MODELING AND FORECASTING OF INDIA’S DEFENSE EXPENDITURES USING BOX-JENKINS ARIMA MODEL

Sapana Sharma ¹, Sanju Karol ²

¹ Ph. D research scholar, Dept. Of Economics, Himachal Pradesh University, Shimla, 171005, India
² Professor of Economics, ICDEOL, Himachal Pradesh University, Shimla-171005, India

ABSTRACT

Many developed and developing countries are at the core of the security and peace agenda concerning rising defense expenditure and its enduring sustainability. The unremitting upsurge in defense expenditure pressurizes the government to rationally manage the resources so as to provide security and peace services in the most efficient, effective and equitable way. It is necessary to forecast the defense expenditure in India which leads the policy makers to execute reforms in order to detract burdens on these resources, as well as introduce appropriate plan strategies on the basis of rational decision making for the issues that may arise. The purpose of this study is to investigate the appropriate type of model based on the Box–Jenkins methodology to forecast defense expenditure in India. The present study applies the one-step ahead forecasting method for annual data over the period 1961 to 2020. The results show that ARIMA (1,1,1) model with static forecasting being the most appropriate to forecast the India’s defense expenditure.

1. INTRODUCTION

The global military expenditure was estimated $1917 billion in 2019 as per the report of Trends in World Expenditure, 2020. In total world’s share India contributed around 3.7% and spending was $71.1 billion in 2019 and became the third largest military spender in the world following US and China. Comparing this with prior year (2018), India moved from fourth position to third position, surpassing Saudi Arabia (SIPRI, 2020). Considering the past records, there was an increasing spending trend since the previous years due to conflicts and hostile relation with her neighboring countries especially with China and Pakistan. While exploring the SIPRI reports for military expenditure, India’s expenditure grew by 259 per cent over the 30 year periods 1990 to 2019.

Upsurge in defense expenditure must lead to GDP growth. Otherwise, the unremitting rise in defense expenditure puts an exertion on resources and hoodoo to fulfill other requirements and requisitions of the population. Upsurge in defense expenditure expects prudent management in order to endow security in the most efficient and effective way. To ponder these facts, it becomes a needful piece of work to forecast the defense expenditure for planning the strategies efficiently in advance. The requirement for more precise forecast of defense expenditure in order to avert the risk of uncertainty has led to the advancement and upgrading of time series models
over the many years. Spacious endeavors have been made in the research community to advance and improve these models. Box-Jenkins (1976) is one of the most vital and broadly used methodologies in time series models.

The purpose of this paper is to build the most appropriate model to investigate and forecast defense expenditure in India. For this purpose, the ARIMA models and the Box-Jenkins methodology are used. The remaining sections of this research paper are structured as follows: Section 2 outlines a theory and evidence, followed by a research methodology in Section 3. The empirical results and discussion of the paper are analyzed in Section 4. Finally, conclusions of this study are presented in Section 5.

2. THEORY AND EVIDENCE

Few time series models have been proposed in the literature to project defense expenditure over time. Cheng, Chang & Lin (2013) used autoregressive integrated moving average (ARIMA) models and artificial neural networks (ANNs) models to forecast China’s military spending for forecasting the short term (1 year), the medium term (3 year), the medium-long term (5 year) and the long term (10 years). To measure accuracy of forecasting the mean absolute percentage error (MAPE) approach is applied. The results indicated that these single variable ARIMA models showed higher accuracy and stability than those used by the single variable ANNs models across the four time periods. For the ANNs models of multiple variables, the accuracy of each forecasted models with different variables had advantages in different time periods.

Recently, a similar study was conducted to forecast the military expenditure of India in forthcoming times. This study used the Box-Jenkins ARIMA model for time series forecasting over the period 1960 to 2019. This study highlighted the minimum AIC value and involved ADF testing (Augmented Dickey-Fuller) to transform expenditure data into stationary form for model generation. The results from this study showed that ARIMA (0,1,6) model was the appropriate model to forecast the India’s defense expenditures with an accuracy of 95.7%. The model, thus, acted as a Moving Average (MA) model and predicted the steady-state exponential growth of 36.94% in military expenditure of India by 2024 (Sharma & Phulli, 2020). Additionally, over the past years, there have been many researches in the field of time series forecasting modeling such as Church & Curram (1996), Zhang and Michael (1998), Omer (2010), Khashei & Bijari (2011), Cheng, Chang & Lin (2013). Time series forecasting is beneficial to develop correct vision about unknown future and plan appropriate strategies on the basis of rational decision making. Many researchers have actively dedicated their studies in designing an optimized model for forecasting using ARIMA such as Stergiou (1989), Raymond (1997), Meyler, Kenny and Terry (1998), Koutroumanidis, Konstantinos and Arabatzis (2009), Omer (2010), Yue, Shengnan and Yuan (2015), Peijun (2016), Fattah et al. (2018), Klazoglou & Dritsakis (2018). Notwithstanding, the little amount of research has been done for the forecasting of the defense expenditure in India using Box-Jenkins ARIMA approach.

3. RESEARCH METHODOLOGY

Considering the above applications, a similar attempt has been made in this study to forecast the defense expenditure in India for the forthcoming five years (2021-25). This study used the ARIMA model for the annual time series data from 1961 to 2020. The data are collected from Defense Expenditure Statistics (DES) published by Ministry of Finance and Handbook of Statistics of Reserve Bank of India. The expenditure is expressed in billion rupees according to the calendar year. This forecasting would benefit at analyzing the defense expenditures ahead of time.

3.1. BOX-JENKINS (ARIMA) MODEL: BASICS

Box and Jenkins (1976) method is also known as the ARIMA model and used for constructing time series models. Box-Jenkins ARIMA models (1976) usage the historical values of a single variable to forecast its future values; hence they are classified as univariate methods. The Box-Jenkins methodology or ARIMA models is based on the following steps: (1) Model Identification, (2) Model Estimation and Selection, (3) Diagnostic Checking and (4) Model Forecasting.
**Model Identification**

To build ARIMA model in model identification it involves selecting the p-order of autoregressive term, d-order of differencing and q-order of moving average (p,d,q). To identify the ARIMA(p,d,q) model as well as unit root tests the Box- Jenkins (1976) method uses the autocorrelation function, ACF, and the partial autocorrelation function, PACF, as well as ADF (Augmented Dickey–Fuller) (1979, 1981) test and PP Phillips–Perron (1998) test. First step of ARIMA model is to recognize whether the variable, which is being forecasted, is stationary in time series or not. The level of stationarity determines the order of differencing. Once stationarity has been checked, the next step is to identify the order of p for the autoregressive process and the order of q for the moving average process. (Dritsakis & Klazoglou, 2019).

**Model estimation and selection**

Model identification is followed by the model estimation of orders p and q, from the parameters. The estimation methods determine the number of parameters which depends on the type and order of the model. There are numerous methods to estimate model parameters such as The Method of Least Squares, The Method of Moments, The method of Maximum Likelihood, The method of Conditional Least Squares. The Method of least square is considered as an easiest way which is more widely used to estimate model parameters. The purpose of Least Squares method is to select parameter estimators which minimize the sum of squared residuals between real values and the estimated values of the dependent variable (Alfaki, 2016).

**Diagnostic Checking**

This step examines for the residuals of ARIMA models obtained from ACF and PACF graphs in the previous stage to be independent and identically distributed. Diagnostic checking involves not only evaluating residual behavior and the order of the model but also testing the significance of the coefficients. It testifies whether they satisfy the characteristics of a white noise process. The Q statistic of Ljung–Box (1978), is used to check the residual, whereby the statistical significance of the autocorrelation coefficients is tested.

**Model forecasting**

After selection of adequate and valid estimated model, n period further forecasts of defense expenditure in India (DE) have been generated by using following equations:

\[
\text{DE}_{t+n} = \alpha + \sum_{i=1}^{p+d} \alpha_i \text{DE}_{t+n-i} + \epsilon_{t+n} - \sum_{i=1}^{q} \epsilon_i \epsilon_{t+n-i} 
\]

Where \( \alpha_0, \alpha_1 \) and \( \epsilon_i \) are unknown parameters

4. **EMPIRICAL RESULTS AND DISCUSSION**

To avoid heteroskedasticity as well as asymmetric distributions problems the data is transformed into natural logarithms. The time plot of the defense expenditure data in Figure 1 and Figure 2 correlogram evidently indicates that the data is not stationary. In Figure 1, it is observed that India’s defense expenditure increased throughout the period of study. Figure 2 represents the autocorrelation (ACF) and partial autocorrelation (PACF) plots to confirm whether stationarity exists or not and it is observed that the autocorrelation coefficients decline slowly, indicating that the series is non-stationary.
The stationarity of defence expenditure has also been tested through Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) unit root test. The outcomes of ADF and PP unit root test have been presented in table 1. Figure 3 shows the trend of defense expenditure in the first difference.
Table 1: Results of Augmented Dickey–Fuller and Phillips–Perron Unit Root Tests

| Variable | ADF Intercept & Trend | PP Intercept & Trend |
|----------|-----------------------|----------------------|
| LDE      | (0.229)               | (0.0154)             |
| DLDE     | (0.000)*              | (0.000)*             |

Note: * indicates significant level at 5% level.

The outcomes of ADF test discloses that at level defense expenditure of India is non-stationary but becomes stationary after taking first difference. The results have also been supported by PP test (Table 1). In Figure 3 and 4, the autocorrelation and partial autocorrelation plot in first difference shows that the autocorrelation coefficients decreased at a rapid rate, representing that the series is stationary. The results from Table 1 and Figure 3 indorse that the time series is stationary in first difference. Therefore, for the ARIMA (p,d,q) model under investigation differencing ‘d’ is reflected as one.

The form of ARMA(p,q) model is defined from the results of Figure 4. The parameters p and q of the ARMA model can be determined by remarking the partial autocorrelation and autocorrelation coefficients, respectively, comparing them with the critical value

\[ \pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{59}} = \pm 0.260 \]

The plot of auto-correlation function (ACF) and PACF (partial auto-correlation function) correlogram for lags 1 to 24 of the first order differenced time series of the defense expenditure in India are presented in the Figure 4.
The above correlogram indorses that the auto-correlation at lag 1 does not increase the significance limits and auto-correlations tail off to zero from lag 1. Although, the all coefficients between lag 1 and 24 are well within the upper and lower confidence limits. The partial correlogram, showed above in Figure 4, also indicates that partial autocorrelation Coefficient also does not increase the significant limits at lag 1 and after lag 1 partial autocorrelation tails off to zero and also lag within the significant limits.

Figure 4 shows the value of the autocorrelation and partial autocorrelation coefficients for the p and q is between 0<p<1 and 0<q<1 respectively. Therefore, based upon the conditions study identified only following three tentative ARIMA (p,d,q) models:

ARIMA(p,d,q): ARIMA(0,1,1), ARIMA(1,1,1), and ARIMA(1,1,0)

All possible identified three models have been estimated and the results of estimated parameters are presented in Table 2. Perusal of table indicates that ARIMA (0,1,1) has constant equal to 0.120 and moving average term MA(1) equal to 0.260. ARIMA (1,1,0) has constant equal to 0.110 AR(1) equal to 0.194. The estimated parameters constant, autoregressive term AR (1), moving average term MA (1) for ARIMA (1,1,1) comes out to be 0.107, 0.434 and -0.549 respectively.

Table 2: Estimated Parameters of Identified Model Using ARIMA

| Model | C | AR(1) | MA(1) | RSS | Adj.R² | AIC | SC | Sig. coefficient | Log likelihood |
|-------|---|-------|-------|-----|--------|-----|----|------------------|----------------|
| 011   | 0.120 | 0.260 | 0.5812 | 0.0337 | 1.7144 | 1.6440 | 2 | 52.577 |
| 110   | 0.110 | 0.194 | 0.4479 | 0.0319 | 1.9566 | 1.8855 | 1 | 58.742 |
| 111   | 0.107 | 0.434 | 0.5490 | 0.3771 | 1.7000 | 1.9875 | 3 | 63.729 |

Source: e-views output based on formulated DE

The results from Table 2 indicates that according to the Akaike Information Criterion (AIC), Schwartz criterion (SC), adjusted R square, significant coefficient and RSS the ARIMA (1,1,1) model is the most appropriate. Choosing this model because of minimum criteria of RSS, AIC and SC i.e. 0.3771, -2.0941 and -1.9875 respectively and
Sapana Sharma, and Sanju Karol

According to adjusted R2 and log likelihood which is highest, and most significant coefficient model in this ARIMA (1, 1, 1).

After identifying the most suitable model which is ARIMA (1, 1, 1), and equation 1 displays the specific form of the model.

\[
LDE_t = 0.106677 + 0.434098 LDE_{t-1} - 0.549873 LDE_{t-2}
\]

(2)

Table 3: Estimated Parameters of Identified Model ARIMA (1,1,1)

| Variable        | Coefficient | Std. Error | t-Statistic | Prob. |
|-----------------|-------------|------------|-------------|-------|
| C               | 0.106677    | 0.009009   | 11.84136    | 0.0000|
| AR(1)           | 0.434098    | 0.130139   | 3.335655    | 0.0015|
| MA(1)           | -0.549873   | 0.175290   | -3.136935   | 0.0027|
| R-squared       | 0.199204    | Mean dependent var | 0.112489 |
| Adjusted R-squared | 0.170084  | S.D. dependent var | 0.090904 |
| S.E. of regression | 0.082813 | Akaike info criterion | -2.094115|
| Sum squared resid | 0.377193 | Schwarz criterion | -1.987540|
| Log likelihood  | 63.72933    | Hannan-Quinn criter. | -2.052602|
| F-statistic     | 6.840827    | Durbin-Watson stat | 1.762525|
| Prob(F-statistic) | 0.002223 |                      |        |
| Inverted AR Roots | 43      |                      |        |
| Inverted MA Roots | 55       |                      |        |

Source: E-Views

Here the data in below the equation is the t-test statistic of the corresponding estimate value.

\[
DLDE = 0.106677 + 0.434098 LDE_{t-1} - 0.549873 LDE_{t-2}
\]

(2)

The estimated value of the variance of the corresponding error term is \( \sigma^2 = 0.082813 \)

It can be seen from the coefficients of t-statistic of the model and its P value that the parameter estimates of all explanatory variables are significant at the 5 per cent level of significance.

After the identification and estimation of the model, the test of appropriated of the estimated model was performed. First of all, to confirm that autocorrelation is not present in residuals of estimated models ARIMA (1,1,1), the residuals correlogram for model ARIMA (1,1,1) was plotted in Figure 5. Furthermore, absence of autocorrelation can also be confirmed from the Ljing-Box Q statistics that was plotted in Figure 6.

Perusal of figure 5 clearly indicates that coefficient of autocorrelation and partial autocorrelation of residual term for estimated model does not possess any significant pattern and lie in between lower and upper critical limits indicating absence of significant autocorrelation among residuals. Moreover, the values for
**Figure 5:** Autocorrelation and Partial Autocorrelation Functions of Residual Terms of Identified Model (1,1,1)

Source: E-Views

Ljung-Box Q statistics are insignificant for estimated model presented in Figure 6 and indicating that the model is valid.
Forecasts of defense expenditure has been generated using best selected ARIMA model, that is, ARIMA(1,1,1) and have been presented in Table 6.7. Table 6.7 shows that forecasted in India in 2021-22 is Rs 3633.04 billion and expected to rise to 4060.94 billion in 2022-23 and to Rs 4527.26 billion in 2023-24 and finally expected to reach 5041.26 billion in 2024-25 and 5611.02 billion in 2025-26.

However, year to year growth rate of forecast levels of defense expenditure in India has shown a declining trend.

**Table 4: Forecast of Defense Expenditure in India**

| Year     | Forecasts | Year to year growth rate (in percentage) |
|----------|-----------|----------------------------------------|
| 2021-22  | 3633.04   | 11.08                                  |
| 2022-23  | 4060.94   | 10.54                                  |
| 2023-24  | 4527.26   | 10.30                                  |
| 2024-25  | 5041.26   | 10.19                                  |
| 2025-26  | 5611.02   | 10.15                                  |

Source: E-views 9 outputs based on Defense Expenditure

It is clear from the Table 4 that forecasted defense expenditure in 2021-22 has shown an increase at the rate of 11.08 per cent in comparison to actual level of defense expenditure in 2020-21 which was 5.82 per cent. Later, defense expenditure has shown a continuous declining growth rate for defense expenditure after 2022-23. In 2022-23 growth rate was 10.54 per cent and declined to 10.15 per cent in 2025-26.
5. CONCLUSION

The continuous rise in defense expenditure pressurizes the government to rationally manage the resources so as to provide security and peace services in the most efficient, effective and equitable way. It is necessary to forecast the defense expenditure in India to expect forthcoming changes viz. opportunities as well as threats (Fauziah & Gunaryati, 2017), consequently that planners can adapt to these changes for successful implementation of planning (Kurzak, 2012). Forecast is beneficial to develop correct vision about unknown future and plan appropriate strategies on the basis of rational decision making. An endeavor has been made in this study to forecast the defense expenditure in India. In order to forecast the defense expenditure in India, Autoregressive integrated moving average method has been employed. For this the ARIMA (1,1,1) has the best selected model for making forecasts for up to 5 years from 2021-22 to 2025-26 for the defense expenditure in India using a 59 years’ time series data from 1961-62 to 2020-21. The reason for choosing the ARIMA model is that its reliability to make forecasts using a time series data with any kind of trend and autocorrelation to forecast the future successive values in the time series. The study is also statistically tested and validated the results of defense expenditure forecast because it confirms that autocorrelation is not present in residuals of estimated models ARIMA (1,1,1).

Hence, it can be concluded that the selected ARIMA(1,1,1) seems to endow a sufficient predictive model for defence expenditure in India. The ARIMA(1,1,1) model forecasts an increased defense expenditure from 2021-22 to 2025-26 in absolute terms. However, in terms of annual growth rate it has shown a continuous decline in defense expenditure during the 2021-22 to 2025-26.

5. SOURCES OF FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

5. CONFLICT OF INTEREST

The author have declared that no competing interests exist.
ACKNOWLEDGMENT

None.

REFERENCES

[1] Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis. Forecasting and control. San Francisco: Holden-Bay.
[2] Church, Keith B., and Stephen P. Curram. (1996). Forecasting consumers' expenditure: A comparison between econometric and neural network models. International journal of forecasting 12(2), 255-267.
[3] Dickey, D. A., & Fuller, W. A. (1979). Distributions of the estimators for autoregressive time series with a unit root. Journal of American Statistical Association, 74(366), 427–431.
[4] Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49(4), 1057–1072. Koutroumanidis, Theodoros, Konstantinos Ioannou, and Garyfallos Arabatzis. (2009). Predicting fuel wood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA–ANN model. Energy Policy 37(9), 3627-3634.
[5] Kuo, Kuo-Cheng, Chi-Ya Chang, and Wen-Cheng Lin. (2013). To predict military spending in China based on ARIMA and artificial neural networks models. Przegląd Elektrotechniczny 89(3b), 176-181.
[6] Faruk, Durdu Ömer. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. Engineering applications of artificial intelligence, 23(4), 586-594.
[7] Meyler, Aidan; Kenny, Geoff and Quinn, Terry (1998). Forecasting Irish inflation using ARIMA models, Central Bank and Financial Services Authority of Ireland Technical Paper Series, 1998 (3/RT/98), 1-48.
[8] Peijun, C. (2016). Predictive Analysis of Chinese Total Health Expenditure Base on ARIMA Model. Medicine and Society, (03).
[9] Phillips, P. C. B., & Perron, P. (1998). Testing for a unit root in time series regression. Biometrika, 75(2), 335–346.
[10] Raymond Y.C. Tse (1997). An application of the ARIMA model to real-estate prices in Hong Kong. Journal of Property Finance, 8(2), 152 – 163.
[11] Sharma Deepanshu & Kritika Phulli (2020 ). Forecasting and Analyzing the Military Expenditure of India Using Box-Jenkins ARIMA Model. arXiv:2011.06060 [econ.GN] General Economics (econ.GN).
[12] SIPRI (2020). SIPRI Military Expenditure Database. Stockholm International Peace Research Institute (SIPRI) published by Oxf ord University Press. www.sipri.org/databases/milex.
[13] Stergiou, K. I. (1989). Modeling and forecasting the fishery for pilchard (Sardina pilchardus) in Greek waters using ARIMA time-series models. ICES Journal of Marine Science, 46(1), 16-23.
[14] Yue Z., W. Shengnan and L. Yuan (2015). Application of ARIMA Model on Predicting Monthly Hospital Admissions and Hospitalization Expenses for Respiratory Diseases. Chinese Journal of Health Statistics, (02), 197-200.