*Research Article*

# Modelling Customs Revenue in Ghana Using Novel Time Series Methods

**Diana Ayorkor Agbenyega, John Andoh ⓘD, Samuel Iddi ⓘD, and Louis Asiedu ⓘD**

*Department of Statistics & Actuarial Science, School of Physical and Mathematical Sciences, University of Ghana, Legon, Accra, Ghana*

Correspondence should be addressed to Louis Asiedu; lasiedu@ug.edu.gh

Governments across the world rely on their Customs Administration to provide functions that include border security, intellectual property rights protection, environmental protection, and revenue mobilisation amongst others. Analyzing the trends in revenue being collected from Customs is necessary to direct government policies and decisions. Models that can capture the trends being purported from the nominal (nonreal) tax values with respect to the trade volumes (value) over the period are indispensable. Predominant amongst the existing models are the econometric models (the GDP-based model, the monthly receipts model, and the microsimulation model), which are laborious and sometimes unreliable when studying trends in time series data. In this study, we modelled monthly revenue data obtained from the Ghana Revenue Authority-Customs Division (GRA-CD) for the period January 2010 to December 2019 using two traditional time series models, ARIMA model and ARIMA Error Regression Model (ARIMAX), and two machine learning time series models, Bayesian Structural Time Series (BSTS) model and a Neural Network Autoregression model. The Neural Network Autoregression model of the form NNAR (1, 3) provided the best forecasts with the least Mean Squared Error (MSE) of 53.87 and relatively lower Mean Absolute Percentage Error (MAPE) of 0.08. Generally, the machine learning models (NNAR (1, 3) and BSTS) outperformed the traditional time series models (ARIMA and ARIMAX models). The forecast values from the NNAR (1, 3) indicated a potential decline in revenue and this emphasizes the need for relevant authorities to institute measures to improve revenue generation in the immediate future.

## 1. Introduction

The Ghana Revenue Authority Customs-Division (GRA-CD) mandates are revenue collection, border protection, collection of international trade statistics, and trade facilitation [1].

The revenue of the Government of Ghana is largely generated from direct taxes, taxes on international trade, indirect taxes, and other taxes which are collected by the VAT Service, the Internal Revenue Service, and the Customs Excise and Preventive Service [2].

In the twenty-first (21st) century, the reliance on Customs services for international trade facilitation and a seamless process flow in the supply chain, without compromising on the security of such trade, cannot be overemphasized.

The GRA-CD still has revenue collection as a top priority, as it continues to complement taxes generated from a limited domestic tax base [3]. The Ghana Revenue Authority (GRA) is constantly devising ways to widen the tax net, automate and simplify its processes, minimize revenue leakages and ensure maximum revenue is mobilised from both its Domestic and Customs Divisions. GRA envisions to achieve a 17.5% tax collection to GDP by the end of 2022. In 2018, Ghana's tax-to-GDP was 14.1% which was lower than the regional rate of 16.5% in Africa [4]. However, as Ghana migrates to an upper middle-income economy, its trade (or customs taxes) as a proportion of total taxes has been dwindling gradually especially in the last 10 years [4]. According to Abrokwah et al. [5], recent years have seen substantial falls in customs revenue collections, with the share of taxes collected at the country's ports falling from 42% in 2017 to 30% in 2019.

Some of the existing literature which sought to assess or investigate the causes and impact of the reduction in revenue from trade tax focused on issues, such as the impact of trade liberalization on reduction in revenue [6] and the impact of tax revenues on the development of countries [7].

Forecasting trade tax or customs revenue has been proven to be one sure way to mitigate the reduction in revenue. According to Anderson and Johnson [2], reliable and trusted revenue predictions provide the foundation for fiscal discipline and for the adoption of an executable public budget.

Whitfield and Duffy [8] described revenue forecasting as an important topic required to track performance and support-related decision-making processes. In the literature, Makridakis et al. [9] described four approaches used for forecasting revenue by both state and local governments, as follows:

(i) Qualitative method: This is usually a combination of expert judgment and formal methods of forecasting.

(ii) Time-series methods: The traditional approaches are simple trend approach, the autoregressive approach, and the mixed technique approach.

(iii) Econometric modelling: This is usually represented by a single equation; a set of equations solved recursively; and sets of equations solved simultaneously.

(iv) Microsimulation modelling: Revenue forecasting for attribute data (Personal Income tax, Particular Commodity) can be predicted and the policy impacts on certain groups or commodities can be predicted using the microsimulation model.

Anderson and Johnson [2] assessed revenue collection of the Customs Excise and Preventive Service (CEPS) for the period 2008–2012 to devise a reasonably accurate projection for the individual tax components. They indicated that the projection will assist in the design of an appropriate expenditure profile as a means of averting any future fiscal deficit in the country.

Molapo et al. [10] used a Bayesian Vector Autoregression model, an ARIMA model, and a State Space Exponential Smoothing Model to forecast quarterly tax revenue data in South Africa. All three times series models outperformed existing forecasting techniques with the Bayesian Vector Autoregression Model having the best forecast accuracy amongst the three.

Chen et al. [11] in their work proposed an automatic approach by harnessing the power of machine learning techniques to relieve the burden of customs targeting officers who set revenue targets. They introduced a novel model based on an off-the-shelf embedding encoder to identify the correctness of a Harmonized Commodity Description and Coding System (HS Code) without any human effort.

Currently, the GRA-CD uses the Duke University's GDP-Based model and Monthly Receipts model, as well as an in-house GDP-Based model and the microsimulation model (see Jenkins et al. [12] for more information about the models). These are econometric and microsimulation

models and hence can be inadequate or deficient for forecasting due to their inability to treat serially correlated time series data.

In quest to study the causes of decline in revenue from GRA-CD, this study will identify the important variables for forecasting revenue. The study will also assess the performance of some traditional and machine learning time series techniques which are suitable for time series data in modelling the revenue collected by the GRA-CD, as the selection of an appropriate forecasting technique is critical to the success of any new initiative or the decision-making process.

The rest of the paper is organised as follows: Section 2 (Material and Methods) discusses the source of data, the mathematical underpins of the adopted time series models and their implementation. In Section 3 (Results and Discussion), we compare the results of the traditional and machine learning time series models used to forecast revenue collected by the GRA-CD. Finally, we conclude by summarising the overall achievements of the study with some recommendations and directions for future developments in Section 4 (Conclusion and Recommendation).

## 2. Methods and Materials

*2.1. Data Acquisition.* Secondary data obtained from the GRA-CD and from the Bank of Ghana historical exchange rates records [13] were used for the study. The Customs data comes from a database containing financial indices of performance and growth of the Customs Division. The data captures records on the following tax variables: the value of imports (CIF) data, the Zero and Exempt portion of value of imports (ZE), the Taxable or Effective value portion of CIF (ECIF), the Effective Duty Rate (EDR), the import duty revenue (IDREV), Total Customs Revenue (TOTREV), and the average monthly Exchange Rate (XCHR) from 2010 to 2019.

*2.2. Definition of Study Variables.*

Cost, Insurance, and Freight (CIF): This can be defined as the value of imports up to the port or location where the goods would be entered for clearance and customs purposes. This includes the cost of purchase of the item (FOB), freight payable/paid to country of destination (Ghana), and insurance (usually calculated as 0.875% of cost plus freight). In Ghana, the tax base for calculating Import Duty is CIF. Thus, CIF is the value of all goods imported under the customs regime "Home consumption." This is the total value of goods imported under the import duty rates 0%, 5%, 10%, 20%, and 35%.

Zero and Exempt (ZE): This connotes the value of imported goods whose import duties are zero-rated (0%) and the value of goods whose import duties have been exempted from payment of taxes.

Effective CIF (ECIF): This represents CIF for goods whose tax rates are five percent (5%), ten percent (10%),

TABLE 1: Sample of the study data.

| Month | CIF | ZE | ECIF | EDR | IDREV | TOTREV | XCHR |
|---|---|---|---|---|---|---|---|
| Jan-10 | 943.19 | 352.88 | 590.31 | 6.63 | 68.33 | 143.25 | 1.42 |
| Feb-10 | 1229.56 | 659.64 | 569.92 | 4.68 | 67.85 | 142.04 | 1.42 |
| Mar-10 | 1368.67 | 670.93 | 697.74 | 5.18 | 83.46 | 172.98 | 1.42 |
| Apr-10 | 1134.18 | 543.99 | 590.19 | 5.27 | 73.44 | 155.29 | 1.41 |
| May-10 | 1078.69 | 481.63 | 597.06 | 5.79 | 71.67 | 154.25 | 1.41 |
| Jun-10 | 1226.05 | 530.87 | 695.18 | 5.83 | 78.24 | 166.14 | 1.41 |
| Jul-10 | 1172.36 | 451.14 | 721.22 | 6.42 | 83.7 | 180.08 | 1.42 |
| Aug-10 | 1398.91 | 610.02 | 788.89 | 5.9 | 89.96 | 191.62 | 1.42 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Oct-11 | 1830.22 | 671.21 | 1159.01 | 6.6 | 138.93 | 290.17 | 1.53 |
| Nov-11 | 1814.09 | 686.87 | 1127.22 | 6.64 | 146.74 | 308.04 | 1.54 |
| Dec-11 | 1974.15 | 912.32 | 1061.83 | 5.89 | 152.06 | 317.41 | 1.56 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Oct-19 | 3725.87 | 1346.75 | 2379.12 | 7.27 | 432.25 | 840.9 | 5.32 |
| Nov-19 | 3200.34 | 983.87 | 2216.48 | 8.19 | 438.93 | 855.19 | 5.39 |
| Dec-19 | 3510.3 | 1102.26 | 2408.04 | 8.2 | 481.24 | 934.93 | 5.53 |

The revenue figures presented in the table are in millions of Ghanaian cedis.

twenty percent (20%), and thirty-five percent (35%) and are also not exempted from import duty taxes. In other words, this is CIF for which import duty taxes have been paid.

Effective Duty Rate (EDR): This is the sum of the ratio of value of import per import duty rates to the total value of import (CIF). Mathematically, EDR can be represented as

$$EDR = \frac{\text{Zero rated CIF} \times 0\% + 5\% \text{ rated CIF} \times 5\% + \cdots + 35\% \text{ rated CIF} \times 35\%}{\text{Total CIF}} \times 100. \tag{1}$$

Average Monthly Exchange Rate (XCHR): This is the historical monthly Ghanaian cedi to US dollar exchange rate.

Table 1 contains a sample of the study data (revenue in millions of Ghanaian cedis and their respective Ghanaian cedis to US dollar exchange rates) collected by Customs Division in Ghana from January 2010 to December 2019. The obtained data from January 2010 to December 2019 was used to fit the study models (training and testing of the models). In this study, 70% of the study data was used for the model building (training) and the remaining 30% was used for testing the fitted models. The fitted models were then used to forecast the revenue for the year 2020.

*2.3. Autoregressive Integrated Moving Average (ARIMA) Models.* The ARIMA model is a linear model in which the independent variables are the lags of the dependent variable and/or the lags of the errors from forecast. The general form of an ARIMA model using backshift notation is

$$\phi_p(B)(1 - B)^d X_t = \theta_q(B)\varepsilon_t, \tag{2}$$

where $X_t$ is the univariate time series data, $d$ is the order of differencing, $B$ is the backshift notation, $p$ and $q$ are the

number of lags of $X_t$ and the error, respectively, and $\varepsilon_t$ is the error at time $t$.

The Box–Jenkins methodology for fitting time series models is then used for model identification through to the estimation of model parameters.

*2.4. ARIMA Error Regression Model (ARIMAX).* The assumption of the errors underpinning the ordinary least-squares (OLS) models (i.e., the errors are independent and identically distributed (iid)) is often violated as the residuals from time series analysis are usually autocorrelated. The ARIMAX model fixes this by modelling the error with an ARIMA process.

$$Y_t = x_t'\beta + \eta_t, \tag{3}$$

$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \phi_1 B)\varepsilon_t, \tag{4}$$

where $\beta$ is the vector of coefficients of $x_t$, $\eta_t$ is the error for the regression model in (3) and modelled as an ARIMA with error $\varepsilon_t$ with $\varepsilon_t \sim WN(0, \sigma^2)$.

*2.5. Bayesian Structural Time Series (BSTS) Model.* The Bayesian Structural Time Series Model can be defined as

$$y_t = \mu_t + \tau_t + \beta' x_t + \varepsilon_t,$$
$$\mu_{t+1} = \mu_t + \delta_t + \eta_{0t},$$
$$\delta_{t+1} = \delta_t + \eta_{1t}, \tag{5}$$
$$\tau_{t+1} = -\sum_{s=1}^{S-1} \tau_t + \eta_{2t}.$$

The trend component is the local level model plus a $\delta$ term that can be defined as the extra expected $\mu$ as $t \longrightarrow t + 1$ so that it is the slope of the local linear trend. The $\eta$ values are the errors from the local linear trend and the seasonal component.

### 2.6. Neural Network Autoregression (NNAR) Model.
Neural networks are nonlinear models that can be trained on data to learn patterns. They have been applied as successful machine learning models for solving various tasks like pattern recognition. The three main types of neural network architecture are the Feedforward, Recurrent [14], and Convolutional Neural Networks [15].

The architecture is called feedforward because the flow of information takes place in the forward direction. A feedforward network defines a mapping $y = f(x; k)$ and learns the value of the parameters $k$ that result in the best function approximation [16].

According to Chapman-Wardy et al. [17], a feedforward can either be single or multiple layered. The single-layer network consists of only one hidden layer. The multiple-layer neural network is also known as the deep learning network. The distinguishing feature of this network is the fact that it has multiple hidden layers for complex processing.

The data is first sent to an input layer where it is processed and forwarded to each neuron in the next layer until it reaches the output layer. Each neuron receives an input vector $x = (x_1, x_2, \ldots, x_n)$ and multiplies it by a specific weight, and a bias is added to this product resulting in an output that is fed to the next layer with the function

$$z = \sum_{i=1}^{n} x_i w_i + b, \tag{6}$$

where $x_i$ is the $ith$ input feature from the data or a neuron, $z$ is the net input of each neuron, $w_i$ is the weight of the $ith$ input, and $b$ is the bias. The net input is then passed through an activation function $g$, to produce the output:

$$g(z) = g\left(\sum_{i=1}^{n} x_i w_i + b\right). \tag{7}$$

An activation function is chosen by the researcher according to the purposes of the research. In this study, the sigmoid activation function was chosen as it has the tendency to reduce the effect of extreme values, hence, robust to outliers.

$$g(z) = \frac{1}{1 + e^{-z}}. \tag{8}$$

### 2.7. Evaluating Time Series Models.
The residuals of the fitted models are important in selecting the appropriate time series model. The more the residuals look like white noise, the better the fit of the model. A large $p$-value ($p > 0.05$ at 5% level of significance) corresponding to a Ljung–Box portmanteau test value from analysis of the residual suggests the tendency of the data to be white noise.

Another important statistic or measure of good fit for a time series model is the Akaike Information Criteria (AIC), which is defined as

$$\text{AIC} = 2k - \ln(\widehat{L}), \tag{9}$$

where $k$ is the number of parameters in the model and $\widehat{L}$ signifies the maximum value of the likelihood estimates of the function. The AIC seeks to increase the likelihood or goodness of fit of the model, $L$, while reducing the complexity of parameters, $k$. Models with lower AIC values are preferred. The Bayesian Information Criterion (BIC) also known as the Schwarz Criterion is a similar statistic that is usually used hand in hand with AIC. The BIC, however, penalises the addition of more parameters than the AIC.

$$\text{BIC} = -2\widehat{L} + k \ln(N), \tag{10}$$

where $N$ is the sample size of the training set.

A model with relatively small AIC or BIC value is always preferred.

Also, the Mean Squared Error (MSE) and the Mean Absolute Percentage Error (MAPE) could be used to assess all the model forecasts to judge the best performing model.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i\right)^2,$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|Y_i - \widehat{Y}_i|}{Y_i}, \tag{11}$$

where $Y_i$ is the actual value and $\widehat{Y}_i$ is the predicted value.

## 3. Results and Discussion

### 3.1. Time Series Plots of Variables.
A time series plot of the variables CIF, ZE, ECIF, IDREV, and TOTREV is shown in Figure 1. It can be seen from Figure 1 that there was a gradual increase in revenue for all the variables under study prior to 2018. There was a spike in 2018 and a tardy decline in revenue collected after 2019.

### 3.2. Results of the ARIMA Model.
Table 2 contains a summary of the most suitable ARIMA model for forecasting revenue collected by GRA-CD.

From Table 2, the fitted ARIMA (0, 1, 1) model is given as

$$X_t = 8.6917 + \varepsilon_t - 0.4344\varepsilon_{t-1}. \tag{12}$$

The AIC and BIC of the model are 1133.50 and 1141.52, respectively. The $p$-value of the Ljung–Box test of residual is 0.01469 ($p < 0.05$). This means the residuals of the fitted model are correlated. As stated earlier, if the residuals are white noise (uncorrelated), then the fitted model fits the data
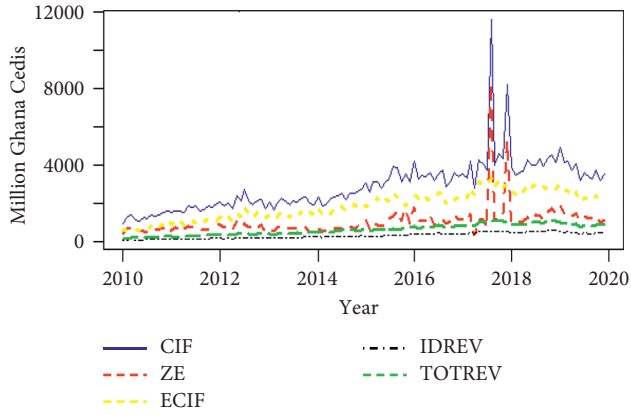
FIGURE 1: Monthly time series of data.

TABLE 2: Summary of ARIMA (0, 1, 1) model.

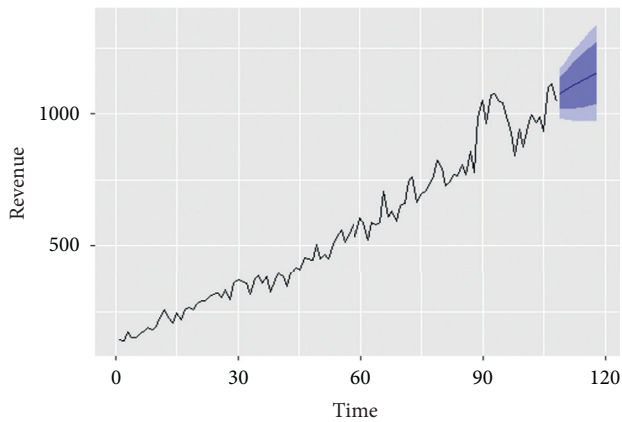| ARIMA (0, 1, 1) with drift | Coefficients | SE |
|---|---|---|
| MA 1 | −0.4344 | 0.0933 |
| Drift | 8.6917 | 2.5850 |
| Sigma$^2$ estimate | 2245.00 | |
| Log likelihood | 563.75 | |
| AIC | 1133.50 | |
| BIC | 1141.52 | |
| Ljung–Box test $p$-value of residuals | 0.01469 | |



FIGURE 2: A plot of revenue forecast from ARIMA model.

well. By inference, ARIMA (0, 1, 1) is not very suitable for forecasting the revenue collected by GRA-CD.

Figure 2 shows a graph of the revenue forecasted from the ARIMA model.

It can be seen from Figure 2 that the forecast points show an increasing trend for the forecast period. The confidence band around the forecast is wide, which confirms the unreliability of the ARIMA model for forecasting the revenue collected by GRA-CD.

### 3.3. Results of the ARIMAX Model.
The results of the most suitable ARIMA Error regression model (ARIMAX) with statistically significant covariates are presented in Table 3.

TABLE 3: Summary of ARIMA error regression model.

| Regression with ARIMA (1, 1, 1) errors | Coefficients | SE |
|---|---|---|
| AR 1 | 0.3871 | 0.1250 |
| MA 1 | −0.8652 | 0.0661 |
| Drift | 7.1504 | 1.0025 |
| EDR variable | 19.3621 | 6.1251 |
| CIF variable | 0.0314 | 0.0067 |
| $\sigma^2$ estimate | 1845.00 | |
| Log likelihood | 551.95 | |
| AIC | 1115.91 | |
| BIC | 1131.95 | |
| Ljung–Box test $p$-value of residuals | 0.3019 | |

Target: revenue generated by GRA-CD.
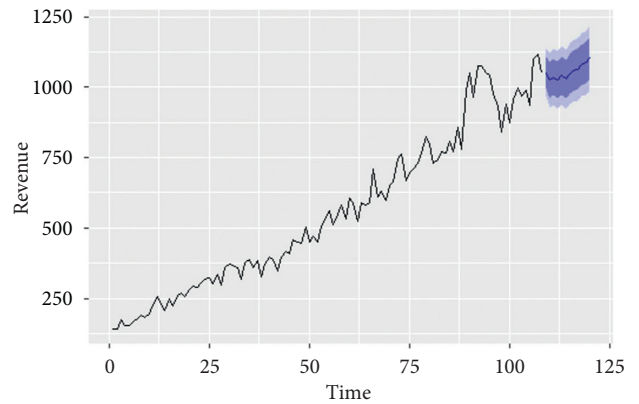


FIGURE 3: A plot of revenue forecast from ARIMAX model.

From Table 3, the resulting regression with ARIMA (1, 1, 1) error model is given by

$$Y'_t = 19.3621X'_{1,t} + 0.0314X'_{2,t} + \eta'_t,$$
$$\eta'_t = \phi_0 + 0.3871\eta_{t-1} + \varepsilon_t - 0.8652\varepsilon_{t-1}, \qquad (13)$$
$$\varepsilon_t \sim WN(0, 1845).$$

From Table 3, the significant variables required for forecasting revenue collected by GRA-CD using the ARIMAX model is Effective Duty Rate (EDR) and Cost, Insurance, and Freight (CIF). The most suitable ARIMAX model was the regression with ARIMA (1, 1, 1) errors. The $p$-value of the Ljung–Box test is 0.3019 ($p > 0.05$). This means that the residuals of the regression with ARIMA (1, 1, 1) errors are uncorrelated (white noise). This signifies that the fitted model is somewhat suitable for forecasting the revenue collected by GRA-CD. The AIC and BIC of the regression with ARIMA (1, 1, 1) errors model are 1115.91 and 1131.95, respectively.

Figure 3 shows a forecast plot of revenue using the regression with ARIMA (1, 1, 1) errors model.

It can be seen from Figure 3 that the forecast points of the ARIMAX model also show an increasing trend for the forecast period. The confidence band around the forecast is narrower as compared to that of the ARIMA model. This means the ARIMAX model gives more precise estimates/forecasts than the ARIMA model.

TABLE 4: Comparison of various BSTS components.

|  | Residual SD | $R^2$ | Harvey's goodness of fit |
|---|---|---|---|
| BSTS Model 1 | 24.78 | 0.992 | 0.27 |
| BSTS Model 2 | 23.13 | 0.993 | 0.27 |
| BSTS Model 3 | 14.01 | 0.997 | −0.46 |
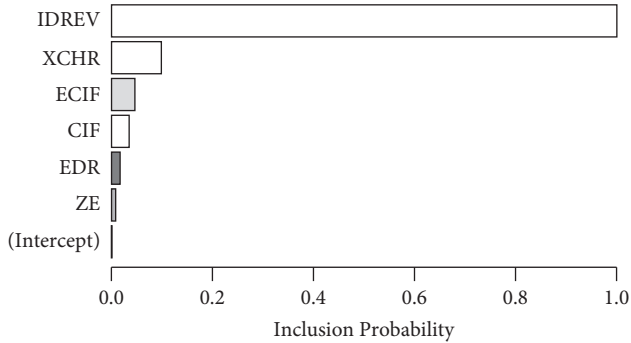


FIGURE 4: Important regression variables of the BSTS model.



FIGURE 5: Forecast of revenue using the BSTS model.

TABLE 5: Neural network autoregression model summary.

| NNAR (1, 3) model | |
|---|---|
| Neuron network | $5 - 3 - 1$ |
| $\lambda$ | 1 |
| $\sigma^2$ estimate | 0.002699 |



FIGURE 6: Forecast of revenue data with NNAR (1, 3) model.

*3.4. Results of the BSTS Model.* Three BSTS models of different compositions are first considered: the local level model, which considers local linear trend and annual seasonality (BSTS Model 1); the local level model with added autoregression (BSTS Model 2); and a BSTS model with a local linear trend, annual seasonality, and a regression component (BSTS Model 3) (see Scott and Varian [18] for further details). Table 4 contains the summary results of the various BSTS models considered in this study.

From Table 4, the fitted Bayesian Structural Time Series models (BSTS Model 1, BSTS Model 2, and BSTS Model 3) had coefficients of determination of 0.992, 0.993, and 0.997, respectively. The BSTS model with a local linear trend, annual seasonality, and a regression component (BSTS Model 3) had the smallest residual standard deviation of 14.01 and the highest coefficient of determination of 0.997. This makes the BSTS Model 3 the best amongst the Bayesian Structural Time Series models for forecasting the revenue collected by GRA-CD. The model as fitted explains 99.7% of the variability in the response variable (revenue data).

The spike-and-slab technique was used to identify the most important variables for modelling the data using their inclusion probabilities. Figure 4 shows some input variables ranked according to their level of importance for the BSTS model.

From Figure 4, Import Duty Revenue (IDREV) had the highest inclusion probability, followed by the average monthly Exchange Rate (XCHR) and Taxable or Effective value portion of CIF (ECIF). This makes IDREV, XCHR, and ECIF the top 3 most important variables required to fit the BSTS model. The Zero Exempt (ZE) had the least inclusion probability in fitting a BSTS model for forecasting revenue. Figure 5 shows the forecast diagram of the best BSTS model (BSTS Model 3).

It is evident from Figure 5 that the forecast shown in blue has a narrower confidence band (green boundaries) than the ARIMA and ARIMAX models. This means the estimates/
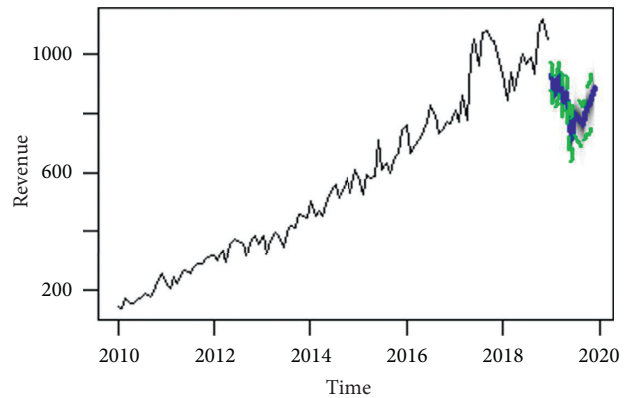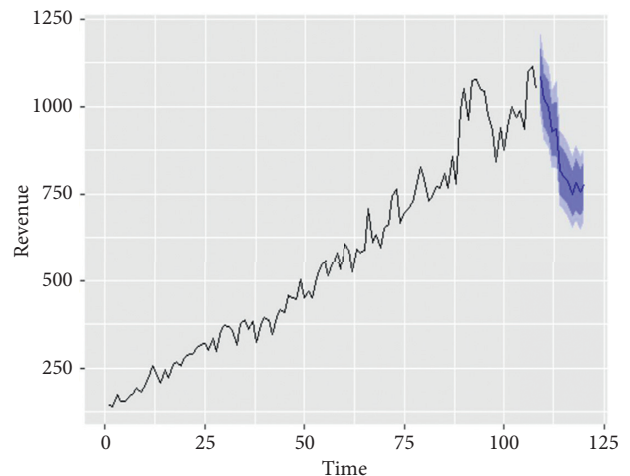
forecasts from the BSTS model are more precise than those from the traditional time series models (ARIMA and ARIMAX). Also, there was a gradual increase in revenue generated from 2010 till 2018 and a steady decline after 2018 that was followed by a gradual increase after 2019.

*3.5. Results of Neural Network Autoregression (NNAR) Model.* The most suitable NNAR model using the study data is an NNAR (1, 3) model. The results of NNAR (1, 3) model are summarised in Table 5.

Table 5 contains the results of a Neural Network Autoregression model with 5 input variables, 3 hidden layers, and 1 output variable (revenue data). The forecast of the NNAR (1, 3) model for the next annual period is illustrated in Figure 6.

TABLE 6: Assessment of study models.

| Model | MSE | MAPE |
| --- | --- | --- |
| ARIMA | 75016.21 | 0.3089 |
| ARIMA error regression model (ARIMAX) | 40162.14 | 0.2223 |
| BSTS | 1319.989 | 0.0390 |
| NNAR (1, 3) | 53.87 | 0.0861 |

The thin band of confidence interval shown in Figure 6 indicates that the NNAR (1, 3) model is reliable in forecasting revenue. It can be seen from Figure 6 that there was an increase in revenue collected over time and a sudden decline at the forecast period (2019 to 2020).

*3.6. Evaluation of Forecasts.* Comparisons of the forecast ability of the study models, using a cost function of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE), are summarised in Table 6.

Table 6 contains the MSE and MAPE values of the study models. These errors measure the deviation of the various model forecasts from the test data. Models with relatively lower MSE and MAPE are preferred. It is evident from Table 6 that NNAR (1, 3) model produces the smallest MSE (53.87) and an appreciable MAPE (0.0861) whereas the BSTS model gives the smallest MAPE (0.0390) and an appreciable MSE (1319.989).

From the above results, the NNAR (1, 3) is adjudged the most suitable model for forecasting revenue collected by GRA-CD as it outperformed the other models with the lowest MSE of 53.87 and relatively better MAPE of 0.0861. It can also be seen from Table 6 that the machine learning time series algorithms (NNAR (1, 3) and BSTS model with a local linear trend, annual seasonality, and a regression component outperformed the traditional time series models (ARIMA and ARIMAX).

## 4. Conclusion and Recommendations

Comparing the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) of the forecasts, it was evident the NNAR (1, 3) model was the best fitted model with the least MSE of 53.87 and a comparatively lower MAPE of 0.0861. This makes forecast estimates from the NNAR (1, 3) the most precise as compared to estimates from the other models considered in the study. For this model, the important variables were the Effective Duty Rate (EDR) and the total value of imports (CIF). The forecast values from the NNAR (1, 3) indicated a potential decline in revenue. This result calls for policies geared at improving revenue generations in the immediate future.

The BSTS model also performed better than the traditional ARIMA and ARIMAX models. Although the model gave a comparatively higher MSE of 1319.989, it was better than the traditional time series models (ARIMA and ARIMAX). Also the BSTS model gave the lowest MAPE of 0.039, making it also suitable for forecasting the revenue collected by GRA-CD. One advantage of the BSTS model is its ability to incorporate external data (covariates) to obtain the posterior distributions for the forecasts. The Import Duty factors (IDREV), Exchange Rate (XCHR), and Taxable or Effective value portion of CIF (ECIF) were the top 3 most important variables required for modelling using BSTS model.

Overall, the two machine learning models (BSTS and NNAR) outperformed the traditional time series models (ARIMA and ARIMAX) in forecasting the revenue collected by GRA-CD. This could be attributed to the fact that the ARIMA and ARIMAX models are relatively simpler models that require some further modifications to be suitable for the analysis of the study data. Also, in cases where the data is nonlinear, simple econometric models like the ones currently used by the GRA-CD for forecasting might fail.

The machine learning time series models considered in this study are robust and can give more reliable forecasts or estimates when updated with new data.

The pace of statistical research and data science has led to advancements in machine learning techniques like the ones considered in this study. Exploring these new methods and constantly modifying them to suit real-world big data should be a desired quest for future research.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Authors' Contributions

Diana Ayorkor Agbenyega contributed to conceptualization, methodology, writing the original draft. John Andoh carried out data curation and formal analysis. Samuel Iddi and Louis Asiedu performed supervision, validation, and reviewing and editing.

## References

[1] I. Ayuma, *The Impact of Emerging Technology on International Trade: A Case Study of the Three-Dimensional (3D) Printing and Customs Authorities in the East African Community*, Ph.D. Thesis, University of Nairobi, Nairobi, Kenya, 2018.

[2] F. N. Anderson and J. K. A. Johnson, "Revenue forecasting using trend analysis," *Mathematical Theory and Modeling*, vol. 4, pp. 166–172, 2014.

[3] W. Prichard, *Taxation and Development in Ghana: Finance, Equity and Accountability*, Tax Justice Network, London, UK, 2009.

[4] Y. Deli, A. G. Rodriguez, I. Kostarakos, and P. Varthalitis, "Dynamic tax revenue buoyancy estimates for a panel of OECD countries," ESRI Working Paper, No. 592, Economic and Social Research Institute, 2018.

[5] E. Abrokwah, E. Bob-Eshun, H. Conron, A. M. Iddrisu, D. Nuer, and R. Warwick, *Customs Revenue in Ghana: Recent Trends and Their Causes*, Economic and Social Research Council, Swindon, UK, 2021.

[6] E. Crivelli, "Trade liberalization and tax revenue in transition: an empirical analysis of the replacement strategy," *Eurasian Economic Review*, vol. 6, no. 1, pp. 1–25, 2016.

[7] A. Owusu-Gyimah, "Tax revenue generation and the economic development of Ghana," *European Journal of Business and Management*, vol. 7, pp. 78–89, 2015.

[8] R. I. Whitfield and A. H. B. Duffy, "Extended revenue forecasting within a service industry," *International Journal of Production Economics*, vol. 141, no. 2, pp. 505–518, 2013.

[9] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting Methods and Applications*, John Wiley & Sons, Hoboken, NJ, USA, 2008.

[10] M. A. Molapo, J. O. Olaomi, and N. O. Ama, "Bayesian vector auto-regression method as an alternative technique for forecasting south african tax revenue," *Southern African Business Review*, vol. 23, 2019.

[11] H. Chen, B. Van Rijnsoever, M. Molenhuis, D. van Dijk, Y.-H. Tan, and B. Rukanova, "The use of machine learning to identify the correctness of hs code for the customs import declarations," in *Proceedings of the 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–8, IEEE, Porto, Portugal, 2021.

[12] G. P. Jenkins, C.-Y. Kuo, and G. Shukla, *Tax Analysis and Revenue Forecasting*, Harvard Institute for International Development, Harvard University, Cambridge, MA, USA, 2000.

[13] Bank of Ghana, "Historical interbank fx rates—bank of Ghana," 2018, https://www.bog.gov.gh/treasury-and-the-markets/historical-interbank-fx-rates/.

[14] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[15] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0588–0592, IEEE, Chennai, India, 2017.

[16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, MA, USA, 2016.

[17] C. Chapman-Wardy, L. Asiedu, K. Doku-Amponsah, and F. O. Mettle, "Modeling the amount of waste generated by households in the greater accra region using artificial neural networks," *Journal of Environmental and Public Health*, vol. 2021, Article ID 8622105, 12 pages, 2021.

[18] S. L. Scott and H. R. Varian, *Bayesian Variable Selection for Nowcasting Economic Time Series*, University of Chicago Press, Chicago, IL, USA, 2015.