SLIIT
2023 SICASH
SLIIT INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SCIENCES & HUMANITIES

Forecasting of Constant GDP per capita of Sri Lanka using ARIMA model

[1]Wijesinghe, W.R.A.N.M, [2]Umayangi, K.A.S, [3]De Silva, H.D.K, [4] Mundigalage, S.D.M., [5]Peiris, T. S. G

[1,2,3,4,5]Department of Mathematics and Statistics, Faculty of Humanities and Sciences,
SLIIT, Malabe, Sri Lanka

Corresponding author - HS21939082@my.sliit.lk

ARTICLE INFO

Article History:
Received: 10 September 2023
Accepted: 01 November 2023

Keywords:
ARIMA models; Forecasting; GDP per capita

Citation:

Wijesinghe, W.R.A.N.M, Umayangi, K.A.S, De Silva, H.D.K, Mundigalage, S.D.M., Peiris, T. S. G (2023). Forecasting of Constant GDP per capita of Sri Lanka using ARIMA model. Proceedings of SLIIT International Conference on Advancements in Sciences and Humanities, 1-2 December, Colombo, pages 269-274.

ABSTRACT

GDP per capita is a global measurement for assessing the economic prosperity of nations. Constant (Real) GDP per capita eliminates the effects of inflation which allows for a more accurate comparison of GDP per capita over time. However, no statistical models have been developed to predict annual constant GDP per capita (CGDPC) in Sri Lanka. In this study, ARIMA (1,1,0) model was developed using past data from 1961 to 2018 to forecast CGDPC. The best-fitted model was identified based on three possible models using sample ACF and sample PACF of the stationary series and comparing statistics such as AIC, BIC, maximum log-likelihood, and volatility. The residuals of the fitted model were white noise. The training dataset has percentage errors ranging from -6.50% to 3.80%. The model was validated for observed data in 2019, 2020, and 2021. The percentage error for the three points were -3.49, -6.10, and 1.49 respectively. The forecasted values for 2022, 2023, and 2024 obtained were 4506.728, 4653.895, and 4810.505 respectively showing that Sri Lanka's economy is expected to grow due to the increase in CGDPC. The GDP per capita growth rates of 2.99%, 3.27%, and 3.37% for the next 3 years also confirm this. The results obtained from this study can be effectively used for better planning. However, it is recommended to improve

HS21939082@my.sliit.lk

https://doi.org/10.54389/SHPE7237

the model further to reduce the percentage of errors using the ARIMAX approach.

1. INTRODUCTION

Sri Lanka, a vibrant island nation located in South Asia, boasts a diverse economy that has evolved over the years. Over the decades, Sri Lanka has witnessed both periods of robust economic growth and challenges. To understand and analyze Sri Lanka's economic health and overall development, the Gross Domestic Product (GDP) serves as a vital indicator by providing valuable insights into the nation's economic performance. Amongst macroeconomic variables, the topic of GDP per capita which is a measure of a country's economic output per individual, has emerged as highly significant (The Investopedia Team, 2023).

Various professions like economists, legislators, central bankers, and governments utilize GDP per capita to gain insights into the well-being of a nation, make fiscal policy decisions, and understand population growth (The Investopedia Team, 2023). Various factors such as inflation, population growth, technological advancements, and economic geography play significant roles in influencing GDP per capita. Thus, for a meaningful comparison of GDP per capita, the impact of inflation is removed using constant prices which allows to determine whether the growth in GDP is due to rising prices or an actual increase in the population. This is known as Real (Constant) GDP per capita (Tramplin, 2023).

Prior knowledge of CGDPC has many advantages such as to achieve a deeper understanding of the country's economic performance, and to improve people's living standards. Thus, many authors in different countries have attempted to develop models to forecast the GDP and CGDPC using various methods such as ARIMA, SARIMA, Multivariate time series models, Dynamic factor models, Seasonal autoregressive fractional integrated moving average (SARFIMA) and non-linear models like SETAR and Artificial Neutral Network (ANN). Out of these methods, the ARIMA model is widely utilized in Sri Lanka and other countries as the preferred method for GDP forecasting. A literature survey (Dilhani, 2017; Eissa, 2020; Maity & Chatterjee, 2012; Ranasinghe & Suriyaarachchi, 2015; Saha & Khanam, 2022) found that no study has been conducted to predict CGDPPP in Sri Lanka. Accordingly, the objective of this study is to develop a forecasting model for Sri Lanka's CGDPC, validate the model, and predict future CGDPC.

2. MATERIALS AND METHODS

Univariate ARMA (p,d,q) model suggested by Box-Jenkins (Box, Jenkins, Reinsel, & Ljung, 2015) is the approach used for developing a time series model for the observed data series. The ARMA (p,d,q) model for the time series {Yt} is represented as:

$$Y_t = \mu + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots\dots\dots + \varphi_p Y_{t-p} \cdot e_t - \theta_1 e_1 - \theta_2 e_2 - \dots\dots - \theta_q e_{t-q}$$

(1)

Where: is a purely random process with mean zero and constant variance,

four steps that are used to determine the ARIMA model are identification of possible models, determining the best-fitted model, model diagnostic, and forecasting future values.

2.1. DATA

GDP per Capita data (constant 2010 US dollars) of Sri Lanka was extracted annually from 1961 to 2021 from the Federal Reserve Economic Data (FRED) with 61 observations. This statistical analysis was performed using Eviews-12 software.

3. RESULTS AND DISCUSSION

3.1. TEMPORAL VARIABILITY OF THE ORIGINAL SERIES

Figure 1 which illustrates the temporal variability of GDP per capita of Sri Lanka varies from a minimum of 645.8449 and a maximum of 4495.711, with a mean of 1731.921 and a median of 1279.813. The Jarque-Bera test statistic is 11.934 (p = 0.002) confirmed that the original series is significantly deviated from normality. The original series is not stationary as the ADF test is not significant (test statistic = 3.65 p = 1.0). Since the original series is exponential, a log transformation was used to transform into linear data making it easier to identify the trend.

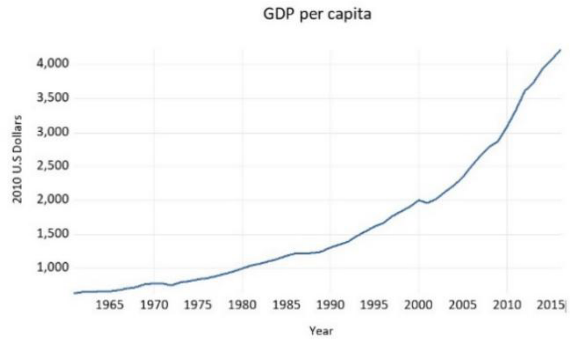
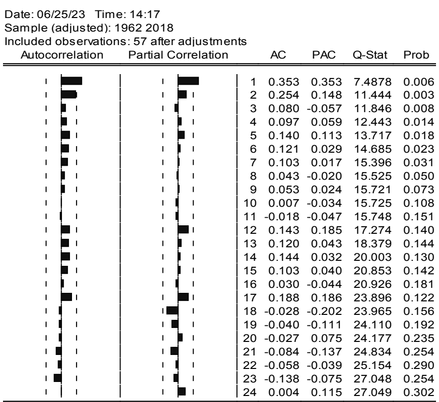


Figure 1. The time series plot of GDP per capita of Sri Lanka from 1961 to 2018

Since the original series is not stationary to achieve stationarity, first-order differencing was taken. The ADF test for the 1st differenced series confirmed stationarity (test statistic = -5,037, p = 0.0001). Therefore, to find suitable models ACF and PACF were obtained for the stationary series in Figure 2.

Identification of the possible models



Date: 06/25/23 Time: 14:17
Sample (adjusted): 1962 2018
Included observations: 57 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.353	0.353	7.4878	0.006
		2	0.254	0.148	11.444	0.003
		3	0.080	-0.057	11.846	0.008
		4	0.097	0.059	12.443	0.014
		5	0.140	0.113	13.717	0.018
		6	0.121	0.029	14.685	0.023
		7	0.103	0.017	15.396	0.031
		8	0.043	-0.020	15.525	0.050
		9	0.053	0.024	15.721	0.073
		10	0.007	-0.034	15.725	0.108
		11	-0.018	-0.047	15.748	0.151
		12	0.143	0.185	17.274	0.140
		13	0.120	0.043	18.379	0.144
		14	0.144	0.032	20.003	0.130
		15	0.103	0.040	20.853	0.142
		16	0.030	-0.044	20.926	0.181
		17	0.188	0.186	23.896	0.122
		18	-0.028	-0.202	23.965	0.156
		19	-0.040	-0.111	24.110	0.192
		20	-0.027	0.075	24.177	0.235
		21	-0.084	-0.137	24.834	0.254
		22	-0.058	-0.039	25.154	0.290
		23	-0.138	-0.075	27.048	0.254
		24	0.004	0.115	27.049	0.302

Figure 2. ACF and PACF of the stationary series

To identify the possible models ACF and PACF plots show that the first lag is statistically significant. Therefore, three possible models are ARIMA (1,1,0), ARIMA (1,1,1), and ARIMA (0,1,1).

3.2. COMPARISON OF SELECTED MODELS

The indication used to compare the 3 selected models are shown in Table 1. When comparing the three models, ARIMA(1,1,1) and ARIMA(1,1,0) are significantly better. ARIMA(1,1,1) has the lowest volatility and the maximum log-likelihood whereas ARIMA(1,1,0) has the lowest AIC and SBIC values. When comparing the parameters, both the models' constant is significant, however, in ARIMA(1,1,1) the MA(1) parameter is not significant. Therefore, ARIMA(1,1,0) was chosen as the best-fitted model.

Table 1. Comparison of estimated models

	ARIMA (1,1,1)	ARIMA (0,1,1)	ARIMA (1,1,0)
Parameter AR(1)	Significant	-	Significant
Parameter MA(1)	Not Significant	Significant	-
Volatility	0.000394	0.000419	0.000402
AIC	-4.857916	-4.833256	-4.873522
SBIC	-4.714544	-4.725727	-4.765993
Log-likelihood	142.4506	140.7478	141.8954

3.3. BEST-FITTED MODEL

Table 2 depicts the statistics of the best-fitted model ARIMA (1,1,0). Since the p-values are less than 5% significance level, all the parameters are significant. The equation of the model is

$$\text{Log GDP}_t = \text{LGDP}_t$$

$$\text{LGDP}_t = 0.033638 + 1.359137\ \text{LGDP}_{t-1} - 0.359137\text{LGDP}_{t-2} \quad (2)$$

Table 2. ARIMA (1,1,0)

Dependent Variable: D(LGDP)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 06/25/23 Time: 14:25
Sample: 1962 2018
Included observations: 57
Convergence achieved after 11 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.033638	0.004372	7.693269	0.0000
AR(1)	0.359137	0.099616	3.605199	0.0007
SIGMASQ	0.000402	7.27E-05	5.527794	0.0000

R-squared	0.129079	Mean dependent var	0.034041
Adjusted R-squared	0.096823	S.D. dependent var	0.021675
S.E. of regression	0.020599	Akaike info criterion	-4.873522
Sum squared resid	0.022913	Schwarz criterion	-4.765993
Log likelihood	141.8954	Hannan-Quinn criter.	-4.831733
F-statistic	4.001676	Durbin-Watson stat	2.054187
Prob(F-statistic)	0.023956		

Inverted AR Roots	.36

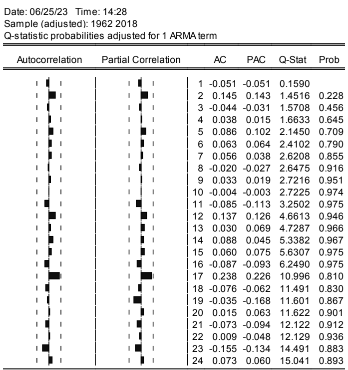
3.4. MODEL DIAGNOSTICS

It was found that the AR root is inside the unit circle, thus the ARIMA process is stationary and invertible. Diagnostics tests are performed after fitting the ARIMA (1,1,0) model. According to Figure 3, it can be seen the Q-statistic probabilities are not statistically significant, with the correlogram indicating that each lag lies within a 95% confidence interval. Therefore, the residuals are random. (JB test statistic= 3.267, p = 0.195) is not significant. Thus, it can be concluded with 95% confidence that the residuals are not significantly deviated from normality. Thus, the errors are white noise.

3.5. FORECASTING

Comparison of observed and predicted values – training and validation data sets

The root mean squared error value (RMSE) is 44.196 and the mean absolute error is 29.474. The Theil inequality coefficient is 0.0106 which is close to 0 meaning that the model's predictions almost match the observed values. Thus, the predicting power of the model is good.


Date: 06/25/23 Time: 14:28
Sample (adjusted): 1962 2018
Q-statistic probabilities adjusted for 1 ARMA term

Lag	AC	PAC	Q-Stat	Prob
1	-0.051	-0.051	0.1590	
2	0.145	0.143	1.4516	0.228
3	-0.044	-0.031	1.5708	0.456
4	0.038	0.015	1.6633	0.645
5	0.086	0.102	2.1450	0.709
6	0.063	0.064	2.4102	0.790
7	0.056	0.038	2.6208	0.855
8	-0.020	-0.027	2.6475	0.916
9	0.033	0.019	2.7216	0.951
10	-0.004	-0.003	2.7225	0.974
11	-0.085	-0.113	3.2502	0.975
12	0.137	0.126	4.6613	0.946
13	0.030	0.069	4.7287	0.966
14	0.088	0.045	5.3382	0.967
15	0.060	0.075	5.6307	0.975
16	-0.087	-0.093	6.2490	0.975
17	0.238	0.226	10.996	0.810
18	-0.076	-0.062	11.491	0.830
19	-0.035	-0.168	11.601	0.867
20	0.015	0.063	11.622	0.901
21	-0.073	-0.094	12.122	0.912
22	0.009	-0.048	12.129	0.936
23	-0.155	-0.134	14.491	0.883
24	0.073	0.060	15.041	0.893

Figure 3. Correlogram of residuals ARIMA (1,1,0)

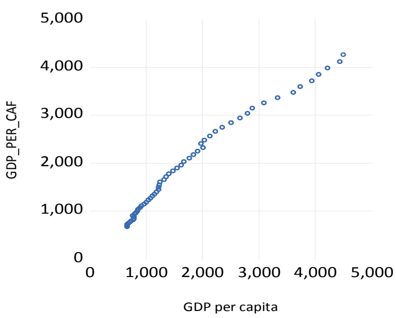


Figure 4: Actual vs Predicted GDP per capita from 1961 to 2018 (training set).

There is a very strong linear relationship between actual and predicted GDP per capita (r=0.999). Therefore, it can be concluded that the predicting power of the model is good.

Year	GDP	GDPF	%error
2019	4458.435	4614.105	-3.49
2020	4280.849	4541.988	-6.10
2021	4375.948	4310.746	1.49

Table 3. Percentage errors of independent data.

For the training set (1961-2018), the percentage error varies from -6.50% to 3.80% whereas for the independent data set percentage error varies from -3.49% to 1.49%.

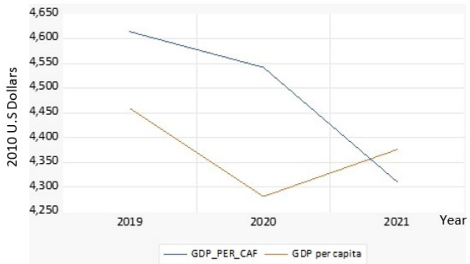


Figure 5. Plot of predicted and actual GDP for independent data (2019 to 2021).

As shown in Figure 5, the forecasted and actual value is much closer in 2021 compared to the other two years since there is a considerable gap. It can be observed that from 2019 to 2020 there has been an overestimation in the forecasted values of GDP. Especially in 2020, there was a huge gap this is because the Covid-19 pandemic had a detrimental impact on Sri Lanka's economy, leading to a decline in economic activities and a decrease in GDP per capita. The country faced challenges such as disruptions in global trade, a significant decline in the tourism industry, and widespread job and income losses.

Year	GDPF	GDP per capita growth %
2022	4506.728	2.99
2023	4653.895	3.27
2024	4810.505	3.37

Table 4. Predicted values of GDP per capita

It can be observed that the GDP per capita value of Sri Lanka is expected to increase in the upcoming years with GDP per capita growth also increasing simultaneously.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. CONCLUSIONS

The aim of this study was to model and forecast Sri Lanka's constant GDP per capita (CGDPC) by applying the Box Jenkins ARIMA approach, based on annual data from 1961 to 2021, as the literature survey confirmed no model has been developed to forecast CGDPC. The best-fitted model developed for the log of the CGDPC series is ARIMA (1,1,0). The error series of the best-fitted model is random and has a constant variance. Furthermore, the autocorrelation function of the series of squared errors was also found to be random. The model was tested for an independent data set and it was found that the percentage of errors is low. The model is very flexible and easy to use. The CGDPC of Sri Lanka is forecasted to experience a steady upward trend in the years 2022 to 2024 with GDP per capita rates for the years 2022, 2023, and 2024 being 2.99%, 3.27%, and 3.37% respectively. The results of this research will benefit policy-makers in formulating strategies and provide valuable insights for business executives making decisions on business expansion.

4.2. RECOMMENDATIONS

It is important to note that the forecasted values may not always be able to accurately reflect reality due to various external factors that can impact the country's economy. Further research could focus on incorporating suitable continuous explanatory variables. Moreover, during the period 1961-2018 there may have been some structural breaks. Thus, it is recommended that the model can be improved by considering dummy variable for structural breaks.

REFERENCES

Box, G., Jenkins, G., Reinsel, G., & Ljung, G. (2015). *Time Series Analysis: Forecasting and Control* . John Wiley & Sons.

Dilhani, E. (2017, November).*Forecasting Gross Domestic Production (GDP) of Sri Lanka from 2016-2025 an Empirical Study.* https://www.researchgate.net/publication/348431527_Forecasting_Gross_Domestic_Production_GDP_of_Sri_Lanka_from_2016-2025_an_Empirical_Study

Eissa, N. (2020, March 30). *Forecasting the GDP per Capita for Egypt and Saudi Arabia using ARIMA Models*. https://www.sciedu.ca/journal/index.php/rwe/article/view/17182/10875

Khashei, M., & Bijari, M. (2010, January). *An artificial neural network (p, d, q) model for timeseries forecasting.* https://www.sciencedirect.com/science/article/abs/pii/S0957417409004850

Liu, H. F. (2002, October). *A SETAR Model for Canadian GDP: Non-linearities and Forecast Comparisons* https://www.uvic.ca/socialsciences/economics/assets/docs/econometrics/ewp0206.pdf

Maity, B., & Chatterjee, B. (2012). F*orecasting GDP Growth Rates of India: An Empirical Study* https://www.hilarispublisher.com/open-access/forecasting-gdp-growth-rates-of-india-an-empirical-study-2162-6359-1-082.pdf

Ranasinghe, P., & Suriyaarachchi, D. (2015).

Forecasting GDP in Sri Lanka: http://graduate.sjp.ac.lk/icma/wp-content/uploads/2019/10/iCMA_abstract_2015_P152.pdf

Saha, A., & Khanam, M. (2022, July). *Forecasting the Growth Rate of GDP and its Indicators Remittance and Labor Force Participation of Bangladesh using VECM and ARIMA Models* https://www.banglajol.info/index.php/DUJS/article/view/62605

The Investopedia Team. (2023, May). *GDP Per Capita Defined; Applications and Highest Per Country* https://www.investopedia.com/terms/p/per-capita-gdp.asp

Tramplin, T. (2023, March 29). *GDP per capita*. https://www.financestrategists.com/wealth-management/macroeconomics/gdp-per-capita/

Urasawa, S. (2014, December). *Real-time GDP forecasting for Japan: A dynamic factor model approach*. https://www.sciencedirect.com/science/article/abs/pii/S0889158314000422

Voumik, L., & Smrity, D. (2020, September). *View of Forecasting GDP Per Capita In Bangladesh: Using ARIMA model.* https://www.ejbmr.org/index.php/ejbmr/article/view/533/304