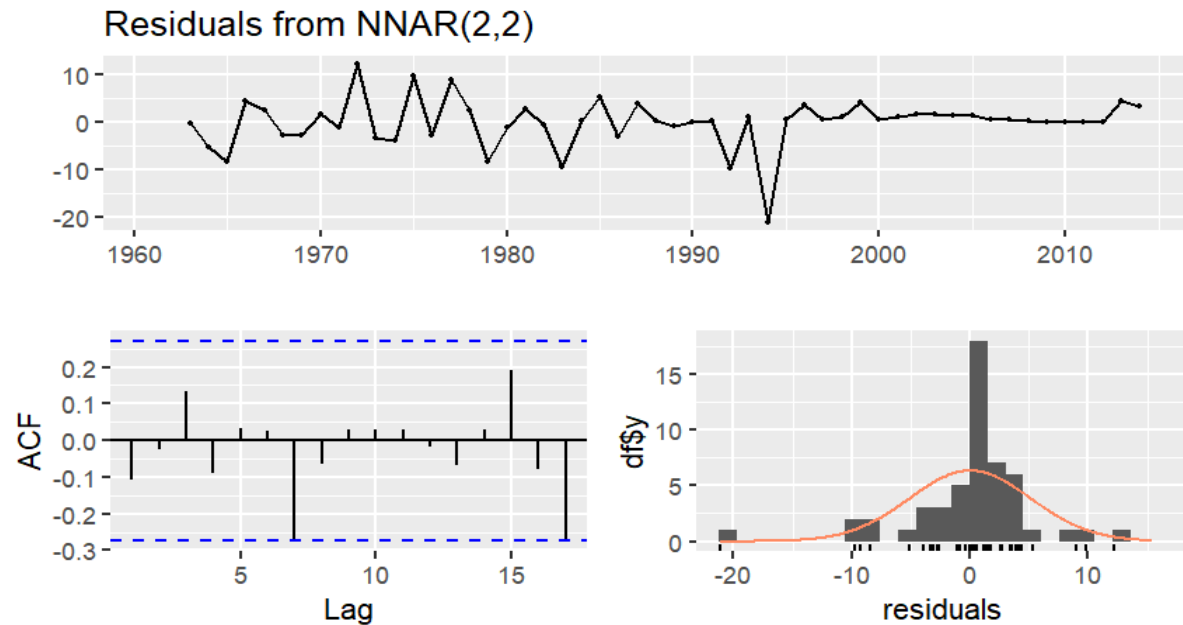
Table 14. Specification of the Selected NNAR (2,2) Model

Model Component	Estimated Specification
Model	NNAR(2,2)
Input lags (p)	2
Hidden neurons (k)	2
Network architecture	2–2–1
Number of networks averaged	20
Number of weights	9
Output unit	Linear
Estimated residual variance (σ^2)	25.94

The selected NNAR (2,2) specification consists of two lagged GDP-growth observations as inputs, two neurons in a single hidden layer, and one linear output unit, producing a compact 2–2–1 architecture with nine estimated weights. To improve stability, the model averages the results of 20 independently initialized neural networks, reducing the influence of any single set of starting weights. The estimated residual variance of 25.94 indicates that some unexplained variation remains, which is expected given the substantial volatility in Somalia’s annual GDP-growth series. Overall, the selected NNAR structure provides a parsimonious nonlinear forecasting framework that is flexible enough to capture potential nonlinear dependence in past GDP growth while remaining sufficiently simple for a relatively small annual time-series sample.

4.3.2 Residual Diagnostic Evaluation and White-Noise Adequacy of the NNAR (2,2) Model.

Figure 7. Residual Diagnostics of the Selected NNAR(2,2) Model.



Residual Diagnostic Evaluation. The residual diagnostics of the NNAR (2,2) model show that the forecast errors fluctuate around zero without a clear systematic pattern. Most residual autocorrelations remain within the 95% confidence bounds, indicating limited remaining serial dependence. Although a few relatively large residuals are observed, particularly during periods of substantial GDP-growth volatility, the overall residual pattern suggests that the NNAR (2,2) model adequately captures the main temporal structure of the series.

4.3.3 Residual Autocorrelation Diagnostic for the NNAR (2,2) Model.

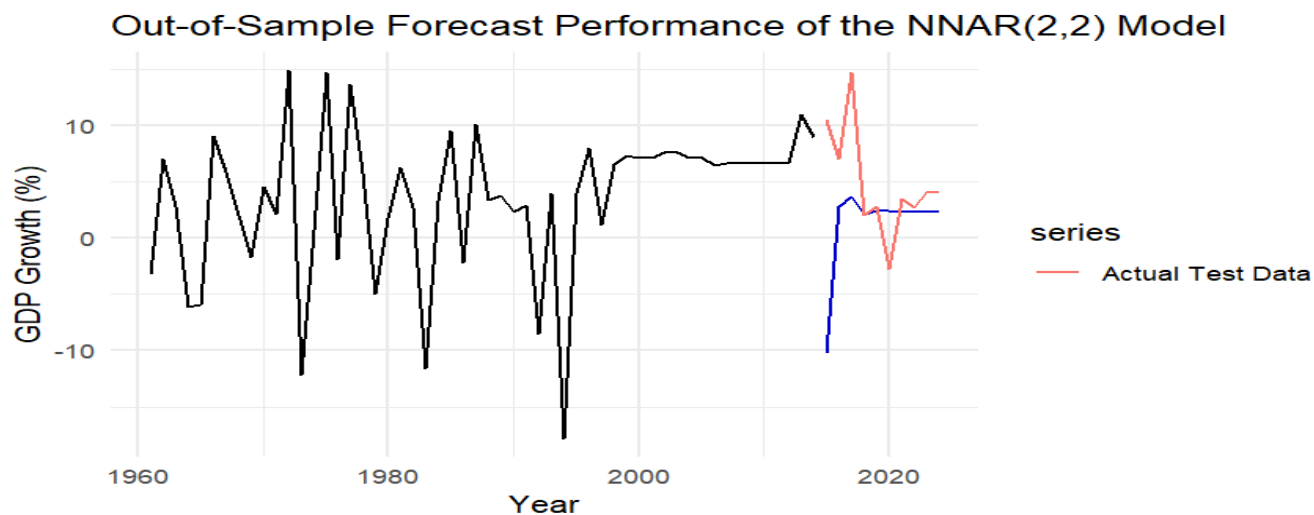
Table 15. Ljung–Box Test for Residual Autocorrelation in the NNAR(2,2) Model.

Model	Test	Q-statistic	df	p-value	Decision at 5%	Conclusion
NNAR(2,2)	Ljung–Box Q Test	7.4071	10	0.6865	Fail to reject H0	No significant residual autocorrelation

The Ljung–Box test confirms that the selected NNAR (2,2) model leaves no statistically significant serial correlation in its residuals. The test produced a Q-statistic of 7.4071 with 10 degrees of freedom and a p-value of 0.6865, which is well above the 5% significance level. Accordingly, the null hypothesis of no residual autocorrelation cannot be rejected, indicating that the remaining forecast errors are not systematically related across the examined lags. This result is consistent with the residual ACF pattern and provides further evidence that the NNAR (2,2) specification has adequately captured the principal temporal dependence in Somalia’s annual GDP-growth series, supporting its suitability for subsequent out-of-sample forecasting.

4.3.4 Out-of-Sample Forecast Performance of the NNAR (2,2) Model.

Figure 8. Out-of-sample NNAR(2,2) forecasts and observed GDP growth, 2015–2024



The out-of-sample forecasts from the NNAR (2,2) model indicate that the model captures the broad pattern of Somalia's GDP growth but has limited ability to reproduce the sharp fluctuations observed during the test period. Following an initial adjustment, the forecasts stabilize around a positive growth rate of approximately 2–3%, whereas the actual observations exhibit considerably greater volatility, including pronounced expansions and a temporary contraction. This divergence suggests that the NNAR model is more effective in representing the general growth tendency than in anticipating abrupt changes and turning points. Overall, the graphical evidence indicates reasonable forecast performance, which is further evaluated using the out-of-sample MAE, RMSE, and MASE measures.

4.3.5 Out-of-Sample Forecast Accuracy of the Selected NNAR (2,2) Model.

Table 16. Out-of-Sample Forecast Accuracy of the Selected NNAR(2,2) Model, 2015–2024

Model	MAE	RMSE	MASE
NNAR (2,2)	4.65846	7.760362	0.7018179

The out-of-sample accuracy measures indicate that the selected NNAR (2,2) model achieves a reasonable level of predictive performance for Somalia's annual GDP growth. The model records an MAE of 4.6585, implying that the forecasts deviate from the observed growth rates by approximately 4.66 percentage points on average. The larger RMSE of 7.7604 suggests that several substantial forecast errors occurred during the evaluation period, particularly when the series experienced abrupt movements. Nevertheless, the MASE value of 0.7018, which is below one, demonstrates that the NNAR model performs better than the naïve benchmark. Overall, the findings suggest that the model captures the general underlying growth pattern, although its forecasting accuracy weakens when confronted with sharp and unexpected fluctuations in GDP growth.

4.3.6 Comparative Out-of-Sample Forecast Accuracy of ARIMA, ETS, and NNAR Models.

Table 17 Comparative Out-of-Sample Forecast Accuracy of ARIMA, ETS, and NNAR.

Model	MAE	RMSE	MASE
ARIMA(0,1,1)	4.109335	4.829134	0.6190897
ETS(A,N,N), $\alpha = 0.0678$	3.830792	4.664092	0.577126
NNAR (2,2)	4.65846	7.760362	0.7018179

The comparative out-of-sample forecast results indicate that ETS(A,N,N) is the best-performing model among the three candidates, as it records the lowest MAE (3.8308), lowest RMSE (4.6641), and lowest MASE (0.5771). This means that, on average, ETS produces the smallest absolute forecast errors, is less affected by large forecasting misses, and provides the strongest scaled accuracy relative to the naïve benchmark. ARIMA (0,1,1) ranks second, with MAE = 4.1093, RMSE = 4.8291, and MASE = 0.6191, indicating satisfactory predictive ability but slightly weaker accuracy than ETS across all evaluation criteria. In contrast, NNAR (2,2) shows the weakest out-of-sample performance, with MAE = 4.6585, RMSE = 7.7604, and MASE = 0.7018, suggesting that although it captures the general movement of GDP growth, it is less successful in tracking sharp year-to-year fluctuations and is more vulnerable to larger forecast errors. Overall, the results demonstrate that the ETS model offers the most accurate, stable, and reliable forecasts for Somalia's annual GDP growth, implying that the underlying series is better represented by a relatively simple exponential-smoothing structure than by the more complex nonlinear NNAR specification.

4.4 Discussion and Study Limitations

ETS(A,N,N) produced the lowest MAE, RMSE and MASE in this comparison, followed by ARIMA(0,1,1), while NNAR(2,2) had the highest errors. This finding shows that, for Somalia's annual GDP-growth series and the test period examined, the relatively simple ETS model forecast more accurately than the selected linear and neural network alternatives. It is consistent with the broader empirical literature's finding that greater model complexity does not necessarily improve forecast accuracy. Other studies reviewed in this paper identified different leading models for Somalia's GDP and other indicators (Osman and Muse, 2024; Ibrahim, Mohamed et al., 2026). Their rankings cannot be compared directly with this study's rankings because the forecast variables, models and evaluation procedures differ.

The findings have several limitations. The analysis concerns one country and one annual GDP-growth series, so its model ranking should not be generalized to other countries or economic indicators. Accuracy was assessed using only ten test observations, representing forecast horizons from one to ten years from the end of 2014; unusual growth shocks in those years could strongly affect the error measures. Annual data also conceal changes within each year, while revisions to historical GDP estimates could alter results in a later data release. Because the models use past GDP growth alone, they cannot directly anticipate new droughts, policy changes or other unforeseen events. The results therefore support ETS(A,N,N) as the best-performing model among those tested under this evaluation design, rather than as a universally superior forecasting method.

5.0 Conclusion, Recommendations, and Sustainability Implications.

The study compared ARIMA (0,1,1), ETS(A,N,N), and NNAR (2,2) for forecasting Somalia's annual GDP growth using a chronological training sample from 1961–2014 and a test sample from 2015–2024. The results show that all three models were statistically adequate for forecasting, with residual diagnostics indicating no significant serial correlation. However, the models differed clearly in out-of-sample predictive accuracy, reflecting their different abilities to represent Somalia's volatile growth dynamics.

Among the three approaches, ETS(A,N,N) performed best, recording the lowest MAE of 3.8308, RMSE of 4.6641, and MASE of 0.5771. ARIMA (0,1,1) ranked second, while NNAR (2,2) produced the highest forecast errors. These findings indicate that the relatively simple level-based ETS structure represented the underlying GDP-growth process more effectively than the linear ARIMA and nonlinear NNAR alternatives during the test period.

Accordingly, ETS(A,N,N) is recommended as the preferred model for short- to medium-term forecasting of Somalia's annual GDP growth, while ARIMA may be retained as a useful alternative for comparison and robustness. Forecasting models should nevertheless be updated regularly as new observations become available, since Somalia's historical growth series remains highly volatile and exposed to abrupt changes. Forecast results should therefore be interpreted together with uncertainty rather than treated as exact future outcomes.

From a sustainability perspective, more reliable GDP-growth forecasts can strengthen fiscal planning, public-investment decisions, and the allocation of scarce development resources. The strong performance of ETS provides a practical forecasting tool for supporting economic resilience and progress toward sustainable development, particularly SDG 8 on decent work and economic growth. At the same time, the persistent volatility identified in the series highlights the importance of flexible, shock-responsive planning and stronger national statistical systems to support evidence-based and sustainable economic decision-making.

Future Research

Future research can extend this study in several important directions. First, subsequent studies could use higher-frequency data, such as monthly or quarterly observations, to better capture the immediate effects of climate shocks, displacement, and conflict on food prices. Annual data may conceal short-term price responses that occur immediately after droughts, floods, conflict outbreaks, or large population movements.

Second, future studies could examine the effects of different types of climate shocks separately. For example, droughts, floods, extreme temperatures, and rainfall variability may affect agricultural production, population displacement, and food prices through different channels. Distinguishing these shocks could provide a clearer understanding of which climatic events create the greatest pressure on food-price inflation in IGAD countries.

Third, further research could investigate the transmission mechanisms connecting displacement to food prices. In particular, future studies may examine whether the effects operate through reductions in agricultural production, disruption of transport and markets, increased food demand

in receiving areas, loss of productive assets, or increased dependence on imported food. Such analysis would help explain not only whether displacement affects food prices, but also how this effect occurs.

Fourth, future research could distinguish between internally displaced persons and cross-border refugees. The economic effects of these two forms of displacement may differ considerably because internal displacement mainly reallocates population within a country, whereas refugee movements can create additional economic and food-market pressures across national borders.

Fifth, spatial and subnational analysis could provide further insights. Climate shocks and conflict are often concentrated in particular regions rather than affecting an entire country equally. Using district- or region-level data could therefore identify geographical differences in the relationship between climate shocks, displacement, and food-price inflation, particularly between food-producing areas, conflict-affected areas, and major displacement-receiving locations.

Future studies could also investigate whether institutional quality, humanitarian assistance, food-import dependence, irrigation infrastructure, market accessibility, and social protection moderate the effects of climate- and conflict-induced displacement on food prices. In addition, comparative research beyond the IGAD region could determine whether the relationships identified in this study also occur in other climate-vulnerable and conflict-affected regions of Africa.

Finally, future research could employ alternative econometric approaches, including spatial panel models, panel vector autoregression, nonlinear models, and causal identification strategies where suitable data are available. These extensions could help capture cross-border spillovers, dynamic interactions, possible threshold effects, and the complex relationships among climate shocks, conflict, displacement, agricultural production, and food-price inflation.

Contributions of Authors.

Abdimajjid Ali Abdi: Had a pivotal role in formulating the research topic, Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing and managing the overall project.

Abdihakim Ahmed Hassan: Was accountable for devising the methodology, Writing – original draft, validating data and performing formal analysis.

Shukri Yusuf Salad: was essential in data validation, inquiry, text reading and editing, visual presentation, and general supervision.

Funding Information:

No funding

Data Availability

The data used in this study were obtained from publicly available sources, mainly the World Bank's World Development Indicators (<https://databank.worldbank.org/source/world-development-indicators>) and SESRIC (<https://oicstat.sesric.org/query>). The cleaned and

processed datasets that support the results can be provided by the corresponding author, Mr. Abdihakim Ahmed Hassan (Sheqow125@gmail.com), upon reasonable request.

Clinical Trial Number:

Not applicable.

Declarations

Ethical approval

Not applicable. This study uses publicly available secondary macroeconomic data and involves no human participants, animals, or identifiable personal data; therefore, institutional ethics approval was not required.

Consent to participate

Not relevant. No human subjects were involved.

Consent to publish

Not applicable. The manuscript contains no individual person's data in any form (including individual details, images, or videos).

Human or animal rights

This study does not involve any human or animal participants, clinical interventions, or identifiable personal data that would require ethical clearance or consent.

Competing interests

The authors declare no competing interests.

REFERENCES

Abdelrahim, N. A. M. and S. Jin (2026). "Eco-Phase: A Phase-Space Trajectory Framework for Unsupervised Drought Detection and LSTM-Based Forecasting in the Horn of Africa." Journal of Geophysical Research: Biogeosciences **131**(7): <https://doi.org/10.1029/2025JG009470>**Digital Object Identifier (DOI)**

Abdi, A. M., et al. (2026). "Comparative Performance of Mechanistic, Statistical, and Hybrid Models of Forecasting Dengue Fever Incidence in Somalia. A Retrospective Time Series Analysis." Health Science Reports **9**(8): e72822. <https://doi.org/10.1002/hsr2.72822>

Abdi, M. H. A., Nur Mohamud; Osman, Abdirahman Omer (2026). Modeling and Forecasting Meat Production in Somalia, 1961–2024: A Comparative Analysis of Single and Hybrid Time Series Models. Research Square. R. Square. <https://doi.org/10.21203>

Abdi, M. K., et al. (2026). "Forecasting annual high school completers in Somaliland using univariate and hybrid time series models." Discover Computing **29**(1): 303.

Abdullahi Hassan, M., et al. (2026). "Comparative Time Series Forecasting of Unemployment in Somalia (1991–2024): Evidence from Single and Hybrid Models [version 1; peer review: awaiting peer review]." Open Research Africa **9**(24).

Abdullahi, M. A., et al. (2026). "Hybrid time series modeling of human fatalities in somalia forecasts critical trends for sustainable humanitarian planning." Discover Sustainability.

Abdullahi, M. O., et al. (2024). "Revolutionizing Somali agriculture: harnessing machine learning and IoT for optimal crop recommendations." Discover Applied Sciences **6**(3): 77.

Ahmed, A., et al. (2026). "Forecasting cereal production and its implications for food security in Somalia: a comparative analysis of ARIMA, ETS and Neural Network models."

Alasow, A. A., et al. (2025). "Climate Rhythms in a Fragile Region: Decoding Somalia's 120-Year Rainfall Patterns Through Spatial Econometrics." Earth Systems and Environment **9**(3): 2045–2060.

Alexandrino, F., et al. (2026). "A Novel Method for Time Series Prediction with Small Data: Integrating Data Augmentation, Normalization Techniques, and Machine Learning." Computational Economics.

Ali, J. A., et al. (2025). "Beyond health: A machine learning analysis of structural barriers to school attainment in Somalia." Plos one **20**(10): e0326958.

Amin Megat Ali, M. S., et al. (2024). "Short-term Gini coefficient estimation using nonlinear autoregressive multilayer perceptron model." Heliyon **10**(4).

Appiah Takyi, K. (2024). "PREDICTION OF POPULATION SIZES OF AFRICAN COUNTRIES BY RECURRENT NEURAL NETWORK."

Beck, N., et al. (2025). "Mind the naive forecast! a rigorous evaluation of forecasting models for time series with low predictability." Applied Intelligence **55**(6): 395.

Billah, B., et al. (2006). "Exponential smoothing model selection for forecasting." International Journal of Forecasting **22**(2): 239–247.

Box, G. E. P. and G. M. Jenkins (1968). "Some Recent Advances in Forecasting and Control." Journal of the Royal Statistical Society Series C: Applied Statistics **17**(2): 91–109.

Caushi, A. (2025). "Assessing the impact of EU integration on economic growth forecasts using artificial neural network models." Jurnal Ilmiah Ilmu Terapan Universitas Jambi **9**(4): 1378–1390.

Chattopadhyay, S. (2007). "Feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India." Acta Geophysica **55**(3): 369–382.

Dahir, H. M., et al. (2025). "Forecasting Urban Population Growth in Somalia: Using ARIMA, TBATS, NNAR, and Hybrid Models."

Demissie, Z., et al. (2026). "Predicting food security in Ethiopia using spatiotemporal integration of heterogeneous data." Journal of Agriculture and Food Research **26**: 102674.

Dhaqane, A. I., et al. (2026). "Integrated Water Allocation Modeling Under Future Development Scenarios in the Juba River Basin, Somalia: A WEAP-Based Approach for Sustainable Water Management." Earth Systems and Environment **10**(3): 2751–2777.

Donizette, A. C., et al. (2025). "Predicting leishmaniasis outbreaks in Brazil using machine learning models based on disease surveillance and meteorological data." Operations Research for Health Care **44**: 100453.

Ettensperger, F. (2020). "Comparing supervised learning algorithms and artificial neural networks for conflict prediction: performance and applicability of deep learning in the field." Quality & Quantity **54**(2): 567–601.

- Gardner Jr, E. S. (1985). "Exponential smoothing: The state of the art." Journal of Forecasting **4**(1): 1–28.
- Hewamalage, H., et al. (2023). "Forecast evaluation for data scientists: common pitfalls and best practices." Data Mining and Knowledge Discovery **37**(2): 788–832.
- Holt, C. C. (2004). "Forecasting seasonals and trends by exponentially weighted moving averages." International Journal of Forecasting **20**(1): 5–10.
- Hornik, K., et al. (1989). "Multilayer feedforward networks are universal approximators." Neural Networks **2**(5): 359–366.
- Hyndman, R. J. and Y. Khandakar (2008). "Automatic Time Series Forecasting: The forecast Package for R." Journal of Statistical Software **27**(3): 1 – 22.
- Ibrahim, A. A. (2026). "Modeling and Forecasting the Declining Trajectory of Agricultural Land Per Capita in Somalia (1961–2050): A Comparative Time Series Approach."
- Ibrahim, A. D., et al. (2026). "Comparative performance of machine learning models in forecasting Somalia's GDP growth using macroeconomic and environmental indicators." Discover Sustainability.
- Iglesias, A., et al. (2004). "A comparison between functional networks and artificial neural networks for the prediction of fishing catches." Neural Computing & Applications **13**(1): 24–31.
- Karagöz, E., et al. (2026). "Multivariate Time-Series Forecasting of Youth Unemployment in Turkey: A Comparison of Deep Learning and Econometric Models." Symmetry **18**(1): 79.
- Luo, X. (2026). A Comparative Study of Traditional Statistical and Machine Learning Models for Forecasting Mainland Chinese Tourist Arrivals in Macao. Proceedings of the 2025 2nd International Conference on Economic Data Analytics and Artificial Intelligence, Association for Computing Machinery: 659–665.
- Mahamoud Yousuf, I., et al. (2026). "Neonatal Mortality Forecasting in Somalia: A Comparative Time-Series Analysis of Single and Hybrid Models [version 1; peer review: 1 approved with reservations]." Open Research Africa **9**(32).
- Makridakis, S., et al. (2018). "Statistical and Machine Learning forecasting methods: Concerns and ways forward." PLOS ONE **13**(3): e0194889.
- Mohamed, A. A., et al. (2026). "Understanding unemployment trends in Somalia through ARIMA modeling: insights for shaping economic policy." Quality & Quantity **60**(1): 1473–1489.
- Mohamed, A. A., et al. (2026). "Forecasting urbanization and electricity consumption dynamics in Somalia using ARIMA modeling." Discover Sustainability **7**(1): 577.
- Mohamed, A. M. and A. D. Ibrahim (2026). "Forecasting unemployment in Somalia using machine learning models." Discover Sustainability.
- Mohamed, A. O. (2022). "Modeling and Forecasting Somali Economic Growth Using ARIMA Models." Forecasting **4**(4): 1038–1050.
- Mohamed, M. A., et al. (2026). "Effects of external financing and public expenditure on economic growth in Somalia." Discover Sustainability **7**(1): 1308.
- Mou, S., et al. (2025). "An enhanced bat algorithm based intelligent inspired architecture for resilient macroeconomic prediction." Scientific Reports **15**(1): 44977.

Moussa, S. A., et al. (2025). Bahrain's GDP Forecasts through Econometric and Machine Learning Models. 2025 IEEE International Conference on Emerging Trends in Engineering and Computing (ETECOM), IEEE.

Mwanga, M. (2025). "A trade-off analysis of deforestation and economic growth in East Africa." Trees, Forests and People **21**: 100980.

Noa-Yarasca, E., et al. (2024). "Deep Learning Model Effectiveness in Forecasting Limited-Size Aboveground Vegetation Biomass Time Series: Kenyan Grasslands Case Study." Agronomy **14**(2): 349.

Nor, M. I. and S. N. Hussein (2025). "Modeling residential property prices in emerging climate-responsive urban markets: a hybrid modeling framework for Baidoa City-Somalia." Frontiers in Built Environment **11**: 1615229.

Obwoye, F. K. (2025). Forecasting Kenya's public debt using time series analysis, UoEm.

Osman, A. A., et al. (2026). "Factors influencing unemployment rate of women in Somalia: an application of machine learning models." SN Social Sciences **6**(3): 112.

Osman, B. M., et al. (2026). "Analyzing deforestation drivers in Somalia using machine learning and 2SLS with economic growth, population dynamics, renewable energy, and globalization." Discover Sustainability **7**(1): 194.

Osman, B. M., et al. (2025). "The impact of climate change on economic growth in Somalia: using random forest and Bayesian approach." Cogent Economics & Finance **13**(1): 2496684.

Osman, B. M. and A. M. S. Muse (2024). "Predictive analysis of Somalia's economic indicators using advanced machine learning models." Cogent Economics & Finance **12**(1): 2426535.

Oukhouya, H. (2024). "Forecasting international stock market trends: XGBoost, LSTM, LSTM-XGBoost, and Backtesting XGBoost models." Statistics, Optimization and Information Computing.

Pérez-Rosero, D. A., et al. (2025). "An Explainable Framework Integrating Local Biplots and Gaussian Processes for Unemployment Rate Prediction in Colombia." Computation **13**(5): 116.

Qi, M. and G. P. Zhang (2001). "An investigation of model selection criteria for neural network time series forecasting." European Journal of Operational Research **132**(3): 666–680.

Quang, P. D., et al. (2026). "A Comparison of Deep Learning and Time Series Models in Forecasting Agricultural Exports: Evidence From Vietnam's Pepper Exports to the United States." Journal of Applied Mathematics **2026**(1): 3278035.

Seiman, S. M. K., et al. (2026). "Comparative analysis of single and hybrid time series models for forecasting of deaths from cancer in Somalia." Discover Public Health **23**(1): 1107.

Tannehill, B. R. (2008). "Forecasting Instability Indicators in the Horn of Africa Region."

Tashman, L. J. (2000). "Out-of-sample tests of forecasting accuracy: an analysis and review." International Journal of Forecasting **16**(4): 437–450.

Valiyev, T., et al. (2025). "A comparative analysis of GDP determinants in Germany and Poland: Integrating econometric and machine learning perspectives." National Accounting Review **7**(4): 501–521.

Woldmariam, M. F. (2025). Forecasting Drought from Satellite Imagery: Comparing Stochastic, Classical Machine Learning, and Deep Learning Approaches. Advanced Geospatial Intelligence and AI for

Environmental Resilience and Sustainable Development. M. O. Sarif and A. Sharifi. Cham, Springer Nature Switzerland: 395–416.

Yousuf, I. M. and S. M. Kahie (2026). "Time Series Analysis and Forecasting of Maternal Mortality in Somalia." JSM 6(1): 1017.